Artículo **Revista de Tecnología y Educación**
Junio, 2020 Vol.4 No.11 1-7

Redes neuronales para predecir el abandono académico en ingeniería

Neural networks to predict academic dropout in engineering

CENTURIÓN-CARDEÑA, Humberto José[1]†*, CANO-BARRÓN, Danice Deyanira[1], SANDOVAL-GÍO[2], Jesús y ZAPATA-GONZÁLEZ, Alfredo[3]

[1]Tecnológico Nacional de México - Instituto Tecnológico Superior de Motul
[2] Tecnológico Nacional de México - Instituto Tecnológico de Mérida
[3]Universidad Autónoma de Yucatán – Facultad de Educación

ID 1er Autor: *Humberto José, Centurión-Cardeña* / **ORC ID**: 0000-0003-3446-2185

ID 1er Coautor: *Danice Deyanira, Cano-Barrón* / **ORC ID**: 0000-0002-3988-9308

ID 2do Coautor: *Jesús, Sandoval-Gío* / **ORC ID**: 0000-0001-5847-3669

ID 3er Coautor: *Alfredo, Zapata-González* / **ORC ID**: 0000-0001-5087-6244

DOI: 10.35429/JTAE.2020.11.4.1.7 Recibido: 24 de Abril 2020; Aceptado: 30 de Junio, 2020

Resumen

Este trabajo describe el proceso de diseño de un modelo basado en una red neural para predecir a estudiantes de ingeniería en situación de riesgo de abandono académico basado en sus perfiles socioeconómicos, académicos y personales. Se basa en la metodología CRISP-DM para la generación de un modelo predictivo utilizando redes neuronales artificiales de retropropagación de una capa y se parte del análisis de datos de contexto y académicos de los estudiantes provenientes de un cuestionario elaborado por CENEVAL en su examen de ingreso y su estado académico después de un año en la institución. Para el entrenamiento de la red neuronal se consideraron las últimas cuatro generaciones de ingreso del Instituto Tecnológico Superior de Motul y la base de datos considera 48 atributos de los casi 120 que considera el instrumento original y cerca de 781 registros para este modelo. El resultado del uso de la arquitectura neuronal es un modelo predictivo con un nivel de significancia del 75.42% y un índice F de 0.6027. Se trabaja en montar este conocimiento en el Sistema Integral de Tutorías que se ha puesto en marcha para poder hacer el seguimiento.

Red Neuronal Artificial, Educación Superior, Abandono

Abstract

This project describes the design process of an artificial neural network model to predict the risk of dropping out of engineering students throughout their socioeconomic, academic, and personal data sing CRISP-DM methodology. The neural network used in the project considers backpropagation functionality with one hidden layer on data from a context questionnaire and academic data from the students in its CENEVAL's entrance exam and their academic status after one year in the institution. The data used to train the neural network it's from 781 records of the last four generations of freshmen year students at the Technological Institute of Motul organized in 48 attributes out of the almost 120 included in the original instrument. The result is a predictive model with a significance level of 75.42% and an F index of 0.6027. This model will be included in the comprehensive tutoring system that is been develop within the organization to monitor the student's academic performance.

Artificial Neural Network, Higher Education, Dropout

Citación: CENTURIÓN-CARDEÑA, Humberto José, CANO-BARRÓN, Danice Deyanira, SANDOVAL-GÍO, Jesús y ZAPATA-GONZÁLEZ, Alfredo. Redes neuronales para predecir el abandono académico en ingeniería. Revista de Tecnología y Educación. 2020. 4-11: 1-7

* Correspondencia del Auto (Correo electrónico: humberto.centurion@itsmotul.edu.mx)
† Investigador contribuyendo como primer autor.

Introducción

En la actualidad con la gran cantidad de datos disponibles, las organizaciones están enfocándose en como explotar esos datos para tener una ventaja competitiva. Si a esto se le aúna que el poder computacional, que las redes ahora son ubicuas y que los algoritmos que se han desarrollado permiten conectar conjuntos de datos permite el análisis más amplio y profundo. La unión de todos estos fenómenos ha propiciado la aplicación de los principios de ciencia de datos y de las técnicas de minería de datos (Provost & Fawcett, 2013).

