

Educational Data Mining: **Análisis de sentimientos en un** **dominio universitario durante la** **pandemia**

Educational Data Mining: Sentiment analysis in a university domain during the pandemic

Mineração de dados educacionais: análise de sentimentos em um domínio universitário durante a pandemia

—

Cristina BENAVIDES-MORALES

Ecuador

Universidad Central del Ecuador

acbenavides@uce.edu.ec

Boris HERRERA-FLORES

Ecuador

Universidad Central del Ecuador

bherrera@uce.edu.ec

Chasqui. Revista Latinoamericana de Comunicación

N.º 151, diciembre 2022 - marzo 2023 (Sección Monográfico, pp. 217-236)

ISSN 1390-1079 / e-ISSN 1390-924X

Ecuador: CIESPAL

Recibido: 05-09-2022 / Aprobado: 06-12-2022 / Publicado: 21-08-2022

Resumen

La nueva normalidad provocada por la emergencia sanitaria del COVID 19 obligó a los gobiernos a adoptar normas de aislamiento obligatorio, las medidas modificaron la interacción social en casi todos los escenarios. En el campo educativo, los establecimientos de enseñanza en general cerraron sus instalaciones y adoptaron herramientas tecnológicas para suplir el aula física, se inauguró así un período de *educación remota de emergencia* (Hodges et al., 2020) (Portillo, S., Castellanos, L., Reynoso, O., & Gavotto, O., 2020) (Castañeda y Vargas, 2021).

La presente investigación analiza la percepción estudiantil sobre la implementación de la educación remota de emergencia en la Carrera de Comunicación Social de la Universidad Central del Ecuador en el período académico 2020 – 2020, para lo cual se aplicó un sondeo de opinión a estudiantes. El instrumento cualitativo recogió la opinión, y el corpus se procesó en una librería de Phytón. En general, se evidenció una valoración positiva frente a la nueva realidad educativa, se resaltó el uso de plataformas y redes sociales, así como innovadoras prácticas educativas en: clases magistrales, experimentación, tutorías y evaluaciones; entre los aspectos negativos, se cuestionó la escasa interacción docente/estudiante, la falta de didáctica en entornos digitales, y la profundización de la brecha tecnológica.

Palabras clave: análisis automático de textos, educación remota de emergencia, COVID 19

Asbtract

The new normality caused by the COVID 19 health emergency forced governments to adopt mandatory isolation rules, the measures modified social interaction in almost all scenarios. In the educational field, educational establishments in general closed their facilities and adopted technological tools to replace the physical classroom, thus inaugurating a period of emergency remote education (Hodges et al., 2020) (Portillo, S., Castellanos, L., Reynoso, O., & Gavotto, O., 2020) (Castañeda and Vargas, 2021).

The present investigation analyzes the student perception about the implementation of emergency remote education in the Social Communication Career of the Central University of Ecuador in the academic period 2020 - 2020, for which an opinion poll was applied to students. The qualitative instrument collected the opinion, and the corpus was processed in a Python bookstore. In general, a positive assessment of the new educational reality was evidenced, the use of platforms and social networks was highlighted, as well as innovative educational practices in: master classes, experimentation, tutorials and evaluations; Among the negative aspects, the low teacher/student interaction, the lack of didactics in digital environments, and the deepening of the technological gap were questioned.

Keywords: automatic text analysis, emergency remote education, COVID 19

Resumo

A nova normalidade causada pela emergência de saúde COVID 19 obrigou os governos a adotar regras de isolamento obrigatório, as medidas modificaram a interação social em quase todos os cenários. No campo educacional, os estabelecimentos de ensino em geral fecharam suas instalações e adotaram ferramentas tecnológicas para substituir a sala de aula física, inaugurando assim um período de educação remota emergencial (Hodges et al., 2020) (Portillo, S., Castellanos, L., Reynoso, O., & Gavotto, O., 2020) (Castañeda e Vargas, 2021).

A presente investigação analisa a percepção do aluno sobre a implementação da educação remota de emergência na Carreira de Comunicação Social da Universidade Central do Equador no período acadêmico 2020 - 2020, para o qual foi aplicada uma pesquisa de opinião aos alunos. O instrumento qualitativo coletou a opinião, e o corpus foi processado em uma livreria Python. De uma forma geral, evidenciou-se uma avaliação positiva da nova realidade educativa, destacando-se a utilização de plataformas e redes sociais, bem como práticas educativas inovadoras em: master classes, experimentação, tutoriais e avaliações; Entre os aspectos negativos, foram questionados a baixa interação professor/aluno, a falta de didática em ambientes digitais e o aprofundamento da lacuna tecnológica.

Palavras-chave: análise automática de texto, educação remota de emergência, COVID 19

1. Introducción

La pandemia mundial abrió un escenario insospechado en el uso de plataformas digitales para la interacción social. Los efectos en la salud y la vida de gran parte de la población mundial fueron devastadores, el colapso de los servicios de salud, la falta de empleo o la escasez de comida, alteraron las dinámicas sociales, políticas y económicas de la población mundial. Las medidas de confinamiento, toques de queda, y restricción de movilidad, limitaron sobre todo la comunicación humana.

El mensaje: *Quédate en casa*, revelaba la necesidad de mantener a la población en sus hogares, y, de hecho, esta medida resultó ser efectiva para evitar el contagio y la propagación del COVID 19, aunque implicó pérdidas económicas en el corto plazo, con la consecuente caída de la producción, la demanda interna y el aumento del desempleo. La sociabilidad se redujo y la interacción se limitó al uso de canales virtuales y plataformas digitales.

En educación, de acuerdo a un informe presentado por la UNESCO (2020), 166 países cerraron sus escuelas y universidades, a escala mundial el 87% de la población estudiantil se vio afectada por estas medidas; es decir, unos 1.520 millones de estudiantes. En todo el mundo, alrededor de 63 millones de maestros dejaron de laborar en las aulas, y es posible que más de 3 millones de niños, niñas y adolescentes podrían no regresar a la escuela (UNICEF, 2020).

