

Minería de Datos Aplicada a la Educación: Modelo de Deserción Universitaria en la Universidad Tecnológica Nacional Facultad Regional Rosario

Luciano Valía, José Rostagno, Juan Miguel Moine,
Cristian Bigatti, Fabiana María Riva, Eduardo Amar

*Departamento de Ingeniería en Sistemas de Información
Facultad Regional Rosario
Universidad Tecnológica Nacional*

*lucianovalia@hotmail.com, joserostagno@hotmail.com, juanmiguelmoine@gmail.com, cristianbigatti@gmail.com,
fabianamriva@gmail.com, eduardoamar000@gmail.com*

Resumen

La deserción universitaria es uno de los mayores problemas que atraviesan las Universidades de nuestro país. La Minería de Datos es una disciplina que puede colaborar con esta situación, brindando un conjunto de técnicas y métodos para detectar patrones de comportamiento en los alumnos y accionar en consecuencia. Por este motivo, abordamos un estudio cuyo objetivo es el descubrimiento de patrones de deserción universitaria en la carrera Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Rosario.

Palabras clave: *deserción universitaria, minería de datos, estudiantes universitarios, patrones de comportamiento.*

Contexto

Este trabajo fue realizado en el marco del Proyecto de Investigación y Desarrollo: Modelo de Minería de Datos para la identificación de perfiles de deserción universitaria en alumnos de ISI (PID-UTN 3802).

Introducción

En la actualidad, la deserción universitaria representa un problema importante que enfrentan las universidades de nuestro País y del exterior. Se estima que en la Argentina alrededor del 80% de los alumnos que ingresan a la universidad desertan [1]. La Universidad Tecnológica Nacional Facultad

Regional Rosario, al igual que la mayoría de las universidades de nuestro País, no está exenta a esta situación en su carrera Ingeniería en Sistemas de Información.

Los sistemas de información en la universidad almacenan día a día una gran cantidad de datos, tanto del estado académico como de la situación socioeconómica de los alumnos, los cuales pueden ser explorados por disciplinas como la Minería de Datos para obtener nuevo conocimiento. Minería de Datos (también llamado Explotación de Información) es el proceso de extraer conocimiento útil, comprensible y novedoso de grandes volúmenes de datos, siendo su principal objetivo encontrar información oculta o implícita que no es posible obtener mediante métodos estadísticos convencionales. El resultado del proceso es un conjunto de patrones (modelos), los cuales pueden ser convertidos en información valiosa para la toma de decisiones [2].

En general, el problema de deserción universitaria ha sido analizado en función de dos grandes grupos de variables que caracterizan a los alumnos. En primer lugar desde el punto de vista académico, tomando variables como el grado de avance en la carrera, medido como el número de materias aprobadas y el promedio de calificaciones del alumno [3,4] y en segundo lugar trabajando exclusivamente con variables socioeconómicas como el máximo nivel de estudios alcanzados por los padres, situación laboral y edad [5,6,7].

Si bien algunos incorporan el estudio simultáneo de variables académicas y socioeconómicas los mismos fueron realizados con información proveniente del sistema de gestión académica, señalando sus autores la mala calidad de la misma (desactualizada e incompleta), especialmente aquella referida a aspectos sociales y económicos [8,9]. En [10] se demuestra que a partir de un relevamiento de 8 estudios previos, sólo 2 de ellos fueron realizados con datos de buena calidad. Esto se muestra gráficamente en la Figura 1, donde se ve que el 75% de las investigaciones han utilizado datos de mala calidad, incompletos y desactualizados.

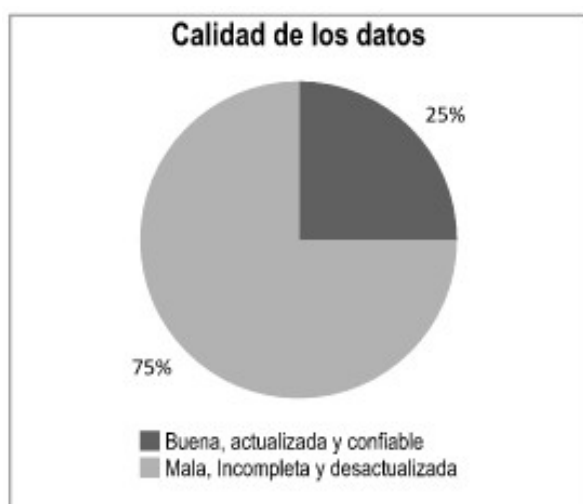


Fig. 1. Calidad de los datos utilizados en estudios previos

En este proyecto se propone el empleo de técnicas y métodos de Minería de Datos para la identificación de patrones de deserción en alumnos de los primeros años de la carrera Ingeniería en Sistemas de Información en la Universidad Tecnológica Nacional Facultad Regional Rosario, utilizando información actualizada proveniente de encuestas presenciales.

