Fernando, Ong, Santos, Triaca

Approved Big Data Proposal

Big Data Problem Proposal Big data Problem: Given historical data on a specific stock, how do you find the best dates to enter and exit a long position using MACD for indicators?

Source of data Stock data can be downloaded from finance.yahoo.com Example (full historical data of Apple stock, 9000+ lines in excel): https://finance.yahoo.com/quote/AAPL/history?period1=345398400&period2=1510675200&interval=1d&filter=history&frequency=1d

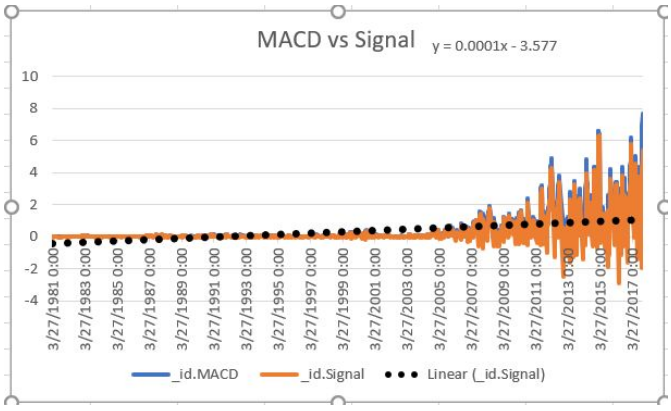
Discussion of Source Dataset

The dataset downloaded from yahoo finance shows the historical stock data of apple (AAPL) stock. It ranges from 1980-2017. It displays date, open price, high price, low price, and closing price. There are 9314 rows of data.

Discussion of output and explanation

Before the mapreduce could take place, some set up of data would need to happen first. The needed values would be (using closing price) moving average window 12, moving average window 26, difference of the two values (which would comprise of the MACD value) and the moving average window 9 of the MACD values, which comprise the signal line.
The output of the first Mapreduce is meant to show the MACD values that are greater than 0. If the MACD value is greater than 0 it means that the shorter moving average window of 12 is higher than the longer moving average window of 26. This could possibly indicate upward trend.
The second mapreduce takes the MACD values that are already reduced to the MACDs that are greater than signal line. If MACD is greater than signal, it usually indicates a buy signal. After the 2 mapreduces, it is reduced to 3150 rows.

Visualization



The graph shows that early on, the MACD and signal were almost the same. The same is indicated by the trendline because of its low value of slope. (.0001). The MACD only overtook the signal clearly towards the latter part of the dataset.