

Reconociendo factores resilientes en alumnos de informática, mediante la aplicación de TIC

Alejandra Malberti¹, Graciela Beguerí¹, Raúl Klenzi¹

¹ Departamento e Instituto de Informática-Facultad de Ciencias Exactas, Físicas y Naturales-Universidad Nacional de San Juan, San Juan, Argentina

Resumen

Una investigación preliminar, desarrollada en la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan (FCEFN-UNSJ), reveló que factores económicos, sociales, culturales y demográficos no fueron determinantes ni causantes de rezago y deserción en el grupo de estudiantes del primer año de las carreras de la mencionada institución.

La temática sobre rezago y deserción tiene múltiples planteamientos en la literatura. Considerando que la mayoría estudia las causas, confecciona modelos pero las tasas de deserción y rezago no disminuyen, es que se plantea en una nueva etapa reflexionar sobre cómo modelar el perfil resiliente del alumno del Departamento de Informática (DI-FCEFN), mediante técnicas de análisis multivariante y minería de datos.

Esta investigación está basada, principalmente, en el instrumento "Escala de Resiliencia", constructo propuesto por Wagnild y Young, 1993.

Los factores resilientes que describen el perfil resiliente de los alumnos "académicamente exitosos" fueron encontrados mediante el uso de TIC, a partir de los valores recabados en encuestas. Los resultados obtenidos, tanto del análisis de clasificación como del análisis multivariado permitieron reconocer que los factores: autoconcepto, confianza en sí mismo, independencia, satisfacción personal y ecuanimidad son determinantes al momento de evaluar la resiliencia.

Palabras clave: Deserción, Rezago, Resiliencia, Minería de Datos, Análisis multivariante, Estudiantes universitarios.

Abstract

A preliminary research developed at the Faculty of Exact, Physical and Natural Science at the National University of San Juan (FCEFN-UNSJ, for its acronym in Spanish) revealed that economic, social, cultural and

demographic factors were neither determining nor causing dropout or attrition among the students attending the first year classes at careers of the above mentioned institution.

The thematic on dropout and attrition has multiple approaches in the literature. Considering that most of literature studies factors, and models are being built but dropout and attrition rates do not decrease, a new effort on how to model the resilient profile of the student of the Department of Informatics (DI-FCEFN) has to be made, through techniques of multivariate analysis and data mining.

This research is mainly based on the instrument "Resilience Scale", a construct proposed by Wagnild & Young, 1993.

The resiliency factors that describe the resilient profile of students considered "academically successful" were found through the use of TICS, based on the values collected by means of surveys.

The results obtained both with classification and multivariate analysis, made it possible to realize that the factors: self, independence, personal satisfaction, equanimity and self-confidence are determining when evaluating resilience.

Key words: Dropout, Lag, Resilience, Datamining, Multivariate analysis, University students.

1. Introducción

La reciente instancia de acreditación a la que se vieron sometidas las diferentes titulaciones del área informática del país motivó, en el ámbito de la FCEFN-UNSJ, al análisis de factores que inciden en una problemática común de las distintas carreras, como son el rezago, la cronicidad y la deserción de estudiantes universitarios. 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. Ya en 2007 alcanzó el 60%, de los cuales

el 40 % de los jóvenes abandonan en el primer año de sus estudios superiores, cifras que parecen no cambiar. Con el propósito de conocer las características ecológicas de los estudiantes universitarios del primer año de las carreras que se dictan en el ámbito de la FCEF-UNSA, se analizaron los resultados obtenidos de una encuesta sociodemográfica realizada en la mencionada facultad. Los resultados alcanzados mostraron disidencia con los presentados en estudios previos, que han coincidido en señalar que factores socioeconómicos inciden en la deserción de estudiantes universitarios [1,2].

Del análisis realizado se concluyó que los factores encontrados no son adecuados o no son suficientes para evaluar la sintomatología de la emergente pandemia en la educación, como son la deserción y rezago estudiantil. Esto motivó incursionar en el análisis de factores que describen la capacidad de resiliencia. El estudio de la Resiliencia ha sido abordado por numerosos autores y diferentes organizaciones a nivel nacional e internacional [3,4,5,6,7,8,9].

Esta línea de investigación, actualmente, se enmarca en el proyecto "MINERÍA DE DATOS EN LA DETERMINACIÓN DE PATRONES DE USO Y PERFILES DE USUARIO" código 21/E889 que se desarrolla en el ámbito de la FCEF-UNSA, aprobado por el Consejo de Investigaciones Científicas Técnicas y de Creación Artística (CICITCA), financiado por la propia universidad y ajustado a evaluación externa.

Entre los objetivos del mencionado proyecto, se persigue contribuir en la implementación de programas tendientes a aumentar la resiliencia de los estudiantes en riesgo de deserción y a su vez se intenta aportar conocimiento tendiente a potenciar las cualidades resilientes descubiertas, en los alumnos académicamente exitosos, a los alumnos ingresantes. Según Jairo Villasmil Ferrer [10], en su tesis doctoral, para promocionar rasgos resilientes en entornos educativos se debería poner énfasis en las fortalezas y cualidades personales, familiares, escolares y comunitarias.

1.1 Minería de Datos

La Minería de Datos, también conocida como Descubrimiento de Conocimiento en Bases de Datos – Knowledge Discovery in Databases: KDD-, significa un proceso de extracción no trivial, de información implícita, útil, previamente desconocida, desde los datos de una base de datos [11,12]. De acuerdo con el Gartner Group [13] "La minería de datos es el proceso de descubrir nuevas correlaciones significativas, patrones y tendencias, tamizando grandes cantidades de datos almacenados en depósitos, usando tecnologías de reconocimiento de patrones, así como técnicas estadísticas y matemáticas". La minería de datos no es una aplicación aleatoria de métodos estadísticos o de

aprendizaje de máquina, sino que es un proceso cuidadosamente planificado.

