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) enables human groups to form real-time systems modeled on natural swarms. Prior studies have shown that by thinking together in "swarms," networked human groups can significantly amplify their collective intelligence and produce more accurate forecasts than traditional methods. The present study explores whether the real-time behavioral data collected during "swarming" can be used to further increase the accuracy of forecasts. To do this, a dense neural network was used to process the deliberation data collected during swarming and generate a Conviction Index (CI) that reflects the expected accuracy of each forecast. 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 213 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 estimate the strength of each forecast and wager only on forecasts of sufficient conviction. The CI-selected wagers yielded a 48% net ROI against Vegas Odds. This is a significant gain, equivalent to more than a 171% increase in ROI. The probability that the CI-selected wagers outperformed by chance was low ($p < 0.01$).**

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

I. Introduction

The technology of Artificial Swarm Intelligence (ASI) has been shown to amplify the predictive accuracy of networked human groups [1, 2]. Prior studies have shown that real-time "human swarms" can produce more accurate forecasts than traditional "Wisdom of Crowd" methods such as votes, polls, and surveys [3]. For example, a 2015 study tested the ability of human swarms to forecast the outcome of college football games. A swarm comprised of 75 amateur sports fans was tasked with predicting 10 college bowl games. As individuals, the participants averaged 50% accuracy when predicting outcomes against the spread. When thinking together in real-time swarms, those same participants achieved 70% accuracy against the spread [2]. Similar 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 were connected in swarms [4]. The human swarms also outperformed the BBC's machine-model known as "SAM," which achieved 64% accuracy over the same 50 games [11]. Human swarms have also been shown to outperform largescale betting markets, such as one 20-week study where human swarms predicted the outcome of 200 National Hockey League games. The swarms were shown to reduce the expected error rate in Vegas Odds by 61% on a subset of games [12].

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 213 games predicted.

The study is organized as follows: in Section II, we introduce human swarms as intelligent systems and discuss biological models of swarm-based decision-making. In Section III, a technology platform for real-time human swarming (swarm.ai) is introduced, and examples of swarms are provided. The method behind the study is described in Section IV, and the results of the study are analyzed in Section V.

II. Swarms As Intelligent Systems

The primary difference between "crowds" and "swarms" is that in crowd-based methods, individual participants provide their input in isolation (for statistical aggregation after the fact), while in swarm-based methods, groups "think together" as real-time systems governed by intelligence algorithms and converge on solutions in synchrony. The swarming process is generally modeled after biological systems such as schools of fish and swarms of bees. The present research uses Swarm AI technology from Unanimous A.I. Inc, which is modeled largely on honeybee swarms. This model was chosen for the current study because honeybee swarms are known to significantly amplify the accuracy of critical decisions by enabling members to form real-time systems – i.e. *"hive minds"* – that can solve problems as a unified and amplified intelligence.

The decision-making processes that govern the behavior of honeybee swarms have been studied since the 1950s and have been shown to be remarkably similar to the decision-making processes in neurological brains [5,6]. Both employ large populations of simple excitable units (i.e., bees and neurons) that work in parallel to integrate noisy evidence, weigh competing alternatives, and converge on decisions in synchrony. In both, outcomes are arrived at through a real-time competition among sub-populations of excitable units. When one sub-population exceeds a threshold level of support, the corresponding alternative is chosen. In honeybees, this enables the group to converge on optimal decisions, picking the best solution to complex problems (i.e. selecting a new home location) over 80% of the time [7,8,9].

The similarity between "brains" and "swarms" becomes even more apparent when comparing decision-making models that represent each. For example, the decision process in primate brains is often modeled as mutually inhibitory leaky integrators that aggregate incoming evidence from competing neural populations [10]. A common framework for primate decision is the Usher-McClelland model in Figure 1 below.

