ISBN: 978-628-96613-1-6. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2025.1.1.876

Powering a non-structured text search application with natural language processing

Carlos Agudelo-Santos[1]; Jose Isaac Zablah[2]
[1,2]Universidad Nacional Autónoma de Honduras, Honduras, *carlos.agudelo@unah.edu.hn, jose.zablah@unah.edu.hn*

Abstract*– Many organizations find data reduction and analysis complex and costly. The Teacher Performance Assessment (VDD) process evaluates student satisfaction through open-ended questionnaires, using an algorithm with regular expressions for searching. However, this method may be outdated due to recent advancements in natural language processing (NLP) and artificial intelligence. These technologies can improve the analysis of unstructured text in the VDD by applying computational hermeneutics and NLP techniques. Tests indicate that NLP enhances the contextual search of relevant terms, yielding more accurate teacher evaluations while minimizing false positives and negatives. However, NLP implementation is more costly and time-consuming, making it suitable only for larger datasets and complex grammatical structures. Traditional algorithms remain effective for smaller datasets and simpler structures with limited computational resources.

Keywords*-- Hermeneutics, algorithms, natural language processing, unstructured text

Mejora de una aplicación para búsqueda de texto no estructurado mediante procesamiento de lenguaje natural

Carlos Agudelo-Santos[1]; Jose Isaac Zablah[2]
[1,2]Universidad Nacional Autónoma de Honduras, Honduras, *carlos.agudelo@unah.edu.hn, jose.zablah@unah.edu.hn*

Resumen– ***La reducción y análisis de datos es una tarea compleja y costosa para muchas organizaciones, en este articulo se ha considerado el proceso de Valoración del Desempeño Docente (VDD), la cual es una herramienta que evalúa la satisfacción de los estudiantes mediante un cuestionario de preguntas abiertas en la Universidad Nacional Autónoma de Honduras (UNAH). El procesamiento se realiza con un algoritmo que utiliza expresiones regulares para la búsqueda de patrones previamente definidos en categorías. Sin embargo, esta solución podría estar desfasada debido a los avances en procesamiento de lenguaje natural (PLN) y la revolución de la inteligencia artificial en los últimos cinco años. Estas tecnologías se proponen para mejorar el análisis de texto no estructurado del VDD, aplicando técnicas de hermenéutica computacional y PLN. Las pruebas han demostrado que el PLN permite ampliar la búsqueda de palabras en su contexto, ofreciendo resultados más sensibles en la evaluación docente y reduciendo falsos positivos y negativos. No obstante, su implementación es más costosa y requiere datos de entrada más grandes y con estructuras gramaticales complejas. El uso de algoritmos sigue siendo viable con datos cortos y estructuras limitadas, combinados con pocos recursos computacionales.***

Palabras clave-- Hermenéutica, algoritmos, lenguaje natural, texto no estructurado.

I. Introducción

Con el advenimiento de la computación en la nube y el uso de servicios informáticos sobre Internet, monetizados en base a su consumo y adaptándose a las necesidades de los usuarios; esta característica se le conoce como elasticidad en el consumo bajo demanda [1]. Estos servicios pueden incluir infraestructura (IaaS), plataforma (PaaS) o software (SaaS), se pueden acceder desde cualquier lugar con una conexión a Internet óptima, lo que se le conoce como la ubiquidad de la nube [2]. Las características propias de este modelo permiten ahorrar costos, aumentar la eficiencia y la seguridad de los datos, ya que se enmarcan en lo que se conoce como acuerdos a nivel de servicio (SLA), que establecen la interacción usuario y proveedor. La nube esta soportada en los conceptos de arquitectura orientada a servicios (SOA) y la tecnología de virtualización [3][4].

En el año 2016, se publicó el artículo "*Performance limitations of a text search application running in cloud instances*" [5], donde se describe el rendimiento de una solución en MySQL ejecutada en una nube construida utilizando hardware común, con el fin de identificar los cuellos de botella en la ejecución de un conjunto de *scripts* desarrollados sobre el estándar SQL. Los mismos se diseñaron para realizar búsquedas de términos que representan fragmentos o palabras completas agrupadas por categoría predefinidas, los *scripts* operaban sobre texto no estructurado almacenado en registros de tablas en el gestor de base datos mencionado.

