

**Modelo de Relación Docente –
Asignatura Aplicando Minería de Texto**

**Teacher-Subject Relationship Model
Applying Text Mining**

Leslie Paullette Macías-Veliz¹
Universidad Técnica de Manabí - Ecuador
lmacias8367@utm.edu.ec

Gabriel Agustín Cotera-Ramírez²
Universidad Técnica de Manabí - Ecuador
gabriel.cotera@utm.edu.ec

doi.org/10.33386/593dp.2023.5.2022

V8-N5 (sep-oct) 2023, pp. 982-998 | Recibido: 10 de julio de 2023 - Aceptado: 24 de agosto de 2023 (2 ronda rev.)

¹ Soy estudiante de la Facultad Ciencias Informática, estudió la carrera de Ing. en Sistema de Información. Con respecto a la presente Investigación considero que ha sido desafiante, fructífera y enriquecedora que me ha generado nuevos conocimientos y aprendizajes significativos.

² Título de Magister en Informática de Gestión Y nuevas Tecnologías por la Universidad Técnica Federico Santa María de Chile en 2008. Docente titular auxiliar de la Facultad de Ciencias Informáticas de la Universidad Técnica de Manabí. Línea de investigación Machine Learning, Procesamiento de datos, IoT.
ORCID: <https://orcid.org/0000-0003-2726-8317>

Descargar para Mendeley y Zotero

RESUMEN

El estudio se enfoca en la relación entre los títulos profesionales de los docentes universitarios y las asignaturas que se les asignan. La metodología utilizada en el estudio se basa en la minería de textos y el análisis de frecuencias de palabras. Se empleó una metodología documental descriptiva que utilizó métodos empíricos y herramientas comparativas. Se recopiló documentos relacionados con las asignaturas y los perfiles de los docentes y se realizaron procesos de limpieza, vectorización y similitud en los textos. Los resultados obtenidos mostraron una correlación significativa entre el perfil docente y los contenidos de las asignaturas. Se encontró un promedio del 80% de correlación entre los docentes y los contenidos de las asignaturas de la carrera de Sistemas de Información. Se realizó el mismo análisis para los docentes de la carrera de Tecnologías de la Información, y se encontró un promedio del 82% de correlación. Además, se analizó la correlación en mayor detalle para docentes y asignaturas específicas, encontrando altos niveles de correlación, con porcentajes que oscilan entre el 82% y el 88%. El estudio utilizó técnicas como word2Vec y similitud coseno para calcular la aproximación entre las palabras y los conceptos relacionados. Los resultados respaldan la importancia de utilizar métodos de minería de texto para analizar y comprender la relación entre los docentes y las asignaturas, y proporcionan una base sólida para la mejora continua de la gestión educativa en el campo de las tecnologías informáticas. El estudio destaca la importancia de implementar una cultura de toma de decisiones basada en los datos generados en las universidades.

Palabras clave: algoritmos, asignatura, docente, minería de texto, similitud.

ABSTRACT

The study focuses on the relationship between the professional titles of university professors and the subjects assigned to them. The methodology used in the study is based on text mining and frequency analysis of words. A descriptive documentary methodology was employed, using empirical methods and comparative tools. Documents related to the subjects and the profiles of the teachers were collected, and processes of cleaning, vectorization, and similarity were performed on the texts. The results obtained showed a significant correlation between the teaching profile and the contents of the subjects. An average correlation of 80% was found between the professors and the contents of the subjects in the Information Systems program. The same analysis was conducted for the professors in the Information Technology program, and an average correlation of 82% was found. Additionally, the correlation was analyzed in greater detail for specific teachers and subjects, revealing high levels of correlation, with percentages ranging from 82% to 88%.

The study utilized techniques such as word2Vec and cosine similarity to calculate the proximity between words and related concepts. The results support the importance of using text mining methods to analyze and understand the relationship between professors and subjects, providing a solid foundation for continuous improvement in educational management in the field of information technology. The study highlights the importance of implementing a decision-making culture based on the data generated in universities.

Key words: algorithms, subject, teacher, text mining, similarity.

Introducción

La digitalización de los procesos académicos en las universidades ha generado una gran cantidad de datos electrónicos relacionados con estudiantes, asignaturas y otros aspectos. Esta información proporciona valiosa ayuda a docentes y autoridades, al tiempo que contribuye al avance en la calidad de los procesos de aprendizaje, brindando testimonios oportunos a las partes interesadas. Los métodos de minería de datos tienen como objetivo extraer conocimientos significativos de estos datos. Cuando se aplica a los datos generados en el sector educativo, se denomina Minería de Textos Educativos (Vargas Pulliquitín, 2018).

En este estudio, el enfoque se da en la relación entre los títulos de tercer y cuarto nivel profesional de los docentes universitarios y las asignaturas que se les asignan para el proceso de enseñanza - aprendizaje de los estudiantes de la Facultad de Ciencias Informáticas de la Universidad Técnica de Manabí. El objetivo principal es comprender cómo se relacionan estos elementos y cómo influyen en la carga académica del docente.

La minería de textos se encarga de desarrollar métodos para analizar datos de un sistema educativo con el fin de descubrir patrones en conjuntos de datos masivos que, de otra manera, serían muy difíciles o imposibles de analizar debido al volumen de datos. Como resultado, la minería de textos se utiliza para tomar decisiones sobre cómo mejorar el proceso de enseñanza y aprendizaje, así como para diseñar o rediseñar entornos de aprendizaje (Florio, 2022).

El rendimiento obtenido de la relación entre el docente y la asignatura que imparte puede ser crucial en la educación superior, ya que el historial académico es uno de los criterios básicos considerados por las universidades de alta calidad. Bin Mat et al. (2013) señalaron que el rendimiento de los estudiantes se puede medir mediante la evaluación de la relación docente-asignatura, pero esta evaluación puede estar dirigida a aspectos específicos del periodo de

estudio durante la permanencia en la institución universitaria (Shahiri et al., 2015). Además, es importante para la institución educativa que el nivel de éxito alcanzado por los estudiantes refleje la calidad de enseñanza brindada (York et al., 2015).

La minería de textos es el proceso de descubrir patrones interesantes de gran cantidad de datos. Es considerado como el resultado natural de la evolución de las tecnologías de información y de la industria de base de datos que lograron desarrollar funcionalidades críticas como recolección y creación de datos, administración de datos y técnicas avanzadas de análisis. La gran abundancia de datos generados mediante la evolución de las tecnologías de internet creó la necesidad de contar con bases de datos heterogéneos e interconectados a escala mundial junto con herramientas poderosas para el análisis de datos. Dada la gran velocidad en la evolución de la generación de datos, se crea una situación que se considera como rica en datos, pero pobre en información. Se busca acercar esta brecha mediante el desarrollo sistemático de herramientas de minería de textos para poder convertir estos datos en fuentes de información confiables (Han et al., 2012).

Estos cambios también tienen un impacto significativo en el sector educativo a nivel nacional, donde el uso de software educativo, Internet y la creación de repositorios de datos han aumentado considerablemente la cantidad de datos disponibles. Es importante explorar y aprovechar esta información para mejorar el proceso de aprendizaje. Uno de los principales desafíos que enfrenta este sector es analizar el crecimiento exponencial de datos educativos y utilizarlos para la toma de decisiones informadas (Romero & Ventura, 2013).

