



Facultad de Energía, las Industrias y los Recursos Naturales no Renovables

CARRERA DE INGENIERÍA EN SISTEMAS

“Modelo de lenguaje utilizando RoBERTuito, para identificar tweets con contenido de violencia política de género hacia las asambleístas electas en Ecuador para el periodo 2021-2025”

“Language model using RoBERTuito, to identify tweets with gender-based political violence content towards female assemblywomen elected in Ecuador for the 2021-2025 term”

Línea de investigación: Sistemas Inteligentes

| |
|--|
| PROYECTO DE TITULACIÓN EN INGENIERÍA EN SISTEMAS. |
|--|

Autores:

◇ ORCID, Edmundo José Pezantes Urrego, edmundo.j.pezantes@unl.edu.ec

Tutor:

- Pablo F. Ordoñez-Ordoñez, Mg.Sc.

Cotutor especialista:

- Oscar M. Cumbicus-Pineda, Mg.Sc



Carrera de Ingeniería en
Sistemas / Computación

LOJA - ECUADOR
2022

Certificación de Tutoría

En calidad de Tutor y Cotutor del Proyecto de Trabajo de Titulación PTT, certificamos la tutela a Edmundo José Pezantes Urrego, con el tema **“Modelo de lenguaje utilizando RoBERTuito, para identificar tweets con contenido de violencia política de género hacia las assembleístas electas en Ecuador para el periodo 2021-2025” - “Language model using RoBERTuito, to identify tweets with gender-based political violence content towards female assemblywomen elected in Ecuador for the 2021-2025 term”**, quien ha cumplido con todas las observaciones requeridas. Es todo cuanto puedo decir en honor a la verdad, facultando al interesado hacer uso de la presente, así como el trámite de pertinencia del presente proyecto.

Loja, 19 de septiembre de 2022

Atentamente,

Pablo F. Ordoñez-Ordoñez, Mg.Sc.
TUTOR

Oscar M. Cumbicus-Pineda, Mg.Sc
COTUTOR

Certificación de Autoría del Proyecto

Yo, Edmundo José Pezantes Urrego , estudiante de la Universidad Nacional de Loja, declaro en forma libre y voluntaria que el presente Proyecto de Trabajo de Titulación que versa sobre **“Modelo de lenguaje utilizando RoBERTuito, para identificar tweets con contenido de violencia política de género hacia las asambleístas electas en Ecuador para el periodo 2021-2025”- “Language model using RoBERTuito, to identify tweets with gender-based political violence content towards female assemblywomen elected in Ecuador for the 2021-2025 term”**, así como la expresiones vertidas en la misma son autoría del compareciente, quien ha realizado en base a recopilación bibliográfica primaria y secundaria. En consecuencia, asumo la responsabilidad de la originalidad de la misma y el cuidado al remitirse a las fuentes bibliográficas respectivas para fundamentar el contenido expuesto.

Atentamente,

Edmundo José Pezantes Urrego

Índice general

| | |
|---|-----------|
| 1. Problemática | 1 |
| 1.1. Situación Problemática | 1 |
| 1.2. Problema de Investigación | 3 |
| 2. Justificación | 4 |
| 3. Objetivos | 6 |
| 3.1. General | 6 |
| 3.2. Específicos | 6 |
| 4. Alcance | 7 |
| 5. Marco Teórico | 8 |
| 5.1. Antecedentes | 8 |
| 5.1.1. Violencia política de género | 8 |
| 5.2. Fundamentación Teórica | 9 |
| 5.2.1. Minería de textos | 9 |
| 5.2.2. Descubrimiento de Conocimiento en Textos (KDT) | 9 |
| 5.3. Modelos de Lenguaje | 11 |
| 5.3.1. Bert | 11 |
| 5.3.1.1. Arquitectura del modelo BERT | 11 |
| 5.3.2. ROBERTA | 13 |
| 5.3.2.1. Arquitectura del modelo Roberta | 13 |
| 5.3.3. RoBERTuito | 14 |
| 5.3.3.1. Arquitectura del modelo RoBERTuito | 14 |
| 5.4. Trabajos Relacionados | 15 |
| 5.4.1. Conclusión | 21 |
| 6. Metodología | 22 |
| 7. Cronograma | 23 |
| 8. Presupuesto | 24 |
| 8.1. Recursos Humanos | 24 |
| 8.2. Recursos de Software, Hardware, TICs | 25 |
| 8.3. Insumos | 25 |

| | |
|---|-----------|
| 8.4. Recursos Totales | 25 |
| Bibliografía | 27 |
| Lista de Acrónimos y Abreviaturas | 30 |
| A. Anexo I | 31 |
| A.1. Revisión Sistemática de Literatura | 31 |
| B. Anexo II | 37 |
| B.1. Entrevista 1 | 37 |
| C. Anexo III | 40 |
| C.1. Entrevista 2 | 40 |
| D. Anexo IV | 42 |
| D.1. FORMATO PARA EL REGISTRO DE LAS ACTIVIDADES DE TU- TORÍA PARA LA ELABORACIÓN DEL PROYECTO DE TRABAJO DE TITULACIÓN | 42 |

Índice de figuras

| | |
|--|----|
| 5.1. Arquitectura BERT BASE y BERT LARGE | 12 |
| 5.2. Salida BERT como incrustaciones | 12 |
| 5.3. Hiperparámetros del modelo RoBERTuito pre-entrenado | 14 |
| 6.1. Metodología de PTT | 22 |
| 7.1. Cronograma de PTT | 23 |

Índice de tablas

| | |
|--|----|
| 5.1. Trabajos relacionados | 15 |
| 8.1. Recurso humano involucrado para el desarrollo del proyecto. | 24 |
| 8.2. Recursos Software, Hardware, TICs a utilizar para el desarrollo del proyecto. | 25 |
| 8.3. Insumos a utilizar para el desarrollo del proyecto. | 25 |
| 8.4. Presupuesto total para la elaboración del TT | 26 |

**“Modelo de lenguaje utilizando
RoBERTuito, para identificar
tweets con contenido de violencia
política de género hacia las
asambleístas electas en Ecuador
para el periodo 2021-2025”**

Línea de investigación: Sistemas Inteligentes

1. Problemática

1.1. Situación Problemática

Con la aparición de la Web 2.0 se dio el acceso al trabajo colaborativo, a la democratización de los contenidos y a la creación de las redes sociales donde son los usuarios los que crean y comparten su propio contenido [1]. Twitter se ha convertido en una de las plataformas más importantes de Microblogging. A través de ella, millones de usuarios pueden compartir intereses, actividades, contenidos e intercambiar experiencias y opiniones [2]. Twitter tiene otra esencia, permite describir lo que se está haciendo y la difusión de noticias; se caracteriza por la crítica en masas, en donde expresan su opinión y puntos de vista acerca de temas sociales, políticos, humanitarios, etc., generando información de gran interés para ser analizada.

Twitter tiene un inmenso potencial para la comunicación puesto que permite obtener información instantánea y de primera mano de algún suceso que esté pasando. Como Twitter es un espacio abierto, libre y público, da la posibilidad a que las personas pueden publicar cualquier información o hacer comentarios que no sean del agrado de los usuarios [3]. Twitter se constituye como nicho informativo en el que la ironía, los chistes, los insultos o los ataques confluyen de una manera naturalizada, en el que las mujeres se sitúan como un perfil habitual sobre el que se vuelcan todos ellos. Muchas mujeres, famosas o anónimas, sufren acoso y violencia en Twitter, por el mero hecho de participar en la red o por cualquier otro motivo asociado a un ideario patriarcal anclado en el imaginario social de parte de las/os usuarias/os. Silenciarlas es uno de los objetivos que pretenden este tipo de conductas violentas, consiguiendo en muchos casos que las usuarias accedan a borrar sus perfiles como medida preventiva contra la violencia [4].

El Instituto Nacional Demócrata (IND), por mucho tiempo ha liderado el apoyo a las mujeres de todo el mundo para superar las barreras a su participación política. Con demasiada frecuencia, la violencia contra las mujeres en la política (VAW-P), en todas sus manifestaciones, crea un efecto negativo que aleja a las mujeres políticamente activas y en algunos hasta las saca del ámbito político por completo [5].

En Ecuador, según datos de la Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres – ENVIGMU 2019, 65 de cada 100 mujeres en el país, han experimentado por lo menos un hecho de algún tipo de violencia en alguno

de los distintos ámbitos a lo largo de su vida. Uno de estos ámbitos es la esfera política, dentro de la cual las mujeres están expuestas a diferentes formas de violencia ya sea al interior de los partidos políticos, familias, comunidades e inclusive en las redes sociales [6].

Respecto a estos indicadores, Ecuador es un país expuesto a la violencia política de género contra las mujeres políticas, así como lo enfatiza el Abg. Luis Hernán Cisneros, Director del CNE Loja, a través de una entrevista (Anexo II), menciona que *“la violencia política de género, influye en la elección de todas las autoridades y también de manera específica en la toma de decisiones de las autoridades y de las mujeres que se presentan a estas dignidades, puesto que ese es el fin de la violencia política y especialmente de la violencia política de género, por lo cual considera que evidentemente si influye y es por eso que la norma establece límites y restricciones acerca de ese tipo de actos”*.

Si bien es cierto, una de las formas de prevenir estos delitos de odio, es detectando a tiempo los discursos de violencia política de género en la población, siendo esta una tarea difícil de conocer a través de las técnicas tradicionales como las encuestas debido a que demandaría de bastantes recursos y tiempo. Sin embargo, de acuerdo al estudio realizado por Santiago-Gonzales [7], mediante la comparación de rendimiento entre el modelo BERT (Bidirectional Encoder Representations from Transformers) y métodos más tradicionales de PLN (Procesamiento del Lenguaje Natural) en distintas tareas de clasificación de PLN, en la cual los experimentos demuestran el gran rendimiento de BERT en tareas de clasificación de PLN. El rendimiento de BERT ha sido mejor en todos los experimentos que han realizado que el de los métodos tradicionales empleados, dando una pauta de que mediante la implementación de modelo de lenguaje se posibilita determinar la existencia de la violencia política de género presente en las opiniones.

Además, es importante mencionar que Damian-Furman, Manuel-Pérez, Laura-Alemaný [8], indican que desde que apareció BERT, los modelos de lenguaje de Transformer y el aprendizaje por transferencia se han convertido en lo último en tareas de procesamiento de lenguaje natural. Actualmente, se realizan trabajos orientados hacia el pre-entrenamiento de modelos especialmente elaborados para dominios particulares, como los textos generados por usuarios. Se ha demostrado que estos modelos específicos de dominio mejoran significativamente el rendimiento en la mayoría de las tareas; sin embargo, para idiomas distintos del inglés, dichos modelos no están ampliamente disponibles. Además, Damian-Furman, Manuel-Pérez, Laura-Alemaný [8], en su trabajo presentan RoBERTuito, un modelo de lenguaje preentrenado para texto generado por usuarios en español, entrenado en más de 500 millones de tuits. Los experimentos en un punto de referencia de tareas que involucran texto generado por el usuario mostraron que RoBERTuito superó a otros modelos de lenguaje previamente entrenados en español.

