

Artificial Prediction Markets as a tool for Syndromic Surveillance

Fatemeh Jahedpari¹, Julian Padget¹, Marina De Vos¹, Benjamin Hirsch²

¹ Department of Computer Science, University of Bath, UK
{f.jahedpari, j.a.padget, m.d.vos}@bath.ac.uk

² EBTIC, Khalifa University, United Arab Emirates
Benjamin.Hirsch@kustar.ac.ae

Abstract. A range of data sources across the internet, such as google search terms, twitter topics and Facebook messages, amongst others, can be viewed as kinds of sensors from which information might be extractable about trends in the expression of matters of concern to people. We focus on the problem of how to identify emerging trends after the original textual data has been processed into a quantitative form suitable for the application of machine learning techniques. We present some preliminary ideas, including an agent-based implementation and some early results, about the application of artificial prediction markets to such data, taking the specific domain of syndromic surveillance (early stage recognition of epidemics) as an example, using publicly available data sets.

1 Introduction

This paper outlines some early stage research into the application of prediction markets to syndromic surveillance. Prediction markets are seen as a mechanism to forecast the outcome of future events by aggregating public opinion, in which market participants trade so-called securities that represent different probabilities about the (expected) outcome of a scenario. We describe prediction markets in more detail in section 2 and compare them with alternative approaches in section 5.

Syndromic surveillance monitors population health indicators which are apparent before confirmatory diagnostic tests become available, in order to predict a disease outbreak within a society at the earliest possible moment, with the aim of protecting community health. Clearly, the earlier a health threat within a population is detected, the lower the morbidity and the higher the number of lives that may be saved. Syndromic surveillance data sources include, but are not limited to, coding of diagnoses at admission or discharge emergency department, chief complaints, medical encounter pre-diagnostic data, absentee rates at schools and workplaces, over-the-counter pharmacy sales, Internet and open source information such as people post in social media. Each of these types of data can generate a signal during a disease development. Therefore, given

the vast amount of these data sources, a proper mechanism is necessitated to integrate them as soon as they become available.

In this research, we focus on developing a novel syndromic surveillance technique by integrating different data sources inspired by the crowd-sourcing behaviour of prediction markets. To achieve our goal, we train a multiagent system in an artificial prediction market in a semi-supervised manner.

2 Prediction Markets

Prediction markets have been used to forecast accurately the outcome of political contests, sporting events, and economic outcomes [19]. In this research, we use an artificial prediction market as a mechanism to integrate several syndromic surveillance data sources to predict a level of disease activity within a population on a specific date. This section briefly explains the preliminaries of prediction markets.

The prediction market, also known as an information market, originated at the Iowa Electronic Marketplace (IEM) in 1988 as a means to bet on presidential elections. A prediction market aims to utilise the aggregated wisdom of the crowd in order to predict the outcome of a future event [16]. In these markets, traders' behaviour has the effect of externalising their private information and beliefs about the possible outcomes, and can hence be used to forecast an event accurately [12]. Prediction markets are increasingly being considered as approaches for collecting, summarising and aggregating dispersed information by governments and corporations [10].

In prediction markets, traders bet on the outcome of future events by trading securities. A security is a financial instrument, like a financial stock, that pays a profit (or makes a loss) based on the outcome of the event. Each outcome of an event has a security associated with it. Traders can buy or sell any number of securities before the expiry time of the security. A security expires when the outcome of the event is realised. To illustrate a simple case, a prediction market can be used to predict "if candidate 'X' will win the election" by offering two securities of 'Yes' and 'No'. Assuming the market finally ends with candidate 'X' winning the election, all traders will receive \$1 payoff for each 'Yes' security they own and \$0 for their 'No' securities, losing the money they spent on buying them.

The aggregated monetary bets made by market traders dynamically determine the price of each security before the market ends. The market price of a security represents the price at which the security can be bought or sold. Also, it can be interpreted as representing the probability of that outcome occurring by fusing the beliefs of all the market participants. Arguably, the price that an agent would pay to buy a security indicates how confident s/he is in the outcome of the event. For example, if a trader believes that the chance of candidate 'X' winning is 80%, s/he then would be willing to buy a 'Yes' security at any price up to \$0.80.

