Modeling Prediction Markets with Dynamic Bayesian Networks

Cole Winstanley, Ethan Shen

Motivation

Prediction markets are exchange-traded markets created for the purpose of trading the outcome of events, and the market prices indicate what the crowd thinks the probability of an event is [1]. They are a **distributed**, **scalable**, and **self-incentivized** means of aggregating information about a future event [2].

Problem Statement

- Can we model prediction markets as a Dynamic Bayesian Network (DBN)?
- Can we use this DBN to model robustness of prediction markets to different types of agents?
- Can we infer future market prices using this DBN?

Data

- Polls and prediction market prices for 2 years before 2012 presidential election

Models

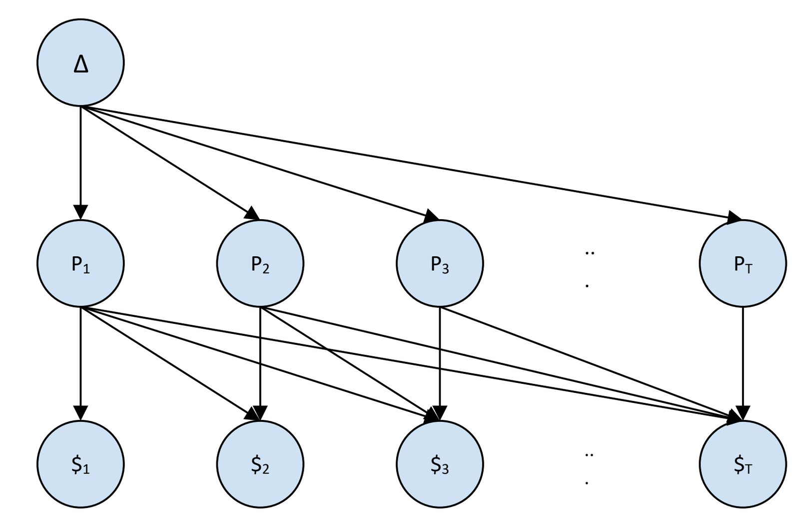


Figure 1: Bayesian Updating (BU) model.

Prices are simply Bayesian updates on the emissions from Δ, discretized using the CDF of the normal distribution up to a tied vote share.

$$\Delta \sim \mathcal{N}(\delta, 1/h)$$
$$P_t = \Delta + e_t \text{ where } e_t \sim \mathcal{N}(0, 1/\sqrt{N_t})$$