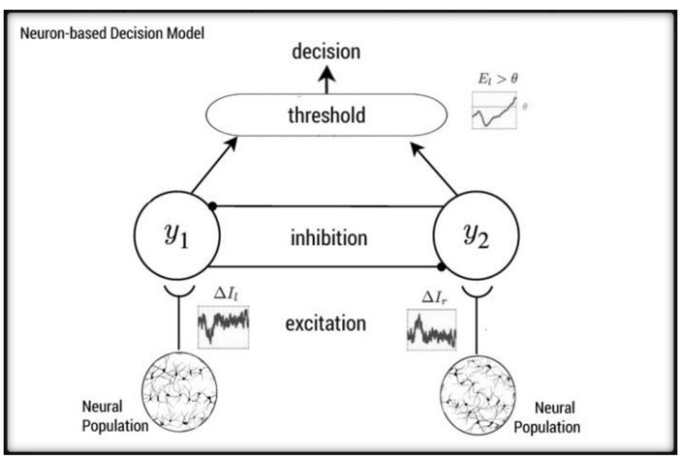

Fig. 1. Usher-McClelland model of neurological decision-making

This neurological decision model can be directly compared to swarm-based decision models, for example the honey-bee model represented in Figure 2 below. As shown, swarm-based decisions follow a very similar process, aggregating input from sub-populations of swarm members through mutual excitation and inhibition, until a threshold is exceeded.

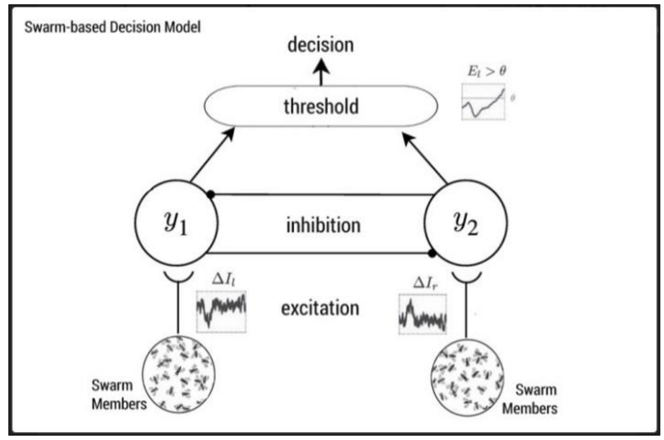

Fig. 2. Mutually inhibitory decision-making model in bee swarms

When viewed in this context, it becomes apparent that while brains are systems of neurons structured so intelligence emerges, swarms are systems of brains structured so amplified intelligence emerges. Thus, the objective of the current study is to connect human sports enthusiasts into synchronous systems that are structured so that an amplified intelligence emerges.

III. ENABLING "HUMAN SWARMS"

Unlike many other social species, humans have not evolved the natural ability to form closed-loop systems that enable real-time swarming. That's because we lack the subtle connections that other organisms use to establish high speed feedback-loops among members. Schooling fish detect vibrations in the water around them. Flocking birds detect subtle motions propagating through the population. Swarming bees use complex body vibrations called a "waggle dance." To enable real-time swarming among groups of networked humans, specialized user interfaces, intelligence algorithms, and networking paradigms are required to close the loop among all members.

To address this need, a technology called Swarm AI was developed to enable human groups to congregate online as real-time swarms, connecting synchronously from anywhere in the world [1]. Modeled after the decision-making process of honeybee swarms, the online system allows groups of distributed users to work in parallel to (a) integrate noisy evidence, (b) weigh competing alternatives, and (c) converge on decisions in synchrony, while also allowing all participants to perceive and react to the changing system in real-time, thereby closing a feedback loop around the full population of participants.

As shown in Figure 3, swarms answer questions by moving a graphical puck to select among a set of alternatives. Each participant provides input by manipulating a graphical magnet with a mouse or touchscreen. By positioning their magnet with respect to the moving puck, real-time participants express and 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. Because the full population of users can adjust their intent continuously in real-time, the swarm moves, not based on the input of any individual, but based on the dynamics of the full system. This enables a complex negotiation among all members at once, empowering the group to collectively explore the decision-space and converge on the most agreeable solution in synchrony.

It is important to note that participants do not only vary the direction of their intent, but also modulate the magnitude of their 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.

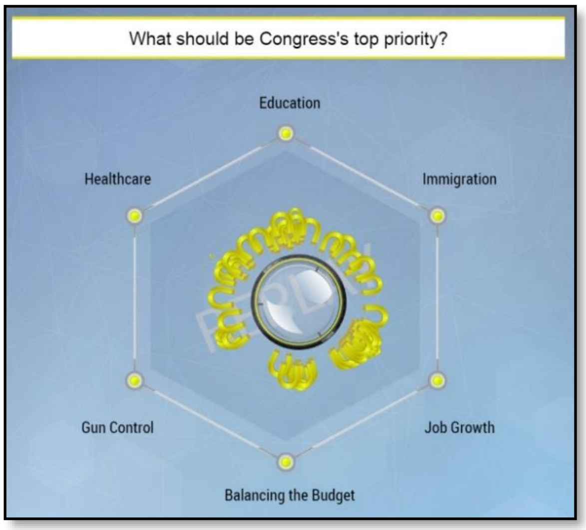

Fig. 3. A human swarm answering a question in real-time

Thus, like bees vibrating their bodies to express sentiment in a biological swarm, or neurons firing activation signals to express conviction levels within a biological neural-network, the participants in an artificial swarm must continuously update and express their changing preferences during the decision process, or lose their influence over the collective outcome. 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. This reveals a range of behavioral characteristics within the swarm population and weights their contributions accordingly, from entrenched participants to flexible participants to fickle participants.

