

ERNANE ROSA MARTINS
(ORGANIZADOR)

A CIÊNCIA DA COMPUTAÇÃO E O DESENVOLVIMENTO DE CONTEÚDO TECNOLÓGICO RELEVANTE PARA A SOCIEDADE

ERNANE ROSA MARTINS
(ORGANIZADOR)

A CIÊNCIA DA COMPUTAÇÃO E O DESENVOLVIMENTO DE CONTEÚDO TECNOLÓGICO RELEVANTE PARA A SOCIEDADE

2020 by Atena Editora

Copyright © Atena Editora

Copyright do Texto © 2020 Os autores

Copyright da Edição © 2020 Atena Editora

Editora Chefe: Profª Drª Antonella Carvalho de Oliveira

Diagramação: Natália Sandrini de Azevedo

Edição de Arte: Lorena Prestes

Revisão: Os Autores



Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição Creative Commons. Atribuição 4.0 Internacional (CC BY 4.0).

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores. Permitido o download da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Conselho Editorial

Ciências Humanas e Sociais Aplicadas

Profª Drª Adriana Demite Stephani – Universidade Federal do Tocantins

Prof. Dr. Álvaro Augusto de Borba Barreto – Universidade Federal de Pelotas

Prof. Dr. Alexandre Jose Schumacher – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso

Profª Drª Angeli Rose do Nascimento – Universidade Federal do Estado do Rio de Janeiro

Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná

Prof. Dr. Antonio Gasparetto Júnior – Instituto Federal do Sudeste de Minas Gerais

Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília

Prof. Dr. Carlos Antonio de Souza Moraes – Universidade Federal Fluminense

Prof. Dr. Constantino Ribeiro de Oliveira Junior – Universidade Estadual de Ponta Grossa

Profª Drª Cristina Gaio – Universidade de Lisboa

Profª Drª Denise Rocha – Universidade Federal do Ceará

Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia

Prof. Dr. Edvaldo Antunes de Farias – Universidade Estácio de Sá

Prof. Dr. Elio Martins Senhora – Universidade Federal de Roraima

Prof. Dr. Fabiano Tadeu Grazioli – Universidade Regional Integrada do Alto Uruguai e das Missões

Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná

Profª Drª Ivone Goulart Lopes – Istituto Internazionale delle Figlie di Maria Ausiliatrice

Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense

Profª Drª Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso

Profª Drª Lina Maria Gonçalves – Universidade Federal do Tocantins

Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte

Prof. Dr. Marcelo Pereira da Silva – Universidade Federal do Maranhão

Profª Drª Miranilde Oliveira Neves – Instituto de Educação, Ciência e Tecnologia do Pará

Profª Drª Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa

Profª Drª Rita de Cássia da Silva Oliveira – Universidade Estadual de Ponta Grossa

Profª Drª Sandra Regina Gardacho Pietrobon – Universidade Estadual do Centro-Oeste

Profª Drª Sheila Marta Carregosa Rocha – Universidade do Estado da Bahia

Prof. Dr. Rui Maia Diamantino – Universidade Salvador

Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará

Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande

Prof. Dr. Willian Cleber Domingues Silva – Universidade Federal Rural do Rio de Janeiro

Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

Ciências Agrárias e Multidisciplinar

Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano

Prof. Dr. Antonio Pasqualetto – Pontifícia Universidade Católica de Goiás

Profª Drª Daiane Garabeli Trojan – Universidade Norte do Paraná

Profª Drª Diocléa Almeida Seabra Silva – Universidade Federal Rural da Amazônia
Prof. Dr. Écio Souza Diniz – Universidade Federal de Viçosa
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul
Prof. Dr. Fágner Cavalcante Patrocínio dos Santos – Universidade Federal do Ceará
Profª Drª Gílrene Santos de Souza – Universidade Federal do Recôncavo da Bahia
Prof. Dr. Júlio César Ribeiro – Universidade Federal Rural do Rio de Janeiro
Profª Drª Lina Raquel Santos Araújo – Universidade Estadual do Ceará
Prof. Dr. Pedro Manuel Villa – Universidade Federal de Viçosa
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará
Profª Drª Talita de Santos Matos – Universidade Federal Rural do Rio de Janeiro
Prof. Dr. Tiago da Silva Teófilo – Universidade Federal Rural do Semi-Árido
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

Ciências Biológicas e da Saúde

Prof. Dr. André Ribeiro da Silva – Universidade de Brasília
Profª Drª Anelise Levay Murari – Universidade Federal de Pelotas
Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás
Prof. Dr. Edson da Silva – Universidade Federal dos Vales do Jequitinhonha e Mucuri
Profª Drª Eleuza Rodrigues Machado – Faculdade Anhanguera de Brasília
Profª Drª Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina
Prof. Dr. Ferlando Lima Santos – Universidade Federal do Recôncavo da Bahia
Prof. Dr. Fernando José Guedes da Silva Júnior – Universidade Federal do Piauí
Profª Drª Gabriela Vieira do Amaral – Universidade de Vassouras
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria
Profª Drª Iara Lúcia Tescarollo – Universidade São Francisco
Prof. Dr. Igor Luiz Vieira de Lima Santos – Universidade Federal de Campina Grande
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará
Profª Drª Magnólia de Araújo Campos – Universidade Federal de Campina Grande
Profª Drª Mylena Andréa Oliveira Torres – Universidade Ceuma
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Paulo Inada – Universidade Estadual de Maringá
Profª Drª Renata Mendes de Freitas – Universidade Federal de Juiz de Fora
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande

