Visualization of stock data prediction results using stock price dataset

Group 6
Jin Woo Kim, jkim80@kent.edu
Yuan Chen, ychen135@kent.edu
Khushboo Yadav, kyadav1@kent.edu
Priyanka Vyas, pvyas2@kent.edu,

Motivation

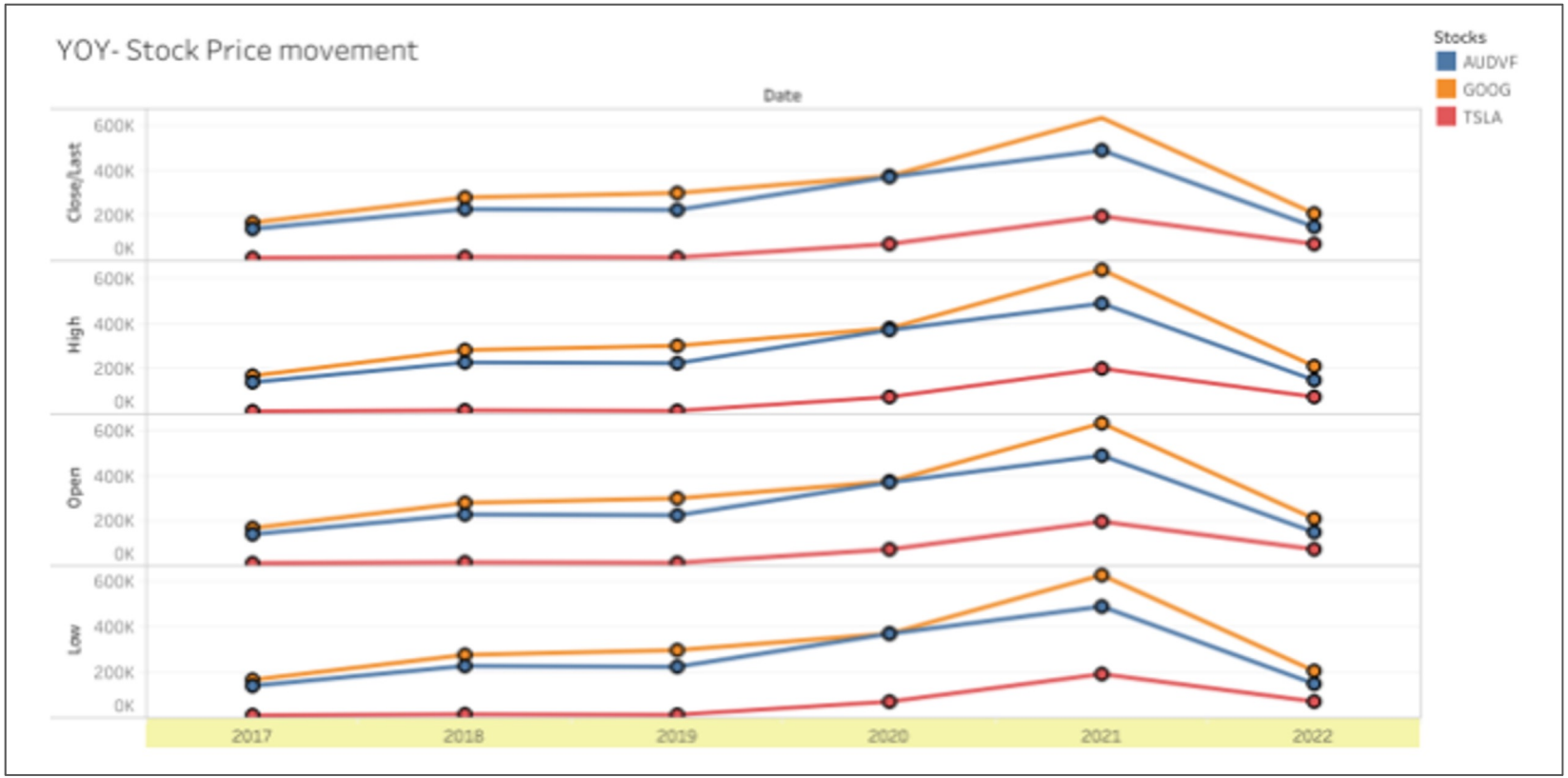
- Traditional approach v/s Exploratory data analysis
- Monitor stock price movement