Dentro de la minería de datos se pueden distinguir una serie de tareas, y cada una de ellas refiere a un tipo de problema a ser resuelto por un algoritmo de minería de datos [14]. Estas tareas pueden ser predictivas o descriptivas. Las primeras pretenden estimar valores futuros o desconocidos de variables de interés, mientras que las descriptivas identifican patrones que explican o resumen los datos.

La clasificación es posiblemente la tarea predictiva más usada. En ella cada instancia o registro de la base de datos pertenece a una clase, la cual se indica por medio del valor de uno de los atributos de instancia. El resto de los atributos de la instancia se usan para predecir la clase. El objetivo de la clasificación es predecir la clase de nuevas instancias. Para esta tarea existen una serie de técnicas, entre las que se encuentran las Redes Neuronales, Naïve Bayes, métodos basados en árboles de decisión y métodos basados en reglas.

El agrupamiento, también llamado segmentación, es la principal tarea descriptiva. En esta tarea los datos son agrupados basándose en el principio de maximizar la similitud entre los elementos de un grupo minimizando la similitud entre los distintos grupos; los elementos de un grupo son similares entre sí y muy diferentes a los elementos de otros grupos.

Otra tarea descriptiva es el descubrimiento de reglas de asociación. Esta tarea tiene por objetivo identificar relaciones no explícitas entre atributos. Un caso particular son las reglas de asociación secuenciales, que tratan de descubrir patrones secuenciales en los datos. Estas reglas se diferencian de las reglas de asociación en que las relaciones entre los datos se basan en el tiempo.

Algoritmos

Cuando se plantea el análisis de datos por medio de minería de datos, lo que inicialmente se plantea en este proceso es la elección de la tarea a realizar para resolver un problema específico. Luego, cada una de las tareas requiere métodos, técnicas o algoritmos para resolverla. Ahora bien, una tarea puede tener muchos métodos diferentes para resolverlas, y puede ocurrir que el mismo método o el mismo tipo de técnica pueda resolver diversas tareas. En este sentido [14] sintetiza la relación entre tareas y técnicas o algoritmos.

Algunas tareas pueden ser resueltas por diversas técnicas, y algunas técnicas pueden aplicarse a varias tareas. Por esta razón [14] sostiene que es importante conocer las capacidades de cada técnica, los ámbitos donde suele funcionar mejor, su eficiencia, robustez, etc., en resumen conocer las características funcionales de cada una. En particular, y atendiendo al alcance del presente trabajo, la clasificación puede ser tratada por técnicas basadas en árboles de decisión. Estas técnicas construyen un modelo de representación a partir de la

regularidad que se presenta en los datos. Un árbol de decisión es un conjunto de condiciones excluyentes organizada en una estructura jerárquica. Los algoritmos ID3 y C4.5, entre otros, construyen árboles de decisión, y pertenecen al tipo divide y vencerás. C4.5, es uno de los algoritmos de minería de datos más utilizado pues permite obtener árboles de decisión bastante precisos y comprensibles.

1.2 Análisis multivariante

La finalidad del análisis multivariante es analizar simultáneamente conjuntos de datos correspondientes a varias variables medidas en cada individuo u objeto estudiado.

A diferencia de otros métodos, su aplicación conlleva a un mejor entendimiento del fenómeno objeto de estudio.

Cualquier análisis simultáneo de más de dos variables puede ser considerado aproximadamente como un análisis multivariante. Muchas técnicas multivariantes son extensiones del análisis univariante y del análisis bivariante. Otras técnicas multivariantes, sin embargo, están diseñadas exclusivamente para tratar con problemas multivariantes. Tal es el caso del análisis factorial, que sirve para identificar la estructura subyacente de un conjunto de variables.

Para algunos investigadores, multivariante significa simplemente examinar relaciones entre más de dos variables. Otros usan el término sólo para problemas en los que se supone que todas las variables múltiples tienen una distribución normal multivariante. Sin embargo, para ser considerado verdaderamente multivariante, todas las variables deben ser aleatorias y estar interrelacionadas de tal forma que sus diferentes efectos no puedan ser interpretados, con algún sentido, separadamente.

Análisis factorial

El análisis factorial es un tipo de análisis de interdependencia. Se utiliza para analizar las interrelaciones existentes entre una serie de variables. Los antecedentes del análisis factorial se encuentran en las técnicas de regresión lineal, iniciadas por Galton. Pero es K. Pearson (1901) quien presentó la primer propuesta del “método de componentes principales”; primer paso para el cálculo del análisis factorial.

El origen del análisis factorial suele atribuirse a Spearman (1904), por su investigación sobre inteligencia [15,16].

La falta de concordancia entre consideraciones teóricas expuestas por diversos autores, junto con la variedad de las bases matemáticas, son a menudo causa de diversidad en las denominaciones.

El análisis factorial, orientado a la deducción de capacidades y cualidades generales, se basa en la correlación de características del grupo de personas investigadas.

Los factores que se obtienen se fundamentan en la semejanza “funcional” entre personas y constituyen las tipologías.

Análisis de las componentes principales

El Análisis Factorial y el Análisis de Componentes Principales están muy relacionados. Algunos autores consideran el segundo como una etapa del primero y otros los consideran como técnicas diferentes.

Su objetivo es reducir la dimensión de un conjunto de variables a otro conjunto con un menor número, perdiendo la menor cantidad de información posible, con el fin de mejorar la interpretación o explicación de los datos.

