

Personalización de la Educación a través de la creación de Perfiles dinámicos de los alumnos

Javier Díaz¹, Laura Lanzarini², María Emilia Charnelli¹,

¹Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata
javierd@info.unlp.edu.ar,
mcharnelli@linti.unlp.edu.ar

²Instituto de Investigación en Informática LIDI (III-LIDI).
Facultad de Informática. Universidad Nacional de La Plata
laural@lidi.info.unlp.edu.ar

Resumen

La línea de investigación actual continúa con el estudio y desarrollo en el área de la Minería de Datos Educativa. Se ha puesto énfasis en el análisis y aplicación de diferentes técnicas no supervisadas y de técnicas de visualización de datos masivos con el objetivo de identificar las características más relevantes de los alumnos de la Facultad de Informática de la UNLP. A su vez, se espera poder contribuir en el área educativa a través de la determinación de perfiles dinámicos de los alumnos de la Facultad de Informática en lo que se refiere a su interacción con recursos educativos de acceso libre y su interacción con las redes sociales. Dichos perfiles podrán ser utilizados para caracterizar su comportamiento actual y asistirlo en forma automática a fin de que pueda alcanzar el comportamiento esperado.

Palabras clave: analítica del aprendizaje, minería de datos educativa, sistemas de recomendación

Contexto

Este trabajo de investigación se encuadra en el Proyecto de Incentivos acreditado “Internet del futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro” desarrollado en el Laboratorio de Investigación en Nuevas Tecnologías Informáticas, LINTI.

Introducción

La Analítica del Aprendizaje (Learning Analytics) es un campo de investigación emergente que analiza información referida a personas que están aprendiendo y sus contextos, con el propósito de comprender y optimizar el aprendizaje y los entornos en los que dicho aprendizaje sucede.

En los últimos años, las instituciones educativas se han embarcado en su propia exploración de grandes conjuntos de datos para mejorar los índices de retención y proporcionar una experiencia personalizada y de mayor calidad para los estudiantes. La aplicación de técnicas de

Minería de Datos en el ámbito educativo ha permitido caracterizar a los distintos actores que intervienen en los procesos de enseñanza y aprendizaje [1].

A medida que el aprendizaje se traslada a la Web, crece la cantidad de datos que pueden utilizarse para mejorar y personalizar este proceso.

A continuación se describen las tareas realizadas recientemente.

Análisis de relevancia de atributos

A partir de las líneas de investigación presentadas en WICC 2015 [2], se continuó trabajando sobre los datos académicos y personales de los alumnos de la Facultad de Informática de la UNLP con el objetivo de mejorar la caracterización de los alumnos que pierden la regularidad.

La información utilizada provino del sistema SIU-Guaraní e involucró a 5268 alumnos, comprendidos entre los años 2002 y 2012.

Para la representación de los datos del problema se tuvo en cuenta el trabajo realizado en [3]. Dicha representación, si bien permite incluir los distintos aspectos que caracterizan los motivos de pérdida o no de regularidad por parte del alumno, dan como resultado una cantidad de atributos o variables elevado dificultando la identificación de patrones o relaciones existentes. Por lo tanto, fue preciso aplicar técnicas de Selección de Características y Visualización para encontrar la información más relevante.

A partir de los atributos seleccionados se construyeron tres tipos de modelos diferentes que permitieron clasificar a los alumnos en regulares y no regulares utilizando los siguientes métodos como el C4.5, PART y un multiperceptrón entrenado con el algoritmo de backpropagation. De los modelos obtenidos se pudo afirmar que los

atributos seleccionados son adecuados para predecir la condición de no regularidad de un estudiante. También dejan de manifiesto la no incidencia de la situación laboral actual de los alumnos en lo que se refiere a su rendimiento académico. Los resultados de esta investigación fueron publicados en [4].

También se estudiaron y aplicaron técnicas de aprendizaje no supervisado para evaluar la performance de los alumnos a lo largo de los primeros cinco años de la carrera. Se aplicaron técnicas de clustering para obtener los grupos de alumnos con un desempeño similar en la carrera. Los resultados de esta investigación fueron publicados en [5].

En lo que se refiere al uso de técnicas de visualización como herramienta para seleccionar atributos relevantes consideramos que, en educación, resultan de suma utilidad ya que permiten a los docentes, independientemente del área en que se desarrollen, hacer uso de las mismas para analizar a sus alumnos sin necesidad de contar con conocimientos específicos de la minería de datos.

En esta dirección, se analizaron técnicas de visualización para representar la información de los alumnos de la Facultad de Informática. Se utilizó una visualización de conjuntos paralelos para identificar tendencias en las respuestas [4]. Uno de los resultados obtenidos muestra que de los alumnos egresados, la mayoría de ellos en 5to año sólo tenían aprobado el 50% de la carrera. Así también se realizaron diagramas de desviación sobre los clústeres obtenidos. Como resultado se observó que se produce un punto de inflexión en el segundo año. A partir de ese momento una gran cantidad de alumnos detienen su progreso en la carrera.

Así también, se utilizaron técnicas de visualización para obtener los atributos

más representativos para construir un modelo de clasificación que permite describir y caracterizar a los alumnos según su condición de regularidad. Los resultados de este trabajo fueron publicados en [5].

