International Journal For Technological Research In Engineering
Volume 4, Issue 9, May-2017 *ISSN (Online): 2347 - 4718*

EMOTION RECOGNITION: A LITERATURE SURVEY

Shilpi Goyal[1], Nirupama Tiwari[2]
RGPV Bhopal India

As one of the most prosperous applications of text analysis and understanding emotions(sentiments) and short messaging text apperception has recently received consequential attention especially during the past several years. This is corroborated by the emergence of Web 2.0, social networking services, micro blogging, blogs, chats, online reviews, forums, discussions and systematic evaluations of text analysis(emotion analysis) techniques. There are five major aspects for this trend: first is the wide range of commercial and social marketing applications, second is the understanding of one's feelings (sentiments), third human-computer interaction , fourth text to speech generation and fifth the availability of natural language and machine learning approaches and technologies after 20 years of research. This paper provides an up-to-date critical survey of emotions, emoticons and short messaging text research.

• Information systems → Information Retrieval → Retrieval tasks and goals → Sentiment analysis

I. INTRODUCTION

Emotions are now extensively studied in the area of psychology, computer sciences, neuro-science, and cognitive sciences as they play significant role in human nature. Most of the devotion of researchers in computer science is in the area of textual emotions, particularly in the field of recognition of emotions. Emotion reveals itself in the form of facial expressions, vocal expressions, writings, and in movements and actions. Subsequently, scientific research in emotion(sentiments) has been followed along multiple dimensions and has drawn upon research from several fields. Mainly used form of communication on social network is in textual form, contributed a platform for computer systems to behave more smartly based on the user's feelings. Enormous amounts of text data are available in the form of blogs, micro blogging sites like Facebook, Twitter, emails, SMS etc. This textual data is beneficial to generate better human interaction system which needs to be able to analyze the text and conclude the emotion of the user. Even though the system can discover the user's emotional states, intricacy of language makes it hard for researchers to distinguish emotional states from pure textual data.

Table 1 Typical applications of textual emotions.

Areas	Specific applications	Sources
Sentiment Analysis / Opinion Mining	Focus: Information retrieval & knowledge discovery from text. Goal: To make computer able to identify & express emotions. Application: Companies are concerned about consumer's opinion about products and their services, issues and events to find the worth choices.	chats, web forums, blogs, discussion groups, tweets, reviews
Text – to – Speech Generation	Focus: classify the emotional matching of sentences in the storyline text, for appropriate communicative representation of Text- to-Speech synthesis. Goal: In spoken communication, spokespersons well express emotions by modifying mode of speech or communication, including intensity, pitch & durational signals. Application: A machine can read for us as human's read.	Text, emotional state, emotions
Human Computer Interaction	Focus: machine learning techniques,natural language processing. Goal: artificial intelligence, robotics, psychology blogs, product reviews, CRM and service oriented companies, customer emotion. Applications: automatic answering systems, dialogue systems, and human like robots.	Multi-Language Text.
Emoticons(Emotion Icons)	Focus: Multilingual Sentiment Analysis. Goal: To find emotions in foreign language text Applications:as language is not understandable, still then we are able to fetch the emotions behind that text.	Text containing symbols, punctuation marks, etc
Computer Assisted Creativity	Goal: The automated generation of evaluative expressions with a bias on certain polarity orientation. Application:automatic personalized advertisement and persuasive	Text on search engine

	communication.	
Recommendation systems	Focus: information retrieval and find the valence whether it is positive, negative or neutral. Goal: a system not recommending items that have an abundance of negative feedback. Application: recommend items whether to purchase or not.	Feedback, comments, reviews
Short Messages	Focus: spelling correction Goal: convert it to plain text. Application: easy to apply algorithms to find emotions.	Chats, reviews, tweets

A general statement of the problem can be formulated as follows: Given a line of text from any online source, identify emotions from that text using some standard algorithms. Available collateral information such as subjectivity, valence, granularity, context may be used to enhancing recognition. The solution to the problem involves transformation of short texts (if any) to plain text, tokenization, identification of emotion words using categorization, calculate the emotion of given message based on dictionary (lexicon, graph) using some classifier using in that context.

II. ISSUES RELEVANT TO EMOTION RECOGNITION

Difficulty and Intricacy of Language Use

—Text Informality: Users write tweets, messages, chats in social media containing grammatical mistakes.

—Acronyms: social networking sites have a lot of abbreviations that are publicly popular and used by most social users. Unfortunately, they are increased day to day in text data, changing dramatically, make the automated system difficult.

—Combination of Languages: Hindi-English is a language used by Indians, to express their feelings or texts with their friends.

—Emotion Icons: Most of social networking users make use of graphical symbol that states specific emotion, ":)" and ":(" are examples of such icons.

—Applicability: Researchers still tackle with the problem of non related topic like advertisements etc.

—Negation and Repetition of terms: Certain terms like negations and modals impact emotion of the sentence, without having strong emotion relations. For example, was good, was not good, and may be good interprets same emotion and executed differently.

—Words used in different contexts convey different emotions.

—statement may carry more than one emotion (and to varying degrees). There exist a gap between multiple object entities.

—statement express emotional mood of user without implicitly or explicitly present. Sense of text conveys negative or positive emotions without explicitly stated. For example: Again the Monday has come. shows frustration.

Creative and Non-Standard Language Uses

—It is difficult to interpret creative uses of language for Automatic natural language systems such as sarcasm, paradox, funniness, and simile. However, these are common in language use.

