

# MINERÍA DE DATOS EN LA DETERMINACIÓN DEL PERFIL DE UN POTENCIAL ALUMNO UNIVERSITARIO DESERTOR

Mag. Graciela Beguerí, Mag. Alejandra Malberti, Mag. Raúl O. Klenzi  
Facultad de Ciencias Exactas, Físicas y Naturales – Universidad

Nacional de San Juan

Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas", San Juan  
{grabeda, amalberti, rauloscarklenzi}@gmail.com

## Resumen

La presente propuesta versa sobre cómo construir un prototipo mediante el cual se pueda reconocer si un alumno universitario es un posible desertor. Para este fin, se propone encontrar y modelar el perfil de un alumno por medio de ciertas características que permitan reconocerlo como posible desertor en el marco de una carrera universitaria. En este sentido, se pretende medir la capacidad de resiliencia valiéndose de instrumentos de recolección de datos, a los que se aplicarán análisis estadísticos y estrategias de minería de datos

**Palabras clave:** Análisis Estadístico - Minería de Datos – Resiliencia – Educación Superior

## Contexto

La determinación automática de perfiles de usuarios por medio de la aplicación de herramientas estadísticas y estrategias de minería de datos, es un tópico de sumo interés en distintas áreas de aplicación en general y en educación en particular. El proyecto de investigación “MINERÍA DE DATOS EN DETERMINACIÓN DE PATRONES DE USO Y PERFILES DE USUARIOS”, desarrollado en la Facultad de Ciencias Exactas Físicas y Naturales – FCFN- de la Universidad Nacional de San Juan - UNSJ, aborda entre sus líneas de investigación, el descubrimiento de patrones que correspondan al perfil de un potencial

alumno desertor. En una primera instancia la matrícula de alumnos bajo estudio, pertenece a las carreras del Departamento de Informática. A partir de ello se pretende aportar conocimiento que permita tomar acciones preventivas por parte de las autoridades de la institución.

## Introducción

Esta propuesta se presenta, en parte, como una continuación de lo tratado en el proyecto 21/E824 **-Búsqueda de conocimiento estratégico en los datos de alumnos y biblioteca de la FCFN** y anteriores relacionados. En ellos se aplicaron diferentes instancias de minería de datos a datos de circulación de material bibliográfico, como así también a encuestas tomadas a usuarios alumnos y docentes, lo que permitió inferir la percepción de la imagen que los usuarios tienen de su biblioteca y a partir de ello elevar propuestas que favorecieron la toma de decisión por parte de la dirección de la biblioteca.

La Minería de Datos es *el proceso de identificar, desde los datos, conocimiento válido, potencialmente útil, y finalmente entendible para usarse en cruciales decisiones de negocios.* (G. Piatecki - Shapiro, [kdnuggets.com](http://kdnuggets.com). Synonym: Knowledge Discovery.). En este proceso se realiza una sucesión de tareas de descubrimiento de conocimiento en los datos, que permiten

arribar a un resultado exitoso y se detallan a continuación:

- **Limpieza de datos:** Elimina ruido y datos inconsistentes.
- **Integración de datos:** Múltiples fuentes de datos pueden ser combinados.
- **Selección de datos:** Sólo los datos pertinentes, a la tarea de análisis, se recuperan desde el reservorio de datos.
- **Transformación de datos:** Los datos se transforman o consolidan en formas apropiadas para la tarea de minería.
- **La minería de datos:** Proceso esencial en el que se aplican métodos inteligentes, con el fin de extraer patrones desde los datos.
- **Evaluación de Patrones:** Permite extraer conocimiento verdaderamente útil.
- **Presentación del conocimiento:** Formas de visualización y técnicas de representación del conocimiento, amigables y entendibles al usuario).

Existen numerosos estudios que señalan alto índice de deserción estudiantil de jóvenes universitarios a causa de problemas sociales, económicos, personales, etc. La no existencia de solución, o la no capacidad de los actores involucrados para afrontarlos, simplemente pueden conllevar a la renunciación del estudio por no poder sobrellevar la carga o la causa de sus problemas.

