Dense Neural Network used to Amplify the Forecasting Accuracy of real-time Human Swarms

Gregg Willcox
Unanimous AI
San Francisco, CA, USA
Gregg@Unanimous.ai

Louis Rosenberg
Unanimous AI
San Francisco, CA, USA
Louis@Unanimous.ai

Rory Donovan
Unanimous AI
San Francisco, CA, USA
Rory@Unanimous.ai

Hans Schumann
Unanimous AI
San Francisco, CA, USA
Hans@Unanimous.ai

***Abstract*— Artificial Swarm Intelligence (ASI) is a hybrid AI technology that enables distributed human groups to "think together" in real-time systems modeled on natural swarms. Prior research has shown that by forming "human swarms," networked groups can substantially amplify their combined intelligence and produce significantly more accurate forecasts than traditional methods. The present study explores whether the rich behavioral data collected during "swarming" can be used to further increase the accuracy of swarm-based forecasts. To do this, a dense neural network was used to process the data collected during a set of swarm-based forecasts and generate a Conviction Index (CI) for each forecast that estimates its expected accuracy. This method was then tested in an authentic forecasting task – wagering on sporting events against the Vegas odds. Specifically, groups of sports fans, working as real-time swarms, were tasked with predicting the outcome of 238 NBA games over 25 consecutive weeks. As a baseline, the swarms achieved an impressive 25% net return on investment (ROI) against the Vegas Odds. This was compared to an enhanced method that used Conviction Index to (a) estimate the strength of each forecast and then (b) wager only on forecasts of sufficient strength. The CI-selected wagers yielded a 57% net ROI against Vegas Odds. This is a significant gain, equivalent to more than doubling the ROI of the naïve swarm betting strategy.**

Keywords*— Swarm Intelligence, Artificial Swarm Intelligence, Collective Intelligence, Human Swarming, Artificial Intelligence, Collaborative Intelligence, Machine Learning, Sports Forecasting, Wisdom of Crowds.

I. INTRODUCTION

The technology of Artificial Swarm Intelligence (ASI) has been shown to amplify the predictive accuracy of networked human groups across a variety of tasks [1-6]. Prior studies have shown that real-time "human swarms" can produce more accurate forecasts than traditional "Wisdom of Crowd" methods such as votes, polls, surveys, and markets.

For example, a 2015 study tested the ability of human swarms to forecast the outcome of NCAA college football games against Vegas betting markets. A swarm of 75 average sports fans was tasked with predicting a set of 10 college bowl games. As individual forecasters, the participants averaged 50% accuracy when predicting game outcomes against the Vegas spread. When forecasting together as real-time human swarms, those same participants achieved 70% accuracy against the Vegas spread [2].

Similar performance increases have been found in other studies, including a five-week study that tasked human participants with predicting a set of 50 soccer matches in the English Premier League. Results showed a 31% increase in accuracy when participants worked in swarms [4]. Human swarms were also shown to outperform Vegas betting markets in a 20-week study that involved predicting the outcome of 200 National Hockey League games. By using ASI technology, human swarms were shown to reduce the expected error rate by 61% on a subset of games [6].

While prior studies have documented the ability of artificial human swarms to amplify the predictive ability of human populations and outperform individual forecasters, statistical aggregations from large crowds of forecasters, computer models, and largescale betting markets, no formal study has studied the estimation of expected accuracy of swarm forecasts with machine learning. Such a machine learning model would allow deeper insights into human swarm behavior, paving the way for the optimization of ASI systems and the widespread application of swarm-based forecasting to diverse problems, such as financial, geopolitical, or sports forecasting.

To address this, the current study develops a machine learning model that processes the behavioral data from human swarms, generates a Conviction Index (CI) that reflects the expected accuracy of the swarm, and predicts the expected ROI of placing a bet on the game against a largescale betting market (i.e. the published Vegas odds). The study then pits the machine learning model against Vegas, computing financial returns for theoretical bets placed against the real-world odds and payouts in a full season of the National Basketball Association (NBA). The model's betting success against Vegas is compared to a naïve model of betting on all games. The present study considered 25 consecutive weeks of NBA games, requiring human swarms to forecast between six and eleven games per week, for a total of 238 games predicted.

