**ANÁLISIS COMPARATIVO DE LA CLASIFICACIÓN DE UNIVERSIDADES GLOBALES UTILIZANDO REDES NEURONALES**

Oscar Alcides Choquehuallpa Hurtado(1)

161139@unamba.edu.pe(1)

Eli Serrano Pérez(2)

121183@unamba.edu.pe(2)

Alexander Chipa Cárdenas(3)

[162141@unamba.edu.pe](mailto:162141@unamba.edu.pe)(3)

Jonathan Bustinza Mendoza(4)

[162134@unamba.edu.pe](mailto:162134@unamba.edu.pe)(4)

Marcial Salazar Cahuana(5)

121179[@unamba.edu.pe](mailto:162134@unamba.edu.pe)(5)

Nilton Rojas Carrasco(6)

[121178@unamba.edu.pe(6)](mailto:121178@unamba.edu.pe(6))

**COMPARATIVE ANALYSIS OF THE CLASSIFICATION OF GLOBAL UNIVERSITIES USING NEURAL NETWORKS**

**RESUMEN**

El siguiente proyecto “ranking de universidades” se desarrolla con la finalidad de determinar en la actualidad que universidad a nivel mundial se considera la mejor, respecto a todas las características que influyen para su determinada clasificación en cual los integrantes de este grupo tomaron un data set de tipo “csv” del periodo 2017-2022. lo cual ya teniendo este archivo se realizó los diferentes procedimientos con el lenguaje de Python como limpieza de datos, categorización, normalización y creación de capas. para el entrenamiento de redes neuronales en cual se consideró los diferentes campos research output, internacional students, type y score.

De esta forma obteniendo un determinado gráfico en cual se pudo predecir la posible universidad con mejor ranking a futuros años

**PALABRA CLAVE**

Inteligencia Artificial, redes neuronales, ranking de universidades, Predicción, Pronóstico

**ABSTRACT**

The following project "ranking of universities" is developed with the purpose of determining at present which university worldwide is considered the best, with respect to all the characteristics that influence its certain classification in which the members of this group took a dataset of type "csv" for the period 2017-2022. which, having this file, the different procedures were carried out with the python language such as data cleaning, categorization, normalization and creation of layers. for the training of neural networks in which the different fields research output, international students, type and score were considered. In this way, obtaining a certain graph in which the possible university with the best ranking in future years could be predicted.

Keywords: Artificial intelligence, neural networks, university ranking, Prediction, Forecast

**INTRODUCCIÓN**

el análisis del desempeño y la evaluación comparativa de las universidades han atraído un creciente interés de los investigadores, los encargados de formular políticas y los medios de comunicación. las clasificaciones universitarias son utilizadas para diferentes propósitos por varios actores de la sociedad. los futuros estudiantes y sus familias son la audiencia más interesada en las clasificaciones universitarias, dado lo que está en juego al elegir una universidad

adecuada en términos de inversión de tiempo, perspectivas profesionales futuras y recursos financieros. las clasificaciones brindan información importante, especialmente para aquellos estudiantes que planean ir al extranjero para recibir educación superior. Además de esto, los administradores universitarios ven cada vez más las clasificaciones como un medio para desarrollar estrategias para fomentar el crecimiento y desarrollo de sus instituciones. a nivel nacional, los formuladores de políticas pueden emplear la información proporcionada por las clasificaciones para evaluar las tendencias en los sistemas de educación superior en el mundo, así como en sus propios países. finalmente, los medios de comunicación utilizan los rankings universitarios como un medio para informar a la sociedad sobre el estado de las instituciones de educación superior en el país y en el mundo. las noticias sobre los resultados del ranking atraen el interés de una amplia variedad de lectores cada año [1].

**CARACTERÍSTICAS GENERALES**

**MATERIAL Y MÉTODOS**

El siguiente material a Utilizar será descargado de la siguiente página https://www.kaggle.com/ titulado (QS World University Rankings 2017 - 2022)

es una publicación anual de rankings universitarios globales de Quacquarelli Symonds. El ranking QS recibe la aprobación del International Ranking Expert Group (IREG) y se considera uno de los tres rankings universitarios más leídos del mundo. QS publica sus clasificaciones universitarias en colaboración con Elsevier.