— Texts in Social media are mundane with terms not present in dictionaries like misspellings (gud), jargon of letters (sooooo), emoticons, abbreviations (143), hashtagged words (#TeenChoice), etc. Most of them convey emotions.

Lack of Massive Amounts of tagged Data

—Machine learning algorithms for sentiment analysis involves significant amounts of training data. There are numerous set of emotions that humans can recognize and express. So to recognize emotions we have to use limited resources of some valence categories and pre defined set of emotions.

Subjective and Multicultural Differences

— Recognize emotions in text can be hard even for humans as humans perspicacity is circumscribed to some languages, gestures etc. Most of the research elaborate that the amount of acquiescent between taggers is lower to identify valence or emotions, as compared to tasks such as identifying part-of-verbalization and identifying entities.

— There can be major dissimilarities in emotions associated with events and demeanors across different cultures. For example, imbibing may be considered as more negative in some components of the world than in others.

A complete review of relevant studies in the field of Emotion Recognition on social networking sites is illustrated using a timeline from late 1990's to till now.

III. TIMELINE OF EMOTION RECOGNITION

Timeline from early 90's to 1992: Researcher developed systems that are capable of manually excerpt sentiments from the text. They gave various universal models to identify the emotions on the basis of different dimensions and valence. This is the time where emoticons are recognized in digital world.

1966	General Inquirer system	[Stone et. al.]	first milestone for extracting textual emotion. Input texts are matched with manual database to recognizegnize their class such as positive, pstv, negative, feel, vigorous, puissance, impotent, pleasure
1970	set of emotions	[Ekman]	[Ekman] defined six rudimentary emotions: joy, sadness, anger, fear,

 ISSN (Online): 2347 - 4718

			disgust and surprise.
1975	set of emotions	[Osgood et. al.]	understanding emotion expression in text, used multidimensional scaling to visualize the affective words to compute kindred attribute ratings. The dimensions were “evaluation”, “potency” and “activity”.
1979	Russells circumplex model	[Russell]	which utilizes the dimensions of arousal and valence to identify 150 affective labels.
1980	set of emotions	[Pluchik]	Eight basic emotions: anger, sadness, disgust, fear, surprise, anticipation, joy and trust.
1982	first Emoticon	[Fahlman et. al.]	Emoticons are apperceived in the digital era and [Fahlman et al.] proposed to utilize :¬) and :¬(to differentiate jests from more earnest posts.
1987	Hand Crafted Models	[Dyer]	[Dyer] use models to understand particular text deeply in order to mine for emotions. These systems are involute and their results are arduous to simplify to other texts.
1987	set of emotions	[Shaver et. al.]	shows how the prototype approach is useful to investigate the processing of information about emotional events, cross-cultural differences in emotion concepts, and the development of emotion knowledge.
1888	set of emotions	[Frijda]	Define emotions based on some laws.
1990	set of emotions	[Ortony et, al.]	OCC specifies about 22 emotion categories and consists of five processes that define the complete system.
1992	set of emotions	[Ekman]	study of emotions and their relation to more than 10,000 facial expressions and expand the list of basic emotions, The newly included emotions are: Amusement, Embarrassment, Excitement, Guilt, Pride in achievement, Relief,Satisfaction, Sensory, pleasure , and Shame.

Timeline from 1992 to 2000: This is the time where lexicons are mostly develop. It mainly focuses on semantic orientation. Researchers manually annotate the text document with part-of-speech structure using adjectives, adverbs, interjections, POS-Tagging.

1990	WordNet	[Miller et. al.]	[Miller et. al.] engendered a lexicon dictionary kenned as semantic lexicon where words are accumulated into sets of synonyms (called “synsets”)
1992		[Hearst]	[Hearst] proposed a sentence interpretation model that endeavors to answer queries predicated on the argumentative structure of the document.
1994		[Tagger]	[Brill Tagger] represented the semantic orientation for verbs, adverb, entity and adjective.
1994		[Wiebe]	subjectivity analysis is the apperception of opinion-oriented language in order to distinguish it from

			objective language
1997	sentiment-based classification	[Hatzivassiloglou et.al.]	developed an algorithm for automatically apperceiving the semantic orientation of adjectives and relegating the semantic orientation of individual words or phrases, utilizing linguistic heuristics, a pre-culled set of seed words, or by human labeling.
1999		[Berland et al.]	describes a procedure that aims at extracting part-of features, utilizing possessive constructions and prepositional phrases, from news corpus.
1999	ANEW lexicon(manual lexicon)	[Bradley and Lang]	developed the Affective Norms of English Words (ANEW) which lists emotional ratings for 1034 English words.
2000		[Wiebe]	explained the concept of subjective adjectives in an information retrieval to explain two genres subjective and objective.
2000		[Hatzivassiloglou and Wiebe]	examined the effects of adjective orientation and gradability on sentence subjectivity. The goal was to tell whether a given sentence is subjective or not judging from the adjectives appearing in that sentence.

Timeline from 2000-2004: Researchers analyse the text automatically on the basis of polarity, subjectivity and objectivity. They mainly works on extracting polarity (orientation), positive, negative, neutral, scale rating(Rating scores are ordinal, this problem is tackled by regression.). Many papers published in 2002 and the subsequent years together explain the popularity of analysis of sentiments focused on natural language processing.

