# Findings and questions

# Questions:

# Predicting next day price may not be the best thing to do. May want to predict price in 5 days, or even much more ahead?

# Should I predict volume?

# Should I include other features to predict price?

# Findings:

# Best results is with window 15 (about 3 weeks), epoch 2000, and batch-size 50-250.

# There may be upward bias, meaning all that we have seen is upward trend (most of the time).

# When selecting hyperparameters, training predict time, testing predict time, training error eval time, and testing error eval time are negligible.

# The larger the window, the longer the training time. The higher the epoch, and the smaller the batch size, the longer the training time.

# Best training error achieved is with window 15: 2.5616e-06

# Best testing error achieved is with window 10: 7.58083e-05

# Equity Price Prediction with LSTM

Anthony Chi

March 23rd, 2018

## I. Definition

### Project Overview

### Problem Statement

To complete the exercise, here’s the steps that I took:

* Explore data –
* Clean data –
  + Outliers:
  + Missing data:
* Prepare data –
* Feature treatments –
* Model selection –

### Metrics

## II. Analysis

### Data Exploration

### Exploratory Visualization

**Target**

**Outliers**

**Missing Data**

**Delete Feature**

**Feature Types Transformed**

**One-hot Features**

**New Features created**

**Features Transformed**

**PCA**

### Algorithms and Techniques

### Benchmark

**III. Methodology**

**Data Preprocessing**

**Target**

**Outliers**

**Missing Data**

**Delete Feature**

.

**Feature Types Transformed**

**One-hot Features**

**New Features created**

**Features Transformed**

**PCA**

### Implementation

### Refinement

### **DecisionTreeRegressor (benchmark)**

### **KNeighborsRegressor:**

## IV. Results

### Model Evaluation and Validation

### Justification

## V. Conclusion

### Free-Form Visualization

### Reflection

### Citation and Sources

**Relevant Files and Folders**