A Review on Sentiment Analysis of Twitter*

IN4252 Web Science and Engineering

Nivedita Prasad†
Student ID: 4712099 Delft University of Technology
Delft, Netherlands
N.Prasad@student.tudelft.nl

ABSTRACT

The research community is showing profound interest in Sentiment Analysis (SA) which is conducted on a corpus of information. Therefore it is necessary to review the scientific publications made in this research area. The exponential growth of information in Twitter has developed a need for potential applications on opinion mining for understanding customer preferences in areas of businesses, recommendation systems for movies, research articles or music, systems that help gain followers for public figures. This paper aims to review and summarize the features, classification algorithms, techniques and results published in the past researches in the area of Sentiment Analysis of Twitter to help future research in the same.

KEYWORDS

Sentiment Analysis, Opinion Mining, Supervised learning, Unsupervised learning, Polarity

1 INTRODUCTION

Web is an active element where users of the web are generators and consumers of the information on the web. The information generated are generally via blog posts, status updates of their state of mind on social networks, opinion about the products they use, general information on any topic etc,. In this paper we analyze the impact of opinions or review of individuals published on social media have strong influences on decisions of potential customers of businesses who may want to buy the product or use the service. In the article [15] the author has cited that according to his study of on-line shopping, he showed that 81 percent of the users research about the product they are about to buy on-line and then later proceed on their state of mind that has been influenced by opinions. Hence businesses want to leverage the concept of Sentiment Analysis (SA) also known as Opinion Mining (OM) to analyze trends and have a competitive advantage by developing business intelligent applications. Similarly, SA provides business intelligent applications to recommendation systems as well; this is told in [28] that when labeling is provided through classification techniques, it provides users summaries of the reviews which contain normalized ratings.

*
†

,

The intuitive feeling created within a user to share more information has made sharing of information a habit. The social networks like Twitter [1] which is a micro-blogging website allows millions of users to publish their thoughts about any topic to share it with their 'followers'. Over a period of time this social networking site has become a source of meaningful continuous flow of information. This flow of information intrigued many researchers to concentrate on sentiment of the users to find more about their opinions on various topics ranging from business, politics, movies, electronic gadgets, current crisis scenarios and every little thing happening around them etc,.Hence Sentiment Analysis as a research area has developed interest in the community of natural language processing to concentrate on automatic text categorization. This problem faced by Sentiment Analysis is solved by the computational treatment of feelings and subjectivity of texts as cited in [23] or by machine learning techniques as told in [24]. This is because the spectrum of information is huge and manual search and selection of opinions on the web is strenuous. In this paper, I have presented a review of the relevant research done on sentiment analysis based specifically on the data retrieved from the micro-blog Twitter .

The rest of the paper is arranged as follows : Section 2 contains the technical details of twitter, Section 3 contains the Influence of Opinion, Section 4 contains detailed documentation on Sentiment Analysis : its importance, research and application , Section 5 contains Results and conclusions and finally section 6 contains a tabulation of all the research papers that have been reviewed in this paper.

