Enhancing simulation using Machine Learning based strategies

Ben Hellmann & Nerya Hadad

Abstract

In recent years, deep neural networks have overshadowed classic machine learning models in many applications. This paper explores a hybrid approach that leverages weak machine learning models to boost the predictive power of Long Short-Term Memory (LSTM) networks for forecasting human behavior. By training weak classifiers on a set of human behavioral data, we aim to improve the simulation of human choices, and thereby improve the accuracy of LSTM predictions. Our experiments show that this method yields a modest but significant improvement in prediction accuracy, with an increase of up to 0.3% from a baseline of 83.6%.

1 Intro

This article builds on the foundation established by (Shapira et al., 2023), which presents an off-policy evaluation (OPE) setup for experts (agents) interacting with human decision-makers (DMs) in a non-cooperative game. The aim is to predict DMs' choices when playing with a set of partners (agents) in a given game, based on their previous interactions with various other artificial agents in the same game. This approach allows optimization of the artificial agents.

Each game consists of 10 rounds, in which the experts provide a review of a hotel, and the DM has to decide whether to go to the hotel or not.

A key issue identified in the work is the lack of sufficient human data for training. To address this, the authors created a series of simulated games utilizing three strategies simulating the decision maker's choice: The first is a "trustful" strategy - 'trusts' the artificial agent's review. Calculates if the review contains more positive or negative keywords and chooses to go or not accordingly. The second is an LLM-based strategy - given the artificial agent's review, an LLM tries to predict whether the hotel is good or bad. The third is a random strategy - randomly chooses whether to go or not to go to the hotel. Each of the strategies receives a weight of 1, giving them an identical probability to start with (which changes randomly as the game progresses). In addition, in order to simulate human learning process, as the game progresses, the probability of making the correct decision grows. The authors of the article then train a LSTM based classifier to predict the human choices.

Given the improvement observed due to the simulation, our aim was to further enhance its effectiveness. We identified a potential limitation: the simulations may not accurately reflect the behavior of actual human decision-makers. Our hypothesis centered on the notion that aligning the simulation more closely with human behavior could improve the accuracy of the LSTM based classifier. To this end, we pursued the training of weak and classic machine learning classifiers on pre-collected data from human decision-makers. We then evaluated their effectiveness in predicting human decisions, and integrated the best-performing strategies into the simulation. We also explored the potential of combining multiple classifiers as multiple strategies to further boost performance.

2 Related Works

Previous attempts have been made to simulate human behavior using machine learning methods, as demonstrated in the article by (Rocha and Duarte, 2019). In their work, they developed ML-based simulations to mimic human decision-making processes in gaming environments. These simulations generated additional training data that enhanced the predictive models' accuracy and robustness. In another article (Plonsky et al., 2017), the authors tried different machine learning models such as random forest, SVM, neural networks and KNN to predict human behaviour. They found that the models, particularly random forest, performed very well. Another related article by (Busogi and Kim,

2017) uses machine learning methods such as linear regression and SVM to predict human choices using 'human in the loop' method.

3 Data

We use the data collected in (Shapira et al., 2023). The writers collected a dataset that consists of 87k decisions from 245 DMs who played against 12 different automatic expert bots (each DM played against six bots). The data consists of 37 engineered features chosen by (Apel et al., 2022) which are passed to the classifiers trying to predict if the human will choose to go to the hotel (1) or to continue to the next round (0).

4 Model

The idea in this study is to add a strategy that takes the 37 engineered features as input and classifies the decision to go or not go to the hotel. For this strategy, we employed 18 different classifiers using the scikit-learn library. The classifiers we used are:

- Random Forest (RFC)
- Support Vector Machine (SVC)
- Multi-Layer Perceptron (MLPC)
- Gradient Boosting (GBC)
- AdaBoost (ABC)
- Bagging (BC)
- K-Nearest Neighbors (KNNC)
- Decision Tree (DTC)
- Extra Trees (ETC)
- Gaussian Naive Bayes (GNBC)
- Bernoulli Naive Bayes (BNBC)
- Multinomial Naive Bayes (MNBC)
- Stochastic Gradient Descent (SGDC)
- Passive Aggressive (PAC)
- Perceptron (PC)
- Ridge Regression (RC)
- Quadratic Discriminant Analysis (QDAC)
- Logistic Regression (LRC)

First, we pre-trained each classifier on the human choice data training set. We tested each classifier as explained in the experiments and results section. We then chose the 10 best-performing classifiers, which were utilized as an additional strategy for the simulations.

