

**ANALÍTICA DE DATOS DE CANALES
DIGITALES, SU INFLUENCIA PREVIO AL
ESTABLECIMIENTO DE LA RADIO ONLINE
UPEC MINDALAE**

**DATA ANALYSIS OF DIGITAL CHANNELS, ITS
INFLUENCE PRIOR TO THE ESTABLISHMENT OF UPEC
MINDALAE ONLINE RADIO**

Recibido: 29/04/2022 - **Aceptado:** 13/01/2026

Marco Antonio Yandún Velasteguí

Docente de la Universidad Politécnica Estatal del Carchi
Tulcán - Ecuador

Magister en Auditoria de Tecnologías de la Información
Universidad Espíritu Santo

marco.yandun@upec.edu.ec

<https://orcid.org/0000-0001-5627-9838>

Jorge Humberto Miranda Realpe

Docente de la Universidad Politécnica Estatal del Carchi
Tulcán - Ecuador

Magister en Ingeniería de Software
Universidad Técnica del Norte

jorge.miranda@upec.edu.ec

<https://orcid.org/0000-0003-1653-7059>

Luis Homero Viveros Almeida

Docente de la Universidad Politécnica Estatal del Carchi
Tulcán - Ecuador

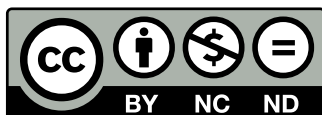
Magister en Administración de empresas
Pontificia Universidad Católica del Ecuador Sede Ibarra

luis.vivero@upec.edu.ec

<https://orcid.org/0000-0003-3612-1798>

Como citar este artículo:

Yandún, M., Miranda, J., & Viveros, L. (Enero – junio 2026). Analítica de datos de canales digitales, su influencia previo al establecimiento de la radio online UPEC MINDALAE. *Sathiri*, 21 (1), 211 – 224. <https://doi.org/10.32645/13906925.1453>



Resumen

Dentro del Proyecto de Investigación “Canales digitales para la investigación y formación continua, avanzada, la innovación y transferencia del conocimiento” de la Universidad Politécnica Estatal del Carchi–UPEC, el presente documento tiene como objetivo analizar los datos recolectados mediante una encuesta aplicada al personal académico y a los estudiantes de la UPEC.

A partir de los datos obtenidos, se generó un conjunto de datos original compuesto por 16 atributos y 1911 registros, los cuales constituyen en el material de análisis. Sobre este conjunto se aplicaron técnicas de Aprendizaje Automático, específicamente el algoritmo K-Nearest-Neighbors (KNN) como metodología principal de análisis; además, se emplearon árboles de decisión, Redes Neuronales Artificiales y diversas bibliotecas de análisis de datos incorporadas en el Lenguaje de Programación Python.

Los resultados permitieron identificar las instancias rol (estudiantes, docente) y el género de los participantes, a partir de las cuales se elaboraron histogramas con algunos atributos, así como la matriz de confusión, matriz de correlación, con el fin de determinar la importancia de los campos rol y género en relación con el resto de los atributos del conjunto de datos utilizados,

Como conclusiones, al aplicar el algoritmo K-Nearest Neighbors (KNN) para el campo rol, se obtuvo una precisión (Accuracy) del 48,3263 %, mientras que para el campo género se alcanzó el mismo valor. En cambio, al aplicar el algoritmo de árbol de decisiones, se obtuvo una precisión del 95,82245% para el atributo rol y del 55,09138% para el atributo género

Palabras claves: Canales digitales, análisis de datos, K-Nearest Neighbors, Aprendizaje automático, redes neuronales artificiales.

Abstract

within the Research Project “Digital channels for research and continuous training, advanced innovation and knowledge transfer” of the State Polytechnic University of Carchi–UPEC, this document aims to perform the analysis of the data collected through a survey of UPEC academic staff and students.

The original dataset is generated with the data, it is composed of 16 attributes and 1911 records that constitute the analysis material to which Automatic Learning is applied with the K-Nearest-Neighbors algorithm as an analysis methodology, in addition the trees are used. of decision, the Artificial Neural Networks and some libraries of data analysis incorporated in the Python Programming Language.

