

MEMORIA FINAL¹

Compromisos y Resultados

Proyectos de Innovación y Mejora Docente 2023/2024

Identificación del proyecto	
Código	sol-202300257146-tra
Título	Predicción del rendimiento de los estudiantes mediante la aplicación de técnicas de aprendizaje automático
Responsable	Esther L. Silva Ramírez

1. Describa los resultados obtenidos a la luz de los objetivos y compromisos que adquirió en la solicitud de su proyecto. Incluya tantas tablas como objetivos contempló.

Objetivo nº 1	<i>Creación de la base de datos</i>
Actividades que había previsto en la solicitud del proyecto:	<i>Desarrollo de encuestas que se llevarán a cabo entre el alumnado para poder recopilar suficiente información sobre los factores que pueden influir en el proceso de aprendizaje de los estudiantes.</i>
Actividades realizadas y resultados obtenidos:	<i>En esta primera fase, recopilación de datos, se diseñaron y distribuyeron encuestas para recopilar información relevante sobre los estudiantes. Las encuestas cubrieron una amplia variedad de temas, desde datos demográficos (edad, sexo, familia numerosa,...), hasta aspectos académicos y personales (trabajo, motivación por la asignatura, estudios de los padres, etc.). Sin embargo, el nivel de participación fue relativamente bajo, con un 12% de los estudiantes completando la encuesta. Esto puede haber afectado al análisis posterior, ya que una mayor participación habría proporcionado más cantidad de datos, lo que hubiese permitido un análisis más preciso de la información. No</i>

¹ Esta memoria no debe superar las 6 páginas.

	<p><i>obstante, las preguntas formuladas cubrieron suficientemente la información necesaria para la modelización posterior del modelo.</i></p> <p><i>Se recopiló un total de 21 variables a partir de las respuestas de los estudiantes. Esta base de datos, aunque limitada en términos de tamaño, fue suficiente para proceder a las siguientes fases del proyecto.</i></p>
--	---

Objetivo nº 2	<i>Selección de las variables que intervendrán en el estudio.</i>
Actividades que había previsto en la solicitud del proyecto:	<i>A partir de la información recogida en la fase anterior se realizará un exhaustivo análisis estadístico que permita determinar las variables que formarán parte del estudio a realizar.</i>
Actividades realizadas y resultados obtenidos:	<i>En esta segunda fase, análisis de datos, se realizó un análisis exploratorio de los datos recopilados para comprender la naturaleza de las variables y su posible influencia en el rendimiento académico. Dada la cantidad limitada de datos, se decidió trabajar con la totalidad de las variables recogidas para no perder ninguna información que pudiese ser potencialmente valiosa.</i>

Objetivo nº 3	<i>Aplicación de distintos algoritmos de machine learning para predecir el rendimiento de los estudiantes.</i>
Actividades que había previsto en la solicitud del proyecto:	<i>A partir de la selección previa de características se pasará a aplicar diferentes algoritmos de aprendizaje automático que permita obtener la predicción del rendimiento de los estudiantes.</i>
Actividades realizadas y resultados obtenidos:	<i>Se aplicaron varios modelos de aprendizaje automático, entre ellos Random Forest, para realizar la predicción del rendimiento de los estudiantes a partir de un conjunto de datos que incluye diversas variables académicas y contextuales. Al tener una</i>

base de datos con una participación limitada, como se indicó en la fase anterior, se decidió utilizar todas las variables disponibles. Las variables incluyeron tanto características relacionadas directamente con el rendimiento académico (calificaciones de tareas), como variables contextuales (estudios y ocupación de los padres, entre otras).

Durante esta fase, además de entrenar los modelos de machine learning, se utilizaron técnicas de explicabilidad de modelos. Estas herramientas proporcionaron una visión detallada de la contribución de cada variable al resultado del modelo, permitiendo entender qué factores tienen un mayor impacto en la predicción del rendimiento académico de los estudiantes.

A partir del estudio realizado, se puede concluir que las variables más influyentes en el modelo son aquellas relacionadas directamente con el rendimiento académico de los estudiantes, los entregables individuales a mitad de cuatrimestre, que presentan los valores más altos. Estas tareas tienen un impacto significativo en el resultado del modelo, lo que indica que un mejor desempeño en estas actividades está fuertemente correlacionado con un mejor rendimiento general del estudiante.

Otras variables, como la nota de admisión y la entrega de un proyecto desarrollado en equipo a comienzos del cuatrimestre, tienen también un impacto importante en la predicción, pero en menor medida, en comparación con las tareas individuales. Estas variables refuerzan la idea de que los aspectos académicos son claves para predecir el rendimiento de los estudiantes.

Por otro lado, variables contextuales como los estudios de la madre y la

ocupación de los padres también tienen algún grado de impacto en las predicciones, pero su influencia es considerablemente menor en comparación con las variables académicas.

En resumen, los resultados sugieren que el rendimiento académico medido a través de entregas de tareas y calificaciones continuas es el principal predictor del rendimiento de los estudiantes. Las variables relacionadas con el contexto familiar, aunque relevantes, no son tan determinantes en este modelo específico.

2. Realice una breve valoración sobre la influencia del proyecto ejecutado en la evolución de las asignaturas implicadas.