Las estrategias orientadas no sólo al almacén eficiente de datos sino a la búsqueda de patrones no evidentes entre ellos se han convertido en un área importante de desarrollo tecnológico, esto se debe al acceso a grandes volúmenes de información que se almacenan en bases de datos centralizadas y distribuidas en diversos dominios, por lo que se considera de gran importancia para interpretar la información y el conocimiento distribuidos por todo el mundo (Riquelme, Ruiz, & Gilbert, 2006). De aquí que el reto principal actualmente consista en ser capaz de trabajar con grandes volúmenes de información y a través de técnicas y herramientas adecuadas, analizarlos para identificar patrones y extraer conocimiento novedoso y útil para aquellos que toman decisiones, para llevar a cabo esta tarea se recurre a la ciencia de datos.

La minería de datos educativa (*Educational Data Mining*), área en la que se enmarca este trabajo, surge como un paradigma orientado al diseño de modelos, tareas, métodos y algoritmos para explorar datos de entornos educativos buscando encontrar patrones y hacer predicciones que caractericen el comportamiento y logro de los estudiantes, el dominio del conocimiento, evaluación funcionalidad y aplicaciones educativas en entornos convencionales, abiertos y a distancia (Luan, 2002).

Es decir, busca establecer métodos y procesos específicos para ambientes educativos que permitan la toma de decisiones efectiva y eficiente en una amplia gama de interacciones y fenómenos que suceden dentro y fuera del aula.

En general, la predicción del éxito de los estudiantes resulta crucial para las instituciones de educación superior debido a que la calidad del proceso de enseñanza y aprendizaje está fuertemente relacionada con la habilidad de responder a las necesidades de formación de los estudiantes (Al-Twijri & Noamanb, 2015). Aunque se pueden encontrar diversos esfuerzos orientados a desarrollar modelos teóricos para explicar los factores que influyen, la retención, permanencia y las tasas de graduación se mantiene con poco cambio a lo largo de décadas, por lo que resulta una característica que una proporción significativa de estudiantes abandonan la universidad en el primer año.

Una de las principales motivaciones de este trabajo es la necesidad de poder enfocar el trabajo con los estudiantes en etapas tempranas de su formación, este modelo pretende poder identificar a los estudiantes con propensión al abandono escolar tan pronto como se inscriban a la institución, esto permitirá no sólo el poder prestarles atención especial sino la creación de programas de apoyo diseñados específicamente a sus necesidades

Este estudio se enfoca en el análisis de los datos de contexto y los resultados de la evaluación de habilidades del EXANI-II, que es un instrumento estándar utilizado en México para identificar el estado académico y personal de los aspirantes a una institución educativa de nivel superior. La información se obtiene al momento que los estudiantes se inscriben a su prueba. Los datos funcionarán como elementos descriptivos de aspectos personales y académicos que faciliten el poder predecir la probabilidad de que el aspirante abandone sus estudios. Este trabajo se enfoca en los estudiantes de nuevo ingreso del Instituto Tecnológico Superior de Motul de las generaciones 2015 al 2018.

Este modelo se basará en los datos proporcionados por las generaciones de estudiantes que ingresaron durante el 2015 y hasta el 2018 para predecir el comportamiento de las generaciones venideras. Una de las bondades de esta herramienta, es que con cada generación que ingrese se puede ajustar el modelo una vez concluido el primer año de estudio, lo que permite ir refinando el modelo para las nuevas generaciones.