2 UNDERSTANDING TWITTER

In the work of [21] , the author has said that, micro blogging sites acts like a platform where every user can share short messages of 140 characters , images, links to websites, graphical interface formats (GIF) which are short videos displaying an emotion, a poll which has multiple choices and can be live for a maximum of 7 days, geographical location of the post and an emoticon pallet which helps the user to display an emotion clearly without keeping a limit on the character length. The messages written by an user is read or commented with opinions by the user's 'followers'.The author states that popular users do not value the information but value popularity by the number of followers as a contrary to the importance we give to the information. As cited in [2], since the day of inception in the March 2006 the number of tweets have increased from five thousand to 628 million and counting tweets per day. According to the [3], analytics of twitter growth showed that there are 328 million monthly active users as per quarter one of 2017 .These statistics is to reiterate the fact that the research conducted in this field on sentiment analysis is indeed important as it gives a real sentiment or feeling of the millions of users. According to [4] users update messages with limit of 140 characters called "tweets" where the user's thought about a topic is shared as written messages. Live-tweet happenings such as Olympics or television series has become a way for users to be involved with other users on-line while sharing their thoughts on current experiences. When a user refers to another user with @username it is called "mention". These references of user-names are visible to author of tweets and his followers. The "What's Happening?" text is displayed on the area where a tweet can be created acts like a cue which instigates the user to share messages of what was happening around them which has original sentiment or feeling. The tool of re-tweet helps the user to re-tweet interesting tweets of the people they follow and make it visible to their followers. In order to streamline the information flow, "hash-tags" a way of labeling a topic followed. Here the character "#" is followed by "topic-name" (#topic-name).The trending topics (TT) in the current period of time is used with hash-tag. Twitters success mainly depends on the businesses that value the information posted on twitter related to them. The businesses can understand if the public has original good or bad opinion about their products and services.This is where Sentiment Analysis of tweets plays a major role.

3 INFLUENCE OF OPINION

Opinion is judgment or a perspective formed on any topic not necessarily based on facts or knowledge by an user. In [21] the author cites that, Opinions can be considered as word of the mouth but in terms of written text in social media. This can be acronym-ed as EWOM or OWOM i.e electronic word of mouth and on-line word of mouth respectively. We trust the opinion more when it is shared by people of our innermost circle. He says that in twitter we have a wide variety of opinions on many topics and hence the necessity of a tool that monitors the opinions of their consumers was necessary and leverage the same for a competitive advantage against their competitors. In [16] the importance of electronic work of mouth has been discussed. In their research findings they have proved that 19% of micro-blogs contain mention of a brand. Of the branding micro-blogs, the statistics shown by the author is that nearly 20% contained some expression of brand sentiments. Of those 20%, more than 50% were positive and 33% were critical of the company or product. The truth of these opinions are often a question of discussion. This is because, it could have been faked by competitors giving fake reviews. The information credibility of the tweets are discussed in [9], here they have classified tweets as news and subjective based on which the new tweets have been classified as true and false.They have shown statistics in the classified the tweets with a precision and recall in the range of 70% to 80%. Their main hypothesis was that there are signals available in the social media environment itself which enable users to assess information credibility. They have trained their learning algorithms considering subsets of features. The 4 subsets were (1) Text subset which are characteristics of the text of the messages consisting 20 features (2) Network subset containing the characteristics of network of users consisting 7 features (3) Propagation subset containing propagation features, re-tweets and total number of tweets consisting 6 features (4) Top element subset has most frequently used hash-tag, user mention with 4 features. The corpus of data required for SA research is mainly acquired using the Twitter API [5] for the research papers.

4 SENTIMENT ANALYSIS

Sentiment is defined as the view or opinion that is held or expressed or the self-indulgent feeling of tenderness or nostalgia about an issue [6]. There are six "universal" emotions [13] such as anger, disgust, happiness, surprise, sadness and fear. These emotions are displayed as the sentiment of the text the user has posted in twitter. In simple words, The computational treatment of sentiment, opinion and subjectivity in text is known as Sentiment Analysis. According to [7] Sentiment Analysis (SA) also known as Opinion mining or emotional AI refers to the use of natural language processing, text analysis, computational linguistics and machine learning techniques to identify and extract the subjective information. SA determines the attitude of user with respect to some topic or the overall contextual polarity. The attitude may be a judgment or evaluation, the emotional state or the emotional effect of the user. The sentence be subjective or objective; as an end result we want to analyze what other people think. For better understanding let us look at an example of subjective sentence: "This Laptop is good" whereas, "Long battery life" is objective.

4.1 Types of Approaches in SA

In, [23] the author has cited a range of approaches for different types of corpus of data extracted from twitter.