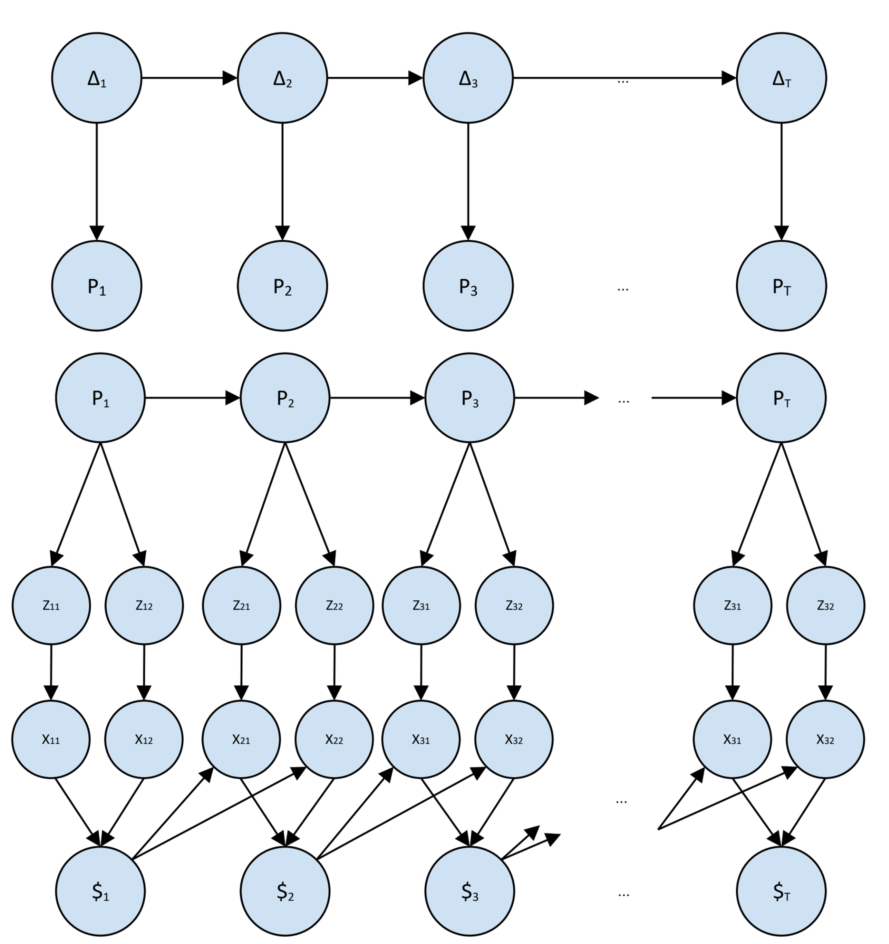


Figure 2: Dynamic Bayesian Network (DBN) model.

- Our model is based on the BU model but with richer representation for time and a model for agents' beliefs.
- We implicitly model the sequence of polls as a HMM over a changing political climate Δ.
- Agents receive noisy versions of the poll data and use this information, along with past prices, to affect future market prices

$$\mathbf{Pr}[Z_{ti} = A|P_t] = (1-\gamma)\,\mathbf{Pr}[P_t = A] + \gamma\,\mathbf{Pr}[P_t \neq A]$$
$$\begin{aligned}\mathbf{Pr}[X_{ti} = A|Z_{ti}, P_{t-1}] &= \mathbf{Pr}[P_{t-1} = X_{ti} = A] \\ &+ (1-e)\,\mathbf{Pr}[P_{t-1} = A \neq Z_{ti}] \\ &+ (e)\,\mathbf{Pr}[Z_{ti} = A \neq P_{t-1}]\end{aligned}$$
$$\mathbf{Pr}[P_t = A] = \text{majority}(\{X_{ti} = A\}_{i \leq n})$$

Learning/Validation

We define a loss function to learn parameters and validate our model using poll data and market prices.

$$\text{Loss}_{KL}(P||Q) = \sum_t \sum_i P_t(i) \log\left(\frac{P_t(i)}{Q_t(i)}\right)$$

Figure 3: KL divergence loss.