El triángulo didáctico tradicional, en el que el profesor, el alumno y el contenido son los vértices del triángulo, es una representación que se utiliza para conceptualizar la enseñanza y el aprendizaje (Kansanen, 2003). Aunque esta figura puede parecer simplificar la complejidad de la enseñanza y el aprendizaje, puede ayudar a

interpretar y discutir cada elemento en relación con los demás.

La aplicación de la minería de textos en las universidades permitirá desarrollar una cultura de toma de decisiones basada en los datos generados, involucrando a los departamentos de tecnología de la información en la creación de nuevas políticas para la recolección y uso de datos. Además, permitirá a los principales actores (docentes, estudiantes y autoridades) convertirse en consumidores inteligentes de la nueva información generada y formular preguntas críticas sobre la propuesta de asignaturas a los docentes en la universidad (Lang et al., 2017).

En Ecuador, los estudios e investigaciones relacionadas con el análisis del rendimiento académico en las universidades son limitados, especialmente cuando se utiliza Minería de Textos Educativos (Contreras Barrera, 2015). En esta investigación sobre la relación entre docentes y asignaturas, se seleccionarán diversos indicadores que se analizarán utilizando estadísticos paramétricos.

La analítica de texto (minería de texto o text mining) comprende un conjunto de técnicas que permiten estructurar la información heterogénea presente en los textos con el objetivo de identificar patrones, como el uso de palabras, para extraer nueva información (Ruiz-Cornejo et al., 2022).

Los avances tecnológicos de los últimos años han acelerado la producción de información en formato digital, lo que plantea la necesidad de resolver problemas durante su procesamiento. Por esta razón, se requieren métodos y tecnologías para el procesamiento, organización y clasificación de dicha información (Kumar & Saeed, 2018).

Revisión de la Literatura

Duarte et al. (2019) en su publicación identifica estrategias de aprendizajes con un estudio transversal mediante la aplicación de una encuesta a estudiantes relacionada a la educación por competencias, divididas en 3 grupos de

estudiantes y que es analizada mediante minería de texto, esta utiliza una metodología que se divide en cuatro etapas. La primera etapa se relaciona con la construcción de la base de datos a analizar, la segunda etapa condensa las actividades necesarias para realizar la minería de texto aplicando el Kmeans (KM), en la tercera etapa se explica el método para determinar el número de clústeres y en la cuarta se expone el análisis de resultados por grupo. Como resultado se tiene que el grupo 1 esperan que se implementen estrategias para que su proceso de aprendizaje en el curso esté orientado a capacitarlos para así obtener resultados positivos a la hora de presentar las evaluaciones. El grupo 2 esperan que se implementen estrategias con las cuales se apliquen los conceptos vistos durante el curso en un ámbito laboral como son: casos de estudio, manejo de herramientas ofimáticas, visitas técnicas o talleres en distintos formatos (virtuales, en Excel, etc.). El grupo 3 esperan que se implementen estrategias básicas de aprendizaje teórico, en general esperan más talleres sin articular ningún tipo de enfoque en el mundo real. Como conclusión expresaron que la identificación de tendencias y patrones en documentos relacionados con la educación permite ajustar las estrategias de enseñanza-aprendizaje para adaptarse a las necesidades y preferencias de los estudiantes.

En el trabajo de Pérez y Cardoso (2010) expresan que la clasificación de documentos de texto es una aplicación de la minería de textos que pretende extraer información de texto no estructurado. Su interés se justifica porque se estima que entre el 80% y el 90% de los datos de las organizaciones son no estructurados. Por otro lado, la búsqueda semántica permite al usuario especificar en una consulta no solamente términos que deben aparecer en el documento, sino conceptos y relaciones, que pueden detectarse mediante el análisis de texto. En la cual implementaron un buscador semántico que aproveche el resultado de algoritmos de aprendizaje automático para la clasificación de documentos. Mediante la clasificación automática de documentos utilizando el modelo aprendido es de bastante calidad comparada con

las etiquetas asignadas manualmente. En base a los experimentos realizados, el algoritmo semi-supervisado tiene un rendimiento bastante bueno en cuanto a los resultados de la clasificación y requiere menos ejemplos etiquetados para aprender al poder aprovechar los ejemplos no etiquetados.

En otra investigación, Espíritu-Sandoval (2016) ha demostrado la viabilidad de crear una memoria organizacional y extraer información de texto no estructurado. En este caso, se construyó una memoria organizativa alimentada por los documentos de Proyecto de Aplicación Profesional (PAP) y se extrajo información utilizando técnicas y herramientas de minería de texto. Esto ha facilitado a la Coordinación de Proyectos de Aplicación Profesional (CPAP) la organización de los documentos PAP y la extracción ordenada y clara de información sobre temas y competencias específicas.

Durante el desarrollo de un trabajo de investigación, Barrio (2014) llevó a cabo un análisis exhaustivo de varias herramientas de minería de texto y se evaluaron diferentes algoritmos. Se ha agregado una nueva funcionalidad a la herramienta SIENA, que permite realizar un análisis rápido y visual de chats colaborativos. Esta adición fue posible gracias a la implementación de la herramienta de minería de texto SOBEK. Para poder utilizar SOBEK, se establecieron conversaciones con los desarrolladores de dicha herramienta, lo que implicó interactuar con diferentes equipos (SIENA y SOBEK) de manera simultánea. Toda la comunicación con los desarrolladores de SOBEK se llevó a cabo en inglés y español.

Rivero et al. (2020) en su investigación implementan técnicas de Minería de Textos (Text Mining) en conjunto con técnicas de Minería Web (Web Mining) para realizar un estudio de los Patrones de Escritura utilizados en la confección de documentos digitales científicos. Tanto la Minería de Textos como la Minería Web se engloban dentro de las técnicas de Minería de Datos y permiten descubrir patrones utilizados en grandes volúmenes de texto. Además, el proyecto incluirá una investigación sobre la

aplicación de técnicas o algoritmos orientados al Procesamiento del Lenguaje Natural en el análisis de textos o documentos obtenidos de Redes Sociales. A través de su aplicación, se busca obtener prototipos de sistemas que faciliten el análisis en cuestión.

En un estudio experimental de Betancourt et al. (2016) en el cual se formaron tres equipos de investigadores para llevar a cabo una investigación cualitativa. Los equipos participaron en debates, observación y recopilación de información, y se familiarizaron con los conceptos de minería de datos. Se logró capacitar a estudiantes de Medicina en la realización de investigaciones cualitativas en el marco de sus actividades educativas, con el objetivo de caracterizar elementos de sexualidad en adolescentes y analizarlos mediante técnicas de minería de textos, lo que facilita su trabajo con los jóvenes.

Otro experimento aplicando minería de texto de Garrido et al. (2014) han identificado una lista de doce competencias docentes que deben adquirir y desarrollar los docentes encargados de la formación en competencias de los estudiantes del nuevo Espacio Europeo de Educación Superior.

En un estudio de Navarro (2021) se exploró la efectividad de las técnicas de minería de texto para la clasificación multiclase de textos de tamaño reducido con conjuntos de datos limitados. Se utilizaron métodos simples como Bag-of-Words (BoW) y Term Frequency (TF), combinados con clasificadores ampliamente utilizados, como Support Vector Machine (SVM) y árbol de decisión. Los resultados obtenidos en los experimentos revelaron que el sistema de minería de textos que emplea estas técnicas básicas de clasificación obtuvo los mejores resultados, logrando una precisión del 100% en la predicción utilizando árbol de decisión.

