An Expert-Model & Machine Learning Hybrid Approach to Predicting Human-Agent Negotiation Outcomes

Johnathan Mell
USC Institute for Creative Technologies
Los Angeles, CA, USA
mell@ict.usc.edu

Markus Beissinger
Microsoft
San Francisco, CA, USA
mabeis@microsoft.com

Jonathan Gratch
USC Institute for Creative Technologies
Los Angeles, CA, USA
gratch@ict.usc.edu

ABSTRACT

We present the results of a machine-learning approach to the analysis of several human-agent negotiation studies. By combining expert knowledge of negotiating behavior compiled over a series of empirical studies with neural networks, we show that a hybrid approach to parameter selection yields promise for designing -more effective and socially intelligent agents. Specifically, we show that a deep feedforward neural network using a theory-driven three-parameter model can be effective in predicting negotiation outcomes. Furthermore, it outperforms other expert-designed models that use more parameters, as well as those using other, more limited techniques (such as linear regression models or boosted decision trees). We anticipate these results will have impact for those seeking to combine extensive domain knowledge with more automated approaches in human-computer negotiation.

1 Background

Negotiation is a complex human social task that requires a diverse set of skills: from strategic planning to rhetorical argument. And while negotiation has been traditionally seen as a human problem, that perception is quickly changing. As technological tools continue to evolve into ever-more sophisticated artificial agents, humans find themselves relying on increasingly human-aware agents to interact with the world around them. Designing agents that are capable of engaging in human-like negotiation has become a challenging problem for researchers, since it involves a variety of techniques, from social awareness to user modeling to learning [2].

Currently, many of these agents are designed based on existing psychological/behavioral models of user behavior in negotiation (for example, see [13]). These agents draw from an immense literature on negotiation strategies in the business, psychology, and economics corpora. As these automated agents continue to negotiate with humans, however, they generate a massive amount of behavioral data. As such, analysis of human-agent interaction through machine learning approaches is becoming increasingly feasible.

But even when the stated goal is merely to predict the outcomes of negotiation, machine learning approaches are not panaceas. Human-agent negotiating datasets have a tendency to be “wide and short”, with hundreds of behavioral and process variables being tracked, but relatively few subjects (due to the difficulty in conducting massive user studies). These problems lead to very noisy inputs into traditional machine learning algorithms, and make feature selection a chancy proposition at best.

IVA '19, July 2–5, 2019, PARIS, France

ACM ISBN 978-1-4503-6672-4/19/07. https://doi.org/10.1145/3308532.3329433

We therefore propose a hybrid approach for the analysis and development of agent in human-agent negotiation. By inputting expert knowledge of the domain into machine learning algorithms, we effectively create “priors” that allow these algorithms to more accurately account for noise without needing massive amounts of data. From the side of model-driven AI, this also allows for us to quickly and effectively evaluate a variety of potentially relevant behavioral parameter sets, while also circumventing some of the limitations of traditional evaluation approaches (such as reliance on regressions).

The following work examines data from three human-agent negotiating experiments conducted on the Interactive Arbitration Guide Online (IAGO) negotiation platform [14]. We show that a theoretically-sound and minimal-parameter neural network outperforms other models that use more simplistic approaches (linear regression) or more parameters (including those that are supersets).

Much of the study of negotiation in general, and human-agent negotiation specifically, has focused on the application of certain techniques that will be effective in creating or claiming value in negotiations. This can entail such techniques as strategically withholding or sharing private information about preferences [16], using positive or negative emotion to manipulate other parties [7], or accurately modeling opponent preferences and crafting offers which “grow the pie” by finding integrative potential [6].