4.1.1 Polarities in Sentiment. The task of classifying a text as negative or positive opinion is called sentiment polarity classification. The opinions are classified between the negative, neutral and positive polarity labels in sentiment-related classification, regression or ranking. The context of reviews can be "thumbs up" and "thumbs down" or evaluative like "like" and "dislike" . This can be used to classify a movie review or give predictive opinions in forums that discuss about which team is likely to "win" or "lose" elections. In [18] the authors suggest that "pro" and "cons" expressions differ from positive and negative polarity and they help us to extract on why the reviewers liked or disliked a product and we can give supported reasons for evaluative judgments. Data is otherwise classified based on the comparative degree of the polarity and additionally each of the neutral (lack of opinion), positive and negative classes has its own distinct vocabulary.

4.1.2 Detection of Subjectivity and Opinion. The problem of identifying subjective texts impacting sentiment classification is more difficult than polarity classification. The adjectives appearing in the sentence are judged based on their orientation and then the subjectivity of the sentences are graded.

4.1.3 Topic based Filtering. In [26], the author cites that Topic based filtering and subjectivity filtering are complementary as the corpus of data are based on topic of discussion.

4.1.4 View point classification. Here the subjectivity of texts are classified generally and not concentrated on a single topic or subject. Here binary and n-ary classification are used which classifies bundled attitudes and beliefs as cited in [23].

4.2 Features Used for Classification

As cited in [23] a piece of text is converted into a feature vector or any other representation for classification process is one of the primary steps in data driven approaches to text processing. Feature selection is vital for machine learning techniques and other tailored approaches of specific classic information extraction and categorization.

4.2.1 Importance of Presence of term versus its Frequency. In [24] the authors have proved that the performance is better using "presence of words" than frequency of words. They have used binary-valued feature vectors which indicates where the entry of a word in the text is present in the class or not, by assigning values of 1 and 0 respectively. This method was better than polarity classification. The author has cited that importance of a topic's sentiment may not be because a keyword that is repeated many times in the message.

4.2.2 Position Information. The position of a term at the middle or the end of the document can have impact on the over all sentiment or subjectivity of the tweet or message as mentioned in [18]. Thus position information is considered as feature in the feature vector.

4.2.3 Uni, Bi or Trigram. The n-grams are uni,bi and tri grams. In [24] they report unigrams outperforms bi-grams when classifying movie reviews by polarity classification. In [11] they have proved that bi-grams and trigrams yield better polarity classification. And in, [25] subsumption hierarchy is defined where they use complex features for opinion analysis.

4.2.4 Negation. Handling Negation is an important concern in opinion mining. A simple word "not" differentiates a sentence into another class even though it has same set of words that could classify it into positive polarity. Sarcasm and irony are forms of negation that are highly difficult to detect. The word "avoid" acts as a polarity reverser.

4.2.5 Parts of Speech. The parts of speech such as adjectives and adverbs have high correlation between themselves [14]. Adjectives are a good indicators of sentiment and works well for unsupervised classification. In contrast [24], the authors found that unigrams perform better than adjective as features.

4.3 Research on Twitter Sentiment Analysis

The first research on twitter was related to sociology than computer science. Amongst the first few articles published, particularly in [17], the author has cited the topological and geographical properties of social network like twitter is a scale free network displaying small-world phenomenon and follows power law distribution for number of followers made and updates done. In [10], the author has cited that sarcastic posts which seem opposite of what is actually means to convey about the topic of discussion are growing in number. The study of "unfollowing" another user occurrence is strongly related to social concepts like homophily which is friendship, common interests, attention status of the user is discussed in [19].