1.2. Problema de Investigación

De acuerdo al planteamiento de la problemática, surgen interrogantes acerca de que si a través de la creación de un modelo de lenguaje a partir del modelo preentrenado RoBERTuito, será posible detectar opiniones con contenido de violencia política de género en Twitter y determinar su existencia en Ecuador hacia las asambleístas electas para el periodo 2021-2025, considerando esto se puede deducir la siguiente pregunta de investigación:

¿Cómo identificar la existencia de violencia política de género hacia las asambleístas electas en Ecuador para el periodo 2021-2025 en twitter, utilizando el modelo de lenguaje RoBERTuito?

2. Justificación

El presente trabajo de titulación (TT), busca determinar la existencia de contenido de violencia política de género en Ecuador en la red social Twitter. Aportando información relevante que puede ser utilizada por algunas organizaciones como: El Consejo Nacional Electoral (CNE), el Instituto de la Democracia (IDD) y el Consejo para la Igualdad de Género que han venido trabajando conjuntamente en campañas para fortalecer el sistema democrático desde la perspectiva de la igualdad y la no violencia política basada en el género [9]. De igual manera, la ONU Mujeres mantiene el compromiso de apoyar en el fortalecimiento de las capacidades institucionales de las autoridades que atienden la violencia política de género para que promuevan, defiendan y garanticen los derechos políticos-electorales de las mujeres en el Ecuador[10].

En el Ecuador, la Ley Orgánica Integral para Prevenir y Erradicar la Violencia contra las Mujeres, reconoce en su artículo 10, literal literal f), a la violencia política como un tipo de violencia y la define como aquella violencia cometida por una persona o grupo de personas, directa o indirectamente, en contra de las mujeres que sean candidatas, militantes, electas, designadas o que ejerzan cargos públicos, defensoras de derechos humanos, feministas, lideresas políticas o sociales, o en contra de su familia.

Por otra parte, se destaca que este TT mediante la evaluación de las opiniones en Twitter a través del uso de un Modelo de Lenguaje, pretende determinar la existencia de contenido de violencia política de género en nuestro país, información que puede llegar a contribuir a aquellas organizaciones que defienden la participación política de la mujeres, para sus diferentes actividades e incluso para la toma de decisiones, siendo esto favorable para garantizar la participación política de las mujeres, ya que al determinar a tiempo los sentimientos de violencia política de género y su nivel de presencia en las personas se puede evitar el que se vea limitada su participación política a una mayor escala y por ende se pueda llegar a producir una violencia física o limitantes en cuanto a funciones políticas concierne y el recibir sanciones, ya que según, La Ley Orgánica Electoral, Código de la Democracia, en el artículo 70.- En la quinta función, manifiesta que el Tribunal Contencioso Electoral tendrá la función de sancionar el incumplimiento de las normas sobre financiamiento, propaganda, gasto electoral, no discriminación o violencia política de género, paridad de género, inclusión de jóvenes y demás vulneraciones de normas electorales [11].

El Abg. Luis Hernán Cisneros, Director del CNE Loja, a través de una entrevista Anexo II, menciona que *“el presente TT aportaría como estudio el poder aplicar esta tecnología y esta metodología que nos facilita o que nos presenta la inteligencia artificial que a través de un análisis de sentimientos determinar la existencia de conductas violencia política de género y más aun siendo temas académicos científicos en violencia política de género, sería un estudio que aportaría de mucho para temas de conocimiento de respuestas, de reacciones y de soluciones en todo tipo de materias en este sentido. Además, enfatiza dado que a la inteligencia artificial ya se la ocupa con la finalidad de plantear estrategias de campaña a través del análisis de sentimientos con el fin de poder concretar un mensaje o discurso, del mismo modo se la debe utilizar para poder identificar actos de violencia política de género. Sería un aporte fundamental puesto que nos permitiría conocer sobre la base de números reales la incidencia”*.

Así mismo, la Ing.Neli Virmania Troya Troya, Concejal del Gobierno Autónomo Descentralizado del Cantón Calvas, de acuerdo a una entrevista Anexo III, manifiesta que *“ al presente TT se lo puede tomar como un punto de partida para tomar medidas o acciones para tratar de erradicar el contenido de violencia política de género en las redes sociales, en twitter en este caso. Además, menciona que todas la herramientas o medios que sirvan para identificar este tipo de contenido de violencia política de género es necesario, de modo que se pueda tratar de controlar o erradicar la violencia política de género, por lo cual sería de mucha utilidad.”*

Por otra parte, el presente Proyecto de Trabajo de Titulación se enmarca a nivel internacional a los Objetivos de Desarrollo Sostenible ODS N.º 5 “Lograr la igualdad entre los géneros y empoderar a todas las mujeres y las niñas” y N.º 9 “Construir infraestructuras resilientes, promover la industrialización sostenible y fomentar la innovación” propuesto por la Organización de las Naciones Unidas (ONU)[12], y localmente en la línea de investigación de Sistemas Inteligentes de la Carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja (UNL) [13].

3. Objetivos

3.1. General

Construir un modelo de lenguaje utilizando RoBERTuito, para identificar tweets con contenido de violencia política de género hacia las asambleístas electas en Ecuador para el periodo 2021-2025.

Build a language model using RoBERTuito, to identify tweets with gender-based political violence content towards elected assemblywomen in Ecuador for the period 2021-2025.

3.2. Específicos

- Construir un conjunto de datos proveniente de la red social Twitter, basado en las menciones a las cuentas oficiales de las Asambleístas ecuatorianas electas para el periodo 2021-2025.

To build a dataset from the social network Twitter, based on the official accounts of the Ecuadorian Assemblywomen elected for the period 2021-2025.

- Construir la arquitectura del modelo utilizando el modelo de lenguaje RoBERTuito.

Build the model architecture using the RoBERTuito language model.

- Evaluar la precisión del modelo de lenguaje construido mediante una muestra de usuarios.

Evaluate the accuracy of the language model constructed using a sample of users.

4. Alcance

El presente TT contempla el descubrir tuits con contenido de violencia política de género hacia las asambleístas electas en Ecuador en el año 2021 mediante el uso de un modelo de lenguaje. A continuación, se detalla el alcance por etapas con las respectivas actividades a realizar:

1. **Fase:** Construir un conjunto de datos proveniente de la red social Twitter, basado en las menciones a las cuentas oficiales de las Asambleístas ecuatorianas electas para el periodo 2021-2025.
 - a) Recolección de los datos
 - b) Limpieza de datos
 - c) Preprocesamiento y transformación del conjunto de datos
2. **Fase:** Construir la arquitectura del modelo utilizando el modelo de lenguaje RoBERTuito.
 - a) Diseñar la arquitectura del modelo.
 - b) Entrenamiento del modelo de lenguaje creado utilizando RoBERTuito.
3. **Fase:** Evaluar la precisión del modelo de lenguaje construido mediante una muestra de usuarios.
 - a) Interpretar los resultados obtenidos de la muestra.
 - b) Dar contestación a la pregunta de investigación.
 - c) Elaborar el documento final.

5. Marco Teórico

En esta sección se presentan las definiciones de aquellos conceptos que aportan al cumplimiento de los objetivos y contribuyen al conocimiento teórico de los mismos; junto con el establecimiento de teorías y hallazgos de trabajos relacionados al presente TT.

5.1. Antecedentes

5.1.1. Violencia política de género

Según el Mecanismo de Seguimiento de la Convención de Belém do Pará (MESECVI - OEA) la violencia política contra las mujeres debe entenderse como: Cualquier acción, conducta u omisión, realizada de forma directa o a través de terceros que, basada en su género, cause daño o sufrimiento a una o varias mujeres, y que tenga por objeto o por resultado menoscabar o anular el reconocimiento, goce o ejercicio de sus derechos políticos. La violencia política contra las mujeres puede incluir entre otras, violencia física, sexual, psicológica, moral, económica o simbólica [14].

En el Ecuador, el Consejo Nacional Para La Igualdad De género señala que la Ley Orgánica Integral para Prevenir y Erradicar la Violencia contra las Mujeres en el Art. 10. Literal f), define a la Violencia política como aquella violencia cometida por una persona o grupo de personas, directa o indirectamente, en contra de las mujeres que sean candidatas, militantes, electas, designadas o que ejerzan cargos públicos, defensoras de derechos humanos, feministas, lideresas políticas o sociales, o en contra de su familia. Esta violencia se orienta a acortar, suspender, impedir o restringir su accionar o el ejercicio de su cargo, o para inducirla u obligarla a que efectúe en contra de su voluntad una acción o incurra en una omisión, en el cumplimiento de sus funciones, incluida la falta de acceso a bienes públicos u otros recursos para el adecuado cumplimiento de sus funciones [15].

5.2. Fundamentación Teórica

5.2.1. Minería de textos

La minería de textos es considerada como una de las áreas con mayor potencial dentro de la minería de datos, debido a que la manera habitual de almacenar información es en forma de textos. La minería de textos es, sin embargo, una disciplina mucho más compleja que la minería de datos, que trata con bases de datos estructurados. Esta tarea, por el contrario, requiere tratar con datos textuales que son propiamente desestructurados y mas confusos [16].

A la extracción, análisis y representación de información proveniente de la interacción, entre los usuarios en redes sociales, se le denomina minería de medios sociales. Varias técnicas de minería de datos son aplicadas en la minería de texto, siendo esta última una prolongación de la primera técnica. Minar texto, en conclusión, se fundamenta en extraer información relevante de textos no estructurados, sin importar el tamaño de estos, con el fin de obtener modelos que proporcionen información relevante y posterior conocimiento [17].

5.2.2. Descubrimiento de Conocimiento en Textos (KDT)

Según Justicia de la Torre [18], el descubrimiento de conocimiento se ha venido realizando sobre los datos almacenados en bases de datos estructuradas. No obstante, la mayoría de la información disponible a tratar está en formato textual. Se define al descubrimiento de conocimiento como la extracción de información previamente desconocida y potencialmente útil a partir de los datos, el presente TT adopta esta definición considerando que los datos son textuales. Los datos textuales no estructurados son los que hacen emerger una serie de problemas que necesitan un tratamiento adicional. Dichos problemas son:

- La falta de estructura del texto.
- La naturaleza heterogénea y distribuida de los documentos; las fuentes de textos pueden ser tan diversas como podemos imaginar: bases de datos documentales, redes sociales, censos, páginas web, informes empresariales, publicaciones, correos electrónicos, etc.
- El multilingüismo o plurilingüismo puede estar presente tanto en diferentes conjuntos de documentos como dentro de una misma colección de textos.

Por lo tanto, el proceso KDT agrupa a diferentes ámbitos de conocimiento, como son:

- Recuperación de Información (para filtrar y reunir documentos adecuados).
- Extracción de Información (que selecciona hechos específicos sobre tipos de entidades y relaciones de interés).
- Procesamiento del Lenguaje Natural (para realizar el preprocesamiento y etiquetado de los texto).
- Minería de Datos (para descubrir asociaciones desconocidas entre hechos desconocidos).

Del mismo modo, Justicia de la Torre [18], plantea una serie de fases a seguir para la realización del proceso KDT.

