

UNIVERSIDAD PERUANA UNIÓN

ESCUELA DE POSGRADO

Unidad de Posgrado de Ingeniería y Arquitectura



**Modelo machine learning para predicción de deserción
estudiantil**

Tesis para obtener el Título de Segunda Especialidad Profesional de
Ingeniería: Estadística Aplicada para Investigación

Autor:

Walter Yataco Cañari

Johnny Prudencio Jacha Rojas

Asesor:

Mg. Johann Alexis Ospina Galindez

Lima, mayo de 2024

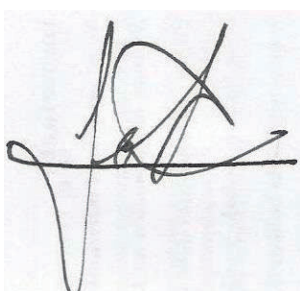
DECLARACIÓN JURADA DE ORIGINALIDAD DE TESIS

Yo (Mag. Johann Alexis Ospina Galindez), docente de la Unidad de Posgrado de Ingeniería y Arquitectura, Escuela de Posgrado de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: **“Modelo machine learning para predicción de deserción estudiantil”** de los autores Walter Yataco Cañari y Johnny Prudencio Jacha Rojas tiene un índice de similitud de 9 % verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de Lima, a los 31 días del mes de mayo del año 2024



Mag. Johann Alexis Ospina Galindez)

ACTA DE SUSTENTACIÓN DE TESIS

En Lima, Ñaña, Villa Unión a 14 días del mes de mayo del año 2024, siendo las 8:40 horas, se reunieron de forma online sincrónica, bajo la dirección del presidente del jurado Dra. Ethel Altez Ortiz, secretario Dr. Josué Edison Turpo Chaparro; los demás miembros: PhD. Javier Linkolk López Gonzales, Mg. Lizeth Geanina Huanca Lopez y el asesor Mg. Johann Alexis Ospina Galindez, con el propósito de administrar el acto académico de sustentación de Tesis de la Segunda Especialidad titulada "Modelo machine learning para predicción de deserción estudiantil", conducente a la obtención del Título de Segunda Especialidad Profesional de Ingeniería: Estadística Aplicada para Investigación.

El presidente inició el acto académico de sustentación invitando a los candidatos a hacer uso del tiempo determinado para su exposición. Concluida la exposición, el presidente invitó a los demás miembros del jurado a efectuar las preguntas, cuestionamientos y aclaraciones pertinentes, los cuales fueron absueltos por los candidatos. Luego, se produjo un receso para las deliberaciones y la emisión del dictamen del jurado. Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Candidatos: Walter Yataco Cañari y Johnny Prudencio Jacha Rojas

| Calificación | | | | Mérito |
|--------------|-----------|---------|--------------------------|---------------|
| | Vigesimal | Literal | Cualitativa | |
| Aprobado | 18 | A- | Con nominación muy bueno | Sobresaliente |

Finalmente, el presidente del jurado invitó a los candidatos a ponerse de pie, para recibir la evaluación final. Además, el presidente concluyó el acto académico de sustentación, procediéndose a registrar a registrar las firmas respectivas.



Presidente



Secretario



Asesor



Miembro



Miembro



Candidato



Candidato

INDICE

| | |
|---|----|
| Portada | 1 |
| Declaración Jurada de Originalidad de tesis | 2 |
| Acta de sustentación | 3 |
| Índice | 4 |
| Resumen | 5 |
| Abstract | 5 |
| Introducción | 6 |
| Metodología | 9 |
| Resultados | 12 |
| Conclusiones | 18 |
| Referencias | 19 |

Modelo machine learning para predicción de deserción estudiantil

Walter Yataco Cañari –
Johnny Prudencio Jacha Rojas

Abstract:

The persistent issue of student dropout negatively impacts the educational sector and society at large. This study presents a machine learning model that leverages data from the National Household Survey to predict student dropout in Peru, integrating a wide range of socio-demographic variables. The research fills a gap in existing literature by providing a model that incorporates socio-demographic variables, an area not fully explored in previous studies. The predictive model aims to identify factors associated with student dropout, aiding educational stakeholders in implementing effective interventions. The findings underscore the model's potential to enhance educational outcomes by enabling early identification of at-risk students, thereby facilitating targeted support. This work contributes to refining predictive models of university dropout rates and suggests the use of ensemble methods to improve the accuracy of single-model predictions. Future research could further explore computational methodologies like deep learning and hybrid models to predict dropout rates and their comparison with this study's outcomes, considering additional influential factors not covered in this research.

Keywords:

Student Dropout, Machine Learning, Predictive Modeling, Socio-Demographic Factors, National Household Survey, Educational Outcomes, Ensemble Methods, Hybrid Models.