As results, the Role instances (students, teacher) and the Gender of the participants are determined, to make histograms with some attributes, the confusion matrix, the correlation matrix and determine the importance of the Role and Gender fields in relation to the fields of the used dataset,

As conclusions, applying the K-Nearest Neighbors algorithm for the Role field, the Accuracy (precision) of 48.3263% is determined, the Gender field of 48.3263%, on the other hand, applying the decision tree is 95.82245% compared to 55 .09138% to the Gender attribute

Keywords: Digital channels, data analysis, K-Nearest Neighbors, Machine Learning, artificial neural network

Introducción

Previo al establecimiento del proyecto de investigación “Canales digitales para la investigación y formación continua, avanzada la innovación y transferencia del conocimiento”, en la Universidad Politécnica Estatal del Carchi, y considerando el contexto social, económico y político, que obligó al planeta entero a buscar nuevas estrategias para llevar a cabo las actividades como consecuencia de la pandemia por el COVID-19, se generaron restricciones y confinamientos, por lo que actividades que en su mayoría se desarrollaban de forma presencial debieron trasladarse a entornos virtuales alternativos, así como a nuevas formas de comunicación y conexión en tiempo real desde diferentes ubicaciones físicas.

Como antecedentes teóricos, se toman como referencia los estudios realizados por Szukala, A., & Davies, I. (2020). En su artículo “The interconnections between research and impact in social sciences education” se destacan aspectos filosóficos, ontológicos y tecnológicos para la generación de los conocimientos. Así también el trabajo de Estrella, Á. M. C., Pareja, E. M. D., & Ortega-Tudela, J. M. (2018). “Service-learning at university: Helping schools respond to diversity using ICT” en el cual se enfatiza el uso de propuestas sociales y el compromiso con la comunidad mediante la aplicación del método de aprendizaje-servicio. Tomando en cuenta lo anterior, surgen nuevas oportunidades para generar otras formas de conexión y difusión del conocimiento, por lo que la implementación de canales digitales enfocados a radio en línea y televisión online se volvió viable, gracias a la innovación constante de las nuevas tecnologías, las plataformas de streaming y la modalidad de teletrabajo implementada. Previamente, se identificó la necesidad de consultar a la comunidad universitaria sobre aspectos relacionados con el rol estudiantil o docente, el género, el dispositivo electrónico utilizado, la frecuencia de sintonización, el interés en escuchar una nueva propuesta de comunicación y difusión, entre otros aspectos.

Para analizar los 1911 registros y sus 16 atributos (dataset), se realizó el análisis mediante herramientas de inteligencia artificial, aprendizaje automático (Machine Learning), árboles de decisión y algoritmos basados en redes neuronales artificiales, con el fin de establecer correlación entre los atributos del conjunto de datos. A continuación, se presentan algunos conceptos utilizados en la investigación.

El aprendizaje automático (Machine Learning) tiene como finalidad lograr el aprendizaje autónomo o automático en los equipos o sistemas, tal como indican Hsieh, K., Phanishayee, A., Mutlu, O., y Gibbons, P.B. (2020). Este tipo de aprendizaje se clasifica en tres categorías principales: aprendizaje supervisado, no supervisado y de refuerzo.

- **Aprendizaje supervisado:** consiste en etiquetar los datos, los cuales se encuentran agrupados, y determinar patrones de detección para ser utilizados en el etiquetado de nuevos conjuntos de información. Adicionalmente, se conoce como un aprendizaje guiado por un algoritmo previamente entrenado, basado en el aprendizaje autónomo (Canto, F.J., 2020).
- **Aprendizaje no supervisado:** utiliza similitudes entre eventos y etiquetas de conjuntos de conjuntos de datos, en los que el aprendizaje se basa en los resultados devueltos por los agentes inteligentes (Booth, J., Margetts, B.K., Bryant, W.A., Issitt, R.W., Hutchinson, C., Martin, N., y Sebire, N.J., 2021).
- **Aprendizaje de refuerzo:** es presente cuando los datos del sistema no se encuentran etiquetados, pero tras la realización de varias acciones o entrenamientos, y luego de

un determinado período de tiempo, el sistema recibe retroalimentación mediante actualizaciones, con la intervención de agentes externos (Sánchez-Galán, J., 2020).

Algoritmos de inteligencia artificial.

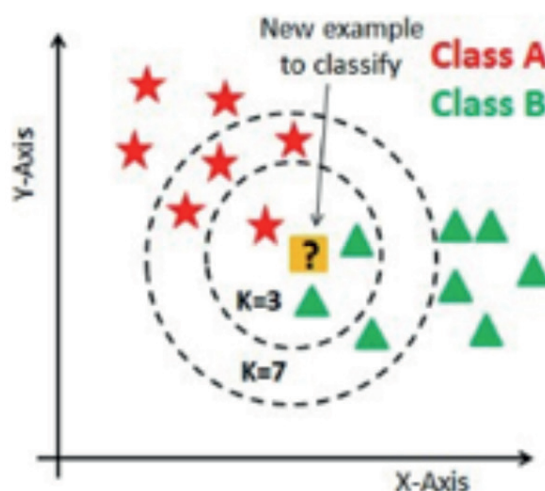
- **Árbol de decisión.** La representación de opciones y sus posibles resultados se estructura en los denominados árboles de decisión, caracterizados por su forma gráfica (Patel, H. H., y Prajapati, P., 2018). Los autores indican que los nodos del árbol representan una elección, mientras que los bordes corresponden a las reglas o condiciones aplicables en una decisión. Cada árbol está compuesto por nodos y ramas, donde cada nodo representa atributos del grupo a clasificar y cada rama representa un valor posible del atributo (Namazkhan, M., Albers, C., y Steg, L., 2020).