A prediction market is run by a market-maker who is the company or individual that interacts with traders to buy and sell securities. The market-maker determines the market price using a market trading protocol. The logarithmic market scoring rule (LMSR) designed by Hanson [9] is an automated market maker. Using LMSR, the price and cost of a security is calculated as follows:

$$C(q_i) = b * \log\left(\sum_{i=1}^m e^{q_i/b}\right)$$

and

$$P(q_i) = \frac{\exp(q_i/b)}{\sum_{j=1}^m \exp(q_j/b)}$$

respectively, where m is number of securities that market offers, each for one possible outcome and $q_i \in (q_1, q_2, \dots, q_m)$ represents the number of units of security i held by market traders. The larger the value of b , the more money the market maker can lose. It also means that traders can purchase additional quantities of a security without causing significant price swings. Note that the price of a security only applies for buying an infinitesimal number of shares and the price of the security immediately changes as soon as a traders start trading. In order to calculate the cost of a trading X securities, the market makers must calculate $C(q + X) - C(q)$.

3 Agent-Based Simulation Architecture

In order to explore empirically the application of artificial prediction markets to syndromic surveillance, we have developed an agent-based simulation, which we now describe, followed by some preliminary results in section 4.

3.1 Agents and Strategies

Our model integrates data and beliefs of different data streams by simulating an artificial prediction market to predict the outcome of an event, which in this case is the disease activity level on a specific date. Each data stream includes the quantitative value of a particular disease activity level for a specific place for different period of time. Each agent is responsible for one data stream and trades securities in various prediction markets based on its capital and belief about the disease activity level of the market event date. Trading agents will in due course (see below) learn from each market based on the revenue they receive and the losses they make when the market closes. Consequently, they can update their strategy, beliefs and confidence for the future markets.

The system has a market-maker that uses a scoring rule to calculate the market price for each security in the market, and a data distributor that provide agents with the data stream for which they are responsible, and trading agents. The simulation mechanism,

specified in Algorithm 1, is as follows. At the beginning of the simulation, all the trading agents are awarded an equal amount of initial capital. For each training example, a prediction market (let us say prediction market for week T) is established. At this time, the Data Distributor Agent provides available data to the trading agents, according to their role. Then, the trading agents participate in the market according to their available capital, beliefs, and trading strategies. Agents can trade any number of securities before the deadline for the closing of the market. Once the market deadline is reached, the market-maker reveals the winning security and rewards the winning security holders with \$1 for each winning security they own. These revenues are added to their capital. However, the agents who own losing securities, lose capital equal to the amount spent on purchasing them.

During the simulation, agents with superior data, strategy and analysis algorithms are likely to accumulate greater capital and hence affect market prices and eventually the outcome. In other words, important – by these metrics – agents are identifiable and have greater influence in predicting the outcome of the event. This increased influence of the more successful agents should increase the accuracy of the system overall: the agents are not in competition *per se*, so we do not care which agents are better, but we do want the better ones to have more effect on the prediction mechanism.

The first agent strategy is based on zero intelligence [7], and so has no scope for learning: buy-sell and security choices are random subject to the constraint of not trading at a loss. For the second strategy, we add a basic learning mechanism, following the design of zero intelligence plus [3], in which agents update their trading strategy and beliefs based on the reward they received from the market in order to improve their reward in future markets. This is achieved by incorporating a simple machine learning mechanism (Widrow-Hoff) to adapt their individual behaviour to the market trend.

3.2 Market Instantiation

Different data sources have different timeliness in detecting a disease outbreak. For example, some data sources such as social media data can signal a disease activity level perhaps two weeks earlier than physician data. Therefore, for the system to be capable of forecasting the outcome at the earliest possible moment and not wait for all the agents' data to arrive to start prediction, we will run multiple concurrent markets for consecutive prediction weeks. For example, if the simulation week number is 1, then 4 further markets for weeks 2 to 5 will also be open. Once the deadline of the first market (week 2) is reached, then that market (week 2) closes and another market after the last market is opened (week 6). Consequently, the agents who have data for those further markets can start trading earlier in those markets and take advantage of cheaper prices, which will lead to the updating of market prices as early as possible. In addition, all data with different timeliness ranging from 4 weeks to one day before the event date will be incorporated in each market and at the same time agents can use their knowledge achieved from the previous market when predicting the outcome of a given market.