The study is organized as follows: in Section II, we introduce the concept of "human swarming" and discuss biological basis for optimized swarm-based decision-making. In Section III, a cloud-based technology platform for real-time human swarming (swarm.ai) is introduced, and examples of swarms are provided. In Section IV the experimental methodology behind this forecasting study is described. Finally, the results of the study are analyzed in Section V.

II. SWARMS AS INTELLIGENT SYSTEMS

Given a population in which each individual has a unique set of information about the world, how do we best combine their perspectives and reach an optimal solution? Researchers have been trying to solve this problem for centuries using techniques that are now commonly referred to as harnessing the "Wisdom of Crowds" or simply crowdsourcing [7,8,9]. These methods generally involve taking votes, conducting polls, collecting surveys, or running information markets. Most crowd-based methods capture input from each human participant in isolation (or near isolation) from other members and then combine the data from the full set of members through statistical aggregation, either over time or post-hoc. In other words, these "crowds" are not actually groups of people interacting freely as real-time collaborative systems, but instead are statistical constructs for mathematical analysis.

Mother Nature has been working on methods to harness the diverse perspectives of populations, having explored this issue across many millions of years of biological evolution. The successful solutions that evolved in nature do not involve taking votes, conducting polls, collecting surveys or running prediction markets – they involve forming dynamic systems in which the full population is enabled interact in real-time and converge together on optimal solutions. Biologists refer to this phenomenon as Swarm Intelligence. It's one of the primary reasons why birds flock, fish school, and bees swarm – they're able to combine their insights in optimal ways, becoming significantly smarter together than alone.

The most researched form of Swarm Intelligence in nature is the honeybee swarm. Studied since the 1950s, the decision-making abilities of honeybee swarms have been shown to be very similar to the decision-making processes in neurological brains [10,11]. Both employ large populations of simple excitable units (i.e., bees and neurons) that work in unison to integrate noisy information about the world, weigh competing alternatives, and converge on unified decisions in real-time synchrony. In both brains and swarms, outcomes are arrived at through a competition among sub-populations of excitable units. When one sub-population exceeds threshold support, the corresponding alternative is chosen. In honeybees, this enables the large colonies to converge on optimal decisions to highly complex problems such as selecting an optimal home location from among a large set of alternatives [12,13,14].

III. ENABLING "HUMAN SWARMS"

Unlike birds, bees and fish, we humans have not evolved the natural ability amplify our combined intelligence by forming real-time swarms. That's because we lack the subtle connections that other organisms use to form feedback loops among members. Schooling fish detect vibrations in the water around them. Flying birds detect motions propagating through the flock. Swarming bees use complex body vibrations called a "waggle dance." To enable real-time swarming among groups of networked people, specialized user interfaces and algorithms are required to close the loop among all members.

To address this need, a software platform called *swarm.ai* was developed to enable human groups to link online as real-time synchronous systems, connecting from anywhere in the world [15-18]. Modeled on the decision-making process employed by honeybee swarms, the system allows groups of distributed users to work in parallel to (a) integrate noisy evidence, (b) weigh competing alternatives, and (c) converge on in synchrony on optimized solution, all while allowing participants to react to the changing system in real-time, thereby closing a feedback loop around the full population.

As shown in Figure 3, the software used in this study enables human swarms to answer questions by collaboratively moving a graphical pointer depicted as a glass puck. Answers are reached when the swarm moves the puck from the center of the screen to one of set of answer options. Each participant provides input by manipulating a graphical magnet with a mouse or touchscreen. By positioning their magnet with respect to the moving puck, participants impart their personal intent on the swarm as a whole. The input from each user is not a discrete vote, but a stream of vectors that varies freely over time, enabling the swarm to move, not based on the input of any individual, but based on the dynamics of the full system. In this way, the group to explores the decision-space and converges on the most agreeable solution in synchrony.