En las instituciones de educación superior el panorama fue más esperanzador, las universidades tuvieron y tienen un escenario favorable en varios sentidos (Portillo, Castellanos, Reynoso, & Gavotto, 2020) (Castañeda, K y Vargas, A, 2021) (García, 2021), por ejemplo, existían centros de educación superior que ya contaban con plataformas digitales para el ejercicio de la docencia presencial, y con docentes más preparados en este tipo de estrategias metodológicas y tecnológicas, desde la década del setenta se contaba con universidades a distancia, semipresenciales o en línea, cuya experiencia fue vital en ese momento. De hecho, se aprovecharon las inmensas posibilidades de las plataformas digitales y las redes sociales para desarrollar la cátedra online de forma creativa; asimismo, la edad de las y los estudiantes hizo viable una enseñanza/aprendizaje remota.

Para García (2021), uno de los principales problemas fue la adaptación de los docentes menos capacitados al nuevo modelo, quienes optaron por la transmisión en directo de una sesión idéntica al tiempo presencial anterior, o la grabación de esa sesión para su posterior recepción por parte de los estudiantes y que, en ambos casos, hacen las veces en remoto de las secuencias presenciales en el aula.

Al momento, existen investigaciones que indagan sobre el rol del docente y las buenas prácticas en el aula (Castañeda y Vargas, 2021), otras intentan construir modelos y escenarios de educación remota en emergencia (Portillo, S., Castellanos, L., Reynoso, O., & Gavotto, O., 2020), o incluso se analizan las condiciones digitales de base en los sistemas educativos (Lugo, M. & Loíacono, L., 2020).

En esta línea, el objetivo general de la investigación pretender estudiar, desde la perspectiva de los estudiantes, la experiencia de la virtualidad y la aplicación de la normativa en el aprendizaje remoto de emergencia, así también, de manera específica, prevé evidenciar la transformación de las prácticas educativas mediadas por la tecnología, y el sentimiento de los estudiantes sobre el nuevo proceso emergente de enseñanza.

El trabajo intentará responder las siguientes cuestiones:

- ¿Qué condiciones han influido, y en qué medida, en el proceso de transición a la educación remota de emergencia?
- ¿El uso de las herramientas virtuales institucionales contribuyeron con el proceso de aprendizaje?
- ¿Cómo cambió la interacción docente/estudiante en el nuevo escenario virtual?

Como parte de la metodología se hizo uso de las técnicas computacionales de procesamiento del lenguaje natural para el tratamiento del texto como un dato, y por medio de la minería de opinión determinar el sentimiento expresado en dichos textos en el dominio estudiantil universitario. Para ello, se aplicó un

sondeo de opinión a 832 estudiantes de la Facultad de Comunicación Social - Universidad Central del Ecuador del período Julio – Octubre 2020.

La técnica arrojó como resultado principal una valoración general satisfactoria en las 11648 opiniones válidas extraídas del sondeo de opinión, se clasificaron 5738 comentarios positivos sobre la incorporación de la educación remota de emergencia en el universo de estudio, lo cual abre un escenario alentador por la capacidad resiliente de la comunidad universitaria en situaciones de crisis. Por otro lado, la investigación contribuye con las investigaciones sobre el futuro de la educación y el uso de herramientas digitales, no sólo como apoyo si no como parte de la construcción de un nuevo modelo de “aprendizaje híbrido” (Lugo, M. & Loíacono, L., 2020).

2. Marco teórico

2.1. Educación remota en emergencia

Con la pandemia mundial, las instituciones de educación superior se enfrentaron al reto de cambiar el proceso de enseñanza/aprendizaje, se trasladó el aula tradicional de clases a un espacio virtual mediado por herramientas tecnológicas. Se diseñó sobre la marcha un modelo pedagógico que cambió la forma de enseñar, aprender, evaluar y tomar acciones acertadas en el aula (González, 2019). Y se comprendió el desafío impuesto por la condición sanitaria, buscando soluciones que ayuden a los estudiantes universitarios en esta nueva realidad (Lizárraga Juárez, A., López Ramírez, E., & Martínez Iñiguez, J. E. 2020).

La primera tarea fue definir el nuevo modelo de educación, para Castañeda y Vargas (2021) no se puede hablar estrictamente de una educación a distancia o educación virtual, pues el centro del proceso de enseñanza está en el uso de plataformas tecnológicas: videoconferencias, gestores educativos como el moodle o redes sociales, aunque los contenidos y la forma de impartirlos siguen siendo los mismos.

El presente trabajo parte de la definición de *Educación remota en emergencia ERT* (Hodges et al., 2020) (Portillo, S., Castellanos, L., Reynoso, O., & Gavotto, O., 2020) (Castañeda y Vargas, 2021), como un tipo de instrucción alternativa y temporal que proporciona acceso a materiales de enseñanza, incorporando actividades sincrónicas y asincrónicas.

En este nuevo modelo existe un peso significativo de las actividades sincrónicas, que emulan la condición anterior pudiendo convertirse en una nueva forma colectiva de aprender pero que también supone varios peligros (Marotias, 2021). Por otra parte, Roatta, S. and Tedini, D. (2021) plantean que “el aislamiento social obligatorio ha favorecido el aprendizaje colaborativo y consolidado el papel que en este paradigma llevaban a cabo docentes y alumnos”, superando a su vez el modelo pasivo o incluso el modelo activo de aprendizaje.

La implantación de la *educación remota en emergencia* requiere de una serie de recursos que garanticen la correcta vinculación de estudiantes, la organización de procesos administrativos y académicos de calidad, se necesita de forma particular una normativa específica y clara, la publicación de instructivos y manuales ha sido abundante y de diversa naturaleza en las universidades, las normas buscaron mejorar y unificar el proceso de enseñanza/aprendizaje a través de fijar procesos, unificar mecanismos de evaluación y proponer dispositivos de prevención sanitaria (García-de-Paz, S. & Santana Bonilla, P., 2021).

