UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURÍMAC

FACULTAD DE INGENIERÍA

ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA CIVIL

Tesis

Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025

Presentado por:

Julio Cesar Riveros Ramos

Para optar el título de Ingeniero Civil

Abancay, Perú

2026

Repositorio Institucional - UNAMBA Perú

UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURÍMAC

FACULTAD DE INGENIERÍA

ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA CIVIL

TESIS

Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025

Presentado por **Julio Cesar Riveros Ramos**, para optar el título de Ingeniero Civil

Sustentado y aprobado el 11 de agosto del 2026 ante el jurado evaluador:

Presidente: *Dr. Calixto Cañari Otero*

Primer miembro: *Ing. Milson Morán Moreno*

Segundo miembro: *Ing. Fanny Silva Noriega*

Asesores: *Mag. Diomedes Napoleón Ferrel Sarmiento*

Ing. Edison Ponce Torres

Repositorio Institucional - UNAMBA Perú

Año de la Esperanza y el Fortalecimiento de la Democracia"

CONSTANCIA DE ORIGINALIDAD
Nº 229-2026

La Universidad Nacional Micaela Bastidas de Apurímac, a través de la Unidad de Investigación de la Facultad de Ingeniería declara que, el Informe de Tesis titulada: **Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurimac - 2025**, presentado por el Br. **Julio Cesar Riveros Ramos**, para optar el Título de **Ingeniero Civil**; ha sido sometido a un mecanismo de evaluación y verificación de similitud, a través del Software Turnitin, siendo el índice de similitud ACEPTABLE de **(19%)** por lo que, cumple con los criterios de originalidad establecidos por la Universidad.

Abancay, 07 de agosto del 2026

Atentamente,

C. c.
Archivo
WHC/ DUI-FI
Yovana/sec
REG. N° 801

Repositorio Institucional - UNAMBA Perú

Agradecimientos

Quiero agradecer a mi universidad y a mis docentes por los conocimientos y orientaciones brindadas a lo largo de mi formación.

A mis padres: Alejandro Riveros Bautista y Constantina Ramos Tomaylla quienes me apoyaron durante el desarrollo de este trabajo, ya sea con consejos, apoyo técnico o palabras de ánimo.

A mi asesor Ing. Diomedes N. Ferrel Sarmiento y coasesor Ing. Edison Ponce Torres por apoyarme y lograr este anhelado objetivo.

Finalmente, agradezco a todas las personas que confiaron en mí y contribuyeron, directa o indirectamente, a la culminación de este trabajo.

Julio Cesar

Repositorio Institucional - UNAMBA Perú

Dedicatoria

Dedico este trabajo a Dios, por darme la fuerza y constancia necesarias para seguir adelante.

A mi familia, en especial a mis padres: Alejandro Riveros Bautista y Constantina Ramos Tomaylla, así como a mis hermanos: Jaime, Christopher, John y Sheyla Milagros, por su apoyo permanente, paciencia y confianza, que fueron fundamentales para alcanzar este objetivo.

Este esfuerzo también está dedicado a todas las personas que, de una u otra manera, formaron parte de este proceso y me alentaron a no rendirme.

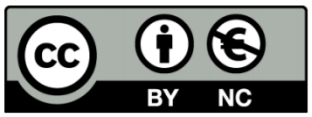
Julio Cesar

Repositorio Institucional - UNAMBA Perú

Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025

Línea de Investigación: Ingeniería de materiales

Esta publicación está bajo una Licencia Creative Commons.



Repositorio Institucional - UNAMBA Perú

ÍNDICE

		Pág.
INTRODUCCIÓN		1
RESUMEN		3
ABSTRACT		4
CAPÍTULO I		5
PLANTEAMIENTO DEL PROBLEMA		5
1.1	Descripción del problema	5
1.2	Enunciado del problema	7
1.2.1	Problema general	7
1.2.2	Problemas específicos	7
1.3	Justificación de la investigación	7
CAPÍTULO II		9
OBJETIVOS E HIPÓTESIS		9
2.1	Objetivos de la investigación	9
2.1.1	Objetivo general	9
2.1.2	Objetivos específicos	9
2.2	Hipótesis de la investigación	9
2.2.1	Hipótesis general	9
2.2.2	Hipótesis específicas	9
2.3	Operacionalización de variables	10
CAPÍTULO III		11
MARCO TEÓRICO REFERENCIAL		11
3.1	Antecedentes	11
3.1.1	Antecedentes internacionales	11
3.1.2	Antecedentes nacionales	16
3.2	Marco teórico	21
3.2.1	Machine Learning	21
3.2.2	Bloques de construcción de un proyecto de aprendizaje	23
3.2.3	Machine Learning vs. programas informáticos tradicionales	24
3.2.4	Tipos de sistemas de Machine Learning	26
3.2.4.1	Según el tipo de supervisión	26
3.2.4.2	Según cómo generalizan el aprendizaje	27
3.2.4.3	Según cómo comparan los datos de entrada	28
3.2.5	Pruebas y validación en Machine Learning	29

Repositorio Institucional - UNAMBA Perú

3.2.6 Diseño del modelo de Machine Learning 29
3.2.7 Precisión del modelo 30
3.2.8 Validación del modelo 32
3.2.9 Desajuste de datos 33
3.2.10 Modelos de entrenamiento en Machine Learning 34
3.2.11 Expansión volumétrica 37
3.2.12 Proctor 38
3.2.12.1 Tipos de ensayo Proctor 39
3.2.13 Parámetros geotécnicos 41
3.2.13.1 Contenido de humedad - ASTM D 2216 42
3.2.13.2 Análisis granulométrico - ASTM D 422 43
3.2.13.3 Clasificación de suelos 45
3.2.13.4 Límites de Atterberg 47
3.2.14 Respuesta mecánica 50
3.2.14.1 Capacidad de soporte del suelo (CBR) 50
3.2.14.2 Rigidez de la subrasante (Módulo resiliente, Mr) 51
3.3 Marco conceptual 53
CAPÍTULO IV 57
METODOLOGÍA 57
4.1 Tipo y nivel de investigación 57
4.1.1 Tipo de investigación 57
4.1.2 Nivel de investigación 57
4.2 Diseño de la investigación 57
4.3 Descripción ética de la investigación 58
4.4 Población y muestra 59
4.4.1 Población 59
4.4.2 Muestra 60
4.5 Procedimiento 60
4.6 Técnica e instrumentos 62
4.6.1 Técnica 62
4.6.2 Instrumentos 63
4.7 Estadístico de investigación 64
4.7.1 Estadística descriptiva 64
4.7.2 Estadística inferencial 65
CAPÍTULO V 68
RESULTADOS Y DISCUSIONES 68
5.1 Análisis de resultados 68
5.1.1 Análisis de los resultados de variable dependiente: Respuesta mecánica 69
5.1.1.1 Capacidad de soporte del suelo 69

Repositorio Institucional - UNAMBA Perú

5.1.1.2 Rigidez de la subrasante 79
5.1.2 Análisis de los resultados de variable independiente: Modelo predictivo basado en Machine Learning 81
5.1.2.1 Modelo predictivo 81
5.1.2.2 Desempeño del modelo de Machine Learning 82
5.1.3 Análisis estadístico descriptivo 85
5.1.3.1 Variable dependiente (Capacidad de soporte del suelo) 85
5.1.3.2 Variable dependiente (Rigidez de la subrasante) 87
5.1.3.3 Variable independiente (Modelo predictivo) 88
5.1.3.4 Variable independiente (Desempeño del modelo) 89
5.1.4 Prueba de normalidad por cada variable 91
5.2 Contrastación de hipótesis 93
5.2.1 Contrastación de hipótesis general 93
5.2.2 Contrastación de hipótesis especificas 95
5.3 Discusión 99
CAPÍTULO VI 102
CONCLUSIONES Y RECOMENDACIONES 102
6.1 Conclusiones 102
6.2 Recomendaciones 103
REFERENCIAS BIBLIOGRÁFICAS 105
ANEXOS 110

Repositorio Institucional - UNAMBA Perú

ÍNDICE DE TABLAS

Pág.

Tabla 1 — Operacionalización de variables 10
Tabla 2 — Clasificación de suelos expansivos 38
Tabla 3 — La diferencia entre ambos ensayos 40
Tabla 4 — Principales parámetros geotécnicos 42
Tabla 5 — La cantidad mínima de espécimen de material húmedo 42
Tabla 6 — Tamices de malla cuadrada 43
Tabla 7 — Estimados de precisión LL 49
Tabla 8 — Estimados de precisión LP 49
Tabla 9 — Correlaciones empíricas para la estimación del Módulo Resiliente (Mr) 52
Tabla 10 — Ubicación de las calicatas en evaluación 68
Tabla 11 — Clasificación y plasticidad de suelos 70
Tabla 12 — Clasificación del suelo según el Índice de Plasticidad (IP). 70
Tabla 13 — Compactación Proctor y capacidad de soporte CBR 77
Tabla 14 — Comparación de valores de CBR de laboratorio y CBR predicho por modelo de Machine Learning 78
Tabla 15 — Diferencia absoluta y porcentual entre CBR de laboratorio y CBR predicho 78
Tabla 16 — Comparación de Mr obtenido por correlaciones empíricas y Machine Learnig 79
Tabla 17 — Resultados de laboratorio y Machine Learning 80
Tabla 18 — Métricas por tipo de modelo 81
Tabla 19 — Desempeño del modelo predictivo 83
Tabla 20 — Estadística de capacidad de soporte del suelo 86
Tabla 21 — Estadística de la rigidez de la subrasante 87
Tabla 22 — Estadística del modelo predictivo 88
Tabla 23 — Estadística del desempeño del modelo 90
Tabla 24 — Prueba de normalidad Machine Learning 91
Tabla 25 — Prueba de normalidad Machine Learning 92
Tabla 26 — Prueba de correlación de Spearman para la hipótesis general 94
Tabla 27 — Prueba de correlación Spearman para la hipótesis especifica 1 96
Tabla 28 — Prueba de correlación Spearman para la hipótesis especifica 2 98
Tabla 29 — Matriz de consistencia 111
Tabla 30 — Validación de instrumentos (Alfa de Cronbach) 148
Tabla 31 — Base de datos geotécnica empleada para el desarrollo del modelo predictivo basado en Machine Learning 153

Repositorio Institucional - UNAMBA Perú

ÍNDICE DE FIGURAS

Pág.

Figura 1 — Un proyecto típico de aprendizaje automático en la práctica real 24
Figura 2 — El enfoque tradicional de Machine Learning 25
Figura 3 — Enfoque de aprendizaje automático 25
Figura 4 — Adaptarse automáticamente al cambio 25
Figura 5 — Machine Learning puede ayudar a los humanos a aprender 26
Figura 6 — Detección de anomalías 28
Figura 7 — Regresión 28
Figura 8 — Curvas de aprendizaje 32
Figura 9 — Predicciones del modelo de regresión polinómica 33
Figura 10 — Matriz de dispersión de datos 34
Figura 11 — Curva de compactación 40
Figura 12 — Molde cilíndrico de 4,0 pulg 41
Figura 13 — Molde cilíndrico de 6,0 Pulg 41
Figura 14 — Curva de distribución de tamaño de partícula 44
Figura 15 — Definición de D10, D30 y D60 44
Figura 16 — Límites de separación de tamaño de suelo 45
Figura 17 — Sistema unificado de clasificación de suelo 46
Figura 18 — Clasificación de suelos según AASHTO 47
Figura 19 — Límites de Atterberg. 48
Figura 20 — Gráfica de plasticidad 48
Figura 21 — Comparación de los límites de Atterberg (LL, LP e IP) de las calicatas evaluadas 71
Figura 22 — Clasificación de los suelos en la Carta de Plasticidad de Casagrande (SUCS) 72
Figura 23 — Comparación de las propiedades índice mediante gráfico radar 73
Figura 24 — Perfil estratigráfico de la calicata C-01 74
Figura 25 — Perfil estratigráfico de la calicata C-02 75
Figura 26 — Perfil estratigráfico de la calicata C-03 76
Figura 27 — Resultados del ensayo Proctor 77
Figura 28 — Comparación del módulo resiliente (Mr) estimado por métodos empíricos y Machine Learning 80
Figura 29 — Comparación entre valores reales y predichos de CBR_95 84
Figura 30 — Comparación entre valores reales y predichos de CBR_100 85
Figura 31 — Comparación de la respuesta mecánica 94
Figura 32 — Comparación de la capacidad de soporte del suelo (CBR) 97

Repositorio Institucional - UNAMBA Perú

Figura 33 — Comparación del Mr 99
Figura 34 — Instrumento contenido de humedad 113
Figura 35 — Instrumento análisis granulométrico 114
Figura 36 — Instrumento límites de consistencia 115
Figura 37 — Instrumento ensayo de compactación 116
Figura 38 — Instrumento CBR 117
Figura 39 — Instrumento perfil estratigráfico 118
Figura 40 — Contenido de humedad C-01 119
Figura 41 — Análisis granulométrico C-01 120
Figura 42 — Clasificación de suelos C-01 121
Figura 43 — Limites de consistencia C-01 122
Figura 44 — Ensayo de compactación C-01 123
Figura 45 — Ensayo CBR C-01 124
Figura 46 — Ensayo CBR-Gráfico C-01 125
Figura 47 — Perfil estratigráfico C-01 126
Figura 48 — Contenido de humedad C-02 127
Figura 49 — Análisis granulométrico C-02 128
Figura 50 — Clasificación de suelos C-02 129
Figura 51 — Limites de consistencia C-02 130
Figura 52 — Ensayo de compactación C-02 131
Figura 53 — Ensayo CBR C-02 132
Figura 54 — Ensayo CBR-Gráfico C-02 133
Figura 55 — Perfil estratigráfico C-02 134
Figura 56 — Contenido de humedad C-03 135
Figura 57 — Análisis granulométrico C-03 136
Figura 58 — Clasificación de suelos C-03 137
Figura 59 — Limites de consistencia C-03 138
Figura 60 — Ensayo de compactación C-03 139
Figura 61 — Ensayo CBR C-03 140
Figura 62 — Ensayo CBR-Gráfico C-03 141
Figura 63 — Perfil estratigráfico C-03 142
Figura 64 — Exploración C-01 143
Figura 65 — Exploración C-02 143
Figura 66 — Exploración C-03 143
Figura 67 — Ensayo Análisis granulométrico 144
Figura 68 — Ensayo Análisis granulométrico C-01 144
Figura 69 — Ensayo Análisis granulométrico C-02 145
Figura 70 — Ensayo Análisis granulométrico C-03 145
Figura 71 — Ensayo de compactación 146

Repositorio Institucional - UNAMBA Perú

Figura 72 — Ensayo de compactación – Proctor modificado 146
Figura 73 — Ensayo de compactación – Moldes 146
Figura 74 — Ensayo CBR C-01 147
Figura 75 — Ensayo CBR C-02 147
Figura 76 — Ensayo CBR C-03 147
Figura 77 — Valoración del experto 01 149
Figura 78 — Valoración del experto 02 150
Figura 79 — Valoración del experto 03 151
Figura 80 — Valoración del Experto 04 152

Repositorio Institucional - UNAMBA Perú

SIGLAS Y ACRÓNIMOS

AASHTO — American Association of State Highway and Transportation Officials.

ASTM — ASTM International (anteriormente American Society for Testing and Materials).

CBR — California Bearing Ratio (Relación de Soporte de California).

CV-R² — Coeficiente de determinación mediante validación cruzada (Cross Validation).

LL — Límite Líquido.

LP — Límite Plástico.

DMS —Máxima Densidad Seca

MAE — Mean Absolute Error (Error Absoluto Medio).

ML — Machine Learning (Aprendizaje Automático).

Mr — Módulo Resiliente.

OMC — Optimum Moisture Content (Contenido Óptimo de Humedad).

R² — Coeficiente de Determinación.

RF — Random Forest (Bosque Aleatorio).

RMSE — Root Mean Squared Error (Raíz del Error Cuadrático Medio).

SUCS — Sistema Unificado de Clasificación de Suelos (Unified Soil Classification System).

Repositorio Institucional - UNAMBA Perú

INTRODUCCIÓN

La evaluación de la respuesta mecánica de los suelos constituye un aspecto fundamental en el diseño y construcción de pavimentos, debido a que la subrasante es la capa encargada de soportar las cargas transmitidas por el tránsito vehicular. Entre los parámetros más utilizados para caracterizar su comportamiento se encuentran el Índice de Soporte de California (CBR) y el módulo resiliente (Mr), los cuales permiten determinar la capacidad de soporte y la rigidez del suelo, siendo indispensables para el diseño estructural de pavimentos seguros y duraderos.

Tradicionalmente, la obtención de estos parámetros se realiza mediante ensayos de laboratorio que demandan tiempo, recursos económicos, equipos especializados y personal capacitado. En los últimos años, el desarrollo de la inteligencia artificial y, particularmente, del Machine Learning ha permitido el desarrollo de modelos predictivos capaces de estimar propiedades geotécnicas a partir de bases de datos previamente obtenidas, constituyéndose en una herramienta complementaria para apoyar la toma de decisiones en la ingeniería civil. Diversas investigaciones nacionales e internacionales han demostrado resultados favorables en la predicción del CBR y del módulo resiliente mediante algoritmos de aprendizaje automático, evidenciando su potencial aplicación en el ámbito geotécnico.

En este contexto, la presente investigación se desarrolló en los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, con el propósito de desarrollar un modelo predictivo basado en Machine Learning para estimar su respuesta mecánica. Para ello, se emplearon como variables de estudio las propiedades físicas y mecánicas del suelo obtenidas mediante ensayos geotécnicos, considerando como indicadores de la respuesta mecánica el CBR y el módulo resiliente.

Metodológicamente, la investigación corresponde a un enfoque cuantitativo, de tipo aplicada, nivel correlacional con alcance predictivo y diseño no experimental, transversal y retrospectivo. El desarrollo del modelo se realizó utilizando una base de datos conformada por 90 registros geotécnicos, mientras que los resultados obtenidos en tres calicatas fueron empleados para contrastar el desempeño del modelo. Asimismo, se evaluaron los algoritmos Gradient Boosting,

Repositorio Institucional - UNAMBA Perú

Extra Trees, Random Forest e HistGradientBoosting, comparando su desempeño mediante las métricas MAE, RMSE, coeficiente de determinación (R^2) y validación cruzada.

Los resultados obtenidos evidenciaron que el algoritmo Gradient Boosting presentó el mejor desempeño para estimar la respuesta mecánica de los suelos evaluados, alcanzando una alta correlación entre los valores experimentales y los estimados por el modelo. Estos hallazgos demuestran que el Machine Learning constituye una alternativa técnicamente viable como herramienta de apoyo para la estimación de parámetros geotécnicos, sin reemplazar los ensayos de laboratorio establecidos por la normativa vigente.

El presente trabajo tiene como alcance el desarrollo y evaluación de un modelo predictivo aplicado a las condiciones geotécnicas de la vía de evitamiento de Abancay, utilizando la información disponible para la investigación. Entre sus principales limitaciones se encuentra el tamaño de la base de datos empleada y la validación del modelo únicamente para las condiciones del área de estudio, por lo que futuras investigaciones podrán ampliar la base de datos y evaluar el comportamiento del modelo en otras condiciones geológicas y geotécnicas.

Finalmente, la tesis se encuentra estructurada en seis capítulos. El Capítulo I desarrolla el planteamiento del problema, la justificación y la formulación de los objetivos; el Capítulo II comprende las hipótesis y la operacionalización de variables; el Capítulo III presenta los antecedentes, el marco teórico y el marco conceptual; el Capítulo IV describe la metodología empleada; el Capítulo V expone los resultados, la contrastación de hipótesis y la discusión; y el Capítulo VI contiene las conclusiones y recomendaciones derivadas de la investigación. Asimismo, se incluyen las referencias bibliográficas y los anexos que respaldan el desarrollo del estudio.

Repositorio Institucional - UNAMBA Perú

RESUMEN

La evaluación de la respuesta mecánica de los suelos de subrasante constituye una etapa fundamental en el diseño estructural de pavimentos, debido a que parámetros como el Índice de Soporte de California (CBR) y el módulo resiliente (Mr) determinan la capacidad del terreno para soportar las cargas repetidas del tránsito vehicular. Sin embargo, su determinación mediante ensayos de laboratorio demanda tiempo, recursos y equipamiento especializado, lo que limita su aplicación en proyectos viales. En ese contexto, el objetivo de la presente investigación fue desarrollar un modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

La investigación tuvo un enfoque cuantitativo, fue de tipo aplicada, de nivel correlacional con alcance predictivo y de diseño no experimental, transversal y retrospectivo. Se utilizó una base de datos conformada por 90 registros geotécnicos, empleándose tres muestras representativas para la validación experimental. Se evaluaron los algoritmos Gradient Boosting, Extra Trees, Random Forest e HistGradientBoosting, comparando su desempeño mediante las métricas MAE, RMSE y R^2, y la contrastación de la hipótesis se realizó mediante el coeficiente de correlación de Spearman.

Los resultados evidenciaron que el modelo desarrollado alcanzó una correlación de Spearman de $\rho = 0.9562$ ($p < 0.05$) entre los valores experimentales y los estimados, demostrando una elevada capacidad predictiva para estimar el comportamiento mecánico de los suelos estudiados. Se concluye que el modelo predictivo desarrollado constituye una herramienta confiable para estimar la respuesta mecánica de los suelos de subrasante, contribuyendo a optimizar la evaluación geotécnica y apoyar la toma de decisiones en el diseño de pavimentos.

Palabras clave: *Machine Learning, respuesta mecánica, subrasante, CBR, módulo resiliente.*

Repositorio Institucional - UNAMBA Perú

ABSTRACT

The evaluation of the mechanical response of subgrade soils is a fundamental stage in pavement structural design, since parameters such as the California Bearing Ratio (CBR) and the resilient modulus (Mr) determine the soil's capacity to withstand repetitive traffic loads. However, their determination through laboratory testing requires considerable time, resources, and specialized equipment, which limits their application in road infrastructure projects. In this context, the objective of this research was to develop a Machine Learning-based predictive model to estimate the mechanical response of subgrade soils along the Abancay bypass road, Apurímac – 2025.

The research adopted a quantitative approach and was classified as applied research with a correlational level and predictive scope, using a non-experimental, cross-sectional, and retrospective design. A database consisting of 90 geotechnical records was used, with three representative samples employed for experimental validation. The Gradient Boosting, Extra Trees, Random Forest, and HistGradientBoosting algorithms were evaluated by comparing their performance using the MAE, RMSE, and R^2 metrics, while hypothesis testing was conducted using Spearman's rank correlation coefficient.

The results showed that the developed model achieved a Spearman correlation coefficient of $\rho = 0.9562$ ($p < 0.05$) between the experimental and predicted values, demonstrating a high predictive capability for estimating the mechanical behavior of the studied soils. It is concluded that the developed predictive model constitutes a reliable tool for estimating the mechanical response of subgrade soils, contributing to the optimization of geotechnical evaluation and supporting decision-making in pavement design.

Keywords: *Machine Learning, mechanical response, subgrade, California Bearing Ratio (CBR), resilient modulus.*

Repositorio Institucional - UNAMBA Perú

CAPÍTULO I
PLANTEAMIENTO DEL PROBLEMA

1.1 Descripción del problema

El diseño y construcción de infraestructura vial requiere una adecuada evaluación de las propiedades mecánicas del suelo que conforma la subrasante, ya que esta capa constituye el soporte fundamental de las estructuras de pavimento. La respuesta mecánica del suelo, representada principalmente por parámetros como la capacidad de soporte (CBR) y el módulo resiliente (Mr), permite estimar la capacidad del terreno para resistir cargas repetidas provenientes del tránsito vehicular (Giordani y Leone, 2023). Una estimación adecuada de estos parámetros resulta esencial para garantizar el desempeño estructural y la durabilidad de los pavimentos.

Tradicionalmente, la determinación de estas propiedades se realiza mediante ensayos de laboratorio y campo, los cuales requieren tiempo, equipamiento especializado y condiciones controladas para obtener resultados confiables. Entre los ensayos más utilizados se encuentra el ensayo de CBR (AASHTO, 1997), ampliamente aplicado para evaluar la capacidad de soporte de suelos en proyectos viales, así como el módulo resiliente, parámetro fundamental en metodologías modernas de diseño mecanicista de pavimentos (AASHTO, 2008). Sin embargo, la obtención de estos parámetros implica procesos experimentales que pueden resultar costosos y limitados en contextos donde no se dispone de suficiente infraestructura de laboratorio o recursos técnicos.

En los últimos años, el avance de las tecnologías de análisis de datos y de la inteligencia artificial ha permitido el desarrollo de modelos predictivos basados en técnicas de aprendizaje automático (Marathe y Sheshadri, 2025), los cuales permiten estimar propiedades geotécnicas a partir de conjuntos de datos previamente analizados. Estas metodologías han demostrado ser herramientas prometedoras en el campo de la ingeniería geotécnica, ya que permiten identificar patrones complejos entre variables y generar predicciones con un alto nivel de precisión (Marathe y Sheshadri, 2025).

Repositorio Institucional - UNAMBA Perú

Diversas investigaciones han demostrado que los modelos basados en aprendizaje automático pueden utilizarse para estimar parámetros geotécnicos como el CBR, el módulo resiliente y otros indicadores mecánicos del suelo, utilizando variables como la granulometría, el contenido de humedad, la densidad seca y los límites de consistencia (Ahmad et al., 2025). Estas herramientas permiten optimizar el proceso de análisis de datos y constituyen un complemento importante para los métodos tradicionales de caracterización del suelo.

A pesar de los avances en la aplicación de técnicas de Machine Learning en ingeniería geotécnica, su uso aún es limitado en contextos locales como el departamento de Apurímac, donde los estudios de suelos continúan basándose principalmente en métodos convencionales. En el caso del distrito de Abancay, las características geotécnicas de los suelos presentan variaciones asociadas a factores geológicos, climáticos y geomorfológicos propios de la región andina, lo que hace necesario contar con herramientas que permitan mejorar la estimación de sus propiedades mecánicas.

En este contexto, surge la necesidad de explorar nuevas metodologías que permitan aprovechar la información disponible y desarrollar modelos predictivos capaces de estimar la respuesta mecánica del suelo a nivel de subrasante. La aplicación de técnicas de Machine Learning podría contribuir a mejorar la eficiencia en la estimación de parámetros geotécnicos, reducir tiempos de análisis y fortalecer la toma de decisiones en proyectos de infraestructura vial.

Por lo tanto, la presente investigación plantea el desarrollo de un modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos a nivel de subrasante, utilizando información experimental proveniente de ensayos geotécnicos y técnicas de análisis de datos que permitan evaluar la capacidad de estos modelos para reproducir el comportamiento mecánico del suelo.

En atención a la problemática descrita y al contexto de estudio considerado en la presente investigación, se formula la siguiente pregunta de investigación:

Repositorio Institucional - UNAMBA Perú

1.2 Enunciado del problema

1.2.1 Problema general

¿De qué manera el modelo predictivo basado en Machine Learning permite estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025?

1.2.2 Problemas específicos

- ¿De qué manera el modelo predictivo basado en Machine Learning permite estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025?

- ¿De qué manera el modelo predictivo basado en Machine Learning permite estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025?

1.3 Justificación de la investigación

La presente investigación se justifica por la necesidad de contar con alternativas que contribuyan a optimizar la estimación de la respuesta mecánica de los suelos a nivel de subrasante, debido a que parámetros como la capacidad de soporte del suelo (CBR) y el módulo resiliente (Mr) constituyen variables fundamentales para el análisis geotécnico y el diseño de pavimentos. La determinación de estos parámetros generalmente requiere la ejecución de ensayos de laboratorio especializados, así como la disponibilidad de recursos técnicos y tiempo para el procesamiento e interpretación de los resultados. En ese contexto, resulta pertinente evaluar el uso de herramientas basadas en Machine Learning que permitan estimar dichas propiedades a partir de información geotécnica disponible, contribuyendo a mejorar los procesos de análisis y apoyo a la toma de decisiones en proyectos de infraestructura vial.

Asimismo, la investigación adquiere relevancia debido a la creciente incorporación de técnicas de inteligencia artificial en diferentes áreas de la ingeniería civil, particularmente en la geotecnia y el diseño de pavimentos. El desarrollo de modelos predictivos representa una oportunidad para complementar los métodos convencionales de caracterización de suelos, permitiendo aprovechar bases de datos existentes para generar estimaciones confiables de parámetros mecánicos asociados al comportamiento de la subrasante.

Repositorio Institucional - UNAMBA Perú

La originalidad del estudio radica en la aplicación de algoritmos de Machine Learning para la estimación de la respuesta mecánica de los suelos de la vía de evitamiento de la ciudad de Abancay, ámbito en el que existe limitada evidencia de investigaciones orientadas a la predicción de la capacidad de soporte del suelo y del módulo resiliente mediante técnicas de aprendizaje automático. En ese sentido, la investigación incorpora herramientas tecnológicas emergentes en un contexto geotécnico local, contribuyendo a ampliar el conocimiento disponible sobre su aplicación en condiciones similares.

Finalmente, el aporte técnico-científico de la investigación consiste en el desarrollo y evaluación de un modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos a nivel de subrasante. Los resultados obtenidos permitirán generar una referencia metodológica para futuras investigaciones relacionadas con la aplicación de inteligencia artificial en estudios geotécnicos y diseño de pavimentos, fortaleciendo la integración entre las técnicas de análisis de datos y la ingeniería civil.

Repositorio Institucional - UNAMBA Perú

CAPÍTULO II
OBJETIVOS E HIPÓTESIS

2.1 Objetivos de la investigación

2.1.1 Objetivo general

Desarrollar un modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

2.1.2 Objetivos específicos

- Determinar el desempeño del modelo predictivo basado en Machine Learning para estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

- Determinar el desempeño del modelo predictivo basado en Machine Learning para estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

2.2 Hipótesis de la investigación

2.2.1 Hipótesis general

El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

2.2.2 Hipótesis específicas

- El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

- El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

Repositorio Institucional - UNAMBA Perú

2.3 Operacionalización de variables

Tabla 1 — Operacionalización de variables

Variable	Definición conceptual	Definición operacional	Dimensiones	Indicadores	Item/und	Instrumentos
Independiente: **X:** Modelo predictivo basado en Machine Learning	Un modelo predictivo basado en M*achine Learning* emplea algoritmos que aprenden a partir de datos para generar estimaciones de variables objetivo; su evaluación considera tanto el algoritmo utilizado como el desempeño del modelo de Machine Leaning (Goodfellow, Bengio & Courville, 2016; Bishop, 2006).	Se implementa un modelo de Machine Learning utilizando datos del estudio; se registra el algoritmo empleado y se evalúa su desempeño mediante métricas estadísticas obtenidas durante la validación del modelo.	X1. Desempeño del modelo predictivo	– Error Absoluto Medio (MAE) – Coeficiente de determinación (R^2)	Valor numérico R^2 %	Ficha de evaluación del modelo predictivo
Dependiente: **Y:** Respuesta mecánica	La respuesta mecánica del suelo comprende el conjunto de propiedades que describen su comportamiento frente a la aplicación de cargas, entre ellas la **capacidad de soporte del suelo** y la **rigidez de la subrasante**, las cuales son fundamentales para el análisis y diseño geotécnico y de pavimentos (Das & Sobhan, 2018; Huang, 2004).	Se determinan valores representativos de la capacidad de soporte y de la rigidez de la subrasante a partir de ensayos de laboratorio o campo; dichos valores constituyen las variables de salida que serán estimadas por el modelo predictivo.	Y1. Capacidad de soporte del suelo	– Índice de CBR	%	Ficha de registro de capacidad de soporte del suelo
			Y2. Rigidez de la subrasante.	– Módulo resiliente (Mr)	MPa	Ficha de registro de rigidez de la subrasante

CAPÍTULO III
MARCO TEÓRICO REFERENCIAL

3.1 Antecedentes

3.1.1 Antecedentes internacionales

a) Tado et al. (2025), en su investigación titulada *"Forecasting California bearing ratio (CBR) of soil using machine learning algorithms: A review "*, tuvo como objetivo analizar y sintetizar las investigaciones recientes relacionadas con la aplicación de algoritmos de Machine Learning para la predicción del California Bearing Ratio (CBR) en suelos naturales y estabilizados, identificando los modelos de mayor precisión, las variables geotécnicas más utilizadas y las tendencias actuales en la ingeniería geotécnica. La investigación se fundamentó en la aplicación de técnicas de inteligencia artificial como Artificial Neural Networks (ANN), Random Forest (RF), Support Vector Machine (SVM), Deep Neural Networks (DNN), Decision Tree (DT), Extreme Gradient Boosting (XGBoost), LightGBM, Genetic Expression Programming (GEP) y otros modelos híbridos, empleados para establecer relaciones entre las propiedades geotécnicas del suelo y el valor del CBR. La metodología consistió en una revisión sistemática de estudios científicos sobre la predicción del CBR mediante Machine Learning, comparando los algoritmos utilizados, las variables de entrada y los principales indicadores estadísticos de desempeño, como el coeficiente de determinación (R^2), el error cuadrático medio (RMSE) y el error absoluto medio (MAE). Los autores identificaron que las variables de entrada más utilizadas fueron el Contenido Óptimo de Humedad (OMC), la Máxima Densidad Seca (MDD), el Límite Líquido (LL), el Límite Plástico (PL) y el Índice de Plasticidad (PI), siendo OMC y MDD las más empleadas para estimar el CBR. Los resultados evidenciaron que las Redes Neuronales Artificiales (ANN) constituyen el algoritmo más utilizado y uno de los más precisos para la predicción del CBR, mientras que modelos como RF, DNN, RFR, AB-DT, GEP y ELM-CSO también alcanzaron un alto desempeño según el tipo de suelo y la base de datos utilizada. Finalmente, los autores concluyeron que el Machine Learning constituye una alternativa eficiente y confiable frente a los ensayos

Repositorio Institucional - UNAMBA Perú

convencionales para estimar el CBR, recomendando el desarrollo de modelos híbridos, el uso de bases de datos más amplias y la incorporación de técnicas de aprendizaje profundo para mejorar la capacidad predictiva.

b) Gao (2024) en su investigación titulada "The application of machine learning in geotechnical engineering" tuvo como objetivo explorar las aplicaciones de las técnicas de Machine Learning en la ingeniería geotécnica, centrándose especialmente en su potencial para mejorar la predicción de parámetros geotécnicos del suelo, tales como la resistencia al corte, la permeabilidad y la capacidad de soporte. Asimismo, el estudio buscó demostrar cómo los métodos de Machine Learning pueden optimizar la toma de decisiones en proyectos de ingeniería civil que involucren suelos complejos y comportamientos no lineales. La investigación se fundamentó en el uso de algoritmos de Machine Learning como Redes Neuronales Artificiales (ANN), Máquinas de Soporte Vectorial (SVM) y Regresión de Procesos Gaussianos (GPR), empleados para modelar y predecir el comportamiento de los suelos. La población estuvo conformada por datos obtenidos de ensayos geotécnicos de diversas muestras de suelo, mientras que la muestra comprendió 200 muestras provenientes de proyectos de infraestructura desarrollados en diferentes partes del mundo. La metodología consistió en la aplicación de algoritmos SVM, ANN y GPR para la predicción de parámetros críticos del suelo, utilizando como variables de entrada el contenido de humedad, la granulometría, la densidad seca, la plasticidad y el límite líquido. Para el entrenamiento de los modelos se empleó el 70% de los datos disponibles y el 30% restante fue destinado a la validación. El desempeño de los modelos fue evaluado mediante el coeficiente de determinación (R^2) y la raíz del error cuadrático medio (RMSE), comparándose además los resultados obtenidos mediante Machine Learning con los generados por métodos tradicionales. Los resultados mostraron que los modelos de Machine Learning superaron a los métodos convencionales en términos de precisión y eficiencia, destacando el modelo basado en Redes Neuronales Artificiales, el cual alcanzó un coeficiente de determinación de $R^2 = 0.94$ en la predicción de la resistencia al corte y la capacidad de soporte del suelo. Finalmente, el autor concluyó que los algoritmos de Machine Learning, especialmente las Redes Neuronales Artificiales y las Máquinas de Soporte Vectorial, constituyen herramientas eficaces para modelar suelos con características complejas, permitiendo obtener

Repositorio Institucional - UNAMBA Perú

predicciones más precisas y rápidas que los enfoques tradicionales, reduciendo costos y tiempos de análisis en proyectos de ingeniería civil.

c) Khatti y Singh Grover (2021), en la investigación titulada "Prediction of geotechnical properties of soil using artificial intelligence framework", tuvieron como objetivo predecir propiedades geotécnicas de los suelos, tales como el límite líquido (LL), límite plástico (PL), índice de plasticidad (PI), contenido óptimo de humedad (OMC) y densidad máxima seca (MDD), mediante la aplicación de técnicas de inteligencia artificial. Los autores emplearon métodos de aprendizaje automático como análisis de regresión (RA), máquinas de soporte vectorial (SVM), regresión de procesos gaussianos (GPR), redes neuronales artificiales (ANN) y máquinas de relevancia vectorial (RVM) para modelar la relación existente entre la composición granulométrica del suelo y sus propiedades geotécnicas. La investigación se sustentó en la premisa de que propiedades como la plasticidad y la capacidad de soporte dependen significativamente del contenido de partículas finas, principalmente limo y arcilla, por lo que la aplicación de algoritmos de inteligencia artificial permite superar las limitaciones de los métodos tradicionales de laboratorio. La población estuvo conformada por registros geotécnicos recopilados de diversas investigaciones previas, mientras que la muestra comprendió 166 muestras de suelo que incluían variables de entrada como grava, arena, limo y arcilla, así como variables de salida correspondientes a LL, PL, PI, OMC y MDD. La metodología consistió en el procesamiento de datos mediante MATLAB R2020a para los modelos SVM, GPR, ANN y RVM, y Microsoft Excel 2021 para los modelos de regresión. Asimismo, se realizó el entrenamiento y validación de los modelos predictivos, complementándose con un análisis de sensibilidad global de Monte Carlo para evaluar la influencia de las variables de entrada sobre las propiedades geotécnicas estudiadas. Los resultados demostraron que el modelo basado en regresión de procesos gaussianos (GPR) presentó el mejor desempeño predictivo, superando a los modelos SVM, ANN y RVM. En particular, la predicción del límite líquido alcanzó un coeficiente de determinación de R^2 = 0.9556, evidenciando una elevada concordancia entre los valores predichos y los obtenidos experimentalmente. Finalmente, los autores concluyeron que los métodos de inteligencia artificial, especialmente el modelo GPR, constituyen herramientas eficaces para predecir propiedades geotécnicas del suelo,

Repositorio Institucional - UNAMBA Perú

permitiendo mejorar la eficiencia de los análisis y reducir los costos asociados a los ensayos convencionales en proyectos de ingeniería civil.

d) Davies y Jánošík (2024), en su investigación titulada *" Enhanced Prediction of California Bearing Ratio (CBR) Values in Geotechnical Engineering Using Decision Tree Algorithm and Meta-Heuristic Optimizations "*, tuvieron como objetivo desarrollar un enfoque basado en Machine Learning para predecir con mayor precisión los valores del California Bearing Ratio (CBR) mediante el algoritmo Decision Tree (DT) optimizado con técnicas metaheurísticas, con la finalidad de mejorar la evaluación de la capacidad portante de los suelos utilizados en proyectos de ingeniería geotécnica. La investigación se fundamentó en el empleo del algoritmo Decision Tree, complementado con los algoritmos de optimización Northern Goshawk Optimization (NGO) y Ebola Optimization Search (EOS) para incrementar la precisión de las predicciones. La muestra estuvo conformada por 109 registros experimentales, considerando como variables de entrada el límite líquido (LL), límite plástico (PL), índice de plasticidad (PI), máxima densidad seca (MDD), contenido óptimo de humedad (OMC), porcentaje de ceniza de aserrín (SDA), polvo de cantera (QD) y cemento Portland ordinario (OPC), siendo el CBR la variable de salida. La metodología consistió en dividir la base de datos en 70 % para entrenamiento, 15 % para validación y 15 % para prueba, desarrollando los modelos DT, DTNG y DTEO, evaluados mediante el coeficiente de determinación (R^2), la raíz del error cuadrático medio (RMSE), el error absoluto medio (MAE), el error absoluto relativo medio (MARE) y el índice de desempeño (PI). Los resultados demostraron que el modelo DTEO obtuvo el mejor desempeño predictivo, alcanzando un $R^2 = 0.996$ y un RMSE = 0.732 durante el entrenamiento; asimismo, en el conjunto total de datos registró un $R^2 = 0.993$ y un RMSE = 0.987, superando a los modelos DT y DTNG. Finalmente, los autores concluyeron que la integración del algoritmo Decision Tree con técnicas metaheurísticas mejora la precisión en la predicción del CBR, constituyendo una herramienta confiable para apoyar la caracterización geotécnica y el diseño de obras de infraestructura.

e) Głuchowski (2024), en su investigación titulada *" Prediction of Resilient Modulus Value of Cohesive and Non-Cohesive Soils Using Artificial Neural*

Repositorio Institucional - UNAMBA Perú

Network ", tuvo como objetivo desarrollar un modelo basado en Redes Neuronales Artificiales (ANN) para predecir el módulo resiliente (Mr) de suelos cohesivos y no cohesivos destinados a capas de subrasante y subbase, con la finalidad de proporcionar una alternativa confiable a los ensayos convencionales utilizados en el diseño de pavimentos. La investigación se fundamentó en el empleo de las Redes Neuronales Artificiales para modelar las relaciones no lineales existentes entre las propiedades físicas, las condiciones de esfuerzo y el módulo resiliente del suelo, incorporando además el análisis SHAP (SHapley Additive Explanations) para interpretar la influencia de cada variable en las predicciones del modelo. La muestra estuvo conformada por 1683 observaciones del módulo resiliente recopiladas de diversas investigaciones publicadas, considerando como variables de entrada el porcentaje que pasa por los tamices N.° 40 y N.° 200, índice de plasticidad (PI), peso unitario seco (DUW), relación de vacíos (e), grado de saturación (Sr), esfuerzo desviador (σd) y presión de confinamiento (σ'_3), mientras que el módulo resiliente (Mr) fue considerado como variable de salida. La metodología consistió en desarrollar una arquitectura de Red Neuronal Artificial multicapa (MLP) con ocho variables de entrada, dos capas ocultas y una neurona de salida, utilizando el algoritmo de optimización quasi-Newton y la función de activación tangente hiperbólica. Para el entrenamiento del modelo, la base de datos se dividió en 60 % para entrenamiento, 20 % para selección y 20 % para prueba, evaluando su desempeño mediante el coeficiente de determinación (R^2), el error cuadrático medio (MSE), la raíz del error cuadrático medio (RMSE) y el error cuadrático normalizado (NSE). Los resultados evidenciaron un elevado desempeño predictivo, alcanzando un $R^2 = 0.9613$, además de bajos valores de error en los conjuntos de selección y prueba, demostrando una adecuada capacidad de generalización del modelo. Finalmente, el autor concluyó que las Redes Neuronales Artificiales constituyen una herramienta eficaz para estimar el módulo resiliente de suelos de subrasante y subbase, permitiendo obtener predicciones precisas que pueden contribuir al diseño mecanístico-empírico de pavimentos y reducir la dependencia de los ensayos tradicionales de laboratorio.

f) Gowda et al. (2024), en su investigación titulada *"Prediction of the Subgrade Soil California Bearing Ratio Using Machine Learning and Neuro-Fuzzy Inference System Techniques: A Sustainable Approach in Urban Infrastructure*

Repositorio Institucional - UNAMBA Perú

Development ", tuvo como objetivo desarrollar modelos predictivos basados en Machine Learning y en un Sistema de Inferencia Neuro-Difusa Adaptativa (ANFIS) para estimar el California Bearing Ratio (CBR) de suelos de subrasante, con la finalidad de mejorar la precisión de las predicciones frente a los ensayos convencionales. La investigación se fundamentó en la aplicación de los algoritmos Regresión Lineal Multivariable (MLR), Redes Neuronales Artificiales (ANN) y Adaptive Neuro-Fuzzy Inference System (ANFIS) para modelar la relación entre las propiedades del suelo y el CBR. La muestra estuvo conformada por 2191 muestras de suelo de subrasante provenientes del estado de Andhra Pradesh (India), considerando como variables de entrada el tipo de suelo, el índice de plasticidad (PI) y la máxima densidad seca (MDD), mientras que el CBR fue la variable de salida. La metodología consistió en desarrollar y comparar los modelos MLR, ANN y ANFIS, utilizando 70 % de los datos para entrenamiento, 15 % para prueba y 15 % para validación, evaluando su desempeño mediante el coeficiente de determinación (R^2) y la raíz del error cuadrático medio (RMSE). Los resultados evidenciaron que el modelo ANFIS presentó el mejor desempeño predictivo, alcanzando valores de R^2 de 0.81, 0.82 y 0.82, con RMSE de 2.26, 2.29 y 2.23 para las etapas de entrenamiento, prueba y validación, respectivamente, superando a los modelos ANN y MLR. Asimismo, el análisis de sensibilidad determinó que el índice de plasticidad (PI) fue la variable de mayor influencia sobre el CBR (46 %), seguido por el tipo de suelo (34 %) y la máxima densidad seca (20 %). Finalmente, los autores concluyeron que el modelo ANFIS constituye una herramienta precisa y confiable para la predicción del CBR de suelos de subrasante, favoreciendo el diseño de pavimentos y la planificación de infraestructura vial

3.1.2 Antecedentes nacionales

a) Matto pablo (2025), en su investigación titulada “Modelo de inteligencia artificial con fines de predicción del CBR en obras viales en la región Huánuco, 2024”, tuvo como objetivo determinar qué modelo de inteligencia artificial permite la predicción del California Bearing Ratio (CBR) en obras viales de la región Huánuco. La investigación se fundamentó en la aplicación de técnicas de Machine Learning, específicamente Redes Neuronales Artificiales (ANN), LSTM, GRU, Random Forest (RF) y Support Vector Machine (SVM) para establecer relaciones entre las propiedades físicas, mecánicas y

Repositorio Institucional - UNAMBA Perú

litoestratigráficas del suelo con los valores del CBR. La población estuvo conformada por suelos de subrasante de obras viales de la macro región centro, mientras que la muestra comprendió 2270 registros de caracterización geotécnica, de los cuales 1365 contenían información completa. La metodología consistió en desarrollar modelos predictivos mediante Python, realizando la limpieza y estandarización de datos, así como la división de la base de datos en 80 % para entrenamiento, 10 % para validación y 10 % para prueba, evaluando su desempeño mediante R^2, MAE, MSE y RMSE. Los resultados evidenciaron que el modelo ANN obtuvo el mejor desempeño, alcanzando un $R^2 = 0,9848$, MAE = 0,8195, MSE = 1,9198 y RMSE = 1,3856, identificándose además que la máxima densidad seca, el contenido óptimo de humedad y el porcentaje de arena fueron las variables más influyentes. Finalmente, el autor concluyó que las Redes Neuronales Artificiales constituyen el modelo más preciso para predecir el CBR a partir de las propiedades físicas y mecánicas de los suelos de subrasante.

b) Moreno (2024), en su trabajo de investigación "Redes neuronales artificiales para la determinación antelada de estabilidad de taludes en carreteras Junín 2022", tuvo como objetivo predecir la estabilidad de los taludes ubicados en la carretera Santa Rosa de Ocopa – Comas, en la región Junín, mediante la aplicación de Redes Neuronales Artificiales (RNA). La investigación buscó estimar parámetros geotécnicos como el factor de seguridad, la cohesión y el ángulo de fricción a partir de propiedades físicas del suelo. La base teórica se sustentó en la aplicación de Redes Neuronales Artificiales como modelos capaces de representar relaciones complejas entre variables geotécnicas y predecir parámetros del suelo. La población estuvo conformada por los taludes de las carreteras de la región Junín, específicamente aquellos localizados en la carretera Santa Rosa de Ocopa – Comas, mientras que la muestra correspondió al talud ubicado en el km 2+500, del cual se obtuvieron los datos utilizados para el entrenamiento y validación del modelo neuronal. La metodología consistió en el entrenamiento de Redes Neuronales Artificiales empleando variables de entrada como contenido de humedad, granulometría y límites de Atterberg. Para el desarrollo del modelo se utilizó el algoritmo Bayesian Regularization, destinándose el 80% de los datos al entrenamiento y el 20% a la validación. Asimismo, el desempeño del modelo fue evaluado mediante el error cuadrático

Repositorio Institucional - UNAMBA Perú

medio (MSE) y coeficientes de correlación entre los valores predichos y los resultados experimentales. Los resultados mostraron una precisión de 99.6% en la predicción del factor de seguridad de los taludes. De igual manera, se obtuvieron coeficientes de correlación de Pearson de 0.985 para la cohesión, 0.957 para el ángulo de fricción y 0.993 para el factor de seguridad, evidenciando una alta concordancia entre los resultados estimados y los obtenidos mediante ensayos geotécnicos. Finalmente, el autor concluyó que las Redes Neuronales Artificiales permiten predecir con alta precisión los parámetros asociados a la estabilidad de taludes, alcanzando correlaciones superiores a 0.95 y constituyéndose en una herramienta confiable para anticipar el comportamiento geotécnico de taludes en proyectos de infraestructura vial.

c) Callupe Navarro y Gago Morales (2023), en su investigación titulada "Evaluación del CBR y la deflexión del terreno natural usando redes neuronales en el mejoramiento de la carretera Oyón–Ambo tramo II, provincia Daniel Alcides Carrión – Pasco", tuvo como objetivo desarrollar un modelo predictivo basado en Redes Neuronales Artificiales (ANN) para estimar el valor del CBR y la deflexión del terreno natural a partir de las propiedades físicas del suelo, con la finalidad de optimizar la evaluación geotécnica en proyectos de pavimentación. La investigación se sustentó en la aplicación de Redes Neuronales Artificiales como herramientas capaces de modelar relaciones no lineales entre las propiedades físicas y los parámetros mecánicos del suelo, utilizando como variables de entrada el porcentaje de grava, porcentaje de finos, humedad natural, contenido óptimo de humedad, índice de plasticidad y máxima densidad seca. La población estuvo conformada por el terreno natural del proyecto de mejoramiento de la carretera Oyón–Ambo, mientras que la muestra comprendió el tramo ubicado entre las progresivas km 203+000 y km 206+000, empleándose 30 muestras de suelo obtenidas mediante calicatas y 30 registros de deflexión medidos en campo. La metodología correspondió a una investigación aplicada, de diseño no experimental y enfoque cuantitativo, donde los datos fueron divididos en 75 % para entrenamiento y 25 % para validación del modelo, evaluando su desempeño mediante el coeficiente de determinación (R^2). Los resultados evidenciaron que la máxima densidad seca fue la variable con mayor influencia en la predicción del CBR, alcanzando una participación del 77,9 % y un $R^2 = 0{,}835$; mientras que las propiedades físicas del suelo

Repositorio Institucional - UNAMBA Perú

presentaron una baja capacidad para predecir la deflexión, obteniéndose un R^2 = 0,024. Finalmente, los autores concluyeron que las Redes Neuronales Artificiales constituyen una herramienta eficiente para la estimación del CBR; sin embargo, la predicción de la deflexión requiere incorporar variables adicionales que permitan mejorar el desempeño del modelo.

d) Purizaca, et al., (2021), en su investigación titulada "Desarrollo de un modelo predictivo de las propiedades mecánicas del suelo usando redes neuronales artificiales", tuvo como objetivo desarrollar modelos predictivos basados en Redes Neuronales Artificiales (RNA) para estimar la máxima densidad seca (MDD), el contenido óptimo de humedad (OMC), el valor del California Bearing Ratio (CBR) al 95 % de la MDD y el CBR al 100 % de la MDD, con la finalidad de reducir el tiempo y costo de los ensayos de laboratorio requeridos en el diseño de pavimentos. La investigación se fundamentó en la aplicación de Redes Neuronales Artificiales tipo perceptrón multicapa como herramientas capaces de modelar relaciones no lineales entre las propiedades físicas y mecánicas del suelo, empleando como variables de entrada el contenido de humedad, contenido de sales, granulometría y límites de consistencia. La población estuvo conformada por 285 muestras de suelo provenientes de expedientes técnicos de obras de carreteras ejecutadas en las provincias de Chiclayo, Ferreñafe y Lambayeque, obtenidas del Sistema Electrónico de Contrataciones del Estado (SEACE), utilizándose la totalidad de los registros para el desarrollo de los modelos predictivos. La metodología consistió en el entrenamiento de cuatro modelos de Redes Neuronales Artificiales utilizando optimización bayesiana y validación cruzada para definir la arquitectura e hiperparámetros, destinando el 80 % de los datos al entrenamiento y el 20 % a la prueba, evaluando el desempeño mediante el coeficiente de correlación (R) y el error cuadrático medio (MSE). Los resultados evidenciaron coeficientes de correlación de R = 0,90 para la predicción de la máxima densidad seca, R = 0,87 para el contenido óptimo de humedad, R = 0,92 para el CBR al 95 % de la MDD y R = 0,89 para el CBR al 100 % de la MDD, demostrando un adecuado desempeño de los modelos desarrollados. Finalmente, los autores concluyeron que las Redes Neuronales Artificiales constituyen una herramienta eficiente para predecir las propiedades mecánicas del suelo, permitiendo obtener resultados confiables en

Repositorio Institucional - UNAMBA Perú

menor tiempo y con menores costos respecto a los procedimientos convencionales de laboratorio.

e) Castillo Delgado (2021), en su investigación titulada "Redes neuronales artificiales para la predicción de propiedades mecánicas de los suelos", tuvo como objetivo predecir las propiedades mecánicas de los suelos mediante un software basado en algoritmos de Redes Neuronales Artificiales, con la finalidad de reducir el tiempo y costo de los estudios geotécnicos en proyectos viales. La base teórica se sustentó en la aplicación de Redes Neuronales Artificiales de tipo GRNN y MLFN, como herramientas capaces de modelar relaciones no lineales entre las propiedades físicas y mecánicas del suelo. La población estuvo conformada por registros de ensayos de mecánica de suelos provenientes de proyectos viales, mientras que la muestra estuvo integrada por 289 registros recopilados de Provías Nacional, correspondientes a granulometría, límites de Atterberg, máxima densidad seca, óptimo contenido de humedad y CBR. La metodología correspondió a un enfoque cuantitativo, de tipo aplicada, nivel correlacional y diseño no experimental de corte transversal. Se desarrollaron modelos neuronales utilizando como variables de entrada el porcentaje de grava, arena, finos, límite líquido, límite plástico e índice de plasticidad, y como variables de salida la máxima densidad seca, el óptimo contenido de humedad y los valores de CBR al 95 % y 100 % de la máxima densidad seca. El desempeño de los modelos fue evaluado mediante el coeficiente de determinación (R^2), el coeficiente de correlación de Pearson y el error cuadrático medio (RMS). Los resultados evidenciaron coeficientes de determinación de 0.7487 para la máxima densidad seca, 0.7771 para el óptimo contenido de humedad, 0.7890 para el CBR al 95 % y 0.8110 para el CBR al 100 % durante el entrenamiento; en la validación se obtuvieron valores de 0.7376, 0.7815, 0.7935 y 0.8232, respectivamente. Finalmente, los autores concluyeron que las Redes Neuronales Artificiales constituyen una herramienta confiable para predecir las propiedades mecánicas del suelo, obteniendo un buen nivel de ajuste y reduciendo el tiempo requerido para la estimación de dichos parámetros en proyectos de ingeniería civil.

f) Valderrama (2020), en su trabajo de investigación "Desarrollo de un modelo de redes neuronales artificiales para predecir las propiedades mecánicas del suelo", tuvo como objetivo desarrollar modelos de Redes Neuronales Artificiales

Repositorio Institucional - UNAMBA Perú

(ANN) para predecir propiedades geotécnicas del suelo en la región de Lambayeque, tales como la máxima densidad seca (MDD), el contenido óptimo de humedad (OMC) y la Relación de Soporte de California (CBR), con la finalidad de obtener resultados rápidos y precisos para su aplicación en proyectos de infraestructura vial. La investigación se fundamentó en el uso de Redes Neuronales Artificiales como herramientas capaces de modelar la relación existente entre las propiedades físicas y mecánicas del suelo, superando algunas limitaciones de los métodos tradicionales mediante procesos de predicción más eficientes y precisos. La población estuvo conformada por 285 muestras de suelo provenientes de proyectos de pavimentación ejecutados en la región Lambayeque, cuyos registros fueron obtenidos de expedientes técnicos descargados del Sistema Electrónico de Contrataciones del Estado (SEACE). Asimismo, la totalidad de las muestras fue utilizada para el entrenamiento y evaluación de los modelos desarrollados. La metodología consistió en el entrenamiento de modelos de Redes Neuronales Artificiales utilizando variables de entrada relacionadas con el contenido de humedad, análisis granulométrico, contenido de sales y límites de consistencia, mientras que las variables de salida estuvieron representadas por parámetros geotécnicos de interés para el diseño de pavimentos. El desempeño de los modelos fue evaluado mediante indicadores estadísticos como el error cuadrático medio (MSE) y el error absoluto medio (MAE), permitiendo medir la precisión de las predicciones obtenidas. Los resultados evidenciaron un comportamiento favorable de los modelos desarrollados, obteniéndose un error cuadrático medio de 0.003 g/cm³ para la predicción de la máxima densidad seca (MDD), 2.42% para el contenido óptimo de humedad (OMC), 6.35% para el CBR al 95% de la MDD y 21.80% para el CBR al 100% de la MDD. Finalmente, el autor concluyó que las Redes Neuronales Artificiales constituyen una herramienta eficaz para la predicción de propiedades geotécnicas del suelo, ofreciendo una alternativa rápida, precisa y económicamente viable frente a los procedimientos convencionales de laboratorio, pudiendo ser aplicadas en proyectos reales de infraestructura vial.