It is important to note that participants do not only vary the direction of intent, but also modulate the magnitude of intent by adjusting the distance between their magnet and the puck. Because the puck is in continuous motion across the decision-space, users need to continually move their magnet so that it stays close to the puck's outer rim. This is significant, for it requires participants to be engaged continuously throughout the decision process, evaluating and re-evaluating their intent as they convey their contribution. If they stop adjusting their magnet with respect to the changing position of the puck, the distance grows and their applied sentiment wanes.

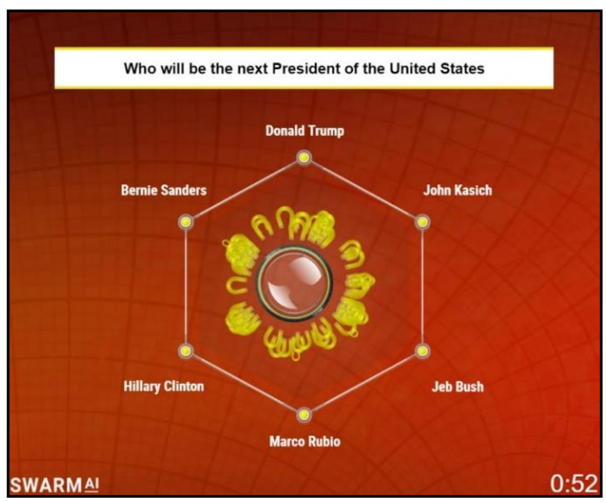

Fig. 3. A human swarm answering a question in real-time

Thus, like bees vibrating their bodies to express sentiment in a biological swarm, or neurons firing to express conviction within a biological neural-network, the human participants in an artificial swarm must continuously update their intent during the ongoing decision process or lose influence. In addition, intelligence algorithms monitor the behaviors of all swarm members in real-time, inferring their implied conviction based upon their relative motions over time.

IV. SWARM CONVICTION STUDY

To assess whether the behavioral patterns within the deliberation data from human swarms can be used to estimate the expected accuracy of forecasts, a formal study was conducted using groups of randomly selected human subjects from a pool of self-reported NBA enthusiasts. Each weekly group consisted of 28 to 43 participants, all of whom logged in remotely to the Swarm system. Each subject was paid $4.00 for their participation in each weekly session, which required them to predict of the outcome of all of the basketball games being played that night, first as (a) individuals on a standard online survey, and then (b) as part of a real-time swarm comprised of the full population.

Across the 25-week period, predictions were generated by for between six and eleven games per week for a total of 213 games. For each game, participants were required to work together as an ASI system to forecast the winner of each game, and converge on their collective level of confidence in this forecast ("Low Confidence" or "High Confidence"). Participants were then asked to predict, by working together as a swarm, how much the team they picked would win by on a scale from "1" to "15+" points.

Figure 4 shows a snapshot of a human swarm comprised of 32 participants in the process of predicting the outcome of a typical NBA game: Washington vs San Antonio. As shown, four options are provided to choose from, enabling the swarm to identify which team will win, as well as express a level of confidence in that outcome. Participants are not voting, but behaving – continuously expressing their views in real-time. The Swarm AI system processes the participants' behaviors and controls the motion of the full system. The confidence indicator is helpful as it causes the swarm to split into multiple different factions and then converge over time on a single solution that maximizes their collective confidence and conviction. It's important to note that Figure 4 shows a snapshot of the swarm as it moves over time towards a final answer. The full process of converging upon a solution generally required between 10 and 30 seconds of real-time interaction within the swarm.

