

Análisis de información generada en las plataformas educativas a partir de herramientas de Inteligencia Artificial



viu

**Universidad
Internacional
de Valencia**

00 Mes 2026

Titulación:

MIAR – Máster universitario
en inteligencia artificial

Curso académico

2025 – 2026

Alumno/a:

Michael Laudrup Luis González

Director/a de TFM: Irma
Zoraida Sanabria Cárdenas

De:
 Planeta Formación y Universidades

Índice

Resumen	6
Abstract.....	7
1. Introducción.....	8
2. Objetivos.....	11
2.1. Objetivo General	11
2.2. Objetivos Específicos	11
3. Marco teórico y estado del arte.....	14
3.1. Marco teórico.....	14
3.1.1. Minería de Datos Educativos (EDM) y Learning Analytics.....	14
3.1.2. Heterogeneidad de datos y estándares	15
3.1.3. Deserción frente a bajo rendimiento	17
3.1.4. Aprendizaje No Supervisado: Descubrimiento de Patrones (Clustering)	17
3.1.5. Aprendizaje Supervisado y Modelos de Ensamble	18
3.1.6. Síntesis del marco teórico.....	19
3.2. Estado del arte.	20
3.2.1. Aprendizaje no supervisado en EDM.....	20
3.2.2. Aprendizaje supervisado en EDM.	21
3.2.3. El enfoque híbrido.....	23
3.2.4. Inteligencia Artificial Explicable (XAI) en Educación.	23
3.2.5. Síntesis del estado del arte	25
4. Referencias.....	27
Anexos I.....	29





Índice de ilustraciones

Ilustración 1 Proceso de preparación del dataset OULAD. Referencia: (Kuzilek, Hlostá, & Zdrahal, 2017).....	16
Ilustración 2 Cronograma de tareas definidas. Elaboración propia¡Error! Marcador no definido.	



Índice de tablas

Tabla 1 Descripción de datos de estudiantes;Error! Marcador no definido.

Tabla 2 Descripción de entrega de actividades;Error! Marcador no definido.

Tabla 3 Descripción datos de interacción de estudiantes;Error! Marcador no definido.

Tabla 4 Ficha técnica de repositorio código fuente....;Error! Marcador no definido.

Resumen

En la actualidad, los Sistemas de Gestión del Aprendizaje (LMS) se han consolidado como la infraestructura fundamental del entorno educativo digital. Si bien estas plataformas han facilitado la generación de un volumen masivo de datos, persiste una brecha crítica entre la recolección de estos y su aprovechamiento pedagógico. En la práctica, los LMS operan frecuentemente como "cajas negras", almacenando registros que los docentes no pueden interpretar, lo que impide la detección temprana de estudiantes en riesgo de fracaso.

El presente Trabajo de Fin de Máster (TFM) aborda esta problemática mediante el desarrollo y validación de un modelo híbrido de Minería de Datos Educativos (EDM). La investigación se fundamenta en la explotación del conjunto de datos OULAD (*Open University Learning Analytics Dataset*), uno de los repositorios de referencia más completos a nivel global, que integra información demográfica, resultados de evaluaciones y registros detallados de interacción de miles de estudiantes en diversos cursos universitarios.

Tras una fase de ingeniería de características a partir de estos datos para transformar los logs en métricas de comportamiento, se aplicaron técnicas de aprendizaje no supervisado (agrupamiento) para segmentar a la población estudiantil e identificar arquetipos de aprendizaje y patrones de conducta. Estos hallazgos se incorporaron posteriormente a modelos de aprendizaje supervisado orientados a predecir el rendimiento final del estudiante (pass, fail, withdraw, distinction). Finalmente, se integraron técnicas de Inteligencia Artificial Explicable (XAI) para identificar las variables con mayor influencia en el riesgo académico, proporcionando información interpretable y accionable para apoyar intervenciones docentes personalizadas.

Palabras clave: Minería de Datos Educativos, Moodle, Aprendizaje Supervisado, Aprendizaje no supervisado, Analítica de Aprendizaje, OULAD, Inteligencia Artificial Explicable.

Abstract

Nowadays, Learning Management Systems (LMS) have become the fundamental infrastructure of the digital educational environment. Although these platforms have enabled the generation of a massive volume of data, a critical gap remains between data collection and its pedagogical use. In practice, LMS often operate as “black boxes”, storing records that teachers cannot interpret, which prevents the early detection of students at risk of failure.

This Master’s Thesis (TFM) addresses this problem through the development and validation of a hybrid Educational Data Mining (EDM) model. The research is based on the exploitation of the OULAD dataset (Open University Learning Analytics Dataset), one of the most comprehensive reference repositories worldwide, which integrates demographic information, assessment results, and detailed interaction logs from thousands of students across various university courses.

After a feature engineering phase to transform logs into behavioural metrics, unsupervised learning techniques (clustering) were applied to segment the student population and identify learning archetypes and behavioural patterns. These findings were subsequently incorporated into supervised learning models aimed at predicting students’ final performance (pass, fail, withdraw, distinction). Finally, Explainable Artificial Intelligence (XAI) techniques were integrated to identify the variables with the greatest influence on academic risk, providing interpretable and actionable information to support personalised teaching interventions.

Keywords: Educational Data Mining, Moodle, Supervised Learning, Unsupervised Learning, Learning Analytics, OULAD, Explainable Artificial Intelligence.

