ANALYSIS ON FACTORS AFFECTING CANADIAN DOLLAR
Madhur Chopra, Simranjit Singh Bhatia, Udit Gupta

Motivation and Background

Predicting exchange rates is not an easy task as some experts may suggest. Many factors combined contribute towards the fluctuations of the exchange rate; economic fundamentals are only a part of the whole equation. There are a number of different methods for predicting exchange rates, likely because none of them have been proven to be better than the other. This speaks to the difficulty of generating a quality forecast. Our motivation for predicting future trends in the CAD-USD exchange rate was arisen from the above specified scenario.

Problem Statement

- Examine which factors affect the value of the Canadian dollar i.e. inflation, population, commodities, stock prices, housing price.
- Which factors have maximum influence and what kind of correlation do they have?
- Build a model to predict the future CAD-USD exchange rate.
- Explore which currency has the strongest correlation with the CAD.

Data Processing Pipeline

1. Data Collection:

o Data is gathered for years ranging from 1990 – 2016. 20 features were selected:
o World Bank:
- Metals($US/lb): Gold, Silver, Nickel, Copper, Aluminum, Zinc, Led, Potash
- Forestry($US/tonne): Pulp, Lumber, Newsprint
- Fish: Finfish, Shellfish
- Agriculture($US/100lb): Potatoes, Cattle, Wheat, Barley, Corn, Hogs, Canola
- Energy: Coal, Oil, Natural Gas
o Bank of Canada: Lending Rate, Deposit Rate
o Worldometers: Population, Migrants, World Population, Urban Population
o World Bank: Exports, GDP, Imports ($US in millions)
o Inflation.eu: Inflation
o Investing.com: CAD_US, AUS_US, GBP_US, CNY_US, EUR_US, JPY_US

2. Feature Engineering

o Data Integration: Date/Year is used to join data from various sources
o Missing Values: Data had different intervals
o Population – Yearly, Exports/Imports – Monthly, Currency – Daily
o Interpolation technique was used to fill the missing values after join
o Scaling: MinMaxScaler was used to bring the features in the range of (0,1)

Year	Population
2017	3,66,26,083
2016	3,62,86,378
2015	3,59,39,927
2010	3,41,26,173
2005	3,22,56,333
2000	3,07,01,903
1995	2,92,99,478
1990	2,76,62,440
1985	2,58,48,173
1980	2,45,15,788
1975	2,31,40,609
1970	2,14,39,200
1965	1,96,93,538
1960	1,79,09,232
1955	1,57,33,858

Year	Month	Inflation Rate
1990	Jan	5.50%
1990	Feb	5.46%
1990	Mar	5.30%
1990	Apr	5.01%
1990	May	4.42%
1990	Jun	4.41%
1990	Jul	4.11%
1990	Aug	4.11%
1990	Sep	4.23%
1990	Oct	4.74%
1990	Nov	5.12%
1990	Dec	4.99%
1991	Jan	6.91%
1991	Feb	6.22%
1991	Mar	6.19%
1991	Apr	6.19%
1991	May	6.16%
1991	Jun	6.27%
1991	Jul	5.99%
1991	Aug	5.98%
1991	Sep	5.46%

Date	CAD_US Exchange Rate
19900102	0.8619203586
19900103	0.8618460743
19900104	0.8611780916
19900105	0.8614748449
19900108	0.8619946556
19900109	0.8616232983
19900110	0.8634087377
19900111	0.8651267411
19900112	0.8641548566
19900115	0.8588851671
19900116	0.8593280055
19900117	0.8580744809
19900118	0.8534607835
19900119	0.8449514153
19900122	0.8478168716
19900123	0.8479606546
19900124	0.8437394533
19900125	0.8413967186
19900126	0.8363301832
19900129	0.8413967186
19900130	0.8419634588
19900131	0.8422471153