The goal of much of human-agent negotiation work is to predict outcomes using variables found within the negotiation. To this end, hundreds of variables may be tracked in an average human-agent negotiation (in this dataset, over 200). These variables include:

- Process measures – number of messages sent by each party, emotional expressions detected, offer numbers and types, etc.
- Strategy variables – policies used by negotiating agents, such as whether they use emotional manipulation or attempt to withhold key information

In a classical behavioral study, one or more strategy variables may be manipulated experimentally in order to see the resulting change on outcomes. There are a number of papers in this vein, which have discovered notable results [1][4][7][8]. Multiple regression is then performed to determine if there are any first-order or interaction effects on the dependent variables. However, this approach has limitations—traditional regression becomes untenable beyond a few independent inputs, as the statistical power quickly becomes weak. Machine learning provides alternatives to this procedure through the construction of neural nets, but suffers from a sensitivity to noise in the data. Furthermore, in datasets that contain hundreds of potential input variables, a brute force approach to analysis (even with feature selection techniques) becomes absurd.

In our approach, we aim to predict outcome metrics of a negotiation based on a series of theory-driven agent models. For each agent parameters set, we wish to predict several outcome metrics of the negotiation: scalar targets, such as agent points, user points, Nash points, and total points, and percentage targets, such as agent point %, user point %, total point %.

2 Experimental Design

The data used in this review comprises 485 subjects collected over a series of 3 different studies. These studies were all conducted on the IAGO platform, a system for facilitating human-agent negotiation data collection [14]. In all studies, the participants engaged in a standard multi-issue bargaining task with an IAGO agent. Several different types of agents were used; their behavior varied according to a number of strategy variables (see below).

Based on existing theories of negotiation, we designed 7 parameter sets that served as basic models in order to predict negotiation outputs. The parameters in each set are listed in Table 1. Our first three parameter sets explored different combinations of information about the player and the agent. *KnowThineEnemy* focused on user variables that agents could track about the human, while being completely agnostic about the agent's own behavior. *KnowAll* included all the information from *KnowThineEnemy*, but also included parameters that defined how the agent acted—in short, the agent was aware of its own behavior. This included information about how the agent acted in its use of emotion (**nice**), information revelation strategy (**withholding**), and general offers (**competitive**). Thirdly, *Self-Reflection* included only information on the agent itself.

Set Name	Parameters
KnowThineEnemy	numUserOffers, numUserMsgOnly, numUserCombined, numUserHappy, numUserAngry
KnowAll	nice, withholding, competitive, numUserOffers, numUserMsgOnly, numUserCombined, numUserHappy, numUserAngry
Self-Reflection	nice, withholding, competitive, numAgentOffers, numAgentHappy, numAgentAngry, numAgentMsg
Emotional	nice, numUserHappy, numUserAngry, numAgentHappy, numAgentAngry
Strategic	competitive, numUserOffers, numAgentOffers
Chatty	withholding, numUserMsg, numAgentMsg
Everything-MakesSense	all previous parameters

Table 1: Models and Parameters

The second set of three parameter sets included models that focused on one particular channel of communication. *Emotional*, for example, looks at variables relating to human and agent affective choices, like the use of anger and happiness. It also included **nice**. *Chatty* focuses on the idea that the messages exchanged in negotiation may be predictive due to their effect on rapport between the human the user, and thus includes both message quantity variables as well as the agent strategy variable **withholding**. Finally, *Strategic* focuses on examining agent strategy (**competitive**), and the quantities of offers exchanged by both parties. The final model, *EverythingMakesSense* simply included all the previous variables—this is the most likely model to be attempted by someone with no particular awareness of negotiation theory, and serves as a reasonable baseline for "traditional" machine learning approaches.

We compared the models above by using three machine learning methods on the data gathered from experiments. For each of the models, we trained a linear regression baseline, an XGBoost [5], and a Deep Neural Network (DNN) to compare predictive performance of the specified input columns to the target columns.