La Red Iberoamericana para la Acreditación de la Calidad de la Educación Superior (RIACES); menciona que la deserción es sinónimo de abandono. Esta Red define también el concepto de mortalidad escolar a aquella franja de estudiantes que suspenden, repiten, cambian de carrera, o abandonan antes de obtener el título. Para ello se abordarán diferentes aspectos vinculados con el análisis de resiliencia.

Otra instancia de reservorio de datos no menos importante, surge del acumulado en función de la acreditación de carreras a la que se encuentra sometido el Departamento Informática de la FCFN-UNSJ a instancias de la Comisión Nacional de Evaluación y Acreditación Universitaria-CONEAU. En el ciclo lectivo 2010 se generó una ingente cantidad de datos que ha permitido reconocer las etapas de las carreras en las que los alumnos comienzan a presentar dificultades en el normal cursado, generando estancamiento, cronicidad, y posible deserción.

De acuerdo a la literatura existente, la cual es muy amplia por los cientos de libros que hay, las miles de páginas en la Web e incomparables grupos y asociaciones, el concepto de resiliencia no es nuevo, aunque posee varias interpretaciones y aplicaciones. Por ejemplo, para las Ciencias Sociales, la resiliencia correspondería a *“la capacidad humana de hacer frente a las adversidades de la vida, superarlas y salir de ellas fortalecida o, incluso, transformada”* (Grotberg, 1996; en Bertrán, Noemí, Romero, 1998). Para la Ingeniería, la Física y la Metalurgia, la Resiliencia es *“La capacidad de ciertos materiales de recobrar su forma original luego de ser sometidos a algún proceso de presión deformadora”*, o *“La capacidad de los metales de resistir a los golpes y recuperar su estructura interna”*. Para la Osteología, es *“la capacidad del tejido óseo de crecer en el sentido correcto luego de una fractura”*. El término resiliencia tiene su origen en el idioma latín, resilio, que significa volver atrás, volver de un salto, resaltar, rebotar.

Algunos autores señalan que no se nace con resiliencia, sino que se desarrolla en la interacción del individuo con su medio; dicho de otra manera, la resiliencia surge de la interacción de los factores personales y sociales y se manifiesta de manera específica en cada

individuo (Feldman, 2001), a diferencia de la tendencia actual, que la reconoce como una cualidad innata.

Ana Cecilia Salgado Lévano, (Salgado, 2005), expone diversos métodos de medición aludiendo que la descripción de ellos fue realizada por Luthar & Cushing (1999), e instrumentos entre los que se destacan:

- Medición de adversidad
- Medición de adaptación positiva
- Medición del proceso de resiliencia
- Cuestionario de Perfil de vida del estudiante a través de las actitudes comportamentales, construido por el Search Instituto -1989- (Citado por Del Aguila, 2003).
- La Escala de Resiliencia para determinar el nivel de un individuo en tres facetas de la resiliencia Jew - 1992- (Citado por Bennett et al., 1998).
- La Escala de Resiliencia (ER), construida por Wagnild & Young (1993), con el propósito de identificar el grado de resiliencia individual.
- Medidas de resiliencia desarrolladas por Grotberg (1995).
- Evaluación de Resiliencia para estudiantes excepcionales (RAES), escala construida por Perry & Bard (2001).
- Inventario de Factores Protectores de Baruth (BPGF). Baruth & Carroll (2002).
- Escala de Resiliencia Adolescente. Oshio et al. (2002)

Muchos estudios se han centrado en analizar qué factores individuales podían promover la resiliencia. Fruto de ello, se han recogido algunas características individuales de las personas resilientes (Wolin y Wolin, 1993; Benson, 1997; Masten y Coasworth, 1998; Infante, 2004 y Suárez, 2004). En este sentido, los establecimientos educativos y todos los elementos que forman parte de la educación, pueden ser claves para

generar procesos de promoción de resiliencia.

El fenómeno de la deserción impacta negativamente en las instituciones de educación superior. La Secretaría de Políticas Universitarias del Ministerio de Educación de la Nación Argentina, reveló que el porcentaje de abandono de los estudiantes fue creciendo en los últimos años y que en 2007 alcanzó el 60%, de los cuales el 40% de los jóvenes abandonan en el primer año de sus estudios superiores.