1. Definición de conceptos: Estos conceptos son escogidos según el objetivo de cada investigador.
 2. Recopilación de información: El objetivo de esta etapa es recopilar la información o datos que serán objeto de estudio, esta recopilación se da a través de herramientas de recopilación, como la extracción de textos de páginas web a través de API's o Web Scraping.
 3. Pre-Procesamiento: El texto no tiene una estructura adecuada para el descubrimiento de conocimiento, por lo que se necesita asegurar la homogeneidad del texto, es decir, que no existan conceptos o caracteres que distorsionen posteriormente la detección de patrones.
 4. Transformación: Se procede a dar la estructura necesaria al texto para la posterior detección de patrones o minería de texto, a continuación, se mencionan los procesos clave:
 - Lematizado: Es el proceso de transformar las variaciones de las palabras con morfemas a su raíz (lexema), por ejemplo, el caso de la palabra “comidita” en la cual el morfema es “ita” su lexema por ende será “comida” debe ser transformada a su lexema, con el fin de evitar tener variantes de las palabras.
 - Etiquetado: el etiquetado gramatical o desambiguación, es el proceso de clasificar una palabra en función tanto de su definición como de su contexto, es decir, su relación con las palabras adyacentes y relacionadas en una frase,
-

oración o párrafo. Este se logra a través de clasificar palabras como sustantivos, verbos, adjetivos, adverbios, etc.

- Tokenizado: Es la acción de dividir el texto en elementos llamados tokens, éstos son elementos indivisibles, por ejemplo, las frases se separan en palabras (tokens) con el fin de construir una matriz de palabras, en donde los encabezados de cada columna serán las palabras existentes dentro del corpus de los tweets.
5. Minado de texto: Esta es una de las etapas más amplias, caracterizándose en el descubrimiento de conocimiento, el cual se puede dar a través de detección de patrones, representaciones vectoriales, modelos de aprendizaje supervisado o no supervisado entre otros métodos, la elección del método dependerá del objetivo del investigador, en este capítulo se realizó una revisión de algunos métodos que fueron aplicados en la investigación.
 6. Interpretación de resultados: En esta etapa se da paso a interpretar y validar el conocimiento obtenido tras realizar el proceso, sin embargo, puede no ser la última etapa, esto dependerá de la validación del contenido en caso de ser rechazado por el investigador podría resultar en el retroceso en el proceso.

5.3. Modelos de Lenguaje

5.3.1. Bert

BERT está diseñado para entrenar previamente representaciones bidireccionales profundas a partir de texto sin etiquetar mediante el condicionamiento conjunto del contexto izquierdo y derecho en todas las capas. Como resultado, el modelo BERT preentrenado se puede ajustar con solo una capa de salida adicional para crear modelos de última generación para una amplia gama de tareas, como la respuesta a preguntas y la inferencia de lenguaje, sin modificaciones sustanciales en la arquitectura específica de la tarea [19].

5.3.1.1. Arquitectura del modelo BERT

BERT se lanza en dos tamaños BERT BASE y BERT LARGE . El modelo BASE se utiliza para medir el rendimiento de la arquitectura comparable a otra arquitectura y el modelo LARGE produce resultados de vanguardia que se informaron en el artículo de investigación. BERT es básicamente una pila de codificadores de arquitectura

de transformadores. Una arquitectura de transformador es una red de codificador-descodificador que utiliza la atención propia en el lado del codificador y la atención en el lado del descodificador. BERT BASE tiene 12 capas en la pila del codificador mientras que BERT GRANDE tiene 24 capas en la pila del codificador. BERT BASE contiene 110 M de parámetros mientras que BERT LARGE tiene 340M de parámetros [20].

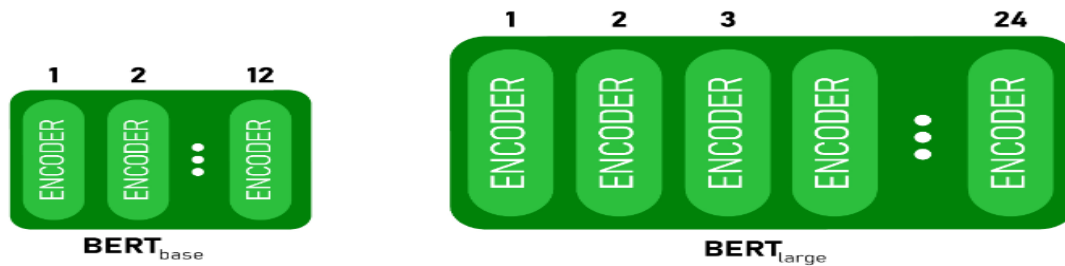


Figura 5.1: Arquitectura BERT BASE y BERT LARGE

Este modelo toma el token CLS como entrada primero, luego es seguido por una secuencia de palabras como entrada. Aquí CLS es un token de clasificación. Luego pasa la entrada a las capas anteriores. Cada capa aplica atención propia, pasa el resultado a través de una red de avance y luego pasa al siguiente codificador.

El modelo genera un vector de tamaño oculto (768 para BERT BASE). Si queremos generar un clasificador de este modelo, podemos tomar la salida correspondiente al token CLS [20].

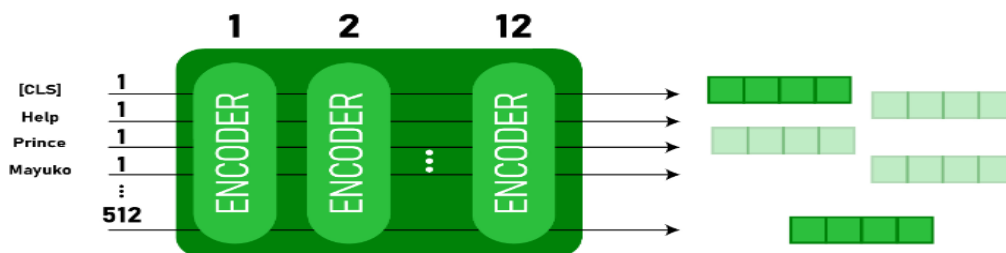


Figura 5.2: Salida BERT como incrustaciones

5.3.2. ROBERTA

ROBERTA tiene un enfoque de preentrenamiento BERT robustamente optimizado. Fue presentado por investigadores de Facebook y la Universidad de Washington. El objetivo de este artículo fue optimizar el entrenamiento de la arquitectura BERT para tomar menos tiempo durante el entrenamiento previo [20].

5.3.2.1. Arquitectura del modelo Roberta

ROBERTA tiene una arquitectura casi similar en comparación con BERT , pero para mejorar los resultados en la arquitectura BERT, se le realizaron algunos cambios de diseño simples en su arquitectura y procedimiento de capacitación [21]. Estos cambios son:

- Eliminación del objetivo de predicción de la siguiente oración (NSP) : en la predicción de la siguiente oración, el modelo está entrenado para predecir si los segmentos de documentos observados provienen del mismo documento o de documentos distintos a través de una pérdida auxiliar de predicción de la siguiente oración (NSP). Los autores experimentaron con la eliminación / adición de la pérdida de NSP a diferentes versiones y concluyeron que la eliminación de la pérdida de NSP coincide o mejora ligeramente el rendimiento de las tareas posteriores [21].
 - Entrenamiento con tamaños de lote más grandes y secuencias más largas: Originalmente, BERT está entrenado para pasos de 1M con un tamaño de lote de 256 secuencias. En este artículo, los autores entrenaron el modelo con 125 pasos de secuencias de 2K y 31K pasos con secuencias de 8k de tamaño de lote. Esto tiene dos ventajas, los lotes grandes mejoran la perplejidad sobre el objetivo de modelado del lenguaje enmascarado y la precisión de la tarea final. Los lotes grandes también son más fáciles de paralelizar mediante el entrenamiento paralelo distribuido[21].
 - Cambio dinámico del patrón de enmascaramiento : en la arquitectura BERT, el enmascaramiento se realiza una vez durante el preprocesamiento de datos, lo que da como resultado una sola máscara estática. Para evitar el uso de una sola máscara estática, los datos de entrenamiento se duplican y se enmascaran 10 veces, cada vez con una estrategia de máscara diferente durante 40 épocas, por lo que se obtienen 4 épocas con la misma máscara. Esta estrategia se compara con el enmascaramiento dinámico en el que el enmascaramiento diferente para cada vez que pasamos datos al modelo[21].
-

5.3.3. RoBERTuito

RoBERTuito, un modelo transformador a gran escala para texto generado por el usuario entrenado en tuits españoles. RoBERTuito es un modelo muy competitivo en entornos multilingües y de cambio de código, incluido el español y el inglés. Tres versiones de RoBERTuito fueron entrenadas: una versión en caja que conserva el caso encontrado en los tweets originales, una versión sin caja y una versión sin caja y sin acento, que pone en minúsculas y elimina los acentos en tuits [8].

5.3.3.1. Arquitectura del modelo RoBERTuito

Se utilizó una arquitectura base RoBERTa en RoBERTuito, con 12 capas de auto-atención, 12 cabezas de atención y un tamaño oculto igual a 768, de la misma manera que BERTweet. Utiliza un objetivo de lenguaje enmascarado, sin tener en cuenta la tarea de predicción de la siguiente oración utilizada en BERT u otras tareas de orden de tweets [8].

Teniendo en cuenta los hiperparámetros exitosos de RoBERTa y BERTweet, se utiliza un tamaño de lote grande para el entrenamiento. Si bien RoBERTa recomienda un tamaño de lote de 8K, debido a las limitaciones de recursos, se equilibra la cantidad de actualizaciones utilizando un tamaño de 4K. Para verificar la convergencia, primero se entrena un modelo sin carcasa para pasos de 200K. Después de esto, se ejecuta en pasos de 600K para los tres modelos. Esto es aproximadamente la mitad de la cantidad de actualizaciones que se usan para entrenar a BETO (y también a BERTweet), pero esta diferencia se compensa con el tamaño de lote más grande que se usa para entrenar a RoBERTuito [8].

| Hyperparameter | Value |
|-------------------|-----------------|
| #Heads | 12 |
| #Layers | 12 |
| Hidden Size | 768 |
| Intermediate Size | 3072 |
| Hidden activation | GeLU |
| Vocab. size | 30,000 |
| MLM probability | 0.15 |
| Max Seq length | 128 |
| Batch Size | 4,096 |
| Learning Rate | $3.5 * 10^{-4}$ |
| Decay | 0.1 |
| β_1 | 0.9 |
| β_2 | 0.98 |
| ϵ | 10^{-6} |
| Warmup steps | 36,000 (6%) |
| Training steps | 600,000 |

Figura 5.3: Hiperparámetros del modelo RoBERTuito pre-entrenado

5.4. Trabajos Relacionados

A través de la realización de una revisión sistemática de literatura entorno a los temas de interés en relación al TT, se ha obtenido un conjunto de trabajos o estudios relacionados al objeto de estudio (Tabla 5.1), los mismos que fundamentan el presente TT y sirven de ayuda para comprender la forma en la que se desarrollan temas similares. El proceso de su obtención se detalla en el Anexo I.