Líneas de investigación y desarrollo

En el marco de este proyecto se han definido las siguientes líneas de investigación:

- La evaluación y comparación de diferentes estudios previos realizados en materia de deserción universitaria utilizando técnicas de

Minería de Datos, con el fin de establecer una discusión sobre un criterio de deserción.

- La definición de variables socioeconómicas y académicas involucradas en el proceso de deserción de los alumnos. La recolección de datos y su posterior procesamiento mediante un análisis exploratorio que permita conocer la naturaleza de los mismos.
- La creación y evaluación de modelos de Minería de Datos, tanto predictivos como descriptivos, que permitan mejorar la toma de decisiones relacionadas con esta problemática.

Resultados obtenidos/esperados

En [10] hemos presentado un estudio comparativo de trabajos previos sobre la problemática de deserción universitaria en diferentes universidades de nuestro País y del exterior.

Además, a partir de una recolección de datos realizada en el año 2015, se ha logrado elaborar un estudio de las variables socioeconómicas y académicas que caracterizan a los alumnos de la UTN Facultad Regional Rosario [11]. Entre los resultados más interesantes se pudo detectar que los hábitos de estudio (cantidad de horas de estudio y asistencia a clase de consulta) varían en función del año de cursado y de la localidad de procedencia (Figura 2). También se encontró un patrón interesante en los alumnos que tienen becas, donde sus rendimientos son inferiores a quienes no la tienen o reciben subsidio del Estado (Figura 3). Además, se pudo observar que el desempeño académico (cantidad de materias aprobadas) varía según el nivel de educación de los padres (Figura 4), y no resulta significativamente influenciado por la situación laboral o el tipo de estudio secundario del alumno.

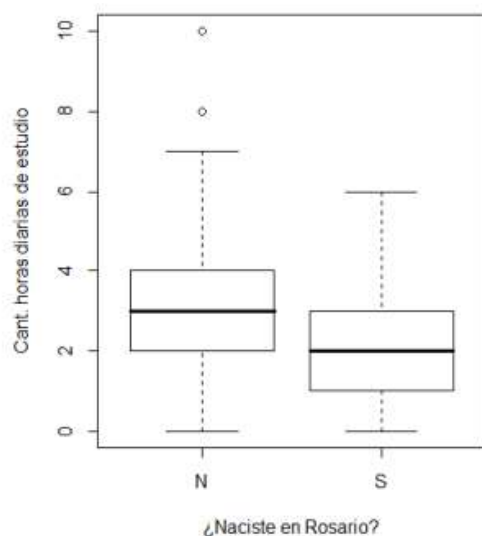


Fig. 2. Nació en Rosario vs Cantidad de horas de estudio.

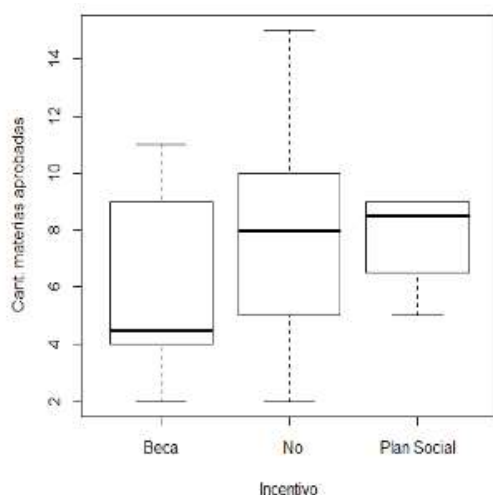


Fig. 3. Recibe incentivo vs Cantidad de materias aprobadas.