Conceptos Para dar solución al problema identificado es necesario conocer el concepto de Inteligencia Artificial, Google collaboratory, las Principales librerías utilizadas en Python.

Inteligencia Artificial. Una de las definiciones que se puede considerar más ajustada a la realidad es la reflejada en la Encyclopedia Of Artificial Intelligence:

google colaboratory. es un entorno de desarrollo en la nube en este

Principales librerías de Python para Machine Learning

● Pandas: Librería más utilizada para el tratamiento de datos en Python, una de sus grandes virtudes que tiene esta librería es la carga de datos como los archivos de texto plano como CSV “Comma Separated Values”.

● Numpy: Librería que por excelencia tiene su virtud en el procesamiento de arrays. Debido a que contiene una gran colección de funciones que permite realizar cálculos matemáticos complejos sobre arrays multidimensionales.

● Matplotlib: Esta librería es importante en las tareas de visualización y entre sus cualidades destacan que es open source y trabaja a bajo nivel [2].

● Keras: Es una interfaz de alto nivel para manipular las redes neuronales, con keras podemos ver fácilmente si nuestras ideas darán buenos resultados inmediatos. [3]

.

## GRÁFICOS, FOTOGRAFÍAS Y TABLAS

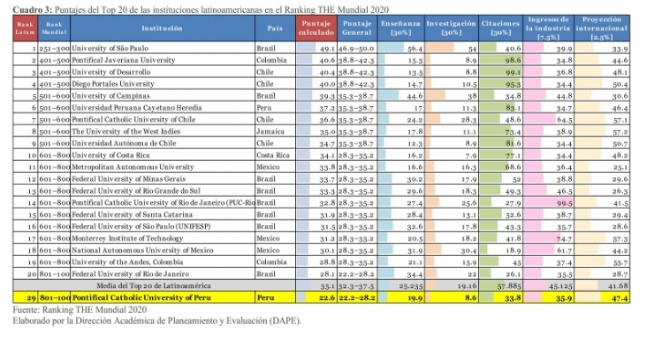


Figura 1. Puntaje del Top 20 de las instituciones latinoamericanas en el Ranking THE Mundial 2020

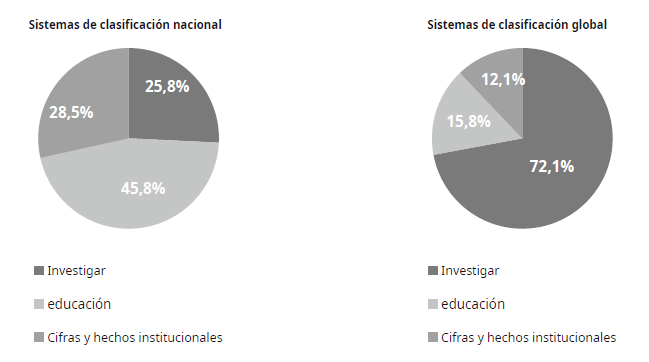
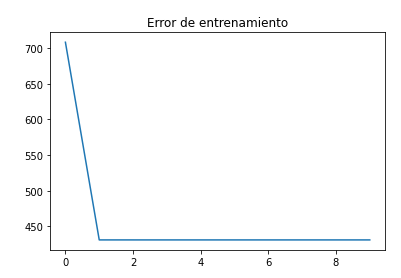


Figura 2.Una comparativa entre sistema de clasificación nacional o globales tomando en cuenta 3 sectores .

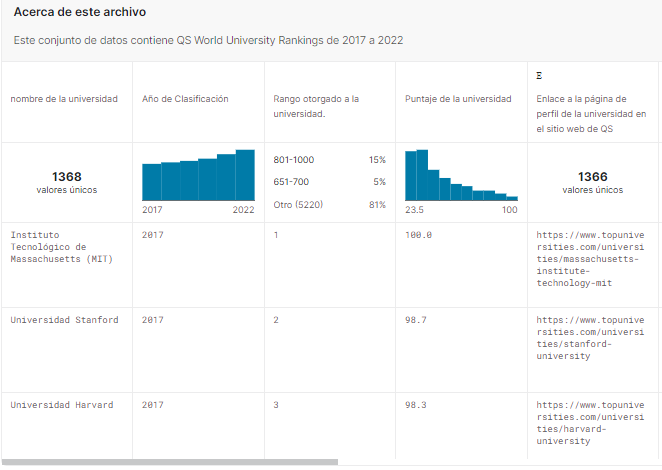


Figuras 3. Errores de entrenamiento mínimo y óptimo con 3 capas.

