

Learning Analytics para el personal académico

Enric Florit, Laura Igual, Eloi Puertas
Dept. de Matemàtiques i Informàtica
Universitat de Barcelona

efz1005@gmail.com, ligual@ub.edu, epuertas@ub.edu

Resumen

Hoy en día, el Learning Analytics (LA), siendo una disciplina relacionada con la ciencia de datos, se ha convertido en una herramienta muy útil para la gestión universitaria. En nuestro caso, proponemos un sistema de LA para ayudar al personal académico en la gestión de grados universitarios. El objetivo principal de nuestro sistema de LA es recopilar información de diferentes fuentes, que comúnmente están disponibles para el personal universitario, y facilitar su comprensión gracias a las visualizaciones y predicciones. En particular, proponemos técnicas de aprendizaje automático para extraer tres indicadores clave para la gestión y la evaluación de la calidad de los títulos universitarios, y que son difíciles de analizar con herramientas estándar. Este sistema de LA se presenta como un conjunto de paneles fáciles de usar dirigidos a ayudar en el proceso de toma de decisiones para gerentes universitarios, decanos, directores o coordinadores de cursos.

Abstract

The field of Learning Analytics (LA), as a data science related discipline, has become a very useful tool for auditing and managing. In this work we propose a LA system for academic personnel to help them in university degrees management. The main goal of our LA system is to gather information from different sources –which are commonly available for university personnel– and make it easier to understand by means of prediction and visualization. In particular, we propose machine learning techniques to extract three key indicators for the management and quality assurance of the university degrees, which are difficult to analyze with off-the-shelf tools. This LA system is presented as a set of user-friendly dashboards addressed to help in the decision making process for university managers, such faculty deans, headmasters or course coordinators.

Palabras clave

Learning Analytics, predicción de calificaciones, predicción del abandono, detección de polaridad, análisis del plan de estudios, análisis de redes.

1. Introducción

En un contexto educativo, el análisis de los datos disponibles relacionados con el proceso de aprendizaje se conoce como *Learning Analytics* (LA). El análisis de dichos datos no solo es interesante para la mejora del proceso de aprendizaje durante un solo curso, sino que también es importante para garantizar la calidad de todo el plan de estudios. En el caso de las instituciones de educación superior, parte del personal académico son profesores encargados de velar por la buena evolución de los estudios ofrecidos y de proporcionar pautas para los cursos futuros. Varias universidades ya han aplicado LA utilizando datos curriculares de estudiantes de diferentes grados, obteniendo resultados muy positivos [11, 13, 10]. Sin embargo, también se sabe que el progreso académico de un estudiante está estrechamente relacionado con la evolución de los cursos.

En este trabajo, presentamos un sistema de LA basado en datos para obtener algunos indicadores relacionados con la evolución de los cursos: 1) predicción de calificaciones y abandono de los estudiantes; 2) detección de polaridad en comentarios de estudiantes en encuestas, y 3) análisis de competencias de asignaturas a partir de los planes de estudios y colaboración docente. Estos indicadores son complejos de extraer directamente a partir de los datos y, por lo tanto, las técnicas de aprendizaje automático pueden usarse para automatizar el proceso y permitir la escalabilidad al obtener una imagen general de las relaciones entre estudios, asignaturas, estudiantes y personal docente.

La predicción de abandono indica qué estudiantes necesitan más asesoramiento y orientación personal. Las predicciones de calificaciones son útiles para proporcionar orientación personalizada. Estas también son formas de estimar cuántos estudiantes continuarán sus

estudios el año siguiente y en qué asignaturas se matricularán. La detección automática de polaridad en preguntas abiertas en encuestas de satisfacción es interesante para complementar las preguntas cuantitativas (respuestas del 1 al 10 y sí / no). Por lo general, existe una correlación entre estos dos tipos de respuestas en una encuesta, pero solo con valores numéricos no siempre es posible inferir la polaridad de los comentarios (positivo, neutro o negativo). Finalmente, el análisis gráfico entre las competencias del curso y la colaboración docente ayuda a comprender la evolución de las asignaturas en un grado y a mantener los planes docentes coherentes con las pautas generales de los estudios.

Estos indicadores se recopilarán en un tablero de control de LA que permitirá su visualización y análisis a nivel de un grado o un grupo de grados y másteres (a nivel global).

Este artículo está organizado de la siguiente forma: La Sección 2 describe el análisis de las calificaciones para la predicción del abandono y las calificaciones futuras, así como el tablero de control desarrollado. La Sección 3 explica el trabajo realizado en la detección de polaridad para analizar los comentarios de los estudiantes sobre los profesores y los cursos. La Sección 4 explica el tablero de control para el análisis de competencias de cursos y sus visualizaciones. Finalmente, la Sección 5 concluye el trabajo sugiriendo trabajo futuro y comentarios finales.