Además, en Temesio, et al. (2021) muestran un estudio sobre los factores que influyen en el rendimiento estudiantil. Los resultados mostraron que los estudiantes asociaron su rendimiento principalmente con su participación

activa y las dificultades encontradas durante su participación. En relación con el diseño de la enseñanza, se identificaron aspectos importantes como el tiempo destinado a las evaluaciones, la diversidad de actividades propuestas, los espacios de interacción entre docentes y estudiantes, la definición clara de objetivos y una planificación coherente. Además, el análisis de los textos de preguntas abiertas reveló la necesidad de una interacción fluida con los equipos docentes, así como respuestas y retroalimentación adecuada a las tareas asignadas.

Por otro lado, Benítez (2016) presentó una herramienta informática basada en minería de texto para la caracterización psicosocial de los estudiantes. Esta herramienta automatiza el análisis de los datos obtenidos a través de la técnica psicológica de los Diez Deseos en una encuesta realizada en la facultad. Su objetivo es proporcionar información relevante para la toma de decisiones orientadas a mejorar la formación de los estudiantes universitarios.

Metodología

La minería de textos se enfoca en la extracción de información útil y significativa de diversos formatos de documentos heterogéneos, como páginas web, correos electrónicos, medios sociales y artículos de revistas, entre otros. Esta extracción se logra mediante la identificación de patrones en los textos, como tendencias en el uso de palabras y estructura sintáctica (Vargas Guarín et al., 2021).

En este estudio, se empleó una metodología documental de carácter descriptivo, utilizando métodos empíricos y herramientas comparativas, incluyendo tanto hardware como software, para un diseño óptimo del estudio. El proceso de recopilación de información documental y estudios relacionados se realizó mediante una cadena de búsqueda basada en la propuesta planteada. Los términos de búsqueda se definieron como: (minería de texto OR aplicación de minería de texto) AND (aprendizaje OR asignatura). La búsqueda se realizó en Google académico, IEEE y Scielo. De

estas fuentes se obtuvieron y seleccionaron los estudios más relevantes.

La técnica de análisis seleccionada fue el análisis de frecuencias de palabras, que permitió identificar los términos o conceptos más recurrentes en un conjunto de datos. Esta técnica resulta útil para analizar opiniones de clientes o conversaciones en redes sociales. Por ejemplo, si términos como “demasiado caro” o “sobrevalorado” aparecen con frecuencia, el análisis puede sugerir que el producto es considerado costoso y podría ser necesario ajustar su precio (Digión et al., 2021).

Además, se planteó una investigación descriptiva para detallar los aspectos de las asignaturas, los docentes y la asignación, elementos del problema de estudio. Como método utilizado en esta investigación, se empleó el enfoque experimental para analizar la relación entre docentes y asignaturas, utilizando la minería de texto como técnica de extracción. El primer paso consistió en seleccionar la documentación que contenía los elementos esenciales de las asignaturas, así como los perfiles docentes de las carreras de Ingeniería de Sistemas de Información e Ingeniería de Tecnologías de la Comunicación de la Facultad de Ciencias Informáticas de la Universidad Técnica de Manabí que incluían los títulos de tercer y cuarto nivel profesional. En el segundo proceso, se estructuró la información en archivos de texto con formato utf8. Finalmente, se llevaron a cabo procesos de limpieza, vectorización y similitud en los textos. Como herramientas, se utilizó una laptop como parte del hardware y se trabajó con aplicaciones como Notepad++, Jupyter (con Python 3.9.12) y el acceso a internet para el desarrollo de la investigación.

La minería de texto es una disciplina que emplea algoritmos de aprendizaje automático y procesamiento de lenguaje natural para extraer información útil y significativa de grandes volúmenes de texto. Una forma de relacionar dos textos es mediante el análisis de similitud de texto, que mide el grado de similitud entre ambos en función de la frecuencia de palabras o términos compartidos. A continuación, se detalla

de manera exhaustiva el proceso realizado para el modelo de relación docente-asignatura.

Figura 1

Modelo de relación docente - asignatura



Fase 1: Recolección de datos

Se llevó a cabo el proceso de recolección de documentos proporcionados por el coordinador de carrera de la Facultad de Ciencias Informáticas, que incluían los contenidos de asignaturas y el perfil docente. La documentación se presentó en formato xlsx, donde se identificaron los contenidos mínimos de las asignaturas y los títulos de tercer y cuarto nivel profesional de los docentes. Posteriormente, se procedió a convertir los archivos en formato xlsx a archivos de texto con formato utf8. En total, se obtuvieron 60 archivos de texto para los contenidos de las asignaturas y 27 archivos de texto para el perfil docente de la carrera de Tecnologías de la Información. Para la carrera de Sistemas de Información, se generaron 47 archivos de texto correspondientes a los contenidos mínimos de las asignaturas y 29 archivos de texto para el perfil docente.

Siguiendo las etapas para la elaboración del algoritmo a aplicar en la relación de textos utilizando la minería de texto, se inició con la importación de las bibliotecas necesarias para el proceso. A continuación, se detalla el código correspondiente a dichas bibliotecas:

```
from contexto.lectura import leer_texto
from contexto.limpieza import limpiar_texto,
lista_stopwords, remover_stopwords
from contexto.escritura import escribir_texto
from contexto.vectorizacion import
VectorizadorFrecuencias
from contexto.comparacion import Similitud
from contexto.exploracion import grafica_
```

```
barchart_frecuencias
from contexto.exploracion import obtener_
ngramas, par_nubes, nube_palabras
from contexto.utils.auxiliares import verificar_
crear_dir
from contexto.comparacion import Similitud,
Distancia, DiferenciaStrings
from contexto.vectorizacion import *
from docentes_SI import *
from asignatura_SI import *
import pandas as pd
import numpy as np
import os
from IPython.display import display
from openpyxl import Workbook
import matplotlib.pyplot as plt
```

Tras la importación de las bibliotecas, se procede a la declaración de variables para llevar a cabo la lectura, limpieza, vectorización y medición de similitud de los textos generados, así como una variable para almacenar los resultados obtenidos. A continuación, se muestra el código correspondiente a la declaración de estas variables:

```
wb = Workbook()
ws = wb.active
ws.cell(row=1, column=1, value="Docente")
ws.cell(row=1, column=2, value="Similitud")
ruta_doc1 = 'Archivos/ContenidosSW.txt'
texto_1 = leer_texto(ruta_doc1)
tamano=29
ruta_docente= []
texto= []
```

Es importante destacar que los detalles específicos de la implementación del código dependerán de las bibliotecas y herramientas utilizadas, así como de los requisitos y objetivos del estudio en particular. A continuación, se presenta el código correspondiente para llevar a cabo la lectura de los archivos de texto generados:

```
for i in range(tamano):
    ruta_docente.append('Archivos/Docentes_SI/
docente_'+str(i)+'.txt')
i=0
for i in range(tamano):
    texto.append(leer_texto(ruta_docente[i]))
Con la ejecución de este código, se logrará la lectura de los archivos de texto generados, lo
```

cual es un paso fundamental para realizar las etapas posteriores del análisis de minería de texto.