  -Yearly trend visualization of price movement, Stocks: Tesla, Audi and Google
- Factors affecting stock price movement
- Tesla Competitors Comparative study of Stock price movement

  -Direct Competitors

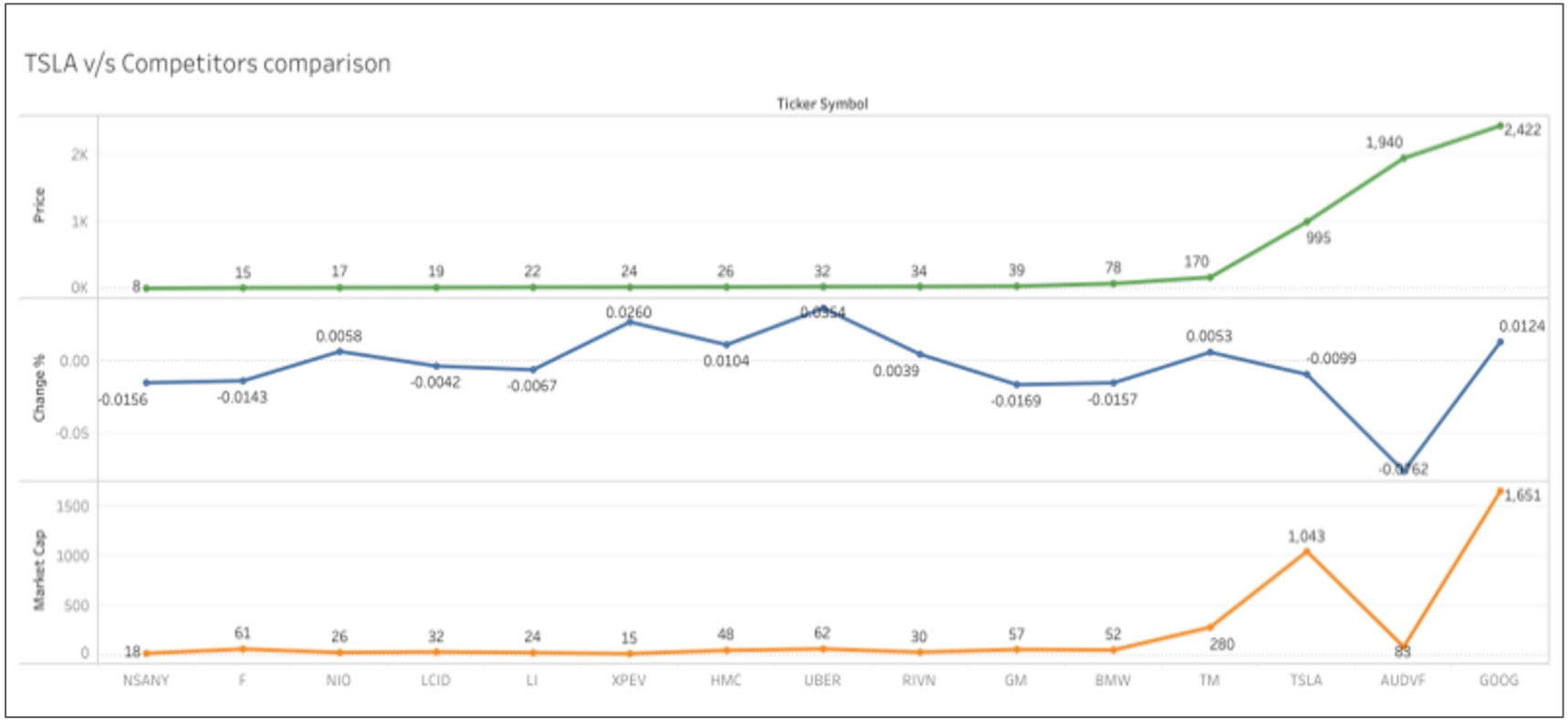
  -Indirect Competitors
- Problem Statement: Supply-Demand balance is driven by market sentiment which in turn affects the stock price trend. One can increase the ROI and business opportunities in Stock market by devising an algorithm to predict the short term price of individual stock.
- Solution Proposed: Novel method to predict and visualise stock price for Tesla stocks using a Neural network.