Ciências Exatas e da Terra e Engenharias

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto
Prof. Dr. Alexandre Leite dos Santos Silva – Universidade Federal do Piauí
Prof. Dr. Carlos Eduardo Sanches de Andrade – Universidade Federal de Goiás
Profª Drª Carmen Lúcia Voigt – Universidade Norte do Paraná
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará
Prof. Dr. Juliano Carlo Rufino de Freitas – Universidade Federal de Campina Grande
Profª Drª Luciana do Nascimento Mendes – Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte
Prof. Dr. Marcelo Marques – Universidade Estadual de Maringá
Profª Drª Neiva Maria de Almeida – Universidade Federal da Paraíba
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

Conselho Técnico Científico

Prof. Me. Abrão Carvalho Nogueira – Universidade Federal do Espírito Santo
Prof. Me. Adalberto Zorzo – Centro Estadual de Educação Tecnológica Paula Souza
Prof. Dr. Adaylson Wagner Sousa de Vasconcelos – Ordem dos Advogados do Brasil/Seccional Paraíba
Prof. Me. André Flávio Gonçalves Silva – Universidade Federal do Maranhão

Prof^a Dr^a Andreza Lopes – Instituto de Pesquisa e Desenvolvimento Acadêmico
Prof^a Dr^a Andrezza Miguel da Silva – Universidade Estadual do Sudoeste da Bahia
Prof. Dr. Antonio Hot Pereira de Faria – Polícia Militar de Minas Gerais
Prof^a Ma. Bianca Camargo Martins – UniCesumar
Prof^a Ma. Carolina Shimomura Nanya – Universidade Federal de São Carlos
Prof. Me. Carlos Antônio dos Santos – Universidade Federal Rural do Rio de Janeiro
Prof. Ma. Cláudia de Araújo Marques – Faculdade de Música do Espírito Santo
Prof. Me. Daniel da Silva Miranda – Universidade Federal do Pará
Prof^a Ma. Dayane de Melo Barros – Universidade Federal de Pernambuco
Prof. Me. Douglas Santos Mezacas -Universidade Estadual de Goiás
Prof. Dr. Edwaldo Costa – Marinha do Brasil
Prof. Me. Eiel Constantino da Silva – Universidade Estadual Paulista Júlio de Mesquita
Prof^a Ma. Fabiana Coelho Couto Rocha Corrêa – Centro Universitário Estácio Juiz de Fora
Prof. Me. Felipe da Costa Negrão – Universidade Federal do Amazonas
Prof^a Dr^a Germana Ponce de Leon Ramírez – Centro Universitário Adventista de São Paulo
Prof. Me. Gevair Campos – Instituto Mineiro de Agropecuária
Prof. Me. Guilherme Renato Gomes – Universidade Norte do Paraná
Prof^a Ma. Jaqueline Oliveira Rezende – Universidade Federal de Uberlândia
Prof. Me. Javier Antonio Albornoz – University of Miami and Miami Dade College
Prof^a Ma. Jéssica Verger Nardeli – Universidade Estadual Paulista Júlio de Mesquita Filho
Prof. Me. José Luiz Leonardo de Araujo Pimenta – Instituto Nacional de Investigación Agropecuaria Uruguay
Prof. Me. José Messias Ribeiro Júnior – Instituto Federal de Educação Tecnológica de Pernambuco
Prof^a Ma. Juliana Thaisa Rodrigues Pacheco – Universidade Estadual de Ponta Grossa
Prof. Me. Leonardo Tullio – Universidade Estadual de Ponta Grossa
Prof^a Ma. Lilian Coelho de Freitas – Instituto Federal do Pará
Prof^a Ma. Liliani Aparecida Sereno Fontes de Medeiros – Consórcio CEDERJ
Prof^a Dr^a Lívia do Carmo Silva – Universidade Federal de Goiás
Prof. Me. Luis Henrique Almeida Castro – Universidade Federal da Grande Dourados
Prof. Dr. Luan Vinicius Bernardelli – Universidade Estadual de Maringá
Prof^a Ma. Marileila Marques Toledo – Universidade Federal dos Vales do Jequitinhonha e Mucuri
Prof. Me. Rafael Henrique Silva – Hospital Universitário da Universidade Federal da Grande Dourados
Prof^a Ma. Renata Luciane Polsaque Young Blood – UniSecal
Prof^a Ma. Solange Aparecida de Souza Monteiro – Instituto Federal de São Paulo
Prof. Me. Tallys Newton Fernandes de Matos – Faculdade Regional Jaguaribana
Prof. Dr. Welleson Feitosa Gazel – Universidade Paulista

**Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)**

C569 A ciência da computação e o desenvolvimento de conteúdo tecnológico relevante para a sociedade [recurso eletrônico] / Organizador Ernane Rosa Martins. – Ponta Grossa, PR: Atena, 2020.

Formato: PDF
Requisitos de sistema: Adobe Acrobat Reader
Modo de acesso: World Wide Web
Inclui bibliografia
ISBN 978-65-86002-68-3
DOI 10.22533/at.ed.683202003

1. Computação – Pesquisa – Brasil. 2. Sociedade e tecnologia.
I. Martins, Ernane Rosa.

CDD 004

Elaborado por Maurício Amormino Júnior – CRB6/2422

Atena Editora

Ponta Grossa – Paraná - Brasil

www.atenaeditora.com.br

APRESENTAÇÃO

A Ciência da Computação estuda as técnicas, metodologias e instrumentos computacionais, visando automatizar os processos e desenvolver soluções com o uso de processamento de dados. Este livro, se propõe a permitir que seus leitores venham a conhecer melhor o panorama atual da Ciência da Computação no Brasil, assim como, os elementos básicos desta ciência, por meio do contato com alguns dos conceitos fundamentais desta área, apresentados nos resultados relevantes dos trabalhos presentes nesta obra, realizados por autores das mais diversas instituições.