Por medio de los árboles de decisión, también es posible aplicar técnicas de minería de datos y aprendizaje automático, utilizando modelos predictivos que asigna las observaciones de un elemento a un valor objetivo determinado.

- **Algoritmo K-Nearest-Neighbor (KNN).** El algoritmo del vecino más cercano K es uno de los algoritmos de clasificación más utilizados para la búsqueda y selección de datos (Fan, A., Gardent, C., Braud, C., y Bordes, A., 2021). Se trata de un algoritmo de aprendizaje automático supervisado, empleado para la resolución de problemas de clasificación y regresión. Su implementación es relativamente sencilla; sin embargo, a medida que aumentan el volumen de datos y el número de registros, el algoritmo puede presentar una disminución en la velocidad de procesamiento. Tal como se muestra en la figura 1, se ilustra la clasificación que existe en la clase A y clase B, las cuales poseen valores céntricos, así como un valor desconocido, que de acuerdo con el algoritmo se determinará a qué clase pertenece.

Figura 1

Representación del algoritmo k vecinos más cercano. Rubiales A. (2020).

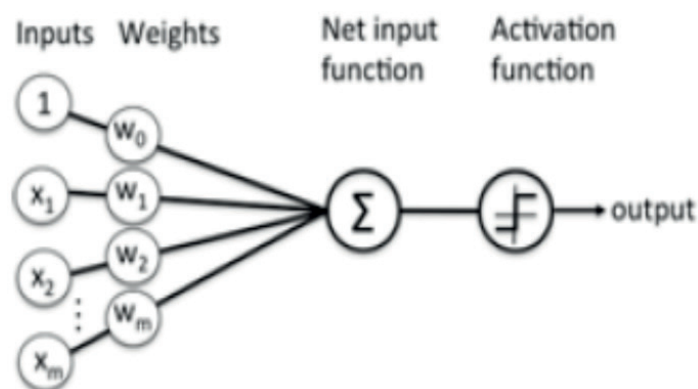


- **Red neuronal artificial** es una serie de algoritmos que relacionan los datos por medio de procesos que imitan la conexión neuronal del cerebro humano. Bahramian, A.,

Parastesh, F., Pham, V., Kapitaniak, T., Jafari, S., y Perc, M. (2021) mencionan, además, que las redes neuronales artificiales pueden adaptarse a cambios en las entradas de datos por medio de agentes inteligentes, lo que permite que la red neuronal genere resultados mejorados sin necesidad de rediseñar los criterios de salida. Asimismo, Hidalgo, J., Guevara, C., y Yandún, M. (2020) señalan que el concepto de redes neuronales, cuyas bases se encuentran en la inteligencia artificial, está ganando popularidad de manera acelerada en el desarrollo de sistemas comerciales.

Figura 2

Red neuronal. Bravo, M., Molero, A., Rojas, J., Prieto, K., y Gómez, Á. (2017).



- **Lenguajes de programación Python.** Conforme al criterio de Vidal-Silva, C. L., Sánchez-Ortiz, A., Serrano, J., y Rubio, J. M. (2021), Python es uno de los lenguajes de programación con los cuales se pueden desarrollar aplicaciones web mediante el marco de trabajo Django, ya que facilita la implementación del patrón Modelo Plantilla Vista. Además, cuenta con estructuras de datos eficientes y de alto nivel, sus librerías incorporadas y su sintaxis clara lo convierten en un lenguaje de programación ideal para escribir y desarrollar aplicaciones en diversas áreas, así como en la mayoría de las plataformas.

Bibliotecas de Python.