ISSN 2523-0360

Situación Problemática

En México para estudiantes de ingeniería, de acuerdo con la directora de planeación y evaluación del Tecnológico Nacional de México sólo el 58% de los que ingresan logran terminar con la carrera, lo que implica un gasto de 30 mil pesos promedio por estudiante (Poy, 2017). Para este subsistema que para el 2018 tenía una matrícula de 591,771 estudiantes, la pérdida económica y de personas es inmensa, por lo que poder contar con sistemas que monitoricen de manera automática el riesgo resulta vital para poder actuar en etapas tempranas.

En particular, esta investigación se desarrolla en el Instituto Tecnológico Superior de Motul, perteneciente al Tecnológico Nacional de México. El instituto inició sus operaciones el 18 de septiembre de 2000 con 2 carreras, 67 alumnos y 7 profesores, en la actualidad se cuenta con 5 programas educativos, más de 850 estudiantes y una planta docente de más de 40 profesores que atienden las necesidades formativas de manera permanente.

Debido al crecimiento de su matrícula y a la diversidad de necesidades de los estudiantes de nuevo ingreso se han implementado una serie de actividades de apoyo para facilitar no sólo su ingreso sino principalmente su egreso y titulación como parte de las funciones sustantivas de la organización, para mantener e inclusive mejorar sus indicadores de calidad.

Los programas de apoyo a su formación incluyen cursos de formación inicial, asesorías académicas para los que lo requieran, atención a estudiantes en situación de riesgo, etc., sin embargo se ha detectado la necesidad de identificar desde el ingreso aquellos elementos que puedan ayudar a identificar a los estudiantes que requieran de apoyos específicos como un medio de trabajo inicial con los profesores y tutores en conjunto para hacer un seguimiento oportuno de los casos y más adelante poder predecir el comportamiento durante su proceso formativo lo que permitiría tomar acciones preventivas más que reactivas como se acostumbra.

Objetivo

El objetivo general de este trabajo es modelar una red neuronal de retropropagación que permita predecir el abandono escolar en estudiantes de ingeniería con base sus perfiles socioeconómicos, académicos y personales.

El propósito de este estudio es poder generar un modelo predictivo que facilite la identificación temprana de estudiantes en situación de riesgo de abandono escolar que permita a la administración tomar decisiones más informadas y eficientes.

Este modelo podría incluirse en la plataforma institucional de tutorías que centraliza el monitoreo del desarrollo académico de los estudiantes de manera que la información llegara de manera directa a los responsables y puedan monitorear de manera específica los casos identificados.

Red Neuronal Artificial

Las redes neuronales actualmente están dando resultados inesperados en temas como reconocimiento de imágenes, escritura manual, comprensión de textos, segmentación de imágenes, manejo autónomo de vehículos y muchas otras cosas (McClure, 2017). Es por ello que su uso es cada vez más extendido para una gran cantidad de problemas relacionados con la predicción.

Las redes neuronales son una forma de emular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. Matich (2001) las describe como "un sistema para el tratamiento de la información, cuya unidad básica se inspira en modelos biológicos" (pág. 2). En este sentido destaca sus principales ventajas que se encuentran son un aprendizaje adaptativo, auto organizado, tolerancia a fallos, operación en tiempo real y de fácil inserción dentro de la tecnología existente.

Finalmente, las redes neuronales artificiales son modelos no lineales para clasificación y regresión que usan diversas estrategias para sobreponerse a las limitaciones de los perceptrones. Son descritas por tres componentes: la arquitectura, la función de activación y el algoritmo de aprendizaje.

ISSN 2523-0360

La arquitectura o topología describe el número de capas de neuronas y las conexiones entre ellas. La función de activación es el que describe el tipo de salida esperado. El algoritmo de aprendizaje es el que trabaja en encontrar los pesos óptimos (Hackeling, 2014).

Metodología a desarrollar

Para el desarrollo de esta investigación se utilizará la metodología CRISP-DM (del inglés Cross Industry Standard Process for Data Mining) que establece una descripción del ciclo de vida de un proyecto estándar de análisis de datos cubriendo las fases, tareas respectivas y las relaciones entre estas tareas (véase la Figura 1). La metodología CRISP-DM contempla el proceso de análisis de datos como un proyecto profesional, considera el hecho de que el proyecto no acaba una vez que se halla el modelo idóneo, sino que está relacionado con otros proyectos, y es preciso documentarlo de forma exhaustiva para que otros equipos de desarrollo utilicen el conocimiento adquirido y trabajen a partir de él (Chapman, Khabza, & Shearer, 2000).