These are the years, when social networking sites are filled with unstructured, informal text so there is a need of emotion mining techniques that work on these unstructured text.

2001		[Das and Chen]	automatic analysis of evaluative text and tracking of the predictive judgments in analyzing market sentiment.
2002		[Corney et. al.]	presents the list of an informal and low structured language standard to communicate on social networking sites.
2002	PMI algorith m	[Turney]	Turney use an algorithm Point wise Mutual Information(PMI) to automate the classification of semantic polarity i.e.,positive and negative opinion, as Thumbs Up and Thumbs Down for electronic documents.

2002	Product Reputation Miner	[Morinaga et. al.]	extracts characteristic words, co-occurrence words, and typical sentences for individual target categories and identify positive or negative opinions based on a dictionary.
2002		[Pang et. al.]	Build lexicon for movie reviews as a dataset to indicate negative and positive statements and well suited for other machine learning approaches like Support Vector Machine, Maximum Entropy and Naive Bayes. The classification of sentiments is done at document level using syntactic approach of N-Grams and have a criticized view on the use of POS-TAGS as they do not provide valuable information for the classification of polarity in Twitter.
2003	Sentiment Analyzer	[Yi et al.]	Extracting sentiments about a given topic using natural language processing techniques.
2003	Separating facts from opinions	[Yu and Hatzivassiloglou]	Naive Bayes classifier on a corpus consisting of Wall Street Journal articles give best results achieving high accuracy(97%), where the task is to distinguish two different news articles about fact and opinions.
2003		[Turney and Littman]	Proposed a less influenced supervised algorithm that can predict the tendency of a word to go which direction, with a small set of positive seed words or with a small set of negative seed words.
2003	opinion mining, ReviewSeer	[Dave et al.]	a document level opinion classifier that uses mainly statistical techniques and some POS tagging information for some of their text term selection algorithms. It achieved high accuracy on review articles but performance degrades for the general web documents.
2003	Latent Dirichlet Allocation (LDA)	[Blei et al.]	is a machine learning method of topic modeling and it's a way of automatically discovering topics that sentences contain.
2003	An improvement on hand-crafted models	[Liu et al.]	Predict accurate emotions at sentence level by using multiple corpus-based linguistic analysis, an approach for graphically visualizing the affective structure of a text document , real world knowledge to extract affect from sentences as it assumes that all individuals feel the same way about a certain life event.
2004	Probabilistic model on machine learning algorithms	[Wilson et al.]	This approach is based on a human annotated text corpus that can identify the attitude, sentiments from the text present on social networking sites, classification is done on the basis of strength of the emotion and a major limitation is the difference in strength annotations measures between annotators.
2004	identify subjectivity of adjectives in Word Net.	[Kamps et al.]	identify subjectivity of adjectives in WordNet, classified adjectives into classes and find their relative distance to another word depending on the class.
2004	Bing Liu Lexicon	[Hu and Liu]	Bing Liu's Lexicon gives a list of positive and negative words

			manually tagged by user reviews.
2004	WordNet -Affect	[Strapparava and Valitutti]	developed a manual linguistic resource for lexical representation of affective knowledge named WordNet-Affect . It annotates the synsets that have an affective content. Emotion Classification is then done by mapping emotional keywords that exist in the input sentence to their corresponding WordNet-Affect concepts.

Timeline from 2005 to 2010: During this incipient stage of research on sentiment analysis of reviews, some of them focus on emotion categorization of the entire documents, which are based on the construction of discriminate-word dictionaries manually or semi-manually.

2005		[Aue and Gamon]	Identification of subject is a context dependent and domain dependent problem which replaces the earlier parable of using sentiwordnet or subjectivity word list etc. as prior knowledge database
2005		[Alm et al.]	explored text-based emotion prediction using supervised learning approaches.
2005		[Read]	Read explore different problems in the area of sentiment classification like Time, Domain and Topic dependency of sentiment orientation and use emoticons such as ":-)" and ":-(" and text-based emoticons to form a training set for the sentiment classification. The dataset was divided into "positive" and "negative " samples. Emoticons-trained classifiers: SVM and Naive Bayes, were able to obtain up to 70% of an accuracy on the test set.
2005		[Wiebe et al.]	present a comprehensive survey of subjectivity recognition using different clues and features.
2005	Lexical Affinity	[Ma et al.]	WordNet and WordNet-Affect are used to recognize whole sense in different context and the number of emotional senses to find suitable lexical affinity.
2005	MPQA subjectivity (Lexicon based method)	[Wilson et al.]	The MPQA Subjectivity Lexicon contains words assigned with their prior polarity and a discrete strength of evaluative intensity
2005		[Gammon et al.]	Uses Machine Learning Techniques with input of some seed words. This classifier is based on assumption that the words with same polarity co-occur in one sentence but words with different polarity cannot.
2005		[Niu et al.]	determine the polarity of outcomes (improvement vs. death, say) described in medical texts.
2006	sentiment analyzer	[Yi et al.]	introduced sentiment analyzer for world wide web text documents.
2006		[Andreevskaia et al.]	apply bootstrapping techniques to reduce the cost of building sentiment lexicons by adding words to an initial subset or seeds.
2006		[Mao et al.]	trained CRF classifier on sequential sentiments.
2006		[Wiebe and Mihalcea]	find relations between word sense disambiguation and subjectivity.
2006		[Wu et al.]	proposed approach for sentence level emotion mining based on detecting predefined semantic labels and attributes of the sentence, then classify only one emotion "happy" based on psychological patterns of human emotions.

