

Construction of a model based on neural networks to determine the duration of Engineering studies in a public university in Peru

Paul Tocto, Doctor¹, Gloria Teresita Huamani, Doctora², Luis Zuloaga, Magister³,

¹Universidad Nacional de Ingeniería, Perú, ptocto@uni.edu.pe

²Universidad Nacional de Ingeniería, Perú, gloria.huamani.h@uni.edu.pe

³Universidad Nacional de Ingeniería, Perú, zuloaga_luis@uni.edu.pe

Abstract- Public universities face various problems, including undergraduate length beyond five years and student dropouts, these affect the prestige and the budget of the university. To predict, planners apply classical statistical analysis techniques, however there is a trend to improve the prediction using Machine Learning techniques. In this research, a supervised regression neural network model is created to make the prediction of the duration of the years of studies. Neural networks were trained and validated with data from 2010 to 2020 of graduates of the Faculty of Industrial and Systems Engineering (FIIS), of the National University of Engineering (UNI). A model has been obtained with an error rate of less than 1%. This model contributes to better decision-making by the authorities.

Keywords—Machine Learning; Neural Network; Regression; undergraduate length, academic performance.

Digital Object Identifier (DOI):

<http://dx.doi.org/10.18687/LACCEI2021.1.1.397>

ISBN: 978-958-52071-8-9 ISSN: 2414-6390

Construcción de un modelo basado en redes neuronales para determinar la duración de los estudios de Ingeniería en una universidad pública en el Perú

Paul Tocto, Doctor¹, Gloria Teresita Huamaní, Doctora², Luis Zuloaga, Magister³,

¹Universidad Nacional de Ingeniería, Perú, ptocto@uni.edu.pe

²Universidad Nacional de Ingeniería, Perú, gloria.huamani.h@uni.edu.pe

³Universidad Nacional de Ingeniería, Perú, zuloaga_luis@uni.edu.pe

Resumen— Las universidades públicas se enfrentan a diversos problemas, entre ellos la duración de la licenciatura, más allá de cinco años, asimismo el abandono de los estudios, estos afectan el prestigio, el presupuesto de la universidad. Para predecir, los planificadores aplican técnicas de análisis estadísticos clásicos, sin embargo, hay una tendencia a mejorar la predicción utilizando técnicas de “aprendizaje automático”. En esta investigación se crea un modelo de red neuronal supervisada de regresión, para realizar la predicción de la duración de los años de estudios. Las redes neuronales se entrenaron y validaron con los datos del 2010 al 2020 de los egresados de la Facultad de Ingeniería Industrial y de Sistemas (FIIS), de la Universidad Nacional de Ingeniería (UNI). Se ha obtenido un modelo con una tasa de error menor al 1%. Este modelo contribuye a una mejor toma de decisiones por parte de las autoridades

Palabras clave. Aprendizaje automático, red neuronal, regresión, duración de los estudios, rendimiento académico

Abstract— Public universities face various problems, including undergraduate length beyond five years and student dropouts, these affect the prestige and the budget of the university. To predict, planners apply classical statistical analysis techniques, however there is a trend to improve the prediction using Machine Learning techniques. In this research, a supervised regression neural network model is created to make the prediction of the duration of the years of studies. Neural networks were trained and validated with data from 2010 to 2020 of graduates of the Faculty of Industrial and Systems Engineering (FIIS), of the National University of Engineering (UNI). A model has been obtained with an error rate of less than 1%. This model contributes to better decision-making by the authorities.

Keywords—Machine Learning; Neural Network; Regression; undergraduate length, academic performance.