Los nuevos componentes principales o factores son una combinación lineal de las variables originales y además son independientes entre sí.

Un aspecto clave en Análisis de Componentes Principales es la interpretación de los factores, ya que ésta debe ser deducida luego de observar la relación de los factores con las variables iniciales. Tal es el caso en el presente estudio.

1.3 Deserción y Rezago

La Deserción, entendida como abandono de los estudios superiores, es tratada por diferentes autores.

Romo López [17] reconoce la existencia de modelos y teorías que buscan explicar el fenómeno de la deserción y rezago. En su trabajo recopila diferentes teorías, que son clasificadas en cinco tipos: psicológico, social, económico, organizacional e interaccional.

Por su parte, Erika Himmel [18], también presenta una síntesis de diversos enfoques teóricos que destacan factores psicológicos, económicos, sociológicos, organizacionales o aspectos de las interacciones entre el estudiante y la institución, que han sido empleados en investigaciones sobre deserción durante los últimos 25 años. A partir de ello la autora sostiene que los trabajos sobre deserción han evolucionado, abarcando variables pertenecientes a diferentes ámbitos, susceptibles de ser modificadas por las instituciones para reducir la deserción.

Adicionalmente en [19] se listan una serie de factores, señalados por distintos autores como atribuibles a la deserción, entre los que se encuentran:

- “● Las condiciones económicas desfavorables del estudiante.
- El deficiente nivel cultural de la familia a la que pertenece.
- La incompatibilidad del tiempo dedicado al trabajo y a los estudios.
- La responsabilidad que implica el matrimonio.
- Las características académicas previas del estudiante, como los bajos promedios obtenidos en la educación media superior que

reflejan la insuficiencia de los conocimientos y habilidades con que egresan los estudiantes, en relación con los requeridos para mantener las exigencias académicas del nivel superior.

- La deficiente orientación vocacional recibida antes de ingresar a la licenciatura, que provoca que los alumnos se inscriban en las carreras profesionales sin sustentar su decisión en una sólida información sobre la misma”.

En el “Estudio sobre Causas de la Deserción Universitaria”, realizado por el Centro de Microdatos del Departamento de Economía de la Universidad de Chile [20] a solicitud del Ministerio de Educación de ese país, se afirma “*las tres causas más determinantes en la deserción de estudiantes en primer año universitario son: problemas vocacionales, situación económica de sus familias, y rendimiento académico.*” Por su parte [2], afirma que “los estudiantes de sexo masculino, la vinculación de los estudiantes al mercado laboral y los estudiantes provenientes de otras regiones, tienen mayor riesgo de deserción. Además, la edad del estudiante incrementa el riesgo, sin embargo su efecto decrece marginalmente al aumentar la edad”. Adicionalmente, [1] considera que “el factor relacionado con la educación de los padres y el tipo de colegio secundario al que asistió el alumno son importantes en la explicación de las diferencias de riesgo de deserción y probabilidades condicionales de graduación... El género también juega un papel significativo... Las demás características demográficas y personales (residir en Rosario, vivir con la familia, ser soltero y tener más edad al iniciar la carrera) resultaron también significativas afectando positivamente la probabilidad condicional de abandonar...”

1.4 Resiliencia

En estudios sobre deserción de estudiantes universitarios, aparece otro factor que genera en el alumno un bajo rendimiento, es el problema del auto-concepto. Según Carl Rogers [21], el autoconcepto 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. Por su parte Arancibia, Maltes y Álvarez [10] identificaron al autoconcepto académico como una variable estrechamente ligada al rendimiento.

En esta línea de investigación, diferentes organizaciones, instituciones, autores, entre otros han abordado el estudio de la Resiliencia [6,7,8,9].

El término resiliencia tiene su origen en el idioma latín, y fue adaptado a las ciencias sociales para caracterizar aquellas personas que, a pesar de nacer y vivir en

situaciones de alto riesgo, se desarrollan psicológicamente sanos y exitosos. Algunas de las definiciones que, desde este campo, han desarrollado diversos autores en torno a este concepto son [4]:

“→ Habilidad para surgir de la adversidad, adaptarse, recuperarse y acceder a una vida significativa y productiva (ICCB, 1994).

→ Historia de adaptaciones exitosas en el individuo que se ha visto expuesto a factores biológicos de riesgo o eventos de vida estresantes; además, implica la expectativa de continuar con una baja susceptibilidad a futuros estresores (Luthar y Zingler, 1991; Masten y Garmezy, 1985; Werner y Smith, 1982 en Werner y Smith, 1992).

→ Enfrentamiento efectivo ante eventos y circunstancias de la vida severamente estresantes y acumulativos (Lösel, Blieneser y Köferl en Brambling et al., 1989).

→ Capacidad humana universal para hacer frente a las adversidades de la vida, superarlas o incluso ser transformado por ellas. La resiliencia es parte del proceso evolutivo y debe ser promovido desde la niñez (Grotberg, 1995).

→ La resiliencia distingue dos componentes: la resistencia frente a la destrucción, esto es, la capacidad de proteger la propia integridad bajo presión; por otra parte, más allá de la resistencia, la capacidad para construir un conductismo vital positivo pese a circunstancias difíciles (Vanistendael, 1994). Según este autor, el concepto incluye además, la capacidad de una persona o sistema social de enfrentar adecuadamente las dificultades, de una forma socialmente aceptable.