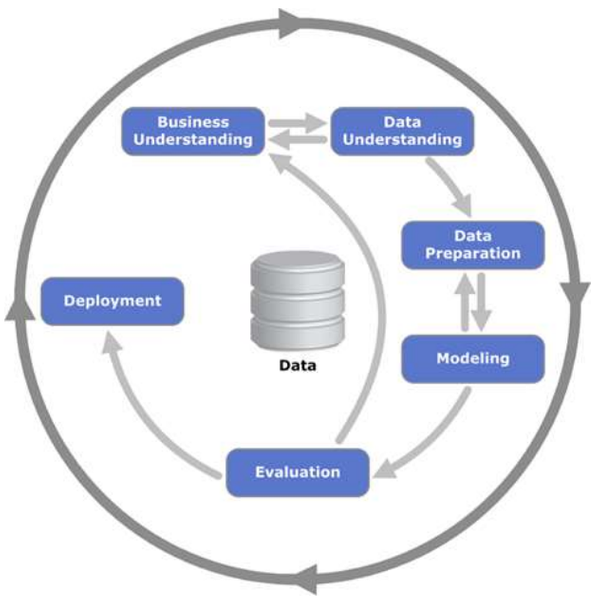

Figura 1 Ciclo de vida del modelo CRISP-DM
Fuente: (Chapman, Khabza, & Shearer, 2000)

En la primera fase, entendimiento del negocio, se busca comprender las necesidades de la organización y las motivaciones que dan origen al proyecto para poder establecer parámetros como la precisión requerida del modelo, las fuentes de información de las que dispone, entre otras cosas.

Una vez identificada las fuentes de información se procede a revisar con detenimiento la calidad de dato que almacena, se generar y/o analizan diccionarios de datos para poder determinar de manera apropiada el tratamiento que se dará a la información. Para el análisis, preparación y modelado de datos se utilizó Python y R que son dos herramientas de uso libre que permiten el tratamiento y el análisis de grandes volúmenes de datos, además de contar con herramientas que facilitan la generación de gráficas orientadas a la visualización de la información de manera efectiva.

Una vez integradas las bases de datos a utilizar y como parte de la preparación de los datos, se requiere de seleccionar un número reducido de variables para mejorar el desempeño de las herramientas de predicción. Actualmente se considera vital el poder contar con un sistema que sea capaz de seleccionar los atributos más representativos para maximizar la exactitud de los procesos de clasificación (Malhi & Gao, 2004) no sólo por la gran cantidad de datos con los que se cuenta sino para evitar fenómenos como el *overfitting* que lejos de ayudar a tener un modelo más eficiente, complica el poder predecir ante nuevos escenarios. Por lo que para esta parte se llevó a cabo un análisis de componentes principales (PCA, de las siglas en inglés de *principal component analysis*) que tiene como objetivo reemplazar atributos redundantes con nuevos atributos que de manera adecuada integren la información contenida originalmente (Zheng & Casari, 2018).

Una vez integrada la base de datos final, existen tres aspectos importantes a considerar al diseñar el modelo de la Red Neuronal: el número de capas, el número de neuronas que habrá en cada capa y la tasa de aprendizaje de la red. A la fecha no existe una metodología definida que permita establecer cada uno de los parámetros del modelo, más bien se ajusta en función de la experiencia del investigador y en los resultados que se obtengan en cada una de las corridas. En cuanto al número de capas, existen algunas recomendaciones que suelen ser usadas como un primer acercamiento al modelado de las redes neuronales, de acuerdo con Heaton (2008) el número de capas que deben considerarse son una o dos, pues si es una puede aproximar cualquier función que mapee de manera continua de un espacio finito a otro, mientras que sin son dos puede aproximar cualquier mapeo con cierto nivel de eficiencia.