We utilized k-folds cross-validation to compare machine learning methods across models, where k = 10 (based on [10]). This means the dataset was randomly split into 10 subsections, where 9 subsections are combined as the training set and 1 is left as the validation set for calculating root mean square error (RMSE) and filling in the predicted values. This training process is repeated 10 times such that the machine learning method is able to make test predictions for every row in our dataset. XGBoost was chosen as a comparison due to its successes in Kaggle competitions [9]. We created separate XGBoost regressions for each scalar target column, and separate XGBoost classifiers for each percentage target column to keep the output in the range [0,1]. We used a DNN to compare model performance [3]. Our DNN consisted of Feedforward layers interleaved with Dropout noise to reduce overfitting. The DNN used 7 layers, and the final layer differed for scalar vs. percentage targets. In both cases, we trained the DNN with stochastic gradient descent.[1]

3 Results & Discussion

For each of the 7 theoretical models, there are three variants: linear regression (Linear), XGBoost (Boost), and deep neural network (DNN). The top-performing subset of the resulting 21 models is listed in Table 2, along with the root mean squared errors (RMSE) for each of the 7 targets. The table also highlights the best performing model for each output in blue, and the second best in yellow.

Generally, the models explained a reasonable amount of the variance in the results. The RMSE of *ChattyDNN* for **agent points** was 6.31. Since the negotiations involved between 65 and 70 points, this prediction implies a 95% confidence interval within 12.62 points.

ChattyDNN outperforms its baseline linear regression counterparts in every category (except **agent points**, where *ChattyDNN* is worse than *ChattyLinear*). *ChattyDNN* also performs adequately in one of the categories in which it is not in the top two—specifically *ChattyDNN* is the top 3 models for **agent point %**. Indeed, *ChattyDNN* performs quite well even accounting for its relative shortcomings in predicting agents' points.

[1] Further detail on the parameters of the individual layers and other reproducibility data is available from the authors by request.

The implications of the results in this work are relevant to the negotiation domain in that they indicate the importance of message exchange to negotiated outcomes. Yet, they also have methodological implications for the design of human-agent systems and studies. Social computing problems are plagued by uncertain inputs, massive numbers of input variables, and relatively small datasets. These problems make them tenuous targets for much of the current work in machine learning. Furthermore, the domains in which social problems are most relevant are these where model-driven approaches are most well-studied. Therefore, approaches which can both leverage current machine learning approaches to process data, but can also "initialize their priors" using expert knowledge are of particular interest. Expert-knowledge is useful; the performance of automatic feature selection is not perfect. Indeed, per Lucas et al [12], domain knowledge can assist with feature selection.

The results of this work are largely in line with psychological intuitions. The best-performing model is *ChattyDNN*. This model uses only three parameters (**withholding, numUserMsg,** and **numAgentMsg**). *Chatty* relies on the observation that communication is key to building rapport in negotiation, and this increased rapport and understanding of opponent preferences can lead to an increase in joint value. That theory-driven intuition about communication indeed appears to be on point, given *Chatty*'s good performance at predicting **Nash points**, **total points**, and **total point %**.

Most notably, *ChattyDNN* outperforms its strict superset models in most categories. *EverthingMakesSenseDNN*, performs worse than *ChattyDNN* over most of its outputs. Given that *Chatty* is an expert-designed model, rather than one that was directly learned, this lends credence to the idea of *hybrid* approaches. In a (common) example of "more parameters are not better", it outperforms a number of more complicated models, while having a theoretical groundwork based on the idea of the importance of information exchange.

However, *ChattyDNN* also outperforms its more pedestrian counterparts, *ChattyBoost* and *ChattyLinear*. This speaks to the benefit of using machine learning algorithms to construct optimized models (even if those models are originally expert-designed).

In this result, we show that a hybrid approach to analysis can yield actionable results, especially in domains that have datasets with features similar to human-agent negotiation. Purely learned approaches may simply contain too many parameters, such as the "kitchen sink" approach of *EverythingMakesSense*. The hybridized approach benefits from prior knowledge of the domain, and ensures we get the most "value" for each parameter that is added.

ACKNOWLEDGMENTS

The authors want to thank our colleague Dr. Gale M. Lucas for her insights on the analytical approach to these results. This research was supported by the US Army. Statements and opinions expressed do not necessarily reflect the position or the policy of the US Government, and no official endorsement should be inferred.

