3/20/2016

**MTH 9898 – Big DATa in finance**

**Assignment C – “Using Twitter to Forecast Stock Prices”**

Eldar urmanov

# Introduction

In this assignment we are using Apache Spark framework on AWS EC2 cluster to analyze the large set of tweet data to extract information that will allow us to predict whether the stock price will go up or down.

# Decription of METHODOLOGY

In my analysis, I only looked at one stock, Google, to make things simple. Apache Spark has a steep learning curve and, in my opinion, a program as simple as this one, is good enough to demonstrate the familiarity with Spark.

I further filtered my tweet data on English-language tweets. I also kept only those tweets where Google’s major products, Android and Chrome, are mentioned.

Assumption here is that the tweets that show high frequency of appearance of these two keywords indicate the tendency for Google’s stock to appreciate in value.

# RESULT

The program shows how many tweets contain word “Android” and word “Chrome” separately. Based on this information, further analysis can be done to predict the stock price behavior.