→ La resiliencia se ha caracterizado como un conjunto de procesos sociales e intrapsíquicos que posibilitan tener una vida sana, viviendo en un medio insano. Estos procesos tendrían lugar a través del tiempo, dando afortunadas combinaciones entre atributos del niño y su ambiente familiar, social y cultural. De este modo, la resiliencia no puede ser pensada como un atributo con que los niños nacen, ni que los niños adquieren durante su desarrollo, sino que se trataría de un proceso interactivo (Rutter, 1992) entre éstos y su medio.

→ La resiliencia habla de una combinación de factores que permiten a un niño, a un ser humano, afrontar y superar los problemas y adversidades de la vida (Suárez, 1995).

→ Concepto genérico que se refiere a una amplia gama de factores de riesgo y los resultados de competencia. Puede ser producto de una conjunción entre los factores ambientales, como el temperamento y un tipo de habilidad cognitiva que tienen los niños cuando son muy pequeños (Osborn, 1993).

En el caso presentado en la figura 2 se observa el entorno de RM, el cual permite la construcción modular de la aplicación. Allí se aprecia: la lectura de datos desde una planilla de cálculo, el filtrado de datos o registros cuyos atributos tienen determinado valor (filas de la planilla de entrada), la selección y/o filtrado de atributos completos (columnas de la planilla de entrada), la definición del rol que cada atributo tiene en el flujo de datos y, finalmente, el algoritmo de clasificación utilizado en esta oportunidad, con la posibilidad de particularizarlo a la aplicación desde el ajuste de los parámetros asociados al algoritmo. El entorno, a su vez, permite la visualización de diferentes aspectos de la aplicación y que son seleccionados por el propio usuario. En este caso se observa que aparecen sobre el margen izquierdo, los repositorios de datos en “**Repositorios**” y los diferentes operadores seleccionables por el usuario en “**Operators**”. En la parte inferior central se presenta el consumo de memoria en la pestaña “**Monitor System**” y los errores que puede haber producido la última ejecución de la aplicación en “**Problems**”. En el extremo derecho de la interfaz, en “**Parameters**” se aprecian los parámetros asociados a cada uno de los módulos, y con ello en “**Process**”, una explicación o ayuda de la tarea que realiza el módulo seleccionado.

Existen operadores para transformación de datos, así como para realizar las principales tareas de minería de datos, esto es: operadores de clasificación, de segmentación, de descubrimiento de reglas de asociación entre otros. En situaciones de análisis más complejo, pueden construirse super-operadores, los cuales encierran subprocesos completos [24,25].

Lenguaje R

R es un lenguaje de programación y un entorno para el análisis estadístico. Debido a su naturaleza, es fácilmente adaptable a una gran variedad de tareas. Fue inicialmente escrito por Robert Gentleman y Ross Ihaka del Departamento de Estadística de la Universidad de Auckland - Nueva Zelanda.

Actualmente R es uno de los lenguajes de mayor difusión en la comunidad científica y es fruto de la colaboración de personas del todo el mundo.

El código de R está disponible como software libre bajo las condiciones de la Licencia Pública General (General Public License GNU-GPL). Además está disponible precompilado para una multitud de plataformas. La página principal del proyecto es [26]. Posee varias interfaces gráficas muy amigables para el usuario, entre las que cabe destacar a R-commander (Rcmdr).

Rcmdr simplifica la forma de interactuar con R y pretende, en primera instancia, facilitar el manejo de R y, en segundo lugar, servir como generador de instrucciones R.

DYANE

DYANE (Diseño y Análisis de Encuestas) es un programa informático, realizado en entorno Windows, para diseñar y realizar encuestas y estudios de investigación social y de mercados.

El programa está estructurado de forma modular, con menús que facilitan el acceso a las distintas funciones y técnicas que incluye. Con DYANE 2.0 se puede:

- Diseñar cuestionarios en formato tradicional o electrónico, en formularios MS Word, con casillas de marcate y campos de escritura.
- Grabar los datos obtenidos en la encuesta, bien con plantillas específicas diseñadas por el programa, bien mediante lectura electrónica de formularios MS Word.
- Aplicar una gran variedad de técnicas estadísticas de análisis de datos y generar gráficos, con el fin de mejorar la presentación de los resultados.

Todos los resultados y gráficos de los análisis se pueden imprimir, enviar a formularios MS Word, exportar a otros programas o guardar en archivos, tal como lo expresa su autor Miguel Santesmases Mestre, en su sitio web[27].

2. Desarrollo

El proceso de investigación realizado hasta el momento contiene dos fases. La primera, motivada por la presunción de que factores socioeconómicos inciden en la deserción universitaria, consistió en el análisis de las características ecológicas de los alumnos de primer año de las carreras que se dictan en la FCFN-UNSJ. Según la Real Academia Española, se entiende a la ecología como parte de la Sociología que estudia la relación entre los grupos humanos y su ambiente, tanto físico como social.

Los resultados alcanzados en la primera fase promovieron un nuevo camino a seguir. En este caso el proceso se centró en descubrir los factores que describen la capacidad de resiliencia de los alumnos de las carreras del DI.

2.1 Fase 1- Características ecológicas

Por iniciativa de las autoridades de la FCFN- UNSJ, se confeccionó una encuesta destinada a alumnos de primer año de las carreras Licenciatura en Geología, Licenciatura en Biología, Licenciatura en Astronomía, Licenciatura en Geofísica, Licenciatura en Sistemas de Información y Licenciatura en Ciencias de la Computación. La mencionada encuesta recabó, aparte de datos generales, información relativa a factores económicos, académicos y socioculturales, tales como “Fuente de Ingreso” “Ingreso Mensual”; “Estudios

alcanzados” así como diferentes motivaciones que han influido en la elección de la carrera.