For each classifier, we trained the LSTM model using simulations including it as a strategy. We gave the classifier-based strategy an identical initial weight to the rest of the strategies (which were explained in the intro), meaning there were 4 strategies. Then, we tested the accuracy achieved by the LSTM trained on the simulations with the added strategy and checked the difference from the null model.

Next, we tried combining the best performing strategies. We used the 4 best performing classifiers and combined them in all possible ways: all combinations of 2, 3 and 4 classifiers. In each combination run, each of the classifier-based strategy got an initial weight of 1 (i.e, when running a combination of 2 classifiers, we had 5 strategies with equal probability of being used).

5 Experiments and Results

Note: the null model is the run with the 3 strategies described in the article (Shapira et al., 2023). It is represented in the graphs by the name 'None'.

5.1 Pre-Training

First, we ran a 10-fold cross-validation in which each classifier was trained on the human engineered features training set. We calculated the average accuracy achieved by each classifier in predicting the human choices as can be seen in Figure 1.

5.2 Stage 1: Individual Classifier Performance

We chose the 10 best-performing classifiers from the pre-training stage. Each was used as a different strategy in addition to the original 3. For each of the 10 classifiers, we ran the LSTM with the updated simulation of four strategies, using 6 different seeds of randomness, and averaged the best epoch. The best epoch's average is shown in Figure 2.

5.3 Stage 2: Verifying The Best Individual Classifiers

We decided to run the eight strategies which produced the best LSTM accuracies on another six seeds. The average full of all 12 seeds can be seen in Figure 3, and the average of the best epoch can

