Graph Neural Networks for Enhancing Sentiment Analysis

Yujia Wang yuw103@ucsd.edu

Yuxin Guo y5guo@ucsd.edu

Yunfan Long yulong@ucsd.edu

Gal Mishne gmishne@ucsd.edu

Yusu Wang yusuwang@ucsd.edu

Abstract

In this study, we want to introduce a novel approach to sentiment analysis by integrating Graph Neural Networks (GNNs), specifically Graph Attention Networks (GAT), with traditional Natural Language Processing (NLP) techniques. By focusing on the challenges of interpreting long-range dependencies within text, particularly in the context of social media interactions, we propose a model that leverages the relational and textual information present in data from college SubReddit forums before and during the COVID-19 pandemic(Yan and Liu 2021). Our methodology employs RoBERTa for text embedding, followed by a GAT to incorporate the structural relationships between messages, enhancing the sentiment analysis capability beyond the limitations of conventional models. The results demonstrate a statistically significant improvement in identifying and relating sentiments across distant text parts, underscoring the potential of GNNs in capturing complex interactions within text data. This research contributes to the evolving field of sentiment analysis by showcasing the effectiveness of combining GNNs with existing NLP frameworks, offering insights into the nuanced emotional landscape of digital communications during a global crisis.

Website: https://yuxinguo13.github.io/DSC_capstone_website
Code: https://github.com/yunfanlong/DSC180B-B12

1 Introduction

In the evolving domain of Natural Language Processing (NLP), sentiment analysis emerges as a crucial technique for extracting and interpreting the complex emotional undertones in textual content. This process is especially significant when we want to analyze data from social media, where textual exchanges are not just abundant but are also woven with intricate contextual and relational threads. Despite significant advancements in NLP, traditional models often encounter limitations when dealing with long-range dependencies in text. These limitations hinder the ability to fully grasp the sentiment landscape, particularly when sentiments span across distant parts of the text or are embedded in complex social interactions.

To bridge this gap, our study proposes a pioneering approach that marries the capabilities of Graph Neural Networks (GNNs), with a spotlight on Graph Attention Networks (GAT), with the linguistic prowess of advanced NLP techniques such as BERT embeddings (Devlin et al. 2018). This hybrid model is engineered to transcend the conventional boundaries of sentiment analysis by adeptly navigating both the textual and relational dimensions of social media data. Our focus is particularly drawn to the discourse surrounding the COVID-19 pandemic within college SubReddit forums, offering a rich tapestry of emotional expressions ripe for analysis.

Graph Attention Networks (GAT) stand at the forefront of this endeavor, heralded for their ability to astutely focus on pivotal parts of a graph structure. This is achieved through a dynamic weighting of the influences between nodes, or in the context of our study, between different pieces of textual content. Such a mechanism is invaluable for identifying long-range interactions within texts, a feat that traditional models like RoBERTa might not fully accomplish due to potential oversights of subtle, yet crucial, contextual signals. By weaving together the textual embeddings generated by RoBERTa with the relational insights gleaned through GAT, our model not only retains the semantic richness provided by RoBERTa but also imbues it with a deeper understanding of the contextual and relational fabric within which these sentiments are embedded.

The training of the GAT model involves the intricate process of integrating relational information from the collected data into a graph structure, where each message or textual fragment is treated as a node, and the interactions between these messages, such as replies or thematic connections, form the edges. This setup allows GAT to leverage the attention mechanism to enrich the embeddings from RoBERTa with relational context, significantly enhancing the model's ability to discern sentiments that are influenced by complex intertextual dynamics.

Our approach aims to redefine the benchmarks in sentiment analysis by offering a more nuanced and comprehensive understanding of sentiments in social media. By harnessing the synergistic potential of GNNs and NLP, we aspire to illuminate the intricate web of human emotions and interactions in the digital age, particularly against the backdrop of a global crisis like the COVID-19 pandemic.

