



UNIVERSIDAD SANTO TOMÁS  
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA  
T U N J A

### **TÍTULO PROYECTO**

Aplicación de machine learning para la enseñanza – aprendizaje de competencias ciudadanas en educación media del Colegio de Boyacá.

### **PROPONENTE(S)**

Harvey Nicolas Echavarria Ortiz  
1049654983  
2296120

### **DIRECTOR**

Henry Alfonso Guío Ávila

Tunja  
Día 08 08 2022

## CONTENIDO

|                                                           |                  |
|-----------------------------------------------------------|------------------|
| <b><u>1. FICHA TÉCNICA DEL PROYECTO.....</u></b>          | <b><u>3</u></b>  |
| <b><u>2. PLANTEAMIENTO DEL PROBLEMA .....</u></b>         | <b><u>4</u></b>  |
| <b><u>3. JUSTIFICACIÓN.....</u></b>                       | <b><u>7</u></b>  |
| <b><u>4. OBJETIVOS.....</u></b>                           | <b><u>9</u></b>  |
| <b><u>5. ESTADO DEL ARTE .....</u></b>                    | <b><u>10</u></b> |
| <b><u>6. METODOLOGÍA: MODELO DE DESARROLLO .....</u></b>  | <b><u>29</u></b> |
| <b><u>7. DESARROLLO DEL PROYECTO .....</u></b>            | <b><u>32</u></b> |
| 7.1. FASE 1. DEFINICIÓN DE REQUISITOS.....                | 32               |
| 7.1.1. DESCRIPCIÓN GENERAL .....                          | 32               |
| 7.1.2. FUNCIONALIDAD DEL PRODUCTO .....                   | 33               |
| 7.1.3. PRODUCT BACKLOG E HISTORIAS DE USUARIO .....       | 33               |
| <b>7.2. FASE 2. SELECCIÓN DE TECNOLOGÍA .....</b>         | <b>37</b>        |
| 7.2.1. HERRAMIENTAS UTILIZADAS .....                      | 37               |
| <b>7.3. FASE 3. DISEÑO DEL SISTEMA .....</b>              | <b>39</b>        |
| <b>7.4. FASE 4. DESARROLLO DE LA APLICACIÓN .....</b>     | <b>40</b>        |
| 7.4.1. IA SIN CONTEXTO CONVERSACIONAL .....               | 40               |
| 7.4.2. IA CON CONTEXTO CONVERSACIONAL .....               | 49               |
| <b>7.5. FASE 5. PRUEBAS.....</b>                          | <b>52</b>        |
| <b><u>8. ANÁLISIS Y DISCUSIÓN DE RESULTADOS .....</u></b> | <b><u>59</u></b> |
| <b><u>9. CONCLUSIONES .....</u></b>                       | <b><u>60</u></b> |
| <b><u>10. REFERENCIAS.....</u></b>                        | <b><u>61</u></b> |
| <b><u>11. ANEXOS.....</u></b>                             | <b><u>63</u></b> |

## 1. FICHA TÉCNICA DEL PROYECTO

|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Título                | Aplicación de machine learning para la enseñanza – aprendizaje de competencias ciudadanas en educación media del Colegio de Boyacá.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Autor (es)            | Harvey Nicolas Echavarria Ortiz                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Director del proyecto | Henry Alfonso Guío Ávila                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Palabras claves       | PLN, Machine Learning, ICFES, Inteligencia Artificial.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Descripción           | <p>En los últimos años, se ha visto la dificultad por parte de los estudiantes de educación media del Colegio de Boyacá, en la comprensión de saberes del módulo de Competencias Ciudadanas del examen ICFES, aspecto evidenciado por el bajo puntaje obtenido en dichas pruebas. De otro lado, se ha demostrado que el uso de las Tecnologías de la Información y las Comunicaciones TIC para apoyar procesos de enseñanza, ha sido altamente efectivo, sobre todo por los nuevos paradigmas en cuanto al uso de medios digitales. Así pues, una alternativa es el uso de tecnologías de Inteligencia Artificial para facilitar estos procesos. Una de estas tecnologías es el Machine Learning (En español, máquinas de aprendizaje), en la cual los computadores son capaces de desarrollar procesos de aprendizaje. Este proyecto tiene como objetivo, implementar un sistema Machine Learning de Procesamiento de Lenguaje Natural (PLN<sup>1</sup>) para que, mediante el reconocimiento de preguntas por parte de los usuarios y la capacidad de dar respuestas de forma autónoma, apoye la enseñanza del módulo de competencias ciudadanas del examen ICFES. Este trabajo está dirigido a los estudiantes de educación media del Colegio de Boyacá. Para el cumplimiento de este objetivo, se plantean seis fases en las que se realizan los procesos de documentación, diseño, desarrollo, entrenamiento, testing e implementación del sistema.</p> |

---

<sup>1</sup>PNL, de las iniciales en español de Procesamiento de Lenguaje Natural.

## 2. PLANTEAMIENTO DEL PROBLEMA

La educación puede considerarse como un arte evolutivo que permite la aprehensión del conocimiento por parte de los estudiantes a medida que su proceso de formación avanza, siendo este más eficiente cuando las instituciones educativas cuentan con diferentes recursos que fortalecen su desarrollo, entre ellos están los recursos tecnológicos que son necesarios para innovar las dinámicas de enseñanza – aprendizaje dentro y fuera de las aulas. En Colombia, dicho arte aún hace uso en su mayoría de las estrategias tradicionales al momento de impartir el conocimiento, por lo cual, la implementación de las TIC se ve reducida y la innovación tecnológica en la enseñanza no encuentra una manera de verse reflejada, es así que algunas causas que se pueden tener en cuenta para que esta situación se mantenga vigente son el uso escaso de herramientas TIC, el uso predominante de estrategias de enseñanzas tradicionales y el acceso limitado a herramientas TIC.

Es por ello, que Colombia implementa las pruebas saber 11 o pruebas de estado que son “la evaluación del nivel de Educación Media que se alinea con las evaluaciones de las Educación Básica con el objetivo de proporcionar información a la comunidad educativa en el desarrollo de las competencias básicas que un estudiante debería desarrollar durante el paso de su vida escolar” (ICFES: ¿Qué son las pruebas saber?2018-12-21 11:00:00 -0500), en otras palabras, es una manera de medir la calidad educativa de las instituciones del país teniendo en cuenta un número específico de competencias, sin embargo se encuentra que existe una brecha entre los resultados obtenidos por los estudiantes de los colegios oficiales y privados, esto se ve reflejado en la publicación del Ranking Col-Sapiens (Research, 2019) que realiza Sapiens Research en el cual, se da a conocer un listado de las mejores instituciones educativas de acuerdo a los resultados del ICFES (saber 11) y en el plan de desarrollo nacional para el período de 2018-2022 llamado “pacto por Colombia, pacto por la equidad” en la sección “Educación con calidad para un futuro con oportunidades para todos”, en la que se da a conocer una diferencia del 34.5% entre colegios oficiales y privados en los resultados de las pruebas saber 11. (Guío Ávila Henry Alfonso, 2021)

Tunja la capital del departamento de Boyacá, no es ajena a dicha situación puesto que en el ranking del año 2019 de 42 colegios sólo el 33% logró clasificar al listado de las mejores instituciones educativas (Research, 2019), al analizar detenidamente se encuentra que en el top 10 del listado, el Colegio de Boyacá es la única institución de carácter oficial ocupando el puesto número 9 (el listado se puede consultar en el enlace

<https://www.srg.com.co/losmejorescolegios-colsapiens/>), siendo el resto ocupado por instituciones de carácter privado; con ello se evidencia que las instituciones educativas de carácter privado tienen una tendencia a obtener mejores resultados en las pruebas saber 11 y por ende, ofrecen mejor calidad en los procesos educativos respecto a las instituciones educativas de carácter oficial.

Es por ello que se evidencia la necesidad de implementar una estrategia que permita a los estudiantes apropiarse de los conocimientos necesarios para que ellos estén en la capacidad de responder con mayor asertividad las pruebas de estado y así obtener mejores puntajes, ya que al hacer uso de las nuevas tecnologías la institución educativa provee nuevas herramientas que incrementan la calidad en el proceso de enseñanza - aprendizaje. Para dicho fin se requiere realizar un sistema computacional basado en machine learning, utilizando la técnica del procesamiento del lenguaje natural orientado a estudiantes de grado décimo y once del Colegio de Boyacá de la ciudad de Tunja.

Así pues, el sistema computacional se construirá con la técnica de machine learning, procesamiento del lenguaje natural y estará enfocada en el área de las competencias ciudadanas dado que como componente es esencial para la convivencia armónica de la persona en sociedad, para dar respuesta a preguntas relacionadas a cuatro temas: Principios fundamentales de la Constitución Política Colombiana de 1991, organización del estado colombiano, los modelos de participación ciudadana y los derechos humanos. Con esto, se pretende que los estudiantes puedan tener una ayuda innovadora y tecnológica a la hora de prepararse para presentar las pruebas de estado (pruebas saber ICFES) y así tener una mayor apropiación de los temas relacionados con las competencias ciudadanas.

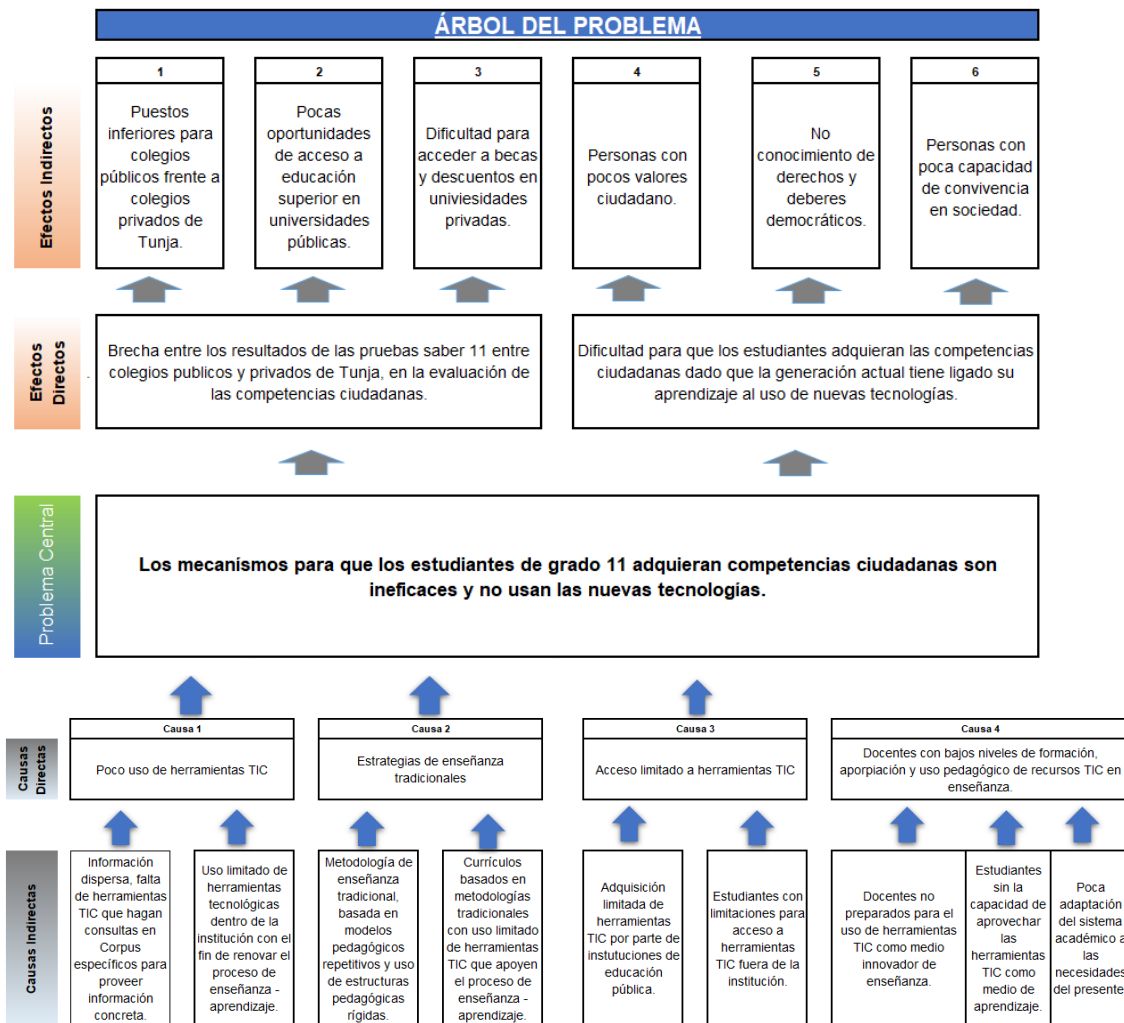


Figura 1. Árbol del problema, elaboración propia.

### 3. JUSTIFICACIÓN

Se ve la necesidad de desarrollar e implementar el proyecto propuesto, ya que con él se provee un recurso tecnológico con el que los estudiantes del colegio de Boyacá de la ciudad de Tunja puedan prepararse para presentar las pruebas ICFES o pruebas de estado en el área de las competencias ciudadanas, de manera que las instituciones educativas de carácter oficial también puedan llegar a hacer uso de las TIC como medio de refuerzo para las metodologías de enseñanzas tradiciones que aún persisten.

Con lo anterior, se busca que los estudiantes tengan un medio de consulta virtual relacionado con cuatro aspectos en el área de las competencias ciudadanas: Principios fundamentales de la constitución política colombiana de 1991, organización del estado colombiano, los modelos de participación ciudadana y los derechos humanos, reconociendo a la persona como ser social por naturaleza y por ende, con la necesidad de la comprensión de las responsabilidades (derechos y deberes) civiles que se deben cumplir. Así mismo, el uso de las TIC y herramientas revolucionarias como el Machine Learning y el PLN se ven necesarias en un campo en el que las dinámicas de enseñanza – aprendizaje cambian de manera constante (Rodríguez Mena García & García Montero, 2003). En un principio era necesaria la presencialidad, un tablero y una tiza, luego las tizas fueron reemplazadas por marcadores para llegar a la modernidad con clases virtualizadas y pantallas digitales.

Es por ello que la tecnología en el ámbito educativo “tiene la capacidad potencial para hacer frente a algunos de los desafíos mayores de la educación de hoy en día, de elaborar prácticas de enseñanza y aprendizaje innovadoras” (La inteligencia artificial en la educación.2019). Así mismo, es importante el desarrollo del proyecto puesto que abre las puertas a nuevas posibilidades que permiten reinventar el entorno educativo, ya que al realizar una búsqueda se tiene como resultado la inexistencia de desarrollos que relacionen el área de las competencias ciudadanas con el machine learning y el PLN en Colombia, dando la posibilidad a padres de familia, estudiantes y docentes de promover el uso de herramientas tecnológicas con el fin de renovar las propuestas educativas que ya se tienen planteadas, para que a futuro las instituciones educativas de carácter oficial también puedan convertir el proceso de enseñanza – aprendizaje en una estrategia pedagógica acorde a las necesidades actuales en la que las TIC<sup>2</sup> son un requisito indispensable para

---

<sup>2</sup> TIC, de las iniciales en español Las Tecnologías de la Información y las Comunicaciones.

desenvolverse día a día, ya sea en el ámbito del aprendizaje o en cualquier otro que involucre una interacción social.

Este proyecto se fundamenta en la necesidad creciente en el ámbito educativo de proponer nuevas estrategias tecnológicas que ayuden al proceso de enseñanza – aprendizaje a convertirse en una estrategia pedagógica dinámica, que pueda ser adaptada a los requisitos que la población educativa requiera, teniendo en cuenta dos aspectos: la persona como ser social y pilar fundamental de la sociedad que necesita la educación como un proceso de formación constante, y la institución como un ente que provee las herramientas necesarias para que el individuo pueda formarse.

Finalmente se entiende el proyecto como una solución tecnológica que al ser propuesta desde la perspectiva de una ingeniería busca crear una comunicación entre el ámbito educativo, la informática y las competencias ciudadanas, de manera que los estudiantes puedan prepararse para presentar las pruebas saber y así reducir de manera eficiente la brecha existente entre las instituciones de carácter oficial y privado, con herramientas tecnológicas innovadores que adapten el proceso de enseñanza – aprendizaje a los requerimientos de la sociedad moderna.

## 4. OBJETIVOS

### 4.1. Objetivo General

Desarrollar un sistema computacional basado en machine learning, utilizando la técnica del procesamiento del lenguaje natural (PLN), que permita responder las preguntas generadas por los estudiantes de grado 11 del Colegio de Boyacá en la ciudad de Tunja, con el fin de que ellos se apropien de los conocimientos asociados al módulo de competencias ciudadanas del examen ICFES.

### 4.2. Objetivos específicos

A continuación, se presentan los objetivos específicos.

Tabla 1. Objetivos específicos

| Nro. | Objetivo específico                                                                                                                                                                                                                                                                                                         |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1    | Generar el corpus con el cual la red neuronal PLN soporte la respuesta dada al estudiante, mediante la consulta y selección de documentación relevante sobre competencias ciudadanas en Colombia y reuniones con expertos en el área, para poder implementar el sistema de acuerdo a los objetivos de aprendizaje deseados. |
| 2    | Implementar el modelo pre-entrenado con base en el corpus generado, por medio de sentencias de código Python para el análisis semántico del proceso de pregunta-respuesta entre el sistema de PLN y el estudiante.                                                                                                          |
| 3    | Codificar el backend de la red PLN de acuerdo a la implementación del modelo pre-entrenado, por medio de la herramienta web Google Colaboratory y el lenguaje de programación Python para establecer los procedimientos y rutinas de configuración.                                                                         |
| 4    | Validar el desarrollo del sistema PLN propuesto por medio de una serie de pruebas con el fin de comprobar su correcto funcionamiento.                                                                                                                                                                                       |

Fuente: Autor

## 5. ESTADO DEL ARTE

La tecnología como herramienta ha sufrido diversos cambios a lo largo del tiempo, por ejemplo, las primeras computadoras o medios de almacenamiento eran tan sólo los cimientos que llevarían al hombre a crear las innovaciones futuristas del siglo XXI; uno de estos grandes desarrollos es la inteligencia artificial (IA) que da inicio con autores como el matemático, lógico e informático Alan Turing que propone mediante un test una forma de comprobar que una máquina puede llegar a ser “Inteligente”. Es así que la IA<sup>3</sup> se puede definir como el “modo de simular las capacidades de inteligencia del cerebro humano” (Ocaña-Fernández, Valenzuela-Fernández, & Garro-Aburto, 2019) asumiendo “que la IA es parte de las Ciencias de la Computación que se ocupa del diseño de sistemas inteligentes, esto es sistemas que exhiben características que asociamos con la inteligencia en las conductas humanas” (Ocaña-Fernández et al., 2019).

La inteligencia artificial es la tecnología creada por el hombre que permite replicar ya sean los comportamientos racionales o algunas funciones del cerebro humano; sin embargo, si bien es cierto que se han logrado avances importantes como la creación de asistentes virtuales por voz entre ellos Alexa (Amazon), Siri (Apple) y Google Assistant (Google), androides capaces de conversar o simuladores que replican los sentidos del cuerpo comandados por el cerebro como la visión, aún no se ha logrado implantar la esencia humana en una máquina, y por ello es que ésta se debe considerar más que un reemplazo de su creador; una herramienta que potencializa las habilidades humanas.

Así mismo, la IA hace uso de múltiples herramientas que permiten su aplicación en distintos campos del conocimiento, como por ejemplo las ingenierías, la medicina o las humanidades, y en este caso la educación en la que las aplicaciones “varían, básicamente, en función de los objetivos que se persiguen con el diseño del programa así como del tipo de conocimiento que la aplicación debe manipular” (Begoña, 1992) , en otras palabras, el uso de las herramientas, programas y equipos de hardware que se utilicen para el desarrollo de un producto con base en IA enfocada al área académica, va a variar de acuerdo a los contextos, a las necesidades que se presentan y al personal a quien vaya dirigido, ya sea a estudiantes, docentes, directores o planta administrativa.

---

<sup>3</sup> IA, de las iniciales en español Inteligencia Artificial.

Entre las herramientas que pueden utilizarse para el modelamiento de una IA se encuentra el machine learning y el procesamiento del lenguaje natural (PLN). El primero se define como “campo de estudio que dota a los ordenadores de la capacidad de aprender sin ser explícitamente programados” (Management Solutions, 2018) que en otras palabras es el área de la IA encargada de proponer patrones que permiten a la máquina reconocer relaciones y así “entender” un nuevo concepto de manera autónoma, teniendo en cuenta que la precisión de este proceso dependerán de algunos factores externos como las fuentes de información, el propio algoritmo y el equipo tanto de software y de hardware que se tenga (Management Solutions, 2018). Dicho lo anterior, es necesario tener en cuenta que existen distintos métodos de aprendizaje: el supervisado que tiene en cuenta los modelos predictivos, el no supervisado que tiene en cuenta un modelo descriptivo y el reforzado que es un aprendizaje basado en hallazgos (Valdez Alvarado, 2019), cada uno enfocado a aspectos claves que permiten un desarrollo y crecimiento adecuado de la IA.

Para definir el PLN se debe partir del concepto de lenguaje desde dos puntos de vista: desde una perspectiva funcional lingüística “como una función que expresa pensamientos y comunicaciones entre la gente” (Augusto Cortez Vásquez, Hugo Vega Huerta, & Jaime Pariona Quispe, 2013) ya sea mediante medios escritos, visuales o por voz y desde una perspectiva formal “como un conjunto de frases, que generalmente es infinito y se forma con combinaciones de elementos tomados de un conjunto (usualmente infinito) llamado alfabeto, respetando un conjunto de reglas de formación (sintácticas o gramaticales) y de sentido (semánticas)” (Augusto Cortez Vásquez et al., 2013), teniendo en cuenta que los lenguajes pueden ser naturales o formales.