2006		[Wang and McCallum]	Explains emotions of latent topics as time goes on. It means it adds a factor of time to capture variance of topic with respect to time.
2006	SentiWordNet	[Esuli and Sebastiani]	It annotates the term with prior sentiments.
2006		[Bethard et. al.]	have introduced the automatic identification of opinions from the process of question answering session.
2006		[Eguchi and Lavrenko]	Use sentiment classification in categorization, regression, and ranking to point out that the polarity . This assigned may be used for summarizing the content of opinionated text units on a topic, whether they be positive or negative, or for only retrieving items of a given sentiment orientation.
2007		[Yang et al.]	used Yahoo! Kimo Blog as corpora to build emotion lexicons. In their studies, emoticons were used to identify emotions associated with textual keywords.
2007	SemEval	[Strapparava et al.]	Tells all words can potentially convey affective meaning, even neutral also, can evoke pleasant or painful experiences because of their semantic relation with emotional concepts or categories. SemEval explains “affective text”, aiming to tag short headline texts with a predefined list of emotions and polarity orientation, the Emotion-Term model is based on Naive Bayes, estimate term-emotion associations using their co-occurrence counts.
2007		[Hancock et al.]	Explain that +ve and –ve emotions are expressed using exclamation and affective words, using content analysis
			Linguistic Inquiry and Word Count (LIWC) to classify emotions as positive or negative.
2007		[Redondo et al.]	adapted the ANEW into Spanish. This approach requires human translators to ensure the quality of the localized resource and therefore is cost expensive and not scalable.
2007	supervised emotion classification	[Yang et al.]	examine the classification of emotion of blogs using machine learning techniques. CRF classifier executes better than the SVM classifier at the sentence level and at the document level, the tactic of picking the last sentence's emotion as the answer outperforms all other strategies.
2007	Upar7, knowledge based system for headline sentiment tagging	[Chaumartin]	developed a linguistic rule-based system UPAR7,by combining WordNet, SentiWordNet and WordNet- Affect lexical resources. This uses dependency graph taken from the Stanford POS tagger. It is important to note that the classification is based on synsets, not on words
2007		[Mei et al.]	Proposed Topic Sentiment Mixture model for analysis of sentiment(emotion) on the topic level.
2008	emotion prediction	[Gill et al.]	explored the emotion rating activities of 65 judges from short blog texts
2008		[Tokuhisa et al.]	proposed a two step model for emotion classification using emotion-provoking event instances extracted from the web.
2008		[Titov et al.]	described a new statistical model called the Multiaspect Sentiment model (MAS), which consisted of two independent components. Differently, the model

			proposed in this paper unifies the process of generating topics and associating emotions with texts.
2008	Latent Semantic Analysis(LSI)	[Strapparava and Mihalcea]	developed a system that used several variations of Latent Semantic Analysis and evaluated several knowledge-based and corpus-based methods for the automatic identification of six emotions in text when no affective words exist. However their approach achieved a low accuracy because it is not context sensitive and lacks the semantic analysis of the sentence.
2008		[Zhao et al.]	uses Support Vector Machine (SVM), Conditional Random Field(CRF) algorithms to cluster opinions of same type.
2008	Statistical Models	[Pang et al.]	semantic information is highly considered as features. These models require annotated corpus, which is often limited for online texts.
2008		[Ganesan et al.]	presents a system for adding the graphical emoticons to text as an illustration of the written emotions.
2008	Opinion spam and analysis	[Jindal and Liu]	malicious users expressing offensive opinions, using their comments for the purpose of advertising, or even spreading rumors and fraudulent reviews. Considering this issue, Opinion Spam Detection is essential to detect and filter out irrelevant information in reviews, which is an important subtask when performing sentiment analysis.
2009		[Vo et al.]	adapted the ANEW into German. This approach requires human translators to ensure the quality of the localized resource and therefore is cost expensive and not scalable.
2009		[Bao et al.]	1. The emotion-term method was formulated by improving the Naive Bayes classifier . Different from traditional Naïve Bayes, the method takes into account emotional ratings when calculating the probability of a category and the probability of a term given an emotion label.2.As a joint emotion-topic model for social emotion mining, the Emotion-Topic Model (ETM) introduced an intermediate layer into LDA, in which a topic acts as an important component of an emotion. Infor-mative and coherent topics are extracted and grouped under different emotions.
2009		[Go et al.]	For the first time, Go et al. investigated tweet sentiment in which they utilized emoticons to annotate tweet with sentiment label and the presumption in the construction of the corpus is that the query ":)" returns tweets with positive smileys, and the query ":(" retrieves negative emotions.
2009		[Denecke]	introduced uses of SentiWordNet in terms of prior polarity scores. The author proposed two methods: rule-based and machine learning based. Accuracy of rule-based is 74% which is less than 82% accuracy of machine learning based. Finally, it is concluded that there need more sophisticated techniques of NLP, for better accuracy
2009		[Das and Bandyopadhyay]	explained the techniques for subjectivity based on Rule-based, Machine learning and Hybrid