3.2 Marco teórico

3.2.1 Machine Learning

El aprendizaje automático (Machine Learning, ML) es un subcampo de la inteligencia artificial (IA) que permite a las computadoras aprender patrones a partir

Repositorio Institucional - UNAMBA Perú

de datos sin necesidad de ser programadas explícitamente. Se basa en el desarrollo de algoritmos capaces de identificar relaciones complejas dentro de los conjuntos de datos y realizar predicciones o clasificaciones en función de la información procesada.(Konasani y Kadre, 2021).

La principal ventaja del Machine Learning en comparación con los métodos tradicionales es su capacidad para manejar grandes volúmenes de datos y mejorar su precisión a medida que recibe más información. En la ingeniería geotécnica, ML se ha convertido en una herramienta clave para optimizar el análisis de suelos, reducir la necesidad de ensayos costosos y mejorar la toma de decisiones en proyectos de infraestructura.

El Machine Learning es una rama de la inteligencia artificial que permite a las computadoras aprender patrones a partir de datos y mejorar su rendimiento en tareas específicas sin ser programadas explícitamente. A través de algoritmos matemáticos y estadísticos, el sistema analiza grandes volúmenes de información, identifica relaciones y genera predicciones o clasificaciones de manera autónoma.

(Dark, 2022) define Machine Learning como el proceso mediante el cual un sistema mejora su desempeño en una tarea con base en la experiencia y una métrica de rendimiento. Es decir, mientras más datos analiza el modelo, mayor es su precisión. Existen tres tipos principales de Machine Learning: supervisado, donde se entrenan modelos con datos etiquetados; no supervisado, que encuentra patrones sin resultados predefinidos; y por refuerzo, donde el algoritmo aprende por ensayo y error (DARK, 2022).

(Konasani y Kadre, 2021). La inteligencia artificial (AI) ha evolucionado rápidamente en los últimos 25 años, transformando la manera en que las máquinas perciben, razonan y actúan. Según Winston (1992), AI es el estudio de los cálculos que permiten estas capacidades. Desde los primeros humanoides que realizaban tareas básicas hasta robots modernos capaces de ejecutar acrobacias complejas, AI ha avanzado significativamente. Sophia, un robot humanoide, incluso recibió ciudadanía saudí y ha participado en entrevistas y conferencias globales sin intervención humana.

Repositorio Institucional - UNAMBA Perú

Sin embargo, la mayoría de los robots industriales y domésticos no tienen forma humana, sino que su diseño responde a sus funciones específicas. En todos ellos, AI y el aprendizaje automático son fundamentales, con algoritmos avanzados como el deep learning impulsando su desarrollo. AI es un campo multidisciplinario que integra filosofía, matemáticas, neurociencia, psicología cognitiva y cibernética, buscando replicar la inteligencia humana. Aunque aún estamos lejos de lograrlo por completo, la investigación sigue avanzando.

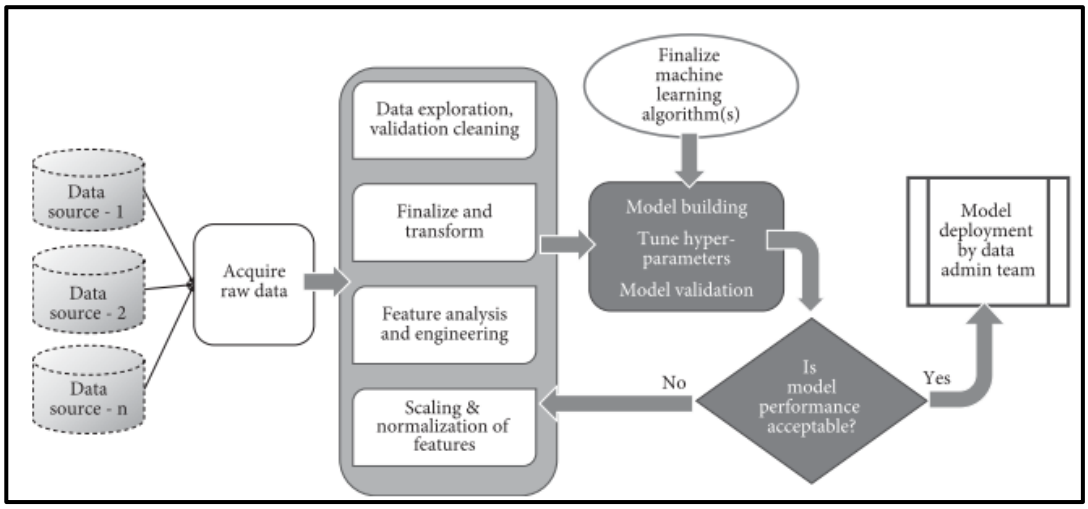
3.2.2 Bloques de construcción de un proyecto de aprendizaje

El machine Learning (ML), también conocido como análisis aumentado, es una rama de la inteligencia artificial (AI) estrechamente vinculada a la estadística computacional. Su función principal es desarrollar algoritmos que aprenden de la experiencia para mejorar la precisión en la realización de tareas específicas. Para ello, estos modelos deben ser entrenados con grandes volúmenes de datos antes de ofrecer predicciones confiables y valiosas. (Konasani y Kadre, 2021)

Aunque comparte técnicas con la estadística, su objetivo es diferente: mientras que la estadística descriptiva se enfoca en analizar datos para inferencias poblacionales, el Machine Learning busca patrones predictivos. Su implementación abarca diversas industrias, optimizando procesos empresariales y operacionales. Antes de hacer predicciones, es necesario realizar un trabajo extenso de recopilación, limpieza y transformación de datos, lo que representa hasta el 70% del esfuerzo en un proyecto de Machine Learning. (Konasani y Kadre, 2021).

(Konasani y Kadre, 2021) . El desarrollo de un modelo de ML requiere equipos multidisciplinarios, incluyendo analistas de negocios, arquitectos de soluciones, científicos de datos y administradores de bases de datos. Una vez entrenado y evaluado, el modelo se implementa en producción para ser utilizado en entornos reales. Grandes empresas, como bancos y plataformas de comercio electrónico, dependen de estos sistemas para mejorar sus operaciones diarias.

Repositorio Institucional - UNAMBA Perú



FUENTE: Konasani y Kadre, 2021, p. 26.

Figura 1 — Un proyecto típico de aprendizaje automático en la práctica real

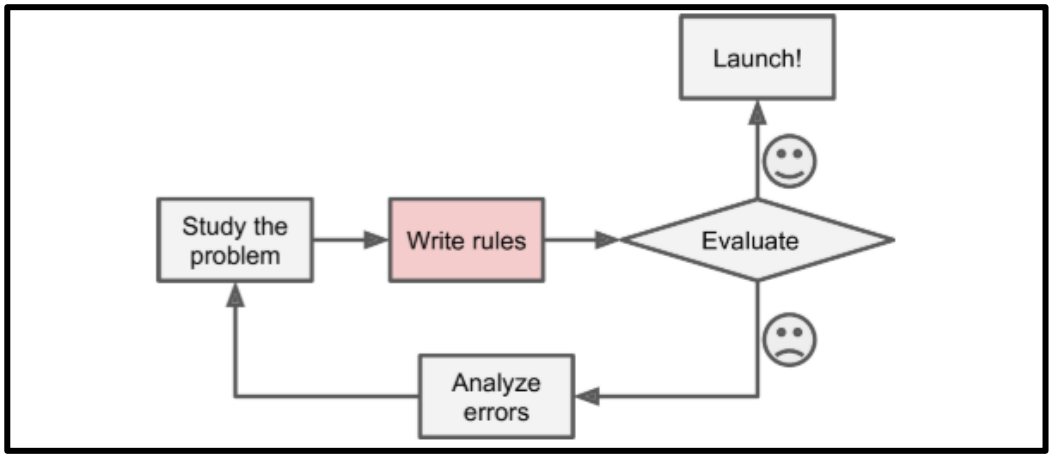
3.2.3 Machine Learning vs. programas informáticos tradicionales

La programación tradicional ha existido por décadas y se basa en la escritura manual de código por parte de un programador. En este enfoque, el código establece reglas específicas en un lenguaje de programación, las cuales la computadora ejecuta para procesar datos de entrada y generar resultados conforme a las instrucciones predefinidas (Dark, 2022).

(Dark, 2022). En contraste, los algoritmos de Machine Learning no requieren una programación explícita para cada tarea. En su lugar, construyen modelos matemáticos a partir de datos de entrenamiento. A medida que el modelo recibe más información, su capacidad para realizar predicciones o tomar decisiones mejora progresivamente, similar a cómo los humanos aprenden con la experiencia. La precisión de estos modelos se valida comparando sus predicciones con datos históricos conocidos.

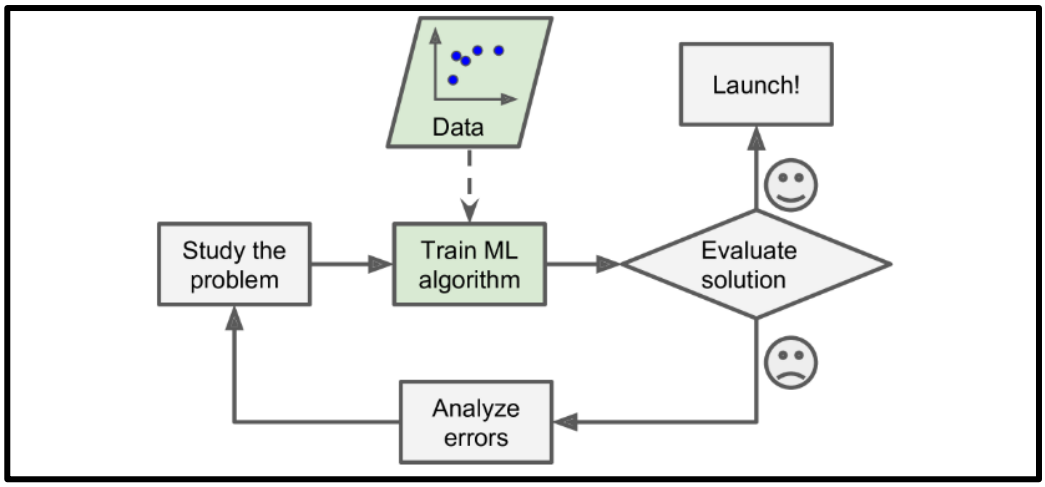
(Konasani Y Kadre, 2021, p. 26). El Machine Learning tiene aplicaciones en múltiples sectores. Por ejemplo, en el ámbito financiero, los bancos utilizan modelos entrenados con datos históricos para evaluar la probabilidad de que un nuevo cliente incumpla un préstamo. En este caso, las variables de entrada pueden incluir ingresos, historial crediticio y número de tarjetas de crédito. Gracias a este enfoque, las instituciones pueden tomar decisiones más precisas y automatizadas sin depender exclusivamente de reglas programadas manualmente.

Repositorio Institucional - UNAMBA Perú



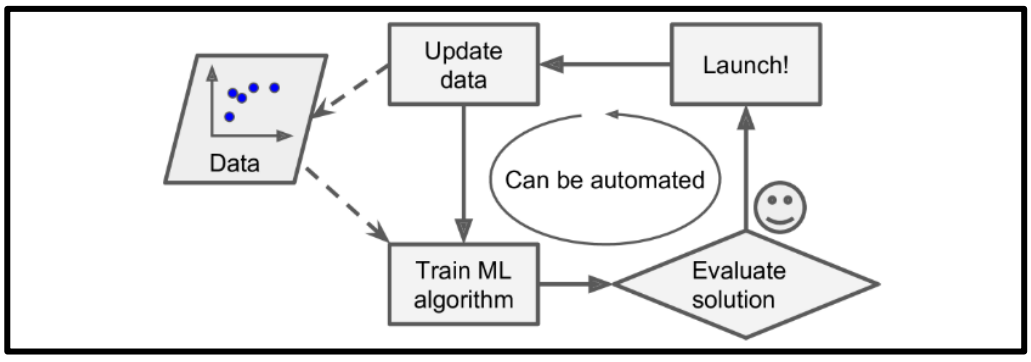
FUENTE: Géron, 2019, p. 31.

Figura 2 — El enfoque tradicional de Machine Learning



FUENTE: Géron, 2019, p. 32.

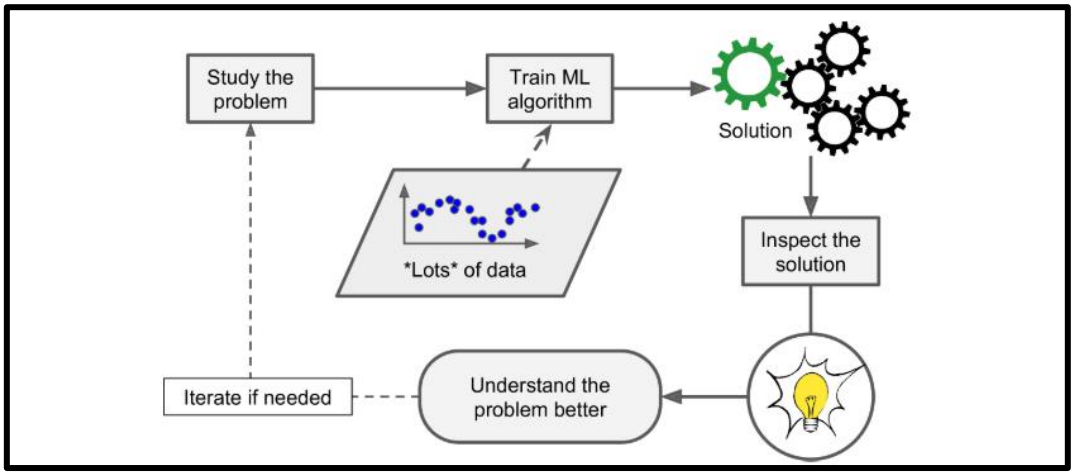
Figura 3 — Enfoque de aprendizaje automático



FUENTE: Gérón, 2019, p. 32.

Figura 4 — Adaptarse automáticamente al cambio

Repositorio Institucional - UNAMBA Perú



FUENTE: Gérón, 2019, p. 32.

Figura 5 — Machine Learning puede ayudar a los humanos a aprender

3.2.4 Tipos de sistemas de Machine Learning

(Géron, 2019). Los sistemas de Machine Learning pueden clasificarse según diferentes criterios, como el tipo de supervisión durante el entrenamiento, la forma en que aprenden y se adaptan, y el método de comparación con los datos de entrada. A continuación, se presentan las principales categorías:

3.2.4.1 Según el tipo de supervisión

a) Aprendizaje supervisado

En este tipo de Machine Learning, el modelo se entrena con un conjunto de datos etiquetados, es decir, donde las entradas están asociadas con salidas conocidas. El algoritmo ajusta sus parámetros para minimizar el error y mejorar la precisión en futuras predicciones.

Regresión Lineal y Logística (predicción de valores continuos o probabilidades).

Árboles de Decisión y Random Forest (clasificación y regresión).
Redes Neuronales Artificiales (procesamiento de imágenes, voz y texto).

b) Aprendizaje no supervisado

Aquí, el modelo no tiene acceso a etiquetas o respuestas correctas. Su objetivo es encontrar estructuras ocultas en los datos, agrupando

Repositorio Institucional - UNAMBA Perú

elementos similares o reduciendo la dimensionalidad para facilitar la interpretación.

c) Aprendizaje Semi-Supervisado

(Géron, 2019). Este enfoque combina elementos del aprendizaje supervisado y no supervisado. El modelo se entrena con una pequeña cantidad de datos etiquetados y una gran cantidad de datos sin etiquetar, aprovechando ambos para mejorar su rendimiento.

Algoritmos que entrenan redes neuronales con imágenes etiquetadas y luego refinan su precisión con imágenes sin etiquetas.

d) Aprendizaje por Refuerzo

Aquí, un agente aprende mediante prueba y error, interactuando con su entorno y recibiendo recompensas o penalizaciones según sus acciones. Con el tiempo, optimiza sus estrategias para maximizar la recompensa acumulada.

Q-Learning: Algoritmos para optimizar estrategias en videojuegos.
Deep Q Networks (DQN): Modelos que combinan redes neuronales con refuerzo.

3.2.4.2 Según cómo generalizan el aprendizaje

a) Aprendizaje Offline (Batch learning)

El modelo se entrena con todos los datos disponibles antes de ser desplegado. No se actualiza con nueva información, por lo que requiere ser reentrenado periódicamente.

b) Aprendizaje Online (Online learning)

El modelo se actualiza continuamente con nuevos datos sin necesidad de ser reentrenado desde cero, lo que le permite adaptarse a cambios en tiempo real.

Repositorio Institucional - UNAMBA Perú

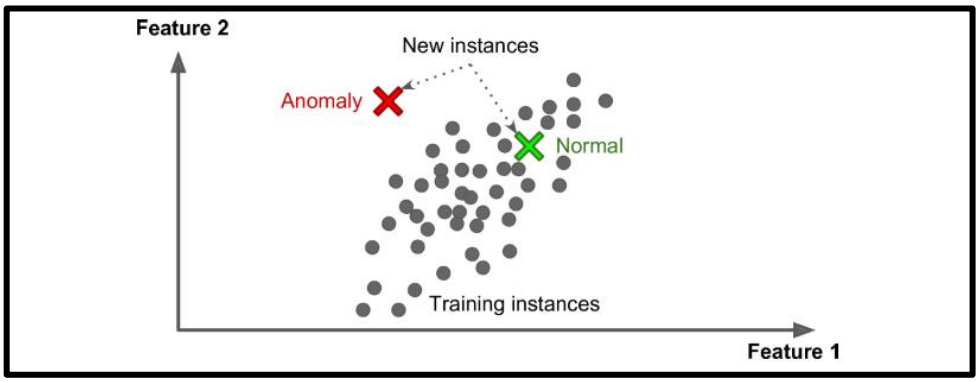
3.2.4.3 Según cómo comparan los datos de entrada

a) Modelos Basados en Memoria (Instance-Based Learning)

El modelo memoriza ejemplos del conjunto de entrenamiento y toma decisiones comparando nuevos datos con casos previos.

K-Nearest Neighbors (KNN).

Sistemas de recomendación y detección de anomalías.



FUENTE: Géron, 2019, p. 39.

Figura 6 — Detección de anomalías

b) Modelos Basados en Parámetros (Model-Based Learning)

Estos modelos extraen relaciones a partir de los datos de entrenamiento y crean una función matemática generalizada para hacer predicciones.

Regresión Lineal, Redes Neuronales Profundas.

Predicción de precios de viviendas, diagnóstico médico, sistemas de predicción meteorológica.



FUENTE: Géron, 2019, p. 35.

Figura 7 — Regresión

Repositorio Institucional - UNAMBA Perú

3.2.5 Pruebas y validación en Machine Learning

Para evaluar qué tan bien un modelo de Machine Learning puede generalizar a nuevos datos, es fundamental probarlo en ejemplos que no ha visto antes. Una opción es implementarlo directamente en un entorno real y monitorear su desempeño, pero si el modelo tiene un rendimiento deficiente, los usuarios podrían verse afectados negativamente (Géron, 2019).

Una mejor alternativa es dividir el conjunto de datos en dos partes: un conjunto de entrenamiento y un conjunto de prueba. El modelo se entrena con el primero y luego se evalúa con el segundo. La precisión obtenida en los datos de prueba proporciona una estimación del error de generalización, que mide qué tan bien el modelo manejará datos desconocidos (Kubát, 2023).

Si el modelo tiene un bajo error en el entrenamiento, pero un alto error en la prueba, significa que ha memorizado los datos de entrenamiento en lugar de aprender patrones generales, lo que indica un problema de sobreajuste (overfitting). Esto sugiere que el modelo no será efectivo para predecir nuevos casos y requiere ajustes para mejorar su capacidad de generalización (Géron, 2019).

3.2.6 Diseño del modelo de Machine Learning

- **Selección del algoritmo de Machine Learning:** Puede incluir Regresión lineal, random forest, redes neuronales artificiales (anns) o support vector regression (svr), dependiendo de la complejidad de los datos.

- **División de datos:** Se separan los datos en conjunto de entrenamiento y conjunto de prueba para evaluar el rendimiento del modelo.

- **Ajuste de hiperparámetros:** Se optimizan los valores del modelo para mejorar su desempeño.

(Géron, 2019).en su libro Hands-on machine learning with scikit-learn, keras, and tensorflow concepts, tools, and techniques to build intelligent systems plantea:
(Gopal, 2019). Evaluar un modelo en Machine Learning puede parecer sencillo utilizando un conjunto de prueba, pero elegir el mejor modelo entre varias opciones,

Repositorio Institucional - UNAMBA Perú

como un modelo lineal o polinómico, requiere una estrategia adecuada. Una opción es comparar su capacidad de generalización en los datos de prueba.

Sin embargo, ajustar hiperparámetros para mejorar el rendimiento puede generar problemas. Si un modelo es ajustado repetidamente en el mismo conjunto de prueba, se corre el riesgo de que solo funcione bien en esos datos específicos y no en datos nuevos, reduciendo su capacidad de generalización.

Para evitar esto, se emplea la técnica de validación por retención (holdout validation), donde se reserva una parte del conjunto de entrenamiento como conjunto de validación. Se entrenan varios modelos con distintos valores de hiperparámetros en el resto del conjunto de entrenamiento y se elige el modelo con mejor desempeño en la validación. Luego, el modelo seleccionado se entrena con el conjunto de entrenamiento completo y se evalúa en el conjunto de prueba para estimar su rendimiento real (Tsoukalas, 2024).

Si el conjunto de validación es demasiado pequeño, la selección del modelo puede ser imprecisa; si es muy grande, la cantidad de datos para el entrenamiento se reduce, afectando la calidad del modelo final. Para mitigar este problema, se usa la validación cruzada, donde se evalúan múltiples subconjuntos de validación en distintas iteraciones. Esto mejora la precisión de la evaluación, aunque incrementa significativamente el tiempo de entrenamiento.(Kalyan, Diaz y Carrasquillo, 2025)

3.2.7 Precisión del modelo

La precisión del modelo se mide mediante métricas estadísticas que determinan qué tan bien se ajusta el modelo a los datos reales.

- **Error Cuadrático Medio (MSE - Mean Squared Error):**

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 \tag{1}$$

donde

MSE	es el error cuadrático medio;
m	es el número total de observaciones;
i	es el índice correspondiente a cada observación;
y_i	es el valor observado de la i-ésima observación;
$\hat{y}_i$	es el valor predicho por el modelo para la i-ésima observación;

Repositorio Institucional - UNAMBA Perú

$(y_i - \hat{y}_i)$ es el error o residuo de predicción.

- **Coeficiente de Determinación (R^2):**

  Indica qué porcentaje de la variabilidad de los datos es explicado por el modelo.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (2)$$

donde

R^2	es el coeficiente de determinación;
m	es el número total de observaciones;
y_i	es el valor observado de la i-ésima observación;
$\hat{y}_i$	es el valor predicho por el modelo para la i-ésima observación;
$\bar{y}$	es la media aritmética de los valores observados;
$\sum_{i=1}^{m} (y_i - \hat{y}_i)^2$	es la suma de los cuadrados de los residuos;
$\sum_{i=1}^{m} (y_i - \bar{y})^2$	es la suma total de cuadrados.

- **Error Cuadrático Medio de Raíz (RMSE - Root Mean Squared Error)**

  El RMSE es la raíz cuadrada del MSE, lo que permite interpretar el error en las mismas unidades

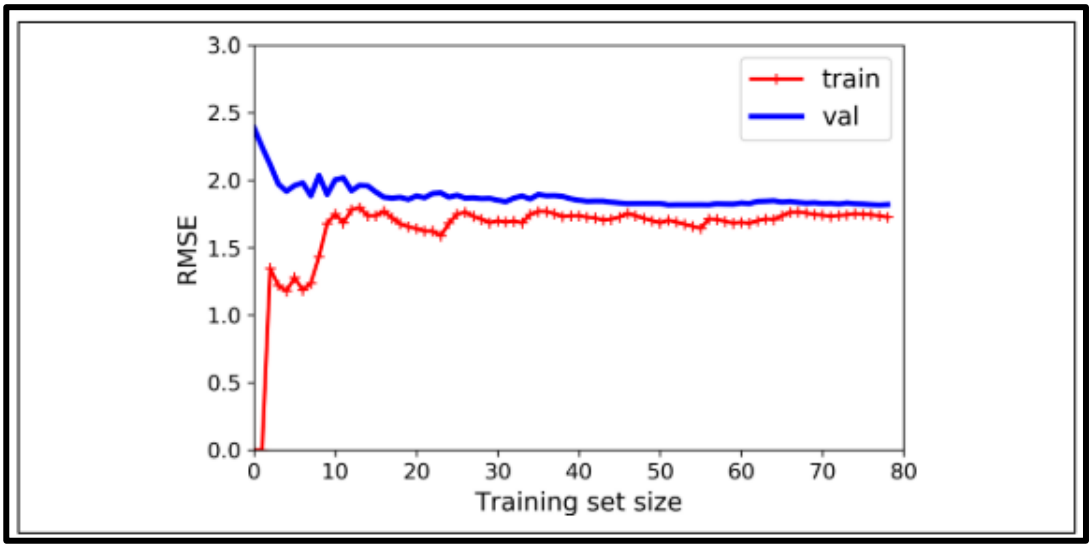
$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2} \quad (3)$$

donde

$RMSE$	es la raíz del error cuadrático medio;
MSE	es el error cuadrático medio;
m	es el número total de observaciones;
y_i	es el valor observado de la i-ésima observación;
$\hat{y}_i$	es el valor predicho por el modelo para la i-ésima observación;
$(y_i - \hat{y}_i)^2$	es el cuadrado del error de predicción.

Un RMSE bajo indica que el modelo predice valores cercanos a los reales con menor dispersión del error.

Repositorio Institucional - UNAMBA Perú



FUENTE: Géron, 2019, p. 160.

Figura 8 — Curvas de aprendizaje

3.2.8 Validación del modelo

La validación del modelo es crucial para determinar su capacidad de generalización, es decir, su habilidad para hacer predicciones precisas en nuevos datos.

Métodos de validación utilizados:

- Comparación entre valores predichos y valores experimentales; Se analizan las diferencias entre los valores generados por el modelo y los resultados obtenidos en ensayos geotécnicos reales.

- Validación por retención (Holdout Validation); Se divide el conjunto de datos en entrenamiento, validación y prueba.

- Validación cruzada (Cross-Validation); Se divide el conjunto de entrenamiento en múltiples subconjuntos para evaluar el modelo repetidamente.

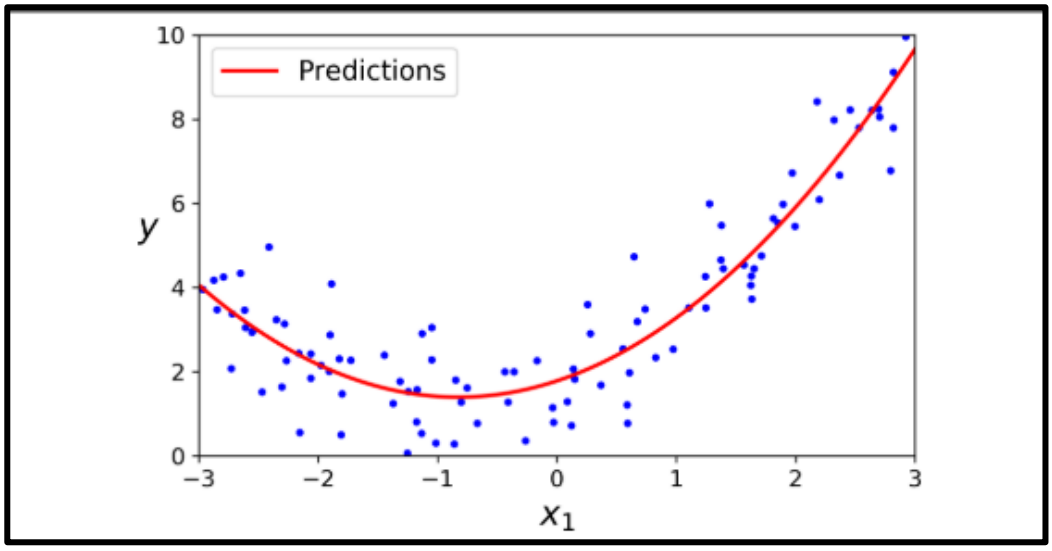
- Unidad de medida; Diferencia porcentual (%) entre predicciones y valores experimentales.

Un modelo bien diseñado garantiza que se seleccione el algoritmo adecuado y se usen variables relevantes.

Repositorio Institucional - UNAMBA Perú

Evaluar la precisión del modelo permite detectar problemas como sobreajuste o subajuste y ajustar el modelo en consecuencia.

La validación del modelo asegura que el sistema pueda aplicarse en escenarios reales con alta confiabilidad.



FUENTE: Géron, 2019, p. 158.

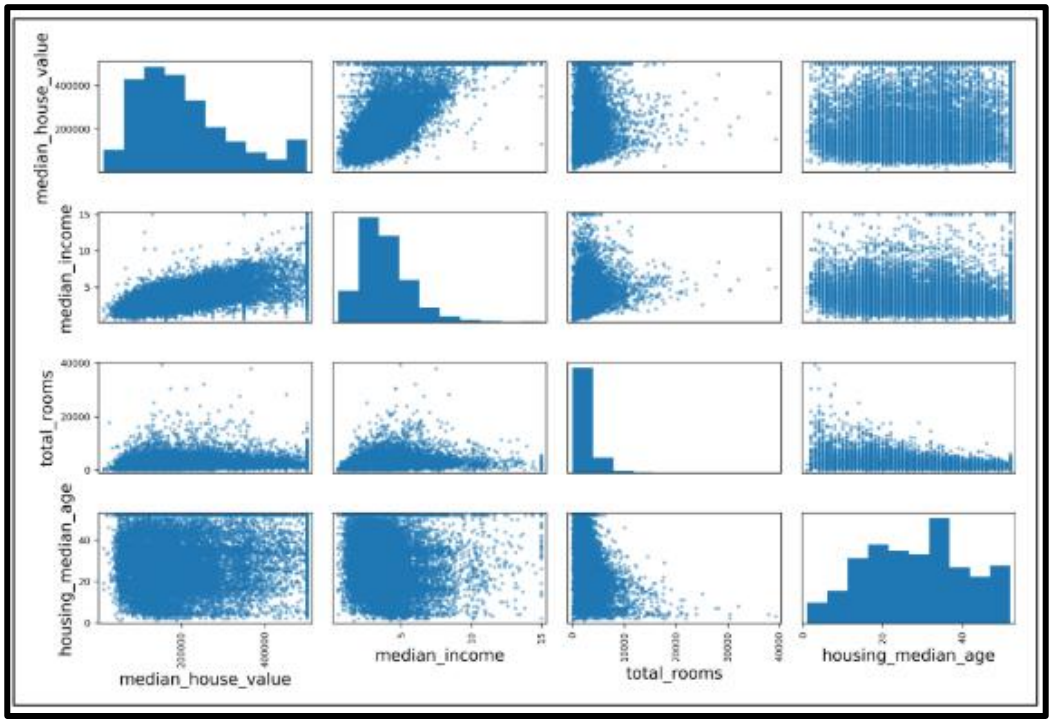
Figura 9 — Predicciones del modelo de regresión polinómica

3.2.9 Desajuste de datos

(Géron, 2019).En su libro Hands-on machine learning with scikit-learn, keras, and tensorflow concepts, tools, and techniques to build intelligent systems plantea:

En algunos casos, se puede disponer de una gran cantidad de datos para entrenar un modelo, pero estos no siempre representan fielmente los datos que se utilizarán en producción. Por ejemplo, si se desarrolla una aplicación móvil para identificar especies de flores a partir de fotografías, se podrían descargar millones de imágenes de internet. Sin embargo, estas no reflejan exactamente las condiciones en las que los usuarios tomarán fotos con la aplicación, lo que genera un desajuste entre los datos de entrenamiento y los reales (Kubát, 2023).

Repositorio Institucional - UNAMBA Perú



FUENTE: Géron, 2019, p. 35.

Figura 10 — Matriz de dispersión de datos

Para minimizar este problema, el conjunto de validación y prueba debe ser representativo de los datos que se usarán en producción. Si el rendimiento del modelo en la validación es bajo, puede ser difícil determinar si se debe a un sobreajuste o a la diferencia entre los datos de entrenamiento y los datos reales (Gonzáles, 2019).

(Géron, 2019). Una solución propuesta por Andrew Ng es incluir un train-dev set, que es un subconjunto del conjunto de entrenamiento que se reserva sin ser utilizado en el entrenamiento. Si el modelo tiene buen desempeño en el train-dev set, pero mal rendimiento en la validación, el problema es un desajuste de datos.

Para corregirlo, se pueden preprocesar las imágenes de entrenamiento para que se asemejen más a las imágenes reales. Si el modelo también tiene un mal desempeño en el train-dev set, significa que hay sobreajuste, lo que se puede abordar simplificando el modelo, regularizándolo o mejorando la calidad de los datos de entrenamiento.

3.2.10 Modelos de entrenamiento en Machine Learning

Hasta ahora, los modelos de Machine Learning y sus algoritmos de entrenamiento han sido tratados como cajas negras, lo que permite obtener buenos resultados sin

Repositorio Institucional - UNAMBA Perú

conocer sus detalles internos. Es posible optimizar modelos de regresión, mejorar clasificadores de imágenes y crear detectores de spam sin comprender completamente su funcionamiento. En muchos casos, no es necesario conocer a fondo su implementación para utilizarlos eficazmente (Kubát, 2023; Gopal, 2019).

Sin embargo, comprender cómo funcionan estos modelos y sus algoritmos puede ser de gran utilidad. Un conocimiento más profundo facilita la selección del modelo adecuado, la elección del algoritmo de entrenamiento óptimo y la configuración de hiperparámetros eficientes. Además, permite depurar errores y analizar problemas con mayor precisión, lo que mejora el desempeño en tareas de optimización.

a) Regresión lineal

La regresión lineal es un modelo de aprendizaje supervisado utilizado para predecir valores numéricos a partir de una o más variables independientes. Se basa en la suposición de que existe una relación lineal entre las variables de entrada (features) y la variable objetivo (output).(Tsoukalas, 2024)

Un modelo de Regresión Lineal predice un valor $y^{\wedge}$ a partir de una combinación lineal de las variables de entrada más un término de sesgo (bias o intercepto):

$$y^{\wedge}=\theta_0+\theta_1X_1+\theta_2X_2+\cdots+\theta_nx_n \quad (4)$$

donde

- $y^{\wedge}$ es el valor predicho;
- x_i son las variables de entrada o características (features);
- n es el número de características en el modelo;
- θ_j son los coeficientes del modelo (pesos) que determinan la influencia de cada variable en la predicción;
- θ_0 es el término de sesgo o intercepto, que ajusta la predicción cuando todas las variables de entrada son cero.

b) Entrenamiento del modelo

El objetivo es encontrar los valores óptimos de los parámetros θ que minimicen el error de predicción. Para ello, se utilizan dos métodos principales:

Repositorio Institucional - UNAMBA Perú

- **Mínimos cuadrados ordinarios (OLS - Ordinary Least Squares)**

  Este método ajusta los parámetros del modelo minimizando la suma de los errores al cuadrado:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{5}$$

  donde

m	es el número total de observaciones;
i	es el índice correspondiente a cada observación;
$\hat{y}_i$	es el valor predicho por el modelo para la i-ésima observación;
y_i	es el valor observado de la i-ésima observación;
$(\hat{y}_i - y_i)^2$	es el cuadrado del error de predicción;
$\sum^{m}_{i=1} (\hat{y}_i - y_i)^2$	es la suma de los cuadrados de los errores.

Ventaja: Preciso en datasets pequeños.

Desventaja: Ineficiente con grandes volúmenes de datos.

- **Descenso del gradiente (Gradient Descent)**

  (Kalyan, Diaz y Carrasquillo, 2025). Algoritmo iterativo que ajusta los parámetros θ\thetaθ gradualmente en función del error de predicción. La actualización de los parámetros sigue la regla:

$$\theta_j := \theta_j - \alpha \frac{\partial J}{\partial \theta_j} \tag{6}$$

  donde

θ_j	es el j-ésimo parámetro o coeficiente del modelo;
α	es la tasa de aprendizaje, que determina el tamaño del ajuste;
J	es la función de costo que se busca minimizar;
$\partial J/\partial \theta_j$	es la derivada parcial de la función de costo J con respecto al parámetro θ_j.

Ventaja: Escalable para grandes volúmenes de datos.

Desventaja: Sensible a la elección de la tasa de aprendizaje (α).

Repositorio Institucional - UNAMBA Perú

3.2.11 Expansión volumétrica

La expansión volumétrica en suelos, especialmente en suelos arcillosos, es un fenómeno importante que afecta tanto la estabilidad de las estructuras como el comportamiento de los suelos bajo condiciones cambiantes de humedad. La expansión volumétrica se refiere al cambio en el volumen de un suelo debido a la variación en su contenido de humedad. Este fenómeno es fundamental en ingeniería civil, especialmente en la construcción de pavimentos, cimientos y otras infraestructuras que se ubican sobre suelos expansivos (MTC, 2014).

- **Causas de la expansión volumétrica:**
  - **Absorción de agua:** Los suelos arcillosos, debido a su alta concentración de minerales como la montmorillonita, son muy propensos a absorber agua. Al hacerlo, las partículas de arcilla se separan, lo que provoca una expansión volumétrica.
  - **Variación estacional de humedad:** Los cambios en el clima, como las lluvias en invierno y la sequedad en verano, causan fluctuaciones de humedad en el suelo, lo que resulta en procesos cíclicos de expansión y contracción.
  - **Interacción con estructuras:** La expansión y contracción pueden afectar la estabilidad de infraestructuras, como carreteras, cimientos y pavimentos, provocando fisuras y deformaciones.

- **Métodos de medición de la expansión volumétrica:**
  - **Ensayo de expansión libre:** Se utiliza para medir la capacidad de un suelo de expandirse cuando se satura con agua. Este ensayo determina la expansión porcentual en volumen de una muestra de suelo cuando se satura completamente.
  - **Ensayo de succión:** Mide el comportamiento de los suelos cuando pierden agua, evaluando su capacidad de contracción.

Repositorio Institucional - UNAMBA Perú

– **Prueba de cambio volumétrico:** Se mide el volumen inicial de una muestra de suelo y su volumen después de ser sometido a diferentes niveles de humedad.

Tabla 2 — Clasificación de suelos expansivos

Potencial de expansión Ep	Expansión en consolidometro, bajo presion vertical de 7 kPa (0,07 kgf/cm2)	índice de plasticidad IP	Porcentaje de partículas menores que dos micras
%	%	%	%
Muy alto	>30	>32	>37
Alto	20-30	23-45	18-37
Medio	10-20	12-34	12-27
Bajo	<10	<20	<17
FUENTE: NORMA E 0.50, 2020, p. 27.			

3.2.12 Proctor

El Ensayo Proctor es un procedimiento de laboratorio fundamental en ingeniería geotécnica que se utiliza para determinar la densidad seca máxima y el contenido óptimo de humedad de un suelo. Fue desarrollado en 1933 por Ralph Proctor, con el objetivo de establecer una relación entre la compactación del suelo y la humedad que optimiza su resistencia y estabilidad.(Geotecnia, 2017; Braja M, 2013)

(MTC Manual ensayo de materiales, 2016). El ensayo consiste en compactar una muestra de suelo en un molde cilíndrico bajo una energía de compactación controlada, variando el contenido de humedad. Se realizan múltiples pruebas con distintos niveles de humedad para identificar el punto en el que el suelo alcanza su máxima densidad seca.

La relación entre la densidad seca (ρ_d) y el contenido de humedad (w) se representa en una curva de compactación, donde el punto más alto de la curva define:

- **Densidad seca máxima (ρ_{dmax})**
  La mayor densidad alcanzada bajo la energía de compactación aplicada.

Repositorio Institucional - UNAMBA Perú

$$\rho_d = \frac{100\,\rho_h}{w + 100} \tag{7}$$

donde

ρ^d es la densidad seca del suelo compactado, expresada en kg/m³;

ρ_h es la densidad húmeda del suelo compactado, expresada en kg/m³;

w es el contenido de humedad del suelo compactado, expresado en porcentaje (%).

- **Contenido Óptimo de Humedad (w_{opt})**

  La cantidad de agua que permite lograr la máxima compactación.

3.2.12.1 Tipos de ensayo Proctor

a) Ensayo proctor estándar (ASTM D698)

- Se aplica una energía de compactación baja, utilizando un martillo de 2.5 kg y una altura de caída de 305 mm en tres capas compactadas dentro del molde.

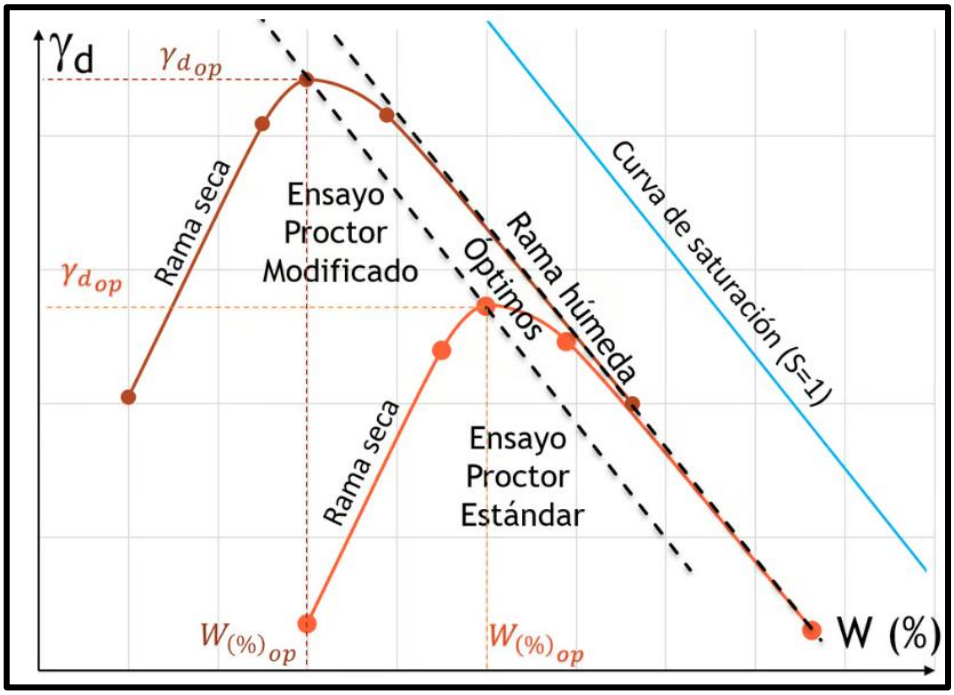
- Representa condiciones de compactación para materiales con cargas ligeras o pavimentos secundarios.

b) Ensayo Proctor Modificado (ASTM D1557)

- Se aplica una mayor energía de compactación, con un martillo de 4.5 kg y una altura de caída de 457 mm en cinco capas.

- Se usa en suelos que requieren una compactación más exigente, como bases de carreteras y aeropuertos.

Repositorio Institucional - UNAMBA Perú



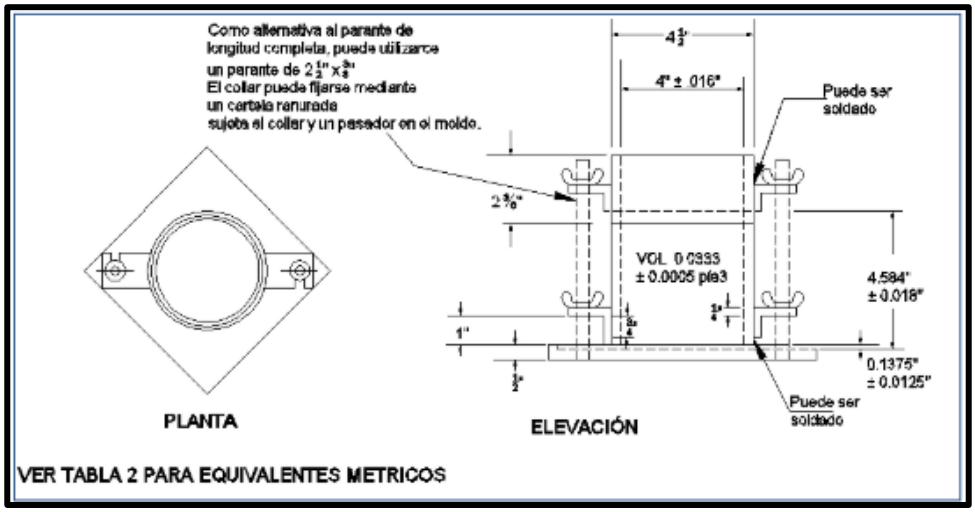
FUENTE: (Ingeniería geotécnica, 2023)

Figura 11 — Curva de compactación

Tabla 3 — La diferencia entre ambos ensayos

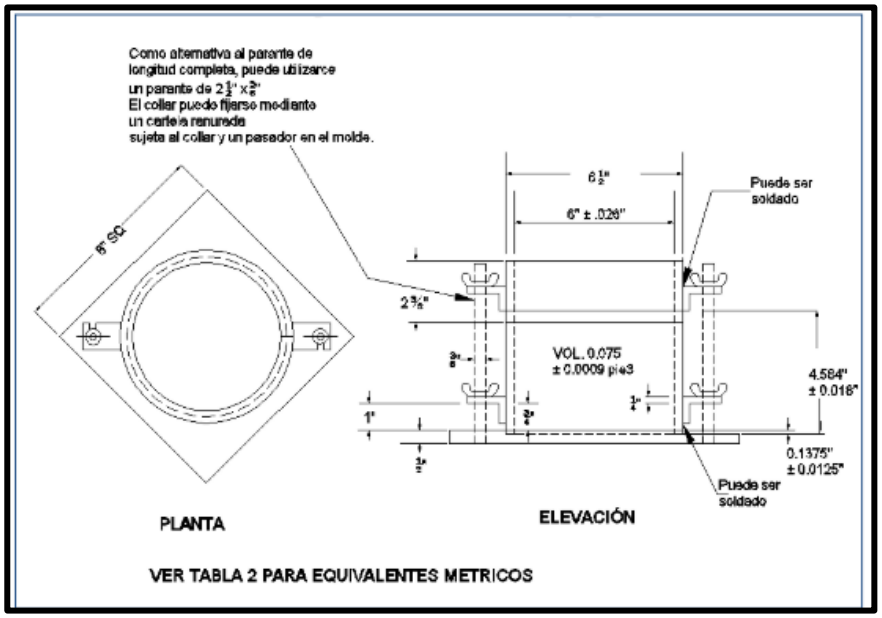
Descripción	Estándar	Modificado
Diámetro del molde [mm]	100 150	100 150
Volumen [cm^3]	945 2.145	945 2.145
Número de capas	3	5
Número de golpes	25 | 56	25 | 56
Peso del martillo [Kg]	2.5	45
Altura de caída del martillo [mm]	305 ±2	460 ±2
FUENTE: (Ingeniería geotécnica, 2023)		

Repositorio Institucional - UNAMBA Perú



FUENTE: MTC, 2014, p. 45.

Figura 12 — Molde cilíndrico de 4,0 pulg



FUENTE: MTC, 2014, p. 45.

Figura 13 — Molde cilíndrico de 6,0 Pulg

3.2.13 Parámetros geotécnicos

Los parámetros geotécnicos son las propiedades físicas, mecánicas e hidráulicas del suelo que determinan su comportamiento bajo diferentes condiciones de carga. Estos parámetros son fundamentales en la ingeniería geotécnica, ya que permiten caracterizar los suelos y predecir su capacidad de carga, deformabilidad y estabilidad en estructuras como carreteras, cimentaciones y pavimentos (Bowles, 1997; Rodríguez, 2018).