Algorithm 1: Agent-Based Simulation Architecture Algorithm

```
1 Give start up capital to each agent
2 Simulation-Current-Week C;
3 Market-Date T;
4 for  $T \leftarrow 1$  to end do
5     Data Distributor disseminate data, which are accessible by week C, to each agent
        according to agent expertise;
6     Start Prediction Market for Week-T;
7     while Market deadline is not reached do
8         Wait();
9         In here, agents will decide the level of disease activity in week T and trade
            security according to their belief and strategy;
10    end
11    End Prediction Market;
12    Reveal the winning security (Based on the label of training examples);
13    Each agent new capital  $\leftarrow$  previous capital + revenue gained in this market – amount
        spent for purchasing securities;
14    Now, according to the utility received in this market, agents should update their
        trading strategy and beliefs ;
15     $T=T+1$ ;
16     $C=C+1$ ;
17 end
```

For the sake of simplicity, Algorithm 1 considers just the one market, as does our current implementation.

4 Preliminary Results

The configuration of the controlling parameters of our system need thorough investigation through running a large number of simulation experiments. These settings include but are not limited to constraints on number of market participants, required time for each market, initial capital for each agent, type of monitored diseases, constraints and requirements for agents to trade, and the minimum required number of training examples. We have only just begun to explore this parameter space.

As discussed in Section 1, there is a vast number of syndromic surveillance data sources. Much research has been done to compare these data sources with the actual value of disease activity level for a specific disease in a particular place. For example, Culotta [5]

stated that he could track influenza rates in the United States using Twitter messages with 95% correlation. Corley [4] could track flu rates in the United States with a correlation of 76% by examining the proportion of blogs containing the two keywords of “influenza” and “flu”. Google Flu Trend [6] can predict ’flu activity level with a 97% correlation by analysing queries sent to the Google search engine.

In the first two batches of experiments, we have tried two well-known trading strategies for trading agents: Zero Intelligence (ZI) [7] and Zero Intelligence Plus (ZIP) [3], the former to provide a baseline behaviour and the latter to investigate the effect of a simple learning mechanism on the trading decision. We now discuss each of these in more detail.

4.1 Preliminary Results with ZI

In the first set of experiments, agents use a Zero Intelligence strategy (adapted from [7]) when trading securities in the market. In this model, agents consider a limit price, according to their data, for each security of the market. In each day of the market, they choose one security randomly and purchase a random quantity of that security, if its limit price is higher than its market price or they sell a random quantity of that security if its limit price is less than its market price. In both situations, the agents considers the maximum number of securities that can be traded, based on their available capital and the securities they own. LMSR, described in Section 2, is used as market scoring rule, as it provides infinite liquidity [8] and does not suffer in thin markets where the number of traders are small.

Each agent in the experiment is awarded \$10 at start up and one market is established for each training example. The winning security is chosen based on the United State influenza-like illnesses rate from 30 September 2002 to 01 September 2003 ³. Each market offers eight securities, corresponding to one security for each standard deviation from the mean, covering from -4 to $+4$ standard deviations.

Figure 1 shows how extending the period of a prediction market can help agents to predict the outcome of the event better. It demonstrates that accuracy goes up and the mis-classification rate falls as the duration of the market increases. As can be seen from the figure, accuracy increased from 82% for 10-day-long markets to 93% for 90-day-long markets, after which it is almost flat. From this, we conclude that 110 days seems a sufficient period for each market and hence our subsequent experiments at this stage use this market length. Longer duration markets provide agents with sufficient time to trade enough numbers of the desired securities and approach more closely the equilibrium price.

The purpose of the experiments reported in this paper is primarily to establish confidence in the behaviour of the simulation, by providing tailored data feeds with known

³ <http://www.cdc.gov/flu/weekly/fluviewinteractive.htm>

properties and then observing whether the agents achieve their expected level of performance, given that data and their (known) strategy. These experiments have a population of 20 agents, each receiving data from data streams with a specified correlation with the United States influenza-like illnesses rate. The agent names in the following figures represents the type of data the agent is receiving. For instance, a95 denotes an agent that receives data with a 95% correlation with the United States influenza-like illnesses rate.