			method.
2009		[Mohammad et al.]	to increase the scope of sentiment lexicon, it includes the identification of individual words as well as multi-word expressions with the support of a thesaurus and a list of affixes.It can be implemented by two methods: antonymy generation and Thesaurus based. Hand-crafted rules are used for antonymy generation. Thesaurus method is based on the seed word list which means if a paragraph has more negative seed words than the positive ones, then paragraph is marked as negative.
2009	sentim ent 140	[Go et al.]	use classifiers built from machine learning algorithms to avoid the problems of simpler keyword-based approach, which may have higher precision, but lower recall. They classify individual tweets.
2010	build sentim ent classifi er using multin omial Naïve bayes classifi er	[Pak and Paroubek]	Present a method for an automatic collection of corpus(Twitter) that can be used to train a sentiment classifier using syntactic structures. They use TreeTagger for POS Tagging as POS Tags are strong indicators of emotional text
2010	SentiSt rength	[Thelwall et al.]	proposed SentiStrength, a lexicon-based method for sentiment exposure on the Social Web. SentiStrength overcomes the problem of ill-formed language by applying several lexical rules, such as the existence of emoticons, intensifiers, negation and booster words like extremely, to compute the average sentiment strength of an online post. They identify

			emotions on two scales: the valence of the emotion indicating if the feeling is positive or negative and the arousal level indicating the energy level associated with the emotion and consider also variation of emotion of the gender.
2010	Develo p Lexico ns for unstruc tured langua ge like emotic ons, social Acrony ms,etc	[Yassine and Hajj]	The purpose was to make out whether the writer articulate his emotions and thoughts in his writings. The processed data was then used to spot the strength between two persons based on the subjectivity of the texts they share online. The main challenge for the model proposed is the free language of online social networks; in this perspective, we developed new lexicons that cover common expressions used by online users, including emoticons, social acronyms, Arabic expressions transliterated into English, etc…
2010	EmoHe art	[Neviarouska ya et al.]	Developed EmoHeart, a lexical rule-based system that identify emotions from text and envision the emotion expressions in a virtual environment. The system starts by looking for emotional abbreviations and emoticons. If not found, it processes the sentence on different levels to generate an emotional vector of the sentence, where each element in the vector represents an emotional class strength. At word level, each word in the sentence is mapped to its emotional vector, where they manually build a dataset of emotional vectors for many words. At the phrase and sentence levels, they combine the

			emotional vectors collected from the words by either performing summation or maximization among the vectors. The emotion of the sentence is the maximum strength of the vector. They achieved an average accuracy of 75% when tested on a manually annotated dataset.
2010	SentiWordNet 3.0 (Lexicon based method)	[Baccianella et al.]	lexical resource construction by applying a random walk algorithm, based on the well known lexicon resource, WordNet. It provides additional information on synsets related to sentiment orientation and returns from every synset a set of three scores and their polarity.
2010		[Batra and Rao]	use probabilistic representation measuring the sentiment of an entity as an combination of the sentiment of all tweets that are associated with that entity.
2010	NRC Emotion Lexicon	[Mohammad and Turney]	The NRC Emotion Lexicon encompass many frequent languages like English, French, etc. annotated for eight emotions (joy, sadness, anger, fear, disgust, surprise, trust, and anticipation) as well as for positive and negative sentiment.
2010		[Das et al.]	Sentiment Analysis explained till now is not sufficient to satisfy the needs of end user, because a user is not interested in binary output in terms of positive or negative but interested in aspectual sentiment classification. Aspectual can be explained as relative information. For example, a social worker may be interested to know the change in the society before and after implementation of his scheme. So, a sentiment analysis system should be understand and identify the aspectual sentiments present in the text. For this problem, Das propose sentiment structurization technique which is based on 5W (Why, Where, When, What, Who).The drawback of 5Ws is that it may lead to label bias problem which is solved by Maximum Entropy Model (MEMM) .
2010	multilingual sentiments	[Boyd-Graber et al.]	give idea for multilingual sentiment analysis is to translate languages into a well-studied language (e.g. English); hence traditional methods can be applied. Cross-language dictionaries work as bridges between different languages.
2010		[Davidov et al.]	Emoticons can also be exploited to extend the more common features used in text mining, such as sentimentcarrying words. A small set of emoticons has already been used as additional features for polarity classification so emoticonlabeled sets are used to automatically train the sentiment classifiers.
2010		[Joshi et.al.]	used two lexical resources: English-Hindi Word Net Linking and English SentiWordNet and created H-SWN(Hindi-SentiWordNet)

Timeline from 2011 to till now: This time mainly focuses on extracting emotions as contextual and conceptual semantic from sentence, reducing the features, extracting emotions from emoticons, emotions from multilingual corpora.

2011		[Kouloumpis et al.]	use certain seed hashtag words such as #cute and #sucks as labels of positive and negative sentiment.