Tabla 5.1: Trabajos relacionados

| Código | Trabajo relacionado | Ref. | Buscador |
|--------|--|------|----------------|
| TR01 | Análisis de sentimiento de comentarios en español en Google Play Store usando BERT | [22] | Google Scholar |
| TR02 | Estudio del rendimiento de BERT frente a métodos clásicos de procesamiento de lenguaje natural | [7] | Google Scholar |
| TR03 | RoBERTuito: a pre-trained language model for social media text in Spanish | [8] | Google Scholar |
| TR04 | “Procesamiento del lenguaje natural con BERT: Análisis de sentimientos en tuits” | [23] | Google Scholar |
| TR05 | A Review of Natural Language Processing Techniques for Sentiment Analysis using Pre-trained Models | [24] | IEEE |
| TR06 | Aspect-Based Sentiment Analysis of Social Media Data With Pre-Trained Language Models | [25] | IEEE |
| TR07 | Comparing pre-trained language models for Spanish hate speech detection | [26] | Science Direct |
| TR08 | A Comparative Study of Using Pre-trained Language Models for Toxic Comment Classification | [27] | Google Scholar |
| TR09 | Detection of extreme sentiments on social networks with BERT | [28] | Springer |
| TR10 | Transformer models for text based emotion detection: a review of BERT based approaches | [29] | Springer |
| TR11 | Pre-trained transformers: an empirical comparison | [30] | Science Direct |
| TR12 | Pre-trained Language Models with Limited Data for Intent Classification | [31] | IEE |
| TR13 | Using Transfer-based Language Models to Detect Hateful and Offensive Language Online | [32] | Google Scholar |
| TR14 | Minado de texto aplicado en twitter para obtener la polaridad en opiniones de usuarios acerca del nuevo proyecto de ley que regula la migración en Chile | [33] | Google Scholar |

| | | | |
|------|--|------|----------------|
| TR15 | Análisis de Sentimientos en Twitter para Descubrir Contenido Xenófobo hacia los Inmigrantes Venezolanos en Ecuador | [34] | Google Scholar |
| TR16 | Detección de lenguaje ofensivo en redes sociales | [35] | Google Scholar |

En el estudio TR01, el autor manifiesta que debido a la gran cantidad de datos generados en las plataformas virtuales, el procesamiento del lenguaje natural se ha vuelto más frecuente. El análisis de sentimientos es una de las aplicaciones de clasificación de texto que se puede utilizar para analizar estos comentarios. El análisis de sentimientos se ha llevado a cabo en inglés, mientras que, en otros idiomas como el español, no ha habido muchos intentos. En este trabajo se usa un modelo pre-entrenado BERT para el análisis de sentimiento de comentarios en español de la tienda de aplicaciones de Google Play Store, los resultados experimentales indican que después del pre-procesamiento adecuado puede mostrar resultados prometedores el conjunto de datos en español, utilizando el modelo BERT se puede alcanzar una precisión de 0.81 en promedio, incluso con la limitación de datos.

En el trabajo TR02, se realiza una comparación de rendimiento entre el modelo BERT (Bidirectional Encoder Representations from Transformers) y métodos más tradicionales de PLN (Procesamiento del Lenguaje Natural) en distintas tareas de clasificación de PLN. Para representar los métodos tradicionales utilizamos un primer paso de pre-procesamiento, en el que vectoriza los datos, seguido por un segundo paso de generación del modelo. Contrasta, con distintos experimentos, la hipótesis de que usar BERT como método predeterminado en problemas de PLN es mejor que usar TF-IDF junto con un algoritmo de aprendizaje automático. En total, realiza cuatro experimentos, y, en todos ellos, han utilizado el modelo BERT y un método tradicional, el rendimiento de BERT ha sido mejor en todos los experimentos que han realizado que el de los métodos tradicionales empleados.

En el estudio TR03, presenta RoBERTuito, un modelo de lenguaje preentrenado para texto generado por usuarios en español, entrenado en más de 500 millones de tuits. Los experimentos en un punto de referencia de tareas que involucran texto generado por el usuario mostraron que RoBERTuito superó a otros modelos de lenguaje previamente entrenados en español. Además de esto, el presente modelo tiene algunas habilidades multilingües, logrando los mejores resultados para las tareas de inglés-español del parámetro de evaluación de cambio de código lingüístico (LinCE) y también un rendimiento competitivo frente a modelos monolingües en tareas de Twitter en inglés.

En el trabajo TR04, sus objetivos se dividen en función de si están centrados en el análisis de sentimientos o en comprender las capacidades de BERT. En primer lugar y por lo que respecta al análisis de sentimientos, el objetivo es realizarlo sobre tuits en

español, además el análisis es multi etiqueta sobre el dataset que ha sido proporcionado. Dentro del análisis de sentimientos se divide en dos partes, analizar con éxito en español y hacerlo sobre tuits que presentan sus particularidades. Por otro lado, se encuentran los objetivos que respectan a BERT, en primer lugar, se analiza la viabilidad de BERT para aplicarlo al procesamiento del lenguaje natural en español, y en segundo lugar la facilidad o dificultad que entraña hacer uso de la herramienta, analizando sus capacidades y distintas formas de trabajar con la misma.

En el estudio TR05, se analizan varios métodos de incrustación de palabras utilizados para el análisis de sentimientos, seguidos de una visión general de los modelos preformados de última generación utilizados para el procesamiento del lenguaje natural, que se utilizan habitualmente en el proceso de análisis de sentimientos. También se analizan los resultados experimentales de dos modelos preentrenados de última generación.

En el presente estudio se presentan diferentes modelos pre-entrenados como ULM-FiT, Transformer, GPT-2 de OpenAI, BiGRU, BERT de Google, Transformer-XL y XLNet. Aunque XLNet supera en más tareas a otros modelos, su complejidad computacional es muy alta. Requiere más tiempo de entrenamiento debido a su naturaleza bidireccional autorregresiva y también necesita un mejor hardware. La evaluación del rendimiento con los dos modelos más recientes, BiGRU y BERT utilizando el conjunto de datos IMDB muestra una mejora en la precisión con el número de épocas en las que BERT muestra un rendimiento superior.

En el trabajo TR06, presenta un nuevo método para la tarea de detección de categorías de aspecto (ACD) que no depende de los datos etiquetados mediante el uso de expresiones regulares (Regex). Para la tarea de clasificación de aspectos (ASC) se explora una técnica de aprendizaje de aprendizaje semi-supervisado. Además, se incorporan a las predicciones las etiquetas Part-of-Speech (POS) se incorporan a las predicciones. Los resultados muestran que Regex es una solución para eliminar la dependencia de los datos etiquetados para el ACD. En el caso de ASC, el ajuste de BERT en un pequeño subconjunto de datos fue el método más preciso para reducir la dependencia de los datos de sentimiento.

En el trabajo TR07, abordan la tarea de identificación del discurso de odio en español en las redes sociales y proporcionamos una comprensión más profunda de las capacidades de las nuevas técnicas basadas en el aprendizaje automático. En particular, compara el rendimiento de los métodos de Deep Learning con modelos lingüísticos recientemente preentrenados basados en Transfer Learning, así como con modelos de aprendizaje automático tradicionales. Se aplican modelos lingüísticos preentrenados multilingües y monolingües preentrenados como BERT, XLM y BETO.

La evaluación de los modelos recientemente pre-entrenados basados en el mecanismo Transformer para la detección de HS en español obteniendo resultados prometedores

para este idioma. En particular, compara el desempeño de dos modelos pre-entrenados multilingües (mBERT y XLM) actualmente disponibles con uno monolingüe (BETO) entrenado específicamente en español. Los resultados obtenidos muestran que el LM preentrenado monolingüe (BETO) superó a mBERT y XLM y por lo tanto concluye que es necesario entrenar un modelo en español, ya que el sistema es capaz de modular con mayor precisión el vocabulario.

En el estudio TR08, se realiza un estudio de cómo hacer un mejor uso de los métodos basados en modelos de lenguaje previamente entrenados para la clasificación de comentarios tóxicos y el desempeño de diferentes modelos de lenguaje previamente entrenados en estas tareas. Los resultados obtenidos muestran que, de los tres modelos de lenguaje más populares, es decir, BERT, Roberta y XLM, BERT y Roberta generalmente superan a XLM en la clasificación de comentarios tóxicos. También se demuestra que el uso de una estructura descendente lineal básica supera a las complejas como CNN y BiLSTM.

En el artículo TR09, se propone un enfoque no supervisado para detectar automáticamente opiniones/publicaciones extremas en las redes sociales. El rendimiento del enfoque propuesto se evalúa en cinco conjuntos de datos de medios y redes sociales diferentes. En este trabajo, utiliza un enfoque semisupervisado conocido como BERT para evaluar la precisión del enfoque y el conjunto de datos clasificados obtenidos. El experimento demuestra que en estos conjuntos de datos, las publicaciones que previamente se clasificaron como extremas negativas o positivas son extremadamente negativas o positivas en muchos casos mientras se usa BERT. Además, se indica que BERT muestra la capacidad de clasificar los sentimientos extremos cuando se ajusta con un conjunto de datos de sentimientos extremos apropiado.

El estudio TR10, analiza modelos basados en transformadores para tareas de PNL. Destaca los pros y los contras de los modelos identificados. Los modelos discutidos incluyen el preentrenamiento generativo (GPT) y sus variantes, Transformer-XL, Crosslingual Language Models (XLM) y las representaciones de codificador bidireccional de los transformadores (BERT). Teniendo en cuenta la fuerza y la popularidad de BERT en la detección de emociones basada en texto, el documento analiza trabajos recientes en los que los investigadores propusieron varios modelos basados en BERT. La encuesta presenta sus contribuciones, resultados, limitaciones y conjuntos de datos utilizados.

El estudio TR11, presenta un doble propósito: primero describe cinco modelos populares de transformadores y examina su uso típico en la literatura anterior, centrándose en la reproducibilidad; luego, realiza comparaciones en un entorno controlado sobre una amplia gama de tareas de PNL. Su análisis revela que solo una minoría de los artículos recientes de NLP que usan transformadores pre-entrenados informaron múltiples ejecuciones (20%). desviación estándar o significación estadística (10 %) y otra información

crucial, lo que perjudica gravemente la replicabilidad y la reproducibilidad. A través de una amplia comparación empírica de conjuntos de datos y puntos de referencia del mundo real, también muestra cómo los hiperparámetros y la semilla inicial impactan en los resultados, y destaca la solidez de los modelos bajos. Exponen la desviación estándar y el rango promedio; el rango promedio se calcula clasificando el rendimiento de los modelos en los conjuntos de datos individuales (1 a 5, cuanto más bajo mejor) y luego promediando los resultados en todos los conjuntos de datos. Los modelos RoBERTa y XLNet, que desarrollaron varias estrategias para mejorar la precisión del BERT simple, son los dos métodos que logran, en promedio, mejores resultados (2.17 de rango promedio). DistilBERT y ALBERT, que se enfocan en mejorar la eficiencia de BERT, logran resultados considerablemente más bajos (4.67 y 3.00 de rango promedio).

En el estudio TR12, realiza un ajuste fino del modelo BERT a través del aprendizaje por transferencia inductiva e investigamos métodos para superar los desafíos debido a la disponibilidad limitada de datos al proponer un enfoque novedoso de aumento de datos semánticos. Esta técnica genera oraciones sintéticas mientras preserva la compatibilidad de etiquetas utilizando el significado semántico de las oraciones, para mejorar la precisión de la clasificación de intenciones. Por lo tanto, con base en las consideraciones para el ajuste fino y el aumento de datos, se presenta una metodología paso a paso sistemática y novedosa para aplicar el modelo lingüístico BERT para la clasificación de intenciones con datos limitados disponibles. Nuestros resultados muestran que el lenguaje preentrenado se puede usar de manera efectiva con datos de redes sociales ruidosos para lograr una precisión de vanguardia en el análisis de intenciones bajo un régimen de datos etiquetados bajos.