En esa misma publicación, se compararon dos tipos de plataformas: una máquina física que sirve como anfitrión y una instancia dentro de una infraestructura de una nube privada. Los resultados muestran que el uso intensivo de una base de datos relacional presenta una mayor pérdida de rendimiento en una instancia en la nube por las limitaciones del sistema de almacenamiento primario; que, al ser comparado con en el almacenamiento local, presenta pérdidas sustanciales de rendimiento. Lo anterior hace que el proceso de ejecución de los scripts requiera más tiempo para completar su tarea. Los datos de entrada fueron más de un millón y medio de registros; con un tiempo medio de procesamiento de 6,295 segundos por cada una de las cinco categorías evaluadas.

La hermenéutica es un método de interpretación de textos que se centra en el significado y el contexto del escrito, este concepto sostiene que el significado de un texto no es estático, sino que está determinado por el entorno en el que se produce el texto. La hermenéutica se ha utilizado durante siglos para interpretar textos religiosos, literarios y filosóficos, en los últimos años, se ha utilizado para desarrollar herramientas de análisis de texto no estructurado [6] [7].

Los algoritmos de procesamiento de lenguaje natural (PLN) suelen utilizar estadísticas para identificar temas y patrones en los textos junto con la hermenéutica. Sin embargo, las estadísticas pueden llegar a ser engañosas en el caso de los textos no estructurados. Por ejemplo, un texto puede contener muchas palabras que están relacionadas con un tema, pero el texto en sí puede no tratar realmente sobre ello. Los algoritmos de PLN también suelen utilizar reglas predefinidas

para comprender el significado de los textos; sin embargo, estas pueden ser limitadas en su capacidad para comprender su origen. En este sentido, un texto puede utilizar un lenguaje figurado o una metáfora, que no pueden ser entendidas por las reglas predefinidas [8][9][10].

El modelo PLN utilizado para procesar las opiniones de los estudiantes en este artículo se llama RoBERTa, acrónimo de "*A Robustly Optimized BERT Pretraining Approach*". Es definido como un modelo de procesamiento de lenguaje natural (PLN) que se basa en la arquitectura de BERT (*Biceived Encoder Representations from Transformers*) pero con mejoras significativas en su entrenamiento y rendimiento. BERT fue desarrollado en 2018 por Google AI Languages y ha demostrado ser una herramienta útil en diversas tareas de análisis de sentimientos y procesamiento de textos [11].

La utilidad de RoBERTa para el análisis de sentimientos radica en su capacidad para comprender y representar el contexto, es decir, para poder utilizar con precisión el significado de las palabras en un texto. A través de un proceso de entrenamiento más exhaustivo y una mayor cantidad de datos, RoBERTa logra una comprensión más profunda del lenguaje natural y las relaciones semánticas entre palabras, ya que sus autores han utilizado más de 161 GB de datos de entrenamiento inicial [12].

El objetivo es tener una base apta para comparar la eficacia y viabilidad computacional de un enfoque tradicional basado en algoritmos con expresiones regulares frente a técnicas avanzadas de procesamiento de lenguaje natural (PLN), específicamente mediante el uso del modelo XLM-RoBERTa, para el análisis de texto no estructurado en el contexto de la prueba de Valoración del Desempeño Docente (VDD) de la Universidad Nacional Autónoma de Honduras (UNAH), considerando métricas de sensibilidad, especificidad, tiempo de procesamiento y requerimientos de infraestructura tecnológica.

II. MÉTODOS

A. *Procesamiento de lenguaje natural*

Para estudiar la satisfacción de los estudiantes con sus docentes, se utilizará una técnica conocida como análisis de sentimientos que es una aplicación del procesamiento de lenguaje natural (PLN). Al emplear RoBERTa se puede identificar matices y tonos emocionales en el texto con gran precisión. Se podrá distinguir entre sentimientos previamente definidos, así como comprender la intensidad de ellos. Su mayor utilidad, se basa en que fue diseñado para evaluar la satisfacción del cliente a partir de reseñas típicamente obtenidas en línea, así como para el uso por investigadores que buscan analizar la percepción del público sobre ciertos temas o para cualquier aplicación que requiera la comprensión y clasificación de opiniones basadas en sentimientos en textos [13].

La capacidad de RoBERTa para captar el contexto y la información implícita en el lenguaje lo convierte en una herramienta eficaz para procesar tareas complejas, como ser la detección del sarcasmo, identificar sentimientos encontrados y comprender textos largos. Además, se cuenta con varios modelos pre-entrenados y maduros, que están disponibles en diferentes idiomas. Esta versatilidad ha generado versiones específicas, lo que han convertido a RoBERTa en una herramienta versátil para aplicaciones de análisis de sentimientos en todo el mundo. Para este estudio se utilizó la versión pre-entrenada de RoBERTa llamada *Twitter-roBERTa-base for Sentiment Analysis* actualizada al 2022, entrenada con más de 124 millones de Tweets entre enero de 2018 y diciembre de 2021; esto se ha ajustado para el análisis de sentimientos como referencia la variante TweetEval [14][15].

