

Caracterización preliminar de los aspirantes a ingresar a la Licenciatura en Ciencias de la Computación

Carlos García Garino^{1,2}, Marisa Haderne¹, Lucía Cortés¹, Carlos Catania^{1,2}, Elina Pacini^{1,2,3}

¹ Facultad de Ingeniería. Universidad Nacional de Cuyo.

² ITIC, Universidad Nacional de Cuyo.

³ CONICET, Carrera del Investigador Científico Técnico

cgarcia@itu.uncu.edu.ar, hadernemarisa@yahoo.com.ar; luciacortes5519@gmail.com, ccatania@itu.uncu.edu.ar, epacini@uncu.edu.ar

RESUMEN

La deserción de los estudiantes universitarios en los primeros años de la carrera es una preocupación presente en todas las instituciones. Por este motivo, es importante analizar el problema en una etapa previa al cursado de la carrera, cuando el estudiante está transitando su ingreso a la institución seleccionada. Los resultados obtenidos pueden luego ser contrastados con el rendimiento de los estudiantes en los primeros años de la Universidad.

El presente trabajo apunta a caracterizar el comportamiento de los aspirantes a ingresar en la Licenciatura en Ciencias de la Computación, mediante técnicas de minería de datos. Para ello, se dispone de datos censales y rendimiento durante el cursado del ingreso de las cohortes 2017 y 2018.

Para explicar el problema se han propuesto diversas variables: situación socioeconómica, condición cultural, institución en la cual cursó estudios medios, rendimiento en los exámenes de ingreso, motivación personal, etc. Todas ellas comprendidas dentro de las características y comportamientos propios del aspirante universitario.

En este trabajo se presentan los resultados obtenidos y se discuten los resultados esperados.

Palabras clave: 1) Deserción estudiantil 2) Minería de Datos Educativa 3) Rendimiento de los estudiantes

CONTEXTO

En la Figura 1 se muestran las líneas de investigación que se llevan a cabo en el Instituto Universitario para las Tecnologías de la Información y las Comunicaciones (ITIC), y que se comparten con la Carrera de Licenciatura en Ciencias de la Computación, que se dicta en la Facultad de Ingeniería de la Universidad Nacional de Cuyo (UNCuyo).

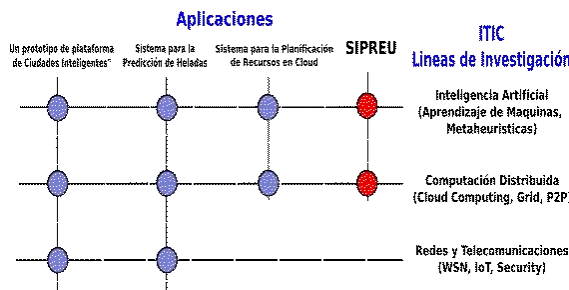


Figura 1: líneas de investigación del ITIC y carrera de Licenciatura en Ciencias de la Computación

En este trabajo se presentan los avances del proyecto “SIPREU: Un sistema inteligente para caracterizar y predecir el rendimiento de estudiantes universitarios” [7], cuyos lineamientos se comunicaron en el congreso WICC 2017 [9] y en la conferencia ADNTIIC [8]. En el proyecto PIN “La Minería de Datos como herramienta para estudiar el rendimiento de estudiantes de la Licenciatura

en Ciencias de la Computación” [2], también se estudian algunos aspectos de interés del problema planteado.

1. INTRODUCCIÓN

El problema planteado en el presente trabajo es relevante y de especial interés para la Facultad de Ingeniería de la Universidad Nacional de Cuyo. Es fundamental para la Facultad conocer al aspirante de la Licenciatura en Ciencias de la Computación, su problemática académica, trayectos educativos anteriores y otros componentes sociales que pueden afectar su rendimiento al momento del ingreso.

La carrera es reciente, su primera cohorte comenzó en el 2017 y acaba de iniciarse el cursado de la segunda cohorte. Para las autoridades es muy importante analizar y conocer las características de sus ingresantes desde el inicio de la carrera.

Por otra parte, el problema posee mucha relevancia local ya que se enmarca dentro del Objetivo 1 del Plan Estratégico 2021 de la UNCuyo [11].

En un trabajo anterior [7] se discutieron con detalle trabajos relacionados con la presente propuesta. A manera de síntesis puede mencionarse el trabajo de Garbanzo Vargas [6] que propone agrupar a los diferentes factores que caracterizan el comportamiento de un estudiante en tres grupos: a) Determinantes personales; b) Determinantes sociales y c) Determinantes institucionales.