1. Introducción

La educación contemporánea se encuentra inmersa en una transformación impulsada por la digitalización de los procesos de enseñanza y aprendizaje. En el centro de este ecosistema digital se sitúan los sistemas de gestión del aprendizaje (LMS, por sus siglas en inglés), plataformas que han evolucionado de ser repositorios estáticos de documentos hasta convertirse en entornos dinámicos y complejos donde se orquesta la experiencia educativa. Entre estos sistemas, Moodle (Modular Object-Oriented Dynamic Learning Environment) se ha consolidado como uno de los LMS más extendidos en el ámbito educativo global, gracias a su arquitectura de código abierto y su flexibilidad pedagógica, siendo adoptado por instituciones para gestionar cursos que van desde la presencialidad hasta la educación a distancia (Gamage, Ayres, & Behrend, 2022). No obstante, a pesar de su uso extendido, el potencial de Moodle para la analítica de aprendizaje sigue estando infroutilizado, lo que hace necesario el diseño de marcos predictivos que aprovechen su analítica interna para identificar dificultades de aprendizaje (Soepriyanto, Nugroho, Nahri, Kesuma, & Setiasih, 2025).

Esta adopción ha traído una generación masiva y continua de datos, conocida como Big Data Educativo. Cada interacción que un estudiante realiza dentro de la plataforma —desde el acceso a un material de lectura, la participación en un foro de discusión, la descarga de una tarea, e incluso la navegación entre menús o la realización de un cuestionario— deja una huella digital registrada en los archivos de registro (*logs*) del sistema. Investigaciones recientes subrayan que estos *logs* por defecto de Moodle son fundamentales para analizar el comportamiento y los hábitos del estudiante, permitiendo modelar desde la retención hasta la carga de trabajo y los estilos de aprendizaje (Soepriyanto, Nugroho, Nahri, Kesuma, & Setiasih, 2025).

Sin embargo, a pesar de esta gran abundancia de datos, la infraestructura tecnológica ha avanzado más rápido que la capacidad pedagógica para explotarla. Existe una brecha significativa entre lo que la tecnología registra y lo que el docente conoce realmente sobre el comportamiento de sus estudiantes. Esta situación genera un escenario donde el LMS se convierte en una "caja negra" de información no

procesada, donde la complejidad de las estructuras de datos impide una interpretación directa por parte del profesorado (Khosravi, y otros, 2022).

Los docentes carecen de herramientas para identificar a tiempo qué alumnos están en riesgo de fracaso o qué perfiles de aprendizaje coexisten en su aula, y a menudo detectan las dificultades cuando ya es demasiado tarde para intervenir. Existe, por tanto, una necesidad crítica de desarrollar sistemas inteligentes que actúen como soporte, procesando masivamente estos registros para revelar patrones ocultos y ofreciendo información que permita tomar decisiones pedagógicas informadas y oportunas.

El presente Trabajo de Fin de Máster (TFM) se propone abordar esta problemática mediante la aplicación de técnicas avanzadas de Inteligencia Artificial (IA) y Minería de Datos Educativos (EDM). El propósito principal es transformar los datos brutos generados en plataformas de aprendizaje (LMS) en conocimiento accionable, explorando cómo la modelización computacional puede identificar perfiles de comportamiento y predecir el riesgo académico. De este modo, se busca dotar al docente de herramientas analíticas que superen la mera gestión administrativa y faciliten una toma de decisiones pedagógicas fundamentada y proactiva.

Para dar respuesta a esta problemática y cumplir con el propósito de la investigación, la presente memoria se ha organizado en cinco capítulos que estructuran el desarrollo del trabajo. A continuación, se describe el contenido de cada uno de ellos:

Tras esta introducción, el Capítulo 2 define los objetivos del proyecto, estableciendo tanto la meta general como los hitos específicos que guían la investigación. Posteriormente, el Capítulo 3 aborda el estado del arte y marco teórico, proporcionando una revisión exhaustiva de la literatura actual y fundamentando los conceptos clave sobre Minería de Datos Educativos y Analítica de Aprendizaje que sustentan la propuesta técnica.

El núcleo del trabajo se desarrolla en el Capítulo 4, titulado desarrollo del proyecto y resultados. Esta sección desglosa la metodología empleada para abordar el estudio y formaliza el planteamiento del problema específico a resolver. A su vez, detalla las fases



de desarrollo del proyecto, desde la extracción de los datos hasta la implementación de los algoritmos, culminando con la exposición y análisis de los resultados obtenidos tras la experimentación.

Finalmente, el Capítulo 5 presenta las Conclusiones y trabajos futuros, donde se sintetizan los hallazgos principales, se discute el impacto de la solución en la labor docente y se proponen nuevas líneas de investigación para dar continuidad al estudio. El documento cierra con el apartado de Referencias, que recoge las fuentes bibliográficas y recursos académicos citados a lo largo de la memoria.

2. Objetivos

Tras haber analizado la brecha existente entre la recolección masiva de datos en los entornos virtuales y su limitado aprovechamiento pedagógico, es imperativo formalizar las metas que rigen esta investigación. El propósito de este capítulo es delimitar el alcance del proyecto, estableciendo una hoja de ruta que transite desde la fundamentación teórica hasta la implementación de una arquitectura de *Deep Learning* capaz de transformar registros brutos en conocimiento accionable.

