Refining Decision Prediction in Persuasion Games: A Hybrid Approach Using Pre-trained Transformer

Hadar Sugarman
Faculty of Data and Decision Sciences
Technion— Israel Institute of Technology
HADAR.SUG@CAMPUS.TECHNION.AC.IL

Tomer Grinberg
Faculty of Data and Decision Sciences
Technion— Israel Institute of Technology
GRTOMER@CAMPUS.TECHNION.AC.IL

Abstract

Recent advances in Large Language Models (LLMs) have spurred interest in designing agents capable of interacting with both human and artificial agents. This study focuses on a critical aspect of these agents: predicting human decisions in off-policy evaluation (OPE) within language-based persuasion games, where the agent aims to influence its partner's decisions through verbal messages. We used a dataset of 87,000 decisions collected by Shapira et al.(2024). from humans engaging in a repeated decision-making game with artificial agents. Our study aims to refine the process of predicting decisions made by decision-makers (DMs) in persuasion games. By fine-tuning pre-trained transformer models for this specific task and adopting a data-centric approach, we achieved a maximal accuracy score of 79.5

1 Introduction

Understanding human choice in language-based persuasion games is a critical area of research in natural language processing and human-computer interaction. This study focuses on predicting human behavior when faced with unobserved opponents in such games. Instead of determining optimal policies, our goal is to predict the choices made by human agents based on their interactions with various artificial agents in the same game. This setup constitutes an off-policy evaluation (OPE) scenario, where expert agents interact with human decision-makers (DMs) in a non-cooperative game environment.

In our study, we fine-tuned DistilGPT-2 by Sanh et al.(2019), a distilled version of GPT-2, for the task of text classification. The data used for training was transformed into multiple versions of tabular text-based formats, each varying with respect to specific criteria: the review content, the user ID, the bot ID, the explanation of the bot strategy, and a brief description of the game. By creating these diverse datasets, we aimed to capture different aspects of the decision-making context and enhance the model's predictive capabilities.

Although transformer models like DistilGPT-2 are the state of the art and generally outperform traditional models such as LSTMs in many NLP tasks, our results did not show significant improvements in prediction accuracy.

2 Related Works

2.1 Persuation in NLP

We adopted the experimental setup of Apel et al. (2022), where an expert persuades a decision-maker (DM) to select a hotel using a single textual review from several available ones. The DM then decides whether to visit the hotel based on this review.

Shapira et al. (2024) applied ML and NLP strategies to predict decisions in a similar framework. This approach aligns with extensive research in NLP on persuasion. Notable studies include Tan et al. (2016) on online persuasion using Reddit's ChangeMyView, which provided a dataset for analyzing persuasive arguments, and Hidey et al.(2017), who focused on classifying arguments and examining the impact of argument sequencing on persuasive success. Other significant works include Wang et al. (2019), who explored persuasive dialogue systems aimed at social good, and Yang et al. (2019a), who created a text repository to identify and analyze persuasive strategies.

2.2 Fine-Tuning

Fine-tuning for text classification has seen significant advancements with the introduction of pre-trained language models. Radford et al. (2018) demonstrated the efficacy of generative pre-training followed by discriminative fine-tuning, achieving state-of-the-art results on several NLP benchmarks. Their approach involves pretraining a language model on a large, diverse corpus of unlabeled text, followed by task-specific fine-tuning using minimal architectural changes. This method outperforms previous models specifically crafted for each task, highlighting the potential of transfer learning in improving text classification performance. Our work builds on these advancements by employing task-specific input transformations to enhance fine-tuning effectiveness.

3 Data

We used data collected using a mobile application that simulated a realistic language-based persuasion game environment. This dataset includes 87,000 decisions made by 245 decision-makers (DMs) who interacted with 12 different automated expert bots, with each DM playing against 6 bots. To transform the raw data into various datasets for fine-tuning, we extracted the following components and created different datasets using subsets of these components:

3.1 Aggregated Round Information

We transformed the tabular data from each round into text format, incorporating details such as the hotel score, bot ID, user ID, reaction time, bot score, and user score. For example: *Round 5:* $\{$*hotel score: 7.37, current reaction time: 4553, user points in this round: 3, bot points in this round: 4*$\}$. To ensure that decisions are based on previous rounds, we aggregated data such that round i includes information from all rounds$j$ such that $j < i$.

3.2 Bot Strategy

We added the relevant text explanation of the corresponding bot strategy to each row to provide context on the bot's approach.

3.3 Attempt Count vs. Current Bot

We added the count of attempts against the current bot, as we believe this influences decision-making. Users might make more rash decisions out of frustration from not passing a level, or they may become more cautious.

3.4 Contextual Sentence

We included a contextual sentence to help the model understand the bigger picture, which may improve the precision of its predictions: *In this task, players in a mobile game make decisions to select good hotels or avoid bad ones, aiming to achieve a target payoff. Each correct choice earns 1 point, helping the player advance to the next level by reaching the predefined target payoff over 10 rounds, competing against rule-based expert bots.*

4 Model

According to Shapira et al.(2024). The article shows that LSTM is preferable to Transformer for performing the prediction; however, we hypothesized that a more state-of-the-art model like a Transformer, and more specifically a pretrained Transformer, may perform better than an LSTM.