INTERPOLATION

| Date	CAD_US Exchange Rate	Inflation	Population
19900102	0.8619203586	0.055	27662440
19900103	0.8618460743	0.0549818182	27663697.3256528
19900104	0.8611780916	0.0549636364	27664954.6513057
19900105	0.8614748449	0.0549454545	27666211.9769585
19900108	0.8619946556	0.0549272727	27667469.3026114
19900109	0.8616232983	0.0549090909	27668726.6282642
19900110	0.8634087377	0.0548909091	27669983.953917
19900111	0.8651267411	0.0548727273	27671241.2795699
19900112	0.8641548566	0.0548545455	27672498.6052227
19900115	0.8588851671	0.0548363636	27673755.9308756
19900116	0.8593280055	0.0548181818	27675013.2565284
19900117	0.8580744809	0.0548	27676270.5821813
19900118	0.8534607835	0.0547818182	27677527.9078341
19900119	0.8449514153	0.0547636364	27678785.2334869
19900122	0.8478168716	0.0547454545	27680042.5591398
19900123	0.8479606546	0.0547272727	27681299.8847926
19900124	0.8437394533	0.0547090909	27682557.2104455
19900125	0.8413967186	0.0546909091	27683814.5360983
19900126	0.8363301832	0.0546727273	27685071.8617511
19900129	0.8413967186	0.0546545455	27686329.187404
19900130	0.8419634588	0.0546363636	27687586.5130568
19900131	0.8422471153	0.0546181818	27688843.8387097

```
X_std = (X - X.min(axis=0)) / (X.max(axis=0) - X.min(axis=0))
X_scaled = X_std * (max - min) + min
```

3. Feature selection

- Features which were highly correlated between themselves were removed.
- Features which were highly correlated to the Canadian Dollar were kept.
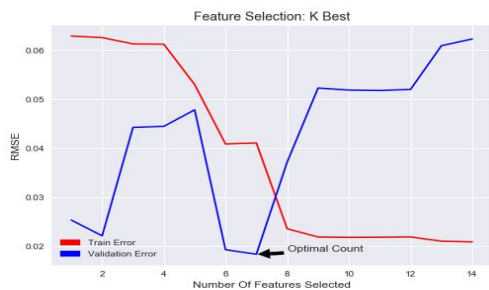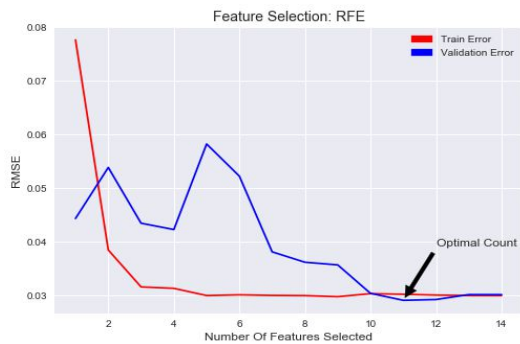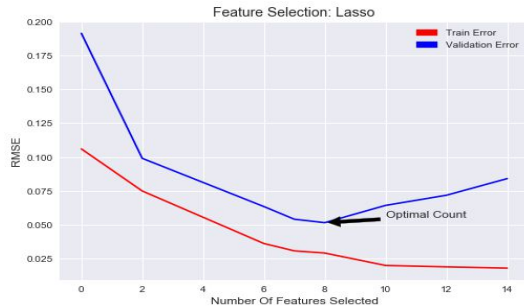
Features Removed:
Urban Population
World Population
Gold_Price
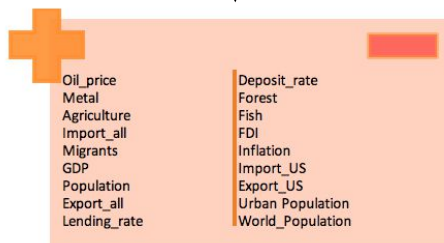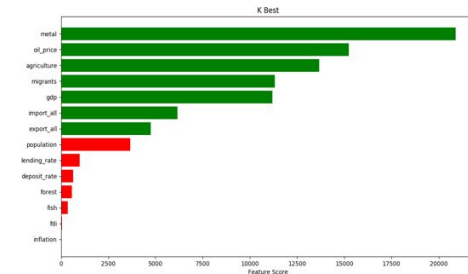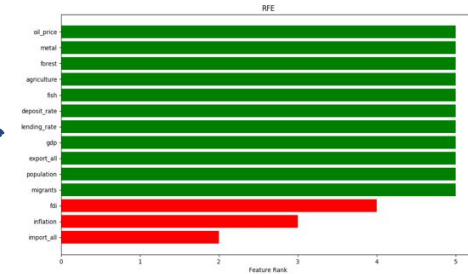Energy
Export_US
Import_US
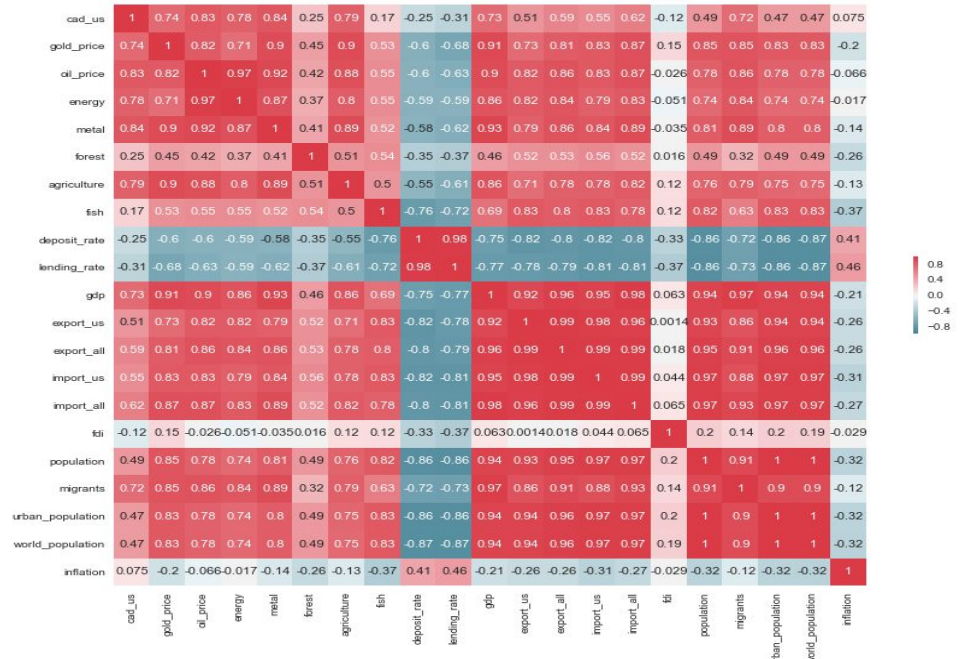
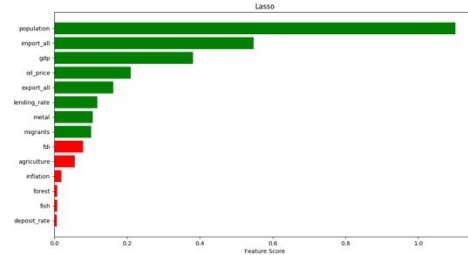- Used different feature selection methods: Lasso, K-Best and RFE