En el trabajo TR13, se realiza cuatro estudios profundos basados en las Representaciones de codificador bidireccional de Transformers (BERT), con modelos de lenguaje generales o específicos de dominio, se probaron contra dos conjuntos de datos que contenían tweets etiquetados como 'Odio', 'Normal' u 'Ofensivo'. Los resultados indican que los modelos basados en la atención confunden profundamente el discurso de odio con el lenguaje ofensivo y normal. Sin embargo, los modelos pre-entrenados superan los resultados de última generación en términos de predicción precisa de las instancias odiosas.

En el trabajo TR14, se utiliza la metodología KDT para la realización de su trabajo, el proceso de descubrimiento de conocimiento en texto, implica dominio en diferentes áreas de conocimiento, métodos de recuperación de información, extracción de la información, procesamiento del lenguaje natural y minería de datos.

Para la etapa de recopilación de información dentro del proceso KDT, las redes sociales son una gran oportunidad para la recopilación de datos, por todo el contenido textual que vierten los usuarios en ellas, la recopilación de este contenido se podría realizar a través de webscraping, una técnica utilizada para la extracción de datos en

páginas web que simula navegación humana, sin embargo las redes sociales en general utilizan API's para facilitar la extracción de datos a los desarrolladores.

En el trabajo TR15, realiza una revisión bibliográfica para determinar las herramientas y metodologías a utilizar, en la cual selecciona la metodología KDT, Python como lenguaje de programación fue seleccionado principalmente por el gran número de librerías de uso libre que están disponibles para este lenguaje, además que están orientadas a la realización de proyectos de machine learning, también destacó por la simplicidad en su sintaxis, en comparación con otros lenguajes. Google Colab fue considerado como un entorno para ejecutar código Python, principalmente por las prestaciones que ofrece en cuanto a recursos de hardware se refiere, específicamente, contribuyó en la tarea de Machine Translation. En cuanto a las Bibliotecas o Librerías utilizadas se listan contención:

- Twitter Scraper: Se consideró utilizar esta librería principalmente por las facilidades que ofrece para la extracción de tuits, pues no tiene límites en la antigüedad de los mismos, además de que no requiere disponer de una cuenta en Twitter.
- Pandas: Se la consideró por ser una poderosa herramienta que permite manipular los datos de una forma rápida y eficaz.
- NLTK: Este conjunto de herramientas para el procesamiento de lenguaje natural fue considerado por su amplio conjunto de librerías que ofrece para la tokenización, derivación, etiquetado, análisis y razonamiento semántico, además que proporcionó una lista de stopwords para el preprocesamiento de los textos
- RegEx: Se consideró emplear esta librería para facilitar la limpieza y preprocesamiento de los datos, pues, mediante el uso de expresiones regulares fue capaz de eliminar los símbolos especiales, espacios, signos de puntuación, etc.
- Matplotlib: Para la representación gráfica de los resultados del presente TT, se consideró utilizar esta poderosa librería que es capaz de crear visualizaciones estáticas, animadas e interactivas.

En el trabajo TR16, se aplican técnicas de procesamiento del lenguaje para desarrollar modelos de aprendizaje supervisados capaces de detectar comentarios ofensivos en redes sociales. Para obtener el mejor rendimiento en cada modelo se probaron distintos tratamiento de los textos que irán desde no aplicar ningún tratamiento hasta el truncado, eliminación y corrección de palabras. Posteriormente se entrenarán modelos de aprendizaje, empezando por modelos clásicos como clasificadores lineales o bosques aleatorios. Luego se pasarán a modelos neuronales como redes convolucionales y recurrentes y, por último, se entrenarán modelos tipo BERT (RoBERTuito y BETO) en el cual RoBERTuito obtuvo mejores resultados.

En el presente trabajo se ha empleado exclusivamente Python como lenguaje de programación que se ha combinado con el uso de librerías y otras herramientas como

base de datos y entornos de programación. Utiliza como entorno de desarrollo Google Colab, en el se entrenaron los modelos basados en redes neuronales y aprendizaje profundo, la ventaja de esta plataforma es la disponibilidad de GPUs para ejecutar estos modelos en menos tiempo que empleando CPU. Librerías Utilizadas:

- Pandas: es la herramienta para el análisis y manipulación de datos más popular de Python. Se integran con otras librerías como Sciki-learn o Keras.
- Spacy: librería de procesamiento de lenguajes natural basada en el uso de redes neuronales.
- NLTK: es otra librería de procesamiento de lenguajes natural simbólico y estadístico.
- Scikit-learn: es una las principales librerías de Python para aprendizaje automático incluye algoritmos para clasificación, regresión y agrupamiento. También incluye funciones para tratamiento de texto como por ejemplo la vectorización con TF-IDF.
- Keras/Tensorflow: es una librería que ofrece una interfaz para la creación de redes neuronales. Es capaz de trabajar sobre Tensorflows y otros backend como Microsoft Cognitive Toolkit o Theano.
- Huggingface: es una librería que da acceso a través de una interfaz común a una gran cantidad de modelos preentrenados como BERT o RoBERTa.

5.4.1. Conclusión

Acorde al análisis de cada uno de los trabajos relacionados al objeto de estudio, se ha encontrado que mediante la aplicación de Modelos de Lenguaje preentrenados, es posible analizar y clasificar datos textuales como por ejemplo los comentarios realizados en Twitter. Entre los Modelos de Lenguaje mas utilizados se encuentra Bert, sin embargo se evidencia que RoBERTa es un modelo creado para la clasificación de tuits en español por lo cual obtiene mejores resultados en dicho idioma. Entre los lenguajes de programación el que mayor destaca es Python por el gran número de librerías de uso libre que están disponibles, también destacó por la simplicidad en su sintaxis. En cuanto a las técnicas utilizadas para la recopilación de la información, la mayoría de los trabajos utilizan API's de twitter y Twitter Scraper, en las cuales las limitantes de las API's hacen que Twitter Scraper sea la mejor.

El conocimiento de estas técnicas para el PLN, mediante su correcta aplicación permitirán una adecuada realización del presente TT.

6. Metodología

En la figura 6.1, se describe la metodología a abordar en el presente Proyecto de Trabajo de Titulación.

| Objetivo | Tareas/Actividades | Métodos | Materiales | Producto | Lugar | Responsables |
|---|--|---|---|---|------------------------------|------------------------------|
| Construir un conjunto de datos proveniente de la red social Twitter, basado en las menciones a las cuentas oficiales de las Asambleístas ecuatorianas electas para el periodo 2021-2025 | Recolección de los datos. | Web scraping, Técnicas de Minería de Téxto (KDT). | Twitter Scraper, Python, Computadora, OpenRefine. | Dataset inicial. | Universidad Nacional de Loja | Edmundo José Pezantes Urrego |
| | Limpieza del conjunto de datos. | | | Dataset Limpio. | | |
| | Preprocesamiento y transformación del conjunto de datos. | | | Dataset Final. | | |
| Construir la arquitectura del modelo utilizando el modelo de lenguaje RoBERTuito. | Diseñar la arquitectura del modelo. | Técnicas de Minería de Téxto (KDT), Fine Tuning . | Computadora, Python, Google Colab, Huggingface. | Arquitectura del modelo | Universidad Nacional de Loja | Edmundo José Pezantes Urrego |
| | Entrenamiento del modelo de lenguaje creado utilizando RoBERTuito. | | | Modelo para la clasificacion de violencia política de género. | | |
| Evaluar la precisión del modelo de lenguaje construido mediante una muestra de usuarios. | Interpretar los resultados obtenidos de la muestra. | Técnicas de Minería de Téxto (KDT). | Computadora, Python, Google Colab, Huggingface. | Estadística de clasificacion de los tweets. | Universidad Nacional de Loja | Edmundo José Pezantes Urrego |
| | Dar contestación a la pregunta de investigación. | | | Modelo de lenguaje funcional capaz de indetificar tweets con contenido de violencia política de género. | | |
| | Elaborar el documento final. | | | Documento final. | | |

Figura 6.1: Metodología de PTT

7. Cronograma

en la figura 7.1 se presenta la planificación de cronograma a fin de evitar retrasos en la entrega del producto final.

| Objetivo | Tareas/Actividades | Producto | Responsables | Semanas | | | | | | | | | | | | | | | | | | | |
|---|--|---|------------------------------|---------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|
| | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Construir un conjunto de datos proveniente de la red social Twitter basado en las cuentas oficiales de las Asambleístas electas en en Ecuador en el año 2021. | Recolección de los datos. | Dataset inicial. | Edmundo Jose Pezantes Urrego | | | | | | | | | | | | | | | | | | | | |
| | Limpieza de datos | Dataset Limpio. | | | | | | | | | | | | | | | | | | | | | |
| | Preprocesamiento y transformación del conjunto de datos | Dataset Final. | | | | | | | | | | | | | | | | | | | | | |
| Construir el modelo utilizando el modelo de lenguaje preentrenado RoBERTuito | Diseñar la arquitectura del modelo. | Arquitectura del modelo | Edmundo Jose Pezantes Urrego | | | | | | | | | | | | | | | | | | | | |
| | Entrenamiento del modelo de lenguaje creado utilizando RoBERTuito. | Modelo para la clasificacion de violencia política de género. | | | | | | | | | | | | | | | | | | | | | |
| Evaluar la precisión del modelo de lenguaje construido mediante una muestra de usuarios. | Interpretar los resultados obtenidos de la muestra. | Estadística de clasificacion de los tweets. | Edmundo Jose Pezantes Urrego | | | | | | | | | | | | | | | | | | | | |
| | Dar contestación a la pregunta de investigación. | Modelo de lenguaje funcional capaz de indetificar tweets con contenido de violencia política de género. | | | | | | | | | | | | | | | | | | | | | |
| | Elaborar el documento final. | Documento final. | | | | | | | | | | | | | | | | | | | | | |

Figura 7.1: Cronograma de PTT

8. Presupuesto

El desarrollo y cumplimiento del presente TT, implica una inversión económica, necesaria para alcanzar los objetivos propuestos, para esto es preciso utilizar algunos recursos, los cuales se detallan en el siguiente presupuesto establecido por ciertas categorías tales como Recursos Humanos (Tabla 8.1), Recursos Técnicos y Tecnológicos (Tabla 8.2), y por último de Insumos (Tabla 8.3).