**PROCEDIMIENTO**

El primer paso a realizar es la obtención de la base de datos de [QS World University Rankings 2017 - 2022 | Kaggle](https://www.kaggle.com/datasets/padhmam/qs-world-university-rankings-2017-2022) de la pagina de kaggle donde se observará los campos de la base de datos en la parte inferior de la página web como se muestra en la Fig. 3, para poder realizar la descarga se tiene que registrarse en Kaggle completando los espacios vacíos Fig. 4.

### IMÁGENES A COLOR



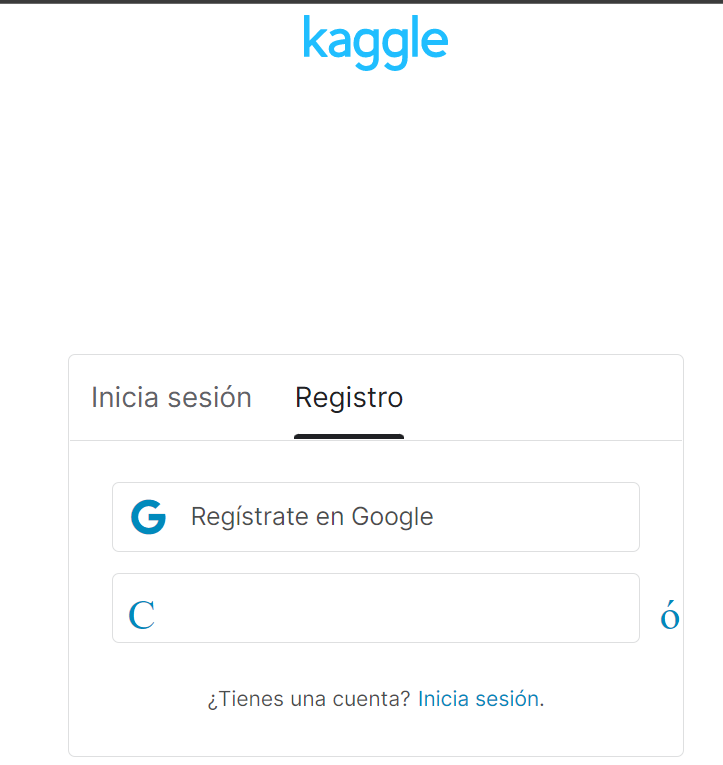


Figura 4. Registro en kaggle

Una vez hecho el registro pasamos a la descarga del archivo con extensión .csv que estará comprimido en .zip Fig.5.

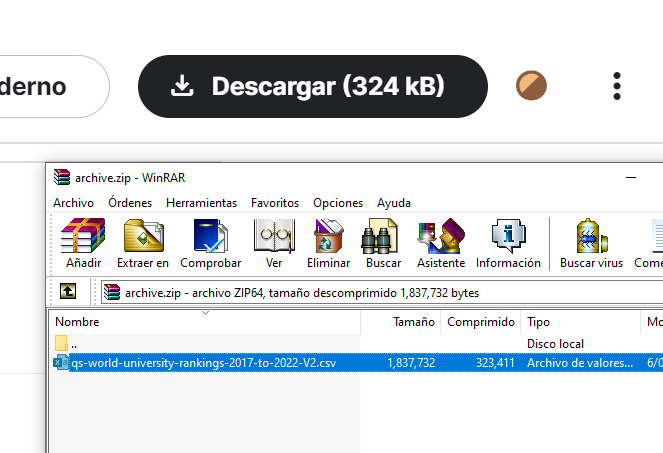


Figura 4. Descarga del archivo con extensión .csv

Realizada la descarga, pasaremos a trabajar con el lenguaje Python en google collaboratory y como es un entorno de trabajo en la nube no utilizamos muchos recursos de nuestro ordenador. Crearemos un nuevo cuaderno donde realizaremos el trabajo, y estará dividido en partes.