2.2. Procesamiento del Lenguaje Natural - PLN

El PLN es un campo de las Ciencias de la Computación, Inteligencia Artificial (IA) y Lingüística que estudia las interacciones entre las computadoras y el lenguaje humano. El PLN no trata la comunicación por medio del lenguaje natural de una forma abstracta, por el contrario, diseña mecanismos para comunicarse que sean eficaces computacionalmente y que se puedan realizar por medio de programas en donde se ejecuta o simula una comunicación (Vásquez et al., 2009).

Entre las múltiples tareas de las que se ocupa el PLN se encuentra la clasificación de textos, que consiste en la asignación de un conjunto de categorías a una colección de documentos, resolviéndose de esta forma la clasificación objetiva de textos. El procesamiento de textos puede realizar análisis de sentimientos al clasificar los textos en función de la polaridad de la opinión que expresa su autor, combinando PLN y minería de textos (Martínez Cámara et al., 2011)

Como parte del procesamiento de textos, el Análisis de Sentimientos o también llamado Minería de Opiniones se refiere a una serie de aplicaciones de técnicas del PLN y lingüística computacional, que tienen como objetivo la extracción de información subjetiva a partir de contenidos generados por los usuarios de blogs, formularios electrónicos de opinión, redes sociales, o emails direccionados. Con este tipo de tecnologías se tiene la capacidad de extraer un valor tangible y directo, que pueda ser positivo o negativo, a partir de un comentario textual.

Se aplica la técnica de aprendizaje supervisado para determinar la polaridad, aun cuando hasta el momento obtiene mejores resultados que las no supervisadas, las cuáles presentan varias desventajas, entre ellas las sujetas al sobre entrenamiento y que son altamente dependientes de la calidad, tamaño y dominio de los datos de entrenamiento. Por otra parte, las aproximaciones no supervisadas no requieren de la existencia de colecciones textuales previamente clasificadas, son menos dependientes del lenguaje, pero se basan en recursos externos, y actualmente existen pocos recursos disponibles para todos los

idiomas y culturas, siendo mayormente dependientes del idioma, como lo describe Amores Fernández (2016).

El procesamiento automático de opiniones no es una tarea sencilla, algunos de los problemas presentes en el tratamiento de las opiniones son: el uso de lenguaje informal, las abreviaturas, los errores ortográficos y tipográficos (Mosquera et al., 2012), el lenguaje irónico y sarcástico (Ghanem et al., 2020), el nivel de conocimiento del lenguaje (Díaz-Galiano et al., 2019), el nivel cultural, tratamiento de la negación (Zafra et al., 2017), entre otros.

2.3. Minería de datos aplicada a la educación superior

Los procesos periódicos de evaluación y acreditación en las universidades exigen el cumplimiento de estándares mínimos, como la eficiencia académica medido por el índice de retención estudiantil, que consiste en la cantidad de estudiantes matriculados por primera vez al primer semestre y su permanencia en la institución después de dos años.

La eficiencia académica es un indicador cuya dimensión cuantitativa refleja aspectos cualitativos que van desde la calidad del sistema educativo, las políticas de ingreso a la institución, hasta una amplia gama de factores que intervienen en el proceso educativo que garanticen la permanencia de los estudiantes hasta su titulación.

En el dominio educativo, la deserción escolar es un tema de estudio para los investigadores, en la línea de predecir un posible fracaso o un sistema de alerta temprana que permita a los profesionales de la educación apoyar o intervenir en cualquier situación que amerite. La minería de datos en dominios educativos ofrece numerosas ventajas al utilizar los algoritmos de clasificación que pueden usarse para obtener una buena predicción del rendimiento académico de los estudiantes.

3. Metodología

El objetivo de este apartado es definir una metodología de trabajo para extraer opinión del estudiante en un ambiente controlado, corriendo una encuesta de texto abierto con preguntas enfocadas a comprender las experiencias de aprendizaje en el dominio educativo.

El resultado esperado es la obtención de una alta carga textual (corpus) bajo un modelo de valoración subjetiva que permita, de esta manera, su uso para el entrenamiento de sistemas de análisis de sentimientos. Entre las opiniones estudiantiles se espera que se revelen aspectos de experiencias diarias que no son fácilmente recogidas en el ambiente formal de un aula de clase, que usualmente no son documentadas, y que revelen la opinión sobre la educación remota de emergencia.

3.1. Método: Educational Data Mining (EDM)

La minería de datos aplicada a la educación o Educational Data Mining (EDM) es una disciplina que desarrolla y aplica métodos de minería de datos sobre información que proviene del ambiente educativo, y los usa para entender mejor a los estudiantes y los entornos en los que aprenden (Romero & Ventura, 2010). EDM es un campo de investigación que se centra en la aplicación de minería de datos, aprendizaje automático y estadística con métodos para detectar patrones en grandes colecciones de datos educativos.

Las dificultades de los estudiantes universitarios son identificadas mediante técnicas computacionales de procesamiento del lenguaje natural (PLN), para luego clasificar y analizar los datos usando técnicas específicas de EDM relacionadas a la minería de datos que agrupe datos educativos y análisis de sentimientos (Pérez-Suasnavas, A. et al., 2022)

En las diferentes experiencias de minería de datos en educación EDM, la identificación de las experiencias estudiantiles está basada no solamente en aspectos académicos, sino también a personales e institucionales. Dicha experiencia, conduce al desarrollo de una metodología cualitativa, para identificar y predecir las dificultades de los estudiantes universitarios (Ahadi et al., 2022), utilizando técnicas relacionadas con la lingüística computacional, la estadística y otras disciplinas informáticas (Hernández-Blanco et al., 2019); integrando diferentes técnicas computacionales y estrategias didácticas, con el propósito de apoyar la labor académica de los docentes.

Altrabsheh et al. (2013) plantea que el análisis de sentimientos aún no se ha explotado a profundidad en el sector educativo, al combinar estas áreas, la educación puede mejorar sus estándares de calidad considerando la opinión de los educandos en tiempo real.