I. INTRODUCCIÓN

La alta dirección universitaria debe tomar decisiones en un contexto cambiante (en este último año los dos periodos académicos han sido virtuales). Asimismo, existe una nueva ley universitaria (2014), que da mayor valor a la investigación y se exige obtener el grado de bachiller con una tesis antes de la titulación profesional. Para facilitar la pronta titulación se dictan cursos que coadyuvan al desarrollo de una tesis de investigación, también se promueve la formación de grupos de investigación y el desarrollo de proyectos formativos. Para

mejorar el rendimiento académico y disminuir el tiempo de permanencia de los estudiantes, se han hecho esfuerzos como el diseñar programas de tutoría para los denominados “estudiantes en riesgo”, quienes han desaprobado una misma materia por dos veces y que, según la nueva ley, si son desaprobados por una tercera vez se toma la decisión de su separación temporal y de persistir en ello una separación definitiva de la universidad. Un gran porcentaje es considerado estudiante regular, quién solo se dedica a los estudios. Sin embargo, no se tiene patrones de comportamiento de los estudiantes para lograr terminar en 5 años una carrera de ingeniería, tal como se establece en la oferta universitaria. Por ello, es necesario determinar patrones e identificación de variables más relevantes como son los sociodemográficos, el rendimiento académico, la complejidad de las materias, entre otros, a partir de la data histórica.

Según la Oficina de Registro Central y Estadísticas ORCE (2019-2) [1], en el segundo semestre del año 2019, la UNI tenía una población de 12,066, distribuida en 11 facultades. Asimismo, de estas facultades, la Facultad de Ingeniería Industrial y de Sistemas de la UNI tiene una matrícula del 12.26%. Según edades, un 64.89% de los estudiantes UNI son menores a 25 años.

El estudio cobra mayor importancia por ser la primera casa de estudios del Perú dedicada a la formación en ingeniería y la oferta de vacantes para los nuevos ingresantes es relativamente pequeña, pese a haber crecido en infraestructura y número de profesores, la oferta de nuevas especialidades requiere de un mayor presupuesto. En esas condiciones determinar el número de años de permanencia para terminar la carrera universitaria de ingeniería industrial y de ingeniería de sistemas resulta un trabajo de interés como apoyo a las decisiones de la dirección universitaria, quienes requieren datos que a su vez se convierten en conocimiento.

Con mayor frecuencia se introduce en la gestión académica las buenas prácticas de la inteligencia de negocios, inteligencia artificial, y Machine Learning. Para lo cual se van identificando soluciones en base a plataformas tecnológicas, tal como afirma Arce [2] “En un mundo basado en datos, la capacidad de integrar y analizar datos de manera confiable y precisa es primordial, es lo que posiciona a las organizaciones delante de la competencia. En Prediq, diseñamos soluciones basadas en la plataforma analítica Qlik para ayudar a nuestros clientes a transformar datos en

conocimientos clave que permiten una toma de decisiones mejor informada, lo que da como resultado un negocio más saludable y rentable”, así como las recomendaciones de Tegmark [3], “De lo único que estoy seguro es de que será lo mejor o lo peor que haya sucedido nunca. Si la tecnología va a avanzar, también debemos asegurarnos de que crezcan los conocimientos con los que vamos a manejar esa tecnología de forma adecuada. Los conocimientos deben ser lo primero” Aunado a Hattie (2009) citado por Visccher [4] es importante la formación de docentes en la gestión de base de datos para la toma de decisiones, “la formación del profesor en data-based decision making (DBM) podría ser crucial para mejorar el rendimiento de los estudiantes, ya que los profesores tienen un impacto considerable sobre el aprendizaje de los estudiantes.”

El uso de inteligencia artificial es una línea de investigación que se viene trabajando en la Facultad de Ingeniería Industrial y de Sistemas (FIIS) de la Universidad Nacional de Ingeniería (UNI). Algunos investigadores se dedican a la aplicación de modelos matemáticos en la gestión universitaria, a quienes les resulta fácil la implementación de proyectos de uso de datos. Tradicionalmente se realizan pronósticos bajo dos enfoques generales [5]. Un enfoque es el análisis cuantitativo; el otro es el enfoque cualitativo. Los pronósticos cuantitativos utilizan una variedad de modelos matemáticos que se apoyan en datos históricos y/o en variables causales para pronosticar la demanda o resultados. Los pronósticos cualitativos o subjetivos incorporan factores como la intuición, las emociones, las experiencias personales y el sistema de valores de quien toma las decisiones para llegar a un pronóstico. En la práctica, la combinación de ambos resulta más efectiva en la mayoría de los casos.