The sentiment analysis tool is built with some challenges. In the review paper of [23], the authors have discussed the building of OM

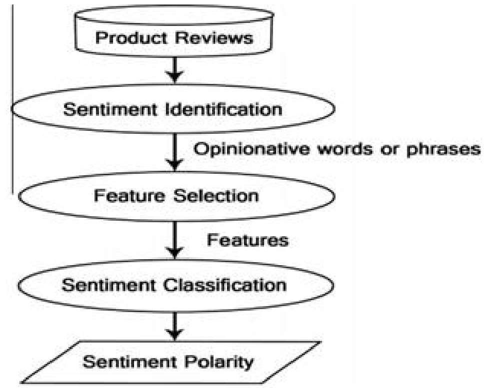

Figure 1: General Process of Sentiment Analysis

or SA system. The first challenge is to find important and prevalent information needed and if the application is integrated with general purpose search engine. Here the queries will have indicator terms such as "review(s)", "opinions" ; the users of this application can use check-box to find reviews is desired. Query classification in general is difficult. The second challenge is to find which message or which portions of the message contains subjective information. The third challenge is to identify the overall sentiment expressed by the tweet regardless of the particular features and aspects of the topic under discussion. The final challenge is to present the sentiment information gathered in a summarized manner. This final challenge involves clustering of votes which may be in the form of stars or letter grades, highlighting selective opinions, points of consensus and disagreement, identifying communities of opinion holders and the corresponding different level of authority in the community. Figure 1 shows the general process adopted for sentiment analysis. The steps of the figure 1 is explained as follows: product reviews are collected from twitter using its API, then the sentiment of the sentence are to be identified, the features have to be selected and then classified by using one of the approaches as mentioned in Section 4.1 and finally find the polarity of the new incoming tweets.

In [27], the author has suggested approaches to identify a particular subjective text which displays abusive emotion (flames) which are often unwanted by the users. The prototype Smokey builds 47 elements based feature vector, that is based on syntax and semantics by combining sentences of each message. The author has developed a training set of 720 messages and has supplied this as input to Quinlan's C4.5 decision-tree generator to determine feature based rules that were able to correctly categorize 64% of the flames and 98% of the non-flames in a separate test set of 460 messages according to the statistics of his paper. He has used to approach of converting all the messages into one format of one sentence per line with delimiters, then the text is run through a parser of Microsoft's NLP, then the output of the parser is converted sed and awk scripts into Lisp expression and the s-expressions are used to generate feature vector for each message and finally the output is evaluated with simple rules produced by the decision tree generator C4.5. Here the messages were classified as "okay", "maybe" and "flames". The "okay" messages were discarded as they tend to mislead the decision tree generator. The author has used supervised classification method to classify the texts which was very successful and also experimented with linear regression which was less successful. For example, the ordered rules of C4.5 are the generated rules made of "if" statements with conditions of imperativeness, condescension-somewhat, site specific-insult which are joined by an OR operator and AND-ed with other similar conditions. The limitation of Smokey was it could not detect unusual text messages with unnecessary spaces which in-turn could not trigger rules.

In [12], the author has cited that the sentiment value greatly enhances the information extraction tasks like review summarization. He says that majority of the sentiment extraction is based on polarity identification and pre-specified mood labels. The SA system is manually supplied with sentiment words or tailored sentiment specific dataset which is created using user-defined hash-tags that are added to the user's tweets in twitter which in-turns helps in the easy identification and classification of incoming tweets. The author has proposed classification of wide variety of sentiment types from text by using "tagged" twitter data. The SA system designed is supplied with 50 twitter tags and 15 smileys as sentiment labels that will be used to train a classifier, that classifies dozens of sentiment types of the tweets. Using this developed SA system, untagged tweets are classified based on their sentiment types. The author has also explored the dependencies and overlap between the different sentiment types represented by smileys and twitter tags. The four features used for this classifier are punctuation, single word features, n-grams and patterns and finally their contribution is evaluated for sentiment classification. Here each word appearing in the tweet acts a binary feature or a sequence of two to five words as binary n-gram feature both of which is given weight equal to inverted count of the word occurrence in the whole of the twitter corpus extracted. The n-grams and rare words have higher weight than common words. Words or n-grams appearing in less than 0.5% in the training set were not considered as feature and the consecutive punctuation symbols which are ASCII smileys are used as single word features. Pattern-based feature is the main feature type of this paper. The words are classified into High Frequency words (HFW) and Content Words (CW). The HFW are punctuations, URL, REF and hasghtags. For the pattern to be meaningful the start to end of the pattern should be HFW. Many patterns are bound to match and overlap and patterns with less than 0.5% is not considered as a feature. Since exact pattern match is not common, the concept of partial match is leveraged to reduce sparsity in the pattern matches. This feature is useful in sarcasm detection. The punctuation feature contained sentence length in terms of words, number of "?", "!", quotes, capital words divided by the average maximum number of times of observance of the same word. The maximum weight of each feature is equal to average weight of the single pattern/word/n-gram feature. The classification algorithm used is k-NN (k Nearest neighbors). Here each example in the training and test set is a feature vector. The euclidean distance is found between the new feature vector of the test set and the feature vectors of the training set and a matching vector is one which shares a patter/n-gram/word feature with the new feature vector. The dataset used consists of 475 million tweets provided by Bredan O'Connor. The evaluation is done via cross-validation.