Para garantizar el éxito del sistema predictivo, se han definido metas que equilibran la potencia computacional de los modelos híbridos con la transparencia necesaria para el entorno docente. A continuación, se detalla el fin último del estudio y los hitos operativos que permitirán validar la eficacia y explicabilidad del modelo propuesto.

2.1. Objetivo General

Diseñar, implementar y evaluar una arquitectura híbrida de deep learning aplicada a datos de interacción generados en un LMS, utilizando el dataset OULAD, que integre aprendizaje no supervisado (autoencoders y clustering) para construir representaciones latentes y perfiles de estudiantes, y redes Transformer para modelar la evolución temporal de dicha interacción, con el fin de predecir de forma temprana el riesgo de fracaso y abandono académico, incorporando además técnicas de inteligencia artificial explicable (XAI) para interpretar y justificar las predicciones del modelo.

2.2. Objetivos Específicos

- Analizar la evolución del estado del arte en minería de datos educativos, contrastando las limitaciones de los métodos lineales clásicos (PCA, Regresión) frente a las capacidades de las arquitecturas profundas (deep Learning) para modelar la complejidad del comportamiento estudiantil.

- Consolidar y preprocesar el conjunto de datos OULAD, transformando los registros brutos de interacción (logs) en secuencias temporales estructuradas que permitan modelar la evolución de la actividad del estudiante en el LMS y sus patrones de interacción a lo largo del curso, más allá de métricas estáticas acumuladas. Esta representación temporal resulta especialmente adecuada para su tratamiento mediante redes neuronales de tipo transformer y se alinea con propuestas recientes que destacan la utilidad de los “default logs” de Moodle para analizar el comportamiento del estudiante (Soepriyanto, Nugroho, Nahri, Kesuma, & Setiasih, 2025).
- Implementar una arquitectura de Autoencoders (Aprendizaje de Representaciones) para transformar los datos de alta dimensionalidad en un espacio latente comprimido, capaz de capturar relaciones no lineales que escapan a técnicas tradicionales como el PCA.
- Aplicar diferentes técnicas de aprendizaje no supervisado mediante algoritmos de clustering sobre distintas representaciones comprimidas de los datos —espacio latente aprendido por autoencoder y espacio reducido mediante PCA— con el fin de segmentar la población estudiantil, identificar arquetipos de aprendizaje emergentes y comparar hasta qué punto cada representación preserva o resalta estructura útil.
- Implementar modelos de predicción secuencial basados en Transformers (Aprendizaje Supervisado), aprovechando el mecanismo de Self-Attention para detectar dependencias a largo plazo en el curso y superar las limitaciones de memoria de las redes recurrentes (LSTM) en la detección temprana de riesgo.
- Desplegar una estrategia de Explicabilidad Dual (XAI) para mitigar el problema de la "Caja Negra":
 - Intrínseca: Visualización de Mapas de Atención para identificar en qué momentos del curso se focaliza el modelo.

- Agnóstica: Aplicación de valores SHAP (SHapley Additive exPlanations)—un método basado en la teoría de juegos que permite descomponer la predicción de cualquier modelo para entender el peso de cada factor— con el fin de cuantificar la contribución individual y global de cada variable en la predicción final del éxito o riesgo del estudiante.
- Comparar empíricamente el modelo transformers propuesto con líneas base tradicionales (Ensembles como XGBoost y reducciones lineales con PCA), analizando sus métricas de precisión (Accuracy, F1-Score) y el grado de explicabilidad, identificando en qué condiciones el enfoque propuesto aporta ventajas o limitaciones.
- Diseñar un marco de "Explicaciones Accionables", traduciendo los hallazgos técnicos (pesos de atención y valores SHAP) en recomendaciones pedagógicas concretas que permitan al docente realizar intervenciones personalizadas.

3. Marco teórico y estado del arte

La construcción de un sistema predictivo robusto exige un sólido anclaje en la literatura científica y pedagógica actual. Este capítulo establece los cimientos teóricos de la Minería de Datos Educativos (EDM) y la Analítica de Aprendizaje (LA), revisando las investigaciones de los últimos cinco años que sustentan el paso hacia el deep learning.

A través de este análisis, se justifica la adopción de arquitecturas híbridas — basadas en Autoencoders, técnicas de clustering y Transformers — como la solución óptima para capturar la complejidad dinámica del comportamiento estudiantil en entornos virtuales. Asimismo, se posiciona la Inteligencia Artificial Explicable (XAI) como el puente necesario para transformar estos modelos complejos en herramientas pedagógicas transparentes, permitiendo al docente realizar intervenciones fundamentadas y accionables.

3.1. Marco teórico

A continuación, se desarrolla el marco conceptual que define la Minería de Datos Educativos (EDM) y la Analítica de Aprendizaje (LA), para posteriormente analizar el estado del arte de los últimos cinco años. Esta revisión sistemática permite identificar las limitaciones de los métodos tradicionales y justifica la elección de arquitecturas híbridas y técnicas de explicabilidad (XAI) como respuesta a la complejidad del comportamiento estudiantil en plataformas como Moodle.