En este trabajo se ha utilizado el Portal de Microsoft Azure Machine Learning Studio, por su fácil manejo interactivo en la construcción de modelos, como afirman Tocto y Canelo [6] “Para el experimento se evaluó a los líderes en el servicio de procesamiento analítico de datos en la nube: Amazon AWS, IBM Watson Studio y Microsoft Azure Machine Learning Studio. Considerando a Microsoft por su fácil manejo interactivo en la construcción del modelo”.

El objetivo general de la investigación es crear un modelo de red neuronal supervisada de regresión, para realizar la predicción de la duración de los años de estudios. En el experimento se determina el tiempo de permanencia del estudiante desde que ingresa hasta que finaliza sus estudios aplicando la Red Neuronal de Regresión, los datos extraídos corresponden a un periodo de 11 años.

Las redes neuronales artificiales utilizadas en los modelos han demostrado tener buen potencial para la predicción.

Este artículo tiene la siguiente estructura: en la sección 2 se hace una revisión bibliográfica para dar cuenta del “aprendizaje automático”, el flujograma de entrenamiento de la red neuronal y la arquitectura de la red neuronal. En la sección 3, se explica la metodología utilizada. En la sección 4 se presenta el experimento y las pruebas. Por último, en la

sección 5 se presenta las conclusiones y recomendaciones para futuras investigaciones.

II. REVISIÓN DE LA LITERATURA

Para la predicción del tiempo de permanencia en la universidad, hay escasas referencias al respecto y las que se abordan, se focalizan principalmente en la deserción estudiantil. Por eso se hace referencia a algunas investigaciones que respaldan el uso de las redes neuronales artificiales para la predicción de otras situaciones.

La eficiencia del modelo basado en redes neuronales se mide con la tasa de error obtenido y la tasa de precisión. Se entrenan tres modelos de redes neuronales utilizando nueve esquemas de aprendizaje. “Los tres modelos neuronales difieren en la topología, y en particular en el número de neuronas de la capa oculta. Los nueve esquemas de aprendizaje difieren en las proporciones de casos para el entrenamiento de la red y la validación. Se determina el esquema de aprendizaje con el cual el modelo tiene la mejor performance” [7]. Por otra parte, en [8] se propone un método de descubrimiento del conocimiento utilizando redes neuronales basados en el enfoque del perceptrón multicapa (MLP), orientado al análisis del riesgo de crédito (CRA) de las pequeñas y medianas empresas (PYMES) en Turquía.

Sobre Aprendizaje automático y Red Neuronal. El aprendizaje automático, se realiza mediante datos históricos que permiten definir el modelo, mediante un número de iteraciones considerados en la configuración [9.10].

Entre las muchas ventajas de utilizar una red neuronal se incluyen: *Aprendizaje Adaptativo* por su capacidad de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial; *Autoorganización*, porque puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje; y la *Tolerancia a fallos*. “La destrucción parcial de una red conduce a una degradación de su estructura; sin embargo, algunas capacidades de la red se pueden retener, incluso sufriendo un gran daño”. [11]

Sin embargo, el aprendizaje de la red neuronal requiere seguimiento. Muy de acuerdo con [12], ya que el valor del aprendizaje automático es que le permite continuamente aprender de los datos y predecir el futuro. Este poderoso conjunto de algoritmos y modelos se está utilizando en todas las industrias para mejorar procesos y obtener información sobre patrones y anomalías dentro datos. Pero el aprendizaje automático no es un esfuerzo solitario, es un proceso de equipo que requiere científicos de datos, ingenieros de datos, analistas de negocios, y líderes empresariales para colaborar.

