|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Programa Curricular** |  | Administración de Sistemas Informáticos |  | **Fecha** | 17 | 11 | 2023 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Datos del estudiante** | | | | | | | |
| Nombres: Daniel | | | | | | | Apellidos: Cortés Pérez |
| Tipo de identificación: | T.I. |  | C.C. | x | C.E. |  | Número: 3057593661 |
| Correo institucional: dancortespe@unal.edu.co | | | | | | | Teléfono: 3057593661 |

|  |
| --- |
| **Formulación Trabajo de Grado** |

|  |
| --- |
| **1. MODALIDAD/SUBMODALIDAD DE TRABAJO DE GRADO**  Modalidad: Trabajos Investigativos  Submodalidad: Trabajo Monográfico |

|  |
| --- |
| **2. TÍTULO DEL TRABAJO DE GRADO**  Análisis de Sentimientos en Ambientes con Comentarios Mediante NLP. |

|  |
| --- |
| **3. INTRODUCCIÓN Y JUSTIFICACIÓN**  En la era de la información, las organizaciones y los individuos generan una gran cantidad de datos. Estos datos pueden utilizarse para comprender las opiniones y emociones de las personas, lo que puede ser de gran utilidad para una variedad de propósitos, como la mejora de la satisfacción del cliente, la detección de tendencias de mercado y la toma de decisiones más informada.  El análisis de sentimientos es una técnica que se utiliza para identificar la emoción subyacente en el texto. Se puede aplicar a una variedad de fuentes de datos, como comentarios en línea, reseñas de productos y redes sociales. El análisis de sentimientos puede proporcionar información valiosa sobre las opiniones y emociones de las personas, lo que puede ayudar a las organizaciones a comprender mejor a sus clientes.  En ambientes con comentarios, el análisis de sentimientos puede utilizarse para una variedad de propósitos, como:   * ***Identificar tendencias de comportamiento:*** El análisis de sentimientos puede utilizarse para identificar tendencias en las opiniones y emociones de las personas. Por ejemplo, una empresa podría utilizar el análisis de sentimientos para identificar si sus clientes están satisfechos con su producto o servicio. * ***Mejorar la toma de decisiones:*** El análisis de sentimientos puede utilizarse para proporcionar información para la toma de decisiones. Por ejemplo, una empresa podría utilizar el análisis de sentimientos para identificar qué productos o servicios son más populares entre sus clientes. * ***Personalizar la experiencia del cliente:*** El análisis de sentimientos puede utilizarse para personalizar la experiencia del cliente. Por ejemplo, una empresa podría utilizar el análisis de sentimientos para mostrar a sus clientes productos o servicios que les puedan interesar.   Este trabajo de grado se enfoca en el análisis de sentimientos en ambientes con comentarios. El objetivo de este trabajo es investigar y documentar un algoritmo que automatice la recopilación y el análisis de comentarios para facilitar la comprensión de las emociones y opiniones expresadas en los comentarios en línea.  **Justificación**  El análisis de sentimientos en ambientes con comentarios es un área de investigación emergente con un gran potencial de aplicación (Liu, B). La búsqueda, documentación y posteriormente desarrollo de un algoritmo que automatice la recopilación y el análisis de comentarios podría tener una serie de beneficios, como:   * ***Mejorar la eficiencia:*** El algoritmo podría automatizar tareas que actualmente se realizan manualmente, lo que podría liberar tiempo y recursos para otras actividades. * ***Mejorar la precisión:*** El algoritmo podría utilizar técnicas de aprendizaje automático para mejorar la precisión de los resultados del análisis de sentimientos. * ***Ampliar el alcance:*** El algoritmo podría utilizarse para analizar grandes cantidades de datos, lo que podría ampliar el alcance de las investigaciones sobre análisis de sentimientos.   Este trabajo de grado tiene el potencial de contribuir a la investigación sobre análisis de sentimientos en ambientes con comentarios. El desarrollo de un algoritmo que automatice la recopilación y el análisis de comentarios podría ser una herramienta valiosa para las organizaciones y los investigadores que buscan comprender mejor las opiniones y emociones de las personas. |

