

Can Twitter Sentiment predict Tezos price change ?

Penny Johnston - MSc in Big Data - 2018 -  Tezos (XTZ)

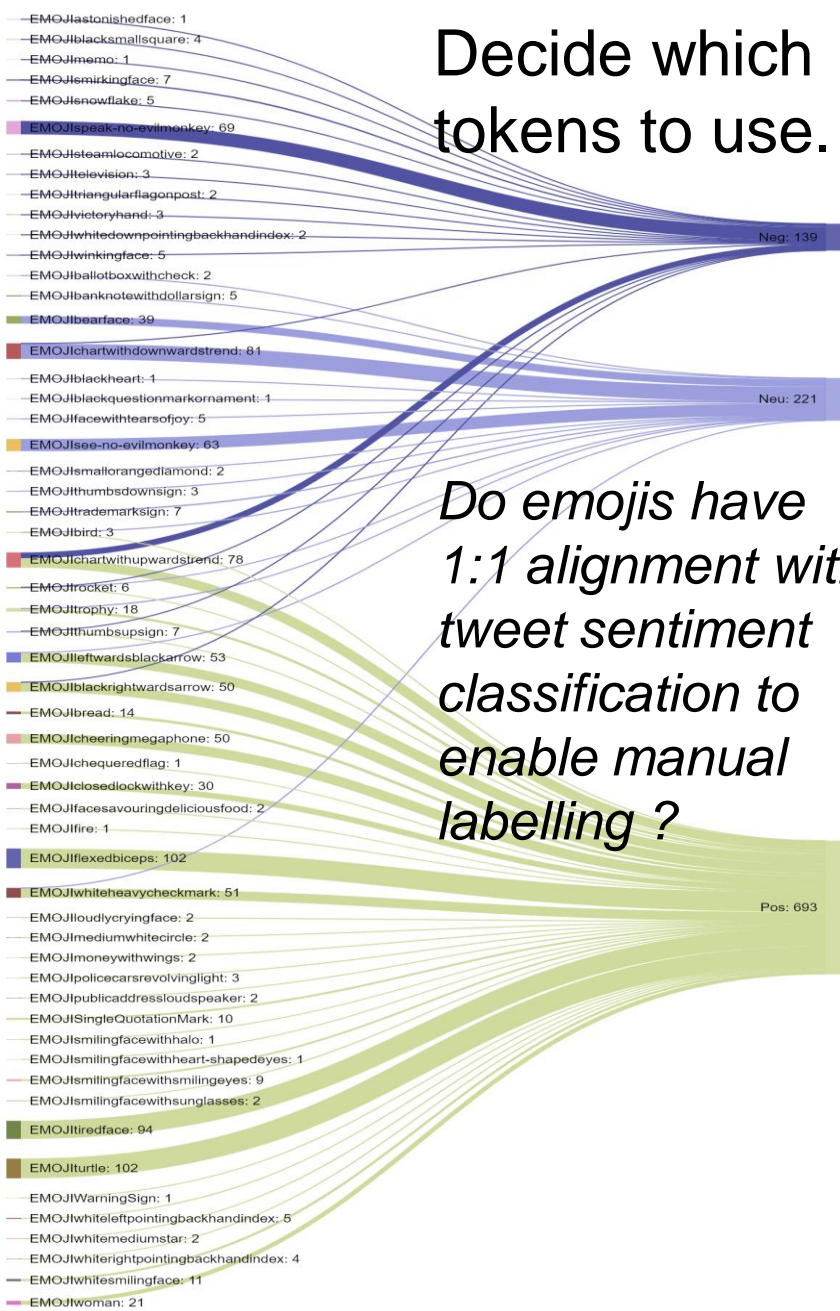
Machine Learning Pipeline

Step 1: Tokenisation

>Sankeymatic.com

Decide which tokens to use.

Do emojis have 1:1 alignment with tweet sentiment classification to enable manual labelling ?



Step 3: Classify each tweet with either positive or negative sentiment.

Supervised or Unsupervised Learning

Naïve Bayes Classifier.

