The Wisdom of the Market: Using Human Factors to Design Prediction Markets for Collective Intelligence

Lorenzo Barberis Canonico
Clemson University, SC
lorenzb@g.clemson.edu

Christopher Flathmann
Clemson University, SC
cflathm@g.clemson.edu

Dr. Nathan McNeese
Clemson University, SC
mcneese@clemson.edu

There is an ever-growing literature on the power of prediction markets to harness "the wisdom of the crowd" from large groups of people. However, traditional prediction markets are not designed in a human-centered way, often restricting their own potential. This creates the opportunity to implement a cognitive science perspective on how to enhance the collective intelligence of the participants. Thus, we propose a new model for prediction markets that integrates human factors, cognitive science, game theory and machine learning to maximize collective intelligence. We do this by first identifying the connections between prediction markets and collective intelligence, to then use human factors techniques to analyze our design, culminating in the practical ways with which our design enables artificial intelligence to complement human intelligence.

INTRODUCTION

Prediction markets are mechanisms that enable participants to bet upon the occurrence of particular events (Hanson, 2003). The objective of a prediction market is to create an incentive structure to coordinate a sophisticated forecasting process that can enable organizations, communities, and countries to better deal with uncertainty about the future. To that end, they have been studied in many social psychology contexts to identify their connections to collective intelligence (Tetlock, 2016). However, despite their relevance to cognitive science, prediction markets have been mostly analyzed in the context of economics and computer science, where the objective is the optimization of the underlying procedures as opposed to collective intelligence.

We approach prediction markets from a human factors perspective to identify the key cognitive features that enables collective intelligence to emerge. Subsequently, we propose a new human-factors-based prediction market whose design enhances collective intelligence. Furthermore, we go one step further and integrate an artificial intelligence component to set the foundation for much higher degrees of collective intelligence. Thus, our model enables designers to leverage all the recent advancements in machine learning.

PRACTICE INNOVATION

Prediction Markets

At their core, prediction markets extend the dynamics of the stock market, where traders buy and sell stocks in anticipation of corporate announcements, to broader events such as political elections and box office performance. A basic example would be an election, where the value of a candidate's "stock" becomes $1 if the candidate wins, and $0 if the candidate loses, thus enabling participants to buy and sell the stock until all trading ends and yields a price that inherently reflects the probability of the candidate winning (70c would imply a 70% chance).

Prediction markets are remarkably effective at forecasting events, and are often better than pundits and experts alike. For instance, in the case of sports, real-money prediction markets were found to be more accurate than expert polls (Goel et al, 2010). In the realm of politics, a study of the Iowa Electronic Market's (IEM) performance over the course of the presidential elections between 1988 and 2000 shows that the IEM's market price on the day each of the 596 different polls were released was more accurate than the polls themselves 75% of the time (Hanson, 2003, Surowiecki, 2005). These results carry over into geopolitical forecasts as well, where IARPA's DAGGRE prediction market accuracy was about 38% greater than the baseline system at over 400 geopolitical questions (Laskey et al, 2015). Recent research also suggests that prediction markets outperform even AI-based and big data approaches. For instance, IEM outperformed a highly advanced machine learning model analyzing 40 million unique tweets in the 2012 election (Attarawala et al, 2017). These results are not isolated, but rather are consistent with a broader pattern of prediction markets being systematically more effective than expert and collective judgements.

The key to the success of prediction markets lies in their ability to aggregate and combine diverse opinions to parse signal from noise and produce a single consistent probability distribution (Hanson, 2003). Markets also provide strong economic incentives for individuals to correct systematic biases, such as overconfidence and underconfidence. A rational trader would only place bets that are expected to be profitable, realigning prices with historical base rates (Atanasov, 2016). Furthermore, they are able to handle more complexity than an individual or centralized body could grasp because "knowledge that is implicit, dispersed, and inaccessible by traditional, conscious methods can be organized through markets to create more rational calculation than can elite experts" (Marcus, 2004, par. 11).

In essence, prediction markets generate a strong financial incentive for participants to express their opinions in precise and informative ways, which are then aggregated and calibrated by the trading mechanism to reflect the most reliable estimate for a particular event. This transition from the local information of each participant to the global information of the crowd speaks to the powerful emergent property of prediction markets: collective intelligence.

Collective Intelligence

Collective intelligence is the result of the proper aggregation of local information in generating a global solution to a problem that is more optimal than what any individual could have provided (Watkins, 2007). However, collective intelligence should not be confused with "groupthink": it is not merely the sum product of group opinions, but is instead a weighed and calibrated end-product of an information exchange between a group of thinkers. Just because a group convenes and votes on an issue, it does not mean that the "wisdom of the crowd" is occuring (Malone, 2018).

