Zhengyang Zuo <zzuo@andrew.cmu.edu>

Foreign Exchange Directionality Prediction

11-676 Big Data Analytics

Project Purpose

- Problem:
  - Foreign exchange market is rapidly changing and difficult to find pattern or trend
- Purpose:
  - Use machine learning methodology to predict trend in foreign exchange
- Value proposition if solution found:
  - Find factors affecting bid price changes
  - Help traders to make decision on exchanging currency at a point

Analytic Approach

- Predict the directionality of the bid price of two currencies
- Features:
  - Average bid price in a time period (5 minutes)
  - Difference of max and min of bid price in 5 minutes
  - Difference of last bid price and bid price 5 minutes ago
  - Difference of last 2 bid prices
  - Difference between last bid price and ask price
- Features transformed into binary values

Analytic Approach

- Used Java and Cassandra for preparing data.
- Used Spark Pipeline with Scala to:
  - Transform features
  - Train the model
  - Make predictions (classifications)
- Used Random Forest as the model (classifier)

Results

- 50 trees in the random forest
- Used sample data of EUR/USD exchange in Oct. 2010
- Test on the test data set with 3960 instances
- Classification accuracy: 0.5735
- Confusion Matrix:

TRUE	FALSE	<- classified as
29	1677	TRUE
12	2242	FALSE

- TRUE means positive directionality, FALSE means negative or neutral directionality

Error Analysis

- Confusion matrix again:

TRUE	FALSE	<- classified as
29	1677	TRUE
12	2242	FALSE

- Most of the instances (99%) are classified as FALSE
- In dataset, 56.9% of instances has label as FALSE
- Analysis:
  - The number of features is not many, and not distinguishable enough
  - Each decision tree has few features, distribution of label has high effect
  - Most decision trees predict FALSE for most times
  - Random forest predicts FALSE due to majority vote

Future Work

- Following steps on the project:
  - Extract more features
  - Try different models
    - Do feature selection
    - Do parameter tuning
- For next project, we can do more sophisticated predictions
  - Which pair of currencies is most profitable to exchange on
  - How much is the profit for exchange on a given currencies pair

Thanks!