



FACULTAD
DE CIENCIAS
ECONÓMICAS



Universidad
Nacional
de Córdoba

REPOSITORIO DIGITAL UNIVERSITARIO (RDU-UNC)

Inteligencia computacional y TIC: estrategias para facilitar el aprendizaje en la Universidad Nacional de Córdoba

Laura C. Díaz, José María Las Heras, Carlos A. Bartó, Rodolfo Marangunic, Agustín Martino, José Galoppo, Aldo Algorry, Adolfo Vignoli, Andrea Laluf, Leandro Saavedra, Daniel Sandín, Pablo Morales

Ponencia presentada en II Jornadas Nacionales. IV Jornadas de la UNC. Experiencias e Investigación en Educación a Distancia y Tecnología Educativa realizado en 2015 en Córdoba. Córdoba, Argentina



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)



**II Jornadas Nacionales
IV Jornadas de la UNC
Experiencias e Investigación
en Educación a Distancia y
Tecnología Educativa**

Entornos culturales y tecnológicos:
tensiones, cruces y desafíos

**18,19 y 20
de noviembre
de 2015**



Area de Tecnología,
Educación y
Comunicación



UNC

UNIVERSIDAD
NACIONAL DE
CÓRDOBA

SAA

SECRETARÍA
DE ASUNTOS
ACADÉMICOS

Inteligencia computacional y TIC: Estrategias para facilitar el aprendizaje en la Universidad Nacional de Córdoba

Laura C. Díaz, José María Las Heras, Carlos A. Bartó, Rodolfo Marangunic, Agustín Martino, José Galoppo, Aldo Algorry, Adolfo Vignoli, Andrea Laluf, Leandro Saavedra, Daniel Sandín y Pablo Morales

Facultad de Ciencias Exactas, Físicas y Naturales y Facultad de Ciencias Económicas – Universidad Nacional de Córdoba
lcd_ic@yahoo.com.ar, laserasjosemaria@gmail.com, cbarto@gmail.com

Eje 5: Inclusión, tecnologías y derechos ciudadanos

Resumen

El objetivo de esta presentación es socializar los avances del proyecto de investigación “Inteligencia computacional y TIC: Estrategias para facilitar el aprendizaje en la Universidad Nacional de Córdoba” que abarca dos grandes líneas aplicadas en asignaturas masivas de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFyN) y de la Facultad de Ciencias Económicas (FCE), ambas pertenecientes a la Universidad Nacional de Córdoba (UNC).

La primera línea implementa procesos de Inteligencia Artificial con dos objetivos: mejorar el conocimiento que se tiene del estudiante, protagonista principal del escenario de la Educación Superior y, construir modelos de redes conceptuales y ontologías para este tipo de asignaturas. Ambos objetivos están orientados a facilitar la apropiación del conocimiento. La segunda, se aboca a incorporar Tecnologías de la Información y la Comunicación (TIC) en los procesos de aprendizaje y evaluación de estudiantes de carreras de Ingeniería.

Las acciones previstas incluyen, además, estudios desde la metodología cualitativa de las ciencias sociales con el fin de indagar alrededor de las intervenciones y hallazgos de los trabajos realizados como así también la formación de docentes, investigadores y estudiantes en las temáticas abordadas a través del desarrollo de seminarios y talleres. En apretada síntesis, se muestran los resultados publicados, las dificultades encontradas en el desarrollo y las líneas de acción a futuro.

Este proyecto es sucesor del denominado “Sistemas Inteligentes Aplicados a la Enseñanza de Programación en Ingeniería”, desarrollado durante el bienio 2012-2013.

Palabras Clave

Educación Superior, Estudiantes, Aspectos Socioeconómicos, Rendimiento Académico

Introducción

Los lineamientos del desarrollo de la investigación para el periodo 2014-2015 se centraron en la contribución de información y acciones, para facilitar el aprendizaje en la UNC:

1° El uso Tecnologías de Inteligentes de Explotación de la Información (TIEI) para proveer información de estudiantes relativa a características socioeconómicas, académicas y cognitivas. De utilidad para: el diseño de políticas públicas, las diversas modalidades e-learning en continua evolución, las bases de conocimiento de Sistemas Tutores Inteligentes, etc.; en contextos de masividad.