A Ciência da Computação, proporciona inúmeros benefícios para a sociedade moderna, tais como: a criação de empregos, o desenvolvimento de novos equipamentos, o ganho de produtividade nas empresas e o acesso à informação. Os estudos desta área são aplicados em diversas outras áreas do conhecimento, proporcionando a resolução de diferentes problemas da sociedade, sendo assim, cada vez mais estes profissionais são valorizados e prestigiados no mercado de trabalho. As empresas enxergam atualmente a necessidade de profissionais cada vez mais qualificados nesta área, a fim de que possam promover ainda mais inovação, desenvolvimento e eficiência.

Dentro deste contexto, este livro aborda diversos assuntos importantes para os profissionais e estudantes desta área, tais como: a utilização das Tecnologias de Informação e Comunicação (TIC's), a acessibilidade na web, a simulação por eventos discretos, as metodologias ativas, as técnicas de Data Mining, os Objetos Digitais de Aprendizagem (ODA), o uso do *Facebook* como interface didático-pedagógica, a aprendizagem colaborativa, os Sistemas de Informação Social, e a avaliação de softwares educativos, como por exemplo, a ferramenta Alice.

Sendo assim, os trabalhos apresentados nesta obra, permitem aos leitores analisar e discutir os relevantes assuntos abordados, tendo grande importância por constituir-se numa coletânea de trabalhos, experimentos e vivências de seus autores. Espera-se que esta venha a ajudar tanto aos alunos dos cursos de Ciência da Computação quanto aos profissionais atuantes nesta importante área do conhecimento, a enfrentarem os mais diferentes desafios da atualidade. Por fim, agradeço a cada autor, pela excelente contribuição na construção deste livro, e desejo a todos os leitores, uma excelente leitura, repleta de boas, novas e significativas reflexões sobre os temas abordados, e que estas possam contribuir fortemente no aprendizado.

CAPÍTULO 5

MINERÍA DE DATOS PARA LA DETERMINAR LOS PERFILES DE RENDIMIENTO ACADÉMICO DE LOS ALUMNOS EN LA UNNE

Data de aceite: 18/03/2020

Julio César Acosta

Facultad de Ciencias Exactas y Naturales y
Agrimensura - Facultad de Ciencias Agrarias.
Universidad Nacional del Nordeste
Corrientes, Argentina.
julioaforever@hotmail.com

David Luis La Red Martínez

Facultad de Ciencias Exactas y Naturales y
Agrimensura. Universidad Nacional del Nordeste
Corrientes, Argentina.
julioaforever@

en base a las cuales clasificaremos los diferentes perfiles de alumnos para poder implementar acciones proactivas que contribuyan a mejorar el rendimiento de los alumnos y disminuir la deserción. Describimos el modelo a implementar con el uso de Data Warehouse para determinar los perfiles de rendimiento académico en las asignaturas Algebra de la carrera Licenciatura en Sistemas de Información (LSI) de la Facultad de Ciencias Exactas y Naturales y Agrimensura (FaCENA) de la Universidad Nacional del Nordeste (UNNE) y Matemática I de la carrera Ingeniería Agronómica (IA) de la Facultad de Ciencias Agrarias (FCA) de la UNNE. Este trabajo se realiza en el marco del Proyecto de Investigación PI 16F002 acreditado por Res. N° 970/16 CS-UNNE.

PALABRAS CLAVE: rendimiento académico; almacenes de datos; minería de datos; modelos predictivos.

RESUMEN: Describimos el contexto en el que se realiza una experiencia en la que evaluaremos el rendimiento de los estudiantes mediante técnicas de Minería de Datos analizando el perfil de cada estudiante desde otras variables, además de las ya clásicas de: calificaciones y desempeño académico. Trabajamos desde el modelo metodológico propuesto de Matriz de Datos y Sistemas de Matrices de Datos que se adecúa al uso que le damos al Data Warehouse para procesar datos y principalmente determinar las variables que intervienen. Buscamos encontrar las variables en estudio entre otras en: factores socioeconómicos, demográficos, actitudinales;

DATA MINING TO DETERMINE THE
ACADEMIC PERFORMANCE PROFILES OF
STUDENTS AT UNNE

ABSTRACT: We describe the context in which an experience is carried out in which we will evaluate student performance through Data

Mining techniques by analyzing the profile of each student from other variables, in addition to the classic ones of: grades and academic performance. We work from the proposed methodological model of Data Matrix and Data Matrix Systems that is adapted to the use we give to the Data Warehouse to process data and mainly determine the variables involved. We seek to find the variables under study among others in: socioeconomic, demographic, attitudinal factors; based on which we will classify the different profiles of students to be able to implement proactive actions that contribute to improve student performance and decrease attrition. We describe the model to be implemented with the use of Data Warehouse to determine the academic performance profiles in the Algebra subjects of the Degree in Information Systems (LSI) of the Faculty of Exact and Natural Sciences and Surveying (FaCENA) of the National University of the Northeast (UNNE) and Mathematics I of the Agricultural Engineering (IA) of the Faculty of Agricultural Sciences (FCA) of the UNNE. This work is carried out within the framework of Research Project PI 16F002 accredited by Res. N ° 970/16 CS-UNNE

KEYWORDS: academic performance; data stores; data mining; predictive models.

1 | INTRODUCCION

Nuestro proyecto de investigación surge de la necesidad de adoptar acciones proactivas frente al desgranamiento y el bajo rendimiento académico de los alumnos de primer año en la Universidad.