2. Análisis de abandono y calificaciones

Uno de los indicadores más importantes para la educación superior es la tasa de abandono, y más específicamente, la tasa de abandono al final del primer curso. Una tasa alta puede sugerir que el proceso de inscripción no fue el más adecuado para la institución, y que la formación previa de los estudiantes no fue suficiente o que sus expectativas no se cumplieron por completo. En España, la tasa de abandono se sitúa entre 25 % y 29 % [14] y, en la Universitat de Barcelona, la tasa de abandono de 2009 a 2014 fue de alrededor del 20 % [1].

En esta sección tratamos de responder las siguientes preguntas: ¿es posible predecir si un estudiante se matriculará nuevamente en los años siguientes?, y ¿es posible predecir qué calificaciones obtendrá el estudiante en los años siguientes?

En la literatura se pueden encontrar trabajos que consideran estas preguntas, pero se basan principalmente en datos personales diferentes al rendimiento académico (por ejemplo, datos LMS) [5, 15], que pueden ser costosos de recopilar. En nuestro enfoque, pre-

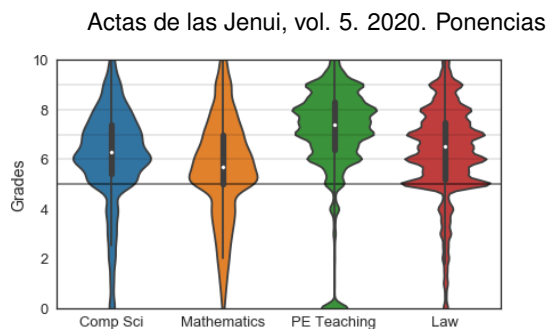


Figura 1: Distribución de las calificaciones del curso (primer intento) entre los grados. Las diferencias son visibles entre los cuatro grados.

sentado en [16, 7], entrenamos nuestros modelos utilizando información de las calificaciones del primer año académico de un estudiante, porque esta información ya se monitoriza por las universidades y se actualiza periódicamente. Además, cualquier otro tipo de información valiosa se puede agregar fácilmente a los modelos en caso de que sea proporcionada por la universidad. Véase [16] para obtener más referencias de trabajos relacionados. Por otro lado, los modelos son fácilmente ampliables, ya que se les puede agregar otros tipos de datos relacionados en caso de que sean proporcionados por la universidad.

2.1. Datos

Para este trabajo recopilamos 134 571 muestras de datos (calificaciones) de un total de 7 557 estudiantes de los grados de Derecho (3 444), Matemáticas (490), Ingeniería Informática (422) y Educación Primaria (3 201) en la Universitat de Barcelona (UB) entre los años 2009 y 2016. La Figura 1 muestra la distribución de las calificaciones del curso (primer intento) entre los grados. Hay diferencias visibles entre las ciencias y las humanidades. El resto de la sección resume los resultados de los tres primeros grados, presentados en [16]¹. Para ver más resultados, consúltese también [7].

2.2. Predicción del abandono

Consideramos aquí el problema del abandono como un problema de clasificación binaria desequilibrada, que puede abordarse mediante el siguiente procedimiento de dos pasos:

Paso 1: Definición del vector de características y preprocesamiento de datos. Cada estudiante en el conjunto de datos se describe utilizando un vector n -dimensional que consiste en las calificaciones de cada asignatura de un año académico determinado. Para los

¹El conjunto de datos está disponible en <https://doi.org/10.1371/journal.pone.0171207>

grados de Informática y Matemáticas en la UB, $n = 10$ para todos los años académicos; para Derecho, $n = 10$ para el primer año académico; y $n = 8$ para el resto. De entrada descartamos incluir otra información en el vector de características, como la nota de admisión, para que el sistema sea menos dependiente de factores externos (como se explica en [16]). Para capacitar adecuadamente a los clasificadores, utilizamos Synthetic Minority Over-sampling Technique (SMOTE) [4] para equilibrar el conjunto de datos.

Paso 2: Clasificación. Comparamos técnicas de vanguardia (Regresión logística, Gaussian Naive Bayes, Support Vector Machine, Random Forest y AdaBoost) que utilizan diferentes enfoques para resolver un problema de clasificación y elegimos el que tiene el mejor rendimiento, a saber, AdaBoost. La puntuación F1 obtenida es del 76 % para Derecho, Matemáticas e Informática. Véase [16] para obtener más detalles sobre la comparativa, las referencias y los resultados.