Since all the agents in these experiments are essentially identical in terms of strategy, the difference in their data sources should lead to them obtaining different amounts of revenue in each market. Figure 2 shows their revenue when each market ends. As the Figure shows, a100 agent, which has complete information about all the events of the experiment, earns high revenue in all markets with the exception of 31/07/2003. In this case, it was not making the wrong choice – it cannot – but was trying to purchase a large number of securities and since the price of a security increases as a result of its purchase, the agent (as explained in Section 2), did not have sufficient capital to complete the deal. In other words, for agents without perfect information, the enforced random choice of security to trade (the ZI strategy) means the agent cannot select the most appropriate one, but rather the one that chance dictates and hence it makes a loss.

Figure 3 shows the capital held by each agent at the end of each market in one experiment and Figure 4 shows the average capital of agents over 50 runs. The main observation from these figures is that, as expected, agents with higher quality data are able to achieve higher levels of revenue. As can be seen from these figures, agent a100 accumulates more capital than other agents even those with high quality data such as a99. The reason for this is that a100 never makes a mistake while the other agents do and as soon as one agent predicts an outcome, it dedicates most of its capital to purchase the corresponding security. Therefore, once an agent predicts a wrong outcome, it loses all its capital, while agent a100 keeps earning revenue in each market and accumulates more capital and hence invests more on upcoming markets and earns more revenue again. Also, agent a100 causes the price of the correct security to increase rapidly as it purchase a large quantity of it and, therefore, makes it difficult for other agents to buy significant quantities of that security due its high price.

Clearly, more comprehensive experimentation is necessary, backed up with appropriate statistical confidence tests. In this section, we have only used the most basic of strategies and one that has known flaws [3]. However, it provides both a useful baseline performance, as well as a setting in which initial hypotheses about the effectiveness of the prediction market model can be validated (such as the agent with 100% correlated data dominating the market and all others losing all their investments).

4.2 Preliminary Results with ZIP

In the second set of experiments, we have changed the agent strategy from ZI to ZIP, by which the agent uses data about trends in the market in order to adjust their behaviour to

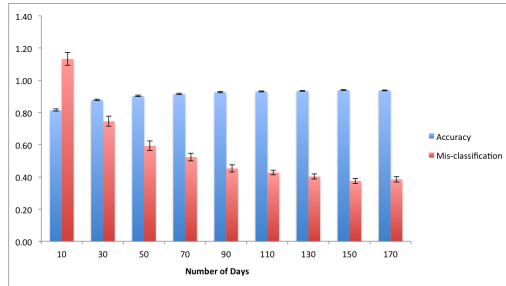


Fig. 1. ZI: Comparing accuracy (s.d.: 0.017–0.047) and mis-classification (s.d.: 0.101–0.289) on the y-axis vs. duration of the prediction market (x-axis). Each data point is the average of 50 experiments with same parameter settings

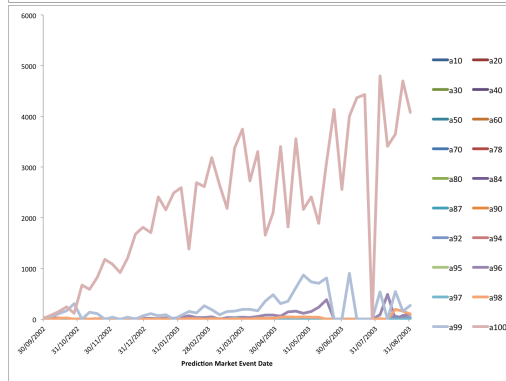


Fig. 2. ZI: Comparing revenue of agents (y-axis) at the end of each market (x-axis) for an example run chosen at random.

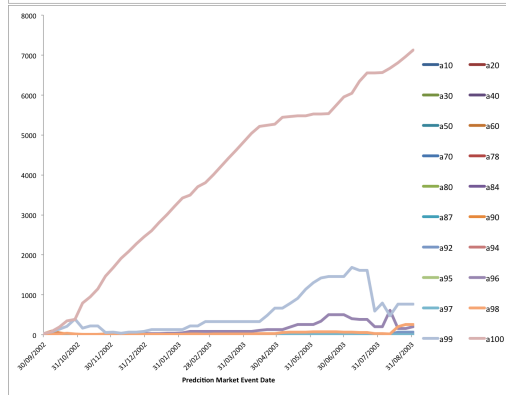


Fig. 3. ZI: Comparing capital of agents (y-axis) at the end of each market (x-axis), for an example run chosen at random.

