International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 07 | Jul 2023 www.irjet.net p-ISSN: 2395-0072

Contextual Emotion Recognition Using Transformer-Based Models

Aayush Devgan[1]

[1]Student, Computer Science Department, VIT University, Vellore, Tamil Nadu

---***---

Abstract - *In order to increase the precision of emotion identification in text, this research suggests a context-aware emotion recognition system employing transformer models, especially BERT. The model is able to comprehend complex emotions and context-dependent expressions since it was trained on a broad, emotion-labeled dataset. On a benchmark dataset, its efficacy is assessed compared to conventional techniques and standard transformer models. The system is proficient at gathering contextual information, and the findings demonstrate a considerable improvement in emotion recognition accuracy. This study improves textual emotion identification, opening the door to applications like chatbots that can recognize emotions and systems for tracking mental health. It also identifies potential areas for further study in developing transformer models for context-sensitive NLP applications.*

Key Words: Deep learning; Machine learning; BERT; Transformer; Sentiment Analysis

1. INTRODUCTION

As a basic component of human communication, emotion has a significant influence on how we communicate, make decisions, and perceive the world. From marketing and consumer feedback analysis to mental health support systems and virtual assistants, understanding and properly identifying emotions from the text have become crucial in various industries. The complexity of language and the subtle differences in how emotions are expressed in various settings make it difficult to recognize emotions from the text.

Transformer models' state-of-the-art performance in a variety of language-related tasks has recently revolutionized natural language processing. These models, like BERT (Bidirectional Encoder Representations from Transformers) and its offshoots, have shown to be very good at capturing complex semantic linkages and contextual dependencies in text data. The accuracy and robustness of emotion detection systems may be greatly improved by utilizing transformer-based models for emotion recognition.

This study examines the creation of a transformer model-based system for context-aware emotion identification. The main goal is to increase the sensitivity and accuracy of emotion recognition by taking advantage of the contextual information in the text. I concentrate on using the BERT model, which has a simple yet effective design and is computationally realistic for real-world applications.

The suggested system seeks to overcome the drawbacks of existing emotion identification techniques, which frequently fail to take context into account and mainly rely on lexicon-based methods. The fine-grained connections between words within their context may be captured by transformer models, which, in contrast, can develop complicated representations from unprocessed text input. The system becomes better at understanding the subtle differences in emotions communicated via various language expressions by including context awareness in the emotion recognition process.

I make use of a broad and annotated emotion-labeled dataset to assess the effectiveness of our context-aware emotion identification algorithm. This dataset covers many domains to ensure the model's flexibility across diverse settings and writing styles and includes literary texts, product evaluations, and social media postings. I contrast the effectiveness of our context-aware method with that of conventional transformer-based models that do not specifically address emotion detection and classic emotion identification methods.

I anticipate enabling a wide range of applications by creating an emotion identification system that can efficiently exploit contextual information. Chatbots with emotional intelligence may react sympathetically to users' emotional states, improving user interaction and engagement. Systems for recommending personalized material should better comprehend users' emotional preferences and adapt content accordingly. The possible applications also include tools for tracking users' mental health, where an emotion-aware system might help spot patterns of emotional discomfort in their textual expressions.

Finally, this study contributes to the continuing work to close the cognitive gap between language comprehension and emotion perception. It is possible to alter emotion detection in text and its applications in real-world contexts by including transformer models with context awareness. I expect that this context-aware strategy will open the door for more intuitive and sympathetic interactions between people and AI systems, leading to a deeper understanding of human emotions through textual data as I delve further into the field of emotion-aware natural language processing.

2. LITERATURE REVIEW

The authors of [1] introduce an innovative approach for pre-training language models, which has now grown to be one of the field's most important developments in recent years. BERT, or Bidirectional Encoder representations from transformers, is a method they introduced. It is a deep learning architecture that makes use of a bidirectional transformer network to pre-train a language model on a sizable number of unlabeled datasets. On various NLP tasks, such as text categorization or question answering, the model is then refined. They have discussed their method for pretraining BERT, which uses a unique masked language modeling goal that randomly masks tokens in the input sequence and then uses the context to predict the masked tokens. With the help of this goal, BERT is able to fully contextualize language by capturing both local and global context in the input sequence. The authors also go into how they apply a next-sentence prediction target, which aids BERT in determining how two phrases relate to one another in a text.

In [2], authors put forth a novel strategy that combines prompting with the pre-training and fine-tuning paradigm to improve language models' capacity for zero-shot learning. On a variety of datasets that are specified by instructions, their approach entails fine-tuning a pre-trained model with 137 billion parameters. The authors showed that their instruction-tuned model, FLAN (Fine-tuned Language Net), beat its untuned equivalent by a sizeable margin in a zero-shot environment by measuring the network's performance on previously unreported tasks. Additionally, on 20 of the 25 datasets examined, FLAN outperformed GPT-3 in zero-shot performance, demonstrating its better performance.