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.

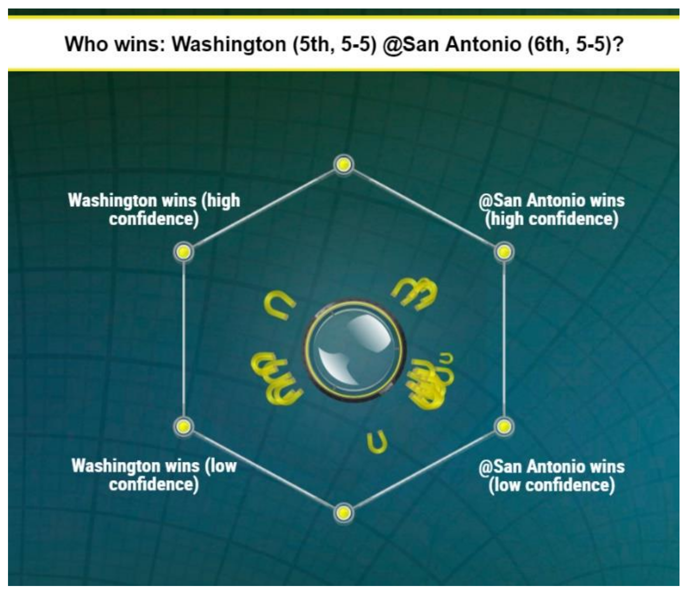

Fig. 4. Human Swarm in the process of forecasting NHL game

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 1,065 swarm predictions of NFL, MLB, and NHL games. The range of reasonable probabilities for each sport differs greatly (e.g. the distribution of Vegas Odds for MLB 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 a program was developed to place simulated bets on the outcomes of matches. To do so, an ROI Estimator was developed to use the CI and the Vegas Odds of the swarm's chosen outcome to predict the expected ROI of betting on the swarm's chosen outcome. 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 218 swarm NHL forecasts, each of which had an associated CI and Vegas Odds. When the expected ROI from this model is positive (>0%), the betting on the chosen outcome is expected to be profitable.

The experiment 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 on 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. Games were selected from the pool of NBA games each week using one of two strategies: (a) betting on the swarm's pick in all games, and (b) betting on the swarm's pick in all games with a positive expected ROI. Simulated bets were placed each week, and the simulated return on the investment was calculated given the outcome of the bet (win / loss) and the Vegas Odds, added to the gambling pool for the next week.

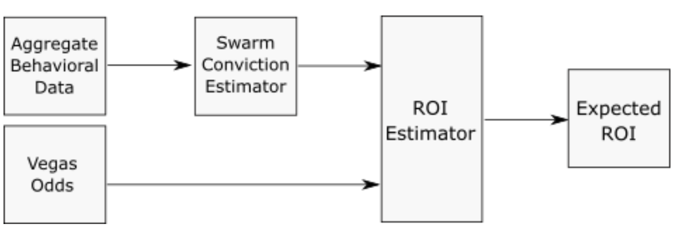

Fig. 5. System Diagram of ROI Estimation from Human Swarm Behavior and Vegas Odds

V. RESULTS

The results of the experiment are discussed in three 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 method is discussed and compared to the Vegas Odds. Finally, the accuracy and betting performance of each of the game selection methods is compared.

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%.

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.

