Sentiment And Sentiment Analysis

Sentiment analysis (aka opinion mining) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials.
In other words it is the process of detecting the contextual polarity of text to find whether a piece of writing is positive, negative or neutral.

- Sentiment comes from Feelings (Emotion, Opinion)
  - Emotion (Joy, Sadness, Anger, Fear, Surprise, Disguise)
  - Opinion (A binary opposition, Good/Bad, Like/Dislike, For/Against)

- Sentiment Analysis is conducted using Non-Linear programming (NLP), Statistics or Machine Learning method to extract & identify the pattern of human sentiment and to characterize the sentiment behind the text unit.
- Text mining and finding out the emotion or opinion of the text unit is the common practice for Sentiment Analysis and often used in Call Center data, Reviews, Twitter extracts etc.

Process Flow & Categorization

Process Flow:

- Identify the source of data
- Establish the connection with the source using authentication methods
- Extract the data in the form of text, voice converted to text, image mapped to data form etc. Common data sources includes Twitter extracts, FB posts, Comments and Reviews on multiple sites etc.
- Perform Sentiment analysis using specified techniques
- Analyze and Present results as per business needs

Categorization Methods:

- Data Driven- Key word/ phrase Identification
- Smiley Recognition
- Likes Count
- Rating count method

Techniques and Functions of Sentiment Analysis

Techniques:

- Classification of emotions
- Polarity of words
- Text based sentiment classification
- Incorporate shallow linguistic/heuristic

Functions:

1. classify_emotion()
2. classify_polarity()

Functions of Sentiment Analysis

Functions:

- classify_emotion(textColumns,algorithm="bayes",prior=1.0,verbose=FALSE,...)

- Part of Sentiment package.
- Classifies the emotion (e.g. anger, disgust, fear, joy, sadness, surprise) of a set of texts using a naive Bayes classifier trained on specified emotions lexicon.
- Arguments

  textColumns - A dataframe of text documents listed one per row.

  Algorithm - A string indicating whether to use the naive bayes algorithm or a simple voter algorithm.

  Prior - A numeric specifying the prior probability to use for the naive Bayes classifier.

  Verbose - A logical specifying whether to print detailed output regarding the classification process.

  Additional parameters to be passed into the create_matrix function.
- Values

  Returns an object of class dataframe with seven columns and one row for each document.

Functions of Sentiment Analysis-Contd…

Functions:

- classify_polarity(textColumns,algorithm="bayes",pstrong=0.5,pweak=1.0,prior=1.0,verbose=FALSE,...)

- Part of Sentiment package.
- Classifies the polarity (e.g. positive or negative) of a set of texts using a naive Bayes classifier trained on Janyce Wiebe's subjectivity lexicon
- Arguments

  textColumns - A dataframe of text documents listed one per row.

  Algorithm - A string indicating whether to use the naive bayes algorithm or a simple voter algorithm.

  Pstrong/Pweak -A numeric specifying the probability that a strongly/weakly subjective term appears in the given text.

  Prior - A numeric specifying the prior probability to use for the naive Bayes classifier.

  Verbose - A logical specifying whether to print detailed output regarding the classification process.

  Additional parameters to be passed into the create_matrix function.
- Values

  Returns an object of class dataframe with four columns and one row for each document.

Steps Used for Performing Twitter Data Sentiment Analysis

- Load necessary R packages
- Create twitter App to get consumer key and consumer secret key
- Establish R and twitter connection for searching the required tweets
- Extract tweets containing the desired term
- Prepare the text for sentiment analysis of tweets
- Performing Sentiment Analysis
- Create data frame with the results and obtain some general statistics
- Plotting to view results

