BERT TFG

Introducción

Problema Abordado

Qué es BERT?

Article Classification

Sentiment Análisis

Query Answer

RESUMEN

Uso de modelos ya preentrenados y realizadas pruebas con ellos…

Indice tablas y esquemas

INTRODUCCIÓN

1.1 Antecedentes

El procesamiento del lenguaje natural, se puede encontrar en muchos de los servicios, que hoy en día la tecnología nos ofrece.

Podemos, encontrar este tipo de procesamientos, en algoritmos como el que utiliza el propio Google en su buscador para resolver las Query.

En este proyecto, hemos empleado, una red neuronal, previamente entrenada por Google.

1.2 Contexto y Justificación del trabajo

En la actualidad existen infinidad de aplicaciones para el procesamiento del lenguaje natural, a partir de aquí mencionado con la nomenclatura NLP (Natural Language Processing).

El uso de los sistemas NLP, ha ido ganando fuerza en aplicaciones orientadas a la interacción con el usuario y a la recopilación de información, es el caso de los ChatBots y los buscadores de páginas web.

Gracias a la red pre-entrenada por Google, llamada Google Bert, los sistemas NLP, han dado un gran salto, permitiendo, que las máquinas, sean capaces de extraer y entender el contexto en el que se encuentran.

Con todo esto, se pretende, realizar la implementación de Google Bert, revisar su funcionamiento, otorgarle nuevas funcionalidades y experimentar con nuevos modelos y tratamientos, todo ello, unificado dentro de una interfaz gráfica, está interfaz será un prototipo que podrá ser mejorada añadiendo más funcionalidades y mejores diseños.

1.3 Objetivos del trabajo

Los principales objetivos del proyecto son:

* Obtener un clasificador de Noticias por cabecera de la noticia utilizando Bert

Adicionalmente, como objetivos secundarios:

* Implementación de una Interfaz Gráfica
* Implementación de un Analizador de Sentimientos
* Implementación de un ChatBot
* Revisión del funcionamiento de Google Bert
* Mejorar resultados obtenidos
* Implementación de técnicas para aprendizaje automático

1.4 Expectativas

Este proyecto tiene como intención, aprender más acerca de las redes neuronales y el aprendizaje automático, además de aprender y entender cómo funciona el NLP.

Así mismo, aplicar las técnicas aprendidas en el grado de informática y los conocimientos obtenidos en la rama de computación.

Adicionalmente, otra de las expectativas del proyecto, es, obtener conocimientos sobre librerías como matplotlib, numpy, pandas y pythorch.

1.5 Contenido que se va a ver en este proyecto

…

2. Qué es Bert?

2.1 Historia

Bert, cuyas siglas significan Bidirectional Encoder Representations for Transformers, el objetivo de este algoritmo consiste en interpretar nuestro lenguaje de búsqueda de una manera mucho más natural, mediante ***PNL*** (Programación Neuro Lingüística). Para ello, Bert utiliza una red neuronal de código abierto, cabe destacar que hasta la llegada de Google Bert, Google, no tenía respuesta para aproximadamente el 15% de las consultas (Dato proporcionado por etudios …).

Como predecesor a Bert aplicado por Google en la tarea de proporcionar resultados ante consultas en el buscador , podemos encontrar el algoritmo llamado RankBrain (Primer algoritmo programado por Google utilizando técnicas de inteligencia artificial)

2.2 Bidireccionalidad

Como mencionan las siglas, Bert utiliza un método de bidireccional para procesar palabra, esa bidireccionalidad consiste en analizar las palabras situadas tanto a la derecha como a la izquierda de la palabra clave que se está analizando.

Gracias a esto, permite al algoritmo entender en profundidad el contexto en el que se encuentra la palabra clave.

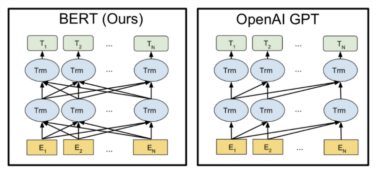


Figura X.X – Comparación entre la arquitectura Bert y la arquitectura GPT

Si se compara el preprocesamiento de otra arquitectura dedicada al NLP, como es la GPT desarrollada por OpenAI, se puede observar como en la arquitectura Bert, las entradas se encuentran totalmente conectadas, este tipo de arquitecturas se denominan arquitecturas de autoatención simple.

Mientras que, si se observa la otra arquitectura, en la GPT se puede observar cómo solamente hay conexión con los tokens predecesores al token clave, esta arquitectura se denomina autoatención casual.

En este caso, la arquitectura de atención casual es menos eficiente en el procesamiento de testo, dado que no puede recibir información sobre tokens futuros. Aunque, por otro lado, la atención casual es necesaria para poder realizar generación de texto, mientras que la arquitectura de atención simple que utiliza Bert resulta útil para procesarlo.

2.3 Redes Transformer

Como indican las siglas, Bert hace uso de redes neuronales del tipo Transformer.

Como predecesoras de estas redes, podemos encontrar las redes neuronales del tipo Recurrente, nacidas con el auge del Deep Learning, surgieron las técnicas de embedding, estas técnicas consisten en representaciones matriciales del texto, hasta finales del año 2017, este tipo de redes se convirtió en el estándar.

\*INSERTAR IMAGEN RED NEURONAL RECURRENTE\*

A finales del 2017 Google, presentó la arquitectura Transformer en una presentación conocida como –“Attention is All You Need”-, este modelo, tenía como innovación, la sustitución de las capas recurrentes por las denominadas capas de atención.

Estas capas, codifican cada palabra de una frase en función del resto de la secuencia, permitiendo así, introducir el contexto en la representación matemática del texto.

\*INSERTAR IMAGEN RED NEURONAL TRANSFORMER\*

La arquitectura del tipo transformer, además trae otras innovaciones como los embeddings posicionales, permiten al algoritmo conocer la posición relativa de cada palabra.

Este tipo de arquitectura se trabaja realizando un pre-entrenamiento previo, y posteriormente un fine-tuning, diseñando un modelo que permita adaptar los modelos a tareas específicas.

Esta arquitectura hace referencia a una arquitectura del tipo Transformers,

ANEXO

Qué es Bert?

[¿Qué es BERT y cómo funciona? - DXmedia](https://dxmedia.net/algoritmo-bert-google/)

[Explicación e implementación paso a paso de BERT Parte 1 - Preprocesamiento (ichi.pro)](https://ichi.pro/es/explicacion-e-implementacion-paso-a-paso-de-bert-parte-1-preprocesamiento-241867215631106)

[Uso de redes neuronales tipo transformer: Novedades en Wolfram Language 12](https://www.wolfram.com/language/12/neural-network-framework/use-transformer-neural-nets.html.es?footer=lang)

[Transformers en Procesamiento del Lenguaje Natural - IIC (uam.es)](https://www.iic.uam.es/innovacion/transformers-en-procesamiento-del-lenguaje-natural/)

CLASIFICACIÓN DE ARTÍCULOS

Datos Usados

Para la implementación, hemos utilizado el modelo bert-base-cased