2 Methods

2.1 Feature Extraction

In the context of sentiment analysis from social media platforms, feature extraction plays a pivotal role in interpreting the nuanced and often informal language used in such environments. Our methodology seeks to refine the process of feature extraction by employing state-of-the-art natural language processing techniques, specifically tailored to analyze the content of Reddit comments and submissions. At the heart of our approach is the utilization of the RoBERTa (Robustly Optimized BERT Pretraining Approach) model, celebrated for its proficiency in grasping the context and semantics embedded in textual data. The selection of RoBERTa is informed by its demonstrated capability to unravel the intricate contextual meanings within texts, a vital attribute for decoding the ambiguous and emotionally charged discourse typical of social media interactions.

Our feature extraction journey begins with the deployment of a RoBERTa model that has undergone specialized pre-training on an extensive corpus of around 58 million Twitter messages. This pre-training step ensures that the model is adept at handling internet slang, abbreviations, and non-standard spelling variations, which are prevalent in online communications.

In the feature extraction phase, we channel each piece of text through the RoBERTa model (Liu et al. 2019) to derive high-dimensional embeddings as shown in Figure 2.1. These embeddings encapsulate the text's semantic and syntactic intricacies in a dense vector format, providing a rich representation of the underlying content. This process of transforming raw text into fixed-length vector embeddings is pivotal, as it standardizes the input for subsequent analytical procedures. These vector representations are then poised for integration into more complex machine learning frameworks, such as Graph Attention Networks (GAT)(Veličković et al. 2017) and ensemble methodologies, which are elaborated upon in subsequent sections of our analysis.

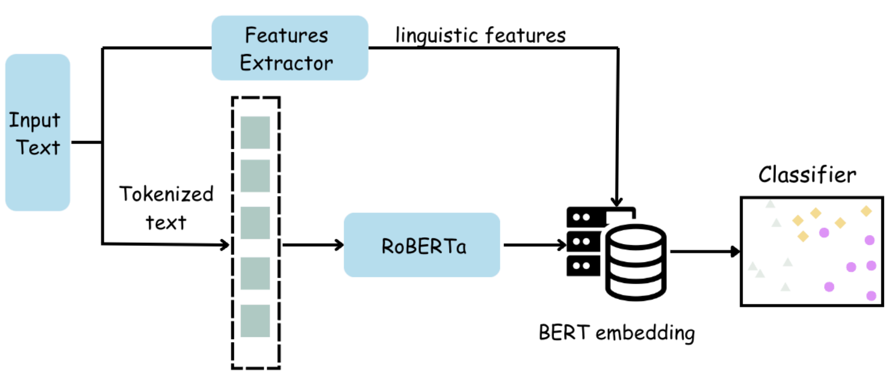

Figure 1: RoBERTa Feature Extraction

2.2 Graph Architecture

Our graph architecture is designed to capture the relationships between messages within our dataset, integrating the semantic feature of textual data with the ability of graph theory. This section details the construction of our graph, highlighting the innovative integration of node and edge features derived from advanced NLP embeddings and relational context.

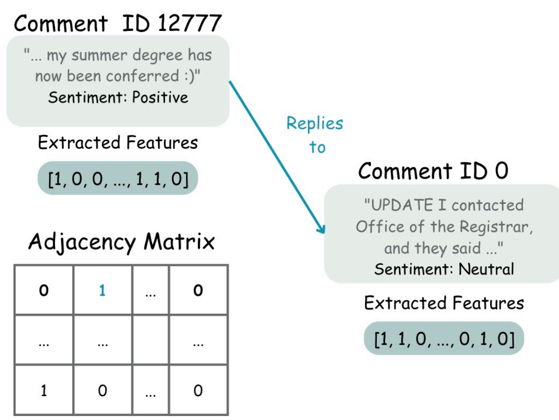

Figure 2: Example of Weighting Two Model Outputs

2.2.1 Node Feature (Roberta Embedding)

The cornerstone of our graph architecture is the utilization of RoBERTa embeddings as node features. Each node in the graph represents a comment in our dataset, and its corresponding feature vector is a embedding representation extracted from the RoBERTa model. These embeddings are not mere numerical representations but carry the contextual richness of the text, including semantic nuances, syntactic information, and inherent meanings captured through RoBERTa's self-attention mechanisms.