According to [8], the author has introduced part of speech and polarity combined features and tree kernel classification which uses old and the proposed new features in the paper, making the model perform better than the baseline unigram model. There are 2 models built where the first one is a 2 class classification of positive and negative classes and the second is a 3 class classifier of positive, negative and neutral classes. The models used are unigram using 10,000 features, feature based model using only 100 features which are assigned a value which may be a natural or real or boolean value and tree model performing better than previous models. The combination models of new features with tree and unigram separately yielded 4% better performance than baseline unigram model. The dataset used is tailored twitter data that are collected in streamline fashion. The preprocessing of tweets are done using an acronym dictionary developed by the authors consisting of 5184 acronyms eg : "lol" is translated as "laughing out loud" and an emoticon dictionary which labels emoticons and additionally a stop word dictionary is used for identifying the stop words and lastly Wordnet dictionary of 8000 words created previously by the authors that assigns a pleasantness score for polarity and discards any word with a score lesser than 0.5. The tree kernel is a combination of many features in tree representation, which uses recursive calculation to compare subtrees. Here the whole tweet is tokenized as emoticon, punctuation marks, negation and a leaf node is added to the root node. The features are first classified as polar and non polar features and then further subdivided as parts of speech and other types of features. The 3 class classifier, 2 class classifier, unigram, tree kernel, 100 senti features, combination of kernel and senti features and the combination of unigram and senti features models all use Support Vector Machine (SVM) classifier and use 5-fold cross validation for evaluation. In the 5-fold test the tree kernels outperform the unigram and the Senti-features by 2.58% and 2.66% respectively. The combination of unigrams with Senti-features outperforms the combination of kernels and Senti-features by 0.78% as per the statistics mentioned in the paper. The polarity classification SA system developed by the authors for the positive versus negative classes performs 4.04% more than the base line unigram model.

In [22], the corpus consists of positive (happy), no sentiment or neutral emotion and negative emotions (sad) retrieved using twitter API for popular newspapers and magazines. For objective texts they collected data from 44 newspapers. The corpus is validated using Zipf's Law. The features are extracted via filtering, tokenization, removing stopping words and finally constructing n-grams out of consecutive words which plays a crucial role in accuracy. The authors have experimented with SVM, CRF, NaÃŕve Bayes classifier using n-gram and Bayes classifier using parts of speech. The accuracy of the classifiers are improved using entropy and discarding the common n-grams not contributing a sentiment and by calculating the log likelihood. Hence accuracy is the fraction of correct classifications out of all the classifications. The Naive Bayes Classifier has yielded the best result using n-gram and parts of speech tags as features.