8.1. Recursos Humanos

Tabla 8.1: Recurso humano involucrado para el desarrollo del proyecto.

| Talento Humano | Justificación | Número de Horas | Valor por Hora | Subtotal |
|--------------------|--|-----------------|----------------|------------|
| Autor del proyecto | Estudiante a cargo de la ejecución del proyecto | 400 | \$2,50 | \$1,000.00 |
| Personal de Apoyo | Docente director en la elaboración y supervisión del proyecto. | 40 | \$10,48 | \$419,00 |
| | Docente guía de la asignatura PTT de CIS/C-UNL | 80 | \$10,48 | \$838,00 |
| TOTAL | | | | \$2,257.00 |

8.2. Recursos de Software, Hardware, TICs

Tabla 8.2: Recursos Software, Hardware, TICs a utilizar para el desarrollo del proyecto.

| Recursos HW/SW | Subtotal |
|--|------------|
| Computadora | \$1,000.00 |
| Mendeley Desktop | \$0,0 |
| Lenguaje de programación para el PLN | \$0,0 |
| Herramientas para el preprocesamiento, minería de textos y visualización | \$0,0 |
| TOTAL | \$1,000,00 |

8.3. Insumos

Tabla 8.3: Insumos a utilizar para el desarrollo del proyecto.

| Insumos | Subtotal |
|-----------------------|----------|
| Materiales de Oficina | \$50 |
| Transporte | \$60 |
| Internet | \$125 |
| TOTAL | \$235,00 |

8.4. Recursos Totales

En síntesis, para cada uno de los presupuestos categorizados, se presenta el Presupuesto Total del Proyecto (Tabla 8.4), en donde se contempla la sumatoria total de los valores presupuestados, añadiendo un valor del 10% para gastos imprevistos.

Tabla 8.4: Presupuesto total para la elaboración del TT

| Presupuesto Total del Proyecto | |
|---------------------------------------|----------------------|
| Recursos | Subtotal (\$) |
| Recursos Humanos | \$2,257.00 |
| Recursos de Software, Hardware, TICs | \$1,000,00 |
| Insumos | \$235,00 |
| SUBTOTALES (\$) | \$3,492.00 |
| Imprevistos (10%) | 349,2 |
| TOTALES (\$) | 3.841,2 |

El costo total para la ejecución del proyecto será de \$ 3.841,2 cabe recalcar que este proyecto es de naturaleza académica, por lo cual todos los gastos presupuestados correrán por parte del ejecutor del proyecto, a excepción de los gastos que demandan y el docente de PTT, estos los asume la institución para la que trabajan, es decir, la Universidad Nacional de Loja.

Bibliografía

- [1] J. María, L. Salinas, C. Internacional, and D. Fundación, “La web 2.0,” pp. 1–6, 2018.
- [2] J. P. Anturi-Martínez, J. L. Paz-Realpe, J. A. Timaná-Peña, and C. Cobos, “Classifiers for sentiment analysis on twitter: A review,” *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, vol. 2019, no. E23, pp. 353–366, 2019.
- [3] Q. P. HACER, “En twitter?,” 2015.
- [4] P. Fernández, “Violencias de género en twitter: análisis desde el trabajo social,” 2018.
- [5] K. Zeiter, S. Pepera, and M. Middlehurst, “Tweets escalofrantes : Análisis de la violencia en línea contra las mujeres en la política,” 2019.
- [6] Instituto Nacional de Estadística y Censos, “Boletín Encuesta Nacional sobre las Relaciones Familiares Y Género Contra Las Mujeres (ENVIGMU),” *Instituto Nacional de Estadística y Censos (INEC)*, pp. 1–92, 2019.
- [7] G. Santiago and C. Carvajal, “Estudio del rendimiento de BERT frente a métodos clásicos de procesamiento de lenguaje natural,” 2020.
- [8] J. M. Pérez, D. A. Furman, L. A. Alemany, and F. Luque, “RoBERTuito: a pre-trained language model for social media text in Spanish,” 2021.
- [9] Concejo Nacional Electoral, “GESTIÓN PARA EL FORTALECIMIENTO DE LA DEMOCRACIA,” *Paper Knowledge . Toward a Media History of Documents*, vol. 7, no. 2, pp. 107–15, 2019.
- [10] ONU Mujeres – Ecuador, “Violencia política ,” 2019.
- [11] L. F. Badillo, R. Oficial, and D. E. L. A. Democracia, “Ley orgánica electoral, código de la democracia,” 2020.
- [12] Organización de las Naciones Unidas, *La Agenda 2030 y los Objetivos de Desarrollo Sostenible Una oportunidad para América Latina y el Caribe Gracias por su interés en esta publicación de la CEPAL*. 2018.

-
- [13] Universidad Nacional de Loja (UNL), “Investigación | Universidad Nacional de Loja,” 2014.
- [14] Mecanismo de Seguimiento de la Convención de Belém do Pará (MESECVI - OEA), *Ley Modelo Interamericana*. 2017.
- [15] Consejo Nacional Para La Igualdad De Género, “LEY ORGÁNICA INTEGRAL PARA PREVENIR Y ERRADICAR LA VIOLENCIA CONTRA LAS MUJERES,” 2018.
- [16] E. A. Doctor Bracho, “Técnicas estadísticas en minería de textos,” 2018.
- [17] P. C. IVÁN SEBASTIÁN, “EVALUACIÓN DE LAS HERRAMIENTAS DE MINERÍA DE TEXTOS EN LOS MENSAJES DE LA RED SOCIAL TWITTER.,” *UNIVERSIDAD DEL AZUAY FACULTAD DE CIENCIAS DE LA ADMINISTRACIÓN*, vol. 7, pp. 1–25, 2018.
- [18] M. D. C. J. D. L. TORRE, *Nuevas Tecnicas De Mineria De Textos: Aplicaciones*. 2017.
- [19] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, vol. 1, no. Mlm, pp. 4171–4186, 2019.
- [20] Acervo Lima, “Explicación del modelo BERT - PNL ,” 2022.
- [21] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, “RoBERTa: A Robustly Optimized BERT Pretraining Approach,” 2019.
- [22] J. J. López Condori and F. O. Gonzales Saji, “Análisis de sentimiento de comentarios en español en Google Play Store usando BERT,” *Ingeniare. Revista chilena de ingeniería*, vol. 29, no. 3, pp. 557–563, 2021.
- [23] M. Ángel, P. Guisado, and A. Berlanga De Jesús, “Procesamiento del lenguaje natural con BERT: Análisis de sentimientos en tuits,” 2019.
- [24] L. Mathew and V. R. Bindu, “A Review of Natural Language Processing Techniques for Sentiment Analysis using Pre-trained Models,” *Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020*, no. Iccmc, pp. 340–345, 2020.
- [25] A. Troya, R. G. Pillai, C. R. Rivero, Z. Genc, S. Kayal, and D. Araci, “Aspect-Based Sentiment Analysis of Social Media Data with Pre-Trained Language Models,” *ACM International Conference Proceeding Series*, pp. 8–17, 2021.
-

-
- [26] F. M. Plaza-del Arco, M. D. Molina-González, L. A. Ureña-López, and M. T. Martín-Valdivia, “Comparing pre-trained language models for Spanish hate speech detection,” *Expert Systems with Applications*, vol. 166, no. October 2020, p. 114120, 2021.
- [27] Z. Zhao, Z. Zhang, and F. Hopfgartner, “A Comparative Study of Using Pre-Trained Language Models for Toxic Comment Classification,” *The Web Conference 2021 - Companion of the World Wide Web Conference, WWW 2021*, pp. 500–507, 2021.
- [28] M. L. Jamil, S. Pais, J. Cordeiro, and G. Dias, “Detection of extreme sentiments on social networks with BERT,” *Social Network Analysis and Mining*, vol. 12, no. 1, 2022.
- [29] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, *Transformer models for text-based emotion detection: a review of BERT-based approaches*, vol. 54. Springer Netherlands, 2021.
- [30] S. Casola, I. Lauriola, and A. Lavelli, “Pre-trained transformers: an empirical comparison,” *Machine Learning with Applications*, vol. 9, no. May, p. 100334, 2022.
- [31] B. Kasthuriarachchy, M. Chetty, G. Karmakar, and D. Walls, “Pre-trained Language Models with Limited Data for Intent Classification,” *Proceedings of the International Joint Conference on Neural Networks*, 2020.
- [32] V. Isaksen and B. Gambäck, “Using Transfer-based Language Models to Detect Hateful and Offensive Language Online,” pp. 16–27, 2020.
- [33] J. A. Mansilla, “Minado de texto aplicado en Twitter para obtener la polaridad en opiniones de usuarios acerca den nuevo Proyecto de ley que regula la migración en Chile,” 2018.
- [34] R. R. Vega Romero, “Análisis de Sentimientos en Twitter para Descubrir Contenido Xenófobo hacia los Inmigrantes Venezolanos en Ecuador,” p. 181, 2021.
- [35] J. M. Molero Alonso, “Detección de lenguaje ofensivo en redes sociales,” 2022.
-

Lista de Acrónimos y Abreviaturas

| | |
|--------------|--|
| API | Application Programming Interface.. |
| BERT | Modelo basado en redes neuronales para la generación de representaciones vectoriales de textos.. |
| CNE | Concejo Nacional Electoral. |
| IA | Inteligencia Artificial. |
| IDD | Instituto de la Democracia. |
| IND | Instituto Nacional Demócrata. |
| KDT | Descubrimiento de Conocimiento en Textos. |
| ODS | Objetivos de Desarrollo Sostenible. |
| ONU | Organización de las Naciones Unidas. |
| TT | Trabajo de Titulación. |
| Tweet | Publicación en la red social Twitter. Puede tener un máximo de 280 caracteres. |
| UNL | Universidad Nacional de Loja. |

A. Anexo I

A.1. Revisión Sistemática de Literatura

REVISIÓN SISTEMÁTICA DE LITERATURA SOBRE MODELOS DE LENGUAJE, PARA DETERMINAR CONTENIDO DE VIOLENCIA POLITICA DE GÉNERO

Metodología

La finalidad de la presente Revisión Sistemática de Literatura (RSL), es identificar los diferentes Modelos de Lenguaje, herramientas y metodologías existentes para la clasificación de tuits en español. Esta revisión es un estudio secundario de una investigación que permite identificar, analizar e interpretar toda la información disponible en base al tema de estudio, basados en las características principales de una RSL. Según lo sugerido por Kitchenham y Charters [1], la actividad para realizar una RSL implica planificar, ejecutar y reportar la revisión, estos parámetros son adaptados a la presente revisión, este proceso se detalla en los siguientes puntos [2][3].

Proceso de la RSL:

1. Planificación

En esta etapa se identifica cuáles son las necesidades de la presente revisión, para lo cual es necesario plantear y responder algunos puntos claves descritos a continuación.

1.1. Objetivo

El objetivo planteado para la realización de la RSL es:

- Identificar los diferentes modelos de lenguaje existentes para encontrar contenido de violencia política de género en twitter.

1.2. Preguntas de Investigación

El planteamiento de las preguntas de investigación es la sección de mayor importante para la presente revisión, puesto que las preguntas dirigen toda la metodología. Con el fin de identificar los modelos de lenguaje existentes para encontrar contenido de violencia política de género en twitter, en la *Tabla I* se plantean las preguntas de investigación de acuerdo con el objetivo de la Revisión.

