Sentiment Analysis of Bank Customer Reviews using Bidirectional Encoder Representations from Transformers

Thanapoom Phatthanaphan
Stevens Institute of Technology
tphattha@stevens.edu

Abstract

This research delves into the realm of sentiment analysis, focusing on the application of Convolutional Neural Networks (CNN) with Bidirectional Encoder Representations from Transformers (BERT) embeddings to decipher the sentiments embedded in a substantial dataset comprising 10,000 recent customer reviews from 48 diverse US banks. The primary objective is to develop an advanced sentiment analysis system capable of accurately classifying customer feedback as positive or negative. The utilization of CNN with BERT embeddings aims to enhance the precision of sentiment classification, offering the banking industry actionable insights to elevate service quality and effectively address customer concerns.

Introduction

In the ever-evolving landscape of the banking industry, the invaluable role of customer feedback cannot be overstated. In an era dominated by digital communication, customer reviews and textual comments wield a considerable influence over consumer choices, shaping the overall reputation and performance of financial institutions. The multifaceted nature of language and the sheer volume of textual data, however, present a formidable challenge when attempting to extract meaningful insights from these reviews.

This research embarks on a pioneering journey to navigate this challenge by adopting a sophisticated approach to sentiment analysis. The traditional methods of sentiment analysis fall short when faced with the complexity and nuance inherent in customer reviews. To overcome these limitations, we embrace cutting-edge technologies, specifically Convolutional Neural Networks (CNN) augmented with BERT embeddings.

The central objective of this research is to construct an advanced sentiment analysis system that goes beyond mere classification of customer reviews as positive or negative. We delve into the nuances of customer sentiments and preferences, aiming to unravel the intricate layers of feedback intricately woven into the vast tapestry of textual data. This ambitious project focuses on a substantial dataset comprising 10,000 recent customer reviews sourced from 48 diverse banks across the United States.

The integration of Convolutional Neural Networks (CNN) with BERT embeddings signifies a departure from conventional sentiment analysis methods. CNNs, known for their prowess in image recognition, are adapted to capture intricate patterns and relationships within the textual data. BERT embeddings, a state-of-the-art language representation model, further enhance the contextual understanding of the reviews, capturing subtleties that elude traditional methods. The primary goal is not merely to categorize reviews but to provide the banking industry with nuanced and precise insights into customer sentiments. By leveraging the power of CNN with BERT embeddings, we aim not only to overcome the challenges inherent in textual feedback analysis but also to elevate the analysis to a level where it becomes a strategic tool for enhancing service quality. This research aspires to empower the banking industry, offering actionable insights that transcend the binary classification of positive or negative sentiments. The vision extends beyond classification; it encompasses a comprehensive understanding of the underlying sentiments and preferences expressed by customers. Ultimately, this endeavor seeks to catalyze a paradigm shift in how the banking industry interprets and responds to customer feedback, fostering an environment of continuous improvement and heightened customer satisfaction.

Dataset

The dataset utilized in this research comprises over 10,000 customer reviews sourced from 48 distinct banking institutions within the United States. This publicly available dataset can be accessed through the Kaggle community platform, offering a comprehensive collection of customer sentiments towards various banks. The dataset's primary objective is to serve as a resource for training and evaluating the sentiment analysis system proposed in this research. For easy accessibility, the dataset can be downloaded using the following link: https://www.kaggle.com/datasets/trainingdatapro/20000-customers-reviews-on-banks/data. The richness of this dataset not only allows for a detailed exploration of customer feedback but also ensures a diverse and representative sample, crucial for the effectiveness and generalizability of the sentiment analysis model.

Implementation method

Building a Sentiment analysis system using BERT to analyze bank customer reviews to classify customer reviews between positive reviews and negative reviews. The sentiment analysis system will be implemented as the steps below,

1. Importing the Dataset

To begin the analysis, it is essential to load the bank customer reviews dataset into the analysis environment, ensuring compatibility with widely used data structures such as CSV. This initial step sets the foundation for exploring and understanding the dataset's structure, which encompasses various features and sentiment labels. Familiarizing oneself with the intricacies of the dataset is crucial for subsequent analyses, enabling a comprehensive grasp of the information contained within and paving the way for effective sentiment analysis.

2. Data preprocessing

A critical phase in preparing the bank customer reviews dataset for analysis involves data cleaning to enhance the overall quality. This process revolves around the systematic removal of irrelevant information, aiming to streamline the dataset and ensure that only pertinent data is retained. By eliminating extraneous details, the dataset becomes more refined, facilitating a more focused and accurate analysis. Data cleaning serves as a crucial step in optimizing the dataset for subsequent tasks, contributing to the reliability and effectiveness of the forthcoming analyses, particularly in the context of sentiment analysis.