Yearly Stock Price movement



- Visualization tool: Tableau
- Price movement compared for AUDI, Google and Tesla Stocks

TSLA v/s Competitors Comparison



Visualization tool: Tableau

Stock data is dated 04/24/2022

Project Background

- Literature review:

  -Prediction using machine learning model based on PCA, PSO, SVM, SSA and CNN

  -XGBoost time series prediction model based on chaos theory

  - Burton’s Hypothesis
- Financial time series data

  -Sequential stock price data. Ref: Nasdaq and Kaggle

  -Compiled and cleaned data for selective field.
- Preferred approaches: Neural Networks, CNN, RNN, LSTM etc
- Big Data Visualization

  - Graphs

  -Tableau, Power BI, Python

Project Description

- **Stocks:**

  - Represents a company's ownership stake

  - Tool for investors to grow their money and surpass inflation

- **Tesla:**

  -Leading technological firm

  -Promising Future

  -Tough Forecasting

Problem Statement

To assist the user in training the model by allowing them to select datasets for training and testing sets and choosing from a variety of modeling methods and as well as compare the outcomes.

Proposed Solution: Building Dashboard

- Selecting the training set and test set by controlling a slider.
- Designing AI using deep learning method for prediction.
- The prediction will be displayed with the test set in the dashboard and accuracy

Data

Nasdaq

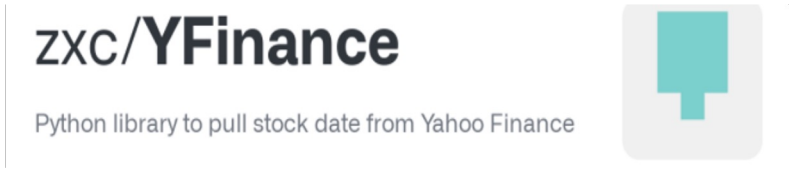
kaggle

Attributes :

- Date
- Open, Close
- High, Low
- Adj Close
- Volume

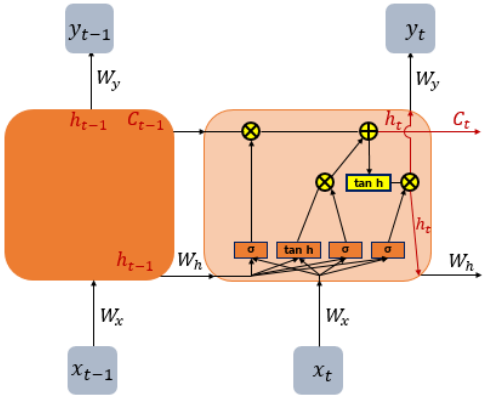
Date	Open	High	Low	Close	Adj Close	Volume
6/29/2010	19	25	17.54	23.889999	23.889999	18766300
6/30/2010	25.79	30.42	23.3	23.83	23.83	17187100
7/1/2010	25	25.92	20.27	21.959999	21.959999	8218800
7/2/2010	23	23.1	18.71	19.200001	19.200001	5139800
7/6/2010	20	20	15.83	16.110001	16.110001	6866900