- **Numpy:** Es una librería incorporada en el lenguaje de programación Python, especializada en aplicaciones de cálculo numérico y el análisis de datos, particularmente cuando se dispone de grandes volúmenes de información (Harris, C.R., Millman, K.J., Walt, S.V., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M.H., Brett, M., Haldane, A., R'io, J.F., Wiebe, M., Peterson, P., G'erard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., y Oliphant, T.E., 2020).
- **Pandas:** Es una librería de Python especializada en el manejo y análisis de estructuras de datos. De acuerdo con Hagedorn, S. (2020), esta librería permite la lectura y escribir archivos en formato CSV, Excel y bases de datos SQL; además, facilita el acceso a los datos mediante índices o la asignación de nombres a los campos y registros. Asimismo, posibilita el trabajo con series temporales, ofreciendo métodos para combinar, reordenar y dividir los conjuntos de datos.

- **Matplotlib:** Es una biblioteca de Python que permite la generación de gráficos a partir de datos contenidos en arreglos (arrays) (Lemenkova, P., 2020).
- **Google Colab:** Es una plataforma que permite la escritura y ejecución de código en el lenguaje de programación Python a través de un navegador web, resultando útil para aplicaciones de aprendizaje automático y análisis de datos. Además, proporciona un entorno de desarrollo con recursos de GPU, almacenamiento en disco y memoria, provistos por Google (Collins, J.M., Collins, M.R., McKenzie, T., y Marino, M.J., 2021).

Materiales y métodos

Materials

Se utilizó la base de datos original, compuesta por 1911 registros o respuestas, generada a partir de una encuesta en línea aplicada a la comunidad universitaria en el marco del proyecto de investigación “Canales digitales para la investigación y formación continua, avanzada la innovación y transferencia del conocimiento”. Con el fin de aumentar la calidad de los resultados, se realizó una revisión y depuración de datos, tal como se muestra en la figura 3. En dicha figura se presenta el conjunto de datos resultante del preprocesamiento, el cual consistió en corrección de datos a) erróneos, b) redundantes y c) inconsistentes, especialmente aquellos que fueron transformados a datos numéricos para facilitar el tratamiento y análisis de la información. Asimismo, se utilizó la totalidad de la base de datos, sin aplicar un proceso de balanceo de datos.

Figura 3

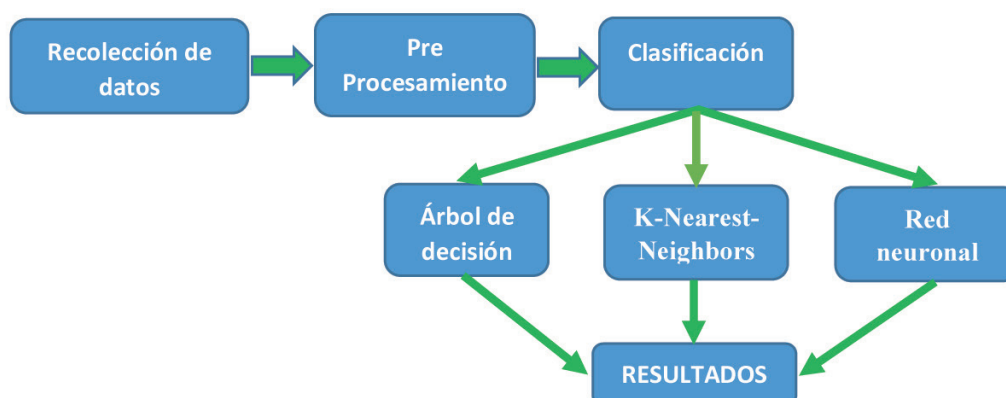
Preprocesamiento de la información, eliminación de ruido de la base de datos.

Rol	Género	Dispositivo Electrónico	Medios Digitales	Frecuencia Uso Medios D	Uso Medios Conoce Infor	Investigación formativa difi	Carreas Difund
	1 Masculino	1	2	4 Si	Si		Si
	1 Masculino	1	2	4 Si	Si		Si
	1 Masculino	1	2	1 Si	Si		Si
	1 Femenino	1	2	3 Si	Si		Si
	1 Femenino	1	2	4 Si	Si		Si
	1 Masculino	2	2	4 Si	Si		Si
	1 Femenino	1	2	4 Si	Si		Si
	1 Femenino	2	2	1 Si	Si		Si
	1 Masculino	1	2	4 Si	Si		Si
	1 Femenino	1	2	4 Si	Si		Si
	1 Femenino	2	2	3 Si	Si		Si
	1 Femenino	1	2	2 Si	Si		Si
	1 Masculino	1	2	4 Si	Si		Si
	1 Femenino	1	2	4 Si	No		Si
	1 Femenino	1	2	4 Si	Si		Si
	1 Femenino	4	2	2 Si	No		No
	1 Femenino	2	2	4 Si	Si		Si
	1 Masculino	1	2	3 Si	Si		Si
	1 Femenino	1	2	3 Si	Si		Si
	1 Masculino	1	4	4 Si	No		Si
	1 Masculino	1	2	4 Si	Si		Si
	1 Masculino	1	2	1 Si	Si		Si
	1 Masculino	1	2	4 Si	Si		Si
	1 Masculino	1	2	1 No	Si		Si
	1 Masculino	2	5	4 Si	Si		Si
	1 Femenino	1	2	3 Si	Si		Si
	1 Femenino	1	2	4 Si	Si		Si
	1 Masculino	2	2	4 Si	Si		Si
	1 Femenino	1	4	3 Si	Si		Si
	1 Femenino	1	2	4 Si	Si		Si