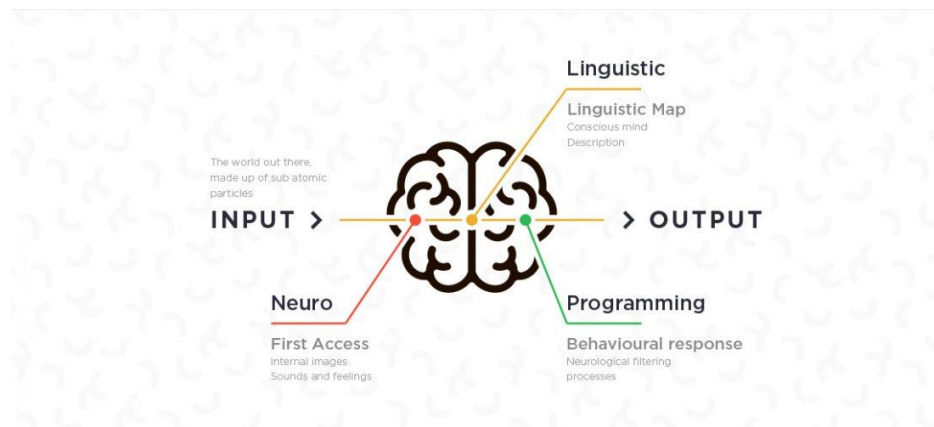


Figura 2. Procesamiento del Lenguaje Natural, tomada de <https://cdn.datacentric.es/wp-content/uploads/2018/12/31125832/procesamiento-del-lenguaje-natural.jpg>

Ahora bien, el lenguaje natural “es aquel que ha evolucionado con el tiempo para fines de comunicación humana, como el español o alemán” (Augusto Cortez Vásquez et al., 2013), por lo cual se puede comprender como la evolución de la comunicación utilizada para el entendimiento de un mensaje entre dos o más personas siendo así el procesamiento del lenguaje natural el análisis del medio utilizado para la comunicación entre personas. En otras palabras, es hacer que la máquina entienda el mensaje que recibe por parte de una persona para así generar una respuesta lógica y coherente. Para que esto sea posible se debe tener en cuenta la arquitectura del PLN está dividida en cinco niveles del lenguaje natural, estos son el nivel fonológico, el nivel morfológico, el nivel sintáctico, el nivel semántico y el nivel pragmático, y el lenguaje que se utiliza para programar la inteligencia artificial, dando como resultado la creación de algunas aplicaciones que comúnmente se utilizan, como por ejemplo la traducción automática, tutorías inteligentes o el reconocimiento de voz. En conclusión, el PLN se define como “un conjunto de instrucciones que un sistema recibe en un lenguaje de programación dado, que permita comunicarse con un humano en su propio lenguaje”. (Cañón, 2007)

Para realizar las configuraciones necesarias en la construcción de desarrollos basados en machine learning que hacen uso del PLN, es necesario hacer uso de sentencias (líneas de código) que representan las órdenes planteadas en lenguaje natural, dichas ordenes están agrupadas en un lenguaje de programación que proporcionan los medios para que la maquina entienda las instrucciones dadas. En este ámbito Python se ha convertido por excelencia en una de las mejores alternativas para construir las órdenes del sistema, brindando dentro de sus ventajas una gran comunidad y facilidad de adaptabilidad por parte de las personas, que utilizan esta herramienta para el desarrollo de sistemas informáticos basados en machine learning.

Python “es un lenguaje de programación interpretado y multiparadigma (soporta orientación a objetos, programación imperativa y programación funcional), cuya filosofía hace hincapié en una sintaxis que favorezca un código legible. Además, es multiplataforma” (Gago Utrera, 2017). Dicho lo anterior, Python es un lenguaje que se ha utilizado desde tiempo atrás en el ámbito de la IA, puesto que provee facilidades al programador al ser multiparadigma y

---

<sup>4</sup> Lenguaje formal que facilita a las personas la redacción de una serie de instrucciones en forma de algoritmos para controlar el funcionamiento de un sistema.

tener una sintaxis que es de fácil y rápida comprensión, lo que conlleva a que la curva de aprendizaje sea eficiente, además ofrece otras ventajas como proveer librerías potentes para distintas tareas, una gran comunidad, facilidad al momento de realizar procesamiento de grandes cantidades de datos (Big Data) y al ser multiplataforma su código puede ser desarrollado en cualquier entorno que tenga intérprete.

Ahora bien, en cuanto a los proyectos que toman en cuenta los conceptos definidos anteriormente, se realizó una búsqueda que da como resultado la inexistencia de desarrollos que relacionen el área de las competencias ciudadanas con el machine learning y el PLN en Colombia. Sin embargo, se encuentran otros desarrollos que se refieren al uso del machine learning y el PLN en otros contextos, entre ellos:

- Automatic Detection of Offensive Language for Urdu and Roman Urdu

Este documento es de importancia para el desarrollo del proyecto puesto que realiza un análisis de lenguaje ofensivo en el idioma Urdu y Urdu Romano, proporcionando como contribución la primera base de datos en un lenguaje que debido a su sintaxis no se ha abarcado constantemente con el fin de implementar estudios relacionados con el área de la inteligencia artificial. Así mismo, muestra el procesamiento del lenguaje natural como la estructuración de ideas que al ser conceptualizadas adquieren un significado que puede ser analizado y categorizado (en este caso como positivo o negativo) por medio de distintas herramientas, entre ellas el estudio proporciona los modelos bayesianos, vecino más cercano, árbol y la regresión, además de utilizar 17 clasificadores. Por último, da a conocer los pasos que se deben tener en cuenta desde el momento en que se recopilan hasta que son utilizados en el área de Machine Learning, entre los pasos se encuentran establecer los “Labels”, generar los archivos CSV, la normalización, la extracción de los n-gram, entre otros. (Akhter, 2020 )

- Ensuring inclusion and diversity in research and research output: A case for a language-sensitive NLP crowdsourcing platform

Este documento es de importancia puesto que su enfoque social permite entender que el procesamiento del lenguaje natural no tiene porqué limitarse tan solo al análisis sintáctico para la comunicación entre dos personas en un solo idioma, por el contrario, en una época en la que todos nos consideramos “ciudadanos del mundo” la tecnología debería ser un motor que impulse la comunicación más allá de las limitaciones que pueden llegar a presentarse dentro del proceso tradicional. En este caso el artículo propone la plataforma Tashkeel como plataforma de crowdsourcing con el fin de realizar la respectiva

investigación del procesamiento del lenguaje natural en idioma árabe debido a las muchas especificaciones o complejidades que este tiene. (Alahmadi, 2020)

- An effecient fake news detection system using machine learning

Este documento es de importancia para la elaboración del proyecto puesto que hace uso de varios temas que se tratan a lo largo de la propuesta de trabajo de grado, entre ellos esta Python como lenguaje de programación (desarrollo y pruebas) y el procesamiento del lenguaje natural, además de enfocar la investigación a un análisis que valide si la información que se procesa en verdadera o falsa, teniendo en cuenta que dentro de los idiomas existen diferenciaciones en cuanto a la sintaxis. Por último, cabe resaltar que este artículo indica que se hace uso de los clasificadores de aprendizaje automático SVM, K-Neighbours, árbol de decisión y bosque aleatorio con el finde realizar los respectivos análisis de la información ya sea acompañados de operaciones matemáticas (fórmulas y operaciones aritméticas) y matriciales que se ven reflejadas en las gráficas presentadas en el documento. (Lakshmanarao, 2019)

- Text classification techniques: A literature review

Este documento es de importancia para el proyecto propuesto ya que trata sobre las técnicas de clasificación de texto para posteriores análisis con ayuda de la inteligencia artificial, además de encontrar las fortalezas, debilidades y posibles campos de aplicación que cada una de ellas posee. Con este artículo el proyecto encuentra referentes que establecen bases teóricas que enfaticen en las técnicas de clasificación de los datos que se puedan extraer y que son necesarias cuando se trabajan temas como el procesamiento del lenguaje natural (PLN) ya que en este caso, la eficacia de la inteligencia artificial que desarrollará el grupo de trabajo dependerá de la calidad de la información recolectada en los distintos documentos (contextos) referentes al área de las competencias ciudadanas. Por último, el documento hace una distinción de los distintos entornos (aprendizaje supervisado, no supervisado y semi-supervisado) que existen en el machine learning a la hora de segmentar la data (aplicables dentro del proceso de entrenamiento de una IA) y los métodos clasificatorios que cada uno de ellos abarcan. (Thangaraj, 2018)

- Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the pólice

Este documento es de importancia para el proyecto puesto que realiza análisis de texto permitiendo una segunda opinión en cuanto a la validez de la información entregada en distintos documentos, en el caso del artículo informes policiales, así mismo se da a conocer el proceso del desarrollo de algunos aspectos de la herramienta VirePol (Modelo de detección de denuncias de falsos robos) como la construcción del modelo, el entrenamiento, el corpus utilizado, el procesamiento de la data, entre otros. Así mismo, hace uso de algunas herramientas que son relevantes en el proyecto propuesto como el procesamiento del lenguaje natural y el machine learning enfocadas en la toma de decisiones y basado en estrategias probabilísticas con el fin de detectar la posibilidad de informes fraudulentos. Por último, este proyecto brinda un enfoque distinto de la aplicación de técnicas de inteligencia artificial que teniendo como soporte el análisis de texto puede contribuir de manera positiva al desarrollo del proyecto propuesto, ya que en los resultados sobresalientes del artículo se encuentra que la tasa de éxito de VirePol es del 91%, resultado más que aceptable. (Quijano-Sánchez, 2018)

Así mismo, se hizo una investigación teórica enfocada en las áreas de Machine Learning, Inteligencia artificial, competencias ciudadanas que para este proyecto está distribuida en 4 áreas y los procesos de enseñanza y aprendizaje. Dicho esto, se hace una búsqueda de autores en distintos motores de investigación como Scopus y Google Scholar, con ayuda de una plantilla en Excel brindada por un docente de una materia ya vista en la maestría, para así crear la ecuación de búsqueda con los parámetros y conectores necesarios.

Como resultado de esta búsqueda, se filtran los datos obtenidos para así seleccionar un total de 13 documentos o referentes teóricos que pueden apoyar y sustentar la investigación seleccionada como trabajo de grado para esta Maestría.



## ECUACIÓN DE BÚSQUEDA

("Artificial intelligence" OR "Google Bert" OR Chatbot OR "Machine Learning") AND (citizens OR Society) AND "Natural language processing"

Restricciones: (LIMIT-TO ( OA , "all" ) ) AND ( LIMIT-TO ( PUBYEAR , 2021 ) OR LIMIT-TO ( PUBYEAR , 2020 ) OR LIMIT-TO ( PUBYEAR , 2019 ) OR LIMIT-TO ( PUBYEAR , 2018 ) ) AND ( LIMIT-TO ( SUBJAREA , "COMP" ) OR LIMIT-TO ( SUBJAREA , "ENGI" ) )

Figura 3. Cadena de búsqueda en Scopus, Fuente: Autor.

| Ecuación de búsqueda                                                             |                     |                     |           |          |  |  |  |  |  |
|----------------------------------------------------------------------------------|---------------------|---------------------|-----------|----------|--|--|--|--|--|
| 1                                                                                | "Machine learning " |                     |           |          |  |  |  |  |  |
| 2                                                                                | chatbot             |                     |           |          |  |  |  |  |  |
| 3                                                                                | community           |                     |           |          |  |  |  |  |  |
| 4                                                                                | learning            |                     |           |          |  |  |  |  |  |
| 5                                                                                |                     |                     |           |          |  |  |  |  |  |
| 6                                                                                |                     |                     |           |          |  |  |  |  |  |
| 7                                                                                |                     |                     |           |          |  |  |  |  |  |
| 8                                                                                |                     |                     |           |          |  |  |  |  |  |
| 9                                                                                |                     |                     |           |          |  |  |  |  |  |
| 10                                                                               |                     |                     |           |          |  |  |  |  |  |
| Escribe hasta cinco sinónimos que correspondan a cada una de tus palabras clave. |                     |                     |           |          |  |  |  |  |  |
|                                                                                  | "Machine learning " | chatbot             | community | learning |  |  |  |  |  |
| S1                                                                               | IA                  | "Virtual assistant" | Citizen   |          |  |  |  |  |  |
| S2                                                                               |                     |                     | Politics  |          |  |  |  |  |  |
| S3                                                                               |                     |                     |           |          |  |  |  |  |  |
| S4                                                                               |                     |                     |           |          |  |  |  |  |  |
| S5                                                                               |                     |                     |           |          |  |  |  |  |  |

Figura 4. Matriz de Excel para ecuación de búsqueda. Elaboración Propia.

☒ Documents ☐ Authors ☐ Affiliations [Advanced](#)

Search  
"Machine learning" × Article title, Abstract, Keywords ▼ +  
E.g., "Cognitive architectures"AND robots

AND ▼

Search  
chatbot × All fields ▼ - +

AND ▼

Search  
community × All fields ▼ - +

AND ▼

Search  
learning × All fields ▼ - +

Figura 5. Búsqueda en Scopus con palabra diferenciadora community. Elaboración Propia.

☒ Documents ☐ Authors ☐ Affiliations [Advanced](#)

Search  
"Machine learning" × Article title, Abstract, Keywords ▼ +  
E.g., "Cognitive architectures"AND robots

AND ▼

Search  
chatbot × All fields ▼ - +

AND ▼

Search  
citizen × All fields ▼ - +

AND ▼

Search  
learning × All fields ▼ - +

Figura 6. Búsqueda en Scopus con palabra diferenciadora citizen. Elaboración Propia.

☒ Documents ☐ Authors ☐ Affiliations **Advanced**

Search "Machine learning"  Article title, Abstract, Keywords

E.g., "Cognitive architectures" AND robots

AND

Search chatbot  All fields    AND

Search politics  All fields    AND

Search learning  All fields    AND

Figura 7. Búsqueda en Scopus con palabra diferenciadora politics. Elaboración Propia.

Google Académico Machine learning for learning citizenship skills

Aproximadamente 194.000 resultados (0,13 s)

Artículos

Cualquier momento  
Desde 2020  
Desde 2019  
Desde 2016  
Intervalo específico...

Ordenar por relevancia  
Ordenar por fecha

Cualquier idioma  
Buscar sólo páginas en español

☒ incluir patentes  
☒ incluir citas

☐ Crear alerta

**Teaching citizen scientists to categorize glitches using machine learning guided training**  
C. Jackson, C. Osterlund, K. Crowston, M. Harandi - Computers in Human ... - 2020 - Elsevier  
The Gravity Spy project couples human volunteers with machine learning to develop a catalog of ... Prototypes may help learners categorize new stimuli by severing as a reference to the ... Another form of category learning emerges from the presentation of many references for the ...  
☆  Citado por 57 Artículos relacionados Las 4 versiones [\[HTML\] sciencedirect.com](https://www.sciencedirect.com)

**Open data as open educational resources: Towards transversal skills and global citizenship**  
J. Alenx, L. Havemann, E. Priego - Open praxis, 2015 - learntechlib.org  
... by a commitment to make data available publicly in non-proprietary, machine-readable formats at ... Critical thinking For Weinberger & Fischer (2006), learners are supposed to ... within the academy, using the same research methodologies that students are learning to master ...  
☆  Citado por 57 Artículos relacionados Las 29 versiones [\[PDF\] learntechlib.org](https://www.learntechlib.org)

**Building Citizenship Skills in Students.**  
C. Blozis, R. Scalise, C.E. Waterman, M. Wells - 2002 - ERIC  
... of participation in community service and a need for development of basic citizenship skills including more ... in the selection of three major categories of intervention: (1) service learning, (2) character ... an increased understanding of their role and responsibilities as a citizen, and a ...  
☆  Citado por 10 Artículos relacionados Las 3 versiones [\[PDF\] ed.gov](https://www.ed.gov)

**How the machine 'thinks': Understanding opacity in machine learning algorithms**  
J. Burrell - Big Data & Society, 2016 - journals.sagepub.com  
How the machine 'thinks': Understanding opacity in machine learning algorithms. Jenna Burrell. Big Data & Society 2016 3:1. Download Citation ... How the machine 'thinks': Understanding opacity in machine learning algorithms. Jenna Burrell. Big Data & Society 2016 3:1. Share ...  
☆  Citado por 793 Artículos relacionados Las 3 versiones [\[PDF\] sagepub.com](https://www.sagepub.com)

Figura 8. Búsqueda en Google Scholar. Elaboración Propia.

Los documentos son los siguientes:

- Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy.

Este documento es importante puesto que reúne el conocimiento de algunos expertos contribuyentes en el área de la inteligencia artificial con el fin de resaltar las oportunidades

significativas, la evaluación realista del impacto, desafíos y agenda de investigación potencial que plantea la rápida aparición de la inteligencia artificial en sectores como las empresas, el gobierno, el sector público y la ciencia y tecnología. Además, ofrece una visión significativa y oportuna de la tecnología de inteligencia artificial y su impacto en el futuro de la industria y la sociedad en general, reconociendo los factores sociales e industriales que influyeron en el desarrollo de la misma.

Tabla 1. Datos bibliográficos de documento

| Autores                                                                      | Año  | Fuente                                          | Temáticas                                                                                                    | Citaciones | Open Access |
|------------------------------------------------------------------------------|------|-------------------------------------------------|--------------------------------------------------------------------------------------------------------------|------------|-------------|
| Dwivedi, Y.K., Hughes, L., Ismagilova, E., (...), Walton, P., Williams, M.D. | 2019 | International Journal of Information Management | Inteligencia Artificial<br>Aprendizaje automático<br>Visión significativa y oportuna de la tecnología de IA. | 74         | No          |

*Fuente: Autor*

- Scenario technique to elicit research and training needs in digital government employing disruptive technologies.

Este documento es importante puesto que analiza la forma en que las tecnologías disruptivas podrían implantarse en diversas áreas del gobierno digital, ya que esto requiere una investigación sistemática para comprender las necesidades emergentes de investigación y capacitación. Además, propone una técnica de escenarios futuros para identificar las necesidades de investigación y formación junto con la introducción de nuevas tecnologías disruptivas en el sector público. También, describe la metodología de este enfoque de escenarios y una ejemplificación de la identificación de las necesidades de investigación y formación relacionadas con la implementación de Internet de las cosas.

Tabla 2. Datos bibliográficos de documento

| Autores               | Año  | Fuente                                         | Temáticas                                                                                                             | Citaciones | Open Access |
|-----------------------|------|------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|------------|-------------|
| Ronzhyn, A., Spitzer, | 2019 | ACM International Conference Proceeding Series | <ul style="list-style-type: none"> <li>• Tecnologías disruptivas.</li> <li>• Implementación de tecnologías</li> </ul> | 2          | No          |

| Autores          | Año | Fuente    | Temáticas                                           | Citaciones | Open Access |
|------------------|-----|-----------|-----------------------------------------------------|------------|-------------|
| V., Wimmer, M.A. |     | pp. 41-47 | disruptivas en diversas áreas del gobierno digital. |            |             |

Fuente: Autor

- Artificial intelligence as a disruptive technology: Economic transformation and government regulation.

Este documento es importante puesto que, al tomar la inteligencia artificial como la última evolución tecnológica, cubre el significado, tipos, subcampos y aplicaciones de la misma. Además de tratar una discusión sobre la superinteligencia y la cuestión y aplicabilidad de la conciencia en las máquinas. En conclusión, permite tener una base teórica sobre la inteligencia artificial y una base ética sobre lo que implica darle capacidad de razonamiento a la máquina.

Tabla 3. Datos bibliográficos de documento

| Autores    | Año  | Fuente                                                                                                       | Temáticas                                                                                                                                                                                                       | Citaciones | Open Access |
|------------|------|--------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|-------------|
| Girasa, R. | 2020 | <i>Artificial Intelligence as a Disruptive Technology: Economic Transformation and Government Regulation</i> | <ul style="list-style-type: none"><li>• Inteligencia Artificial.</li><li>• Análisis (político, económico y social) en distintos campos con base en la aplicación de la IA y el Internet de las cosas.</li></ul> | 1          | No          |

Fuente: Autor

- On the interplay between requirements, engineering, and artificial intelligence.

Este trabajo es importante puesto que se presentan reflexiones sobre los problemas que enfrenta la disciplina y practica académica de la ingeniería de requisitos. También se habla

del cambio transformador en la forma en que se lleva a cabo la ingeniería de requisitos, ya que la realidad actual exagera la inteligencia artificial y el aprendizaje autónomo, dichos cambios se reflejan en dos aspectos: las herramientas de Inteligencia Artificial se utilizan cada vez más durante el proceso de Ingeniería de Requisitos y el proceso de Ingeniería de Requisitos para sistemas que incluyen Inteligencia Artificial es diferente.

Tabla 4. Datos bibliográficos de documento

| Autores                              | Año  | Fuente                         | Temáticas                                                                                                                                    | Citaciones | Open Access |
|--------------------------------------|------|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|------------|-------------|
| Kostova, B., Gürses, S., Wegmann, A. | 2020 | CEUR Workshop Proceedings 2584 | <ul style="list-style-type: none"> <li>Inteligencia Artificial.</li> <li>Aprendizaje autónomo.</li> <li>Ingeniería de Requisitos.</li> </ul> | 0          | No          |

*Fuente: Autor*

- Application-agnostic chatbot deployment considerations: A case study.

Este trabajo es importante puesto que toma relevante los avances en el aprendizaje autónomo para la interacción entre humanos y maquinas, además trata específicamente a los chatbots como una tecnología que hace uso de los últimos avances en el procesamiento del lenguaje natural con el fin de comprender texto (idioma) y producir una respuesta. Además, repasa un hecho histórico que tuvo lugar en 2016 con el propósito de extraer lecciones importantes.

Tabla 5. Datos bibliográficos de documento

| Autores                                       | Año  | Fuente                                                                                                                      | Temáticas                                                                           | Citaciones | Open Access |
|-----------------------------------------------|------|-----------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|------------|-------------|
| Rivas, P., Chelsi, C., Nishit, N., Ravula, L. | 2019 | Proceedings - 6th Annual Conference on Computational Science and Computational Intelligence, CSCI 2019 9071112, pp. 361-365 | <ul style="list-style-type: none"> <li>Machine Learning</li> <li>Chatbot</li> </ul> | 0          | No          |

*Fuente: Autor*

- Personalisation the artificial intelligence way.

Este trabajo es importante puesto que describe cinco tipos diferentes de IA (sonido, series de tiempo, texto, imagen y video) e ilustra varias formas en que se puede utilizar la IA, entre ellas está la gestión de relaciones con los clientes, comercio electrónico, recomendaciones de clientes, seguridad, asistencia por voz y lenguaje natural. Además, este trabajo, sostiene que la inteligencia artificial se convertirá en la base de un nivel de personalización del cliente que no solo será reconocido, sino que pronto será demandado por clientes en todas partes.

Tabla 6. Datos bibliográficos de documento

| Autores     | Año  | Fuente                                                             | Temáticas                                                                                          | Citaciones | Open Access |
|-------------|------|--------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|------------|-------------|
| Pearson, A. | 2019 | Journal of Digital and Social Media Marketing<br>7(3), pp. 245-269 | Inteligencia artificial.<br>El aprendizaje automático.<br>El aprendizaje profundo.<br>Tipos de IA. | 0          | No          |

*Fuente: Autor*

- Mind the gap: responsible robotics and the problem of responsibility.