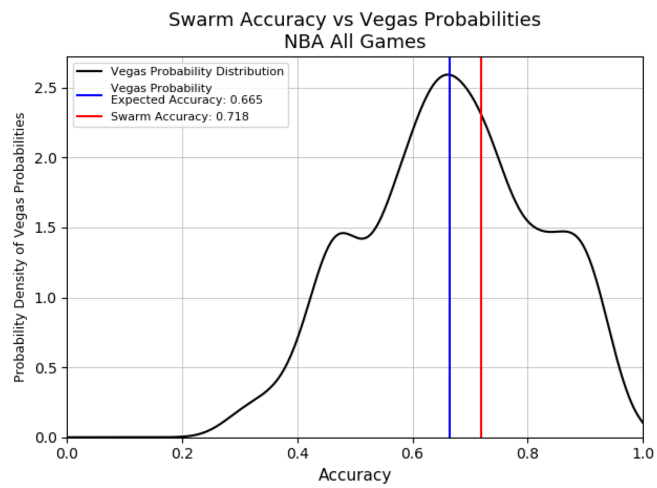

Fig 6. Vegas vs Swarm accuracy across all games predicted

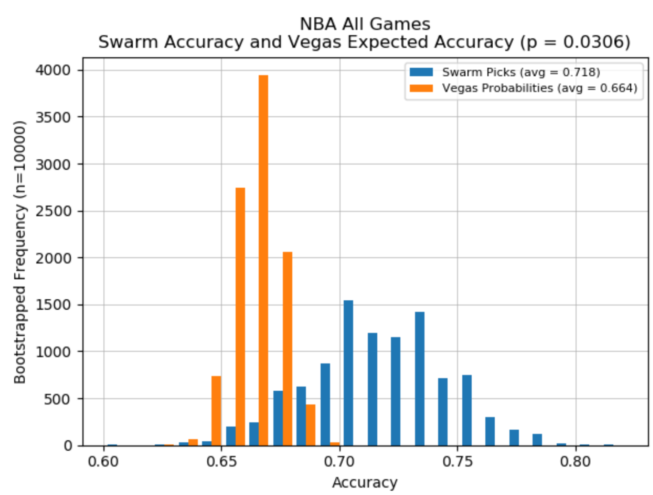

Fig 7. Bootstrapped average accuracy for Vegas vs Swarm

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 $117.71, or an ROI of 17.7%.

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 [-7.21%, 40.30%], indicating that we are not confident that betting on all swarm picks would return a positive ROI.

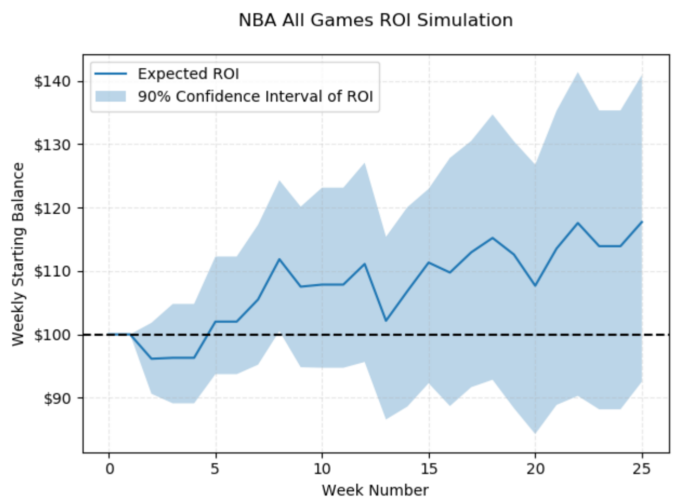

Fig 8. Cumulative simulated betting performance of fixed bets on all games predicted

So, while the swarm was significantly more accurate at predicting the outcomes of games than Vegas Odds, we cannot be confident that betting on the swarm outcomes would return a positive ROI. Two reasons could have contributed to this difference: (a) Vegas Odds includes a 2-5% "Bookie's Cut" in all outcomes to allow the sportsbooks to make money which applies to the ROI simulation, but is averaged out for the Accuracy analysis, or (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 the CI-selection method to the performance of Vegas Odds over the selected games. To do so, the Expected ROI of each of the 213 games was calculated using the Swarm Conviction Estimator and ROI Estimation machine learning programs. 134 games out of the 213 total games were expected to produce a net positive expected ROI and were selected to be bet on. This selection of games, referred to as the CI-selected games, was then compared against Vegas' expected win rate over the same games.

Figure 9 shows the distribution of Vegas Odds for the CI-selected games, and Vegas' expected win rate: 68.7%. The swarm, on the other hand, had a win rate of 79.1% across these same games. This is a significant improvement, equivalent to selecting games to bet on with a 10% higher accuracy than would be expected given the Vegas Odds. To examine the significance of this result, the average accuracy of each system over the full season was bootstrapped 10,000 times.