PRIMERA PARTE: carga de datos

Importamos la librería Pandas y subiremos el archivo de la base de datos qs-world-university-rankings-2017-to-2022-V2 s a Google collaboratory, para posteriormente visualizar la cantidad de campos que tiene la data.

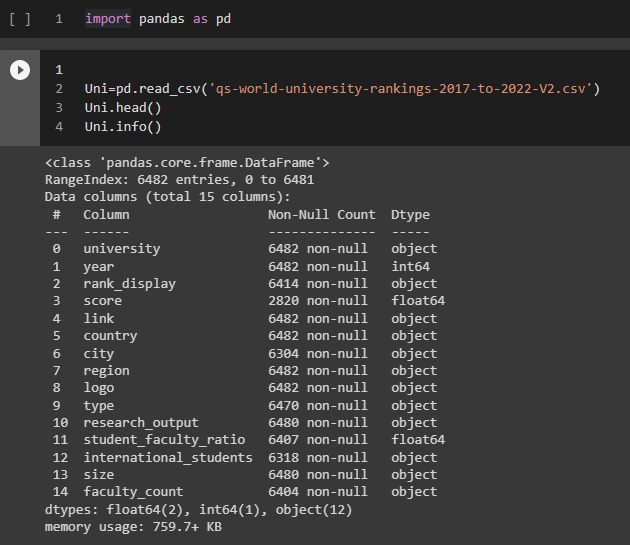


Figura 5. Código y cantidad de campos.

Como se muestra en la Fig. 5 tenemos 15 campos:

1. universidad
2. año
3. rank display
4. puntuación
5. enlace
6. país
7. ciudad
8. región
9. logotipo
10. tipo
11. investigación resultado
12. student\_faculty\_ratio
13. estudiantes internacionales
14. tamaño
15. Cantidad de facultades

SEGUNDA PARTE: Limpieza de Datos

Seleccionando los campos con los que se va a trabajar que serán los campos de research\_output, international students, type, score (resultado de investigación, estudiantes internacionales, escribe, puntaje) y llenamos los campos vacíos de score con su media

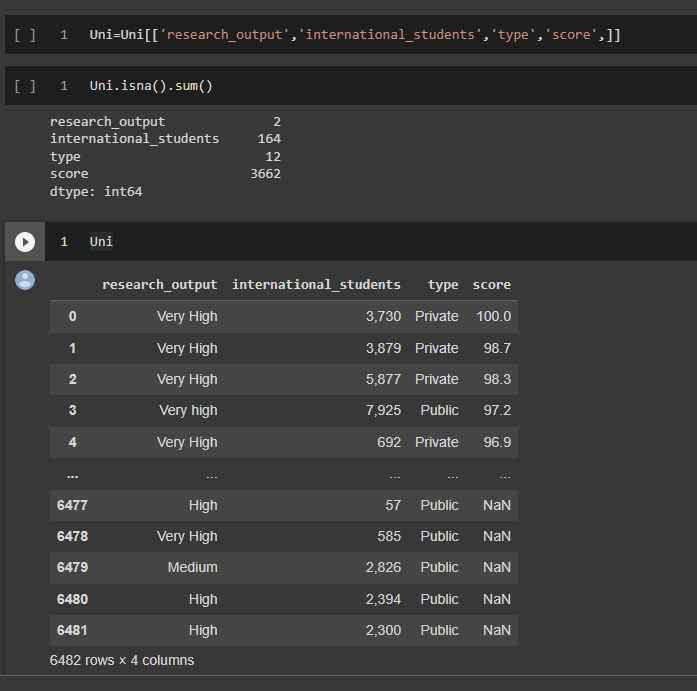


Figura 5. Seleccionar campos y llenar campos vacíos.

Llenamos el tipo de universidad con privada o pública con un rango de 0 y 1 Fig.6.