En nuestro caso, para el idioma castellano, se utilizó el modelo basado en XLM-roBERTa entrenado con alrededor de 198 millones de tweets y ajustado para el análisis de emociones en más de cien idiomas incluyendo el español, con la suficiente capacidad de clasificar sentimientos como ser ira, disgusto, miedo, alegría, tristeza, sorpresa y otras. Con mayor precisión, XLM-roBERTa se le entrenó previamente para realizar Modelado de Lenguaje Enmascarado (MLM). Durante este proceso, el modelo selecciona aleatoriamente el 15% de las palabras de una oración y las enmascara; al ejecutarse este modelo debe predecir las palabras ocultas y sus aciertos afectan el índice de precisión. A diferencia de las redes neuronales recurrentes tradicionales (RNN), que suelen analizar las palabras de manera secuencial, o de los modelos autorregresivos como GPT; que utilizan tokens. En cambio, los MLM emplean un enfoque que permite que el modelo adquiera una representación bidireccional de la oración [16].

B. *Pruebas de rendimiento*

Los *benchmark* o pruebas de rendimiento en ciencias de la computación, así como en otras áreas del conocimiento, son una serie de pruebas estandarizadas que se utilizan para medir los valores máximos de operación de un sistema o en alguno de sus componentes. El objetivo de un benchmark es proporcionar una forma objetiva y repetible de comparar el rendimiento de diferentes sistemas o componentes frente a otros. Esto permite a los usuarios y fabricantes tomar decisiones informadas sobre la elección de hardware y software [17].

La prueba de rendimiento del sistema de archivos utilizada es IOzone, el cual proporciona información crucial sobre las tasas de transferencia del sistema de archivos o del dispositivo de almacenamiento sometido a pruebas de lectura secuencial, lectura aleatoria, escritura secuencial y escritura aleatoria; entre otros. Estos resultados son esenciales para evaluar la eficiencia y la capacidad de respuesta del sistema de almacenamiento en diferentes escenarios de uso, donde los valores mayores indican mejores resultados. La tasa de

transferencia se mide en kilobytes por segundo (KBytes/s) [18].

Para las pruebas de lectura y escritura secuencial, una tasa de transferencia alta es beneficiosa para operaciones que requieren acceso secuencial a grandes conjuntos de datos, como es el caso de videos o imágenes crudas. En lo referente a las pruebas aleatorias de lectura y escritura estas evalúan la velocidad a la que el sistema puede escribir o leer datos que están dispersos en el dispositivo de almacenamiento. Las mediciones de transferencia de datos aleatorias indican la capacidad del sistema para trabajar con archivos almacenados de forma dispersa en la unidad de almacenamiento, ya que esto es común en archivos pequeños como ser archivos de texto, bases de datos relacionales, archivos en tiempo real, entre otros [18].

Para medir la capacidad de cálculo del anfitrión, se utilizó la prueba de rendimiento sintética Linpack, esta es una biblioteca de software utilizada para realizar cálculos numéricos intensivos, especialmente en el ámbito de la computación científica y técnica. Su implementación más conocida es HPL (*High Performance Linpack*), que evalúa el rendimiento de sistemas informáticos en el ámbito de la computación de alto rendimiento (HPC). Linpack mide el rendimiento de la computación de coma flotante de un sistema informático, esto es evaluar la velocidad con la que una computadora puede realizar cálculos que involucran números con decimales. Lo anterior importante de conocer, porque es el procesamiento más común en problemas científicos. Los resultados se expresan en múltiplos de FLOPS (*Floating-point Operations Per Second*), que corresponde a la cantidad de operaciones de coma flotante que la computadora puede realizar por segundo. A mayor puntaje de FLOPS, mayor capacidad de cómputo del sistema [19][20].

C. *Datos y herramientas de software*

Los datos utilizados se originaron del proceso de Valoración del Desempeño Docente (VDD) de la Universidad Nacional Autónoma de Honduras (UNAH) al finalizar el tercer período académico del 2015, estos son los mismos datos utilizados por artículo de "*Performance Limitations of a Text Search Application Running in Cloud Instances*" [5]. La muestra es obtenida por conglomerado, tomando solamente las respuestas del período de interés que coinciden con el artículo mencionado.