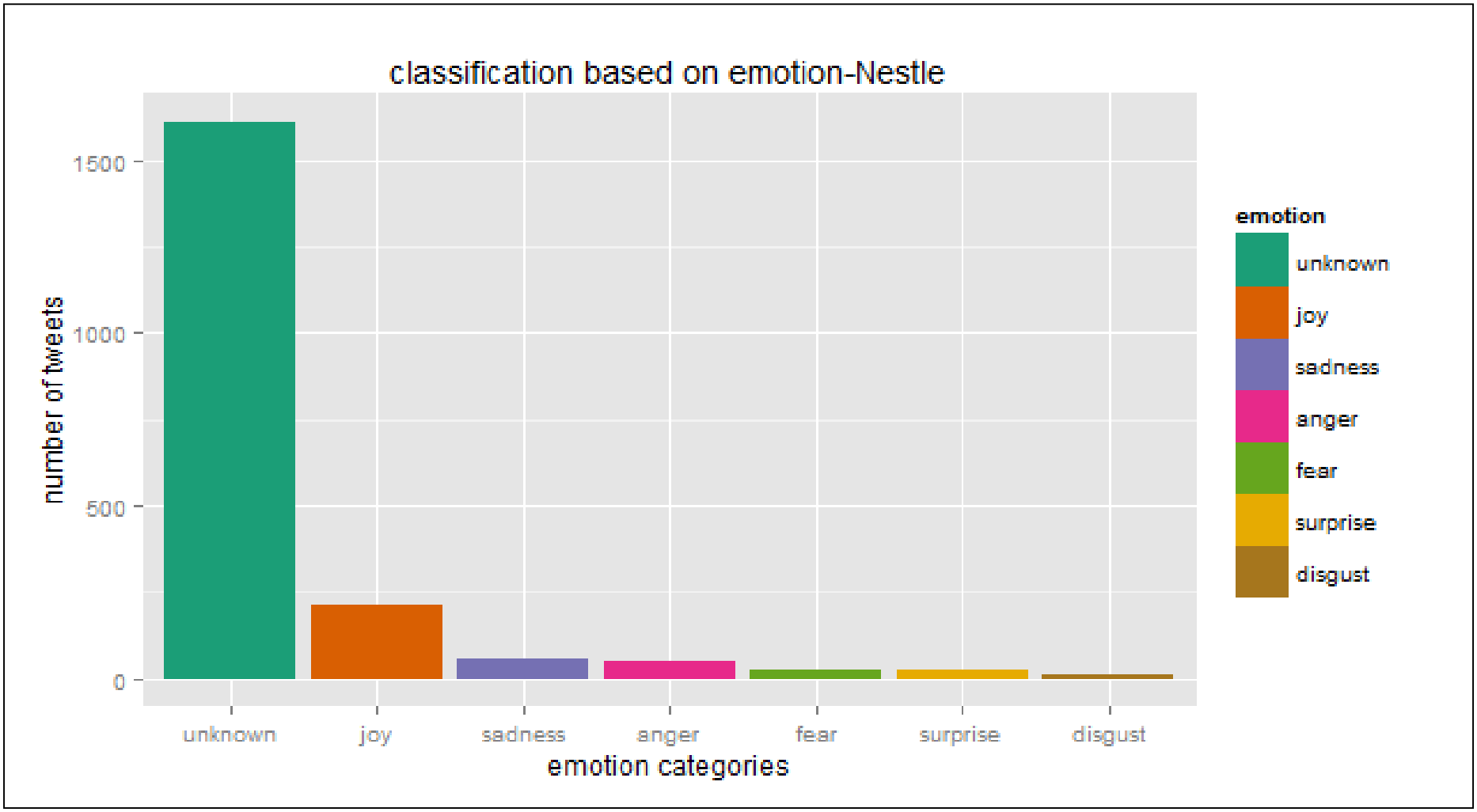
Sample Data Snapshot – Nestle Tweets as downloaded

ysis.R × | getTwitterOAuth.R × | new_sentiment_analysis.R* × | Twitter_sent_anlysis.R × | sent_df ×