Specifically, not all groups are good knowledge generators (Surowiecki, 2005). At the extreme, a crowd morphs into a mob: a dangerous and inefficient arrangement to distribute knowledge to members. Even at a micro-level, teams often fail to integrate all relevant information about a problem before making a decision due to the kind of pressure towards conformity inherent to group interactions. Social norms can pressure individuals with distinct perspectives to alter their behavior in order to assimilate, which undermines the kind of diversity that lies at the core of the accuracy gains in collective intelligence (Watkins, 2007).

Prior research has identified four conditions that enable the emergence of collective intelligence in a crowd (Surowiecki, 2005):

1. Diversity of opinion: each person should have some private information, even if it's just an eccentric interpretation of the known facts
2. Independence: people's opinions are not determined by the opinions of those around them
3. Decentralization: people are able to specialize and draw on local knowledge
4. Aggregation: some mechanism exists for turning private judgments into a collective decision

Prediction markets succeed because their nature supports all four factors, as participants have a financial incentive to research and grain private information that is then implicitly shared once they begin trading in the market. Decentralization is especially apparent, as any individual gets to immediately trade with every other participant in the market, enabling information to flow very rapidly because it does not have to go through a hierarchy.

Independence is also extremely important to collective intelligence because the underlying reality of any crowd effort is that no individual has perfect access to all information, and that the estimate of all individuals is always flawed in some way. Independence guarantees that errors in individual judgment won't wreck the group's collective judgment as long as those errors aren't systematically pointing in the same direction. One of the quickest ways to make people's judgments systematically biased is to make them dependent on each other for information. Furthermore, independent individuals are more likely to have new information rather than the same old data everyone is already familiar with (Surowiecki, 2005).

In essence, the reason why the average of all a classroom's estimates for how many jelly beans are in a jar is only a few percentage points away from the actual number is because the overconfident estimates and underconfident estimates offset each other, thereby distilling signal from the noise and yielding an estimate that is superior to that of any individual participant.

Challenges and Limitations of Prediction Markets

Just like financial markets, prediction markets are not immune to problems. Forecasting future events is such a challenging task that is prone to errors of all types, that can potentially be magnified by the macro-nature of a prediction markets.

Prior research has identified three major types of errors in prediction markets (Dudik et al, 2017):

1. Sampling error, which arises from traders possessing noisy estimates that dilute the truth-value of their information.
2. Market-maker bias, arising from a particular cost function being used to generate an opportunity for profit to facilitate trading actually inducing particular biases on overshooting or undershooting the estimate.
3. Convergence error, arising from huge market fluctuations caused by all the trading before the price stabilizes arising because, at any point in time, market prices may still be in flux

These problems are exacerbated by the fact that redesigning the aggregation function, the primary technical solution often discussed in the literature, is often not enough. Chen et al (2005) analyzed data from football games and found that linear, logarithmic, absolute distance, and quadratic scoring did not differ significantly as aggregation functions in their overall accuracy. Purely technical approaches have thus not shown major improvements in addressing some of the variability factors in prediction markets.

An alternative that has been proposed is to move away from prediction markets altogether, and focus on better prediction polls to elicit and aggregate estimates from individuals. The results however have not been supportive of such a claim, as prior research has indicated that simple aggregate of prediction polls tends towards not just underconfidence (despite the well-known tendency for people to be overconfident) but also less meaningful as the average forecasts converge towards 50% probability for two-option questions (Satopää et al, 2014).

Lastly, prediction markets rely on financial incentives to motivate participants. This makes sense, for no participant would trade if there was not an opportunity to profit from someone else's lack of information (Ex. buying a candidate's stock that sells for 50c to resell it at 70c). Thus prior researchers have highlighted the importance for the manager of the prediction market to subsidize trading for new events in order to catalyze the trading process among participants (Chen et al, 2010, Hanson, 2003). This kind of solution creates a barrier to entry to the implementation of prediction markets in areas beyond geopolitics or elections, where a thick market of many participants can be expected.

We thus seek to address these concerns (prediction errors, suboptimal aggregation functions, behavioral biases) by designing a new type of prediction market that not only harnesses collective intelligence in a sustainable way, but also enables artificial intelligence to address many of the limitations of traditional design.