Este trabajo es importante puesto que trata de responder la pregunta sobre los robots y la responsabilidad: responder por la forma en que entendemos, debatimos y decidimos quién o qué puede responder por las decisiones y acciones emprendidas, mediante mecanismos cada vez más interactivos, autónomos y sociables. Además, realiza un análisis dividió en diferentes etapas, comienza examinando críticamente la teoría instrumental de la tecnología y luego considera tres casos en los que las innovaciones (tecnología autónoma, aprendizaje automático y robots sociales) recientes en robótica desafían este procedimiento operativo estándar al abrir brechas en la forma habitual de asignar responsabilidades.

Tabla 7. Datos bibliográficos de documento

| Autores      | Año  | Fuente                                                  | Temáticas                                                                              | Citaciones | Open Access |
|--------------|------|---------------------------------------------------------|----------------------------------------------------------------------------------------|------------|-------------|
| Gunkel, D.J. | 2020 | Ethics and Information Technology<br>22(4), pp. 307-320 | Teoría instrumental de la tecnología.<br>Ética de la máquina y responsabilidad híbrida | 14         | No          |

*Fuente: Autor*

- Recurrent convolutional neural network for answer selection in community question answering.

Este trabajo es importante puesto que propone una red convolucional recurrente (RCNN) para dar respuesta a preguntas de la comunidad, para ello combina la red neuronal convolucional (CNN) con la red neuronal recurrente (RNN) para capturar la coincidencia semántica entre pregunta y respuesta. El clasificador softmax se utiliza para identificar la calidad de coincidencia de las respuestas para una pregunta determinada.

Además, se diseña una estrategia de aprendizaje de dos frases para entrenar el modelo, que ajusta los RNN con las representaciones aprendidas dependientes del contexto.

Tabla 8. Datos bibliográficos de documento

| Autores                              | Año  | Fuente                          | Temáticas                                                                     | Citaciones | Open Access |
|--------------------------------------|------|---------------------------------|-------------------------------------------------------------------------------|------------|-------------|
| Zhou, X., Hu, B., Chen, Q., Wang, X. | 2018 | Neurocomputing<br>274, pp. 8-18 | Red neuronal convolucional recurrente (RCNN)<br>Red neuronal recurrente (RNN) | 28         | No          |

*Fuente: Autor*

- Developing enterprise chatbots: Learning linguistic structures (Book).

Este trabajo es importante puesto que enfoca el tema al desarrollo de un chatbot que sea capaz de soportar una conversación cohesiva y coherente y tenga conocimientos, convirtiéndolo en un sistema inteligente avanzado y complejo. También, menciona que los diseñadores deben aprender a combinar lo intuitivo, enfoques explicables de comprensión y razonamiento del lenguaje con tecnologías de aprendizaje profundo y estadísticas de alto rendimiento.

Es importante porque menciona los paradigmas populares que existen para la construcción de un chatbot, el primero consiste en construir una plataforma de bot con capacidades universales de PNL y ML, y el segundo en acumular un gran conjunto de datos de diálogo de entrenamiento para enviarlo a una red de aprendizaje profundo y esperar que el chatbot capacitado aprenda automáticamente "cómo chatear". Por último, el libro proporciona una fuente completa de algoritmos y arquitecturas para construir chatbots para varios dominios basados en las tendencias recientes en lingüística computacional y aprendizaje automático.

Tabla 9. Datos bibliográficos de documento

| Autores     | Año  | Fuente                                                                   | Temáticas                                                             | Citaciones | Open Access |
|-------------|------|--------------------------------------------------------------------------|-----------------------------------------------------------------------|------------|-------------|
| Galitsky, B | 2019 | Developing Enterprise Chatbots: Learning Linguistic Structures pp. 1-559 | Chatbot. tecnologías de aprendizaje profundo Construcción de chatbot. | 9          | No          |

*Fuente: Autor*

- Measuring semantic coherence of a conversation.

Este trabajo es importante puesto que, propone una medición en la coherencia semántica de una conversación con respecto a un conocimiento previo que se basa en la identificación de relaciones semánticas entre conceptos introducidos durante la misma, teniendo en cuenta que los sistemas de conversación se han vuelto cada vez más populares como una forma para que los humanos interactúen con las computadoras. También propone y evalúa el aprendizaje autónomo y la coherencia semántica utilizando gráficos de conocimiento, sus

incrustaciones de espacio vectorial y modelos de incrustación de palabras, como fuentes de conocimiento previo.

Tabla 10. Datos bibliográficos de documento

| Autores                                                             | Año  | Fuente                                                                                                                                                          | Temáticas                                                                                       | Citaciones | Open Access |
|---------------------------------------------------------------------|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------|-------------|
| Vakulenko, S., de Rijke, M., Cochez, M., Savenkov, V., Polleres, A. | 2018 | Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)<br>11136 LNCS, pp. 634-651 | Estructura y semántica de una conversación.<br>Coherencia semántica.<br>Aprendizaje automático. | 5          | No          |

*Fuente: Autor*

- An Overview of Chatbot Technology.

Este trabajo es importante puesto que abarca la evolución historia de los chatbots, las motivaciones que impulsan a su uso en distintas áreas y su utilidad. También, destaca el impacto de los estereotipos sociales en el diseño de chatbots, aclara algunos conceptos tecnológicos y da a conocer la clasificación de los chatbot basada en distintos criterios. Por último, presenta la arquitectura general de los chatbots modernos y también mencionamos las principales plataformas para su creación.

Tabla 11. Datos bibliográficos de documento

| Autores           | Año  | Fuente                                               | Temáticas                             | Citaciones | Open Access |
|-------------------|------|------------------------------------------------------|---------------------------------------|------------|-------------|
| Eleni Adamopoulou | 2020 | Artificial Intelligence Applications and Innovations | Chatbot.<br>Utilidad de los chatbots. | 4          | Si          |

| Autores             | Año | Fuente | Temáticas                                                                                              | Citaciones | Open Access |
|---------------------|-----|--------|--------------------------------------------------------------------------------------------------------|------------|-------------|
| Lefteris Moussiades |     |        | Impacto de los estereotipos sociales en el diseño de chatbots.<br>Concepto y clasificación de chatbot. |            |             |

*Fuente: Autor*

- An overview of machine learning in chatbots.

Este trabajo es importante puesto que proporciona una descripción general de los chatbot y las tendencias de investigación relacionadas con el desarrollo de los mismos de apariencia humana. Se realiza una revisión bibliográfica presentada en un mapa mental y se da a conocer que los chatbots operan en tres etapas: comprender la entrada del lenguaje natural, generar una respuesta automática y relevante, y la construcción de respuestas de lenguaje natural realistas y fluidas.

Por último, el trabajo dice que el cuello de botella actual en el diseño de chatbots con inteligencia artificial radica en la falta de procesamiento del lenguaje natural, por lo cual carece de la capacidad de comprender correctamente el contenido y el contexto de la entrada de un usuario.

Tabla 12. Datos bibliográficos de documento

| Autores                                                    | Año  | Fuente                                              | Temáticas                                                     | Citaciones | Open Access |
|------------------------------------------------------------|------|-----------------------------------------------------|---------------------------------------------------------------|------------|-------------|
| Prissadang Suta, Xi Lan2, Biting Wu2 Pornchai Mongkolnam1, | 2020 | International Journal of Mechanical Engineering and | Chatbots.<br>Lenguaje natural.<br>Tendencias de investigación | 1          | Si          |

| Autores          | Año | Fuente                              | Temáticas                                                                                                       | Citaciones | Open Access |
|------------------|-----|-------------------------------------|-----------------------------------------------------------------------------------------------------------------|------------|-------------|
| Jonathan H. Chan |     | Robotics Research<br>Vol. 9, No. 4, | relacionadas con el desarrollo de chatbots. Construcción de respuestas de lenguaje natural realistas y fluidas. |            |             |

*Fuente: Autor*

- How the machine ‘thinks’: Understanding opacity in machine learning algorithms.

Este trabajo es importante puesto que considera el tema de la opacidad como un problema para los mecanismos de clasificación socialmente consecuentes, como filtros de spam, detección de fraudes con tarjetas de crédito, motores de búsqueda, tendencias de noticias, segmentación del mercado y publicidad, calificación de seguros o préstamos y calificación crediticia.

También se realiza una distinción entre tres formas de opacidad: la primera es opacidad como secreto corporativo o estatal intencional, la segunda es opacidad como secreto técnico y la tercera es una opacidad que surge de las características de los algoritmos de aprendizaje automático.

Tabla 13. Datos bibliográficos de documento

| Autores       | Año  | Fuente                        | Temáticas                                                                                                   | Citaciones | Open Access |
|---------------|------|-------------------------------|-------------------------------------------------------------------------------------------------------------|------------|-------------|
| Jenna Burrell | 2016 | Big Data & Society<br>January | Opacidad como un problema para los mecanismos de clasificación.<br>Formas de opacidad.<br>Machine learning. | 793        | Si          |

| Autores | Año | Fuente | Temáticas                     | Citaciones | Open<br>Access |
|---------|-----|--------|-------------------------------|------------|----------------|
|         |     |        | Opacidad en una red neuronal. |            |                |

*Fuente: Autor*

## 6. METODOLOGÍA: MODELO DE DESARROLLO

Para este proyecto que tiene un enfoque investigativo mixto (Hernández Sampieri, 2014) se pretende trabajar la metodología ágil SCRUM, puesto que facilita el proceso de desarrollo y documentación del software que en este caso es el sistema computacional basado en machine learning, además permite una revisión constante de los avances del proyecto ya que por medio de diferentes estrategias como actas de entregas o reuniones entre el tutor y el estudiante, se puede validar el cumplimiento de metas y objetivos con el fin de que el sector educativo (Colegio de Boyacá de la ciudad de Tunja) tenga un impacto tecnológico positivo con el que los estudiantes se apropien de los conocimientos relacionados con las competencias ciudadanas.

Con dicha metodología se tienen en cuenta las siguientes fases que facilitarán el diseño, programación y testing del sistema computacional que se pretende realizar:

1. **Recopilación y adaptación de la información:** En esta fase se realiza un proceso de investigación con el fin de crear una biblioteca de información relacionada con los procesos de enseñanza – aprendizaje para las competencias ciudadanas en Colombia que permita nutrir la IA, con ello se asegura que la coherencia de las respuestas tenga un mayor porcentaje de acierto, por lo cual se espera como resultado de esta primera etapa crear una línea base y un modelo con el que se obtenga y filtre la información. Así mismo, dicha biblioteca de información será utilizada para crear el corpus que tendrá como objetivo ser la base de datos para el entrenamiento de modelos de machine learning.
2. **Implementación del modelo pre-entrenado (Codificación del Backend):** En esta fase se realiza el proceso de selección e implementación de las técnicas de machine learning y algoritmos necesarios para diseñar y programar la IA, de manera tal que términos mencionados anteriormente como el PLN entran a pasar de la teoría a la práctica, además de algunos otros conceptos que son necesarios para que por ejemplo, la inteligencia artificial sea capaz de generar una respuesta coherente en un lenguaje específico. Así mismo, se hace la respectiva configuración de archivos, paquetes, librerías y codificación por medio de distintas herramientas y lenguajes como Python, con el fin de crear una inteligencia artificial capaz de responder preguntas relacionadas con el área de las competencias ciudadanas.
3. **Testing:** En esta fase se realizan las pruebas necesarias para comprobar que la inteligencia artificial es capaz de responder de manera coherente a las preguntas

relacionadas al área de las competencias ciudadanas propuestas por el usuario, por lo cual en caso de encontrar algún error se debe realizar la respectiva corrección y retomar la fase.

A continuación, se da a conocer de manera detallada cada una de las fases mencionadas anteriormente:

Tabla 14. Fase 1

| Tipo                        | Elementos                                                                                                                                                                                                                                                                                                                                                                   |
|-----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Recopilación de información | <ul style="list-style-type: none"><li>• Recolección de documentación sobre competencias ciudadanas en Colombia para uso de educación a nivel media.</li><li>• Normalización de los documentos. (Filtrado y clasificación de información)</li><li>• Creación del corpus para los cuatro enfoques conversacionales</li><li>• Preparación y configuración de la data</li></ul> |
| Recursos                    | <ul style="list-style-type: none"><li>• Sitio de trabajo</li><li>• Conexión a internet</li><li>• Computador</li><li>• Navegador Web</li><li>• Office</li></ul>                                                                                                                                                                                                              |
| Resultados                  | Corpus que será utilizado como base de datos para alimentar la inteligencia artificial.                                                                                                                                                                                                                                                                                     |

*Fuente: Autor*

Tabla 15. Fase 2

| Tipo                                                                | Elementos                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Implementación del modelo pre-entrenado (Codificación del Backend): | <ul style="list-style-type: none"><li>• Configuración de entorno de desarrollo</li><li>• Selección de algoritmo de machine learning (Ventajas que provee)</li><li>• Selección e implementación del modelo pre-entrenado para análisis semántico (criterios de porcentaje de fiabilidad)</li><li>• Configuración y codificación de los archivos necesarios para el funcionamiento de la IA</li></ul> |
| Recursos                                                            | <ul style="list-style-type: none"><li>• Sitio de trabajo</li><li>• Conexión a internet</li><li>• Computador</li><li>• Navegador Web</li></ul>                                                                                                                                                                                                                                                       |

| Tipo       | Elementos                                                                                                                                     |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
|            | <ul style="list-style-type: none"><li>• Aplicaciones Online</li><li>• Programas del computador</li><li>• Librerías</li><li>• Office</li></ul> |
| Resultados | Creación de la inteligencia artificial a nivel funcional                                                                                      |

*Fuente: Autor*

Tabla 16. Fase 3

| Tipo       | Elementos                                                                                                                                                                                                      |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Testing    | <ul style="list-style-type: none"><li>• Validación de la inteligencia artificial</li></ul>                                                                                                                     |
| Recursos   | <ul style="list-style-type: none"><li>• Sitio de trabajo</li><li>• Conexión a internet</li><li>• Computador</li><li>• Navegador Web</li><li>• Aplicaciones Online</li><li>• Programas del computador</li></ul> |
| Resultados | Comprobación del correcto funcionamiento de la inteligencia artificial.                                                                                                                                        |

*Fuente: Autor*

## 7. DESARROLLO DEL PROYECTO

Para el desarrollo de este proyecto se utilizó la metodología ágil SCRUM, que permite por medio de distintas etapas repetitivas denominadas Sprint realizar cambios incrementales en el producto, así mismo se aprovecharon los artefactos que ésta posee, entre ellos el product backlog que se desglosará en la sección 6.1, las scrum meetings teniendo como evidencia una serie de actas y utilizando para ello espacios presenciales y virtuales y las historias de usuario.



Figura 9. Descripción de la metodología ágil SCRUM. Tomada de <https://www.digite.com/es/agile/que-es-scrum/>

La implementación del modelo pre-entrenado y el proceso de codificación se realizó por medio de la herramienta web de libre uso Google Colaboratory, permitiendo codificar directamente con Python. Dicho esto, se tiene como resultado un proceso evolutivo representado en una serie de archivos que evidencian el incremento de acuerdo a los objetivos planteados.

A continuación, se presenta la descripción de cada una de las actividades realizadas:

### 7.1. Fase 1. Definición de requisitos.

En ese capítulo se incluyen las funcionalidades y características a tener en cuenta en la implementación de un modelo pre-entrenado, para el sistema de pregunta – respuesta en el área de competencias ciudadanas.

#### 7.1.1. Descripción general

Se desarrolló la implementación de un modelo pre-entrenado por medio de la plataforma web Google Colaboratory con el lenguaje de programación Python, con el fin de realizar la

dinámica del proceso pregunta – respuesta teniendo como base 4 contextos (Principios fundamentales de la Constitución Política Colombiana de 1991, organización del estado colombiano, los modelos de participación ciudadana y los derechos humanos) del área de las competencias ciudadanas.

### 7.1.2. Funcionalidad del producto

- ✓ Implementación de un modelo pre-entrenado en Google Colaboratory.
- ✓ Conexión con Google Drive para lectura de contextos.
- ✓ Validación del modelo por medio de funciones.

### 7.1.3. Product Backlog e Historias de Usuario

Para esta fase, se creó el product backlog que consiste en una tabla que identifica cada una de las tareas realizadas, el sprint a las que pertenecen y el estado en que se encuentran. Cabe resaltar que el product backlog está en constante cambio puesto que evoluciona al mismo tiempo que el proyecto avanza sin embargo, con este artefacto se permitió construir las historias de usuario del proyecto agregando así otras características a las tablas que las representan, como por ejemplo la prioridad y el responsable de su realización.

| Proceso Transversal                                | Product Backlog |                                                  |                          |
|----------------------------------------------------|-----------------|--------------------------------------------------|--------------------------|
|                                                    | Sprint          | Descripción                                      | Estado                   |
|                                                    |                 |                                                  | En proceso    Finalizada |
| Creación del Corpus de las competencias Ciudadanas | 1               | Levantamiento de Requisitos                      | x                        |
|                                                    | 2               | Documentación de Requisitos                      | x                        |
|                                                    | 3               | Selección de arquitectura IA                     | x                        |
|                                                    |                 | Selección de Algoritmo IA                        | x                        |
|                                                    | 4               | Configuración Entorno de Desarrollo              | x                        |
|                                                    |                 | Selección de modelo pre-entrenado                | x                        |
|                                                    |                 | Creación de archivos Jupiter (Google Colab)      | x                        |
|                                                    |                 | Configuración de Archivos Jupiter (Google Colab) | x                        |
|                                                    |                 | Cambio de entorno a GPU                          | x                        |
|                                                    |                 | Configuración de GPU - Reconocimiento de GPU     | x                        |
|                                                    |                 | Instalación de librerías                         | x                        |
|                                                    |                 | Tokenización                                     | x                        |
|                                                    |                 | Carga de contextos desde Google Drive            | x                        |
|                                                    |                 | Tokenización de Contextos                        | x                        |
|                                                    |                 | Pipeline                                         | x                        |
|                                                    |                 | Creación de funciones                            | x                        |
|                                                    |                 | Ejecución de IA                                  | x                        |
|                                                    | 5               | Exportar Modelo                                  | x                        |
|                                                    | 6               | Realizar pruebas de IA con Corpus                | x                        |

Figura 10. Product Backlog. Elaboración Propia

En la figura 4 se observan las tareas ejecutadas que nacen a partir de una serie de reuniones entre el docente experto en el área de las competencias ciudadanas, los

directores de la tesis y el estudiante desarrollador de la tesis. Con esto, se provee la información necesaria para crear las historias de usuario que permiten contextualizar y definir en lenguaje natural las necesidades planteadas en el desarrollo del proyecto.

Tabla 17. Historia de Usuario 1

| Historia de Usuario                                                                                                                                                                                                           |                                                     |                           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|---------------------------|
| <b>Nombre de Historia:</b> Yo como usuario quiero hacer uso de una herramienta informática (IA) para incrementar la motivación del aprendizaje.                                                                               |                                                     |                           |
| <b>Número:</b> 1                                                                                                                                                                                                              | <b>Cliente:</b> Usuario - Docente                   |                           |
| <b>Prioridad:</b> Alta                                                                                                                                                                                                        | <b>Estimación:</b>                                  | <b>Versión:</b> 1.0       |
| <b>Sprint:</b> 1                                                                                                                                                                                                              | <b>Responsable:</b> Harvey Nicolás Echavarría Ortíz |                           |
| <b>Hora Asignadas:</b> 1/2                                                                                                                                                                                                    | <b>Fecha Inicio:</b> 5/02/21                        | <b>Fecha Fin:</b> 5/02/21 |
| <b>Descripción:</b>                                                                                                                                                                                                           |                                                     |                           |
| Como usuario quiero utilizar una inteligencia artificial que permita adaptar los métodos de enseñanza de acuerdo a las necesidades de la sociedad moderna, implementando las TIC como una estrategia de renovación educativa. |                                                     |                           |
| <b>Validación:</b>                                                                                                                                                                                                            |                                                     |                           |
| Tiempo de uso de la herramienta informática (IA) y recepción por parte de los estudiantes de grado 11.                                                                                                                        |                                                     |                           |

Fuente: Autor

Tabla 18. Historia de Usuario 2

| Historia de Usuario                                                                                                                                                                           |                                                     |                           |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|---------------------------|
| <b>Nombre de Historia:</b> Yo como usuario quiero determinar si la exposición de los muchachos con un tipo de conocimiento mediado por la IA puede o no generar un impacto.                   |                                                     |                           |
| <b>Número:</b> 2                                                                                                                                                                              | <b>Cliente:</b> Usuario - Docente                   |                           |
| <b>Prioridad:</b> Alta                                                                                                                                                                        | <b>Estimación:</b>                                  | <b>Versión:</b> 1.0       |
| <b>Sprint:</b> 1                                                                                                                                                                              | <b>Responsable:</b> Harvey Nicolás Echavarría Ortíz |                           |
| <b>Hora Asignadas:</b> 1/2                                                                                                                                                                    | <b>Fecha Inicio:</b> 5/02/21                        | <b>Fecha Fin:</b> 5/02/21 |
| <b>Descripción:</b>                                                                                                                                                                           |                                                     |                           |
| Como usuario quiero utilizar una inteligencia artificial para mejorar los resultados obtenidos por los estudiantes en las pruebas saber 11 (ICFES) en el área de las competencias ciudadanas. |                                                     |                           |
| <b>Validación:</b>                                                                                                                                                                            |                                                     |                           |
| Los resultados que se obtengan en las pruebas saber 11 (ICFES)                                                                                                                                |                                                     |                           |

Fuente: Autor

Tabla 19. Historia de Usuario 3

| Historia de Usuario                                                                                                                                                                                                                       |                                                     |                           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|---------------------------|
| <b>Nombre de Historia:</b> Yo como usuario quiero hacer uso de una herramienta informática (IA) para poner a prueba el conocimiento de las competencias ciudadanas en estudiantes de grado 11.                                            |                                                     |                           |
| <b>Número:</b> 3                                                                                                                                                                                                                          | <b>Cliente:</b> Usuario - Docente                   |                           |
| <b>Prioridad:</b> Alta                                                                                                                                                                                                                    | <b>Estimación:</b>                                  | <b>Versión:</b> 1.0       |
| <b>Sprint:</b> 1                                                                                                                                                                                                                          | <b>Responsable:</b> Harvey Nicolás Echavarría Ortíz |                           |
| <b>Hora Asignadas:</b> 1/2                                                                                                                                                                                                                | <b>Fecha Inicio:</b> 5/02/21                        | <b>Fecha Fin:</b> 5/02/21 |
| <b>Descripción:</b>                                                                                                                                                                                                                       |                                                     |                           |
| Como usuario quiero utilizar una inteligencia artificial que permita a los estudiantes realizar consultas relacionadas con un área específica de las competencias ciudadanas para así validar si la respuesta dada por ellos es correcta. |                                                     |                           |
| <b>Validación:</b>                                                                                                                                                                                                                        |                                                     |                           |
| Los resultados que se obtengan al contratar la respuesta dada por la IA y el estudiante.                                                                                                                                                  |                                                     |                           |

