

## Predicción del Desempeño Académico de Estudiantes

**1. Introducción** El objetivo de este estudio es desarrollar un modelo de clasificación que prediga el desempeño académico de los estudiantes utilizando un conjunto de datos con información académica y socioeconómica. Se analizaron cuatro modelos de clasificación: Random Forest, Árboles de Decisión, K-Nearest Neighbors (KNN) y Support Vector Machines (SVM).

**2. Exploración y Preparación de Datos** Se cargó el dataset y se verificaron valores nulos. Se aplicó la codificación de variables categóricas y la normalización de las variables numéricas. Posteriormente, se dividió el dataset en conjuntos de entrenamiento (80%) y prueba (20%).

### 3. Resultados de los Modelos

- **Random Forest**
  - Precisión: 99.95%
  - Matriz de confusión muestra una clasificación casi perfecta.
  - Posible sobreajuste debido a la alta precisión.
- **Árboles de Decisión**
  - Precisión: 100%
  - No comete errores de clasificación, lo que sugiere un sobreajuste.
- **K-Nearest Neighbors (KNN)**
  - Precisión: 84.07%
  - Mejora en la generalización respecto a los modelos anteriores.
  - Mejor adaptación a la distribución real de los datos.

**4. Análisis de los Resultados** Los modelos de Random Forest y Árboles de Decisión muestran una precisión inusualmente alta, lo que sugiere un sobreajuste. En cambio, el modelo KNN, con una precisión del 84.07%, muestra un comportamiento más realista. Se recomienda ajustar hiperparámetros y aplicar técnicas de regularización para mejorar la capacidad de generalización de los modelos más complejos.

### 5. Conclusiones y Recomendaciones

- Random Forest y Árboles de Decisión requieren ajustes para evitar sobreajuste.
- KNN es un buen punto de partida pero podría mejorarse con optimización de hiperparámetros.
- Se recomienda probar otros modelos como Regresión Logística o Redes Neuronales.
- Se sugiere evaluar el balanceo de clases para evitar sesgos en la clasificación.

Este estudio proporciona una base para mejorar la identificación temprana de estudiantes en riesgo y diseñar estrategias de apoyo académico personalizadas.