By embedding the textual data into a high-dimensional semantic space, we enable the GAT model to perform analyses based on the content of each node (each comment). This method allows for the differentiation between nodes in terms of the subtle textual similarities and distinctions that are crucial for natural language understanding.

Furthermore, the dynamic nature of RoBERTa embeddings, which are sensitive to the context in which words appear, provides a robust foundation for our graph. This means that even if two documents contain similar words, the differences in their usage and context are captured in their respective embedding, allowing our model to distinguish the difference between them effectively.

2.2.2 Edge Feature Representation and Adjacency Matrix Creation

In our graph structure, the adjacency matrix is pivotal for encoding the interactions between nodes, where each node corresponds to a comment within our subreddit dataset. The edges, represented within this matrix, signify the direct message-reply relationships, capturing the essence of interaction.

Our adjacency matrix is a square matrix where each row and column represent a node (comment) in the graph. The elements of the matrix indicate whether a direct interaction (reply) exists between two nodes. For every pair of comments where one is a direct reply to the other, the corresponding element in the adjacency matrix is set to 1. This is determined by examining the parent_id or link_id fields in the dataset, which indicate reply relationships.

Binary Interaction Encoding

- 1 denotes the presence of a direct reply from one comment to another, establishing a directed edge between the two nodes.
- 0 indicates no direct interaction between the nodes.

Symmetric vs. Asymmetric

- Asymmetric adjacency matrix is used for directed graphs, where the direction of the reply matters (from the commenter to the original post).
- Symmetric adjacency matrix may be employed to represent undirected edges, implying mutual influence, though this approach is less common in message-reply contexts.

The adjacency matrix shown in Figure 2.2 thus captures the interaction of the subreddit's conversational threads, highlighting how users engage with one another through replies. This structure allows the Graph Attention Network (GAT) to dynamically weigh these interactions. By connecting the graph's connectivity through this adjacency matrix, our model adeptly navigates the dataset's social interaction, leveraging both the content through the node features and the interaction patterns from the edge representation.

2.3 GAT Model Architecture

The GAT model architecture is central to our approach for analyzing social media interactions, particularly those within university communities on Reddit. In this study, we implemented our GAT model using TensorFlow. Our model involves multiple attention heads operating in parallel to compute attention coefficients. These coefficients are determined through a shared attention mechanism, which includes a learnable linear transformation followed by a LeakyReLU nonlinearity. This setup enables the model to assign varying levels of importance to different nodes' features, based on their relevance to the task at hand.

For the output layer, the GAT model is tasked with classifying the sentiment of each node, represented by comments or submissions, into specific categories. The training process involves adjusting both the attention coefficients and the neural network parameters to minimize the loss function, typically cross-entropy, between the predicted and actual sentiment labels.

Furthermore, to ensure the robustness and generalizability of our model, we employ a K-Folds cross-validation strategy. The data is partitioned into four folds, with each serving as the test set in turn while the model is trained on the remaining data. This approach mitigates potential biases and overfitting, providing a more reliable assessment of the model's performance. Then, we outline a systematic exploration of various parameter combinations, including learning rates and regularization terms, to identify the optimal configuration for sentiment analysis in this context.

Through this refined GAT model architecture, we aim to capture the intricate relational dynamics within Reddit's university communities, providing more deep insights into sentiment trends and patterns that emerge from these online interactions.

(A graph showing the architecture will be added later)

2.4 Ensemble of RoBERTa and GAT

In our ensemble model using Model Stacking technique(Wolpert 1992), we adopt a logistic regression meta-model to integrate the predictive outputs from the Graph Attention Network (GAT) and RoBERTa models. This model stacking strategy harnesses the distinct capabilities of GAT and RoBERTa in analyzing relational graph data and textual content, respectively.