In [29], the authors have designed a new algorithm called "SentiStrength" that simultaneously extracts the positive and negative sentiment from short messages. The algorithm uses a dictionary of sentiment words which is associated with sentiment strength measures and uses a range of non-standard spellings and common textual methods expressing sentiments. The algorithm was trained using a initial set of 2600 human classified MySpace comments and then evaluated on 1041 MySpace comments. The author has suggested a machine learning approach that optimizes the sentiment term weight and method to extract sentiment from frequently used non-standard spelling and other methods of textual writing. The paper has put forward the dual 5-point system for positive and negative sentiment. The author has cited from reference that sentiment is of 2 axes, that is arousal (high to low) and valance (positive to negative).Hence sentiment strength evaluation is a combined scale of positive and negative sentiment. In the dataset, 20 words and 101 characters per comment on average were selected and the sentistrength algorithm judged the positive and negative sentiments from 1 till 5 , giving 1 to no positive or negative emotion and 5 very positive or negative emotion. This algorithm could correct spelling mistakes, ignoring negative emotion in questions, giving minimum strength to frequent punctuation and to sentence with exclamation mark, emoticon list with associated strengths, booster word list with that increases emotion strength, a negating word list inverting subsequent emotional words. The accuracy of SentiStrength algorithm is 2% more than the baseline algorithm. The main success to sentistrength is the procedures it uses to decode non standard spelling and boosting the strength of words.

In [20], the authors have proposed an opinion analysis system to compare consumer opinion to find out the strength and weakness of each product in terms of various product features. The Opinion Observer prototype designed by the authors gives a visual aid to the customers and product manufacturers by displaying side by side and feature by feature comparison of consumer opinions of the products. Additionally the authors have proposed language pattern mining, a technique to extract product features from the pros and cos of the reviews. The visualization helps the user to clearly see which variety of the same product category is superior. According to the author the reviews are available in 3 different formats: pros and cons; pros, cons and detailed review; free format. From the first two formats the positive and negative opinion is easily known as pro is the positive and con is the negative opinion. Their approach is based on an important observation made by the authors themselves. The observation was that "Each sentence segment contains at-most one product feature and the sentence segments are separated by comma, full-stop and conjunctions "and" and "but" ". For example the features obtained for a sentence is as follows: "great photos" will give the feature "photos", "easy to use" gives the feature "use", "good manual" gives the feature "manual", "battery usage" gives the feature "battery" etc.,. The authors have noted that there are implicit and explicit features, synonyms and granularity of features. The authors do a five step pre-processing and labeling. The steps include Parts of speech tagging and discarding numerals, replacing actual feature with general names, using n-grams to reduce segment length, distinguishing duplicate tags and finally perform word stemming. The output is saved for the generation of rules for the supervised pattern matching process and then association rule mining is done to classify sentiments. In the interface of the Opinion Observer developed by the authors, there is an option to read the whole review or retrieve the reviews. The automatic pattern finding is adopted by the authors to extract reviews. The authors have achieved 100% precision and 52% recall with their system and does not work well with context dependent synonyms.

4.4 Application of Sentiment Analysis

In [27], detection of "flames" (abusive,hostile and opposing language) in email and other types of communication, is an application of subjectivity identification and classification. Another example is, review summarization (Give me the positive reviews of product X or show me the articles which explains why movie X is boring) as cited in [12] and Visualizing opinions using opinion observer as cited in [20]. The application of SA in [29] was finding Movie Popularity from multiple on-line reviews and parts of vehicle liked by its owners.

5 CONCLUSION

Although the above discussion has shown the most relevant works in the field of Sentiment Analysis in Twitter, there are some problems that are not resolved yet. Some of them are data sparsity, bad grammar, corpus of all jargons to be used for training, languages other than English for Sentiment Analysis. As conclusion SA is a rapidly expanding area of research and it is necessary to develop SA systems that will explicitly boost business through the development of business applications which indicate ares of improvement , achievements, etc., using the knowledge given by twitter. A tabulation of the review done is provided in Section 6.

6 TABULATION OF STUDIES

The **Figure 2** is a summary of all the research papers reviewed in this paper for clear understanding. The columns used are authors, title, approach or methods, queries posed, features used, application and finally accuracy of each research methodology based on the evaluation process adopted.