Las emociones de los estudiantes son un campo específico de estudio para los investigadores educativos, se indaga entre otros temas, la relación entre el rendimiento académico y el encuentro dialógico entre el educando y el educador, según Altrabsheh et al. (2013), cuando se examina emociones de estudiantes que aprenden por plataformas electrónicas conocidos como e-learners, la investigación se vuelve más compleja pues, la educación remota en emergencia rara vez tiene comunicación cara a cara, a menos que el profesor y el estudiante tengan una cámara web, por tanto, la interacción es escasa. Es necesario comprender los sentimientos de los estudiantes, para determinar cuál es su opinión en relación a diversos temas, pero sobre todo los que refieren a la interacción, los datos contribuirán al cambio significativo en la calidad de la enseñanza.

En el dominio educativo el análisis de sentimientos, denominado SAE (Sentiment Analysis for Education), de los aportes de la comunidad de investigadores se ha trabajado en los comentarios de los estudiantes los cuales son pre procesados para asegurar la calidad de los datos y luego analizados a

través de Naive Bayes (Xu, 2018) y Support Vector Machine (Ledesma et al., 2018) individualmente o combinados. Esto analiza si la publicación o comentario en su conjunto es positiva o negativa (Altrabsheh et al., 2013).

Según Feldman (2013), el análisis de sentimientos abarca las áreas de investigación más importantes en la informática, existen más de 7000 artículos escritos sobre el tema y cientos de nuevas empresas que están desarrollando soluciones de análisis de sentimientos y los principales paquetes estadísticos, como SAS (Statistical Analysis System) y SPSS (Statistical Package for the Social Sciences), que incluyen módulos de análisis de sentimientos.

3.2. Técnica: Procesamiento de lenguaje natural y minería de opiniones

Esta técnica permite determinar si una opinión es positiva o negativa, más allá de una polaridad básica también se puede obtener un valor numérico dentro de un rango determinado, que de alguna forma obtenga una valoración asociada a una determinada opinión.

Dentro del campo de la inteligencia artificial, el proceso de minería de texto está conformado por dos etapas (Montes y Gómez, 2001):

- a) **Procesamiento:** en esta primera etapa, los textos se transforman a una representación estructurada con el propósito de facilitar el análisis de sentimiento sobre la opinión del estudiante. La técnica de procesamiento del lenguaje natural, como tarea en la inteligencia artificial permite la normalización de los textos, y arroja una estructura adecuada para facilitar el análisis.

Entre las principales técnicas de normalización aplicadas en esta investigación esta: la tokenización que consiste en el proceso de dividir en palabras o tokens de un documento de texto; eliminar stopwords (pronombres, conjunciones, preposiciones y las distintas formas del verbo ser, haber, etc.) necesarias para realizar oraciones pero carecen de información; lematización (transformación morfológica de las palabras a lemas que otorguen más precisión); etiquetado gramatical que identifique al sustantivo, verbo, adjetivo, etc.; y finalmente eliminar ambigüedades u opiniones sin valor que definitivamente se descartan.

- b) **Descubrimiento:** en esta segunda etapa las representaciones obtenidas se analizan con algoritmos de entrenamiento, con el propósito de descubrir algunos patrones que sean de interés y detecten polaridad a la opinión. Esta detección de la polaridad se define a la capacidad computacional de determinar una opinión positiva o negativa, como un tratamiento automático a las opiniones del lenguaje informal.

3.3. Instrumento: Lenguaje de programación

Hoy en día existen algunas herramientas que incorporan tecnologías de lenguaje humano para proporcionar analítica de datos confiables y de aceptable rendimiento. Entre ellas se presentan algunas herramientas que se destacan y utilizan diversos algoritmos de aprendizaje automático, como RapidMiner, SAS, y código Python como herramientas que asisten el aprendizaje automático.

Existe cierta dificultad para resolver algoritmos, pues la subjetividad de las opiniones no siempre permite entender de forma clara un sentimiento positivo o negativo, por ello la necesidad de contar con un corpus en español bien nutrido. La fuerza de la opinión también aporta en el cálculo de la polaridad y su tratamiento metodológico contribuye a determinar si una opinión es válida en el procesamiento del texto.

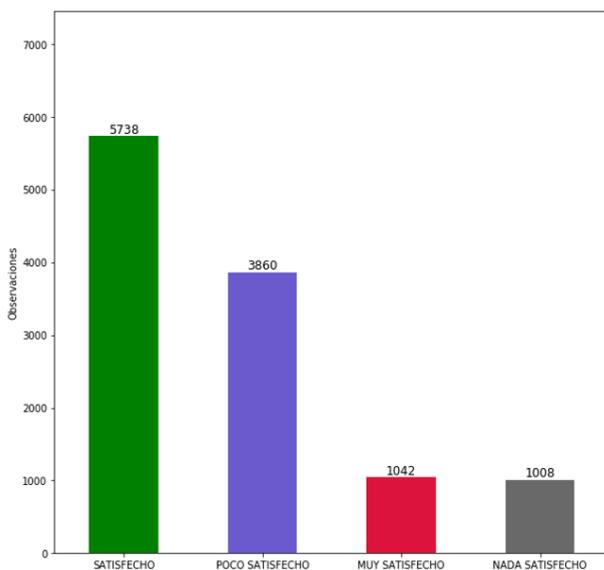
Esta investigación utilizó métodos estadísticos de aprendizaje automático codificados en Python con librerías adicionales de soporte al procesamiento del lenguaje natural y la detección de polaridad de las opiniones, sobre un corpus compilado en español que consume el algoritmo y entrega un resultado.

3.4 Universo y muestra

El universo de la investigación fueron 1521 estudiantes matriculados en la Carrera de Comunicación Social de la Universidad Central del Ecuador en el período julio – septiembre 2020, de los cuáles 832 respondieron a un sondeo de opinión. La muestra fue no probabilística de tipo incidental, voluntaria y anónima, y corresponde al 54,7% del universo. Los encuestados pertenecen predominantemente a un contexto urbano, 39% fueron hombres y 61% mujeres, con un promedio de edad de 22.4 años de edad.

