Sentiment Analysis with BERT

Vaishnavi Kukkala, Hari Krishna Para, Sai Charan Chandu Patla

University of New Haven

Emails: vkukk2@unh.newhaven.edu, hpara2@unh.newhaven.edu, schan30@unh.newhaven.edu

Abstract

This paper explores the application of BERT (Bidirectional Encoder Representations from Transformers) for sentiment analysis. Sentiment analysis is a critical task in natural language processing (NLP) with applications ranging from social media monitoring to customer feedback analysis. Using BERT, we aim to classify text into positive, negative, or neutral sentiments with high accuracy. Our experiments involve fine-tuning a pre-trained BERT model on publicly available datasets, comparing its performance to traditional machine learning and deep learning baselines. We analyze the model's performance in terms of accuracy, precision, recall, and F1-score. Results indicate that BERT significantly outperforms traditional baselines, though challenges in domain adaptation and computational efficiency persist.

1. Introduction

Sentiment analysis, or opinion mining, involves determining the sentiment expressed in textual data. As one of the most popular applications in NLP, it is critical for understanding user sentiment in applications such as product reviews, social media, and customer service.

Traditional approaches to sentiment analysis, such as bag-of-words models or TF-IDF with machine learning classifiers, suffer from limitations in capturing semantic and contextual nuances. Neural networks, particularly RNNs and LSTMs, brought significant improvements but were limited by their sequential nature. Transformer-based models, particularly BERT, have revolutionized NLP by enabling deeper contextual understanding through bidirectional attention mechanisms.

In this paper, we explore the application of BERT for sentiment analysis. We fine-tune a pre-trained BERT model on sentiment datasets and evaluate its performance against traditional baselines. Additionally, we perform ablation studies to understand the contribution of different components to the model's overall performance.

2. Related Work

Numerous studies have explored sentiment analysis using machine learning and deep learning methods. Early methods relied on feature extraction techniques such as n-grams and sentiment lexicons coupled with classifiers like SVM or logistic regression. The advent of deep learning introduced RNNs and LSTMs for sentiment analysis, allowing models to better capture sequential dependencies.

The release of Transformer-based models marked a paradigm shift in NLP. BERT has been extensively applied to a variety of NLP tasks, including sentiment analysis. Studies have shown BERT's superior performance on datasets like SST-2 and IMDB reviews. However, challenges such as domain adaptation and computational requirements remain relevant.

3. Methodology

We use the following datasets for training and evaluation:

1. IMDB Reviews: A large dataset of movie reviews categorized as positive or negative.
2. SST-2 (Stanford Sentiment Treebank): A dataset with fine-grained sentiment labels (positive, neutral, and negative).

Our approach involves fine-tuning a pre-trained BERT model on the sentiment analysis datasets.

The model includes:

- A pre-trained BERT base model (uncased, 12-layer, 768-hidden, 12-heads).
- A classification head comprising a feed-forward layer with softmax activation for sentiment prediction.

Implementation details include preprocessing, training, and baseline comparisons.

4. Results and Discussion

Quantitative results show that BERT consistently outperformed both logistic regression and LSTM baselines across datasets. Its ability to capture bidirectional context and nuanced semantic information significantly enhanced performance.

Error Analysis:

1. Domain-specific Language: The model struggled with domain-specific slang or idiomatic expressions.
2. Neutral Sentiments: Fine-grained classification posed challenges due to subtle contextual variations.
3. Data Imbalance: Imbalanced sentiment distributions affected model generalizability.

5. Conclusion and Future Work

This study demonstrates the effectiveness of BERT for sentiment analysis, achieving state-of-the-art performance on benchmark datasets. Despite its success, challenges in handling domain-specific language and computational efficiency remain.

Future work could involve:

- Exploring domain adaptation techniques for BERT.

- Investigating lightweight transformer variants for efficiency.

- Incorporating external knowledge graphs for improved sentiment understanding.