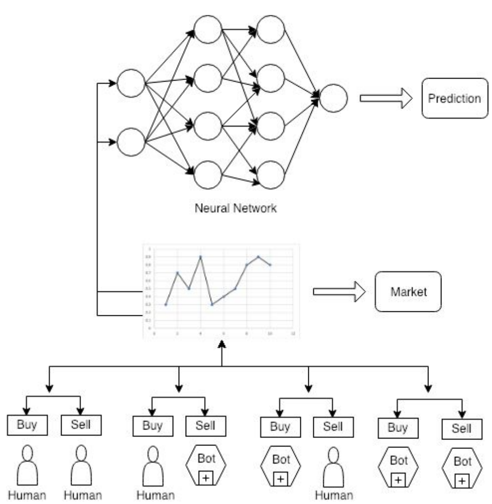

Figure 1
The collective intelligence emerging from humans and AIs in the prediction market is fed as input into a neural network

PRACTICE APPLICATION

Model Design

Our prediction market in Figure 1 relies on 4 components: 1) *human participants* without access to historical trading data, 2) *noisy trading bots* that trade randomly to generate a profit opportunity for other traders, 3) *market making bots* that learn from market patterns and trade in order to stabilize or open up new markets for event, and a 4) *neural network* that receives the data generated by the prediction market to make its estimate.

As Figure 1 shows, the noisy bots create the profit opportunity that motivates the human participants to engage in the trading alongside the market making bots that rapidly smooth price changes and stabilize the market for each event. The data generated by process is then fed as the input to the neural network that calibrates its aggregate estimate of all the participants over time.

Human Interface

The primary objective on the human side of prediction markets is to incentivize only traders with new information to engage. Without such a goal, the market's incentives can lead to speculate behavior observable in stock markets, where some traders specialize in trading on price movements as opposed to changes in fundamentals. Such behavior would dilute the signal of prices in the prediction market for they would be distorted by copycat and speculative traders who are not contributing to the implicit deliberation process. To that end, our design does not include interfaces displaying historical trading data as to induce traders to consider only information relevant to the reality of the event as opposed to its financial counterpart in the prediction market.

Furthermore, we deploy Hanson's market scoring rule market maker (MSR) to generate activity on new events (Hanson, 2003, Chen et al, 2010). This structure induces even just a single trader to reveal their information, which would otherwise not occur under a standard double action in traditional prediction markets. Furthermore, MSR enables the manager of the prediction market to not only reduce the amount of money needed to simulate initial activity in the prediction market, but to also efficiently allocate that capital in fixed amounts set in advance regardless of how active the trading ends up being before the final estimate (Chen et al, 2010).

From a human factors perspective, our design seeks to use these incentive structures to guarantee non-competitive self-selecting mechanism that encourages the type of diversity in information and decision-making needed for prediction markets to avoid the kind of exuberance that hallmarks "bubbles" in financial markets. In our prediction market, each participant is asked to evaluate the uniqueness of their information before entering any trade. This is further reinforced by the reasonosing from the current market price as to whether it reflects the participant's unique information or whether it justifies the buying or selling of shares to direct the price closer to their estimate. Diversity has value in a prediction market, thus participants with diverse information will self-select to become a trader (Watkins, 2007).

Beyond that however, our design incorporates differences in forecaster knowledge and skill. Specifically, the order size and the amount of money at stake on a given trade serve as a useful proxy for the participant's confidence, for their risk-aversion will make the amount invested be proportional to the gap between their expectations and the current market price. On a long enough timeline, successful traders will be highly rewarded and thus wield greater influence future prices. This feature is consistent with the "marginal trader hypothesis" in economics where the efficiency of a stock market is driven by a minority of unbiased and active participants who wield corrective influence (Wolfers, 2009).

Machine Interface

The machine side of the design includes two types of agents: noisy traders and market makers. We designed them because they provide distinct functions within the prediction market, by making it more effective and precise.

The noisy traders are designed to buy and sell shares of each event randomly and selling them within a narrow interval from their current price. Their behavior creates the opportunity for the human participants to profit on a regular basis, preventing a no-trade scenario that would make the price updates given new information sudden and extreme as opposed to gradual and incremental (Chen et al, 2010). This type of agent resolves many of the concerns related to motivation of the human participants.

The market makers are altogether different. They are not designed to behave randomly, but rather learn how trade from the historical trading data of the prediction itself by using simple machine learning algorithms. Their purpose is to stabilize trading by engaging in the type of behavior high frequency trading firms in financial markets engage in to drive out speculation on price movements. Their function is to also efficiently make bets on forecasts two-sided to reduce imbalances, and the nature of machine learning positions them to rapidly improve their usefulness as they receive new data from the human participants over time. Their trading behavior is extremely important to counterbalance the long-shot biases that plague human-only prediction markets.