Dado que muchos de los factores no se pueden medir con facilidad en el momento del acceso del estudiante a la Universidad, se tendrán en cuenta los factores propios de los determinantes sociales, ya que los mismos parecen, según estudios realizados, brindar un buen punto de partida.

En cualquier caso, los factores (atributos) para caracterizar la población inicial surgirán de los estudios mencionados y de una discusión interdisciplinar entre los integrantes del equipo de trabajo, entre quienes se encuentran especialistas en informática, postgraduados en educación, profesores universitarios, etc.

También se ha registrado el rendimiento de los aspirantes (cohortes 2017 y 2018) a

ingresar a la carrera de Licenciatura en Ciencias de la Computación. La información recogida es una muestra reducida, pero suficiente para brindar algunos indicadores útiles para avanzar en el proyecto, como se discute en la sección 3 de este trabajo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Además de la relevancia general y local del problema corresponde analizar la relevancia disciplinar del tema. La Minería de Datos Educativa es una línea que ha cobrado mucha relevancia mundial en los últimos años, como se observa en el portal dedicado al tema que lleva adelante la International Educational Data Mining Society [3]. En dicho sitio se expresa textualmente: “La Minería de Datos en educación es una disciplina emergente, orientada al desarrollo de métodos para explorar el creciente volumen de datos provenientes de los sistemas de educación, y utilizar esos métodos para comprender mejor a los estudiantes y sus entornos de aprendizaje”.

La citada sociedad edita un journal específico: Journal of Educational Data Mining y lleva adelante las conferencias internacionales acerca del tema, cuya próxima edición tendrá lugar en Buffalo, Nueva York en 2018.

3. RESULTADOS OBTENIDOS / ESPERADOS

Datos censales y de rendimiento

Se dispone de datos censales y de rendimiento de los aspirantes a ingresar en la carrera de Licenciatura en Ciencias de la Computación, correspondientes a las cohortes 2017 y 2018. Como se observa en la Figura 2, ingresaron 32 alumnos a la cohorte 2017 de la carrera.

La cohorte 2017 presentó la particularidad que 16 estudiantes (14%) ingresaron a través del otorgamiento de equivalencias, mientras que los otros 15 (13%) de los aspirantes aprobaron el curso de ingreso. El 73% restante se distribuye entre quienes

desaprobaron el ingreso (47%) y quienes no registraron actividad (26%).

Los resultados obtenidos reflejan una tendencia habitual en las carreras nuevas que incorporan muchos estudiantes con experiencias universitarias previas.

Para la cohorte 2018, como se muestra en la Figura 3, de un total de 110 aspirantes, aprobaron el ingreso 32 alumnos (29%).

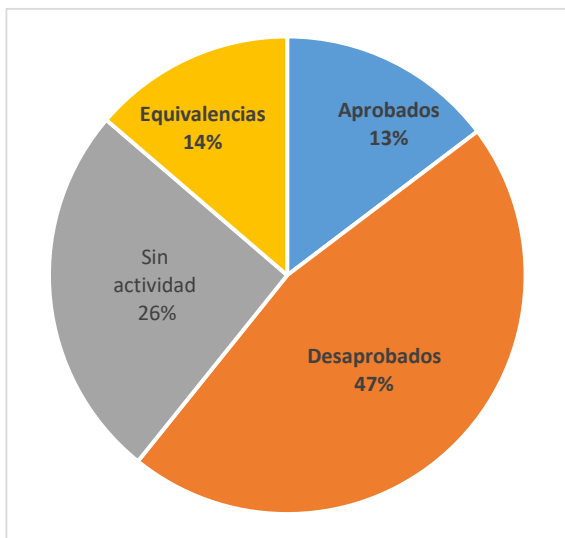


Figura 2: Cohorte 2017. Rendimiento de los aspirantes a ingresar en la Carrera de Licenciatura en Ciencias de la Computación.

