Real time sentiment analysis using python for classification (positive,negative,neutral)

using tweepy ,Pyspark ,keras,nltk

real time sentiment analysis is an AI solution to track mentions of your brand and products, wherever they may appear, and automatically analyze them with almost no human input needed

Examples of Sentiment Analysis

Social Media Monitoring

Almost half of the world's use social media platforms daily and in regular basis . Every minute users send over half a million tweets and Facebook comments , surely lot's of these messages and comments contain business related conversations and a large amount of these messages contain valuable business insights about how customers feel towards products, brands and services.

Sentiment analysis allows businesses to mine this data and extract the sentiments that is found in social media conversations, to understand how people are talking about a given product or topic, and why.

This way of engaging with customers is very effective and rapid mining of this data allows analytics to quickly come up with a decision that benefits the company.

This mined data can also be used to monitor the competition and what the rivals do that attracts customers that way finding new opportunities to enhance our products

Brand Monitoring

Using blogs and forum conversations , companies can detect what customers think of their products also these communication tools can affect a product's reputation so that means that the company may get to stop a rumor or clarify a user-friendly feature in a product that is misunderstood

Customer Support Analysis

Customer service is a key towards products enhancement where more than 90%

Of customer's relay on customer service for reporting issues and solving these issues.

The most important issues and problems can be easily detected using sentiment analysis

Customer Feedback Analysis

Applying sentiment analysis on surveys are very important just like brand monitoring however this time the customer is asked to write his opinion about the services of the company or a product in particular or even give an idea about how to improve the services of the company

Market Research

Let's say the company wants to launch a campaign

It then needs to now what product attracts customers the most and start producing that product extensively or put an offer selling one unwanted product with another that people desire so that people buy both with a cheaper price

Creating a Custom Sentiment Analysis Classifier

Step 1

Start collecting the data using twitter API with the Tweepy library

This can be done by registering into a developer account in twitter and asking for access token ,secret access token, api key and api secret

And then using OAuthHandler and streamlistner from tweepy library to start receiving tweets

Then use cursor in order to choose the keyword to be searched ,language of the tweets and how many tweets to get

Step 2

Find dataset with tweets or sentiment text in general with known label output and train the machine learning models on the dataset

Step 3

Use the pretrained model for predicting whether the collected data from twitter API is positive,negative or neutral

Enhancements:

- Using full dataset in training and validation phases all 1600000 observation
- Using pyspark for machine learning classification on Big data
- Shuffling data for getting optimal distribution
- Applying various data cleaning techniques on text dataset using regular expression and pyspark feature manipulation: HashingTF, Tokenizer, StopWordsRemover

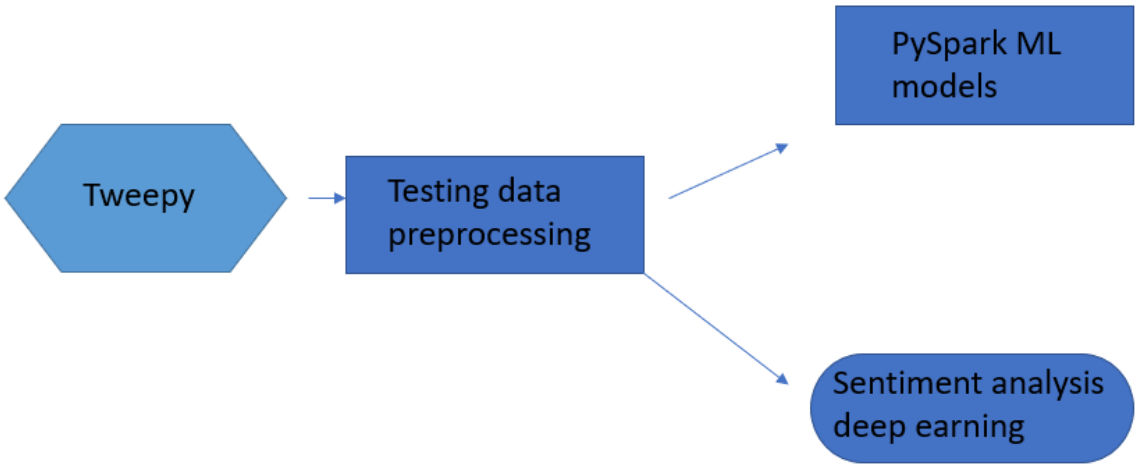
Suggested enhancements:

Applying tuning of neural network for optimal results

Using socket for stream capturing as it updates tweets instantly giving a better chance for a quick decision and modeling

Performance results:

Model	Test accuracy
Logistic regression	74
Naïve bayes	76
GBTclassifier	70
FMclassifier	71
Neural network	76



Script flowchart