Esta muestra se compone de 399,808 registros donde se almacenan las respuestas a tres preguntas específicas de la VDD, como ser la primera (P1), "¿Ha identificado usted en su docente alguna actitud o práctica que no sea apropiada en un profesor universitario? Por favor explique...", la segunda (P2) es "¿Qué cualidades personales y profesionales identifica usted en el docente?"; y finalmente la tercera (P3) "A su criterio, ¿En qué aspectos del desempeño puede mejorar el docente?". Las respuestas han sido capturadas mediante un instrumento tipo cuestionario con respuestas abiertas. No se ha utilizado ningún criterio de exclusión ni se han hecho clasificaciones o categorías por datos demográficos, género, carrera ni otra. Es un tipo de estudio retrospectivo con un método de investigación mixto.

Para realizar las pruebas de PLN se utilizó el lenguaje de programación Python v3.10.3 en conjunto con el editor de texto Notepad++ v7.9.3 para la codificación y parametrización del modelo empleado. Las bibliotecas utilizadas están disponibles públicamente, como ser Matplotlib v3.7, NLTK v3.8.1, NumPy v1.18.; Pandas v2.1, SciPy v1.10.1, Seaborn v0.12.2; Torch v2.0.1, TQDM v4.66.1 y Transformers 4.32.1, entre otros [21][22][23].

El código desarrollado en Python se ejecutó en una instancia computacional virtualizada desplegada en la nube computacional de Akamai [24], esta contaba con el sistema operativo CentOS Stream 9 de 64bit [25], con 2 núcleos de CPU, 4GB de RAM, 80GB de almacenamiento SSD y conectado a Internet a una velocidad de 1Gbps. Al consultarle al sistema operativo en el nodo virtual sobre las especificaciones de CPU nos devuelve que la máquina física cuenta con un AMD EPYC 7713 64-core.

Los datos de entrada utilizados estaban en un archivo de texto separado por comas. Como resultado de su procesamiento se obtienen los valores medios de las evaluaciones de sentimientos como ser alegría, sorpresa, tristeza, enfado, disgusto, miedo y otros; generados al utilizar el modelo *Twitter-roBERTa-base for Sentiment Analysis* en su variante *XLM-roBERTa* [12][14].

III. Resultados

A. *Pruebas de rendimiento*

Sobre la infraestructura virtualizada, se ejecutó IOZone. La prueba de lectura secuencial midió una tasa de transferencia de datos de 10,767,122 KBytes/s. En la prueba de lectura aleatoria los resultados revelaron una tasa de 9,624,297 KBytes/s. En la prueba de escritura secuencial, se obtuvo una capacidad de transferencia de datos de 800,935 KBytes/s. En la prueba de escritura aleatoria se obtuvo como resultados una tasa de 3,142,473 KBytes/s. Es importante mencionar que una mayor tasa de transferencia en cualquiera de las pruebas es un indicador de una capacidad mayor del sistema evaluado para escribir o leer grandes volúmenes de datos. Los resultados a nivel gráfico se observan en la Figura 1.

Los resultados obtenidos al ejecutar la prueba de rendimiento sintética Linpack en el equipo anfitrión estimó una tasa de cálculo de aproximadamente 51 GFLOPS (GigaFLOPS). La prueba se completó en un tiempo de ejecución de 652 segundos, resolviendo un problema de tamaño 22611x22611 y la memoria utilizada se mantuvo abajo de los 4GB.

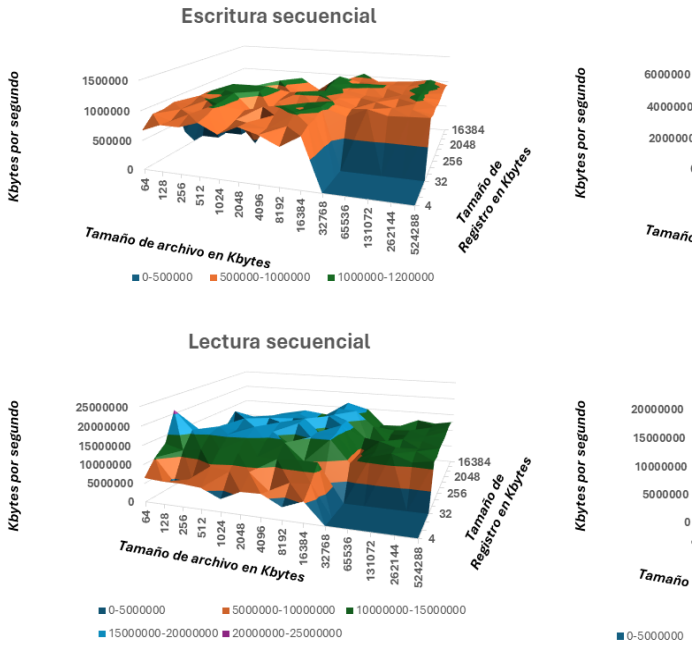

Fig. 1 Representación de las tasas de transferencia de datos (KBytes/s) en operaciones secuenciales y aleatorias de escritura/lectura