ISSN 2523-0360

Para este proyecto se consideró una capa ya que aunque se pretende un buen nivel de eficiencia, los resultados de predicción no son sensibles a los fallos. En cuanto al número de neuronas, se consideró la regla de tomar el promedio de número de entradas más la salida.

Uno de los procesos que se llevó a cabo para aligerar el costo computacional fue el del análisis de componentes principales (PCA de las siglas en inglés de *principal component analysis*) tiene como objetivo reemplazar atributos redundantes con nuevos atributos que de manera adecuada integren la información contenida originalmente (Zheng & Casari, 2018). Este método se enfoca en la noción de independencia lineal, es decir, la mayoría de los atributo son combinaciones de un número reducido de atributos clave, por lo que aquellos que dependen linealmente de ellos son un una pérdida de espacio y poder computacional y por lo tanto se deben de descartar.

Atributo	Clave	Importancia
1	rezago	81.4
2	prom_bac	14.2
3	dan_reqc	8.9
4	fecha_apli	7.5
5	ipan	7.5
6	ipma	7.2
7	iele	6.9
8	dan_malf	6.3
9	dan_eir	6
10	dan_ofi	5.9
11	icne	5.7
12	fre_cde	3.8
13	vac_rm	3.5
14	fre_tsc	3.4
15	dan_mft	3.3
16	icle	3.3
17	edad_ingreso	3
18	fre_sme	2.5
19	hrs_trab	2.5
20	ser	2.1

Tabla 1 Atributos considerados en el estudio y su importancia de acuerdo al PCA
Fuente: Elaboración Propia

De igual manera, debido a que el PCA únicamente enlista la importancia de cada una de las variables se generaron tres escenarios de datos con número de atributos diferentes: 20, 15 y 10, debido a la baja significancia que se pudo observar en las variables ya sea debido a su poca variación o a que varían demasidado (Véase table 1). Con base en estas cantidades de variables de entrada considerada, el número de neuronas ocultas se determinó de 10, 8 y 6, respectivamente.

Finalmente para la tasa de aprendizaje (*learning rate*) que es la cantidad con la que se actualizan los pesos de la red en cada uno de los pasos, se trata de un valor positivo entre 0 y 1 (Goodfellow, Bengio, & Courville, 2016, pág. 86). El valor más común que se le da es el de 0.01, pero pueden ser considerados valores más pequeños como el 0.001, ambos fueron considerados para la calibración de la neurona además del 0.05.

Como puede observarse en la Tabla 2 el proceso completo del diseño de la calibración de la red neuronal que se consideró para este trabajo incluye 3 tasas de aprendizaje, 3 cantidades de variables de entrada y por consiguiente 3 cantidades de neuronas en la capa oculta.

Tasa de aprendizaje	Atributos de entrada	Mínimo de neuronas en la capa oculta
.01	20	11
	15	8
	10	6
.001	20	11
	15	8
	10	6
.05	20	11
	15	8
	10	6

Tabla 2 Número de capas y neuronas en la red neuronal
Fuente: Elaboración Propia

Finalmente, antes de comenzar el proceso de entrenamiento se debe de seleccionar el tipo de umbral que servirá para determinar la convergencia del entrenamiento. Para este caso se utilizó la función sigmoidea o logística ya que sus valores se encuentran en un rango comprendido entre 0 y 1. Aplicando esto a una red neuronal significa que sea cual sea la entrada, la salida estará comprendida entre 0 y 1, lo cual se adecua a la salida esperada (0 si no es propenso el estudiante a abandonar, 1 si lo es). La ecuación que describe el comportamiento de la función es (1).

$$P(t)= 1/(1+ e^{(-t)}) \quad (1)$$

En concreto para este estudio, la red neuronal que se implementó es tipo retropropagada (*backpropagation*) que utiliza a la función sigmoidea como función de activación con 30,000 procesos de aprendizaje para todos los escenarios que se plantearon.