2° Acciones orientadas a los actores más directamente involucrados (Docentes, Investigadores y estudiantes) para la incorporación de recursos TIC en los procesos de aprendizaje y evaluación aplicados a carreras de Ingeniería.

Las contribuciones se materializaron en las acciones de:

1. Detección de relaciones entre las características de los estudiantes de Ingeniería de la FCEFyN de la UNC considerando dimensiones socioeconómicas, académicas y cognitivas, mediante la aplicación de Procesos de TIEI sobre la base de datos de SIU Guarani para las estudiantes de las asignaturas de Informática de la FCEFyN y de la asignatura Administración y Sistemas de Información Gubernamental de la FCE. (Díaz L. Martins S. García Martínez R., 2015)
2. Investigación cualitativa sobre el rol de estudiante y del profesional en el contexto actual, atendiendo a su compromiso social y político (Martino A., Las Heras J., Díaz L., 2015)
3. Formación de los estudiantes miembros de este equipo de investigación, en TIEI, Aprendizaje automático y Sistemas Tutores Inteligentes aplicables a Educación. (Laluf A., Saaavedra L., Díaz L., Bartó C., Algorry A., 2015)
4. El Desarrollo de redes conceptuales para los Sistemas Tutores Inteligentes puestos en funcionamiento.
5. Cursos orientados a la consolidación de las mejoras introducidas sobre la plataforma MOODLE para las Evaluaciones Conceptuales de Informática en el lenguaje Python (Díaz L., Bartó C., 2013) y la puesta a punto del Sistema Tutor Inteligente SIETTE, basado en la teoría de Respuesta al Item, utilizando Test Adaptativos Informatizados que evolucionan mediante Redes Bayesianas (Marangunic R., Bartó C., Díaz L., 2013) y Cursos de Inteligencia Artificial aplicada a Educación.
7. Diseño e implementación de una experiencia "de evaluación formativa" aplicada en alumnos de cursos presenciales de Introducción a la Matemática, con la finalidad de orientar el aprendizaje basado

en competencias, incorporando TIC como estrategia para guiar el desarrollo. (Galoppo J., Díaz L., 2015)

8. Potenciar la comunicación entre alumnos y docentes y apoyar la enseñanza presencial, en orden a lograr un aprendizaje significativo (Ausubel et al., 1978) que permita alcanzar las competencias "de ingreso" fijadas "por CONFEDI" incorporando TIC en los procesos de aprendizaje y evaluación y al desarrollo de un modelo educativo para Ingeniería centrado en competencias.

En esta presentación se profundiza en los resultados obtenidos a partir de las acciones a las que hace referencia el primer apartado, siendo otros también motivo de presentaciones en este evento. Todas las actividades en relación a este trabajo fueron desarrolladas en equipos interdisciplinarios, expertos del dominio y expertos en minería de datos, pertenecientes a este proyecto de investigación y/o al Grupo de Investigación en Sistemas de Información (GISI) de la Universidad Nacional de Lanús. (Díaz L., Martins S., Las Heras J., García Martínez, R., 2015)

El análisis exploratorio con metodologías de las TIEI, el Problema

La Explotación de Información es la subdisciplina de los Sistemas de Información que aporta las herramientas para la transformación de información en conocimiento (García-Martínez et al., 2015). Ha sido definida como la búsqueda de patrones interesantes y de regularidades importantes en grandes masas de información (Britos et al., 2005).

Un Proceso de Explotación de Información se define, como un grupo de tareas relacionadas lógicamente (Curtis et al., 1992) que, a partir de un conjunto de información con un cierto grado de valor para la organización, se ejecuta para lograr otro, con un grado de valor mayor que el inicial (Ferreira et al., 2005).