B. Análisis de texto no estructurado

El procesamiento de los datos mediante PLN, requirió 6.55 horas para completarse. Se procesaron 399,807 registros que corresponden a 133,269 respuestas por pregunta. No se excluyó ningún registro. Se evaluaron a 2,336 docentes que sirven 5,175 secciones; con un instrumento utilizado por más 22,137 estudiantes. No se ha aplicado ningún criterio de exclusión.

Se usó la media aritmética para conocer el valor de cada sentimiento evaluado sobre las respuestas recibidas por los docentes. Los resultados, con fines de este estudio se presentan de forma global sobre todo el conjunto de datos que conforman el texto no estructurado del instrumento VDD, en este sentido hay consideraciones de protección de datos personales.

Los resultados han sido muy diferenciados para cada pregunta, para el caso de la pregunta P1, dominó un sentimiento no determinado (otros), seguido de cerca por enfado, alegría y miedo. Para la pregunta P2, el sentimiento de alegría fue ampliamente dominante y para la pregunta P3; el modelo no ha sido útil ya que no determinó los sentimientos, ya que la categoría otros ha sido la dominante por un margen muy amplio. Estos resultados se pueden observar en la Figura 2.

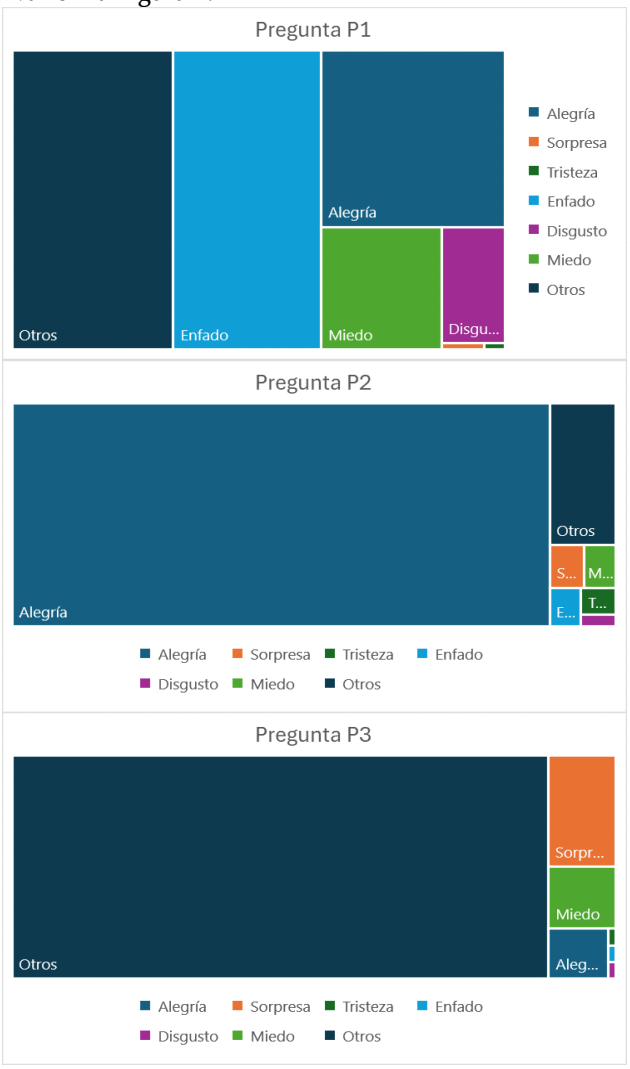

Fig. 2 Los gráficos de rectángulos representan la jerarquía de los sentimientos dominantes en cada pregunta conforme los resultados del PLN

Todos los valores (también llamados *scores*) obtenidos por los sentimientos evaluados por el modelo PLN pueden tomar valores entre cero y uno. Estos al acercarse a la unidad son un indicador que ese texto ha devuelto una sensibilidad por un sentimiento determinado. Si la valoración es cercana a cero, es indicador que ese sentimiento no está presente en el texto. Se evalúan todos los sentimientos, pero se toma el que devuelve un valor más alto como el sentimiento dominante. Para fines de este trabajo se realizó la media de todos sentimientos de cada una de las respuestas para generar una

visión global como la mostrada en la Figura 2. En cuanto al tiempo de ejecución el PLN le tomó 0.059 segundos por registro para completar el análisis.