3.1.1. Minería de Datos Educativos (EDM) y Learning Analytics.

La minería de datos educativos (EDM, Educational Data Mining) y la analítica del aprendizaje (LA, Learning Analytics) son técnicas que utilizan el análisis de datos para mejorar los procesos educativos. Para ello se extraen patrones y tendencias del rendimiento, el comportamiento y las interacciones de los estudiantes. Estas disciplinas optimizan los procesos de aprendizaje, identifican a

los estudiantes con sus dificultades y logran personalizar las experiencias de aprendizaje (Instituto Andaluz Interuniversitario en Ciencia de Datos e Inteligencia Computacional (DASCI), s.f.).

Para materializar estas promesas, es fundamental comprender el origen de la información. Como se detalla en la revisión fundamental sobre EDM publicada en la revista *Modelling*, el primer paso crítico en cualquier ciclo de minería de datos es la recolección de registros provenientes de los "entornos educativos" donde interactúan los estudiantes (Papadogiannis, Wallace, & Karountzou, 2024). Estos entornos, principalmente los Sistemas de Gestión de Aprendizaje (LMS), actúan como la fuente primaria que alimenta los algoritmos.

3.1.2. Heterogeneidad de datos y estándares

En la actualidad existen una gran diversidad de instituciones educativas con plataformas educativas en línea. Cada una de estas plataformas tiene su propio sistema de gestión del aprendizaje, es decir, su propia manera de recolectar datos de los estudiantes, de estructurar la información y de delimitar qué tipo de datos tienen mayor relevancia sobre otros, así como la relación entre estos. Esta falta de estandarización en la captura y estructuración de los datos plantea un desafío metodológico significativo: la dificultad para replicar estudios y generalizar hallazgos entre distintos entornos tecnológicos (Gašević, Dawson, Rogers, & Gasevic, 2016).

A consecuencia de esta situación, se vuelve necesario establecer un marco de referencia común que garantice la validez externa y la comparabilidad de los resultados (benchmarking). En este contexto, el conjunto de datos OULAD (Open University Learning Analytics Dataset) se ha consolidado como un conjunto de datos de referencia ampliamente utilizado en la investigación sobre Minería de Datos Educativos. Este dataset destaca no solo por su amplia adopción en la literatura científica, sino por su riqueza multidimensional y su riguroso proceso de tratamiento y certificación garantizando la anonimidad de los datos (Véase Ilustración 1)

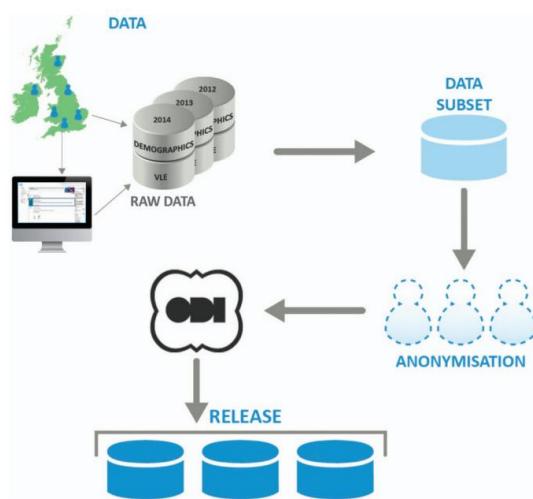


Ilustración 1 Proceso de preparación del dataset OULAD.

Referencia: (Kuzilek, Hlostá, & Zdrahal, 2017)

Integra datos demográficos, resultados de evaluación y registros detallados de interacción (logs) generados en el Entorno Virtual de Aprendizaje (VLE) de la Open University durante los cursos académicos 2013 y 2014. (Kuzilek, Hlostá, & Zdrahal, 2017)

Si bien han surgido conjuntos de datos más recientes en los últimos años, su idoneidad para la presente investigación es limitada. Gran parte de los datasets actuales (procedentes de competiciones tipo Kaggle o plataformas MOOC específicas) presentan inconvenientes críticos: o bien se trata de datos sintéticos generados artificialmente para sortear regulaciones de privacidad —lo que compromete la fidelidad del comportamiento humano real—, o bien carecen de la documentación académica rigurosa necesaria para sustentar una investigación formal. Asimismo, alternativas masivas como EdNet se centran excesivamente en el trazado de conocimiento (Knowledge Tracing) a nivel de ítem, alejándose de la estructura general de interacción típica de Moodle que este trabajo pretende modelar. Por tanto, OULAD se mantiene como un recurso idóneo para validar modelos predictivos en entornos LMS tradicionales, cuya relevancia continúa siendo reconocida en los inventarios de datos de la literatura científica actual (Papadogiannis, Wallace, & Karountzou, 2024).

3.1.3. Deserción frente a bajo rendimiento

En el ámbito del *Learning Analytics* es crucial diferenciar entre la deserción (dropout) y el bajo rendimiento (failure).

La deserción se define como el cese de la actividad en la plataforma y la desvinculación formal del curso antes de su finalización. Este fenómeno suele estar asociado a factores motivacionales, gestión del tiempo o insatisfacción con el entorno. En el dataset OULAD, esto se representa mediante la etiqueta Withdrawn (retirada). Por el contrario, el bajo rendimiento se refiere a aquellos estudiantes que, completando el ciclo del curso y participando en las evaluaciones, no alcanzan los estándares mínimos de conocimiento, etiquetados como Fail.

Distinguir estos dos fenómenos es fundamental para la modelización computacional, ya que los patrones de interacción que preceden a un abandono (ej. disminución progresiva de accesos) difieren sustancialmente de los patrones de un estudiante que suspende (ej. accesos constantes, pero bajo desempeño en cuestionarios). Un sistema robusto debe ser capaz de diagnosticar ambos riesgos por separado.