*Fuente: Autor*

Tabla 20. Historia de Usuario 4

| Historia de Usuario                                                                                                                                                                                                   |                                                     |                           |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|---------------------------|
| <b>Nombre de Historia:</b> Yo como usuario quiero hacer uso de un dispositivo de inteligencia artificial para innovar la forma en que se brindar y evalúa la información relacionada con las competencias ciudadanas. |                                                     |                           |
| <b>Número:</b> 4                                                                                                                                                                                                      | <b>Cliente:</b> Usuario - Docente                   |                           |
| <b>Prioridad:</b> Alta                                                                                                                                                                                                | <b>Estimación:</b>                                  | <b>Versión:</b> 1.0       |
| <b>Sprint:</b> 1                                                                                                                                                                                                      | <b>Responsable:</b> Harvey Nicolás Echavarría Ortíz |                           |
| <b>Hora Asignadas:</b> 1/2                                                                                                                                                                                            | <b>Fecha Inicio:</b> 5/02/21                        | <b>Fecha Fin:</b> 5/02/21 |
| <b>Descripción:</b>                                                                                                                                                                                                   |                                                     |                           |
| Como usuario quiero utilizar una inteligencia artificial que permita modernizar los procesos de enseñanza – aprendizaje en las aulas de clase al hacer uso de herramientas TIC.                                       |                                                     |                           |
| <b>Validación:</b>                                                                                                                                                                                                    |                                                     |                           |
| Recepción de la IA por parte de los estudiantes.                                                                                                                                                                      |                                                     |                           |

*Fuente: Autor*

Tabla 21. Historia de Usuario 5

| Historia de Usuario |
|---------------------|
|---------------------|

|                                                                                                                                                                                                                                                 |                                                     |                           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|---------------------------|
| <b>Nombre de Historia:</b> Yo como usuario quiero hacer uso de una herramienta informática (IA) que permita a los estudiantes de grado 11 formular preguntas y obtener respuestas respecto a un área específica de las competencias ciudadanas. |                                                     |                           |
| <b>Número:</b> 5                                                                                                                                                                                                                                | <b>Cliente:</b> Usuario - Docente                   |                           |
| <b>Prioridad:</b> Alta                                                                                                                                                                                                                          | <b>Estimación:</b>                                  | <b>Versión:</b> 1.0       |
| <b>Sprint:</b> 1                                                                                                                                                                                                                                | <b>Responsable:</b> Harvey Nicolás Echavarría Ortíz |                           |
| <b>Hora Asignadas:</b> 1/2                                                                                                                                                                                                                      | <b>Fecha Inicio:</b> 5/02/21                        | <b>Fecha Fin:</b> 5/02/21 |
| <b>Descripción:</b>                                                                                                                                                                                                                             |                                                     |                           |
| Como usuario quiero utilizar una inteligencia artificial que facilite la búsqueda de información con respecto a un área específica de las competencias ciudadanas por medio de una IA.                                                          |                                                     |                           |
| <b>Validación:</b>                                                                                                                                                                                                                              |                                                     |                           |
| Coherencia de las respuestas de la IA.                                                                                                                                                                                                          |                                                     |                           |

*Fuente: Autor*

Tabla 22. Historia de Usuario 6

| Historia de Usuario                                                                                                                                                                       |                                                     |                     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|---------------------|
| <b>Nombre de Historia:</b> Yo como usuario quiero hacer uso de una herramienta informática (IA) que me permita realizar consultas sobre áreas específicas de las competencias ciudadanas. |                                                     |                     |
| <b>Número:</b> 6                                                                                                                                                                          | <b>Cliente:</b> Usuario - Estudiante                |                     |
| <b>Prioridad:</b> Alta                                                                                                                                                                    | <b>Estimación:</b>                                  | <b>Versión:</b> 1.0 |
| <b>Sprint:</b> 1                                                                                                                                                                          | <b>Responsable:</b> Harvey Nicolás Echavarría Ortíz |                     |
| <b>Hora Asignadas:</b> 1/2                                                                                                                                                                |                                                     |                     |
| <b>Descripción:</b>                                                                                                                                                                       |                                                     |                     |
| Como usuario quiero utilizar una inteligencia artificial que facilite la búsqueda de información con respecto a un área específica de las competencias ciudadanas por medio de una IA.    |                                                     |                     |
| <b>Validación:</b>                                                                                                                                                                        |                                                     |                     |
| Coherencia de las respuestas de la IA.                                                                                                                                                    |                                                     |                     |

*Fuente: Autor*

## 7.2. Fase 2. Selección de tecnología

En ese capítulo se incluye la descripción de cada una de las tecnologías utilizadas en la implementación de un modelo pre-entrenado, para el sistema de pregunta – respuesta en el área de competencias ciudadanas.

| Sprint 2 | Arquitectura                 |                             |                                         |
|----------|------------------------------|-----------------------------|-----------------------------------------|
|          | Nombre                       | Concepto                    | Ventajas                                |
|          | Redes Neuronales Transformer | Basadas en la autoatención. | Se agrega la idea de "contexto".        |
|          |                              | Con enfoque específico.     | Se procesa los datos en paralelo.       |
|          |                              |                             | Capas de codificación y decodificación. |
|          |                              |                             | Análisis bidireccional.                 |
|          | Algoritmo                    |                             |                                         |
|          | Nombre                       | Concepto                    | Ventajas                                |
|          | Bert                         | IA desarrollada por Google  | Uno de los mas usados                   |
|          |                              | Para PLN                    | Mayor eficiencia y eficacia en PLN      |
|          |                              | Arquitectura Transformer    | Múltiples usos                          |
|          |                              |                             | Publicada en 2018 (reciente)            |

Figura 11. Características de Arquitectura y Algoritmo. Elaboración Propia

En la imagen anterior se da a conocer la arquitectura y algoritmo seleccionado describiendo su concepto y las ventajas que cada uno tiene.

### 7.2.1. Herramientas utilizadas

- ✓ **Python:** “Es un lenguaje de programación de alto nivel creado a finales de los 80/principios de los 90 por Guido van Rossum” (Quintero, 2022), cuenta con las siguientes características:
  - “Es un lenguaje de programación interpretado y multiplataforma cuya filosofía hace hincapié en una sintaxis que favorezca un código legible”. (Quintero, 2022)
  - “Se trata de un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional”. (Quintero, 2022)
- ✓ **Google Colaboratory:** “Es una herramienta para escribir y ejecutar código Python en la nube de Google. También es posible incluir texto enriquecido, “links” e imágenes. En caso de necesitar altas prestaciones de cómputo, el entorno permite configurar algunas propiedades del equipo sobre el que se ejecuta el código”. (Baume, 2021)
- ✓ **Algoritmo Bert:** “Es una biblioteca de código abierto creada en 2018 en Google. Es una técnica nueva para PNL y adopta un enfoque de modelos de entrenamiento completamente diferente al de cualquier otra técnica”. (Cardellino, 2021)

- ✓ **Google Drive:** Es una herramienta que permite almacenamiento en la nube, en este caso para guardar los documentos referentes a cada contexto.
- ✓ **Arquitectura Transformer:** “Un Transformer es un tipo de arquitectura de red neuronal. En resumen, las redes neuronales son un tipo de modelo muy eficaz para analizar tipos de datos complejos como imágenes, vídeos, audio y texto”. (Group, 2022)

### 7.3. Fase 3. Diseño del sistema

Para el diseño del sistema se utilizó el entorno de desarrollo Google Colab y el lenguaje de programación Python, ya que por medio de distintas sentencias de código se logró realizar desde la instalación de las respectivas librerías hasta la ejecución de la función que relaciona el modelo pre – entrenado, los contextos que se definieron en la sección de levantamiento de requisitos y cada uno de los pepline establecidos. Cabe resaltar que dicho proceso se describirá detalladamente en el numeral 6.4 referente al desarrollo del aplicativo.

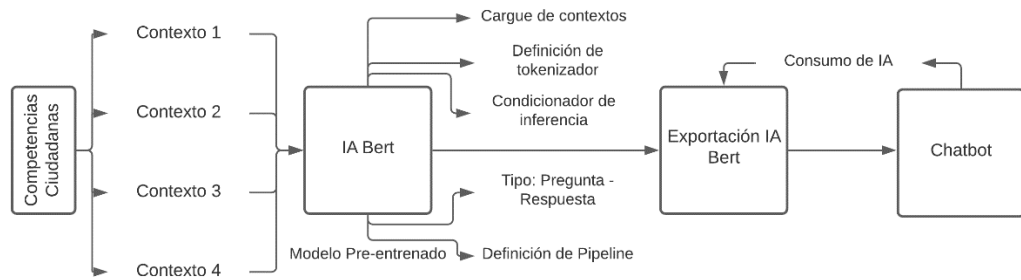


Figura 12. Contexto General IA Competencias Ciudadanas. Elaboración Propia

En la figura 6 se observa el contexto general del sistema teniendo como base la arquitectura Transformer y el algoritmo Bert, fungiendo en conjunto como la red neuronal que permitió incluir el análisis de contextos en forma bidireccional para obtener así una respuesta a las preguntas planteadas por el usuario; esto es posible gracias a la capa de atención (“dar para cada palabra de la entrada un vector de atención propia en el que se captura la relación contextual entre las palabras de la oración” (González, 2020)), la capa de feed forward (“es una red neuronal de feed forward que se aplica a cada uno de los vectores de atención de manera individual, como cada una de las redes de atención es independiente de las demás se puede paralelizar el proceso de feed forward” (González, 2020)) y dos factores más:

- Representaciones del codificador.
- Transformadores.

#### 7.4. Fase 4. Desarrollo de la aplicación

Para el desarrollo del proyecto se utilizó un proceso iterativo e incremental, que dio como resultado una serie de archivos que reflejan los avances obtenidos conforme continuaba la realización del proyecto; así mismo, se tiene en cuenta que como requisito adicional se pidió que existiera interacción “conversacional” entre el usuario y la IA, razón por la cual se hizo uso de la librería Chatterbot y su respectivo corpus. En la figura 7 se puede observar cada uno de los archivos generados además de una carpeta con el resultado final.

| Nombre ↑                                                                                                                                                                                                         | Propietario | Última modificación | Tamaño de archivo |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|---------------------|-------------------|
|  Con Conversacion                                                                                                               | yo          | 7 jul 2022 yo       | —                 |
|  Sin conversacion                                                                                                               | yo          | 7 jul 2022 yo       | —                 |
|  BERTENTREGAR V4.0(BERT+CONVERSACIONENTREGAR).ipynb                                                                             | yo          | 10:22 yo            | 152 kB            |
|  BERTENTREGARV3.ipynb                          | yo          | 10:22 yo            | 22 kB             |
|  Ejercicio BertV1.ipynb                        | yo          | 18 feb 2021 yo      | 60 kB             |
|  Ejercicio BertV2.1.ipynb                                                                                                       | yo          | 18 feb 2021 yo      | 66 kB             |
|  Ejercicio BertV2.2.ipynb                                                                                                       | yo          | 18 feb 2021 yo      | 146 kB            |
|  Ejercicio BertV2(Ejercicios Cuentos).ipynb  | yo          | 18 feb 2021 yo      | 173 kB            |

Figura 13. Versiones archivos Python. Elaboración Propia

A continuación, se dan a conocer los pasos realizados en cada uno de los archivos finales:

##### 7.4.1. IA sin contexto conversacional

###### 7.4.1.1 Reconocimiento y configuración de GPU

En este paso se aseguró que el archivo colab que se estaba trabajando estuviera bajo un entorno de ejecución de GPU, lo cual provee una mayor potencia de procesamiento a la hora de realizar cálculos, ejecuciones de IA y uso de librerías. Mayor razón para utilizar un entorno que provee de almacenamiento en la nube y una potente capacidad en cuanto a RAM y disco, en las figuras 14 y 15 que se muestran a continuación, se evidencia el cambio para el uso de GPU para así obtener una mayor eficiencia mientras se hace uso del archivo colab.

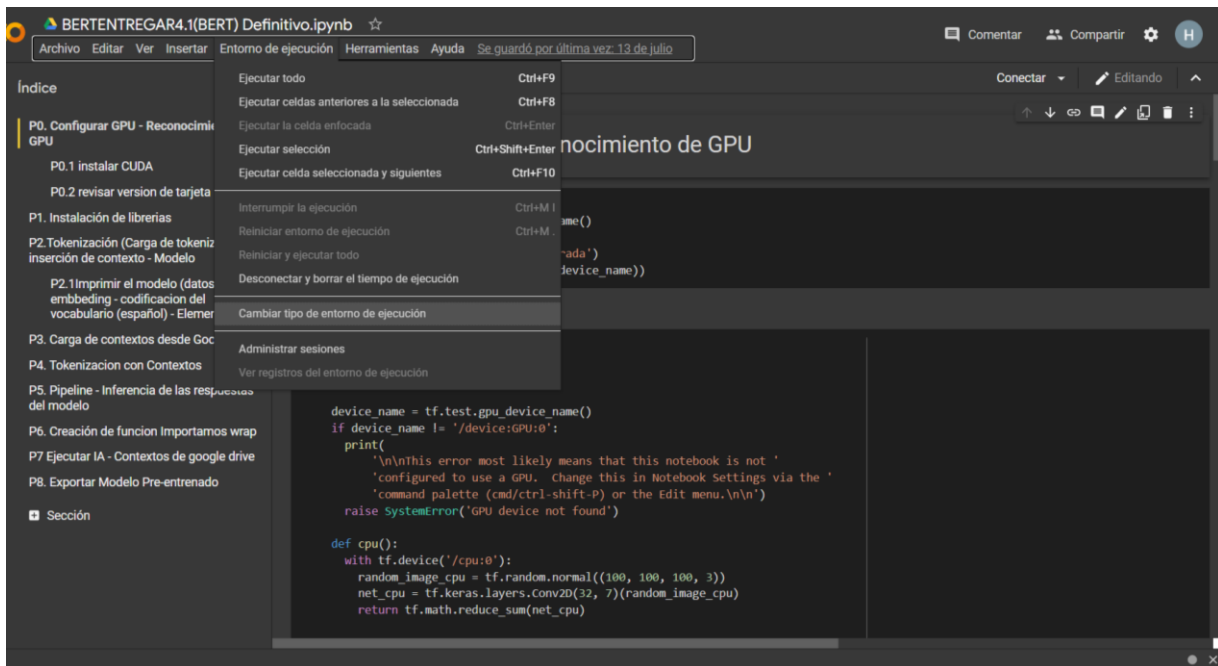


Figura 14. Configuración GPU Google Colab. Elaboración Propia

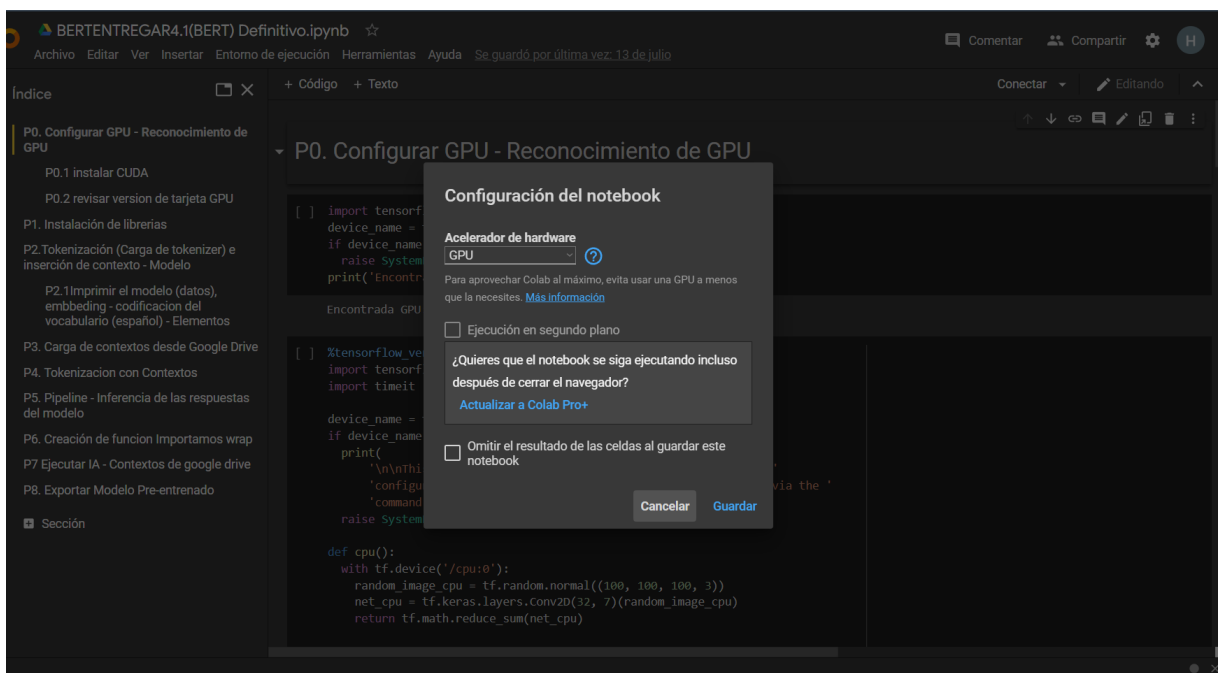
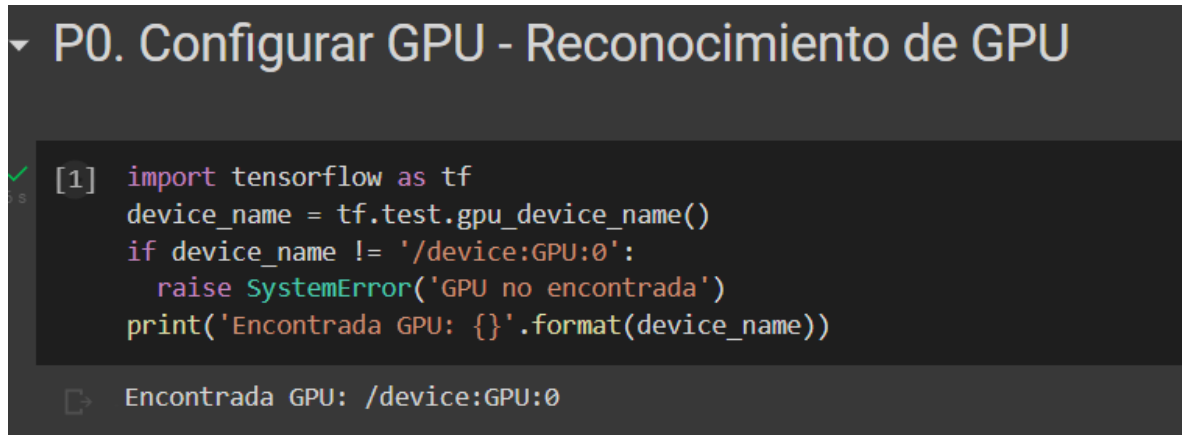


Figura 15. Configuración GPU Google Colab. Elaboración Propia

Para comprobar que la GPU proporciona una mayor eficiencia, primero se realizó un reconocimiento de la misma por medio de código con tensorflow para luego ejecutar una

función que genera una capa de red neuronal convolucional típica sobre una imagen aleatoria, que da como resultado una comparación respecto al tiempo de ejecución con cada entorno. En este resultado se valida que el entorno de ejecución con GPU es superior al entorno de ejecución con CPU.