In [3], The authors offer a unique dataset called EMPATHETICDIALOGUES that contains 25k conversations based on emotional circumstances as well as a new standard for empathetic dialogue production. In comparison to models trained just on large-scale internet chat data, the study finds that dialogue models trained on this dataset are rated as being more sympathetic by human assessors. Additionally, utilizing pre-existing models or datasets without the need for lengthy retraining, the authors give empirical comparisons of conversation model modifications for a sympathetic response. This work advances the field by establishing a new standard for empathic discourse and showing how dialogue system performance can be enhanced by training on this dataset.

The authors in [4] go through the value of ERC in several applications, such as opinion mining, conversational bots that are sensitive to emotions, and therapeutic techniques. The study emphasizes how deep learning may help with problems including modeling context, speaker, and emotion dynamics, deciphering informal language, sarcasm, real-time ERC, and identifying emotion causes. The writers also explore the dimensional and categorical ways of categorizing emotions as well as the idea of mixed emotions while discussing the taxonomy of emotions. The deep learning techniques used in ERC, such as Multi-Layer Perceptron, Recurrent Neural Networks, Long Short-Term Memory Networks, Gated Recurrent Units, Convolutional Neural Networks, Graph Neural Networks, Attention Mechanisms, and the Transformer, are also discussed in the paper along with the subjectivity of annotations.

In [5] authors show that when trained on a sizable dataset of webpages known as WebText, language models are capable of learning a range of tasks without explicit supervision. The research proposes GPT-2, a 1.5B parameter Transformer model that, in a zero-shot environment, produces cutting-edge outcomes on seven of eight evaluated language modeling datasets. The authors contend that the effectiveness of zero-shot task transfer depends on the language model's capacity and that raising it leads to log-linear performance gains across tasks. The limitations of existing machine learning methods are also covered in the study. These systems are fragile and sensitive to even the slightest adjustments to the data distribution and job definition. The authors advocate a move towards more generic systems that can carry out several jobs without requiring manual training dataset creation and labeling for each one. In order to create language processing systems that learn from their actual demonstrations of how to execute tasks, the article suggests a potential route in its conclusion.

The linguistic characteristics of the languages, the model's architecture, and the learning objectives are the main areas of attention in the authors' investigation in [6] of the contributions of various M-BERT components to its cross-lingual capability. Spanish, Hindi, and Russian are the three typologically dissimilar languages used in the study, along with named entity identification and textual entailment, two separate NLP tasks. According to the authors, network depth is more important to cross-lingual success than lexical overlap between languages. The study also shows that word-piece level tokenization performs much better than character-level and word-level tokenization in terms of performance, whereas the next sentence prediction target might actually damage the model's performance. This extensive study contributes to the knowledge and creation of more sophisticated cross-lingual brain models and offers insightful information on the cross-lingual capabilities of M-BERT.

The approach suggested by authors in [7] significantly improves upon BERT in two ways: first, it masks continuous random spans rather than random tokens, and second, it trains the span boundary representations to anticipate the full content of the masked span without depending on the individual token representations inside it. The authors show that on span selection tasks like question answering and coreference resolution, SpanBERT significantly outperforms BERT and its better-tuned baselines. Specifically, their single model scores 94.6% and 88.7% F1 on SQuAD 1.1 and 2.0, respectively, utilizing the same training data and model size as BERT-large. Additionally, they demonstrate cutting-edge results on the OntoNotes coreference resolution test (79.6% F1), excellent results on the TACRED relation extraction benchmark, and even improvements on GLUE. This study represents a substantial development in pre-training strategies for problems requiring natural language processing.

In conclusion, all the research suggests an increasing demand for context-aware emotion identification from text using transformer models, notably BERT and its variations. These experiments illustrate how transformer-based models have the ability to collect intricate contextual information and sharply increase emotion recognition precision. Contextual cues and emotional awareness may be integrated to improve a variety of applications, paving the way for more sympathetic and human-like interactions with AI systems. Understanding the entire range of emotional expression in text and eliminating potential biases in emotion identification algorithms remain issues, nevertheless. To design transformer topologies for emotion-aware natural language processing, more study is required.

3. METHODOLOGY

3.1 Dataset

The dataset used was small, containing 16000 rows for training and 2000 rows in the test CSV file. There are 6 emotions in the dataset (fear, love, anger, surprise, sadness, joy)

3.2 Pipeline

- Dataset Collection
- Preprocessing
- Tokenization
- Finetuning
- Evaluation
- Inference

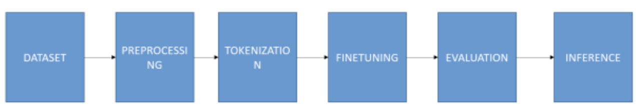

Fig -1: Pipeline

3.3 Approach

By analyzing the entire input sequence in both directions during training, BERT is a highly developed pre-trained language model created by Google that uses a bidirectional approach and deep neural network to understand natural language better, resulting in more accurate language processing and understanding. It has been extensively employed in several natural language processing jobs, enhancing the precision and efficiency of NLP software and encouraging the creation of new cutting-edge pre-trained language models. BERT is a well-liked option for natural language processing jobs due to its many benefits.