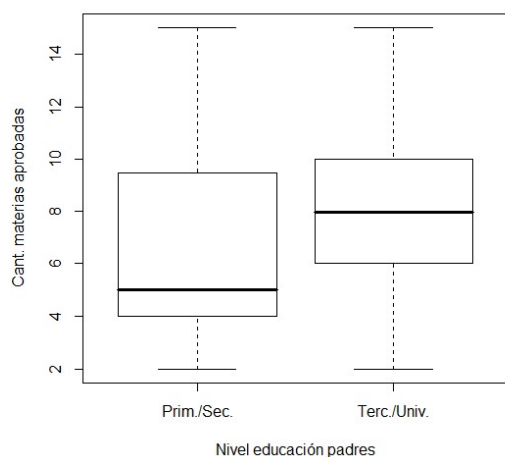


Fig. 4. Nivel de estudio padres vs Cantidad de materias aprobadas

Siguiendo esta línea de trabajo se espera poder recolectar información sobre el estado académico actual de los alumnos relevados, para comenzar con la elaboración de modelos predictivos y descriptivos que permitan estimar la deserción universitaria en nuestra institución. Para esta tarea, se considerará inicialmente que un alumno desertó si luego de dos años de haber sido encuestado (es decir, al inicio del ciclo 2017) no se encuentra inscripto en alguna materia de la carrera.

Formación de los Recursos Humanos

En el marco de este proyecto de investigación se encuentran trabajando alumnos y graduados de la Universidad Tecnológica Nacional Facultad Regional Rosario bajo la coordinación de docentes e investigadores de la institución. Se han realizado investigaciones sobre temáticas actuales de Minería de Datos, para su posterior divulgación tanto en jornadas abiertas a la comunidad como en cátedras de la carrera Ingeniería en Sistemas de Información.

Referencias

1. Lorenzano, C. "La Deserción Universitaria en la Universidad de Tres de Febrero, Universidad Tres de Febrero". Recuperado el 2 de marzo 2015, de <http://www.untref.edu.ar/documentos/AutoevaluacionLadesercion.pdf>
2. Hernández Orallo, Ramirez Quintana y Ferri Ramirez (2004) "Introducción a la Minería de Datos". Editorial Pearson Prentice Hall. España. ISBN 84-205-4091-9.
3. Balestieri, C. & otros (2013) "Modelo neuronal para la estimación del riesgo de deserción de alumnos de grado". EST 2013, 42 JAIIO.
4. Karamouzis, S. T., & Vrettos, A. (2008) "An artificial neural network for predicting student graduation outcomes". World Congress on Engineering and Computer Science (pp. 991-994).

5. Formia, S. (2013) “Evaluación de técnicas de Extracción de Conocimiento en Bases de Datos y su aplicación a la deserción de alumnos universitarios” (Tesis Especialidad). Universidad Nacional de La Plata.
6. Giovagnoli, P. I. (2002) “Determinantes de la deserción y graduación universitaria: una aplicación utilizando modelos de duración”. Documentos de Trabajo. Universidad Nacional de La Plata. Recuperado el 16 de febrero de 2016 de <http://sedici.unlp.edu.ar/handle/10915/3436>.
7. Valero, S., & otros (2010) “Minería de datos: predicción de la deserción escolar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos”. Universidad Tecnológica de Izúcar de Matamoros. Recuperado el 2 de marzo de 2015, de <http://www.utim.edu.mx/~svalero/docs/e1.pdf>.
8. Pereira, R. T., Romero, A. C., & Toledo, J. J. (2013) “Descubrimiento de perfiles de deserción estudiantil con técnicas de minería de datos”. Vínculos, vol10, pág. 373-383.
9. Fischer, E. (2012) “Modelo para la Automatización del Proceso de Determinación de Riesgo de Deserción en Estudiantes Universitarios”. (Tesis doctoral). Universidad de Chile.
10. Rostagno, José; Moine, Juan Miguel; Bigatti, Cristian Germán. (2016) “Abordaje del fenómeno de deserción universitaria con técnicas de minería de datos”. Revista Rumbos Tecnológicos. Volumen 8. Universidad Tecnológica Nacional Facultad Regional Avellaneda.
11. Valía, Luciano; Moine, Juan Miguel; Rostagno, José y Otros (2016) “Hacia un modelo de deserción universitaria: Análisis exploratorio de variables socioeconómicas y académicas de alumnos de primer y segundo año de la carrera Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional Facultad Regional Rosario”. IV Congreso Nacional de Ingeniería Informática y Sistemas de Información. Universidad Católica de Salta.