Métodos

La metodología se desarrolla conforme a lo indicado en el gráfico y comienza con el proceso de recolección de datos por medio de una encuesta online; posteriormente el preprocesamiento de los datos mediante la aplicación de los algoritmos de inteligencia artificial, con los que se experimenta

para determinar el mejor clasificador y el algoritmo que presenta el mayor porcentaje de precisión (Accuracy). Finalmente, se comparan los algoritmos y herramientas descritos anteriormente.



Resultados

Con la aplicación del algoritmo K-Nearest Neighbors (KNN) desde el lenguaje de programación Python, se obtienen resultados preliminares, tales como el conteo de registros, el valor medio, la desviación estándar, el valor mínimo y el máximo, tal como se muestra en la figura 4, considerando que el análisis aplicó a la totalidad del conjunto de datos (dataset).

Figura 4

Resultados de la Aplicación del algoritmo KNN en Python a todo el dataset.

	Rol	Género	Dispositivo Electrónico	Medios Digitales	Frecuencia Uso Medios D	Uso Medios Conoce Infor	Investigación formativa difundir Medios Digitales	Carreas Difundir Medios	Interés formación Continua	Tiempo formación continua
count	1911.000000	1911.000000	1911.000000	1911.000000	1911.000000	1911.000000	1911.000000	1911.000000	1911.000000	1911.000000
mean	1.050759	0.615385	1.459969	2.304029	3.001570	0.948718	0.902145	0.953951	0.899006	1.518053
std	0.219562	0.544515	0.796158	0.831437	1.045808	0.220630	0.297196	0.209647	0.301400	0.792348
min	1.000000	0.000000	1.000000	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	1.000000
25%	1.000000	0.000000	1.000000	2.000000	2.000000	1.000000	1.000000	1.000000	1.000000	1.000000
50%	1.000000	1.000000	1.000000	2.000000	3.000000	1.000000	1.000000	1.000000	1.000000	1.000000
75%	1.000000	1.000000	2.000000	2.000000	4.000000	1.000000	1.000000	1.000000	1.000000	2.000000
max	2.000000	3.000000	4.000000	5.000000	4.000000	1.000000	1.000000	1.000000	1.000000	4.000000

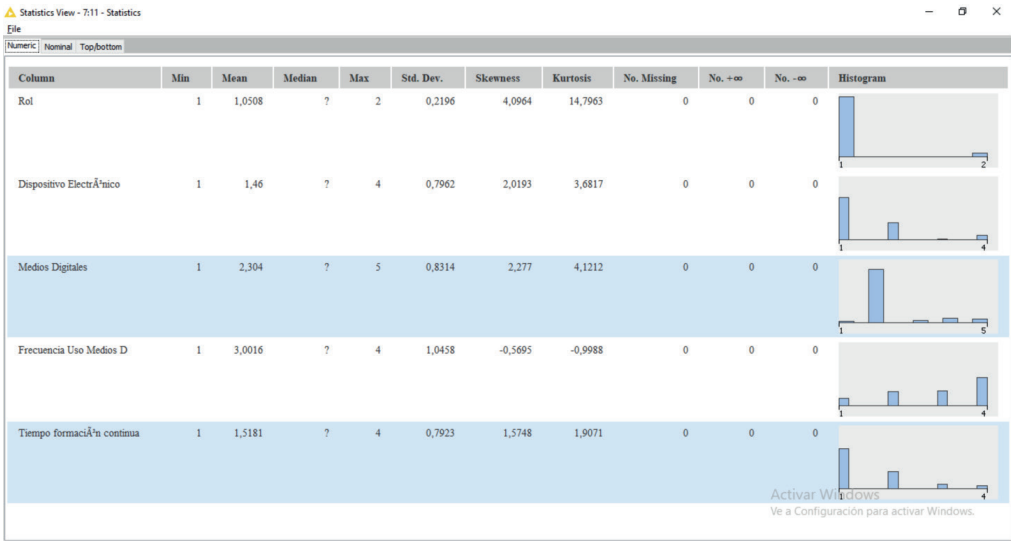
Con el fin de obtener mejores resultados, se procede al reemplazo de valores textuales por valores numéricos para el análisis. En este proceso, se transforman los atributos “Si”, “No”, “Masculino”, “Femenino” y “LGTBI”; en la figura 6 se muestra el reemplazo aplicado a la totalidad del conjunto de datos (dataset).