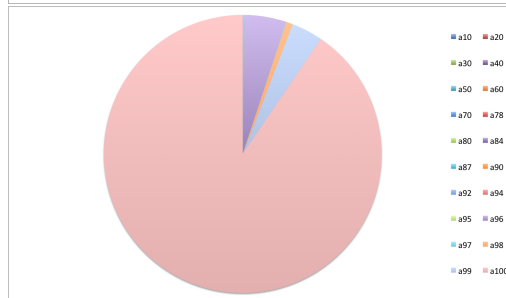


Fig. 4. ZI: Comparing capital of agents (y-axis) at the end of the experiment (averaged over 50 runs).

be less random (as in ZI) and more in line with the market valuation of a given security. Although ZIP and its variants have been shown to be effective strategies in terms of profit making, there are two reasons why this approach is likely to be ineffective in the context of prediction markets:

1. The ZIP strategy depends upon both buyer and seller employing this strategy, but in the prediction market, the two parties are the buyer and the market-maker, of which the latter has no interest in profit and which has no strategy as such. Consequently, only one party in the market is ‘learning’. This still has a positive effect as discussed below, but starts to underline the difference between trading markets (with bilateral strategies) and prediction markets (with unilateral strategies).
2. The ZIP strategy aims at trading for profit regardless of the (financial) instrument being traded, leading to the establishment of an equilibrium price, whereas the point of a prediction market is to choose the *right* instrument, rather than the currently most profitable.

The experiments are run with the same data as for ZI. Thus Figures 6 and 7 show the results from a randomly chosen 110 day market, as was done for ZI. It is notable that ZIP achieved higher accuracy – nearly 97% – with market durations of > 30 days, than in the ZI experiment. However, as before a100 agent dominates the market.

The story in terms of revenue (Figure 6) is much the same as for ZI, although a99 stops making a profit much sooner with ZIP. As can be seen from the figure, the agent obtains revenue until 04/11/2002, which is the first time that it makes a mistake and loses the majority of its capital. Consequently, the agent has little remaining capital, but continues earning money for the following two markets, but its second mistake (on 25/11/2002) bankrupts it, after which it cannot invest further. This scenario applies to all other agents and causes a100 to dominate the market, as it never makes any mistake.

5 Related Work

Many syndromic surveillance systems exist worldwide, each designed for a specific country, region or state [11, 20, 15, 13, 17]. We refer to them as traditional, since they do not utilise internet based data. While these systems can detect an outbreak with high accuracy, they suffer from slow response times. For example, the Centers for Disease Control and Prevention (CDC) publishes USA national and regional data typically with a 1–2 week reporting lag. It monitors over 3,000 health providers nationwide to report the proportion of patients seen that exhibit influenza-like illnesses (ILI) ⁴ [5, 6].

On the other hand, modern syndromic surveillance systems appeal to internet based data such as search engine queries, health news, and peoples’ posts on social networks

⁴ <http://www.cdc.gov/flu/weekly/fluactivity.htm>

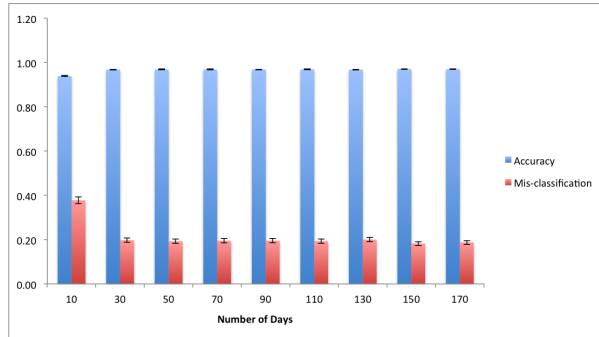


Fig. 5. ZIP: Comparing accuracy (s.d.: 0.010–0.018) and mis-classification (s.d.: 0.062–0.109) on the y-axis vs. duration of the prediction market (x-axis). Each data point is the average of 50 experiments with same parameter settings

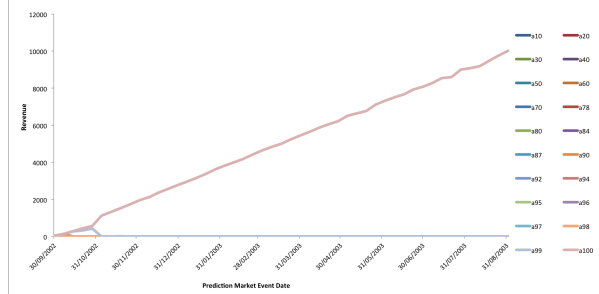


Fig. 6. ZIP: Comparing revenue of agents (y-axis) at the end of each market (x-axis) for an example run chosen at random.

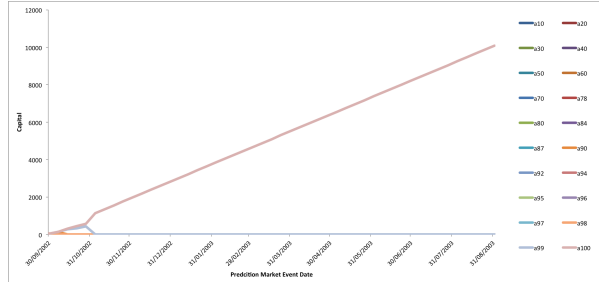


Fig. 7. ZIP: Comparing capital of agents (y-axis) at the end of each market (x-axis) for an example run chosen at random.

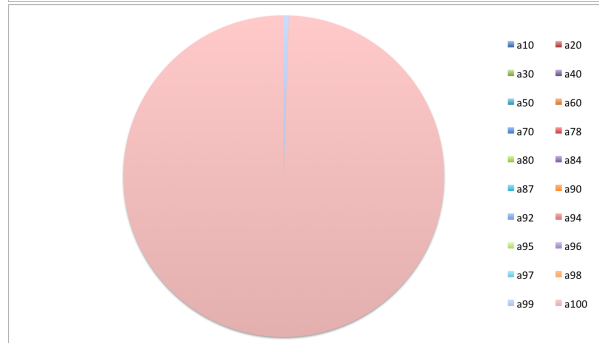


Fig. 8. ZIP: Comparing capital of agents (y-axis) at the end of the experiment (averaged over 50 runs).

to predict an outbreak earlier [18, 2, 4], albeit with necessarily lower precision. While some of them claim that they could achieve high accuracy, they are vulnerable to false alarms [6, 1] due to their dependence on a single data stream and disregarding the benefits from fusing different data sources. Ginsberg et al [6] state, regarding Google Flu Trends, that “Despite strong historical correlations, our system remains susceptible to false alerts caused by a sudden increase in ILI-related queries. An unusual event, such as a drug recall for a popular cold or flu remedy, could cause such a false alert”.

To the best of our knowledge, there is no system that fuses both traditional and internet based data sources. This could be due to the different timescales that these data sources have and the consequent issues of appropriate synchronisation. Prediction markets can overcome this problem as traders can trade securities as soon as they receive new information and impact the price and consequent probability of an event outcome. It is interesting to note that Polgreen et al [14] report on the use a prediction market with human health care expert participants to forecast infectious disease activity 2-4 weeks in advance.

Moreover, internet based system are only suitable for places where sufficient source data is available. For example, twitter-based systems cannot have a high accuracy on places where using twitter is not very common, if even accessible. In addition, even if sufficient data is available, system accuracy cannot be guaranteed worldwide since peoples’ behaviour changes from place to place, reflecting differing (digital) cultures. For example, people in a particular city may seek a physician as soon as they encounter the symptoms of a disease and do not trust online information, while people in another city may defer visiting a doctor and seek out online information in order to cure themselves at the early stages of their sickness. Furthermore, peoples’ behaviour may change over time. For example, a particular social media may become less popular and cede its role to newer technology over the time.