Parameter Learning

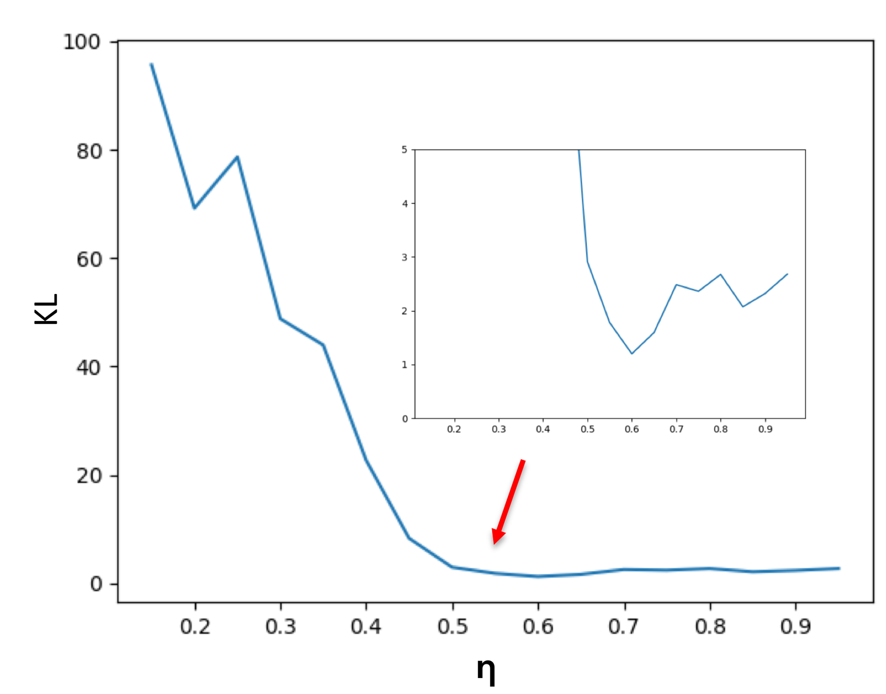


Figure 4: Learning the optimal carryover parameter to minimize KL divergence.

Model Validation

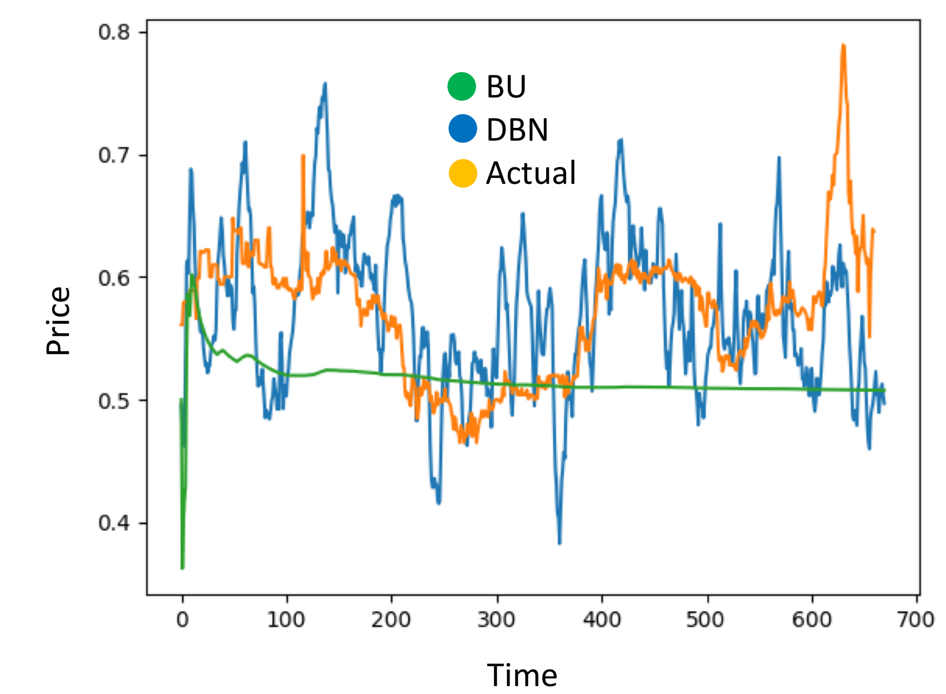
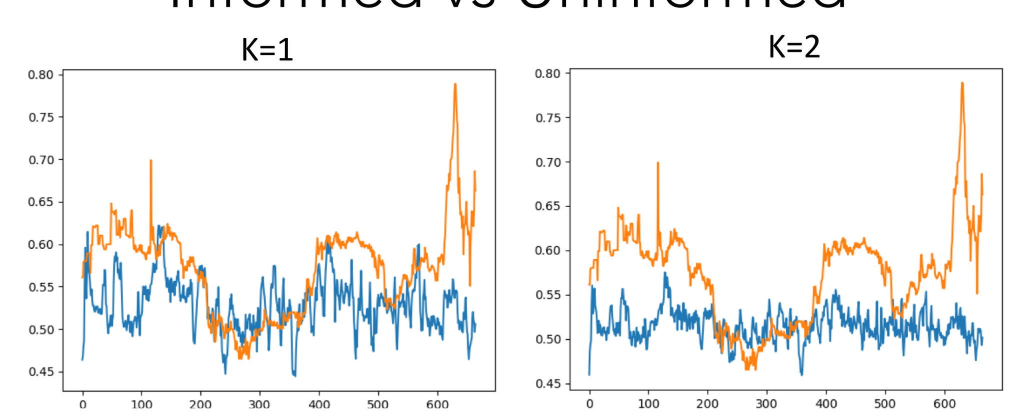