El cuestionario tuvo 14 preguntas abiertas y cada una acompañada por 4 ítems de valoración: Muy satisfecho, satisfecho, poco satisfecho, nada satisfecho. En el cuestionario se trataron aspectos como: sílabo y materia; satisfacción y cumplimiento; respeto, comunicación y material apoyo; y finalmente, normativa, infraestructura y diálogo.

La encuesta en formulario electrónico de Google Forms se envió a través de correo electrónico a los estudiantes, el instrumento estuvo abierto por un período de 7 días, y se obtuvieron 11648 opiniones válidas (datos cualitativos) y se extrajo una valoración general de la opinión (datos cuantitativos). En el Gráfico N°1 se demuestra una satisfacción intermedia en relación a la educación remota de emergencia implementada en el período académico 2020 - 2020:

Gráfico N° 1. Valoración cuantitativa general de la opinión

Fuente: elaboración propia

4. Resultados

4.1. Resultados cualitativos

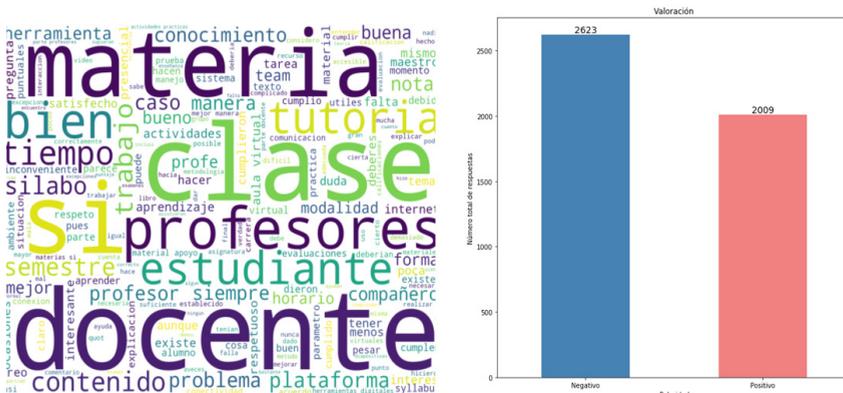
Después de aplicarse la metodología de procesamiento de lenguaje natural se identificaron 4632 opiniones válidas de las 11684, la clasificación se la hizo de forma manual y automática para convertir al texto en un dato sujeto de análisis.

En primer lugar, se extrajeron las palabras utilizadas más frecuentemente a través del programa *nubes de palabras*, en ese proceso se evidencia el uso de ciertos términos repetidos en una opinión, y gráficamente determinar si una tendencia predominante se refleja en la gráfica con el tamaño más grande de las palabras; en el Gráfico N°2 se demuestra el uso global de palabras de la encuesta, en aspectos de interés como la tutoría, el contenido y la temporalidad de uso en plataforma virtual.

En el procesamiento del texto global se identificó una tendencia del estudiante a expresar su opinión cuando colocaba una calificación cuantitativa negativa, como una reacción natural para expresar sentimientos de queja, inconformidad y demanda de atención, cuando la valoración era positiva no comentaban. En el Gráfico N°2 se evidencia esta tendencia de polaridad negativa en los comentarios, pero que si la comparamos con la valoración general de la

opinión media (Gráfico N° 1), podría parecer una contradicción que se responde con la revisión manual de los comentarios.

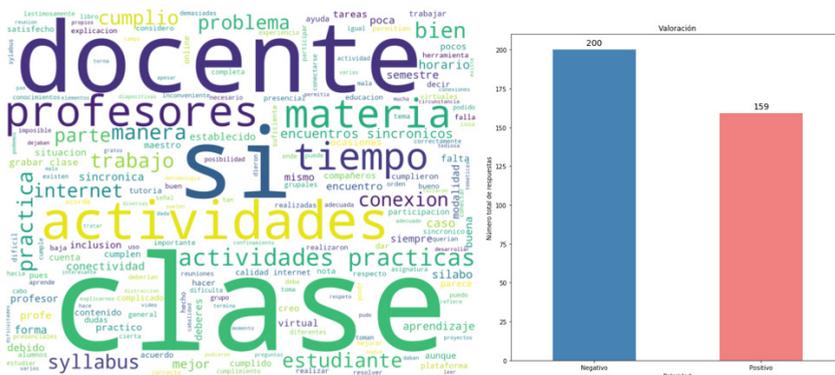
Gráfico N° 2. Nube de palabras y análisis de sentimientos de la opinión en general



Fuente: elaboración propia

Se destaca en el Gráfico N°3, los encuentros sincrónicos con un evidente cumplimiento de la actividad, pero con problemas de conexión y uso de la plataforma virtual respecto al soporte de contenidos. Incluye un sentimiento de inconformidad, expresando opinión que aporta una carga superior a la positiva que tampoco es despreciable.

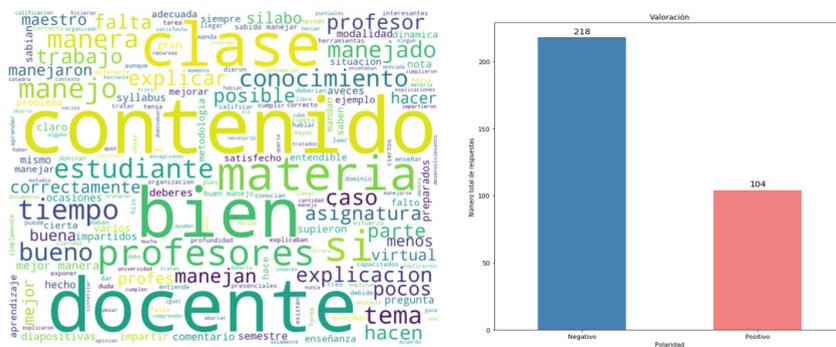
Gráfico N° 3. Nube de palabras y análisis de sentimientos en pregunta, sobre encuentros sincrónicos



En el Gráfico N°4 sobre el manejo de contenidos por parte del Docente, en cantidad de texto destaca un manejo adecuado de transmisión de conocimiento

por la virtualidad, pero desde el sentimiento expresa el estudiante una adicional crítica y sugerencia de mejora.

