Sentiment Analysis on Social Media Data for Brand Monitoring

Jonathan Agustin　　Fernando Calderon　　Juliet Lawton

University of San Diego

1 Introduction

In the contemporary digital environment, companies are increasingly leveraging social media platforms to track public perception of their brands. This process, known as **Brand Identification** and **Sentiment Analysis**, employs **Natural Language Processing** (NLP), text analytics, and computational linguistics to detect and extract subjective information related to specific brands. This project, which intersects the fields of NLP, **Deep Learning**, and **Classification**, aims to devise a model that can accurately categorize social media posts into positive, negative, or neutral sentiments towards specific brands, providing invaluable insights for brand tracking.

2 Problem Statement

The task is to analyze sentiments towards brands from social media text data, a complex and multifaceted classification problem. Given a set $X = \{x_1, x_2, \ldots, x_n\}$ of social media posts, our goal is to decipher the function $f : X \to Y$, where $Y = \{y_1, y_2, y_3\}$ represents the sentiment labels, with $y_1 =$ positive, $y_2 =$ negative and $y_3 =$ neutral respectively. This task requires a deep understanding of human language, including context, sarcasm, and irony, which are often present in social media posts. Machine Learning models, such as **BERT**, **RoBERTa**, and **GPT-2**, are effective in understanding and leveraging context in sentences, making them suitable for this task. However, the noisy nature of social media text data, including misspellings, abbreviations, emojis, URLs, and other non-alphanumeric characters, poses a significant challenge. The problem statement is to develop a model that can accurately identify brands and classify sentiments of a text as positive, negative, or neutral, extracted from various social media sources. This task is crucial for businesses aiming to track their brand sentiment in real-time, letting them respond promptly to customer feedback and market trends. Despite the challenges posed by the complexities of human language and the noisy nature of social media data, this project aims to explore these nuances, address the obstacles, and strive for an effective solution.

3 Proposed Methodology

We will use the Sentiment140 dataset from Stanford University, which includes 1.6 million tweets, for sentiment analysis. The dataset uses emoticons as a proxy for sentiment labels, authentically representing sentiment in social media text. For brand identification, we propose to use the Brand Sentiment dataset from Surge AI, which includes over 1.5 million tweets mentioning over 2000 popular brands. Our methods are grounded on the robust capabilities of pre-trained language models in response to the complex challenge of human language understanding and brand identification. We will harness widely recognized models from Hugging Face's model hub, such as BERT, DistilBERT, RoBERTa, or GPT-2. These transformer-based models, constructed on the foundation of large and diverse text corpora, provide a rich understanding of language nuances and brand references. Our methods consist of three main stages:

1. **Data Preprocessing.** The first stage of our methods is Data Preprocessing, which involves transforming raw text into a structured format ready for modeling. This stage is crucial as it talks to social media data's inherent unstructuredness and noisiness, bridging the gap between human-generated content and machine-understandable language. We undertake certain preprocessing steps to ensure every piece of data feeds into our model in the most ideal form. The first step is **Noise Handling**, which removes irrelevant parts in our dataset, such as URLs, user handles, emojis, and non-alphanumeric characters. These parts rarely contribute to the broader understanding of sentiment and could

introduce noise into our model. The second step is **Tokenization**, where we break down our text data into smaller units called tokens. These tokens, which correspond to words or subwords, represent the working units for our language models. They provide a granularity level at which our model learns the contextual relationship between different text parts.

2. **Model Training.** The second stage involves the training and fine-tuning of our pre-trained language model on our processed dataset. We use the power of TensorFlow to ensure an efficient training process. Distinct from models trained from scratch, our chosen pre-trained models have undergone a phase of initial training on a vast corpus of text. They have captured the rich interplay of words and context in various linguistic scenarios. **Fine-tuning** refers to adjusting this pre-trained model's weights further, this time on our specific task. We thus align the model's capabilities to the nuances and features inherent to the sentiments in a social media text.

3. **Model Evaluation.** The final stage evaluates our fine-tuned model's performance. Evaluation is crucial in measuring how successfully our model has learned to predict sentiments in the social media context. We use standard classification metrics: **Accuracy**, **Precision**, **Recall**, and **F1 Score**. Accuracy gives us a measure of how correctly our model has classified the sentiments. Precision provides an understanding of how likely a positive prediction is positive. Recall tells us what proportion of actual positives is correctly classified. The F1 Score balances Precision and Recall and is best used when the class distribution is unbalanced. These metrics collectively provide a comprehensive view of our model's classification performance, enabling us to gauge its effectiveness and reliability.

4 Expected Behaviors/Outcomes

- **Model Performance.** The primary expectation is that our model will perform well in accuracy, precision, recall, and F1 score. These metrics will provide a comprehensive view of our model's performance. Accuracy will tell us how often our model is correct. Precision will inform us about the proportion of correct identifications. The recall will give us the ratio of actual positives that were identified correctly. The F1 score, the harmonic mean of precision and recall, will provide us with a metric that balances these considerations.

- **Understanding of Language Nuances and Brand Identification.** Beyond the quantitative metrics, we expect our model to show a nuanced understanding of human language and an ability to accurately identify brands. Given the complexity and ambiguity of human language, especially in social media posts, this is a non-trivial expectation. We expect our model to handle sarcasm, irony, and context-specific meanings common in social media posts, and correctly identify brand references within these posts.

- **Real-world Applicability.** Finally, we expect that our model will have real-world applicability. The ultimate goal of this project is to provide businesses with a tool to accurately and efficiently analyze sentiment from social media data. We expect that our model can provide actionable insights to help companies to make more informed decisions. For example, a business could use our model to track public sentiment about their brand in real-time, compare their performance with competitors based on public sentiment, or analyze customer feedback to improve their products or services.

References

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., … & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. *OpenAI Blog, 1(8)*.

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Hutto, C.J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014*.