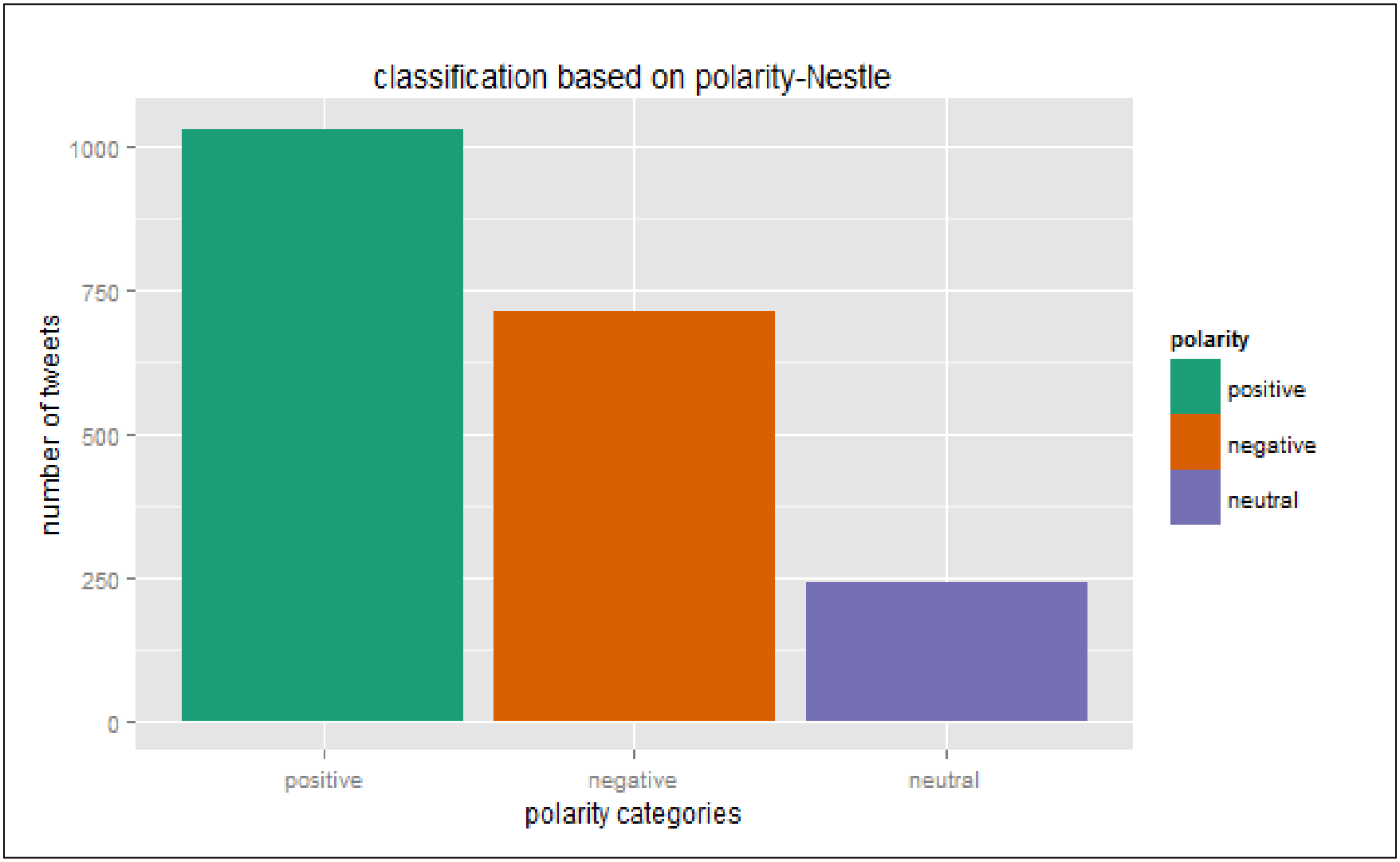
1987 observations of 3 variable

	text
1	i need someone to nestle with
2	indian insurer icici lombard may have to paymillion to nestle
3	helpsyoung people into work recruitment
4	where nestle erred byon
5	did you know nestl\u{FFFD} was a part ofother controversies by read
6	battle of the barsnestle amp cadbury fight over kit kat trademark
7	batches of indian goods including those made by hul britannia nestle haldiram fail to gain entry in us
8	kitkat fourfinger trademark thrown out by eu court in nestlecadbury chocolate waruk ukbiz
9	companies should own every single bit of water on the planet according toformer ceo and nowchairman peter
10	dominos effect post banning nestles maggie in indiahaldirams unfit for consumption says us fda\u{FFFD}
11	word cloud of tweets about maggi nestle in the lasthours
12	kitkat fourfinger trademark thrown out by eu court in nestlecadbury chocolate waruk ukbiz
13	great article fromon nicks campaign to getto stop bottling californias remaining water \u{FFFD}
14	problems likeleadpoisoned maggi noodles dont have to be solved just avoid them itisenough maggibanhtt\u{FFFD}
15	beauty tipsvivonex rtf elemental formulanestle health science new sealedcas\u{FFFD}
16	i think it is in reaction of nestle ban in india
17	dominos effect post banning nestles maggie in indiahaldirams unfit for consumption says us fda
18	then come to my house i get them what kind of milkydo ya want peak lait oribasagariount or nestle come to
19	great new glutenfree cornflakes inbut surely they could be cheaper after all less in them
20	is there anywhere is chicago that i can buy some maxibons if so where
21	nestle crunch
22	sub engit gackt x nestle \u{FFFD}super mario brothers u\u{FFFD}february th gackt gacktitalia http\u{FFFD}

Output-Emotions Cloud in the tweets



Output-Polarity Cloud in the tweets



Words Dispersion on Cloud in the tweets