El uso de TIEI, ofrece la oportunidad de descubrir comportamientos socioeconómicos, académicos, cognitivos, entre otros, de los sujetos en procesos de aprendizaje, que con otras metodologías no serían necesariamente detectados (Kuna et al., 2010).

Primer caso en estudio

Se plantea la pregunta general: ¿Cómo se caracteriza el rendimiento académico de los estudiantes de Ingeniería en sus primeros años, tomando a la asignatura Informática como eje del análisis?

Sobre la base de datos proveniente del sistema SIU_Guaraní de alumnos de las carreras de Ingeniería de la Universidad Nacional de Córdoba, inscriptos en la materia Informática en el primer cuatrimestre de los años 2012-2013 relevada en Julio de 2014. Cuenta con más de 1500 registros y contiene

variada información del estudiante tanto del tipo académico, como socio-económico y de situación geográfica.

Las Variables utilizadas:

- Procedencia del sustento económico del alumno: trabajo propio, familia y/o de beca. Cada una representada como una variable booleana donde 1 indica que su fuente de ingreso proviene de dicha variable.
 - Últimos estudios alcanzados por su padre y madre, representados por una escala de 0 a 4, en la cual 0 indica que el padre o la madre no tienen estudios, 1 que tienen estudios primarios completos o secundarios incompletos, 2 estudios secundarios completos o superiores incompletos, 3 superiores completos y 4 posee estudios de posgrado.
 - Género del estudiante (1 indica que es Masculino, 0 Femenino).
 - Ubicación de procedencia expresada mediante tres variables booleanas: si es argentino, si es de la Provincia de Córdoba, y si es de Córdoba Capital. En caso afirmativo su valor es 1, 0 en caso contrario.
 - Si el alumno aprobó la materia durante el mismo año que realizó la cursada (asignando 1 en caso afirmativo, 0 en caso contrario).
 - Si el alumno realizó la cursada de la materia acorde a lo establecido en el plan de estudios (asignando 1 en caso afirmativo, 0 en caso contrario).
 - Dos variables que determinan el rendimiento del alumno en su primer año de ingreso (categorizados en 4 niveles incrementales desde 0 a 3 acorde a la cantidad de materias que realizan) y su desempeño en el total de años cursados respecto al plan de estudios (categorizados en 5 niveles incrementales de 0 a 4 de acuerdo al grado de materias aprobadas respecto del plan de estudios y del año en el que ingresaron).

A partir de la base de datos proporcionada y la selección de las variables representativas para el dominio de interés, fueron realizadas distintas tareas de: pre-procesado para adecuar la información a las necesidades y requerimientos específicos del proyecto, elección y uso de Procesos de Explotación de Información y aplicación de Tecnologías de Minería de Datos en los Procesos.

Se presenta una descripción estadística de las variables consideradas. Posteriormente, los resultados obtenidos de aplicar los procesos de explotación de información para determinar patrones respecto al desempeño académico a partir de las variables socioeconómicas y de procedencia geográfica. Se han utilizado distintos tipos de algoritmos de agrupamiento con el objetivo de incorporar visualizaciones complementarias de los datos a partir de las características distintivas de cada algoritmo considerado.

Análisis Estadístico del Dominio: Se muestran la media, en adelante Media del Universo (MdU), y el desvío estándar para cada variable seleccionada. En su interpretación se debe tener especial cuidado, atendiendo al pre procesado ya mencionado. Sin embargo, proporcionan una aproximación útil para el análisis. Aproximadamente: el 40% cursa Informática el año en que ingresa, el 31% la aprueba en ese período de cursada, la mayoría cursa simultáneamente entre 3 y 4 materias, el 7% accede a beca, el 14% trabaja, el 98% es de nacionalidad argentina, el 59% es de Córdoba y el 29% de Córdoba Capital, el 92% vive con su familia, el 74% es de sexo masculino, el nivel de estudios de la madre es superior al del padre, ambos oscilan entre dos límites: estudios secundarios completos o universitarios incompletos y, estudios universitarios o superiores completos.