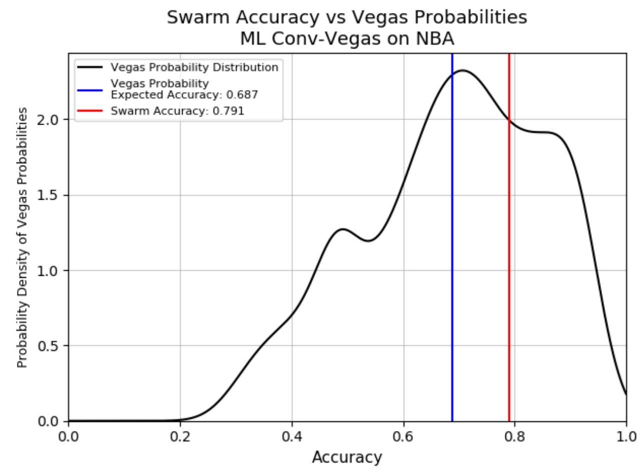

Fig 9. Vegas vs Swarm accuracy for CI-selected games

The average accuracies for each trial are shown in figure 10. We find that the probability that the swarm outperformed Vegas' expectations by chance was low (p=.0017), indicating that we can be confident that the CI-selected swarm forecasts had a higher accuracy than Vegas Odds expected.

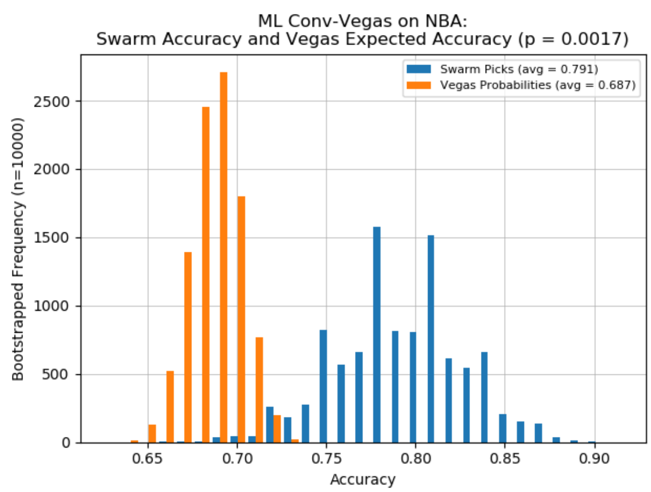

Fig 10. Bootstrapped average accuracy of Vegas odds and swarm picks over CI-selected games

In addition, a betting simulation was run in which 15% of the current bankroll was distributed evenly and bet on each of the CI-selected games that week. The performance of this model when betting against Vegas (including the impact of the Bookie's cut) is seen in figure 11. Starting with $100 and investing each week according to this strategy, the net balance after 25 weeks was $148, or an ROI of 48%. 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, the ROI for each game that week was predicted using the ROI Estimator program, and then the games with positive ROI predictions were selected to bet on. We find that the 90% confidence interval over the ROI of this betting strategy is [7.69%, 96.89%], indicating that the probability that this selection of games produced a positive ROI by chance is less than 5%.

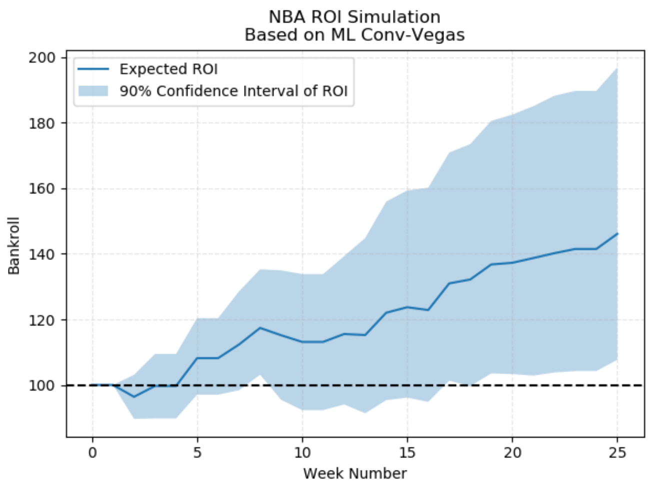

Fig 11. Bootstrapped average accuracy of Vegas odds and swarm picks over CI-selected games

So, we can be confident that by identifying behavioral characteristics within a swarm that indicate the likelihood of the swarm's pick and estimating the ROI of betting on the game against Vegas Odds, we selected outcomes that are both more accurate than Vegas Odds expected and that produced a positive ROI when bet on. This is a notable improvement over the original all-game method, equivalent to a 171% increase in ROI over the course of the 25 weeks, but can we be confident that the CI-selection method outperforms the original method?