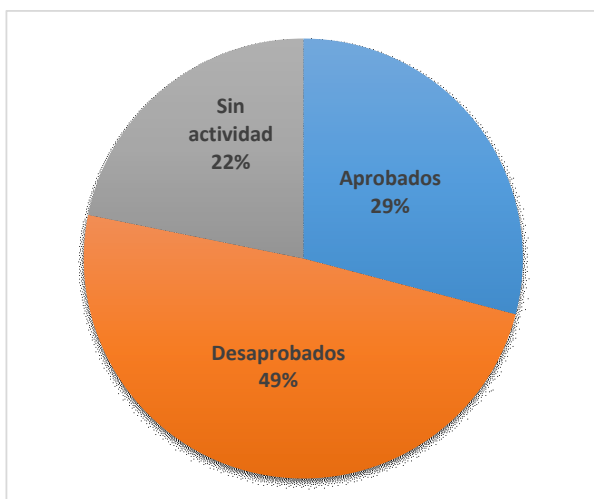


Figura 3: Cohorte 2018. Rendimiento de los aspirantes a ingresar en la Carrera de Licenciatura en Ciencias de la Computación.

Determinación de la muestra

Para determinar la muestra se integraron los archivos de datos personales y de calificaciones obtenidas por los aspirantes a ingresar en 2018. A partir del archivo unificado se hizo un análisis exhaustivo de los datos para determinar los atributos que serán tenidos en cuenta para trabajar en el modelo descriptivo/predictivo.

Del análisis que se llevó a cabo surge:

1. Se descartan los casos que no registran actividad, es decir, aspirantes inscriptos que no se presentaron a ninguna instancia de evaluación.
2. De las calificaciones obtenidas se observa que de los alumnos que aprobaron Matemática también aprobaron Resolución de Problemas, y no viceversa. Por este motivo se ha optado por tener en cuenta como atributo sólo el promedio de calificaciones obtenidas en Matemática.
3. También se descartan atributos repetitivos como País y Provincia de procedencia y Domicilio por protección de datos personales.
4. Existen instancias de las que no se disponen algunos atributos como Departamento y Localidad aunque en algunos casos pudieron completarse tomando en cuenta la característica del número telefónico.
5. En conclusión, la vista minable se integra de 116 instancias con 7 atributos: Sexo, Colegio Secundario, Tipo de escuela secundaria (Gestión privada o gestión pública: UNCuyo, Cens, Técnicas, Otros estatales) Departamento, Localidad, Promedio, Ingresó (SI/NO).
6. De las 116 instancias, en 47 se verifica "Si ingresó" y 69 "No ingresó", por lo tanto es posible a partir de esta muestra conseguir un modelo apropiado.

Resultados obtenidos

A partir de la muestra previamente descrita se realizó un análisis tipo exploratorio de los datos. Para ello, se utilizaron algunas

herramientas visuales propias de la estadística como barplots y boxplots. Este análisis se llevó a cabo con el software R [12].

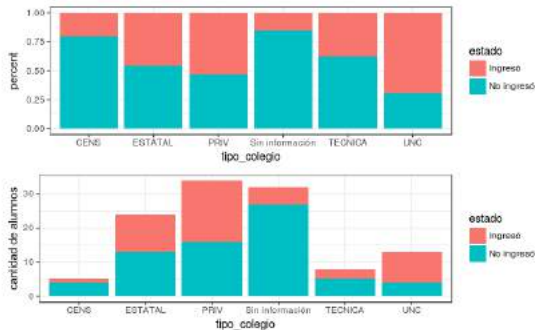


Figura 4: Distribución de la cantidad de aspirantes por tipo de colegio

En la figura 4, a priori no se observa diferencia entre los aspirantes provenientes de escuelas privadas y públicas. Si se observa un alto índice de aprobación en los aspirantes que provienen de colegios secundarios de la Universidad Nacional de Cuyo, indicados en los gráficos como UNC.

Posteriormente, se realizó el modelo de agrupación en clústeres de la muestra con la aplicación de Microsoft Clustering [10]. Este asistente divide los datos en un conjunto de entrenamiento (80%) y en un conjunto de pruebas (20%). Luego, se usa el conjunto de entrenamiento para entrenar los modelos y el conjunto de prueba para probar la precisión.



Figura 5. Modelo de agrupación en clústeres

Se puede observar en la Figura 5 que el cluster 1 posee la mayor cantidad de ingresantes, 95%, siendo el 54% de escuelas privadas, 20% estatales y 16% de la UNCuyo

Resultados esperados

Al final del proyecto se espera contar con mayor información acerca de los estudiantes de las cohortes 2017 y 2018. Para ello se espera enriquecer la información censal disponible mediante entrevistas personales a partir de un cuestionario actualmente en elaboración.