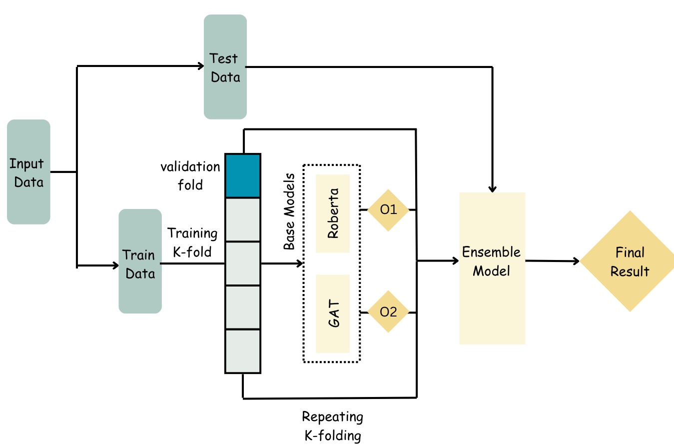

Figure 3: Architecture of Ensemble Model

The ensemble approach begins by scaling the raw prediction probabilities from GAT and RoBERTa. Specifically, for a given message i from school j in year k, the predicted probability of expressing non-negative sentiment by model m (either GAT or RoBERTa) is denoted as $p_{ijk}^{(m)}$. These raw probabilities are then scaled to a common scale using the following transformation:

$$p_{ijk}^{(m)}(1) = 0.5 \times \left(\frac{p_{ijk}^{(m)} - c^{(m)}}{1 - c^{(m)}} + 0.5 \right) \quad (1)$$

Here, $c^{(m)}$ represents the optimal cutoff for model m, determined by maximizing the geometric mean of sensitivity and specificity.

Subsequently, the scaled probabilities $p_{ijk}^{(1)}$ (from RoBERTa) and $p_{ijk}^{(2)}$ (from GAT) serve as inputs to the logistic regression meta-model shown as the Figure 5. The logistic regression is formulated to predict the final sentiment based on these probabilities. The logistic model is represented as:

$$\log\left(\frac{\Pr(\text{non-Negative})}{1 - \Pr(\text{non-Negative})}\right) = \beta_0 + \beta_1 p_{ijk}^{(1)} + \beta_2 p_{ijk}^{(2)} \quad (2)$$

In this equation, Pr(non-Negative) stands for the probability of the message being classified as non-negative. The coefficients $\beta_0, \beta_1, \beta_2$ are estimated during the training process, encapsulating the relative contributions of the scaled probabilities from RoBERTa and GAT to the final sentiment prediction.

The output of the logistic regression meta-model, $\bar{p}_{ijk}$, signifies the consolidated probability that message i is non-negative. The final sentiment classification is determined by applying a threshold of 0.5 to $\bar{p}_{ijk}$; messages with $\bar{p}_{ijk} \geq 0.5$ are categorized as non-negative, while the rest are deemed negative. Through this process, the ensemble model adeptly combines the insights derived from the graph-based and text-based analyses provided by GAT and RoBERTa, leading to enhanced sentiment classification performance (Yan and Liu 2021).

