Analyzing Real-Time Tweets and Their Impact on Stock Prices Using NiFi, Solr and Spark MLlib

Use Case 1 - Exploratory and Visualization

Understanding your Customers' Opinions by Streaming and Summarizing Twitter Data



Twitter has 330 Million monthly active users expressing their opinion on the internet Global IP traffic has reached 1.1 Zettabytes. By 2020 it will grow to 40 Zettabytes





What impact can twitter sentiment have on stock market?

Airlines and hotel industries have adopted twitter based customer service



Spark Streaming



Twitter data is streaming through spark





NiFi connects the streamed twitter data and temporarily stores





Solr analyzes tweets and summarizes them in terms of location, language, word cloud and top tweeters





Interactive Banana
Dashboard conveys live
streaming data

Use Case 2 – Statistical Analysis Analyzing the Impact of Global Twitter Sentiment on Target's Stock Price

Storage



Streaming stock market data through python API



Data is being cleaned and stored as a pandas data frame



Closing rates for every minute are visualized on Matplotlib



Calculating the corelation between sentiment and stock prices with Spark MLlib







Streaming twitter data through spark



Sentiments are analyzed and converted to HDFS format and store in HIVE



Tweets from various languages are translated to English for sentiment analysis

The sentiments are visualized on Matplotlib



Where can I apply these in my company?

Are you a retailer? Are you launching a marketing campaign? You can measure the consumer reaction with our framework without any kind of a survey.

Launched product? Do you want to know how people are receiving it? You can use our framework to gauge customer satisfaction even before they call.

Do you want to stay ahead of vour competitors? Set up live customer service through twitter deploying our framework

Did your company announce the financial report? Do vou want to track market sentiment? Deploy our framework and watch live market sentiment along with stream tweets

Section 1 Team 7

Regression

Badarinath K S (kanak016@umn.edu) Barkha Thakur(thaku048@umn.edu) Devansh Gargh (garg0042@umn.edu) Rachel Wolfe (wolfe384@umn.edu)

Yuyang Zhang (zhan5471@umn.edu)