Repositorio Institucional - UNAMBA Perú

Tabla 4 — Principales parámetros geotécnicos

Parámetro Geotécnico	Descripción	Norma ASTM
Contenido de humedad (w)	Relación entre la cantidad de agua en el suelo y su peso seco.	ASTM D2216-19
Granulometría	Distribución del tamaño de partículas del suelo.	ASTM D 422
Límites de Atterberg	Determinan la plasticidad del suelo (LL, *LP* e *IP).*	ASTM D4318-10

3.2.13.1 Contenido de humedad - ASTM D 2216

Es la cantidad de agua presente en el suelo en relación con su peso seco. Es un parámetro fundamental en la ingeniería geotécnica, ya que afecta la resistencia, la capacidad de carga y la compactación del suelo (MTC - manual de ensayos de materiales, 2016).

Tabla 5 — La cantidad mínima de espécimen de material húmedo

Máximo tamaño de partícula (pasa el 100%)	Tamaño de malla estándar	Masa mínima recomendada de espécimen de ensayo húmedo para contenidos de humedad reportados	
		a ± 0,1%	a ± 1%
2 mm o menos	2,00 mm (Nº 10)	20 g	20 g *
4,75 mm	4,760 mm (Nº 4)	100 g	20 g *
9,5 mm	9,525 mm (3/8")	500 g	50 g
19,0 mm	19,050 mm (3/4")	2,5 kg	250 g
37,5 mm	38,1 mm (1 1/2")	10 kg	1 kg
75,0 mm	76,200 mm (3")	50 kg	5 kg
FUENTE: MTC, 2014, p. 45.			

Según la norma ASTM D2216-19, el contenido de humedad se determina en laboratorio mediante el método de secado en horno, en el cual una muestra de suelo es secada a 105°C - 110°C hasta alcanzar un peso constante.

Repositorio Institucional - UNAMBA Perú

$$W = \frac{M_{CWS} - M_{CS}}{M_{CS} - M_C} \times 100 = \frac{M_W}{M_S} \times 100 \quad (8)$$

donde

Mcws es el peso del contenedor más el suelo húmedo, en gramos;

Mcs es el peso del contenedor más el suelo secado en horno, en gramos;

Mc es el peso del contenedor, en gramos;

Mw es el peso del agua, en gramos;

Ms es el peso de las partículas sólidas, en gramos.

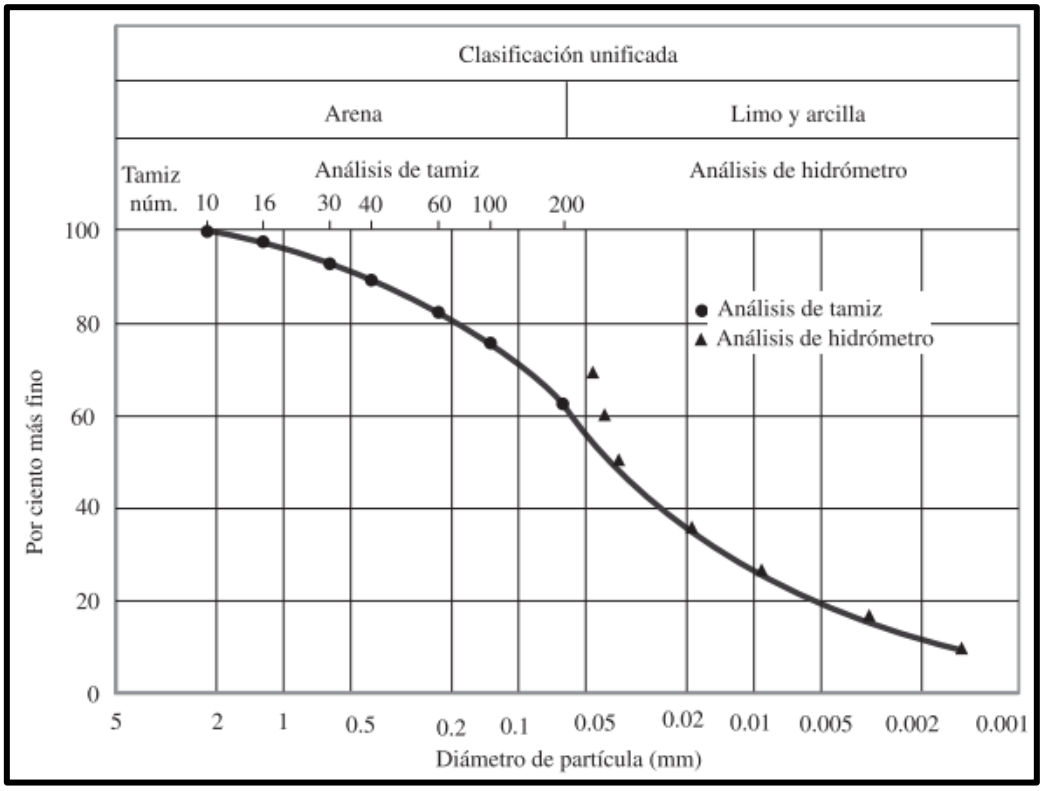
3.2.13.2 Análisis granulométrico - ASTM D 422

(MTC - manual de ensayos de materiales, 2016)La granulometría es el estudio de la distribución del tamaño de partículas en un suelo, permitiendo su clasificación en gravas, arenas, limos y arcillas. Es un parámetro esencial en la ingeniería geotécnica, ya que influye en la compactación, permeabilidad, estabilidad y capacidad de carga del suelo.

Tabla 6 — Tamices de malla cuadrada

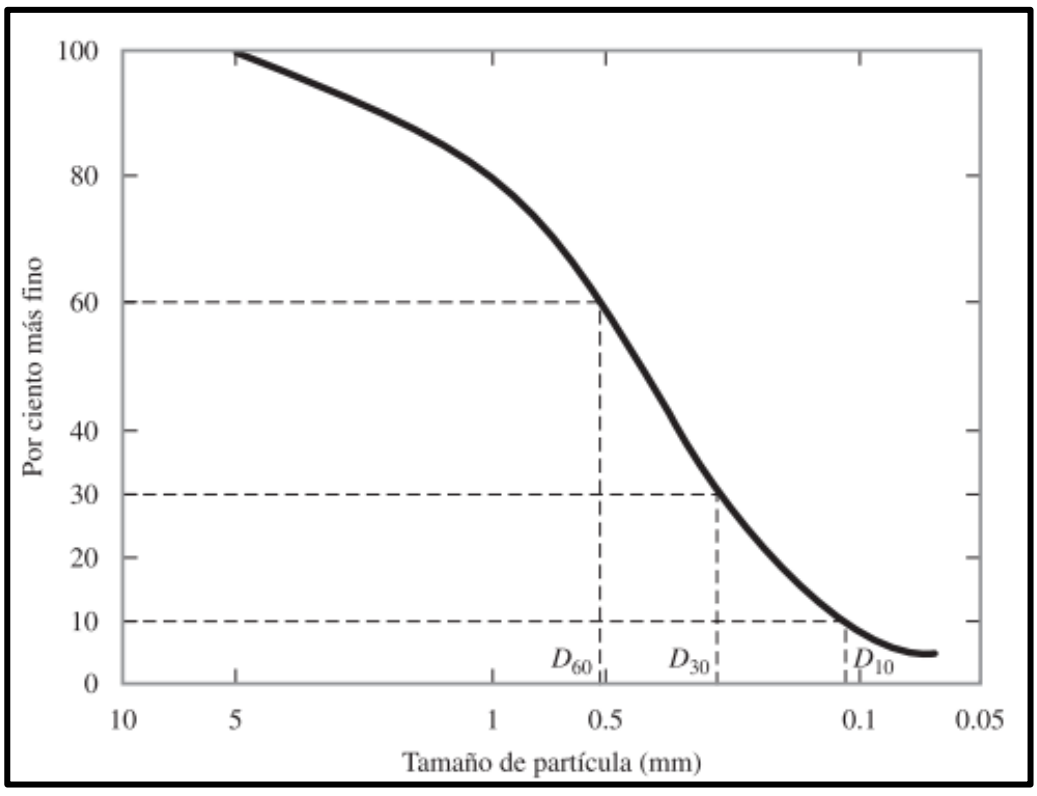
TAMICES	ABERTURA (mm)
3"	75.000
2"	50.800
1 ½ "	38.100
1"	25.400
3/4 "	19.000
3/8"	9.500
N° 4	4.760
N° 10	2.000
N° 20	0.840
N° 40	0.425
N° 60	0.260
N° 140	0.106
N° 200	0.075
FUENTE: MTC, 2014, p. 45.	

Repositorio Institucional - UNAMBA Perú



FUENTE: Braja M, 2013, p. 60.

Figura 14 — Curva de distribución de tamaño de partícula



FUENTE: Braja M, 2013, p. 60.

Figura 15 — Definición de D10, D30 y D60

Repositorio Institucional - UNAMBA Perú

Nombre de la organización	Tamaño de grano (mm)			
	Grava	Arena	Limo	Arcilla
Instituto de Tecnología de Massachusetts (MIT)	>2	2 a 0.06	0.06 a 0.002	<0.002
Departamento de Agricultura de E.U. (USDA)	>2	2 a 0.05	0.05 a 0.002	<0.002
Asociación Americana de Carreteras Estatales y Oficiales del Transporte (AASHTO)	76.2 a 2	2 a 0.075	0.075 a 0.002	<0.002
Sistema Unificado de Clasificación de Suelos (Cuerpo de Ingenieros del Ejército de E.U., Oficina de Reclamación de E.U., Sociedad Americana para Pruebas y Materiales)	76.2 a 4.75	4.75 a 0.075	Finos (p.ej., linos y arcillas) <0.075	

FUENTE: Braja M, 2013, p. 60.

Figura 16 — Límites de separación de tamaño de suelo

3.2.13.3 Clasificación de suelos

La clasificación de suelos es un sistema que permite agrupar los materiales terrestres en función de sus propiedades físicas y mecánicas. Este proceso es fundamental en ingeniería geotécnica, ya que facilita la identificación de suelos adecuados para cimentaciones, pavimentos y otras estructuras (Braja M, 2011).

Para su clasificación, se utilizan diferentes sistemas normalizados, siendo los más reconocidos:

- Sistema Unificado de Clasificación de Suelos (SUCS o USCS, ASTM D2487-17).

- Clasificación de Suelos para Carreteras de AASHTO (AASHTO M 145-91).

a) Sistema unificado de clasificación de suelo SUCS

El Sistema Unificado de Clasificación de Suelos (SUCS o USCS, por sus siglas en inglés) es uno de los métodos más utilizados en

Repositorio Institucional - UNAMBA Perú

ingeniería geotécnica para clasificar suelos en función de su granulometría y plasticidad. Este sistema fue desarrollado por Casagrande en 1948 y posteriormente estandarizado en la norma ASTM D2487-17 (Standard Practice for Classification of Soils for Engineering Purposes).

El SUCS es ampliamente utilizado en cimentaciones, estabilidad de taludes y pavimentos, ya que permite evaluar el comportamiento mecánico del suelo según su tamaño de partículas y cohesión (Braja M, 2011).

Criterio para la asignación de símbolos de grupo				Símbolos de grupo
Suelos de grano grueso Más de 50% retenido en el tamiz núm. 200	**Gravas** Más de 50% de fracción gruesa retenida en el tamiz núm. 4	Gravas limpias Menos de 5% finos[a]	$C_u \geq 4$ y $1 \leq C_c \leq 3^c$	GW
			$C_u < 4$ y/o $1 > C_c > 3^c$	GP
		Gravas con finos Más de 12% finos[a,d]	$PI < 4$ o gráficos por debajo de línea "A" (figura 4.2)	GM
			$PI > 7$ y gráficos en o por encima de línea "A" (figura 4.2)	GC
	Arenas 50% o más de la fracción gruesa pasa tamiz núm. 4	Arenas limpias Menos de 5% finos[b]	$C_u \geq 6$ y $1 \leq C_c \leq 3^c$	SW
			$C_u < 6$ y/o $1 > C_c > 3^c$	SP
		Arenas con finos Más de 12% finos[b,d]	$PI < 4$ o gráficos por debajo de línea "A" (figura 4.2)	SM
			$PI > 7$ y gráficos en o por encima de línea "A" (figura 4.2)	SC
Suelos de grano fino 50% o más pasa a través del tamiz núm. 200	**Limos y arcillas** Límite líquido menor que 50	Inorgánico	$PI > 7$ y gráficos en o por encima de línea "A" (figura 4.2)[e]	CL
			$PI < 4$ o gráficos por debajo de línea "A" (figura 4.2)[e]	ML
		Orgánico	$\frac{\text{Límite líquido: secado}}{\text{Límite líquido: no secado}} < 0.75$; vea la figura 4.2; zona OL	OL
	Limos y arcillas Límite líquido 50 o más	Inorgánico	Gráficos *PI* en o por encima de línea "A" (figura 4.2)	CH
			Gráficos *PI* por debajo de "A" línea (figura 4.2)	MH
		Orgánico	$\frac{\text{Límite líquido: secado}}{\text{Límite líquido: no secado}} < 0.75$; vea la figura 4.2; zona OH	OH
Suelos altamente orgánicos	Materia orgánica principalmente, color oscuro y orgánico			Pt

[a] Gravas con 5 a 12% de finos requieren símbolos dobles: GW-GM, GW-GC, GP-GM, GP-GC.
[b] Arenas con 5 a 12% de finos requieren símbolos dobles: SW-SM, SW-SC, SP-SM, SP-SC.
[c] $C_u = \frac{D_{60}}{D_{10}}$; $C_c = \frac{(D_{30})^2}{D_{60} \times D_{10}}$
[d] Si $4 \leq PI \leq 7$ y gráficos en la zona rayada en la figura 4.2, se usa doble símbolo GC-GM o SC-SM.
[e] Si $4 \leq PI \leq 7$ y gráficos en la zona rayada en la figura 4.2, se usa doble símbolo CL-ML.

FUENTE: Braja M, 2013, p. 104.

Figura 17 — Sistema unificado de clasificación de suelo

b) Clasificación de Suelos según AASHTO

La clasificación de suelos de la AASHTO (American Association of State Highway and Transportation Officials) es un sistema ampliamente utilizado en ingeniería vial y geotécnica para evaluar la aptitud de los suelos en la construcción de carreteras y pavimentos. Se basa en la granulometría y en la plasticidad del suelo, evaluada mediante los Límites de Atterberg.

Repositorio Institucional - UNAMBA Perú

Este sistema está regulado por la norma AASHTO M 145-91, que clasifica los suelos en ocho grupos principales (A-1 a A-8), según su capacidad de soporte y comportamiento estructural.

Clasificación general	**Materiales granulares (35% o menos del total de la muestra pasada por el núm. 200)**						
	A-1			**A-2**			
Grupo de clasificación	A-1-a	A-1-b	A-3	A-2-4	A-2-5	A-2-6	A-2-7
Análisis de tamiz (porcentaje de paso)							
Núm. 10	50 máx.						
Núm. 40	30 máx.	50 máx.	51 mín.				
Núm. 200	15 máx.	25 máx.	10 máx.	35 máx.	35 máx.	35 máx.	35 máx.
Características de la fracción de paso núm. 40							
Límite líquido				40 máx.	41 mín.	40 máx.	41 mín.
Índice de plasticidad	6 máx.		NP	10 máx.	10 máx.	11 mín.	11 mín.
Tipos comunes de materiales significativos constituyentes	Fragmentos de roca, grava y arena		Arena fina		Limo o grava arcillosa y arena		
Clasificación general de la subrasante	Excelente a bueno						

Clasificación general	**Materiales granulares (35% o menos del total de la muestra pasada por el núm. 200)**			
Grupo de clasificación	A-4	A-5	A-6	A-7 A-7-5* A-7-6†
Análisis de tamiz (porcentaje de paso)				
Núm. 10				
Núm. 40				
Núm. 200	36 mín.	36 mín.	36 mín.	36 mín.
Características de la fracción de paso núm. 40				
Límite líquido	40 máx.	41 mín.	40 máx.	41 mín.
Índice de plasticidad	10 máx.	10 máx.	11 mín.	11 mín.
Tipos comunes de materiales significativos constituyentes	Suelos limosos		Suelos arcillosos	
Clasificación general de la subrasante	Regular a malo			

*Para A-7-5, $PI \leq LL - 30$
†Para A-7-6, $PI > LL - 30$

FUENTE: Braja M, 2013, p. 100.

Figura 18 — Clasificación de suelos según AASHTO

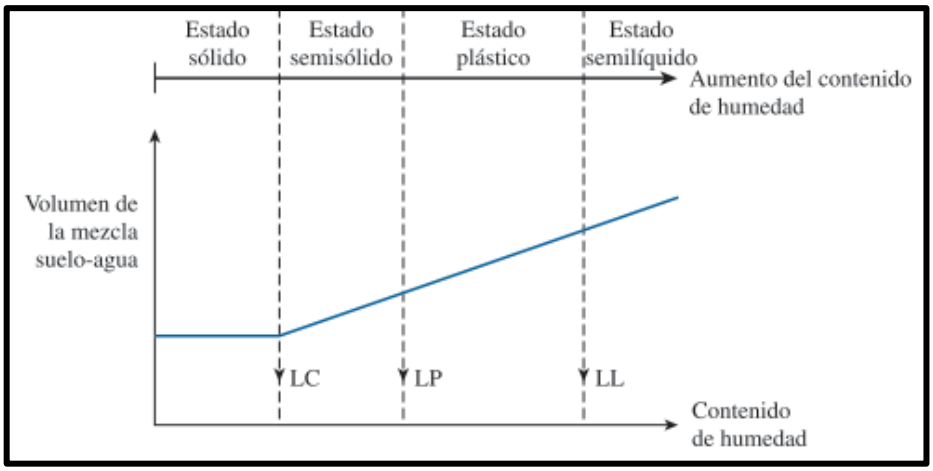
3.2.13.4 Límites de Atterberg

(MTC manual ensayo de materiales, 2016). Los Límites de Atterberg son parámetros fundamentales en la ingeniería geotécnica que permiten

Repositorio Institucional - UNAMBA Perú

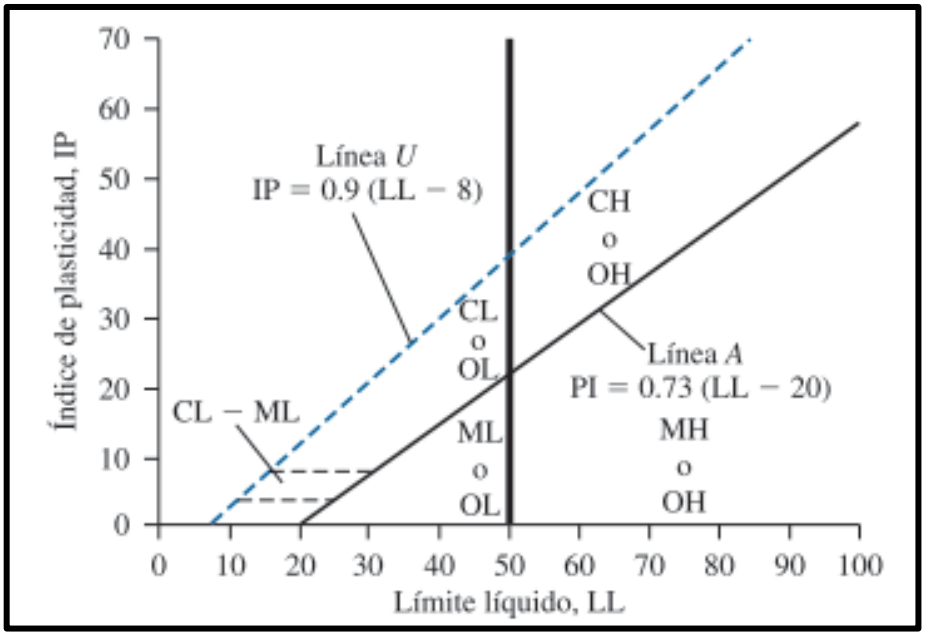
evaluar la plasticidad y el comportamiento de los suelos finos (limos y arcillas) en función de su contenido de humedad. Estos límites fueron desarrollados por Albert Atterberg en 1911 y posteriormente refinados por Casagrande para su aplicación en la clasificación de suelos.

La determinación de los Límites de Atterberg se realiza según la norma ASTM D4318-10 (Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils).



FUENTE: Braja M, 2013, p. 36.

Figura 19 — Límites de Atterberg.



FUENTE: Braja M, 2013, p. 40.

Figura 20 — Gráfica de plasticidad

Repositorio Institucional - UNAMBA Perú

- **Límite Líquido (LL)**
  - Es el contenido de humedad en el cual el suelo pasa del estado plástico al estado líquido.
  - Se determina mediante el Dispositivo de Casagrande o el ensayo de cono de penetración (ASTM D4318-10).
  - Un LL alto indica suelos altamente plásticos y expansivos.

Tabla 7 — Estimados de precisión LL

índice de precisión y tipo de ensayo	Desviación Estándar	Rango Aceptable de dos resultados
Precisión de un operador simple		
Límite Líquido	0.8	2.4
Precisión Multilaboratorio		
Límite Líquido	3.5	9.9
FUENTE: MTC manual ensayo de materiales, 2016, p. 72.		

- **Límite Plástico (LP)**

  Es el contenido de humedad en el cual el suelo pasa del estado semisólido al estado plástico. Se determina moldeando el suelo en hilos de 3 mm de diámetro, verificando el punto en el que se desmorona.

Tabla 8 — Estimados de precisión LP

índice de precisión y tipo de ensayo	Desviación Estándar	Rango Aceptable de dos resultados
Precisión de un operador simple		
Límite Plástico	0.9	2.6
Precisión Multilaboratorio		
Límite Plástico	3.7	10.6
FUENTE: MTC manual ensayo de materiales, 2016, p. 72.		

Repositorio Institucional - UNAMBA Perú

- **Índice de Plasticidad (IP)**
  - Es la diferencia entre el Límite Líquido y el Límite Plástico:
  - IP=LL−LP

3.2.14 Respuesta mecánica

La respuesta mecánica del suelo se refiere al comportamiento que presenta el material frente a la aplicación de cargas externas, particularmente aquellas generadas por el tránsito vehicular en estructuras de pavimento. Este comportamiento está relacionado con la capacidad del suelo para resistir esfuerzos, deformarse y recuperar parcialmente su forma cuando las cargas son removidas.

En el ámbito de la ingeniería geotécnica y del diseño de pavimentos, la respuesta mecánica de la subrasante se evalúa principalmente mediante parámetros como la capacidad de soporte del suelo (CBR) y el módulo resiliente (Mr). El CBR permite estimar la resistencia del suelo frente a la penetración, siendo ampliamente utilizado en metodologías tradicionales de diseño de pavimentos, mientras que el módulo resiliente representa la rigidez del suelo bajo cargas repetidas, constituyendo un parámetro fundamental en los enfoques mecanicistas de diseño estructural (MTC, 2008; Giordani y Leone, 2023; Montejo, 1998).

La adecuada estimación de estos parámetros permite comprender el comportamiento estructural de la subrasante y establecer condiciones adecuadas para el dimensionamiento de las capas del pavimento, contribuyendo a garantizar su desempeño y durabilidad a lo largo de su vida útil.

3.2.14.1 Capacidad de soporte del suelo (CBR)

La capacidad de soporte del suelo es un parámetro geotécnico que permite evaluar la resistencia del suelo frente a la penetración bajo condiciones controladas de carga. Este parámetro se determina mediante el ensayo CBR (MTC - Manual de ensayos de materiales, 2016), el cual consiste en medir la presión necesaria para que un pistón estándar penetre una muestra de suelo compactada a una velocidad constante, comparando dicho esfuerzo con el requerido para penetrar un material patrón.

Repositorio Institucional - UNAMBA Perú

El valor del CBR se expresa en porcentaje y representa la relación entre la resistencia del suelo evaluado y la resistencia del material de referencia. Este parámetro es ampliamente utilizado en el diseño de pavimentos flexibles, ya que permite estimar la capacidad del suelo para soportar las cargas transmitidas por el tránsito vehicular y sirve como base para el dimensionamiento de las capas estructurales del pavimento (Giordani y Leone, 2023; Montejo, 1998).

En términos prácticos, valores bajos de CBR indican suelos con baja capacidad de soporte y mayor susceptibilidad a deformaciones, mientras que valores más altos reflejan suelos con mayor resistencia estructural. Por esta razón, el CBR constituye uno de los parámetros más utilizados en estudios de suelos para proyectos viales.

3.2.14.2 Rigidez de la subrasante (Módulo resiliente, Mr)

El módulo resiliente (Mr) es un parámetro que representa la rigidez elástica del suelo cuando está sometido a cargas repetidas, como las generadas por el tránsito vehicular. Este parámetro describe la capacidad del suelo para deformarse temporalmente bajo la acción de una carga y recuperar su forma original una vez que dicha carga es retirada.

A diferencia del ensayo CBR, que se basa en una prueba de penetración estática, el módulo resiliente se obtiene a partir de ensayos triaxiales repetitivos que simulan las condiciones reales de carga que experimenta la subrasante en servicio. Este parámetro es fundamental en metodologías modernas de diseño mecanicista de pavimentos, ya que permite representar de manera más realista el comportamiento del suelo frente a cargas dinámicas (AASHTO, 2008).

El valor del módulo resiliente se expresa generalmente en unidades de presión, como megapascales (MPa), y su magnitud depende de factores como el tipo de suelo, el grado de compactación, el contenido de humedad y las condiciones de confinamiento. Valores elevados de Mr indican una mayor rigidez de la subrasante, lo que se traduce en una mejor capacidad para distribuir las cargas aplicadas por el pavimento.

Repositorio Institucional - UNAMBA Perú

Tabla 9 — Correlaciones empíricas para la estimación del Módulo Resiliente (Mr)

Expresión	Unidad	Tipo de suelo	Rango recomendado	Fuente
Mr = 10 · CBR	MPa	Subrasante general	CBR < 20%	SHELL (Heukelom & Foster, 1960)
Mr = 38 · CBR^0.711	MPa	Subrasantes finas	2% < CBR < 30%	USACE (Green & Hall, 1975)
Mr = 21 · CBR^0.65	MPa	Subrasante	3% < CBR < 20%	CSIR (Sudáfrica)
Mr = 18 · CBR^0.64	MPa	Subrasante fina	2% < CBR < 25%	TRRL (Lister, 1987)
Mr = B · CBR	MPa	Subrasante	5.25 < B < 21	AASHTO Guide
Mr = 1500 · CBR	psi	Subrasante fina	CBR < 10%	AASHTO 1993
Mr = 4326 ln(CBR) + 241	psi	Granulares	CBR ≤ 10%	FHWA
Mr = 22.1 · CBR^0.55	MPa	Granulares	12% < CBR < 80%	FHWA
Mr = 3000 · CBR^0.65	psi	Suelos finos	7% < CBR < 20%	NCHRP
Mr = 200 · DN^0.5	MPa	Cohesivos	Ensayo DCP	IMAE
Mr = 400 · DN^0.45	MPa	Granulares	Ensayo DCP	IMAE
Mr = k1·Pa·(θ/Pa)^k2	MPa	Ensayo triaxial repetido	General	MEPDG

NOTA

Las expresiones presentadas corresponden a correlaciones empíricas desarrolladas por distintos organismos técnicos (SHELL, USACE, CSIR, TRRL, AASHTO, FHWA, NCHRP e IMAE) para estimar el Módulo Resiliente (Mr) a partir de parámetros indirectos como el CBR o el índice de penetración dinámica (DN). Las unidades se presentan en MPa o psi según la formulación original; cuando fue necesario, se realizó la conversión correspondiente (1 MPa = 145.04 psi). La aplicabilidad depende del tipo de suelo y del rango de CBR indicado por cada autor.

Repositorio Institucional - UNAMBA Perú

El módulo resiliente se estima comúnmente a partir del CBR obtenido al 95% de la densidad máxima Proctor, ya que esta densidad representa el estándar de compactación habitual en obras viales. El valor al 100% se usa para condiciones de diseño especiales o muy exigentes

3.3 Marco conceptual

a) AASHTO (American Association of State Highway and Transportation Officials):

Organización que establece normas para la clasificación de suelos en carreteras y pavimentos (Vial, 2013).

b) Capacidad de soporte del suelo:

Es la aptitud del suelo para resistir las solicitaciones transmitidas por el pavimento. Su caracterización puede efectuarse mediante indicadores de resistencia como el CBR (Christopher, Schwartz y Boudreau, 2006).

c) Clasificación de Suelos:

Sistema que agrupa los suelos según sus propiedades mecánicas y físicas (DAS, 2013).

d) Coeficiente de determinación (R^2):

Es una medida que evalúa el ajuste de las predicciones mediante la comparación entre la suma de los errores cuadráticos y la variabilidad de los valores observados respecto de su media. (Chicco, Warrens y Jurman, 2021).

e) Contenido de Humedad:

Proporción de agua en un suelo en relación con su peso seco, expresada en porcentaje (ASTM International, 2013).

f) Densidad Seca:

Relación entre la masa del suelo seco y su volumen total, utilizada para evaluar su compactación (Braja M , 2013).

g) Desempeño del modelo predictivo:

Es el grado en que un modelo realiza adecuadamente una tarea, evaluado mediante medidas cuantitativas pertinentes. Para valorar su capacidad de generalización, debe

Repositorio Institucional - UNAMBA Perú

examinarse su comportamiento con datos diferentes de los utilizados durante el entrenamiento (Goodfellow, Bengio y Courville, 2016).

h) Ensayo Proctor:

Prueba de compactación utilizada para determinar la densidad seca máxima y el contenido óptimo de humedad de un suelo (Bowles, Joseph E., 1997).

i) Error absoluto medio (MAE)

Es una medida del error de predicción obtenida mediante el promedio de las diferencias absolutas entre los valores observados y los estimados. Se expresa en las mismas unidades de la variable evaluada, por lo que su interpretación depende de la escala de los datos; valores menores representan errores promedio de menor magnitud (Hyndman y Athanasopoulos, 2021).

j) Granulometría:

Distribución del tamaño de partículas del suelo, dividiéndolas en gravas, arenas, limos y arcillas (Das y Sobhan, 2018).

k) Índice de Plasticidad:

Diferencia entre el límite líquido y el límite plástico de un suelo, indicando su sensibilidad a la humedad (Holtz et al., 2011).

l) Índice de soporte de California (CBR):

Es la relación porcentual entre la carga necesaria para producir una penetración determinada en una muestra de suelo y la carga requerida para alcanzar la misma penetración en un material patrón de piedra triturada (Christopher, Schwartz y Boudreau, 2006).

m)Límite Líquido:

Contenido de humedad en el que un suelo pasa del estado plástico al líquido (Braja M, 2013).

Repositorio Institucional - UNAMBA Perú

n) Límite Plástico:

Contenido de humedad en el que un suelo pasa del estado semisólido al estado plástico (Braja M, 2013).

o) Límites de Atterberg:

Parámetros que determinan la plasticidad y consistencia de los suelos finos en función de su contenido de humedad (Holtz et al., 2011).

p) Machine Learning:

Rama de la inteligencia artificial que permite a los modelos computacionales aprender patrones a partir de datos y hacer predicciones (Géron, 2019).

q) Modelo predictivo basado en Machine Learnig:

Modelo computacional que aprende relaciones a partir de datos para estimar una variable de salida utilizando un conjunto de características de entrada. En los problemas de regresión, la salida corresponde a un valor numérico (Goodfellow, Bengio y Courville, 2016).

r) Módulo resiliente (Mr):

Es la relación entre el esfuerzo axial cíclico aplicado y la deformación axial recuperable correspondiente, bajo condiciones de carga repetida. Caracteriza la rigidez de los materiales no ligados utilizados en pavimentos (Christopher, Schwartz y Boudreau, 2006).

s) Normas ASTM:

Estándares internacionales utilizados para la caracterización de suelos en ingeniería geotécnica (ASTM International, 2024).

t) Parámetros Geotécnicos:

Propiedades físicas y mecánicas del suelo utilizadas en el análisis y modelación de su comportamiento (Braja M, 2013).

Repositorio Institucional - UNAMBA Perú

u) Respuesta mecánica del suelo:

Comprende el comportamiento del suelo frente a las cargas, caracterizado mediante propiedades de resistencia, rigidez y deformación (Christopher, Schwartz y Boudreau, 2006).

v) Rigidez de la subrasante:

Es la propiedad que expresa la resistencia del suelo de subrasante a deformarse bajo esfuerzos aplicados. Ante cargas repetidas, se caracteriza mediante el módulo resiliente (Christopher, Schwartz y Boudreau, 2006).

w) RMSE (Root Mean Squared Error):

Métrica utilizada para medir el error en modelos predictivos, indicando la diferencia promedio entre valores predichos y reales (Géron, 2019).

x) Sistema Unificado de Clasificación de Suelos (USCS o SUCS):

Método de clasificación de suelos basado en su granulometría y plasticidad (Holtz et al., 2011).

y) Validación del Modelo:

Proceso de evaluación del rendimiento de un modelo predictivo utilizando datos de prueba.(Wang et al., 2022).

Repositorio Institucional - UNAMBA Perú

CAPÍTULO IV
METODOLOGÍA

4.1 Tipo y nivel de investigación

4.1.1 Tipo de investigación

La presente investigación es de tipo aplicada, debido a que busca utilizar conocimientos científicos y herramientas tecnológicas para resolver un problema específico relacionado con la estimación de la respuesta mecánica de los suelos mediante modelos predictivos basados en Machine Learning. Según Hernández, Fernández y Baptista (2014), la investigación aplicada se orienta a utilizar el conocimiento científico para resolver problemas prácticos en contextos reales.

En este estudio, el enfoque aplicado se evidencia en el desarrollo de un modelo predictivo capaz de estimar parámetros geotécnicos como la capacidad de soporte del suelo (CBR) y la rigidez de la subrasante representada por el módulo resiliente (Mr), los cuales forman parte de la variable dependiente denominada respuesta mecánica del suelo.

4.1.2 Nivel de investigación

El nivel de la investigación es correlacional, ya que se analiza la relación entre el modelo predictivo basado en Machine Learning (variable independiente) y la respuesta mecánica del suelo (variable dependiente). Según (Hernández y Mendoza, 2018), los estudios correlacionales permiten identificar el grado de relación entre variables dentro de un fenómeno de estudio.

Asimismo, el estudio presenta un alcance predictivo, ya que el modelo desarrollado permite estimar el comportamiento de la variable dependiente a partir de datos disponibles mediante técnicas de aprendizaje automático.

4.2 Diseño de la investigación

El diseño de la investigación es no experimental. Es no experimental porque las variables no fueron manipuladas deliberadamente, sino que se analizaron datos geotécnicos tal como

Repositorio Institucional - UNAMBA Perú

se presentan en su contexto natural (Hernández, Fernández & Baptista, 2014). Es transversal porque el análisis de los datos se realizó en un único periodo de estudio (Bernal, 2016). Asimismo, presenta un carácter retrospectivo, ya que se emplearon registros y resultados de ensayos previamente obtenidos para el desarrollo del modelo predictivo.

4.3 Descripción ética de la investigación

La ética de la investigación comprende los principios y las buenas prácticas que orientan la conducta responsable del investigador durante las distintas fases de la actividad científica. Implica actuar con honestidad, rigurosidad, imparcialidad, transparencia, respeto y responsabilidad en la obtención, análisis y comunicación de los resultados (CONCYTEC, 2024).

La presente investigación se desarrolló respetando los principios de integridad científica, ética en la investigación y responsabilidad académica establecidos por la normativa nacional e institucional. En primer lugar, se tomó como referencia el Código Nacional de la Integridad Científica, aprobado por el Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC), el cual establece los principios de honestidad, responsabilidad, imparcialidad, respeto y justicia que deben regir toda actividad de investigación científica. Asimismo, dicho código prohíbe expresamente prácticas como la fabricación, falsificación y plagio de datos, promoviendo la transparencia y la confiabilidad de los resultados obtenidos (CONCYTEC, 2024).

De igual manera, la investigación se desarrolló en concordancia con el Reglamento de Ética en la Investigación de la Universidad Nacional Micaela Bastidas de Apurímac (UNAMBA), el cual establece los lineamientos para garantizar la integridad científica durante la planificación, ejecución y difusión de las investigaciones realizadas en la institución. Este reglamento promueve principios como la honestidad, transparencia, responsabilidad y respeto, además de asegurar que los trabajos de investigación se desarrollen conforme a la normativa vigente y con compromiso hacia la sociedad y el medio ambiente.

En ese sentido, durante el desarrollo de la presente investigación, los datos obtenidos de los ensayos geotécnicos fueron procesados y analizados de manera objetiva, sin alteración, manipulación o sesgo que pudiera afectar la confiabilidad de los resultados. Asimismo, se describió de forma transparente el procedimiento seguido para la recopilación,

Repositorio Institucional - UNAMBA Perú

procesamiento y análisis de los datos, así como la implementación de los modelos predictivos basados en Machine Learning. De igual manera, se respetaron los derechos de autor mediante la adecuada citación de las fuentes de información utilizadas, evitando cualquier forma de plagio y garantizando el cumplimiento de los principios de integridad académica. Finalmente, la investigación busca contribuir al desarrollo de la ingeniería geotécnica mediante la aplicación de técnicas de Machine Learning para estimar la respuesta mecánica de los suelos de subrasante, proporcionando una herramienta de apoyo para la toma de decisiones en proyectos de infraestructura vial.

4.4 Población y muestra

4.4.1 Población

La población corresponde al conjunto total de unidades de análisis que comparten características comunes y constituyen el objeto de estudio de una investigación (Hernández, Fernández & Baptista, 2014). En la presente investigación, la población estuvo conformada por 90 registros geotécnicos relacionados con la respuesta mecánica de los suelos de subrasante, los cuales contenían información de la capacidad de soporte del suelo (CBR), módulo resiliente (Mr) y demás parámetros geotécnicos requeridos para el desarrollo del modelo predictivo basado en Machine Learning. Estos registros corresponden a la totalidad de la información disponible que cumplió con las características requeridas para la investigación.

- **Criterios de inclusión**

  Se consideraron registros geotécnicos que presenten información completa de los parámetros asociados a la respuesta mecánica del suelo, particularmente valores de CBR y módulo resiliente, así como condiciones de compactación que permitan evaluar el comportamiento de la subrasante. Como resultado de la aplicación de estos criterios, se seleccionaron los 90 registros, los cuales conforman la muestra de estudio.

- **Criterios de exclusión**

  Se excluyeron de la investigación los registros geotécnicos que no se encontraban disponibles o que no contaban con la información necesaria para cumplir los objetivos del estudio. En consecuencia, no se incorporaron registros adicionales, debido a que la totalidad de la información disponible que cumplía los criterios establecidos estuvo conformada por 90 registros.

Repositorio Institucional - UNAMBA Perú

4.4.2 Muestra

La muestra es un subgrupo de la población de interés sobre el cual se recolectan los datos y que debe delimitarse previamente con precisión (Hernández, Fernández & Baptista, 2014).

La muestra de la investigación estuvo conformada por 90 registros geotécnicos, seleccionados mediante un muestreo no probabilístico intencional, debido a que se consideraron únicamente aquellos registros que presentaban información completa y útil para el desarrollo del modelo predictivo basado en Machine Learning.

No se aplicó una fórmula estadística para determinar el tamaño muestral, debido a que la investigación no tuvo un diseño experimental ni realizó selección aleatoria de unidades de análisis. Por el contrario, se trabajó con la totalidad de registros disponibles que cumplían los criterios de inclusión establecidos, por lo que la muestra corresponde a una muestra censal por disponibilidad.

Los registros considerados incluyeron parámetros geotécnicos como humedad natural, límites de Atterberg, granulometría, clasificación SUCS, densidad seca máxima, humedad óptima y valores de CBR al 95% y 100% de compactación. Estos datos fueron empleados para el entrenamiento, validación y evaluación del modelo predictivo.

Adicionalmente, se utilizaron los resultados de tres calicatas ejecutadas en la vía de evitamiento de Abancay, las cuales sirvieron como casos de contraste para comparar los valores obtenidos en laboratorio con los valores estimados por el modelo.

4.5 Procedimiento

(Hernández, Fernández & Baptista, 2014). explican que la investigación cuantitativa sigue un proceso secuencial y probatorio, cuyas etapas se encuentran relacionadas. Sobre esta base, el procedimiento de la presente investigación describe las actividades realizadas para recopilar, organizar, procesar y analizar la información necesaria para alcanzar los objetivos del estudio.

El desarrollo de la investigación se llevó a cabo mediante las siguientes etapas:

Repositorio Institucional - UNAMBA Perú

a) Etapa 1. Recopilación de información geotécnica

Se recopilaron los registros geotécnicos correspondientes a suelos de subrasante obtenidos de estudios de mecánica de suelos e informes de laboratorio. Inicialmente se revisó la totalidad de la información disponible, seleccionándose únicamente aquellos registros que contenían información completa de las variables requeridas para la investigación.

b) Etapa 2. Organización y conformación de la base de datos

Los registros seleccionados fueron organizados en una base de datos estructurada, conformada por 90 registros geotécnicos, incluyendo las variables de entrada (humedad natural, límites de Atterberg, granulometría, clasificación SUCS, densidad seca máxima, humedad óptima, entre otras) y las variables de salida correspondientes a la respuesta mecánica del suelo (CBR al 95 %, CBR al 100 % y módulo resiliente estimado).

c) Etapa 3. Preprocesamiento de los datos

La base de datos fue sometida a un proceso de revisión y depuración para identificar registros incompletos, errores de digitación e inconsistencias. Posteriormente, se realizó el procesamiento necesario para garantizar la calidad de los datos utilizados durante el entrenamiento del modelo de Machine Learning.

d) Etapa 4. Selección de variables

Se definieron las variables independientes y dependientes de acuerdo con la literatura especializada y los objetivos de la investigación, considerando únicamente aquellas variables geotécnicas con influencia directa sobre la respuesta mecánica de los suelos de subrasante.

e) Etapa 5. Entrenamiento del modelo predictivo

Con la base de datos preparada, se entrenaron los algoritmos de Machine Learning seleccionados para la investigación. Durante esta etapa se ajustaron los parámetros necesarios de cada algoritmo para obtener el mejor desempeño predictivo.

f) Etapa 6. Validación del modelo

El desempeño de los modelos fue evaluado mediante validación cruzada K-Fold, lo que permitió estimar su capacidad de generalización utilizando diferentes particiones

Repositorio Institucional - UNAMBA Perú

de la base de datos. Para la evaluación se emplearon indicadores de desempeño como el coeficiente de determinación (R^2), el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE).

g) Etapa 7. Contraste con datos reales

Finalmente, el modelo seleccionado fue contrastado utilizando los resultados experimentales obtenidos de tres calicatas ejecutadas en la vía de evitamiento de Abancay, comparando los valores predichos por el modelo con los obtenidos mediante ensayos de laboratorio.

h) Etapa 8. Análisis e interpretación de resultados

Los resultados obtenidos fueron analizados e interpretados considerando el desempeño de los modelos predictivos, la precisión alcanzada y su aplicabilidad para la estimación de la respuesta mecánica de los suelos de subrasante.

4.6 Técnica e instrumentos

4.6.1 Técnica

La técnica de investigación es el procedimiento o la forma particular utilizada para obtener datos o información. Su selección debe guardar correspondencia con el problema, los objetivos y el diseño de la investigación (Arias, 2012).

En la presente investigación se empleó la técnica de análisis documental para recopilar y analizar la información geotécnica utilizada en el desarrollo del modelo predictivo basado en Machine Learning. En primer lugar, se empleó la observación directa, mediante la revisión de los registros provenientes de ensayos geotécnicos, con la finalidad de verificar la calidad, consistencia e integridad de la información correspondiente a las propiedades físicas y mecánicas de los suelos de subrasante. Asimismo, se utilizó la observación indirecta, consistente en la revisión de estudios geotécnicos, informes técnicos y bibliografía especializada que permitieron sustentar el marco teórico y complementar la base de datos utilizada en la investigación.

Por otro lado, se aplicó el análisis experimental a partir de los resultados de ensayos de laboratorio disponibles, obteniéndose los valores de las propiedades geotécnicas empleadas como variables de entrada y salida para el desarrollo del

Repositorio Institucional - UNAMBA Perú

modelo predictivo, principalmente la capacidad de soporte del suelo (CBR), el módulo resiliente (Mr) y demás parámetros asociados a la respuesta mecánica del suelo. Finalmente, se utilizaron técnicas computacionales mediante la implementación de algoritmos de Machine Learning, los cuales permitieron entrenar, validar y evaluar los modelos predictivos para estimar la respuesta mecánica de los suelos de subrasante, comparando su desempeño mediante indicadores estadísticos de precisión.

4.6.2 Instrumentos

Los instrumentos de recolección de datos son los recursos, dispositivos o formatos, físicos o digitales, utilizados para obtener, registrar o almacenar información (Arias, 2012).

En la presente investigación se utilizó como instrumento de recolección de datos la "Ficha de recolección de datos geotécnicos", mediante la cual se sistematizó la información obtenida de los ensayos de laboratorio, incluyendo parámetros como granulometría, contenido de humedad, límites de Atterberg, densidad seca, capacidad de soporte del suelo (CBR) y módulo resiliente (Mr). Posteriormente, los registros fueron organizados en una base de datos digital en Microsoft Excel, permitiendo verificar la consistencia de la información y facilitar su procesamiento.

La ficha fue sometida a validación mediante juicio de cuatro expertos: Dr. Edwar Ilasaca Cahuata, Mtr. Américo Roldán Juárez, Mtr. Katherine Astrid Llerena Torreblanca e Ing. Máximo Villegas Vargas. Los especialistas evaluaron diez criterios: pertinencia, coherencia, congruencia, objetividad, validez técnica, suficiencia, organización, claridad, formato y aplicabilidad. Las puntuaciones obtenidas fueron de 44, 49, 49 y 47 sobre 50, respectivamente, equivalentes a coeficientes de validación de 0,88; 0,98; 0,98 y 0,94. Los cuatro expertos emitieron opinión favorable para la aplicación del instrumento.

Adicionalmente, a partir de las calificaciones otorgadas por los expertos a los diez criterios de evaluación, se calculó el coeficiente alfa de Cronbach, obteniéndose un valor de 0,76, según se presenta en la tabla 30 del Anexo E. Este análisis se refiere a la consistencia interna de las valoraciones de los criterios y complementa

Repositorio Institucional - UNAMBA Perú

la evaluación por juicio de expertos. Las respectivas constancias de validación se incluyen en el mismo anexo.

Finalmente, como herramientas informáticas para el desarrollo del modelo predictivo, se utilizaron el lenguaje de programación Python y la biblioteca Scikit-learn, los cuales permitieron implementar, entrenar y evaluar los algoritmos de Random Forest, Gradient Boosting, Extra Trees e HistGradient Boosting, empleados para estimar la respuesta mecánica de los suelos de subrasante. Asimismo, el desempeño de los modelos fue evaluado mediante indicadores estadísticos como el coeficiente de determinación (R^2), el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE) y el error cuadrático medio (MSE).

4.7 Estadístico de investigación

En toda investigación científica, el análisis estadístico desempeña un papel importante al transformar los datos recopilados en información útil para su interpretación. Mediante este proceso es posible describir las características de los datos y fundamentar las decisiones derivadas de los resultados obtenidos (Hernández y Mendoza, 2018). Se divide en:

- Estadística descriptiva
- Estadística inferencial

4.7.1 Estadística descriptiva

Es la rama de la estadística que permite organizar, resumir y describir las características de los datos mediante tablas, gráficos y medidas estadísticas (Hernández y Mendoza, 2018).

En el presente trabajo permitirá analizar las propiedades geotécnicas de los suelos y los valores obtenidos de la capacidad de soporte (CBR) y módulo resiliente (Mr), proporcionando una visión clara de los datos recolectados antes de aplicarlos al modelo de Machine learning.

a) Medidas de tendencia central:

- **Media (promedio):** Indica el valor medio de los parámetros geotécnicos.

Repositorio Institucional - UNAMBA Perú

- **Mediana:** Representa el punto medio de la distribución de los valores para evitar sesgos por valores extremos.

- **Moda:** Identifica el valor más frecuente en las mediciones.

b) Medidas de dispersión:

- **Desviación estándar:** Mide cuánto varían los valores geotécnicos respecto a la media.

- **Coeficiente de variación:** Relaciona la dispersión de los valores en función de su media.

c) Gráficos a emplear:

- **Histogramas:** Para visualizar la distribución de los valores de los parámetros geotécnicos.

- **Diagramas de dispersión (Scatter plots):** Para analizar relaciones entre las variables independientes.

- **Boxplots (diagramas de caja y bigotes):** Para detectar valores atípicos y evaluar la dispersión de los datos.

4.7.2 Estadística inferencial

La estadística inferencial permite contrastar hipótesis poblacionales y estimar parámetros a partir de los datos obtenidos de una muestra (Hernández y Mendoza, 2018).

La estadística inferencial permitió contrastar las hipótesis de investigación mediante pruebas estadísticas acordes con los objetivos del estudio y con la naturaleza de los datos. Para evaluar la capacidad predictiva del modelo desarrollado se emplearon métricas de desempeño como el coeficiente de determinación (R^2), el coeficiente de determinación mediante validación cruzada (CV-R^2), el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE). Asimismo, debido a que la variable respuesta mecánica no presentó distribución normal, se utilizó el coeficiente de

Repositorio Institucional - UNAMBA Perú

correlación de Spearman para analizar la relación entre los valores obtenidos experimentalmente y los estimados por el modelo predictivo basado en Machine Learning.

a) Hipótesis general

- **Prueba estadística a utilizar:**
  - Coeficiente de correlación de Spearman, para evaluar la relación entre los valores experimentales y los valores estimados por el modelo predictivo.
  - Coeficiente de determinación (R^2) para poder estimar la respuesta mecánica de los suelos a nivel de subrasante.
  - Error Cuadrático Medio (MSE) y Raíz del Error Cuadrático Medio (RMSE) para medir la precisión del modelo.

- **Criterios de decisión:**
  - Si el coeficiente de correlación de Spearman presenta significancia estadística ($p < 0.05$), se rechaza H_0 y se determina que existe una relación estadísticamente significativa entre los valores experimentales y los valores estimados por el modelo predictivo.
  - Complementariamente, valores altos de R^2 y bajos valores de MAE y RMSE evidencian un mejor desempeño predictivo del modelo.
  - Si la relación no resulta estadísticamente significativa ($p \geq 0.05$), no se rechaza H_0.

b) Hipótesis Específica 1

- **Prueba estadística a utilizar:**
  - Coeficiente de correlación de Spearman, para evaluar la relación entre los valores de capacidad de soporte del suelo (CBR) obtenidos experimentalmente y los valores estimados mediante el modelo predictivo basado en Machine Learning.

Repositorio Institucional - UNAMBA Perú

- **Criterios de decisión:**
  - Si los coeficientes de correlación de spearman son significativos ($p<0.05$), se rechaza H0 y se confirma que los parámetros geotécnicos influyen en la predicción de la expansión volumétrica.

  - Si los coeficientes no son significativos, se acepta H0, indicando que estos parámetros no tienen un impacto relevante en el modelo.

c) Hipótesis Específica 2

- **Prueba estadística a utilizar:**
  - Comparación del modelo predictivo con un modelo estadístico base (regresión lineal múltiple).

  - Evaluación de métricas de desempeño del modelo de Machine Learning:
  - R2 (Coeficiente de determinación).

  - MSE y RMSE.

  - Gráficos de residuales y predicción vs observado.

- **Criterios de decisión:**
  - Si el modelo de Machine Learning tiene un R2 alto y errores bajos, se rechaza H0 y se valida que el modelo puede estimar la expansión volumétrica con precisión.

  - Si los errores son elevados y el modelo no predice correctamente, se acepta H0.

Repositorio Institucional - UNAMBA Perú

CAPÍTULO V
RESULTADOS Y DISCUSIONES

5.1 Análisis de resultados

Se presenta los resultados obtenidos a partir del entrenamiento y aplicación del modelo predictivo basado en Machine Learning, cuyo propósito es estimar la respuesta mecánica de los suelos a nivel de subrasante en la ciudad de Abancay, Apurímac. El análisis se fundamenta en una base de datos construida a partir de ensayos de laboratorio que incluyen propiedades, índices, parámetros de compactación y valores de capacidad de soporte (CBR), los cuales constituyen los insumos principales del modelo.