En la figura 1 se muestra la metodología usada, desde obtener los datos de los egresados de la FIIS (Facultad de Ingeniería Industrial y de Sistemas, de la Universidad Nacional de Ingeniería), desde hace 10 años atrás, que se sube al portal de Microsoft Azure Machine Learning Studio, a continuación se

identifican las variables que tienen correlación con la variable objetivo, luego normalizando los datos, se entrenan el modelo de red neuronal con n-neuronas de una capa (desde una neurona hasta cincuenta neuronas), con el 80% de los datos obtenidos y se validan con el 20% de los datos, seleccionando el modelo de n-nodos de menor error absoluto medio. Este flujograma corresponde a una capa a diferencia de la propuesta de [13].

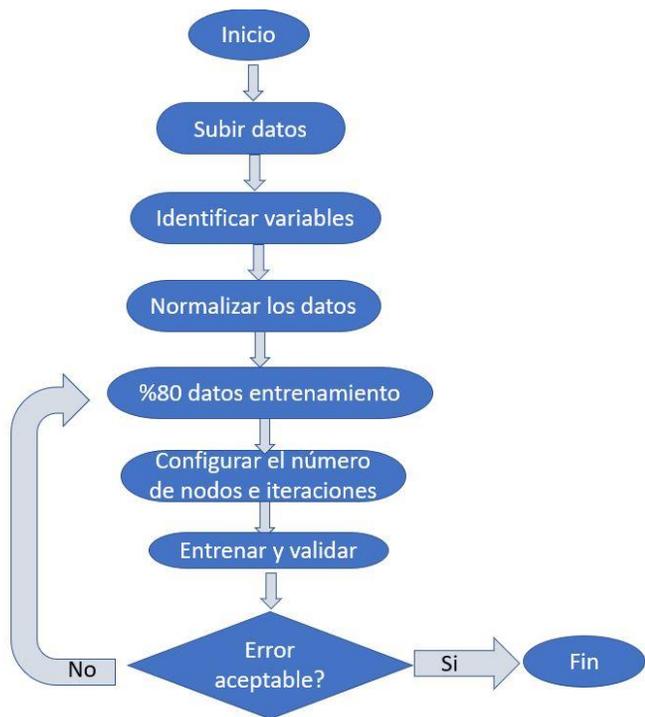


Fig. 1. Flujograma de entrenamiento de la red neuronal y validación

Con mayor frecuencia el uso de computación de la nube facilita los cálculos y almacenamiento, los predictores más importantes en la adopción de la computación en la nube, según señalan en [14] son: "confianza, el riesgo de seguridad de TI percibido y el estilo de gestión". Por otra parte, en cuanto a la metodología se toma en cuenta lo indicado en el Portal Microsoft Azure Machine Learning Studio [15]. Para nuestro experimento se resume en los siguientes pasos: Paso 1. Subir los datos al entorno Azure en la nube. Paso 2. Seleccionar las variables a considerar en el modelo, el cual permite identificar que variables tienen mayor correlación con la variable resultado de la evaluación. Con este criterio se seleccionan las variables que tienen mayor coeficiente de correlación.

Se utiliza una Red Neuronal de Regresión, por su estructura flexible y su distribución libre del manejo de los datos, que permite considerarlo como una potencial herramienta de predicción [16], de la duración en años de los estudios de un alumno en la FIIS. La arquitectura considerada es de una red

neuronal de una capa oculta, que es generalmente suficiente [17] (ver Fig. 2).

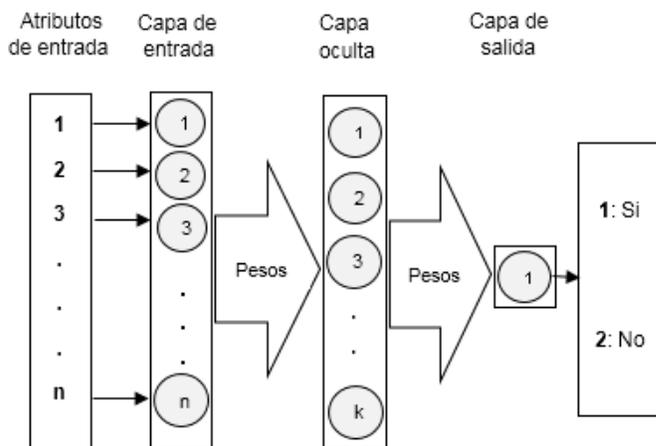


Fig. 2 Arquitectura de la red neuronal para hallar la duración de los estudios en la FIIS [6]