2011	Lexicon based method	[Taboada et al.]	focus on sentiment strength varies from -5 to +5
2011		[Jiang et al.]	study the target-dependent sentiment classification of tweets by using SVM and General Inquirer. They classify the sentiments of the tweets as positive, negative or neutral according to a given query. Thus, the query serves as the target of the sentiments. In addition, they also apply a context-aware approach in order to incorporate the context of tweets into the classification.
2011	hybrid method(problem exist : lexicons (low recall) and machine learning techniques(depend on availability of labeled datasets))	[Zhang et al.]	They added hashtags to preprocessed data that provides a subjective meaning, special rules apply for the treatment of comparative judgments, the treatment of negation, and the treatment of expressions that can change the orientation of a phrase. To identify a greater number of words indicative of subjective content, using Chi-square test, with the idea that if a term is more likely to appear in a positive or negative judgment it is more likely to be a subjective content identifier.
2011		[Das et al.]	Genetic Algorithm achieved a good success for the subjectivity detection for Multiple Objective Optimization
2011	sentic computing	[Cambria et al.]	Concept Net, a semantic network was introduced with approx 10000 concepts and more than 72000 features extracted from Open mind corpus and developed Sentic Computing. This research is based on a common sense and emotion representation consisting of four dimensions as basis to classify the affective states: Sensitivity, Attention, Pleasantness and Aptitude. It have been used for short texts to infer emotional states over the web.
2011	Latent Dirichlet Allocation (LDA) methodology	[Hernandez and Sallis]	propose an unsupervised method of reducing features based on the Latent Dirichlet Allocation (LDA) methodology. The method is evaluated with a corpus of 10,000 tweets in English on the iPad tablet. They uses vector space model and using the TF-IDF metric to weight the terms to reduce features.
2011		[Agarwal and Xie]	used manually annotated tweets with sentiment and perform unigram model to do classification.
2011		[wang et al].	Utilized hashtag to perform graph-based classification.
2011		[Burget et al.]	proposed a framework that depends heavily on the pre-processing of the input data (Czech Newspaper Headlines) and labeling it using a classifier. The pre-processing was done at the word and sentence levels, by applying POS tagging, lemmatization and removing stop words. Term Frequency – Inverse Document Frequency (TF-IDF) was used to calculate the relevance between each term and each emotion class. They achieved an average accuracy of 80% for 1000 Czech news headlines using SVM with 10-fold cross validation. However their method was not tested on English dataset. Also it is not context sensitive as it only considers emotional keywords as features.
2011	multilingual twitter messages	[Cui et al.]	analyzed the emoticon of tweets with graph propagation algorithm for emoticon weighting.
2011		[Kolya et al.]	identified event and emotional expressions at word level from the sentences of TempEval-

			2010 corpus, in which the emotional expressions are also identified simply based on the sentiment lexicons, e.g., Subjectivity Wordlist, WordNet-Affect and SentiWordNet.
2012	SenticNet(concept based lexicon)	[Cambria et al.]	1. proposed an affective categorisation model primarily inspired by Plutchik's studies on human emotions. Such model represents affective states both through labels and through four independent but concomitant affective dimensions (Pleasantness, Attention, Sensitivity, Aptitude). In total, he identified 24 emotion labels. 2.SenticNet was proved valuable for sentiment detection in conventional text (e.g., product reviews) is a concept-based lexicon for sentiment analysis. It contains 14k fine-grained concepts collected from the Open Mind corpus and coupled with their sentiment orientations.
2012		[Shenghua et al.]	present two baseline models: 1) emotion-term model that uses Naive Bayes to model social emotion and affective terms via their co-occurrences and 2) a LDA topic model which utilizes the term co-occurrence information within a document and discovers the inherent topics within affective text.
2012	using Hidden Markov Model	[Dung et al.]	make use the idea that emotions are related to human mental states which are caused by some emotional events. This idea is implemented using Hidden Markov Model where each sentence consists of many sub-ideas and each idea is treated an event that causes a transition to a certain state. The sequence of events in
			the sentence is followed by the system and determines the emotion of the text.
2012		[Mohammad et al.]	developed a classifier to detect emotions using tweets with emotion word hashtags (e.g., #anger, #surprise) as labeled data as they are good indicators that the tweet as a whole (even without the hashtagged emotion word) is expressing the same emotion. 2. use the pointwise mutual information to measure the association between a word and a given emotion. So he builds a word emotion association lexicons which are lists of words and associated emotions. For example, the word victory may be associated with the emotions of joy and relief.
2013		[Cambria]	the sentiment of a word is implicitly associated with the semantics of its context
2013	Random walk analysis of the concepts	[Montejo-Raez et al.]	A new unsupervised approach to the problem of polarity classification in Twitter posts is resolved by combining a random walk algorithm that weights synsets from the text with polarity scores provided by SentiWordNet, it is possible to build a system comparable to a SVM based supervised approach in terms of performance. They present a new approach to the scoring of posts according to the positive or negative degree of the opinions expressed in the text.
2013		[Narayanan et al.]	worked on a fast and accurate sentiment classsification using an Naive Bayes model by combination of methods like effective negation handling, word n grams and feature selection by