Visualization Techniques

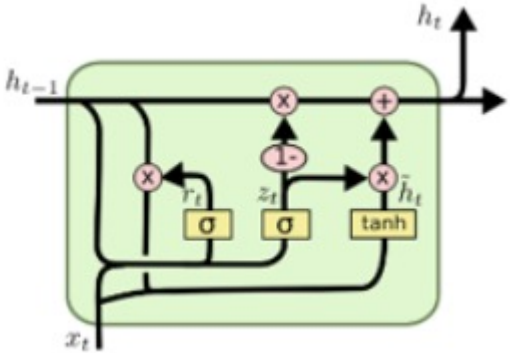


Method - Stock Prediction based LSTM,GRU

LSTM model

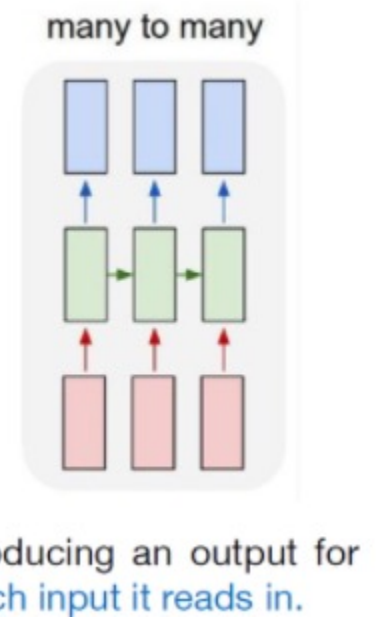


GRU mode



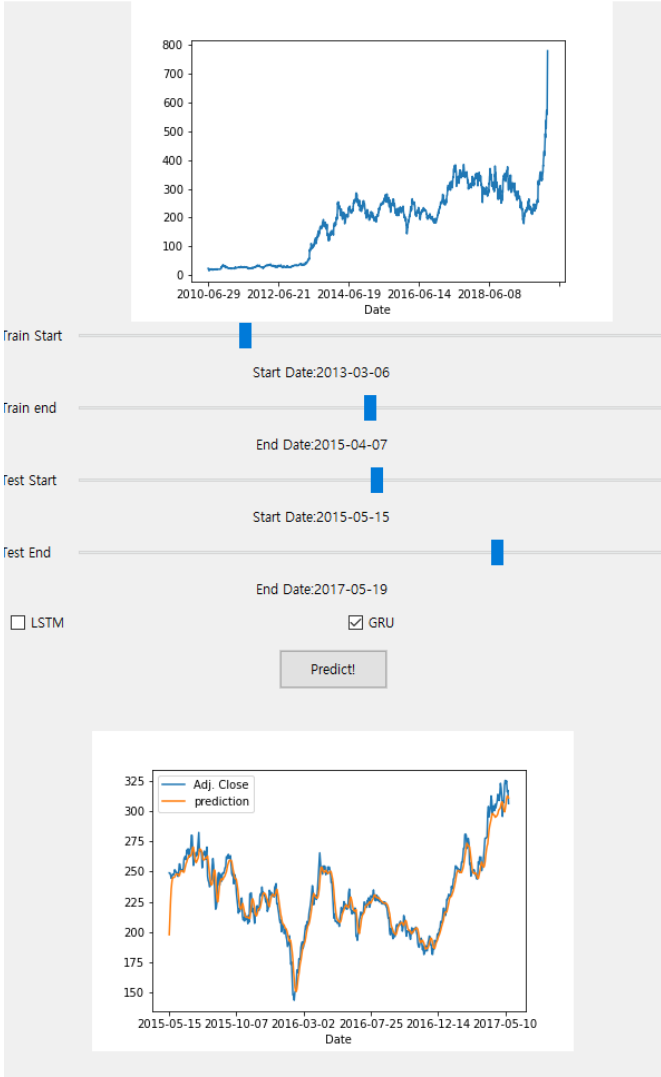
Details of the models

- Many to Many structure.
- 1 feature(“adj close”), 5 past time steps
- data Normalization: Min Max scaler
- LSTM,GRU model
  - 50 units, 2 hidden layers
  - Activation function: tanh
  - Dense unit 2
  - Optimizer: SGD
  - Epoch: 40