Los orígenes de la deserción universitaria identificados son los siguientes:

Factores de carácter psico-educativo, son los que mayor peso explicativo o influencia tienen en el abandono o prolongación de los estudios superiores en estos momentos.

La percepción del ambiente universitario y la auto-eficacia académica tienen una gran influencia en la persistencia académica por parte de los estudiantes universitarios de primer año, durante el primer semestre de curso.

Otro factor que se da a través de la deserción de estudiantes universitarios es el problema del auto-concepto y esto genera en el alumno un bajo rendimiento escolar.

Según C. Rogers, el auto-concepto es un conjunto organizado y cambiante de percepciones que se refieren al sujeto. Como ejemplo de estas percepciones citamos: las características, atributos, cualidades y defectos, capacidades y límites, valores y relaciones que el sujeto reconoce como descriptivos de sí y que él percibe como datos de su identidad.

En resumen, la deserción se da básicamente por los antecedentes personales y familiares de los estudiantes; los factores institucionales de tipo académico, administrativo y social, así como factores económicos que pueden provocar este fenómeno.

El estudio de la resiliencia es trascendental hoy en día por las

implicancias que tiene tanto a nivel personal como social. Por ello urge concentrar esfuerzos en construir instrumentos que sean capaces de medir y evaluar con precisión, ya que sólo así podremos ser capaces de identificar las características y atributos resilientes de las personas, así como de evaluar la eficacia y efectividad de los programas de intervención (toma de decisiones) y el impacto que tienen en los estudiantes (de la comunidad universitaria).

Esto muestra la necesidad de abordar el tema de la resiliencia en los estudiantes del Departamento Informática de la FCFN, con la finalidad de poder encontrar los factores que determinen su adaptación y continuidad, o no, en el sistema universitario.

## Líneas de investigación y desarrollo

Se intentará encontrar el perfil del Alumno Desertor de las carreras del Departamento de Informática de la FCFN, con el propósito de aportar a la autoridad de información que le ayude a encarar acciones preventivas, y así disminuir la tasa de deserción existente en las carreras a tratar.

Las tareas de reconocimiento de perfiles de alumnos se realizarán, a partir de la aplicación de herramientas estadísticas y de minería de datos, a los datos obtenidos de encuestas a alumnos. Este proceso permitirá definir, modelar y validar la caracterización de un alumno desertor.

El tipo de diseño de la investigación propuesto es de campo. Para ello, la población a estudiar corresponde a:

- Datos de alumnos de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan.
- Datos obtenidos a partir de instrumentos de recolección de

información, desde un constructo validado.

## Resultados y Objetivos

En esta propuesta se pretende:

- Aplicar estrategias de minería de datos a datos contenidos en repositorios pertenecientes a alumnos de las carreras del Departamento Informática.
- Diseñar y validar constructos destinados a descubrir características que permitan modelar a alumnos desertores.
- Aplicar estrategias de minería de datos, a datos obtenidos a partir de diferentes constructos.

La aplicación correspondiente a la recopilación de datos desde el constructo validado se observa en la figura 1.

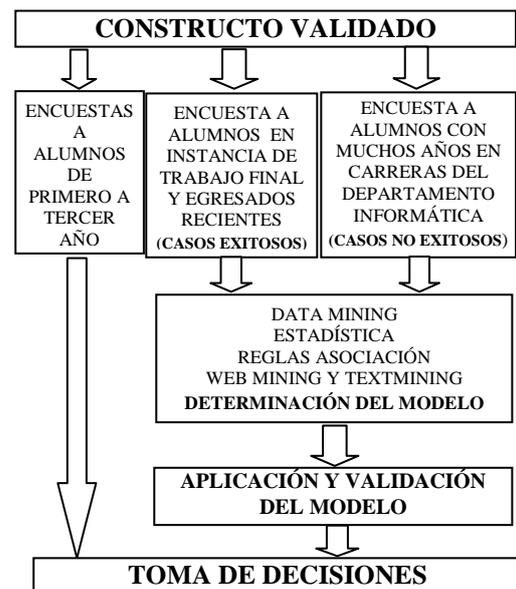


Figura 1: Esquema para la determinación del modelo de resiliencia.