			mutual information results in a significant improvement in accuracy.
2013		[Petz et al.]	researchers declare mathematical definition for opinion, they define an opinion as a quintuple (ei, aij, sijkl, hk, tl), when the opinion is expressed. An entity is the target object of an opinion. The aspects represent parts or attributes of an entity (part-of-relation). The sentiment is positive, negative or neutral or can be expressed with intensity levels.
2013		[Ortega et al.]	proposed a technique with three phases; pre-processing, polarity identification and classification. WordNet and SentiWordNet based approach is used for the purpose of polarity detection and rule-based classification is performed.
2013	multilingual twitter messages	[Cui et al.]	Mainly focus on building emotion tokens, including emotion symbols (e.g. emoticons), irregular forms of words and combined punctuations using emotion tokens are extracted automatically from tweets, emotion tokens are helpful for both English and non-English Twitter sentiment analysis, and are independent with the tweets in different time periods to build the lexicon with the help of a graph propagation algorithm
2013	word-level and topic-level emotional dictionaries	[Mao et al.]	1. Algorithms of building the word-level and topic-level emotional dictionaries are proposed, which are totally automatic, and no human resource is needed.2. The approach is compared with the state-of-the-art algorithms by the means of social emotion classification. In addition, qualitative investigation is conducted to analyze the generated emotional dictionary. 3. Compared with the existing emotional lexicons, the constructed emotional dictionary is language-independent, fine-grained, and can be updated constantly.
2014		[Kiritchenko et al.]	1. implemented a variety of features based on surface form and lexical categories by describing the process of creating the automatic, tweet-specific lexicons and demonstrate their superior predictive power over several manually and automatically created general-purpose lexicons and high-coverage, tweet-specific lexicons that we generated from tweets with sentiment-word hashtags and from tweets with emoticons. 2. created a supervised statistical sentiment analysis system that detects the sentiment of short informal textual messages such as tweets and SMS (message-level task) as well as the sentiment of a term (a word or a phrase) within a message (term-level task).
2014		[Shaheen et al.]	Propose a approach for emotion classification in English sentences where the emotions are treated as concepts extracted from the sentence. Concepts can be expressed as nouns, adjectives, adverbs, and verbal phrases or as a combination of different phrases. For example, consider the sentence “I found a solution to a problem”. This sentence represents an emotional concept extracted from the semantic relations between its words. The sentence indicates the emotion “Happiness”, as the concept of solving a problem will trigger the emotion “Happiness”.

2014	Concept level sentimen tal Analysis: EmoSent icSpace	[Poria et al.]	introduce concept level sentiment analysis, common-sense computing, and machine learning for improving the accuracy of tasks by the use of sentic patterns and dependency based rules
2014	sentimen t topic models	[Rao et al.]	propose two sentiment topic models to associate latent topics with evoked emotions of readers. The first model which is an extension of the existing Supervised Topic Model, generates a set of topics from words firstly, followed by sampling emotions from each topic. The second model generates topics from social emotions directly. Both models can be applied to social emotion classification and generate social emotion lexicons.
2014	unifying model	[Fraisse et al.]	presents a logical formalization of a set 20 semantic categories related to opinion, emotion and sentiment. Our formalization is based on the BDI model (Belief, Desire and Intetion) and constitues a first step toward a unifying model for subjective information extraction.
2014	twitter multiling ual affective lexicons	[Fraisse et al.]	presented a novel approach based on Twitter as a comparable corpus to extract automatically affective lexicons in seven langages (English, French, German, Italian, Spanish, Portuguese and Russian), this is motivated by the fact, that non englishspeaker's, usually, use bilingual terms in their messages. So, this is based on the co-occurence between the English and the target affective terms to generate multilingual affective lexicons.
2014	TOM	[Khan et al.]	1. Introduces and implements a hybrid approach for determining the sentiment of each tweet.2.Demonstrates the value of pre-possessing data using detection and analysis of slangs/abbreviations, lemmatization, correction and stop words removal. 3.Resolves the data sparsity issue using domain independent techniques.
2014		[Cambria et al.]	implemented the semantic multidimensional scaling for open domain sentimental analysis. In this work, the largest existing taxonomy of common knowledge is blended with a natural language based semantic network of common sense knowledge and multi dimensional scaling is applied on the resulting knowledge base for open domain opinion mining and sentimental analysis.
2015		[Koto et al.]	propose POS sequence as feature to investigate pattern or word combination of tweets in two domains of Sentiment Analysis: subjectivity and polarity, utilize Information Gain to extract POS sequence in three forms: sequence of 2-tags, 3-tags, and 5-tags. The results reveal that there are some tendencies of sentence pattern which distinguish between positive, negative, subjective and objective tweets. This shows that feature of POS sequence can improve Sentiment Analysis accuracy.
2015		[Amalanath an and Anouncia]	manually mapped the emoticons from Unicode 8.0 to nine emotional categories and performed the sentiment classification of tweets, using both emoticons and bag-of--words as features.

2015	Emoji Sentiment Ranking	[Novak et al.]	Emoji Sentiment Ranking, the first emoji sentiment lexicon of 751 emojis, is a valuable resource for helping humans during the annotation process, or even to automatically label the tweets with emojis for sentiment. In a lexicon-based approach to sentiment analysis, the emoji lexicon can be used in combination with a lexicon of sentiment-bearing words. Alternatively, an emoji with alreadyknown sentiment can act as a seed to transfer the sentiment to the words in proximity. Such a corpusbased approach can be used for an automated corpus construction for feature generation , and then applied to train a sentiment classifier.
2016	SentiCircles	[Saif et al.]	propose a semantic sentiment representation of words, called SentiCircle, which is able to assign context-specific sentiment orientation to words at both entity-level and tweet-level using different methods. SentiCircle representation effectively updated the sentiment strength of many terms dynamically based on their contextual semantics in tweets.

IV. NEED OF DEFINING EMOTIONS IN DIFFERENT DOMINIONS.

The social network on Internet provide a coherent medium through which people can interact and socialize. Almost every second person is a Internet user worldwide. In India also, over 30% people use Internet for their social need[source: Internet WordStats Usage and population Statistics]. The percentage increases exponentially day by day. All tasks performed online and people's are more comfortable to use this. This includes filling a form, comments, feedback, opinions, suggestions, reviews from user. Now this extends to chats, blogs, discussions, forums, promotion of products and services, Micro blogging sites like Twitter, Facebook. So here we are interested to recognize emotions of user's in all that sources.