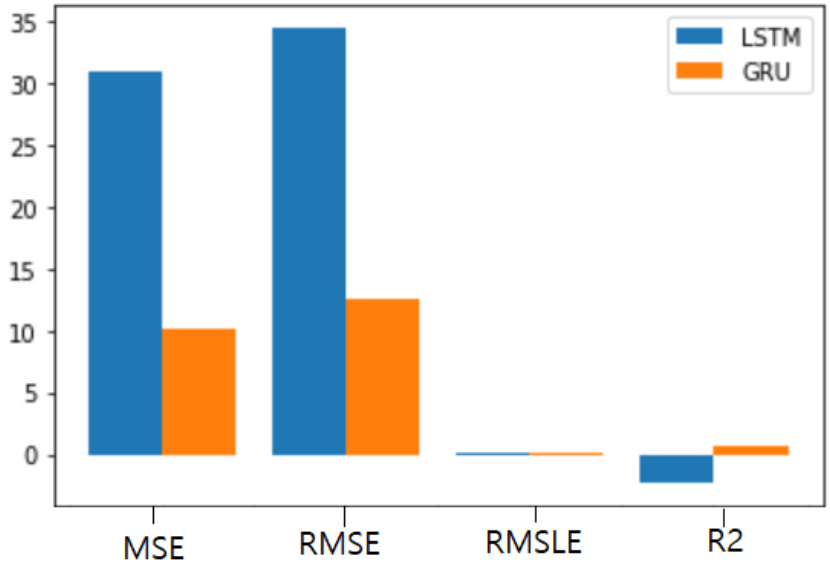
Method/Technique - GUI: Matplotlib + tkinter

- Plotting : matplotlib
- GUI: tkinter
- Plot on the top: Tesla stock from 2010 to 2020
- Use slider to select train and test data period
- Select AI model(LSTM or GRU)
- Press the button to make prediction
- Plot on the bottom shows the comparison between original data and prediction.