>NLtk.TextBlob >NLtk.NaiveBayesClassifier

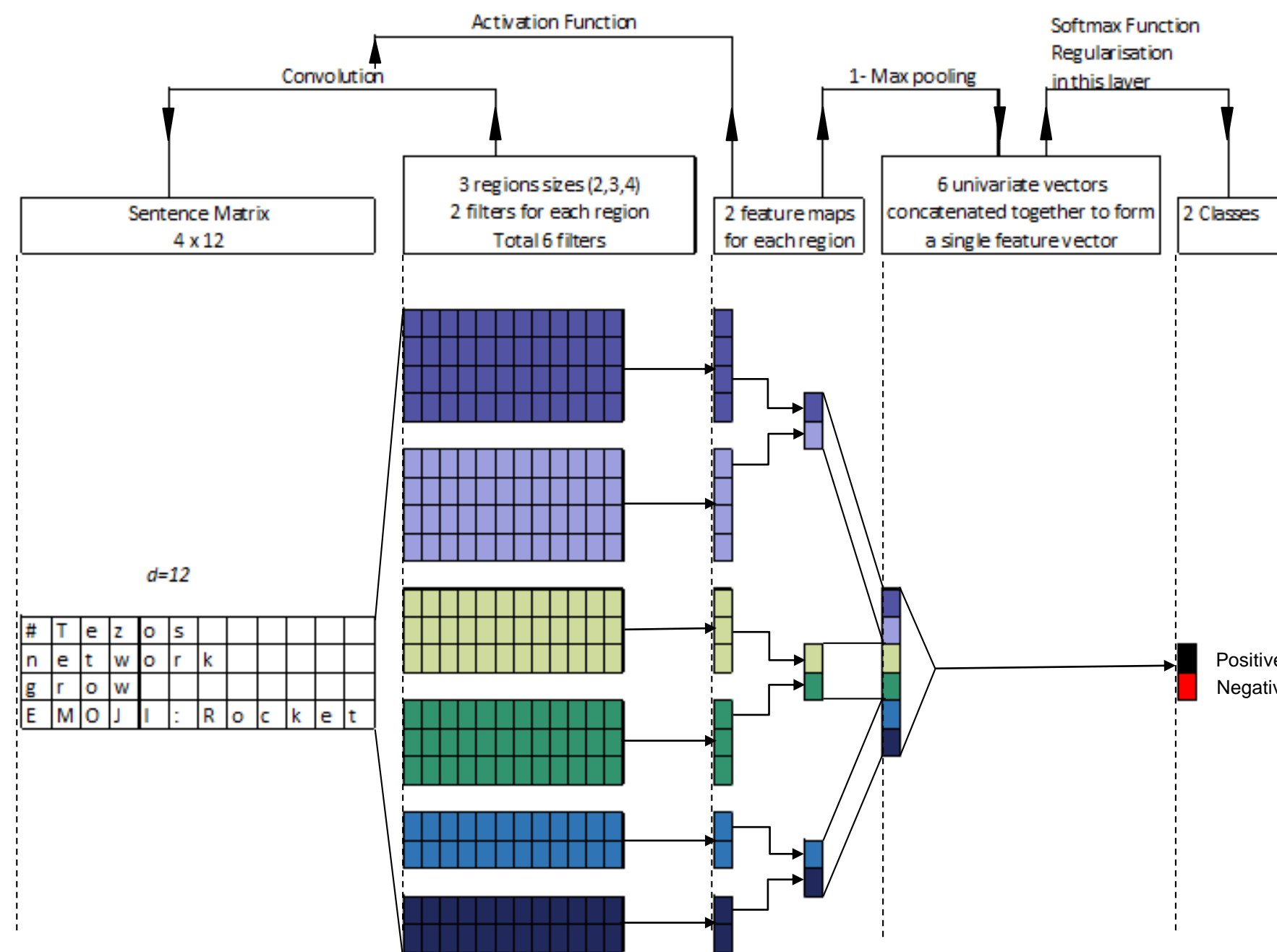
There are many classifiers, some require labelling (supervised learning) which involves labelling a dataset and training the model such as Naïve Bayes Classifiers.