III METODOLOGÍA

Se aplica el flujograma de entrenamiento de la red neuronal (Fig.1) y Arquitectura de la red neuronal (Fig. 2) se describe la colección de datos, definición de variables, entrenamientos y validación de la red neuronal, para obtener el modelo de red neuronal con el menor error.

A. Colección de datos y definición de variables

En este estudio se considera los datos de los egresados en la FIIS desde el año 2010 hasta la actualidad, en las especialidades de Ingeniería Industrial y de Ingeniería de Sistemas. La base de datos está conformada por 1,787 registros de egresados. Luego de realizar la limpieza de los datos tenemos 1745 registros. La Tabla I especifica las 14 variables independientes y la variable dependiente "años de estudio".

TABLA I
VARIABLES DEL MODELO

Nro	Atributo	Descripción
1	ESPECIA	Especialidad (I1, I2)
2	SEXO	Sexo (MASCULINO, FEMENINO)
3	DEPNAC	Departamento de nacimiento
4	PROVNAC	Provincia de nacimiento.
5	DISTNAC	Distrito de nacimiento
6	TIPCOLEGIO	Tipo de colegio
7	DEPUBICOLE	Departamento de ubicación del colegio
8	PROVUBICOLE	Provincia de ubicación del colegio
9	DISTUBICOLE	Distrito de ubicación del colegio
10	DATOSLABOR	Datos laborales (No trabaja, Trabaja, Practicante)
11	DISTRESID	Distrito de residencia
12	VIVECONPADR	Vive con sus padres (SI, NO)
13	EDAD INICIAL	Edad al iniciar sus estudios
14	PROMEDIOPRIMER	Promedio de notas en el primer ciclo
15	AÑOS DE ESTUDIO	Años de estudio (Variable en estudio)

B. Construcción del modelo

Se crea el modelo mediante los siguientes pasos:

Paso 1. Subir los datos al entorno de Microsoft Azure Machine Learning Studio en la nube.

Paso 2. Utilizando el módulo “Filter Based Feature Selection”, sobre todos los datos, identificamos las variables que tienen mayor correlación con la variable resultado de la evaluación: AñosEstudio. La Fig.3 muestra las variables con la mayor correlación.

AñoEstudio	PROMEDIOPRIMER	AñoNac	AñoInic	EdadInic	ESPECIA	SEXO
1	0.316922	0.125315	0.122995	0.095656	0	0

Fig. 3 Coeficiente de correlación de las variables

Paso 3. Se normaliza los datos de la variable AñosEstudio, extremos mayores y menores, remplazándolos por el promedio de los datos, utilizando el módulo “Clip Values”, considerando (Valor mínimo 4 y Valor máximo 11). La Fig. 4 muestra el histograma de dicha variable dependiente. La Fig. 5 muestra los estadísticos para esta variable.

La Tabla II contienen los estadísticos para las 4 variables que presenta el mayor coeficiente de correlación con la variable dependiente AñosEstudio, los que son usados en la red neuronal.