Denominación de la variable	Media	DesvSt
‘Aprobó Inf. en cursada’	0,31	0,46
‘Ritmo Inicial’	1,68	0,95
‘Cumple Plan’	3,13	0,94
‘Demora en Cursarla’	0,38	0,74
‘Beca’	0,07	0,25
‘Trabaja’	0,14	0,35
‘Familia’	0,92	0,37
‘Sexo’	0,74	0,44
‘Madre últimos Estudios’	2,42	0,87
‘Padre últimos Estudios’	2,20	0,94
‘Argentino’	0,98	0,13
‘Córdoba’	0,59	0,49
‘Capital Córdoba’	0,29	0,45

Tabla. 1. Resultados del procesamiento estadístico de las Variables usadas para la descripción del problema

A continuación se describen las reglas de pertenencia a los grupos identificadas para cada uno de los algoritmos utilizados con el fin de caracterizar los estudiantes por sus similitudes. Además se proporciona una descripción de la composición del grupo, para aquellos casos en los que los hallazgos son considerados significativos por el experto del dominio ya sea porque confirman sus percepciones, porque las refutan o por constituir un aporte novedoso. Por último, se describen las percepciones del experto del dominio.

El algoritmo HAC particiona en cuatro conjuntos:

C1	Estudiantes que no viven con su familia de origen y no tienen beca
C2	Están cursando la asignatura en un período posterior al año en que ingresaron, viven con su familia y no poseen beca.

C3	Los beneficiarios de beca
C4	Viven con su familia y cursan informática el año en que ingresan

El algoritmo SOM, los caracteriza como se detalla a continuación:

C1_1	Cursan la materia en el primer año, trabajan y casi en su totalidad no alcanza a aprobarla
C1_2	Demoran en cursarla
C2_1	Cursan la asignatura el primer año, un gran número (el mayor) aprueba en la cursada y no trabajan
C2_2	Cursan la materia en su primer año, no aprueban durante la cursada, no trabajan y no poseen beca

El algoritmo KMeans Strengthening ha proporcionado otra perspectiva del análisis de la formación de grupos, tomando como característica objetivo a la variable "Cumple Plan", como una expresión del desempeño general del estudiante.

K5	Cursan el primer año pero no aprueban la asignatura y viven con su familia
K4	Cursan el primer año pero no aprueban la asignatura y viven con su familia
K1	No aprueban en Cursada y no viven con su familia
K2	No cursan la asignatura conforme al plan de carrera y además no aprueban durante la Cursada, viven con su familia y son varones
K3	No cursan la asignatura conforme al plan de carrera y además no la aprueban en la Cursada, viven con su familia, todas mujeres

Otras relaciones halladas, poco significativas, involucran nivel de estudios de los padres, si vive o no con su familia y otros atributos que sería interesante su consideración para realizar indagaciones con metodologías cualitativas de investigación y profundizar desde el paradigma interpretativo.

Las reglas de pertenencia del algoritmo HAC participaron 'Beca', 'Vive con la familia' y 'Demora en cursar la asignatura', mientras que SOM involucró a 'Demora en cursarla', 'Aprobó Informática en cursada', 'Beca' y 'Trabaja', por último, para Kmeans 'Aprobó en cursada', 'Vive con la familia', 'Demora en cursarla' y 'Mujer/Varón'.

Con relación a la procedencia geográfica, los resultados de la aplicación de HAC:

K1	Son argentinos, no tienen beca y trabajan
K2	Argentinos, beneficiarios de becas

K3	Argentinos, que no trabajan y no tienen beca, comenzaron con el más alto ritmo su carrera, sin embargo el grupo anterior mejores porcentajes de aprobación y de mantener el plan de carrera
K4	Los estudiantes extranjeros
K5	Son argentinos, no trabajan ni tienen beca, comenzaron con el más bajo ritmo

Las principales variables que participan en la caracterización de los grupos son: 'Beca' (si/no), 'Vive con la familia' (si/no), 'Ritmo inicial' (si/no), 'Trabaja' (si/no) y 'Argentino' (si/no). Este algoritmo clasifica grupos muy eficientemente, e invita a profundizar en su composición.