IV. Discusión

El análisis de texto no estructurado realizado mediante algoritmos implementados con scripts y ejecutados al interior del sistema de gestión de bases de datos [5] toma 0.021 segundos procesar cada registro frente al lenguaje natural que utiliza 0.059 segundos. Así mismo no se ha medido el tiempo de implementación y puesta en marcha, ya que en el PLN se requiere de la instalación de múltiples librerías y configuraciones específicas del entorno de ejecución, en cambio para la ejecución al interior un gestor de base de datos no requiere seteos extensos; más que la copia de los scripts [26].

El tiempo requerido para completar la reducción de datos con PLN fue 6.55 horas (393 minutos) frente a lo reportado por el algoritmo de VDD que fue de 6.11 horas (367 minutos). Esto es un indicado que la tarea de reducción de datos con el algoritmo es más rápida, pero menos flexible. En cambio, al utilizar PLN se obtiene mayor especificidad porque evalúa sentimientos, pero a un costo de capacidad computacional y de tiempo mayor [27].

Al comparar la plataforma anterior utilizada para la VDD frente a la de este artículo se encontró que existe una diferencia entre los resultados de la prueba de rendimiento del sistema de archivos IOZone. Se encontró que la plataforma nueva cuenta con una capacidad de transferencia más alta; en la lectura secuencial y aleatoria es mil veces superior, en escritura secuencial es setecientas veces superior; en escritura aleatoria es doscientas ochenta veces mayor a la utilizada por el instrumento VDD. Esto es una evidencia del cambio tecnológico en un período de casi una década, desde la publicación del artículo original que se tomó como referencia para este trabajo [5].

Los resultados de la prueba sintética Linpack devolvió una capacidad de cálculo promedio de 51 GFLOPS usando dos núcleos virtualizados. En cuanto a la capacidad de microprocesador al menos a nivel teórico hay una diferencia superior a 35 GLOPS teóricos que daba el procesador de la instancia física empleada en el artículo de referencia, el cual era un Intel Core i7 Q740; con una capacidad de 16 GFLOPS [28].

Una posible limitación del modelo de procesamiento de lenguaje natural (PLN) utilizado radica en la dependencia del corpus de entrenamiento, que puede no reflejar con precisión el contexto cultural o lingüístico específico de los estudiantes hondureños. Esto puede generar sesgos en la clasificación de sentimientos, especialmente ante expresiones coloquiales, ambigüedades o usos figurativos del lenguaje. Asimismo, la predominancia del sentimiento "otros" en una de las preguntas evaluadas sugiere una posible incapacidad del modelo para capturar categorías emocionales relevantes no contempladas en su entrenamiento original. Finalmente, el uso de modelos pre-entrenados limita la adaptabilidad del PLN a dominios altamente específicos sin un proceso adicional de ajuste fino ("*fine-tuning*") o ajuste contextualizado.

IV. Conclusión

El PLN requiere una cantidad de recursos computacionales altos para poder implementarse. Una alternativa sería la utilización de infraestructuras compartidas como la usada en este artículo y que coincide con lo planteado en el trabajo previo de análisis de texto mediante algoritmos. El costo de implementación es directamente proporcional a las capacidades computaciones contratadas, siendo esto la característica elástica del paradigma de la computación en la nube.

La sensibilidad del procesamiento mediante algoritmos es menor, ya que el PLN permite agrupar por sentimientos específicos, en cambio en el método anterior se hace un cálculo aritmético con una alta posibilidad de verse afectado por valores extremos y entregar falsos positivos y falsos negativos al momento de evaluar las respuestas al evaluar los docentes.

Al comparar el método mediante algoritmo y PLN, concluimos que el primero debería de utilizarse en condiciones de limitación de recursos y con datos de entrada con muy poca estructura gramatical y extensión. El segundo debe considerarse para textos de mayor longitud y con estructuras gramaticales más definidas; las diferencias en recursos de procesamiento y tiempo son grandes; por lo que el usuario que lo implemente debe tomar esta decisión al verificar los datos de entrada.