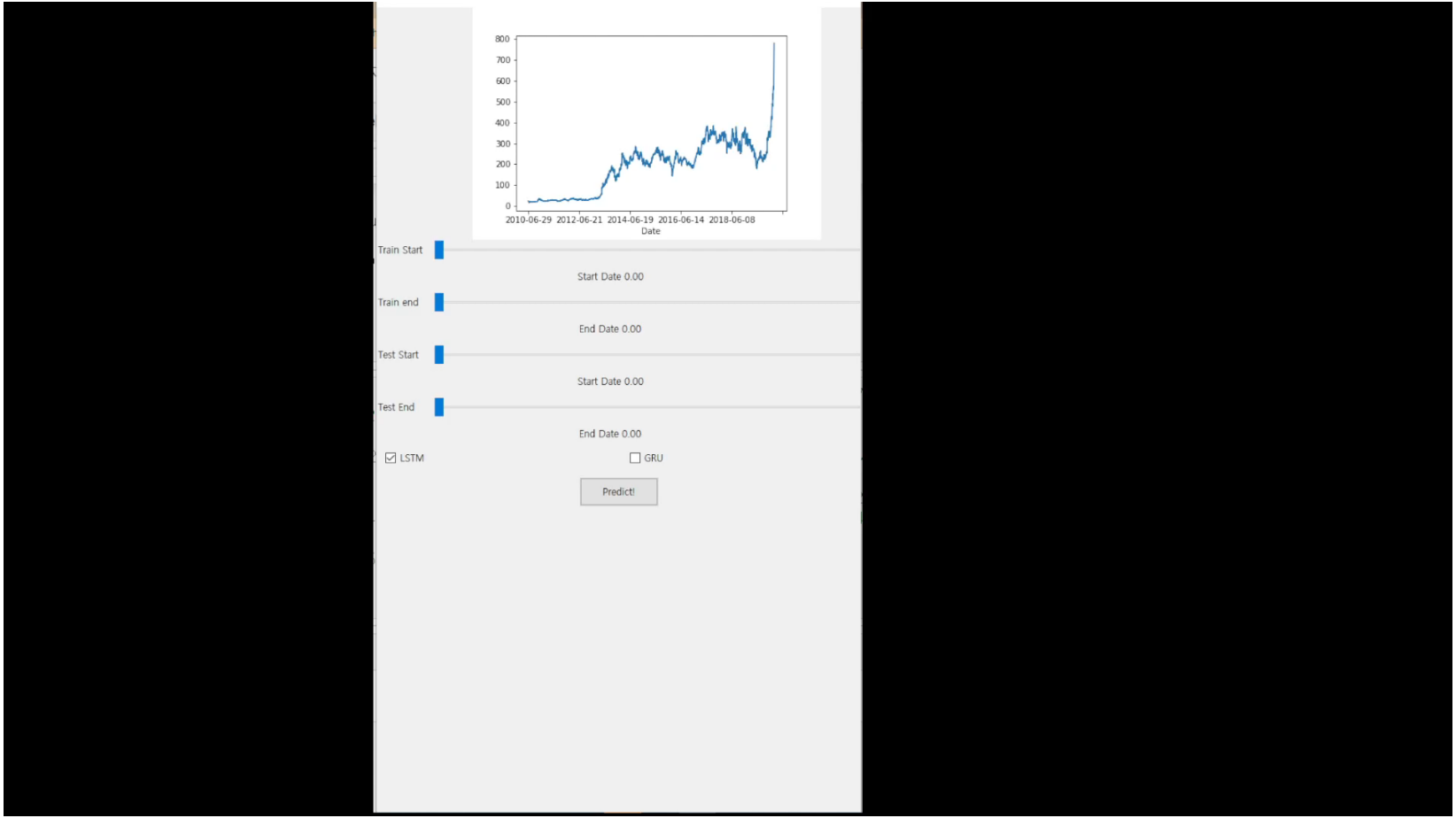
Result

- 10 trials with random period of test and train data.
- Mean of MSE, RMSE, RMSLE, R2 of 10 trials for LSTM and GRU
- GRU showed better result
- The visualization showed the error of the prediction much better than just numbers.