Los estudiantes universitarios de la facultad, conformaron la población y el grupo objeto de estudio estuvo constituido por 289 alumnos, quienes respondieron la encuesta y cursaban al menos alguna asignatura de primer año de las carreras citadas, durante el ciclo lectivo 2009.

La encuesta fue implementada en DYANE 2.0 para facilitar la carga de datos y su posterior análisis.

Los métodos de análisis y tratamiento de la información fueron los típicos dentro de la estadística inferencial y descriptiva.

Resultados alcanzados

De los desertores, pertenecientes al grupo de estudiantes tratado,

- Más del 95% de los estudiantes eran solteros, no tienen personas a cargo y su fuente de ingreso provenía de su familia.
- Más del 80% no trabajaba y además habitaban con quienes poseen un vínculo de primer grado.
- Más del 70% eligieron la carrera por vocación o expectativas laborales, vivían en departamentos colindantes o en el que se sitúa la facultad y residían en viviendas propias.
- En cuanto a la formación académica alcanzada por las personas que les proporcionaban recursos económicos, todos habían finalizado algún tipo de estudio.

Se puede observar que los factores económicos, sociales, culturales y demográficos no han sido determinantes ni causantes del rezago y deserción. Estos resultados mostraron disidencia con los obtenidos en otros estudios realizados, relacionados a esta problemática, los cuales consideran que los factores económicos y culturales, entre otros, son los responsables del rezago y la deserción universitaria, especialmente en los primeros años.

2.2 Fase 2- Factores resilientes

Motivados por los resultados alcanzados en la fase previa, se inició un nuevo camino. La hipótesis inicial en esta parte de la investigación, en un principio, era corroborar que la CR estaba directamente relacionada con el buen rendimiento académico de los estudiantes universitarios. Con este objetivo se indagó y adaptó una herramienta por medio de la cual determinar qué cualidades conforman el perfil de los estudiantes “académicamente exitosos” y potenciarlas en los alumnos del primer año.

Se optó por el constructo de Wagnild y Young por ser éste válido y fiable, además de su brevedad y pertinencia para una población de jóvenes. Algunos ítems, variables, fueron acondicionados debido a las particularidades de los estudiantes.

Esta experiencia se inició en el año 2011, en el ámbito de la FCEF N con los alumnos ingresantes y con una muestra de alumnos avanzados que se encontraban cursando al menos una asignatura del tercer, cuarto o quinto año de las carreras del DI.

En este trabajo se considera que todos los alumnos avanzados son académicamente exitosos, por ser alumnos de las carreras, Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información, las cuales cuentan con títulos intermedios a nivel de Tecnicaturas al finalizar tercer año.

La encuesta final incluyó variables de clasificación de los alumnos, tales como carrera, edad, sexo, año de ingreso a la carrera y preguntas tales como: ¿En el último año ha tenido problemas desagradables? ¿Has pensado que la carrera elegida no fue la correcta? Esta última solo para los alumnos avanzados.

Luego del relevamiento de los datos de la encuesta se analizaron los niveles de resiliencia, tanto en los alumnos avanzados como en los del primer año, para luego ser clasificados en 3 categorías: Mayor – Moderada – Escasa.

Las tablas 1 y 2 muestran una síntesis descriptiva de algunas de las variables tratadas en los alumnos avanzados.

Capacidad de Resiliencia	Frecuencia		Total (n=41)
	Masculino	Femenino	
Mayor	9	3	12
Moderada	17	4	21
Escasa	7	1	8

Tabla 1 - Número de alumnos del DI según CR y sexo.

¿En el último año ha tenido problemas desagradables?	Frecuencia		Total (n=41)
	Masculino	Femenino	
Nunca	14	4	18
A veces	18	4	22
Frecuentemente	1	0	1
Siempre	0	0	0
	-	-	
	-	-	
	-	-	
¿Has pensado que la carrera elegida no fue la correcta?	Frecuencia		Total (n=41)
	Masculino	Femenino	
Nunca	22	7	29
A veces	9	1	10
Frecuentemente	1	0	1
Siempre	0	0	0
No contesta	1	0	1

Tabla 2 - Número de alumnos del DI según sexo y tipo de respuestas.

Con el fin de analizar las posibles interrelaciones entre las variables y así identificar los diversos factores que subyacen a la CR, se efectuó un análisis factorial exploratorio.

Los índices de ajuste, prueba de Bartlett significativa e índice KMO (Kaiser-Meyer-Olkin) superior a 0.8, evidencian la aptitud de los factores encontrados y por medio del estudio realizado, se comprueba una estructura compuesta por cinco factores. Tabla 5.

Cabe mencionar que los cinco factores obtenidos permiten explicar un 76,059% de la varianza contenida en las variables iniciales. Los factores extraídos, podrían ser interpretados como: autoconcepto, independencia, satisfacción personal, ecuanimidad y confianza en sí mismo.

A partir de los datos de CR obtenidos desde la encuesta, rotulados como Escasa - Moderada - Mayor se llevó adelante, por medio de la utilización del software de minería de datos RapidMiner (versión 5.3.008), un algoritmo de clasificación que pone, como variable a clasificar al atributo CR obtenido desde el constructo elaborado por Wagnild-Young (1993), adaptado por Beguerí-Malberti [28].

En esta propuesta se utiliza el algoritmo de clasificación en árboles W-J48, implementado en WEKA y ejecutado desde la interface de RM. El algoritmo W-J48 de WEKA es una implementación del algoritmo C4.5. La particularidad de W-J48, que en este caso será utilizado como descriptor de los datos y que permite atributos con valores numéricos o valores desconocidos o faltantes, representando en esta propuesta a respuestas no asignadas por los encuestados, motivo su elección.