ISSN 2523-0360

Resultados

Una vez decididos los parámetros para las pruebas, se generaron las bases de datos y se corrieron los procesos de aprendizaje, de igual manera se determinó que la base de datos de prueba que consistiría en el 15% de la base de datos original cuyos registros se seleccionan de manera aleatoria. Es importante destacar que la selección de registros se mantiene la misma base de datos para el entrenamiento y para las pruebas del modelo ya calibrado.

En la Tabla 3 se pueden observar los resultados de la implementación de los diferentes elementos de calibración de la neurona considerados para este estudio en términos de la precisión con la que la neurona puede clasificar correctamente a los estudiantes y la matriz de confusión correspondiente. Los mejores resultados en términos de precisión se obtienen con 15 o 10 atributos de entrada con una tasa de aprendizaje de .01 y 10 atributos con una tasa de aprendizaje de .05, por lo que se analizarán sus matrices de confusión ya que logra clasificar la mayor cantidad correcta de elementos (se suman los elementos en negros para determinar cuántos están bien clasificados en cada escenario).

Tasa de aprendizaje	Atributos de entrada	Matriz de confusión		
			0	**1**
.001	20	**0**	**61**	27
		1	7	**23**
	15	**0**	**72**	4
		1	39	**3**
	10	**0**	**88**	1
		1	29	**0**
.01	20	**0**	**48**	36
		1	9	**25**
	15	**0**	**61**	24
		1	6	**27**
	10	**0**	**67**	23
		1	6	**22**
.05	20	**0**	**61**	21
		1	9	**27**
	15	**0**	**53**	28
		1	10	**27**
	10	**0**	**58**	31
		1	3	**26**

Tabla 3 Calibración de la red neuronal
Fuente: Elaboración Propia

De acuerdo con la exactitud que se presenta en la Tabla 4, los modelos a considerar son los cuatro marcados en negro cerca del 75%, siendo el más eficiente el de 10 neuronas con una tasa de aprendizaje de .01.

Sin embargo al revisar el indicador F que pondera el número de verdaderos aciertos con aquellos casos en los que se mal clasificó un caso, los modelos más eficientes resultan ser el de 15 neuronas con una tasa de aprendizaje de .01 y el 20 neuronas con tasa de .05. Computacionalmente hablando, resulta más eficiente trabajar con un modelo que considera menos variables por lo que el más eficiente resultaría ser el de 15 neuronas

En conclusión el modelo elegido se caracteriza por una buena exactitud y una buena cantidad de elementos identificados correctamente, es decir logra identificar de mejor manera a los estudiantes propensos a abandonar sus estudios.

Tasa de aprendizaje	Atributos de entrada	Exactitud	Precisión	Recall	F
.001	20	0.7119	0.4600	0.7667	0.5750
	15	0.6356	0.4286	0.0714	0.1224
	10	**0.7458**	0.0000	0.0000	**0.0000**
.01	20	0.6186	0.4098	0.7353	0.5263
	15	**0.7458**	0.5294	0.8182	**0.6429**
	10	**0.7542**	0.4889	0.7857	**0.6027**
.05	20	**0.7458**	0.5625	0.7500	**0.6429**
	15	0.6780	0.4909	0.7297	0.5870
	10	0.7119	0.4561	0.8966	0.6047

Tabla 4 Eficiencia de la red neuronal
Fuente: Elaboración Propia

Conclusiones

El desarrollo de este proyecto relacionado con la predicción temprana de casos de abandono escolar resulta importante debido a la cantidad de recursos invertidos en cada uno de los estudiantes que ingresa a una institución educativa de nivel superior. Este proyecto se enfocó en el dar uso a información que se encuentra a disposición de las autoridades pero que raramente se utilizan para caracterizar a los estudiantes.