Casi una década después, las capacidades computacionales son muy superiores a las de la publicación "*Performance limitations of a text search application running in cloud instances*" [5]; las capacidades computacionales han aumentado, pero han sido proporcionales a la complejidad de implementación; la diferencia primordial es que el PLN permite mejoras incrementales según se amplían las capas de la red neuronal utilizada. Los análisis mediante algoritmos son difíciles de ajustar a nuevas necesidades.

El uso de PLN para el análisis de las respuestas del instrumento del VDD puede devolver resultados favorables en cuanto a la especificidad de los resultados, ayudando a la interpretación de estos con los objetivos de evaluación. Aunque su implementación es compleja y costosa, la evolución tecnológica juntos los continuos desarrollos en inteligencia artificial son la clave para modelos de análisis y procesamiento más avanzados en el futuro requiriendo una menor intervención humana. Debería de considerarse en trabajo futuro entrenar un modelo PLN para las necesidades de evaluación del desempeño docente a nivel universitario.

El análisis de texto no estructurado mediante scripts que contienen algoritmos o que usan técnicas monolíticas no ofrecen la flexibilidad que se requiere en la actualidad, su mantenimiento y mejora resulta compleja. Aunque estos se

ejecutan mejor en infraestructuras modestas, los resultados requieren de una mayor sensibilidad y están sujetos a interpretación.

Las universidades se pueden beneficiar de la adopción de modelos avanzados de PLN como XLM-RoBERTa, a través de una interpretación más precisa de respuestas abiertas en procesos de evaluación de diversa índole, al identificar matices afectivos y reducir errores en la clasificación semántica en encuestas. Esta mejora metodológica contribuye a una retroalimentación más específica y útil para optimizar el desempeño académico y el control de calidad.

Para su integración institucional, se recomienda una implementación progresiva que considere las capacidades computacionales disponibles, la capacitación del personal técnico y los aspectos éticos relacionados con el manejo automatizado de información sensible. La vinculación del PLN con sistemas existentes de gestión académica podría consolidar procesos más eficientes y con mayor valor analítico para la toma de decisiones. Sería un proceso estratificado, de preferencia iniciado en facultades y carreras de menor a mayor complejidad, afinando sobre todo el proceso y las métricas conforme se avanza en su uso.

Como línea de trabajo futuro, se plantea el desarrollo de un modelo personalizado de PLN entrenado con datos locales provenientes de procesos de evaluación docente. Esta estrategia permitiría ajustar el modelo a las particularidades lingüísticas y culturales del entorno universitario nacional, mejorando la precisión en la detección de sentimientos y la interpretación contextual del discurso estudiantil. El entrenamiento con datos propios facilitaría la incorporación de categorías emocionales relevantes y reduciría sesgos asociados a modelos pre-entrenados en contextos ajenos. Su implementación requerirá la construcción de un "corpus anotado", infraestructura computacional especializada y el cumplimiento de criterios éticos para el manejo de información sensible.

AGRADECIMIENTOS

Este articulo ha recibido el apoyo de la Dirección de Investigación Científica Humanística y Tecnológica de la Universidad Nacional Autónoma de Honduras (DICIHT/UNAH), quienes por medio de un estipendio permitieron cubrir los costos generados de esta investigación.

REFERENCIAS

[1] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, y I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility", *Future Gener. Comput. Syst.*, vol. 25, no. 6, pp. 599–616, 2009. doi: 10.1016/j.future.2008.12.001

[2] P. Mell y T. Grance, "The NIST definition of cloud computing", NIST [En línea]. Disponible en: https://faculty.winthrop.edu/domanm/csci411/Handouts/NIST.pdf. [Consultado: 28-ene-2025].

[3] A.B. Letaifa, A. Haji, M. Jebalia, y S. Tabbane, "State of the Art and Research Challenges of new services architecture technologies: Virtualization, SOA and Cloud Computing", *Int. J. Grid Distrib. Comput.*, vol. 3, no. 4, pp. 69-87, 2010.

[4] L.J. Zhang y Q. Zhou, "CCOA: Cloud Computing Open Architecture", 2009 IEEE Int. Conf. on Web Services. doi: 10.1109/icws.2009.144

[5] J.I. Zablah, I.X. Corrales, J.M. Aguilar, A. Garcia, F. Gomez, y M.T. Medina, "Performance limitations of a text search application running in cloud instances", *IEEE Lat. Am. Trans*. 14(3), 1499–1506. 2016. doi: 10.1109/tla.2016.7459641

[6] J. Bleicher, *Contemporary Hermeneutics*, Routledge, 2017. ISBN: 9781351622370

[7] R. Rosenberg, "Health-related quality of life between naturalism and hermeneutics", *Soc. Sci. Med.*, vol. 41, no. 10, pp. 1411-1415, 1995.