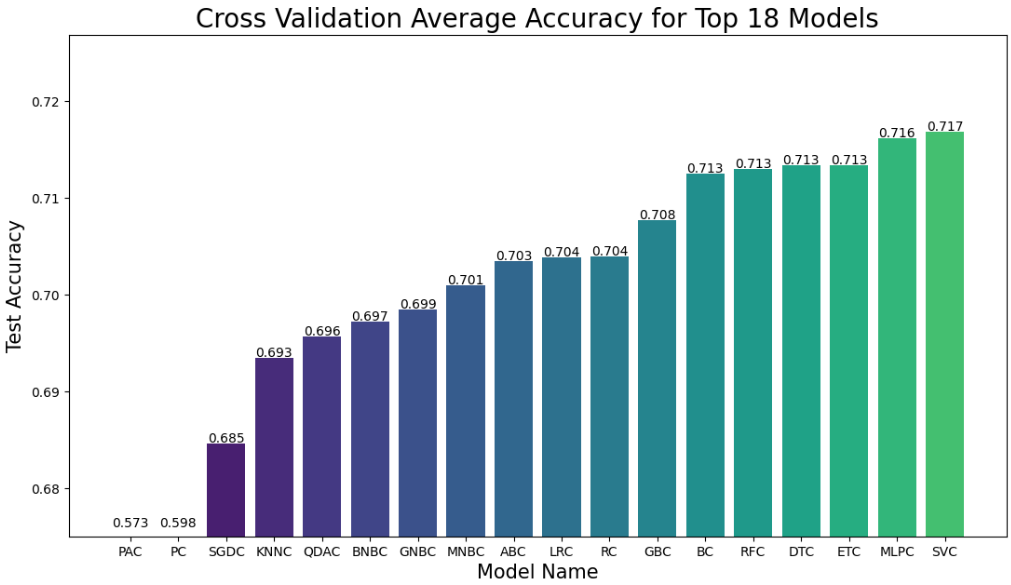

Figure 1: Pre-Training Results

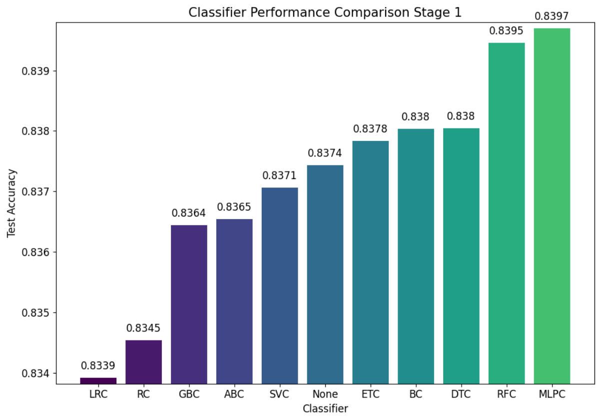

Figure 2: Stage 1 - 6/12 seeds

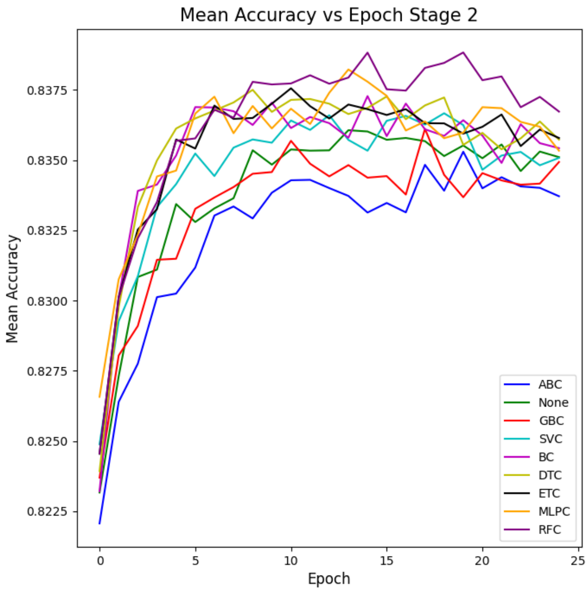

Figure 3: Stage 2 - 12/12 seeds

be seen in Figure 4. We found that the best model is the Random Forest Classifier with a result of 83.88% accuracy. Notice that in 7 out of the 8 cases, the score improved due to our additional strategy ("None" represents the null model).

Additionally, we tested whether the classifiers that achieved the highest accuracy in the pre-training stage also achieved the highest improvement as a strategy for the LSTM, as seen in Figure 5. The results were mainly positive with 6 of the classifiers appearing in the same order that they appeared in the pre-train with the exception of SVM performing significantly worse, and Random Forest performing significantly better.

5.4 Stage 3: Combining Classifiers

We used multiple strategies - all combinations of the 4 best performing classifiers on stage 2 with six seeds. Then we calculated the average accuracy of the best epoch in each as can be seen in Figure 6.

5.5 Stage 4: Verifying Best Combinations

We ran the 2 best performing combinations on stage 3 on additional 6 seeds. Then we calculated the average accuracy of the best epoch in each as can