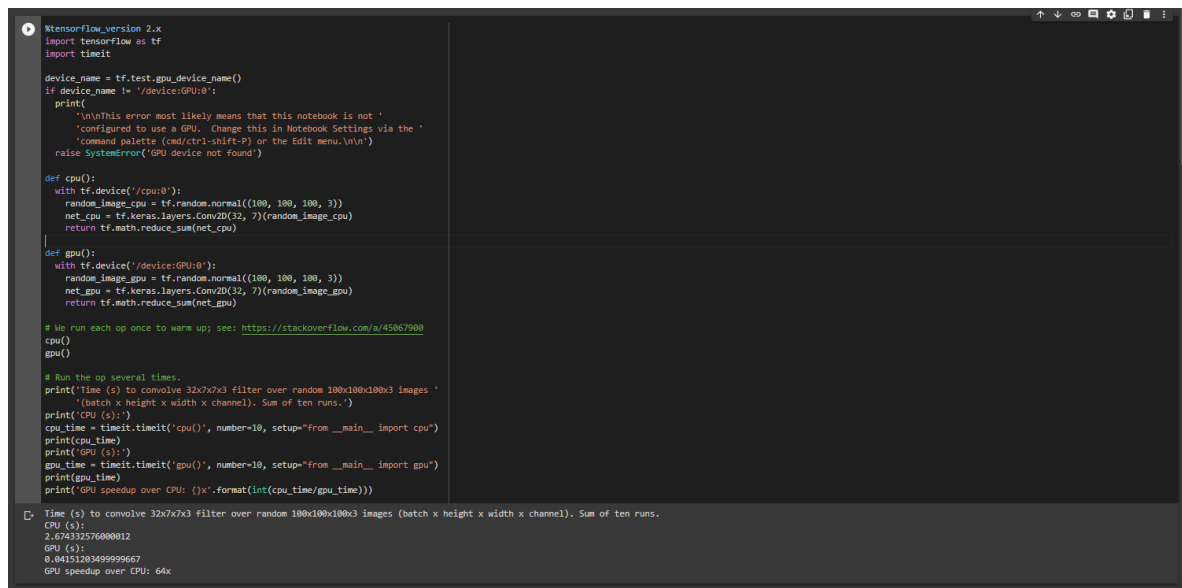


```
▼ P0. Configurar GPU - Reconocimiento de GPU

[1] import tensorflow as tf
    device_name = tf.test.gpu_device_name()
    if device_name != '/device:GPU:0':
        raise SystemError('GPU no encontrada')
    print('Encontrada GPU: {}'.format(device_name))

Encontrada GPU: /device:GPU:0
```

Figura 16. Reconocimiento de GPU. Elaboración Propia



```
Tensorflow_version 2.x
import tensorflow as tf
import timeit

device_name = tf.test.gpu_device_name()
if device_name != '/device:GPU:0':
    print(
        '\n\nThis error most likely means that this notebook is not '
        'configured to use a GPU. Change this in Notebook Settings via the '
        'command palette (cmd/ctrl-shift-P) or the Edit menu.\n\n')
    raise SystemError('GPU device not found')

def cpu():
    with tf.device('/cpu:0'):
        random_image_cpu = tf.random.normal([100, 100, 3])
        net_cpu = tf.keras.layers.Conv2D(32, 7)(random_image_cpu)
        return tf.math.reduce_sum(net_cpu)

def gpu():
    with tf.device('/device:GPU:0'):
        random_image_gpu = tf.random.normal([100, 100, 3])
        net_gpu = tf.keras.layers.Conv2D(32, 7)(random_image_gpu)
        return tf.math.reduce_sum(net_gpu)

# We run each op once to warm up; see: https://stackoverflow.com/a/45867908
cpu()
gpu()

# Run the op several times.
print('Time (s) to convolve 32x7x7x3 filter over random 100x100x100x3 images '
      '(batch x height x width x channel). Sum of ten runs.')
print('CPU (s):')
cpu_time = timeit.timeit('cpu()', number=10, setup='from __main__ import cpu')
print('CPU time')
print('GPU (s):')
gpu_time = timeit.timeit('gpu()', number=10, setup='from __main__ import gpu')
print('GPU time')
print('GPU speedup over CPU: {}'.format(int(cpu_time/gpu_time)))

Time (s) to convolve 32x7x7x3 filter over random 100x100x100x3 images (batch x height x width x channel). Sum of ten runs.
CPU (s):
2.674332576000012
GPU (s):
0.04151203499999667
GPU speedup over CPU: 64x
```

Figura 17. GPU vs CPU. Elaboración Propia

#### 7.4.1.2 Instalación de librería

En este paso se realizó la instalación de la librería Transformers por medio de sentencias de código Python, con lo cual se puede hacer uso de la arquitectura propuesta para este proyecto.

```
[ ] !pip install transformers

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting transformers
  Downloading transformers-4.20.1-py3-none-any.whl (4.4 MB)
    4.4 MB 30.1 MB/s
Collecting huggingface-hub<1.0,>=0.1.0
  Downloading huggingface_hub-0.8.1-py3-none-any.whl (101 kB)
    101 kB 12.9 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-packages (from transformers) (3.7.1)
Collecting tokenizers!=0.11.3,<0.13,>=0.11.1
  Downloading tokenizers-0.12.1-cp37m-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (6.6 MB)
    6.6 MB 61.0 MB/s
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.7/dist-packages (from transformers) (4.64.0)
Requirement already satisfied: importlib-metadata in /usr/local/lib/python3.7/dist-packages (from transformers) (4.12.0)
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from transformers) (2.23.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.7/dist-packages (from transformers) (21.3)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.7/dist-packages (from transformers) (1.21.6)
Requirement already satisfied: regex<=2019.12.17 in /usr/local/lib/python3.7/dist-packages (from transformers) (2022.6.2)
Collecting pyyaml>=5.1
  Downloading PyYAML-6.0-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux2010_x86_64.whl (596 kB)
    596 kB 66.7 MB/s
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.7/dist-packages (from huggingface-hub<1.0,>=0.1.0->transformers) (4.
Requirement already satisfied: pyrsingl=3.0.5,>=2.0.2 in /usr/local/lib/python3.7/dist-packages (from packaging>=20.0->transformers) (3.0.9)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata->transformers) (3.8.0)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests->transformers) (1.24.3)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests->transformers) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests->transformers) (2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests->transformers) (3.0.4)
Installing collected packages: pyyaml, tokenizers, huggingface-hub, transformers
  Attempting uninstall: pyyaml
    Found existing installation: PyYAML 3.13
    Uninstalling PyYAML-3.13:
      Successfully uninstalled PyYAML-3.13
Successfully installed huggingface-hub-0.8.1 pyyaml-6.0 tokenizers-0.12.1 transformers-4.20.1
```

Figura 18. Instalación de librería Transformer. Elaboración Propia

#### 7.4.1.3 Tokenización.

En este paso se realizó la importación del tokenizer, pipeline y la estructura del modelo (AutoModelQuestionAnswering) con el fin de tener las herramientas que facilitaron el desarrollo del proyecto; así mismo, se descargó el modelo pre – entrenado capaz de tokenizar (asignar valores a cada una de las palabras de un contexto) tanto palabras mayúsculas como minúsculas, con ello la librería transformer va a buscar en internet los archivos correspondientes al modelo y al tokenizer.

```
[ ] # Importar
from transformers import AutoTokenizer, AutoModelForQuestionAnswering, pipeline
the_model = 'mrms488/distill-bert-base-spanish-wmt-cased-finetuned-spa-squad2-es'
tokenizer = AutoTokenizer.from_pretrained(the_model, do_lower_case=False) #Modelo sea capaz de tokenizar tanto palabras mayúsculas como minúsculas
model = AutoModelForQuestionAnswering.from_pretrained(the_model) #Modelo pre-entrenado para responder preguntas, libreria transformer va a buscar en internet los archivos correspondientes al modelo y al tokenizer

Downloading: 100% 135/135 [00:00<00:00, 1.20kB/s]
Downloading: 100% 465/465 [00:00<00:00, 14.0kB/s]
Downloading: 100% 237k/237k [00:00<00:00, 284kB/s]
Downloading: 100% 112/112 [00:00<00:00, 3.59kB/s]
Downloading: 100% 419M/419M [00:07<00:00, 60.9MB/s]
```

Figura 19. Configuración GPU Google Colab. Elaboración Propia

```
[ ] print(model)

BertForQuestionAnswering(
  (bert): BertModel(
    (embeddings): BertEmbeddings(
      (word_embeddings): Embedding(31002, 768, padding_idx=1)
      (position_embeddings): Embedding(512, 768)
      (token_type_embeddings): Embedding(2, 768)
      (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
      (dropout): Dropout(p=0.1, inplace=False)
    )
    (encoder): BertEncoder(
      (layer): ModuleList(
        (0): BertLayer(
          (attention): BertAttention(
            (self): BertSelfAttention(
              (query): Linear(in_features=768, out_features=768, bias=True)
              (key): Linear(in_features=768, out_features=768, bias=True)
              (value): Linear(in_features=768, out_features=768, bias=True)
              (dropout): Dropout(p=0.1, inplace=False)
            )
            (output): BertSelfOutput(
              (dense): Linear(in_features=768, out_features=768, bias=True)
              (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
              (dropout): Dropout(p=0.1, inplace=False)
            )
          )
          (intermediate): BertIntermediate(
            (dense): Linear(in_features=768, out_features=3072, bias=True)
            (intermediate_act_fn): GELUActivation()
          )
          (output): BertOutput(
            (dense): Linear(in_features=3072, out_features=768, bias=True)
            (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
            (dropout): Dropout(p=0.1, inplace=False)
          )
        )
        (1): BertLayer(
          (attention): BertAttention(
```

Figura 20. Modelo Impreso. Elaboración Propia

#### 7.4.1.4 Carga de contextos

En este paso se realizó la conexión con Google Drive, encargado de almacenar los contextos (documentos) de cada uno de las cuatro áreas de las competencias ciudadanas; así mismo, cada documento es leído, codificado y almacenado en una variable para su posterior uso.

```
[ ] # -*- encoding: utf-8 -*-
from google.colab import files
from google.colab import drive
drive.mount('/content/gdrive')

data_path = '/content/gdrive/My Drive/cuentos_DATASETS/Contexto1DEFBOT.txt'
with open(data_path, 'rb') as f:
    text = f.read().decode(encoding='utf-8', errors='replace')
    print(text)

data_path = '/content/gdrive/My Drive/cuentos_DATASETS/Contexto2DEFBOT.txt'
with open(data_path, 'rb') as f:
    textDos = f.read().decode(encoding='utf-8', errors='replace')
    print(textDos)

data_path = '/content/gdrive/My Drive/cuentos_DATASETS/Contexto3DEFBOT.txt'
with open(data_path, 'rb') as f:
    textTres = f.read().decode(encoding='utf-8', errors='replace')
    print(textTres)

data_path = '/content/gdrive/My Drive/cuentos_DATASETS/Contexto4DEFBOT.txt'
with open(data_path, 'rb') as f:
    textCuatro = f.read().decode(encoding='utf-8', errors='replace')
    print(textCuatro)

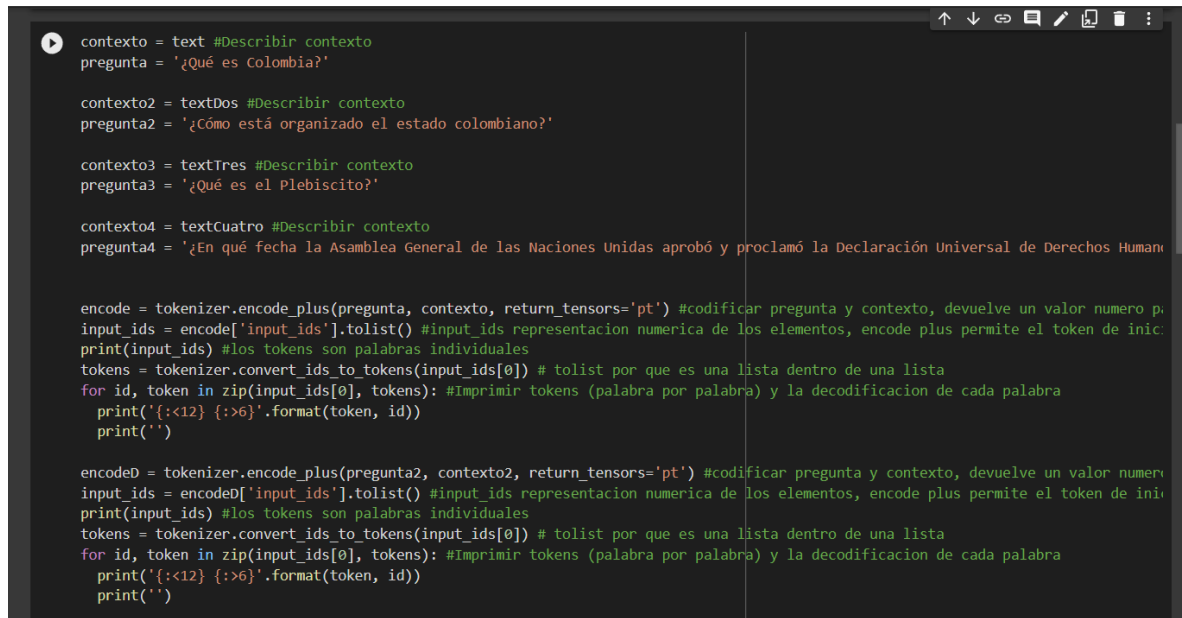
Mounted at /content/gdrive
----- CONSTITUCIÓN POLÍTICA DE COLOMBIA, PRINCIPIOS FUNDAMENTALES
Título 1 - De los principios fundamentales
Artículo 1
Artículo 10. Colombia es un Estado social de derecho, organizado en forma de República unitaria, descentralizada, con autonomía de sus entidades ter
Artículo 2
Artículo 20. Son fines esenciales del estado: servir a la comunidad, promover la prosperidad general y garantizar la efectividad de los principios, d
Artículo 3
Artículo 30. La soberanía reside exclusivamente en el pueblo, del cual emana el poder público. El pueblo la ejerce en forma directa o por medio de s
Artículo 4
Artículo 40. La Constitución es norma de normas. En todo caso de incompatibilidad entre la Constitución y la ley u otra norma jurídica, se aplicará
Artículo 5
```

Figura 21. Carga de Contextos. Elaboración Propia

#### 7.4.1.5 Tokenización de contextos

En este paso se asignaron los contextos y las preguntas a cada uno de los tokenizadores que al codificar cada uno de los parámetros mencionados devuelven un valor numérico a cada elemento del contexto y la pregunta. Así mismo, estas representaciones se almacenaron en la variable `inputs_ids` que junto con el encode permite generar el token de inicio y de separación.

Por último, se imprimen los tokens generados para cada contexto (palabra por palabra) y la decodificación de cada palabra.

A screenshot of a code editor with a dark background and light-colored text. The code is written in Python and demonstrates the tokenization of contexts and questions. It defines four context-question pairs, encodes them using a tokenizer, and prints the resulting tokens and their decodings. The code is as follows:

```
contexto = text #Describir contexto
pregunta = '¿Qué es Colombia?'

contexto2 = textDos #Describir contexto
pregunta2 = '¿Cómo está organizado el estado colombiano?'

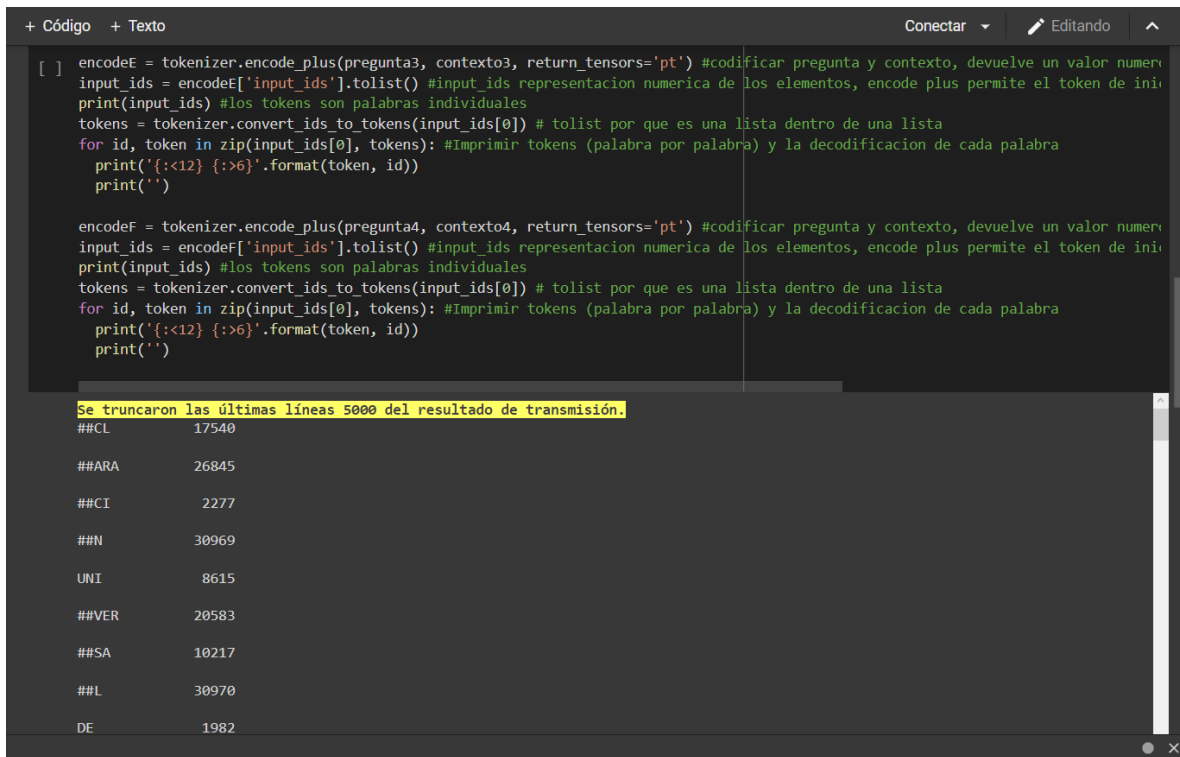
contexto3 = textTres #Describir contexto
pregunta3 = '¿Qué es el Plebiscito?'

contexto4 = textCuatro #Describir contexto
pregunta4 = '¿En qué fecha la Asamblea General de las Naciones Unidas aprobó y proclamó la Declaración Universal de Derechos Humanos?'

encode = tokenizer.encode_plus(pregunta, contexto, return_tensors='pt') #codificar pregunta y contexto, devuelve un valor numerico p
input_ids = encode['input_ids'].tolist() #input_ids representacion numerica de los elementos, encode plus permite el token de inicio
print(input_ids) #los tokens son palabras individuales
tokens = tokenizer.convert_ids_to_tokens(input_ids[0]) # tolist por que es una lista dentro de una lista
for id, token in zip(input_ids[0], tokens): #Imprimir tokens (palabra por palabra) y la decodificacion de cada palabra
    print('{:<12} {:>6}'.format(token, id))
print('')

encode0 = tokenizer.encode_plus(pregunta2, contexto2, return_tensors='pt') #codificar pregunta y contexto, devuelve un valor numerico
input_ids = encode0['input_ids'].tolist() #input_ids representacion numerica de los elementos, encode plus permite el token de inicio
print(input_ids) #los tokens son palabras individuales
tokens = tokenizer.convert_ids_to_tokens(input_ids[0]) # tolist por que es una lista dentro de una lista
for id, token in zip(input_ids[0], tokens): #Imprimir tokens (palabra por palabra) y la decodificacion de cada palabra
    print('{:<12} {:>6}'.format(token, id))
print('')
```

Figura 22. Tokenización de Contextos. Elaboración Propia



```
[ ] encodeE = tokenizer.encode_plus(pregunta3, contexto3, return_tensors='pt') #codificar pregunta y contexto, devuelve un valor numerico
input_ids = encodeE['input_ids'].tolist() #input_ids representacion numerica de los elementos, encode plus permite el token de inicio
print(input_ids) #los tokens son palabras individuales
tokens = tokenizer.convert_ids_to_tokens(input_ids[0]) # tolist por que es una lista dentro de una lista
for id, token in zip(input_ids[0], tokens): #Imprimir tokens (palabra por palabra) y la decodificacion de cada palabra
    print('{:<12} {:>6}'.format(token, id))
    print('')

encodeF = tokenizer.encode_plus(pregunta4, contexto4, return_tensors='pt') #codificar pregunta y contexto, devuelve un valor numerico
input_ids = encodeF['input_ids'].tolist() #input_ids representacion numerica de los elementos, encode plus permite el token de inicio
print(input_ids) #los tokens son palabras individuales
tokens = tokenizer.convert_ids_to_tokens(input_ids[0]) # tolist por que es una lista dentro de una lista
for id, token in zip(input_ids[0], tokens): #Imprimir tokens (palabra por palabra) y la decodificacion de cada palabra
    print('{:<12} {:>6}'.format(token, id))
    print('')
```

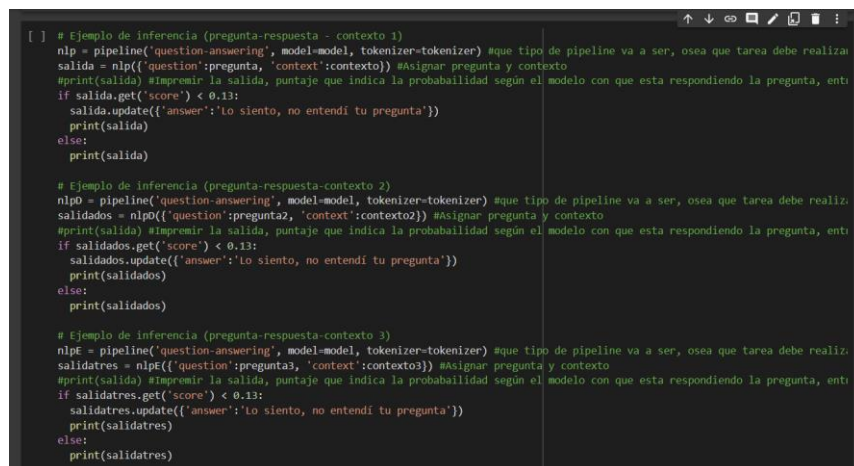
Se truncaron las últimas líneas 5000 del resultado de transmisión.

|       |       |
|-------|-------|
| ##CL  | 17540 |
| ##ARA | 26845 |
| ##CI  | 2277  |
| ##N   | 30969 |
| UNI   | 8615  |
| ##VER | 20583 |
| ##SA  | 10217 |
| ##L   | 30970 |
| DE    | 1982  |

Figura 23. Tokenización de Contextos. Elaboración Propia

#### 7.4.1.6 Pipeline

En este paso se realizó la implementación del pipeline como medio de inferencia para el modelo, en el proceso se relacionó el modelo con el tokenizer especificando la tarea que este último debe realizar (responder preguntas); así mismo, en la variable *salida* se asigna la pregunta y el contexto, teniendo en cuenta que en caso de que el porcentaje de acierto de la respuesta sea menor a un 13% se genere un mensaje de error.



```
[ ] # Ejemplo de inferencia (pregunta-respuesta - contexto 1)
nlp = pipeline('question-answering', model=model, tokenizer=tokenizer) #que tipo de pipeline va a ser, osea que tarea debe realizar
salida = nlp({'question':pregunta, 'context':contexto}) #Asignar pregunta y contexto
#print(salida) #Imprimir la salida, puntaje que indica la probabilidad según el modelo con que esta respondiendo la pregunta, entonces
if salida.get('score') < 0.13:
    salida.update({'answer':'Lo siento, no entendí tu pregunta'})
    print(salida)
else:
    print(salida)

# Ejemplo de inferencia (pregunta-respuesta-contexto 2)
nlp0 = pipeline('question-answering', model=model, tokenizer=tokenizer) #que tipo de pipeline va a ser, osea que tarea debe realizar
salidados = nlp0({'question':pregunta2, 'context':contexto2}) #Asignar pregunta y contexto
#print(salida) #Imprimir la salida, puntaje que indica la probabilidad según el modelo con que esta respondiendo la pregunta, entonces
if salidados.get('score') < 0.13:
    salidados.update({'answer':'Lo siento, no entendí tu pregunta'})
    print(salidados)
else:
    print(salidados)

# Ejemplo de inferencia (pregunta-respuesta-contexto 3)
nlpE = pipeline('question-answering', model=model, tokenizer=tokenizer) #que tipo de pipeline va a ser, osea que tarea debe realizar
salidatres = nlpE({'question':pregunta3, 'context':contexto3}) #Asignar pregunta y contexto
#print(salida) #Imprimir la salida, puntaje que indica la probabilidad según el modelo con que esta respondiendo la pregunta, entonces
if salidatres.get('score') < 0.13:
    salidatres.update({'answer':'Lo siento, no entendí tu pregunta'})
    print(salidatres)
else:
    print(salidatres)
```

Figura 24. Pipeline. Elaboración Propia

```
# Ejemplo de inferencia (pregunta-respuesta-contexto 4)
nlpF = pipeline('question-answering', model=model, tokenizer=tokenizer) #que tipo de pipeline va a ser, osea que tarea debe realiz.
salidacuatro = nlpF({'question':pregunta4, 'context':contexto4}) #Asignar pregunta y contexto
#print(salida) #Imprimir la salida, puntaje que indica la probabilidad según el modelo con que esta respondiendo la pregunta, ent
if salidacuatro.get('score') < 0.13:
    salidacuatro.update({'answer':'Lo siento, no entendí tu pregunta'})
    print(salidacuatro)
else:
    print(salidacuatro)

{'score': 0.3269334137439728, 'start': 170, 'end': 197, 'answer': 'un Estado social de derecho'}
{'score': 0.32726404070854187, 'start': 149, 'end': 162, 'answer': 'en dos partes'}
{'score': 0.504113495349884, 'start': 2980, 'end': 3053, 'answer': 'el pronunciamiento del pueblo convocado por el presidente de la
{'score': 0.6976566314697266, 'start': 77, 'end': 100, 'answer': '10 de diciembre de 1948'}
```