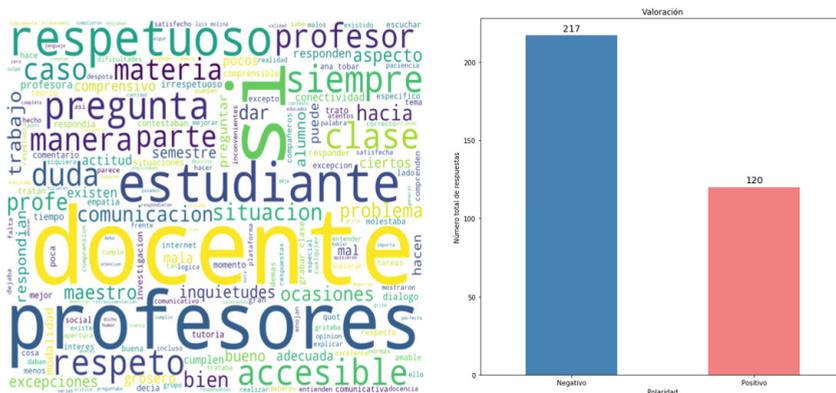
Gráfico N° 4. Nube de palabras y análisis de sentimientos en pregunta, sobre manejo del contenido



Fuente: elaboración propia

Respecto a la relación entre el estudiante con el docente en aspectos de respeto, comunicación y accesibilidad en el Gráfico N°5, hay una clara tendencia a una respuesta adecuada y suficiente a inquietudes de contenido en clase, así como también una queja y malestar a situaciones de trato y comunicación que evidencia en el texto abierto con carga emocional.

Gráfico N° 5. Nube de palabras y análisis de sentimientos en pregunta, sobre relación estudiante - docente



Fuente: elaboración propia

Así también, es importante analizar la riqueza léxica hallada en el dato de texto investigado. Y no sólo por tratarse de estudiantes de una carrera de

comunicación social, sino porque al encontrarse mejor porcentaje de utilización de términos diferentes, contribuye con la detección de la polaridad.

Tabla N°1. Riqueza léxica por semestres

Semestre	Palabras Totales (Tokens)	Palabras Diferentes (Types)	Riqueza Léxica	Porcentaje
Primero	250	159	0,64	64%
Segundo	250	154	0,62	62%
Tercero	250	169	0,68	68%
Cuarto	250	157	0,63	63%
Quinto	250	166	0,66	66%
Sexto	250	147	0,59	59%
Séptimo	250	164	0,66	66%
Octavo	250	154	0,62	62%
Noveno	250	153	0,61	61%

Fuente: elaboración propia

Finalmente, hacia el análisis de los resultados, la encuesta cualitativa demuestra que con técnicas computacionales es posible capturar sentimientos, que la encuesta tradicional no lo hace con datos cuantitativos, la información que arrojan las técnicas cualitativas de extracción de información favorece a la toma de decisiones en la gestión de la educación.

5. Discusión

En la actualidad las instituciones educativas, gracias a la informatización de sus procesos educativos y gestión digital, cuentan con enormes cantidades de información sobre diferentes áreas del proceso educativo, el EDM desarrolla métodos para explorar estos datos, los transforma para entender y atender mejor a los estudiantes, y como consecuencia, las instituciones educativas podrían realizar ajustes pertinentes en determinados aspectos del mejoramiento y aseguramiento de la calidad.

Las investigaciones que involucren análisis de sentimientos en entornos educativos son escasas, se encontraron estudios referentes a experiencias de aprendizajes que utilizan opiniones de estudiantes en plataformas digitales, en el campo del análisis cualitativo está el trabajo de Ingoles et al. (2018) sobre el “Flujo de trabajo para comprender los problemas de aprendizaje de los estudiantes utilizando sus tweets”, otro estudio que utiliza una técnica de

análisis de datos (contenido) se encuentra en el paper “Sistema clasificador de tweets estudiantiles, para comprender sus problemas y aspectos educativos positivos” (Greeshma y Veigas, 2020), y un análisis morfológico y semántico de textos lo hallamos en el documento “Sistema de clasificación de etiquetas sobre los problemas en su aprendizaje” (Abdelhamid et al., 2020).

En el presente estudio se identificó una tendencia del estudiante a expresar su opinión cuando colocaba una calificación cuantitativa negativa, como una reacción natural para expresar sentimientos de queja, inconformidad y demanda de atención, cuando la valoración era positiva no comentaban. Por ello se concluye que la polaridad cualitativa del estudio es negativa, pero que la valoración general de la opinión cuantitativa es satisfactoria (escala 2 de Likert).

Entre las principales ideas se evidenció cierta incertidumbre en relación a la conectividad, problemas con la plataforma, uso de laboratorios, tutorías e interacción docente – estudiante. Ven como positivo el cumplimiento de los contenidos propuestos en el syllabus, inclusión de actividades prácticas, la asistencia regular de los docentes, el manejo de los contenidos audiovisuales que les permite repetir y repasar.

El nuevo proceso educativo es *remoto* por la utilización de plataformas digitales, hecho que contribuye con el aprendizaje, sin embargo, los problemas de conexión y congestión en la plataforma de la Universidad impidieron una realización plena en la transmisión de los contenidos. Así también, la falta o escasa conectividad de los estudiantes impidió su conexión en actividades sincrónicas, por tanto, el problema no radica en el manejo de las plataformas sino más bien en la brecha digital que existe en el Ecuador.

Dicho proceso educativo es emergente porque la clase es una copia idéntica de la sesión presencial, transmitida en vivo por medio de video conferencia en Teams, la disertación del docente demuestra su buen manejo de los contenidos, pero al mismo tiempo queda expuesta la carencia de una pedagogía educativa mediada por aparatos tecnológicos. Por ejemplo, aunque la conexión no podía exceder los 40 minutos consecutivos, las actividades asincrónicas no consideraron el tiempo real para su ejecución y muchas veces demandaban más tiempo para su cumplimiento, este hecho responde a la poca experiencia de los docentes en educación a distancia o incluso semi presencial.

