**WEEK 10 REPORT**

**Group Name:**

|  |  |  |
| --- | --- | --- |
|  | Member 1 | Member 2 |
| Name | Keilor Fallas Prado | Ky Dang |
| Email | kfallasprado@gmail.com | Keith.dang1610@gmail.com |
| Country | Costa Rica | Vietnam |
| College/ Company |  |  |
| Specialization | NLP | NLP |

**Problem description:**

The dataset consists of two distinct subsets: a **Train** dataset and a **Test** dataset. The **Train** dataset includes three features along with a label column. The **Test** dataset, however, does not contain the label column, as its purpose is to evaluate the model's performance once trained.

The task at hand is to build a deep learning model capable of accurately detecting **hate speech** from textual data. Given the nature of the task, it is crucial to select an appropriate model for processing and analyzing textual information.

There are a variety of deep learning models available, each with its strengths and weaknesses, including but not limited to:

* **Long Short-Term Memory (LSTM)**: LSTMs are recurrent neural networks (RNNs) designed to capture long-range dependencies in sequential data. They are often used for tasks like text classification, sentiment analysis, and speech recognition.
* **Gated Recurrent Units (GRU)**: GRUs are a variant of LSTMs that also handle sequential data. While they are computationally more efficient than LSTMs, they offer similar performance for many natural language processing (NLP) tasks.
* **Convolutional Neural Networks (CNN)**: CNNs are primarily known for image processing but have also been shown to perform well for text classification tasks, where they can extract features from sequences of words.
* **Transformers (e.g., BERT, RoBERTa)**: Transformers, particularly models like **BERT (Bidirectional Encoder Representations from Transformers)** and **RoBERTa (Robustly optimized BERT approach)**, have revolutionized NLP. These models are based on the transformer architecture, which uses self-attention mechanisms to capture complex relationships within the data. BERT and RoBERTa, in particular, have been pre-trained on large corpora and have shown remarkable results across various NLP tasks, including sentiment analysis, named entity recognition, and hate speech detection.

In the context of **hate speech detection**, the use of a **transformer-based model** such as **BERT or RoBERTa** may be highly beneficial. These models have demonstrated superior performance on a wide range of NLP tasks due to their ability to understand the nuances of language, including sarcasm, context, and subtle linguistic features often present in hate speech.

**Project lifecycle**

|  |  |  |
| --- | --- | --- |
| **Weeks** | **Due date** | **Plan** |
| Week 8 | 11/26/2024 | Review data source and ensure it is representative of hate speech contexts. |
| Week 9 | 12/02/2024 | Remove duplicates, nulls, and irrelevant data. |
| Week 10 | 12/09/2024 | Evaluate and select models such as Logistic Regression, SVM, or Transformers (e.g., BERT). |
| Week 11 | 12/16/2024 | Tokenization - Identify relevant linguistic and contextual features |
| Week 12 | 12/23/2024 | Training and evaluation model |
| Week 13 | 12/30/2024 | Document the challenge |

**Github Repo link:**

* Individual GitHub links:
  + KyDang: <https://github.com/KeithDang1610/NLP_HateSpeech-Detection>
  + Keilor: