Enhancing Conversational Analysis in the DKTC Dataset

TaeHoon Lee
Aiffel

September 2023

1 Abstract

Conversational analysis in text-based dialogues is a critical area of study, and the DKTC dataset offers a unique context for examining various conversational dynamics. This dataset includes threat conversations, workplace harassment conversations, other harassment conversations, and extortion conversations. In this paper, we propose a novel approach to enhance conversational analysis, specifically focusing on identifying perpetrators and victims within the dialogues. Our approach involves segmenting conversations based on speaker changes, aided by the '\n' delimiter, and utilizing the OpenAI API for classification. While our approach reduces the average dialogue length and shows promise, it also reveals challenges associated with multi-speaker conversations and underscores the importance of retaining victim data for improved classification. This research contributes to the field of dialogue analysis, offering insights into the complexities of conversational context and the roles of perpetrators and victims.

2 Introduction

Conversational analysis plays a pivotal role in understanding and extracting valuable insights from text-based dialogues. In the realm of dialogue analysis, the DKTC dataset has emerged as a valuable resource for studying various conversational dynamics, including threat conversations, workplace harassment conversations, other harassment conversations, and extortion conversations. This dataset presents a unique set of challenges and opportunities for researchers and practitioners alike.
In this paper, we embark on a journey to enhance the performance of conversational analysis models, with a specific focus on the DKTC dataset. Our primary goal is to accurately identify the roles of perpetrators and victims within the dialogues. To achieve this, we employ a novel data preprocessing approach. We segment the conversations based on speaker changes, leveraging

the '\n' unit as a delimiter. This strategy has the potential to enhance the model's understanding of conversational context.
To further refine our analysis, we integrate the OpenAI API into our workflow. The API assists in the classification of the dialogues, aiding in the identification of perpetrators and victims. We configure the system with precise instructions, such as "Based on the messages, identify the perpetrator and the victim between A and B," and process the results in a structured JSON format.
Our research journey unfolds with intriguing findings and challenges. While our data preprocessing approach successfully reduces the average dialogue length from 400 to 200, it raises complex issues associated with multi-speaker conversations. The presence of conversations involving more than two speakers introduces ambiguities in dialogue segmentation, making it challenging to determine the roles of perpetrators and victims.
Additionally, we discover the significance of victim data, especially in cases of "other harassment" conversations, where role distinctions are less clear-cut. Victim data provides valuable context that aids in classification.
As our dataset size decreases, we also observe a blurring of class boundaries, impacting the model's performance. This issue is closely tied to the importance of retaining victim data for more accurate analysis.
In summary, this paper navigates through the intricacies of conversational analysis, highlighting the complexities of multi-speaker dialogues and the importance of context in determining the roles of perpetrators and victims. Our findings pave the way for further research in dialogue analysis and contribute to the development of more accurate conversational models.

3 Method

There are several ways to improve the model's performance. You can introduce new data or enhance performance by replacing some of the models used in the training process. However, I have opted for data preprocessing. The data consists of conversations between perpetrators and victims, and by extracting and using only the perpetrator's data, we can improve the model's performance. At the same time, this reduces the amount of data but enhances the training speed.

3.1 Data

In the DKTC dataset, there are four classes: threat conversations, workplace harassment conversations, other harassment conversations, and extortion conversations. The class distribution is balanced, with approximately 4,000 data instances. Approximately 99.5 of the data is in Korean. The data follows a human conversation format, with conversations segmented by \n units. The number of participants in each conversation is not fixed, but there are at least two individuals engaged in the conversation.

3.2 Precondition

We assume the presence of both perpetrators and victims in the conversation. In this conversation, we assume that only two individuals, one perpetrator and one victim, are engaged in the conversation. In reality, there may be more people participating in the conversation.

3.3 Experiment Method

We preprocess the data by segmenting it based on '\n' units, assuming a change of speaker. For example, the first sentence corresponds to person A, the second sentence to person B, and so on. This preprocessing step separates the dialogues of individuals, such as A's conversation and B's conversation.
We leverage the OpenAI API to classify the dialogues of A and B. We seek to determine who the perpetrator and who the victim is within the conversation. To achieve this, we need to configure the system for the OpenAI API with the following instructions: "Based on the messages, identify the perpetrator and the victim between A and B. If it's unclear who the perpetrator is, please label it as 'unknown.' The desired output format is in JSON, for instance: 'Perpetrator': A, 'Victim': B."
When using the API, the return value will be in the format of 'Perpetrator': A, 'Victim': B.
Through these steps, the data has been reduced from an average length of 400 to 200. Consequently, we can reduce the maxlen length for input to the model.

4 Result

The performance was measured using accuracy as the metric. The average accuracy for the data without using the OpenAI API was 0.89, while the data processed with the OpenAI API resulted in an average accuracy of 0.84. This represents a decrease in performance from 0.89 to 0.84. Let's analyze the possible reasons for this decline:
Multi-Speaker Conversations: One contributing factor is the assumption of two-person conversations, while the dataset contains many conversations with more than two speakers. For example, cases with multiple perpetrators, such as "Perpetrator 1: Give me some money!", "Perpetrator 2: Should I check your pockets?", "Victim: I'm sorry!" create issues during preprocessing. After segmentation, you get dialogues like "A: Give me some money!", "B: Should I check your pockets?", "A: I'm sorry!" This can result in conversations that are confusing and ambiguous even for humans, making it difficult to determine who the perpetrator and victim are. Extracting only perpetrator conversations from such cases would lead to nonsensical dialogues like "Give me some money, I'm sorry!" and negatively affect model training.
Information in Victim Data: Another reason is that victim data contains meaningful information. Among the target classes, there is a category called "other harassment" where the distinction between perpetrator and victim is

not clear-cut, and context from surrounding sentences must be used to infer roles. Therefore, the victim's data can be helpful in classifying labels.
Decreased Data and Class Boundaries: Lastly, as the data decreased, the boundaries between classes became less distinct. This is related to the second reason. With the reduction of victim data, the class boundaries in the conversations become less clear, making it challenging to differentiate between the classes accurately.
In summary, the decrease in performance may be attributed to the complexities arising from multi-speaker conversations, the importance of victim data, and the blurred class boundaries resulting from data reduction. Further refinements and strategies may be needed to address these issues and improve model performance.

5 Conclusion

In conclusion, our study focused on enhancing the performance of a conversational analysis model, specifically aimed at detecting perpetrators and victims in text-based dialogues. We employed a data preprocessing approach, segmenting conversations based on speaker changes using \n units, and utilized the OpenAI API to classify the roles of individuals in the dialogues. Our findings reveal several crucial insights:
Firstly, the presence of multi-speaker conversations, where more than two individuals are engaged, posed challenges during preprocessing. This resulted in ambiguous and confusing dialogue segments, making it difficult to determine the roles of perpetrators and victims, even for human annotators. Extracting only perpetrator conversations from such cases led to nonsensical dialogue segments that could negatively impact model training.
Secondly, victim data contained valuable information, especially in cases of "other harassment" conversations, where role distinctions were not clear-cut. Utilizing victim data proved to be beneficial in the classification of labels, as it helped resolve contextual ambiguities.
Lastly, the reduction in data volume led to blurred class boundaries within conversations, making it challenging to accurately distinguish between different classes.
While our approach succeeded in reducing the average dialogue length from 400 to 200, it also resulted in a decrease in model accuracy, from 0.89 to 0.84. This decline in performance underscores the complexity of the task, especially in multi-speaker conversations, and the importance of retaining valuable victim data for classification. Future work in this area should focus on refining preprocessing strategies and addressing these challenges to improve model performance further.
In summary, our study sheds light on the intricacies of processing multi-speaker conversations for conversational analysis and underscores the importance of context, especially in cases where the roles of perpetrators and victims are not immediately clear. These insights can guide future research in

the field of dialogue analysis and contribute to the development of more accurate and effective conversational models.

6 References

[1] L. Perez and J. Wang, “The effectiveness of data augmentation in image classification using deep learning,” arXiv preprint arXiv:1712.04621, 2017.