Furthermore, prior research with Football forecasts indicates that even though such bots have no understanding of the underlying event being analyzed, they are on average more accurate than the human participants and thus force the participants to refine their analysis which improves their forecasts (Malone, 2018). This type of research suggests that prediction markets would benefit from the type of machine interface we include in our design to mitigate human biases and create a more robust guarantee that prediction markets focus on forecasting as opposed to speculation.

Neural Network Layer

A major contribution of this paper is the layer that sits above the human participants and the trading bots. We include a neural network whose input is the data generated by participants to produce an aggregate estimate of the probability of the event occurring. Artificial neural networks, much like the human brain, use neurons in a that made up of collections of nodes that function as processing units with weighted connections to each other (Kaur & Wasan, 2006). A neural network has a very basic architecture: it has an input layer of neurons that accepts input, a customizable number of hidden layers performing calculations and transformations over the data, and an output layer of neurons that outputs predictions in the selected format (Kaur & Wasan, 2006). Whenever a neural network makes a prediction, the error rate is measured so that the network can adjust the weights of its neurons in order to calibrate its model and achieve better accuracy over time (Simon & Eswaran, 1997).

As mentioned earlier, there is no real consensus on how to translate fluctuating prices in the prediction market into a sensible probability estimate for the event being forecasted. We believe the issue lies in the prior literature emphasis on algorithmic solutions. Instead, we seek to integrate the recent advances in AI to create dynamic as opposed to static solution to the problem, and one that evolves with the prediction market over time.

Prior research has been able to show at a theoretical level how the market scoring rules of prediction markets connect to the equations underlying the value of function of many machine learning models (Chen & Vaughan, 2010). Beyond the theory however, artificial prediction markets can fuse the predictions of trained classifiers into contract prices on all possible outcomes (Barbu & Lay, 2012). The results indicate that such systems outperform cutting edge algorithms that ensemble a variety of models in the healthcare domain, which is attributed to the market mechanism's ability to aggregate specialized classifiers that participate only on specific instances (Barbu & Lay, 2012; Jahedpari et al, 2017)).

This type of result however extends beyond artificial prediction markets where machine learning models are the participants. Tetlock's research on forecasting for IARPA also showed that an extremizing algorithm that took the probabilistic estimates of "superforecasters" as its input actually outperformed 99% of the individual super-forecasters, by aggregating their opinions and weighing them based on track record and diverse POV (Tetlock, 2016).

We go multiple steps further with our design. Not only do we accommodate a hybrid prediction market where humans and AIs alike participate, but we also reject static algorithms in favor of a neural network which is much more effective and capable of aggregating and weighing different the different estimates and viewpoints emerging from the interactions of the prediction markets. The neural network will not merely be learning from price fluctuations, but from the trading behavior of the human participants as well, potentially identifying talented forecasters without having to rely or give undue influence to any of them. Furthermore, our design generalizes to different settings because the neural network obviates the need to perfect knowledge from the group as it detects patterns among participants who themselves may not be expert.

In essence, the machine learning layer enables our prediction market to transcend the limitations inherent to traditional prediction markets by having a neural network learn and grow from the data generated by the human traders.

DISCUSSION & CONCLUSION

Prediction markets have been proven useful in forecasting geopolitical events, sports outcomes, and elections. Yet, they have failed to become ubiquitous despite their success. We believe this is due to the many flaws of traditional prediction markets that privilege algorithms over human effort. By putting collective intelligence at the center of the discussion, our design shows us how a human factors perspective can not only enhance human intelligence, but also make room for artificial intelligence.

On multiple fronts, AI can be the key to overcome many of the challenges of prediction markets. It can downplay the influence of human biases in the market by checking and balancing activity to suppress bubbles before they form. It can induce the participants to think probabilistically and more precisely about their understanding of the event. It can also make the market more adaptive to new information by forcing every participant to update their beliefs based on new information. Through these interventions, the machines in our prediction market smooth out the often erratic behavior of human traders and thus provide a more reliable forecasting mechanism that can extend beyond the few areas where prediction markets have been tried.

Beyond that, however, our model forces us to reconsider how we think about prediction markets in the first place. By de-emphasizing the computation and game theoretical perspective in favor of the cognitive science one, we no longer think about prediction markets as an auction mechanism but rather as a platform for decentralized collaboration between humans trying to tackle the challenges of uncertainty.

Furthermore, our design opens up a new branch of research where the prediction market is seen as a coordination mechanism to enable a different type of cognition: artificial intelligence emerging from collective intelligence. The wisdom of the crowd becomes the input the neural network learns from, whose emergent property is an altogether different type of artificial intelligence that is worth exploring in future research.

