

PROYECTO DE INVESTIGACIÓN INTERNO SIN FINANCIAMIENTO PII-DM-2020-01
"Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado automatizado en R"

En la ciudad de Quito D.M., a los dos días del mes de agosto de dos mil veintitrés, comparecen a la celebración de la presente Acta de Finalización del Proyecto de Investigación Interno sin Financiamiento **PII-DM-2020-01 "Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado automatizado en R"**, por una parte, la **Dra. Alexandra Patricia Alvarado Cevallos** en calidad de **Vicerrectora de Investigación, Innovación y Vinculación** de la Escuela Politécnica Nacional, y por otra el **Dr. Miguel Alfonso Flores Sánchez** en calidad de **Director del Proyecto de Investigación Interno Sin Financiamiento PII-DM-2020-01**, al tenor de lo siguiente:

1. ANTECEDENTES:

- a) Mediante Memorando EPN-DM-2020-0126-M del 9 de marzo de 2020, la Jefatura del Departamento de Matemática - DM, remite al Vicerrectorado de Investigación y Proyección Social, la propuesta de Proyecto de Investigación Interno sin Financiamiento "**Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado automatizado en R**" presentada por el Dr. Miguel Flores.
- b) Mediante Memorando EPN-DI-2020-0189-M del 23 de marzo de 2020, la Dirección de Investigación notifica a la Jefatura del Departamento de Matemática, el registro del Proyecto de Investigación Interno sin Financiamiento "**Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado automatizado en R**" dirigido por el Dr. Miguel Flores, con código PII-DM-2020-01, fecha de inicio 24 de marzo de 2020 y fecha de finalización 23 de marzo de 2021.

2. DATOS GENERALES DEL PROYECTO:

Código de Proyecto	PII-DM-2020-01
Nombre del Proyecto	Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado automatizado en R
Director del Proyecto	MIGUEL ALFONSO FLORES SANCHEZ
Colaboradora técnica	KAREN PRISCILLA CALVA YAGUANA
Departamento	Matemática (DM)
Línea de Investigación	<ul style="list-style-type: none"> • Modelos estadísticos
Objetivo	El objetivo de este proyecto es explicar y predecir el rendimiento académico, en función de variables de rendimiento y factores sociodemográficos de los estudiantes del curso de nivelación en la Escuela Politécnica Nacional mediante modelos de aprendizaje supervisado; cuyos resultados estén disponibles a través de una aplicación web
Duración del Proyecto	<ul style="list-style-type: none"> • Fecha de Inicio: 24 de marzo de 2020 • Fecha de fin: 23 de marzo de 2021 • Duración total: 12 meses
Presupuesto	\$ 0,00 USD
Entrega del Informe Final	11 de mayo de 2021

3. INFORME FINAL:

Mediante Memorando EPN-DM-2021-0324-M del 11 de mayo de 2021, el Dr. Miguel Flores, Director del Proyecto de Investigación Interno sin Financiamiento PII-DM-2020-01, presenta el Informe Final del proyecto que dirige, mismo que es revisado por la Dirección de Investigación, que emite observaciones mediante Memorando EPN-DI-2022-0335-M del 22 de marzo de 2022.

Mediante Memorando EPN-DM-2022-0352-M del 24 de marzo de 2022, el Dr. Miguel Flores responde a las observaciones realizadas. Esta información es revisada por la Dirección de Investigación, se anexa y forma parte integrante del Acta de Finalización, cuyas conclusiones y productos generados son:

CONCLUSIONES:

- El algoritmo proporciona una buena precisión predictiva.
- Mucha flexibilidad: se puede optimizar las diferentes funciones de pérdida y los hiperparámetros para que la función se ajuste de mejor forma.
- No se requiere procesamiento previo de datos: a menudo funciona muy bien con valores categóricos y numéricos tal como están.
- Maneja datos faltantes: no se requiere imputación. Es recomendado en bases de datos con datos atípicos.
- El algoritmo GB continuará mejorando para minimizar todos los errores. Esto puede enfatizar a los valores atípicos y causar un sobreajuste. Se incorporó validación cruzada para neutralizar esta desventaja.
- Computacionalmente caro: el algoritmo GB a menudo requiere muchos árboles (> 1000) que pueden ser exhaustivos en tiempo y memoria.
- La alta flexibilidad da como resultado muchos parámetros que interactúan e influyen fuertemente en el comportamiento del enfoque (número de iteraciones, profundidad del árbol, parámetros de regularización, etc.). Esto requiere una búsqueda exhaustiva durante el ajuste.

PRODUCTOS:

- **Artículo publicado:** *"Modelo de Predicción del Rendimiento Académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado"*; Karen Calva, Miguel Flores, Hugo Porras, Ana Cabezas Martínez; Latin American Journal of Computing (Indexado Latindex); ISSN: 1390-9134; mayo 2021.
- **Artículo en formato de la Revista Politécnica:** *"Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado automatizado en R"*; Miguel Flores, Karen Calva; Revista Politécnica (Indexado SCOPUS); ISSN: 2477-8990.
- **Presentación en congreso para difusión a la comunidad politécnica:** *"Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado"*; Karen Calva; Congreso de Investigación Aplicada a Ciencia de Datos – II Congreso Nacional de R Users Group; Facultad de Ciencias; Escuela Politécnica Nacional; enero 2022.

4. LIQUIDACIÓN ECONÓMICA:

El Proyecto de Investigación Interno sin Financiamiento PII-DM-2020-01 no contó con asignación presupuestaria.

5. FINALIZACIÓN:

Con la presente Acta se declara finalizado y cerrado el Proyecto de Investigación Interno sin Financiamiento PII-DM-2020-01 *"Modelo de predicción del rendimiento académico para el curso de nivelación de la Escuela Politécnica Nacional a partir de un modelo de aprendizaje supervisado automatizado en R"*.

Para constancia de lo ejecutado y por estar de acuerdo con el contenido de la presente Acta, las partes libre y voluntariamente suscriben la misma, en tres ejemplares de igual contenido, tenor y valor legal.

Dado en la ciudad de Quito, D.M. a los dos días del mes de agosto de dos mil veintitrés.

Dra. Alexandra Alvarado
**Vicerrectora de Investigación,
Innovación y Vinculación**

sp/cc

Dr. Miguel Flores
**Director del Proyecto
PII-DM-2020-01**