La Universidad y las cátedras en estudio han adoptado diversas medidas tendientes a mejorar los resultados expuestos cuantitativos, tales como un Programa de Tutorías, donde un equipo de docentes y alumnos tutores ejecutan un seguimiento y acompañamiento a los alumnos que se detectan han fracasado en el primer examen parcial y planes de clases de apoyo y consultas extraordinarias en vísperas de parciales y durante el transcurso del dictado de las asignaturas; medidas éstas que no han tenido los impactos deseados.

Las carreras en las que se cursan las asignaturas en cuestión tienen un plan de estudio donde se prevé un régimen de correlatividades, que les pueden generar a los alumnos algunas restricciones para el cursado normal de la carrera; Algebra (LSI) tiene la correlativa Cálculo Diferencial e Integral en el primer cuatrimestre de segundo año y Matemática I (IA) tiene la correlativa

Matemática II en el segundo trimestre de primer año.

Ambas asignaturas requieren que la capacidad de razonamiento puro esté involucrada en la enseñanza-aprendizaje, y en el caso de las Matemáticas aplicadas, se procura el conocimiento matemático para usarlos en la aplicación de soluciones concretas y reales de la vida práctica profesional; los alumnos se enfrentan en muchos casos por primera vez, al problema de adquirir conocimiento de modelos

matemáticos, para luego aplicarlos en problemas concretos y luego interpretar desde la situación problemática planteada, los resultados obtenidos de los modelos matemáticos usados.

La cantidad de alumnos que regularizan y/o que aprueban las asignaturas involucradas en este proyecto no es satisfactoria, consideramos que esa situación puede contribuir al desgranamiento y deserción de los alumnos en los primeros niveles de sus carreras. Es importante, por tanto, *estudiar y determinar cuáles son las variables que inciden en el rendimiento académico a fin de poder establecer estrategias de acción pedagógicas que permitan mejorar dicho rendimiento*.

El tema de nuestra investigación tiene plena vigencia y actualidad en nuestra Universidad, que tiene políticas definidas de atención y contención a la demanda masiva de parte de los alumnos (principalmente en los primeros años).

El mejoramiento de la calidad académica en la Universidad, no necesariamente debe enfocarse sólo en el sistema de enseñanza-aprendizaje, sino que se debe atender otras variables, como por ejemplo, la sistematización de procesos de evaluación permanentes que permitan monitorear cuestiones ligadas a la calidad académica y retroalimente la propuesta de mejora para la Universidad (BRIAND, DALY y WUST, 1999). El *rendimiento académico* es uno de los factores más críticos que debe evaluarse continuamente.

Definimos rendimiento académico como la productividad del sujeto, matizado por sus actividades, rasgos y la percepción más o menos correcta de los cometidos asignados (MALETIC, COLLARD y MARCUS, 2002). Evaluaremos elementos que influyen en el desempeño como: los factores socioeconómicos, la amplitud de programas de estudio, las metodologías de enseñanza, los conocimientos previos del alumno (MARCUS, 2003); por esta razón, no resulta adecuado evaluar el desempeño general de los alumnos a través de porcentajes de aprobación, notas obtenidas, etc., ya que esos procesos de evaluación no brindan toda la información necesaria para detectar, y corregir problemas cognitivos, de aprehensión, de discernimiento, actitudinales.

Determinaremos las características propias del estudiante, analizando patrones de comportamiento y de condiciones que posibiliten la definición de los perfiles de los alumnos.

Se presentan varios métodos para determinar y clasificar patrones que se utilizan en Inteligencia Artificial (del inglés Artificial Intelligence - AI) y Aprendizaje de Máquinas (del inglés Machine Learning – ML). La Minería de Datos (del inglés Data Mining - DM), son procesos de descubrimiento de nuevas y significativas relaciones, patrones y tendencias en grandes volúmenes de datos utilizando técnicas de AI y ML. Con estas técnicas se extraen patrones y tendencias para describir, comprender mejor los datos y predecir los comportamientos futuros (MARCUS y MALETIC,

2003). Se define DW (del inglés Data Warehouse) como una colección de datos orientada a un dominio, integrada, no volátil y variante en el tiempo para ayudar a tomar decisiones (SALTON, 1989). Se expone que la necesidad de proporcionar una fuente única de datos limpia y consistente para propósitos de apoyo para la toma de decisiones y la necesidad de hacerlo sin afectar a los sistemas operacionales son las razones por las que surgen los DW (MOLINA LOPEZ y GARCIA HERRERO, 2006).

Esperamos contribuir a encontrar una respuesta al histórico bajo rendimiento académico de los alumnos. Los modelos predictivos que buscamos, permitirán tomar acciones tendientes a evitar el fracaso académico, detectando los alumnos con perfil de riesgo de fracaso académico de manera temprana, a poco del inicio del cursado de las asignaturas; lo que permitirá concentrar en ellos los esfuerzos de tutorías y apoyos especiales.

2 | MATERIALES Y MÉTODOS

Buscamos detectar grupos de estudiantes en riesgo de fracaso en sus estudios, con la finalidad de tomar medidas proactivas frente al desgranamiento y el bajo rendimiento académico de los alumnos de primer año en la Universidad.

Si bien ambas asignaturas donde se realiza la experiencia tienen régimen de acreditación similar, difieren en la carga horaria y los tiempos de dictado a saber: Algebra (LSI) tiene 128 (ciento veintiocho) horas reloj de dictado de las cuales el 50% corresponde a teoría y el 50 % a trabajos prácticos en la modalidad cuatrimestral (corresponde al primer cuatrimestre de primer año de la carrera), mientras Matemática I (IA) tiene 96 (noventa y seis) horas reloj de dictado con idéntica distribución porcentual de tiempos de dictado de teoría y de trabajos prácticos, pero en la modalidad trimestral (corresponde al primer trimestre de primer año de la carrera).