|  |
| --- |
| **4. OBJETIVO GENERAL**  Implementar un algoritmo para el análisis de sentimientos automático para comentarios en línea en diversos contextos y dominios, que permita la identificación de tendencias y la extracción de insights valiosos de los datos recopilados. |

|  |
| --- |
| **5. OBJETIVOS ESPECÍFICOS**   * Revisar la literatura sobre análisis de sentimientos en ambientes con comentarios. * Identificar y seleccionar las técnicas y herramientas adecuadas para el desarrollo del algoritmo. * Diseñar y desarrollar el algoritmo de análisis de sentimientos. * Evaluar el rendimiento del algoritmo en un conjunto de datos de comentarios en línea. * Aplicar el algoritmo a un conjunto de datos de comentarios en línea en un contexto específico. |

|  |
| --- |
| **6. ANTECEDENTES**  El procesamiento del lenguaje natural (NLP) es un campo en constante evolución que se ocupa de la interacción entre las computadoras y el lenguaje humano. En los últimos años, ha habido un crecimiento significativo en el desarrollo de librerías y herramientas de NLP que brindan funcionalidades avanzadas para analizar y comprender el texto.  Ahora, el análisis de sentimientos es una técnica que se utiliza para identificar la emoción subyacente en el texto. Se puede aplicar a una variedad de fuentes de datos, como comentarios en línea, reseñas de productos y redes sociales. El análisis de sentimientos puede proporcionar información valiosa sobre las opiniones y emociones de las personas, lo que puede ayudar a las organizaciones a comprender mejor a sus clientes. (Liu, B.)  El análisis de sentimientos en ambientes con comentarios es un área de investigación emergente con un gran potencial de aplicación. El desarrollo de algoritmos de análisis de sentimientos automáticos puede ayudar a las organizaciones a:   * Identificar tendencias en las opiniones y emociones de las personas. * Mejorar la toma de decisiones. * Personalizar la experiencia del cliente.   **Trabajos previos**  Se han realizado numerosos trabajos de investigación sobre análisis de sentimientos en ambientes con comentarios. Estos trabajos se han centrado en el desarrollo de nuevos algoritmos, la evaluación de algoritmos existentes y la aplicación de análisis de sentimientos a diferentes contextos.  Algunos de los trabajos previos más relevantes son los siguientes:   * ***[Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 3(1), 1-167.]*** Este trabajo proporciona una revisión exhaustiva de la literatura sobre análisis de sentimientos. * ***[Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. Proceedings of the ACL 2002 conference on Empirical methods in natural language processing, 79-86.]*** Este trabajo presenta un algoritmo de análisis de sentimientos basado en aprendizaje automático. * ***[Mohammad, S. M., & Kiritchenko, S. (2013). Sentiment analysis of twitter data. arXiv preprint arXiv:1307.5102.]*** Este trabajo presenta un conjunto de datos de comentarios en línea de Twitter que se han etiquetado con su sentimiento.   Estos trabajos han contribuido a sentar las bases para el desarrollo de algoritmos de análisis de sentimientos automáticos. Sin embargo, todavía existen una serie de desafíos que deben abordarse, como la subjetividad del lenguaje humano, la falta de datos etiquetados y la necesidad de algoritmos que sean capaces de identificar diferentes tipos de emociones. |