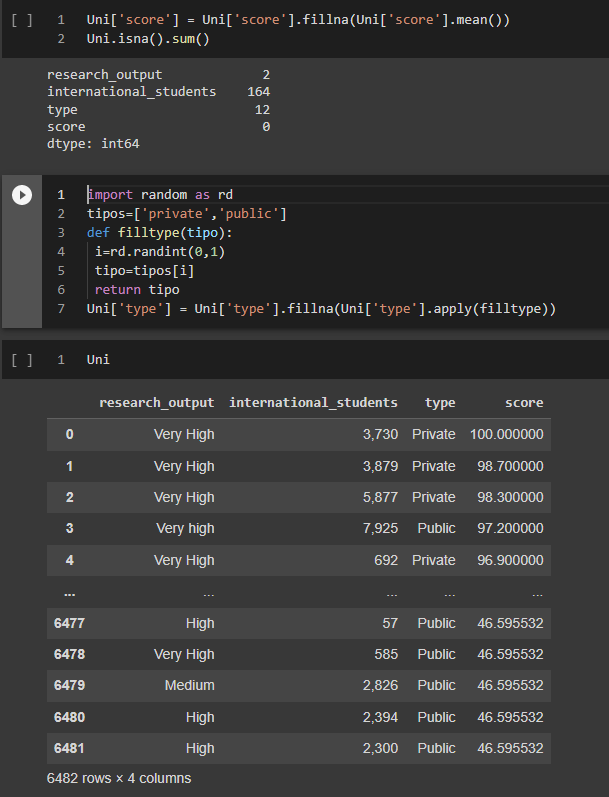


Figura 6. Cambio de dato

Pasamos luego a la estandarización de los datos "reasearch\_output y type" a minúscula Fig.7,para su mejor manejo.

Limpiamos los puntos y comas del campo international\_students siguiendo con el llenado del campo international\_students con valores de rango [1000-25000], y podemos dar una breve visualización de los campos a utilizar con las nuevas modificaciones Fig.8 y por último realizamos la conversión del campo international students de string a int Fig.9.