AñosEstudio
Histogram

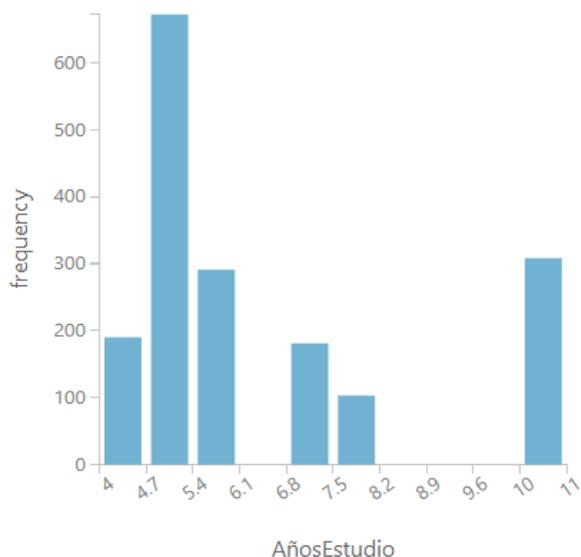


Fig. 4 Histograma de la variable AñosEstudio

Statistics

Mean	6.5014
Median	6
Min	4
Max	11
Standard Deviation	2.3075
Unique Values	6
Missing Values	0
Feature Type	Numeric Feature

Fig. 5 Estadísticos de la variable AñosEstudio

TABLA II
ESTADÍSTICOS DE LAS VARIABLES

Estadístico	Variable			
	PROMEDIO PRIM	AÑONAC	AÑOINIC	EDADINIC
Mean	11.5576	1984.6315	2003.275	23.1862
Median	12	1990	2008	18
Min	2	5	11.634	14
Max	22	1999	2016	1992
Std Dev	1.7374	95.036	95.5419	94.2927
Unique Val	18	37	36	23
Missing	0	0	0	0
Val Feature type	Numeric Feature	Numeric Feature	Numeric Feature	Numeric Feature

Paso 4. Se normalizan todos los datos numéricos utilizando el módulo “Normalize Data”, mediante el método de transformación ZScore, ver Fig. 6.

Normalize Data

Transformation method

Use 0 for constant ...

Columns to transform
Selected columns:
Column type: Numeric, All

Fig. 6 Normalización de los datos

Paso 5. Se separa el 80% de los datos para la fase de entrenamiento mediante el módulo “Split Data”, ver Fig. 7.

Split Data

Splitting mode

Fraction of rows in the f...

Randomized split

Random seed

Stratified split

Fig. 7 Separación de datos para entrenamiento 80%

Paso 6. Se configura el módulo de “Neural Network. Regression”, con 27 nodos ocultos y 100 iteraciones como se muestra en la Fig. 8.

Neural Network Regression

Create trainer mode

Hidden layer specification

Number of hidden nodes

Learning rate

Number of learning iter...

The initial learning wei...

Fig. 8 Configuración del módulo “Neural Network Regression” con 27 nodos ocultos y 100 iteraciones.

Paso 7. Se realiza el entrenamiento utilizando el módulo “Train Model” que utiliza los datos de entrenamiento del paso 5 y el módulo configurado del paso 6.

Paso 8. Se valida el modelo con los datos de verificación, que son el 20% del total, utilizando el módulo “Score Model”.

Paso 9. Se evalúa el modelo con el módulo “Evaluate Model”.

IV RESULTADOS

Se muestra lo que se obtiene en la experimentación, al momento de crear los modelos desde un nodo hasta cincuenta nodos neuronales.

A. Experimentación

El experimento se realiza utilizando los parámetros definidos en el paso 4, con 80% de datos para el entrenamiento, y variando los nodos ocultos desde 1 hasta 50, obteniendo los mejores resultados con los siguientes números de nodos: 18, 23, 16, 10 y 14. Ver figuras 9, 10, 11, 12 y 13.

Nodos Internos	18
Mean Absolute Error	0.57120902
Root Mean Squared Error	0.71686772
Relative Absolute Error	0.73572796
Relative Squared Error	0.55542858
Coefficient of Determination	0.44457142
Tiempo de entrenamiento	0:01:01.161

Fig. 9 Resultado de la evaluación para 18 nodos ocultos y 25,000 iteraciones en el modelo (Valor mínimo 4 y Valor máximo 11) y con 80% de los datos para entrenar

Nodos Internos	23
Mean Absolute Error	0.57127448
Root Mean Squared Error	0.7170477
Relative Absolute Error	0.73581227
Relative Squared Error	0.55570752
Coefficient of Determination	0.44429248
Tiempo de entrenamiento	0:01:03.550