2.3. Predicción de calificaciones

Consideramos el problema de predecir calificaciones futuras para un estudiante en particular dados los anteriores. Para resolver este problema, implementamos un sistema recomendador de filtro colaborativo con ajuste de línea base [17], debido a su escalabilidad y flexibilidad para manejar los valores nulos. Este enfoque se ha comparado con modelos de regresión como Regresión lineal [20] y SVM [18], dando resultados menos precisos que los conseguidos por el sistema de recomendación [16]. El vector de características utilizado es el mismo que en el problema de predicción de abandono.

En general, un sistema de recomendación funciona encontrando similitudes entre las filas de una matriz *sparse* y prediciendo los valores nulos utilizando datos de la misma matriz [8]. En nuestro caso particular, las filas de la matriz corresponden a las calificaciones de los estudiantes para cada curso de un grado. Los sistemas de recomendación funcionan bajo el supuesto de que si dos asignaturas son similares y un estudiante no tiene una calificación para una de ellas, entonces es válido predecir este valor desconocido utilizando la calificación de un estudiante similar.

Después de realizar una validación cruzada quíntuple con los datos de los grados en Matemáticas, Derecho e Ingeniería Informática, obtenemos un Error Absoluto Medio (MAE) de 1,215, 1,321 y 1,344, respectivamente. Visualizamos cada calificación predicha para los cursos de segundo año en comparación con las calificaciones reales de nuestro conjunto de datos en la Figura 2 (grado de Matemáticas) para ver cómo se distribuye el error a lo largo de las predicciones. Esta figura se compone de dos partes: 1) un área central que muestra un diagrama de dispersión de la predicción de

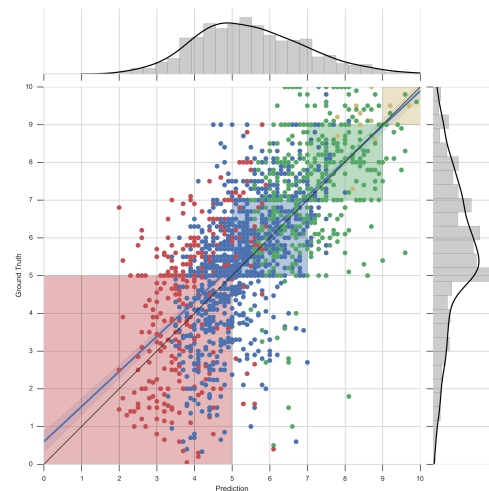


Figura 2: Gráfico de dispersión y distribución de predicciones de calificaciones para el grado de Matemáticas. Cada punto corresponde a una puntuación de un estudiante para una asignatura en particular. Los puntos se colorean de acuerdo con la puntuación media obtenida por el estudiante en el año académico anterior. Los gráficos de histograma muestran las distribuciones de las calificaciones predichas (eje X) y las calificaciones reales (eje Y).

calificaciones (eje X) y calificaciones reales (eje Y) de todos los cursos de segundo año, junto con una recta de predicción perfecta (línea negra) y una recta de regresión (línea azul), y 2) dos histogramas que muestran las distribuciones de las calificaciones predichas (eje X) y las calificaciones reales (eje Y). Las cuatro regiones sombreadas del diagrama de dispersión corresponden a las áreas donde las calificaciones pronosticadas cuantificadas serían aceptadas como correctas. Observamos que la recta de regresión es cercana a la recta de predicción perfecta. Los puntos que caen en las áreas blancas de la trama son aquellos que se han pronosticado de forma incorrecta tras la cuantización. Para proporcionar más información visual, cada punto se colorea de acuerdo con el promedio obtenido por el estudiante en el año académico anterior.

Todo el código utilizado en este trabajo está disponible públicamente en un repositorio de GitHub ².

2.4. El tablero de control

Finalmente, compilamos e importamos los resultados de la predicción en un tablero de control en forma de gráficos que muestran las calificaciones obtenidas y pronosticadas de cada estudiante. Hay una demostración del tablero de control disponible públicamente ³.

²<https://github.com/pid-ub/pid-UB>

³<http://pid-ub.github.io/demo/>

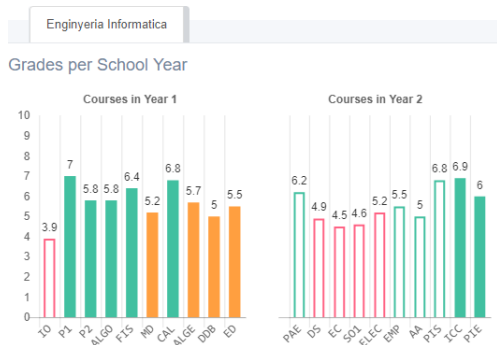


Figura 3: Vista en el tablero de control de las calificaciones pasadas y predichas de un estudiante.

