Breakthrough Markets

Crowdsourcing Scientific Breakthroughs with Hybrid Prediction Markets

Dr. Lorenzo Barberis Canonico, RomBot Inc.
Andres Castaneda, Stripe Inc.

1. INTRODUCTION

One of the major challenges plaguing the scientific community recently is the replication crisis. All the way back to 2015, the Reproducibility Project revealed that less than half of the major published results being examined ended up being replicated [Mann 2016]. Specifically, out of over 100 experiments in three high-ranking psychology journals, between a half and two-thirds did not display the original findings once a replication study was completed [Collaboration et al. 2015]. Not only does the replication crisis undermine trust in the scientific community, but it also impairs scientific progress as unreliable results nudge future research in the wrong direction.

However, the replication crisis is only one component of a much larger problem of scientific stagnation. In the private sector and corporate RD, not only is basic research on the decline (basic research by corporate scientists fell by 60% from 1980 to 2006), but commercialized innovation has also slowed down (patent registration was higher in 1966 than in 1993) [Gruber and Johnson 2019; Cowen 2011]. This deceleration in research productivity has been attributed, especially in the pharmaceutical sector, to the free-rider problem inherent to R&D [Evangelatos et al. 2016]. Another explanation has focused on the diminishing return of research capital past a certain stage in the scientific development process; for example, despite Moore's Law, the number of researchers required in order to double chip density today is more than eighteen times the number who were needed in the early 1970s [Gruber and Johnson 2019].

The Replication Crisis and the decelerating rate of research productivity both emerge from the same problem: a distorted incentive structure that directs researchers away from high-risk scientific breakthroughs and towards lower-risk incremental progress. However, collective intelligence can inform the design of mechanisms that can more effectively coordinate the allocation of research capital towards breakthroughs. Specifically, in this extended abstract we discuss the design for a new type of prediction market that leverage both human and machine intelligence to coordinate scientists and researchers towards scientific breakthroughs occur.

2. BACKGROUND

2.1 Crowdsourced Scientific Breakthroughs

Whether in math, biochemistry, or micro-biology, there have been several examples of crowdsourcing being used effectively to produce a major breakthroughs at a much more rapid pace than anticipated. For instance, in 2009 the mathematician Timothy Gowers and Terence Tao used their blogs to coordinate a crowdsourcing effort to find a new combinatorial proof to the density version of the Hales-Jewett theorem [Gowers and Nielsen 2009]. After just 7 weeks, the project announced that the problem was officially solved [Malone 2018]. Not only did the crowdsourcing effort result in a major mathematical breakthrough, but it also delivered such impressive results in such a short timeline.

However, such examples are far from limited to mathematics: crowdsourcing has also led major advancements in biochemistry. Specifically, in 2010 a team of scientists developed a gamified platform called Foldit where researchers and players could attempt to predict the shape of proteins from their genetic sequence and other general information. In just 3 weeks, the community discovered the protein structure of an AIDS-related enzyme that had eluded the scientific community for over 15 years [Cooper et al. 2010]. Similarly, in 2019 a team of cell biologists created a videogame where players had to identify the malaria species of infected cells shown on a screen. After 500,000 assessments, the aggregated decisions of non-expert players reached 99% accuracy even though each decision was made in less 3s [Linares et al. 2019].

These examples show that we have only begun to scratch the surface of how collective intelligence mechanisms can accelerate scientific breakthroughs.

2.2 Crowdsourcing Forecasts with 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 (the Iowa Electronic Markets beats presidential polls 75% of the time), project management (both Microsoft and Google have used prediction markets to more effectively forecast project completion time), and box office performance (the Hollywood Stock Exchange correctly picked 35 out of the eventual 40 Oscar nominees in 2002) [Surowiecki 2005; Malone et al. 2009]. The key to the success of prediction markets lies in their ability to aggregate implicit, dispersed, and broadly inaccessible knowledge and combine diverse opinions to parse signal from noise by incentivizing the correction of systemic biases, such as overconfidence and underconfidence, in order to outperform elite experts [Hanson 2003; Atanasov et al. 2016; Marcus 2004].