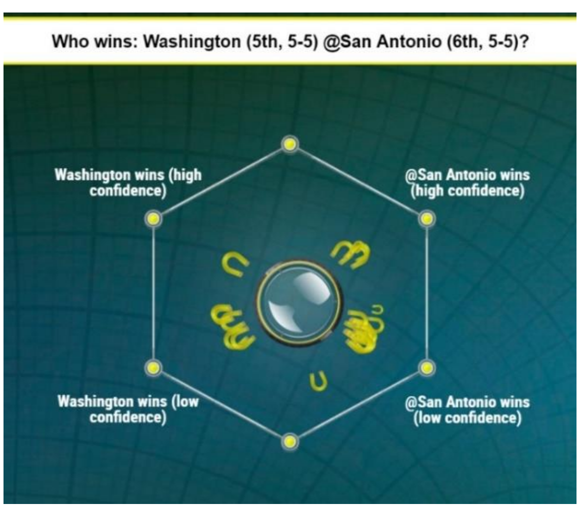

Fig. 4. Human Swarm in the process of forecasting NHL game

To estimate the relative expected accuracy for each forecast generated by the ASI system, a dense neural network (the Swarm Conviction Estimator) was trained using the behavioral deliberation data captured during each swarm and used that data to predict the probability that the swarm's forecast was correct. This behavioral deliberation data includes (i) the percentage of users pulling for each target sampled at various times throughout the swarm, (ii) the total number of users in the swarm, and (iii) the time the swarm took to converge on a forecast, among other behavioral indicators.

The network is trained using the time-varying behavioral deliberation data from a historical database of 424 swarm predictions of NFL and NHL games. The range of reasonable probabilities for each sport differs greatly (e.g. the distribution of Vegas Odds for NHL is much narrower than the same distribution for NFL), so the network's outputted probabilistic forecast cannot be considered a calibrated probability for a given sport, but rather a relative measure of the swarm's conviction in the chosen outcome. Each relative conviction, referred to as a Conviction Index (CI), can therefore be used in a single sport, such as NBA, to rank forecasts from lowest to highest expected accuracy.

To validate the accuracy and precision of the Swarm Conviction Estimator in a real-world environment, the conviction scores were compared to Vegas Odds, and simulated bets were placed on the outcomes of games. To decide which games to bet on, an ROI Estimator was developed to predict the expected ROI of betting on the swarm's chosen outcome based on the CI and Vegas odds of the match. The Vegas Odds were sourced from Sportsbook, a widely-used online bookie. This ROI Estimator is a random forest that was trained on a database of 243 swarm NHL and 181 swarm NFL forecasts, each of which had an associated CI and Vegas Odds.

When the expected ROI from the ROI Estimator is positive (>0%), betting on the chosen outcome is expected to be profitable. Games were selected from the pool of NBA games each week using one of four strategies: (a) betting on the swarm's pick in all games, (b) betting on the swarm's pick in all games with a positive expected ROI, (c) betting on the swarm's pick in all games with an expected ROI above 10%, and (d) betting on the swarm's pick in all games with an expected ROI above 20%. These strategies were designed to simulate progressively more aggressive betting strategies, from betting on all games to betting on only a select few games that are expected to return a significant payout.

The experimental simulations started with a mock wager pool of $100, and a betting rule directing that a total of 15% of the gambling pool would be bet each week, regardless of the games selected to bet on that week. The expected ROI for betting on each of the swarm's forecasted outcomes was calculated using the Swarm Conviction Estimator and the ROI Estimator, as shown in Figure 5. Simulated bets were placed each week on each strategy's selected games, and the simulated return on the investment was calculated given the outcome of the bet (win / loss) and the Vegas Odds. The net return on investment was then added to that strategy's gambling pool for the next week.

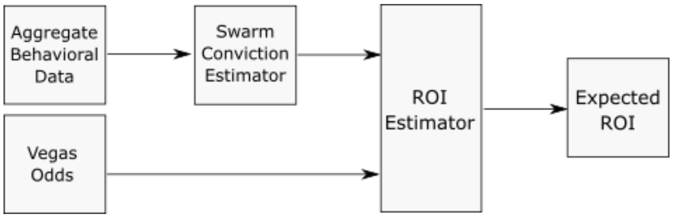

Fig. 5. System Diagram of ROI Estimation from Human Swarm Behavior and Vegas Odds

V. RESULTS

The results of the experiment are discussed in two parts. First, the accuracy and betting performance of the human swarms over all games is discussed and compared to the Vegas Odds. Next, the accuracy and betting performance of the CI-selection methods are discussed and compared to the uninformed all swarm picks method. To assess whether human swarms were able to more accurately forecast all NBA outcomes than Vegas, the swarm's raw forecasts for all games each week were compared against the Vegas Odds for the corresponding game for each of the 25 weeks of the testing period. Vegas' expected win rate for these selected games was calculated as the average Vegas Odds over all games that the swarm selected as Pick of the Week. Figure 6 shows the distribution of Vegas Odds for the selected games, and Vegas' expected win rate: 66.5%. The swarm, on the other hand, had a win rate of 71.8% across these same games. This is a valuable improvement, equivalent to outperforming Vegas' expectations by more than 5%.

