

# DetECCIÓN TEMPRANA DEL RENDIMIENTO ESTUDIANTIL MEDIANTE ANALÍTICA PREDICTIVA Y VIBE CODING

Carrizo Claudio, Javier Saldarini, Raúl Navarro Peláez, Daiana Viotti  
Facultad Regional San Francisco – Universidad Tecnológica Nacional

## RESUMEN

La deserción universitaria constituye uno de los principales desafíos de las instituciones de educación superior. En el ámbito de una Facultad de Ingeniería de una Universidad Pública Argentina, se detectaron problemas de rendimiento académico en materias del área de programación del primer nivel de la carrera de Ingeniería en Sistemas de Información.

Se viene desarrollando un proceso de analítica académica predictiva orientado a anticipar el rendimiento estudiantil en instancias de evaluación parcial. El presente trabajo aborda el análisis comparativo de nueve algoritmos de aprendizaje automático supervisado, con el objetivo de identificar los que ofrecen mayor poder predictivo. Se introduce además el paradigma Vibe Coding como metodología de asistencia a la codificación, utilizando GitHub Copilot como herramienta de IA Generativa.

## OBJETIVO PRINCIPAL

Realizar un estudio comparativo de nueve algoritmos de aprendizaje supervisado (Regresión Logística Multinomial, Árbol de Decisión, Random Forest, Gradient Boosting, XGBoost, k-Nearest Neighbors, SVM, Naive Bayes y Redes Neuronales), asistido por Vibe Coding, para identificar el modelo predictivo más adecuado para anticipar el desempeño académico de estudiantes.

## CONTEXTO

Proyecto de Investigación y Desarrollo (PID): "Analítica Predictiva aplicada para la mejora del desempeño académico en UTN Facultad Regional San Francisco"

- Homologado por la Secretaría de Ciencia, Tecnología y Posgrado de la UTN
- Código: TEECSF0010247 — Disp. SCTyP N° 32-2023
- En ejecución desde inicios del año 2024
- Continuación de la línea presentada en WICC 2025

## RESULTADOS OBTENIDOS

- Construcción del perfil del estudiante de primer año (datos personales, geográficos, académicos y socioeconómicos).
- Dataset de 2335 registros y 33 atributos de cursado en materias del área de programación.
- Descripción detallada de cada atributo (tipo de dato y valores asociados).
- Exploración inicial de datos mediante herramientas de visualización.
- Pre-procesamiento del dataset en Python: limpieza, transformación y verificación de calidad.
- Análisis exploratorio de datos mediante estadística descriptiva.
- Selección del conjunto de atributos para el modelo predictivo.

## EJES DE LA LÍNEA DE INVESTIGACIÓN

1. Estudio en profundidad de los principales algoritmos de aprendizaje supervisado para predicción del desempeño académico.
2. Incorporación del paradigma Vibe Coding como metodología de asistencia a la codificación, mediante IA generativa (GitHub Copilot).
3. Diseño del esquema de evaluación comparativa: métricas de desempeño (Exactitud, Precisión, Recall, F1-Score, AUC-ROC) y validación cruzada.
4. Implementación de cada algoritmo sobre el dataset pre-procesado, usando scikit-learn de Python.
5. Evaluación y comparación de resultados mediante métricas definidas.
6. Selección del modelo predictivo con mejor desempeño.
7. Documentación del proceso Vibe Coding: prompts utilizados, herramientas de IA, calidad del código y correcciones.
8. Comunicación de resultados al Área de Gestión Académica y docentes del Área de Programación.

## 9 ALGORITMOS COMPARADOS

Reg. Logística Multinomial • Árbol de Decisión • Random Forest  
Gradient Boosting • XGBoost • k-Nearest Neighbors  
Support Vector Machine • Naive Bayes • Redes Neuronales

## FORMACIÓN DE RECURSOS HUMANOS

Grupo de Investigación y Desarrollo

- 1 Director de Proyecto
- 1 Co-Director
- 1 Docente investigador de apoyo
- 1 Estudiante de nivel avanzado (Ing. en Sistemas de Información)

En el marco del proyecto se desarrolla un Trabajo Final de Carrera de un estudiante de Ingeniería en Sistemas de Información. El equipo cuenta con antecedentes en PID de UTN vinculados a "Minería de datos educacional" y presentaciones en jornadas y congresos de la especialidad.

## RESULTADOS ESPERADOS

- Informe comparativo de los algoritmos de aprendizaje supervisado evaluados, con análisis de métricas de desempeño por cada modelo.
- Selección y justificación del algoritmo más apropiado para la construcción del modelo predictivo definitivo.
- Documentación metodológica del proceso de Vibe Coding aplicado en la investigación, como aporte a la comunidad científica.
- Construcción, entrenamiento y evaluación del modelo predictivo seleccionado.
- Comunicación de resultados al Área de Gestión Académica y docentes, mediante herramienta de visualización.
- Publicación de resultados en congresos y jornadas científicas.