Minería de Datos Masivos

Además de la información personal y académica obtenida del sistema de gestión y los entornos de enseñanza y aprendizaje, resulta de interés obtener información adicional sobre la interacción de los alumnos con las redes sociales [6]. Actualmente se está realizando un trabajo que analiza las relaciones existentes entre las publicaciones realizadas por los alumnos de la Facultad de Informática de la UNLP en las redes sociales. En este caso, las publicaciones son de texto libre y no están estructuradas por lo que las técnicas aplicadas en la etapa de extracción de conocimiento deben ser diferentes a las aplicadas con los datos personales y académicos de los alumnos. En un principio, se buscará la relevancia de cada uno de los tópicos y la posibilidad de descubrir agrupamientos subyacentes entre ellos. Estos resultados ayudarán a entender mejor el perfil de los alumnos de nuestra facultad.

Modelado de Perfiles de Alumnos y Sistemas de Recomendación

Uno de los puntos centrales de esta línea de investigación consiste en modelar perfiles dinámicos a través de información provista por los entornos de enseñanza y aprendizaje que usan las diferentes cátedras de la facultad, información de los préstamos de la biblioteca, actividad en las redes sociales, entre otros [7] [8] [9] [10].

Estos perfiles permitirán proporcionar información para un sistema de recomendación [11] [12] [13].

El sistema recomendador se desarrollará como una extensión de los entornos de aprendizaje, que utilizando la información de los perfiles obtenidos pueda recomendar materiales educativos [14].

Líneas de Investigación, Desarrollo e Innovación

- Estudio de técnicas de agrupamiento aplicables a datos masivos.
- Estudio de distintas técnicas de preprocesamiento aplicables a Minería de Textos.
- Estudio, análisis y comparación de diferentes técnicas de visualización para grandes volúmenes de datos.
- Revisión y análisis de técnicas específicas de Learning Analytics.
- Estudio y desarrollo sobre modelado de perfiles dinámicos de los alumnos según sus características más relevantes.
- Estudio de algoritmos de recomendación.
- Desarrollo de un sistema recomendador de materiales educativos utilizando la información de los perfiles obtenidos.

Resultados y Objetivos

- Detección de las características más relevantes de los estudiantes aportando información útil en relación a los datos personales y el rendimiento académico.
- Desarrollo de una prueba de concepto que arroja información preliminar adecuada para predecir la condición de no

regularidad de un estudiante y la incidencia de su situación laboral.

- Estudio y aplicación de técnicas de Minería de Datos para datos masivos con el objetivo de analizar información de las redes sociales.
- Estudio de las técnicas existentes para diseñar un sistema recomendador a partir de perfiles dinámicos de los alumnos.

Formación de Recursos Humanos

Dentro de los temas involucrados en esta línea de investigación, actualmente hay 2 investigadores realizando su doctorado, y uno de ellos a su vez, finalizando una maestría.

Referencias

[1] Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(6), (pp. 601-618).

[2] Díaz, Francisco Javier; Lanzarini, Laura Cristina; Charnelli, María Emilia; Baldino, Guillermo; Schiavoni, María Alejandra; Amadeo, Ana Paola. Analítica del aprendizaje y la personalización de la educación. *XVII Workshop de Investigadores en Ciencias de la Computación*. 2015, Salta.

[3] Charnelli, E. Lanzarini, L. Baldino, G. Diaz, F. Determining the profiles of young people from Buenos Aires with a tendency to pursue computer science studies. *XX Congreso Argentino de*

Ciencias de la Computación. La Matanza, 2014.

[4] Lanzarini, L.; Charnelli, M.E.; Diaz, J. Academic performance of university students and its relation with employment. *Latin American Computing Conference (CLEI), Perú 2015. IEEE Conference Publications*.

[5] Lanzarini, Laura Cristina; Charnelli, María Emilia; Baldino, Guillermo; Díaz, Javier F. Selección de atributos representativos del avance académico de los alumnos universitarios usando técnicas de visualización. Un caso de estudio. *Revista: TE & ET*; no. 15 . ISSN: 1850-9959. p. 42-50. 2015

[6] KAYA, Tugberk; BICEN, Huseyin. The effects of social media on students' behaviors; Facebook as a case study. *Computers in Human Behavior*, 2016, vol. 59, p. 374-379.

[7] W. Paireekreng and T. Prexawanprasut, "An integrated model for learning style classification in university students using data mining techniques," *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2015 12th International Conference on*, Hua Hin, 2015, pp. 1-5.

[8] Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Kloos, C. D. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior*, 47, 139-148.

- [9] Chrysafiadi, K., & Virvou, M. (2015). Student Modeling for Personalized Education: A Review of the Literature. In *Advances in Personalized Web-Based Education* (pp. 1-24). Springer International Publishing.
- [10] Mota, P., Melo, F., & Coheur, L. (2015, May). Modeling Students Self-Studies Behaviors. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (pp. 1521-1528). International Foundation for Autonomous Agents and Multiagent Systems.
- [11] Drachsler, H., Verbert, K., Santos, O. C., & Manouselis. Panorama of recommender systems to support learning. In *Recommender systems handbook* (pp. 421-451). Springer US, 2015.
- [12] Tejada-Lorente, Á., Bernabé-Moreno, J., Porcel, C., Galindo-Moreno, P., & Herrera-Viedma, E. (2015). A Dynamic Recommender System as Reinforcement for Personalized Education by a Fuzzly Linguistic Web System. *Procedia Computer Science*, 55, 1143-1150.
- [13] LI, Yanyan, et al. Designing a Learning Recommender System by Incorporating Resource Association Analysis and Social Interaction Computing. *En State-of-the-Art and Future Directions of Smart Learning*. Springer Singapore, 2016. p. 137-143.
- [14] J. Díaz, A. Schiavoni, P. Amadeo, M. E. Charnelli, Búsqueda personalizada de recursos educativos abiertos basada en el perfil del alumno dentro de un entorno educativo, *XVI Workshop de Investigadores en Ciencias de la Computación*, Tierra del Fuego, 2014.