

## PROJECT SPECIFICATION

## Sentiment Analysis with Neural Networks

## Importing Twits

CRITERIA	MEETS SPECIFICATIONS
Length of Data	Print the number of twits in the dataset.

## Preprocessing the Data

CRITERIA	MEETS SPECIFICATIONS
Pre-Processing	The function <code>preprocess</code> correctly lowercases, removes URLs, removes ticker symbols, removes punctuation, tokenizes, and removes any single character tokens.
Preprocess All the Twits	Preprocess all the twits into the <code>tokenized</code> variable.
Bag of Words	Create a bag of words using the tokenized data.
Frequency of Words Appearing in Message	Remove most common and rare words by defining the following variables: <code>freqs</code> , <code>low_cotoff</code> , <code>high_cutoff</code> , <code>K_most_common</code> .