Fase 2: Preprocesar los datos

Para llevar a cabo el preprocesamiento del texto, se aplican diversas técnicas, como la eliminación de palabras comunes (stopwords), la conversión a minúsculas y la eliminación de signos de puntuación, entre otras. Asimismo, en la etapa de vectorización del texto, se utiliza la técnica de word2vec para convertir el texto en vectores numéricos. A continuación, se presenta el código correspondiente para realizar la limpieza y vectorización de los datos:

```
i=0
for i in range(tamano):
    lista_textos = [texto_1, texto[i]]
    sw = lista_stopwords()
    texto_limpio = [limpieza_texto(q, lista_
palabras=sw, quitar_numeros=False, n_min=3,
quitar_acentos=True) for q in lista_textos]
A continuación, se procede a la vectorización del texto utilizando la técnica de word2vec. Se crea una instancia de la clase Word2Vec pasando como argumento la lista de textos preprocesados.
# Definir algunos vectorizadores para hacer
diferentes pruebas
v_word2vec = VectorizadorWord2Vec()
s_word2vec = Similitud(v_word2vec)
```

Con la ejecución de este código, se logra realizar el preprocesamiento del texto, incluyendo la eliminación de palabras comunes y la vectorización utilizando word2vec, lo cual es esencial para obtener representaciones numéricas adecuadas de los textos analizados.

Fase 3: Explorar datos

En la etapa de comparación de vectores, se llevó a cabo el cálculo de la similitud entre los vectores numéricos correspondientes a los dos textos. Para esta tarea, se utilizó una medida comúnmente empleada, como la similitud coseno, que permite evaluar el grado de similitud obtenido y determinar la correlación entre los

textos analizados. A continuación, se presenta el código que implementó dicha operación:

```
coseno_word2vec = s_word2vec.coseno(texto_
limpio[0], texto_limpio[1])
print('Similitudes coseno entre los textos de
contenido mínimo y perfil de '+ mostrar_
docente(i) +' (Word2Vec):')
print(coseno_word2vec)
```

Al ejecutar este código, se logra comparar los vectores y calcular la similitud entre los textos, brindando una medida cuantitativa de la correlación entre los mismos. Esto permite obtener información relevante acerca de la relación existente y contribuye al análisis y comprensión de los datos textuales en el ámbito de la minería de texto.

Fase 4: Obtención de resultado

En la fase siguiente, se exportaron los valores obtenidos en una matriz en formato xlsx, lo cual permitió realizar un análisis exhaustivo de los resultados obtenidos. A continuación, se presenta el código utilizado para la exportación de los datos:

```
ws.cell(row=i+2, column=1, value=mostrar_
docente(i))
ws.cell(row=i+2, column=2, value=str(coseno_
word2vec))
wb.save(r'Archivos/Asignatura_SI/Archivo_
Docente_carrera_SI.xlsx')
texto_limpio=""
```

Al exportar los resultados en formato xlsx, se facilita la realización de análisis posteriores mediante herramientas especializadas para el procesamiento y visualización de datos. Esto permite examinar detalladamente los resultados obtenidos y extraer conclusiones valiosas en el contexto de la minería de texto. Es necesario mencionar que el proceso de exportación de datos es fundamental para la documentación y replicabilidad de los resultados obtenidos en la investigación, lo que promueve la transparencia y la posibilidad de realizar análisis comparativos o complementarios en futuros estudios.

En este estudio, se seleccionaron cuidadosamente los elementos para determinar

la similitud entre los títulos de tercer y cuarto nivel profesional del docente y los contenidos mínimos de las asignaturas de las carreras de Sistema de Información y Tecnología de la Información y Comunicación. Para este propósito, se optó por utilizar los vectorizadores densos, como word2vec, en lugar de los vectorizadores de frecuencia TF-IDF. Esta elección se basó en la consideración de que los vectorizadores densos son menos radicales y permiten encontrar similitudes entre textos con significados similares, incluso si no comparten muchas palabras en común.

Además, se empleó la técnica de similitud coseno en combinación con la vectorización word2vec. La utilización de esta técnica posibilita medir el grado de alineación entre los dos vectores mediante un rango de valores entre -1 y 1. Un valor de 1 indica una similitud perfecta entre los textos, mientras que un valor de -1 indica una similitud negativa. Un valor de 0, por su parte, señala que los vectores son ortogonales o completamente diferentes.

La adopción de la técnica de similitud coseno en conjunto con la vectorización word2vec permitió obtener resultados más precisos y significativos en la medición de la similitud entre los docentes y las asignaturas. Esta combinación proporciona una representación numérica de los textos que captura su similitud semántica, trascendiendo la mera coincidencia de palabras.

Es importante destacar que esta elección metodológica permite una evaluación más precisa y sutil de la relación entre los docentes y las asignaturas, considerando no solo la presencia o ausencia de palabras clave, sino también su significado y contexto. Esto resulta especialmente relevante en el ámbito de la minería de texto, donde la comprensión y el análisis de la semántica son fundamentales para extraer información valiosa de grandes volúmenes de texto. Al utilizar los vectorizadores densos y la técnica de similitud coseno, se establece una base sólida para la evaluación de la relación entre los docentes y las asignaturas, abriendo nuevas oportunidades para comprender y mejorar la

calidad de la educación en el contexto específico estudiado.

Resultados

Una vez completado el tratamiento de los documentos, se procedió a la elaboración del algoritmo en el lenguaje de programación Python. A continuación, se presentan los resultados obtenidos mediante la aplicación del algoritmo al modelo de relación docente-asignatura utilizando la técnica de minería de texto. La implementación de la minería de texto en este estudio permitió analizar y extraer información valiosa a partir de los documentos examinados. Mediante el algoritmo desarrollado, se logró procesar y organizar eficientemente los datos extraídos de 107 asignaturas y 56 docentes, lo cual brindó una base sólida para la comprensión y el análisis de la relación entre los docentes y las asignaturas estudiadas.

En la Tabla 1 se presentan los valores obtenidos de la similitud entre los perfiles docentes y las carreras tras aplicar el algoritmo de minería de texto. Dichos resultados fueron exportados a un archivo con extensión xlsx. A partir de los datos generados, se elaboraron cuatro gráficos que proporcionan una descripción concisa del nivel de correlación porcentual entre el docente y la asignatura, así como la relación con las carreras de Sistemas de Información (SI) y Tecnologías de la Información (TI).