En la primer fila de la tabla 3, se observan los atributos inductores resultantes de la aplicación del mencionado algoritmo, sin poda y admitiendo como mínimo un registro por hoja, a todas las encuestas relevadas de las carreras del DI y en la segunda fila, el árbol propiamente dicho.

Las variables que describen la encuesta desde la aplicación del W-J48 (inductores) fueron posteriormente procesadas mediante un algoritmo del tipo One-Rule, con el objetivo de describir con el menor error posible y desde un único atributo la mayor cantidad de las respuestas asociadas a CR, y sus resultados son mostrados en la tercera fila.

	TOTALES AVANZADOS	TOTALES INGRESANTES
Atributos Inductores	<p>1 Planes 23 estar 3 depend 2 arreglo 8 Me sient 18 pto vis 17 confian 7 preocupa 20 vida 5 molesto (10)</p>	<p>2 arreglo 4 interesI 5 molesto 20 vida 17 confian 15 entretan 16 atravasa 14 interes (8)</p>
Arbol de Clasificación W-J48	<p>W-J48 unpruned tree 1 Planes <= 3 23 estar <= 2 3 depend P <= 2: Escasa (4.0) 3 depend P > 2 2 arreglo <= 3 8 Me sient <= 3 18 pto vis <= 2: Moderada (2.0) 18 pto vis > 2 17 confian <= 2: Moderada (1.0) 17 confian > 2 7 preocupa <= 1: Moderada (1.0) 7 preocupa > 1: Escasa (3.0) 8 Me sient > 3: Moderada (2.0) 2 arreglo > 3: Moderada (12.0) 23 estar > 2 20 vida <= 3: Escasa (1.0) 20 vida > 3 1 Planes <= 2: Moderada (1.0) 1 Planes > 2 8 Me sient <= 3: Moderada (1.0) 8 Me sient > 3 5 molesto <= 1: Moderada (1.0) 5 molesto > 1: Mayor (6.0) 1 Planes > 3: Mayor (6.0) Number of Leaves : 13 Size of the tree : 25</p>	<p>W-J48 unpruned tree 2 arreglo <= 2: Escasa (4.0) 2 arreglo > 2 4 interesI <= 3 5 molesto <= 2 20 vida <= 3: Escasa (1.0) 20 vida > 3: Moderada (5.0) 5 molesto > 2: Escasa (6.0) 4 interesI > 3 17 confian <= 1: Escasa (1.0) 17 confian > 1 15 entretan <= 3: Moderada (13.0) 15 entretan > 3 16 atravasa <= 3: Moderada (2.0) 16 atravasa > 3 14 interes <= 3 17 confian <= 3: Moderada (2.0) 17 confian > 3: Mayor (2.0) 14 interes > 3: Mayor (3.0) Number of Leaves : 10 Size of the tree : 19</p>
Rule Model Single Rule Induction (Single Attribute)	<p>if 1 Planes <= 3.5 then Moderada (21/6/8) else Mayor (0/6/0) correct: 27 out of 41 training examples.</p>	<p>if 4 interesI <= 3.5 then Escasa (0/10/5) else Moderada (5/2/17) correct: 27 out of 39 training examples.</p>

Tabla 3 - Algoritmos de Clasificación en datos de alumnos avanzados e ingresantes de carreras del DI-FCEFN.

Entre los atributos inductores, que describen en cada caso la mayoría de los registros de la encuesta, se observan cuatro atributos coincidentes (atributos subrayados en la tabla 3). A su vez y en el análisis del árbol propiamente dicho, para el resultado de la etiqueta CR= Escasa, tanto para alumnos avanzados como para ingresantes, aparecen los atributos 2 Arreglo, 17 Confian y 20 Vida. Estos atributos, deberían ser los principales a considerar y tratar al momento de abordar la problemática de abandono o deserción.

<u>2</u>	<u>Arreglo</u>	Cuando tengo problemas me las arreglo de alguna forma
<u>17</u>	<u>Confian</u>	Me parece que las personas confían en mí
<u>20</u>	<u>Vida</u>	Mi vida tiene sentido
<u>1</u>	<u>Planes</u>	Cuando hago planes los realizo
<u>4</u>	<u>InteresI</u>	Mantengo el interés en las cosas que considero importante

Tabla 4 - Etiqueta de atributos con sus correspondiente significados en la escala de resiliencia.

La respuesta en cada grupo se observa en la tabla 3 y allí se aprecia que para el grupo de alumnos avanzados el atributo **1 Planes** es el que mejor describe los datos con 27 registros correctos sobre los 41 que conforman la muestra (27/41, 65,85%). Cuando el valor de este atributo **1 Planes** es menor que 3,5 el modelo concluye CR= **Moderada** en 21 de 35 casos (21/(21+6+8), 60%) y caso contrario CR=**Mayor** en los restantes 6 casos (100%). Como se asume que los alumnos avanzados son los casos “académicamente exitosos”, debería ser **1 Planes**, el atributo a analizar en los alumnos ingresantes.