3. Tokenization

To prepare the textual data for analysis using BERT, it is imperative to implement WordPiece tokenization, a fundamental step in leveraging BERT's capabilities. This process involves breaking down words into subword tokens, a technique that allows for a more nuanced understanding of language nuances and context. By adopting BERT's tokenization mechanism, the textual data is transformed into a format that aligns with the model's input requirements. This crucial step ensures that the intricate details and relationships within the text are effectively captured, enabling BERT to process and analyze the information with its advanced natural language processing capabilities. The utilization of WordPiece tokenization and BERT's tokenization mechanism collectively plays a pivotal role in optimizing the input data, enhancing the model's ability to discern subtle linguistic nuances during the subsequent stages of analysis.

4. Model Training (BERT)

To facilitate the effective training and optimization of the sentiment analysis model using BERT, the dataset is systematically divided into three distinct sets: training, validation, and test sets. The training set serves as the foundation for instructing the BERT model, enabling it to learn patterns and relationships within the data. Simultaneously, the validation set plays a crucial role in fine-tuning the model by serving as a means to experiment with hyperparameter configurations. This iterative process involves adjusting key parameters to enhance the model's performance, with a particular focus on optimizing for the objectives of sentiment analysis. Additionally, the pre-trained BERT model undergoes a tailored fine-tuning process, specifically honing its capabilities for the sentiment analysis task at hand. The experimentations with hyperparameter configurations are guided by the insights gained from the validation set, ensuring that the model is refined and aligned with the desired sentiment analysis objectives. This comprehensive approach to dataset division, model training, and hyperparameter tuning forms a strategic framework for the optimization of the sentiment analysis model using BERT.

5. Classification

Once the BERT model is successfully trained, the next crucial step involves its application to classify new and previously unseen bank customer reviews into distinct positive or negative sentiments. The trained model, equipped with contextual embeddings, proves instrumental in capturing nuanced sentiment nuances within the reviews, allowing for a more fine-grained analysis of the expressed sentiments. It is essential to harness the contextual embeddings provided by BERT to grasp the subtle contextual variations in sentiment, thereby enhancing the model's accuracy in classifying diverse and nuanced expressions within the textual data. Additionally, depending on the nature of the sentiment analysis task, consideration should be given to the potential inclusion of a neutral sentiment class. Acknowledging the possibility of neutral sentiments contributes to a more comprehensive and realistic classification scheme, accommodating instances where customer reviews may not strongly lean towards either positive or negative sentiments, thereby offering a more nuanced and accurate sentiment analysis outcome.

6. Evaluation

The evaluation framework for this project is designed to meticulously assess the performance of the Convolutional Neural Networks (CNN) with BERT embeddings in the complex task of sentiment analysis. The primary metrics for evaluation are below,

- **Accuracy**: Measure overall correctness in sentiment predictions.

- **AUC-ROC**: Consider this metric to evaluate the model’s ability to discriminate between positive and negative sentiments, particularly relevant in imbalanced datasets.

The dataset will be split into training, validation, and test sets to facilitate a comprehensive evaluation process. During training, the model's efficiency and scalability will be monitored, with a focus on training time, resource utilization, and memory usage.

In the testing phase, accuracy will gauge the model's overall correctness in predicting sentiments, while the AUC score will provide nuanced insights into its ability to discriminate between positive and negative sentiments, especially in scenarios with class imbalances.

Qualitative analysis will complement quantitative metrics, with a keen eye on instances of misclassification to uncover patterns or contextual nuances that may inform refinements. This holistic evaluation aims to not only measure performance but also to illuminate the intricacies of sentiment classification within the dynamic landscape of customer reviews in the banking industry.

Tools & Technologies

The implementation will be carried out using relevant software/library packages, including the TensorFlow or PyTorch deep learning frameworks for CNN, and the Hugging Face Transformers library for BERT embeddings. The research can be conducted on platforms like Google Colab or any environment with GPU support for efficient model training.

Problem

Embarking on the adventure of working alone, especially when diving into a new research topic like studying feelings with BERT, can be pretty tough. Understanding this new way of thinking takes a lot of time and brainpower. Trying to learn new ways of doing things while also running a research project is like walking a tightrope. It needs to be really good at adapting, managing the time, and being flexible with thoughts to grasp all the details of what I am studying and the tools I am using.