La Figura 3 muestra una captura de pantalla de ejemplo. Las barras verdes representan asignaturas aprobadas. Esta visualización, accesible solo para el tutor de cada uno de los alumnos, permite ver las calificaciones históricas y futuras de un alumno de un vistazo. Esto le da al tutor la oportunidad de tomar decisiones informadas sobre el asesoramiento estudiantil. Además, el tablero de control envía alertas a los tutores cada vez que un alumno al que asesoran corre el riesgo de abandonar los estudios. La herramienta permite cargar y procesar nuevos datos para actualizar los resultados de predicción en línea.

3. Análisis de sentimiento en encuestas a estudiantes

El análisis de sentimiento es un área de investigación en rápido crecimiento que se centra en la extracción automática de sentimientos de documentos textuales. Las aplicaciones más comunes son la extracción de sentimientos de Tweets, reseñas de productos y comentarios en Internet, con el objetivo de tomar mejores decisiones de marketing, realizar recomendaciones, mejorar los diferentes servicios, así como avanzar en la interacción persona-ordenador [2]. El análisis de sentimiento y la subtarea de detección de polaridad pueden abordarse como una tarea de clasificación binaria (“positiva” o “negativa”), una tarea de clasificación múltiple al agregar una clase “neutral”, o una tarea de estimar positividad y negatividad en una escala numérica discreta o continua. En este trabajo, la detección de polaridad se usa para determinar si los comentarios de los estudiantes sobre las asignaturas o los profesores son positivos, negativos o neutros. La detección automática de polaridad en los comentarios de texto abierto de los estudiantes proporciona una nueva fuente de información valiosa para el personal académico.

3.1. Descripción del corpus

Para realizar esta tarea de detección de polaridad, en primer lugar, creamos un corpus a partir de nuestros datos. El corpus es una colección de dos tipos de encuestas de estudiantes: encuestas sobre profesores (relacionados con una asignatura dada) y encuestas sobre las asignaturas (que pueden tener más de un profesor). Ambas encuestas están formuladas en catalán, pudiéndose contestar en catalán, español o inglés. Como había un número muy pequeño de comentarios en inglés, estos fueron descartados para los experimentos. Cada encuesta tiene dos tipos de preguntas: preguntas cuantitativas que se responden con calificaciones de 0 a 10 y preguntas de texto abierto.

El corpus que utilizamos en este trabajo se basa en las preguntas de texto abierto. Las respuestas a las preguntas de texto abierto se han anonimizado y aleatorizado, y actualmente no es posible hacer coincidir las respuestas de texto abierto con las respuestas cuantitativas. Sin embargo, estamos explorando la posibilidad de obtener este mapeo para obtener un conjunto de datos semi-supervisados.

Las preguntas abiertas sobre profesores son:

1. qué aspectos de la actividad docente de este profesor han ayudado más a los estudiantes en el proceso de aprendizaje en la asignatura, y
2. qué aspectos de la actividad docente de este profesor deberían mejorar.

La única pregunta abierta con respecto a las asignaturas es simplemente escribir cualquier comentario, sugerencia u observación.

3.2. Creación del corpus

Para la creación del corpus, primero extraemos los datos, luego limpiamos y preprocesamos los datos y, finalmente, anotamos la polaridad de cada respuesta.

Para limpiar y preprocesar los datos, realizamos varios pasos. Primero, reemplazamos todos los nombres de los profesores dentro de los comentarios por un pseudónimo para hacer que todos los datos sean anónimos. En segundo lugar, detectamos el idioma de cada comentario y los agrupamos por idioma. Los pocos errores de la detección automática se corrigen manualmente durante el proceso de anotación. El promedio de palabras por comentario es de 27,9 palabras para comentarios en español y 25,5 para comentarios en catalán, y el promedio de palabras por oración es 16,7 para oraciones en español y 15,4 para oraciones en catalán. Finalmente, cada oración se procesa con un etiquetador de categorías gramaticales, y es dividida por palabras y lematizada. Usamos *Freeling*⁴, una herramienta de análisis de lenguaje de código abierto que

⁴<http://nlp.lsi.upc.edu/freeling/index.php/>

cubre la mayor parte de la extracción de características para español y catalán. Cada oración en el corpus se anota manualmente (por dos anotadores para el catalán, y tres anotadores para el español) para obtener un conjunto de datos supervisados. El acuerdo promedio entre anotadores es de 0,87 para catalán y 0,75 para español. Las oraciones en las que los anotadores no estaban de acuerdo fueron descartadas.