El uso de la plataforma Moodle (<https://uvirtual.uce.edu.ec/>) y la cuenta de correo institucional anclada al programa Microsoft Teams entre otras herramientas virtuales institucionales, contribuyeron con el proceso de aprendizaje, pues operaron como dispositivos organizadores del aprendizaje; si antes el modelo pedagógico estaba basado en competencias, ahora el modelo está caracterizado por técnicas de transmisión de la información, lo cual, saca a la luz una nueva brecha entre los docentes nativos digitales y los migrantes digitales. Los docentes nativos digitales proponen a sus estudiantes mejores y más alternativas para construir aprendizaje colaborativo y no se limitan a la clase magistral por video conferencia.

La discusión sobre el conectivismo es una de esas “teorías emergentes” que trata de explicar el aprendizaje en la era digital. El estudiante no solo consume conocimiento, sino que produce conocimiento. Esta situación tiene un significado especial en el caso de la enseñanza a distancia o semi presencial (Solis, P. et al., 2020).

Por otra parte, la interacción docente/estudiante en el nuevo escenario virtual se modificó radicalmente. Además de un problema de presencialidad, la diferencia fundamental ocurre por el nivel de interacción con el docente, la misma se restringió a las actividades sincrónicas y tutorías. Los encuestados afirman que, aunque existió en la mayoría de los casos respeto entre educandos y educadores, las posibilidades de abrir un debate de ideas eran casi nulas. Ahí pues, si el deber de la universidad no sólo se restringe a la formación de profesionales, el debate de ideas y la confrontación de posiciones quedan excluidas de este nuevo modelo. Si caminamos hacia un modelo híbrido, este no debería prescindir del diálogo multidireccional.

6. Conclusiones

La aplicación de técnicas de análisis de sentimientos en los estudios de carácter cualitativo educacionales, es aún un dominio poco explorado, y por tanto fértil para el desarrollo de nuevas investigaciones. La minería de datos educacional EDM y en particular la aplicación de técnicas de análisis de sentimientos ofrece numerosas ventajas a las instituciones educativas permitiendo en modo automático gestionar contenido social y obtener criterios subjetivos sobre calidad de los procesos académicos que coexisten con el estudiantado. De esta forma sería posible poder anticiparnos a fenómenos tan delicados como la deserción educativa o el descenso paulatino de la calidad docente. Investigaciones en esta línea revelarían anticipadamente situaciones adversas o condiciones favorables en el sistema educativo, y por lo tanto construir una herramienta de alerta temprana que permita a los profesionales de la educación garantizar el índice de eficiencia terminal.

Existe una carencia de estudios relacionados con un trabajo multiinstitucional en EDM y el análisis de aprendizaje. El análisis de sentimientos aún no se explota a profundidad en el sector educativo como una herramienta que apoye a mejorar los estándares de calidad, se pueden analizar los comentarios de estudiantes en tiempo real, en consecuencia, el profesor puede cambiar su estilo de enseñanza de acuerdo con los resultados.

Las emociones de los estudiantes son un campo amplio de estudio para los investigadores. Se puede indagar sobre la importancia de entender si dichos estudiantes tienen dificultades en el curso, y qué les disgusta del mismo, determinando un vínculo entre los emociones y los niveles de aprendizaje; también se podría investigar las emociones de los estudiantes, como un escenario todavía menos explorado, sobre todo después del surgimiento de las nuevas

plataformas electrónicas de aprendizaje en línea y la metodología e-learning, en este caso se trata de entender y analizar las distancias de emociones entre el estudiante y el profesor, jugando un papel importante la expresión facial en la comprensión de la emotividad.

Se recomienda el uso de técnicas computacionales para la detección de polaridad en la opinión, pues se logra capturar el sentir del estudiante al expresar aún más su queja, inconformidad o sugerencia que la encuesta tradicional de preguntas cerradas no logra identificar.

La técnica de minería de opinión o análisis de sentimiento es recomendable, al aplicar en grupos de educación superior que por su formación académica son críticos, además de su natural uso de la tecnología encuentran un espacio privado para expresar en la intimidad de su dispositivo electrónico y la confidencialidad de hacerlo cuando y donde quiera.

7. Referencias

- Abdelhamid, S., Aly, M., y Katz, A. (2020). Harvesting tweets for a better understanding of Engineering Students' First-Year Experiences. 2020 First-Year Engineering Experience.
- Ahadi, A., Singh, A., Bower, M., y Garrett, M. (2022). Text Mining in Education—A Bibliometrics-Based Systematic Review. *Education Sciences*, 12(3), 210.
- Altrabsheh, N., Gaber, M. M. and Cocea, M. (2013). Sa-e: sentiment analysis for education. In *International conference on intel ligent decision technologies*, volume 255, pages 353–362.
- Amores Fernández, M. A. (2016). Detección de la polaridad de las opiniones basada en nuevos recursos léxicos. PhD thesis, Universidad Central “Marta Abreu” de Las Villas.
- Arriaga, P. K. G., Avilés, V. A. F., Avilés, K. P. R., Arcentales, A. M. A., and Gómez, G. C. (2018). Aplicación de técnicas de minería de datos para el análisis de la eficiencia académica. *Revista Científica Hallazgos*21.
- Castañeda, K & Vargas, A (2021). En tiempos de pandemia: una mirada retrospectiva sobre la educación a distancia, virtual y remota de emergencia, así como sobre las buenas prácticas docentes. *Revista Academia y Virtualidad*, ISSN 2011-0731, Vol. 14, N°. 1, págs. 13-22.
- Díaz-Galiano, M. C., Vega, M. G., Casasola, E., Chiruzzo, L., Cumbereras, M. A. G., Cámara, E. M., Moctezuma, D., Montejo-Ráez, A., Cabezudo, M. A. S., Tellez, E. S., et al. (2019). Overview of tass 2019: One more further for the global spanish sentiment analysis corpus. In *IberLEF@ SEPLN*, pages 550–560
- Dubiau, L. and Ale, J. M. (2013). Análisis de sentimientos sobre un corpus en español: Experimentación con un caso de estudio. In *XIV Argentine Symposium on Artificial Intel ligence (ASAI)-JAIIO 42* (2013).
- Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4):82–89.
- García, L. (2021) COVID-19 y educación a distancia digital: preconfi namiento, confi namiento y posconfi namiento. *Revista Iberoamericana de Educación a Distancia*, vol. 24, núm. 1, Enero-Junio DOI: <https://doi.org/10.5944/ried.24.1.28080>
- García-de-Paz, S., & Santana Bonilla, P. J. (2021). La Transición a entornos de educación virtual en un contexto de emergencia sanitaria: Estudio de caso de un equipo docente