El poder seleccionar las variables más significativas para proceder a calibrar el modelo, que incluyó un proceso que incluía cambios en el número de variables y tasas de aprendizaje para determinar el modelo óptimo.

ISSN 2523-0360

Fue durante esta fase que se encontró que las variables del cuestionario de contexto no resultan significativas de manera decisiva para predecir el abandono, a pesar de ser de diferente naturaleza (académica, psicológica y económica), pero ciertamente tienen un nivel de eficiencia aceptable que permitiría identificar casos potenciales de abandono y tratarlos de manera temprana.

El modelo más eficiente resultó el de 15 variables con tase de aprendizaje de .01, con una efectividad del 74.58% y un índice F de .6429, lo que hace del modelo lo suficientemente sólido pero que requerirá de introducir nuevas variables buscando que la identificación pudiera ser más eficiente de manera que pudiera ayudar los interesados de manera más puntual. Es importante recordar que no se trata de un caso en donde la efectividad sea tan importante como en otros escenarios, por lo que permite a la organización a entender el comportamiento de sus estudiantes. Se espera poder montar este modelo de conocimiento en el Sistema Integral de Tutorías que el Instituto ha puesto en marcha en meses previos para proveer a los Tutores y Coordinación de Tutorías un panorama académico esperado de los estudiantes y poder hacer el seguimiento más eficiente y puntual de aquellos casos que resulten más propensos a abandonar.

En cuanto a las recomendaciones, se considera agregar preguntas que permitan determinar cambios en el contexto de los estudiantes como si contraen matrimonio, tienen hijos, empiezan a trabajar, entre otras cosas para poder determinar su impacto en sus estudios. Se espera poder agregar este cuestionario en la plataforma antes mencionada para que los estudiantes puedan responderla al inicio de cada semestre y poder correr el modelo de nuevo para identificar nuevos casos potenciales de abandono. De igual manera es importante hacer crecer la base de datos de registros para poder entrenar el modelo de mejor manera, ya que aunque se consideraron 4 generaciones se pierden casos por no contar con la información completa debido al proceso de ingreso que contempla la institución.

Referencias

Al-Twijri, M., & Noamanb, A. (2015). A New Data Mining Model Adopted for Higher Institutions. *Procedia Computer Science*(65), 836 – 844.

Chapman, P., Khabza, T., & Shearer, C. (2000). *CRISP-DM 1.0: Step by step dataminig guide.* Obtenido de The modeling agency: https://www.the-modeling-agency.com/crisp-dm.pdf

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* Londrés: MIT Press.

Hackeling, G. (2014). *Mastering Machine Learning with scikit-learn.* Birmingham: Packt Publishing Ltd.

Heaton, J. (2008). *Introduction to neural networks with Java* (Segunda ed.). Chesterfield, Estados Unidos: Heaton Research Inc.

Luan, J. (2002). Data mining and its applications in higher education. *Journal of New Directions for institutional Research*, 17-36.

Malhi, A., & Gao, R. (2004). PCA-Based Feature Selection Scheme for Machine Defect Classification. *IEEE Transactions on Instrumentation and Measurement, 53*(6), 1517-1525.

Matich, D. (marzo de 2001). *Universidad Tecnológica Nacional.* Obtenido de Facultad Regional del Rosario: https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monograias/matich-redesneuronales.pdf

McClure, N. (2017). *TensorFlow Machine Learning Cookbook.* Birmingham: Packt Publishing Ltd.

Poy, L. (3 de junio de 2017). Sólo el 58% concluye estudios en el Tecnológico Nacional de México. *La Jornada*.

Provost, F., & Fawcett, T. (2013). Data Science for Business. Sebastopol: O'Reilly Media, Inc.

Riquelme, J., Ruiz, R., & Gilbert, K. (2006). Minería de Datos: Conceptos y Tendencias. *Revista Iberoamericana de Inteligencia Artificial, 10*(10), 11-18.

Zheng, A., & Casari, A. (2018). *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientist.* Sebastopol: O'Reilly.

ISSN 2523-0360