Figura 5
Resultados del proceso de reemplazo de valores texto a numéricos.

	Rol	Género	Dispositivo Electrónico	Medios Digitales	Frecuencia Uso Medios D	Uso Medios Conoce Infor	Investigación formativa difundir Medios Digitales	Carreas Difundir Medios	Interés formación Continua	Tiempo formación continua
0	1	0	1	2	4	1	1	1	1	2
1	1	0	1	2	4	1	1	1	1	1
2	1	0	1	2	1	1	1	1	0	1
3	1	1	1	2	3	1	1	1	1	3
4	1	1	1	2	4	1	1	1	1	1

Los histogramas se generan mediante el lenguaje de programación Python. En la figura 6 se observan los histogramas correspondientes a los atributos rol, dispositivos electrónicos, medios digitales, frecuencia de uso y tiempo de formación continua. Asimismo, para cada atributo la media, el valor mínimo, el valor máximo y desviación estándar.

Figura 6
Histogramas generados de los campos rol, Dispositivo electrónico, Medios digitales Frecuencia de uso de medios digitales y tiempo de dedicación a la formación continua.



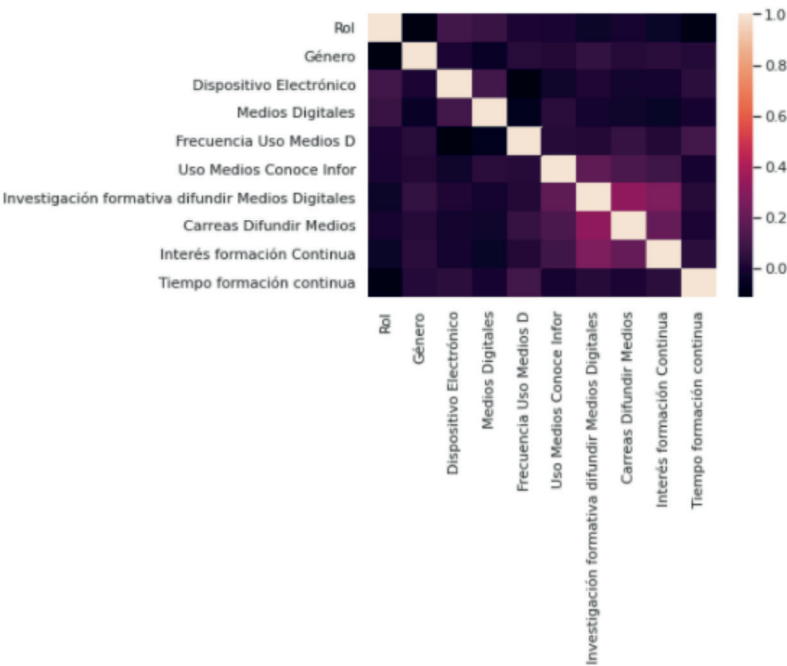
Utilizando el entorno de Google Colab y la ejecución de código en el lenguaje de programación Python se genera la matriz de confusión. En la Figura 8 se observan las etiquetas de datos reales frente a los datos predictivos generados.

(dataset). En la Figura 9 se muestran los resultados, en los cuales se puede apreciar la formación de la diagonal principal con valores iguales a uno. Asimismo, en la Figura 11 se presenta la matriz de correlación utilizando los campos: rol, género, dispositivos electrónico, medios digitales, frecuencia de uso de medios digitales, uso de medios de información, medios de difusión, interés en formación continua y tiempo de formación continua.

Figura 9
Tabla de correlación de campos generada en Colab.