En ambas asignaturas para alcanzar la condición de alumno regular, los alumnos deben asistir al menos al 75% de las clases de trabajos prácticos, que se dictan dos veces por semana en clases de 2 hs. cada una y deben aprobar 2 (dos) exámenes parciales cuyos contenidos son exclusivamente de trabajos prácticos; cada uno de ellos tiene su instancia de recuperación y para aquellos alumnos que hayan aprobado al menos 1 (uno) de los parciales en cualquiera de las 4 (cuatro) instancias disponibles, existe una instancia más para recuperar el examen que queda aún sin aprobar. Cualquiera de los exámenes parciales se aprueba con 60 (sesenta) puntos sobre 100 (cien) puntos posibles. La asistencia a clases de teoría es libre y se dictan dos veces por semana en clases de 2 hs. cada una.

Se acreditan ambas asignaturas con un examen final al que se accede en

condición de alumno regular o de alumno libre; el alumno regular debe rendir en el examen final oral solamente de los contenidos de teoría. El alumno que se presenta al examen final en condición de alumno libre, debe rendir un examen escrito de trabajos prácticos y tras aprobar esa instancia pasa al examen de teoría en condiciones similares a la antes mencionada.

Los porcentuales de los alumnos que regularizan Algebra y Matemática I no son los deseados; en el caso de Algebra, de 320 alumnos inscriptos en los últimos 4 años en promedio, aproximadamente un 20% no alcanza a rendir el primer examen parcial en promedio y al final del cursado, regularizan la asignatura solo un 30% aproximadamente; en el caso de Matemática I el desgranamiento después del primer parcial no es tan evidente y el porcentual aproximado de alumnos regulares al final del cursado es del 40%.

La cantidad de alumnos que regularizan y/o que aprueban las asignaturas involucradas en este proyecto no es satisfactoria, consideramos que esa situación puede contribuir al desgranamiento y deserción de los alumnos en los primeros niveles de sus carreras. Es importante, por tanto, estudiar y determinar cuáles son las variables que inciden en el rendimiento académico a fin de poder establecer estrategias de acción pedagógicas que permitan mejorar dicho rendimiento.

Trabajamos en el desarrollo de métodos que contribuyan a encontrar técnicas para la detección temprana de los alumnos que tendrán dificultades en sus estudios, para ofrecerles contención y acompañamiento especial en el inicio de sus estudios Universitarios. Indagamos en aspectos como:

- a) diferencia del nivel de aprendizajes de contenidos previos en los alumnos,
- b) situaciones particulares personales de los propios alumnos,
- c) la capacidad de las cátedras para el seguimiento del aprendizaje de los alumnos,
- d) escasa motivación para el estudio de ciencias básicas y otros que puedan revelarse como incidentes en la problemática que nos ocupa.

Para recuperar contenidos en los grupos de riesgo detectados trabajaremos con materiales elaborados con nuevas tecnologías de la información (NTIC). Esto no debe desplazar ni sustituir las formas presenciales de enseñanza - aprendizaje, sino más bien ofrecer alternativas diferentes para aquellos alumnos que requieren modelos diferentes para sus estudios y aprendizajes. Se considera que las NTIC tienen el potencial para desempeñar un papel importante en la recuperación de contenidos al permitir un abordaje más eficaz, en el sentido de permitirnos procesos de aprendizaje más profundos y más persistentes, mientras el peso de un aprendizaje efectivo permanece con las personas, sus capacidades y valores interpersonales (MOTSCHING-PITRIK y HOLZINGER, 2002) (DEMLI, HAMPEL y MOTSCHING-

PITRIK, 2011).

Entendemos importante en nuestro trabajo el estudio en dos poblaciones aparentemente diferentes como son los alumnos de las carreras LSI y de IA, para determinar si los perfiles de los estudiantes varían según la elección de la carrera y medir las diferencias en la predisposición y adaptación para el trabajo y aprendizaje mediado con las NTICs.

En los últimos años se han realizado numerosos trabajos relacionados con la producción de contenidos; en se tiene una concepción global e integral del e-learning, en estos nuevos escenarios se incluyen la combinación del aprendizaje cara a cara y el soportado por medios tecnológicos (especialmente la Web), tal que las fortalezas de ambas configuraciones se puedan aprovechar y explotar (NICHOLS, 2003). Este aprendizaje combinado (blended learning o b-learning) se considera de suma utilidad no sólo para las universidades sino también para la sociedad en general.

Se ha corroborado que los docentes del siglo XXI deben incorporar definitivamente las NTICs como recursos didácticos, sin abandonar los tradicionales de tiza y pizarrón, pero deben conocer el uso de las NTICs con al menos en parte del potencial que ellas ofrecen (ACOSTA y LA RED MARTINEZ, 2012); algunas teorías psicológicas y pedagógicas consideran necesaria la inclusión del e-moderator docente con habilidades especiales en las actividades online (SALMON, 2000). La actividad del docente tutor se transforma a veces en un hecho fundamental, la manera en que se usa la tecnología puede transformarse en un factor de gran influencia en la calidad de la EA-EV (enseñanza - aprendizaje en entornos virtuales) (WENGER, WHITE y SMITH, 2009). Se debe trabajar entonces para lograr una forma de EA-EV que tome en cuenta las necesidades individuales, los intereses y estilos.

En este proyecto de investigación, las variables que inciden en el rendimiento académico de los alumnos serán detectadas a fin de establecer, a través de los valores que ellas toman en cada caso, la población de alumnos en riesgo de fracaso, para establecer acciones tendientes a evitar el fracaso de cada uno de los alumnos, con las acciones que correspondan en cada caso particular y/o de cada grupo detectado y disminuir así el posterior desgranamiento.