Figura 25. Pipeline. Elaboración Propia

#### 7.4.1.7 Creación de función

En este paso se creó la función que asocia el modelo, los contextos y el pipeline permitiendo así la interacción con el usuario, cumpliendo así con la dinámica de pregunta – respuesta por medio de la visualización de un menú y un input.

```
[ ] def pregunta_respuestaChatbot(model, contexto, contextodos, cintextotres, contextocuatro, nlp, nlpD, nlpE, nlpF): # funcion que so.
    from textwrap import wrap
    import numpy as np
    import warnings
    warnings.filterwarnings("ignore", category=np.VisibleDeprecationWarning)

    print('Hola, este sistema se realizó con el fin de tener una herramienta de consulta en el área de las competencias ciudadanas, '
    print('Puedes consultar sobre 4 contextos específicos de las competencias ciudadanas, si mi respuesta no es exacta, no te preocu
    print('¡Gracias por brindarme una oportunidad!')
    print('Para salir seleccione un contexto cualquiera y escriba la palabra Adios')

    # Imprimir menú
    print('Menú:')
    print('-----')
    print('Ingrese 1 - Conexto definición de las competencias ciudadanas')
    # Imprimir contexto
    print('\n'.join(wrap(contexto))) #Imprimir texto de referencia

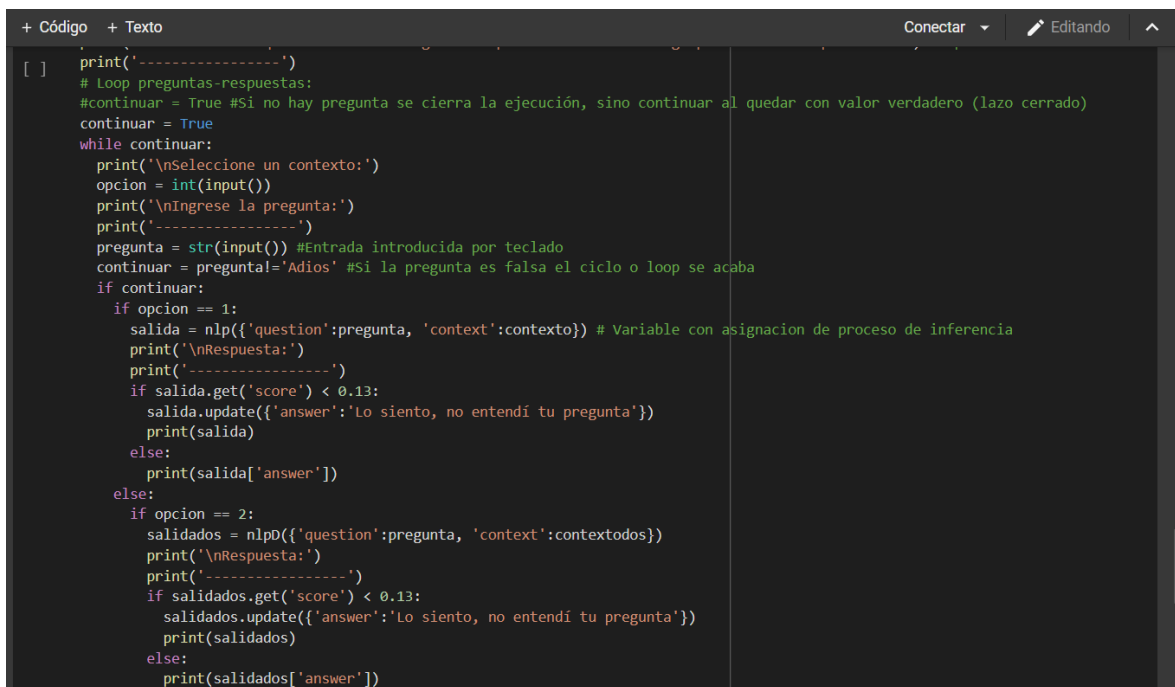
    print('Ingrese 2 - Conexto Normativo')
    print('\n'.join(wrap(contextodos))) #Imprimir texto de referencia

    print('Ingrese 3 - Los modelos de participación ciudadana')
    print('\n'.join(wrap(cintextotres))) #Imprimir texto de referencia

    print('Ingrese 4 - Declaración de los derechos humanos')
    print('\n'.join(wrap(contextocuatro))) #Imprimir texto de referencia

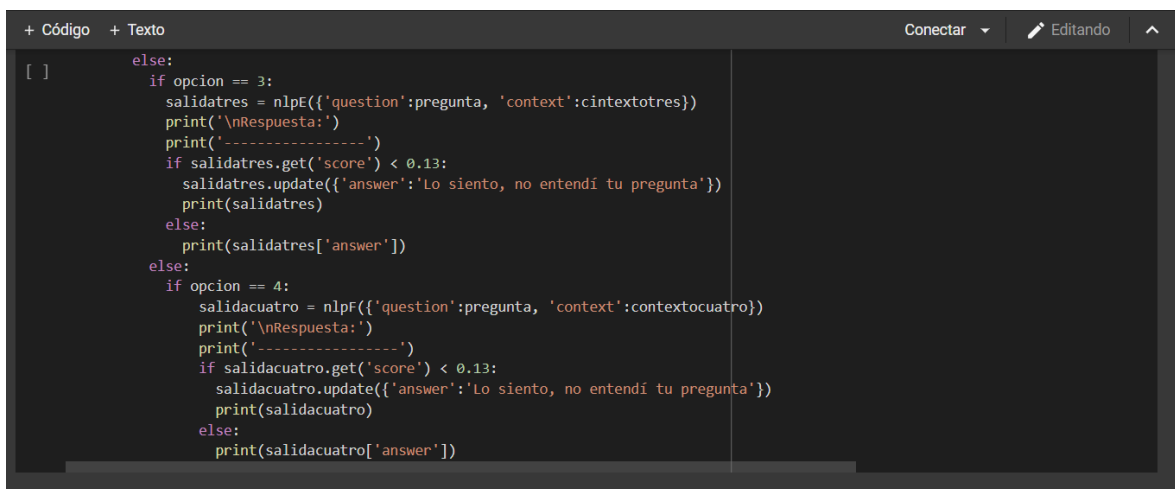
    print('\n')
    print('Seleccione cualquier contexto e ingrese la palabra "Adios" - Luego presione Enter para salir') #Imprimir menu
    print('-----')
```

Figura 26. Función para IA. Elaboración Propia



```
+ Código + Texto Conectar Editando
[ ] print('-----')
    # Loop preguntas-respuestas:
    #continuar = True #Si no hay pregunta se cierra la ejecución, sino continuar al quedar con valor verdadero (lazo cerrado)
    continuar = True
    while continuar:
        print('\nSeleccione un contexto:')
        opcion = int(input())
        print('\nIngrese la pregunta:')
        print('-----')
        pregunta = str(input()) #Entrada introducida por teclado
        continuar = pregunta!='Adios' #Si la pregunta es falsa el ciclo o loop se acaba
        if continuar:
            if opcion == 1:
                salida = nlp({'question':pregunta, 'context':contexto}) # Variable con asignacion de proceso de inferencia
                print('\nRespuesta:')
                print('-----')
                if salida.get('score') < 0.13:
                    salida.update({'answer':'Lo siento, no entendí tu pregunta'})
                    print(salida)
                else:
                    print(salida['answer'])
            else:
                if opcion == 2:
                    salidados = nlpD({'question':pregunta, 'context':contextodos})
                    print('\nRespuesta:')
                    print('-----')
                    if salidados.get('score') < 0.13:
                        salidados.update({'answer':'Lo siento, no entendí tu pregunta'})
                        print(salidados)
                    else:
                        print(salidados['answer'])
```

Figura 27. Función para IA. Elaboración Propia



```
+ Código + Texto Conectar Editando
[ ] else:
    if opcion == 3:
        salidatres = nlpE({'question':pregunta, 'context':cintextotres})
        print('\nRespuesta:')
        print('-----')
        if salidatres.get('score') < 0.13:
            salidatres.update({'answer':'Lo siento, no entendí tu pregunta'})
            print(salidatres)
        else:
            print(salidatres['answer'])
    else:
        if opcion == 4:
            salidacuatro = nlpF({'question':pregunta, 'context':contextocuatro})
            print('\nRespuesta:')
            print('-----')
            if salidacuatro.get('score') < 0.13:
                salidacuatro.update({'answer':'Lo siento, no entendí tu pregunta'})
                print(salidacuatro)
            else:
                print(salidacuatro['answer'])
```

Figura 28. Función para IA. Elaboración Propia

#### 7.4.1.8 Ejecución de IA

En este paso se ejecutó la función creada, evidenciando que se cumple con la dinámica de pregunta – respuesta pues se puede seleccionar un contexto, realizar una pregunta y obtener una respuesta.

```
[ ] pregunta_respuestaChatbot(model,text,textDos,textTres,textCuatro,nlp,nlpD,nlpE,nlpF)

Hola, este sistema se realizó con el fin de tener una herramienta de consulta en el área de las competencias ciudadanas, estoy aquí para ayudarte.
Puedes consultar sobre 4 contextos específicos de las competencias ciudadanas, si mi respuesta no es exacta, no te preocupes estoy aquí para ayudarte.
¡Gracias por brindarme una oportunidad!
Para salir seleccione un contexto cualquiera y escriba la palabra Adios
Menú:
-----
Ingrese 1 - Contexto definición de las competencias ciudadanas
-----
CONSTITUCIÓN POLITICA DE COLOMBIA,
PRINCIPIOS FUNDAMENTALES Título 1 - De los principios fundamentales
Artículo 1 Artículo 10. Colombia es un Estado social de derecho,
organizado en forma de República unitaria, descentralizada, con
autonomía de sus entidades territoriales, democrática, participativa y
pluralista, fundada en el respeto de la dignidad humana, en el trabajo
y la solidaridad de las personas que la integran y en la prevalencia
del interés general. Artículo 2 Artículo 20. Son fines esenciales
del Estado: servir a la comunidad, promover la prosperidad general y
garantizar la efectividad de los principios, derechos y deberes
consagrados en la Constitución; facilitar la participación de todos en
las decisiones que los afectan y en la vida económica, política,
administrativa y cultural de la Nación; defender la independencia
nacional, mantener la integridad territorial y asegurar la convivencia
pacífica y la vigencia de un orden justo. Las autoridades de la
República están instituidas para proteger a todas las personas
residentes en Colombia, en su vida, honra, bienes, creencias, y demás
derechos y libertades, y para asegurar el cumplimiento de los deberes
sociales del Estado y de los particulares. Artículo 3 Artículo 30.
La soberanía reside exclusivamente en el pueblo, del cual emana el
poder público. El pueblo la ejerce en forma directa o por medio de sus
representantes, en los términos que la Constitución establece.
Artículo 4 Artículo 40. La Constitución es norma de normas. En todo
caso de incompatibilidad entre la Constitución y la ley u otra norma
jurídica, se aplican las disposiciones constitucionales. Es deber de
los ciudadanos y de las autoridades cumplir con las disposiciones
constitucionales.
```

Figura 29. Ejecución IA. Elaboración Propia

#### 6.4.1.9. Exportación del modelo Pre - entrenado

```
P8. Exportar Modelo Pre-entrenado

[ ] model.save_pretrained('/content')
```

Figura 30. Exportación IA. Elaboración Propia

### 7.4.2. IA con contexto conversacional

En esta sección se repiten los pasos anteriores realizando los siguientes cambios:

#### 7.4.2.1 Instalación de librerías

Además de la instalación de la librería transformer, se realizó la instalación de la librería chatterbot y chatterbot\_corpus.

```
[ ] !pip install chatterbot

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: chatterbot in /usr/local/lib/python3.7/dist-packages (1.0.8)
Requirement already satisfied: sqlalchemy<1.4,>=1.3 in /usr/local/lib/python3.7/dist-packages (from chatterbot) (1.3.24)
Requirement already satisfied: pytz in /usr/local/lib/python3.7/dist-packages (from chatterbot) (2022.1)
Requirement already satisfied: mathparse<0.2,>=0.1 in /usr/local/lib/python3.7/dist-packages (from chatterbot) (0.1.2)
Requirement already satisfied: python-dateutil<2.9,>=2.8 in /usr/local/lib/python3.7/dist-packages (from chatterbot) (2.8.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-packages (from python-dateutil<2.9,>=2.8->chatterbot) (1.16.0)

[ ] !pip install chatterbot_corpus

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: chatterbot_corpus in /usr/local/lib/python3.7/dist-packages (1.2.0)
Collecting PyYAML<4.0,>=3.12
  Using cached PyYAML-3.13-cp37m-linux_x86_64.whl
Installing collected packages: PyYAML
  Attempting uninstall: PyYAML
    Found existing installation: PyYAML 6.0
    Uninstalling PyYAML-6.0:
      Successfully uninstalled PyYAML-6.0
  Successfully uninstalled PyYAML-6.0
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.
transformers 4.20.1 requires pyyaml>=5.1, but you have pyyaml 3.13 which is incompatible.
huggingface-hub 0.8.1 requires pyyaml>=5.1, but you have pyyaml 3.13 which is incompatible.
Successfully installed PyYAML-3.13
```

Figura 31. Instalación Librerías Chatterbot y chatterbot\_corpus. Elaboración Propia

#### 7.4.2.2 Creación de funciones

Se agrega una función que permitió validar el funcionamiento de la librería chatterbot, así mismo, se modifica la función que ya se tenía para los contextos de competencias ciudadanas agregando el aspecto conversacional.

```
▼ P6.1 Test Funcion para conversación

[ ] from chatterbot import ChatBot
from chatterbot.trainers import ChatterBotCorpusTrainer

chat = ChatBot('cctm')
trainer = ChatterBotCorpusTrainer(chat)
trainer.train("chatterbot_corpus.spanish.greetings")
continuar= True
while continuar:
    peticion = input('Tú: ')
    continuar = peticion!='' #Si la pregunta es falsa el ciclo o loop se acaba
    if continuar:
        respuesta = chat.get_response(peticion)
        print('Bot: ',respuesta)

Training greetings.yml: [#####] 100%
Tú: Hola
Bot: Hola
Tú: Como estas?
Bot: Supongo que bien.
Tú:
```

Figura 32. Función Chatbot. Elaboración Propia

```
[ ] from textwrap import wrap
from chatterbot.trainers import ListTrainer
from chatterbot import ChatBot

def pregunta_respuestaChatbot(model, contexto, contextodos, cintextotres, contextocuatro, nlp, nlpD, nlpE, nlpF): # funcion que solicita un modelo, un contexto, y pipeline

    import numpy as np
    import warnings
    warnings.filterwarnings("ignore", category=np.VisibleDeprecationWarning)
    from chatterbot import ChatBot
    from chatterbot.trainers import ChatterBotCorpusTrainer
    chat = ChatBot('cctmx')
    trainer = ChatterBotCorpusTrainer(chat)
    trainer.train("chatterbot.corpus.spanish.greetings")
    continuar = True
    print('Hola, este sistema se realizó con el fin de tener una herramienta de consulta en el área de las competencias ciudadanas, estoy aquí para ayudarte')
    print('Puedes consultar sobre 4 contextos específicos de las competencias ciudadanas, si mi respuesta no es exacta, no te preocupes estoy aprendiendo con el tiempo')
    print('¡Gracias por brindarme una oportunidad!')
    print('Para ingresar a los contextos ingresa la siguiente frase: Tengo una pregunta')
    print('Para terminar la conversacion escribe la palabra Adios')
    print('Para salir selecciona un contexto cualquiera y escribe la palabra Adios')
    while continuar:
        petición = input('Tú: ')
        continuar = petición != 'Adios' #Si la pregunta es falsa el ciclo o loop se acaba
        if continuar:
            respuesta = chat.get_response(petición)
            if petición == 'Tengo una pregunta':
                # Imprimir menú
                print('Menú:')
                print('-----')
                print('Ingresa 1 - Contexto definición de las competencias ciudadanas')
                # Imprimir contexto
                print('\n'.join(wrap(contexto))) #Imprimir texto de referencia

                print('Ingresa 2 - Contexto Normativo')
                print('\n'.join(wrap(contextodos))) #Imprimir texto de referencia

                print('Ingresa 3 - Los modelos de participación ciudadana')
                print('\n'.join(wrap(cintextotres))) #Imprimir texto de referencia

                print('Ingresa 4 - Declaración de los derechos humanos')
                print('\n'.join(wrap(contextocuatro))) #Imprimir texto de referencia
```

Figura 33. Función Chatbot. Elaboración Propia

```
+ Código + Texto Conectar Editando

[ ] # Loop preguntas-respuestas:
#continuar = True #Si no hay pregunta se cierra la ejecución, sino continuar al quedar con valor verdadero (lazo cerrado)
continuar = True
while continuar:
    print('\nSelecciona un contexto:')
    print('-----')
    opcion = int(input())
    print('\nIngresa la pregunta:')
    print('-----')
    pregunta = str(input()) #Entrada introducida por teclado
    continuar = pregunta != 'Adios' #Si la pregunta es falsa el ciclo o loop se acaba
    if continuar:
        if opcion == 1:
            salida = nlp({'question':pregunta, 'context':contexto}) # Variable con asignacion de proceso de inferencia
            print('\nRespuesta:')
            print('-----')
            if salida.get('score') < 0.13:
                salida.update({'answer':'Lo siento, no entendí tu pregunta'})
                print(salida['answer'])
                print(salida['score'])
            else:
                print(salida)
        else:
            if opcion == 2:
                salidados = nlpD({'question':pregunta, 'context':contextodos})
                print('\nRespuesta:')
                print('-----')
                if salidados.get('score') < 0.13:
                    salidados.update({'answer':'Lo siento, no entendí tu pregunta'})
                    print(salidados['answer'])
                    print(salidados['score'])
                else:
                    print(salidados)
            if opcion == 3:
                salidatres = nlpE({'question':pregunta, 'context':cintextotres})
                print('\nRespuesta:')
                print('-----')
                if salidatres.get('score') < 0.13:
                    salidatres.update({'answer':'Lo siento, no entendí tu pregunta'})
                    print(salidatres['answer'])
                    print(salidatres['score'])
                else:
                    print(salidatres)
```

Figura 34. Función Chatbot. Elaboración Propia

```
print(salidaTres)
if opcion == 4:
    salidaCuatro = nlpf({'question':pregunta, 'context':contextocuatro})
    print('\nRespuesta:')
    print('-----')
    if salidaCuatro.get('score') < 0.13:
        salidaCuatro.update({'answer':'Lo siento, no entendí tu pregunta'})
        print(salidaCuatro['answer'])
        print(salidaCuatro['score'])
    else:
        print(salidaCuatro)
else:
    print('Bot: ',respuesta)
```

Figura 35. Función Chatbot. Elaboración Propia

### 7.4.2.3 Ejecución de IA

En este paso se ejecutó la función creada, evidenciando que se cumple con la dinámica de pregunta – respuesta y el aspecto conversacional pues se puede seleccionar un contexto, realizar una pregunta y obtener una respuesta.

```
[ ] pregunta_respuestaChatbot(model, text, textDos, textTres, textCuatro, nlp, nlpD, nlpE, nlpF)

Training greetings.yml: [#####] 100%
Hola, este sistema se realizó con el fin de tener una herramienta de consulta en el área de las competencias ciudadanas, estoy aquí para ayudarte
Puedes consultar sobre 4 contextos específicos de las competencias ciudadanas, si mi respuesta no es exacta, no te preocupes estoy aprendiendo con el tiempo
¡Gracias por brindarme una oportunidad!
Para ingresar a los contenido ingresa la siguiente frase: Tengo una pregunta
Para terminar la conversacion escribe la palabra Adios
Para salir seleccione un contexto cualquiera y escriba la palabra Adios
Tú: Adios
```

Figura 36. Ejecución IA. Elaboración Propia

## 7.5. Fase 5. Pruebas

En esta sección se anexan las capturas de pantalla de las pruebas realizadas a la IA, para este proceso de se formulan una serie de preguntas relacionadas con los contextos 1, 2, 3 y 4 del área de las competencias ciudadanas, comprobando así que la respuesta es acorde con los contenidos de los textos. Cabe resaltar que si la pregunta no está relacionada con el contexto seleccionado no se obtendrá ninguna respuesta.

```
Seleccione cualquier contexto e ingrese la palabra "Adios" - Luego presione Enter para salir
-----

Seleccione un contexto:
1

Ingrese la pregunta:
-----
¿Qué es la constitución?

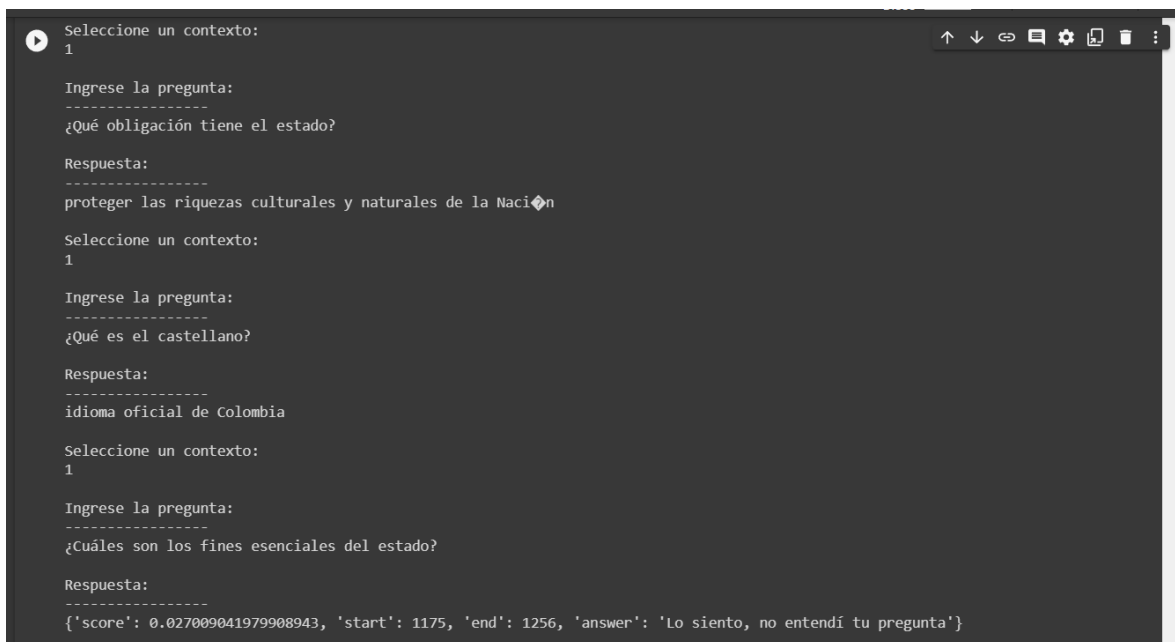
Respuesta:
-----
norma de normas

Seleccione un contexto:
1

Ingrese la pregunta:
-----
¿Qué es Colombia?

Respuesta:
-----
un Estado social de derecho
```