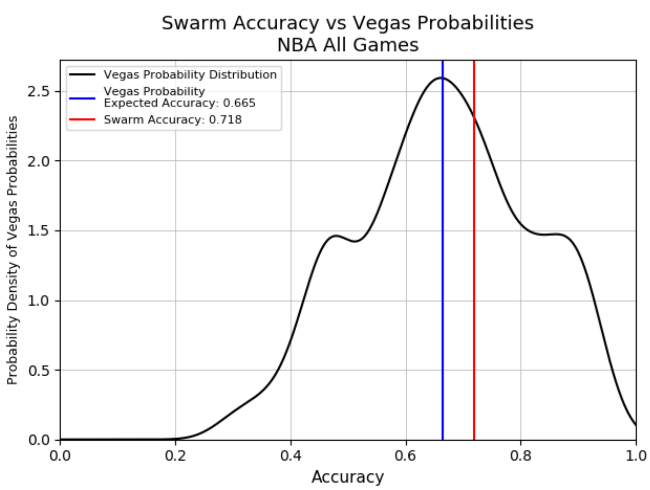

Fig 6. Vegas vs Swarm accuracy across all games predicted

To examine the significance of this result, the average accuracy of each system over the full season was bootstrapped 10,000 times. The average accuracies for each trial are shown in Figure 7. We find that the probability that the swarm had a higher win rate than Vegas Odds due to chance was low (p=.0306), so we can be confident that these swarms were able to predict the outcome of games with higher accuracy than Vegas Odds.

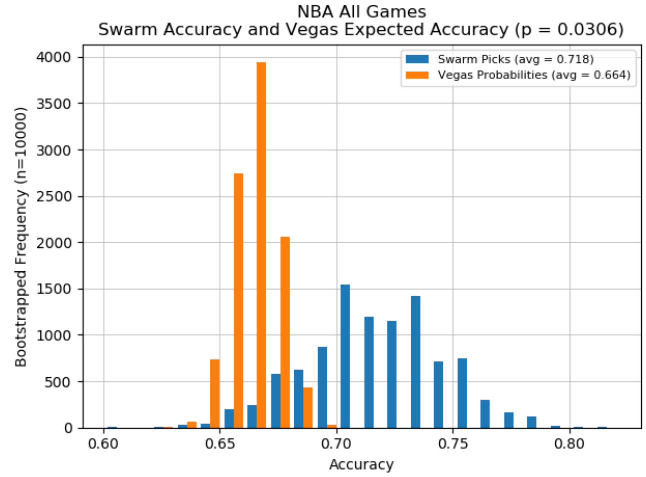

Fig 7. Bootstrapped average accuracy for Vegas vs all Swarm picks

In addition, a betting simulation was run for each prediction set in which 15% of the current bankroll was distributed evenly among bets on each of the swarm's predictions that week. The performance of this model when betting against Vegas (and including the Bookie's cut) is seen in Figure 8. Starting with \$100 and investing each week according to this strategy, the net balance after 25 weeks would be \$124.74, or an ROI of 24.7%.

A bootstrapped simulation was performed to estimate a 90% confidence interval around this result, where 10,000 simulated seasons were generated by randomly selecting with replacement among the games that were seen each week. We find that the 90% confidence interval over the ROI of this betting strategy is [-7.48%, 61.69%], indicating that we are not confident that betting on all swarm picks would return a positive ROI (p=0.112).