Simultáneamente, y desde la aplicación del algoritmo Single Rule Induction (Single Attribute) - algoritmo que aprende una sola regla con un único atributo- a la muestra de alumnos ingresantes, se observa que el atributo que mejor describe la CR de ese conjunto de datos, es **4 interésI**, que describe correctamente 27 de los 39 registros de la muestra (27/39, 69,23%). Cuando el atributo **4 interésI** ha sido respondido con un valor menor o igual a 3,5 concluye CR= **Escasa**, en 10 de 15 casos (10/(10+5), 66,66%), caso contrario describe a CR= **Moderada**, en 17 de 24 casos (17/(5+2+17);70,83%). Se considera que el atributo **4 interésI** debería ser objeto de estudio y seguimiento conforme se pretende fomentar perfiles de alumnos con CR= Moderada y Mayor. La tabla 4 presenta la descripción de cada uno de los atributos mencionados. El Análisis multivariante resume en 5 factores los 23 atributos de la tabla original conforme se visualizan en la tabla 5. En esa tabla y tanto para el conjunto de datos correspondientes a alumnos avanzados como ingresantes se trata de ver la correspondencia entre el análisis multivariante y los atributos descriptores de cada encuesta según el algoritmo de clasificación W-J48. Se observa que respecto de los datos de alumnos avanzados la totalidad de los atributos inductores (10) están contenidos en alguno de los cinco factores. Respecto de los registros asociados a alumnos ingresantes los 8 atributos inductores se encuentran en solamente 4 de los 5 factores. En el caso de alumnos ingresantes, no hay atributos asignados al Factor 2.

Factores	ALUMNOS AVANZADOS		ALUMNOS INGRESANTES	
	Atributos según Análisis Factorial	Atributos Inductores según W-J48	Atributos según Análisis Factorial	Atributos Inductores según W-J48
1	1 - 9 - 6 - 21 - 14 - 16 - 17	1 - 17	14 - 10 - 13 - 12 - 2	2 - 14
2	18 - 5 - 2 - 23 - 10	23 - 2 - 18 - 5	11 - 1 - 19 - 22 - 9	
3	22 - 3 - 12 - 11 - 4	3	7 - 15 - 3 - 5 - 23	5 - 15
4	20 - 15 - 8	8 - 20	6 - 4 - 8	4
5	7 - 13 - 19	7	17 - 20 - 18 - 21 - 16	20 - 17 - 16

Tabla 5 - Asignación de atributos según factores y tipo. de alumno

Conclusiones

Hay quienes consideran que en la Argentina el problema de fracaso y deserción universitaria se plantea desde los inicios de la Universidad Pública, y a la vez sostienen que generalmente se reconoce el problema y se intenta solucionarlo a través de medidas puntuales que raramente producen resultados positivos o satisfactorios.

De acuerdo al análisis realizado, en cuestiones de género, edad, estado civil, situación económica, estudios de sus progenitores, entre otras, se concluye que con los modelos y estudios típicos sobre deserción, rezago y fracaso académico las universidades estatales no pueden realmente disminuir considerablemente, sólo mitigar o atenuar, los índices cada vez más alarmantes.

En general, los factores económicos, sociales, culturales y demográficos no han sido determinantes ni causantes del rezago y deserción, por lo menos del grupo de estudiantes considerado.

Este incipiente estudio, que sin dudas requiere de las encuestas asociadas a mayor cantidad de cohortes, ha permitido describir desde la minería de datos y análisis multivariante, el perfil resiliente, tanto de alumnos avanzados como ingresantes, con tan solo 14 de los 23 atributos originales contenidos en 5 dimensiones o factores. Así mismo se observa una correspondencia total entre los atributos inductores encontrados mediante el algoritmo clasificador y el análisis multivariante para los datos asociados a alumnos avanzados, reflejándose un pequeño error en los datos

de alumnos ingresantes en donde los atributos inductores solamente pertenecen a cuatro de los cinco factores determinados mediante el análisis multivariante.

Es de destacar la relevancia encontrada entre las variables que conforman el factor autoconcepto, ya que el autoconcepto promueve la resiliencia y un estudiante resiliente es responsable con sus obligaciones académicas.

El estudio de cohortes posteriores debería poder validar o rebatir esta forma de modelación y análisis.

Se espera que las conclusiones sirvan de aporte, generando un ambiente de discusión en los ámbitos universitarios, que permitan consolidar las políticas públicas implementadas o colaboren en su diseño.