Figura 37. Pruebas a IA. Elaboración Propia



```
Seleccione un contexto:
1

Ingrese la pregunta:
-----
¿Qué obligación tiene el estado?

Respuesta:
-----
proteger las riquezas culturales y naturales de la Nación

Seleccione un contexto:
1

Ingrese la pregunta:
-----
¿Qué es el castellano?

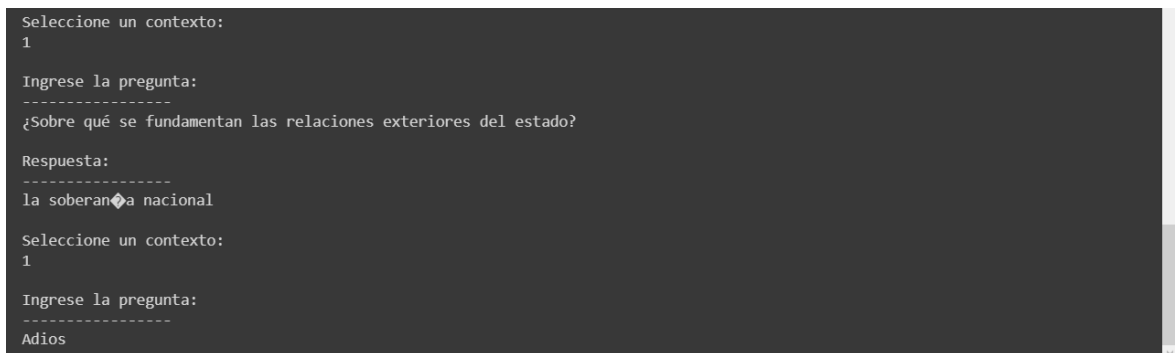
Respuesta:
-----
idioma oficial de Colombia

Seleccione un contexto:
1

Ingrese la pregunta:
-----
¿Cuáles son los fines esenciales del estado?

Respuesta:
-----
{'score': 0.027009041979908943, 'start': 1175, 'end': 1256, 'answer': 'Lo siento, no entendí tu pregunta'}
```

Figura 38. Pruebas a IA. Elaboración Propia



```
Seleccione un contexto:
1

Ingrese la pregunta:
-----
¿Sobre qué se fundamentan las relaciones exteriores del estado?

Respuesta:
-----
la soberanía nacional

Seleccione un contexto:
1

Ingrese la pregunta:
-----
Adios
```

Figura 39. Pruebas a IA. Elaboración Propia

```
Seleccione cualquier contexto e ingrese la palabra "Adios" - Luego presione Enter para salir
-----

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿En cuántas partes está organizado el estado colombiano?

Respuesta:
-----
dos partes

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Cuáles son las ramas del poder público?

Respuesta:
-----
{'score': 0.011041567660868168, 'start': 3097, 'end': 3125, 'answer': 'Lo siento, no entendí tu pregunta'}

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Cuáles son las ramas del poder?

Respuesta:
-----
la rama legislativa, la rama ejecutiva y la rama judicial
```

Figura 40. Pruebas a IA. Elaboración Propia

```
Seleccione cualquier contexto e ingrese la palabra "Adios" - Luego presione Enter para salir
-----

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Cómo está conformada la rama legislativa?

Respuesta:
-----
Congreso de la República

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Cómo está conformada la rama judicial?

Respuesta:
-----
{'score': 0.038585416972637177, 'start': 3100, 'end': 3125, 'answer': 'Lo siento, no entendí tu pregunta'}

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Cómo esta conformada la rema ejecutiva?

Respuesta:
-----
{'score': 0.06377629935741425, 'start': 3256, 'end': 3262, 'answer': 'Lo siento, no entendí tu pregunta'}
```

Figura 41. Pruebas a IA. Elaboración Propia

```
Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Quienes conforman la rama ejecutiva?

Respuesta:
-----
Corte Suprema de Justicia, Consejo de Estado, La Corte Constitucional

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Quienes conforman la rama judicial?

Respuesta:
-----
Corte Suprema de Justicia, Consejo de Estado, La Corte Constitucional

Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Quien conforma la rama ejecutiva?

Respuesta:
-----
{'score': 0.12734703719615936, 'start': 3100, 'end': 3125, 'answer': 'Lo siento, no entendí tu pregunta'}
```

Figura 42. Pruebas a IA. Elaboración Propia

```
Seleccione un contexto:
2

Ingrese la pregunta:
-----
¿Quiénes se encuentran dentro de la rama ejecutiva?

Respuesta:
-----
Orden Nacional, Presidencia de la República, Vicepresidencia de la República

Seleccione un contexto:
2

Ingrese la pregunta:
-----
Adios
```

Figura 43. Pruebas a IA. Elaboración Propia

```
Seleccione cualquier contexto e ingrese la palabra "Adios" - Luego presione Enter para salir
-----

Seleccione un contexto:
3

Ingrese la pregunta:
-----
¿Qué permite el modelo democrático de Colombia?

Respuesta:
-----
participar en las decisiones que definan el rumbo del país

Seleccione un contexto:
3

Ingrese la pregunta:
-----
¿Qué expone la ley 134 de 1994?

Respuesta:
-----
Artículo 2
```

Figura 44. Pruebas a IA. Elaboración Propia

```
Ingrese la pregunta:
-----
¿Qué expone la ley 134 de 1994 en su artículo 2?

Respuesta:
-----
"La iniciativa popular legislativa y normativa ante las corporaciones públicas

Seleccione un contexto:
3

Ingrese la pregunta:
-----
¿Qué es el referendo?

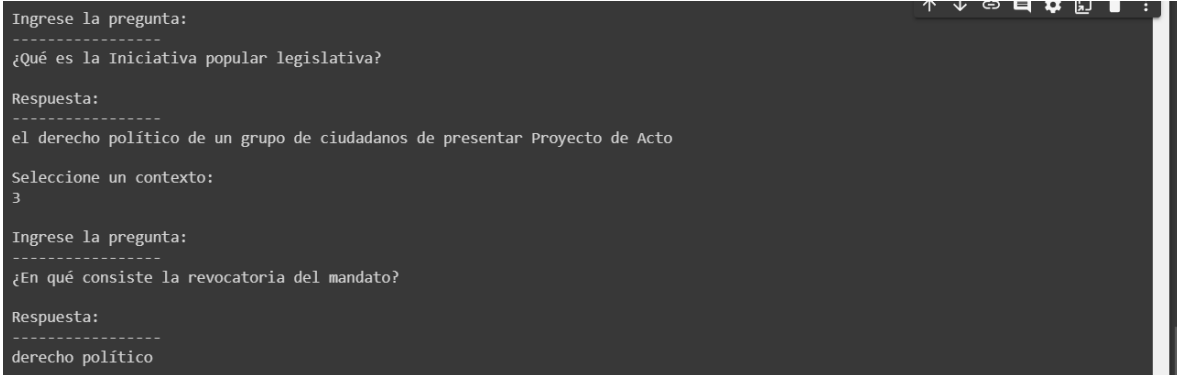
Respuesta:
-----
participación del pueblo en ejercicio de su soberanía

Seleccione un contexto:
3

Ingrese la pregunta:
-----
¿Qué es el Plebiscito?

Respuesta:
-----
el pronunciamiento del pueblo convocado por el presidente de la República
```

Figura 45. Pruebas a IA. Elaboración Propia



```
Ingrese la pregunta:
-----
¿Qué es la Iniciativa popular legislativa?

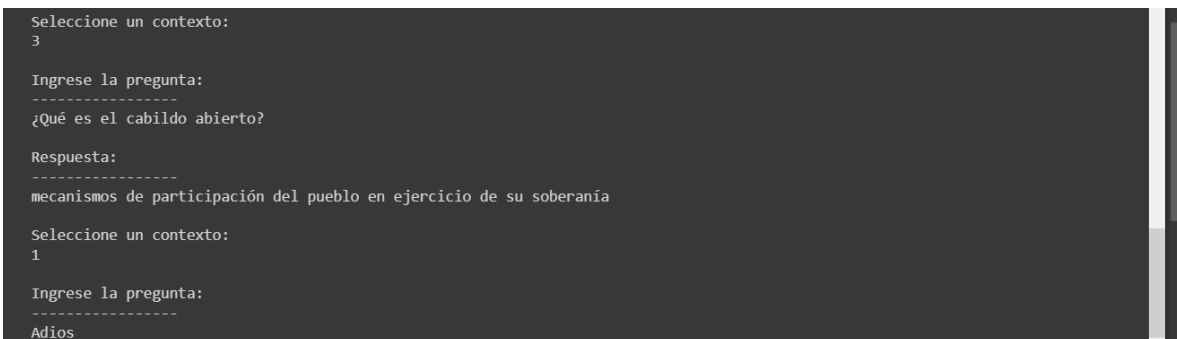
Respuesta:
-----
el derecho político de un grupo de ciudadanos de presentar Proyecto de Acto

Seleccione un contexto:
3

Ingrese la pregunta:
-----
¿En qué consiste la revocatoria del mandato?

Respuesta:
-----
derecho político
```

Figura 46. Pruebas a IA. Elaboración Propia



```
Seleccione un contexto:
3

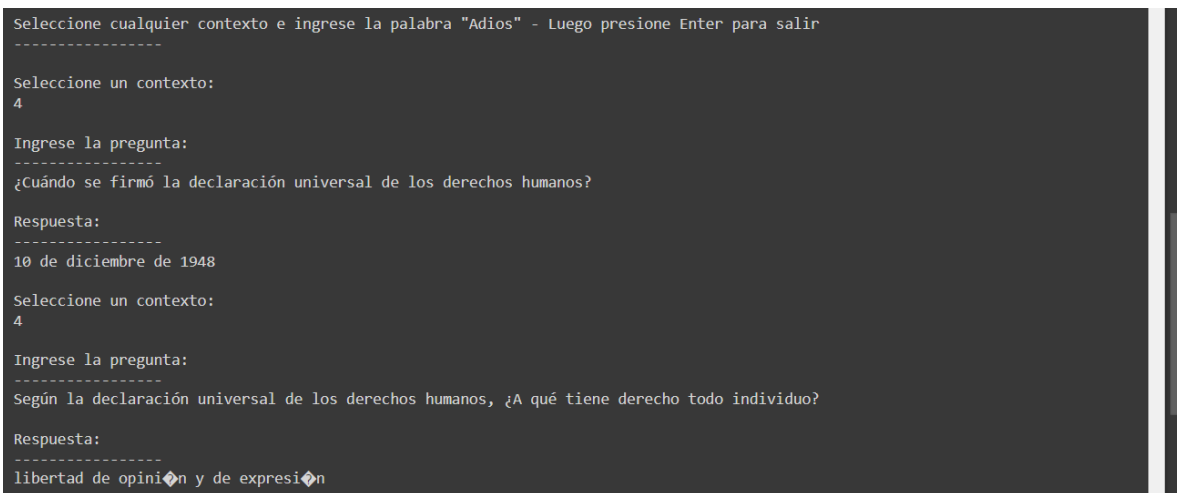
Ingrese la pregunta:
-----
¿Qué es el cabildo abierto?

Respuesta:
-----
mecanismos de participación del pueblo en ejercicio de su soberanía

Seleccione un contexto:
1

Ingrese la pregunta:
-----
Adios
```

Figura 47. Pruebas a IA. Elaboración Propia



```
Seleccione cualquier contexto e ingrese la palabra "Adios" - Luego presione Enter para salir
-----

Seleccione un contexto:
4

Ingrese la pregunta:
-----
¿Cuándo se firmó la declaración universal de los derechos humanos?

Respuesta:
-----
10 de diciembre de 1948

Seleccione un contexto:
4

Ingrese la pregunta:
-----
Según la declaración universal de los derechos humanos, ¿A qué tiene derecho todo individuo?

Respuesta:
-----
libertad de opinión y de expresión
```

Figura 48. Pruebas a IA. Elaboración Propia

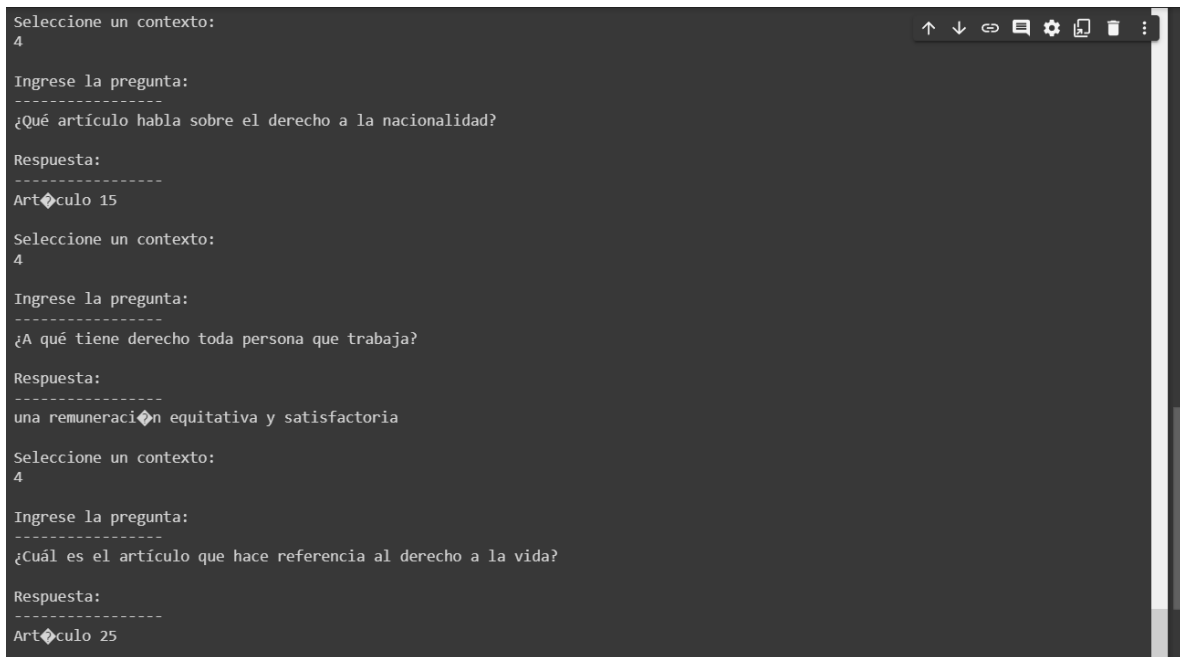


Figura 49. Pruebas a IA. Elaboración Propia

## **8. ANÁLISIS Y DISCUSIÓN DE RESULTADOS**

Así como el ser humano depende de la información que tenga para nutrir la respuesta que pueda proporcionar a una pregunta, la máquina al ser una imitación de las funciones lógico-matemáticas del ser humano depende de la información que tenga como base, que en este caso son los corpus de cada una de las áreas de las competencias ciudadanas, y la claridad que tenga la estructura de la pregunta. Dicho lo anterior, los resultados obtenidos en la sección correspondiente a las pruebas demuestran que el software desarrollado es capaz de proporcionar a los usuarios respuestas a las preguntas realizadas de acuerdo a cada corpus de las competencias ciudadanas.

Con el trabajo realizado se comprobó que la máquina a pesar de no poseer conciencia propia, que es el mayor sueño del hombre, es capaz de replicar los procesos analíticos del ser humano, en este caso, el procesamiento de textos, no obstante, así como el ser humano es imperfecto la máquina también lo es pues su creador no es Dios sino el hombre, razón por la cual al depender de factores externos como la calidad de la información y la exactitud de las preguntas la hace susceptible a errores.

Cabe resaltar que para esta versión se trabajó con los corpus versión 1.0 y conforme se avanzó en la fase de pruebas se realizaron cambios estructurales al banco de preguntas para obtener las respuestas deseadas.

## 9. CONCLUSIONES

- El uso de herramientas TIC en la educación motivan y fortalecen el proceso de enseñanza – aprendizaje, razón por la cual su uso e implementación promueven el análisis crítico de los métodos de enseñanza propuesto para las instituciones educativas.
- El machine learning y el PLN permiten que la máquina por medio de la inteligencia artificial copie funciones propias del ser humano, como en el caso de la comprensión lectora, ejecutando así actividades como dar respuestas a las preguntas formuladas por el usuario.
- Al realizar el proceso de desarrollo se verifica que se da cumplimiento a cada uno de los objetivos propuestos en este proyecto por medio de la entrega de los archivos de texto relacionados con cada una de las áreas de las competencias ciudadanas junto con el repositorio en drive, los archivos Python con la codificación del backend y las evidencias fotográficas de las pruebas realizadas junto con el archivo de texto que se genera con cada ejecución de la IA.

## 10. REFERENCIAS

- Akhter, M. J. (2020 ). Automatic Detection of Offensive Language for Urdu and Roman Urdu. *IEEE Access* 8,9094176, pp. 91213-91226.
- Alahmadi, D. B. (2020). Ensuring inclusion and diversity in research and research output: A case for a language-sensitive NLP crowdsourcing platform. *Applied Sciences (Switzerland)* 10(18),6216.
- Augusto Cortez Vásquez, H. V. (2009). Procesamiento de lenguaje natural. *Revista de Ingeniería de Sistemas e Informática* vol. 6, N.º 2, 10.
- Baume, G. L. (2021). *Breve introducción a Google Colab*. Obtenido de Breve introducción a Google Colab: <http://fcaglp.unlp.edu.ar/~gbaume/grupo/Publicaciones/Apuntes/GoogleColab.pdf>
- Bolaños, X. (30 de Julio de 2020). *Procesamiento del Lenguaje Natural con Machine Learning*. Obtenido de Procesamiento del Lenguaje Natural con Machine Learning: <https://www.encora.com/es/blog/procesamiento-del-lenguaje-natural-con-machine-learning#:~:text=¿En%20qué%20consiste%20el%20Procesamiento,sin%20embar go%20también%20existen%20excepciones.>
- Cañón, P. A. (2007). PROCESAMIENTO DEL LENGUAJE NATURAL EN LA RECUPERACIÓN DE LA INFORMACIÓN. Colombia.
- Cardellino, F. (30 de Marzo de 2021). *freecodecamp*. Obtenido de freecodecamp: <https://www.freecodecamp.org/espanol/news/tutorial-de-google-bert-para-pnl-con-aprendizaje-automatico/>
- Conocimiento, I. d. (s.f.). *Procesamiento del Lenguaje Natural Las máquinas analizan nuestro lenguaje en texto*. Obtenido de Procesamiento del Lenguaje Natural Las máquinas analizan nuestro lenguaje en texto: <https://www.iic.uam.es/inteligencia-artificial/procesamiento-del-lenguaje-natural/>
- González, I. C. (16 de Julio de 2020). *"Procesamiento del lenguaje natural con BERT: Análisis de sentimientos en tuits"*. Obtenido de "Procesamiento del lenguaje natural con BERT: Análisis de sentimientos en tuits": [https://e-archivo.uc3m.es/bitstream/handle/10016/32792/TFG\\_Iago\\_Collarte\\_Gonzalez.pdf?sequence=1&isAllowed=y](https://e-archivo.uc3m.es/bitstream/handle/10016/32792/TFG_Iago_Collarte_Gonzalez.pdf?sequence=1&isAllowed=y)

- Group, I. (18 de Marzo de 2022). *Transformers: explicamos el modelo detrás de GPT-3, BERT y T5*. Obtenido de Transformers: explicamos el modelo detrás de GPT-3, BERT y T5: <https://www.ibidemgroup.com/edu/traduccion-automatica-gpt3-bert-t5/>
- Guío Ávila Henry Alfonso, C. G. (2021). Aplicación de machine learning al proceso de enseñanza - aprendizaje de competencias ciudadanas en educación media del Colegio de Boyacá, usando procesamiento de lenguaje natural.
- J., H. M. (2013). Aplicaciones de Procesamiento de Lenguaje. *Revista Politécnica*, 10. Obtenido de Aplicaciones de Procesamiento de Lenguaje: <https://core.ac.uk/download/pdf/18586869.pdf>
- Lakshmanarao, A. S. (2019). An effecient fake news detection system using machine learning. *International Journal of Innovative Technology and Exploring Engineering* 8(10), pp. 3125-3129.
- Na8. (27 de Diciembre de 2018). *Procesamiento del Lenguaje Natural (NLP)*. Obtenido de Procesamiento del Lenguaje Natural (NLP): <https://www.aprendemachinelearning.com/procesamiento-del-lenguaje-natural-nlp/>
- Pressman, R. (2010). *Ingeniería del Software. Un enfoque práctico*. USA: McGraw-Hill.
- Quijano-Sánchez, L. L.-C.-C. (2018). Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the pólice. *Knowledge-Based Systems* 149, pp. 155-168, pp. 155-168.
- Quintero, S. D. (1 de Abril de 2022). *Aprende Python*. Obtenido de Aprende Python: [https://aprendepython.es/\\_downloads/907b5202c1466977a8d6bd3a2641453f/aprendepython.pdf](https://aprendepython.es/_downloads/907b5202c1466977a8d6bd3a2641453f/aprendepython.pdf)
- RAMÍREZ, D. H. (2018). *EL MACHINE LEARNING A TRAVÉS DE LOS TIEMPOS, Y LOS APORTES A LA HUMANIDAD*. Obtenido de EL MACHINE LEARNING A TRAVÉS DE LOS TIEMPOS, Y LOS APORTES A LA HUMANIDAD: <https://repository.unilibre.edu.co/bitstream/handle/10901/17289/EL%20MACHINE%20LEARNING.pdf?sequence=1&isAllowed=y>
- Rodríguez Mena García, M., & García Montero, I. (2003). El aprendizaje para el cambio. Papel de la educación. *Convergencia*, 317-335.
- Thangaraj, M. S. (2018). Text classification techniques: A literature review. *Interdisciplinary Journal of Information, Knowledge, and Management* 13, pp. 117-135.