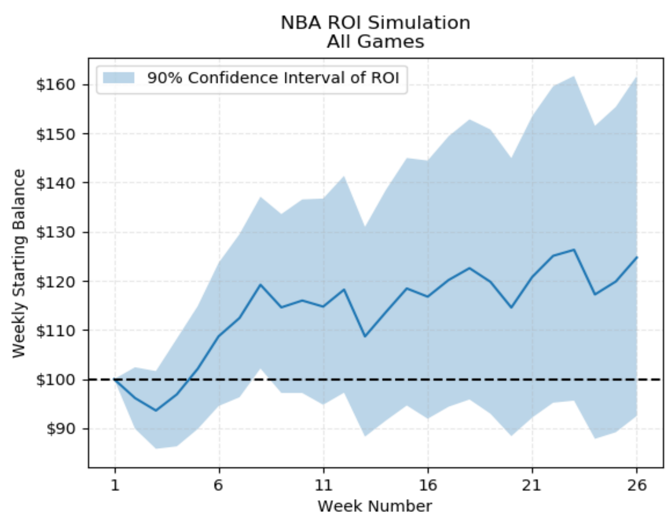

Fig 8. Cumulative simulated betting performance of fixed bets on all games predicted

So, while the swarm was significantly more accurate at predicting outcomes than Vegas Odds, we cannot be confident that betting on the swarm outcomes would return a positive ROI. Two factors could have contributed to this difference: (a) Vegas Odds includes a 2-5% "Bookie's Cut" in all outcomes to allow sportsbooks to make money, impacting the ROI simulation, but averaged out for the Accuracy analysis, and (b) the compounding nature of the simulation's bankroll increases the variability of the success of this betting strategy relative to Vegas Odds.

To assess whether the behavioral patterns in these swarms could be used to precisely forecast the outcome of games, we next compared the performance of CI-selection methods to the performance of Vegas Odds over the selected games. To do so, the Expected ROI of each of the 238 games was calculated using the Swarm Conviction Estimator and ROI Estimation machine learning programs. The Expected ROI of each game was used to determine if the game should be bet upon. Three strategies for betting on these values are compared: (1) betting on all games with an expected positive ROI, (2) betting on games with an expected ROI above 10%, and (3) betting on games with an expected ROI above 20%. Of the total 238 games, these betting strategies selected 202, 137 and 92 games to bet on respectively.

The accuracy and ROI from these selections of games, referred to as the CI-selected games, was compared against the accuracy and ROI from betting on swarm picks over all games. The simulated performance of all models when betting against Vegas (including the impact of the Bookie's cut) is shown in Figure 9. In these simulations, the higher the expected ROI cutoff of the betting strategy, the higher the season-end ROI. The strategy with the highest ROI was the CI-selected 20%+ method, which returned a 56.6% ROI over the 25-week season.

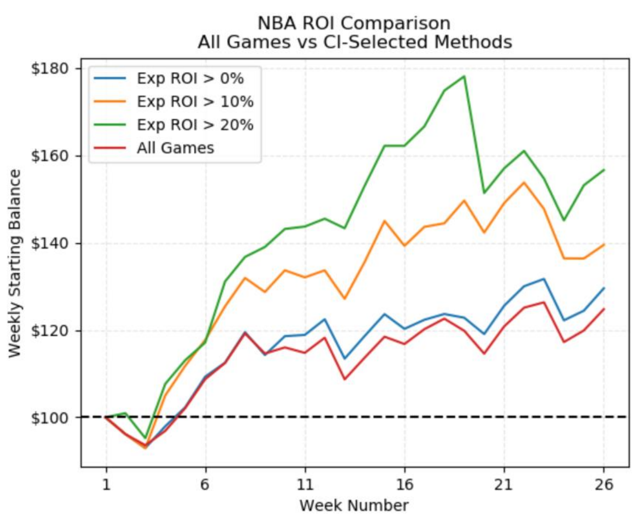

Fig 9. Cumulative simulated betting performance of all vs. CI-selected games at various thresholds

The final ROI of each method is shown in Table 1. To assess whether this amplification of ROI is significant compared to the all-swarm-picks method, the season-end ROI of each selection method were compared over 10,000 bootstrapped season simulations. The CI-selection methods were found to frequently outperform the all-swarm-picks method over a full simulated season (up to 77% of the time), but not frequently enough to be confident that the CI-selection methods were outperforming the all-picks method due to random chance (p=0.23).

The probability that each of these methods returned a positive ROI due to random chance was calculated over these 10,000 bootstrapped seasons as shown in Table 2. Notably, the 20% CI-Selected games generated a positive return on investment 89.11% of the time, meaning that this betting strategy had roughly an impressive 9 in 10 chance of ending the season with a financial gain.