REFERENCES

[1] Anthony, P. and Jennings, N.R., 2003. Developing a bidding agent for multiple heterogeneous auctions. *ACM Transactions on Internet Technology (TOIT)*, *3*(3), pp.185-217.

[2] Baarslag, T., Kaisers, M., Gerding, E., Jonker, C.M. and Gratch, J., 2017. When will negotiation agents be able to represent us? The challenges and opportunities for autonomous negotiators.

[3] Bengio, Y. 2013. Deep learning of representations: Looking forward. In *International Conference on Statistical Language and Speech Processing* (pp. 1-37). Springer, Berlin, Heidelberg.

[4] Castelfranchi, C. and Falcone, R., 1998, July. Principles of trust for MAS: Cognitive anatomy, social importance, and quantification. In *Multi Agent Systems, 1998. Proceedings. International Conference on* (pp. 72-79). IEEE.

[5] Chen, T. and Guestrin, C., 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). ACM.

[6] De Dreu, C.K., Koole, S.L. and Steinel, W., 2000. Unfixing the fixed pie: a motivated information-processing approach to integrative negotiation. *Journal of personality and social psychology*, *79*(6), p.975.

[7] de Melo, C. M., Carnevale, P., & Gratch, J., 2011, May. The effect of expression of anger and happiness in computer agents on negotiations with humans. In The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 3 (pp. 937-944). International Foundation for Autonomous Agents and Multiagent Systems.

[8] Fox, J., Ahn, S.J., Janssen, J.H., Yeykelis, L., Segovia, K.Y. and Bailenson, J.N., 2015. Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. *Human–Computer Interaction*, *30*(5), pp.401-432.

[9] Gerhart, B. and Rynes, S., 1991. Determinants and consequences of salary negotiations by male and female MBA graduates. *Journal of Applied Psychology*, *76*(2), p.256.

[10] Klambauer, G., Unterthiner, T., Mayr, A. and Hochreiter, S., 2017. Self-normalizing neural networks. In *Advances in neural information processing systems* (pp. 971-980).

[11] for Autonomous Agents and Multiagent Systems.

[12] Lucas, G., Stratou, G., Lieblich, S. and Gratch, J., 2016, October. Trust me: multimodal signals of trustworthiness. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction* (pp. 5-12). ACM.

[13] Mell, J., et al, 2018. The Results of the First Annual Human Agent League of the Automated Negotiating Agents Competition, *Intelligent Virtual Agents*.

[14] Mell, J. and Gratch, J., 2017, May. Grumpy & Pinocchio: Answering Human-Agent Negotiation Questions through Realistic Agent Design. In *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems*(pp. 401-409). International Foundation for Autonomous Agents and Multiagent Systems.

[15] and interpreting deep learning models. *arXiv preprint arXiv:1708.08296*.

[16] Thompson, L.L., 1991. Information exchange in negotiation. *Journal of Experimental Social Psychology*, *27*(2), pp.161-179.

Model Name	User Points RMSE	Agent Points RMSE	Nash Points RMSE	Total Points RMSE	User Point % RMSE	Agent Point % RMSE	Total Point % RMSE
ChattyDNN	6.31	6.54	194.92	8.26	12.8	11.1	12.8
SelfReflectionDNN	6.85	7.09	220.23	9.94	12.5	11.5	14.5
StrategicDNN	6.90	7.01	221.55	9.79	12.5	11.5	14.5
EverythingMakesSenseDNN	6.76	8.05	221.89	10.80	12.5	11.1	13.8
ChattyLinear	6.87	5.70	222.58	9.17	13.7	11.4	14.1
StrategicLinear	6.73	5.51	226.20	9.52	13.5	11.0	14.6
EverythingMakesSenseLinear	7.04	5.33	235.39	9.76	14.1	10.7	15.0

Table 2: Root Mean Squared Error for Top Models, Negotiation Outcomes

Blue items are the best values in the column; yellow items are the second best. Models with no top performance in any column are omitted.