El conjunto de datos anotados resultante contiene 883 oraciones negativas, 169 oraciones positivas, 137 oraciones neutrales en español; y 1596 oraciones negativas, 885 oraciones positivas, 444 oraciones neutrales en catalán. La proporción entre oraciones negativas y positivas en catalán es de casi el doble, mientras que en español es más de 5 veces mayor. Esto puede explicarse por el hecho de que no es obligatorio responder las encuestas: solo alrededor de 20% de los estudiantes en cada materia envían una respuesta antes del final del semestre y, generalmente, es porque tienen sentimientos extremos sobre una asignatura en particular o sobre el personal docente.

3.3. Modelo de análisis de sentimiento

Para configurar nuestro modelo, realizamos varios experimentos basados en el artículo [3] en nuestro corpus, tanto en español como en catalán. Como la proporción en catalán del corpus es mayor y está mejor equilibrada en términos de distribución de clases, solo informamos de estos resultados.

Se utilizaron diferentes técnicas de extracción de características para crear el modelo (consúltese [9] para más detalles). Comparamos los resultados experimentales mediante el uso de varios clasificadores de aprendizaje supervisados, siendo los sistemas más exitosos *Support Vector Machines* (SVM) y *Logistic Regression* (LR). Los métodos relacionados con las redes neuronales no funcionan bien en este escenario, debido al escaso número de comentarios que estamos utilizando.

El sistema con el mejor rendimiento utiliza una combinación de frecuencia de términos en *bag-of-words*, bigramas y trigramas (todos con TF-IDF y creados a partir de oraciones lematizadas) que proporciona una precisión de 0,72. En comparación con el rendimiento humano, el acuerdo entre los anotadores era del 0,87 para los comentarios catalanes: nuestro mejor modelo representa un 83%.

4. Análisis de competencias

Queremos detectar agrupaciones y estructuras organizativas en dos conjuntos de datos: el conjunto de asignaturas impartidas en un grado universitario, y el conjunto de profesores que imparten dichas asignaturas. Para hacerlo, establecemos nuestro objetivo de

desarrollar una herramienta para visualizar ambos conjuntos de datos como grafos. Esto puede ser útil para facilitar la reestructuración de los planes de estudios, gracias a la agrupación visual de las asignaturas de un grado.

En [6] se realizó una visualización de varios tipos de relaciones entre asignaturas dentro del mismo plan de estudios. Según los autores, se debe utilizar un análisis de las relaciones entre asignaturas para “evaluar y monitorear los planes de estudio” e “informar las estrategias de renovación del plan de estudios”. Las relaciones entre las asignaturas dentro de un plan de estudios permiten un tratamiento orientado a la red de sus (meta) datos. Además, la forma habitual de representar un plan de estudios en una tabla es limitada, siendo más natural tratarlo desde la perspectiva de su estructura de grafo, como se señala en [19].

Una característica clave de nuestro contexto curricular es que la mayoría de los estudiantes no eligen libremente sus asignaturas durante la mayor parte de sus estudios (el programa está predefinido). Esto sí que sucede en otros sistemas (por ejemplo, en América del Norte), lo que hace que este sea el foco principal de los estudios citados. Nuestro enfoque, por el contrario, se basará en el contenido y observaremos las relaciones de las asignaturas en función de factores de similitud.

La mayoría de las técnicas para interpretar y visualizar datos de relaciones de asignaturas se pueden aplicar al conjunto de datos de los profesores y las materias que pueden haber co-impartido. Esto se propone en gran medida siguiendo el estilo del análisis de datos académicos, un campo que también se ha beneficiado de las técnicas de visualización de datos y grafos [12].

4.1. Asignaturas y competencias

Centrémonos primero en la clasificación de asignaturas. Recopilamos la siguiente información en una base de datos SQL: grados impartidos en la Universitat de Barcelona, asignaturas en cada grado y competencias asociadas a una asignatura determinada. Por competencias nos referimos a descripciones de alto nivel de habilidades que los estudiantes deben adquirir o mejorar después de cursar una asignatura.

Las competencias son una fuente de similitudes entre asignaturas fácil de manejar, porque cada grado tiene una lista fija y cerrada de ellas. Por lo tanto, podemos relacionar dos asignaturas A y B de un grado en base a un criterio cuantitativo: $A \sim_k B$ siempre que compartan al menos k competencias. Para cualquier k positivo, esto nos da un grafo, en el que los vértices son el conjunto de asignaturas, y las aristas están dadas por la relación definida.

Podemos etiquetar cada asignatura de acuerdo con su categoría, a saber: *formación básica* (FB), *obligatoria* (OB), *optativa* (OT), *prácticas* (PR) y *trabajo final*

de grado (TR). Esto nos da una coloración del grafo. Cada arista puede etiquetarse con un número entero correspondiente al número de competencias compartidas entre las asignaturas que conecta, obteniendo un grafo con pesos.

