5th International Conference on

Big Data Analysis and Data Mining

June 20-21, 2018 | Rome, Italy

Stock market real time recommended model in big data framework using Apache Spark

Essam M Ramzy Hamed, Mostafa Mohamed Seif and **A Hegazy**
Arab Academy for Science, Technology and Maritime Transport, Egypt

Stock market is considered as a complicated and nonlinear system. Now stock market prediction is recognized as an attracting point for financial investors. The historical price is not considered as the main factor to predict stock market trend. There are many factors such as politics and natural events that affect social media environments like Twitter and Facebook that generate huge datasets and need data analysis to extract the polarity of the data and its effectiveness on the stock market. On the other hand, this data may be unstructured and may need special handling on storing and processing. This paper proposes a real time forecasting of stock market trends based on news, tweets, and historical price. A supervised machine learning algorithm is used to build this model. Historical price will have to be combined with sentiment analysis to build hybrid model based on Apache Spark and Hadoop HDFS to handle big data (structured and unstructured) generated from social media and news websites. The proposed model works on two modes; offline mode that works on historical data included today's data after ending of stock market session and real time mode that works on real time data. This model increases the accuracy of prediction due to the additional features added by sentiment analysis on StockTwits and market news data. Additionally, the model enhances the performance of handling this data set due to parallel processing occurred on data using Apache Spark.

dodessammisr@gmail.com

Computer Engineering & Information Technology
ISSN: 2324-9307

Data Mining 2018
June 20-21, 2018

Volume 7

Page 79