Tabla 10 — Ubicación de las calicatas en evaluación

Vía	Profundidad	Este	Sur
Calicata C1 Vía de Evitamiento Abancay 03001	1.5 m	731356.21 m E	8494714.84 m S
Calicata C2 Vía de Evitamiento Abancay 03001	1.5 m	731264.63 m E	8494771.30 m S
Calicata C3 Vía de Evitamiento Abancay 03001	1.5 m	730746.12 m E	8494613.26 m S
NOTA Las coordenadas se presentan en metros (m) bajo el sistema de proyección UTM. La profundidad corresponde al nivel de exploración de cada calicata ejecutada en la Vía de Evitamiento Abancay 03001.			

El desempeño del modelo fue evaluado mediante métricas estadísticas que permiten cuantificar su capacidad predictiva y su coherencia respecto a los valores experimentales. Asimismo, se realizó una validación comparativa utilizando ensayos de CBR ejecutados en campo, con el fin de verificar la aplicabilidad práctica del modelo bajo condiciones reales.

Repositorio Institucional - UNAMBA Perú

El análisis de resultados se organiza de acuerdo con las dimensiones de la variable dependiente, iniciando con la capacidad de soporte del suelo y posteriormente la rigidez de la subrasante, incorporando en cada caso una interpretación técnica orientada a explicar el comportamiento mecánico observado.

5.1.1 Análisis de los resultados de variable dependiente: Respuesta mecánica

La variable dependiente denominada respuesta mecánica del suelo a nivel de subrasante fue evaluada a partir de dos parámetros fundamentales en el comportamiento estructural del pavimento: la capacidad de soporte del suelo (CBR) y la rigidez de la subrasante.

Estos parámetros permiten cuantificar la resistencia del suelo frente a cargas verticales y su capacidad de deformación bajo solicitaciones de servicio. Los resultados obtenidos del modelo predictivo fueron comparados con los valores experimentales disponibles, con el propósito de analizar su coherencia, precisión y aplicabilidad en condiciones reales.

5.1.1.1 Capacidad de soporte del suelo

La capacidad de soporte del suelo describe el comportamiento resistente de la subrasante frente a la aplicación de cargas externas, especialmente aquellas generadas por el tránsito vehicular. En esta investigación, dicha propiedad fue determinada mediante el ensayo CBR, el cual permite cuantificar la resistencia del material en función de su grado de compactación. Este parámetro constituye un indicador fundamental en el diseño estructural de pavimentos, ya que condiciona el espesor de las capas superiores y el desempeño global de la estructura vial.

Repositorio Institucional - UNAMBA Perú

Tabla 11 — Clasificación y plasticidad de suelos

CALICATA	HUMEDAD NATURAL	SUCS	AASHTO	LIMITES DE ATTERBERG		
				LL	LP	IP
C -01	9.74	CH	A-6(10)	54.2	27.66	26.54
C -02	8.44	CL	A-6(8)	39.8	15.42	24.38
C -03	8.71	SC	A-6(4)	34.8	17.14	17.66

NOTA

Los valores de humedad natural y límites de Atterberg (LL, LP e IP) fueron determinados mediante ensayos de laboratorio conforme a las normas ASTM D4318 y ASTM D2216. La clasificación SUCS se realizó según el Sistema Unificado de Clasificación de Suelos, mientras que la clasificación AASHTO se determinó conforme a los criterios de la norma M 145. Los resultados corresponden a muestras representativas obtenidas en campo.

Tabla 12 — Clasificación del suelo según el Índice de Plasticidad (IP).

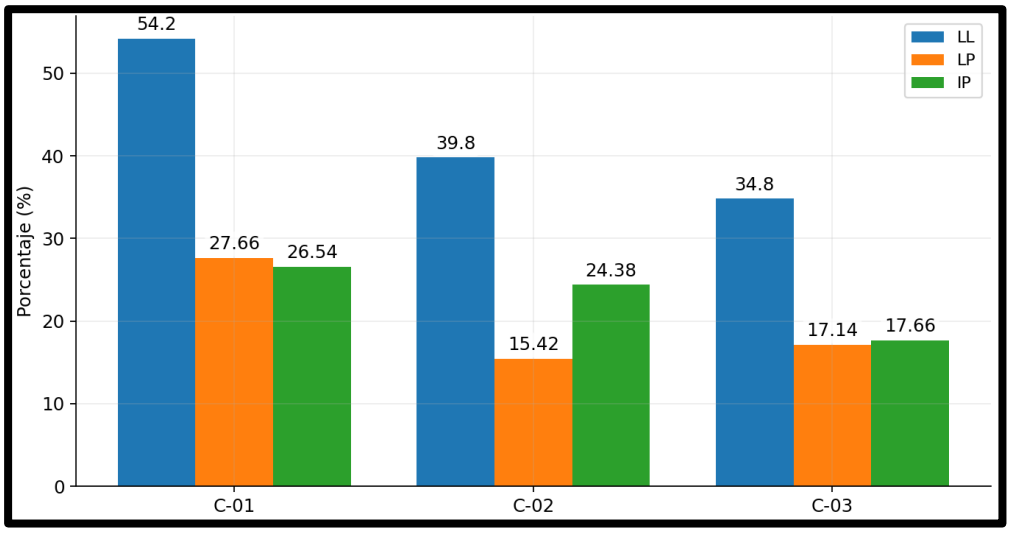
Índice de Plasticidad	Plasticidad	Característica
IP>20	Alta	Suelos muy arcillosos
20 > IP > 10	Media	Suelos arcillosos
10>IP>4	Baja	Suelos poco arcillosos
IP = 0	No Plástico (NP)	Suelos exentos de arcilla

NOTA

La clasificación del grado de plasticidad se realiza en función del Índice de Plasticidad (IP), definido como la diferencia entre el Límite Líquido (LL) y el Límite Plástico (LP), conforme a los criterios establecidos en el Sistema Unificado de Clasificación de Suelos (SUCS). Esta categorización permite identificar el comportamiento cohesivo y el potencial expansivo del material fino.

FUENTE: MTC, 2016.

Repositorio Institucional - UNAMBA Perú

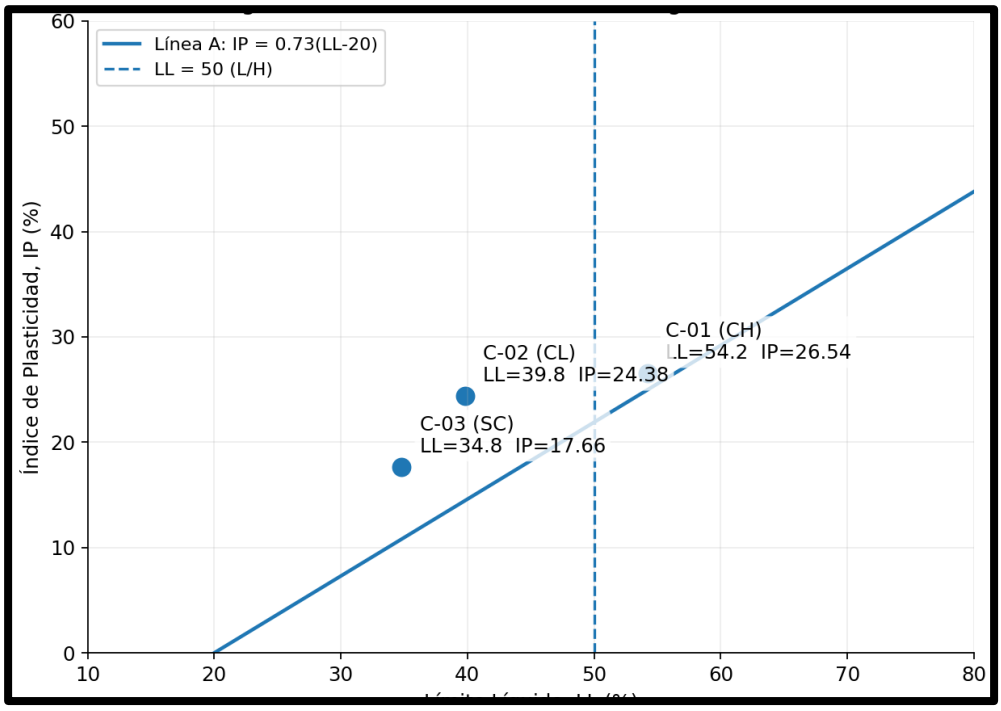


NOTA

La figura muestra los valores del Límite Líquido (LL), Límite Plástico (LP) e Índice de Plasticidad (IP) obtenidos mediante ensayos de laboratorio conforme a la norma ASTM D4318. Se observa que la calicata C-01 presenta el mayor grado de plasticidad (IP = 26.54%), clasificada como CH (arcilla de alta plasticidad), mientras que C-02 y C-03 corresponden a suelos de menor plasticidad relativa (CL y SC, respectivamente). Estos resultados permiten evaluar el comportamiento mecánico y la susceptibilidad volumétrica del suelo como subrasante.

Figura 21 — Comparación de los límites de Atterberg (LL, LP e IP) de las calicatas evaluadas

Repositorio Institucional - UNAMBA Perú

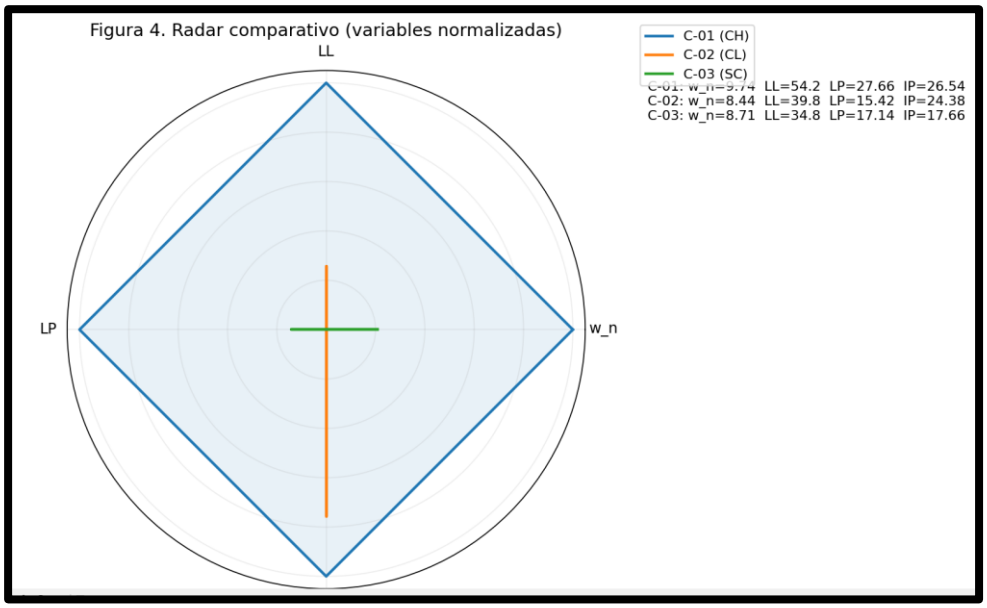


NOTA

La figura presenta la ubicación de las muestras C-01, C-02 y C-03 en la Carta de Plasticidad de Casagrande, considerando los valores del Límite Líquido (LL) y el Índice de Plasticidad (IP). La línea A ($IP = 0.73(LL - 20)$) permite diferenciar suelos arcillosos de limosos, mientras que la línea vertical $LL = 50$ distingue entre suelos de baja y alta plasticidad. Los resultados indican que C-01 se clasifica como CH (arcilla de alta plasticidad), C-02 como CL (arcilla de baja plasticidad) y C-03 como SC (arena arcillosa), conforme al Sistema Unificado de Clasificación de Suelos (SUCS).

Figura 22 — Clasificación de los suelos en la Carta de Plasticidad de Casagrande (SUCS)

Repositorio Institucional - UNAMBA Perú



NOTA

La figura presenta un análisis comparativo de la humedad natural (w_n), límite líquido (LL), límite plástico (LP) e índice de plasticidad (IP) para las calicatas evaluadas, empleando una normalización min–max para facilitar la comparación relativa entre variables de distinta magnitud. Se observa que la calicata C-01 (CH) presenta los valores relativos más altos en todas las variables índice, evidenciando mayor plasticidad y comportamiento cohesivo. Por su parte, C-02 (CL) muestra plasticidad intermedia, mientras que C-03 (SC) presenta menor plasticidad relativa, coherente con su clasificación como suelo granular con finos.

Figura 23 — Comparación de las propiedades índice mediante gráfico radar

Repositorio Institucional - UNAMBA Perú

Proyecto :	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025	**PERFORACION** C-01
UBICACIÓN :	Abancay - Apurímac	
SOLICITA :	Bach. Julio Cesar Riveros Ramos	
TECNICA DE INVESTIGACION : CALICATA		**DIMENSIÓN DE CALICATA :** L: 0.90m A: 1.00m
FECHA: Febrero, 2026		**PROFUNDIDAD FINAL :** 1.50 m
731356.21 m E		**PRESENCIA DE NIVEL FRIÁTICO:** NO SE ENCONTRÓ
8494714.84 m S		

PROF	ESPESOR	COTA M.S.N.M	CANTIDAD DE ESTRATO	SIMBOLO	CLASIFIC SUCS	DESCRIPCION DEL SUELO	OBSERVACIONES Y REGISTRO FOTOGRAFICO
	0.30 m	2450 msnm	1	----------		Material orgánico	
	1.20 m		2		CH	**SUCS:** CH Arcilla densa de alta plasticidad arenosa **AASHTO:** A-6 (10) Contienen partículas finas limosas o arcillosas con un límite líquido bajo **HUMEDAD %** 0.00% **LIMITES DE CONSISTENCIA** % LL : 54.2% % LP : 27.7% % IP : 26.54% **COLOR DEL SUELO** Rojo oscuro	**Tipo de excavación: manual** Profundidad: 1.50 m **Nivel friático:** No presenta **A fecha: Febrero** del 2026 **Hay presencia de particulas con tamaños menores a 1½" pulgadas**

NOTA

La figura muestra el perfil geotécnico correspondiente a la calicata C-01, ejecutada mediante excavación manual hasta una profundidad de 1.50 m. El estrato predominante fue clasificado como CH (arcilla de alta plasticidad) según el sistema SUCS, con LL = 54.2%, LP = 27.7% e IP = 26.54%. No se registró presencia de nivel freático durante la excavación.

Figura 24 — Perfil estratigráfico de la calicata C-01

Repositorio Institucional - UNAMBA Perú

PROYECTO :	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025	**PERFORACION** C-02
UBICACIÓN : Abancay - Apurímac **SOLICITA :** Bach. Julio Cesar Riveros Ramos		
TECNICA DE INVESTIGACION : CALICATA **FECHA:** Febrero, 2026 731264.63 m E 8494771.30 m S	**DIMENSIÓN DE CALICATA :** L: 0.90m A: 1.00m **PROFUNDIDAD FINAL :** 1.50 m **PRESENCIA DE NIVEL FRIÁTICO:** NO SE ENCONTRÓ	

PROF	ESPESOR	COTA M.S.N.M	CANTIDAD DE ESTRATO	SIMBOLO	CLASIFIC SUCS	DESCRIPCION DEL SUELO	OBSERVACIONES Y REGISTRO FOTOGRAFICO
	1.50 m	2450 msnm	1	CL	CL	**SUCS:** CL Arcilla ligera de baja plasticidad arenosa **AASHTO:** A-6 (8) Contienen partículas finas limosas o arcillosas con un límite líquido bajo **HUMEDAD %** 8.44% **LIMITES DE CONSISTENCIA** % LL : 39.8% % LP : 15.4% % IP : 24.38% **COLOR DEL SUELO** Rojo	 **Tipo de excavación: manual** Profundidad: 1.50 m **Nivel friático:** No presenta **A fecha: Febrero** del 2026 **Hay presencia de particulas con tamaños menores a 3/4" pulgadas**

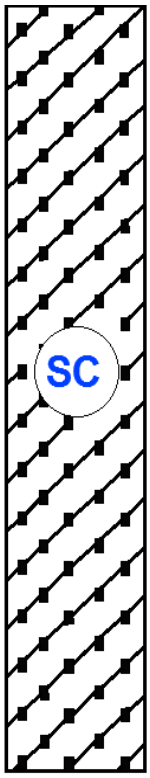
NOTA

La figura presenta el perfil geotécnico correspondiente a la calicata C-02, excavada manualmente hasta una profundidad de 1.50 m. El suelo predominante fue clasificado como CL (arcilla de baja plasticidad) según el sistema SUCS, con LL = 39.8%, LP = 15.4% e IP = 24.38%. No se registró presencia de nivel freático durante la excavación. El material presenta color rojo y partículas menores a 3/4".

Figura 25 — Perfil estratigráfico de la calicata C-02

Repositorio Institucional - UNAMBA Perú

PROYECTO :	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025	**PERFORACION** C-03
UBICACIÓN : Abancay - Apurímac **SOLICITA :** Bach. Julio Cesar Riveros Ramos	**INGENIERO:** EPT	
TECNICA DE INVESTIGACION : CALICATA **FECHA:** Febrero, 2026 730746.12 m E 8494613.26 m S	**DIMENSIÓN DE CALICATA :** L: 0.90m A: 1.00m **PROFUNDIDAD FINAL :** 1.50 m **PRESENCIA DE NIVEL FRIÁTICO:** NO SE ENCONTRÓ **COTA: 3327**	

PROF	ESPESOR	COTA M.S.N.M	CANTIDAD DE ESTRATO	SIMBOLO	CLASIFIC SUCS	DESCRIPCION DEL SUELO	OBSERVACIONES Y REGISTRO FOTOGRAFICO
	0.30 m	2450 msnm	1	----------	AF	**Afirmado**	
	1.20 m		2		SC	**SUCS:** SC Arena arcillosa con grava **AASHTO:** A-6 (4) Contienen partículas finas limosas o arcillosas con un límite líquido bajo **HUMEDAD %** 8.71% **LIMITES DE CONSISTENCIA** **% LL : 34.8%** **% LP : 17.1%** **% IP : 17.66%** **COLOR DEL SUELO** Marrón	**Tipo de excavación: manual** Profundidad: 1.50 m **Nivel friático:** No presenta **A fecha: Febrero** del 2026 **Hay presencia de particulas con tamaños menores a 3/4" pulgadas**

NOTA

La figura presenta el perfil geotécnico correspondiente a la calicata C-03, excavada manualmente hasta una profundidad de 1.50 m. El suelo fue clasificado como SC (arena arcillosa) según el sistema SUCS, con LL = 34.8%, LP = 17.1% e IP = 17.66%. No se registró presencia de nivel freático durante la excavación. El material presenta color marrón y partículas menores a 3/4".

Figura 26 — Perfil estratigráfico de la calicata C-03

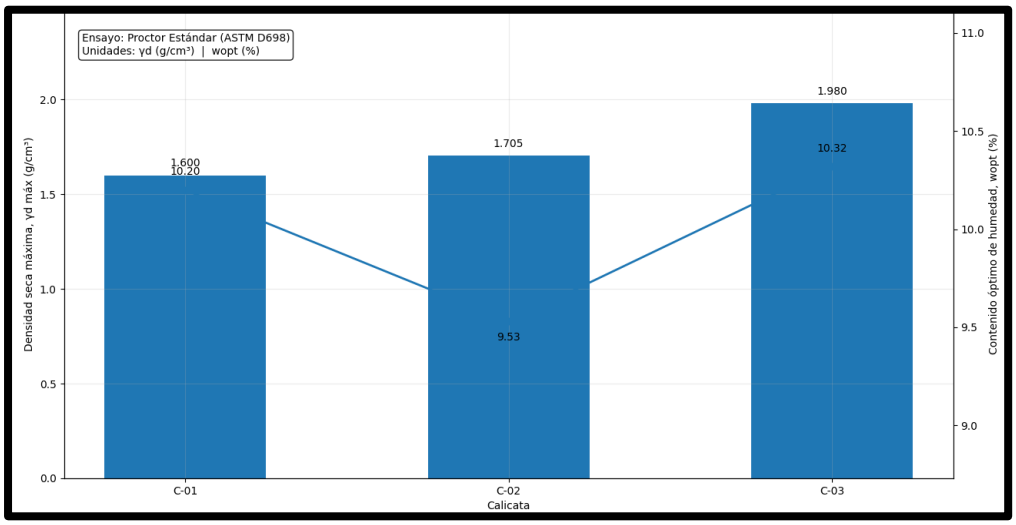
Repositorio Institucional - UNAMBA Perú

Tabla 13 — Compactación Proctor y capacidad de soporte CBR

	PROCTOR		CBR	
	DSM	**COH**	**CBR_100%**	**CBR_95%**
C -01	1.600	10.20	7.81	6.45
C -02	1.705	9.53	9.43	6.50
C -03	1.98	10.32	11.03	8.45

NOTA

La Densidad Seca Máxima (DSM) y el Contenido Óptimo de Humedad (COH) fueron determinados mediante el ensayo Proctor estándar conforme a la norma ASTM D698. Los valores de CBR al 100% y 95% de la densidad máxima corresponden a ensayos realizados bajo condiciones de compactación controlada, conforme a la norma ASTM D1883. Estos resultados permiten evaluar la capacidad estructural del suelo como subrasante.



NOTA

Se presentan la densidad seca máxima (γd máx) y el contenido óptimo de humedad (wopt) obtenidos mediante el ensayo Proctor Estándar (ASTM D698).

Figura 27 — Resultados del ensayo Proctor

Repositorio Institucional - UNAMBA Perú

Tabla 14 — Comparación de valores de CBR de laboratorio y CBR predicho por modelo de Machine Learning

CALICATAS	CBR_LABORATORIO		CBR_PREDICHO	
	CBR_100%	CBR_95%	CBR_100%	CBR_95%
C -01	7.81	6.45	7.47	6.00
C -02	9.43	6.5	9.04	6.21
C -03	11.03	8.45	10.53	7.92

NOTA

La tabla presenta la comparación entre los valores de CBR obtenidos mediante ensayo de laboratorio y los valores estimados por el modelo predictivo basado en Machine Learning para condiciones de compactación al 95% y 100% de la densidad seca máxima. La comparación permite evaluar el grado de ajuste y la capacidad predictiva del modelo desarrollado en la investigación.

Tabla 15 — Diferencia absoluta y porcentual entre CBR de laboratorio y CBR predicho

CALICATAS	Δ CBR		Δ%	
	CBR_100	**CBR_95**	**CBR_100**	**CBR_95**
C -01	0.340	0.45	4.35%	6.98%
C -02	0.390	0.29	4.14%	4.46%
C -03	0.500	0.53	4.53%	6.27%

NOTA

ΔCBR representa la diferencia absoluta entre el valor de CBR obtenido en laboratorio y el valor estimado por el modelo predictivo, mientras que Δ% expresa la variación porcentual respecto al valor experimental. Estos indicadores permiten cuantificar el nivel de error del modelo y evaluar su precisión en la estimación del CBR al 95% y 100% de compactación.

Repositorio Institucional - UNAMBA Perú

5.1.1.2 Rigidez de la subrasante

La rigidez de la subrasante representa la capacidad del suelo para resistir deformaciones elásticas bajo la aplicación repetida de cargas de tránsito. En el presente estudio, dicha propiedad se expresa mediante el módulo resiliente (Mr), parámetro fundamental en el diseño mecanicista–empírico de pavimentos, ya que permite cuantificar el comportamiento estructural del suelo frente a esfuerzos cíclicos. Valores mayores de Mr indican una subrasante con mayor capacidad portante y menor susceptibilidad a deformaciones permanentes, lo cual influye directamente en el desempeño y vida útil del pavimento.

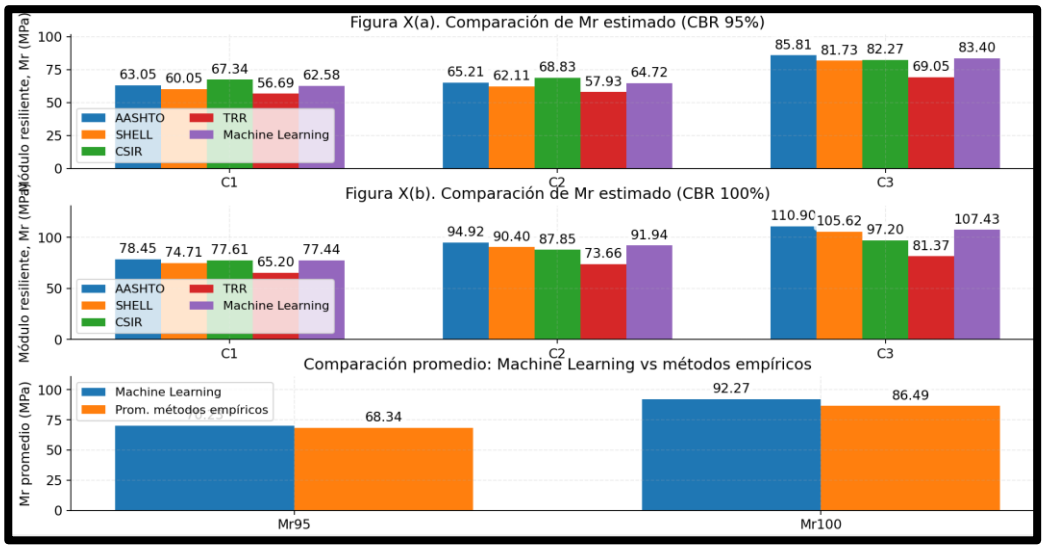
Tabla 16 — Comparación de Mr obtenido por correlaciones empíricas y Machine Learnig

MODULO DE RESILIENCIA	CALICATA	MR_PREDICHO MPA				
		AASHTO	SHELL	CSIR	TRR	MACHINE LEARNING
Mr95%	C1	63.05	60.05	67.34	56.69	62.58
	C2	65.21	62.11	68.83	57.93	64.72
	C3	85.81	81.73	82.27	69.05	83.40
Mr100%	C1	78.45	74.71	77.61	65.20	77.44
	C2	94.92	90.40	87.85	73.66	91.94
	C3	110.90	105.62	97.20	81.37	107.43

NOTA

El módulo resiliente (Mr) fue estimado a partir de los valores de CBR obtenidos al 95% y 100% de compactación, aplicando las correlaciones empíricas propuestas por AASHTO, SHELL, CSIR y TRR, ampliamente utilizadas en el diseño de pavimentos flexibles. Asimismo, se presentan los resultados obtenidos mediante el modelo predictivo basado en aprendizaje automático (Machine Learning), con fines comparativos y para la validación del desempeño del modelo propuesto. Los valores se expresan en MPa.

Repositorio Institucional - UNAMBA Perú



NOTA

El módulo resiliente (Mr) fue estimado a partir de los valores de CBR al 95% y 100% de compactación empleando las correlaciones empíricas propuestas por AASHTO, SHELL, CSIR y TRR, y comparado con los resultados obtenidos mediante el modelo predictivo basado en Machine Learning. La subfigura (a) presenta los resultados para CBR95%, la subfigura (b) para CBR100%, y la subfigura (c) muestra la comparación de los valores promedio entre el modelo de Machine Learning y el promedio de los métodos empíricos. Los valores se expresan en MPa.

Figura 28 — Comparación del módulo resiliente (Mr) estimado por métodos empíricos y Machine Learning

Tabla 17 — Resultados de laboratorio y Machine Learning

DSM	LABORATORIO		MACHINE LEARNING	
	CBR	MR	CBR_ML	MR_ML
95%	6.45	61.7825	6.00	62.58
	6.5	63.52	6.21	64.72
	8.45	79.715	7.92	83.4
100%	7.81	73.9925	7.47	77.44
	9.43	86.7075	9.04	91.94
	11.03	98.7725	10.53	107.43

NOTA

La tabla presenta los valores de capacidad de soporte (CBR) y módulo resiliente (Mr) obtenidos mediante ensayos de laboratorio y su correspondiente estimación mediante el modelo predictivo basado en Machine Learning, para niveles de compactación del 95% y 100%.

Repositorio Institucional - UNAMBA Perú

5.1.2 Análisis de los resultados de variable independiente: Modelo predictivo basado en Machine Learning

5.1.2.1 Modelo predictivo

a) Tipo de modelo

En el presente estudio se implementaron diversos algoritmos de aprendizaje supervisado con el objetivo de desarrollar un modelo predictivo. Los modelos evaluados fueron Random Forest, Extra Trees, Gradient Boosting e HistGradientBoosting, los cuales pertenecen a la familia de métodos de ensamble basados en árboles de decisión.

La selección de estos algoritmos se fundamenta en su capacidad para modelar relaciones no lineales, manejar interacciones entre variables y reducir el riesgo de sobreajuste mediante técnicas de agregación y mejora secuencial. Asimismo, estos modelos son ampliamente utilizados en problemas de regresión dentro del ámbito geotécnico y de ingeniería civil.

Cada algoritmo fue entrenado utilizando validación cruzada con el propósito de evaluar su capacidad de generalización y seleccionar el modelo con mejor desempeño para la predicción del CBR y el posterior del módulo resiliente (Mr).

Tabla 18 — Métricas por tipo de modelo

MODELO		MAE	RMSE
0	GradientBoost	0.263	0.336
1	ExtraTrees	0.623	0.880
2	RandomForest	0.944	1.485
3	HistGB	1.206	1.678

NOTA

Los valores presentados corresponden a las métricas de error obtenidas en la etapa de evaluación del modelo para la predicción del CBR. El MAE (Error Absoluto Medio)

Repositorio Institucional - UNAMBA Perú

representa el error promedio absoluto entre los valores reales y los estimados, mientras que el RMSE (Raíz del Error Cuadrático Medio) penaliza en mayor medida los errores de mayor magnitud. Ambos indicadores se expresan en porcentaje (%) y permiten comparar la precisión relativa de los algoritmos evaluados. Un menor valor de MAE y RMSE indica un mejor desempeño predictivo.

De acuerdo con los valores presentados en la Tabla, el algoritmo Gradient Boosting presenta el mejor desempeño predictivo, al registrar los menores valores tanto de MAE (0.263) como de RMSE (0.336).

Estos resultados indican que, en promedio, el modelo comete el menor error absoluto en la estimación del CBR y, además, muestra la menor penalización frente a errores de mayor magnitud, lo que evidencia una alta precisión y estabilidad en sus predicciones.

El algoritmo Extra Trees ocupa el segundo lugar en desempeño, con errores moderados (MAE = 0.623; RMSE = 0.880), mostrando una capacidad predictiva adecuada pero inferior respecto a Gradient Boosting.

Por su parte, Random Forest y HistGradientBoosting presentan errores progresivamente mayores, siendo HistGB el modelo con menor precisión dentro del conjunto evaluado.

En consecuencia, considerando que el criterio de selección se basa en la minimización del error de predicción, se determina que el algoritmo Gradient Boosting es el modelo óptimo para la estimación del CBR y su posterior aplicación en el cálculo del módulo resiliente (Mr).

5.1.2.2 Desempeño del modelo de Machine Learning

El desempeño del modelo predictivo basado en Machine Learning fue evaluado con el propósito de determinar su capacidad de estimación y su confiabilidad en la predicción del CBR y del módulo resiliente (Mr). Para ello, se emplearon métricas estadísticas que permiten analizar tanto el grado de ajuste del modelo respecto a los datos observados como su estabilidad frente a diferentes particiones del conjunto de datos.

Repositorio Institucional - UNAMBA Perú

El análisis del desempeño constituye una etapa fundamental dentro del proceso de modelamiento, ya que permite verificar que el algoritmo no solo reproduzca adecuadamente el comportamiento de la variable objetivo, sino que también mantenga consistencia y capacidad de generalización al aplicarse a nuevos datos.

Tabla 19 — Desempeño del modelo predictivo

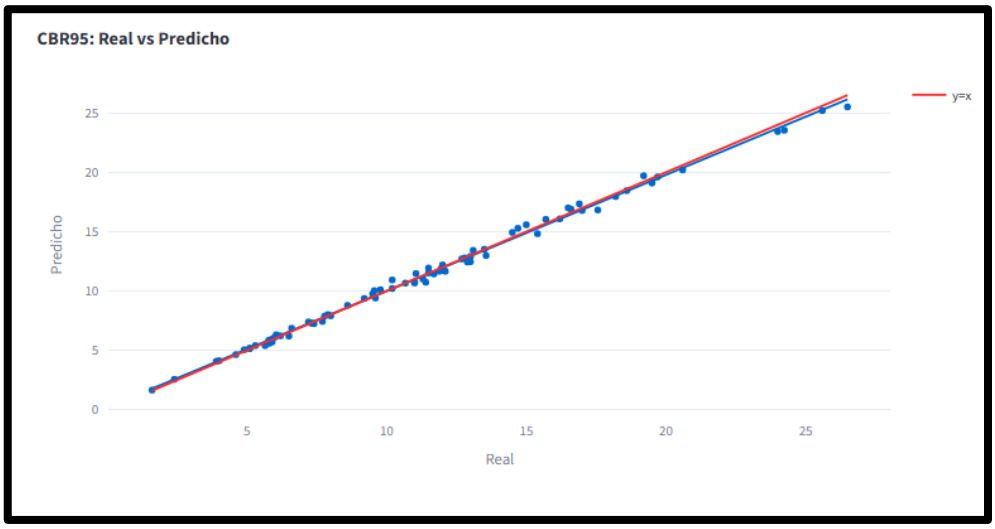
PREDICCIÓN		CV_R^2	R^2
0	CBR_95	0.7367	0.996
1	CBR_100	0.673	0.998
3	MR_95	0.785	0.922
4	MR_100	0.985	0.912

NOTA

CV_R^2 corresponde al R^2 promedio de la validación cruzada K-Fold y R^2 al coeficiente de determinación del conjunto de prueba. Valores cercanos a 1 indican mayor capacidad predictiva del modelo para estimar CBR y Mr.

Los resultados muestran que el modelo presenta valores elevados de coeficiente de determinación en el conjunto de prueba (R^2 = 0.996 para CBR_95 y R^2 = 0.998 para CBR_100), lo que indica una alta capacidad de ajuste a los datos observados.

Repositorio Institucional - UNAMBA Perú

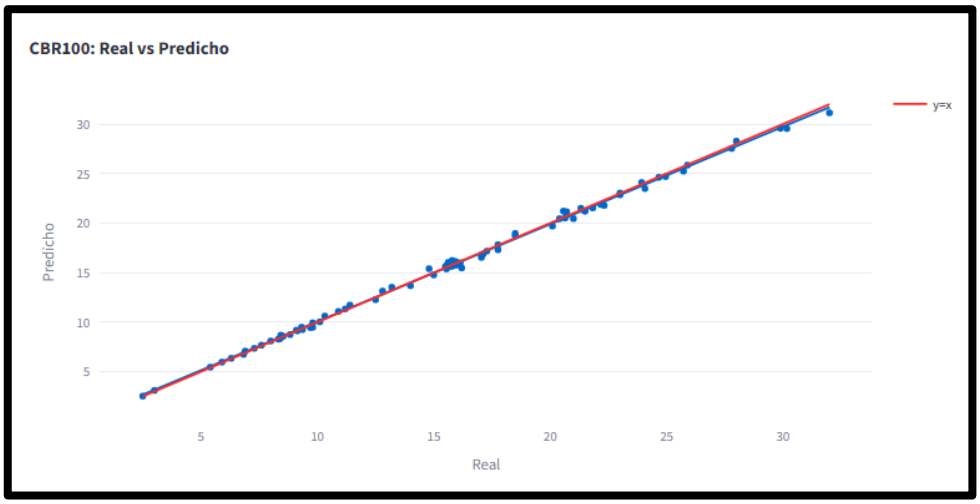


NOTA

La figura muestra la comparación entre los valores reales y los valores predichos de CBR_95 obtenidos mediante el modelo de Machine Learning. La línea roja representa la recta ideal (y = x), que indica una predicción perfecta. La proximidad de los puntos a dicha línea evidencia un alto grado de ajuste y precisión del modelo, confirmando su adecuada capacidad predictiva para la estimación del CBR.

Figura 29 — Comparación entre valores reales y predichos de CBR_95

Repositorio Institucional - UNAMBA Perú



NOTA

La figura presenta la relación entre los valores reales y los valores estimados de CBR_100 mediante el modelo de Machine Learning. La línea roja (y = x) representa el ajuste ideal, donde los valores predichos coinciden exactamente con los observados. La concentración de los puntos cercanos a esta línea indica un alto nivel de precisión y un adecuado ajuste del modelo, evidenciando su confiabilidad en la estimación del CBR al 100% de compactación.

Figura 30 — Comparación entre valores reales y predichos de CBR_100

5.1.3 Análisis estadístico descriptivo

5.1.3.1 Variable dependiente (Capacidad de soporte del suelo)

- **1º. Clasificación de variables**

  La capacidad de soporte del suelo (CBR) es una variable cuantitativa continua, medida en porcentaje (%), que representa la resistencia relativa del suelo frente a la penetración estándar.

  Desde el enfoque estadístico, es una variable de razón, ya que posee cero absolutos y permite realizar operaciones matemáticas como media, varianza y desviación estándar.

Repositorio Institucional - UNAMBA Perú

- **2º. Cálculo de medidas estadística**

Tabla 20 — Estadística de capacidad de soporte del suelo

Descriptivos				
			Estadístico	**Desv. Error**
Capacidad de soporte del suelo.	**Media**		8.07	0.4849
	95% de intervalo de confianza para la media	**Límite inferior**	7.0028	
		Límite superior	9.1372	
	Media recortada al 5%		8.07	
	Mediana		7.865	
	Varianza		2.8211	
	Desv. Estándar		1.6796	
	Mínimo		6.0	
	Máximo		11.03	
	Rango		5.03	
	Rango Inter cuartil		2.65	
	Asimetría		0.4752	
	Curtosis		-0.8428	

NOTA

La tabla presenta los estadísticos descriptivos correspondientes a la capacidad de soporte del suelo, calculados a partir de los ensayos de laboratorio realizados en las muestras de subrasante. Se incluyen medidas de tendencia central (media y mediana), medidas de dispersión (varianza, desviación estándar, rango e intervalo intercuartílico), así como indicadores de forma de distribución (asimetría y curtosis). El intervalo de confianza del 95% para la media fue estimado mediante la distribución t de Student.

La variable presenta valores entre 6.00% y 11.03%, mostrando una media de 8.07% con una desviación estándar de 1.68%, lo que indica una variabilidad moderada en la capacidad de soporte del suelo. La asimetría positiva (0.4752) evidencia una ligera concentración de valores hacia el rango inferior, mientras que la curtosis negativa (-0.8428) sugiere una distribución ligeramente más plana que la normal.

Repositorio Institucional - UNAMBA Perú

En conclusión, el suelo presenta una capacidad de soporte promedio de 8.07%, con variabilidad moderada y comportamiento relativamente uniforme, lo que permite considerarlo representativo para el análisis del modelo predictivo.

5.1.3.2 Variable dependiente (Rigidez de la subrasante)

- **1º. Clasificación de variables**

  La rigidez de la subrasante, expresada mediante el Módulo Resiliente (Mr) en MPa, es una:

  - Variable cuantitativa continua
  - Escala de medición: razón
  - Unidad: MPa
  - Naturaleza: mecánica – comportamiento elástico del suelo
  - Método de estimación: AASHTO, SHELL, CSIR, TRR y Machine Learning

- **2º. Cálculo de medidas estadística**

Tabla 21 — Estadística de la rigidez de la subrasante

Descriptivos			Estadístico	Desv. Error
Rigidez de la subrasante (Mr)	**Media**		78.182	2.764
	95% de intervalo de confianza para la media	**Límite inferior**	72.528	
		Límite superior	83.836	
	Media recortada al 5%		77.7814	
	Mediana		77.525	
	Varianza		15.142	
	Desv. Estándar		15.142	
	Mínimo		56.69	
	Máximo		110.9	
	Rango		54.21	
	Rango Inter cuartil		22.1375	
	Asimetría		0.563	
	Curtosis		-0.503	

NOTA

La tabla presenta los estadísticos descriptivos del módulo resiliente (Mr) expresado en MPa, considerando los valores obtenidos mediante los diferentes métodos de estimación y niveles de compactación evaluados en la subrasante.

Repositorio Institucional - UNAMBA Perú

Los resultados muestran que la rigidez promedio de la subrasante es de 78.18 MPa, con una variabilidad moderada (desviación estándar de 15.14 MPa). El intervalo de confianza al 95% indica que la media poblacional se encuentra entre 72.53 MPa y 83.84 MPa. La ligera asimetría positiva (0.563) evidencia mayor concentración de valores por debajo del promedio, mientras que la curtosis negativa (-0.503) sugiere una distribución ligeramente más plana que la normal.

5.1.3.3 Variable independiente (Modelo predictivo)

- **1º. Clasificación de variables**

  La variable independiente corresponde al modelo predictivo basado en técnicas de Machine Learning, utilizada para estimar la respuesta mecánica del suelo.

  - Variable cualitativa
  - De tipo nominal
  - Las categorías evaluadas fueron:
  - Gradient Boosting
  - Extra Trees
  - Random Forest
  - HistGradient Boosting

  Dado que la variable representa diferentes enfoques algorítmicos, su análisis comparativo se realiza a través de indicadores cuantitativos de desempeño predictivo.

- **2º. Cálculo de medidas estadística**

Tabla 22 — Estadística del modelo predictivo

Descriptivos				
			Estadístico	**Desv. Error**
Modelo predictivo	Media		0.9269	0.1812
	95% de intervalo de confianza para la media	Límite inferior	0.4984	
		Límite superior	1.3553	
	Media recortada al 5%		0.9269	
	Mediana		0.9120	

Repositorio Institucional - UNAMBA Perú

	Varianza	0.2626	
	Desv. Estándar	0.5125	
	Mínimo	0.2630	
	Máximo	1.6780	
	Rango	1.4150	
	Rango Inter cuartil	0.7245	
	Asimetría	0.1307	
	Curtosis	-1.1909	

NOTA

La tabla presenta los estadísticos descriptivos correspondientes a las métricas de error (MAE y RMSE) de los modelos predictivos evaluados para la estimación del módulo resiliente (Mr). El error estándar corresponde a la media muestral calculada a un nivel de confianza del 95%.

Los resultados muestran que el error promedio del modelo predictivo es de 0.9269, con una desviación estándar de 0.5125, lo que indica una variabilidad moderada entre los algoritmos evaluados. El intervalo de confianza al 95% (0.4984 – 1.3553) evidencia estabilidad en la estimación del error. La asimetría cercana a cero (0.1307) indica distribución aproximadamente simétrica, mientras que la curtosis negativa (-1.1909) sugiere una distribución más plana que la normal.

5.1.3.4 Variable independiente (Desempeño del modelo)

- **1º. Clasificación de variables**
  - Variable cuantitativa continua
  - Escala de razón
  - Indicadores:
    - ✓ Coeficiente de determinación (R^2)
    - ✓ Coeficiente de determinación validado (CV_R^2)

Estos indicadores miden la capacidad explicativa y predictiva del modelo.

Repositorio Institucional - UNAMBA Perú

- **2º. Cálculo de medidas estadística**

Tabla 23 — Estadística del desempeño del modelo

Descriptivos				
			Estadístico	**Desv. Error**
Desempeño del modelo	Media		0.8760	0.0450
	95% de intervalo de confianza para la media	Límite inferior	0.7695	
		Límite superior	0.9824	
	Media recortada al 5%		0.8760	
	Mediana		0.9170	
	Varianza		0.0162	
	Desv. Estándar		0.1273	
	Mínimo		0.6730	
	Máximo		0.9980	
	Rango		0.3250	
	Rango Inter cuartil		0.2148	
	Asimetría		-0.6300	
	Curtosis		-1.3530	

NOTA

La tabla presenta el análisis descriptivo del desempeño del modelo predictivo considerando los valores de R^2 y CV_R^2 obtenidos en las diferentes estimaciones. Elaboración propia.

El desempeño global del modelo presenta una media de 0.8761, con una desviación estándar de 0.1180, lo que indica una alta capacidad predictiva y baja variabilidad entre las estimaciones. Los valores de R^2 y CV_R^2 oscilan entre 0.673 y 0.998, evidenciando que el modelo explica adecuadamente la variabilidad de los datos. La asimetría leve sugiere una distribución relativamente equilibrada, confirmando estabilidad en el rendimiento del modelo.

Repositorio Institucional - UNAMBA Perú

5.1.4 Prueba de normalidad por cada variable

a) Variable independiente: Machine Learning

Para evaluar la distribución de las métricas de desempeño del modelo (MAE, RMSE, CV_R^2 y R^2), se aplicó la prueba de normalidad de Shapiro-Wilk, debido al tamaño reducido de la muestra ($n < 50$).

- **Hipótesis**
  - H_0: Los datos presentan distribución normal.
  - H_1: Los datos no presentan distribución normal.

- **Decisión**

Si el valor de Sig. (p-valor) $< 0.05 \rightarrow$ se rechaza H_0 y se acepta H_1.

Si el valor de Sig. (p-valor) $\geq 0.05 \rightarrow$ se acepta H_0 y se rechaza H_1.

Tabla 24 — Prueba de normalidad Machine Learning

Indicadores		Shapiro-Wilk		
		Estadístico	**gl-**	**Sig.**
(Machine Learning)	MAE, RMSE, CV_R^2 R^2	0.9499	16	0.4886

NOTA

La tabla presenta los resultados de la prueba de normalidad Shapiro-Wilk aplicada a los indicadores MAE, RMSE, CV_R^2 y R^2 del modelo predictivo. Datos procesados en Python mediante la librería SciPy (2025). Elaboración propia.

En la tabla se observa que el estadístico de Shapiro-Wilk es 0.9499 con un valor de significancia $p = 0.4886$, el cual es mayor que 0.05. Por lo tanto, no se rechaza la hipótesis nula, concluyendo que los indicadores de desempeño del modelo predictivo presentan distribución normal. Esto permite aplicar procedimientos estadísticos paramétricos en el análisis comparativo del modelo.

Repositorio Institucional - UNAMBA Perú

b) Variable dependiente: Respuesta mecánica

Con la finalidad de analizar la distribución estadística de los datos correspondientes a la respuesta mecánica del suelo (capacidad de soporte – CBR y rigidez de la subrasante – Mr), se aplicó la prueba de normalidad de Shapiro-Wilk, debido a que el tamaño de muestra es inferior a 50 observaciones, condición bajo la cual dicha prueba presenta mayor confiabilidad estadística.

- **Hipótesis**
  - H_0: Los datos presentan distribución normal.
  - H_1: Los datos no presentan distribución normal.

- **Decisión**
  - Si el valor de Sig. (p-valor) < 0.05 → se rechaza H_0 y se acepta H_1.
  - Si el valor de Sig. (p-valor) ≥ 0.05 → se acepta H_0 y se rechaza H_1.

Tabla 25 — Prueba de normalidad Machine Learning

Indicadores		Shapiro-Wilk		
		Estadístico	gl-	Sig.
Respuesta mecánica	CBR Mr	0.8552	42	0.0001

NOTA

La tabla presenta los resultados de la prueba de normalidad Shapiro-Wilk aplicada a la variable respuesta mecánica, considerando de manera conjunta los valores de capacidad de soporte (CBR) y módulo resiliente (Mr). Datos procesados en Python mediante la librería SciPy (2025).

El resultado de la prueba de Shapiro-Wilk muestra un estadístico W = 0.8552 con gl = 42 y un valor de significancia p = 0.0001, lo que indica que los datos de la respuesta mecánica (CBR y Mr) no siguen una distribución normal. Esta desviación se debe principalmente a la diferencia de magnitud entre los valores de CBR (bajos porcentajes) y los valores de Mr (expresados en MPa y de mayor escala), generando una distribución asimétrica y dispersa. En consecuencia, la variable presenta heterogeneidad estadística al analizarse de manera conjunta.

Repositorio Institucional - UNAMBA Perú

- **Conclusión de la prueba de normalidad**

  Los resultados de la prueba de Shapiro-Wilk evidencian que la variable desempeño del modelo predictivo presenta distribución normal ($p > 0.05$), mientras que la variable respuesta mecánica (CBR y Mr) no cumple con el supuesto de normalidad ($p < 0.05$). En consecuencia, para los análisis estadísticos posteriores se consideró el uso de métodos no paramétricos, a fin de garantizar la validez y confiabilidad de los resultados obtenidos.

5.2 Contrastación de hipótesis

5.2.1 Contrastación de hipótesis general

La contrastación de la hipótesis general se realizó mediante el análisis de correlación no paramétrica (Spearman), considerando el nivel de significancia $\alpha = 0.05$.

- **H_0:**

  El modelo predictivo basado en Machine Learning no presenta un desempeño adecuado para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025.

- **H_1:**

  El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025.

  Dado que la respuesta mecánica no es normal:
  Se usará Correlación de Spearman.

  Criterio de decisión:

  - Si $p < 0.05 \rightarrow$ Se rechaza H_0
  - Si $p \geq 0.05 \rightarrow$ No se rechaza H_0

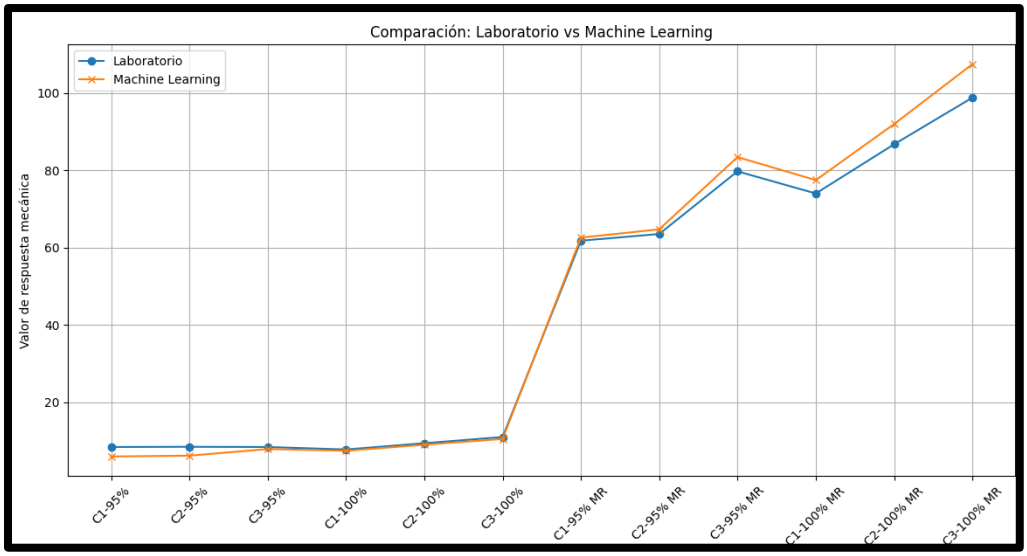
Repositorio Institucional - UNAMBA Perú

Tabla 26 — Prueba de correlación de Spearman para la hipótesis general

Hipótesis General		Machine Learning	Respuesta mecánica
Respuesta mecánica	Correlación de Spearman	0.9562	1
	Sig. (bilateral)		0.000
	N	12	12
Machine Learning	Correlación de Spearman	1	0.9562
	Sig. (bilateral)	0.000	
	N	12	12

NOTA

Coeficiente de correlación no paramétrica de Spearman ($\rho = 0.956$; $p < 0.05$), indicando una relación positiva muy fuerte y estadísticamente significativa entre ambas variables. Datos procesados en Python (SciPy).



NOTA

La figura muestra la comparación de los valores de capacidad de soporte (CBR) obtenidos mediante ensayos de laboratorio y los valores del módulo resiliente (Mr) determinados mediante correlaciones empíricas, con los correspondientes valores estimados mediante el modelo predictivo basado en Machine Learning, para niveles de compactación del 95 % y 100 %.

Figura 31 — Comparación de la respuesta mecánica

Repositorio Institucional - UNAMBA Perú

Los resultados de la prueba de correlación de Spearman evidenciaron un coeficiente de $\rho = 0.956$ con un nivel de significancia $p < 0.05$, lo que indica la existencia de una relación positiva muy fuerte y estadísticamente significativa entre la respuesta mecánica del suelo obtenida en laboratorio y los valores estimados mediante el modelo predictivo basado en Machine Learning.

En consecuencia, se rechaza la hipótesis nula (H_0) y se acepta la hipótesis alternativa (H_1), concluyendo que el modelo predictivo presenta capacidad significativa para estimar la respuesta mecánica del suelo, demostrando coherencia y consistencia en la estimación del CBR y del módulo resiliente (Mr).