## Formación de Recursos Humanos

En el proyecto de investigación “MINERÍA DE DATOS EN DETERMINACIÓN DE PATRONES DE USO Y PERFILES DE USUARIOS” participan ocho docentes

del Departamento de Informática, los que a su vez pertenecen a asignaturas de las áreas Teoría de la Computación, Ciencias Básicas, Algoritmos y Lenguajes, y Arquitectura, Sistemas Operativos y Redes. La ejecución de las tareas proyectadas incidirá directamente en una formación más profunda de los integrantes del equipo de investigación, en las tecnologías de Data Mining, Bibliomining, Textmining, y en herramientas estadísticas aplicadas a las tecnologías de búsqueda automática de conocimiento. Este aspecto beneficiará de manera directa a las carreras del Departamento de Informática, pues las temáticas abordadas están vinculadas de manera directa con las materias en las cuales se desempeñan los integrantes de este proyecto.

Se prevé la generación de Trabajos Finales de Grado para las carreras Licenciatura en Sistemas de Información y Licenciatura en Ciencias de la Computación. Así mismo, y desde la continuidad de proyectos anteriores, se están dirigiendo tres tesinas de grado en temáticas afines a las abordadas en la presente propuesta.

Recientemente se ha creado en dependencias de la FCFN la Maestría en Informática, que tiene entre sus áreas a Inteligencia de Negocios, aplicación sustantiva de esta temática.

Asimismo, se presentarán solicitudes de Becas, de la UNSJ, correspondientes a alumnos avanzados.

## Referencias

- Han, J; Kamber M. *Data Mining: Concepts y Techniques*. Second Editions. Morgan Kaufmann Publisher. 2006.
- Hernández Orallo J., Ramírez Quintana, J, Ferri Ramirez, C. (2008) *Introducción a la Minería de Datos*. Pearson-Prentice Hall.
- Infante, F. *La resiliencia como proceso: una revisión de la literatura reciente*. Melillo, A. & Suárez, N. (Comp.) Resiliencia Descubriendo las propias fortalezas. Buenos Aires, Argentina: Paidós. 2002
- Larose D. *Data Mining. Methods and Models*. Department of Mathematical Sciences Central Connecticut State University Wiley. A John Wiley & Sons, Inc Publication. 2006.
- Manciaux, M. *La resiliencia: Resistir y rehacerse*. Editorial Gedisa Barcelona. 2005.
- Myatt, G.J. *Making Sense of Data. A Practical Guide to Exploratory Data Analysis and Data Mining*. Wiley-Interscience a John Wiley & Sons, Inc., Publication. 2007
- Quintero, Angela María. *La resiliencia: un reto para el trabajo social*. Madrid. 2001.
- Salgado Lévano , Ana Cecilia. 2005. *Métodos e instrumentos para medir la resiliencia: una alternativa peruana*. Revista de Psicología, número 011 pp. 41-48 Liberabit Universidad de San Martín de Porres Lima –Perú. 2005. ISSN (V.I.): 1729-4827.
- <http://www.resiliencycenter.com/bkstore/ResAdv-chap1.shtml> *The Resiliency Advantage* by Al Siebert, PhD Available now at [Amazon.com](http://Amazon.com) or through co-publisher [Practical Psychology Press](http://PracticalPsychologyPress.com). ISBN-13: 978-1-57675-329-3

## Páginas de Internet.

- Atehortúa M. Miguel. *Resiliencia: Otra perspectiva de las experiencias sociales y personales dentro de la empresa* (2002) Obtenido el 6 de noviembre de 2010. En: [http://www.scielo.unal.edu.co/scielo.php?pid=S0123-59232002000100002&script=sci\\_arttext](http://www.scielo.unal.edu.co/scielo.php?pid=S0123-59232002000100002&script=sci_arttext)
- Quinteros Velásquez, Ángela. *Resiliencia: Contexto no clínico para trabajo social* (2005) Obtenido el 6 de noviembre de 2010. En: <http://redalyc.uaemex.mx/pdf/773/77330103.pdf>