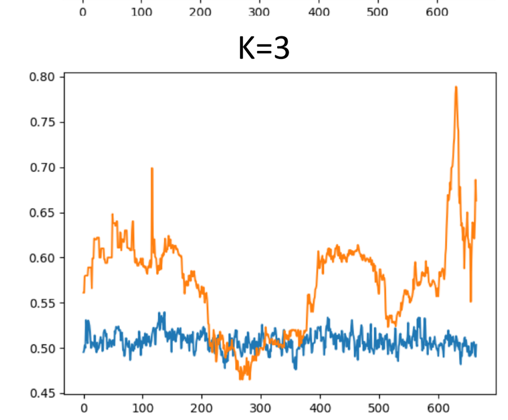
Figure 5: Validating models using real-world polling data to predict market price. BU, DBN, Actual prices are shown.

Model Robustness/Error Analysis

We test the robustness of our model by experimenting with 1) modeling ability with different numbers of informed vs uninformed agents and 2) modeling ability with agents receiving noisy signals of public sentiment.

Informed vs Uninformed





# of uninformed agents	KL Divergence Score
0	6.14087
1	6.42678
2	8.17539
3	9.53749

Figure 6: Comparison of model with K uninformed agents out of 5 total agents.

Noise Amplification

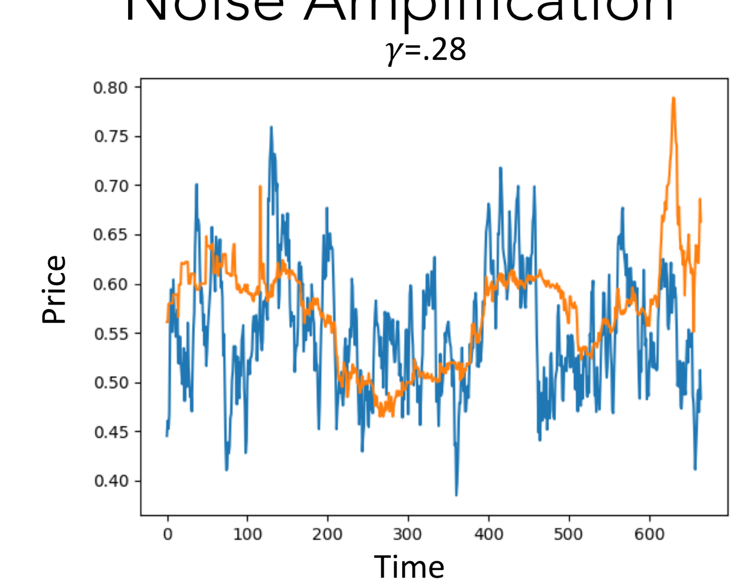


Figure 7: Model comparison with noisy agent signals.

Error Analysis
- Underestimates fluctuations near end of market
- Predictions are noisy because polls don't contain all political climate information

Forward Inference

We can use DBNs to predict future political outcomes without explicit poll or political climate data.

We model the changes in polls over time as a random walk between *P* and *N* with parameter *w*.

Spike in first prediction timestep due to high confidence in election results immediately after last price.

$$\begin{aligned}\mathbf{Pr}[P_t = A|P_{t-1}] &= (1-w)\,\mathbf{Pr}[P_{t-1} = A] \\ &+ w\,\mathbf{Pr}[P_{t-1} \neq A]\end{aligned}$$