3 I METODOLOGIA

Identificamos en la información que alimenta nuestra base de datos el modelo de *sistemas de matrices de datos*, según el cual, si asumimos que todo objeto puede ser analizado desde una matriz de datos (SAMAJA, 2005):

- a) Todos los datos de todas las investigaciones científicas de todas las disciplinas tienen una estructura invariante llamada *matriz de datos* –que se conforma con una *unidad de análisis* (UA), una *variable* (V), una escala de *valores* para las variables (R) y *el indicador*.
- b) Todas las investigaciones científicas contienen datos de distinto tipo y de diferentes niveles de integración, existe *un conjunto de matrices de datos* que guardan entre sí relaciones lógico-metodológicas determinadas.
- c) El lugar del indicador en la conformación del dato, no es un detalle menor, lo pensamos como aquellos procedimientos aplicados a dimensiones relevantes de la variable para efectuar su medición.

Se identifican elementos de diversos tipos y configuraciones en la descripción de cualquier objeto complejo, y en tal sentido aparece necesariamente un “*grupo de matrices*” formado, por lo menos, por tres matrices de datos: una matriz de datos central, en lo que llamaremos “nivel de anclaje” (N_a), que se focaliza en el plano de la investigación; la unidad de análisis de la matriz de datos del N_a tiene atributos que pueden tratarse a su vez como una nueva matriz de datos, pero ahora en un nivel inferior al N_a , al que llamaremos “nivel sub unitario” (N_{-1}) y una matriz constituida por el contexto de las unidades de análisis del N_a , que denominaremos matriz supraunitaria y se encuentra en un “nivel supra unitario” (N_{+1}).

Se enuncia una *ley general del análisis de datos* según la cual el análisis de datos tiene como tarea invariante *la comparación de un estado de cosas existente (o dado empíricamente) con un estado de cosas posibles en el marco de un modelo (o presunción) asumida como necesaria*; así por ejemplo, cualquier valor estadístico (promedio, frecuencia, mediana, etc.) tendrán sentido solamente si pueden ser *comparados* con algún patrón (conocido o inferido), para estimar hasta qué punto los valores de nuestro estudio coinciden o no con lo esperado y a partir de ello poder estimar la situación presente como contingente o necesaria.

Se proponen direcciones para el tratamiento y análisis de datos: La dirección del tratamiento y análisis de datos *centrada en la variable*, informa acerca del comportamiento de la población con respecto a alguno de sus aspectos más relevantes; *variable* es lo que se puede predicar de la unidad de análisis y presenta variaciones (de calidad, de orden, de cantidad, de relación) en cada una de las unidades de análisis o de una misma unidad de análisis en diferentes momentos; este tratamiento se hace con procedimientos de estadística descriptiva, va desde el análisis univariado o bivariado hasta los distintos tipos de análisis multivariado entrega información principalmente sobre la población en estudio, a partir de una muestra y esos valores nos entregarán información de la población, siempre y cuando la muestra sea representativa del universo (YNOUB, 2007).

La otra dirección del análisis es *centrada en la unidad de análisis*, la cual

nos permite caracterizar los diferentes valores de las variables de cada unidad de análisis, de manera tal que las diferentes configuraciones sean “información” a partir de la cual se pueda inferir una dinámica integral, propia del universo en estudio.

Se introduce una tercera dirección de análisis, que denomina *centrado en el valor*; es el análisis en el cual se aplican tratamientos destinados a sistematizar, codificar y/o agregar información, con vistas a la construcción de una variable, la construcción de la variable será un medio más que un fin; ello permite así explicar el tratamiento de la información “desde el origen”.

En investigaciones que tratan objetos usuales, esta dirección de análisis puede no evidenciarse. Pero en nuestro trabajo donde no todas las variables que inciden la determinación de los perfiles de rendimiento académico de los alumnos están determinadas, este modo de abordar el tratamiento y análisis de datos tiene una importancia particular y es aquí donde el DW será un importante auxiliar para descubrir cuáles son esas variables.

Tratamiento y análisis de datos centrado en la unidad de análisis y en el valor. No se debe confundir el análisis centrado en el valor con el análisis centrado en la unidad de análisis, porque ambos tienen “sentido horizontal” en la tabla de datos, pero lo cierto es que operan en diferentes niveles, mientras el tratamiento centrado en la unidad de análisis lo hace en el nivel de anclaje, el tratamiento centrado en el valor lo hace en los niveles subunitarios.

El análisis centrado en el valor consiste sintéticamente en:

- a. idear criterios para clasificar información cualitativa y exploratoria.
- b. ejecutar los procedimientos de resumen que se hayan previsto para sintetizar variables multidimensionales (escalas, índice, tipologías, etc.)
- c. reagrupar valores (para poner de manifiesto alguna heterogeneidad respecto de alguna característica relevante).

Esta dirección del análisis responde a tres problemas:

- a. la confiabilidad de la información obtenida (de cada medición y del conjunto de las mediciones).
- b. la validez de los indicadores elaborados (escalas, índices, tipologías, etc.).
- c. el reagrupamiento de valores como efecto de los resultados obtenidos.

Tratamiento y análisis de datos centrado en la dirección de la variable. Esta es la dirección de análisis destinada a sintetizar la información acerca de una(s) variable(s) en particular. Para ello disponemos de las herramientas de la Estadística descriptiva y de la Estadística inferencial;

Se define el campo de aplicación de cada uno de los tratamientos y análisis de

datos que se hacen usando métodos estadísticos a saber: la estadística descriptiva entiende en la recolección, ordenamiento, análisis y representación de un conjunto de valores de una variable, con la finalidad de describir las características de la variable; mientras la estadística inferencial, a través de determinados métodos y procedimientos, es capaz de inferir las propiedades de una población o de los elementos de ella a partir del estudio estadístico de una porción de la misma, llamada muestra (JOHNSON y KUBY, 2003).