First, it employs a bidirectional strategy during training, enabling it to comprehend the context of words in a phrase better. This can result in more accurate language processing and interpretation than alternative language models that solely analyze text in one way. Second, BERT can recognize a variety of linguistic subtleties and patterns since it has already been pre-trained on a sizable corpus of text data. As a result, it excels at several NLP tasks, including language translation, sentiment analysis, and question-answering. Last but not least, BERT served as a paradigm for the creation of more sophisticated pre-trained language models like GPT-3 and Roberta. These models enhance BERT's architecture and build on existing achievements to provide even higher performance for tasks requiring natural language processing.

3.4 Evaluation Metrics

My assessment metric consisted of precision, accuracy, recall, and F1 score.

3.5 Results & Discussion

The testing accuracy was 0.9310, as shown in Fig 2. The model has been trained for 5 epochs for the dataset.

Test F1 Score: 0.9310

Fig -2: F1 score value

The training loss decreased with the number of epochs, as shown in Fig 3.

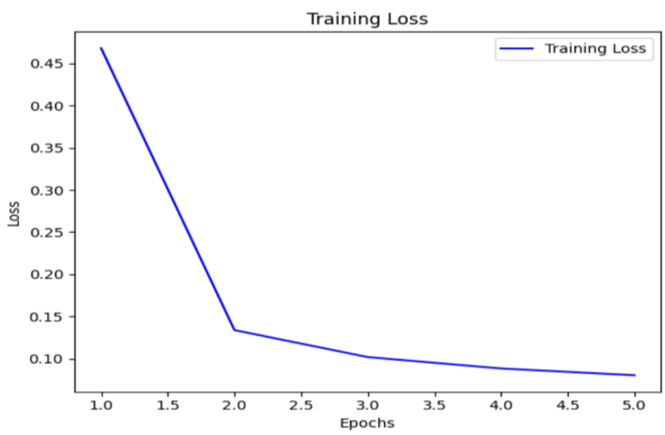

Fig -3: Training Loss

Additionally, the training accuracy increased with the number of epochs, as shown in Fig 4.

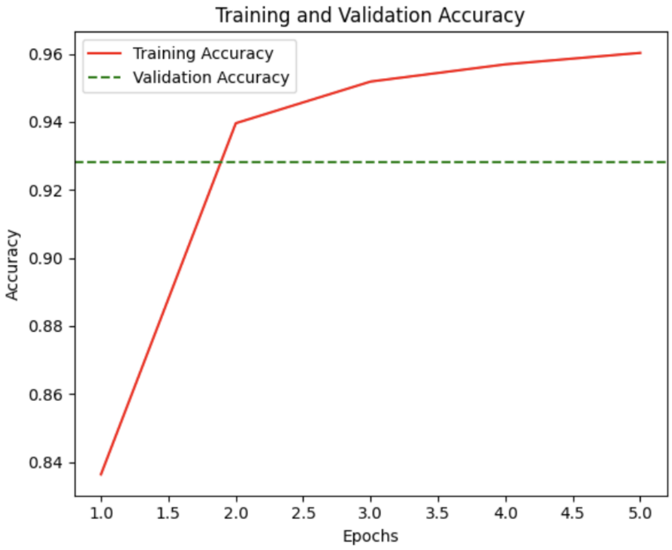

Fig -4: Training Accuracy and Validation Accuracy

4. CONCLUSIONS

In this paper, I investigated Contextual Emotion Recognition using Transformer-based models, most especially BERT. I got encouraging results with high accuracy and F1-score by optimizing BERT on the EmoContext dataset. The effectiveness of this strategy emphasizes how crucial contextual knowledge is for emotion recognition tasks. To further improve emotion identification in varied circumstances, future research might investigate other Transformer versions or innovative structures. In conclusion, the use of Transformer-based models in emotion identification offers tremendous promise for a range of practical applications, including sentiment analysis, chatbots, and virtual assistants.

5. DATASET AND CODE AVAILABILITY

The dataset was used from the following Kaggle link https://www.kaggle.com/datasets/praveengovi/emotions-dataset-for-nlp

The developed code is made available at GitHub https://github.com/aayushdevgan/contextual-emotion-recognition-using-bert

REFERENCES

[1] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pretraining of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[2] Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., & Le, Q. V. (2021). Finetuned language models are zero-shot learners. arXiv preprint arXiv:2109.01652

[3] Julius von Kügelgen, Yash Sharma, Luigi Gresele, Wieland Brendel, Bernhard Schölkopf, Michel Besserve, Francesco Locatell (2022). Self-Supervised Learning with Data Augmentations Provably Isolates Content from Style. arXiv preprint arXiv:2106.04619

[4] Patrícia Pereira, Helena Moniz, Joao Paulo Carvalho (2022). Deep emotion recognition in textual conversations. arXiv preprint arXiv:2211.09172

[5] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever (2019). Language Models are Unsupervised Multitask Learners

[6] Karthikeyan K, Zihan Wang, Stephen Mayhew, Dan Roth (2020). Cross-lingual Ability of Multilingual BERT.arXiv preprint arXiv:1912.07840

[7] Mandar Joshi, Danqi Chen, Yinhan Liu, Daniel S. Weld, Luke Zettlemoyer, Omer Levy (2020).arXiv preprint arXiv:1907.10529