o Combined the outputs do identify the best features

<div align="center">

Optimal Feature Count   Best Features

</div>

Finally 9 out of 20 features were selected and considered to be the most influential ones: oil, metal, agriculture, imports, migrants, gdp, population, exports and lending rate.

3

Model Training

- o Various regression models were trained on the features selected
  - • Polynomial, SGD, Linear Regression, Ridge, Elastic Net, Lasso, Bayesian
- o Training Data: January 1990 to December 2014
- o Validation Data: January 2015 to November 2016
- o Predictions were made for December 2016
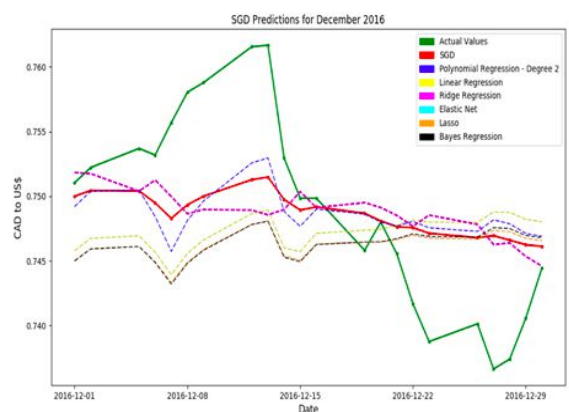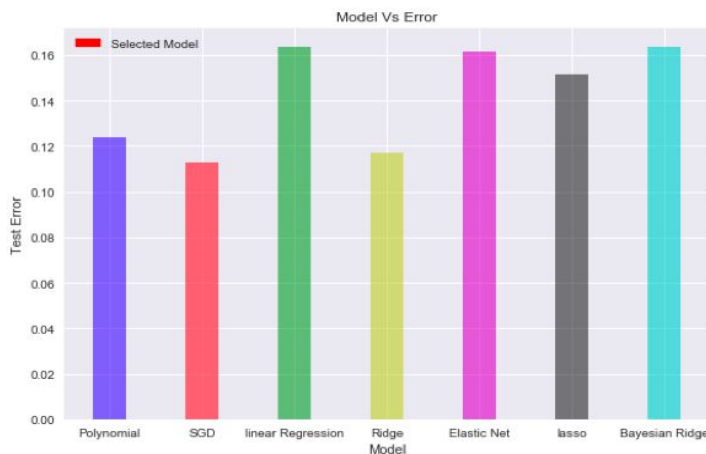
The best Predictions for CAD-USD exchange rate was given by SGD