5.2.2 Contrastación de hipótesis especificas

a) Hipótesis específica 1

- **H_0:**

  El modelo predictivo basado en Machine Learning no presenta un desempeño adecuado para estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

- **H_1:**

  El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

- **Criterio de decisión:**
  - Si $p < 0.05 \rightarrow$ Se rechaza H_0
  - Si $p \geq 0.05 \rightarrow$ No se rechaza H_0

Repositorio Institucional - UNAMBA Perú

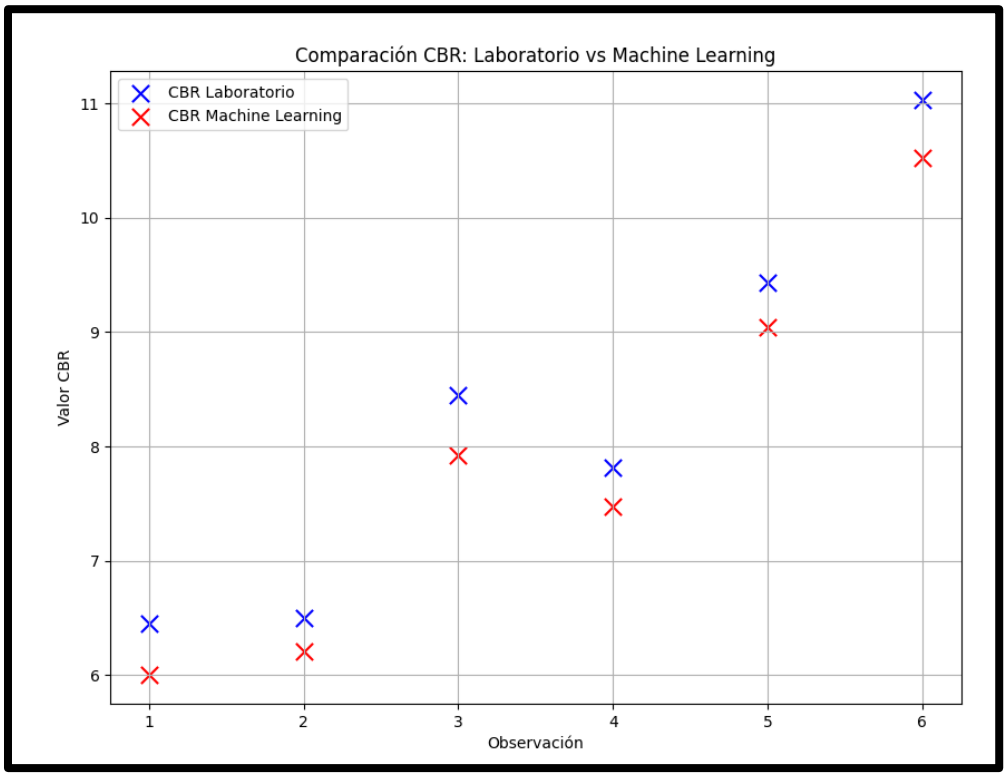
Tabla 27 — Prueba de correlación Spearman para la hipótesis especifica 1

Hipótesis especifica 1		CBR Machine Learnig	CBR Laboratorio
CBR Laboratorio	Correlación de Spearman	0.829	1
	Sig. (bilateral)	0.042	
	N	6	6
CBR Machine Learnig	Correlación de Spearman	1	0.829
	Sig. (bilateral)		0.042
	N	6	6
NOTA La tabla presenta el coeficiente de correlación de Spearman entre los valores de capacidad de soporte del suelo (CBR) obtenidos en laboratorio y los estimados mediante el modelo predictivo basado en Machine Learning. Se consideró un nivel de significancia $\alpha = 0.05$ y un tamaño de muestra $N = 6$.			

Los resultados de la prueba de correlación de Spearman evidencian un coeficiente $\rho = 0.829$, lo que indica una correlación positiva fuerte entre el CBR obtenido en laboratorio y el estimado mediante Machine Learning. Asimismo, el valor de significancia bilateral es $p = 0.042$, el cual es menor al nivel de significancia establecido ($\alpha = 0.05$), demostrando que la relación observada es estadísticamente significativa.

En consecuencia, se rechaza la hipótesis nula (H_0) y se acepta la hipótesis alternativa (H_1), concluyendo que el modelo predictivo presenta una capacidad significativa para estimar la capacidad de soporte del suelo (CBR).

Repositorio Institucional - UNAMBA Perú



NOTA

La figura muestra la comparación de los valores de capacidad de soporte del suelo (CBR) obtenidos mediante ensayos de laboratorio y los estimados por el modelo predictivo basado en Machine Learning, para niveles de compactación del 95% y 100%.

Figura 32 — Comparación de la capacidad de soporte del suelo (CBR)

b) Hipótesis específica 2

- **H_0:**

  El modelo predictivo basado en Machine Learning no presenta un desempeño adecuado para estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

- **H_1:**

  El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.

- **Criterio de decisión:**
  - Si $p < 0.05 \rightarrow$ Se rechaza H_0
  - Si $p \geq 0.05 \rightarrow$ No se rechaza H_0

Repositorio Institucional - UNAMBA Perú

Tabla 28 — Prueba de correlación Spearman para la hipótesis especifica 2

Hipótesis especifica 2		Mr Machine Learnig	Mr Laboratorio
Mr Laboratorio	Correlación de Spearman	0.893	1
	Sig. (bilateral)	0.005	
	N	6	6
Mr Machine earnig	Correlación de Spearman	1	0.893
	Sig. (bilateral)		0.005
	N	6	6

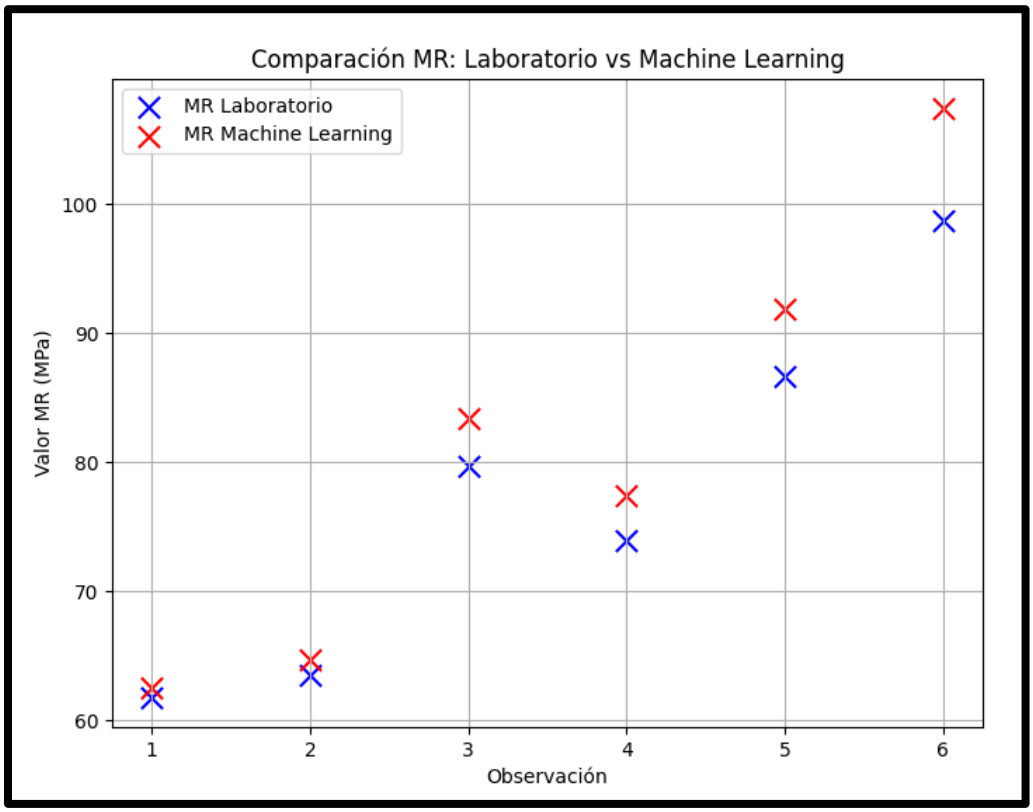
NOTA

La tabla presenta el coeficiente de correlación de Spearman entre los valores del módulo resiliente (Mr) obtenidos mediante fórmulas empíricas y los estimados mediante el modelo predictivo basado en Machine Learning. Se consideró un nivel de significancia $\alpha = 0.05$ y un tamaño de muestra $N = 6$.

Los resultados muestran un coeficiente de correlación de Spearman $\rho = 0.893$, lo que indica una correlación positiva fuerte entre el módulo resiliente obtenido en laboratorio y el estimado mediante Machine Learning. Asimismo, el valor de significancia bilateral es $p = 0.005$, siendo menor al nivel de significancia establecido ($\alpha = 0.05$), lo que demuestra que la relación es estadísticamente significativa.

En consecuencia, se rechaza la hipótesis nula (H_0) y se acepta la hipótesis alternativa (H_1), concluyendo que el modelo predictivo presenta capacidad significativa para estimar la rigidez de la subrasante (Mr).

Repositorio Institucional - UNAMBA Perú



NOTA

La figura presenta la comparación de los valores del módulo resiliente (Mr) obtenidos en laboratorio y los estimados mediante el modelo predictivo, evidenciando el comportamiento de la rigidez de la subrasante en diferentes niveles de compactación.

Figura 33 — Comparación del Mr

5.3 Discusión

El presente estudio tuvo como objetivo desarrollar y evaluar un modelo predictivo basado en Machine Learning para estimar la respuesta mecánica del suelo de subrasante en el distrito de Abancay, Apurímac – 2025. Los resultados obtenidos evidencian un comportamiento altamente consistente entre los valores experimentales de laboratorio y las estimaciones generadas por el modelo.

En la contrastación de la hipótesis general, la prueba de correlación de Spearman mostró un coeficiente $\rho = 0.956$ con un nivel de significancia $p < 0.05$, lo que indica una relación positiva muy fuerte y estadísticamente significativa entre ambas variables. Este resultado confirma que el modelo predictivo no solo mantiene la tendencia de los datos experimentales, sino que reproduce de manera coherente el comportamiento mecánico del suelo.

Repositorio Institucional - UNAMBA Perú

Desde el punto de vista estadístico, los valores de R^2 alcanzaron hasta 0.998 en la predicción del CBR al 100%, y valores superiores a 0.92 en el módulo resiliente (Mr), lo cual evidencia un alto grado de ajuste del modelo. Asimismo, los errores absolutos medios (MAE = 0.263 en el mejor modelo) y los errores cuadráticos medios (RMSE = 0.336) demuestran que la diferencia entre valores reales y estimados es mínima, validando la precisión del sistema predictivo.

Estos hallazgos guardan concordancia con estudios internacionales como el de Gao (2024), quien reportó un R^2 de 0.94 en la predicción de parámetros geotécnicos mediante redes neuronales. De igual manera, Khatti y Singh Grover (2021) obtuvieron un R^2 de 0.9556 utilizando modelos de regresión de procesos gaussianos (GPR) para la predicción de propiedades geotécnicas. En comparación, los resultados obtenidos en Abancay presentan niveles de ajuste similares e incluso superiores, lo que demuestra que el aprendizaje automático puede adaptarse eficazmente a contextos locales con bases de datos adecuadamente estructuradas.

En el ámbito nacional, los resultados son consistentes con Valderrama (2020), quien desarrolló modelos ANN para la predicción del CBR y obtuvo errores favorables en comparación con ensayos de laboratorio. Asimismo, Manrique (2019) reportó coeficientes de correlación entre 0.7 y 0.9 en la predicción de propiedades mecánicas del suelo, valores comparables con los obtenidos en la presente investigación ($\rho = 0.829$ para CBR y $\rho = 0.893$ para Mr). Esto evidencia que el desempeño del modelo desarrollado se encuentra dentro de los rangos considerados aceptables y confiables en investigaciones similares.

Es importante destacar que el modelo mostró un mejor comportamiento en la estimación del módulo resiliente (Mr) que en el CBR. Esto puede explicarse debido a que el Mr representa una propiedad elástica del suelo bajo cargas repetidas, la cual presenta una relación más continua y menos afectada por la variabilidad experimental en comparación con el ensayo CBR, que puede verse influenciado por factores como el control de humedad, compactación y condiciones del molde. Por lo tanto, la ligera diferencia en los coeficientes de correlación entre ambas variables responde a la naturaleza intrínseca de los ensayos.

Desde el punto de vista ingenieril, la aplicación de Machine Learning en la predicción de la respuesta mecánica del suelo permite reducir tiempos de ensayo, optimizar costos y

Repositorio Institucional - UNAMBA Perú

facilitar la toma de decisiones en etapas preliminares de diseño. En contextos como el distrito de Abancay, donde las condiciones geotécnicas pueden variar significativamente, contar con un modelo predictivo confiable representa una herramienta estratégica para proyectos viales.

Asimismo, los resultados obtenidos refuerzan la tendencia actual en la ingeniería geotécnica hacia la digitalización y el uso de modelos predictivos basados en datos, alineándose con los enfoques mecanicistas modernos de diseño de pavimentos, donde el módulo resiliente constituye un parámetro fundamental.

En conjunto, los resultados demuestran que el modelo desarrollado cumple con los objetivos planteados, validando la hipótesis general y específicas, y confirmando que el Machine Learning constituye una herramienta técnicamente viable, estadísticamente válida y aplicable para la estimación de la respuesta mecánica del suelo a nivel de subrasante.

Repositorio Institucional - UNAMBA Perú

CAPÍTULO VI
CONCLUSIONES Y RECOMENDACIONES

6.1 Conclusiones

La presente investigación se desarrolló a partir de la integración de información proveniente de laboratorio, informes, expedientes técnicos u otros, cuyos registros permitieron consolidar una base de datos de 90 registros experimentales utilizados para el entrenamiento y validación del modelo predictivo. Estos datos incluyeron parámetros geotécnicos asociados a la capacidad de soporte (CBR) y al módulo resiliente (Mr), correspondientes a diferentes niveles de compactación. Posteriormente, se realizaron ensayos específicos en 3 calicatas, cuyos resultados fueron empleados para contrastar y validar el desempeño del modelo. Esta integración de datos históricos y ensayos experimentales garantizó consistencia técnica, robustez estadística y confiabilidad en el proceso de aprendizaje del algoritmo.

En relación con el objetivo general, se desarrolló y evaluó satisfactoriamente un modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, empleando una base de datos geotécnica conformada por registros experimentales de laboratorio. Los resultados obtenidos demostraron que el modelo reproduce de manera confiable el comportamiento mecánico del suelo, validando la hipótesis general y evidenciando que esta herramienta constituye una alternativa técnicamente viable para apoyar la caracterización geotécnica y el diseño preliminar de pavimentos.

Respecto al primer objetivo específico, referido a la estimación de la capacidad de soporte (CBR), el modelo presentó un desempeño adecuado al reproducir resultados consistentes con los ensayos de laboratorio, evidenciando una alta concordancia entre los valores experimentales y los estimados. Este comportamiento demuestra que el modelo puede emplearse como una herramienta de apoyo para realizar estimaciones preliminares del CBR, optimizando el proceso de evaluación de la subrasante sin sustituir los ensayos geotécnicos requeridos para el diseño definitivo

Repositorio Institucional - UNAMBA Perú

En cuanto al segundo objetivo específico, referido a la estimación del módulo resiliente (Mr), el modelo obtuvo un desempeño satisfactorio, representando adecuadamente la rigidez de la subrasante a partir de los parámetros geotécnicos de entrada. Los resultados evidenciaron que el aprendizaje automático constituye una metodología capaz de modelar relaciones complejas entre las propiedades físicas del suelo y su respuesta mecánica, aportando una alternativa confiable para estudios relacionados con el diseño mecanicista de pavimentos.

Finalmente, la investigación confirma que la incorporación de técnicas de Machine Learning en la ingeniería geotécnica representa una herramienta complementaria para el análisis de suelos de subrasante, al contribuir con estimaciones confiables que favorecen la optimización del tiempo de procesamiento de información, el aprovechamiento de bases de datos geotécnicas existentes y el apoyo a la toma de decisiones durante las etapas iniciales de planificación y diseño de infraestructura vial. Asimismo, los resultados obtenidos constituyen un antecedente para futuras aplicaciones de inteligencia artificial en el ámbito de la ingeniería civil en la región Apurímac.

6.2 Recomendaciones

Se recomienda evaluar el desempeño del modelo predictivo desarrollado en otros proyectos viales ubicados en diferentes provincias de la región Apurímac, considerando suelos con características geotécnicas distintas a las analizadas en la presente investigación. Esta aplicación permitirá verificar la capacidad de generalización del modelo frente a nuevos escenarios y establecer el grado de confiabilidad de sus estimaciones cuando se emplee en diferentes condiciones de subrasante. Asimismo, contribuirá a determinar las limitaciones y fortalezas del modelo para su utilización como herramienta de apoyo en estudios geotécnicos relacionados con el diseño y rehabilitación de pavimentos.

Ampliar la base de datos utilizada para el entrenamiento del modelo mediante la incorporación de nuevos registros geotécnicos provenientes de campañas de exploración realizadas en diferentes épocas del año y en diversas unidades geológicas de la región. Una mayor cantidad y diversidad de información permitirá mejorar el proceso de aprendizaje del algoritmo, incrementar la representatividad de los datos y fortalecer la capacidad predictiva del modelo frente a condiciones geotécnicas variables. De esta manera, se

Repositorio Institucional - UNAMBA Perú

obtendrán estimaciones más robustas y con mayor potencial de aplicación en futuros proyectos de infraestructura vial.

Comparar el desempeño del modelo desarrollado con otros algoritmos de aprendizaje automático empleados en problemas de regresión, tales como Redes Neuronales Artificiales, XGBoost, CatBoost o Support Vector Machine, utilizando la misma base de datos y los mismos indicadores de evaluación aplicados en esta investigación. Esta comparación permitirá identificar las ventajas y limitaciones de cada algoritmo, así como seleccionar la alternativa que ofrezca un mejor equilibrio entre precisión, estabilidad y capacidad de generalización para la estimación del CBR y del módulo resiliente.

Evaluar el comportamiento del modelo predictivo empleando bases de datos obtenidas en diferentes etapas de ejecución de proyectos viales, incluyendo investigaciones geotécnicas realizadas durante la fase de diseño y registros provenientes del control de calidad en obra. Esto permitirá verificar la estabilidad de las predicciones del modelo frente a información generada en diferentes contextos de aplicación y fortalecer su utilidad como herramienta complementaria para la toma de decisiones en ingeniería de pavimentos.

Finalmente se recomienda, promover la incorporación de técnicas de Machine Learning como herramientas complementarias en los estudios geotécnicos para proyectos de infraestructura vial, especialmente durante las etapas preliminares de caracterización de la subrasante. Su aplicación puede contribuir a optimizar el procesamiento y análisis de información geotécnica disponible, reducir los tiempos requeridos para obtener estimaciones iniciales de parámetros mecánicos y apoyar la toma de decisiones técnicas. No obstante, las predicciones generadas por estos modelos deben emplearse como complemento del criterio profesional del ingeniero y de los ensayos geotécnicos establecidos por las normas técnicas vigentes.

Repositorio Institucional - UNAMBA Perú

REFERENCIAS BIBLIOGRÁFICAS

ABDELKADER, Hassan A. M.; HUSSEIN, Mohamed M. A. y YE, Haiwang. 2021. Influence of waste marble dust on the improvement of expansive clay soils. *Advances in Civil Engineering* [en línea]. Vol. 2021, art. 3192122. ISSN 1687-8094.

AHMED, Abrar; EL-EMAM, Magdi; AHMAD, Naveed y ATTOM, Mousa. 2024. Stabilization of pavement subgrade clay soil using sugarcane ash and lime. *Geosciences* [en línea]. Vol. 14, n.º 6, art. 151. ISSN 2076-3263.

AMERICAN ASSOCIATION OF STATE HIGHWAY AND TRANSPORTATION OFFICIALS (AASHTO). 1993. *AASHTO Guide for Design of Pavement Structures, 1993*. Washington, D. C.: AASHTO. ISBN 1-56051-055-2.

AMERICAN ASSOCIATION OF STATE HIGHWAY AND TRANSPORTATION OFFICIALS (AASHTO). 2008. *Mechanistic-Empirical Pavement Design Guide: A Manual of Practice*. Edición provisional. Washington, D. C.: AASHTO. ISBN 978-1-56051-423-7.

AMERICAN ASSOCIATION OF STATE HIGHWAY AND TRANSPORTATION OFFICIALS (AASHTO). 2021. ***AASHTO T 307-99 (2021):*** *Standard Method of Test for Determining the Resilient Modulus of Soils and Aggregate Materials*. Washington, D. C.: AASHTO.

AHMAD, Umair; JAMIL, Irfan; JAMAL, Hamza; KHAN, Muhammad Bilal; ACCOUCHE, Omar y AZAB, Mourad. 2025. Using machine learning paradigm to predict CBR value of treated expansive soil. *Geotechnical Research* [en línea]. Vol. 12, n.º 4, pp. 186–201.

ASTM INTERNATIONAL. 2017. *ASTM D2487-17e1: Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System)*. West Conshohocken, Pennsylvania: ASTM International.

ASTM INTERNATIONAL. 2019. *ASTM D2216-19: Standard Test Methods for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass*. West Conshohocken, Pennsylvania: ASTM International.

ASTM INTERNATIONAL. 2021a. ***ASTM D698-12(2021):*** *Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Standard Effort (12,400 ft-lbf/ft³ [600 kN-m/m³])*. West Conshohocken, Pennsylvania: ASTM International.

ASTM INTERNATIONAL. 2021b. ***ASTM D1557-12(2021):*** *Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Modified Effort (56,000 ft-lbf/ft³ [2,700 kN-m/m³])*. West Conshohocken, Pennsylvania: ASTM International. Disponible en: https://doi.org/10.1520/D1557-12R21

Repositorio Institucional - UNAMBA Perú

ASTM INTERNATIONAL. 2021c. ***ASTM D1883-21:*** *Standard Test Method for California Bearing Ratio (CBR) of Laboratory-Compacted Soils*. West Conshohocken, Pennsylvania: ASTM International. Disponible en: https://doi.org/10.1520/D1883-21

BERNAL TORRES, César Augusto. 2016. *Metodología de la investigación: administración, economía, humanidades y ciencias sociales*. 4.ª ed. Bogotá: Pearson Educación. ISBN 978-958-699-309-8.

BISHOP, Christopher M. 2006. *Pattern Recognition and Machine Learning*. Nueva York: Springer. ISBN 978-0-387-31073-2.

BOWLES, Joseph E. 1997. *Foundation Analysis and Design*. 5.ª ed. Nueva York: McGraw-Hill. ISBN 978-0-07-912247-6.

BREIMAN, Leo. 2001. Random forests. *Machine Learning* [en línea]. Vol. 45, pp. 5–32. Disponible en: https://doi.org/10.1023/A:1010933404324

CONSEJO NACIONAL DE CIENCIA, TECNOLOGÍA E INNOVACIÓN TECNOLÓGICA (CONCYTEC). 2020. *Código Nacional de la Integridad Científica*. Lima: CONCYTEC.

DAS, Braja M. 2011. *Fundamentos de ingeniería de cimentaciones*. 7.ª ed. México: Cengage Learning.

DAS, Braja M. 2013. *Fundamentos de ingeniería geotécnica*. 4.ª ed. México: Cengage Learning.

DAS, Braja M. y SOBHAN, Khaled. 2018. *Principles of Geotechnical Engineering*. 9.ª ed. Boston: Cengage Learning. ISBN 978-1-305-97093-9.

FRIEDMAN, Jerome H. 2001. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics* [en línea]. Vol. 29, n.º 5, pp. 1189–1232. Disponible en: https://doi.org/10.1214/aos/1013203451

GAO, Wei. 2024. The application of machine learning in geotechnical engineering. *Applied Sciences* [en línea]. Vol. 14, n.º 11, art. 4712. Disponible en: https://doi.org/10.3390/app14114712

GÁLVEZ IDROGO, Dilser. *Estudio comparativo de California Bearing Ratio (CBR) de laboratorio y penetración dinámica de cono (PDC) en la subrasante de la nueva vía de evitamiento alterna de la ciudad de Cajamarca* [en línea]. Tesis para optar el título profesional de Ingeniero Civil. Cajamarca: Universidad Nacional de Cajamarca, 2021 [consulta: 12 noviembre 2026]. Disponible en: https://hdl.handle.net/20.500.14074/4510

GÉRON, Aurélien. 2019. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 2.ª ed. Sebastopol, California: O'Reilly Media. ISBN 978-1-4920-3264-9.

Repositorio Institucional - UNAMBA Perú

GOODFELLOW, Ian; BENGIO, Yoshua y COURVILLE, Aaron. 2016. *Deep Learning*. Cambridge, Massachusetts: MIT Press. ISBN 978-0-262-03561-3.

GOPAL, M. 2019. *Applied Machine Learning* [en línea]. Nueva York: McGraw-Hill Education. ISBN 978-1-260-45684-4. Disponible en: https://www.accessengineeringlibrary.com/content/book/9781260456844

HERNÁNDEZ-SAMPIERI, Roberto; FERNÁNDEZ-COLLADO, Carlos y BAPTISTA-LUCIO, María del Pilar. 2014. *Metodología de la investigación*. 6.ª ed. México: McGraw-Hill Education. ISBN 978-1-4562-2396-0.

HERNÁNDEZ-SAMPIERI, Roberto y MENDOZA TORRES, Christian Paulina. 2018. *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta*. Ciudad de México: McGraw-Hill Education. ISBN 978-1-4562-6096-5.

HOLTZ, Robert D.; KOVACS, William D. y SHEAHAN, Thomas C. 2011. *An Introduction to Geotechnical Engineering*. 2.ª ed. Upper Saddle River, Nueva Jersey: Pearson Education. ISBN 978-0-13-249634-6.

HUANG, Yang H. 2004. *Pavement Analysis and Design*. 2.ª ed. Upper Saddle River, Nueva Jersey: Pearson Prentice Hall. ISBN 978-0-13-142473-9.

KHATTI, Jitendra y GROVER, Kamaldeep Singh. 2021. Prediction of geotechnical properties of soil using artificial intelligence framework. *International Journal of Recent Technology and Engineering* [en línea]. Vol. 10, n.º 4, pp. 218–227. ISSN 2277-3878. Disponible en: https://doi.org/10.35940/ijrte.D6625.1110421

KONASANI, Venkata Reddy y KADRE, Shailendra. 2021. *Machine Learning and Deep Learning Using Python and TensorFlow* [en línea]. Nueva York: McGraw-Hill Education. ISBN 978-1-260-46229-6. Disponible en: https://www.accessengineeringlibrary.com/content/book/9781260462296

KUBÁT, Miroslav. 2023. *Fundamentals of Artificial Intelligence: Problem Solving and Automated Reasoning* [en línea]. Nueva York: McGraw-Hill Education. ISBN 978-1-260-46778-9. Disponible en: https://www.accessengineeringlibrary.com/content/book/9781260467789

LI, Bing; YOU, Zhanping; NI, Kai y WANG, Yaqi. 2024. Prediction of soil compaction parameters using machine learning models. *Applied Sciences* [en línea]. Vol. 14, n.º 7, art. 2716. Disponible en: https://doi.org/10.3390/app14072716

LIAW, Andy y WIENER, Matthew. 2002. Classification and regression by randomForest. *R News* [en línea]. Vol. 2, n.º 3, pp. 18–22. Disponible en: https://CRAN.R-project.org/doc/Rnews/

Repositorio Institucional - UNAMBA Perú

MANRIQUE DÁVILA, Medalith. 2019. *Predicción de propiedades mecánicas de suelos utilizando redes neuronales artificiales* [en línea]. Tesis para optar el título profesional de Ingeniera Civil. Huancayo: Universidad Continental. Disponible en: https://hdl.handle.net/20.500.12394/7191

MARATHE, Shriram y SHESHADRI, Akhila. 2025. Machine-learning-based prediction of California Bearing Ratio (CBR) for lateritic–lithomargic pavement subgrades through the major index properties. *Indian Geotechnical Journal* [en línea]. Vol. 55, n.º 6, pp. 2779–2804. Disponible en: https://doi.org/10.1007/s40098-025-01297-1

MITCHELL, James K. y SOGA, Kenichi. 2005. *Fundamentals of Soil Behavior*. 3.ª ed. Hoboken, Nueva Jersey: John Wiley & Sons. ISBN 978-0-471-46302-7.

MONTEJO FONSECA, Alfonso. 1998. *Ingeniería de pavimentos para carreteras*. 2.ª ed. Bogotá: Universidad Católica de Colombia. ISBN 978-958-96036-2-5.

ÖZKAYNAK, Mustafa İsmail y YILMAZ, Yıldırım. 2024. Prediction of resilient modulus with pre-post experimental data of undisturbed subgrade soils using machine learning algorithms. *Transportation Geotechnics* [en línea]. Vol. 49, art. 101396. Disponible en: https://doi.org/10.1016/j.trgeo.2024.101396

PAHNO, Steve; YANG, Jidong J. y KIM, S. Sonny. 2021. Use of machine learning algorithms to predict subgrade resilient modulus. *Infrastructures* [en línea]. Vol. 6, n.º 6, art. 78. Disponible en: https://doi.org/10.3390/infrastructures6060078

QUISPE CAÑAPATAÑA, Carlos Hector y CAMPOS IBAÑEZ, Juan Jorge. *Análisis de CBR, máxima densidad y resistencia al corte para estabilizar un suelo arcilloso SC utilizando 1%, 3% y 5% con cenizas de saúco en Lima Este distrito Ate–Vitarte* [en línea]. Tesis para optar el título profesional de Ingeniero Civil. Lima: Universidad Privada del Norte, 2022 [consulta: 12 noviembre 2026]. Disponible en: https://hdl.handle.net/11537/30690

RAMOS PEREIRA, Matilde Lorena. *Mejoramiento de la capacidad de soporte del suelo a nivel de subrasante a través de la utilización del Sistema Rocamix entre la prolongación de la Avenida San Antonio de Padua y de la calle 04 de la Asociación de Vivienda 2 de Octubre del Centro Poblado San Antonio–Moquegua, 2017* [en línea]. Tesis para optar el título profesional de Ingeniero Civil. Moquegua: Universidad José Carlos Mariátegui, 2019 [consulta: 12 noviembre 2026]. Disponible en: https://hdl.handle.net/20.500.12819/574

Repositorio Institucional - UNAMBA Perú

MINISTERIO DE TRANSPORTES Y COMUNICACIONES. 2008. *Manual de diseño de carreteras no pavimentadas de bajo volumen de tránsito*. Lima: Ministerio de Transportes y Comunicaciones.

MINISTERIO DE TRANSPORTES Y COMUNICACIONES. 2013. *Glosario de términos de uso frecuente en proyectos de infraestructura vial: versión actualizada a junio de 2013* [en línea]. Lima: Ministerio de Transportes y Comunicaciones. Aprobado mediante Resolución Directoral n.º 18-2013-MTC/14. Disponible en: https://www.proviasdes.gob.pe/Prog_incentivos/Normatividad/Norm_sectorial_vinc_meta_40/Glosario_Terminos_Uso_Frecuente_jun13.pdf

MINISTERIO DE TRANSPORTES Y COMUNICACIONES. 2014. *Manual de carreteras: suelos, geología, geotecnia y pavimentos. Sección suelos y pavimentos. Versión abril de 2014* [en línea]. Lima: Ministerio de Transportes y Comunicaciones. Aprobado mediante Resolución Directoral n.º 10-2014-MTC/14. Disponible en: https://www.gob.pe/institucion/mtc/normas-legales/4441297-10-2014-mtc-14

MINISTERIO DE VIVIENDA, CONSTRUCCIÓN Y SANEAMIENTO. 2020. *Norma técnica E.050: Suelos y cimentaciones* [en línea]. Lima: Servicio Nacional de Capacitación para la Industria de la Construcción. ISBN 978-612-48427-3-3. Disponible en: https://www.gob.pe/institucion/sencico/informes-publicaciones/887225-norma-tecnica-e-050-suelos-y-cimentaciones

ŞENTÜRK, Muharrem Atakan; ORDU, Ertuğrul y TAN, Rabia Korkmaz. 2025. Machine learning-based prediction of soil compaction parameters. *Environmental Earth Sciences* [en línea]. Vol. 84, art. 349.

TASKIRAN, T. 2010. Prediction of California bearing ratio (CBR) of fine-grained soils by AI methods. *Advances in Engineering Software* [en línea]. Vol. 41, n.º 6, pp. 886–892.

TERZAGHI, Karl; PECK, Ralph B. y MESRI, Gholamreza. 1996. *Soil Mechanics in Engineering Practice*. 3.ª ed. Nueva York: John Wiley & Sons. ISBN 978-0-471-08658-1.

VALDERRAMA PURIZACA, Frank Jesús. 2020. *Desarrollo de un modelo de redes neuronales artificiales para predecir las propiedades mecánicas del suelo* [en línea]. Tesis para optar el título profesional. Pimentel: Universidad Señor de Sipán. Disponible en: https://hdl.handle.net/20.500.12802/8395

YODER, Eldon J. y WITCZAK, Matthew W. 1975. *Principles of Pavement Design*. 2.ª ed. Nueva York: John Wiley & Sons. ISBN 978-0-471-97780-3.

Repositorio Institucional - UNAMBA Perú

ANEXOS

Repositorio Institucional - UNAMBA Perú

Anexo A
MATRIZ DE CONSISTENCIA

Tabla 29 — Matriz de consistencia

<table>
<tr><th colspan="7">Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025</th></tr>
<tr><th>PROBLEMA</th><th>OBJETIVOS</th><th>HIPÓTESIS</th><th colspan="3">VARIABLES E INDICADORES</th><th>METODOLOGÍA</th></tr>
<tr><td rowspan="2">Problema general:
¿De qué manera el modelo predictivo basado en Machine Learning permite estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025?</td><td rowspan="2">Objetivo general:
Desarrollar un modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.</td><td rowspan="2">Hipótesis general:
El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.</td><td>DIMENSIÓN</td><td>INDICADORES</td><td>UNIDAD</td><td rowspan="3">Tipo de investigación:
Aplicada

Nivel:
Correlacional-predictivo

Diseño:
no experimental

La Población:
la población está conformada por registros geotécnicos relacionados con la respuesta mecánica de suelos a nivel de subrasante, específicamente valores</td></tr>
<tr><td colspan="3">VARIABLE INDEPENDIENTE: Modelo predictivo basado en Machine learning</td></tr>
<tr><td>Problemas específicos:
• ¿De qué manera el modelo predictivo basado en Machine Learning permite estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la ciudad de</td><td>Objetivos específicos:
• Determinar el desempeño del modelo predictivo basado en Machine Learning para estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la</td><td>Hipótesis específicas:
• El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la capacidad de soporte del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.</td><td>X1. Desempeño del modelo predictivo</td><td>– Error Absoluto Medio (MAE)
– Coeficiente de determinación (R^2)</td><td>Valor numérico

R^2 %</td></tr>
</table>

Abancay, Apurímac – 2025? • ¿De qué manera el modelo predictivo basado en Machine Learning permite estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025?	ciudad de Abancay, Apurímac – 2025. • Determinar el desempeño del modelo predictivo basado en Machine Learning para estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.	• El modelo predictivo basado en Machine Learning presenta un desempeño adecuado para estimar la rigidez del suelo de subrasante de la vía de evitamiento de la ciudad de Abancay, Apurímac – 2025.	***VARIABLE DEPENDIENTE:*** *Respuesta mecánica*			de capacidad de soporte del suelo (CBR) y módulo resiliente (Mr) **Muestra:** - para el desarrollo del modelo predictivo se trabajó con 90 registros geotécnicos, los cuales presentaron información completa y consistente para el entrenamiento del modelo de Machine Learning y el análisis de la respuesta mecánica del suelo.
			Y1. Capacidad de soporte del suelo	Índice de CBR	%	
			Y2. Rigidez de la subrasante.	Módulo resiliente (Mr) (pav.flexible)	MPa	

Anexo B

INSTRUMENTO DE RECOLECCIÓN DE DATOS

CONTENIDO DE HUMEDAD DEL SUELO-MTC E 108-2016

PROYECTO:

SOLICITA:
MUESTRA:
UBICACIÓN:

CALICATA:
PROFUNDIDAD: FECHA:

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)			
Peso de Capsula + Muestra Húmeda (gr)			
Peso de Capsula + Muestra Seca (gr)			
Peso del Agua (gr)			
Peso de la Muestra Seca (gr)			
Contenido de Humedad			

Figura 34 — Instrumento contenido de humedad

Repositorio Institucional - UNAMBA Perú

ANÁLISIS GRANULOMÉTRICO POR TAMIZADO-MTC E 107- 2016

PROYECTO:

SOLICITA:
MUESTRA:
UBICACIÓN:

CALICATA:

PROFUNDIDAD: FECHA: Febrero, 2026

Antes del Lavado
Peso de la Muestra Seca =

Despues del Lavado
Peso de la Muestra Seca =
% de Error en Peso = Ok!

TAMIZ	ABERTURA (mm)	PESO RETENIDO	PESO QUE PASA	% RETENIDO	% RETENIDO ACUMULADO	% QUE PASA
4"	100					
3"	75					
2"	50					
1½"	37.5					
1"	25					
3/4"	19					
3/8"	9.5					
Nº 4	4.76					
Nº 10	2					
Nº 20	0.840					
Nº 40	0.425					
Nº 60	0.260					
Nº 140	0.106					
Nº 200	0.075					
Cazuela	-					
Lavado	-					
Total Fracción Retenida en Lavado =		**0.00**		**0.00%**		

Fracciones Retenidas	De la Fraccion Gruesa	Propiedades
% de Gravas(Ret.n°4) =	% de grava =	T. Max.Absoluto=
% de Arena (Pas N°4-Ret N°200)=	% de arena =	T. Max.Relativo=
% de Finos(Pas N°200)=	Total =	Clasf. SUCS= 0.00%
Total=		Clasif.AASHTO= A-2-6 (0)

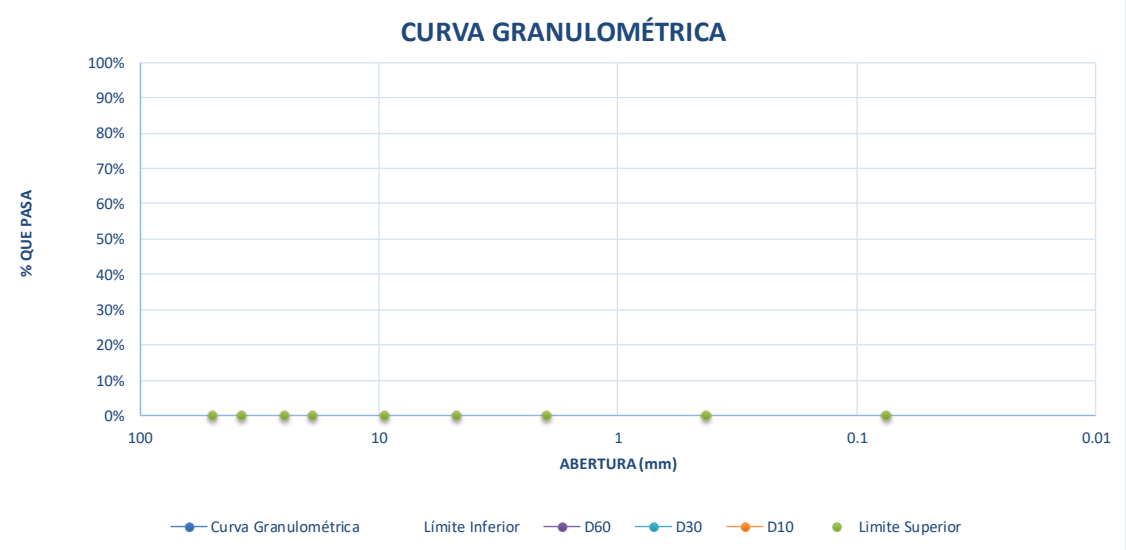


Figura 35 — Instrumento análisis granulométrico

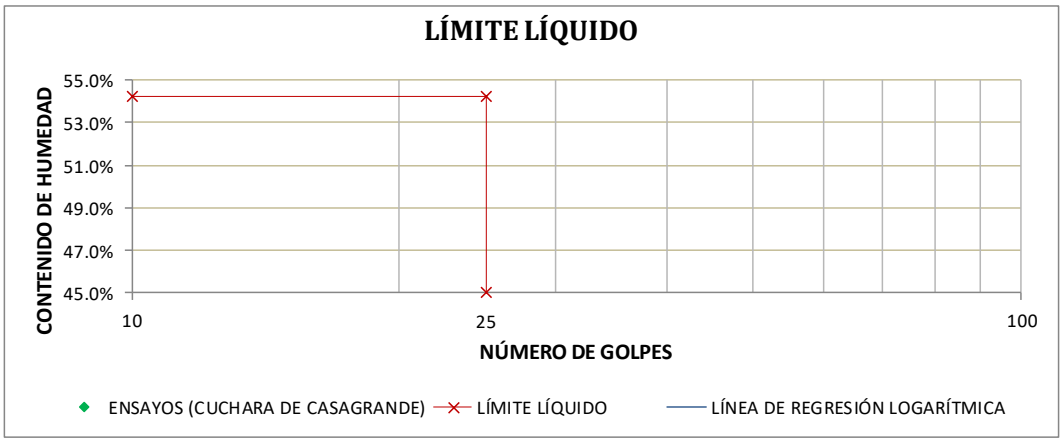
Repositorio Institucional - UNAMBA Perú

LÍMITE DE CONSISTENCIA-MTC E 110 - 201	
PROYECTO:	
SOLICITA: **MUESTRA:** **UBICACIÓN:**	

CALICATA:
PROFUNDIDAD: FECHA:

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03	MUESTRA 04
Peso de Capsula (gr)				
Peso de Capsula + Muestra Húmeda (gr)				
Peso de Capsula + Muestra Seca (gr)				
Peso del Agua (gr)				
Peso de la Muestra Seca (gr)				
Contenido de Humedad				
Número de Golpes				

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)			
Peso de Capsula + Muestra Húmeda (gr)			
Peso de Capsula + Muestra Seca (gr)			
Peso del Agua (gr)			
Peso de la Muestra Seca (gr)			
Contenido de Humedad			



LÍMITE LÍQUIDO =
LÍMITE PLÁSTICO =
ÍNDICE DE PLASTICIDAD =

observaciones: -Muestra tomadas en Campo por el Laboratorio

Figura 36 — Instrumento límites de consistencia

Repositorio Institucional - UNAMBA Perú

ENSAYO DE COMPACTACION - PROCTOR MODIFICADO

ASTM D-1557-70 - AASHTO T-180-70

OBRA: 0

SOLICITA:

MUESTRA:

UBICACIÓN:

CALICATA: 0

PROFUNDIDAD: **FECHA:**

N° de Capas: 5				
N° de golpes por capas	25		Volumen Molde :	2125
Muestra	**1**	**2**	**3**	**4**
Wcap.	45.00	45.00	45.00	45.00
Wcap+sh.				
Wcap+ss.				
Wss.				
Ww.				
%CH.				
Wmolde				
Wsh+molde				
Wsh.				
Densidad Hum.				
Densidad Seca				

Densidad Seca Maxima(g/cm3) :	1.600
Contenido Optimo de Agua(%) :	10.20

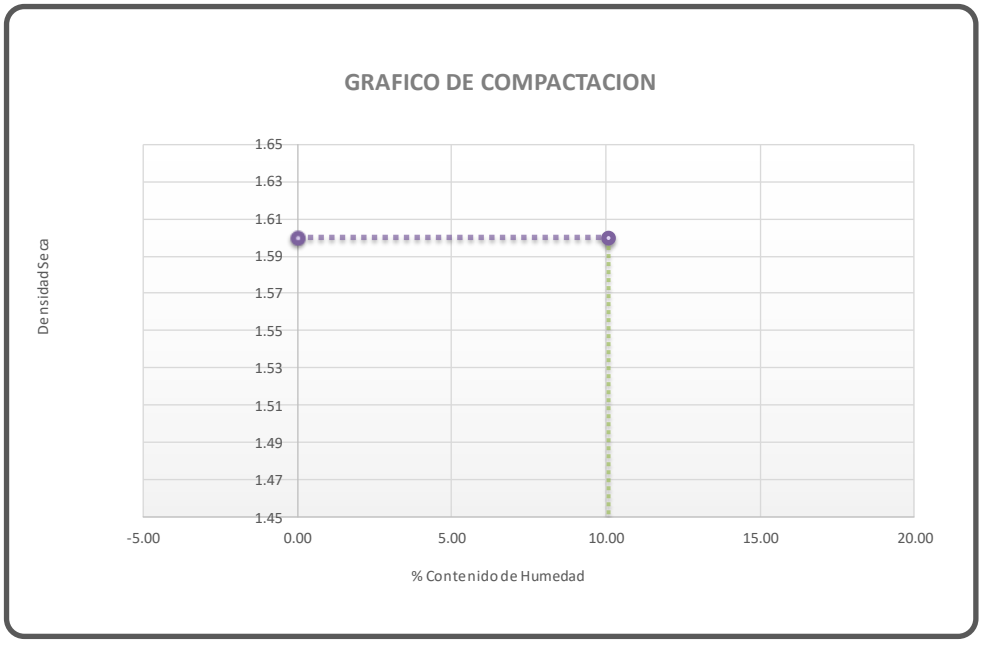


Figura 37 — Instrumento ensayo de compactación

Repositorio Institucional - UNAMBA Perú

RELACION DE SOPORTE DE CALIFORNIA (CBR).

ASTM 1833 - 73

OBRA	0										
SOLICITA:											
MUESTRA:						Humedad Optima(%):					
UBICACIÓN:						Densidad Seca Maxima(g/cm3):					
CALICATA:						95% Densidad Seca Máxima(g/cm3):					
PROFUNDIDAD:											
ene-00						**FECHA:** ene-00					
DIMENSIONES DE MOLDE:			**Diametro**	**Altura**	**Area**	**Diametro**	**Altura**	**Area**	**Diametro**	**Altura**	**Area**
			15.24	11.64	182.42	15.24	11.64	182.42	15.24	11.64	182.42
N° DE GOLPES POR CAPA				12			25			56	
CONDICIONES DE LA MUESTRA			**SIN MOJAR**		**SATURADA**	**SIN MOJAR**		**SATURADA**	**SIN MOJAR**		**SATURADA**
Peso del molde(gr).											
Volumen de la Muestra(cc).											
Muestra Humeda + Molde(gr).											
Muestra Humeda(gr).											
Densidad Humeda(gr/cm3).											
CONTENIDO DE HUMEDAD DE LA MUESTRA			**MEDIO**		**MEDIO**	**MIEDIO**		**MEDIO**	**MEDIO**		**MEDIO**
Peso de la capsula(gr).											
Muestra Humeda + Capsula(gr).											
Muestra Seca+ Capsula(gr).											
Muestra Seca(gr).											
Contenido de Agua(gr)											
Contenido de Humedad(%).											
Densidad Seca(gr/cm3).											
	DIA	**HORA**	**INTERVALO (DIAS)**	**LECT. DEFOR.**	**EXPANSIÓN**	**LECT. DEFOR.**	**EXPANSIÓN**		**LECT. DEFOR.**	**EXPANSIÓN**	
					PULGS		**PULGS**			**PULGS**	
					%		**%**			**%**	
MEDICION DE LA EXPANSION	1.00		0.00								
	2.00		1.00								
	3.00		2.00								
	4.00		3.00								
	5.00		4.00								
	PENETRACION "PULGADAS"	**CARG. PATRON "kg/cm2"**	**FACTOR CARGA**		**CBR**	**FACTOR CARGA**		**CBR**	**FACTOR CARGA**		**CBR**
			DIAL	**kg/cm2**	**12**	**DIAL**	**kg/cm2**	**25**	**DIAL**	**kg/cm2**	**56**
ENSAYO DE PENETRACION	0										
	0.025										
	0.050										
	0.075										
	0.100	70.455									
	0.125										
	0.150										
	0.200	105.682									
	0.300										
	0.400										
	0.500										

Figura 38 — Instrumento CBR

Repositorio Institucional - UNAMBA Perú

Proyecto :							PERFORACION C-01 
UBICACIÓN : SOLICITA :							
TECNICA DE INVESTIGACION : CALICATA FECHA:					DIMENSIÓN DE CALICATA : L: PROFUNDIDAD FINAL : PRESENCIA DE NIVEL FRIÁTICO: 		
PROF	ESPESOR	COTA M.S.N.M	CANTIDAD DE ESTRATO	SIMBOLO	CLASIFIC SUCS	DESCRIPCION DEL SUELO	OBSERVACIONES Y REGISTRO FOTOGRAFICO
					0.00%	SUCS: 0.00% AASHTO: HUMEDAD % LIMITES DE CONSISTENCIA % LL : 0.0% % LP : 0.0% % IP : 0.00% COLOR DEL SUELO Rojo oscuro	

Figura 39 — Instrumento perfil estratigráfico

Repositorio Institucional - UNAMBA Perú

Anexo C

RESULTADOS DE LABORATORIO

LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING , CONSTRUCTION

LABORATORIO

RUC : 20609203570

997374826 / 985190489

MAUCACALLE - SAHUNAY MZ B N°2, ABANCAY, PERU

CONTENIDO DE HUMEDAD DEL SUELO-MTC E 108-2016

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos

MUESTRA: NATURAL

UBICACIÓN: Abancay - Apurímac

731356.21 m E 8494714.84 m S

CALICATA: N° - 01

PROFUNDIDAD: 1.50 mts

FECHA: Febrero, 2026

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)	32.00	32.90	32.00
Peso de Capsula + Muestra Húmeda (gr)	512.60	498.25	530.20
Peso de Capsula + Muestra Seca (gr)	470.00	456.90	486.00
Peso del Agua (gr)	42.60	41.35	44.20
Peso de la Muestra Seca (gr)	438.00	424.00	454.00
Contenido de Humedad	**9.73%**	**9.75%**	**9.74%**

CONTENIDO DE HUMEDAD NATURAL = 9.74%

IZAN

Ing. Johon Nahuirima Cabezas

GERENTE GENERAL

CIP: 296789

Figura 40 — Contenido de humedad C-01

Repositorio Institucional - UNAMBA Perú

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING - CONSTRUCTION
LABORATORIO RUC : 20605203570
997374826 / 905190489
MAUCACALLE - SAHUNAY MZ B LT2, ABANCAY, PERU

ANÁLISIS GRANULOMÉTRICO POR TAMIZADO-MTC E 107- 2016

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

731356.21 m E 8494714.84 m S

CALICATA: N° - 01

PROFUNDIDAD: 1.50 mts

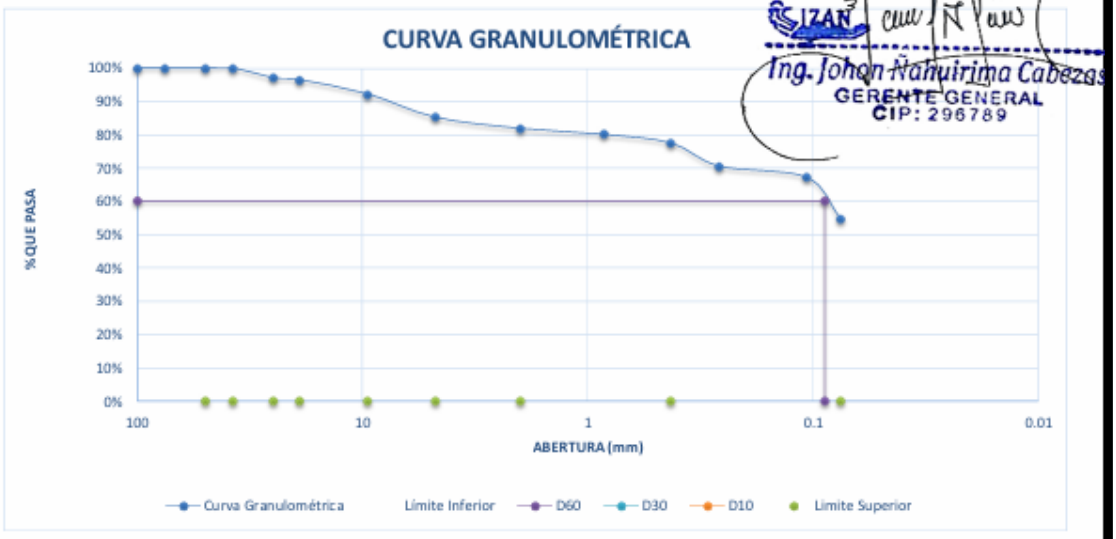
FECHA: Febrero, 2026

Antes del Lavado
Peso de la Muestra Seca = 2747.00 gr

Despues del Lavado
Peso de la Muestra Seca = 1247.00 gr
% de Error en Peso = 0.08% Ok!