3.1.4. Aprendizaje No Supervisado: Descubrimiento de Patrones (Clustering)

El Aprendizaje No Supervisado es una rama de la Inteligencia Artificial donde el modelo trabaja con datos no etiquetados, buscando estructuras ocultas o patrones intrínsecos en la información. En el contexto educativo, su aplicación principal es el clustering o agrupamiento.

A diferencia de la clasificación, donde se le dice al algoritmo "este es un buen alumno", el clustering agrupa a los estudiantes basándose únicamente en la similitud de sus comportamientos (ej. frecuencia de accesos, horarios de conexión, tipos de recursos visitados). Esto permite identificar arquetipos de estudiantes (perfiles) que no son evidentes a simple vista, como "el estudiante intensivo de fin de semana" o "el estudiante que solo lee foros, pero no participa". Algoritmos como K-Means (basado en centroides) o DBSCAN (basado en densidad) permiten segmentar la población

estudiantil, proporcionando al docente una taxonomía del aula que facilita la personalización de la enseñanza.

3.1.5. Aprendizaje Supervisado y Modelos de Ensamble

El Aprendizaje Supervisado implica entrenar algoritmos con un conjunto de datos donde se conoce la respuesta correcta (la variable objetivo, en este caso, el resultado final del alumno), para que el modelo aprenda a predecir dicha variable en nuevos estudiantes basándose en su historial. En el contexto de la minería de datos educativos, la evolución de estos modelos ha pasado por tres fases diferenciadas:

- **Modelos Lineales:** Tradicionalmente, se empleaban técnicas como la Regresión Logística, insuficientes para capturar las relaciones complejas y no lineales del comportamiento estudiantil.
- **Modelos de Ensamble (Ensemble Learning):** Estudios comparativos recientes sobre el rendimiento estudiantil (Hasan, y otros, 2020) han validado que los algoritmos de ensamble, específicamente Random Forest, superan en precisión a las técnicas clásicas y a las redes neuronales simples cuando se trabaja con datos estructurados. Estos combinan múltiples "árboles de decisión" para manejar el desbalance de clases y determinar la importancia de las variables (Feature Importance). Sin embargo, tal como se evidencia en la metodología empleada por Hasan, y otros (2020), estos modelos dependen de características agregadas (ej. recuento total de visualizaciones), lo que reduce la compleja dinámica temporal del aprendizaje a valores estáticos acumulados, perdiendo el contexto secuencial.
- **Deep learning Secuencial y Transformers:** Dado que el aprendizaje es un proceso dinámico que evoluciona en el tiempo, el estado del arte más avanzado se ha desplazado hacia redes neuronales capaces de procesar secuencias (Kusumawardani & Alfarizi, 2023).

Inicialmente, las redes recurrentes (RNN y LSTM) permitieron analizar la evolución temporal del alumno. Más recientemente, la arquitectura transformer ha ganado protagonismo en el modelado secuencial. A diferencia de las anteriores, los transformers se basan en mecanismos de autoatención (self-attention), que permiten ponderar la



relevancia de cada evento del curso en relación con los demás, independientemente de la distancia temporal entre ellos. De este modo, es posible capturar dependencias a largo plazo (por ejemplo, cómo el desempeño en una tarea temprana puede influir en el abandono en semanas posteriores) (Kusumawardani & Alfarizi, 2023; Kuzilek, Hlostá, & Zdrahal, 2017).

3.1.6. Síntesis del marco teórico

En síntesis, el marco teórico expuesto evidencia que tanto las técnicas no supervisadas como las supervisadas poseen fortalezas complementarias. Mientras que el clustering permite descubrir la estructura subyacente de los datos sin sesgos previos, los modelos supervisados capitalizan esa información para realizar predicciones concretas. Esta complementariedad teórica sugiere que la integración de ambas ramas —en un enfoque híbrido— constituye la arquitectura idónea para abordar la complejidad del comportamiento estudiantil en plataformas como Moodle, premisa que guiará la revisión de la literatura y la propuesta metodológica de este trabajo.

3.2. Estado del arte.

La literatura científica reciente en minería de datos educativos ha evolucionado desde la aplicación aislada de algoritmos básicos hacia arquitecturas cada vez más complejas orientadas a la personalización. Para contextualizar la contribución de este TFM, esta sección analiza las investigaciones más relevantes de los últimos cinco años que han utilizado el dataset OULAD. El análisis se estructura focalizándose en la evolución metodológica: partiendo de los enfoques puramente exploratorios (no supervisados) y predictivos (supervisados), hasta llegar a las propuestas híbridas más vanguardistas. Asimismo, se identifican las limitaciones recurrentes en los estudios actuales —específicamente en la reducción de dimensionalidad y la interpretabilidad— que justifican la necesidad de explorar técnicas no lineales de Deep Learning.

3.2.1. Aprendizaje no supervisado en EDM.

En el contexto específico del dataset OULAD, la investigación más reciente no solo valida la eficacia de las técnicas de clustering, sino que las posiciona como herramientas indispensables para la personalización educativa. Un ejemplo es el estudio publicado por El Ghali, Atouf, El Guemmat, Broumi, & Talea (2025), quienes profundizaron en la segmentación estratégica de estudiantes utilizando este conjunto de datos. Su investigación contrastó algoritmos basados en vecindad (KNN) frente al clustering jerárquico, empleando PCA (Análisis de Componentes Principales) para gestionar la alta dimensionalidad de las variables demográficas y de interacción.