REFERENCES

[1] [n. d.]. ([n. d.]). https://twitter.com/
[2] [n. d.]. ([n. d.]). http://www.internetlivestats.com/twitter-statistics/
[3] [n. d.]. ([n. d.]). https://www.statista.com/topics/737/twitter/
[4] [n. d.]. ([n. d.]). https://en.wikipedia.org/wiki/Twitter
[5] [n. d.]. ([n. d.]). https://developer.twitter.com/en/docs/api-reference-index
[6] [n. d.]. ([n. d.]). https://www.google.nl/search?q=opinion+meaning&rlz=1C1NHXL_enNL762NL762&oq=opinion+&aqs=chrome.2.69i57j0l5.7785j0j7&sourceid=chrome&ie=UTF-8
[7] [n. d.]. ([n. d.]). https://en.wikipedia.org/wiki/Sentiment_analysis
[8] Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. 2011. Sentiment analysis of twitter data. In *Proceedings of the workshop on languages in social media*. Association for Computational Linguistics, 30–38.
[9] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web*. ACM, 675–684.
[10] Cristian Danescu-Niculescu-Mizil, Michael Gamon, and Susan Dumais. 2011. Mark my words!: linguistic style accommodation in social media. In *Proceedings of the 20th international conference on World wide web*. ACM, 745–754.
[11] Kushal Dave, Steve Lawrence, and David M Pennock. 2003. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of the 12th international conference on World Wide Web*. ACM, 519–528.
[12] Dmitry Davidov, Oren Tsur, and Ari Rappoport. 2010. Enhanced sentiment learning using twitter hashtags and smileys. In *Proceedings of the 23rd international conference on computational linguistics: posters*. Association for Computational Linguistics, 241–249.
[13] Paul Ekman, Wallace V Friesen, and Phoebe Ellsworth. 2013. *Emotion in the human face: Guidelines for research and an integration of findings.* Elsevier.
[14] Vasileios Hatzivassiloglou and Janyce M Wiebe. 2000. Effects of adjective orientation and gradability on sentence subjectivity. In *Proceedings of the 18th conference on Computational linguistics-Volume 1*. Association for Computational Linguistics, 299–305.
[15] John A Horrigan. 2008. Online shopping. *Pew Internet & American Life Project Report* 36 (2008), 1–24.

,

Serial number	Authors	Title	Approach / Method	Query	Model Used	Features Used	Application	Accuracy
1	Ellen Spertus	Smokey: Automatic Recognition of Hostile Messages	Supervised Classification using NLP	Imperative statements, Noun phrases as appositions,second person rules, profanity, condescion, insults , epithet, polite, praise,Misscellaneous	Decision-tree generator using "if" statements	47 element feature vector	Flame recognition : Smokey	Categorize 64% of the flames and 98% as nonflames in a test set of 460 messages
2	Davidov, Dmitry and Tsur, Oren and Rappoport, Ari	Enhanced sentiment learning using twitter hashtags and smileys	Supervised Sentiment Classification	"Give me all positive reviews of product X/ Show me articles which explain why movie X is boring"	k nearest neighbour (k-NN)	Punctuation, Words, n-grams, Patterns	Recommendation Systems	Identify sentiment types of untagged sentences
3	Agarwal, Apoorv and Xie, Boyi and Vovsha, Ilia and Rambow, Owen and Passonneau, Rebecca	Sentiment analysis of twitter data	Supervised Classification using polarity	Not Applicable	Support Vector Machine (SVM) for 2 class, 3 class, unigram, tree kernel, 100 sent feature, kernel plus senti-feature, unigram plus senti-features models	100 features known as Senti-features which comines parts of speech and prior polarity of words	Not Applicable	Combination of new features and tree model and combination of new features and tree kernel model outperforms unigram baseline by 4%
4	Pak, Alexander and Paroubek, Patrick	Twitter as a corpus for sentiment analysis and opinion mining	Supervised Classification using polarity	Not Applicable	SVM, Conditional Random Fields (CRF), Naïve Bayes classifier, using n-gram, Bayes classifier using parts of speech	n-gram as binary feature	Not Applicable	Naïve Bayes based on Bayes theorem yeilded best result
5	Thelwall, Mike and Buckley, Kevan and Paltoglou, Georgios and Cai, Di and Kappas, Arvid	Sentiment strength detection in short informal text	Sentiment Strength Detection algorithm	love-u-! ,do you listen to rock music? , you rock!!!,whats up? , Just showin love 2 ur page	Comparing sentistrength and machine learning	n-grams of length 1-3 consisting of all terms extracted from the text, including emoticons, spelling-corrected words, repeated punctuation, question marks , exclamation marks	Movie Popularity from multiple online reviews, Parts of vehicle liked by the owners, differnetiaing between emotional and informative social media content	Sentistrength2% more accurate than baseline
6	Liu, Bing and Hu, Minqing and Cheng, Junsheng	Opinion observer: analyzing and comparing opinions on the web	Natural Language Processing and supervised pattern discovery	"great photos" will give the feature"photos", "easy to use" gives the feature "use", "good manual" gives the feature "manual", "battery usage" gives the feature "battery"	Opinion Observer prototype , language pattern mining for extracting product features from the pros and cos of the review	Parts of speech tagging and discarding numerals, replacing actual feature with general names, using n-grams to reduce segment length, distinguishing duplicate tags and finally perform word stemming	Visualizing the strengths and weakness of feature by feature of products useful for both customers and product manufacturers that helps in decision making	Out performs existing methods with 100% precison and 52% recall.