[8] A. Martinez y W. Martinez, "At the interface of computational linguistics and statistics: Computational linguistics and statistics", *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 7, no. 4, pp. 258–274, 2015.

[9] N. E. H. Ben Chaabene y M. Mallek, "Learning statistics from raw text documents", *2018 5th Int. Conf. Control Decis. Inf. Technol. (CoDIT)*, 2018. doi: 10.1109/CoDIT.2018.8394841

[10] C.S. Lee, Y.F. Kao, Y.H. Kuo, y M.H., "Wang, Automated ontology construction for unstructured text documents", *Data Knowl. Eng.*, vol. 60 no. 3, pp. 547-566, 2007.

[11] J. Devlin, M.W. Chang, K. Lee, y K. Toutanova, "*BERT: Pre-training of deep bidirectional Transformers for language understanding*", arXiv [cs.CL], 2018. doi: 10.48550/ARXIV.1810.04805

[12] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, y V. Stoyanov, "*RoBERTa: A robustly optimized BERT pretraining approach"*, arXiv [cs.CL], 2019. doi:10.48550/ARXIV.1907.11692

[13] D. Loureiro, F. Barbieri, L. Neves, L.E. Anke, y J. Camacho-Collados, "*TimeLMs: Diachronic language models from Twitter",* arXiv [cs.CL], 2022. doi: 10.48550/AERXIV.2202.03829

[14] F. Barbieri, L.E. Anke, y J. Camacho-Collados, "*XLM-T: Multilingual Language Models in Twitter for Sentiment Analysis and Beyond*", arXiV [cs.CL], 2022. doi:10.48550/arXiv.2104.12250

[15] F. Barbieri, J. Camacho-Collados, L. Neves, y L. Espinosa-Anke, "*TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification*", arXiv [cs.CL], 2020. doi:10.48550/arXiv.2010.12421

[16] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, y V. Stoyanov, "*Unsupervised Cross-lingual Representation Learning at Scale"* arXiv [cs.CL], 2020. doi: 10.48550/arXiv.1911.02116

[17] J.L. Hennessy, y D.A. Patterson, *Computer Architecture: A Quantitative Approach,* ed. 5ª, Morgan Kaufmann Series in Computer Architecture and Design, ELSEVIER, 2014. ISBN: 9780123838735

[18] IOZone, "Iozone Filesystem Benchmark", [En línea]. Disponible en https://www.iozone.org [Consultado: 28-ene-2025]

[19] C.L. Lawson, R.J. Hanson, D.R. Kincaid, y F.T. Krogh, Basic linear algebra subprograms for Fortran usage. *ACM Trans. Math. Softw. (TOMS)*, vol. 5 , no. 3, pp. 308–323, 1977. doi: 10.1145/355841.355847

[20] J.J. Dongarra, J. Du Croz, S. Hammarling, y R.J. Hanson, An extended set of FORTRAN basic linear algebra subprograms. *ACM Trans. Math. Softw.*, vol. 14, no. 1, pp. 1–17, 1988.

[21] Python, "Python Software Foundation", [En línea]. Disponible en: https://www.python.org [Consultado: 28-ene-2025]

[22] PyPi, "Python Package Installer", [En línea]. Disponible en: https://pypi.org/project/pip [Consultado: 28-ene-2025]

[23] D. Ho, "Notepad++ Programming text editor", [En Linea]. Disponible en: https://notepad-plus-plus.org/ [Consultado: 20-ene-2025]

[24] Linode, "Akamai Connected Cloud", [En Linea]. Disponible en: https://www.linode.com [Consultado: 20-ene-2025]

[25] CentOS, "The CentOS Project", [En línea]. Disponible en: https://www.centos.org [Consultado: 20-ene-2025]

[26] G. Harrison, y S. Feuerstein, *MySQL stored procedure programming*. O'Reilly Media, pp. 159–176, 2006. ISBN: 9780596100896

[27] H. Zhang, y M.O. Shafiq, "Survey of transformers and towards ensemble learning using transformers for natural language processing", *J. Big Data*, vol. 11, no. 1, 2024. doi:10.1186/s40537-023-00842-0

[28] INTEL, "Intel Core i7-740QM", [En Linea]. Disponible en: https://ark.intel.com/content/www/xl/es/ark/products/49024/intel-core-i7-740qm-processor-6m-cache-1-73-ghz.html?q=740 [Consultado: 20-ene-2025]