Tabla I Preguntas de Investigación.

| ID | Preguntas de Investigación |
|-----|---|
| PI1 | ¿Qué modelos de lenguaje existen para la clasificación de texto en español? |
| PI2 | ¿Qué herramientas existen para la aplicación de minería de textos para analizar tuits en español? |
| PI3 | ¿Qué metodologías existen para la aplicación de minería de texto para analizar tuits español? |

1.3. Estrategias de búsqueda

- a) **Scripts de búsqueda.** La estructura semántica de las búsquedas en esta revisión se realizó siguiendo los principios de búsqueda estructurada, comunes en el lenguaje SQL, usando operadores lógicos, que ofrecen, en mayor o menor medida, todas las bibliotecas virtuales. Aplicando operadores lógicos (AND, OR) en conjunto con las

palabras claves en relación al objeto de estudio, se genera la cadena de búsqueda utilizada en inglés y su traducción en español, las cuales se muestran en la Tabla II.

Tabla II Scripts de Búsqueda.

| ID | Script de Búsqueda |
|-----|---|
| SB1 | <p>("political gender violence " OR "Spanish text classification" OR "twitter") AND ("methodologies" OR "techniques" OR "tools") AND ("fine tuning") AND ("Language Models" OR " pretrained language model")</p> <p>("violencia política de género " OR "clasificación de texto en español" OR "twitter") AND ("metodologías" OR "técnicas" OR "herramientas") AND ("modelos lingüísticos" OR modelo lingüístico preentrenado")</p> |

- b) Bibliotecas Virtuales.** En la Tabla III, se muestra las bibliotecas virtuales utilizadas en esta revisión con su dirección Web correspondiente.

Tabla III Bibliotecas virtuales.

| Bibliotecas Virtuales | URL |
|-----------------------|---|
| IEEE Digital Library | https://ieeexplore.ieee.org/ |
| Science Direct | https://www.sciencedirect.com/ |
| Springer | https://www.springer.com/gp |
| Google Scholar | https://scholar.google.com/ |

- c) Criterios de Selección.** A efectos de investigación, en la presente revisión es necesario definir los criterios de selección de artículos relacionados con los objetivos planteados. A continuación, se detallan los criterios de inclusión y exclusión teniendo en cuenta las preguntas de investigación.

- ✓ **Criterios de Inclusión.** Los criterios de inclusión que se utilizó se describen en la TABLA IV.

Tabla IV Criterios de Inclusión.

| Criterio | Descripción |
|-----------------------|--|
| Contenido | Los estudios deben contener relación con el objeto de investigación al menos en su título, resumen, introducción o desarrollo. |
| Fecha de publicación | Estudios publicados desde el año 2018 hasta el presente año. |
| Bibliotecas Virtuales | IEEE Digital Library, Science Direct, Springer, Google Scholar |
| Idioma | Se consideran artículos escritos solo en Inglés y Español |

- ✓ **Criterios de Exclusión.** No serán considerados los estudios que no cuenten con información relacionada al objeto de estudio, aquellos estudios que no ayuden a responder las preguntas de investigación, estudios duplicados y los que no cumplan con los criterios de inclusión que se muestran en la Tabla IV.

2. Ejecución y Resultados

Una vez aplicadas las cadenas de búsqueda y de haber realizado el análisis correspondiente de todos los documentos encontrados efectuando los criterios de inclusión y exclusión se obtuvo un total de 13 trabajos relacionados, seleccionados de todas las fuentes bibliográficas, en la Tabla V se detalla el número de trabajos relacionados más relevantes, encontrados en cada una de las fuentes bibliográficas entorno a las cadenas de búsqueda dadas anteriormente.

Tabla V Trabajos relacionados seleccionados.

| Código | Título | Ref. | Buscador |
|--------|--|------|----------------|
| TR01 | Análisis de sentimiento de comentarios en español en Google Play Store usando BERT | 4 | Google Scholar |
| TR02 | Estudio del rendimiento de BERT frente a métodos clásicos de procesamiento de lenguaje natural | 5 | Google Scholar |
| TR03 | RoBERTuito: a pre-trained language model for social media text in Spanish | 6 | Google Scholar |
| TR04 | Procesamiento del lenguaje natural con BERT: Análisis de sentimientos en tuits | 7 | Google Scholar |
| TR05 | A Review of Natural Language Processing Techniques for Sentiment Analysis using Pre-trained Models | 8 | IEEE |
| TR06 | Aspect-Based Sentiment Analysis of Social Media Data With Pre-Trained Language Models | 9 | Google Scholar |
| TR7 | Comparing pre-trained language models for Spanish hate speech detection | 10 | Science Direct |
| TR08 | A Comparative Study of Using Pre-trained Language Models for Toxic Comment Classification | 11 | Google Scholar |
| TR09 | Detection of extreme sentiments on social networks with BERT | 12 | Springer |
| TR10 | Transformer models for text-based emotion detection: a review of BERT-based approaches | 13 | Springer |
| TR11 | Pre-trained transformers: an empirical comparison | 14 | Science Direct |
| TR12 | Pre-trained Language Models with Limited Data for Intent Classification | 15 | IEEE |
| TR13 | Using Transfer-based Language Models to Detect Hateful and Offensive Language Online | 16 | Google Scholar |
| TR14 | Minado de texto aplicado en twitter para obtener la polaridad en opiniones de usuarios acerca del nuevo proyecto de ley que regula la migración en Chile | 17 | Google Scholar |
| TR15 | Análisisde Sentimientos en Twitter para Descubrir Contenido Xenófobo hacia los Inmigrantes Venezolanos en Ecuador | 18 | Google Scholar |
| TR16 | Detección de lenguaje ofensivo en redes sociales | 19 | Google Scholar |

2.1. Análisis de los datos

Se ha realizado un análisis individual para cada uno de los trabajos relacionados propuestos, los mismos que aportaron con información relevante entorno a la pregunta de investigación definida en el presente TT, dicho análisis se redacta en la sección de Marco Teórico en el apartado de Trabajos Relacionados.

Referencias

- [1] B. A. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report EBSE-2007-01. School of Computer Science and Mathematics, Keele University," no. January, pp. 1–57, 2007.
- [2] F. J. Rodríguez-Sedano, "Uso de herramienta on-line Parsifal para la elaboración de una revisión sistemática de la literatura (SLR)," Mar. 2019, doi: 10.5281/ZENODO.2603914.
- [3] "Acerca de · Parsifal." <https://parsif.al/about/> (accessed May 24, 2022).
- [4] J. J. López Condori and F. O. Gonzales Saji, "Análisis de sentimiento de comentarios en español en Google Play Store usando BERT," Ingeniare. Revista chilena de ingeniería, vol. 29, no. 3, pp. 557–563, 2021.
- [5] G. Santiago and C. Carvajal, "Estudio del rendimiento de BERT frente a métodos clásicos de procesamiento de lenguaje natural," 2020.
- [6] J. M. Pérez, D. A. Furman, L. A. Alemany, and F. Luque, "RoBERTuito: a pre-trained language model for social media text in Spanish," 2021.
- [7] M. Ángel, P. Guisado, and A. Berlanga De Jesús, "Procesamiento del lenguaje natural con BERT: Análisis de sentimientos en tuits," 2019.
- [8] L. Mathew and V. R. Bindu, "A Review of Natural Language Processing Techniques for Sentiment Analysis using Pre-trained Models," Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020, no. Iccmc, pp. 340–345, 2020.
- [9] A. Troya, R. G. Pillai, C. R. Rivero, Z. Genc, S. Kayal, and D. Araci, "Aspect-Based Sentiment Analysis of Social Media Data with Pre-Trained Language Models," ACM International Conference Proceeding Series, pp. 8–17, 2021.
- [10] F. M. Plaza-del Arco, M. D. Molina-González, L. A. Ureña-López, and M. T. Martín-Valdivia, "Comparing pre-trained language models for Spanish hate speech detection," Expert Systems with Applications, vol. 166, no. October 2020, p. 114120, 2021.
- [11] Z. Zhao, Z. Zhang, and F. Hopfgartner, "A Comparative Study of Using Pre-Trained Language Models for Toxic Comment Classification," The Web Conference 2021 - Companion of the World Wide Web Conference, WWW 2021, pp. 500–507, 2021.
- [12] M. L. Jamil, S. Pais, J. Cordeiro, and G. Dias, "Detection of extreme sentiments on social networks with BERT," Social Network Analysis and Mining, vol. 12, no. 1, 2022.

- [13] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, Transformer models for text-based emotion detection: a review of BERT-based approaches, vol. 54. Springer Netherlands, 2021.
- [14] S. Casola, I. Lauriola, and A. Lavelli, "Pre-trained transformers: an empirical comparison," Machine Learning with Applications, vol. 9, no. May, p. 100334, 2022.
- [15] B. Kasthuriarachchy, M. Chetty, G. Karmakar, and D. Walls, "Pre-trained Language Models with Limited Data for Intent Classification," Proceedings of the International Joint Conference on Neural Networks, 2020.
- [16] V. Isaksen and B. Gambäck, "Using Transfer-based Language Models to Detect Hateful and Offensive Language Online," pp. 16–27, 2020.
- [17] J. A. Mansilla, "Minado de texto aplicado en Twitter para obtener la polaridad en opiniones de usuarios acerca den nuevo Proyecto de ley que regula la migración en Chile," 2018.
- [18] R. R. Vega Romero, "Análisis de Sentimientos en Twitter para Descubrir Contenido Xenófobo hacia los Inmigrantes Venezolanos en Ecuador," p. 181, 2021.
- [19] J. M. Molero Alonso, "Detección de lenguaje ofensivo en redes sociales," 2022.

B. Anexo II

B.1. Entrevista 1



Universidad
Nacional
de Loja



UNIVERSIDAD NACIONAL DE LOJA

**“Facultad de la Energía, las Industrias y los Recursos Naturales no Renovables”
Carrera de Ingeniería en Sistemas.**

La presente entrevista se realiza con fines académicos para el Análisis de Sentimientos en Twitter y encontrar contenido de violencia política de género en Ecuador. Los datos proporcionados por los entrevistados se manejarán con estricto fin académico y no representan el pensamiento de la Universidad Nacional de Loja y el Concejo Nacional Electoral.

Los entrevistados aceptan que sus nombres y sus cargos se registren en esta entrevista.

| | |
|------------------------|--------------------------------------|
| Entrevistador | Edmundo José Pezantes Urrego |
| Entrevistado | Abg. Luis Hernán Cisneros |
| Cargo del entrevistado | Director CNE Loja |
| Fecha de entrevista | 19/05/2022 |
| Hora de entrevista | 14:30 |
| Canal de Comunicación | Plataforma de video conferencia ZOOM |

1. ¿Considera que la violencia política de género influye en el número de autoridades mujeres electas en la actualidad y por qué?

Claro que sí, influye en la elección de todas las autoridades y también de manera específica en la toma de decisiones de las autoridades y de las mujeres que se presentan a estas dignidades, ese es el fin de la violencia política y especialmente de la violencia política de género, entonces yo considero que evidentemente sí influye y es por eso que la norma establece límites y restricciones acerca de ese tipo de actos.

2. ¿Qué opinión tiene usted acerca de la violencia política de género en Ecuador?

Pienso que el termino en si violencia de cualquier tipo no puede ser aceptado en una sociedad, en general el termino violencia ya implica un acto fuera de los límites de una sociedad civilizada así que dentro de este concepto todo lo que es violencia no debe ser aceptada.



3. ¿Considera que en el Ecuador las redes sociales influyen en la propagación de contenido de violencia política de género y por qué?