Con la estructura de grafo definida, pasamos a producir la visualización. Queremos renderizar el grafo de una manera “natural”: los clústeres y las partes densamente conectadas del grafo deben mostrarse de forma consecuentemente cercana. Esta disposición natural se logra mediante el módulo *d3-force* de D3⁵.

Es de esperar que, tras establecer un umbral k para el número de competencias compartidas, un grafo tenga clústeres correspondientes a 1) categorías de asignaturas, con asignaturas divididas según el año en que se imparten, o 2) áreas temáticas, con asignaturas agrupadas de acuerdo con sus contenidos.

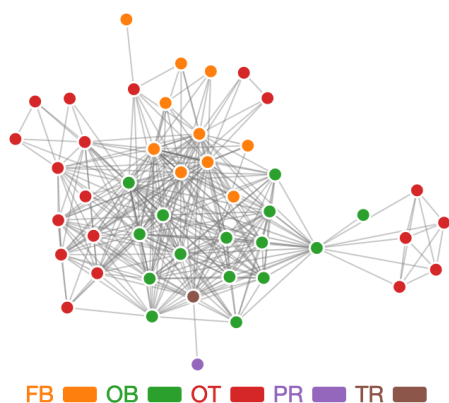


Figura 4: Grafo del grado de Derecho ($k = 1$).

Un ejemplo de lo primero se da en la Figura 4. Obsérvese cómo los nodos centrales son los correspondientes a las asignaturas OB (verde), que contienen la materia principal del grado y generalmente se toman en el segundo y tercer año.

La Figura 5 es un ejemplo de lo segundo. Vemos algunas asignaturas OT (rojo) conectadas solo al clúster principal por unas pocas aristas, que corresponden a las menciones en bioinformática, matemáticas y administración de empresas.

Podemos preguntarnos si estos grafos son indicativos de una buena estructura del plan de estudios. En algunos casos, visualizar el grafo con un umbral bajo de k da como resultado un grafo (casi) completamente conectado, pero el grafo generalmente muestra alguna estructura desde $k = 5$ en adelante. Esto siempre sucederá, a menos que todas las competencias estén presentes en todas las asignaturas, en cuyo caso el plan de estudios probablemente debería revisarse.

Para facilitar la inspección de los grafos, se han implementado algunas opciones de filtrado. La cantidad

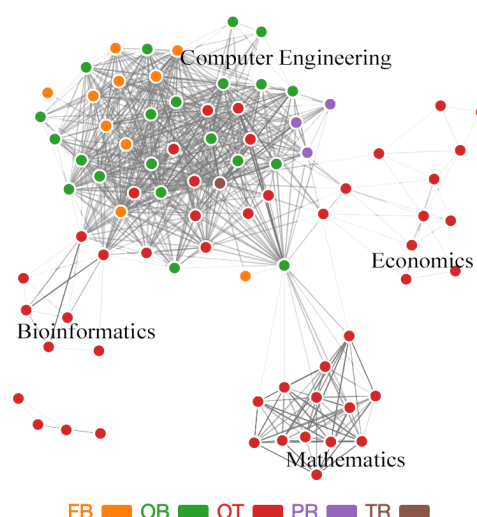


Figura 5: Grafo del grado de Ingeniería Informática ($k = 1$).

de aristas en los grafos es considerablemente grande, por lo que el filtrado debe hacerse desde SQL (al realizar las operaciones JOIN necesarias). El filtro más importante viene dado por la selección de años iniciales y finales. De esta manera, solo se consideran las asignaturas que se imparten durante ese rango de años, lo cual es útil para considerar asignaturas optativas que varían con el tiempo.

El aumento del umbral es útil para identificar clústeres fuertes. Por ejemplo, el grado de Lingüística (Figura 6) con umbral bajo tiene un grafo denso, mientras que elevando a $k = 5$ se producen clústeres naturales que uno puede identificar y etiquetar fácilmente.

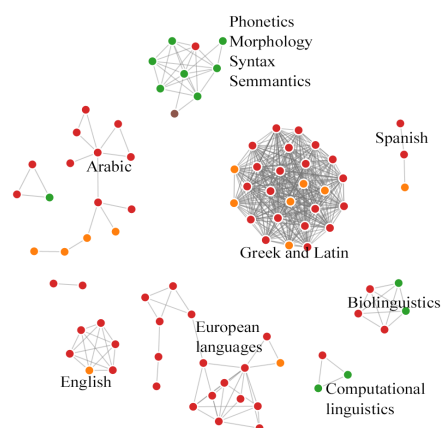


Figura 6: Grafo del grado de Lingüística con umbral alto ($k = 5$).