Por último, basado en el interés de indagar las características de los estudiantes del interior de Córdoba y de Capital, se repitió el proceso con once atributos: 'Aprobó Inf. en cursada', 'Ritmo Inicial', 'Cumple Plan', 'Demora en Cursarla', 'Beca', 'Trabaja', 'Familia', 'Sexo', 'Madre últimos Estudios', 'Padre últimos Estudios' y 'Capital Córdoba'.

Algunas interpretaciones de los Expertos del Dominio para dar luz a los procesos decisionales de las políticas públicas destinadas a este universo de estudiantes:

Cada algoritmo separa distintos grupos de interés: HAC al universo de becarios, SOM a los que demoran en cursar la asignatura escogida como representativa de contextos de masividad en el UNC, y KMeans a los que la aprueban conforme a lo esperado en el plan de estudios de la carrera. Es posible adecuar su profundidad y extensión al propósito que persigue el actor que se interroga.

El mejor desempeño en el plan de carreras lo alcanzan los estudiantes que aprobaron la asignatura durante la cursada, con un bajo porcentaje de trabajadores pero con gran cantidad de beneficiarios de beca y con un alto nivel de estudios alcanzados por ambos padres. El desempeño más bajo mostró estar relacionado con demoras en el inicio de la carrera, la no aprobación de la asignatura durante la cursada, un bajo porcentaje de beneficiarios de beca y un nivel más bajo de estudios alcanzados por los padres. Esta podría resultar información clave para las políticas de inclusión y permanencia. Como así también el que ya se describió en relación a la población de Córdoba Capital. Por último, sorprende que, conforme a la escala utilizada para ordenar el nivel de estudios de ambos padres, se evidencie una tendencia sostenida del nivel de estudios de la madre superior al del padre. Las becas muestran una tendencia favorable a mejorar el rendimiento académico, el género de los estudiantes no se muestra relacionado con su desempeño y a los estudiantes que trabajan se les dificulta más sostener el plan de carrera al día.

Segundo caso de estudio

Se presentan los resultados de aplicar el mismo diseño metodológico sobre otro curso con de la UNC, también masivo, en diferente Unidad Académica con dos finalidades, extender el análisis y observar la posibilidad de generalizar el diseño propuesto.

La pregunta general: ¿Cómo se caracteriza el rendimiento académico de los estudiantes de la carrera de Contador de la Facultad de Ciencias Económicas (FCE) de la Universidad Nacional de Córdoba, tomando a la asignatura Administración y Sistemas de Información Gubernamental (ASIG) como eje del análisis?

Se dispone de una base de datos con más de 6500 alumnos registrados en el sistema SIU_Guaraní para la carrera Contador en la FCE de la UNC. La información corresponde a los estudiantes que se inscribieron a la materia ASIG en los años 2012 a 2014 con las mismas características que para el caso anterior:

Análisis Estadístico del Dominio: Para todas las variables involucradas en el desarrollo, con el fin de facilitar la lectura posterior de los resultados obtenidos.

Denominación de la variable	Media	DesvSt
‘Aprobó Asignatura en cursada’	0,04	0,19
‘Ritmo Inicial’	1,78	1,45
‘Cumple Plan’	0,29	0,91
‘Demora en Cursarla’	1,36	0,75
‘Beca’	0,02	0,15
‘Trabaja’	0,61	0,49
‘Familia’	0,58	0,49
‘Sexo’	0,43	0,50
‘Madre últimos Estudios’	2,17	0,86
‘Padre últimos Estudios’	1,96	0,84
‘Argentino’	1	0,06
‘Córdoba’	0,81	0,39
‘Capital Córdoba’	0,50	0,50

Tabla 2. Resultados del procesamiento estadístico de las Variables usadas para la descripción del problema