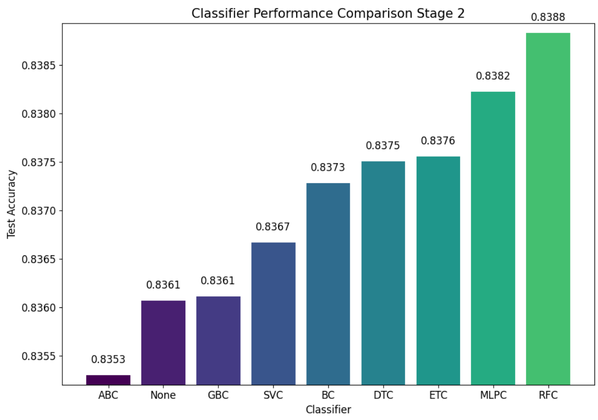

Figure 4: Stage 2 - 12/12 seeds

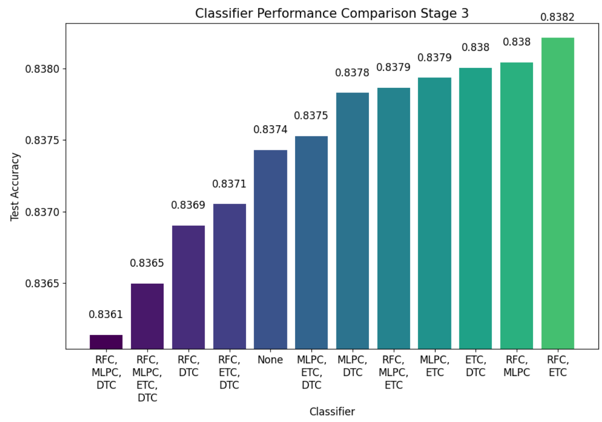

Figure 6: Stage 3 - 6/12 seeds

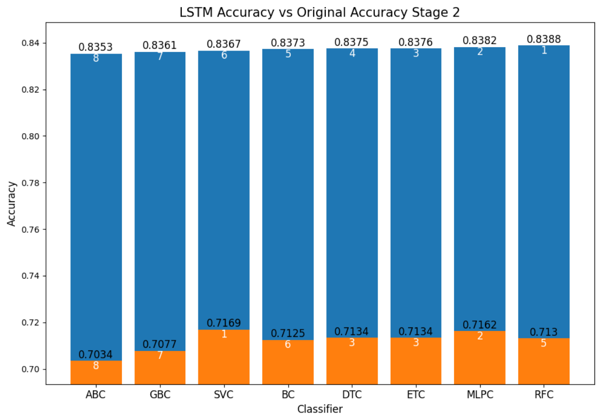

Figure 5: Stage 2 - Comparison To Cross Validation

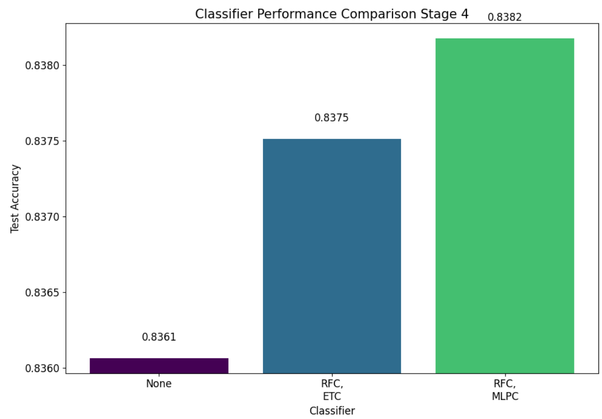

Figure 7: Stage 4 - 12/12 seeds

be seen in Figure 7. We found that the highest result was achieved by the combination of the 2 best performing models in stages 1 and 2 (MLPC and RFC), but matched the result of the second best model out of the 2 rather. i.e. the results were lower than the results achieved by the lone strategies - 83.82% accuracy as oppose to 83.88%.

5.6 Conclusion

We demonstrated that several machine learning models can improve the accuracy of the LSTM trained on simulations including an extra classifier-based strategy. In particular, the Random Forest classifier improved the result from 83.6% in the null model to 83.88%, as seen in Figure 3, consistently outperforming the other models. Additionally, we found that combining different classifier-based strategies could improve accuracy compared to the null model, but achieved less accuracy than using a single strategy. The issue might be due to allocating 40% of the simulation to the classifiers, and decreasing it might improve those results.

References

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.

Moise Busogi and Namhun Kim. 2017. Analytical modeling of human choice complexity in a mixed model assembly line using machine learning-based human in the loop simulation. *IEEE Access*, 5:10434–10444.

Ori Plonsky, Ido Erev, Tamir Hazan, and Moshe Tennenholtz. 2017. Psychological forest: Predicting human behavior. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.

D. de Almeida Rocha and J. C. Duarte. 2019. Simulating human behaviour in games using machine learning. In *2019 18th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, pages 163–172. IEEE.

Eilam Shapira, Reut Apel, Moshe Tennenholtz, and Roi Reichart. 2023. Human choice prediction in non-cooperative games: Simulation-based off-policy evaluation. *arXiv preprint arXiv:2305.10361*.