Referencias

- [1] Giovagnoli, P. Determinantes de la deserción y graduación universitaria: una aplicación utilizando modelos de duración. Documento de trabajo, 37. (2002)
- [2] Lopera Oquendo, C. Determinantes de la deserción universitaria en la facultad de economía de la Universidad del Rosario. Universidad del Rosario, Facultad de Economía. Borrador de investigación, 95 (2008). pp. 26.
- [3] Kotliarenco, María Angélica; Cáceres, Irma; Fontecilla, Marcelo. Estado de Arte en Resiliencia. Centro de Estudios y Atención del Niño y la Mujer . Organización Panamericana de la Salud, Oficina Sanitaria Panamericana, Oficina Regional de la Organización Mundial de la Salud. (1997). <http://resilnet.uiuc.edu/library/resilman/resilman.html> [Nov. 2012]
- [4] Munist, Mabel; Santos, Hilda; Kotliarenco, María Angélica; Suarez Ojeda, Elbio Néstor; Infante, Francisca; Grotberg, Edith. Manual de identificación y promoción de la resiliencia en niños y adolescentes. Organización Panamericana de la Salud- Organización Mundial de la Salud - Fundación W.K. Kellogg - Autoridad Sueca para el Desarrollo Internacional (ASDI) - (1998).
<http://resilnet.uiuc.edu/library/resilman/resilman.html> [Nov. 2012]
- [5] Wagnild, Gail M.; Young, Heather M. Development and Psychometric Evaluation of the Resilience Scale. Journal of Nursing Measurement Vol 1, N° 2, (1993). pp 165-178. Springer Publishing Company.
- [6] ADDIMA- Asociación para la promoción y desarrollo de la Resiliencia. <http://www.addima.org/>
- [7] AMERSE. Asociación Mexicana de Resiliencia Salud y Educación. <http://www.amerse.blogspot.com.ar/>
- [8] Fundación Añil - Centro de Resiliencia. <http://www.fundanil.com.ar/>
- [9] ResilienceNet, información para ayudar a niños y familias a superar adversidades. <http://resilnet.uiuc.edu/espanol/index-sp.html>
- [10] Villasmil Ferrer, J. Autoconcepto académico y resiliencia en universitarios de alto rendimiento: construcción desde una perspectiva biográfica. Anuario del Doctorado en Educación: Pensar la Educación, Norteamérica, 0, May. 2012. <<http://erevistas.saber.ula.ve/index.php/anuariodoctoradoeducacion/article/view/3812>>. [Abr. 2013].
- [11] Chen, M., Han, J. and Yu P. Data Mining: An Overview from a Database Perspective. IEEE Transactions on Knowledge and Data Engineering, (1996). V.8 N.6, pp.866-883.
- [12] Agrawal, R., Imielinski, T. and Swami, A. Mining association rules between sets of items in large databases. In: Proceedings of the ACM SIGMOD International Conference on Management of Data. (1993). Washington D.C., 207-216..
- [13] Larose, D. Discovering Knowledge In Data. An Introduction To Data Mining. Published by John Wiley & Sons, Inc., (2005) Hoboken, New Jersey. .
- [14] Hernández Orallo J., Ramirez Quintana, J, Ferri Ramirez, C. Introducción a la Minería de Datos. Pearson-Prentice Hall. (2008).
- [15] Hair, J. F.; Anderson, R.E.;Tatham, R.L. y Black, W.C. (1999). Análisis Multivariante. 5ª Ed. Prentice Hall Iberia, Madrid.
- [16] Mardia, K.V., Kent, J.T. And Bibby, J.M. (1979). Multivariate Analysis, Academic Press, London.
- [17] Romo, A. Los factores curriculares y académicos relacionados con el abandono y el rezago. En [Libros en línea] ANUIES, Deserción, Rezago y Eficiencia Terminal en las IES: Propuesta metodológica para su estudio. (2002).
- [18] Himmel, Erika. Modelos de Análisis de la Deserción Estudiantil en la Educación Superior. Revista Calidad de la Educación, Chile. (2002)
<<http://www.inacap.cl/tportalvp/index.php?t=102&i=2&cc=8460&tm=2>> [Ago. 2008]
- [19] ANUIES - Asociación Nacional de Universidades e Instituciones de Educación Superior. Programas Institucionales de Tutoría. Una propuesta de la AUNIES para su organización y funcionamiento en las instituciones de educación superior. Colección Biblioteca de la Educación Superior. Serie Investigaciones. (2002).

<http://www.anuies.mx/servicios/d_estrategicos/libros/lib42/0.htm> [Nov. 2009]

[20] Centro de Microdatos. Estudio Sobre Causas De La Deserción Universitaria. Departamento de Economía. Universidad de Chile. (2008).

<<http://www.oei.es/pdf2/causas-desercion-universitaria-chile.pdf>> [Sep. 2010]

[21] Lamoutte, Elsa Milagros Casanova. El desarrollo del concepto de sí mismo en la teoría fenomenológica de la personalidad de Carl Rogers. Revista de psicología general y aplicada: Revista de la Federación Española de Asociaciones de Psicología, 46(2), 177-186. (1993)

[22] Resilience Scale
<http://wilderdom.com/html/ResilienceScale.doc>

[23] Rodríguez, Mariana, et al. Propiedades psicométricas de la escala de resiliencia versión argentina. Revista Evaluar, 2009, vol. 9.

[24] Land, Sebastian; Fischer, Simon. RapidMiner 5- RapidMiner in academic use. (2012).

www.rapid-i.com [Feb. 2013]

[25] North, M. Data Mining for the Masses. A Global Text Project Book. (2012)
<http://resilnet.uiuc.edu/library/resilman/resilman.html>
[Nov. 2012]

[26] R - <http://www.r-project.org/>

[27] DYANE -
http://www.miguelsantesmases.com/dyane_v4.htm

[28] Beguerí, G. ; Klenzi, R.; Malberti, A. Determinación de Factores Resilientes en los Alumnos Universitarios. XV Workshop de Investigadores en Ciencias de la Computación- WICC 2013. Paraná, Entre Ríos.

Datos de Contacto de los Autor/es

María Alejandra Malberti
Av. Ignacio de la Roza 590 (O)
Complejo Universitario "Islas Malvinas"
Rivadavia, San Juan Argentina
e-mail: amalberti@gmail.com

Graciela Beguerí
Av. Ignacio de la Roza 590 (O)
Complejo Universitario "Islas Malvinas"
Rivadavia, San Juan Argentina
e-mail: grabeda@gmail.com

Raúl Oscar Klenzi
Av. Ignacio de la Roza 590 (O)
Complejo Universitario "Islas Malvinas"
Rivadavia, San Juan
Argentina
e-mail: rauloscarklenzi@gmail.com

María Alejandra Malberti. Lic. en Cs. de la Computación. Magister en Educación Psicoinformática (2001) Magister en Informática (2008). Prof. Titular Docente – Investigador FCEFNU-UNSJ.

Graciela Beguerí. Prof de Enseñanza Media y Superior en Matemática. Magister en Logística (2007) – UNCUYO. Prof. Titular Docente – Investigador FCEFNU-UNSJ.

Raúl Oscar Klenzi. Ing. en electrónica. Magister en Informática (2008). Prof. Titular FCEFNU-UNSJ. Prof. Titular Docente – Investigador FCEFNU y FI - UNSJ.