Figure 2: Research in Sentiment Analysis

[16] Bernard J Jansen, Mimi Zhang, Kate Sobel, and Abdur Chowdury. 2009. Twitter power: Tweets as electronic word of mouth. *Journal of the Association for Information Science and Technology* 60, 11 (2009), 2169–2188.

[17] Akshay Java, Xiaodan Song, Tim Finin, and Belle Tseng. 2007. Why we twitter: understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. ACM, 56–65.

[18] Soo-Min Kim and Eduard Hovy. 2006. Automatic identification of pro and con reasons in online reviews. In *Proceedings of the COLING/ACL on Main conference poster sessions*. Association for Computational Linguistics, 483–490.

[19] Funda Kivran-Swaine, Priya Govindan, and Mor Naaman. 2011. The impact of network structure on breaking ties in online social networks: unfollowing on twitter. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1101–1104.

[20] Bing Liu, Minqing Hu, and Junsheng Cheng. 2005. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web*. ACM, 342–351.

[21] Eugenio Martínez-Cámara, M Teresa Martín-Valdivia, L Alfonso Urena-López, and A Rturo Montejo-Ráez. 2014. Sentiment analysis in Twitter. *Natural Language Engineering* 20, 1 (2014), 1–28.

[22] Alexander Pak and Patrick Paroubek. 2010. Twitter as a corpus for sentiment analysis and opinion mining.. In *LREc*, Vol. 10.

[23] Bo Pang, Lillian Lee, et al. 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval* 2, 1–2 (2008), 1–135.

[24] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*. Association for Computational Linguistics, 79–86.

[25] Ellen Riloff, Siddharth Patwardhan, and Janyce Wiebe. 2006. Feature subsumption for opinion analysis. In *Proceedings of the 2006 conference on empirical methods in natural language processing*. Association for Computational Linguistics, 440–448.

[26] Ellen Riloff, Janyce Wiebe, and William Phillips. 2005. Exploiting subjectivity classification to improve information extraction. In *Proceedings of the national conference on artificial intelligence*, Vol. 20. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 1106.

[27] Ellen Spertus. 1997. Smokey: Automatic recognition of hostile messages. In *AAAI/IAAI*. 1058–1065.

[28] Loren Terveen, Will Hill, Brian Amento, David McDonald, and Josh Creter. 1997. PHOAKS: A system for sharing recommendations. *Commun. ACM* 40, 3 (1997), 59–62.

[29] Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. 2010. Sentiment strength detection in short informal text. *Journal of the Association for Information Science and Technology* 61, 12 (2010), 2544–2558.