To investigate whether this amplification of betting success is significant, the accuracy of the selection methods over 10,000 bootstrapped seasons is compared in figure 12. We find that the CI-selected games are significantly more accurate ($p < 0.001$) than the full set of swarm picks.

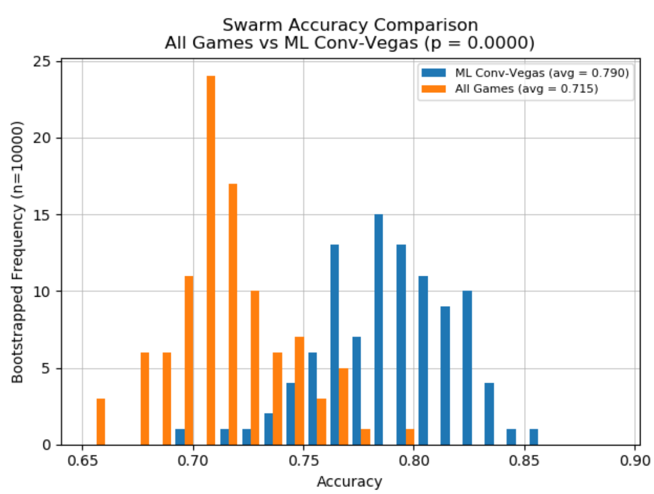

Fig 12. Bootstrapped average accuracy of all swarm picks vs CI-selected picks

Further, it is possible to statistically compare the ROI of the betting methods through a similar bootstrapping simulation over the course of 10,000 bootstrapped 25-week seasons. Below in Figure 13 we compare the season-end ROI of the two selection methods and find that the probability that the CI-selected picks returned a lower ROI than the all swarm picks method due to chance was low (p=0.15), but not significant.

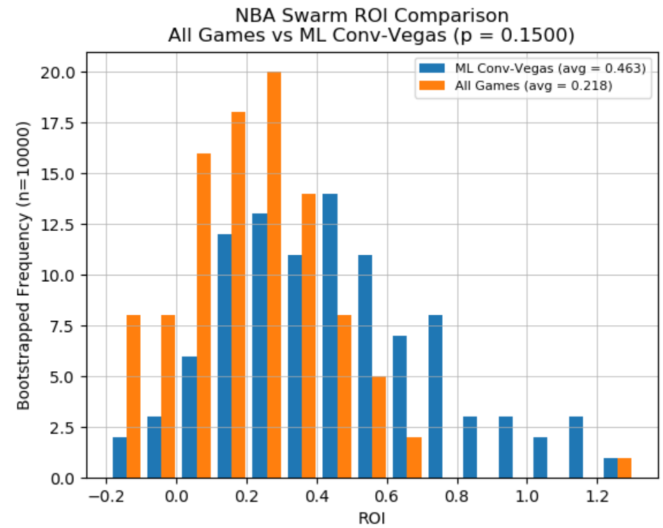

Fig 13. Bootstrapped ROI over whole NBA season of CI-selected games vs all swarm picks.

VI. CONCLUSIONS

Can the unique deliberation behavior captured during real-time swarm forecasts be analyzed to assess the likelihood of forecast accuracy? And can that 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 this is the case. As demonstrated across 25 consecutive weeks of the 2017-2018 NBA season, a machine learning program analyzing the 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 systems were able to beat Vegas – a world-class oddsmaker – at predicting the outcome of select games each week, the machine learning program increased the ROI of the swarm's betting strategy by over 90%.

This study was limited by the availability of training and testing data: only one sport, and one season of that sport, was rigorously studied. Additionally, the games covered in this study were not forecasted probabilistically, due to the lack of suitable data to calibrate the Conviction Indexes to NBA. 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 there appears to be significant 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 swarms that generated predictions. Also, thanks to Unanimous A.I. for the use of the swarm.ai platform for this ongoing work.

REFERENCES

[1] Rosenberg, L.B., “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. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[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 N. Pescetelli, “Amplifying Prediction Accuracy using Human Swarms”, Collective Intelligence 2017. New York, NY ; 2017.

[5] 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

[6] 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.

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

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

[9] 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.

[10] Usher, M. McClelland J.L 2001 “The time course of perceptual choice: the leaky, competing accumulator model.” Psychol. Rev. 108, 550–592

[11] McHale, Ian. Sports Analytics Machine (SAM) as reported by BBC: http://blogs.salford.ac.uk/business-school/sports-analytics-machine/

[12] 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.