⁵<https://github.com/d3/d3-force>

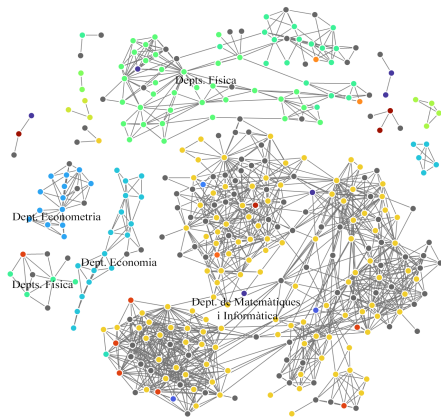


Figura 7: Grafo de colaboración docente para el grado de Matemáticas.

4.2. Colaboración docente

Para analizar la colaboración docente, además de los datos ya mencionados, también reunimos los grupos de clase para cada asignatura, el tipo de grupo, y el profesor a cargo de cada grupo. Como se espera que la mayoría de las asignaturas tengan varios grupos, podemos relacionar a dos profesores si han co-impartido una asignatura determinada. A esto lo llamamos *grafo de colaboración docente*.

A nivel organizativo, esta visualización ayuda a identificar posibles fallos en la estructura departamental (por ejemplo, profesores que imparten muchas asignaturas diferentes o solo una asignatura a lo largo del tiempo). Hemos podido mostrar esto coloreando el grafo de acuerdo con los departamentos asignados, información que también tenemos en nuestra base de datos.

La Figura 7 presenta el grafo de colaboración docente para el grado de Matemáticas. En la componente conexa amarilla, tenemos los tres grupos en el Departamento de Matemáticas e Informática.

4.3. Discusión y limitaciones

Hemos creado un software para visualizar asignaturas y relaciones entre estas⁶, así como patrones de colaboración docente. Como se dijo al comienzo de esta sección, el objetivo principal de estas visualizaciones es ayudar al personal académico a tomar decisiones informadas sobre la planificación y renovación del plan de estudios. Esto se puede lograr detectando patrones al visualizar los datos, lo cual ha sido ilustrado en las diferentes figuras.

Los principales tipos de patrones observados son la agrupación de asignaturas por área temática, y la agru-

pación por categorías (FB, OB, ...), que están relacionadas con el año en que un estudiante típicamente tomaría esa asignatura. Proponemos que un currículum bien estructurado será aquel en el que las asignaturas se agrupen naturalmente según los contenidos y el área temática, ya que esto aclarará varios indicadores: requisitos previos, solapamiento de contenidos y equilibrio general de las materias.

El filtrado nos ha permitido comprender mejor los clústeres: al establecer umbrales de conectividad más altos, podemos detectar las principales agrupaciones de asignaturas.

Otros estudios [6, 19] adoptaron el enfoque de mapear todas las combinaciones de asignaturas que los estudiantes habían tomado. Ese enfoque no tiene sentido en nuestro contexto, ya que todos nuestros estudiantes tienen una gran cantidad de materias obligatorias comunes. Por lo tanto, no podemos incorporar en nuestro caso esta información a la predicción del desempeño de los estudiantes. Esto es una limitación importante, ya que no se puede ver si existe una correlación entre la elección de asignaturas y las calificaciones de un estudiante (cf. Sección 2).

5. Conclusiones y trabajo futuro

Este artículo presenta una amplia gama de técnicas de aprendizaje automático y visualizaciones utilizadas para resolver diferentes problemas de *learning analytics* y extraer automáticamente tres indicadores clave para la gestión y la evaluación de la calidad de los grados universitarios.

Los sistemas basados en datos presentados constituyen una herramienta de LA que puede adaptarse eficientemente a cualquier grado y actualizarse periódicamente. El personal académico puede obtener fácilmente una visión general de cómo están evolucionando los estudios y tener conocimientos extra para tomar mejores decisiones. Dado que esta herramienta revela información sobre estudiantes, profesores y calificaciones, debe ser confidencial y restringida al personal académico.

Hemos mostrado resultados prometedores para la predicción de las calificaciones de los estudiantes y el abandono a partir de los datos de la UB. Sin embargo, es necesario un estudio más general que aumente el número de estudiantes, variedad de grados, y universidades involucradas en el trabajo para validar el sistema en otros escenarios.

Los resultados obtenidos mediante el análisis de sentimiento en las respuestas de texto abierto a las encuestas de estudiantes son alentadores. Recopilando y anotando más datos, podríamos obtener mejores resultados utilizando modelos de redes neuronales. Como trabajo futuro, queremos analizar las respuestas numé-

⁶Código disponible en <https://github.com/3nr1c/grad-db>. Hay una demostración limitada disponible en <https://grad-db-demo.herokuapp.com/>

ricas conjuntamente con las preguntas de texto abierto, estableciendo su correlación con el sentimiento. Una vez validado el modelo, la suma de los comentarios positivos y negativos se puede mostrar en los informes de satisfacción.