Los resultados presentados por El Ghali et al. marcan el estándar del estado del arte actual: la aplicación del PCA no solo redujo el tiempo de cómputo en un 60%, sino que, combinado con clustering jerárquico, permitió alcanzar métricas de calidad excepcionales. Mediante esta metodología, lograron identificar cuatro arquetipos de estudiantes, revelando perfiles complejos como aquellos con "alta participación, pero bajo rendimiento" o patrones de compromiso erráticos (El Ghali, Atouf, El Guemmat, Broumi, & Talea, 2025).

Estos hallazgos son relevantes para la presente investigación, ya que sugieren que el dataset OULAD contiene estructura y regularidades aprovechables. No obstante, el hecho de que el estudio reciente se apoye metodológicamente en PCA también refleja que el campo tiende a priorizar transformaciones lineales por su simplicidad e interpretabilidad. En este contexto, se propone explorar el uso de técnicas no lineales (autoencoders) como una vía complementaria para evaluar si pueden representar aspectos del comportamiento estudiantil que PCA no capture de forma directa. Esta propuesta se plantea de manera exploratoria, sin asumir a priori una mejora garantizada, y su utilidad se determinará empíricamente a partir de comparativas y métricas en las fases posteriores del trabajo.

3.2.2. Aprendizaje supervisado en EDM.

La evolución de los modelos predictivos aplicados a OULAD muestra una transición desde clasificadores lineales hacia métodos de ensamble y, más recientemente, hacia enfoques secuenciales de aprendizaje profundo. Este desplazamiento es coherente con la propia naturaleza del dataset OULAD, que combina variables demográficas, evaluaciones y trazas de interacción en el VLE para más de 32.000 estudiantes, y cuyo uso extendido facilita la comparación metodológica entre estudios (Kuzilek, Hlostá, & Zdrahal, 2017)

En tareas con variables tabulares agregadas, algoritmos de gradient boosting como XGBoost se emplean con frecuencia como línea base competitiva; por ejemplo, Firat (2025) reporta mejores resultados de XGBoost frente a Random Forest en su configuración experimental, destacando además la dependencia de estos modelos respecto a la ingeniería manual de características. En este sentido, parte de la literatura enfatiza que la ganancia de rendimiento en ensambles se apoya en “cómo” se agregan los logs (por semana/día, por tipo de actividad, por curso) y en la selección de variables, lo que puede limitar la captura de patrones temporales finos cuando se reduce la secuencia a totales acumulados.

Para incorporar la dimensión temporal de los registros del LMS, trabajos recientes han evaluado modelos secuenciales: Torkhani & Rezgui (2025) obtienen alrededor de un

83% de accuracy con LSTM, superando a varios enfoques clásicos en su comparación. Este resultado es consistente con la línea de trabajos que tratan los clics/interacciones como series temporales (p. ej., agregación semanal) para anticipar riesgo académico antes de finalizar el curso, priorizando el “cuándo” además del “cuánto” interactúa el estudiante. Además, otros enfoques recientes amplían el foco a escenarios multiclasificación (pass, fail, withdrawn, distinction) y muestran que redes neuronales pueden mejorar la predicción temprana cuando integran diversas fuentes (demografía, evaluaciones y clickstream) con preprocesado y feature engineering más exhaustivos (Junejo, y otros, 2024).

En esta misma línea, estudios como Kusumawardani & Alfarizi (2023) exploran transformer encoders sobre OULAD para predicción secuencial (diaria/semanal), reportando resultados competitivos y ventajas prácticas asociadas a la atención y la parallelización del entrenamiento. De forma complementaria, investigaciones recientes sobre OULAD centradas específicamente en abandono (withdrawal) también han propuesto variantes transformer y reportan un rendimiento alto y estable en predicción temprana cuando se usan secuencias semanales de actividad (Luo, Yu, Xie, Wu, & Xu, 2024). Este tipo de evidencias refuerza la idea de que, cuando se preserva la estructura secuencial de los logs, la autoatención puede competir (y en algunos diseños superar) a enfoques recurrentes en tareas de detección temprana de riesgo.

Más allá de la eficiencia, una posible ventaja de los transformers para este TFM es su utilidad interpretativa como señal complementaria: los mecanismos de autoatención generan pesos que permiten analizar qué partes de la secuencia (por ejemplo, interacciones en semanas concretas) han sido más relevantes para la predicción. No obstante, la literatura suele recomendar tratar la atención como una señal interpretativa “auxiliar” y complementarla con métodos de explicabilidad agnósticos al modelo para obtener justificaciones más robustas. En consecuencia, este TFM propone no solo evaluar modelos transformer en términos de rendimiento, sino también analizar patrones de atención y complementarlos con técnicas agnósticas como SHAP para reforzar la explicabilidad (XAI) del sistema híbrido, facilitando que el docente comprenda el “porqué” del riesgo detectado.

3.2.3. El enfoque híbrido.

Los modelos híbridos representan la vanguardia en la minería de datos educativos, combinando técnicas no supervisadas y supervisadas en una arquitectura secuencial. A diferencia de los modelos monolíticos tradicionales, el enfoque híbrido segmenta primero la población mediante clustering para luego entrenar predictores especializados en cada perfil.