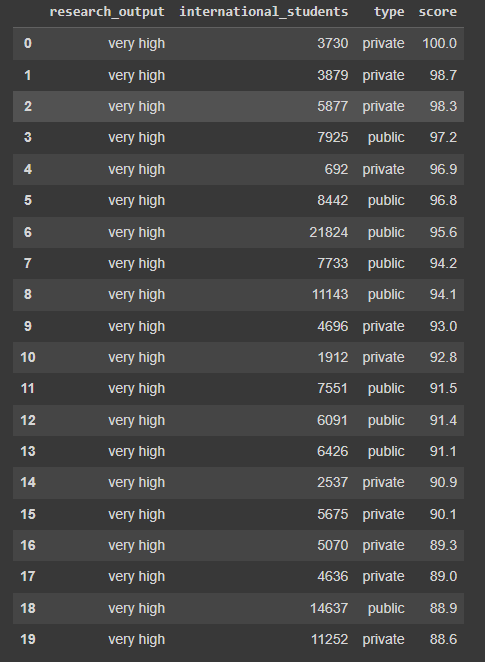


Figura 7. "reasearch\_output y type" a minuscula

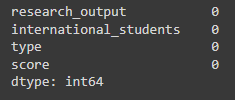


Figura 8. Campos actualizados

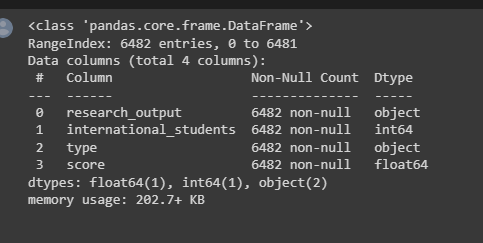


Figura 9. international students de string a int

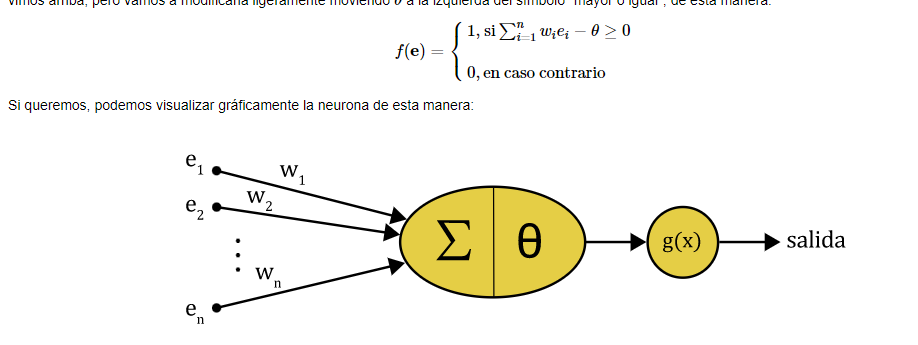
TERCERA PARTE: Normalización de datos

CUARTA PARTE: creación de capas

QUINTA PARTEA: Muestra de grafico

**ECUACIONES**

Volvamos de nuevo a la definición de neurona articial y veamos qué relación tiene con los problemas de clasificación lineal. Recordemos su expresión como la vimos arriba, pero vamos a modificarla ligeramente moviendo θ a la izquierda del símbolo "mayor o igual", de esta manera:



Podemos observar la neurona en la Fig. 4

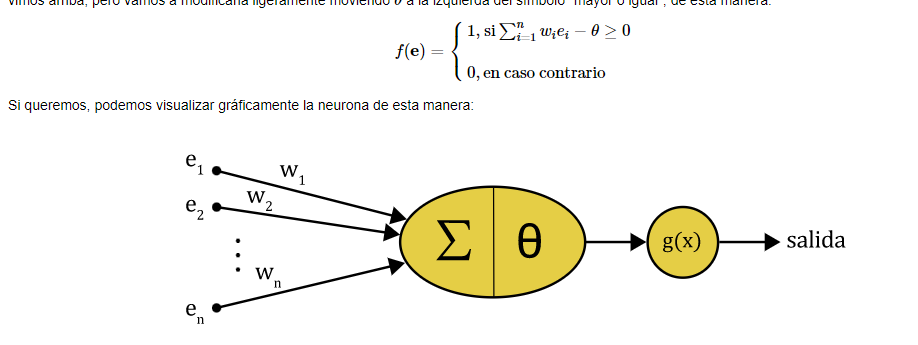


Figura 4.Descripción gráfica de una neurona

## CONCLUSIONES

Con los datos disponibles hasta la fecha se presenta una herramienta para la predicción de las mejores universidades a futuro en redes neuronales artificiales, que complementa a la información proporcionada por el ministerio de educación a través del portal de datos abiertos. Los resultados obtenidos a lo largo del artículo confirman la validez de esta herramienta y la efectividad en la predicción de la cantidad de universidades que desean tener mejor ranking a nivel mundial lo cual se evidencia en la sección de análisis y resultados. Los modelos identificados a lo largo del artículo presentan un horizonte de estimación que depende del lugar donde se encuentran, en aras de obtener una buena predicción.

**RECONOCIMIENTO**

Se agradece al docente por los alcances y conocimientos dados para desarrollar este proyecto. Los Alumnos de la asignatura: “COMPUTACIÓN EN LA NUBE” del ciclo 2022-I, desean expresar su agradecimiento a la Universidad Nacional Micaela Bastidas por todo el apoyo recibido durante el desarrollo del curso.

# Referencias

|  |  |
| --- | --- |
| [1] | C. MURAT PERIT, CENGIZ ACARTURK, OGUZHAN ALASEHIR y CANAN CILINGIR, «Un análisis comparativo de los sistemas de clasificación de universidades,» Akadémiai Kiadó, Budapest, 2015. |
| [2] | C. Mayta Avalos, J. C. Valdivia Mamani, F. Castilla Rosales y M. Gimenes Colana, «Predicción de mortalidad a causa del Covid 19 en Perú utilizando,» Ulasalle, Tacna, 2021. |
| [3] | J. Torres, DEEP LEARNING Introducción práctica con Keras, Ulasalle, 2018. |
| [4] | W. McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, O'Reilly Media, 2012. |