Convolutional Neural Network.

>Anaconda >Tensorflow >Keras

Others, as in the Convolution Neural Net below, require no supervision and extract their features from an unlabelled dataset in order to train their model.

The first step in creating the CNN, is to supply it with an embedded tweet vector. This is a rich distributed token representation of the tweet.



Step 2: Clean & Stem Tweet

>Tweepy >Mongodb
>Python >Pandas

Example of a clean tweet :
#Tezos network grow world node connect network continu increas we begin guy EMOJI:Rocket #X @Name1 @Name2 @Name3 @Name4 @Name5 @Name6 https://t.co/<Code>

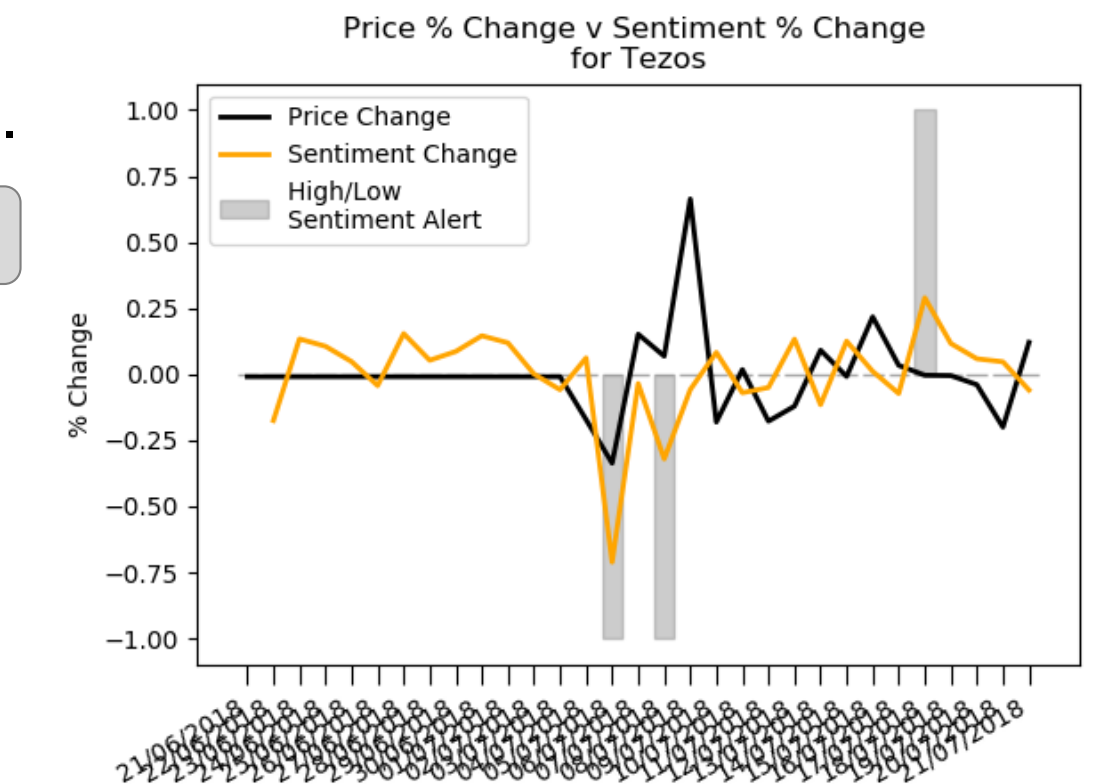


Step 4: Results

>Python >Plotly



Plot the resulting sentiment % change against the price % change downloaded from CoinMarketCap.



Step 5: Discussion

Unsupervised learning saves time and effort. However, by removing the human step we also remove the common sense aspect of a classification. For sentiment analysis, emoji's are hugely important and seem to be underutilised. Cryptocurrencies use very domain specific language, which these methods may not address.

Step 6: Conclusion

The project is still on going with further investigation to be carried out.

Key> The software used in project.