Tabla 1
Resumen de la data

Asignaturas	Docente 1	Docente 2	Docente 3	Docente 4	Docente 5	Docente 6
Desarrollo de aplicaciones Web	0.8300	0.6413	0.6526	0.6815	0.8632	0.6815
Aplicaciones Móviles	0.7929	0.7409	0.7380	0.7950	0.8049	0.7950
Desarrollo Basado en Plataformas	0.7763	0.7691	0.8161	0.8145	0.7611	0.8145
Seguridad Informática	0.7711	0.7487	0.6794	0.7487	0.7527	0.7487
F. de Tecnologías de Información	0.7707	0.7800	0.7518	0.8216	0.7517	0.8216
Control de calidad del Software	0.7654	0.7823	0.7998	0.8423	0.7386	0.8423
Arquitectura de Software	0.7522	0.7392	0.7965	0.7921	0.7416	0.7921
Auditoría Informática	0.7462	0.8016	0.7837	0.7901	0.6728	0.7901
Arquitecturas Empresariales	0.7455	0.7816	0.8141	0.7831	0.6850	0.7831
M. de desarrollo de software	0.7422	0.7847	0.8134	0.8167	0.6886	0.8167
A. de Calidad del Software	0.7415	0.7481	0.7966	0.7947	0.7043	0.7947
Informática Forense	0.7411	0.8003	0.7371	0.7832	0.6916	0.7832
Seminario de Titulación	0.7259	0.7251	0.7977	0.6867	0.6104	0.6867
Gestión de Proyectos de Software	0.7216	0.7442	0.8291	0.7469	0.6552	0.7469
Fundamentos de la Programación	0.7210	0.7157	0.7229	0.7742	0.7037	0.7742
Minería de Datos	0.7202	0.7175	0.6884	0.7679	0.7146	0.7679
Sistemas Operativos	0.7138	0.7792	0.7566	0.8160	0.6740	0.8160
A. de Bases de Datos	0.7134	0.7296	0.7020	0.7618	0.6979	0.7618
I. a la Investigación Científica	0.6990	0.7399	0.7663	0.6924	0.6018	0.6924
Computación en la Nube	0.6870	0.6744	0.6555	0.7172	0.6877	0.7172
Técnicas avanzadas de despliegue	0.6863	0.7483	0.7501	0.8027	0.6622	0.8027
D. del Trabajo de Titulación	0.6834	0.7028	0.7876	0.6686	0.5724	0.6686
Bases de datos	0.6772	0.6607	0.6143	0.7037	0.6824	0.7037

En las figuras 2, 3, 4 y 5 constituyen una representación visual efectiva de los hallazgos, permitiendo Figura 2

Afinidad de docentes con la carrera de SI y TI

Tabla 2
Afinidad de docentes con las carreras de SI y TI

Docente	SI	TI
Docente 1	89%	89%
Docente 2	89%	88%
Docente 3	87%	88%
Docente 4	85%	85%
Docente 5	84%	83%
Docente 6	84%	83%
Docente 7	84%	83%
Docente 8	84%	83%
Docente 9	84%	83%
Docente 10	84%	83%
Docente 11	83%	83%
Docente 12	83%	83%
Docente 13	83%	83%
Docente 14	82%	81%
Docente 15	81%	81%
Docente 16	81%	81%
Docente 17	81%	80%
Docente 18	80%	80%
Docente 19	79%	79%
Docente 20	76%	75%
Docente 21	76%	75%
Docente 22	75%	76%
Docente 23	76%	75%
Docente 24	75%	74%
Docente 25	69%	68%
Docente 26	67%	66%
Docente 27	66%	64%
Docente 28	60%	59%
Docente 29	60%	59%
Docente 30	90%	90%
Docente 31	89%	89%
Docente 32	87%	88%
Docente 33	87%	88%
Docente 34	87%	88%
Docente 35	86%	87%
Docente 36	86%	86%
Docente 37	86%	86%

Docente 38	86%	86%
Docente 39	85%	85%
Docente 40	83%	83%
Docente 41	81%	82%
Docente 42	81%	82%
Docente 43	81%	82%
Docente 44	81%	82%
Docente 45	81%	82%
Docente 46	81%	82%
Docente 47	81%	82%
Docente 48	81%	81%
Docente 49	80%	80%
Docente 50	80%	80%
Docente 51	79%	78%
Docente 52	78%	77%
Docente 53	78%	77%
Docente 54	78%	77%
Docente 55	74%	75%
Docente 56	73%	73%

En base a estos resultados, se puede concluir una fácil interpretación de los niveles de correlación encontrados. Cada gráfico se enfoca en un aspecto específico de la relación docente-asignatura, proporcionando una perspectiva detallada y útil para comprender los resultados obtenidos.

Las figuras 2 y 3 se centran en la relación entre la docencia y las carreras de Sistemas de Información (SI) y Tecnologías de la Información (TI). Estos gráficos proporcionan una visión más específica de cómo se relacionan los docentes y las asignaturas dentro de cada una de estas carreras, destacando los niveles de correlación encontrados.

Las figuras 4 y la figura 5 muestran la correlación porcentual entre el docente y la asignatura, brindando una visión general del grado de correspondencia entre ambos elementos. Esta información resulta fundamental como indicador de la calidad de la enseñanza y determinar posibles áreas de mejora.

En conjunto, estas representaciones gráficas permiten una comprensión más completa y visualmente atractiva de los resultados obtenidos mediante el análisis de minería de

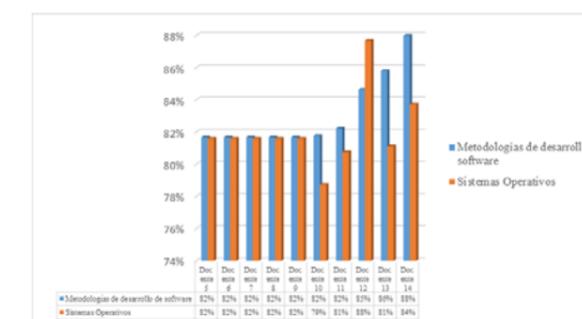
texto. Proporcionan una síntesis clara y concisa de la correlación existente entre el docente y la asignatura, así como su relación con las carreras de Sistemas de Información y Tecnologías de la Información. Estas visualizaciones constituyen una herramienta valiosa para la toma de decisiones informadas y la mejora continua de la calidad educativa en el ámbito investigado.

Para realizar el análisis de los datos obtenidos, se relaciona en base a los títulos de tercer y cuarto nivel profesional de los docentes de la carrera de Sistemas de información y los contenidos de las asignaturas de las carreras de Sistemas de Información y Tecnologías de la Información. En total, se evaluaron a 56 docentes, los resultados los observamos en la tabla 2.

que existe un promedio del 80% de correlación entre los docentes de la carrera de Sistemas de Información y los contenidos de las asignaturas de dicha carrera, así como los contenidos de la carrera de Tecnologías de la Información. Estos hallazgos indican una sólida correspondencia entre los docentes y los temas abordados en las asignaturas, lo cual es un indicador positivo en términos de la calidad educativa y la alineación entre la docencia y los programas de estudio.

Se aplicó el mismo procedimiento descrito anteriormente para relacionar a los docentes de la carrera de TI. En la figura 3 se presenta una representación visual de los resultados obtenidos, por motivos de espacio y tamaño del gráfico se muestran solo 10 docentes de los 56 y se comparan con 2 asignaturas de un total de 107. Los resultados indican que existe un promedio del 82% de correlación entre los docentes de la carrera de TI y los contenidos de las asignaturas de la misma carrera. Esta alta correlación demuestra la coherencia y alineación entre los docentes y los temas abordados en las asignaturas, lo cual es fundamental para garantizar una formación académica de calidad en el campo de la tecnología de la información.

