
Documentación de Caso de Uso 3: Notas en Evaluación Final

1. Contexto y Objetivo del Modelo

En el ámbito educativo, la prevención del fracaso escolar es una de las principales prioridades institucionales. La capacidad de identificar tempranamente qué estudiantes tendrán un bajo rendimiento permite activar tutorías o planes de apoyo antes de los exámenes finales.

Este caso de uso demuestra la aplicación de nuestra plataforma de Inteligencia Artificial para predecir la calificación académica final que obtendrá un estudiante en el segundo cuatrimestre.

- **Objetivo de la predicción (Target):** Curricular units 2nd sem (grade) (Calificación final del alumno en el segundo semestre).
- **Dataset utilizado:** students_parsed.csv

2. Estructura de los Datos

El modelo ha sido entrenado analizando un perfil integral de estudiantes de cursos pasados. Las variables que evalúa la IA van mucho más allá de lo académico, incluyendo:

- **Datos Académicos Previos:** Calificación de acceso al centro, asignaturas matriculadas, calificaciones y evaluaciones aprobadas durante el 1º semestre.
- **Perfil Socioeconómico y Demográfico:** Edad, género, estado civil, ocupación y nivel de estudios de los padres, y necesidad de becas.
- **Datos Financieros del Alumno:** Si está al día con las cuotas de matrícula (Tuition fees up to date) o si tiene deudas pendientes (Debtor).
- **Contexto Macroeconómico:** Tasa de desempleo, inflación y PIB en el momento de la matrícula.

3. Configuración del Algoritmo

Al tratarse de un problema con múltiples variables de naturaleza tan diferente (desde notas numéricas hasta el estado civil), se optó por un modelo de **Random Forest (RF)**. Este algoritmo es especialista en manejar datos mixtos sin confundirse.

Hiperparámetros establecidos:

- **Número de Estimadores (n_estimators):** 100. Se ha utilizado un "bosque" robusto de 100 árboles de decisión. Esto garantiza que la IA evalúe cada expediente desde muchas

perspectivas (ej. un árbol puede fijarse más en el estado financiero del alumno y otro en sus notas del primer semestre).

- **Profundidad Máxima (max_depth):** 10. Un límite de profundidad equilibrado que permite hacer suficientes cruces de variables sin llegar a memorizar los datos de forma excesiva.

4. Análisis de Resultados y Rendimiento

El modelo fue validado con 5 ejecuciones (Folds) para comprobar su fiabilidad, arrojando resultados de nivel de excelencia:

- **R2 Score (~0.974 - 97.4%):** El modelo posee un poder predictivo altísimo. Es capaz de explicar la nota final del segundo semestre con casi un 97.5% de precisión, demostrando que el rendimiento de la primera mitad del curso, sumado al perfil socioeconómico, dibuja un patrón clarísimo del éxito final.
- **MAE - Error Absoluto Medio (~0.571):** En la práctica docente, esto significa que cuando la IA calcula qué nota va a sacar un alumno a final de curso, **su predicción se desvía, de media, solo en medio punto** respecto a la nota real final.
- **RMSE - Raíz del Error Cuadrático Medio (~0.840):** Este indicador confirma que la IA es extremadamente estable y no sufre grandes equivocaciones. Incluso en los peores casos o expedientes atípicos, el error se mantiene por debajo del punto de desviación.

Conclusión:

El modelo "Notas en Evaluación Final" es una herramienta de Inteligencia Artificial de altísimo valor para el sector educativo. Con una tasa de acierto excepcional, este modelo (exportable en .joblib) puede integrarse en el software de la secretaría académica del centro universitario para generar "alertas tempranas", avisando a los tutores de aquellos alumnos que, estadísticamente, están en riesgo de suspender su evaluación final.