4.1 DistilGPT-2

4.1.1 Model Architecture

Due to hardware limitations that make running the full GPT-2 model infeasible, we employ the DistilGPT-2 model, a distilled version of GPT-2 by Sanh et al.(2019). DistilGPT-2 retains most of GPT-2's capabilities while being more computationally efficient. Importantly, we needed a robust model capable of handling a large number of tokens efficiently, and DistilGPT-2 fits this requirement by leveraging the Transformer architecture, which uses self-attention mechanisms to process input sequences in parallel, making it efficient for handling long-range dependencies in text. DistilGPT-2 has fewer layers than GPT-2, typically containing 6 layers compared to GPT-2's 12 layers, significantly reducing the number of parameters and computational requirements. The number of attention heads per layer remains the same as in GPT-2, with each layer containing 12 attention heads to capture different aspects of the input data. The hidden size, or the dimension of the token embeddings and internal representations, is the same as in GPT-2, 768. However, the overall number of parameters is reduced due to the fewer number of layers. DistilGPT-2 is trained using knowledge distillation, where it learns to mimic the behavior of the larger GPT-2 model by minimizing the difference between the outputs of GPT-2 and DistilGPT-2. Due to these architectural modifications, DistilGPT-2 achieves approximately 60% of GPT-2's performance while being 60% faster and using less than half the number of parameters, making it suitable for deployment in environments with limited computational resources.

4.1.2 Training Procedure

The training procedure involves configuring various parameters and strategies to optimize the performance of the DistilGPT-2 model for the specific text classification task. Key parameters include the learning rate, batch size, number of epochs, and weight decay. In this study, we set the learning rate to 2e-5, the batch size per device to 4 due to computational limitations. We performed training of at most 45 epochs, while using an early stopping call back with a patience of three epochs, halting training if the model's performance does not improve, thus saving computational resources and preventing overfitting. This led to about 10 epochs in practice. Gradient accumulation is used to effectively increase the batch size by accumulating gradients over two steps before performing an update, allowing the model to handle larger batch sizes without exceeding memory limits. We also employ weight decay to regularize the model and prevent overfitting.

5 Experiments

In our study, we conducted multiple fine-tuning runs using different datasets. Each dataset was crafted to capture varying aspects of the decision-making context, providing diverse inputs for the model.

5.1 Round Information and strategy explanation without reviews

Initially, we used a dataset containing textual descriptions of the rounds and the bot strategy, but without reviews, to try and focus solely on the structural and contextual elements of the decision scenarios. This was done to understand if the review itself is a crucial part of the decision-making process, as the rest of the features may capture most of the behavioral aspects of decision-making.

5.2 Added reviews and contextual sentence

The next step was adding a contextual sentence and the review to the dataset. The review is meant to provide the full context of the round, while the contextual sentence is aimed at giving the model a better understanding of the task.

5.3 removed bot strategy

This experiment is identical to the previous one, but this data did not include the textual explanation of the bots decision making.

5.4 Added bot strategy removed contextual sentence

We also tried removing the context sentence, speculating that it might divert focus from important details. This approach considered the possibility that the context might not be crucial and could be implicitly learned during training. Simplifying the input in this way might still allow the model to perform effectively.

5.5 Added reviews and removed bot strategy

This experiment included a dataset with the round information, reviews and no bot strategy. Here, we wanted to see if the textual representation of the bot strategy coupled with only the review and round info is advantageous or not.

6 Results

Our experiments yielded mixed results, failing to surpass the accuracy reported in the original paper; Experiment 5.1 (”Round Information and strategy explanation without reviews”) achieved a max accuracy score of 69.5 percent, Experiment 5.2 (”Added reviews and contextual sentence”) increased the accuracy to 76.5 percent. Which indicates that the review is a key elements in the decision process. Experiment 5.3 (”Removed bot strategy”) further improved the accuracy to 79 percent, and the highest accuracy of 79.5 percent was obtained in Experiment 5.4 (”Added reviews and strategy explanation”). Experiment 5.5: (”Added reviews and removed bot strategy”) reached a max accuracy score of 77.4 percent. This experiment proves the textual representation of the bot strategy benefits the model. see results in Figure 1

Thus, we can infer that both the inclusion of reviews and the bot strategy are crucial for enhancing the model's performance. Additionally, the contextual sentence did not contribute as expected and may have introduced irrelevant information, leading to longer run times without significant accuracy improvement.

For future work, we plan to refine the data and experiment with larger models and more fine-tuning strategies. We believe leveraging a pre-trained model holds great potential for this task, although we have not yet identified the optimal approach to maximize its benefits.

References

[] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. *DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter*. In Proceedings of the NeurIPS EMC^2 Workshop. Retrieved from https://arxiv.org/abs/1910.01108

[] Christopher Hidey, Elena Musi, Alyssa Hwang, Smaranda Muresan, and Kathy McKeown. 2017. *Analyzing the Semantic Types of Claims and Premises in an Online Persuasive Forum*. In Proceedings of the 4th Workshop on Argument Mining, pages 11–21, Copenhagen, Denmark. Association for Computational Linguistics.

[] Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. *Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions*. In Proceedings of the 25th international conference on world wide web, pages 613–624.

[] Xuewei Wang, Weiyan Shi, Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang, and Zhou Yu. 2019. *Persuasion for Good: Towards a Personalized Persuasive Dialogue System for Social Good*. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 5635–5649, Florence, Italy. Association for Computational Linguistics.

[] Diyi Yang, Jiaao Chen, Zichao Yang, Dan Jurafsky, and Eduard Hovy. 2019a. *Let's make your request more persuasive: Modeling persuasive strategies via semi-supervised neural nets on crowdfunding platforms*. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 3620–3630.

[] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. *Improving Language Understanding by Generative Pre-Training*. OpenAI.

[] Reut Apel, Ido Erev, Roi Reichart, and Moshe Tennenholtz. 2022. *Predicting Decisions in Language Based Persuasion Games*. Journal of Artificial Intelligence Research, 73:1025–1091.

[] Eilam Shapira, Omer Madmon, Roi Reichart, and Moshe Tennenholtz. 2024. *Can Large Language Models Replace Economic Choice Prediction Labs?* ArXiv:2401.17435 [cs].

[] Eilam Shapira, Reut Apel, Moshe Tennenholtz, and Roi Reichart. 2024. *Human Choice Prediction in Language-based Persuasion Games: Simulation-based Off-Policy Evaluation*. Technion.

Appendix

A Results

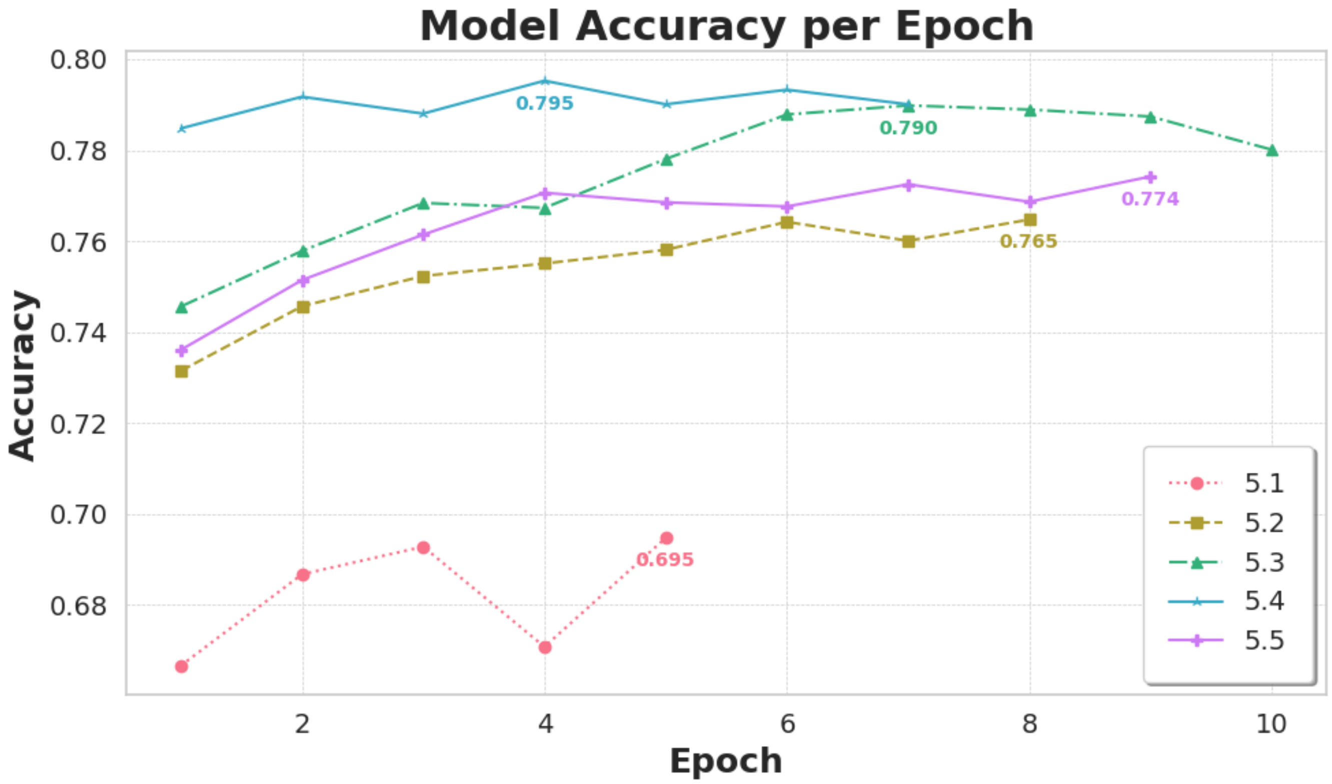

Figure 1: Accuracy per epoch for all models on all test samples