- o Below Graph shows difference in actual and predicted values for each day in December 2016:
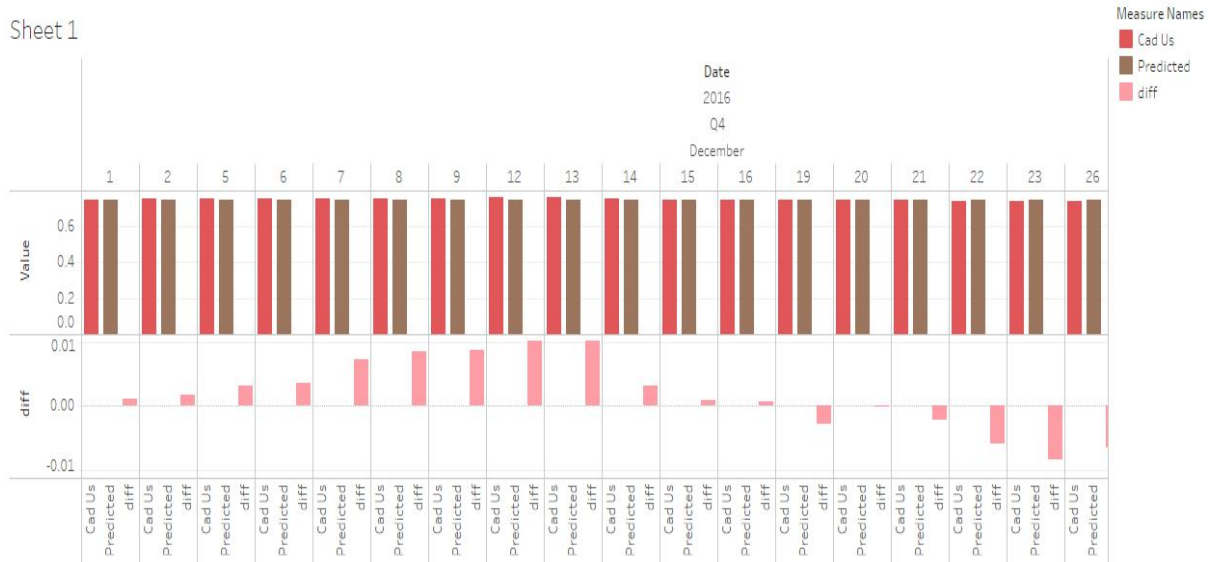
Correlation with Other Currencies

- o Canadian Dollar shows high correlation with Euro and Australian Dollar:

4

o Low correlation with Chinese and Japanese Yen:

Data Product

1. The 9 most influential features on CAD_US exchange rate: oil, metal, agriculture, imports, migrants, gdp, population, exports and lending rate.
2. Predictions made on December 2016 CAD_US exchange rates.

Conclusion and Learning

1. Feature selection should not be left to a single feature selection technique. We should apply multiple feature selection techniques for selecting features for insightful results.
2. For time series data, it is necessary to capture the current trends and thus model should be trained with the most recent data available along with the historical data.
3. Canadian Dollar, European Euro and Australian Dollar are highly correlated whereas Japanese Yen and Chinese Yuan have no correlation with the Canadian Dollar.

Summary

o From our analysis, factors which have the most influence on the Canadian Dollar are: Oil Price, Metal, Agriculture, Import, Migrants, GDP, Population, Exports, Lending Rate.
o Prediction of Canadian Dollar is not easy based on the training from the past data, however training the model on recent data produced better predictions.
o More confidence can be introduced by combining the project with other factors that influence exchange rates like - sentiments.