Es posible afirmar que la mayoría de los estudiantes en su primer cuatrimestre respondía satisfactoriamente al plan de estudio de la carrera, sin embargo al llegar a la asignatura ASIG (tercer año para el plan nuevo vigente desde 2009 y quinto año para el plan viejo) acumulan una demora en de dos años como mínimo, el 2% accede a beca, el 61% trabaja, son casi inexistente los extranjeros, el 81% es de la provincia de Córdoba y el 50% de Córdoba Capital, el 58% costea sus estudios con aportes de la familia, el 57 % es de sexo femenino, el nivel de estudios de la madre es superior al del padre, ambos cercanos a estudios secundarios completos o universitarios incompletos. El experto del

dominio percibe que los estudiantes no aprueban inmediatamente al finalizar la cursada (4%) porque la asignatura les resulta 'ajena' a sus estudios, no posee correlativas superiores y además, es muy extensa. Su sugerencia consiste en reprocesar los datos tomando un año como mínimo para el plazo de aprobación.

A continuación, cada grupo y su regla de pertenencia:

El algoritmo KMeans identifica cinco grupos. Los clasifica eficientemente, e invita a profundizar en su composición.

K1 (2120 individuos)	Proceden de Córdoba, costean sus estudios con aporte familiar, no poseen beca de ayuda económica y no aprobaron la asignatura al finalizar la cursada.
K2 (2200 individuos)	Proceden de la provincia de Córdoba, no aprobaron la asignatura en la cursada y no poseen beca
K3 (1018 individuos)	Proceden de otras provincias, no poseen beca y no aprobaron la asignatura al finalizar la cursada
K4 (212 individuos)	Aprobaron la asignatura al finalizar la cursada
K5 (125 individuos)	Poseen becas y no aprobaron al finalizar la cursada

Las variables que participan en la caracterización son: 'Beca' (si/no), 'Aprobó en cursada' (si/no), 'Córdoba' (si/no), 'Familia' (si/no), 'Trabaja' (si/no), 'Ritmo Inicial' (de 0 a 3) y 'Cumple Plan' (de 0 a 4).

Se proporciona una descripción de su composición en términos de las trece variables usadas en el estudio:

K1	Iniciaron sus estudios con un buen ritmo en la carrera, aunque levemente inferior a los que aprobaron la asignatura (K4) y superior a los becarios (K5), el nivel de estudios de ambos padres es similar a la MdU, el 53% son de Capital, el 33% trabaja, demoraron un poco menos que la MdU en cursar la materia, actualmente cumplen el plan de carreras, aunque muy lentos, algo mejor que la MdU
K2	Todos trabajan, el 71% proceden de Capital, sólo el 10% recibe aporte económico familiar, éstos también trabajan. Todos iniciaron su carrera con el más bajo rendimiento académico. El nivel de estudios del padre y de la madre es bajo, aunque ligeramente superior al de los beneficiarios de beca (K5). Son actualmente los más demorados en su plan de carrera
K3	Ambos padres poseen un alto nivel de estudios, aunque inferior al grupo K4, el 74% recibe

	apoyo económico de la familia, la mitad son mujeres, y el 48% trabaja para costear sus estudios. Este grupo incluye a todos los extranjeros. Comenzaron con un ritmo inicial levemente inferior a la MdU y lo sostienen actualmente en el cumplimiento del plan de carrera y demoran más que la MdU en cursar esta asignatura
K4	Iniciaron la carrera cursando prácticamente todas las materias correspondientes al plan siendo el grupo de mayor 'Ritmo Inicial', cursan esta asignatura en el cuatrimestre que corresponde, sin embargo llevan muy lento su plan de carrera, aunque levemente mejor en relación a la MdU (0,75 vs 0,29). El nivel de estudios de ambos padres es el más alto de la población, el 84 % recibe apoyo económico de su familia, el 34% trabaja, ninguno posee beca, el 79% es de Córdoba y el 42% de Capital, el 54% son mujeres
K5	En su gran mayoría son mujeres (86%), el nivel de estudios de ambos padres es el más bajo del universo, a lo sumo estudios incompletos de nivel secundario, el 94% son de Córdoba, reciben ayuda económica familiar, muy pocos trabajan (20%), su ritmo inicial de carrera fue decayendo con el tiempo y demoran en cursar la asignatura