Mean	MSE	RMSE	RMSLE	R2
LSTM	30.94	34.43	0.13	-2.3
GRU	10.17	12.55	0.05	0.62

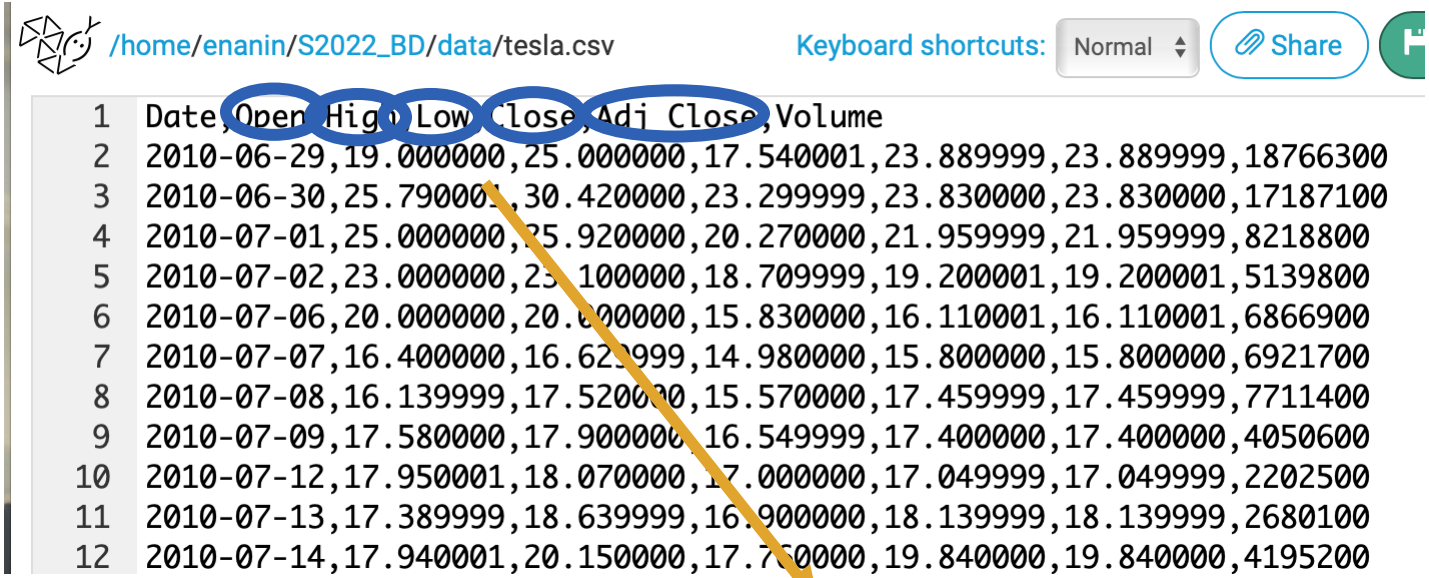
Demo video


800
700
600
500
400
300
200
100
0
2010-06-29 2012-06-21 2014-06-19 2016-06-14 2018-06-08
Date
Train Start
Start Date 0.00
Train end
End Date 0.00
Test Start
Start Date 0.00
Test End
End Date 0.00
LSTM
GRU
Predict!

Experimental Evaluation

- Data Source and Parameter settings
- Comparison of prediction model with existing prediction model
- **Evaluation of prediction model**
- **Summary of experimental Results**

Data Source and Parameter settings



Target value"OPEN_FUTURE"

Comparison of prediction model with existing prediction model

- **VAR(Vector Auto Regression)**

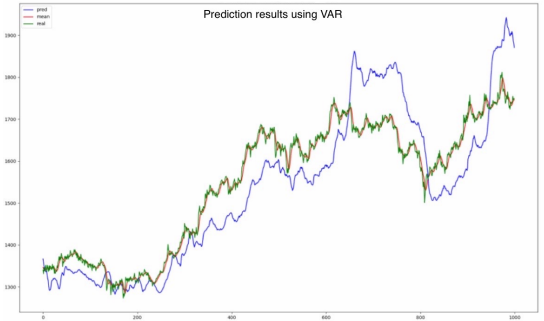
Its core idea does not consider economic theory, but directly considers the relationship between time series of economic variables

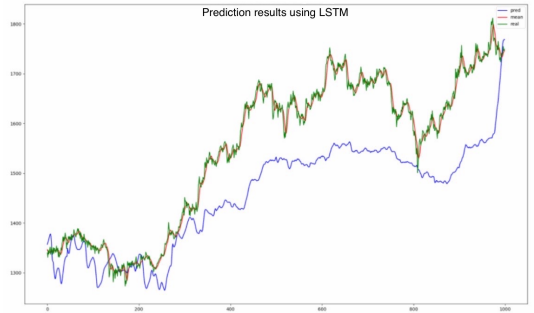
- **LSTM(Long Short- Term Memory)**

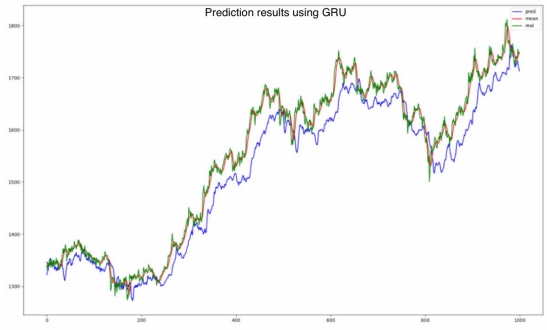
It's suitable for processing and predicting important events with very long intervals and delays in time series.

- **GRU(Gated recurrent Unit)**

GRU has fewer parameters and faster calculation speed, it takes much less time in practice, which can greatly accelerate our iterative process.







Visual comparison of the prediction results

- Prediction result using VAR
- Prediction result using LSTM
- Prediction result using GRU

Evaluation of prediction model

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - z_i)^2}$$

N -- The total number of samples forecasted

di -- The actual value of a sample

zi -- *T*he forecasting value of a sample

MODEL	RMSE
VAR	12.05923
LSTM	9.01871
GRU	7.72115

Summary of experimental Results

- **LSTM model, GRU model and VAR model can predict stock price effectively;**

- **In comparison, GRU has fewer parameters, so its operation speed is faster than others and it performs well in a small amount of data training experience. LSTM performed the most consistently.**

Future Work

- Test the model with more data sets
- Optimize the models
- Consider more factors

Thank You.

Computer Science Department