Fig. 10 Resultado de la evaluación para 23 nodos ocultos y 25,000 iteraciones en el modelo (Valor mínimo 4 y Valor máximo 11) y con 80% de los datos para entrenar

Nodos Internos	16
Mean Absolute Error	0.5713524
Root Mean Squared Error	0.71680704
Relative Absolute Error	0.73591264
Relative Squared Error	0.55533456
Coefficient of Determination	0.44466544
Tiempo de entrenamiento	0:00:58.331

Fig. 11 Resultado de la evaluación para 16 nodos ocultos y 25,000 iteraciones en el modelo (Valor mínimo 4 y Valor máximo 11) y con 80% de los datos para entrenar

Nodos Internos	10
Mean Absolute Error	0.57145445
Root Mean Squared Error	0.71640732
Relative Absolute Error	0.73604407
Relative Squared Error	0.55471538

Coefficient of Determination	0.44528462
Tiempo de entrenamiento	0:01:04.486

Fig. 12 Resultado de la evaluación para 10 nodos ocultos y 25,000 iteraciones en el modelo (Valor mínimo 4 y Valor máximo 11) y con 80% de los datos para entrenar

Nodos Internos	14
Mean Absolute Error	0.5717183
Root Mean Squared Error	0.7169672
Relative Absolute Error	0.7363839
Relative Squared Error	0.5555828
Coefficient of Determination	0.4444172
Tiempo de entrenamiento	0:00:59.065

Fig. 13 Resultado de la evaluación para 14 nodos ocultos y 25,000 iteraciones en el modelo 1 (Valor mínimo 4 y Valor máximo 11) y con 80% de los datos para entrenar

B. ANÁLISIS DE RESULTADOS

Los indicadores de evaluación del modelo de regresión son los siguientes: Error absoluto medio, raíz del error cuadrático medio, error absoluto relativo, error cuadrático relativo y coeficiente de determinación. El “error” representa la diferencia entre el valor predicho y el valor real obtenido. El coeficiente de determinación es un indicativo del nivel de variabilidad explicada en los datos, esta métrica evalúa qué tan bien un modelo explica y predice los resultados futuros, esta métrica nos indica el rendimiento predictivo del modelo, es decir, permite medir la exactitud del modelo, con los resultados obtenidos el modelo que tiene mejor grado de predictibilidad es el Modelo, que considera 18 nodos ocultos y 25,000 iteraciones, con un 80% de los datos para entrenar y 20% de los datos para validar, con un error absoluto medio de 0.57120902 y con un coeficiente de determinación de 0.44457142, ver Figura 9.

V. CONCLUSIONES

A. Conclusiones

- Se creó un modelo de red neuronal de regresión con un error de menos de 1%,
- Las variables que más influyen en la duración de los estudios son el desempeño del estudiante en el primer ciclo, Año de nacimiento, Año de inicio de sus estudios y su edad.
- Con este modelo se podría identificar a los alumnos que tendrían una alta probabilidad de atraso en sus estudios y tomar acciones de prevención como las tutorías y asesorías.

B. Trabajo a futuro

- Aprovechar la disponibilidad de herramientas de Machine Learning en la nube, para crear modelos que sirvan a las organizaciones.
- Considerar el uso de esta herramienta para crear modelos que detecten los futuros abandonos de los estudios, por parte de los estudiantes.
- Comprobar [18] Los estudiantes bajo la modalidad e-learning son proclives a la deserción “se encontró que sí es posible construir modelos predictivos exitosos y que el algoritmo que produjo los mejores modelos fue una red neuronal en tres de los cuatro cursos”.
- Predecir el comportamiento de los ingresantes, haciendo un estudio de los estudiantes de secundaria con predisposición a estudiar ingeniería [19]
- Considerar otros parámetros relacionados a su preparación, antes del ingreso a la universidad.