Esta estrategia ha sido validada empíricamente por Al-Tameemi et al. (2024) quienes demostraron que el uso de clústeres previos mejora significativamente la precisión de la clasificación en entornos educativos. No obstante, existe un patrón recurrente en la literatura de vanguardia, observado tanto en Al-Tameemi et al. como en El Ghali et al. (2025): la omnipresencia del PCA como técnica estándar de reducción de dimensionalidad, lo que implica una presuposición de que la relación entre los datos es lineal.

Dado que el aprendizaje humano es un proceso complejo y dinámico, la asunción de linealidad puede resultar limitante en ciertos contextos y simplificar relaciones potencialmente más ricas en los datos. Por ello, este TFM plantea complementar el enfoque basado en PCA con una exploración de arquitecturas de autoencoders (Deep Learning) como alternativa no lineal de reducción de dimensionalidad. El objetivo es evaluar empíricamente si estas representaciones latentes son capaces de capturar dependencias que PCA no modela de forma directa y, en caso de aportar valor, utilizarlas como base para los modelos predictivos posteriores.

3.2.4. Inteligencia Artificial Explicable (XAI) en Educación.

Tal como se adelantó al finalizar el análisis sobre los modelos supervisados y la arquitectura Transformer, la búsqueda de una mayor precisión predictiva conlleva la creación de "cajas negras", modelos con estructuras complejas que no son fácilmente interpretables (Khosravi, y otros, 2022). En el ámbito educativo, un modelo que predice el fracaso con un 90% de acierto pero que no explica las causas es pedagógicamente estéril, ya que impide al docente diseñar una intervención correctiva fundamentada. La

literatura define esto como la necesidad de generar "explicaciones accionables" (actionable explanations), entendidas como datos que permiten establecer un procedimiento correctivo o bucle de retroalimentación para un conjunto de acciones (Khosravi, y otros, 2022).

Estudios recientes coinciden en que la adopción real de la Minería de Datos Educativos depende de la confianza (Trustworthiness) que el usuario final deposite en el sistema, siendo la explicabilidad el vehículo fundamental para incrementarla. Los modelos de Deep Learning avanzados, como las redes LSTM o los propios Autoencoders, operan mediante transformaciones no lineales complejas que oscurecen la relación directa entre la entrada y la salida; De hecho, la estructura de estas redes profundas dificulta su comprensión teórica incluso para expertos en la materia, debido a la geometría de sus espacios de alta dimensión (Sejnowski, 2020). Esta opacidad genera una barrera crítica: el docente no puede distinguir si una predicción de riesgo se debe a una falta de actividad, a un bajo rendimiento o a un patrón de comportamiento anómalo.

Para mitigar esta problemática sin renunciar a la potencia de los modelos no lineales, este TFM adopta una estrategia de explicabilidad dual, alineada con el marco de trabajo XAI-ED:

- **Explicabilidad Intrínseca (Attention Maps):** Aprovechando la arquitectura Transformer seleccionada, se extraerán y visualizarán las matrices de pesos de atención (Self-Attention Weights). A diferencia de las redes recurrentes, estos mapas permiten observar directamente en qué momentos del curso o actividades específicas se "fijó" el modelo para determinar el riesgo de un estudiante. Este enfoque sigue las recomendaciones de complementar modelos complejos con componentes visuales interpretables para garantizar la transparencia (Khosravi, y otros, 2022).
- **Explicabilidad Agnóstica del Modelo (SHAP):** Dado que la arquitectura propuesta es híbrida, la interacción entre componentes puede ser compleja. Para garantizar una interpretación global, se utilizarán valores SHAP (Shapley Additive exPlanations). Esta técnica, fundamentada en la teoría de juegos, permite asignar una puntuación de contribución a cada variable, desvelando qué

características empujaron la predicción hacia el éxito o el fracaso y distribuyendo equitativamente la importancia entre ellas (Lundberg & Lee, 2017).

3.2.5. Síntesis del estado del arte

La revisión bibliográfica realizada, si bien no pretende abarcar la totalidad de la vasta producción científica en Minería de Datos Educativos, ha permitido identificar ciertas tendencias y patrones predominantes en los estudios recientes (2020-2025) sobre el dataset OULAD. A partir de la muestra seleccionada, se extraen tres conclusiones que fundamentan la propuesta experimental de este TFM:

1. Exploración de la no-linealidad: En los trabajos analizados (como El Ghali et al., 2025), se observa un uso frecuente de técnicas lineales como el PCA para la reducción de dimensionalidad. Si bien estas técnicas son efectivas, cabe la posibilidad de que la complejidad del comportamiento estudiantil contenga matrices no lineales que se pierden en estas proyecciones. Por ello, resulta pertinente explorar si arquitecturas de Deep Learning, como los Autoencoders, pueden generar representaciones latentes más ricas y mejorar el rendimiento de los modelos posteriores en este contexto específico.
2. Evolución hacia modelos secuenciales: Aunque las redes LSTM han mostrado buenos resultados en la literatura consultada, los estudios más recientes comienzan a señalar las arquitecturas basadas en atención (Transformers) como una alternativa prometedora para capturar dependencias a largo plazo. Este trabajo busca contribuir a esta línea de investigación emergente, aplicando y validando la eficacia de los Transformers sobre los datos de la Open University.
3. La integración de la explicabilidad (XAI): Coincidiendo con la visión de autores como Sejnowski (2020) y Khosravi et al. (2022), se detecta una necesidad transversal de dotar de transparencia a los modelos complejos. Más allá de la precisión predictiva, este TFM asume la premisa de que la adopción



real de estas herramientas depende de su interpretabilidad. Por tanto, se propone evaluar una estrategia de explicabilidad dual (SHAP y Attention Maps) para determinar si ofrece información pedagógicamente accionable al docente.