Selection Method	**ROI (end of season)**	**Probability of Outperforming All Swarm Picks**	**Probability of Positive ROI**
All Swarm Picks (238 games)	24.7%	-	88.80%
Expected ROI > 0% (202 games)	29.5%	0.3304	88.87%
Expected ROI > 10% (137 games)	39.5%	0.3059	87.55%
Expected ROI > 20% (92 games)	56.6%	0.2330	89.11%

Table 1. Simulated betting performance of all vs. CI-selected games at various thresholds

To investigate why the ROI can be doubled as compared to the All Swarm Picks method, but statistical significance was not found, the bootstrapped season-end ROI histogram was plotted in Figure 10. The variance of the bootstrapped ROI of the most aggressive strategy (Expected ROI > 20%) was high in comparison to the All Swarm Picks method, likely because of the small sample size of the method: it selected only 40% of games to bet on.

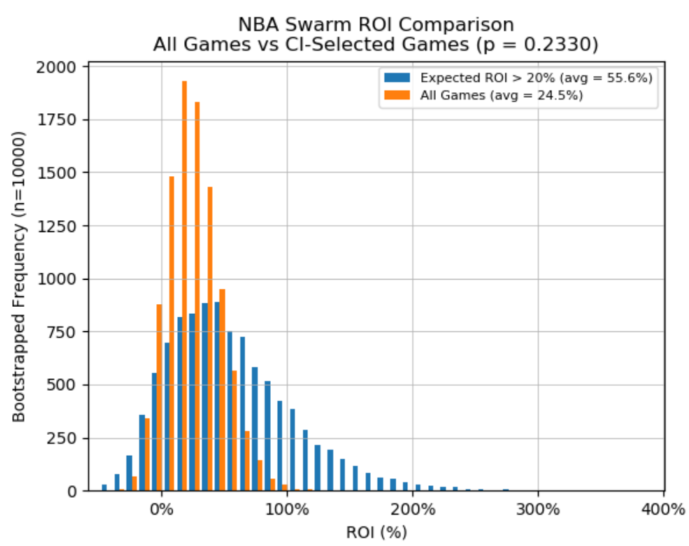

Fig 10. Bootstrapped average accuracy of all swarm picks vs CI-selected picks (expected ROI > 20%)

As games were selected with high expected ROI, the simulation ROI increased. This suggests that the Swarm Conviction Estimator and ROI Estimation programs are translating the swarm behavior into an accurate relative ranking system that can be used to select games where the Vegas Odds are inaccurate, and a positive ROI can be expected. These programs can, in turn, be used to bet on games and improve the ROI of the Swarm Intelligence system.

VI. CONCLUSIONS

Can the unique deliberation behaviors captured from live human participants during real-time swarm-based forecasts be analyzed to assess the likelihood of forecast accuracy? Furthermore, can such an assessment be used to identify the strongest forecasts among a set of forecasts (e.g. the best bets against the Vegas odds)? The results of this study suggest strongly this may be the case. As demonstrated across 25 consecutive weeks of forecasting the 2017-2018 NBA season, a machine learning program, configured to analyze the real-time behavioral characteristics of swarms of approximately 35 typical sports fans, was able to both select outcomes of the games more accurately and outperform the betting success of the swarm itself. In fact, although both swarm-based methods were able to outperform the Vegas betting market at predicting the outcome of select games each week, the machine learning program more than doubled the ROI of the unaided swarm's betting strategy and did so without training on any NBA data.

It's important to note that this study was limited by the availability of training and testing data: only one season of each of the three sports in this study was available for training, and only one sport was used for testing. Future work with more extensive historical datasets may enable even more accurate results. Additionally, the games covered in this study were not forecast probabilistically, due to the lack of suitable data to perfectly calibrate the Conviction Indexes to NBA. Future work aims to generate probabilistic forecasts. In addition, future work will investigate the success of behavioral swarm analysis in different settings, will strive to improve to optimize the CI for general and calibrated settings, and will refine the method in which bets are placed to allow for more sophisticated betting mechanisms (i.e. using the Kelly Criteria), as we believe there remains substantial room for improvement when optimizing a wagering strategy against Vegas Odds based on swarm-based predictive intelligence.