Las percepciones del Experto:

K1	Hay una correlación directa entre nivel económico, mayor disponibilidad de tiempo porque no trabajan y mejor integración con el medio, son cordobeses
K2	Los resultados son muy consistentes y hay una correlación directa entre nivel económico, nivel de estudios formales de sus padres y rendimiento académico
K3	Son estudiantes de nivel socioeconómico medio alto que no han logrado adecuarse al ritmo de estudios universitario, que se ven debilitados en un medio extraño y las características 'impersonales de la masividad'
K4	La característica que a este grupo como: 'los que aprobaron al finalizar la cursada' tienen relación con aspectos socio-económicos que podría verse materializado en el alto nivel de estudios alcanzado por ambos padres
K5	Su nivel socio-económico más bajo se materializa en el nivel de estudios de ambos padres, que reciben un importante apoyo de sus familias, pero que no es suficiente. En una primera mirada, sorprende que sean mujeres en su gran mayoría, tal vez asociado a las condiciones del mercado laboral. Que en casi su totalidad sean de Córdoba, tal vez explicaría que gente de otras provincias en las mismas condiciones, prefiere elegir otras opciones menos adversas, con costos de transporte más accesibles, inmediatez en materia emotiva y sentimental, etc. No

sorprende que el rendimiento académico sea bajo, en su gran mayoría los estudiantes no alcanzan el ritmo académico propuesto por el plan de carrera

Los resultados de la aplicación de HAC clasifican al universo de estudiantes en ocho conjuntos (C1 a C8). Las principales variables que participan en la caracterización de los grupos son: 'Beca' (si/no), 'Aprobó en cursada' (si/no), 'Córdoba' (si/no), 'Trabaja' (si/no), 'Ritmo Inicial' (de 0 a 3) y 'Cumple Plan' (de 0 a 4), los atributos que difieren del anterior son los relativos a la forma en que el estudiante costea su carrera. No se muestran resultados ya que no agregan nuevos patrones sobre el comportamiento poblacional.

El análisis de los datos utilizando el algoritmo SOM, permite caracterizar seis grupos. No se encontraron hallazgos novedosos que contribuyan significativamente al descubrimiento de nuevos patrones de comportamiento.

Se puede observar que en las reglas para la clasificación de las poblaciones identificadas, los atributos que tienen mayor presencia son: 'Aprobó Asignatura al finalizar la cursada', 'Beca', 'Trabaja', 'Córdoba', 'Cumple Plan' y 'Ritmo Inicial'. Otros atributos aparecen raramente, como 'nivel de estudios de la madre' y 'Familia'. Son descartados: 'Sexo', 'nivel de estudios del padre', 'Argentino' y 'Córdoba Capital' por todos los algoritmos en el proceso de descubrimiento de grupos a partir de sus similitudes. Este hallazgo sugiere la necesidad de indagar acerca de cuáles son los atributos que se interrelacionan más significativamente en la definición de estos grupos.

Conclusiones

Las interpretaciones que se realizan en esta presentación responden al objetivo de una primera caracterización de la población en estudio, a instancias de obtener los primeros resultados que dan luz a los procesos decisionales de las políticas públicas destinadas a este universo de estudiantes. Por otra parte, ellas están acotadas y en cierta medida, determinadas, por las profundidades de los árboles de inducción.

Se observa que es necesario contar con conocimiento a priori del experto del dominio y de sus percepciones en las instancias de preparación de la base de datos y selección de variables como a la hora de leer los resultados e interpretarlos para construir conclusiones que contribuyan a mejorar el conocimiento que se tiene acerca de los estudiantes de Educación Superior en contextos de masividad. En este sentido, durante el desarrollo del trabajo se sostuvo el diálogo en forma permanente entre el equipo de Ingeniería de Explotación de Información y los Expertos del Dominio, pertenecientes a distintas Unidades Académicas de la UNC y al GISI de la UNLa.