TAMIZ	ABERTURA (mm)	PESO RETENIDO	PESO QUE PASA	% RETENIDO	% RETENIDO ACUMULADO	% QUE PASA
4"	100	0.00	1246.0	0.00%	0.00%	100.00%
3"	75	0.00	1246.00	0.00%	0.00%	100.00%
2"	50	0.00	1246.00	0.00%	0.00%	100.00%
1½"	37.5	0.00	1246.00	0.00%	0.00%	100.00%
1"	25	82.00	1164.00	2.99%	2.99%	97.01%
3/4"	19	16.00	1148.00	0.58%	3.57%	96.43%
3/8"	9.5	116.00	1032.00	4.22%	7.79%	92.21%
Nº 4	4.76	192.00	840.00	6.99%	14.78%	85.22%
Nº 10	2	92.00	748.00	3.35%	18.13%	81.87%
Nº 20	0.840	50.00	698.00	1.82%	19.95%	80.05%
Nº 40	0.425	71.00	627.00	2.58%	22.53%	77.47%
Nº 60	0.260	190.00	437.00	6.92%	29.45%	70.55%
Nº 140	0.106	93.00	344.00	3.39%	32.84%	67.16%
Nº 200	0.075	344.00	0.00	12.52%	45.36%	54.64%
Cazuela	-	87.00	-	3.17%	48.53%	-
Lavado	-		-	51.47%	100.00%	-
Total Fracción Retenida en Lavado =		**1246.00**		**100.00%**		

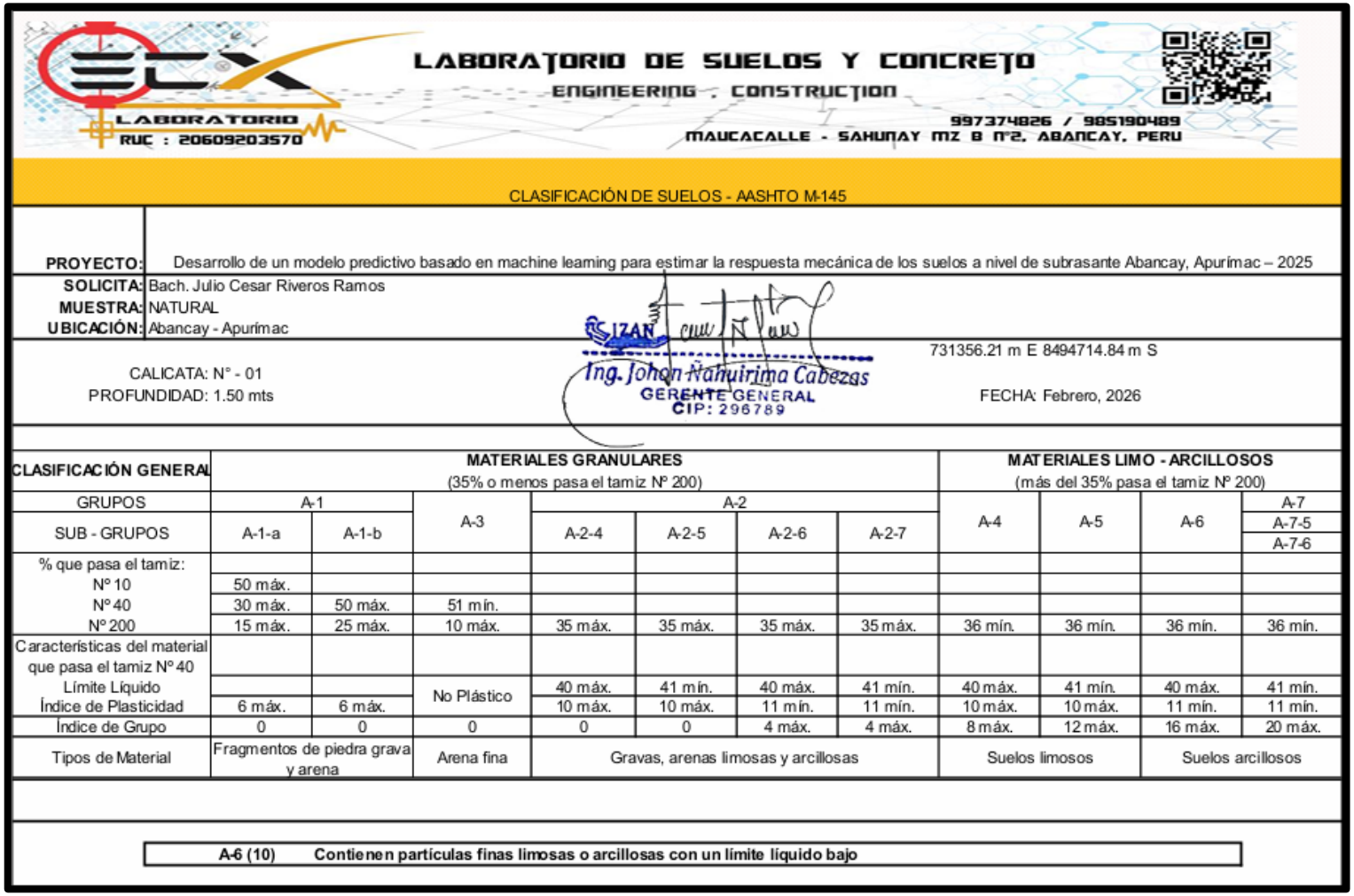
Fracciones Retenidas		De la Fraccion Gruesa		Propiedades
% de Gravas(Ret.n°4) =	14.78%	% de grava =	32.58%	T. Max.Absoluto= 1½"
% de Arena (Pas N°4-Ret N°200)=	30.58%	% de arena =	67.42%	T. Max.Relativo= 3/8"
% de Finos(Pas N°200)=	54.64%	Total =	100.00%	Clasf. SUCS= CH
Total=	100.00%			Clasif.AASHTO= A-6 (10)



Ing. Johon Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 41 — Análisis granulométrico C-01

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING . CONSTRUCTION

LABORATORIO
RUC : 20609203570

997374826 / 985190489
MAUCACALLE - SAHUNAY MZ B N°2, ABANCAY, PERU

CLASIFICACIÓN DE SUELOS - AASHTO M-145

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

IZAN
Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

CALICATA: N° - 01
PROFUNDIDAD: 1.50 mts

731356.21 m E 8494714.84 m S

FECHA: Febrero, 2026

CLASIFICACIÓN GENERAL	MATERIALES GRANULARES (35% o menos pasa el tamiz Nº 200)							MATERIALES LIMO - ARCILLOSOS (más del 35% pasa el tamiz Nº 200)			
GRUPOS	A-1		A-3	A-2				A-4	A-5	A-6	A-7
SUB - GRUPOS	A-1-a	A-1-b		A-2-4	A-2-5	A-2-6	A-2-7				A-7-5 A-7-6
% que pasa el tamiz:											
Nº 10	50 máx.										
Nº 40	30 máx.	50 máx.	51 mín.								
Nº 200	15 máx.	25 máx.	10 máx.	35 máx.	35 máx.	35 máx.	35 máx.	36 mín.	36 mín.	36 mín.	36 mín.
Características del material que pasa el tamiz Nº 40											
Límite Líquido			No Plástico	40 máx.	41 mín.	40 máx.	41 mín.	40 máx.	41 mín.	40 máx.	41 mín.
Índice de Plasticidad	6 máx.	6 máx.		10 máx.	10 máx.	11 mín.	11 mín.	10 máx.	10 máx.	11 mín.	11 mín.
Índice de Grupo	0	0	0	0	0	4 máx.	4 máx.	8 máx.	12 máx.	16 máx.	20 máx.
Tipos de Material	Fragmentos de piedra grava y arena		Arena fina	Gravas, arenas limosas y arcillosas				Suelos limosos		Suelos arcillosos	

A-6 (10) **Contienen partículas finas limosas o arcillosas con un límite líquido bajo**

Figura 42 — Clasificación de suelos C-01

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING - CONSTRUCTION
LABORATORIO RUC : 20605202570
997374826 / 905190489
MAUCACALLE - SAHUNAY MZ B LT2. ABANCAY, PERU

LÍMITE DE CONSISTENCIA-MTC E 110 - 201

PROYECTO:	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025
SOLICITA:	Bach. Julio Cesar Riveros Ramos
MUESTRA:	NATURAL
UBICACIÓN:	Abancay - Apurímac

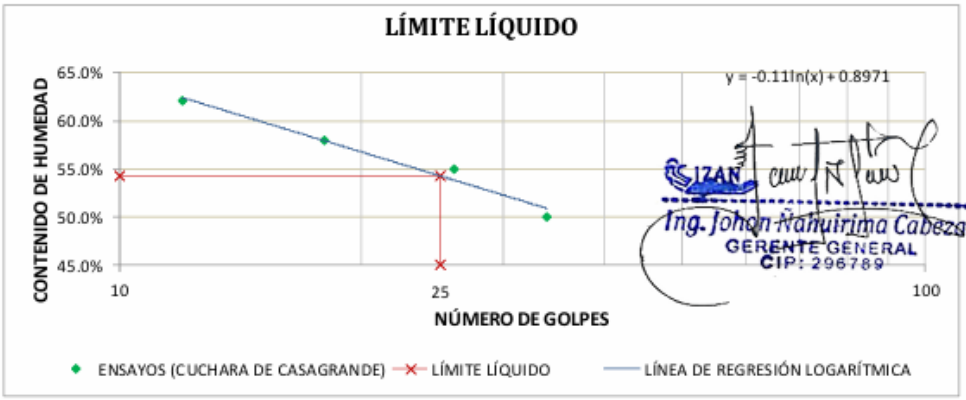
9.74% 731356.21 m E 8494714.84 m S

CALICATA: N° - 01

PROFUNDIDAD: 1.50 mts FECHA: Febrero, 2026

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03	MUESTRA 04
Peso de Capsula (gr)	31.80	33.10	30.95	32.55
Peso de Capsula + Muestra Húmeda (gr)	57.72	57.59	54.51	56.25
Peso de Capsula + Muestra Seca (gr)	47.80	48.60	46.15	48.35
Peso del Agua (gr)	9.92	8.99	8.36	7.90
Peso de la Muestra Seca (gr)	16.00	15.50	15.20	15.80
Contenido de Humedad	**62.00%**	**58.00%**	**55.00%**	**50.00%**
Número de Golpes	**12**	**18**	**26**	**34**

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)	12.90	13.40	13.10
Peso de Capsula + Muestra Húmeda (gr)	19.80	19.95	20.05
Peso de Capsula + Muestra Seca (gr)	18.30	18.54	18.54
Peso del Agua (gr)	1.50	1.41	1.51
Peso de la Muestra Seca (gr)	5.40	5.14	5.44
Contenido de Humedad	**27.78%**	**27.43%**	**27.76%**



IZAN
Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

LÍMITE LÍQUIDO = 54.20%
LÍMITE PLÁSTICO = 27.66%
ÍNDICE DE PLASTICIDAD = 26.54%

observaciones: -Muestra tomadas en Campo por el Laboratorio

Figura 43 — Limites de consistencia C-01

Repositorio Institucional - UNAMBA Perú

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING AND CONSTRUCTION
LABORATORIO
RUC : 20609203570
997374826 / 985190489
AV. EL ARCO S/N

ENSAYO DE COMPACTACION - PROCTOR MODIFICADO
ASTM D-1557-70 - AASHTO T-180-70

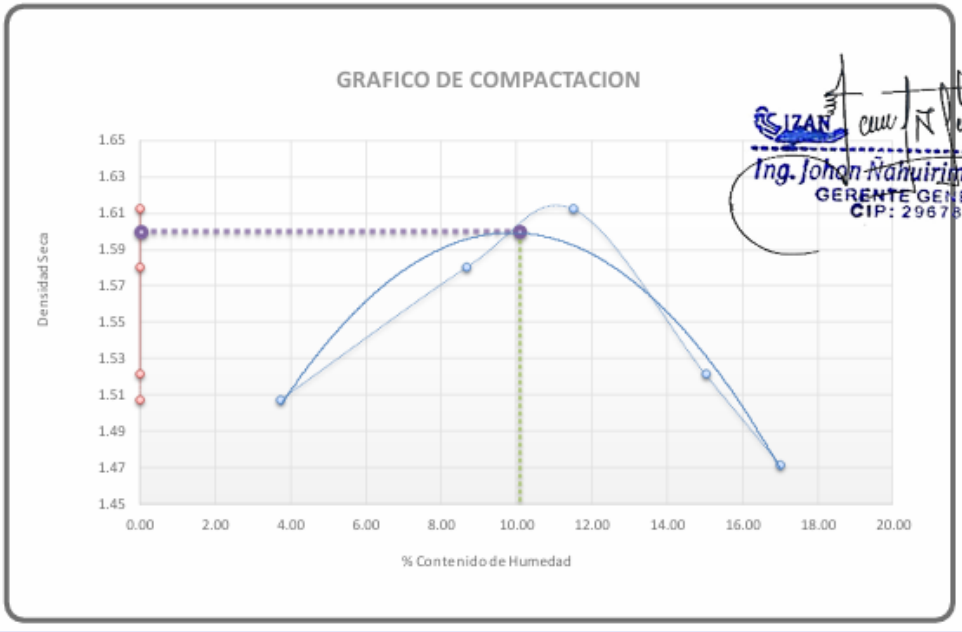
OBRA: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

CALICATA: N° - 01 **731356.21 m E** 8494714.84 m S
PROFUNDIDAD: 1.50 mts **FECHA:** Febrero, 2026

N° de Capas: 5				
N° de golpes por capas	25		Volumen Molde :	2125
Muestra	**1**	**2**	**3**	**4**
Wcap.	45.00	45.00	45.00	45.00
Wcap+sh.	**498.23**	**501.50**	**519.00**	**512.00**
Wcap+ss.	**481.89**	**465.00**	**470.00**	**451.00**
Wss.	436.89	420.00	425.00	406.00
Ww.	16.34	36.50	49.00	61.00
%CH.	3.74	8.69	11.53	15.02
Wmolde	**6379**	**6379**	**6379**	**6379**
Wsh+molde	**9702**	**10029**	**10200**	**10098**
Wsh.	3323	3650	3821	3719
Densidad Hum.	1.564	1.718	1.798	1.750
Densidad Seca	1.507	1.580	1.612	1.522

Densidad Seca Maxima(g/cm3) :	1.600
Contenido Optimo de Agua(%) :	10.20



IZAN
Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296739

Figura 44 — Ensayo de compactación C-01

Repositorio Institucional - UNAMBA Perú

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING AND CONSTRUCTION
LABORATORIO
RUC : 20609203570
997374026 / 985190489
AV. EL ARCO S/N

RELACION DE SOPORTE DE CALIFORNIA (CBR).

ASTM 1833 - 73

OBRA Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subra
Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos

MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

Humedad Optima(%): **10.20**
Densidad Seca Maxima(g/cm3): **1.600**
95% Densidad Seca Máxima(g/cm3): **1.520**

CALICATA: N° - 01
PROFUNDIDAD: 1.50 mts
731356.21 m E **8494714.84 m S**

FECHA: Febrero, 2026

DIMENSIONES DE MOLDE:	Diametro	Altura	Area	Diametro	Altura	Area	Diametro	Altura
	15.24	11.64	182.42	15.24	11.64	182.42	15.24	11.64
N° DE GOLPES POR CAPA	12			25			56	
CONDICIONES DE LA MUESTRA	SIN MOJAR		SATURADA	SIN MOJAR		SATURADA	SIN MOJAR	
Peso del molde(gr).	7785			7785			7790	
Volumen de la Muestra(cc).	2082			2058			2059	
Muestra Humeda + Molde(gr).	11050		11220	11345		11560	11450	
Muestra Humeda(gr).	3265		3435	3560		3775	3660	
Densidad Humeda(gr/cm3).	1.57		1.65	1.73		1.83	1.78	
CONTENIDO DE HUMEDAD DE LA MUESTRA	MEDIO		MEDIO	MIEDIO		MEDIO	MEDIO	
Peso de la capsula(gr).	35.00		35.00	35.00		35.00	35.00	
Muestra Humeda + Capsula(gr).	720.00		760.50	790.00		760.00	780.50	
Muestra Seca+ Capsula(gr).	652.00		670.50	715.60		673.00	709.50	
Muestra Seca(gr).	617.00		635.50	680.60		638.00	674.50	
Contenido de Agua(gr)	68.00		90.00	74.40		87.00	71.00	
Contenido de Humedad(%).	11.02		14.16	10.93		13.64	10.53	
Densidad Seca(gr/cm3).	1.413		1.445	1.559		1.614	1.608	

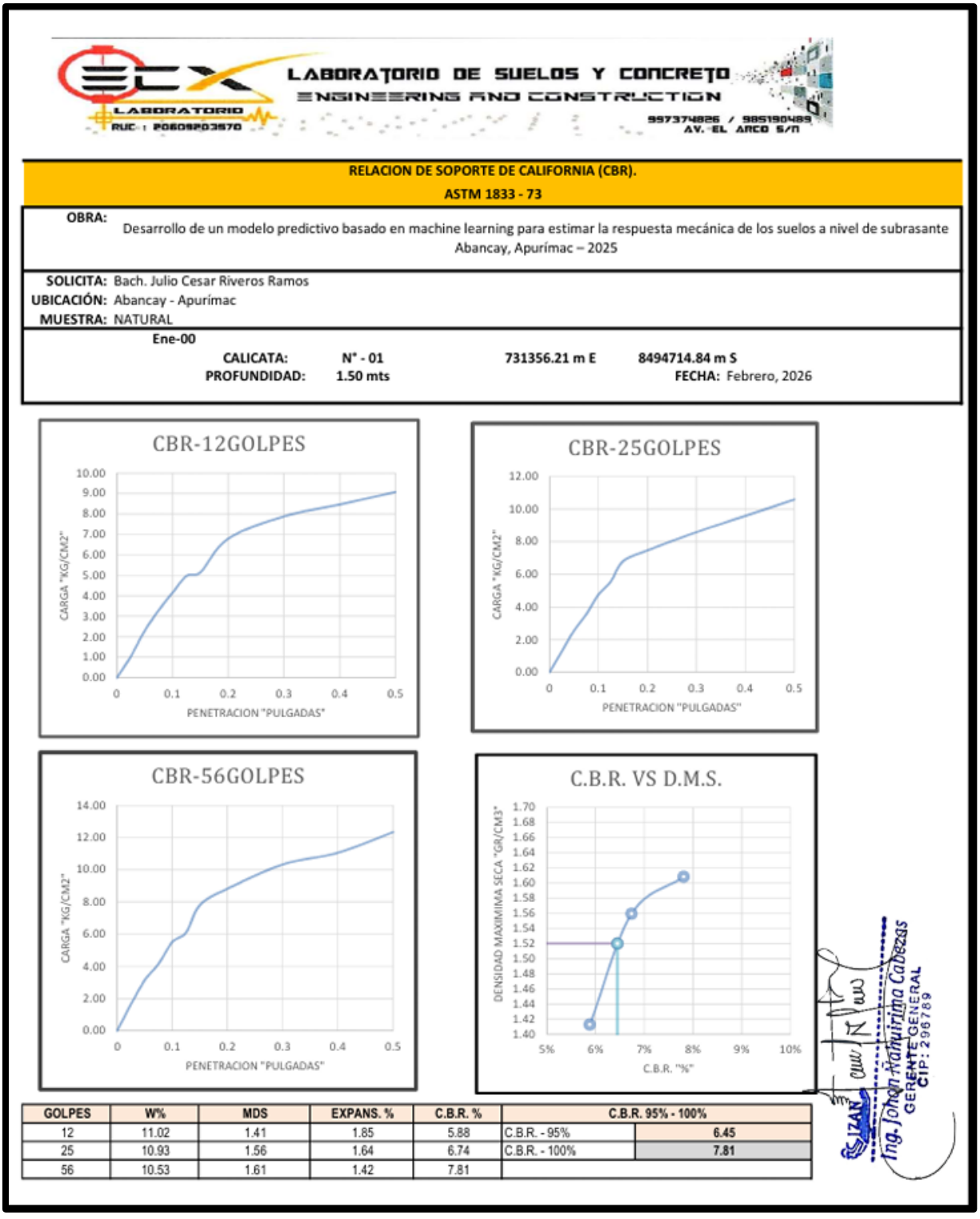
	DIA	HORA	INTERVALO (DIAS)	LECT. DEFOR.	EXPANSIÓN PULGS	EXPANSIÓN %	LECT. DEFOR.	EXPANSIÓN PULGS	EXPANSIÓN %	LECT. DEFOR.	E PULGS
MEDICION DE LA EXPANSION	1.00	13:05	0.00	0.000			0.0000			0.0000	
	2.00	13:02	1.00	0.025	0.0250	0.5454	0.0350	0.0350	0.7635	0.0425	0.0425
	3.00	13:10	2.00	0.050	0.0500	1.0908	0.0500	0.0500	1.0908	0.0500	0.0500
	4.00	13:04	3.00	0.085	0.0850	1.8543	0.0750	0.0750	1.6362	0.0650	0.0650
	5.00	12:58	4.00	0.085	0.0850	1.8543	0.0750	0.0750	1.6362	0.0650	0.0650

	PENETRACION "PULGADAS"	CARG. PATRON "kg/cm2"	FACTOR CARGA DIAL	FACTOR CARGA kg/cm2	CBR 12	FACTOR CARGA DIAL	FACTOR CARGA kg/cm2	CBR 25	FACTOR CARGA DIAL	FACTOR CARGA kg/cm2
ENSAYO DE PENETRACION	0		0	0.00		0	0.00		0	0.00
	0.025		20.5	1.03		25.25	1.27		32.5	1.63
	0.050		45.6	2.29		50.6	2.54		62.5	3.14
	0.075		65.5	3.29		70.5	3.54		82.6	4.15
	0.100	70.455	82.5	4.15	5.88	94.5	4.75	6.74	109.5	5.50
	0.125		99	4.97		110.5	5.55		121.5	6.11
	0.150		102.5	5.15		135	6.78		155.2	7.80
	0.200	105.682	135.25	6.80	6.43	148.5	7.46	7.06	175.5	8.82
	0.300		156.8	7.88		170.8	8.58		205.5	10.33
	0.400		168.5	8.47		190.5	9.57		220	11.06
	0.500		180.5	9.07		210.5	10.58		245.5	12.34

IZAN
Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 45 — Ensayo CBR C-01

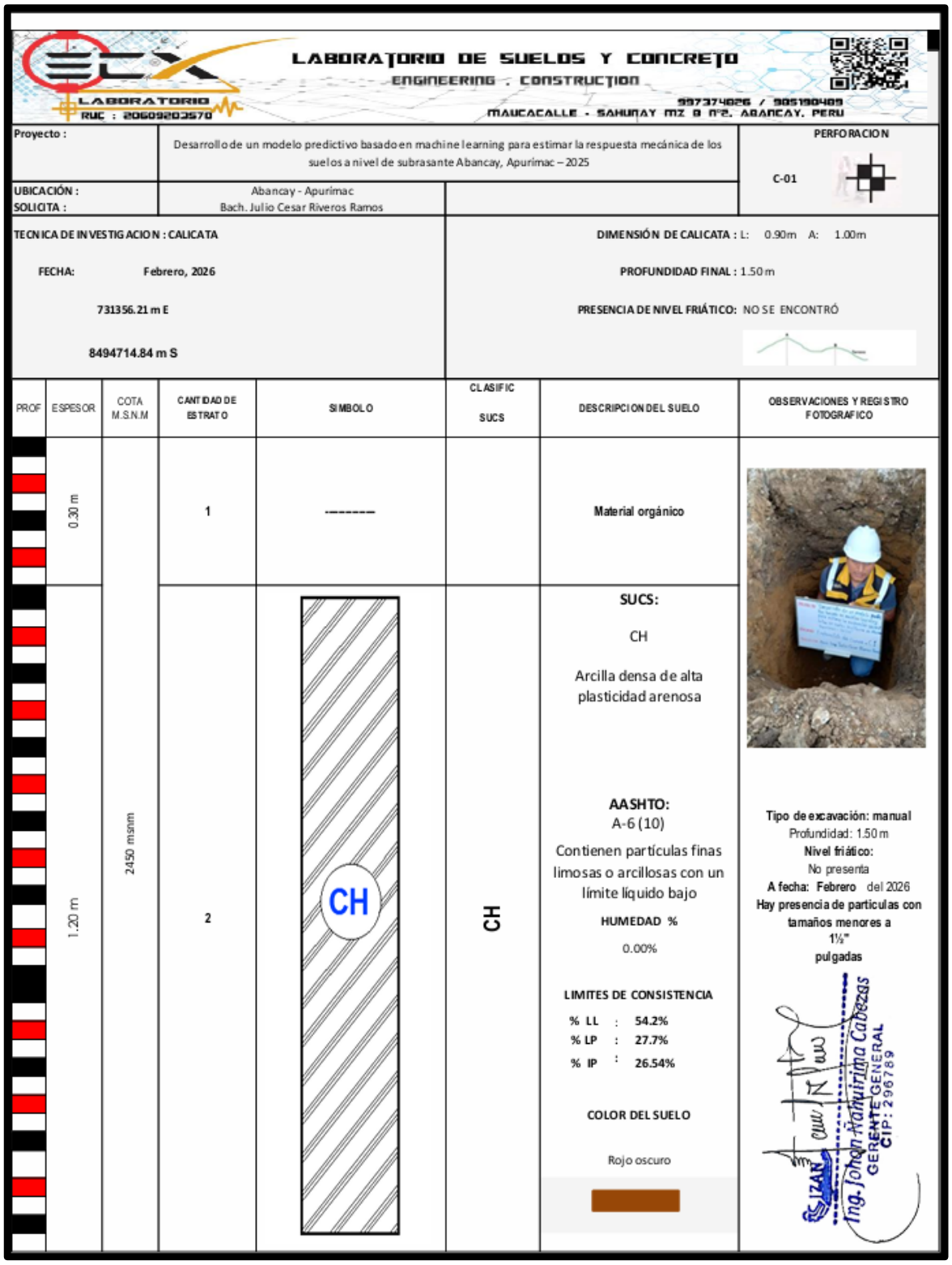
Repositorio Institucional - UNAMBA Perú



GOLPES	W%	MDS	EXPANS. %	C.B.R. %	C.B.R. 95% - 100%	
12	11.02	1.41	1.85	5.88	C.B.R. - 95%	6.45
25	10.93	1.56	1.64	6.74	C.B.R. - 100%	7.81
56	10.53	1.61	1.42	7.81		

Figura 46 — Ensayo CBR-Gráfico C-01

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING . CONSTRUCTION

LABORATORIO RUC : 20609203570

997374026 / 985190489

MAUCACALLE - SAHUNAY MZ B N°2. ABANCAY. PERU

Proyecto :	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025	PERFORACION C-01
UBICACIÓN :	Abancay - Apurímac	
SOLICITA :	Bach. Julio Cesar Riveros Ramos	

TECNICA DE INVESTIGACION : CALICATA

FECHA: Febrero, 2026

731356.21 m E

8494714.84 m S

DIMENSIÓN DE CALICATA : L: 0.90m A: 1.00m

PROFUNDIDAD FINAL : 1.50 m

PRESENCIA DE NIVEL FRIÁTICO: NO SE ENCONTRÓ

PROF	ESPESOR	COTA M.S.N.M	CANTIDAD DE ESTRATO	SIMBOLO	CLASIFIC SUCS	DESCRIPCION DEL SUELO	OBSERVACIONES Y REGISTRO FOTOGRAFICO
	0.30 m	2450 msnm	1	--------		Material orgánico	
	1.20 m		2	CH	CH	SUCS: CH Arcilla densa de alta plasticidad arenosa AASHTO: A-6 (10) Contienen partículas finas limosas o arcillosas con un límite líquido bajo HUMEDAD % 0.00% LIMITES DE CONSISTENCIA % LL : 54.2% % LP : 27.7% % IP : 26.54% COLOR DEL SUELO Rojo oscuro	Tipo de excavación: manual Profundidad: 1.50 m Nivel friático: No presenta A fecha: Febrero del 2026 Hay presencia de particulas con tamaños menores a 1½" pulgadas

Ing. Johan Yahuirima Cabezas GERENTE GENERAL CIP: 296789

Figura 47 — Perfil estratigráfico C-01

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING , CONSTRUCTION

LABORATORIO
RUC : 20609203570

997374826 / 985190489
MAUCACALLE - SAHUNAY MZ B N°2, ABANCAY, PERU

CONTENIDO DE HUMEDAD DEL SUELO-MTC E 108-2016

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

CALICATA: N° - 02 — 731264.63 m E 8494771.30 m S
PROFUNDIDAD: 1.50 mts — FECHA: Febrero, 2026

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)	30.00	30.00	30.00
Peso de Capsula + Muestra Húmeda (gr)	620.80	605.40	590.10
Peso de Capsula + Muestra Seca (gr)	575.00	560.60	546.36
Peso del Agua (gr)	45.80	44.80	43.74
Peso de la Muestra Seca (gr)	545.00	530.60	516.36
Contenido de Humedad	**8.40%**	**8.44%**	**8.47%**

CONTENIDO DE HUMEDAD NATURAL = 8.44%

Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 48 — Contenido de humedad C-02

Repositorio Institucional - UNAMBA Perú

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING - CONSTRUCTION
LABORATORIO RUC : 20609203570
957374826 / 905190489
MAUCACALLE - SAHUNAY MZ B LT2. ABANCAY, PERU

ANÁLISIS GRANULOMÉTRICO POR TAMIZADO-MTC E 107- 2016

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

731264.63 m E 8494771.30 m S

CALICATA: N° - 02

PROFUNDIDAD: 1.50 mts

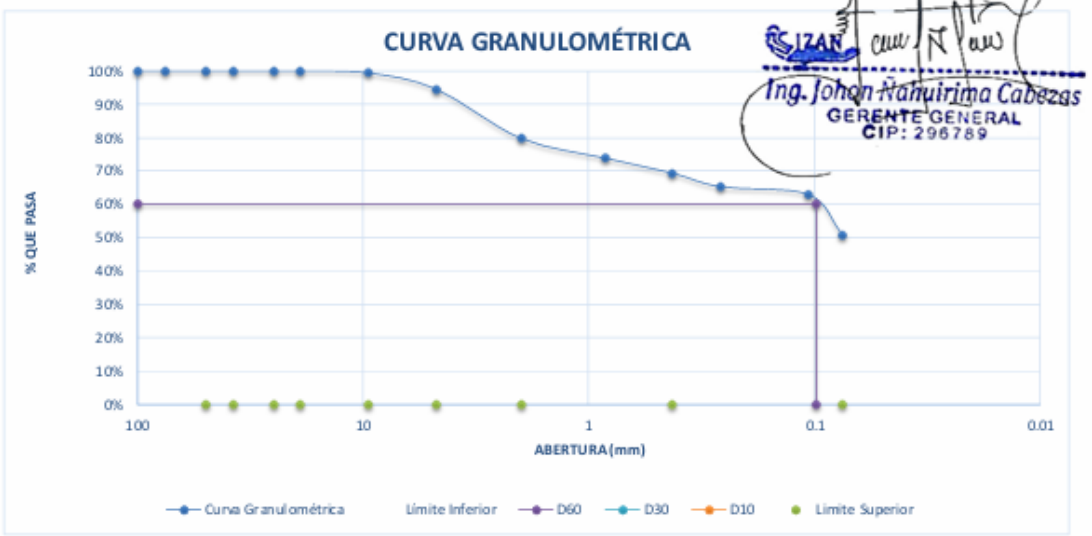
FECHA: Febrero, 2026

Antes del Lavado
Peso de la Muestra Seca = 2755.00 gr

Despues del Lavado
Peso de la Muestra Seca = 1356.00 gr
% de Error en Peso = 0.15% Ok!

TAMIZ	ABERTURA (mm)	PESO RETENIDO	PESO QUE PASA	% RETENIDO	% RETENIDO ACUMULADO	% QUE PASA
4"	100	0.00	1354.0	0.00%	0.00%	100.00%
3"	75	0.00	1354.00	0.00%	0.00%	100.00%
2"	50	0.00	1354.00	0.00%	0.00%	100.00%
1½"	37.5	0.00	1354.00	0.00%	0.00%	100.00%
1"	25	0.00	1354.00	0.00%	0.00%	100.00%
3/4"	19	0.00	1354.00	0.00%	0.00%	100.00%
3/8"	9.5	19.00	1335.00	0.69%	0.69%	99.31%
N° 4	4.76	132.00	1203.00	4.79%	5.48%	94.52%
N° 10	2	406.00	797.00	14.74%	20.22%	79.78%
N° 20	0.840	163.00	634.00	5.92%	26.13%	73.87%
N° 40	0.425	123.00	511.00	4.46%	30.60%	69.40%
N° 60	0.260	113.00	398.00	4.10%	34.70%	65.30%
N° 140	0.106	63.00	335.00	2.29%	36.99%	63.01%
N° 200	0.075	335.00	0.00	12.16%	49.15%	50.85%
Cazuela	-	55.00	-	2.00%	51.14%	-
Lavado	-		-	48.86%	100.00%	-
Total Fracción Retenida en Lavado =		1354.00		100.00%		

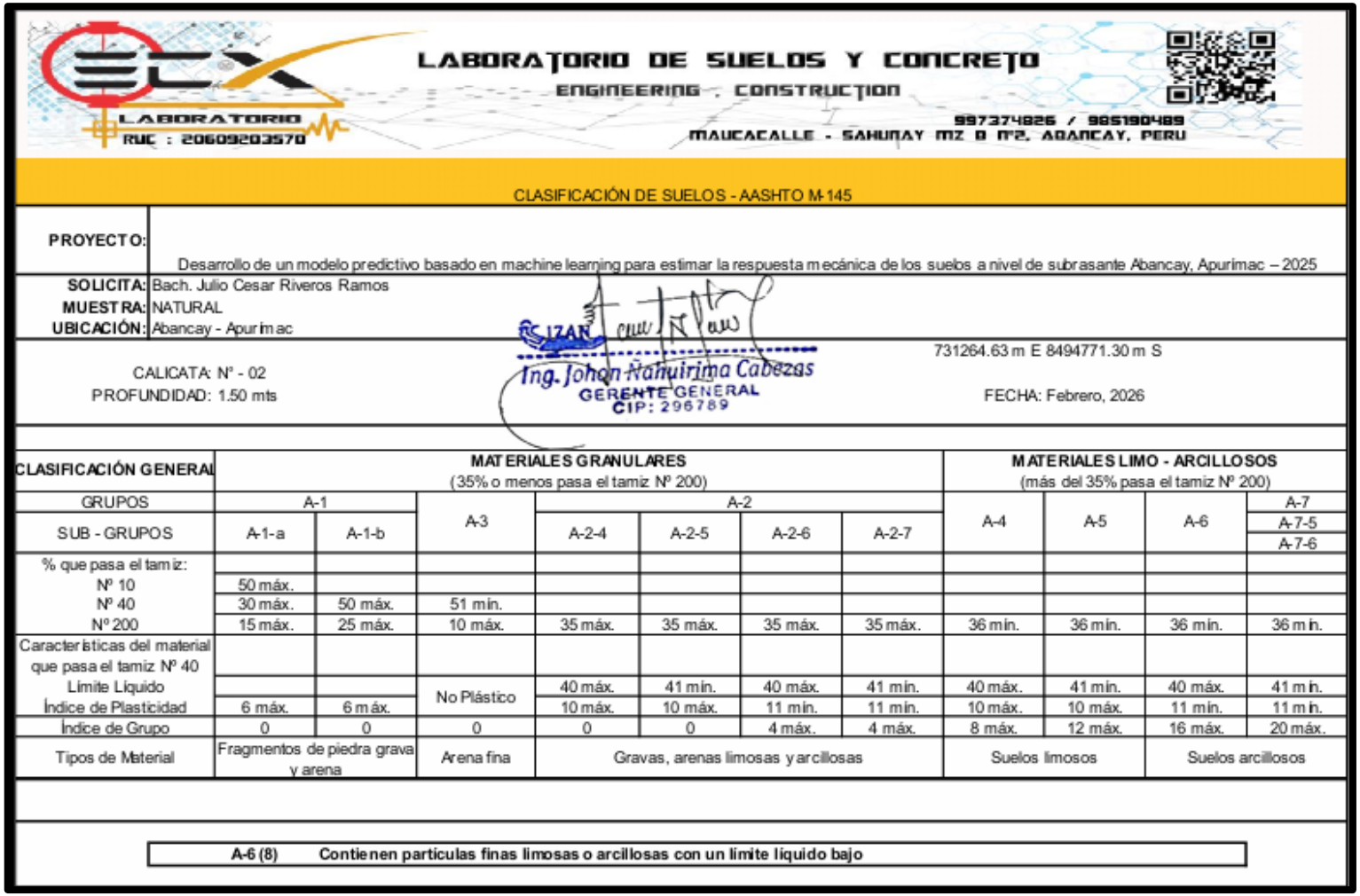
Fracciones Retenidas		De la Fraccion Gruesa		Propiedades
% de Gravas(Ret.n°4) =	5.48%	% de grava =	11.15%	T. Max.Absoluto= 3/4"
% de Arena (Pas N°4-Ret N°200)=	43.67%	% de arena =	88.85%	T. Max.Relativo= N° 4
% de Finos(Pas N°200)=	50.85%	Total =	100.00%	Clasf. SUCS= CL
Total=	100.00%			Clasif.AASHTO= A-6 (8)



IZAN
Ing. Johan Ramirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 49 — Análisis granulométrico C-02

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING . CONSTRUCTION

LABORATORIO RUC : 20609203570

997374826 / 985190489

MAUCACALLE - SAHUNAY MZ B N°2. ABANCAY, PERU

CLASIFICACIÓN DE SUELOS - AASHTO M-145

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos

MUESTRA: NATURAL

UBICACIÓN: Abancay - Apurimac

Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

CALICATA: N° - 02

PROFUNDIDAD: 1.50 mts

731264.63 m E 8494771.30 m S

FECHA: Febrero, 2026

CLASIFICACIÓN GENERAL	MATERIALES GRANULARES (35% o menos pasa el tamiz Nº 200)							MATERIALES LIMO - ARCILLOSOS (más del 35% pasa el tamiz Nº 200)			
GRUPOS	A-1		A-3	A-2				A-4	A-5	A-6	A-7
SUB - GRUPOS	A-1-a	A-1-b		A-2-4	A-2-5	A-2-6	A-2-7				A-7-5 A-7-6
% que pasa el tamiz:											
Nº 10	50 máx.										
Nº 40	30 máx.	50 máx.	51 min.								
Nº 200	15 máx.	25 máx.	10 máx.	35 máx.	35 máx.	35 máx.	35 máx.	36 min.	36 min.	36 min.	36 min.
Caracteristicas del material que pasa el tamiz Nº 40											
Limite Liquido			No Plástico	40 máx.	41 min.	40 máx.	41 min.	40 máx.	41 min.	40 máx.	41 min.
Índice de Plasticidad	6 máx.	6 máx.		10 máx.	10 máx.	11 min.	11 min.	10 máx.	10 máx.	11 min.	11 min.
Índice de Grupo	0	0	0	0	0	4 máx.	4 máx.	8 máx.	12 máx.	16 máx.	20 máx.
Tipos de Material	Fragmentos de piedra grava y arena		Arena fina	Gravas, arenas limosas y arcillosas				Suelos limosos		Suelos arcillosos	

A-6 (8) **Contienen particulas finas limosas o arcillosas con un limite liquido bajo**

Figura 50 — Clasificación de suelos C-02

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING - CONSTRUCTION
LABORATORIO RUC : 20609203570
997374826 / 985190489
MAUCACALLE - SAHUNAY MZ B N°2, ABANCAY, PERU

LÍMITE DE CONSISTENCIA-MTC E 110 - 201

PROYECTO:	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025
SOLICITA:	Bach. Julio Cesar Riveros Ramos
MUESTRA:	NATURAL
UBICACIÓN:	Abancay - Apurímac

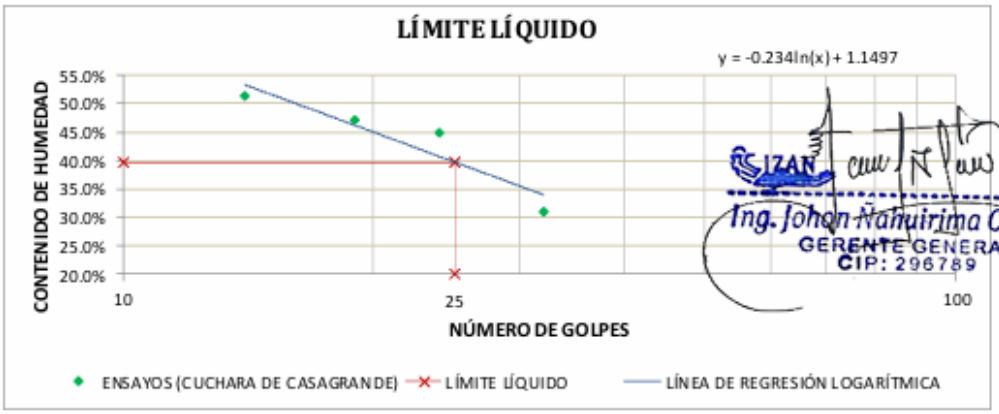
731264.63 m E 8494771.30 m S

CALICATA: N° - 02
PROFUNDIDAD: 1.50 mts

FECHA: Febrero, 2026

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03	MUESTRA 04
Peso de Capsula (gr)	36.70	37.40	38.25	39.10
Peso de Capsula + Muestra Húmeda (gr)	60.00	60.65	61.56	59.93
Peso de Capsula + Muestra Seca (gr)	52.10	53.20	54.35	55.00
Peso del Agua (gr)	7.90	7.45	7.21	4.93
Peso de la Muestra Seca (gr)	15.40	15.80	16.10	15.90
Contenido de Humedad	**51.30%**	**47.15%**	**44.78%**	**31.01%**
Número de Golpes	**14**	**19**	**24**	**32**

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)	36.20	35.80	37.10
Peso de Capsula + Muestra Húmeda (gr)	43.03	42.42	44.08
Peso de Capsula + Muestra Seca (gr)	42.10	41.55	43.15
Peso del Agua (gr)	0.93	0.87	0.93
Peso de la Muestra Seca (gr)	5.90	5.75	6.05
Contenido de Humedad	**15.76%**	**15.13%**	**15.37%**



IZAN
Ing. Johon Nahuirima Cabeza
GERENTE GENERAL
CIP: 296789

LÍMITE LÍQUIDO = 39.80%
LÍMITE PLÁSTICO = 15.42%
ÍNDICE DE PLASTICIDAD = 24.38%

observaciones: -Muestra tomadas en Campo por el Laboratorio

Figura 51 — Limites de consistencia C-02

Repositorio Institucional - UNAMBA Perú

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING AND CONSTRUCTION
LABORATORIO
RUC : 20609203570
997374826 / 985190489
AV. EL ARCO S/N

ENSAYO DE COMPACTACION - PROCTOR MODIFICADO

ASTM D-1557-70 - AASHTO T-180-70

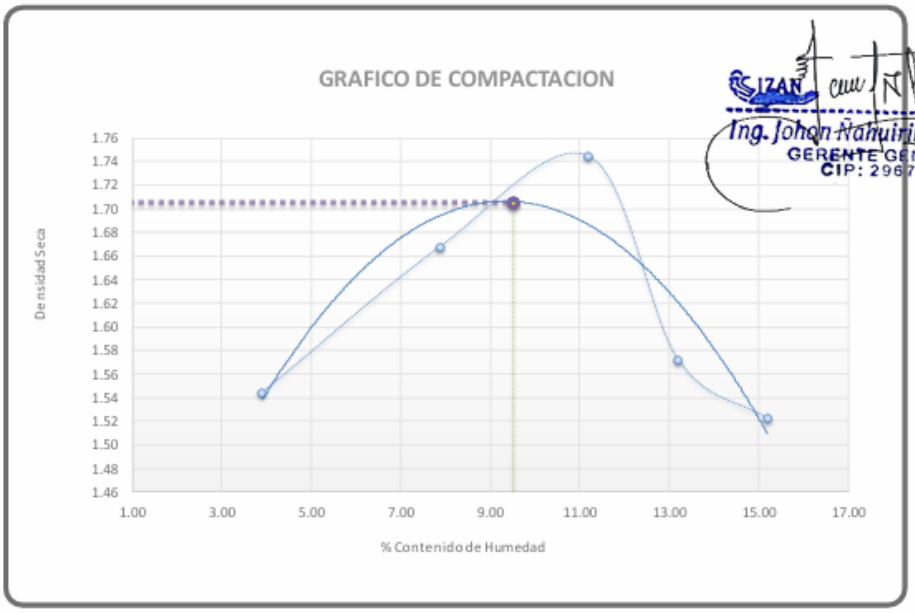
PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

CALICATA: N° - 02 **731264.63 m E** 8494771.30 m S
PROFUNDIDAD: 1.50 mts **FECHA:** Febrero, 2026

N° de Capas: 5				
N° de golpes por caj	25	Volumen Molde :	2124	
Muestra	**1**	**2**	**3**	**4**
Wcap.	46.20	47.10	45.80	48.00
Wcap+sh.	**202.05**	**198.44**	**190.36**	**229.12**
Wcap+ss.	**196.20**	**187.40**	**175.80**	**208.00**
Wss.	150.00	140.30	130.00	160.00
Ww.	5.85	11.04	14.56	21.12
%CH.	3.90	7.87	11.20	13.20
Wmolde	**6580**	**6580**	**6580**	**6580**
Wsh+molde	**9987**	**10400**	**10700**	**10360**
Wsh.	3407	3820	4120	3780
Densidad Hum.	1.604	1.798	1.940	1.780
Densidad Seca	1.544	1.667	1.744	1.572

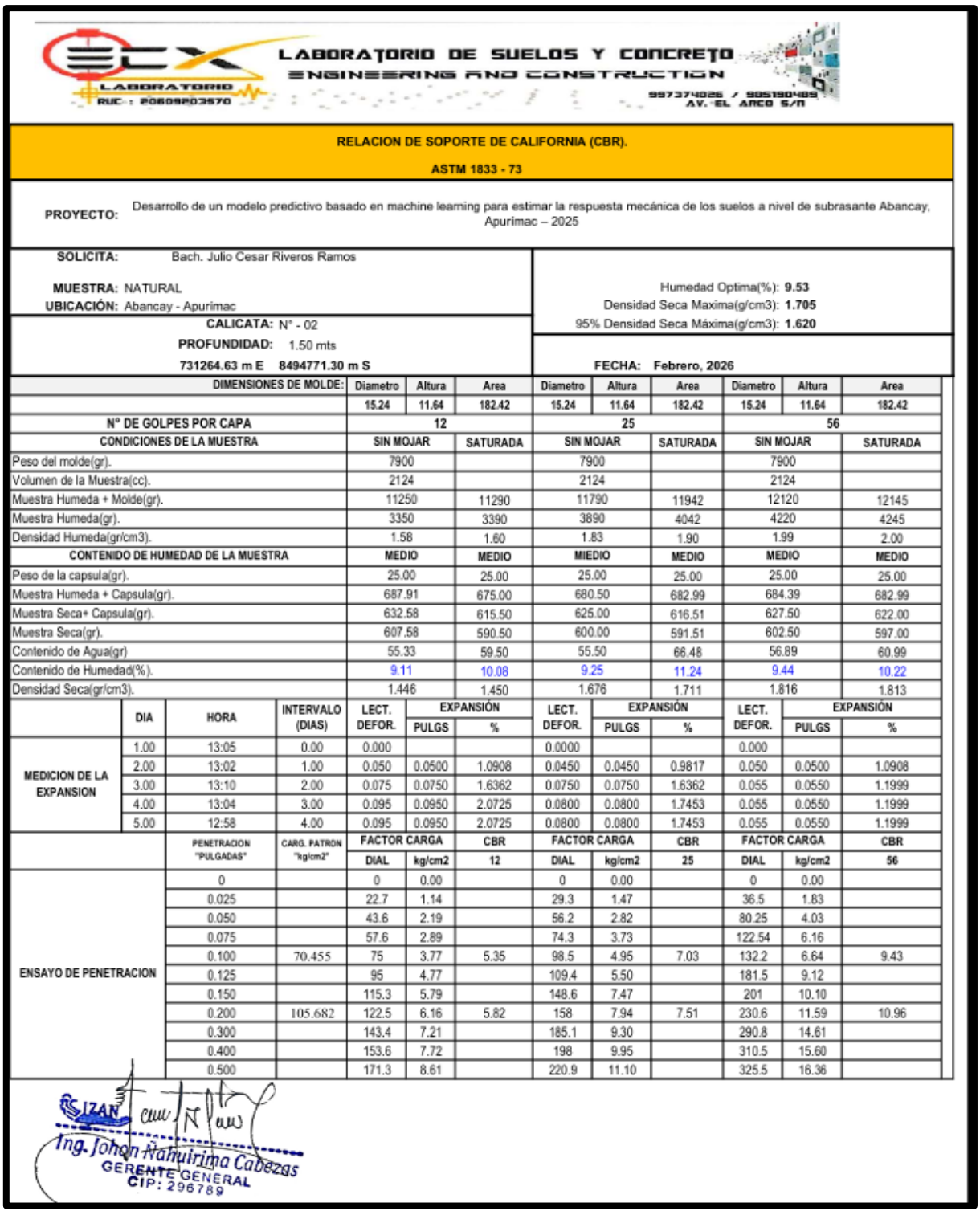
Densidad Seca Maxima(g/cm3) :	1.705
Contenido Optimo de Agua(%) :	9.53



IZAN
Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 52 — Ensayo de compactación C-02

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING AND CONSTRUCTION
LABORATORIO
RUC : 20609203570
957374026 / 985190489
AV. EL ARCO S/N

RELACION DE SOPORTE DE CALIFORNIA (CBR).

