



IDEA

O-o us D down

Contraction

## **Problem Statement**

- Prediction of stock prices
  - Explosion of new retail traders
  - Trading signal generation
- Financial time series
  - Autocorrelation
    - Current price correlated with last timestep
  - Stationarity
    - Variance, or other statistics, change over the course of the time series

## Background

- Many techniques have been used in solving stock price predictions
  - Logistic Regression
  - ANN-ARIMA Technique
  - MCMC Simulation
  - LSTM Models

# Methodology

- Hybrid approach: ARIMA-MCMC
  - <u>Auto-Regressive Integrated Moving Average: ARIMA</u>
    - Removes autocorrelation and creates a stationary time series
  - MCMC
    - Creates many simulations of probability distributions using Bayesian methods
- Technique
  - ARIMA is fit to the previous 100 minutes to predict the next 30
  - Residuals are fed to MCMC to predict future residuals
  - Price prediction is the ARIMA prediction plus MCMC predicted residual

#### **Results - Residual MCMC**



#### Results - Price ARIMA-MCMC



## Benchmark

- Deep Long Short-term Memory (LSTM)
  - Current state of the art
  - Pros:
    - Robust to stationarity and autocorrelation
    - No assumptions about models or their distributions are necessary
  - Cons:
    - Black-box model is not transparent
    - Computationally and time expensive

#### **Benchmark Results - LSTM**



## **Conclusion & Future Work**

- Conclusion
  - MCMC has richer output than LSTM
  - ARIMA-MCMC outperformed LSTM in RMSE
  - ARIMA-MCMC took longer than LSTM when forecasting
- Future Work
  - Streamline the training process to speed up ARIMA-MCMC
  - Train and evaluate on more securities
  - Mine output for buy/sell signals + risk values