En definitiva, basándonos en el alcance de esta revisión, se constata que, si bien estas tecnologías han sido exploradas de manera aislada o parcial, no se ha hallado evidencia en la literatura reciente de una arquitectura que las orqueste simultáneamente sobre el dataset OULAD.

Por consiguiente, la combinación específica de Autoencoders (para la representación latente no lineal), Transformers (para el modelado secuencial) y XAI (para la explicabilidad dual) se presenta no solo como una arquitectura híbrida de alto interés experimental, sino como una propuesta novedosa que busca cubrir el vacío existente. Este enfoque pretende verificar si la integración sinérgica de estas tres técnicas avanzadas ofrece ventajas tangibles y superiores frente a la aplicación fragmentada o tradicional observada en el estado del arte.

4. Referencias

- Al-Tameemi, G., Xue, J., Ali, I. H., & Ajit, S. (2024). A Hybrid Machine Learning Approach for Predicting Student Performance Using Multi-class Educational Datasets. *Procedia Computer Science*, 238, 888-895. doi:10.1016/j.procs.2024.06.108
- El Ghali, M., Atouf, I., El Guemmat, K., Broumi, S., & Talea, M. (2025). ENHANCING E-LEARNING THROUGH STRATEGIC STUDENT SEGMENTATION: INSIGHTS FROM THE OULAD DATABASE . *Theoretical and Applied Information Technology*, 103(4), 1290-1299.
- Firat, M. (2025). Comparative Analysis of Random Forest vs XGBoost Machine Learning Algorithms for Predicting ODL Student Success.
- Gamage, S. H., Ayres, J. R., & Behrend, M. B. (Enero de 2022). A systematic review on trends in using Moodle for teaching and learning. *International Journal of STEM Education*, 9(9), 1-24. doi:<https://doi.org/10.1186/s40594-021-00323-x>
- Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not ignore instructional conditions. *Computers & Education*, 53-66.
- Hasan, R., Palaniappan, S., Mahmood, S., Abbas, A., Sarker, K. U., & Sattar, M. U. (2020). Predicting Student Performance in Higher Educational Institutions Using Video Learning Analytics and Data Mining Techniques. *Applied Sciences*, 10(11), 3894. doi:10.3390/app10113894
- Instituto Andaluz Interuniversitario en Ciencia de Datos e Inteligencia Computacional (DASCI). (s.f.). *DASCI*. Recuperado el 2026, de Minería de datos educativos y análisis del aprendizaje: <https://dasci.es/linea-investigacion/mineria-de-datos-educativos/>
- Junejo, N. U., Nawaz, M. W., Huang, Q., Dong, X., Wang, C., & Zheng, G. (2024). Accurate Multi-Category Student Performance Forecasting at Early Stages of

- Online Education Using Neural Networks. *arXiv preprint arXiv:2412.05938*, 1-27.
- Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., . . . Gašević, D. (2022). Computers and Education: Artificial Intelligence. *Computers and Education: Artificial Intelligence*, 3. Obtenido de <https://doi.org/10.1016/j.caecai.2022.100074>
- Kusumawardani, S. S., & Alfarizi, S. A. (2023). Transformer Encoder Model for Sequential Prediction of Student Performance Based on Their Log Activities. *11*, págs. 18960-18971. IEEE. doi:10.1109/ACCESS.2023.3246122
- Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. *Scientific Data*, 4(170171).
- Luo, H., Yu, Q., Xie, S., Wu, H., & Xu, K. (2024). GLU-Transformer for Predicting Withdrawal in VLE. *ICBAR '24: Proceedings of the 2024 4th International Conference on Big Data, Artificial Intelligence and Risk Management*, 835-840.
- Papadogiannis, I., Wallace, M., & Karountzou, G. (2024). Educational Data Mining: A Foundational Overview. *Encyclopedia*, 4(4), 1644-1664.
- Sejnowski, T. J. (2020). The unreasonable effectiveness of deep learning in artificial intelligence. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1907373117
- Soepriyanto, Y., Nugroho, R. P., Nahri, M. H., Kesuma, D. W., & Setiasih, M. (2025). From logs to insights: A comprehensive framework for data-driven learning insights. *Jurnal Inovasi Teknologi Pendidikan*, 12(1), 40-49. doi:<https://doi.org/10.21831/jitp.v12i1.77432>
- Torkhani, W., & Rezgui, K. (2025). OULAD MOOC Student Performance. (págs. 228-241). Atlantis Press. doi:10.2991/978-94-6463-654-3_18



Anexos I

Los anexos también contienen información adicional que se considera relevante para justificar las conclusiones del trabajo, pero, por lo general, el autor de contenido del anexo es distinto al autor del trabajo. Suele ser un documento independiente del trabajo. Pueden ser tablas de datos, imágenes, etc. Es necesario incluir las referencias de los documentos de donde procedan.