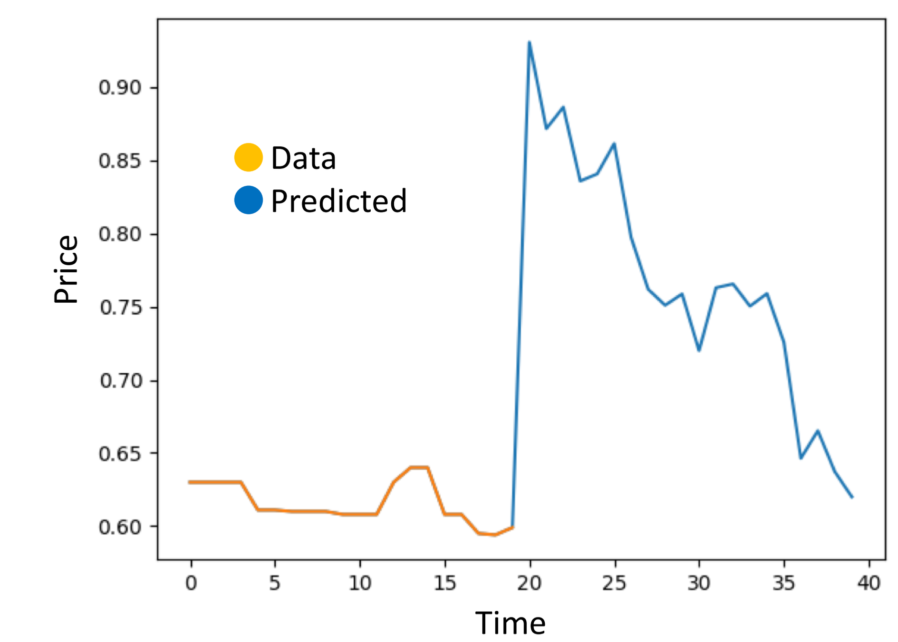
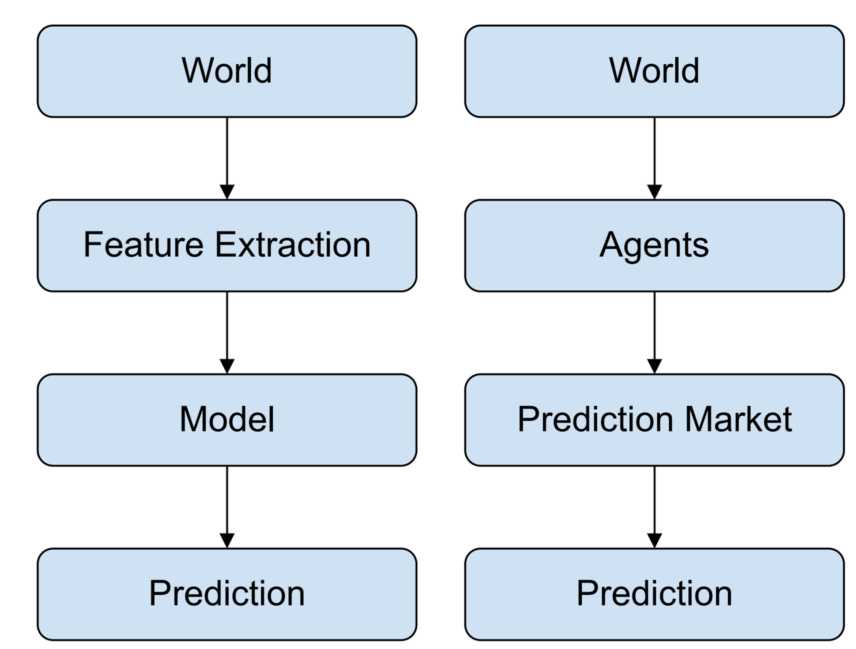


Figure 8: Predicting future results given previous market prices.

Conclusions

1. We have a Bayesian model of a prediction market that achieves a low error given actual behavior of markets in the real world.
2. This model shows that the market is sensitive to agent information but robust to agents with noisy signals of political climate.
3. DBNs take advantage of unique prediction market properties to infer future outcomes and climate without explicit features.

Significance



- Prediction markets aggregate and featurize the world autonomously
- Potential for applications with little data or costly feature engineering

Future Directions/Acknowledgements

Future directions include training multiple RL agents to participate in a market given a DBN representation to maximize profitability. Special thanks to Andrea Shulman for guidance.