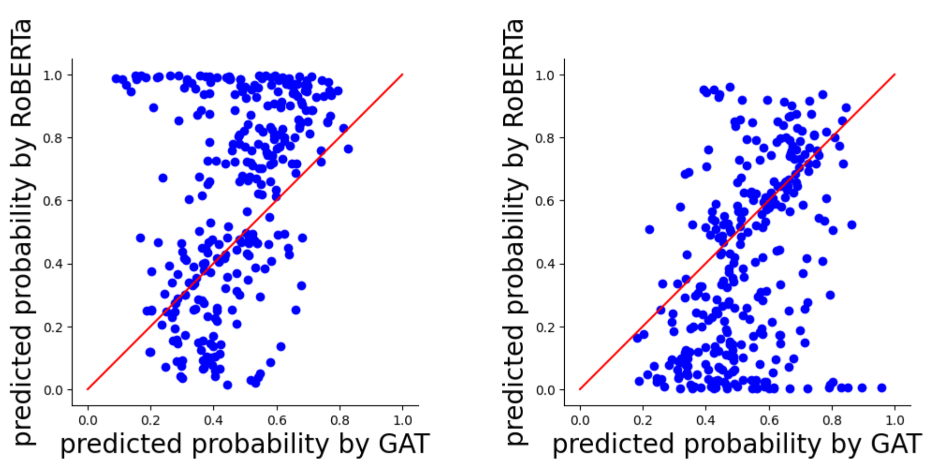

Figure 5: Example of Weighting Two Model Outputs

3 Results

The empirical analysis of sentiment classification across various educational institutions before and during the COVID-19 pandemic reveals insightful trends and patterns. Our methodology, which integrates Graph Attention Network, with RoBERTa embeddings, was applied to a dataset extracted from college SubReddit forums.

Our findings, summarized in Table 1, highlight the nuanced shifts in sentiment at individual schools from 2019 to 2020. For instance, UCLA exhibited an increase in the classification accuracy rate (CAR) from 48.0% in 2019 to 60.0% in 2020, indicating a marked improvement in our model's ability to discern sentiment accurately during the pandemic year. Similar trends were observed at Columbia University, with the CAR increasing from 50.0% to 56.0%. Conversely, some institutions like UCSD and UCB saw a decline in CAR, suggesting varied impacts of the pandemic on sentiment across different university communities. The F1 scores, precision, and recall metrics further delineate the model's performance, underscoring the complexity of sentiment analysis in the context of evolving social dynamics during the pandemic.

Table 1: Prediction Acurracies By Schools

	Year	UCLA	UCSD	UCB	UofM	Harvard	Columbia
CAR*	2019	0.480	0.500	0.540	0.580	0.460	0.500
	2020	0.600	0.400	0.400	0.480	0.480	0.560
F1†	2019	0.500	0.510	0.489	0.588	0.426	0.490
	2020	0.615	0.375	0.444	0.409	0.381	0.522
specificity	2019	0.481	0.500	0.550	0.577	0.455	0.500
	2020	0.593	0.391	0.414	0.474	0.471	0.571
precision‡	2019	0.520	0.520	0.440	0.600	0.400	0.480
	2020	0.640	0.360	0.480	0.360	0.320	0.480
recall††	2019	0.440	0.480	0.640	0.560	0.520	0.520
	2020	0.560	0.440	0.320	0.600	0.640	0.640

In dissecting the performance of different models utilized in our study for the year 2019 (Table 2), we observed a consistency in the CAR across all of our thress models: GAT, RoBERTa, and ESM(ensemble model). The overall accuracy hovered around 50%, with slight variations across different institutions. This uniformity underscores the challenges faced by these models in capturing the intricate sentiment dynamics within collegiate digital discourse during the pre-pandemic period.

Transitioning to the year 2020 (Table 3), the landscape of sentiment classification experienced noticeable shifts. The GAT model, for instance, demonstrated a CAR of 60.0% at UCLA but only 40.0% at UCSD and UCB, reflecting the diverse impacts of the pandemic on student sentiment. BERT and ESM models exhibited similar variability, further emphasizing the nuanced nature of digital communications during a global crisis.

Table 2: Prediction Acurracies By Models in 2019

Model	Metric	UCLA	UCSD	UCB	UofM	Harvard	Columbia	Overall
GAT	CAR	0.48	0.50	0.54	**0.58**	0.46	0.50	0.498
	F1	0.50	0.51	0.49	0.59	0.43	0.49	0.482
BERT	CAR	0.46	0.52	**0.64**	0.56	0.42	0.46	0.497
	F1	**0.53**	**0.56**	**0.65**	**0.61**	0.49	0.49	**0.534**
ESM	CAR	**0.54**	**0.54**	0.48	0.42	**0.56**	**0.52**	**0.505**
	F1	0.49	0.50	0.44	0.36	**0.56**	**0.51**	0.489