|  |
| --- |
| **7. METODOLOGÍA**  El enfoque metodológico que se utilizará para el desarrollo de este trabajo de grado será el enfoque cuantitativo. Este enfoque se caracteriza por utilizar métodos y técnicas de investigación que se basan en la recolección y análisis de datos cuantitativos.  Procedimientos y técnicas  Los procedimientos y técnicas que se utilizarán para alcanzar los objetivos específicos propuestos son los siguientes:  ***Objetivo específico 1:*** Revisar la literatura sobre análisis de sentimientos en ambientes con comentarios.  ***Técnica:*** Revisión bibliográfica  ***Procedimiento:*** Se utilizarán bases de datos de artículos científicos y libros para recopilar información sobre análisis de sentimientos en ambientes con comentarios. La información recopilada se analizará para identificar las tendencias y los desafíos actuales en esta área de investigación.  ***Objetivo específico 2:*** Identificar y seleccionar las técnicas y herramientas adecuadas para el desarrollo del algoritmo.  ***Técnica:*** Revisión bibliográfica  ***Procedimiento:*** Se utilizará la información recopilada en el objetivo específico 1 para identificar las técnicas y herramientas más adecuadas para el desarrollo del algoritmo. Se considerarán factores como la precisión, la eficiencia y la capacidad de identificar diferentes tipos de emociones.  ***Objetivo específico 3:*** Diseñar y desarrollar el algoritmo de análisis de sentimientos.  ***Técnica:*** Desarrollo de software  ***Procedimiento:*** Se utilizarán técnicas de aprendizaje automático para desarrollar el algoritmo. El algoritmo se implementará en un lenguaje de programación de alto nivel.  ***Objetivo específico 4:*** Evaluar el rendimiento del algoritmo en un conjunto de datos de comentarios en línea.  ***Técnica:*** Evaluación de la precisión  ***Procedimiento:*** Se utilizará un conjunto de datos de comentarios en línea etiquetados con su sentimiento para evaluar la precisión del algoritmo.  ***Objetivo específico 5:*** Aplicar el algoritmo a un conjunto de datos de comentarios en línea en un contexto específico.  ***Técnica:*** Aplicación del algoritmo  ***Procedimiento:*** Se utilizará el algoritmo desarrollado para analizar un conjunto de datos de comentarios en línea en un contexto específico. Por ejemplo, el algoritmo podría utilizarse para analizar comentarios de clientes sobre un producto o servicio específico. |

**8. CRONOGRAMA DE ACTIVIDADES**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Semanas de ejecución de cada actividad** | | | | | | | | | | | | | | | |
| **ACTIVIDADES A REALIZAR** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** |
| Actividad 1: Revisión bibliográfica | x | x | x |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Actividad 2: Selección y preparación de herramientas de análisis de datos. |  | x | x | x |  |  |  |  |  |  |  |  |  |  |  |  |
| Actividad 3: Recopilación de Datos. |  |  |  | x | x |  |  |  |  |  |  |  |  |  |  |  |
| Actividad 4: Diseño e implementación del algoritmo. |  |  |  |  | x | x | x | x | x |  |  |  |  |  |  |  |
| Actividad 5: Evaluación y documentación del rendimiento del algoritmo. |  |  |  |  |  |  |  | x | x | x | x | x |  |  |  |  |
| Actividad 6: Preparación del informe con los hallazgos hechos y los resultados preliminares. |  |  |  |  |  |  |  |  |  |  |  | x | x | x | x | x |
| Actividad 7: Presentación de resultados a la comunidad académica. |  |  |  |  |  |  |  |  |  |  |  |  |  |  | x |  |

|  |
| --- |
| **9. RESULTADOS ESPERADOS Y APORTE ESPECÍFICO**  Los resultados esperados del trabajo de grado son los siguientes:   * Un algoritmo de análisis de sentimientos automático para comentarios en línea en diversos contextos y dominios. * Una evaluación del rendimiento del algoritmo en un conjunto de datos de comentarios en línea. * Una aplicación del algoritmo a un conjunto de datos de comentarios en línea en un contexto específico. * Una documentación del algoritmo y las pruebas realizadas. * Una presentación de los resultados del trabajo de grado.   Estos resultados se evidenciarán en forma de productos tangibles, como los siguientes:   * Un informe técnico que describe el algoritmo, su diseño, implementación y evaluación. * Un artículo que publica los resultados del trabajo de grado. * Un prototipo de software que implementa el algoritmo.   **Aporte específico**  El aporte específico del trabajo de grado es el desarrollo de un algoritmo de análisis de sentimientos automático que sea preciso, eficiente y capaz de identificar diferentes tipos de emociones. Este algoritmo puede tener un impacto positivo en la sociedad en el corto, mediano y largo plazo, en los siguientes aspectos:   * ***Corto plazo:*** El algoritmo puede utilizarse para mejorar la comprensión de las opiniones y emociones de las personas en línea. Esto puede ser útil para las empresas para mejorar la satisfacción de sus clientes, para los gobiernos para comprender las preocupaciones de sus ciudadanos y para los investigadores para comprender mejor el comportamiento humano. * ***Mediano plazo:*** El algoritmo puede utilizarse para identificar tendencias en las opiniones y emociones de las personas. Esto puede ser útil para las empresas para identificar oportunidades de mercado, para los gobiernos para anticiparse a los cambios sociales y para los investigadores para comprender mejor los fenómenos sociales. * ***Largo plazo:*** El algoritmo puede utilizarse para desarrollar nuevas aplicaciones que permitan a las personas interactuar de manera más efectiva con la tecnología. Por ejemplo, el algoritmo podría utilizarse para mejorar la comunicación entre humanos y máquinas, para desarrollar sistemas de recomendación más personalizados o para crear experiencias de usuario más envolventes. |