	Rol	Género	Dispositivo Electrónico	Medios Digitales	Frecuencia Uso Medios D	Uso Medios Conoce Infor	Investigación formativa difundir Medios Digitales	Carreas Difundir Medios	Interés formación Continua	Tiempo formación continua
Rol	1.000000	-0.108134	0.096990	0.070292	0.004213	-0.000277	-0.028148	-0.006065	-0.033257	-0.097060
Género	-0.108134	1.000000	0.006131	-0.042255	0.032320	0.018773	0.058484	0.028224	0.037546	0.021563
Dispositivo Electrónico	0.096990	0.006131	1.000000	0.101839	-0.113424	-0.020635	0.008882	-0.014187	-0.011402	0.041191
Medios Digitales	0.070292	-0.042255	0.101839	1.000000	-0.072202	0.036518	-0.013023	-0.018759	-0.038280	-0.001581
Frecuencia Uso Medios D	0.004213	0.032320	-0.113424	-0.072202	1.000000	0.029847	0.019024	0.062417	0.020435	0.102638
Uso Medios Conoce Infor	-0.000277	0.018773	-0.020635	0.036518	0.029847	1.000000	0.178939	0.130025	0.095287	-0.000691
Investigación formativa difundir Medios Digitales	-0.028148	0.058484	0.008882	-0.013023	0.019024	0.178939	1.000000	0.322583	0.269535	0.026404
Carreas Difundir Medios	-0.006065	0.028224	-0.014187	-0.018759	0.062417	0.130025	0.322583	1.000000	0.199792	0.005007
Interés formación Continua	-0.033257	0.037546	-0.011402	-0.038280	0.020435	0.095287	0.269535	0.199792	1.000000	0.041620
Tiempo formación continua	-0.097060	0.021563	0.041191	-0.001581	0.102638	-0.000691	0.026404	0.005007	0.041620	1.000000

Figura 10
Matriz de correlación de campos generada en Colab.



Se toma en cuenta los campos rol (estudiante, docente) y el campo género para determinar su importancia en relación con los demás campos utilizados.

Figura 11

Relación de importancia del campo Rol frente a los demás campos.

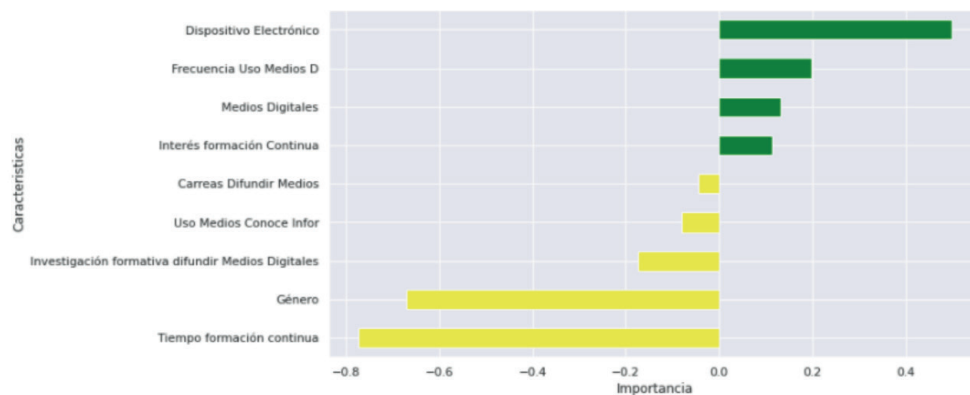
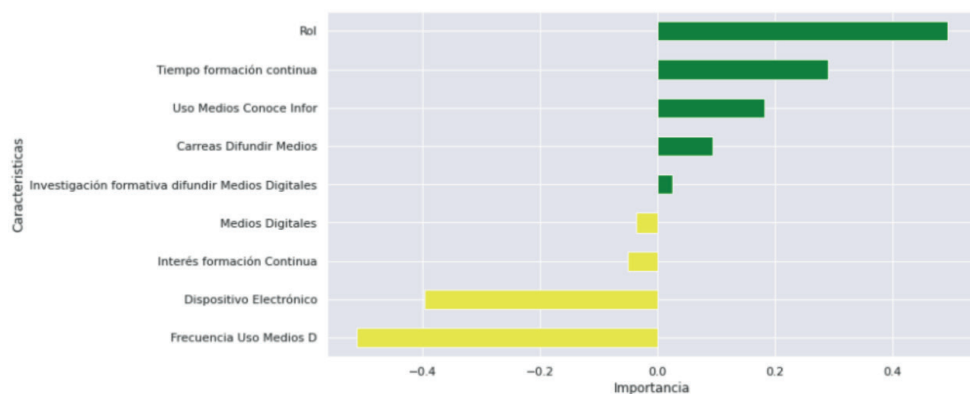


Figura 12

Relación de importancia del campo Género frente a los demás campos.



Conclusiones

Para la presente investigación, mediante el uso de diferentes métodos y herramientas de inteligencia artificial aplicados a los datos originales obtenidos, se logró identificar los factores preponderantes que contribuyen a la toma de decisiones para el establecimiento e implementar la radio online UPEC Mindalae. Entre los resultados más relevantes, se identificaron aquellos relacionados con el rol (docente y estudiante) y el género de las personas encuestadas.