Además, se contrastará el rendimiento de los ingresantes con el correspondiente desempeño en primer año, y se podrán comparar los resultados de este proyecto con otros obtenidos mediante técnicas estadísticas [3,4]. Por otra parte, se ha comenzado a trabajar en la articulación horizontal entre dos asignaturas de primer año, concretamente entre Geometría Analítica e Introducción a la Programación y a contrastar el rendimiento de los estudiantes en dichas asignaturas. Algunos resultados preliminares [13] se comunicaron durante el congreso ENIDI. Finalmente se espera obtener información propia de los alumnos a partir de los trabajos de Brottier [1] que ha estudiado a las nuevas generaciones de estudiantes.

4. FORMACIÓN DE RECURSOS HUMANOS

La modalidad del proyecto SIPREU [5] financiado por la UNCuyo tiene, entre otros objetivos, la formación de Recursos Humanos. Una de las integrantes del proyecto está próxima a presentar su proyecto de tesis doctoral bajo la dirección de los investigadores formados del mismo. Otra integrante del equipo de trabajo actualmente lleva a cabo un proyecto de investigador novel PIN [2], financiado por la Facultad de Ingeniería de la Universidad Nacional de Cuyo, facilitando de esta forma su inicio en actividades de I+D.

Además, distintos integrantes del equipo de trabajo del presente proyecto han trabajado en los últimos años en diversas áreas relacionadas con la caracterización de los alumnos en la Universidad Nacional de Cuyo. Se contará con su experiencia para alcanzar los objetivos propuestos en el proyecto

SIPREU, así como incrementar las publicaciones de sus integrantes.

5. BIBLIOGRAFIA

- [1] L. Brottier: Proyecto 06/B005 Expectativas laborales de las nuevas generaciones. Proyecto bienal, Universidad Nacional de Cuyo, 2013-2015.
- [2] L. Cortés: La Minería de Datos como herramienta para estudiar el rendimiento de estudiantes de la Licenciatura en Ciencias de la Computación. Proyecto de Investigadores Noveles. Facultad de Ingeniería, Universidad Nacional de Cuyo, convocatoria 2017.
- [3] Educational Data Mining Web Site: <http://www.educationaldatamining.org/>
- [4] D. Fernández: Proyecto 06/B188 Una mirada introspectiva de los indicadores académicos que influyen tanto en el desempeño y el desgranamiento de los estudiantes como en la duración real de la carrera. Proyecto bienal, Universidad Nacional de Cuyo, 2009-2011
- [5] D. Fernández: Proyecto 06B/246 Detección, identificación y análisis de los factores fundamentales, externos e internos, que influyen y definen el ingreso de los aspirantes a Ingeniería y Arquitectura de la Universidad Nacional de Cuyo. Proyecto bienal, Universidad Nacional de Cuyo, 2011-13
- [6] G. M. Garbanzo Vargas. Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública: pp. 43-63, vol. 31, núm. 1, 2007, Educación. Universidad de Costa Rica. San Pedro, Montes de Oca. Disponible en: <http://www.redalyc.org/articulo.oa?id=44031103>
- [7] C. García Garino 06/B349 SIPREU: Un sistema inteligente para caracterizar y predecir el rendimiento de estudiantes universitarios. Proyecto Bienal Tipo II 2016 – 2018.
- [8] C. García Garino, M. Haderne, C. Catania, E. Pacini. SIPREU: An Intelligent System for characterization and Prediction of University Students performance. ADNTIIC 2016: Advances in New Technologies, Interactive Interfaces and Communicability. Córdoba.
- [9] M. Haderne, L. Cortés, C. García Garino, C. Catania y Elina Pacini: Caracterización del aspirante a ingresar a la Licenciatura en Ciencias de la Computación, pp. 231-235. Actas del XIX Workshop de investigadores en Ciencias de la Computación, WICC 2017. Buenos Aires, 27 y 28 de Abril de 2017. ITBA y Red UNCI. ISBN 978-987-42-5143-5.
- [10] Complementos de minería de datos de Microsoft SQL Server 2012 SP1 para Microsoft Office
- [11] Universidad Nacional de Cuyo: Plan estratégico 2021. Disponible en: <http://www.uncu.edu.ar/planificacion/upload/plan-estrategico-1.pdf>. 2014
- [12] R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

URL <https://www.R-project.org/>.

- [13] S. Raichman y E. Pacini. S. Raichman and E. Pacini. Intervención educativa de articulación entre las asignaturas Introducción a la Programación y Geometría Analítica. Noveno Encuentro de Investigadores y Docentes de Ingeniería EnIDI 2017, Ciudad de Mendoza, Argentina. En prensa