## 11. ANEXOS

### ANEXO 1. Objetivos vs Productos:

| Tabla de Objetivos Vs Producto                                                                                                                                                                                                                                                                                              |                                                                                                     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| Objetivo                                                                                                                                                                                                                                                                                                                    | Producto                                                                                            |
| Generar el corpus con el cual la red neuronal PLN soporte la respuesta dada al estudiante, mediante la consulta y selección de documentación relevante sobre competencias ciudadanas en Colombia y reuniones con expertos en el área, para poder implementar el sistema de acuerdo a los objetivos de aprendizaje deseados. | Repositorio de Drive.<br>Contextos (Archivos txt).                                                  |
| Implementar el modelo pre-entrenado con base en el corpus generado, por medio de sentencias de código Python para el análisis semántico del proceso de pregunta-respuesta entre el sistema de PLN y el estudiante.                                                                                                          | Archivo Python.                                                                                     |
| Codificar el backend de la red PLN de acuerdo a la implementación del modelo pre-entrenado, por medio de la herramienta web Google Colaboratory y el lenguaje de programación Python para establecer los procedimientos y rutinas de configuración.                                                                         | Archivo Python.                                                                                     |
| Validar el desarrollo del sistema PLN propuesto por medio de una serie de pruebas con el fin de comprobar su correcto funcionamiento.                                                                                                                                                                                       | Capturas de pantalla de este documento<br>Archivo txt que se genera con cada ejecución del sistema. |

- Enlace Videos:

- Fotografías:

```
harvey@harvey-HP-Pavilion-Laptop-15-cw0003la: ~/Escritorio/Maestria
1149
1149
1149
1149
1149
1149
1149
1149
1149
1149
1149
7300
21722
8039
12967
30960
1982
[Link] 3
CT 13769
#AUD 11233
#ADA 12667
#MA 14408
EL 1162
Modelo 4598
democrático 14014
de 1008
Colombia 6634
le 1232

[REPLY] harvey@harvey-HP-Pavilion-Laptop-15-cw0003la: ~/Escritorio/Maestria
Q [ ] - o x

en 1036
esta 1359
Declar 5526
#acn 29231
e 1089
[REP] 5

{'score': 0.3269328474998474, 'start': 179, 'end': 197, 'answer': 'un Estado social de derecho'}
{'score': 0.32726413611559893, 'start': 149, 'end': 162, 'answer': 'en dos partes'}
{'score': 0.5841136741638184, 'start': 2989, 'end': 3053, 'answer': 'el pronunciamiento del pueblo convocado por el presidente de la República'}
{'score': 0.697656948883057, 'start': 77, 'end': 109, 'answer': '10 de diciembre de 1944'}

Hola, este sistema se realizó con el fin de tener una herramienta de consulta en el área de las competencias ciudadanas, estoy aquí para ayudarte
Puedes consultar sobre 4 contextos específicos de las competencias ciudadanas, si ni respuesta no es exacta, no te preocupes estoy aprendiendo con el tiempo
(Gracias por brindarme una oportunidad)
Para salir seleccione un contexto cualquiera y escriba la palabra Adios
Menu:
.....
Ingrese 1 - Contexto definición de las competencias ciudadanas
.....
PRINCIPIOS FUNDAMENTALES Título 1 - De los principios fundamentales
Artículo 1 Artículo 1º. Colombia es un Estado social de derecho,
organizado en forma de República unitaria, descentralizada, con
autonomía de sus entidades territoriales, democrática, participativa y
pluralista, fundada en el respeto de la dignidad humana, en el trabajo
y la solidaridad de las personas que la integran y en la prevalencia
del interés general. Artículo 2 Artículo 2º. Son fines esenciales
del Estado: servir a la comunidad; promover la prosperidad general y
garantizar la efectividad de los principios, derechos y deberes
consagrados en la Constitución; facilitar la participación de todos en
las decisiones que los afectan y en la vida económica, política,
administrativa y cultural de la Nación; defender la independencia
nacional, mantener la integridad territorial y asegurar la convivencia
pacífica y la vigencia de un orden justo. Las autoridades de la
República están instituidas para proteger a todas las personas
residentes en Colombia, en su vida, honra, bienes, creencias, y demás
derechos y libertades, y para asegurar el cumplimiento de los deberes
sociales del Estado y de los particulares. Artículo 3 Artículo 3º.
La soberanía reside exclusivamente en el pueblo, del cual emana el
poder público. El pueblo lo ejerce en forma directa o por medio de sus
representantes, en los términos que la Constitución establece.
Artículo 4 Artículo 4º. La Constitución es norma de normas. En todo
caso de incompatibilidad entre la Constitución y la ley u otra norma
jurídica, se aplican las disposiciones constitucionales. Es deber de
los nacionales y de los extranjeros en Colombia acatar la Constitución
y las leyes, y respetar y obedecer a las autoridades. Artículo 5
Artículo 5º. El Estado reconoce, sin discriminación alguna, la
libertad de los derechos inalienables de la persona y ampara a la
familia como institución básica de la sociedad. Artículo 6 Artículo
6º. Los particulares solo son responsables ante las autoridades por
hechos lícitos e ilícitos. Los servidores públicos lo son por
```

```
harvey@harvey-HP-Pavilion-Laptop-15-cw0005la: ~/Escritorio/Maestria$  
Fundar sindicatos y a sindicarse para la defensa de sus intereses.  
Artículo 24 Toda persona tiene derecho al descanso, al disfrute del  
tiempo libre, a una limitación razonable de la duración del trabajo y  
a vacaciones periódicas pagadas. Artículo 25 Toda persona tiene  
derecho a un nivel de vida adecuado que le asegure, así como a su  
familia, la salud y el bienestar, y en especial la alimentación, el  
vestido, la vivienda, la asistencia médica y los servicios sociales  
necesarios; tiene asimismo derecho a los seguros en caso de desempleo,  
enfermedad, invalidez, vejez u otros casos de pérdida de sus  
medios de subsistencia por circunstancias independientes de su  
voluntad. La maternidad y la infancia tienen derecho a cuidados y  
asistencia especiales. Todos los niños, nacidos de matrimonio o fuera  
de matrimonio, tienen derecho a igual protección social. Artículo 26  
Toda persona tiene derecho a la educación. La educación debe ser  
gratuita, al menos en lo concerniente a la instrucción elemental y  
fundamental. La instrucción elemental será obligatoria. La instrucción  
técnica y profesional habrá de ser generalizada; el acceso a los  
estudios superiores será igual para todos, en función de los méritos  
respectivos. La educación tendrá por objeto el pleno desarrollo de la  
personalidad humana y el fortalecimiento del respeto a los derechos  
humanos y a las libertades fundamentales; favorecerá la comprensión,  
la tolerancia y la amistad entre todas las naciones y todos los grupos  
étnicos o religiosos, y promoverá el desarrollo de las actividades de  
las Naciones Unidas para el mantenimiento de la paz. Los padres  
tendrán derecho preferente a escoger el tipo de educación que habrá de  
darse a sus hijos. Artículo 27 Toda persona tiene derecho a tomar  
parte libremente en la vida cultural de la comunidad, a gozar de las  
artes y a participar en el progreso científico y en los beneficios que  
de él resulten. Toda persona tiene derecho a la protección de los  
intereses morales y materiales que le correspondan por razón de las  
producciones científicas, literarias o artísticas de que sea autora.  
Artículo 28 Toda persona tiene derecho a que se establezca un orden  
social e internacional en el que los derechos y libertades proclamados  
en esta Declaración se hagan plenamente efectivos. Artículo 29 Toda  
persona tiene deberes respecto a la comunidad, puesto que solo en ella  
puede desarrollar libre y plenamente su personalidad. En el ejercicio  
de sus derechos y en el disfrute de sus libertades, toda persona  
debe estar solamente sujeta a las limitaciones establecidas por la ley con  
el único fin de asegurar el reconocimiento y el respeto de los  
derechos y libertades de los demás, y de satisfacer las justas  
exigencias de la moral, del orden público y del bienestar general en  
una sociedad democrática. Estos derechos y libertades no podrán, en  
ningún caso, ser ejercidos en oposición a los propósitos y principios  
de las Naciones Unidas. Artículo 30 Nada en esta Declaración podrá  
interpretarse en el sentido de que confiere derecho alguno al Estado,  
a un grupo o a una persona, para emprender y desarrollar actividades o  
realizar actos tendientes a la supresión de cualquiera de los derechos  
y libertades proclamados en esta Declaración.  
  
Seleccione cualquier contexto e ingrese la palabra "Adios" - Luego presione Enter para salir  
-----  
Seleccione un contexto:  
1  
Ingrese la pregunta:  
-----  
(¿)¿Qué es Colombia?  
Respuesta:  
-----  
un Estado social de derecho  
Seleccione un contexto:  
1  
Ingrese la pregunta:  
-----  
(¿)¿Qué es el castellano?  
Respuesta:  
-----  
Idioma oficial de Colombia  
Seleccione un contexto:  
4  
Ingrese la pregunta:  
-----  
(¿)¿Cuándo se firmó la declaración universal de los derechos humanos?  
Respuesta:  
-----  
10 de diciembre de 1948  
Seleccione un contexto:  
4  
Ingrese la pregunta:  
-----  
(¿)¿Qué artículo habla sobre el derecho a la vida?  
Respuesta:  
-----  
25  
Seleccione un contexto:  
1  
Ingrese la pregunta:  
-----  
Adios  
harvey@harvey-HP-Pavilion-Laptop-15-cw0005la: ~/Escritorio/Maestria$
```

```
harvey@harvey-HP-Pavilion-Laptop-15-cw0005la: ~/Escritorio/Maestria
de 1008
cualquiera 5528
de 1008
los 1065
derechos 2120
y 1042
libertades 10813
proclam 10491
#ados 1168
en 1036
esia 1359
declar 5526
#acin 29231
. 1099
[SEP] 5

{'score': 0.32093284749988474, 'start': 170, 'end': 197, 'answer': 'un Estado social de derecho'}
{'score': 0.32784813015509083, 'start': 149, 'end': 162, 'answer': 'en dos partes'}
{'score': 0.5041130741038184, 'start': 2980, 'end': 3053, 'answer': 'el pronunciamiento del pueblo convocado por el presidente de la República'}
{'score': 0.6976568688883057, 'start': 77, 'end': 100, 'answer': '10 de diciembre de 1948'}
[nltk_data] Downloading package stopwords to /home/harvey/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data] /home/harvey/nltk_data...
[nltk_data] Package averaged_perceptron_tagger is already up-to-
[nltk_data] date!
Training greetings vnl: [#####] 100%
Hola, este sistema se realizó con el fin de tener una herramienta de consulta en el área de las competencias ciudadanas, estoy aquí para ayudarte
Puedes consultar sobre 4 contextos específicos de las competencias ciudadanas, si nt respuesta no es exacta, no te preocupes estoy aprendiendo con el tiempo
¡Gracias por brindarme una oportunidad!
Para ingresar a los contenido ingresa la siguiente frase: Tengo una pregunta
Para terminar la conversacion escribe la palabra Adios
Para salir seleccione un contexto cualquiera y escriba la palabra Adios
Tu: Hola
Bot: Hola
Tu: ¿Cómo estas?
Bot: Supongo que bien.
Tu: ¿Que haces?
Bot: Bien.
Tu:

gratuita, al menos en lo concerniente a la instrucción elemental y
fundamental. La instrucción elemental será obligatoria. La instrucción
técnica y profesional habrá de ser generalizada; el acceso a los
estudios superiores será igual para todos, en función de los méritos
respectivos. La educación tendrá por objeto el pleno desarrollo de la
personalidad humana y el fortalecimiento del respeto a los derechos
humanos y a las libertades fundamentales; favorecer la comprensión,
la tolerancia y la amistad entre todas las naciones y todos los grupos
étnicos o religiosos, y promover el desarrollo de las actividades de
las Naciones Unidas para el mantenimiento de la paz. Los padres
tendrán derecho preferente a escoger el tipo de educación que habrá de
darse a sus hijos. Artículo 27 Toda persona tiene derecho a tomar
parte libremente en la vida cultural de la comunidad, a gozar de las
artes y a participar en el progreso científico y en los beneficios que
de él resulten. Toda persona tiene derecho a la protección de los
intereses morales y materiales que le correspondan por razón de las
producciones científicas, literarias o artísticas de que sea autora.
Artículo 28 Toda persona tiene derecho a que se establezca un orden
social e internacional en el que los derechos y libertades proclamados
en esta Declaración se hagan plenamente efectivos. Artículo 29 Toda
persona tiene deberes respecto a la comunidad, puesto que sólo en ella
puede desarrollar libre y plenamente su personalidad. En el ejercicio
de sus derechos y en el disfrute de sus libertades, toda persona
estará solamente sujeta a las limitaciones establecidas por la ley con
el único fin de asegurar el reconocimiento y el respeto de los
derechos y libertades de los demás, y de satisfacer las justas
exigencias de la moral, del orden público y del bienestar general en
una sociedad democrática. Estos derechos y libertades no podrán, en
ningún caso, ser ejercidos en oposición a los propósitos y principios
de las Naciones Unidas. Artículo 30 Nada en esta Declaración podrá
interpretarse en el sentido de que confiere derecho alguno al Estado,
a un grupo o a una persona, para emprender y desarrollar actividades o
realizar actos tendientes a la supresión de cualquiera de los derechos
y libertades proclamados en esta Declaración.

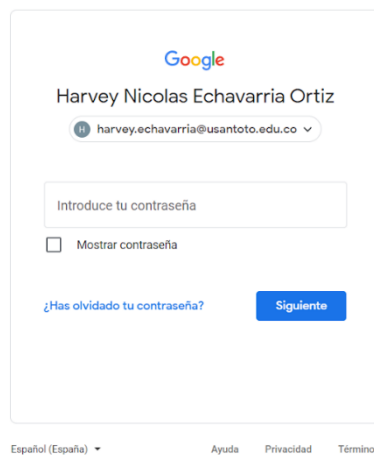
Seleccione un contexto:
-----
1
Ingrese la pregunta:
-----
¿Qué es Colombia?
Respuesta:
-----
un Estado social de derecho
Seleccione un contexto:
-----
1
Ingrese la pregunta:
-----
Adios
harvey@harvey-HP-Pavilion-Laptop-15-cw0005la: ~/Escritorio/Maestria$
```

### ANEXO 3. Manual de Usuario:

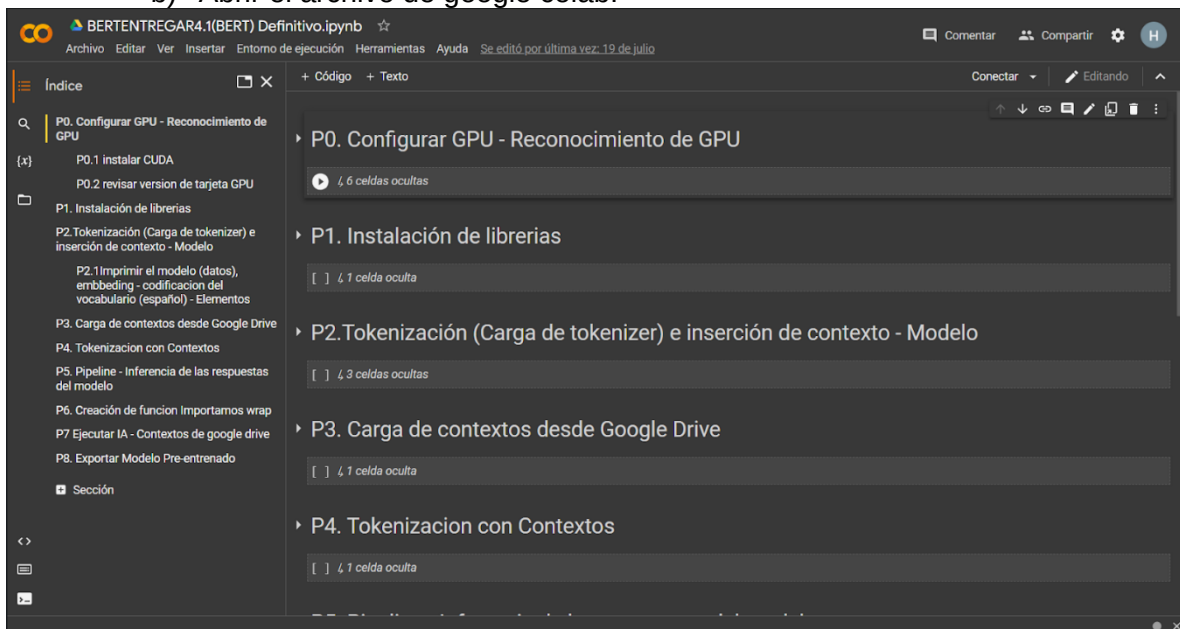
A continuación, se da a conocer una guía que ejemplifica el uso de los archivos a nivel de Google Colab y localmente (equipo con dispositivo Ubuntu - Windos) para su ejecución:

#### 1. Ejecución desde Google Colab:

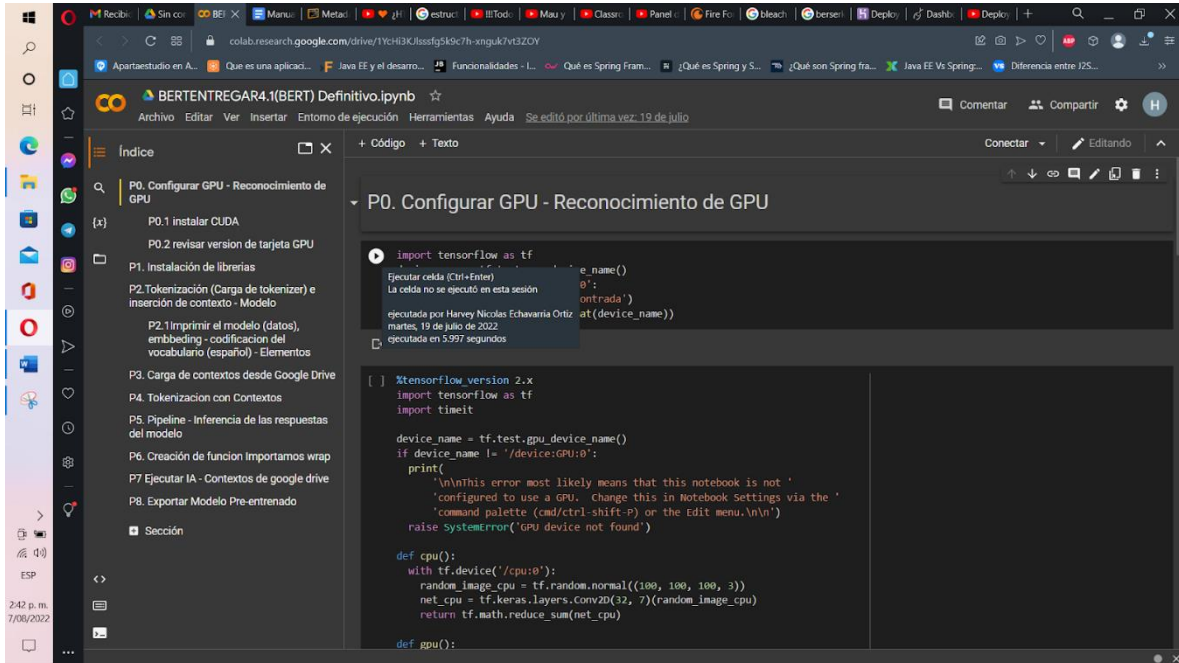
##### a) Acceder a la cuenta de usuario:



##### b) Abrir el archivo de google colab:



- c) En el ícono de play dar clic (cabe resaltar que esto debe hacerse con cada bloque de código de los numerales<sup>5</sup>):



Nota: Hasta que no finalice la ejecución de cada paso no debe ejecutarse el siguiente paso, ya que el proceso de ejecución es secuencial.

## 2. Ejecución a nivel local (Ubuntu - Windows):

- Abrir una terminal
- Acceder por terminal a la carpeta en la que se encuentra el archivo python y los contextos, por medio del comando `cd`.
- Ejecutar en la terminal el archivo python por medio del comando `python3` (nombre del archivo) – Caso Ubuntu
- Ejecutar en la terminal el archivo python por medio del comando `python` (nombre del archivo) – Caso Windows

---

<sup>5</sup> Los numerales pueden visualizarse dando clic al icono de la parte superior izquierda (tres líneas)

ANEXO 4. Documentación Banco Preguntas:

Contexto 1:

1. ¿Qué es la constitución?
2. ¿Qué es Colombia?
3. ¿Qué obligación tiene el estado?
4. ¿Qué es el castellano?
5. ¿Cuáles son los fines esenciales del estado?
6. ¿Sobre qué se fundamentan las relaciones exteriores del estado?

Contexto 2:

1. ¿En cuántas partes está organizado el estado colombiano?
2. ¿Cuáles son las ramas del poder?
3. ¿Cómo está conformada la rama legislativa?
4. ¿Quiénes se encuentran dentro de la rama ejecutiva?
5. ¿Quiénes conforman la rama judicial?

Contexto 3:

1. ¿Qué permite el modelo democrático de Colombia?
2. ¿Qué expone la ley 134 de 1994?
3. ¿Qué expone la ley 134 de 1994 en su artículo 2?
4. ¿Qué es el referendo?
5. ¿Qué es el cabildo abierto?
6. ¿Qué es la Iniciativa popular legislativa?
7. ¿Qué es el Plebiscito?

Contexto 4:

1. ¿Cuándo se firmó la declaración universal de los derechos humanos?
2. Según la declaración universal de los derechos humanos, ¿A qué tiene derecho todo individuo?
3. ¿Qué artículo habla sobre el derecho a la nacionalidad?
4. ¿A qué tiene derecho toda persona que trabaja?
5. ¿Cuál es el artículo que hace referencia al derecho a la vida?

### ANEXO 5. Guía de instalación local:

#### **Ubuntu**

##### a) Instalación de Python

```
sudo apt install python  
  
sudo apt install python3-pip
```

##### b) Instalación de librerías

```
pip install Transformers  
  
pip install chatterbot
```

##### c) Verificar version de python

```
python3 --version
```

##### d) Instalación de torch:

```
pip install torch
```

#### **Windows**

##### a) Verificar version de python

```
python --version
```

#### b) Instalación de librerías

```
pip install Transformers
```

```
pip install chatterbot
```

#### c) Instalación de torch:

```
pip install torch
```

#### d) Ejecución

Abrir Carpeta IA en terminal

Pegar el siguiente comando:

```
python BertLocal.py
```

Anexo 6. Ficha Técnica

| Google Colab |                   |                          |
|--------------|-------------------|--------------------------|
|              | Característica    | Requisito                |
| 1            | Sistema Operativo | Windows/Mac/Linux        |
| 2            | Conectividad      | Cuanta de Google         |
| 3            | Almacenamiento    | Acceso a contextos de IA |
| 4            | Conectividad      | Acceso a internet        |

| Local |                   |                                                                |
|-------|-------------------|----------------------------------------------------------------|
|       | Característica    | Requisito                                                      |
| 1     | Sistema Operativo | Windows/Mac/Linux                                              |
| 2     | Herramientas      | Python y librerías                                             |
| 3     | Almacenamiento    | Espacio para instalación de librerías y guardar los contextos. |