Figura 3
Relación docente – asignatura



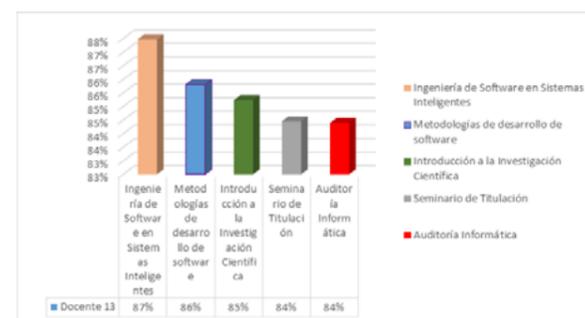
Para obtener los datos presentados en la Figura 3, se seleccionaron 10 docentes que mostraron los porcentajes de relación más altos en las asignaturas de Metodología de Desarrollo de Software y Sistemas Operativos. Estos docentes son una muestra representativa de la correlación existente entre los docentes y dichas asignaturas.

En particular, se encontró que el docente 14 alcanzó un porcentaje de correlación del 88% en la asignatura de Metodología de Desarrollo de Software. El docente 13 obtuvo un nivel de correlación del 86%, mientras que el docente 12 logró un porcentaje del 85%.

Asimismo, los docentes 5, 6, 7, 8, 9, 10 y 11 demostraron un nivel de correlación del 82% en ambas asignaturas, lo que indica una consistencia significativa entre su docencia y los contenidos de Metodología de Desarrollo de Software y Sistemas Operativos.

Podemos observar un promedio del 83% de analogía entre los docentes y la asignatura de Metodología de Desarrollo de Software, y una media del 82% en la asignatura de Sistemas Operativos. Estos resultados indican una correlación moderada a alta, lo que implica que los docentes comparten un porcentaje significativo de similitud en términos de contenido y formación profesional en ambas asignaturas. Esta coherencia entre los docentes y los temas abordados es esencial para una enseñanza efectiva y una formación académica de calidad en el campo de las tecnologías informáticas.

Figura 4
Relación segunda docente – asignatura



Para obtener los datos presentados en la Figura 4, se seleccionaron cuidadosamente 5 asignaturas con los porcentajes de relación más altos entre los docentes. En este caso, se ha tomado en consideración la destacada contribución del docente 13 en diferentes asignaturas.

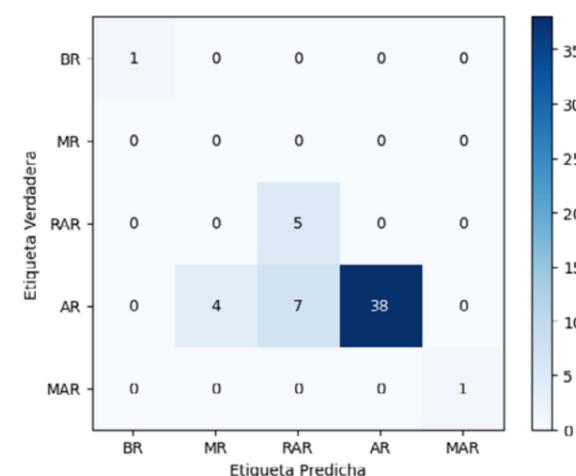
En la asignatura de Ingeniería de Software, el docente 13 obtuvo un porcentaje de relación del 87%. Además, logró un nivel de correlación del 86% en la asignatura de Metodologías de Desarrollo de Software, un 85% en Introducción a la Investigación Científica y un 84% en las asignaturas de Seminario de Titulación y Auditoría Informática. Estos resultados promedian un 85% de correlación con las asignaturas mencionadas, lo que indica una relación moderada a alta.

En cuanto al procedimiento de relación, se utilizó la técnica word2Vec y la similitud coseno. La utilización de word2Vec presenta ventajas significativas, ya que permite capturar información semántica y sintáctica en las representaciones vectoriales de palabras. Esta técnica es capaz de identificar similitudes entre palabras en función de su contexto y significado compartido. Además, se pueden realizar operaciones vectoriales interesantes, como la suma y resta de vectores, que proporcionan resultados relevantes. Por ejemplo, al realizar la operación “rey” - “hombre” + “mujer”, se obtiene el vector que se acerca más al concepto de “reina”. Las representaciones vectoriales obtenidas mediante word2Vec han demostrado su utilidad en diversas aplicaciones de procesamiento del lenguaje natural, como

búsqueda de información, clasificación de documentos, traducción automática y agrupación de textos, entre otras.

La evaluación del algoritmo de minería de texto aplicado en la relación docente-asignatura se llevó a cabo utilizando la matriz de confusión (figura 5). Se obtuvo una precisión del 80,36%, lo que indica que el algoritmo aplicado muestra un alto nivel de precisión.

Figura 5
Matriz de confusión



Cálculo de la precisión

$accuracy = accuracy_score(y_true, y_pred)$

Resultado Precisión:
0.8035714285714286

Discusión

En un estudio llevado a cabo por Strech et al. (2015), se ha demostrado la capacidad de predecir el éxito o fracaso de los estudiantes a través de la minería de texto, utilizándolo variables sociales como edad, sexo, estado civil, nacionalidad, experiencia previa de residencia fuera de la localidad, becas, necesidades especiales, tipo de admisión, tipo de estudiante, años de matrícula, cursos aplazados, tipo de dedicación y posibles situaciones de deuda. Estos factores son de suma importancia y pueden ejercer una influencia significativa en el rendimiento académico de los estudiantes. Aunque nuestros

resultados difieren en el enfoque y las variables analizadas, comparten la idea de que la minería de texto puede proporcionar información valiosa en el ámbito educativo. Estos estudios destacan la importancia de utilizar métodos de análisis de texto para comprender mejor los factores que influyen en el rendimiento académico y la relación entre docentes y contenidos educativos.

Al comparar nuestros resultados con los estudios mencionados, se puede observar que todos ellos se basan en la minería de texto como una herramienta para analizar y comprender diferentes aspectos en el ámbito educativo. Mientras que el estudio de Consuegra et al. (2016) se centra en la evaluación de estructuras argumentativas en trabajos académicos, nuestro estudio se enfoca en la relación entre docentes y asignaturas. Sin embargo, ambos estudios comparten la idea de que la minería de texto puede proporcionar información valiosa para evaluar y comprender la calidad y coherencia de los argumentos presentados en trabajos académicos, así como la relación entre los docentes y los contenidos educativos. Nuestros resultados muestran una relación significativa entre los docentes y las asignaturas analizadas, lo que respalda la importancia de una enseñanza efectiva y una formación académica de calidad en el campo de las tecnologías informáticas.

Los resultados de Zhañay (2016) como nuestro estudio utiliza la minería de texto para analizar datos no estructurados y obtener información relevante. Ambos hacen uso de técnicas avanzadas, como el modelo vectorial y word2Vec, para capturar información semántica y sintáctica en el análisis de los datos. Estos enfoques resultan valiosos para comprender las percepciones, opiniones y relaciones en diferentes contextos, ya sea en proyectos y servicios como el Tranvía Cuatro Ríos de Cuenca o en la relación entre docentes y asignaturas en el campo de las tecnologías informáticas, como se utiliza en nuestro estudio.

Los resultados obtenidos en nuestro estudio y en el estudio de Xiao et al. (2015) respaldan la utilidad de la minería de texto como una herramienta efectiva para obtener

información relevante y comprender la relación entre variables en diferentes contextos.