ASTM 1833 - 73

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos

MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

Humedad Optima(%): **9.53**
Densidad Seca Maxima(g/cm3): **1.705**
95% Densidad Seca Máxima(g/cm3): **1.620**

CALICATA: N° - 02
PROFUNDIDAD: 1.50 mts
731264.63 m E 8494771.30 m S

FECHA: Febrero, 2026

DIMENSIONES DE MOLDE:	Diametro	Altura	Area	Diametro	Altura	Area	Diametro	Altura	Area
	15.24	11.64	182.42	15.24	11.64	182.42	15.24	11.64	182.42
N° DE GOLPES POR CAPA	12			25			56		
CONDICIONES DE LA MUESTRA	SIN MOJAR		SATURADA	SIN MOJAR		SATURADA	SIN MOJAR		SATURADA
Peso del molde(gr).	7900			7900			7900		
Volumen de la Muestra(cc).	2124			2124			2124		
Muestra Humeda + Molde(gr).	11250		11290	11790		11942	12120		12145
Muestra Humeda(gr).	3350		3390	3890		4042	4220		4245
Densidad Humeda(gr/cm3).	1.58		1.60	1.83		1.90	1.99		2.00
CONTENIDO DE HUMEDAD DE LA MUESTRA	MEDIO		MEDIO	MIEDIO		MEDIO	MEDIO		MEDIO
Peso de la capsula(gr).	25.00		25.00	25.00		25.00	25.00		25.00
Muestra Humeda + Capsula(gr).	687.91		675.00	680.50		682.99	684.39		682.99
Muestra Seca+ Capsula(gr).	632.58		615.50	625.00		616.51	627.50		622.00
Muestra Seca(gr).	607.58		590.50	600.00		591.51	602.50		597.00
Contenido de Agua(gr)	55.33		59.50	55.50		66.48	56.89		60.99
Contenido de Humedad(%).	9.11		10.08	9.25		11.24	9.44		10.22
Densidad Seca(gr/cm3).	1.446		1.450	1.676		1.711	1.816		1.813

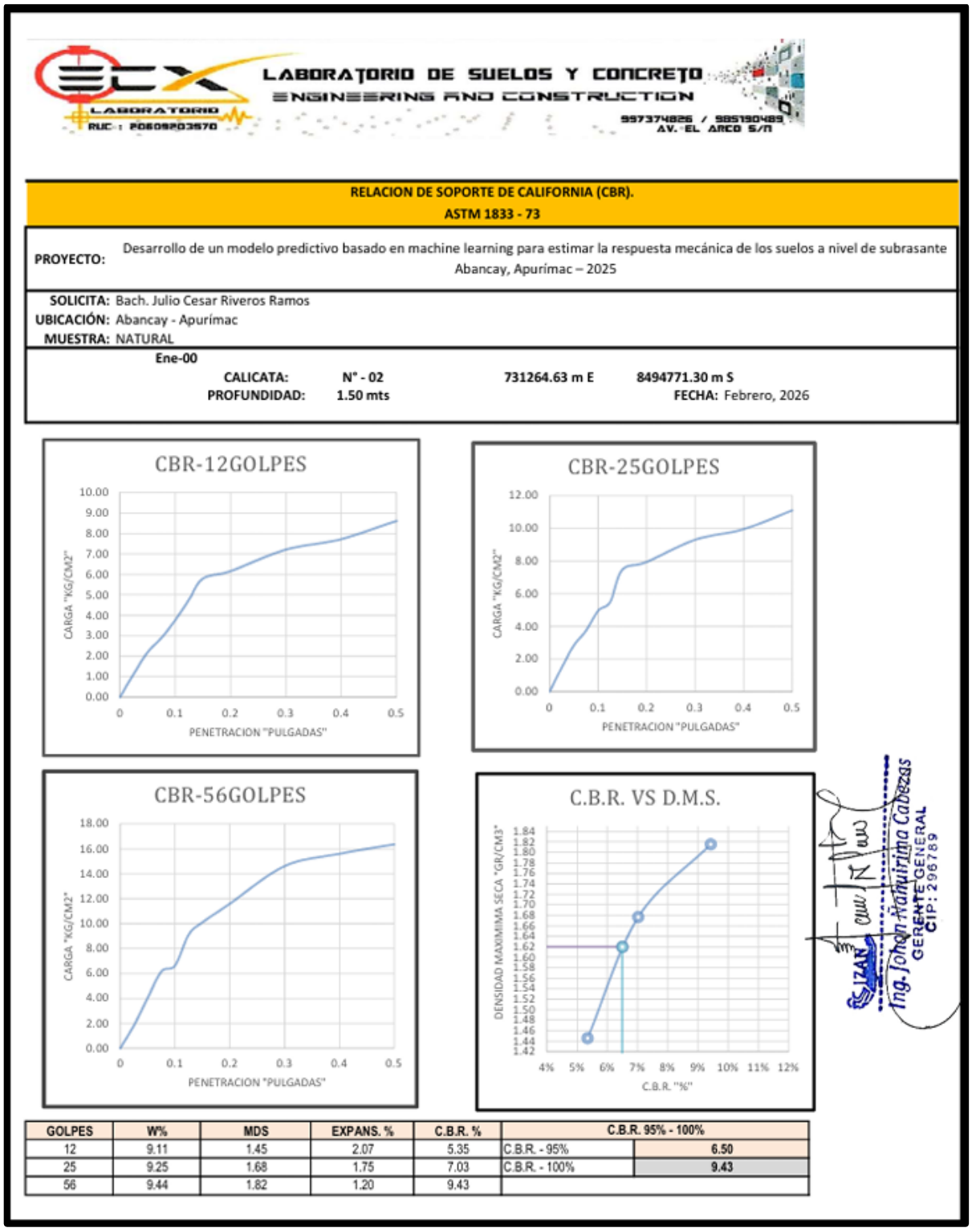
	DIA	HORA	INTERVALO (DIAS)	LECT. DEFOR.	EXPANSIÓN PULGS	EXPANSIÓN %	LECT. DEFOR.	EXPANSIÓN PULGS	EXPANSIÓN %	LECT. DEFOR.	EXPANSIÓN PULGS	EXPANSIÓN %
MEDICION DE LA EXPANSION	1.00	13:05	0.00	0.000			0.0000			0.000		
	2.00	13:02	1.00	0.050	0.0500	1.0908	0.0450	0.0450	0.9817	0.050	0.0500	1.0908
	3.00	13:10	2.00	0.075	0.0750	1.6362	0.0750	0.0750	1.6362	0.055	0.0550	1.1999
	4.00	13:04	3.00	0.095	0.0950	2.0725	0.0800	0.0800	1.7453	0.055	0.0550	1.1999
	5.00	12:58	4.00	0.095	0.0950	2.0725	0.0800	0.0800	1.7453	0.055	0.0550	1.1999

	PENETRACION "PULGADAS"	CARG. PATRON "kg/cm2"	FACTOR CARGA DIAL	FACTOR CARGA kg/cm2	CBR 12	FACTOR CARGA DIAL	FACTOR CARGA kg/cm2	CBR 25	FACTOR CARGA DIAL	FACTOR CARGA kg/cm2	CBR 56
ENSAYO DE PENETRACION	0		0	0.00		0	0.00		0	0.00	
	0.025		22.7	1.14		29.3	1.47		36.5	1.83	
	0.050		43.6	2.19		56.2	2.82		80.25	4.03	
	0.075		57.6	2.89		74.3	3.73		122.54	6.16	
	0.100	70.455	75	3.77	5.35	98.5	4.95	7.03	132.2	6.64	9.43
	0.125		95	4.77		109.4	5.50		181.5	9.12	
	0.150		115.3	5.79		148.6	7.47		201	10.10	
	0.200	105.682	122.5	6.16	5.82	158	7.94	7.51	230.6	11.59	10.96
	0.300		143.4	7.21		185.1	9.30		290.8	14.61	
	0.400		153.6	7.72		198	9.95		310.5	15.60	
	0.500		171.3	8.61		220.9	11.10		325.5	16.36	

IZAN
Ing. Johan Namuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 53 — Ensayo CBR C-02

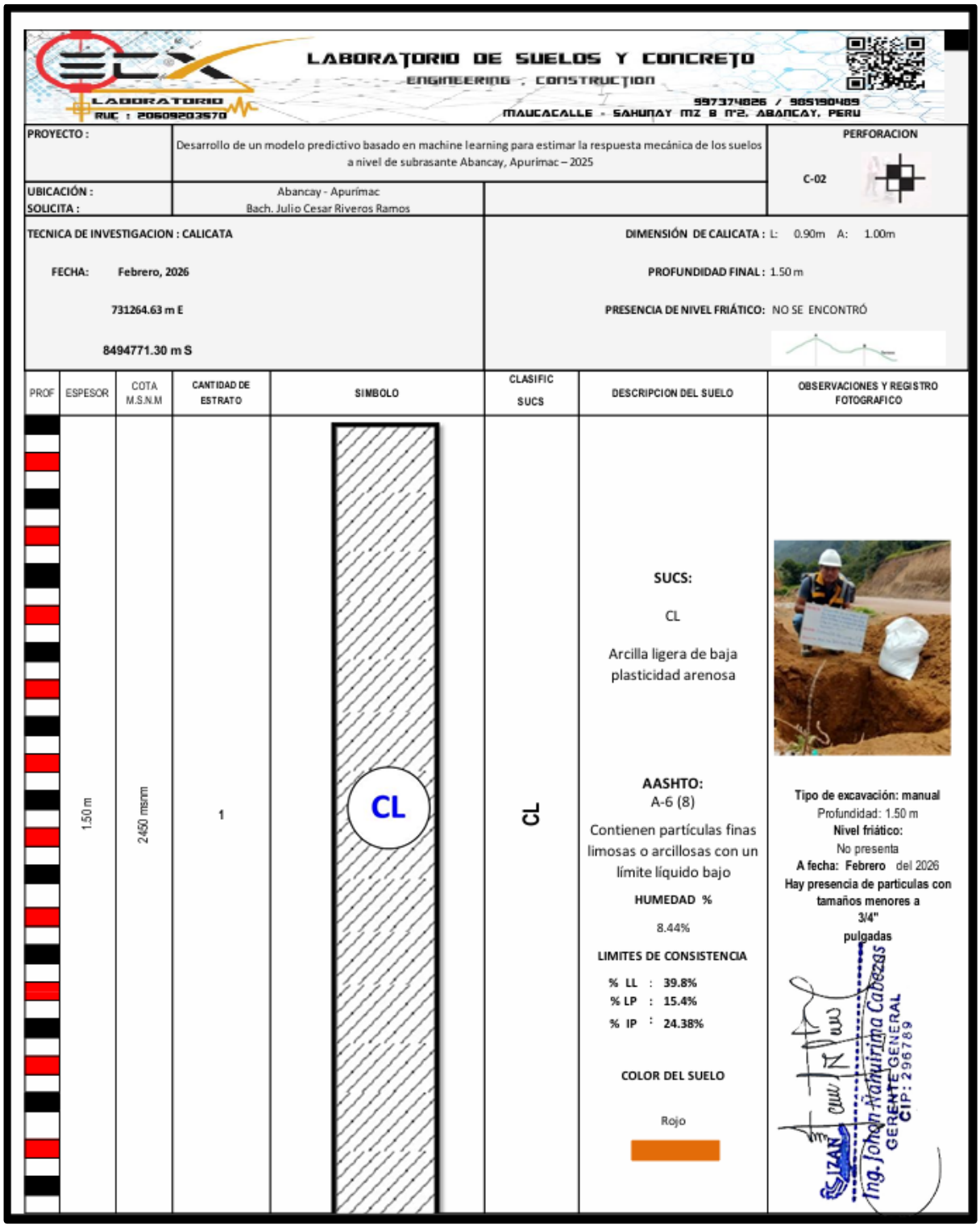
Repositorio Institucional - UNAMBA Perú



GOLPES	W%	MDS	EXPANS. %	C.B.R. %	C.B.R. 95% - 100%	
12	9.11	1.45	2.07	5.35	C.B.R. - 95%	6.50
25	9.25	1.68	1.75	7.03	C.B.R. - 100%	9.43
56	9.44	1.82	1.20	9.43		

Figura 54 — Ensayo CBR-Gráfico C-02

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING - CONSTRUCTION

LABORATORIO RUC : 20609203570

997374826 / 985190489

MAUCACALLE - SAHUNAY MZ B N°2. ABANCAY, PERU

PROYECTO :	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025	PERFORACION C-02
UBICACIÓN :	Abancay - Apurímac	
SOLICITA :	Bach. Julio Cesar Riveros Ramos	

TECNICA DE INVESTIGACION : CALICATA

FECHA: Febrero, 2026

731264.63 m E

8494771.30 m S

DIMENSIÓN DE CALICATA : L: 0.90m A: 1.00m

PROFUNDIDAD FINAL : 1.50 m

PRESENCIA DE NIVEL FRIÁTICO: NO SE ENCONTRÓ

PROF	ESPESOR	COTA M.S.N.M	CANTIDAD DE ESTRATO	SIMBOLO	CLASIFIC SUCS	DESCRIPCION DEL SUELO	OBSERVACIONES Y REGISTRO FOTOGRAFICO
	1.50 m	2450 msnm	1	CL	CL	SUCS: CL Arcilla ligera de baja plasticidad arenosa AASHTO: A-6 (8) Contienen partículas finas limosas o arcillosas con un límite líquido bajo HUMEDAD % 8.44% LIMITES DE CONSISTENCIA % LL : 39.8% % LP : 15.4% % IP : 24.38% COLOR DEL SUELO Rojo	Tipo de excavación: manual Profundidad: 1.50 m Nivel friático: No presenta A fecha: Febrero del 2026 Hay presencia de particulas con tamaños menores a 3/4" pulgadas

Ing. Johan Ivanuirima Cabezas GERENTE GENERAL CIP: 296789

Figura 55 — Perfil estratigráfico C-02

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING , CONSTRUCTION

LABORATORIO

RUC : 20609203570

997374826 / 985190489

MAUCACALLE - SAHUNAY MZ B N°2, ABANCAY, PERU

CONTENIDO DE HUMEDAD DEL SUELO-MTC E 108-2016

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos

MUESTRA: NATURAL

UBICACIÓN: Abancay - Apurímac

730746.12 m E 8494613.26 m S

CALICATA: N° - 03

PROFUNDIDAD: 1.50 mts

FECHA: Febrero, 2026

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)	28.00	28.00	28.00
Peso de Capsula + Muestra Húmeda (gr)	682.30	655.40	701.20
Peso de Capsula + Muestra Seca (gr)	630.00	605.00	647.30
Peso del Agua (gr)	52.30	50.40	53.90
Peso de la Muestra Seca (gr)	602.00	577.00	619.30
Contenido de Humedad	**8.69%**	**8.73%**	**8.70%**

CONTENIDO DE HUMEDAD NATURAL = 8.71%

IZAN

Ing. Johan Nahuirima Cabezas

GERENTE GENERAL

CIP: 296789

Figura 56 — Contenido de humedad C-03

Repositorio Institucional - UNAMBA Perú

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING - CONSTRUCTION
LABORATORIO RUC : 20605203570
997374826 / 905190489
MAUCACALLE - SAHUNAY MZ B LT2. ABANCAY, PERU

ANÁLISIS GRANULOMÉTRICO POR TAMIZADO-MTC E 107- 2016

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

730746.12 m E 8494613.26 m S

CALICATA: N° - 03

PROFUNDIDAD: 1.50 mts

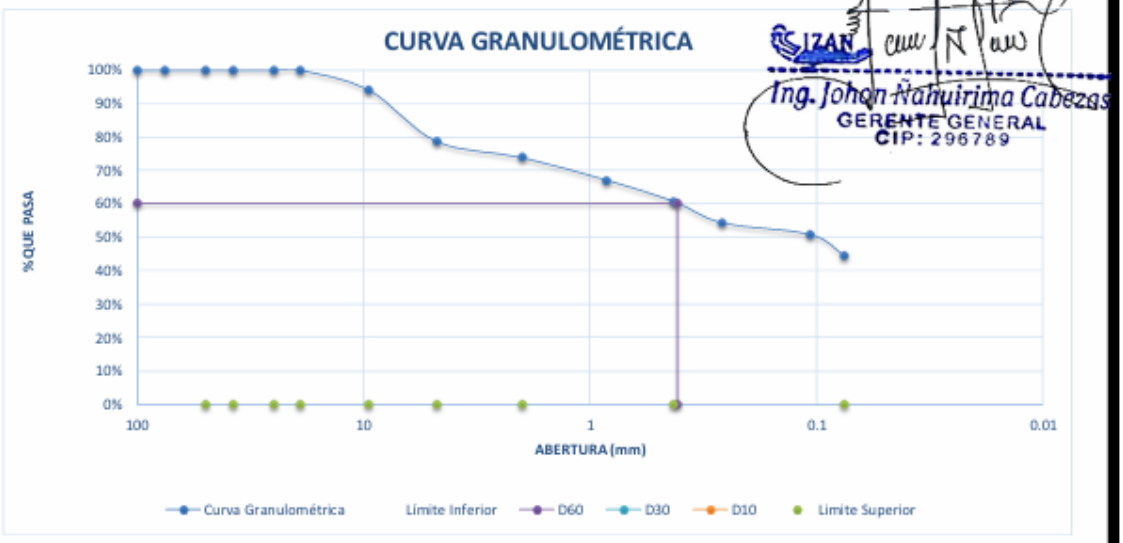
FECHA: Febrero, 2026

Antes del Lavado
Peso de la Muestra Seca = 2700.00 gr

Despues del Lavado
Peso de la Muestra Seca = 1506.00 gr
% de Error en Peso = 0.13% Ok!

TAMIZ	ABERTURA (mm)	PESO RETENIDO	PESO QUE PASA	% RETENIDO	% RETENIDO ACUMULADO	% QUE PASA
4"	100	0.00	1504.0	0.00%	0.00%	100.00%
3"	75	0.00	1504.00	0.00%	0.00%	100.00%
2"	50	0.00	1504.00	0.00%	0.00%	100.00%
1½"	37.5	0.00	1504.00	0.00%	0.00%	100.00%
1"	25	0.00	1504.00	0.00%	0.00%	100.00%
3/4"	19	0.00	1504.00	0.00%	0.00%	100.00%
3/8"	9.5	159.00	1345.00	5.89%	5.89%	94.11%
Nº 4	4.76	419.00	926.00	15.52%	21.41%	78.59%
Nº 10	2	130.00	796.00	4.81%	26.22%	73.78%
Nº 20	0.840	184.00	612.00	6.81%	33.04%	66.96%
Nº 40	0.425	172.00	440.00	6.37%	39.41%	60.59%
Nº 60	0.260	175.00	265.00	6.48%	45.89%	54.11%
Nº 140	0.106	92.00	173.00	3.41%	49.30%	50.70%
Nº 200	0.075	173.00	0.00	6.41%	55.70%	44.30%
Cazuela	-	85.00	-	3.15%	58.85%	-
Lavado	-		-	41.15%	96.85%	-
Total Fracción Retenida en Lavado =		1504.00		100.00%		

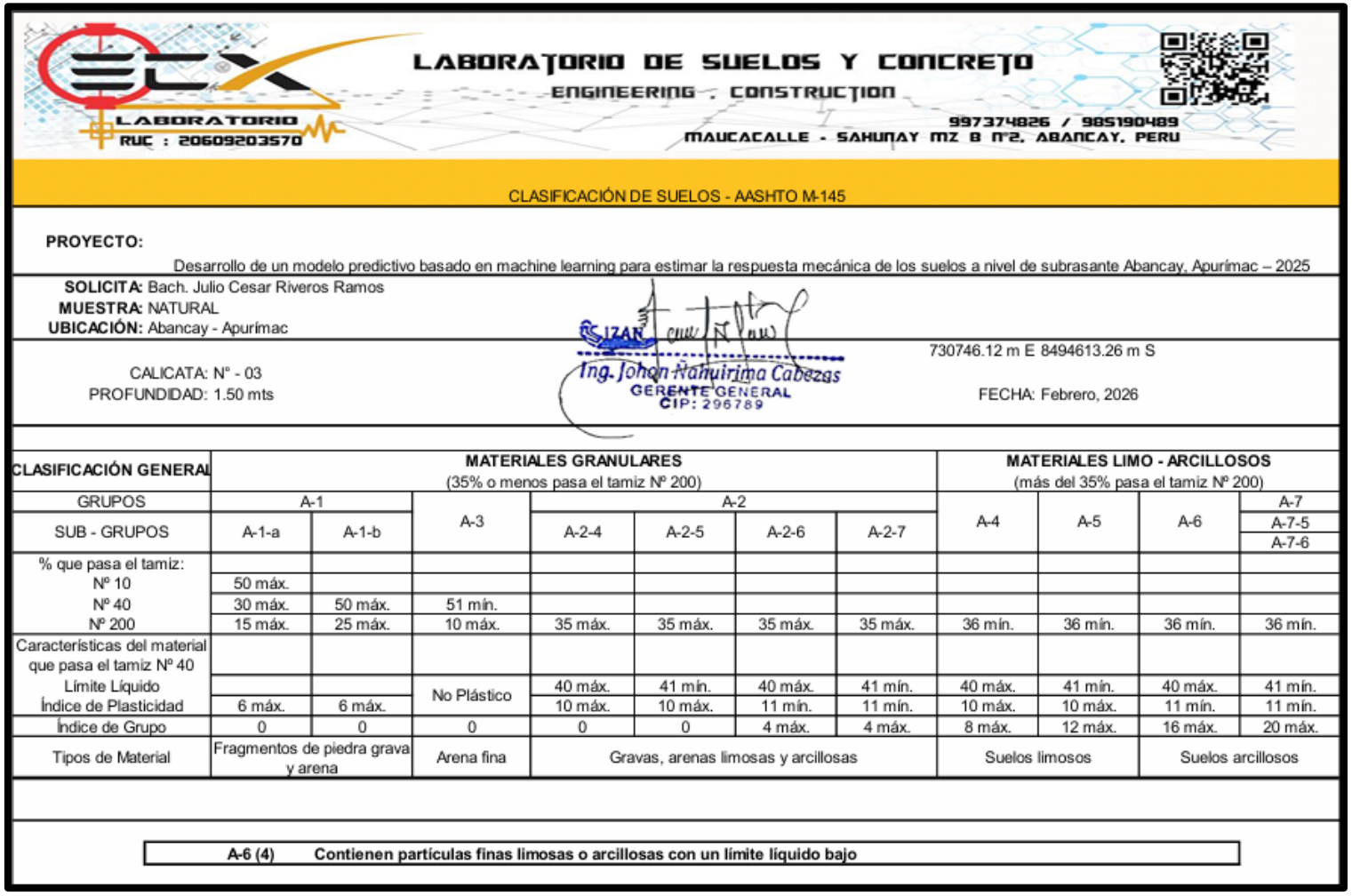
Fracciones Retenidas		De la Fraccion Gruesa		Propiedades
% de Gravas(Ret.n°4) =	21.41%	% de grava =	38.43%	T. Max.Absoluto= 3/4"
% de Arena (Pas N°4-Ret N°200)=	34.30%	% de arena =	61.57%	T. Max.Relativo= 3/8"
% de Finos(Pas N°200)=	44.30%	Total =	100.00%	Clasf. SUCS= SC
Total=	100.00%			Clasif.AASHTO= A-6 (4)



IZAN
Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 57 — Análisis granulométrico C-03

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING - CONSTRUCTION

LABORATORIO
RUC : 20609203570

997374826 / 985190489
MAUCACALLE - SAHUNAY MZ B N°2, ABANCAY, PERU

CLASIFICACIÓN DE SUELOS - AASHTO M-145

PROYECTO:
Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

Ing. Johon Ramirima Cabezas
GERENTE GENERAL
CIP: 296789

730746.12 m E 8494613.26 m S

CALICATA: N° - 03
PROFUNDIDAD: 1.50 mts

FECHA: Febrero, 2026

CLASIFICACIÓN GENERAL	MATERIALES GRANULARES (35% o menos pasa el tamiz Nº 200)							MATERIALES LIMO - ARCILLOSOS (más del 35% pasa el tamiz Nº 200)			
GRUPOS	A-1		A-3	A-2				A-4	A-5	A-6	A-7
SUB - GRUPOS	A-1-a	A-1-b		A-2-4	A-2-5	A-2-6	A-2-7				A-7-5 A-7-6
% que pasa el tamiz:											
Nº 10	50 máx.										
Nº 40	30 máx.	50 máx.	51 mín.								
Nº 200	15 máx.	25 máx.	10 máx.	35 máx.	35 máx.	35 máx.	35 máx.	36 mín.	36 mín.	36 mín.	36 mín.
Características del material que pasa el tamiz Nº 40											
Límite Líquido			No Plástico	40 máx.	41 mín.	40 máx.	41 mín.	40 máx.	41 mín.	40 máx.	41 mín.
Índice de Plasticidad	6 máx.	6 máx.		10 máx.	10 máx.	11 mín.	11 mín.	10 máx.	10 máx.	11 mín.	11 mín.
Índice de Grupo	0	0	0	0	0	4 máx.	4 máx.	8 máx.	12 máx.	16 máx.	20 máx.
Tipos de Material	Fragmentos de piedra grava y arena		Arena fina	Gravas, arenas limosas y arcillosas				Suelos limosos		Suelos arcillosos	

A-6 (4) **Contienen partículas finas limosas o arcillosas con un límite líquido bajo**

Figura 58 — Clasificación de suelos C-03

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING - CONSTRUCTION
LABORATORIO RUC : 20605203570
997374826 / 905190489
MAUCACALLE - SAHUNAY MZ B LT2. ABANCAY, PERU

LÍMITE DE CONSISTENCIA-MTC E 110 - 201

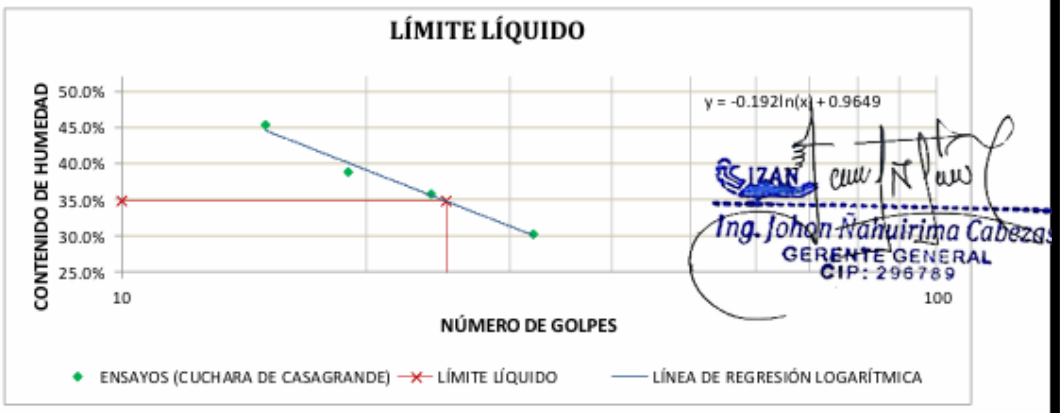
PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

0 730746.12 m E 8494613.26 m S
CALICATA: N° - 03
PROFUNDIDAD: 1.50 mts FECHA: Febrero, 2026

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03	MUESTRA 04
Peso de Capsula (gr)	30.10	30.25	30.05	30.40
Peso de Capsula + Muestra Húmeda (gr)	54.52	54.76	54.30	54.51
Peso de Capsula + Muestra Seca (gr)	46.90	47.90	47.90	48.90
Peso del Agua (gr)	7.62	6.86	6.40	5.61
Peso de la Muestra Seca (gr)	16.80	17.65	17.85	18.50
Contenido de Humedad	**45.36%**	**38.87%**	**35.85%**	**30.32%**
Número de Golpes	**15**	**19**	**24**	**32**

DESCRIPCIÓN	MUESTRA 01	MUESTRA 02	MUESTRA 03
Peso de Capsula (gr)	13.13	13.47	13.62
Peso de Capsula + Muestra Húmeda (gr)	23.81	23.74	23.51
Peso de Capsula + Muestra Seca (gr)	22.25	22.22	22.05
Peso del Agua (gr)	1.56	1.52	1.46
Peso de la Muestra Seca (gr)	9.12	8.75	8.43
Contenido de Humedad	**17.11%**	**17.37%**	**17.32%**



Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

LÍMITE LÍQUIDO = 34.80%
LÍMITE PLÁSTICO = 17.14%
ÍNDICE DE PLASTICIDAD = 17.66%

observaciones: -Muestra tomadas en Campo por el Laboratorio

Figura 59 — Limites de consistencia C-03

Repositorio Institucional - UNAMBA Perú

LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING AND CONSTRUCTION
LABORATORIO
RUC : 20609203570
997374026 / 985190489
AV. EL ARCO S/N

ENSAYO DE COMPACTACION - PROCTOR MODIFICADO
ASTM D-1557-70 - AASHTO T-180-70

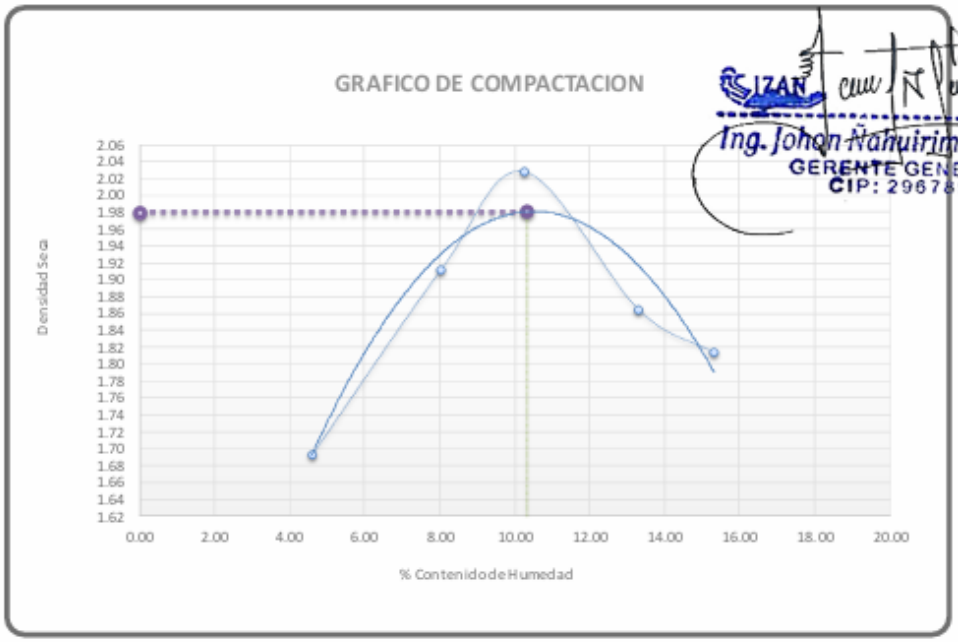
PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos
MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

CALICATA: N° - 03 **730746.12 m E** 8494613.26 m S
PROFUNDIDAD: 1.50 mts **FECHA:** Febrero, 2026

N° de Capas: 5				
N° de golpes por cap	25		Volumen Molde :	2124
Muestra	**1**	**2**	**3**	**4**
Wcap.	55.00	55.00	60.00	60.00
Wcap+sh.	**499.50**	**525.45**	**545.00**	**536.50**
Wcap+ss.	**480.00**	**490.50**	**499.90**	**480.50**
Wss.	425.00	435.50	439.90	420.50
Ww.	19.50	34.95	45.10	56.00
%CH.	4.59	8.03	10.25	13.32
Wmolde	**6370**	**6370**	**6370**	**6370**
Wsh+molde	**10129**	**10756**	**11120**	**10856**
Wsh.	3759	4386	4750	4486
Densidad Hum.	1.770	2.065	2.236	2.112
Densidad Seca	1.692	1.912	2.028	1.864

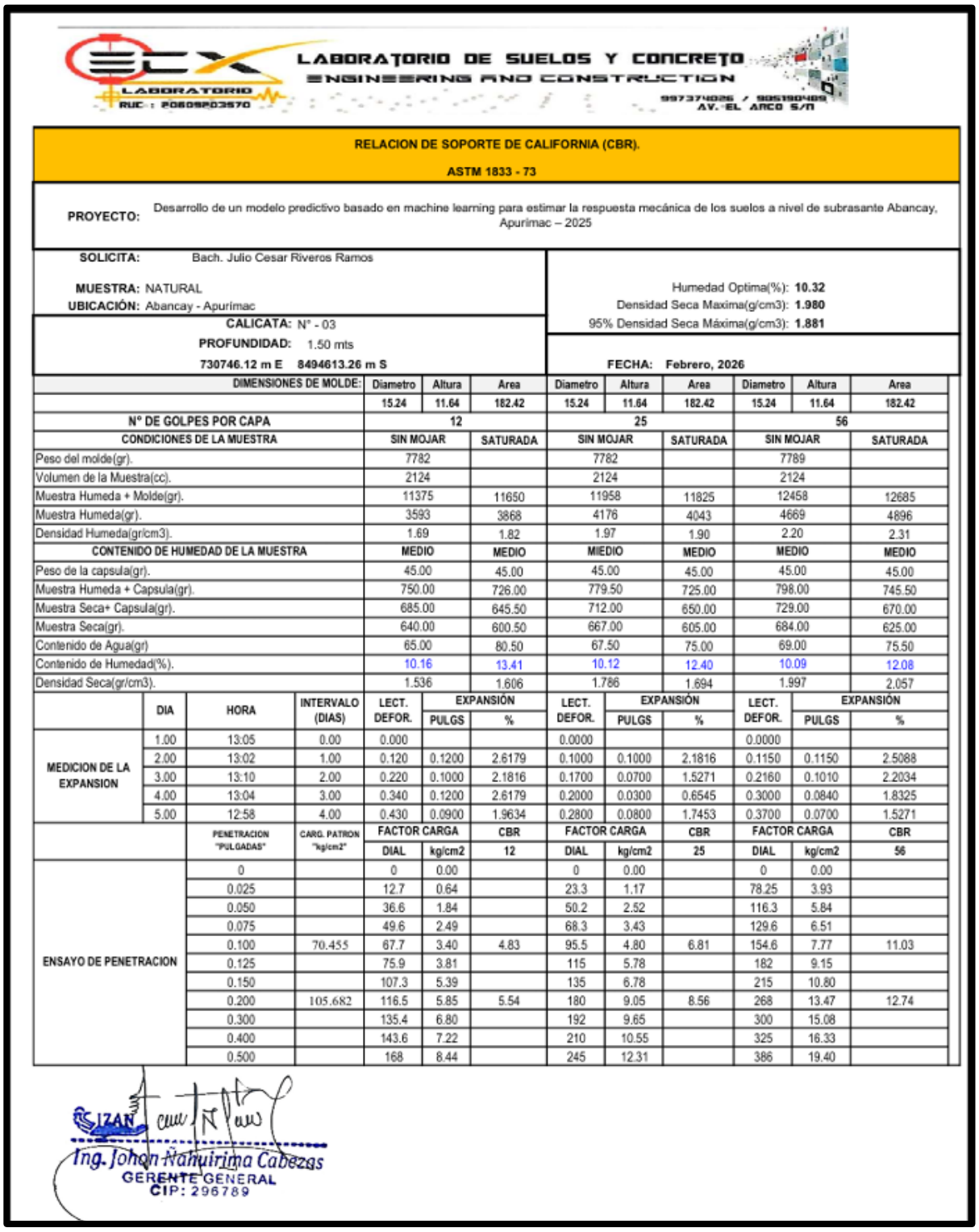
Densidad Seca Maxima(g/cm3) :	1.980
Contenido Optimo de Agua(%) :	10.32



IZAN
Ing. Johan Nahuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 60 — Ensayo de compactación C-03

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO
ENGINEERING AND CONSTRUCTION
LABORATORIO
RUC : 20609203570
997374026 / 905190489
AV. EL ARCO S/N

RELACION DE SOPORTE DE CALIFORNIA (CBR).

ASTM 1833 - 73

PROYECTO: Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025

SOLICITA: Bach. Julio Cesar Riveros Ramos

MUESTRA: NATURAL
UBICACIÓN: Abancay - Apurímac

Humedad Optima(%): **10.32**
Densidad Seca Maxima(g/cm3): **1.980**
95% Densidad Seca Máxima(g/cm3): **1.881**

CALICATA: N° - 03
PROFUNDIDAD: 1.50 mts
730746.12 m E 8494613.26 m S

FECHA: Febrero, 2026

DIMENSIONES DE MOLDE:	Diametro	Altura	Area	Diametro	Altura	Area	Diametro	Altura	Area
	15.24	11.64	182.42	15.24	11.64	182.42	15.24	11.64	182.42
N° DE GOLPES POR CAPA	12			25			56		
CONDICIONES DE LA MUESTRA	SIN MOJAR		SATURADA	SIN MOJAR		SATURADA	SIN MOJAR		SATURADA
Peso del molde(gr).	7782			7782			7789		
Volumen de la Muestra(cc).	2124			2124			2124		
Muestra Humeda + Molde(gr).	11375		11650	11958		11825	12458		12685
Muestra Humeda(gr).	3593		3868	4176		4043	4669		4896
Densidad Humeda(gr/cm3).	1.69		1.82	1.97		1.90	2.20		2.31
CONTENIDO DE HUMEDAD DE LA MUESTRA	MEDIO		MEDIO	MIEDIO		MEDIO	MEDIO		MEDIO
Peso de la capsula(gr).	45.00		45.00	45.00		45.00	45.00		45.00
Muestra Humeda + Capsula(gr).	750.00		726.00	779.50		725.00	798.00		745.50
Muestra Seca+ Capsula(gr).	685.00		645.50	712.00		650.00	729.00		670.00
Muestra Seca(gr).	640.00		600.50	667.00		605.00	684.00		625.00
Contenido de Agua(gr)	65.00		80.50	67.50		75.00	69.00		75.50
Contenido de Humedad(%).	10.16		13.41	10.12		12.40	10.09		12.08
Densidad Seca(gr/cm3).	1.536		1.606	1.786		1.694	1.997		2.057

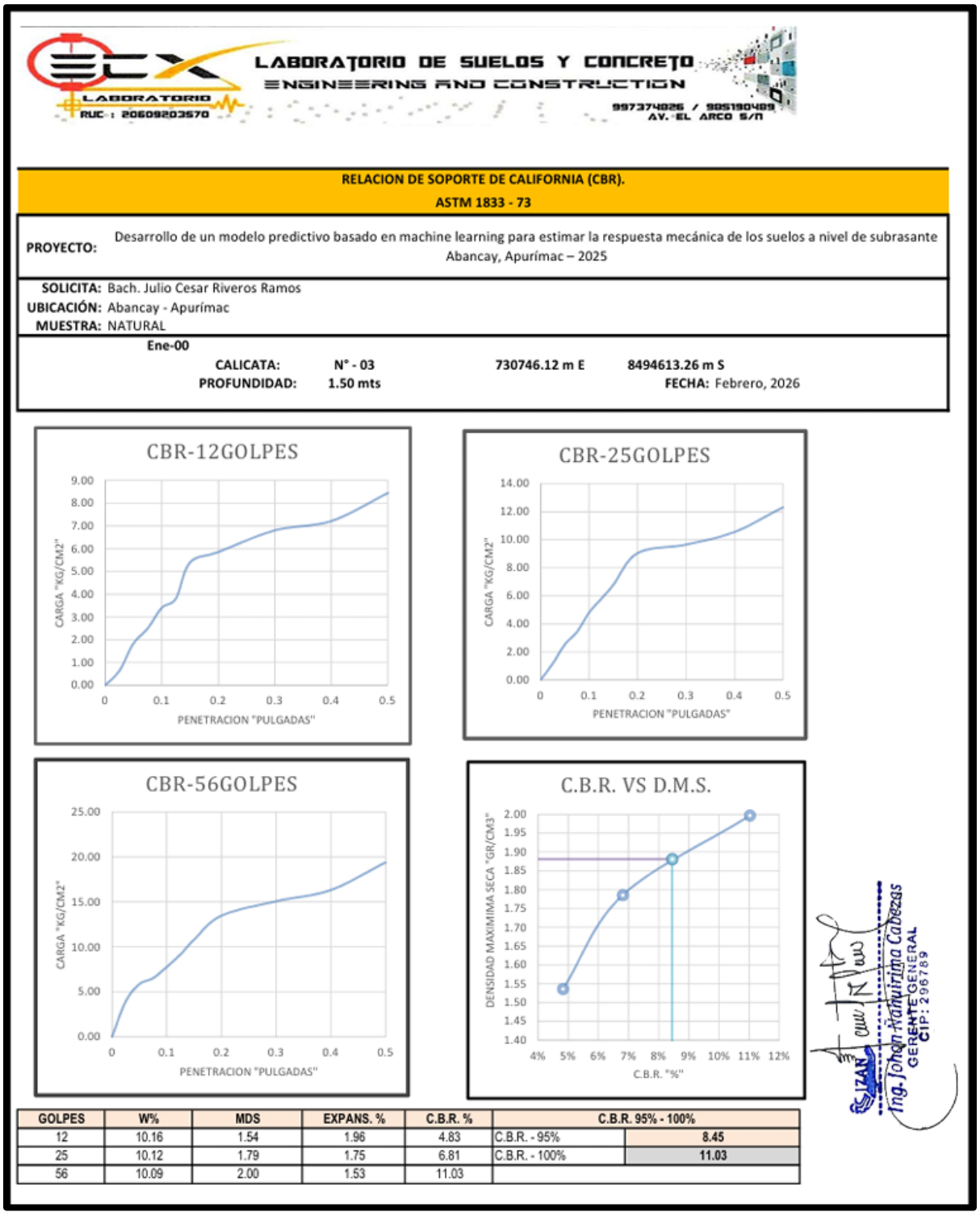
	DIA	HORA	INTERVALO (DIAS)	LECT. DEFOR.	EXPANSIÓN		LECT. DEFOR.	EXPANSIÓN		LECT. DEFOR.	EXPANSIÓN	
					PULGS	%		PULGS	%		PULGS	%
MEDICION DE LA EXPANSION	1.00	13:05	0.00	0.000			0.0000			0.0000		
	2.00	13:02	1.00	0.120	0.1200	2.6179	0.1000	0.1000	2.1816	0.1150	0.1150	2.5088
	3.00	13:10	2.00	0.220	0.1000	2.1816	0.1700	0.0700	1.5271	0.2160	0.1010	2.2034
	4.00	13:04	3.00	0.340	0.1200	2.6179	0.2000	0.0300	0.6545	0.3000	0.0840	1.8325
	5.00	12:58	4.00	0.430	0.0900	1.9634	0.2800	0.0800	1.7453	0.3700	0.0700	1.5271

	PENETRACION "PULGADAS"	CARG. PATRON "kg/cm2"	FACTOR CARGA		CBR	FACTOR CARGA		CBR	FACTOR CARGA		CBR
			DIAL	kg/cm2	12	DIAL	kg/cm2	25	DIAL	kg/cm2	56
ENSAYO DE PENETRACION	0		0	0.00		0	0.00		0	0.00	
	0.025		12.7	0.64		23.3	1.17		78.25	3.93	
	0.050		36.6	1.84		50.2	2.52		116.3	5.84	
	0.075		49.6	2.49		68.3	3.43		129.6	6.51	
	0.100	70.455	67.7	3.40	4.83	95.5	4.80	6.81	154.6	7.77	11.03
	0.125		75.9	3.81		115	5.78		182	9.15	
	0.150		107.3	5.39		135	6.78		215	10.80	
	0.200	105.682	116.5	5.85	5.54	180	9.05	8.56	268	13.47	12.74
	0.300		135.4	6.80		192	9.65		300	15.08	
	0.400		143.6	7.22		210	10.55		325	16.33	
	0.500		168	8.44		245	12.31		386	19.40	

Ing. Johon Namuirima Cabezas
GERENTE GENERAL
CIP: 296789

Figura 61 — Ensayo CBR C-03

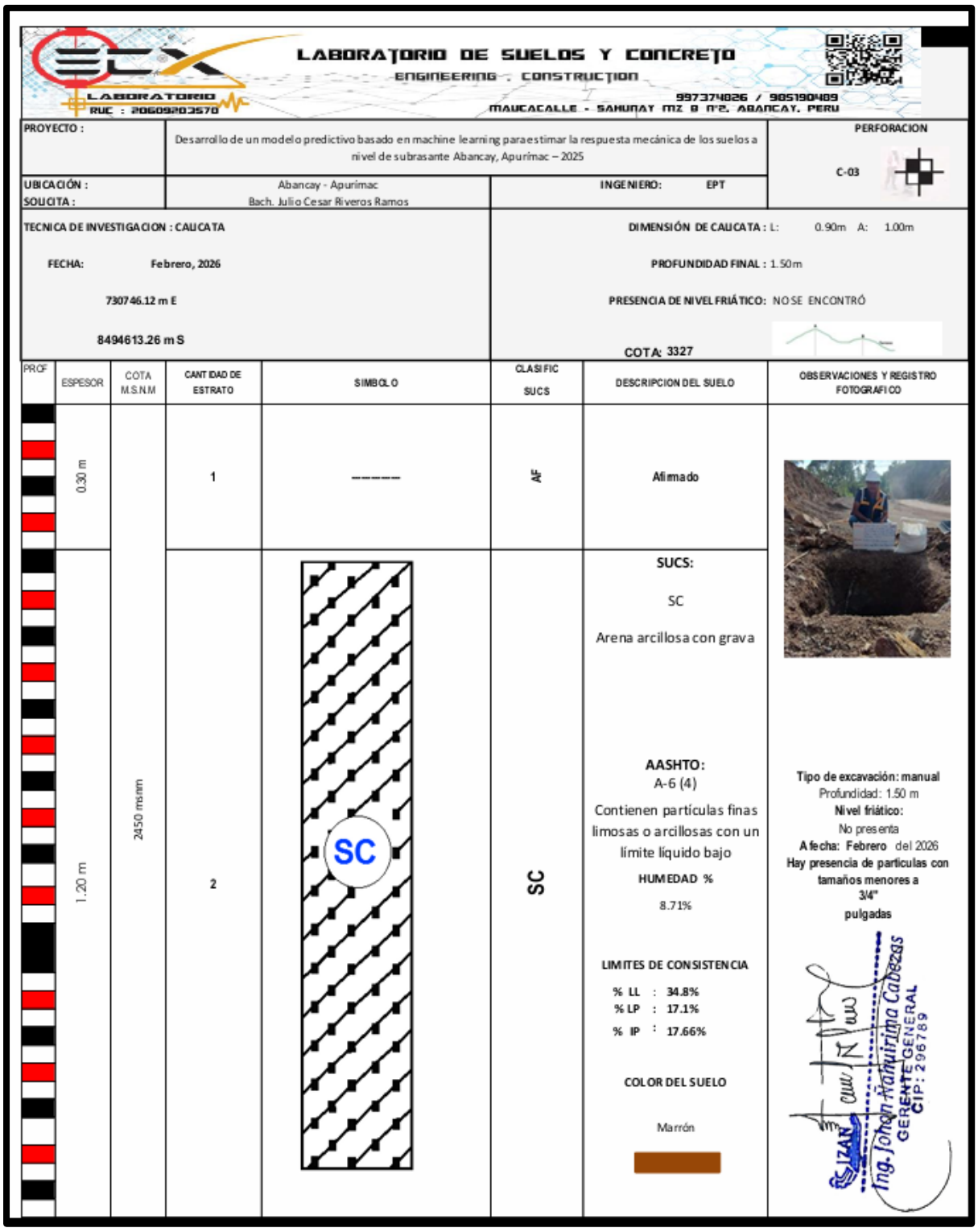
Repositorio Institucional - UNAMBA Perú



GOLPES	W%	MDS	EXPANS. %	C.B.R. %	C.B.R. 95% - 100%	
12	10.16	1.54	1.96	4.83	C.B.R. - 95%	8.45
25	10.12	1.79	1.75	6.81	C.B.R. - 100%	11.03
56	10.09	2.00	1.53	11.03		

Figura 62 — Ensayo CBR-Gráfico C-03

Repositorio Institucional - UNAMBA Perú



LABORATORIO DE SUELOS Y CONCRETO

ENGINEERING . CONSTRUCTION

LABORATORIO RUC : 20609203570

997374026 / 905190489

MAUCACALLE - SAHUMAY MZ B N°2. ABANCAY. PERU

PROYECTO :	Desarrollo de un modelo predictivo basado en machine learning para estimar la respuesta mecánica de los suelos a nivel de subrasante Abancay, Apurímac – 2025	PERFORACION C-03
UBICACIÓN : Abancay - Apurímac	INGENIERO: EPT	
SOLICITA : Bach. Julio Cesar Riveros Ramos		

TECNICA DE INVESTIGACION : CALICATA

FECHA: Febrero, 2026

730746.12 m E

8494613.26 m S

DIMENSIÓN DE CALICATA : L: 0.90m A: 1.00m

PROFUNDIDAD FINAL : 1.50m

PRESENCIA DE NIVEL FRIÁTICO: NO SE ENCONTRÓ

COTA: 3327

PROF	ESPESOR	COTA M.S.N.M	CANTIDAD DE ESTRATO	SIMBOLO	CLASIFIC SUCS	DESCRIPCION DEL SUELO	OBSERVACIONES Y REGISTRO FOTOGRAFICO
	0.30 m	2450 msnm	1	-------	AF	Afirmado	
	1.20 m		2	SC	SC	SUCS: SC Arena arcillosa con grava. AASHTO: A-6 (4) Contienen partículas finas limosas o arcillosas con un límite líquido bajo. HUMEDAD % 8.71%. LIMITES DE CONSISTENCIA % LL : 34.8% % LP : 17.1% % IP : 17.66%. COLOR DEL SUELO Marrón	Tipo de excavación: manual Profundidad: 1.50 m Nivel friático: No presenta A fecha: Febrero del 2026 Hay presencia de particulas con tamaños menores a 3/4" pulgadas

Ing. Johan Atahuirima Cabezas GERENTE GENERAL CIP: 296789

Figura 63 — Perfil estratigráfico C-03

Repositorio Institucional - UNAMBA Perú

Anexo D

PANEL FOTOGRÁFICO



Figura 64 — Exploración C-01



Figura 65 — Exploración C-02



Figura 66 — Exploración C-03

Repositorio Institucional - UNAMBA Perú



Figura 67 — Ensayo Análisis granulométrico

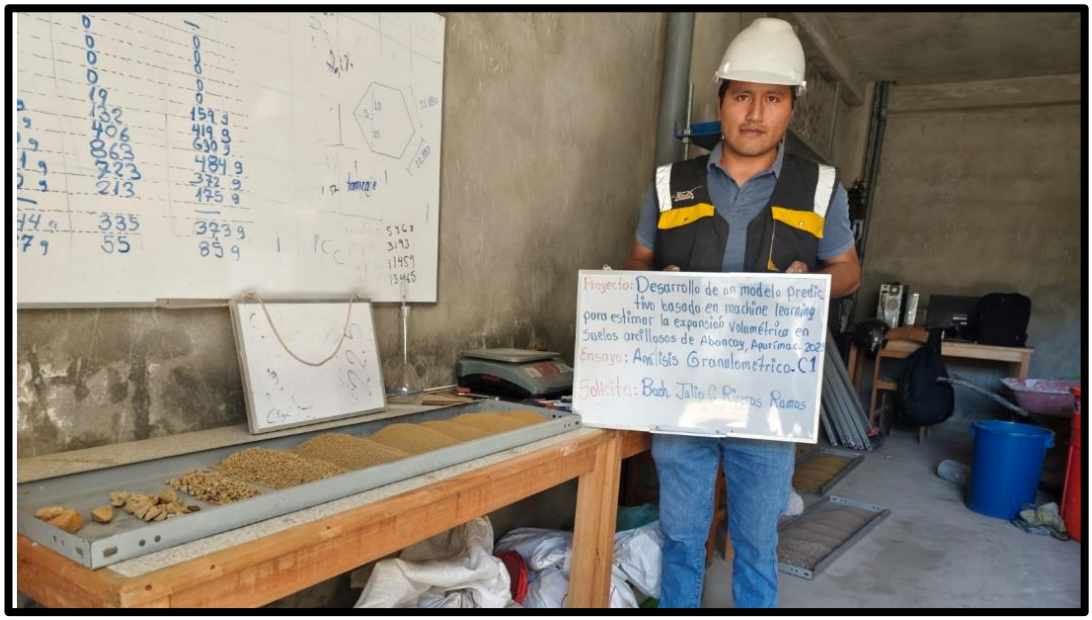


Figura 68 — Ensayo Análisis granulométrico C-01

Repositorio Institucional - UNAMBA Perú



Figura 69 — Ensayo Análisis granulométrico C-02



Figura 70 — Ensayo Análisis granulométrico C-03

Repositorio Institucional - UNAMBA Perú

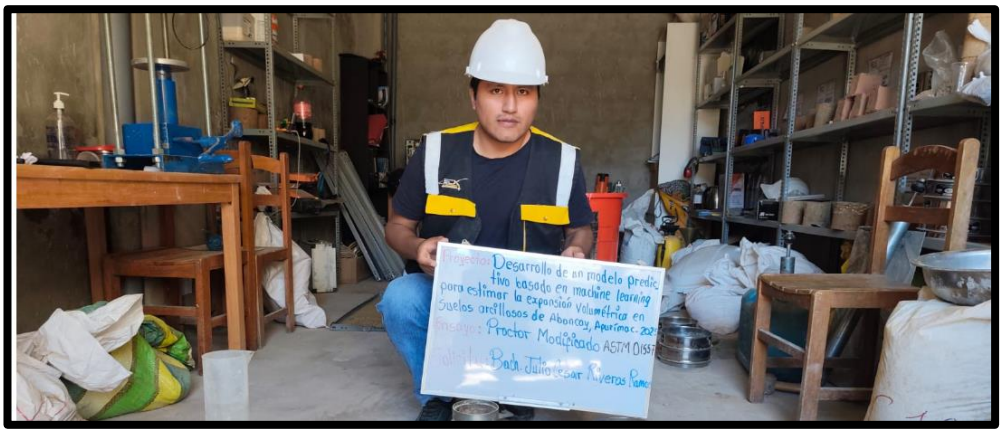


Figura 71 — Ensayo de compactación



Figura 72 — Ensayo de compactación – Proctor modificado



Figura 73 — Ensayo de compactación – Moldes

Repositorio Institucional - UNAMBA Perú



Figura 74 — Ensayo CBR C-01



Figura 75 — Ensayo CBR C-02

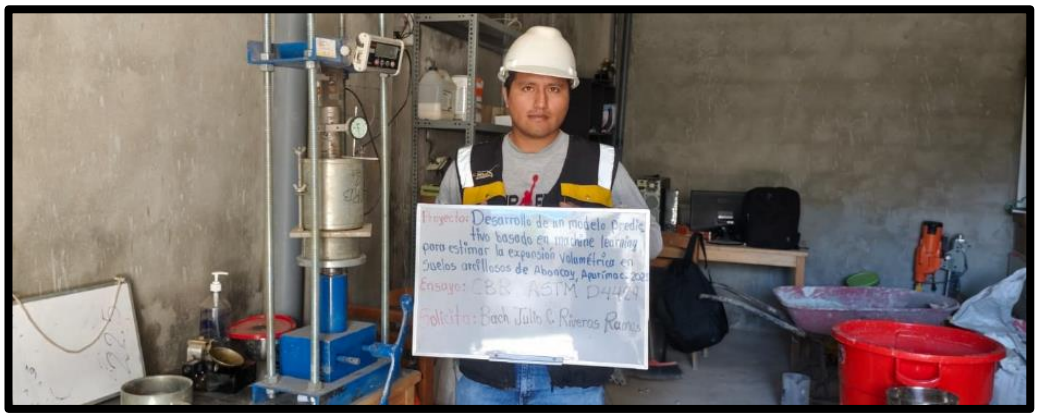


Figura 76 — Ensayo CBR C-03

Repositorio Institucional - UNAMBA Perú

Anexo E

CERTIFICADO DE VALIDACIÓN DE INSTRUMENTOS

Tabla 30 — Validación de instrumentos (Alfa de Cronbach)

Experto	1	2	3	4	5	6	7	8	9	10	Suma
Dr. Edwar Ilasaca Cahuata	5	4	4	5	4	4	5	4	4	5	**44**
Mtr. Américo Roldán Juárez	5	5	5	5	5	4	5	5	5	5	**49**
Mtr. Katherine Astrid Llerena Torreblanca	5	5	5	5	4	5	5	5	5	5	**49**
Ing. Máximo Villegas Vargas	5	4	5	5	5	4	5	5	4	5	**47**
Varianza	*0.00*	*0.25*	*0.19*	*0.00*	*0.25*	*0.19*	*0.00*	*0.19*	*0.25*	*0.00*	
Sumatoria de varianzas	***1.31***										
Varianza de la suma de los ítems	***4.19***										

$$\alpha = \frac{k}{k-1}\left[1 - \frac{\Sigma_{i=1}^{k} S_i^2}{S_T^2}\right] \qquad \text{(E.1)}$$

donde

- α es el coeficiente de Cronbach (0.76);
- k es el número de ítem (10);
- $\sum S_i^2$ es la sumatoria de las varianzas de los ítems (1.31);
- S_T^2 es la varianza de la suma de los ítems (4.19).