- en Formación Profesional Básica. *Revista de Educación a Distancia (RED)*, 21(65). <https://doi.org/10.6018/red.450791>
- Ghanem, B., Karoui, J., Benamara, F., Rosso, P., and Moriceau, V. (2020). Irony detection in a multilingual context. In Jose, J. M., Yilmaz, E.
- González, K. (2019). E-research y tecnologías colaborativas en la educación superior (Firts edit; Redipe- Red Iberoamericana de pedagogía, Ed.). USA.
- Greeshma, A., y Veigas, A. M. (2020). Analyzing Student's Learning Experiences by using Social Media data Through Data Mining. *International Journal of Scientific Research in Computer Science Applications and Management Studies*, 9(1)
- Hernández-Blanco, A., Herrera-Flores, B., Tomás, D., y Navarro-Colorado, B. (2019). A Systematic Review of Deep Learning Approaches to Educational Data Mining. *Complexity*, 2019.
- Hodges, C., Moore, S., Lockee, B., Trust, T. & Bond, A. (2020). The Difference Between Emergency Remote Teaching and Online Learning. *Educause Review*. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>
- Ingole, P., Bhoir, S., y Vidhate, A. V. (2018). Hybrid Model For Text Classification. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 7-15. <https://doi.org/10.1109/ICECA.2018.8474738>
- Ledesma, E. F. R., Díaz, D. A., and García, J. J. G. (2018). Tutor educativo usando máquina de soporte vectorial. *Pistas Educativas*, 35(108).
- Lizárraga Juárez, A., López Ramírez, E., & Martínez Iñiguez, J. E. (2020). Apropiación tecnológica en el manejo de competencias digitales por los profesores de escuelas normales en México. *Revista Boletín Redipe*, 9(6), 157-167. <https://doi.org/10.36260/rbr.v9i6.1009>
- Lugo, M. & Loíacono, L. (2020). Planificar la educación en la pospandemia: de la educación remota de emergencia a los modelos híbridos. *Educación y Tecnología*, 3(1). <https://publicaciones.flacso.edu.uy/index.php/edutic/article/view/2>
- Marotias, A. (2021). La educación remota de emergencia y los peligros de imitar lo presencial. *Hipertextos*, 8(14), 173-177. <https://doi.org/10.24215/23143924e025>
- Martínez Cámara, E., Martín Valdivia, M. T., Perea Ortega, J. M., and Ureña López, L. A. (2011). Técnicas de clasificación de opiniones aplicadas a un corpus en español.
- Montes y Gómez, M. (2001). Minería de texto: Un nuevo reto computacional. 3rd International Workshop on Data Mining MINDAT-2001.
- Mosquera, A., Lloret, E., and Moreda, P. (2012). Towards facilitating the accessibility of web 2.0 texts through text normalisation.
- Pérez-Suasnavas, A. L., Hasperué, W., Bustamante, M. E. M., & Santamaría, J. L. (2022). Predicción de dificultades estudiantiles mediante técnicas de minería de textos. *South Florida Journal of Development*, 3(5), 6128-6137.
- Portillo, S., Castellanos, L., Reynoso, O., & Gavotto, O. (2020). Enseñanza remota de emergencia ante la pandemia Covid-19 en Educación Media Superior y Educación Superior. *Propósitos y Representaciones*, 8 (SPE3), e589. Doi: <http://dx.doi.org/10.20511/pyr2020.v8nSPE3.589>
- Raut, V. B. and Londhe, D. (2014). Survey on opinion mining and summarization of user reviews on web. *International Journal of computer science and information technologies*, 5(2):1026-1030.
- Roatta, S. and Tedini, D. (2021). La pandemia del Covid-19 y el aprendizaje semipresencial en la educación superior. *Revista Iberoamericana de Tecnología en Educación y*

- Educación en Tecnología, no. 28, 318-323, 2021. doi: 10.24215/18509959.28.e39
- Romero, C. and Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6):601–618.
- Singh V., y Thurman A. (2019). How many ways can we define online learning? A systematic literature review of definitions of online learning (1988-2018). *American Journal of Distance Education*, 33(4), 289–306. <https://doi.org/10.1080/08923647.2019.1663082>
- Solis, P. Y. J., Sosa, T. I. P., Palma, M. L. T., Morocho, W. R. B., & Litardo, E. D. M. (2020). La educación virtual un nuevo reto para la formación del comunicador social de la extensión universitaria. *Magazine de las Ciencias: Revista de Investigación e Innovación*, 5 (CISE).
- UNESCO (2020). El coronavirus covid-19 y la educación superior: impacto y recomendaciones. Disponible en <https://www.iesalc.unesco.org/2020/04/02/el-coronavirus-covid-19-y-la-educacion-superior-impacto-y-recomendaciones/>
- Vásquez, A. C., Quispe, J. P., Huayna, A. M., et al. (2009). Procesamiento de lenguaje natural. *Revista de investigación de Sistemas e Informática*, 6(2):45–54.
- Xu, S. (2018). Bayesian naive bayes classifiers to text classification. *Journal of Information Science*, 44(1):48–59.
- Yang, K., Yu, N., and Zhang, H. (2007). Widit in trec-2007 blog track: Combining lexicon-based methods to detect opinionated blogs.
- Zafra, S. M. J., Valdivia, M. T. M., Cámara, E. M., and López, L. A. U. (2017). Studying the scope of negation for spanish sentiment analysis on twitter. *IEEE Transactions on Affective Computing*, 10(1):129–141.