Sin duda que influyen, al estar en un mundo donde la tecnología tiene una fundamental importancia en la toma de decisiones, por ejemplo, atreves de la psicometría y el análisis de sentimientos se puede llegar a conclusiones mediante documentos académicos investigativos científicos, como a través de las redes sociales y de las fake news puede incidir esto en la decisión de los ciudadanos, en todos los aspectos de la vida y más aún en los político que es cuando nosotros entramos en materia, yo creo que efectivamente existe una incidencia en este sentido, pero que al final del día si bien es cierto es la norma la regula, sanciona y la que establece ciertos límites es también la ciudadanía que tiene que poder recibir este tipo de información, analizarla, evaluarla y considerarla y consumirla dentro de los parámetros que nos impone la ética y la moral porque no podríamos como sociedad tomar decisiones sobre la base de comentarios, chismes, supuestos o noticias falsas que es la característica que fundamentalmente se presenta en actos o expresiones dentro de las redes sociales que tratan de menoscabar o menospreciar y minimizar la participación de las mujeres.

4. ¿Conoce de algún medio que permita analizar los tweets de contenido de violencia de género?

La tecnología actualmente, permite a través de la inteligencia artificial y de recursos tecnológicos como metodológicos hacer un análisis de sentimientos del comportamiento de los ciudadanos en redes sociales, entre esas esta justamente twitter y claro que es un medio o metodologías que permiten establecer todo tipo de comportamiento sea de aceptación, rechazo o violencia en las redes sociales.


5. ¿Le aportaría de alguna manera el realizar un estudio en la red social Twitter para tener información acerca de la existencia de mensajes que reflejen la violencia política de género?

Claro que aportaría como estudio el poder aplicar esta tecnología y esta metodología que nos facilita o que nos presenta la inteligencia artificial que a través de un análisis de sentimientos determinar la existencia de conductas violencia política de género y más aun siendo temas académicos científicos en violencia política de género, sería un estudio que aportaría de mucho para temas de conocimiento de respuestas, de reacciones y de soluciones en todo tipo de materias en este sentido.



6. ¿Le gustaría que existiera algún medio para analizar automáticamente los tweets de violencia política de género y por qué?

Por supuesto que sí, es muy factible puesto que no representaría una actividad imposible de realizar, dado que a la inteligencia artificial ya se la ocupa con la finalidad de plantear estrategias de campaña a través del análisis de sentimientos con el fin de poder concretar un mensaje o discurso, del mismo modo se la debe utilizar para poder identificar actos de violencia política de género. Sería un aporte fundamental puesto que nos permitiría conocer sobre la base de números reales la incidencia no solo en la decisión sino también en qué medida se está incurriendo en esta práctica.

| Nombre | Cargo | Firma |
|---------------------------|-------------------|--|
| Abg. Luis Hernán Cisneros | Director CNE Loja |  <p>Escaneo electrónicamente por: LUIS HERNAN CISNEROS JARAMILLO</p> |

C. Anexo III

C.1. Entrevista 2



Universidad
Nacional
de Loja

UNIVERSIDAD NACIONAL DE LOJA

"Facultad de la Energía, las Industrias y los Recursos Naturales no Renovables"
Carrera de Ingeniería en Sistemas.

La presente entrevista se realiza con fines académicos para el Análisis de Sentimientos en Twitter y encontrar contenido de violencia política de género en Ecuador. Los datos proporcionados por los entrevistados se manejarán con estricto fin académico y no representan el pensamiento de la Universidad Nacional de Loja y el Concejo Nacional Electoral.

Los entrevistados aceptan que sus nombres y sus cargos se registren en esta entrevista.

| | |
|--------------------------|---|
| Entrevistador | Edmundo José Pezantes Urrego |
| Entrevistada | Ing. Neli Virmania Troya Troya |
| Cargo de la entrevistada | Concejal del Gobierno Autónomo Descentralizado del Cantón Calvas. |
| Fecha de entrevista | 23/05/2022 |
| Hora de entrevista | 17:00 |
| Canal de Comunicación | Plataforma de video conferencia ZOOM |

1. **¿Considera que la violencia política de género influye en el número de autoridades mujeres electas en la actualidad y por qué?**

Considero que si influye, porque al tratarse de violencia política de genero contra la mujer se está afectando o tratando de limitar la participación política de las mujeres, puesto que ese es el fin de la violencia política de genero contra la mujer.

2. **¿Qué opinión tiene usted acerca de la violencia política de género en Ecuador?**

Si bien es cierto que en la actualidad se tiene una mayor participación de las mujeres en el ámbito político, aún existen casos de violencia política contra las mujeres, por lo cual se debe seguir trabajando para erradicar por completo este tipo de violencia que atenta contra la integridad y capacidad de las mujeres en cuanto a desenvolverse políticamente se trata.



3. ¿Considera que en el Ecuador las redes sociales influyen en la propagación de contenido de violencia política de género y por qué?

Considero que si es influyente, puesto que las redes sociales son un medio de potencial propagación de contenido sea positivo o negativo y las personas podrían dejarse influenciar por dicho contenido, por ende si existe algún mensaje o publicación que sea de contenido de violencia política de género se expandirá rápidamente hacia todas las personas que hagan uso de las redes sociales.

4. ¿Conoce de algún medio que permita analizar los tweets de contenido de violencia de género?


Tengo conocimiento o he sabido acerca del análisis de sentimientos en twitter, pero acerca de un medio que permita determinar si existe violencia política de género como tal no.

5. ¿Le aportaría de alguna manera el realizar un estudio en la red social Twitter para tener información acerca de la existencia de mensajes que reflejen la violencia política de género?

Si, sería muy interesante puesto que se lo puede tomar como un punto de partida para tomar medidas o acciones para tratar de erradicar el contenido de violencia política de género en las redes sociales, en twitter en este caso.

6. ¿Le gustaría que existiera algún medio para analizar automáticamente los tweets de violencia política de género y por qué?

Si, todas la herramientas o medios que sirvan para identificar este tipo de contenido de violencia política de género es necesario, de modo que se pueda tratar de controlar o erradicar la violencia política de género, por lo cual sería de mucha utilidad.

| Nombre | Cargo | Firma |
|--------------------------------|---|--|
| Ing. Neli Virmania Troya Troya | Concejal del Gobierno Autónomo Descentralizado del Cantón Calvas. |  <p>Firmado digitalmente por: NELI VIRMANIA</p> |

D. Anexo IV

D.1. FORMATO PARA EL REGISTRO DE LAS ACTIVIDADES DE TUTORÍA PARA LA ELABORACIÓN DEL PROYECTO DE TRABAJO DE TITULACIÓN

D.1. FORMATO PARA EL REGISTRO DE LAS ACTIVIDADES DE TUTORÍA
PARA LA ELABORACIÓN DEL PROYECTO DE TRABAJO DE TITULACIÓN

43



UNL

Universidad
Nacional
de Loja




UNIVERSIDAD NACIONAL DE LOJA

COMISIÓN DE ARTICULACIÓN DE LAS FUNCIONES SUSTANTIVAS

FORMATO PARA EL REGISTRO DE LAS ACTIVIDADES DE TUTORÍA PARA LA ELABORACIÓN DEL TRABAJO DE INTEGRACIÓN CURRICULAR/TITULACIÓN


| | |
|--|---|
| Facultad: | Facultad de la Energía las Industrias y los Recursos no Renovables |
| Carrera/Programa: | Ingeniería en Sistemas. |
| Nombre y apellidos del docente asesor/director: | Oscar Miguel Cumbicus Pineda, Mg.Sc |
| Tema del trabajo de integración curricular/titulación: | "Modelo de lenguaje utilizando RoBERTuito, para identificar tweets con contenido de violencia política de género hacia las asambleístas electas en Ecuador para el periodo 2021-2025" |
| Nombres y apellidos del/los aspirante/s: | Edmundo José Pezantes Urrego |
| Período académico ordinario: | Abril - Septiembre 2022 |

| Nro. | Fecha | Tiempo empleado en la tutoría | Tema tratado en la tutoría | Recomendaciones del Asesor/Director | Modalidad | | Firma del estudiante/ Registro Virtual |
|------|------------|-------------------------------|---|---|------------|---------|--|
| | | | | | Presencial | Virtual | |
| 1.00 | 28/04/2022 | 2:00:00 | Idea del Proyecto de Trabajo de Titulación. | *Leer documentación acerca de la idea propuesta y de los Modelos de Lenguaje para entender y tener una mayor idea acerca de lo que se va a realizar. * Analizar a quienes les servirá la realización del Trabajo de Titulación. | x | | |
| 2.00 | 12/05/2022 | 2:00:00 | Instrumento de entrevista | * Correcciones a las preguntas planteadas inicialmente, centrarlas en el porque de sus respuestas y en que les beneficia la realización de la idea del presente TT, fijando así las preguntas finales a realizar en la entrevista. * Se define las personas a entrevistar. | x | | |

| | | | | | | | |
|------|------------|---------|---|--|---|--|---|
| 3.00 | 23/06/2022 | 2:00:00 | Revisión de entrevista realizada, Problemática y Justificación planteada. | <ul style="list-style-type: none"> * Se sugiere centrar mas el tema a un determinado periodo y candidatas electas. * Correcciones en fuentes a citar. * Agregar un parrafo en la problemática acerca de los Modelos de Lenguaje con mejores resultados. | x | | |
| 4.00 | 12/07/2022 | 2:00:00 | Revisión de Problemática, Justificación, Objetivos y Alcance. | <ul style="list-style-type: none"> * Corrección en el Objetivo General, especificar el modelo que se va a utilizar y que no difiera del tema. * Corrección en objetivo específico, especificar la proveniencia de los datos a extraer para la construcción del conjunto de datos. * Realizar actualización de en el alcance de acuerdo a las correcciones realizadas en los objetivos. | x | |  |
| 5.00 | 18/07/2022 | 2:00:00 | Revisión de Marco Teórico y Metodología. | <ul style="list-style-type: none"> * En el Marco Teórico agregar contenido acerca de Procesamiento de Lenguaje Natural, una sección acerca de los Modelos de Lenguaje más relevantes y aparte de trabajos relacionados realizados en base a la revisión bibliográfica incluir una conclusión. * En la Metodología se recomienda incluir las técnicas a utilizar para la recolección de los datos y la construcción del modelo. | x | |  |
| 5.00 | 21/07/2022 | 2:00:00 | Revisión Metodología, Cronograma y Presupuesto. | <ul style="list-style-type: none"> * Ajustar los Plazos de las Actividades en el Cronograma. * En el presupuesto se recomienda tener en cuenta solo lo que se va a utilizar. | x | |  |

D.1. FORMATO PARA EL REGISTRO DE LAS ACTIVIDADES DE TUTORÍA
PARA LA ELABORACIÓN DEL PROYECTO DE TRABAJO DE TITULACIÓN

45

| | | | | | | | |
|------|------------|---------|------------------------------|---|---|--|---|
| 6.00 | 27/07/2022 | 2:00:00 | Revision del documento final | * Se recomienda incluir entre comillas y con cursiva las citas textuales de las entrevistas realizadas. *Nombrar los autores cuando se cita algún trabajo. | x | |  |
|------|------------|---------|------------------------------|---|---|--|---|

Fecha de presentación: 22/08/2022

Firma del docente

Firma del Director/a y/o Encargado de la Gestión Académica de la Carrera