Comparando nuestros resultados con el estudio de Dikaiakos et al. (2012), se encontraron algunas similitudes y diferencias. Ambos estudios utilizan la métrica similitud coseno como una medida para evaluar la relación entre elementos, ya sean textos o vectores. En ambos casos, se obtienen valores de similitud que reflejan la correspondencia entre los elementos analizados. Aunque los porcentajes de coincidencia disminuyeron en nuestro estudio son específicos para nuestro contexto, ambos estudios resaltan la importancia de la alineación y correspondencia entre los elementos analizados. Los resultados obtenidos en ambos casos indican una sólida correspondencia entre los docentes y los contenidos de las asignaturas, lo cual es fundamental para una educación de calidad.

Después de llevar a cabo este proceso de recopilación de datos, se realiza el reconocimiento de entidades nombradas. Según Wauer et al. (2011), este procedimiento implica identificar las partes específicas del texto que ofrecen pistas relevantes para alcanzar el objetivo del análisis, al mismo tiempo que se descartan las palabras con múltiples significados que no se ajustan al contexto deseado. En el estudio actual, se emplea la función de similitud de la técnica de vectorización Word2Vec, la cual genera palabras o entidades relevantes para el análisis de la correlación entre el perfil docente y los contenidos mínimos de las asignaturas. Por otra parte, Ayyadevara (2018) indica que, en el contexto del análisis de similitud entre palabras, la clase VectorizadorWord2Vec permite acceder a representaciones pre-entrenadas que capturan información contextual de las palabras. Estas representaciones vectoriales densas son capaces de capturar relaciones semánticas y sintácticas entre palabras en un espacio de alta dimensionalidad. Por ejemplo, las palabras “hombre” y “niño” tendrán representaciones similares en dicho espacio, mientras que “hombre” y “cuchara” estarán más distantes.

Además, Talamé et al. (2019) expresa en su investigación el uso de técnicas como

la identificación de entidades nombradas y la utilización de librerías como Contexto en Python simplifican las tareas y proyectos de análisis de texto, proporcionando herramientas para la limpieza, traducción, análisis de similitudes y generación de visualizaciones, entre otras funcionalidades. Aunque los enfoques y las técnicas utilizadas son diferentes, ambos estudios resaltan la relevancia de utilizar métodos de minería de texto para analizar y comprender la relación entre los docentes y los contenidos educativos. Tanto nuestros resultados como los alcanzados en el estudio de Talamé et al. (2019) demostrar una relación significativa y una destacada relación en los aspectos analizados.

La minería de texto es una técnica valiosa para agrupar y clasificar patrones en grupos, permitiendo el análisis de grandes cantidades de datos textuales. Esta técnica, junto con el procesamiento del lenguaje natural (PLN), es capaz de identificar, comprender y generar texto, lo que resulta útil en diversas aplicaciones de minería de datos, como lo es la relación docente - asignatura. Los estudios y herramientas mencionados en este artículo científico demuestran la aplicabilidad de la minería de texto en diversas áreas, como la predicción del éxito académico, análisis de datos no estructurados, análisis de opiniones en redes sociales y procesamiento del lenguaje natural. Estas técnicas y herramientas brindan una comprensión más profunda de los datos textuales y permiten obtener conocimientos significativos para la toma de decisiones en diferentes contextos.

Conclusiones

La presente investigación constituye una referencia fundamental para abordar conceptualmente todos los elementos necesarios para comprender la minería de textos y desarrollar su implementación. En primer lugar, se destaca la importancia de la información y su potencial uso en este contexto. Además, la utilización de herramientas computacionales y técnicas de minería de texto resulta fundamental, ya que permite manejar grandes volúmenes de información de manera eficiente y automatizada.

También, proporciona la posibilidad de obtener resultados precisos y significativos, que de otra manera requerirían una cantidad considerable de tiempo y esfuerzo humano.

La aplicación de la minería de texto en el presente estudio ha demostrado ser una estrategia altamente efectiva para el procesamiento y análisis de documentos, lo cual ha facilitado la identificación de patrones y tendencias relevantes en la relación entre docentes y asignaturas. Los resultados obtenidos han revelado una conclusión del 80% y 82% entre los docentes y los contenidos de asignaturas de las carreras de Sistemas de Información y Tecnología de la Información, respectivamente. Estos representan un valioso aporte al avance del conocimiento en este ámbito, así como la apertura de nuevas perspectivas para mejorar la calidad educativa y la comprensión de esta relación en el contexto investigado.

Esta investigación proporciona una visión integral de la minería de textos y su aplicación práctica. Además, destaca la importancia de la información, su relevancia en el ámbito académico y su potencial uso en diversas áreas. La minería de textos representa una herramienta valiosa para extraer conocimiento útil y aprovechar al máximo la gran cantidad de datos textuales disponibles. También ha demostrado ser una estrategia efectiva para el procesamiento y análisis de documentos, facilitando la identificación de patrones y tendencias relevantes en la relación al docente – asignatura.

Referencias Bibliográficas

- Ayyadevara, V. K. (2018). Word2vec. En V. K. Ayyadevara, *Pro Machine Learning Algorithms* (pp. 167-178). Apress. https://doi.org/10.1007/978-1-4842-3564-5_8
- Barrio Pareja, E. E. B. (2014). Minería de texto para el análisis de la colaboración en SIENA. 5(10), 1-61.
- Benítez, C. A. F. (2016). Herramienta informática con minería de texto para la caracterización psicosocial de

estudiantes y grupos universitarios. 1-75.

- Betancourt Bethencourt, J. A., Acao Francois, L., & Álvarez Escoda, M. (2016). Entrenamiento de estudiantes de Medicina en investigaciones cualitativas y minería de texto durante el 2015 en Camagüey. *Educación Médica Superior*, 30(3), 669-677.
- Bin Mat, U., Buniyamin, N., Arsad, P. M., & Kassim, R. (2013). An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention. *2013 IEEE 5th Conference on Engineering Education (ICEED)*, 126-130. <https://doi.org/10.1109/ICEED.2013.6908316>
- Consuegra, A. A., Salazar, Y. M., García, J. H., & Vizcaino, D. H. (2016). MINERÍA DE TEXTO COMO UNA HERRAMIENTA PARA LA BÚSQUEDA DE ARTÍCULOS CIENTÍFICOS PARA LA INVESTIGACIÓN. . . pp., 7(1), 8.
- Contreras Barrera, M. (2015). Minería de texto: Una visión actual. *Biblioteca Universitaria*, 17(2), 129-138. <https://doi.org/10.22201/dgb.0187750xp.2014.2.72>
- Digiñón, L. B., Álvarez, M. M., Digiñón, L. B., & Álvarez, M. M. (2021). Experiencia de enseñanza-aprendizaje con aula virtual en el acompañamiento pedagógico debido al Covid-19. *Apertura (Guadalajara, Jal.)*, 13(1), 20-35. <https://doi.org/10.32870/ap.v13n1.1957>
- Dikaiakos, M. D., Katsifodimos, A., & Pallis, G. (2012). Minersoft: Software retrieval in grid and cloud computing infrastructures. *ACM Transactions on Internet Technology*, 12(1), 2:1-2:34. <https://doi.org/10.1145/2220352.2220354>
- Duarte, J., Talero, L., & Arias-Tabares, J. (2019). Identificación de estrategias académicas aplicando minería de texto: El caso de la asignatura Finanzas y Presupuestos. *Espacios*, 40, 9.
- Espíritu-Sandoval, K. Y. E. (2016). MAESTRO EN SISTEMAS COMPUTACIONALES. 2(3), 1-71.
- Florio, N. (2022). Las palabras de la educación española: Una experiencia de minería de textos: Análisis lingüístico diacrónico de las leyes educativas con R. *HUMAN REVIEW. International Humanities Review / Revista Internacional de Humanidades*, 13(1), Article 1. <https://doi.org/10.37467/revhuman.v11.4001>
- Garrido, C. D., Orta, G. del C. L., Rivilla, A. M., & Méndez, E. R. (2014). Las competencias docentes: Diagnóstico y actividades innovadoras para su desarrollo en un modelo de educación a distancia. *REDU. Revista de Docencia Universitaria*, 12(1), Article 1. <https://doi.org/10.4995/redu.2014.6431>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed). Elsevier/Morgan Kaufmann.
- Kansanen, P. (2003). Studying—the Realistic Bridge Between Instruction and Learning. An Attempt to a Conceptual Whole of the Teaching-Studying-Learning Process. *Educational Studies*, 29, 221-232. <https://doi.org/10.1080/03055690303279>
- Kumar, S., & Saeed, M. (2018). *Applications of Educational Data Mining and Learning Analytics Tools in Handling Big Data in Higher Education: Trends, Issues, and Challenges*. https://www.researchgate.net/publication/326554623_Applications_of_Educational_Data_Mining_and_Learning_Analytics_Tools_in_Handling_Big_Data_in_Higher_Education_Trends_Issues_and_Challenges
- Lang, C., Siemens, G., Wise, A., Gašević, D., & Merceron, A. (2017). Handbook of Learning Analytics—Second edition. *Society for Learning Analytics Research (SoLAR)*. <https://www.solaresearch.org/publications/hla-22/>
- Navarro, I. S. (2021). Prototipado Rápido de Minería de Texto Usando Campos Conceptuales Fast Prototyping for Text