Con respecto al análisis de competencias de asignaturas, hemos comenzado una línea en el análisis y visualización de planes de estudios, que se inspira en parte en la literatura existente realizada en universidades norteamericanas. El siguiente paso es analizar en profundidad los grafos generados, con métricas como el grado de conectividad y la centralidad. Estas métricas permitirán una mejor comprensión de los programas académicos. Esperamos poder generar estrategias específicas para mejorar planificaciones de estudios en función de este tipo de visualizaciones y datos.

Agradecimientos

Esta investigación ha recibido soporte parcial de la Universitat de Barcelona: Convocatòria d'Ajuts a la Recerca en Docència Universitària de l'Institut de Ciències de l'Educació de la Universitat de Barcelona REDICE-18 y Proyecto 2014PID-UB/068.

Referencias

- [1] AQU. Indicators for educational development and analysis of qualifications. <http://winddat.aqu.cat>, 2016.
- [2] Dobrescu Alexandra Balahur, Rada Mihalcea y Andrés Montoyo. Computational approaches to subjectivity and sentiment analysis: present and envisaged methods and applications. *Computer Speech & Language*, 28(1):1-6, 2014.
- [3] Jeremy Barnes, Roman Klinger y Sabine Schulte im Walde. Assessing state-of-the-art sentiment models on state-of-the-art sentiment datasets. En *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, páginas 2-12, Copenhagen, Denmark. Association for Computational Linguistics, 2017.
- [4] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall y W. Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321-357, 2002.
- [5] R. Conijn, C. Snijders, A. Kleingeld y U. Matzat. Predicting student performance from lms data: a comparison of 17 blended courses using moodle lms. *IEEE Transactions on Learning Technologies*, 10(1):17-29, 2017.
- [6] Shane Dawson y Harry Hubball. Curriculum analytics: application of social network analysis for improving strategic curriculum decision-making in a research-intensive university. *Teaching & Learning Inquiry: The ISSOTL Journal*, 2(2):59-74, 2014.
- [7] G. Dempere, E. Puertas y L. Igual. Data science for on-the-go prediction of student performance. En *IEEE Data Science Workshop*, 2018.
- [8] Michael D. Ekstrand, John T. Riedl y Joseph A. Konstan. Collaborative filtering recommender systems. *Foundations and Trends® in Human-Computer Interaction*, 4(2):81-173, 2011.
- [9] Indra Ikauniece, Venelin Kovachev y Eloi Puertas. Spanish and Catalan Polarity Detection in Student Satisfaction Surveys, 2019.
- [10] L. Johnson, S. Adams Becker, M. Cummins, V. Estrada, A. Freeman y C. Hall. Nmc horizon report: 2016 higher education edition, 2016.
- [11] Johann Ari Larusson y Brandon White. *Learning analytics: From research to practice*. Springer New York, enero de 2014, páginas 1-195.
- [12] Jiaying Liu, Tao Tang, Wei Wang, Bo Xu, Xiangjie Kong y Feng Xia. A survey of scholarly data visualization. *IEEE Access*, PP:1-1, marzo de 2018.
- [13] University of California Riverside. Optimizing resource allocation for teaching. <https://bit.ly/2zPeTRj>, 2018.
- [14] David Rodríguez-Gómez, Mònica Feixas, Joaquín Gairín y José Luís Muñoz. Understanding catalan university dropout from a cross-national approach. *Studies in Higher Education*, 40(4):690-703, 2015.
- [15] Cristobal Romero, Pedro G. Espejo, Amelia Zafra, Jose Raul Romero y Sebastian Ventura. Web usage mining for predicting final marks of students that use moodle courses. *Computer Applications in Engineering Education*, 21(1):135-146, 2013.
- [16] S. Rovira, E. Puertas y L. Igual. Data-driven system to predict academic grades and dropout. *PLOS ONE*, 2017.
- [17] Guy Shani y Asela Gunawardana. *Evaluating recommendation systems*. En *Recommender Systems Handbook*. Springer US, Boston, MA, 2011, páginas 257-297.
- [18] Alex J. Smola y Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199-222, agosto de 2004.
- [19] Karen E. Willcox y Luwen Huang. Network models for mapping educational data. *Design Science*, 3:e18, 2017.
- [20] Kelly H. Zou, Kemal Tuncali y Stuart G. Silverman. Correlation and simple linear regression. *Radiology*, 227(3):617-622, junio de 2003.