

EECS6895 Adv. Bigdata Analytics HW 2 - Sun-Yi Lin(sl3833)

10 Firms Stock Prices Prediction

The stock prices prediction is the most major part in data analytics, there're tons of models aim to predict the stock prices precisely. In this homework, I'm going to use **Python** and some **Python libraries** to design a simple model to predict the stock prices of **tomorrow**.

The Python libraries I choosed are **matplotlib**, **pandas** and **yahoo-finance**. To install these libraries, use the following commands in Terminal.

```
pip install matplotlib
pip install pandas
pip install yahoo-finance
```

I write two Python applicatinons to predict the price of the 10 firms' stock.

PridictionPlot.py

The first is the full-functional one that accepts the company's code as input, and prints out the **close price**, the **estimate price**, the **last day that the datas been collected**, and even **a plot that shows statisitc diagrams**.

In this application, I collect the prices start from 01/01/2015 to the submission day, and calculate the **moving average** and the **exponential weighted moving average**, calculate the deltas of the last two elements of these two sequences, give these two numbers a specific ratios (in here, the golden ratio is actually used) and finally add the composed value to the last closed price of the company.

The lines in the plot shows the close price labeled as the company's code, the moving average labeled as MAVG, and the exponential weighted moving average labeled as EWMA.

```

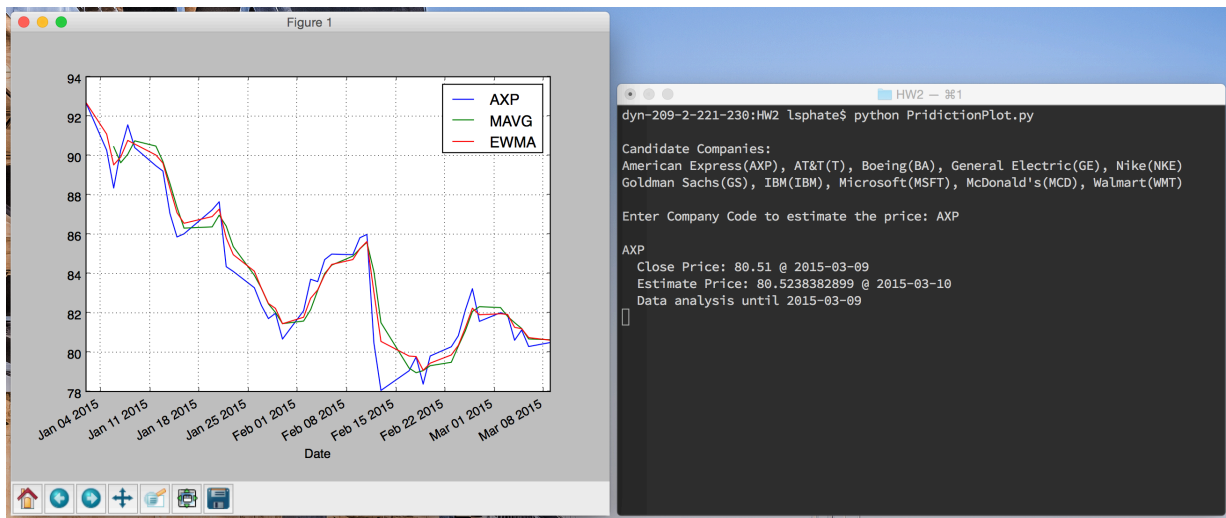
import sys
import datetime as dt
import pandas as pd
import pandas.io.data as pio
from pandas import Series
import matplotlib.pyplot as plot
from yahoo_finance import Share

def plotDiagram(company, today):
    df = pio.get_data_yahoo(
        company,
        start = dt.datetime(2015, 1, 1),
        end = dt.datetime(today.year, today.month, today.day))
    df.loc[today] = (Share(company).get_open(), Share(company).get_days_high(),
Share(company).get_days_low(), Share(company).get_price(), Share(company).get_volume(),
float(Share(company).get_price()))
    cprice = df['Adj Close']
    mavg = pd.rolling_mean(cprice, 10)
    ewma = pd.ewma(cprice, span = 3)
    print "\n" + company
    print (" Close Price: "
        + str(Share(company).get_price())
        + " @ " + str(today))
    print (" Estimate Price: "
        + str(predictPrice(cprice, mavg, ewma))
        + " @ " + str(today + dt.timedelta(days=1)))
    print " Data analysis until " + str(today)
    cprice.plot(label = company)
    mavg.plot(label = 'MAVG')
    ewma.plot(label = 'EWMA')
    plot.legend()
    plot.show()

def predictPrice(CPRICE, MAVG, EWMA):
    gr = (1 - 5 ** 0.5) / 2
    return ((EWMA[-1] - EWMA[-2]) * (gr)
        + (MAVG[-1] - MAVG[-2]) * (1 - gr)
        + CPRICE[-1])

if __name__ == '__main__':
    print "\nCandidate Companies:"
    print "American Express(AXP), AT&T(T), Boeing(BA), General Electric(GE), Nike(NKE)"
    print "Goldman Sachs(GS), IBM(IBM), Microsoft(MSFT), McDonald's(MCD), Walmart(WMT)"
    today = dt.date.today() - dt.timedelta(days=0)
    company = raw_input('\nEnter Company Code to estimate the price: ')
    plotDiagram(company, today)

```



Pridiction10Firms.py

The second application is reduced to only get the ten numbers of the prediction as the homework asked. The ideas of the application are identical to the above one.

```
import sys
import datetime as dt
import pandas as pd
import pandas.io.data as pio
from pandas import Series
from yahoo_finance import Share

def plotDiagram(company, today):
    df = pio.get_data_yahoo(
        company,
        start = dt.datetime(2015, 1, 1),
        end = dt.datetime(today.year, today.month, today.day))
    df.loc[today] = (Share(company).get_open(), Share(company).get_days_high(),
Share(company).get_days_low(), Share(company).get_price(), Share(company).get_volume(),
float(Share(company).get_price()))
    cprice = df['Adj Close']
    mavg = pd.rolling_mean(cprice, 10)
    ewma = pd.ewma(cprice, span = 3)
    print "\n" + "Estimate Price of " + company + " @ " + str(today + dt.timedelta(days=1)) + ": "
    print " " + str(predictPrice(cprice, mavg, ewma))

def predictPrice(CPRICE, MAVG, EWMA):
    gr = (1 - 5 ** 0.5) / 2
    return ((EWMA[-1] - EWMA[-2]) * (gr)
        + (MAVG[-1] - MAVG[-2]) * (1 - gr)
        + CPRICE[-1])

if __name__ == '__main__':
    firms = ['AXP', 'T', 'BA', 'GE', 'NKE', 'GS', 'IBM', 'MSFT', 'MCD', 'WMT']
    today = dt.date.today() - dt.timedelta(days=0)
    for company in firms:
        plotDiagram(company, today)
```

```
dyn-209-2-221-230:HW2 lsphate$ python Prediction10Firms.py
Estimate Price of AXP @ 2015-03-11:
80.5238382899

Estimate Price of T @ 2015-03-11:
33.138253381

Estimate Price of BA @ 2015-03-11:
154.655144245

Estimate Price of GE @ 2015-03-11:
25.6208479796

Estimate Price of NKE @ 2015-03-11:
97.4004200224

Estimate Price of GS @ 2015-03-11:
187.195460246

Estimate Price of IBM @ 2015-03-11:
161.141277834

Estimate Price of MSFT @ 2015-03-11:
42.7179877367
```