However, prediction markets have been shown to have their own set of limitations. Although they have been remarkable at forecasting replicability in psychology (71% accuracy), results have been more disappointing in economics (overestimation throughout all 18 studies being examined) [Dreber et al. 2015; Camerer et al. 2016]. To understand these limitations, prior research has identified three major types of errors in prediction markets that emerge from the behavioral biases exhibited by the human participants: sampling error (traders hold noisy estimates that dilute the truth-value of their information), market-maker bias (the cost function being used to facilitate trading ends up anchoring the participants' estimates), and convergence error (huge market fluctuations destabilizing trading) [Dudik et al. 2017].

In essence, a prediction market can be reframed as a collective intelligence mechanism that financially incentivizes participants to contribute to the crowdsourcing of better-than-expert forecasts regarding uncertainties in the future.

3. HYBRID PREDICTION MARKET

Prior attempts at using prediction markets for scientific research have been limited to mechanisms where the objective was to forecast either the replicability of a pre-existing study (replication markets) or the update to a pre-existing hypothesis given the release of a new dataset or algorithm (Crowdsourced Learning Mechanism) [Mann 2016; Abernethy and Frongillo 2011]. Our model in Figure 1 (reproduced with permission from the author) goes one step further, and repurposes the design of Barberis Canonico et al's hybrid prediction market (where bots trade alongside humans and the trading data is fed into a neural network to de-bias the forecasts) [Barberis Canonico et al. 2019]. However, in their case the objective is to extend Malone et al's research on Football forecasts where the accuracy of the trading bots forces the participants to refine their analysis thereby improving their forecasts, as well as Tetlock's research on Superforecasting where an extremizing algorithm aggregating and track-record-weighing the estimates of the Superforecasters actually outperformed 99% of the individual super-forecasters [Malone 2018; Tetlock and Gardner 2016]. In our case however, the objective is

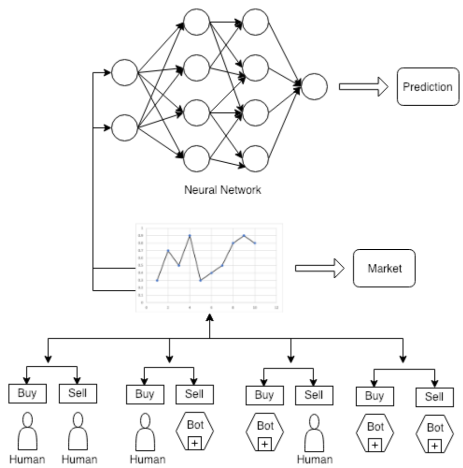

Fig. 1. humans and bots trade against each other until the market price for event stabilizes. The data generated by process is then fed as the input to the neural network that calibrates its aggregate estimate of all the participants over time.

not to estimate the accuracy of specific hypothesis, but rather to crowdsource the weight of the edges of a bayesian network mapping the combinatorial relationships between different concepts within the scientific domain.

Bayesian networks are networks where the weight of the edges are updated to reflect the probability of a relationship between two nodes as new evidence is incorporated. A primary application of Bayesian networks for science is in mathematics, where researchers have plotted theorems, proofs, and conjectures and their perceived relationships to estimate when the P vs. NP conjecture would be solved [Gasarch 2012]. However, this approach alone has fallen short because the estimates based on reference-analysis of prior conjectures and surveys of mathematicians are not as accurate in forecasting future proofs [Hisano and Sornette 2012]. The solution is to crowdsource more accurate estimates of the probabilities that two proofs can be used to prove a conjecture through a combinatorial prediction market, where participants can bet on the relationship between two events as opposed to just single events [Hanson 2003].

The key aspect of our mechanism however is that the focus is less on minimizing making mistakes in estimating the likelihood a particular research path will lead to a breakthrough, and more on generating a real-time map of the most promising combinations of research concepts for scientists and researchers to investigate. Indeed, given the complexity of science in the status quo, prior research has highlighted interdisciplinary work as the primary strategy to promote breakthroughs [Azoulay 2019]. To that end, having a hybrid-prediction-market-backed ranked list of the most promising combinations of scientific concepts and results becomes an accelerator to the formation of interdisciplinary teams, whose efforts in turn are directed at research whose results are intersubjectively expected to lead to the biggest update in the bayesian network.

Overall, as opposed to Replication Markets, which seek to validate prior research, Breakthrough Markets seek to combine the expertise of all scientists and researchers to identify opportunities for interdisciplinary breakthroughs by crowdsourcing a bayesian network for scientific domains through a prediction market that merges human intelligence with artificial intelligence.