Todos los algoritmos contribuyeron a dar respuesta a la pregunta de investigación, quedando acreditado que la metodología utilizada es adecuada para profundizar y extender estas indagaciones.

Las líneas de trabajo sugeridas se relacionan con la necesidad de identificar los atributos de índole socioeconómicos que mayor impactan en el rendimiento académico de los estudiantes e indagar sobre otros aspectos relativos a sus formas de vida.

Bibliografía

- Britos, P., Hossian, A., García Martínez, R. y Sierra, E. (2005). *Minería de Datos Basada en Sistemas Inteligentes*. 876 páginas. Editorial Nueva Librería. ISBN 987-1104-30-8.
- Curtis, B., Kellner, M., Over, J. 1992. *Process Modelling*. *Communications of the ACM*, 35(9): 75-90.
- Díaz, L., Algorry, A., Eschoyey, M., Barto, C., Marangunic, R. (2013). "Actions towards the application of intelligent systems in computer education. *IEEE Latin America Transactions*", 11(1): 591-595. ISSN 1548-0992.
- Díaz L., García Martínez R., (2015): "Hacia Una Praxis Transformadora De La Comprensión Del Estudiante De Educación Superior En Contextos De Masividad". VI Coloquio Intenacional: Estado, Política Pública y Acción Colectiva. Córdoba.
- Galoppo J., Díaz L., Vignoli A., Sandin D., (2015): "Evaluación Formativa de Competencias de Ingreso en los Alumnos de las Carreras de Ingeniería de la Universidad Nacional de Córdoba". EMCI 2015
- Díaz L., Martín S., García Martínez R., (2015): "Descubrimiento De Patrones Socio-Económicos De Población Estudiantil De Carreras De Ingeniería Basado En Tecnologías De Explotación De Información". TE&ET 2015. Corrientes.
- Díaz L., Martins S., Las Heras J., García Martínez R., (2015): "Descubrimiento de Patrones Socio-Económicos de Población Estudiantil de Carreras de Ciencias Económicas Basado en Explotación de Información". Reporte Técnico 24-09-01- GISI-UNLa
- Ferreira, J., Takaj, O., Pu, C. 2005. *Integration of Business Processes with Autonomous Information Systems: A Case Study in Government Services*. *Proceedings Seventh IEEE International Conference on E-Commerce Technology*. Pág. 471-474
- Juarros, M. F. (2006). *¿Educación superior como derecho o como privilegio?: Las políticas de admisión a la universidad en el contexto de los países de la región*. *Andamios*, 3(5): 69-90. ISSN 1870-0063.
- Kohonen, T. 1982. *Self Organized Formation of Topologically Correct Feature Maps*. *Biological Cybernetics*. Vol 43.

- Kuna H, García Martínez R, Villatoro R, (2010): "Identificación de causas de abandono de estudios universitarios. Uso de procesos de explotación de información". *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología, TE&ET*. Art. 5.
- Laluf A., Saavedra L., Díaz L., Algorry A. Bartó C. (2015): "Descubrimiento de Patrones cognitivos en Evaluaciones de Informática basado en Explotación de Información". *TE&ET 2015. Corrientes*.
- Martins, S. 2014. *Derivación del Proceso de Explotación de Información Desde el Modelado del Negocio*. *Revista Latinoamericana de Ingeniería de Software*, 2(1): 53-76. ISSN 2314-2642.
- MacQueen, J. 1967. *Some methods for classification and analysis of multivariate observations*. 5th Berkeley Symposium on mathematics, Statistics and Probability, 1, S. 281-298.
- Martino A., Las Heras J., Díaz L. (2015), "Sola y errante, la voz en búsqueda de alteridad". En *Memorias del VI Simposio de Etica y Filosofía Política*. Facultad de Derecho. Universidad Nacional de Córdoba.
- Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*. San Mateo: Morgan Kaufmann.
- Zabalza M., (2013). *VIII Congreso Iberoamericano de Docencia Universitaria y Nivel Superior*. VIII CIDU <http://www.iberoamericano2014.unr.edu.ar/>