Table 3: Prediction Acurracies By Models in 2020

Model	Metric	UCLA	UCSD	UCB	UofM	Harvard	Columbia	Overall
GAT	CAR	**0.60**	0.40	0.40	0.48	0.48	**0.56**	0.498
	F1	**0.62**	0.38	0.44	0.41	0.38	0.52	0.481
BERT	CAR	0.46	0.46	0.44	0.50	0.48	0.56	0.497
	F1	0.53	0.53	**0.46**	0.44	0.52	**0.59**	**0.533**
ESM	CAR	0.40	**0.64**	**0.52**	**0.51**	**0.48**	0.50	**0.505**
	F1	0.32	**0.63**	0.43	**0.55**	**0.54**	0.51	**0.488**

Our results illuminate the significant role of context and temporal factors in sentiment analysis. The model performance, across schools and years observed, therefore, points out to a multi-faceted sentiment within the academic community, especially when the scenario is to be taken in the wake of unprecedented challenges such as the COVID-19 pandemic. All this not only reasserts the possible effectiveness of an integrated approach such as GNNs combined with NLP techniques in sentiment analysis but also reasserts the potential of such hybrid models for understanding the complex emotional landscapes during crises presented by digital communications.

4 Discussion

In this study, we use a blend of the cutting-edge of Machine Learning (ML) techniques with more traditional statistical modeling in a responsible way to discuss the effect of COVID-19 on the emotional and psychological well-being of the people affiliated with higher education insititutes. Actually, the real accomplishment of the present study is not to prove anew the difficulties that unstructured and observational data pose to the real world, considering that it uses data from Reddit, but the present study exhibits an integrated approach in sentiment analysis: RoBERTa for text embedding and Graph Attention Networks (GAT) to capture relational information.

However, the study is not without its limitations. One important consideration is the study's reliance on data from a select group of R1 doctoral universities, which may not fully represent the diversity of experiences within the broader higher education landscape. Including a wider array of institutions, including those with different academic and research focuses, could offer more insights into how various educational contexts influenced sentiment during the pandemic. Additionally, the focus on university-associated subreddits may limit the generalizability of the findings to broader populations. Expanding the data sources to include a wider range of social media platforms and communities could provide a more comprehensive understanding of the pandemic's impact on public sentiment.

5 Conclusion

Building upon our new approach for sentiment analysis, we integrate Graph Neural Networks(Scarselli et al. 2008) more specifically, Graph Attention Networks (GAT), as the representative model with traditional Natural Language Processing (NLP) techniques, such as RoBERTa embeddings. The conclusion mentions the important contributions of the study and some future work.

The results demonstrate an improvement in sentiment analysis capability, particularly in interpreting long-range dependencies within text data derived from college SubReddit forums before and during the COVID-19 pandemic. This advancement is primarily attributed to the study's novel methodology that leverages the synergistic potential of GNNs with existing NLP frameworks, thereby enriching the model's ability to capture the nuanced emotional landscape inherent in digital communications.

The core achievement of this research lies in its ability to capture the complex interplay of textual content and relational information exchanged within social media interactions in order to afford a more nuanced and comprehensive understanding of the sentiments. This is evidence that the model can effectively identify and associate sentiments from very far-flung parts of the text, thus underscoring the potential of GNN in capturing intricate text data interactions.

Thus, we believe that this study opens up a significant way towards sentiment analysis, and this is also very high importance for understanding the impacts of emotions due to global crises, like in the case of the COVID-19 pandemic, over digital platforms. The methodological approach set by this research really establishes a new benchmark in the field of sentiment analysis and encourages any following research that could further refine and expand the set framework toward contributing to a wider understanding of human emotions in the modern digital era.

6 AI Use Statement

6.1 Tools Used and Their Application

In the development of this project, two primary AI tools were utilized: ChatGPT and Copilot.

6.1.1 Usage of Copilot

Copilot served as an on-the-fly assistant, particularly in reminding me of certain aspects of code, including class methods and attributes. It was instrumental in navigating through complex code structures and understanding confusing concepts.