Análisis del impacto de la innovación en las asignaturas relacionadas con el proyecto

Este proyecto de innovación docente centrado en la predicción del rendimiento académico mediante técnicas de aprendizaje automático ha tenido un impacto significativo en la evolución de la asignatura implicada. A continuación, se detallan los principales aspectos en los que esta innovación influye:

1. Mejora en la Detección Temprana del Riesgo de Abandono

Uno de los mayores retos en esta asignatura del primer curso del Grado en Ingeniería Informática es la alta tasa de abandono, especialmente por ser una de las asignaturas técnicas fundamentales. El uso de herramientas de machine learning permite a los docentes identificar patrones tempranos de rendimiento y anticiparse a posibles casos de abandono. Este conocimiento facilita, en la medida de lo posible, la adopción de medidas preventivas, como tutorías personalizadas y seguimiento más cercano de los estudiantes en riesgo.

2. Optimización de los Esfuerzos de Evaluación Continua

La asignatura implicada en este proyecto, que cuenta con una alta carga de trabajos prácticos y tareas entregables, puede verse beneficiada con la implementación de este proyecto de innovación. Los modelos de aprendizaje automático han permitido observar que el rendimiento en tareas específicas tiene una influencia significativa en el éxito académico general. Esto permite una optimización en la estructura de la evaluación continua, poniendo un mayor énfasis en las tareas que mejor predicen el

rendimiento final, y adaptando las estrategias de intervención según los resultados.

3. Fomento de la Explicabilidad en el uso de modelos predictivos

El uso de la explicabilidad de los modelos ha sido crucial para que los docentes podamos comprender la importancia de cada factor en la predicción del rendimiento. Este enfoque ha incrementado la confianza en el uso de modelos predictivos dentro del entorno académico, al ser transparentes y comprensibles para el profesorado.

4. Mayor Implicación de los Estudiantes

El conocimiento de que sus tareas y rendimiento en actividades prácticas son determinantes en su evaluación final permitirá generar una mayor implicación por parte del alumnado que curse esta asignatura implicada en el proyecto. Permitirá, en los próximos cursos, que los estudiantes muestren más interés en el proceso de evaluación continua, al comprender que su participación activa en las tareas prácticas puede tener un impacto directo en sus resultados académicos.

3. Incluya en la siguiente tabla el número de alumnos matriculados y el de respuestas recibidas en cada opción y realice una valoración crítica sobre la influencia que el proyecto ha ejercido en la opinión de los alumnos.

Opinión de los alumnos al inicio del proyecto				
Número de alumnos matriculados:				
Valoración del grado de dificultad que cree que va a tener en la comprensión de los contenidos y/o en la adquisición de competencias asociadas a la asignatura en la que se enmarca el proyecto de innovación docente				
Ninguna dificultad	Poca dificultad	Dificultad media	Bastante dificultad	Mucha dificultad
Opinión de los alumnos en la etapa final del proyecto				
Valoración del grado de dificultad que ha tenido en la comprensión de los contenidos y/o en la adquisición de competencias asociadas a la asignatura en la que se enmarca el proyecto de innovación docente				
Ninguna dificultad	Poca dificultad	Dificultad media	Bastante dificultad	Mucha dificultad
Los elementos de innovación y mejora docente aplicados en esta asignatura han favorecido mi comprensión de los contenidos y/o la adquisición de competencias asociadas a la asignatura				
Nada de acuerdo	Poco de acuerdo	Ni en acuerdo ni en desacuerdo	Muy de acuerdo	Completamente de acuerdo

En el caso de la participación de un profesor invitado				
<i>La participación del profesor invitado ha supuesto un gran beneficio en mi formación</i>				
Nada de acuerdo	Poco de acuerdo	Ni en acuerdo ni en desacuerdo	Muy de acuerdo	Completamente de acuerdo
<p>Valoración crítica sobre la influencia que ha ejercido el proyecto en la opinión de los alumnos</p> <p><i>El proyecto tenía el potencial de identificar a los estudiantes en riesgo antes de la finalización de la asignatura, ofreciendo intervenciones preventivas. Sin embargo, al tratarse de una asignatura que se imparte en el segundo cuatrimestre, las limitaciones temporales no han permitido que fuera posible aplicar los modelos predictivos de manera efectiva en este curso académico. Esto significa que no se ha podido realizar una detección precoz y, en consecuencia, el proyecto no ha tenido la oportunidad de rescatar a estudiantes antes de que suspendieran la asignatura.</i></p> <p><i>A pesar de estas limitaciones, los alumnos que participaron en la encuesta mostraron interés en los aspectos del proyecto.</i></p>				

4. Describa las medidas de difusión a las que se comprometió en la solicitud y las que ha llevado a cabo².

Descripción de las medidas comprometidas en la solicitud
<i>En la solicitud original del proyecto, no se incluyeron medidas de difusión específicas. Sin embargo, tras observar los resultados obtenidos y la información valiosa que se ha generado, hemos decidido que los resultados del proyecto son lo suficientemente interesantes y relevantes como para ser difundidos más ampliamente.</i>
Descripción de las medidas que se han llevado a cabo
<i>Se está estudiando la posibilidad de publicar los resultados en una revista indexada en JCR (Journal Citation Reports), con el objetivo de compartir el conocimiento adquirido con la comunidad científica y académica. Esta medida contribuirá a aumentar la visibilidad de los resultados del proyecto, beneficiando a otras instituciones que deseen implementar técnicas de aprendizaje automático en la predicción del rendimiento académico.</i>

² Si en la solicitud no indicó compromiso de difusión de resultados este criterio no se tendrá en cuenta en la evaluación