ACKNOWLEDGMENT

Thanks to Chris Hornbostel and David Baltaxe for their efforts in coordinating the weekly sessions that generated NBA basketball predictions. Also, thanks to Unanimous AI for the use of the swarm.ai platform for this ongoing work.

REFERENCES

[1] Rosenberg, Louis, "Human Swarms, a real-time method for collective intelligence." Proceedings of the European Conference on Artificial Life 2015, pp. 658-659

[2] Rosenberg, Louis. "Artificial Swarm Intelligence vs Human Experts," Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE.

[3] Rosenberg, Louis. Baltaxe, David and Pescetelli, Nicollo. "Crowds vs Swarms, a Comparison of Intelligence," IEEE 2016 Swarm/Human Blended Intelligence (SHBI), Cleveland, OH, 2016, pp. 1-4.

[4] Baltaxe, David, Rosenberg, Louis and Pescetelli, Nicollo. "Amplifying Prediction Accuracy using Human Swarms", Collective Intelligence 2017. New York, NY ; 2017.

[5] Halabi, Safwan., et. Al. "Radiology SWARM: Novel Crowdsourcing Tool for CheXNet Algorithm Validation", SiiM Conference on Machine Intelligence in Medical Imaging, 2018.

[6] Rosenberg, Louis and Willcox, Gregg. "Artificial Swarm Intelligence cs Vegas Betting Markets," 2018 11th International Conference on Developments in eSystems Engineering (DeSE), Cambridge, 2018, pp. 155-159.

[7] Galton F, (1907) Vox populi. Nature 75:7.

[8] Lorge I, Fox D, Davitz J, Brenner M (1958) A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. Psychol Bull 55:337–372.

[9] J. Lorenz, H. Rauhut, F. Schweitzer, and D. Helbing. How social influence undermines the wisdom of crowds. Proc Natl Acad Sci U S A. 2011 May 31;108(22):9020-5

[10] Seeley T.D, Buhrman S.C 2001 "Nest-site selection in honey bees: how well do swarms implement the 'best-of-N' decision rule?" Behav. Ecol. Sociobiol. 49, 416–427

[11] Marshall, James. Bogacz, Rafal. Dornhaus, Anna. Planqué, Robert. Kovacs, Tim. Franks, Nigel. "On optimal decision-making in brains and social insect colonies." Soc. Interface 2009.

[12] Seeley, Thomas D., et al. "Stop signals provide cross inhibition in collective decision-making by honeybee swarms." Science 335.6064 (2012): 108-111.

[13] Seeley, Thomas D. Honeybee Democracy. Princeton Univ. Press, 2010.

[14] Seeley, Thomas D., Visscher, P. Kirk. "Choosing a home: How the scouts in a honey bee swarm perceive the completion of their group decision making." Behavioural Ecology and Sociobiology 54 (5) 511-520.

[15] Rosenberg, Louis, Willcox, Gregg, Halabi, Safwan., Lungren, Matt, Baltaxe, David. and Lyons, Mimi. "Artificial Swarm Intelligence employed to Amplify Diagnostic Accuracy in Radiology," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, 2018 (Nov 2-4).

[16] Willcox, Gregg., Rosenberg, Louis., Askay, David., Metcalf, Lynn., Harris, Eric., Amplifying the Social Intelligence of Teams Through Human Swarming, 2018 IEEE International Conference on Artificial Intelligence for Industries. (ai4i 2018). Laguna Hills, CA.

[17] Rosenberg, Louis. Pescetelli, Nicollo, Willcox, Gregg., "Artificial Swarm Intelligence amplifies accuracy when predicting financial markets," 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York City, NY, 2017, pp. 58-62. doi: 10.1109/UEMCON.2017.8248984

[18] Rosenberg, Louis and Willcox, Gregg, "Artificial Swarms find Social Optima: (Late Breaking Report)," 2018 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), Boston, MA, 2018, pp. 174-178. doi: 10.1109/COGSIMA.2018.8423987