6.1.2 Usage of ChatGPT

ChatGPT played a crucial role in converting our understanding into executable code. It provided code implementations based on our instructions, especially helpful given our limited familiarity with PyTorch Geometric (PyG) and Tensorflow. While ChatGPT expedited the coding process and helped overcome learning challenges, it also introduced some issues. For instance, it directly used `data.x` for training on every network for the IMDB dataset, where the feature was not inherently present, necessitating significant debugging from our end.

6.2 Report Writing Assistance

ChatGPT's contribution to report writing was substantial. It assisted in organizing, drafting, and refining the report, enhancing the coherence and flow of the content. The tool was particularly adept at paraphrasing and providing insights into the methodologies and optimization strategies used in our models.

6.3 Response to Feedback

Regarding the feedback on the extent of ChatGPT's involvement in the report writing:

Extent of Original Content: Approximately 70% of the report is composed in our own words. ChatGPT acted more as an assistant than a primary writer.

Examples of Prompts Used:

- *Prompt 1:* "Please polish our introduction and suggest any additions based on your knowledge."
- *Prompt 2:* "Can you help rephrase this paragraph for better clarity?"

- *Prompt 3:* ”I need assistance in summarizing the optimization parameters of our model.”
- *Prompt 4:* ”Could you provide a brief explanation of...”
- *Prompt 5:* ”Help me draft a section comparing different neural network models.”
- *Prompt 6:* ”Suggest some potential improvements or future work based on our project findings.”
- *Prompt 7:* ”Assist in converting our analysis results into a coherent written format.”
- *Prompt 8:* ”Provide guidance on structuring the discussion section of our report.”

References

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805*

Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. "Roberta: A robustly optimized bert pretraining approach." *arXiv preprint arXiv:1907.11692*

Scarselli, Franco, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2008. "The graph neural network model." *IEEE transactions on neural networks* 20 (1): 61–80

Veličković, Petar, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. "Graph attention networks." *arXiv preprint arXiv:1710.10903*

Wolpert, David H. 1992. "Stacked generalization." *Neural networks* 5 (2): 241–259

Yan, Tian, and Fang Liu. 2021. "Sentiment Analysis and Effect of COVID-19 Pandemic using College SubReddit Data." *arXiv preprint arXiv:2112.04351*

Appendices

A.1 Project Proposal

For a domain NLP and machine learning, our project targets a critical challenge in the field: the limitations of current sentiment analysis models in comprehending context and sentiment connections, especially in lengthy and complex text structures. Traditional models often struggle with capturing the sentiment in texts where the expression of emotions or opinions is not straightforward or is distributed over large segments of the text. This issue becomes particularly pronounced in texts such as lengthy social media posts, detailed product reviews, or extensive articles, where sentiments can be context-dependent.

To address this challenge, our project proposes an innovative approach that combines the strengths of BERT (Bidirectional Encoder Representations from Transformers) with Graph Neural Networks (GNNs). BERT, known for its effectiveness in understanding the context of words in a sentence, provides a robust foundation for processing and understanding text data. However, BERT has limitations in capturing relationships and dependencies in texts that extend beyond immediate word sequences. This is where Graph Neural Networks come into play.

GNNs are particularly adept at modeling relational data. By integrating them with BERT, we aim to create a model that not only understands the context within sentences but also effectively maps and interprets the complex, interwoven sentiment structures that span across longer segments of text. This method allows us to capture the broader sentiment landscape within a text, linking disparate yet contextually relevant parts to offer a more comprehensive sentiment analysis.

Our project utilizes data from college subreddits over different time periods. This choice offers a rich, varied, and dynamic dataset, reflecting a wide range of emotions and topics that are ideal for testing the model's capabilities. Subreddits, characterized by their informal and diverse nature of communication, present an excellent opportunity to challenge and refine our model. The temporal aspect of the data also allows us to analyze how sentiments and topics evolve over time, further enriching our understanding and providing a rigorous testbed for our model.

Investing time in this project is crucial as it has the potential to set a new standard in sentiment analysis, particularly in dealing with complex and extended texts. The success of this project would be a big stride in this NLP field, opening the door for more context-aware, fine-grained sentiment analysis tooling.