REFERENCIAS

- [1] Oficina de Registro Central y Estadísticas ORCE (20192)
- [2] Arce, Ricardo. PREDIQT. “América sistemas” noticiero digital 1008 10 febrero 2021 gerencia@americasistemas.com.pe
- [3] Tegmark, Max. La IA puede ser lo mejor o lo peor que haya sucedido nunca. TELOS, “https://telos.fundaciontelefonica.com/max-tegmark-la-ia-puede-ser-lo-mejor-o-lo-peor-que-haya-sucedido-nunca/”
- [4] A. van der Scheer^{1,2} and Adrie J. Visscher Effects of a Data-Based Decision-Making Intervention for Teachers on Students’ Mathematical Achievement Emmelien
- [5] JAY Heyzer y Barry RENDER. Principios de administración de operaciones. Séptima edición PEARSON EDUCACIÓN, México, 2009 ISBN: 978-607-442-099-9 Área: Administración y economía Formato: 21 27 Páginas: 752 p. 107
- [6] Canelo, César y Tocto Paul, 2018. Design of a model based on neural networks in the cloud to determine the credit risk in microcredit applicants in the financial market of Peru. In 16th LACCEI International Multi-Conference for Engineering, Education, and Technology: “Innovation in Education and Inclusion”, 19-21 July 2018, Lima, Peru.
- [7] Adnan Khashman, “Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes”, Expert Systems with Applications 37, 2010, pp. 6233–6239.
- [8] Gülnur Derelioglu & Fikret Gürgen, “Knowledge discovery using neural approach for SME’s credit risk analysis problem in Turkey”, Expert Systems with Applications 38, 2011, pp. 9313– 9318.
- [9] Rubiolo, Mariano. Desarrollo de nuevos modelos y algoritmos basados en redes neuronales para tareas de minería de datos. Argentina [Tesis doctoral] 2014
- [10] Santín Daniel. Eficiencia técnica y redes neuronales: un modelo para el cálculo del valor añadido en educación. Universidad Complutense de Madrid. [Tesis doctoral] 2013
- [11] Matich, Damián (2001) Redes Neuronales: Conceptos Básicos y Aplicaciones. Rosario. Universidad Tecnológica Nacional.
- [12] Hurwitz, J and Kirsh, D. Machine learning for dummies. IBM. Wileys and sons, Inc..2018
- [13] Siyi Wang, Yunpeng Zhang, Chen Zhang, Ming Yang. Improved artificial neural network method for predicting photovoltaic output performance. Global Energy interconnection. Vol 3 Dic. 2020.
- [14] Priyadarshinee P., Raut R., Kumar M., Gardas B., “Understanding and predicting the determinants of cloud computing adoption: A two staged hybrid SEM - Neural networks approach”, Computers in Human Behavior, Nov 2017, pp. 341-362
- [15] Portal Microsoft Azure Machine Learning Studio. Microsoft Azure Machine Learning Studio (classic) (azureml.net)
- [16] Sutradhar R., Barbera L., Comparing an Artificial Neural Network to Logistic Regression for Predicting ED Visit Risk Among Patients With

- Cancer: A Population-Based Cohort Study, *Journal of Pain and Symptom Management*, Volume 60, Issue 1, 2020, Pages 1-9, ISSN 0885-3924, <https://doi.org/10.1016/j.jpainsymman.2020.02.010>.
- [17] Dreiseitl, S., Ohno-Machado, L., Logistic regression and artificial neural network classification models: a methodology review, *Journal of Biomedical Informatics*, Volume 35, Issues 5–6, 2002, Pages 352-359, ISSN 1532-0464, [https://doi.org/10.1016/S1532-0464\(03\)00034-0](https://doi.org/10.1016/S1532-0464(03)00034-0). (<https://www.sciencedirect.com/science/article/pii/S1532046403000340>)
- [18] Urteaga, I, Siri, L., y Garófalo, G. (2020). Predicción temprana de deserción mediante aprendizaje automático en cursos profesionales en línea. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(2), pp. 147-167. doi: 10.5944/ried.23.2.26356
- [19] Cetinkaya, A., Kaan, Omer. Prediction of middle school students' programming talent using artificial neural networks. *Engineering Science and Technology, an International Journal* 23 (2020) 1301-1307.