6 Discussion

Since we are in an early stage of this research, a substantial part of the work is to come. We have numerous ideas that have yet to be implemented, including: (i) the learning capability of agents, (ii) consideration of the confidence of agents, (iii) of the different timeliness of data streams, and (iv) the effect of a heterogeneous population of agents with different trading strategy and risk prediction model –among other characteristics. The very preliminary results we have meet our broad expectations for the behaviour of prediction markets, but it is too early to say whether they can be a general-purpose tool with useful levels of precision and recall across a range of domains. The ZI strategy, being essentially random under the constraint of not making a loss, establishes a useful performance baseline, as well as a framework against which to validate the basic system hypotheses. The ZIP strategy, while appropriate for bilateral markets seeking to establish equilibrium prices, is inappropriate – at least, as conventionally formulated –

for prediction markets, although the dampening effect of the learning mechanism does lead to higher prediction rates and smoother overall behaviour.

We welcome feedback on the appropriateness of the approach and the above directions for development as well as alternative mechanisms that might be incorporated in the prediction market setting.

References

1. Harshavardhan Achrekar, Avinash Gandhe, Ross Lazarus, Ssu-Hsin Yu, and Benyuan Liu. Predicting flu trends using twitter data. In *Computer Communications Workshops (INFOCOM WKSHPs), 2011 IEEE Conference on*, pages 702–707. IEEE, 2011.
2. Herman Anthony Carneiro and Eleftherios Mylonakis. Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases*, 49(10):1557–1564, 2009.
3. Dave Cliff and Janet Bruten. Minimal-intelligence agents for bargaining behaviours in market-based environments. Technical Report HPL-97-91, Hewlett-Packard Laboratories, 1997. Available via <http://www.hpl.hp.com/techreports/97/HPL-97-91.html>, retrieved 20071115.
4. Courtney D Corley, Diane J Cook, Armin R Mikler, and Karan P Singh. Text and structural data mining of influenza mentions in web and social media. *International journal of environmental research and public health*, 7(2):596–615, 2010.
5. Aron Culotta. Detecting influenza outbreaks by analyzing twitter messages. *arXiv preprint arXiv:1007.4748*, 2010.
6. Jeremy Ginsberg, Matthew H Mohebbi, Rajan S Patel, Lynnette Brammer, Mark S Smolinski, and Larry Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, 2008.
7. Dhananjay K Gode and Shyam Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of political economy*, pages 119–137, 1993.
8. Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.
9. Robin Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets*, 1:3–15, 2007.
10. Robin Hanson, Ryan Oprea, and David Porter. Information aggregation and manipulation in an experimental market. *Journal of Economic Behavior & Organization*, 60(4):449–459, 2006.
11. Michael D Lewis, Julie A Pavlin, Jay L Mansfield, Sheilah OBrien, Louis G Boomsma, Yevgeniy Elbert, and Patrick W Kelley. Disease outbreak detection system using syndromic data in the greater Washington DC area. *American journal of preventive medicine*, 23(3):180–186, 2002.
12. Evdokia Nikolova and Rahul Sami. A strategic model for information markets. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 316–325. ACM, 2007.
13. Yasushi Ohkusa, M Shigematsu, K Taniguchi, and N Okabe. Experimental surveillance using data on sales of over-the-counter medications-japan, November 2003–April 2004. *MMWR Morb Mortal Wkly Rep*, 54:47–52, 2005.
14. Philip M. Polgreen, Forrest D. Nelson, and George R. Neumann. Use of prediction markets to forecast infectious disease activity. *Clinical Infectious Diseases*, 44(2):pp. 272–279, 2007.

15. Emergency department syndromic surveillance system: England & Northern Ireland, September 2013.
16. Russ Ray. Prediction markets and the financial “wisdom of crowds”. *Journal of Behavioral Finance*, 7(1):2–4, 2006.
17. Communicable and respiratory disease report for england & wales, October 2013. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317139997178, retrieved 20131216.
18. Alessio Signorini, Alberto Maria Segre, and Philip M. Polgreen. The use of twitter to track levels of disease activity and public concern in the U.S. during the Influenza A H1N1 Pandemic. *PLoS ONE*, 6:e19467, 05 2011.
19. Erik Snowberg, Justin Wolfers, and Eric Zitzewitz. Prediction markets for economic forecasting. Working Paper 18222, National Bureau of Economic Research, July 2012. Available from: <http://www.nber.org/papers/w18222>.
20. Fu-Chiang Tsui, Jeremy U Espino, Virginia M Dato, Per H Gesteland, Judith Hutman, and Michael M Wagner. Technical description of rods: a real-time public health surveillance system. *Journal of the American Medical Informatics Association*, 10(5):399–408, 2003.