4 | DATA WAREHOUSE

Como soporte de los datos trabajaremos con Data Warehouse (DW); en informática, un almacén de datos (DW), es un sistema especial de bases de datos utilizado para el almacenamiento de datos y el procesamiento de los mismos para la presentación de informes y análisis de información, es considerado como un componente central de la inteligencia de organizaciones.

Un DW es un repositorio de datos que proporciona una visión global, común e integrada de los datos que presenta las siguientes características: (CURTO DIAZ, 2010)

- a) Orientado a un tema: organiza una colección de información alrededor de un tema central.
- b) Integrado: incluye datos de múltiples orígenes y presenta consistencia de datos.
- c) Variable con el tiempo: se realizan fotos de los datos basadas en fechas o hechos.
- d) No volátil: sólo de lectura para los usuarios finales.

Detrás de la arquitectura de componentes del DW existe un conjunto de procesos básicos asociados: los ETL (del inglés Extract, Transform, Load – Extracción, Transformación y Carga).

Los procesos ETL hacen referencia a la recuperación y transformación de los datos desde las fuentes orígenes cargándolos en el DW. En primer lugar los datos se analizan desde las fuentes y se extraen aquellos que serán de utilidad para el proceso en ejecución.

Luego de extraer los datos se los carga al DW pero, en muchas ocasiones, éstos requieren pasar por un proceso de transformación. La transformación de los datos significa un formateo y/o estandarización de los mismos convirtiendo ciertos números en fechas, eliminando campos nulos, etc.

Es necesario que antes de completar el DW con los datos se realicen controles para enviar información cualitativamente correcta. Luego se procede a aplicar alguna

técnica para realizar el análisis de los datos almacenados en el DW. El método más utilizado es el proceso de DM que aplica la inteligencia artificial para encontrar patrones y relaciones dentro de los datos permitiendo la creación de modelos, es decir, representaciones abstractas de la realidad.

Existen varias alternativas del DM, por ejemplo: la Minería de Datos en Educación (Educational Data Mining, EDM). El objetivo de la EDM es el desarrollo de métodos para la exploración de tipos de datos únicos provenientes de plataformas educativas y usándolos para entender mejor a los estudiantes en el aprendizaje (FORMIA y LAZARINI, 2013) (PEREIRA, ROMERO y TOLEDO, 2013) (LA RED MARTINEZ, ACOSTA, URIBE y RAMBO, 2012) (LA RED MARTINEZ, GIOVANNINI, BAEZ MOLINAS, TORRES y YACUZZI, 2017)

Modelo propuesto: La estructura del DW, consta de una tabla de hechos y varias tablas de dimensión. Una tabla de hechos o una entidad de hecho es una tabla o entidad que almacena medidas para medir el negocio como las ventas, el coste de las mercancías o las ganancias.

Cada medida se corresponde con una intersección de valores de las dimensiones y generalmente se trata de cantidades numéricas, continuamente evaluadas y aditivas. Se pueden distinguir dos tipos de columnas en una tabla de hechos, columnas de hechos y columnas llaves. Las columnas de hechos almacenan las medidas del negocio que se quieren controlar y las columnas llaves forman parte de la clave de la tabla. Una tabla de dimensiones o entidad de dimensiones es una tabla o entidad que almacena detalles acerca de hechos. Por ejemplo una tabla de dimensión de hora almacena los distintos aspectos del tiempo como el año, trimestre, mes y día. Además incluye información descriptiva sobre los valores numéricos de una tabla de hechos. Las tablas de dimensiones para una aplicación de análisis de mercado, por ejemplo, pueden incluir el tipo de período de tiempo, región comercial y producto. Asimismo las tablas de dimensiones describen los distintos aspectos de un proceso de negocio. Si se desea determinar los objetivos de ventas, se pueden almacenar los atributos de dichos objetivos en una tabla de dimensiones. Cada tabla de dimensiones contiene una clave simple y un conjunto de atributos que describen la dimensión.

En nuestro caso, las columnas de una tabla de dimensiones se utilizan para crear informes o para mostrar resultados de consultas. Por ejemplo las descripciones textuales de un informe se crean desde las etiquetas de las columnas de una tabla de dimensiones. El modelo que se presenta en este trabajo se compone de la tabla de hechos “ALUMNOS” y varias tablas de dimensiones asociadas a la misma que incluyen características que se desean estudiar.

Etapa de recolección de datos: Tal como se planteó, el estudio del desempeño académico de los estudiantes no sólo debe evaluarse teniendo en cuenta los

resultados de las instancias de evaluaciones previstas por la asignatura sino que también deben analizarse otros factores culturales, sociales y/o económicos que afecten el rendimiento del alumno. Por ello para este trabajo resultó determinante la participación directa del estudiante, pues era necesario conocer datos sobre aspectos personales que no se podían obtener de otra manera que no fuera a través de respuestas directas por parte de cada alumno. A tal fin se dispuso la elaboración de una aplicación web que permitió contar con una Encuesta On-Line compuesta por preguntas relacionadas a situación familiar e historial de estudios secundarios, entre otras cuestiones.

Etapa de depuración y preparación de datos: Para la realización de una correcta explotación del DW se debe asegurar que los datos obtenidos en la etapa anterior sean consistentes y mantengan la coherencia entre ellos. Así, en la etapa siguiente, se realizará un proceso de limpieza en los datos, que es la eliminación de aquellos registros con todos sus campos en blanco, corrección de errores tipográficos, llenado de algunos campos nulos, entre otros.