- **Confiabilidad.** El coeficiente alfa de Cronbach obtenido fue de **0,76**, valor que se encuentra dentro del rango de confiabilidad aceptable. Este resultado evidencia que las calificaciones otorgadas por los cuatro expertos presentan una adecuada consistencia interna. En consecuencia, el instrumento denominado Ficha de recolección de datos geotécnicos presenta condiciones favorables para su aplicación en la investigación.

Repositorio Institucional - UNAMBA Perú

E.1 Valoración del experto 01

FORMATO ÚNICO DE VALIDACIÓN DE INSTRUMENTO

I. INFORMACIÓN GENERAL

Investigador	Bach. Julio Cesar Riveros Ramos	DNI	77227392
Título de la investigación	Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025		
Universidad	Universidad Nacional Micaela Bastidas de Apurímac		
Experto	EDWAR ILASACA CAHUATA	DNI	01228290
Grado académico	Docto Ciencias y Tecnologia y Medio Amb.		
Institución laboral	UNAMBA		
Nombre del instrumento	Ficha de recolección de datos geotécnicos		

1 – Deficiente, 2 – Regular, 3 – Aceptable, 4 – Muy buena, 5 – Excelente.

II. ASPECTOS DE VALIDACIÓN

	CRITERIOS	Descripción	1	2	3	4	5	OBSERVACIONES
1	Pertinencia	Los datos registrados son pertinentes para cumplir los objetivos de la investigación.					X	
2	Coherencia	Existe correspondencia entre las variables del instrumento y las variables de la investigación.				X		
3	Congruencia	Los datos registrados por el instrumento guardan coherencia con el marco teórico y la metodología de la investigación.				X		
4	Objetividad	La información registrada puede medirse de manera objetiva, evitando interpretaciones subjetivas.					X	
5	Validez técnica	Los parámetros geotécnicos incluidos corresponden a normas y procedimientos técnicos aplicables.				X		
6	Suficiencia	El instrumento contiene toda la información necesaria para desarrollar el modelo predictivo.				X		
7	Organización	La estructura del instrumento facilita el registro ordenado de la información geotécnica.					X	
8	Claridad	Las instrucciones y los campos del instrumento son claros y comprensibles.				X		
9	Formato	El instrumento presenta una adecuada organización, legibilidad y estructura para facilitar el registro de la información.				X		
10	Aplicabilidad	El instrumento puede utilizarse de manera práctica para recopilar los datos requeridos en la investigación.					X	

	A	B	C	D	E	
Conteo total				24	20	
Coeficiente de validación: $\frac{A+B+C+D+E}{50}$						Resultado: 0.88
Luego de Revisado el instumento procede su aplicación :	SI	NO				

UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURIMAC
Dr. Edwar Ilasaca Cahuata
DOCENTE

Firma del experto:

Fecha: 02 / 02 / 2026

Figura 77 — Valoración del experto 01

Repositorio Institucional - UNAMBA Perú

E.2 Valoración del experto 02

FORMATO ÚNICO DE VALIDACIÓN DE INSTRUMENTO

I. INFORMACIÓN GENERAL

Investigador	Bach. Julio Cesar Riveros Ramos	**DNI**	77227392
Título de la investigación	Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025		
Universidad	Universidad Nacional Micaela Bastidas de Apurímac		
Experto	Mtr. Américo Roldan Juárez	**DNI**	45043748
Grado académico	Maestro		
Institución laboral	Docente – UNAMBA y UTEA		
Nombre del instrumento	Ficha de recolección de datos geotécnicos		

1 – Deficiente, 2 – Regular, 3 – Aceptable, 4 – Muy buena, 5 – Excelente.

II. ASPECTOS DE VALIDACIÓN

	CRITERIOS	Descripción	1	2	3	4	5	OBSERVACIONES
1	Pertinencia	Los datos registrados son pertinentes para cumplir los objetivos de la investigación.					X	
2	Coherencia	Existe correspondencia entre las variables del instrumento y las variables de la investigación.					X	
3	Congruencia	Los datos registrados por el instrumento guardan coherencia con el marco teórico y la metodología de la investigación.					X	
4	Objetividad	La información registrada puede medirse de manera objetiva, evitando interpretaciones subjetivas.					X	
5	Validez técnica	Los parámetros geotécnicos incluidos corresponden a normas y procedimientos técnicos aplicables.					X	
6	Suficiencia	El instrumento contiene toda la información necesaria para desarrollar el modelo predictivo.				X		
7	Organización	La estructura del instrumento facilita el registro ordenado de la información geotécnica.					X	
8	Claridad	Las instrucciones y los campos del instrumento son claros y comprensibles.					X	
9	Formato	El instrumento presenta una adecuada organización, legibilidad y estructura para facilitar el registro de la información.					X	
10	Aplicabilidad	El instrumento puede utilizarse de manera práctica para recopilar los datos requeridos en la investigación.					X	

	A	B	C	D	E	
Conteo total				4	45	
Coeficiente de validación: $\frac{A+B+C+D+E}{50}$						Resultado: 0.98
Luego de Revisado el instumento procede su aplicación :	X	NO				

UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURIMAC
Mtr Américo Roldán Juárez
DOCENTE

Firma del experto:

Fecha: 01 / 02 / 2026

Figura 78 — Valoración del experto 02

Repositorio Institucional - UNAMBA Perú

E.3 Valoración del experto 03

FORMATO ÚNICO DE VALIDACIÓN DE INSTRUMENTO

I. INFORMACIÓN GENERAL

Investigador	Bach. Julio Cesar Riveros Ramos		DNI	77227392
Título de la investigación	Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025			
Universidad	Universidad Nacional Micaela Bastidas de Apurímac			
Experto	Katherine Astrid Llerena Torreblanca		DNI	70761859
Grado académico	Mtro. (Maestro en Gerencia de la Construcción)			
Institución laboral	Consorcio Señor de Huanca - Supervisión			
Nombre del instrumento	Ficha de recolección de datos geotécnicos			

1 – Deficiente, 2 – Regular, 3 – Aceptable, 4 – Muy buena, 5 – Excelente.

II. ASPECTOS DE VALIDACIÓN

	CRITERIOS	Descripción	1	2	3	4	5	OBSERVACIONES
1	Pertinencia	Los datos registrados son pertinentes para cumplir los objetivos de la investigación.					X	
2	Coherencia	Existe correspondencia entre las variables del instrumento y las variables de la investigación.					X	
3	Congruencia	Los datos registrados por el instrumento guardan coherencia con el marco teórico y la metodología de la investigación.					X	
4	Objetividad	La información registrada puede medirse de manera objetiva, evitando interpretaciones subjetivas.					X	
5	Validez técnica	Los parámetros geotécnicos incluidos corresponden a normas y procedimientos técnicos aplicables.				X		
6	Suficiencia	El instrumento contiene toda la información necesaria para desarrollar el modelo predictivo.					X	
7	Organización	La estructura del instrumento facilita el registro ordenado de la información geotécnica.					X	
8	Claridad	Las instrucciones y los campos del instrumento son claros y comprensibles.					X	
9	Formato	El instrumento presenta una adecuada organización, legibilidad y estructura para facilitar el registro de la información.					X	
10	Aplicabilidad	El instrumento puede utilizarse de manera práctica para recopilar los datos requeridos en la investigación.					X	

	A	B	C	D	E	
Conteo total				4	45	
Coeficiente de validación: $\frac{A+B+C+D+E}{50}$						Resultado: 0.98
Luego de Revisado el instumento procede su aplicación :	SI	NO				

Firma del experto: Katherine A. Llerena Torreblanca, INGENIERO CIVIL, CIP 168500

Fecha: 02/02/2026

Figura 79 — Valoración del experto 03

Repositorio Institucional - UNAMBA Perú

E.4 Valoración del experto 04

FORMATO ÚNICO DE VALIDACIÓN DE INSTRUMENTO

I. INFORMACIÓN GENERAL

Investigador	Bach. Julio Cesar Riveros Ramos	**DNI**	77227392
Título de la investigación	Modelo predictivo basado en Machine Learning para estimar la respuesta mecánica de los suelos de subrasante de la vía de evitamiento de Abancay, Apurímac – 2025		
Universidad	Universidad Nacional Micaela Bastidas de Apurímac		
Experto	Ing. Máximo Villegas Vargas	**DNI**	80489943
Grado académico	Ing. Civil - Especialista en Peritaje en Ing. Civil		
Institución laboral	Poder Judicial		
Nombre del instrumento	Ficha de recolección de datos geotécnicos		

1 – Deficiente, 2 – Regular, 3 – Aceptable, 4 – Muy buena, 5 – Excelente.

II. ASPECTOS DE VALIDACIÓN

	CRITERIOS	Descripción	1	2	3	4	5	OBSERVACIONES
1	Pertinencia	Los datos registrados son pertinentes para cumplir los objetivos de la investigación.					X	
2	Coherencia	Existe correspondencia entre las variables del instrumento y las variables de la investigación.				X		
3	Congruencia	Los datos registrados por el instrumento guardan coherencia con el marco teórico y la metodología de la investigación.					X	
4	Objetividad	La información registrada puede medirse de manera objetiva, evitando interpretaciones subjetivas.					X	
5	Validez técnica	Los parámetros geotécnicos incluidos corresponden a normas y procedimientos técnicos aplicables.					X	
6	Suficiencia	El instrumento contiene toda la información necesaria para desarrollar el modelo predictivo.				X		
7	Organización	La estructura del instrumento facilita el registro ordenado de la información geotécnica.					X	
8	Claridad	Las instrucciones y los campos del instrumento son claros y comprensibles.					X	
9	Formato	El instrumento presenta una adecuada organización, legibilidad y estructura para facilitar el registro de la información.				X		
10	Aplicabilidad	El instrumento puede utilizarse de manera práctica para recopilar los datos requeridos en la investigación.					X	
		Conteo total	A	B	C	D	E	
						12	35	
		Coeficiente de validación: $\frac{A+B+C+D+E}{50}$						Resultado: 0.94
		Luego de Revisado el instumento procede su aplicación :	SI	NO				

Maximo M. Villegas Vargas
INGENIERO CIVIL
CIP: N° 77980

Firma del experto:

Fecha: 02 / 02 / 2026

Figura 80 — Valoración del Experto 04

Repositorio Institucional - UNAMBA Perú

Anexo F

BASE DE DATOS DE LOS PARÁMETROS GEOTÉCNICOS UTILIZADOS EN LA INVESTIGACIÓN

Tabla 31 — Base de datos geotécnica empleada para el desarrollo del modelo predictivo basado en Machine Learning

ID	Proyecto	Profundidad_m	Humedad_natural	Limite_Liquido_LL_%	Limite_Plastico_LP_%	Indice_Plasticidad_IP_%	Gravas	Arenas	limos_arcillass	Clasificacion_SUCS	Densidad_Seca_Max_gcm3	Humedad_Optima_%	CBR_100%	CBR_95%
1	P1	1.5	7.53	26.5	20.01	6.36	64.93	14.08	20.99	GC-GM	2.08	7.55	21.48	12
2	P2	1.5	4.65	27.4	20.9	6.52	31.01	35.52	33.47	SC-SM	2.052	8.81	22.18	11.69
3	P3	1.5	5.17	23.8	17.1	6.64	62.54	25.45	12.01	GC-GM	1.91	8.5	10.32	6.05
4	P4	1.5	7.07	24	22	2	33.82	39.64	26.54	SM	2.032	9.43	22.32	12.1
5	P5	1.5	5.25	0	0	0	34.36	25.59	40.05	GM	1.8	9	17.76	9.56
6	P6	1.5	7.71	0	0	0	66.9	19.36	13.74	GM	2.004	7.18	17.12	9.5
7	P7	1.5	5.53	31.2	10.4	20.77	10.3	25.7	64	CL	1.973	15.47	8.33	5.9
8	P8	1.5	6.73	28.2	21.2	7.03	6.7	80.65	12.65	SC	1.91	12	7.29	5.3
9	P9	1.5	5.62	29.8	26.7	3.16	30.7	55.59	13.71	SM	1.937	7.66	15.79	11
10	P10	1.5	6.85	0	0	0	52.34	33.49	14.17	GM	1.927	6.84	15.78	13.5
11	P11	1.5	5.25	0	0	0	53.14	34.62	12.24	GM	1.852	9	15.62	13
12	P12	1.5	5.26	0	0	0	59.04	26.54	14.42	GM	1.895	8.68	15.55	12
13	P13	1.5	5.66	0	0	0	69.54	16.16	14.3	GM	1.923	7.76	15.98	12.89
14	P14	1.5	5.23	25.9	23.3	2.62	85.75	1.25	13	GM	1.962	7.22	16.14	11.89
15	P15	1.5	3.66	32.2	22.6	10.66	21.47	0.91	77.62	CL	1.81	9.1	17.05	6.6
16	P16	1.5	8.7	0	0	0	50.77	9.33	39.9	GM	1.889	7.82	15.62	13
17	P17	1.5	7.51	24.5	20.3	4.18	19.21	65.52	15.27	SC-SM	1.957	7.66	15.79	11.05
18	P18	1.5	12.2	28.2	19.3	8.95	5.04	42.48	52.48	CL	1.976	14.5	6.3	3.9
19	P19	1.5	6.23	26.6	20	6.62	42.31	10.72	46.97	GC-GM	2.03	8.19	15.62	13.56
20	P20	1.5	6.02	24.8	20.6	4.24	85.02	0.42	14.56	GC-GM	1.895	8.68	15.55	12
21	P21	1.5	5.66	29.4	21.7	7.77	26.38	0.24	73.38	CL	1.97	14.95	8.43	5.78
22	P22	1.5	12.41	31.7	22.3	9.38	33.21	2.06	64.73	CL	1.951	16	10.11	5.9
23	P23	1.5	7.48	31.7	19.5	12.2	12.38	0.77	86.85	CL	1.995	15	9.14	4.9
24	P24	1.5	11.64	29	17.8	11.17	1.7	0.47	97.83	CL	1.796	9	17.76	8
25	P25	1.5	12.64	39.9	23.4	10.41	36.62	1.13	62.25	CL	1.967	14.82	8.53	5.8
26	P26	1.5	7.44	27.8	24.3	3.47	54.69	0.52	44.79	GM	1.952	7.53	15.93	11.5
27	P27	1.5	6.13	0	0	0	66.19	13.77	20.04	GM	1.952	7.53	15.93	11.5
28	P28	1.00	3.66	42	32	10	24.27	35.51	40.22	ML	1.783	12.00	9.32	7.78
29	P29	1.00	4.91	44	34	10	22.49	34.03	43.48	ML	1.764	12.70	8.83	7.37
30	P30	1.00	4.41	33	19	14	26.92	34.83	38.25	CL	1.7	13.20	6.83	5.65
31	P31	1.00	7.51	58	30	28	14.88	39.61	45.51	CH	1.674	14.20	5.90	5.10
32	P32	1.00	4	29	18	11	45.55	46.042	8.408	GC	2.126	8.80	13.20	10.67
33	P33	1.00	4.95	36	25	11	40.53	45.51	13.96	GM	2.1	9.40	11.40	9.20
34	P34	1.00	4.91	30	18	12	50.64	42.1	7.26	GC	2.136	8.60	15.70	12.00
35	P35	1.00	4.88	35	26	9	18.55	39.75	41.7	ML	1.71	13.20	9.36	7.30
36	P36	1.00	3.37	29	18	11	59.68	33.88	6.44	GC	2.167	7.80	15.50	12.80
37	P37	1.00	4.02	40	28	12	47.64	35.6	16.76	GM	2.155	8.30	17.05	11.95
38	P38	1.6	8.7	20.76	8.09	12.67	3.53	75.42	21.04	SC	1.904	0	23	15.4
39	P39	1.5	0	28	20	8	0	26.5	73.5	CL	1.812	14.4	5.4	4
40	P40	1.5	0	25	18	7	56	28	15	GC	2.145	7.8	55	46

Repositorio Institucional - UNAMBA Perú

ID	Proyecto	Profundidad_m	Humedad_natural	Limite_Liquido LL_%	Limite_Plastico_LP_%	Indice_Plasticidad_IP_%	Gravas	Arenas	limos_arcillass	Clasificacion_SUCS	Densidad_Seca_Max_gcm3	Humedad_Optima_%	CBR_100%	CBR_95%
41	P41	1.6	4	29	22	7	35.31	57.88	6.81	SP-SC	2.117	8.46	44	14.1
42	P42	1.6	2.2	24	17	7	39.77	52.38	7.85	SP-SC	2.17	6.6	27.8	14.5
43	P43	1.8	0.6	23	16	7	43.27	46.6	10.13	SP-SC	2.181	7.45	42.7	17
44	P44	1.5	15.9	26	18	8	7.36	24.96	67.67	CL	1.886	15.82	15	9.6
45	P45	1.5	26.4	40	31	9	0	29.3	70.7	ML	1.651	18	9.7	7.7
46	P46	1.5	4.56	0	0	0	49.62	48.26	2.12	GP	2.054	7.2	40.8	32.97
47	P47	1.5	12.8	36.6	22.4	14.3	8.39	43.42	48.19	SC	1.698	14.9	3	2.4
48	P48	1.5	2.147	0	0	0	0	4	96	SP - SM	1.76	8.6	12.8	11.3
49	P49	1.5	11.4	38	24	14	0	0	0	CL	1.807	15.429	2.5	1.6
50	P50	1.5	6	24	17	7	43	37	20	GC - GM	2.145	7.8	55	46
51	P51	1.5	5	25	18	7	57	28	15	GC	2.15	7.8	44	33
52	P52	1.5	4.38	0	0	0	0	99.04	0.96	SP	1.779	9.6	24.96	10.2
53	P53	1.5	1.66	0	0	0	8.69	88.72	2.59	SP	1.838	1.8	30.17	12.69
54	P54	1.5	2.36	0	0	0	3.23	95.71	1.06	SP	1.886	11.1	74.73	30.92
55	P55	0.00–1.50	14.2	34	26	8	6	58	36	CL	1.72	13.5	9.8	7.2
56	P56	1.50–3.00	12.8	29	21	8	10	60	30	SM	1.78	11.2	18.5	15
57	P57	0.00–1.20	10.5	27	20	7	20	55	25	SP-SM	1.84	9	32	26.5
58	P58	1.20–2.40	11.9	31	22	9	12	57	31	SC	1.76	12	21	17
59	P59	0.00–1.40	15.1	35	24	11	8	50	42	CL	1.69	14.2	14	11
60	P60	1.40–2.60	9.8	25	18	7	25	60	15	GP	1.9	8.5	45	38
61	P61	0.00–1.80	13.6	30	20	10	10	62	28	SM	1.77	11	23	19.5
62	P62	1.80–3.00	16.8	38	27	11	5	52	43	CH	1.65	15	8	6
63	P63	0.00–1.40	11.4	26	18	8	18	65	17	SP	1.82	9.2	28	24
64	P64	1.40–2.60	13.1	28	19	9	9	61	30	SM	1.76	11.3	21.5	18.2
65	P65	2.8	17.72	44.49	20.34	14.32	36.65	40.24	24.01	CL	1.82	10.31	17.28	9.78
66	P66	1.5	13.2	38	22	16	8	34	58	CL	1.78	14.6	9.8	6.5
67	P67	1.2	15.4	41	24	17	6	29	65	CL	1.74	15.8	7.6	5.1
68	P68	1	11.8	32	20	12	12	46	42	CL	1.82	13.2	11.2	7.9
69	P69	1.5	16.1	45	27	18	5	31	64	CL	1.71	16.9	6.9	4.6
70	P70	1.2	14.7	36	21	15	9	38	53	CL	1.79	14.3	9.1	6.2
71	P71	1.5	10.9	28	18	10	18	52	30	ML	1.9	11.5	14.8	10.2
72	P72	1	12.6	34	21	13	14	44	42	CL	1.84	13.1	12.5	8.6
73	P73	1.5	11.3	30	19	11	20	48	32	ML	1.92	12	16.2	11.4
74	P74	1.2	13.9	39	23	16	7	36	57	CL	1.76	15	8.4	5.8
75	P75	1	12.1	33	20	13	10	42	48	CL	1.81	13.4	10.9	7.4
76	P76	1.5	9.8	27	17	10	24	50	26	ML	1.95	10.8	18.5	13.1
77	P77	1.5	5.88	0	0	0	52.2	33.4	14.4	GM	2.090	7.80	25.73	19.70
78	P78	1.5	6.31	22.11	18.96	3.15	48.2	30.5	21.3	GM	2.070	8.00	24.67	19.20
79	P79	1.5	6.11	23.18	20.65	2.53	41.8	36.8	21.4	GM	2.050	8.20	23.93	18.60
80	P80	1.5	12.95	34.45	30.18	4.27	42.1	24.8	33.1	GM	1.980	8.70	20.64	15.70
81	P81	1.5	9.72	33.16	28.73	4.43	34.1	33.9	32	GM	2.000	8.60	20.40	16.20
82	P82	1.5	13.47	32.93	29.2	3.73	36.1	36	27.9	GM	2.020	8.40	21.83	16.60
83	P83	1.5	9.87	31.68	27.79	3.89	37.7	34.3	28	GM	2.040	8.30	20.57	16.50
84	P84	1.5	8.6	28.93	23.23	5.7	56	29.7	14.3	GM	1.999	7.18	21.32	20.6
85	P85	1.5	12.8	24.93	23.26	1.68	59.79	21.86	18.35	GM	2.001	11.49	20.71	19.5
86	P86	1.5	10.5	0	0	0	20.98	36.53	42.49	SM	2.127	10.06	29.89	25.6

Repositorio Institucional - UNAMBA Perú

ID	Proyecto	Profundidad_m	Humedad_natural	Limite_Liquido_LL_%	Limite_Plastico_LP_%	Indice_Plasticidad_IP_%	Gravas	Arenas	limos_arcillass	Clasificacion_SUCS	Densidad_Seca_Max_gcm3	Humedad_Optima_%	CBR_100%	CBR_95%
87	P87	1.5	4.3	0	0	0	60.23	22.36	17.42	GM	1.985	12.39	20.1	17.56
88	P88	1.5	8.3	0	0	0	56.49	22.61	20.9	GM	1.969	11.83	20.71	14.7
89	P89	1.5	5.3	0	0	0	58.02	25.99	15.99	GM	1.955	10.49	24.07	16.9
90	P90	1.5	12.2	32.44	23.85	8.58	54.54	27.83	17.64	GM	2.176	6.67	25.9	24.24

FUENTE: Adaptado de la compilación de datos de las investigaciones de Gálvez Idrogo (2021), Campos Ibáñez y Quispe Cañapataña (2022) y Ramos Pereira (2019), así como de informes de ensayos de laboratorios de la ciudad de Abancay. El conjunto de datos y los documentos de respaldo se encuentran disponibles en el siguiente enlace: https://drive.google.com/drive/folders/1gQeM352JKqymk3_hnLWxSfLvaETIF2WU

Repositorio Institucional - UNAMBA Perú

Anexo G

FRAGMENTOS DEL CÓDIGO FUENTE DE LA APLICACIÓN DESARROLLADA EN PYTHON

App.py

```
import os
import time
from datetime import datetime

import numpy as np
import pandas as pd
import streamlit as st

from auth import require_login, logout
from ml_model import (
    detect_targets, infer_feature_columns,
    auto_clean_df, coerce_numeric, remove_outliers_iqr,
    compare_models,
permutation_importance_original_cols,
    eval_metrics
)
from mr_methods import mr_robusto_mpa,
mr_empirico_promedio, mr_table_all_methods, pct_diff
from plots import fig_real_vs_pred, fig_residuos,
fig_bar_importance, fig_model_performance
from export_utils import to_excel_bytes

APP_TITLE = "GEOMEC 2.0"
APP_SUBTITLE = "Sistema inteligente para la
estimación de la respuesta mecánica de suelos"
APP_FOOTER = "Machine Learning aplicado a
ingeniería geotécnica"

st.set_page_config(page_title=APP_TITLE,
layout="wide", page_icon="◈")

st.markdown("""
<style>
    :root {
        --geomec-navy: #0B1F36;
        --geomec-blue: #2563EB;
        --geomec-teal: #0D9488;
        --geomec-slate: #334155;
        --geomec-bg: #F7FAFC;
        --geomec-card: #FFFFFF;
        --geomec-border: #D9E6F2;
    }
    .stApp {
        background:
            radial-gradient(circle at 20% 5%,
rgba(13,148,136,0.12), transparent 28%),
            radial-gradient(circle at 95% 15%,
rgba(37,99,235,0.13), transparent 28%),
            linear-gradient(180deg, #F8FBFF 0%, #EEF4F8
100%);
    }
    [data-testid="stSidebar"] {
        background: linear-gradient(180deg, #071A2D 0%,
#102A43 100%);
        border-right: 1px solid rgba(255,255,255,0.08);
    }
    [data-testid="stSidebar"] * { color: #F4F8FB
!important; }
    h1, h2, h3 { color: var(--geomec-navy); font-weight:
800; letter-spacing: .2px; }
    .hero-card {
        border-radius: 24px;
        padding: 28px 32px;
        margin-bottom: 18px;
        background: linear-gradient(135deg,
rgba(11,31,54,0.98), rgba(13,148,136,0.82));
        box-shadow: 0 18px 42px rgba(11,31,54,0.18);
        border: 1px solid rgba(255,255,255,0.12);
        color: white;
    }
    .hero-title { font-size: 38px; font-weight: 900; line-
height: 1.0; margin: 0; }
    .hero-sub { font-size: 16px; opacity: 0.92; margin-top:
8px; max-width: 900px; }
    .hero-tag {
        display:inline-block; margin-top: 14px; padding:
6px 12px; border-radius: 999px;
        background: rgba(255,255,255,0.14); border: 1px
solid rgba(255,255,255,0.18);
        font-size: 13px; font-weight: 700;
    }
    .metric-card {
        background: rgba(255,255,255,0.88);
        border: 1px solid var(--geomec-border);
        border-radius: 18px;
        padding: 14px 16px;
        box-shadow: 0 8px 20px rgba(15,31,51,0.06);
        min-height: 96px;
    }
    .metric-label { color: #64748B; font-size: 13px; font-
weight: 700; text-transform: uppercase; letter-spacing:
.5px; }
    .metric-value { color: var(--geomec-navy); font-size:
27px; font-weight: 900; margin-top: 4px; }
    .metric-note { color: #64748B; font-size: 12px;
margin-top: 2px; }
    .process-panel {
        border-radius: 20px;
        padding: 18px 20px;
        background: linear-gradient(135deg,
rgba(255,255,255,0.96), rgba(235,245,255,0.96));
        border: 1px solid #D9E6F2;
        box-shadow: 0 12px 30px rgba(15,31,51,0.08);
    }
    .step-row {
        display:flex; align-items:center; gap: 10px; padding:
6px 0;
        color: #1E293B; font-weight: 650;
    }
    .step-dot {
        width: 12px; height: 12px; border-radius: 999px;
background: #0D9488;
        box-shadow: 0 0 0 5px rgba(13,148,136,0.12);
    }
    .section-card {
        background: rgba(255,255,255,0.84); border: 1px
solid #D9E6F2;
        border-radius: 18px; padding: 16px; box-shadow: 0
8px 18px rgba(15,31,51,0.05);
    }
    div[data-testid="stMetric"] {
```

Repositorio Institucional - UNAMBA Perú

```
        background: rgba(255,255,255,0.88);
        border: 1px solid #D9E6F2;
        padding: 13px;
        border-radius: 16px;
        box-shadow: 0 6px 16px rgba(15,31,51,0.06);
    }
    .stButton>button { border-radius: 12px; font-weight: 800; }
    .footer-note { color:#64748B; font-size: 12px; text-align:center; margin-top: 28px; }
</style>
""", unsafe_allow_html=True)

def hero():
    st.markdown(f"""
    <div class="hero-card">
        <div class="hero-title">{APP_TITLE}</div>
        <div class="hero-sub">{APP_SUBTITLE}</div>
        <div class="hero-tag">{APP_FOOTER}</div>
    </div>
    """, unsafe_allow_html=True)

def stat_card(label, value, note=""):
    st.markdown(f"""
    <div class="metric-card">
        <div class="metric-label">{label}</div>
        <div class="metric-value">{value}</div>
        <div class="metric-note">{note}</div>
    </div>
    """, unsafe_allow_html=True)

def processing_panel(seconds: int = 3):
    steps = [
        "Lectura y verificación de registros",
        "Depuración y normalización de variables",
        "Construcción del pipeline de aprendizaje",
        "Optimización de hiperparámetros",
        "Validación cruzada y consolidación estadística",
        "Selección del modelo con mejor desempeño",
    ]
    panel = st.empty()
    bar = st.progress(0)
    total = len(steps)
    delay = max(0.05, float(seconds) / total)
    completed = []
    for i, step in enumerate(steps, start=1):
        completed.append(step)
        rows = "".join([f"<div class='step-row'><span class='step-dot'></span>{s}</div>" for s in completed])
        panel.markdown(f"<div class='process-panel'><b>Procesamiento del modelo</b>{rows}</div>", unsafe_allow_html=True)
        bar.progress(int(i * 100 / total))
        time.sleep(delay)
    time.sleep(0.2)
    panel.empty()

hero()
require_login()
st.sidebar.markdown("---")
st.sidebar.caption(f"Usuario autenticado: {st.session_state.user}")
logout()
st.sidebar.divider()

uploaded = st.sidebar.file_uploader("Cargar base de datos (.xlsx)", type=["xlsx"])

st.sidebar.subheader("Preprocesamiento")
use_clean = st.sidebar.checkbox("Normalizar datos (NP→0, vacíos→NaN)", True)
use_outliers = st.sidebar.checkbox("Eliminar valores atípicos (IQR)", True)
iqr_k = st.sidebar.slider("Parámetro IQR k", 1.0, 3.0, 1.5, 0.1)

st.sidebar.subheader("Machine Learning")
model_names = ["ExtraTrees", "RandomForest", "GradientBoost", "HistGB"]
models_sel = st.sidebar.multiselect("Algoritmos a evaluar", model_names, default=model_names)
k = st.sidebar.slider("Validación cruzada K-Fold", 3, 10, 5, 1)
cv_mode = st.sidebar.radio("Tipo de validación", ["KFold", "GroupKFold"])
group_col = st.sidebar.text_input("Columna de grupo", value="Proyecto")
log_target = st.sidebar.checkbox("Transformación log del target", True)
n_iter = st.sidebar.slider("Búsqueda de hiperparámetros", 10, 60, 25, 5)

st.sidebar.subheader("Módulo resiliente Mr")
robust_mode = st.sidebar.radio("Criterio robusto", ["median", "mean"], index=0)

st.sidebar.subheader("Interfaz")
processing_delay = st.sidebar.slider("Actualización del panel (seg)", 1, 10, 3, 1)

if uploaded is None:
    c1, c2, c3 = st.columns(3)
    with c1:
        stat_card("Estado", "Sin base", "Cargue un archivo Excel")
    with c2:
        stat_card("Targets", "CBR 95 / 100", "Detección automática")
    with c3:
        stat_card("Modelos", "4", "Algoritmos disponibles")
    st.info("Cargue una base de datos en formato Excel para iniciar el flujo de análisis.")
    st.stop()

# ============================
# Base de datos
# ============================
df_raw = pd.read_excel(uploaded)
df_raw.columns = [str(c).strip() for c in df_raw.columns]
df = df_raw.copy()

if use_clean:
    df = auto_clean_df(df)

t95, t100 = detect_targets(df)
targets = [t95, t100]

df = coerce_numeric(df, targets)
df = df.dropna(subset=targets).reset_index(drop=True)

rows_before = len(df)
if use_outliers:
    df = remove_outliers_iqr(df, targets, k=iqr_k)
rows_after = len(df)

features, num_cols, cat_cols = infer_feature_columns(df, targets)
X = df[features].copy()
```

Repositorio Institucional - UNAMBA Perú

```
y95 = pd.to_numeric(df[t95], errors="coerce").values
y100 = pd.to_numeric(df[t100], errors="coerce").values

groups = None
if cv_mode == "GroupKFold":
    if group_col in df.columns:
        groups =
df[group_col].astype(str).fillna("NA").values
    else:
        st.warning("GroupKFold fue seleccionado, pero no
existe la columna de grupo indicada.")
        groups = None

st.subheader("Panel de control")
c1, c2, c3, c4 = st.columns(4)
with c1:
    stat_card("Registros", f"{len(df)}", f"Base procesada")
with c2:
    stat_card("Variables", f"{len(features)}",
f"{len(num_cols)} num. / {len(cat_cols)} cat.")
with c3:
    stat_card("Algoritmos", f"{len(models_sel)}",
"seleccionados")
with c4:
    stat_card("Outliers", f"{rows_before - rows_after}",
"removidos por IQR" if use_outliers else "sin filtro")

with st.expander("Base procesada"):
    st.dataframe(df.head(80), use_container_width=True)

st.subheader("Entrenamiento y comparación de
modelos")
colA, colB = st.columns([1, 1])
train_btn = colA.button("Entrenar modelos",
type="primary", use_container_width=True)
clear_btn = colB.button("Reiniciar modelos",
use_container_width=True)

if clear_btn:
    for kx in list(st.session_state.keys()):
        if kx.startswith("RES_") or kx.startswith("BEST_"):
            del st.session_state[kx]
    st.rerun()

if train_btn:
    if not models_sel:
        st.error("Seleccione al menos un algoritmo.")
        st.stop()
    processing_panel(processing_delay)
    with st.spinner("Ejecutando entrenamiento,
optimización y validación cruzada..."):
        res95, best95 = compare_models(X, y95, num_cols,
cat_cols, models_sel, cv_mode, k, groups, log_target,
n_iter)
        res100, best100 = compare_models(X, y100,
num_cols, cat_cols, models_sel, cv_mode, k, groups,
log_target, n_iter)

    st.session_state["RES_95"] = res95
    st.session_state["RES_100"] = res100
    st.session_state["BEST_95"] = best95
    st.session_state["BEST_100"] = best100
    st.success("Modelo entrenado y validado
correctamente")

if "BEST_95" not in st.session_state:
    st.info("Presione **Entrenar modelos** para generar
métricas, gráficos, predicción y exportación.")
    st.stop()

best95_name, best95_model, best95_cv, best95_params,
_ = st.session_state["BEST_95"]
best100_name, best100_model, best100_cv,
best100_params, _ = st.session_state["BEST_100"]

pred95 = best95_model.predict(X)
pred100 = best100_model.predict(X)

m95 = eval_metrics(y95, pred95)
m100 = eval_metrics(y100, pred100)

mr95_ml = np.array([mr_robusto_mpa(v,
robust=robust_mode) for v in pred95])
mr100_ml = np.array([mr_robusto_mpa(v,
robust=robust_mode) for v in pred100])

mr95_emp = np.array([mr_empirico_promedio(v) for v in
pred95])
mr100_emp = np.array([mr_empirico_promedio(v) for v
in pred100])

diff95 = float(np.nanmean(pct_diff(mr95_ml,
mr95_emp)))
diff100 = float(np.nanmean(pct_diff(mr100_ml,
mr100_emp)))

st.subheader("Resultados principales")
r1, r2, r3, r4 = st.columns(4)
with r1:
    stat_card("Modelo CBR95", best95_name, f"CV R²:
{best95_cv:.3f}")
with r2:
    stat_card("Modelo CBR100", best100_name, f"CV R²:
{best100_cv:.3f}")
with r3:
    stat_card("MAE CBR95", f"{m95['MAE']:.3f}",
f"RMSE: {m95['RMSE']:.3f}")
with r4:
    stat_card("MAE CBR100", f"{m100['MAE']:.3f}",
f"RMSE: {m100['RMSE']:.3f}")

summary = pd.DataFrame({
    "Indicador": ["R²", "MAE", "RMSE", "Δ% Mr vs
empírico (promedio métodos)"],
    "CBR_95": [m95["R2"], m95["MAE"],
m95["RMSE"], np.nan],
    "CBR_100": [m100["R2"], m100["MAE"],
m100["RMSE"], np.nan],
    "MR_95": [np.nan, np.nan, np.nan, diff95],
    "MR_100": [np.nan, np.nan, np.nan, diff100],
})
with st.expander("Tabla de indicadores"):
    st.dataframe(summary, use_container_width=True)

st.subheader("Comparación de modelos mediante
validación cruzada")
cc1, cc2 = st.columns(2)
with cc1:
    st.markdown("**CBR95**")
    st.dataframe(st.session_state["RES_95"],
use_container_width=True)
with cc2:
    st.markdown("**CBR100**")
    st.dataframe(st.session_state["RES_100"],
use_container_width=True)

mc1, mc2 = st.columns(2)
with mc1:
```

Repositorio Institucional - UNAMBA Perú

```
    st.plotly_chart(fig_model_performance(st.session_state
["RES_95"], "Desempeño CV - CBR95"),
use_container_width=True)
with mc2:
    st.plotly_chart(fig_model_performance(st.session_state
["RES_100"], "Desempeño CV - CBR100"),
use_container_width=True)

st.subheader("Evaluación gráfica")
g1, g2 = st.columns(2)
with g1:
    st.plotly_chart(fig_real_vs_pred(y95, pred95, "CBR95:
Real vs Estimado"), use_container_width=True)
with g2:
    st.plotly_chart(fig_real_vs_pred(y100, pred100,
"CBR100: Real vs Estimado"),
use_container_width=True)

r1, r2 = st.columns(2)
with r1:
    st.plotly_chart(fig_residuos(y95, pred95, "CBR95:
Residuales"), use_container_width=True)
with r2:
    st.plotly_chart(fig_residuos(y100, pred100, "CBR100:
Residuales"), use_container_width=True)

st.subheader("Importancia de variables")
with st.spinner("Calculando importancias por
permutación..."):
    imp95 =
permutation_importance_original_cols(best95_model, X,
y95, scoring="r2", n_repeats=10)
    imp100 =
permutation_importance_original_cols(best100_model,
X, y100, scoring="r2", n_repeats=10)

i1, i2 = st.columns(2)
with i1:
    st.plotly_chart(fig_bar_importance(imp95,
"Importancias (CBR95) - Top 15", top=15),
use_container_width=True)
with i2:
    st.plotly_chart(fig_bar_importance(imp100,
"Importancias (CBR100) - Top 15", top=15),
use_container_width=True)

st.subheader("Predicción individual")
with st.expander("Formulario editable"):
    input_data = {}
    cols = st.columns(3)
    for idx, col in enumerate(features):
        box = cols[idx % 3]
        if col in num_cols:
            default = float(pd.to_numeric(df[col],
errors="coerce").median()) if col in df.columns else 0.0
            input_data[col] = box.number_input(col,
value=default)
        else:
            opts =
df[col].astype(str).fillna("NA").value_counts().head(25).i
ndex.tolist() if col in df.columns else ["NA"]
            input_data[col] = box.selectbox(col, options=opts,
index=0)

    one = pd.DataFrame([input_data])

    if st.button("Predecir registro", type="primary",
use_container_width=True):
        processing_panel(max(1, processing_delay))
        p95 = float(best95_model.predict(one)[0])
        p100 = float(best100_model.predict(one)[0])
        mr95 = float(mr_robusto_mpa(p95,
robust=robust_mode))
        mr100 = float(mr_robusto_mpa(p100,
robust=robust_mode))

        p1, p2, p3, p4 = st.columns(4)
        with p1:
            stat_card("CBR 95", f"{p95:.3f}", "%")
        with p2:
            stat_card("CBR 100", f"{p100:.3f}", "%")
        with p3:
            stat_card("Mr 95", f"{mr95:.2f}", "MPa")
        with p4:
            stat_card("Mr 100", f"{mr100:.2f}", "MPa")

        st.markdown("**Mr por métodos desde CBR95
estimado**")
        st.dataframe(mr_table_all_methods(p95),
use_container_width=True)

        st.markdown("**Mr por métodos desde CBR100
estimado**")
        st.dataframe(mr_table_all_methods(p100),
use_container_width=True)

st.subheader("Predicción por lotes")
st.caption("Suba un Excel con las mismas columnas de
entrada utilizadas por el modelo.")
up2 = st.file_uploader("Excel para predicción por lotes",
type=["xlsx"], key="batch")

batch_out = None
if up2 is not None:
    batch_df = pd.read_excel(up2)
    batch_df.columns = [str(c).strip() for c in
batch_df.columns]
    batch_df = auto_clean_df(batch_df)

    missing = [c for c in features if c not in
batch_df.columns]
    if missing:
        st.error(f"Faltan columnas: {missing[:15]}{'...' if
len(missing)>15 else ''}")
    else:
        Xb = batch_df[features].copy()
        pb95 = best95_model.predict(Xb)
        pb100 = best100_model.predict(Xb)

        batch_df["CBR_95_pred"] = pb95
        batch_df["CBR_100_pred"] = pb100
        batch_df["Mr_95_robusto_MPa"] =
[mr_robusto_mpa(v, robust=robust_mode) for v in pb95]
        batch_df["Mr_100_robusto_MPa"] =
[mr_robusto_mpa(v, robust=robust_mode) for v in
pb100]

        batch_out = batch_df
        st.dataframe(batch_df.head(80),
use_container_width=True)

st.subheader("Exportación de resultados")
os.makedirs("geomec_figs", exist_ok=True)

fig_images = {}
try:
    p1 = "geomec_figs/cbr95_real_pred.png"
    p2 = "geomec_figs/cbr100_real_pred.png"
```

Repositorio Institucional - UNAMBA Perú

```
    p3 = "geomec_figs/cbr95_imp.png"
    p4 = "geomec_figs/cbr100_imp.png"

    fig_real_vs_pred(y95, pred95, "CBR95: Real vs
Estimado").write_image(p1, scale=2)
    fig_real_vs_pred(y100, pred100, "CBR100: Real vs
Estimado").write_image(p2, scale=2)
    fig_bar_importance(imp95, "Importancias (CBR95) -
Top 15", top=15).write_image(p3, scale=2)
    fig_bar_importance(imp100, "Importancias (CBR100)
- Top 15", top=15).write_image(p4, scale=2)

    fig_images = {"Figuras": [(p1, "A1"), (p2, "A25"), (p3,
"A49"), (p4, "A73")]}
except Exception:
    fig_images = {}

df_pred = df.copy()
df_pred["CBR_95_pred"] = pred95
df_pred["CBR_100_pred"] = pred100
df_pred["Mr_95_robusto_MPa"] = mr95_ml
df_pred["Mr_100_robusto_MPa"] = mr100_ml

sheets = {
    "Datos_limpios": df,
    "Predicciones": df_pred,
    "Metricas": pd.DataFrame([
        {"Target": "CBR_95", "Mejor_modelo":
best95_name, "CV_R2": best95_cv, **m95, "Params":
str(best95_params)},
        {"Target": "CBR_100", "Mejor_modelo":
best100_name, "CV_R2": best100_cv, **m100,
"Params": str(best100_params)},
        {"Target": "Mr_95", "Metodo": "Robusto",
"Dif_%_vs_empiricos": diff95},
        {"Target": "Mr_100", "Metodo": "Robusto",
"Dif_%_vs_empiricos": diff100},
    ]),
    "Importancias_95": imp95,
    "Importancias_100": imp100,
    "Resumen": summary,
    "Info": pd.DataFrame([{
        "Software": APP_TITLE,
        "Version": "2.0",
        "Generado":
datetime.now().isoformat(timespec="seconds"),
        "Registros": len(df),
        "Variables_entrada": len(features),
        "Validacion": cv_mode,
        "K": k,
    }])
}
if batch_out is not None:
    sheets["Pred_lotes"] = batch_out

xls = to_excel_bytes(sheets, images=fig_images)

st.download_button(
    "Descargar reporte Excel",
    data=xls,
    file_name="GEOMEC_2_0_Resultados_CBR_MR.xls
x",
    mime="application/vnd.openxmlformats-
officedocument.spreadsheetml.sheet",
    use_container_width=True,
)

st.markdown("<div class='footer-note'>GEOMEC 2.0 ·
Sistema inteligente para análisis geotécnico mediante
Machine Learning</div>", unsafe_allow_html=True)
```

Repositorio Institucional - UNAMBA Perú

```
                ml_model.pimport numpy as np
import pandas as pd

from sklearn.model_selection import KFold,
GroupKFold, RandomizedSearchCV
from sklearn.compose import ColumnTransformer,
TransformedTargetRegressor
from sklearn.pipeline import Pipeline
from sklearn.metrics import r2_score,
mean_squared_error, mean_absolute_error
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import OneHotEncoder,
StandardScaler, FunctionTransformer
from sklearn.ensemble import (
    RandomForestRegressor,
    ExtraTreesRegressor,
    GradientBoostingRegressor,
    HistGradientBoostingRegressor
)
from sklearn.inspection import permutation_importance

RANDOM_STATE = 42

MODEL_REGISTRY = {
    "RandomForest":
RandomForestRegressor(random_state=RANDOM_STA
TE, n_jobs=-1),
    "ExtraTrees":
ExtraTreesRegressor(random_state=RANDOM_STATE,
n_jobs=-1),
    "GradientBoost":
GradientBoostingRegressor(random_state=RANDOM_S
TATE),
    "HistGB":
HistGradientBoostingRegressor(random_state=RANDO
M_STATE),
}

def detect_targets(df: pd.DataFrame):
    cols = list(df.columns)
    def pick(opts):
        for o in opts:
            if o in cols:
                return o
        return None
    t100 = pick(["CBR_100%", "CBR_100", "CBR100"])
    t95  = pick(["CBR_95%", "CBR_95", "CBR95"])
    if not t100 or not t95:
        raise ValueError("No se detectaron targets:
CBR_95% y CBR_100% (o CBR_95/CBR_100,
CBR95/CBR100).")
    return t95, t100

def infer_feature_columns(df: pd.DataFrame, targets:
list[str]):
    X = df.drop(columns=targets, errors="ignore")
    num_cols =
X.select_dtypes(include=[np.number]).columns.tolist()
    cat_cols = [c for c in X.columns if c not in num_cols]
    return X.columns.tolist(), num_cols, cat_cols

def auto_clean_df(df: pd.DataFrame) -> pd.DataFrame:
    df = df.copy()
    df.columns = [str(c).strip() for c in df.columns]
    df = df.replace({"NP": 0, "np": 0, "N.P.": 0, "": np.nan,
" ": np.nan})
    return df

def coerce_numeric(df: pd.DataFrame, cols: list[str]):
    for c in cols:
        if c in df.columns:
            df[c] = pd.to_numeric(df[c], errors="coerce")
    return df

def remove_outliers_iqr(df: pd.DataFrame, cols: list[str],
k: float = 1.5) -> pd.DataFrame:
    out = df.copy()
    for c in cols:
        if c in out.columns:
            x = pd.to_numeric(out[c], errors="coerce")
            if x.notna().sum() < 10:
                continue
            q1 = np.nanpercentile(x, 25)
            q3 = np.nanpercentile(x, 75)
            iqr = q3 - q1
            lo = q1 - k * iqr
            hi = q3 + k * iqr
            out = out[(x >= lo) & (x <= hi)]
    return out.reset_index(drop=True)

def rmse(y_true, y_pred):
    return float(np.sqrt(mean_squared_error(y_true,
y_pred)))

def eval_metrics(y_true, y_pred):
    return {
        "R2": float(r2_score(y_true, y_pred)),
        "MAE": float(mean_absolute_error(y_true,
y_pred)),
        "RMSE": rmse(y_true, y_pred),
    }

def build_preprocessor(num_cols, cat_cols):
    num = Pipeline(steps=[
        ("imputer", SimpleImputer(strategy="median")),
        ("scaler", StandardScaler()),
    ])
    cat = Pipeline(steps=[
        ("imputer",
SimpleImputer(strategy="most_frequent")),
        ("to_str", FunctionTransformer(lambda X:
X.astype(str), feature_names_out="one-to-one")),
        ("onehot",
OneHotEncoder(handle_unknown="ignore")),
    ])
    return ColumnTransformer([
        ("num", num, num_cols),
        ("cat", cat, cat_cols),
    ])

def build_pipeline(num_cols, cat_cols, model_name: str,
log_target: bool):
    pre = build_preprocessor(num_cols, cat_cols)
    base = MODEL_REGISTRY[model_name]
    if model_name == "HistGB":
        to_dense = FunctionTransformer(lambda x:
x.toarray() if hasattr(x, "toarray") else x)
        pipe = Pipeline([("preprocess", pre), ("to_dense",
to_dense), ("model", base)])
    else:
        pipe = Pipeline([("preprocess", pre), ("model",
base)])
    if log_target:
        return TransformedTargetRegressor(regressor=pipe,
func=np.log1p, inverse_func=np.expm1)
    return pipe
```

Repositorio Institucional - UNAMBA Perú

```
def param_space(model_name: str):
    if model_name in ["RandomForest", "ExtraTrees"]:
        return {
            "regressor__model__n_estimators": [400, 700,
1000, 1400],
            "regressor__model__max_depth": [None, 10, 20,
30, 40],
            "regressor__model__min_samples_leaf": [1, 2, 3,
5],
            "regressor__model__max_features": ["sqrt",
"log2", 0.5, 0.7],
        }
    if model_name == "GradientBoost":
        return {
            "regressor__model__n_estimators": [200, 400,
600],
            "regressor__model__learning_rate": [0.03, 0.05,
0.08, 0.1],
            "regressor__model__max_depth": [2, 3, 4],
            "regressor__model__subsample": [0.7, 0.85, 1.0],
        }
    return {
        "regressor__model__learning_rate": [0.03, 0.05,
0.08, 0.1],
        "regressor__model__max_depth": [None, 5, 10, 15],
        "regressor__model__max_leaf_nodes": [15, 31, 63],
        "regressor__model__min_samples_leaf": [10, 20,
30],
    }

def make_cv(cv_mode: str, k: int, X, y, groups=None):
    if cv_mode == "GroupKFold":
        if groups is None:
            raise ValueError("GroupKFold requiere columna
de grupo.")
        splitter = GroupKFold(n_splits=k)
        return list(splitter.split(X, y, groups=groups))
    splitter = KFold(n_splits=k, shuffle=True,
random_state=RANDOM_STATE)
    return list(splitter.split(X, y))

def fit_best_model(X, y, num_cols, cat_cols,
model_name, cv_mode, k, groups=None,
log_target=True, n_iter=25):
    cv = make_cv(cv_mode, k, X, y, groups=groups)
    reg = build_pipeline(num_cols, cat_cols, model_name,
log_target)
    search = RandomizedSearchCV(
        reg,
        param_distributions=param_space(model_name),
        n_iter=n_iter,
        scoring="r2",
        cv=cv,
        random_state=RANDOM_STATE,
        n_jobs=-1,
        verbose=0,
        error_score="raise"
    )
    search.fit(X, y)
    best = search.best_estimator_
    best.fit(X, y)
    return best, float(search.best_score_),
search.best_params_

def compare_models(X, y, num_cols, cat_cols,
model_names, cv_mode, k, groups=None,
log_target=True, n_iter=25):
    results = []
    best_obj = None
    best_score = -1e9
    for m in model_names:
        est, cv_r2, params = fit_best_model(X, y, num_cols,
cat_cols, m, cv_mode, k, groups, log_target, n_iter)
        pred = est.predict(X)
        met = eval_metrics(y, pred)
        results.append({"Modelo": m, "CV_R2": cv_r2,
**met, "Params": str(params)})
        if cv_r2 > best_score:
            best_score = cv_r2
            best_obj = (m, est, cv_r2, params, met)
    return pd.DataFrame(results).sort_values("CV_R2",
ascending=False).reset_index(drop=True), best_obj

def permutation_importance_original_cols(model, X:
pd.DataFrame, y, scoring="r2", n_repeats=10):
    r = permutation_importance(model, X, y,
n_repeats=n_repeats, random_state=RANDOM_STATE,
                               scoring=scoring, n_jobs=-1)
    imp = pd.DataFrame({
        "Variable": X.columns,
        "Importancia_media": r.importances_mean,
        "Importancia_std": r.importances_std
    }).sort_values("Importancia_media",
ascending=False).reset_index(drop=True)
    return im
```

Repositorio Institucional - UNAMBA Perú