- Mining Using Conceptual Fields. 2(25), 11-38.
- Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*, 41(4, Part 1), 1432-1462. <https://doi.org/10.1016/j.eswa.2013.08.042>
- Pérez Abelleira, M. A., & Cardoso, C. A. (2010). Minería de texto para la categorización automática de documentos. 5, 11-45.
- Rivero, S., David, V., Humberto, J., Ester, M., Alejandro, L., Alejandro, A., Brian, C., Facundo, Á., Nicolás, P., & Noelia, C. (2020). Aplicación de Técnicas Descriptivas de Minería de Textos sobre Contenido Digital Realizando Análisis Inteligente. 2(5), 279-283.
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27. <https://doi.org/10.1002/widm.1075>
- Ruiz-Cornejo, N., Arriagada-Benítez, M., Ruiz-Cornejo, N., & Arriagada-Benítez, M. (2022). Análisis estilométrico en Python aplicado a epistolarios de la premio nobel de literatura Gabriela Mistral. *Ingeniare. Revista chilena de ingeniería*, 30(2), 229-238. <https://doi.org/10.4067/S0718-33052022000200229>
- Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, 72, 414-422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Strech, P., Cruz, L., Soares, C., Mendes, J., & Abreu, R. (2015). *A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance*. 3. https://www.researchgate.net/publication/278030689_A_Comparative_Study_of_Classification_and_Regression_Algorithms_for_Modelling_Students%27_Academic_Performance
- Talamé, L., Cardoso, A., & Amor, M. (2019). *Comparación de herramientas de procesamiento de textos en español extraídos de una red social para Python*. 53-67.
- Temesio, S., GarcÁa, S., & Perez, A. (2021). Rendimiento estudiantil en tiempo de pandemia: Percepciones sobre aspectos con mayor impacto. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, 28, 359-369.
- Vargas Guarín, J. D., Barona Jiménez, M. A., & Garcia Beltrán, F. E. (2021). *Doctor mining: Sistema text mining para localizar información sobre enfermedades pulmonares*. <https://repositorio.unbosque.edu.co/handle/20.500.12495/7146>
- Vargas Pulliquitín, M. F. (2018). *Minería de texto de la web, de opinión pública y hechos referentes al barrio La Floresta* [BachelorThesis, Quito, 2018.]. <http://bibdigital.epn.edu.ec/handle/15000/19574>
- Wauer, M., Meinecke, J., Schuster, D., Konzag, A., Aleksy, M., & Riedel, T. (2011). Semantic Federation of Product Information from Structured and Unstructured Sources. *IJBDCN*, 7, 69-97. <https://doi.org/10.4018/jbdcn.2011040105>
- Xiao, X., Xu, H., & Xu, S. (2015). Using IBM SPSS modeler to improve undergraduate mathematical modelling competence. *Computer Applications in Engineering Education*, 23(4), 603-609. <https://doi.org/10.1002/cae.21632>
- York, T. T., Gibson, C., & Rankin, S. (2015). *Defining and Measuring Academic Success*. 20(5), 1-21. <https://doi.org/10.7275/HZ5X-TX03>
- Zhañay, M. B. A. (2016). *MINERÍA DE TEXTO EN MEDIOS SOCIALES. CASO DE ESTUDIO DEL PROYECTO TRANVÍA DE CUENCA*. Universidad del Azuay.

Generación de Engagement en los Estudiantes a través de Elementos Estratégicos de Rebranding en el Instituto Tecnológico Universitario EuroAmericano

Engagement Generation in Students Through Strategic Elements of Rebranding at the EuroAmerican University Technological Institute

Carlos Alberto Marques-Gutiérrez¹
Instituto Superior Tecnológico Universitario EuroAmericano - Ecuador
cmarques@euroamericano.edu.ec

Valeria Alejandra Almao-Malvacias²
Universidad de Guayaquil - Ecuador
valeria.almaom@ug.edu.ec

Gilma Leonor Vallejo-Piza³
Instituto Superior Tecnológico Universitario EuroAmericano - Ecuador
Tafur Consulting
gvallejo@euroamericano.edu.ec

doi.org/10.33386/593dp.2023.5.2079

V8-N5 (sep-oct) 2023, pp. 999-1009 | Recibido: 7 de agosto de 2023 - Aceptado: 24 de agosto de 2023 (2 ronda rev.)

1 Gerente del Centro Estratégico de Comunicación IntelMarketing. Con más de 10 años de experiencia elaborando proyectos integrales de comunicación con las nuevas tendencias tecnológicas del mercado nacional e internacional. Director de Marketing y Publicidad en Coordinación con el Departamento de Ventas del Tecnológico Universitario EuroAmericano. ORCID: <https://orcid.org/0009-0007-4851-638X>

2 Productora en multimedia. Asistente del área de investigación en Formación Educativa FETAFUR Tafur Consulting S.A. Formador de Formador de Formadores. Miembro de los proyectos del Plan de Investigación 2021 de la Universidad de Guayaquil. Ponente en el 1º Congreso Internacional de Vinculación con la Sociedad Universidad de Guayaquil 2021. 0000- ORCID: <https://orcid.org/0001-8983-2070>

3 Lcda. en Ciencias de la Educación con mención en Artes. Docente en el Instituto Universitario Euroamericano, de las materias Fundamentos y teoría del diseño, Apreciación Artística. Asistente del área de investigación y formación educativa en Tafur Consulting. Formador de Formadores. ORCID: <https://orcid.org/0000-0002-6205-590X>