La Encuesta no permite la carga, por parte de los estudiantes, de calificaciones de la asignatura en estudio. Esto se dispuso así para evitar errores en los datos ya sea por olvido, o confusión al momento de ingresar los valores. Por ello la carga de notas correspondientes al primer parcial, segundo parcial y sus recuperatorios, examen final y situación del alumno (regular, promovido o libre), es realizada por el equipo responsable de este trabajo de investigación.

La información se obtendrá a partir de la base de datos histórica de las cátedras continuarán respecto a las calificaciones de los alumnos. Con esta información depurada se deberá proceder a trabajar en las próximas etapas: - Carga de Datos al DW: Mediante la ejecución del flujo de datos, la información almacenada en la tabla *encuesta* se distribuirá a las tablas pertenecientes al modelo del DW.

5 | RESULTADOS

Hasta el momento se ha completado la primera etapa que implicó el diseño del modelo del DW sobre el cual se implementarán técnicas de DM a fin de encontrar las principales variables que intervienen en el rendimiento académico de los alumnos para así determinar los perfiles de rendimiento académico de los estudiantes, vinculados a su desempeño académico en las asignaturas LSI-FaCENA e IA-FCA UNNE. En el avance que aquí se presenta respecto del Proyecto se pudo comprobar que la etapa de depuración y preparación de los datos ha demandado tiempo y esfuerzo debido principalmente a la poca integridad y coherencia que existía en la información recolectada que luego se usará para realizar la evaluación final. En

etapas sucesivas se continuará con el proceso de minería de datos para evaluar y comparar patrones que se obtengan, incorporando eventualmente nuevas variables detectadas para definir los perfiles de estudiantes. La evaluación, análisis y utilidad de estos patrones con los que se construirá un modelo predictivo de rendimiento académico permitirá soportar la toma de decisiones eficaces por parte del cuerpo docente de las asignaturas involucradas.

REFERENCIAS

- ACOSTA, J. y LA RED MARTINEZ, D., **Un aula virtual no convencional de Algebra en la FaCENA-UNNE**. Saarbrücken: Ed. Académica Española. 2012.
- BRIAND, L.; DALY, J. y WUST, J. A unified framework for coupling measurement in objectorientated systems. **IEEE Transactions on Software Engineering**, 25 (1), 91-121. 1999.
- CURTO DIAZ, J. **Introducción al business intelligence**. Barcelona: UOC. 2010.
- DERRTL, M.; HAMPEL, T.; MOTSCHING-PITRIK, R. y PITNER, T. Inclusive social tagging and its support in Web 2.0 services. **Computers in Hum an Behavior**, 27(4), 1460-1466. 2011.
- FORMIA, S. y L. LANZARINI, L. Caracterización de la deserción universitaria en la UNRN utilizando minería de datos. **Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (TE&ET)**, (11), 92-98. 2013.
- JOHNSON, R. y KUBY, P. **Estadística Elemental. Lo esencial**. México DF: International Thomson Editores. 2003.
- LA RED MARTINEZ, D.; ACOSTA, J.; URIBE, V. y RAMBO, A. Academic Performance: An Approach From Data Mining. **Journal of Systemics, Cybernetics and Informatics**, 10 (1), 66-72. 2012.
- LA RED MARTINEZ, D.; GIOVANNINI, M.; BAEZ MOLINAS, M.; TORRE, J. y YACUZZI, N. Academic performance problems: A predictive data mining-based model. **Academia Journal of Educational Research**, 5, (4), 61-75. 2017.
- MALETIC, J.; COLLARD, M. y MARCUS, A. Source Code Files as Structured Documents. En J. Maletic; M. Collard y A. Marcus (Coords.), 10th IEEE International Workshop on Program Comprehension (IWPC'02) (pp. 289-292). París. 2002.
- MARCUS, A. *Semantic Driven Program Analysis*. Tesis doctoral sin publicar. Kent State University, Ohio. 2003.
- MARCUS, A. y MALETIC, J. Recovering Documentation-to-Source-Code Traceability Links using Latent Semantic Indexing. En A. Marcus y J. Maletic (Coords.), 25th IEEE/ACM International Conference on Software Engineering (ICSE'03) 3(10) (pp. 125-137). Portland, Oregon. 2003.
- MOLINA LOPEZ, J. y GARCIA HERRERO, J. **Técnicas de Análisis de Datos**. Madrid: Universidad Carlos III. 2006.
- MOTSCHING-PITRIK, R. y HOLZINGER, A., Student-centered teaching meets new media: concept and case study. **Journal of Educational Technoplogy and Society**, 5(4), 160-172. 2002.
- NICHOLS, M. A theory for e-Learning. **Journal of Educational Technology and Society**, 2, 1-10.

2003.

PEREIRA, R.; ROMERO, A. y TOLEDO, J. Descubrimiento de perfiles de deserción estudiantil con técnicas de minería de datos. **Vínculos**, (10) 1, 374-383. 2013.

SALMON, G. **E-moderating: The key to teaching and learning online**. London: Kogan Page. 2000.

SALTON, G. **Automatic Text Processing**: The Transformation, Analysis and Retrieval of Information by Computer. Boston: Addison-Wesley Longman Publishing Co. 1989.

SAMAJA, J. **Epistemología y Metodología: elementos para una teoría de la investigación científica**. Buenos Aires: Eudeba. 2005.

WENGER, E.; WHITE, D. y SMITH, J. **Digital habitats. Stewarding technology for communities**. Portland: Cpsquare. 2009.

YNOUB, R. **El proyecto y la metodología de la investigación científica**. Buenos Aires: CENGAGE Learning. 2007.

 Atena
Editora

2 0 2 0