REFERENCES

Jacob D Abernethy and Rafael M Frongillo. 2011. A collaborative mechanism for crowdsourcing prediction problems. In *Advances in Neural Information Processing Systems*. 2600–2608.

Pavel Atanasov, Phillip Rescober, Eric Stone, Samuel A Swift, Emile Servan-Schreiber, Philip Tetlock, Lyle Ungar, and Barbara Mellers. 2016. Distilling the wisdom of crowds: Prediction markets vs. prediction polls. *Management science* 63, 3 (2016), 691–706.

Pierre Azoulay. 2019. Small research teams 'disrupt'science more radically than large ones. (2019).

Lorenzo Barberis Canonico, Nathan J McNeese, and Chris Flathmann. 2019. The Wisdom of the Market: Using Human Factors to Design Prediction Markets for Collective Intelligence. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 63. SAGE Publications Sage CA: Los Angeles, CA.

Colin F Camerer, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, and others. 2016. Evaluating replicability of laboratory experiments in economics. *Science* 351, 6280 (2016), 1433–1436.

Open Science Collaboration and others. 2015. Estimating the reproducibility of psychological science. *Science* 349, 6251 (2015), aac4716.

Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popović, and others. 2010. Predicting protein structures with a multiplayer online game. *Nature* 466, 7307 (2010), 756–760.

Tyler Cowen. 2011. *The great stagnation: How America ate all the low-hanging fruit of modern history, got sick, and will (eventually) feel better: A Penguin eSpecial from Dutton*. Penguin.

Anna Dreber, Thomas Pfeiffer, Johan Almenberg, Siri Isaksson, Brad Wilson, Yiling Chen, Brian A Nosek, and Magnus Johannesson. 2015. Using prediction markets to estimate the reproducibility of scientific research. *Proceedings of the National Academy of Sciences* 112, 50 (2015), 15343–15347.

Miro Dudik, Sébastien Lahaie, Ryan M Rogers, and Jennifer Wortman Vaughan. 2017. A decomposition of forecast error in prediction markets. In *Advances in Neural Information Processing Systems*. 4371–4380.

Nikolaos Evangelatos, Matthias Reumann, Hans Lehrach, and Angela Brand. 2016. Clinical Trial Data as Public Goods: Fair Trade and the Virtual Knowledge Bank as a Solution to the Free Rider Problem-A Framework for the Promotion of Innovation by Facilitation of Clinical Trial Data Sharing among Biopharmaceutical Companies in the Era of Omics and Big Data. *Public health genomics* 19, 4 (2016), 211–219.

William I Gasarch. 2012. Guest column: the second P=? NP poll. *ACM SIGACT News* 43, 2 (2012), 53–77.

Timothy Gowers and Michael Nielsen. 2009. Massively collaborative mathematics. *Nature* 461, 7266 (2009), 879–881.

Jonathan Gruber and Simon Johnson. 2019. *Jump-starting America: How breakthrough science can revive economic growth and the American dream*. Blackstone Audio.

Robin Hanson. 2003. Combinatorial information market design. *Information Systems Frontiers* 5, 1 (2003), 107–119.

Ryohei Hisano and Didier Sornette. 2012. On the distribution of time-to-proof of mathematical conjectures. *arXiv preprint arXiv:1202.3936* (2012).

María Linares, María Postigo, Daniel Cuadrado, Alejandra Ortiz-Ruiz, Sara Gil-Casanova, Alexander Vladimirov, Jaime García-Villena, José María Nuñez-Escobedo, Joaquín Martínez-López, José Miguel Rubio, and others. 2019. Collaborative intelligence and gamification for on-line malaria species differentiation. *Malaria journal* 18, 1 (2019), 21.

Thomas W Malone. 2018. *Superminds: The surprising power of people and computers thinking together*. Little, Brown.

Thomas W Malone, Robert Laubacher, and Chrysanthos Dellarocas. 2009. Harnessing crowds: Mapping the genome of collective intelligence. (2009).

Adam Mann. 2016. The power of prediction markets. *Nature News* 538, 7625 (2016), 308.

BK Marcus. 2004. Can markets predict elections. *Mises institute.(Accessed 2 April 2005)* (2004).

James Surowiecki. 2005. *The wisdom of crowds*. Anchor.

Philip E Tetlock and Dan Gardner. 2016. *Superforecasting: The art and science of prediction*. Random House.