En relación con el atributo rol, la precisión (accuracy) obtenida mediante el algoritmo de árboles de decisiones fue del 95,82245%, mientras que para el atributo género fue del 55,09138%. Por su parte, al aplicar el modelo de regresión lineal, se alcanzó un accuracy del 97% para el campo rol y 62% para el campo género. Finalmente, al aplicar el algoritmo K-Nearest Neighbors (KNN), la precisión obtenida fue del 48,3263 % tanto para el campo rol como para el campo género.

Referencias bibliográficas

- Bahramian, A., Parastesh, F., Pham, V., Kapitaniak, T., Jafari, S., y Perc, M. (2021). Collective behavior in a two-layer neuronal network with time-varying chemical connections that are controlled by a Petri net. *Chaos*, 31 3, 033138.
- Booth, J., Margetts, B.K., Bryant, W.A., Issitt, R.W., Hutchinson, C., Martin, N., y Sebire, N.J. (2021). Machine Learning Approaches to Determine Feature Importance for Predicting Infant Autopsy Outcome. *Pediatric and Developmental Pathology*, 24, 351–360.
- Bravo, M., Molero, A., Rojas, J., Prieto, K., y Gómez, Á. (2017). Desarrollo de un sensor virtual de flujo usando Redes Neuronales Artificiales para Bombas de Cavidades Progresivas en el campo San Diego de Cabrutica
- Canto, F.J. (2020). Eficacia de diferentes reglas hebbianas en el Aprendizaje Supervisado.
- Collins, J.M., Collins, M.R., McKenzie, T., & Marino, M.J. (2021). The Role of CoLab and GAITS in Enabling the RADx Tech Program. *IEEE Open Journal of Engineering in Medicine and Biology*, 2, 119-124.
- Estrella, Á. M. C., Pareja, E. M. D., & Ortega-Tudela, J. M. (2018). Service-learning at university: Helping schools respond to diversity using ICT, Bordon. *Revista De Pedagogia*, 69(3), 73-87.
- Fan, A., Gardent, C., Braud, C., y Bordes, A. (2021). Augmenting Transformers with KNN-Based Composite Memory for Dialog. *Transactions of the Association for Computational Linguistics*, 9, 82-99.
- Hagedorn, S. (2020). When sweet and cute isn't enough anymore: Solving scalability issues in Python Pandas with Grizzly. *CIDR*.
- Harris, C.R., Millman, K.J., Walt, S.V., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M.H., Brett, M., Haldane, A., R'io, J.F., Wiebe, M., Peterson, P., G'erard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., y Oliphant, T.E. (2020). Array programming with NumPy. *Nature*, 585, 357–362.
- Hidalgo, J., Guevara, C., y Yandún, M. (2020). Generation of User Profiles in UNIX Scripts Applying Evolutionary Neural Networks. In *International Conference on Applied Human Factors and Ergonomics* (pp. 56-63). Springer, Cham.
- Hsieh, K., Phanishayee, A., Mutlu, O., y Gibbons, P.B. (2020). The Non-IID Data Quagmire of Decentralized Machine Learning. *ArXiv*, abs/1910.00189.
- Lemenkova, P. (2020). Python Libraries Matplotlib, Seaborn and Pandas for Visualization Geospatial Datasets Generated by QGIS. *Social Science Research Network*.

- Namazkhan, M., Albers, C., y Steg, L. (2020). A decision tree method for explaining household gas consumption: The role of building characteristics, socio-demographic variables, psychological factors and household behaviour. *Renewable and Sustainable Energy Reviews*, 119, 109542.
- Patel, H. H., y Prajapati, P. (2018). Study and analysis of decision tree based classification algorithms. *International Journal of Computer Sciences and Engineering*, 6(10), 74-78.
- Rubiales A. (2020). Vecinos Cercanos (K-Nearest Neighbors) en Sklearn.
- Sánchez-Galán, J. MACHINE LEARNING Y SUS APLICACIONES. In ACTAS DEL V CONGRESO INVESTIGACIÓN, DESARROLLO E INNOVACIÓN DE LA UNIVERSIDAD INTERNACIONAL DE CIENCIA Y TECNOLOGÍA IDI-UNICYT 2020 _ (p. 54).
- Szukala, A., & Davies, I. (2020). The interconnections between research and impact in social sciences education. *Journal of Social Science Education*, 19(2), 1-7. doi:10.4119/jsse-3537
- Vidal-Silva, C. L., Sánchez-Ortiz, A., Serrano, J., y Rubio, J. M. (2021). Experiencia académica en desarrollo rápido de sistemas de información web con Python y Django. *Formación universitaria*, 14(5), 85-94.