PRACTITIONER TAKEAWAYS

Through an interdisciplinary perspective we brought a human factors approach to the design of a more effective prediction market that is not merely optimized for the wisdom of the crowd but also enables a higher level cognitive process that integrates collective intelligence with artificial intelligence.

REFERENCES

Atanasov, P., Rescober, P., Stone, E., Swift, S. A., Servan-Schreiber, E., Tetlock, P., ... & Mellers, B. (2016). Distilling the wisdom of crowds: Prediction markets vs. prediction polls. Management science, 63(3), 691-706.

Attarwala, A., Dimitrov, S., & Obeidi, A. (2017, September). How efficient is Twitter: Predicting, 201 2 US Presidential elections using Support Vector Machine via Twitter and comparing against Iowa Electronic Markets. In, 2017 Intelligent Systems Conference (IntelliSys) (pp. 646-65, 2). IEEE.

Barbu, A., & Lay, N. (2012). An introduction to artificial prediction markets for classification. Journal of Machine Learning Research, 13(Jul), 2177-, 2, 204.

Chen, Y., Chu, C. H., Mullen, T., & Pennock, D. M. (2005, June). Information markets vs. opinion pools: An empirical comparison. In Proceedings of the 6th ACM conference on Electronic commerce (pp. 58-67). ACM.

Chen, Y., Dimitrov, S., Sami, R., Reeves, D. M., Pennock, D. M., Hanson, R. D., ... & Gonen, R. (2010). Gaming prediction markets: Equilibrium strategies with a market maker. Algorithmica, 58(4), 930-969.

Chen, Y., & Vaughan, J. W. (2010, June). A new understanding of prediction markets via no-regret learning. In Proceedings of the 11th ACM conference on Electronic commerce (pp. 189-198). ACM.

Cooke, N. J., Gorman, J. C., Winner, J. L., & Durso, F. T. (2007). Team cognition. Handbook of applied cognition, 2, 239-, 268.

Dudik, M., Lahaie, S., Rogers, R. M., & Vaughan, J. W. (2017). A decomposition of forecast error in prediction markets. In Advances in Neural Information Processing Systems (pp. 4371-4380).

Goel, S., Reeves, D. M., Watts, D. J., & Pennock, D. M. (2010, June). Prediction without markets. In Proceedings of the 11th ACM conference on Electronic commerce (pp. 357-366). ACM.

Hanson, R. (2003). Combinatorial information market design. Information Systems Frontiers, 5(1), 107-119.

Jahedpari, F., Rahwan, T., Hashemi, S., Michalak, T. P., De Vos, M., Padget, J., & Woon, W. L. (2017). Online prediction via continuous artificial prediction markets. IEEE Intelligent Systems, 3, 2(1), 61-68.

Kaur, H., & Wasan, S. K. (2006). Empirical Study on Applications of Data Mining Techniques in Healthcare. Journal of Computer Science, 2(2), 194-, 200. doi:10.3844/jcssp., 2006.194., 200

Laskey, K. B., Hanson, R., & Twardy, C. (2015, July). Combinatorial prediction markets for fusing information from distributed experts and models. In, 2015 18th International Conference on Information Fusion (Fusion) (pp. 189, 2-1898). IEEE.

Malone, T. W. (2018). Superminds: The Surprising Power of People and Computers Thinking Together. Little, Brown.

Marcus, B. K. (2004). Can markets predict elections. Mises institute.(Accessed 2 April 2005).

McNeese, N. J., Reddy, M. C., & Friedenberg, E. M. (2014, September). Towards a team mental model of collaborative information seeking during team decision-making. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 58, No. 1, pp. 335-339). Sage CA: Los Angeles, CA: SAGE Publications.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training.

Satopää, V. A., Baron, J., Foster, D. P., Mellers, B. A., Tetlock, P. E., & Ungar, L. H. (2014). Combining multiple probability predictions using a simple logit model. International Journal of Forecasting, 30(2), 344-356.

Simon, B. P., & Eswaran, C. (1997). An ECG Classifier Designed Using Modified Decision Based Neural Networks. Computers and Biomedical Research, 30(4), 257-, 27, 2. doi:10.1006/cbmr.1997.1446

Surowiecki, J. (2005). The wisdom of crowds. Anchor.

Tetlock, P. E., & Gardner, D. (2016). Superforecasting: The art and science of prediction. Random House.

Wagstaff, K. (2012). Machine learning that matters. arXiv preprint arXiv:1, 206.4656.

Watkins, J. H. (2007). Prediction markets as an aggregation mechanism for collective intelligence.

Wolfers, J. (2009). Prediction markets: The collective knowledge of market participants. In CFA Institute Conf. Proc. Quart. (Vol., 26, No., 2, pp. 37-44).