|  |
| --- |
| **10. BIBLIOGRAFÍA CITADA**  [Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.](https://scirp.org/reference/referencespapers.aspx?referenceid=2442500)  [Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 3(1), 1-167.](Liu,%20B.%20(2012).%20Sentiment%20analysis%20and%20opinion%20mining.%20Synthesis%20lectures%20on%20human%20language%20technologies,%203(1),%201-167.)  [Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. Eighth International AAAI Conference on Weblogs and Social Media.](https://ojs.aaai.org/index.php/ICWSM/article/view/14550)  Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. Proceedings of the ACL 2002 conference on Empirical methods in natural language processing, 79-86.  Mohammad, S. M., & Kiritchenko, S. (2013). Sentiment analysis of twitter data. arXiv preprint arXiv:1307.5102. |

|  |
| --- |
| **11. REFERENCIAS BIBLIOGRÁFICAS ADICIONALES**  [Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1-167.](https://scirp.org/reference/referencespapers.aspx?referenceid=2442504)  [Zhang, L., Wang, S., & Liu, B. (2018). Deep Learning for Sentiment Analysis: A Survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.](https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1253)  [Abdul-Mageed, M., & Ungar, L. (2017). EmoNet: Fine-grained emotion detection with gated recurrent neural networks. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2017)*.](https://aclanthology.org/P17-1067)  [Akhtar, M. S., Kumar, A., Ekbal, A., & Bhattacharyya, P. (2016). A hybrid deep learning architecture for sentiment analysis. In *Proceedings of the International Conference on Computational Linguistics (COLING 2016)*.](https://aclanthology.org/C16-1047)[Gui, L., Hu, J., He, Y., Xu, R., Lu, Q., & Du, J. (2017). A question answering approach to emotion cause extraction. In *Proceedings of the Conference on Empirical Methods on Natural Language Processing (EMNLP 2017)*.](https://aclanthology.org/D17-1167)  [Elgueta Morales, J. A. (2017). Repositorio Digital - Sistema de Bibliotecas Universidad del Bio-Bio. Obtenido de http://repobib.ubiobio.cl/jspui/handle/123456789/1772](http://repobib.ubiobio.cl/jspui/handle/123456789/1772)  [González, C. (2020). Repositorio Institucional académico de la Universidad Andres Bello. Obtenido de https://repositorio.unab.cl/xmlui/handle/ria/15339](https://repositorio.unab.cl/xmlui/handle/ria/15339)  [Maynard, D. (2003). GATE. Obtenido de GATE: https://gate.ac.uk/](https://gate.ac.uk/)  [Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, **35**, 1798–1828.](https://ieeexplore.ieee.org/document/6472238) |

|  |  |  |  |
| --- | --- | --- | --- |
| **Aval del futuro director** | | | |
| Nombre: | Néstor Darío Duque Méndez | Unidad Académica Básica: | Informática y Computación |
| Correo electrónico: | ndduqueme | Teléfono: |  |
| Firma: | ***Nota****: con la firma el profesor manifiesta que participó en la construcción de la propuesta y avala su presentación. Además, se compromete a realizar el proceso de acompañamiento durante el período académico inmediatamente siguiente.* | | |



\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Firma del estudiante