There are many domains in this field.The domain are categorized to content based such as SMS, chats, blogs, and Wikipedia that are all introduced in 1980's to 1990's. then, media based came into existence in 2002 like virtual world and sharing of videos. Then mapping based services like google maps, yahoo maps are introduced in 2008. In all these we are lacking somewhere as if we need some answer from machine like recommendation system or question-answering, so here comes the context based where social search and recommendations attract user's attention since 2009 based on extracting emotions from text.

V. EMOTION RECOGNITION FROM STRUCTURED AND UNSTRUCTURED TEXT.

In this section we survey the state of the art in emotion recognition in the engineering literature.Defining the tasks that are useful in recognizing emotions are discussed in Section 5.1. Extraction of features such as syntactic, semantic and linguistic resources available in market are reviewed in Section 5.2 . Sections 5.3 and 5.4 are detailed reviews of recent work in emotion recognition, including Lexicon based and machine learning approaches.

Emotion Tasks

Sentiment Analysis is considered as a problem of classification. The first step is to extract and select text features. Some of the current features are:

—Presence and frequency of terms: These features are individual words or word n-grams and their frequency counts. It either gives the words binary weighting (zero if the word appears, or one if otherwise) or uses term frequency weights to indicate the relative importance of feature.

—Parts of speech (POS): finding adjectives, as they are important indicator of emotion.

—Opinion words and phrases: these are words commonly used to express opinions including good or bad, like or hate. On the other hand, some phrases express opinions without using opinion words.

—Negations: the appearance of negative words may change the orientation of opinion strength like not good is equivalent to bad.For this purpose, need is to text normalization mainly to handle negation.

Classes

Emotion mining categorize the emotion on different basis

—On the basis of Subjective/Objective texts: subjective means users feeling towards particular thing while objective is a fact.

—On the basis of valence indication: valence should be positive and negative.

—On the basis of tolerance: tolerance means strength or intensity.

In order to confine emotion from text document we require the classification which intends to infer the emotion expressed by the documents based on predefined lists of emotion. Predefined list of emotion contains basic set of emotions as [Ekman] defined six basic emotion named as Joy, Anger, Fear, Disgust, Sad and Surprise.

These are mainly focused on two main tasks.

—The test data aka text document that is collected from corpora required to understand the emotions invoked by words and phrases is huge. This is because a different word evokes different emotions learnt from our day to day experiences. For this purpose, need is to enhanced dictionary with emotion word from WordNet Affect, Sentic-Net, SentiCircles to improve in result.

—Since the scope of words is larger in the scenario, the usage of words and their inflected form is large too and negations are major emotion modifiers. So these problems need to be solved properly. For this purpose, need is to text normalization mainly to handle negation.

Polarity

The sentiment polarity detection means classifying the text document into semantic classes such as positive, negative or neutral. It can be defined in another classes of emotions like anger, sad, happy, surprise etc.. Polarity is assigned using dictionary such as SentiWordNet, WordNet, SenticNet, SentiFul and others.

Subjective/Objective

Sentiment analysis classify the text at subjective and objective nature. Subjective nature means the text having opinion content eg. “The car is uncomfortable” and objective nature means text having no opinion contain but contains some fact like “SBI opens new branch in Delhi”. [Das, 2011] accomplish a good success for the subjectivity detection on Multiple Objective Optimization in Genetic Algorithm.

5.1.2 Level

To analyze how people express emotions. Emotions can be expressed in simple text at different levels, by the use of adjectives or internet language like emoticons, interjections, acronyms and intentional misspelling like repetition of single letter in a word.

Emotions in simple text:

In simple text, we investigated that text is classified into the basic set of different emotions, classified using lexicons(in 1990’s using manual lexicons and now the automatically generated lexicons).

Sentiment Analysis is classified in number of levels. There are seven main classification levels- document-level, sentence-level, phrase-level, tweet-level, word-level, entity-level and feature/aspect/topic/attribute/facet level .

Document-level Sentiment Analysis intend to classify an opinion document as expressing a positive or negative opinion or sentiment. It considers the whole document as one basic information unit.

Sentence-level Sentiment Analysis intend to classify sentiment that can be expressed in each and every sentence. Initially, identify whether the sentence is subjective or objective. If the sentence is subjective, Sentence-level Sentiment Analysis will determine whether the sentence expresses positive or negative opinions. [Wilson et al.] explain that expressing a sentiment is not always subjective in nature. There is no basic difference between document and sentence level classifications because sentences are just short documents.

Tweet-level intends to identify the overall sentiment of individual tweets. As tweets is related to sentence as tweets contain 140 characters, is a single sentence or two. Tweets contains emoticons, informal language.

Word-level/Phrase level intends to identify the sentiment associated to each word as each word is associated with the sharing a sequence of human emotions.

Entity-level SA intends to use probabilistic models that can measure the sentiment of an entity as an combination of the sentiment of all tweets that are co-related with that entity. For example, the tweet, “The new Twitter for iPhone is awesome.”, expresses a positive sentiment for “Twitter”, but not for “ iPhone ”.

Classifying the text at document-level, sentence-level, or at word-level does not provide the required detail needed for opinions on all aspects of the entity, we need to move forward to aspect level. Sentiment Analysis till then, is not sufficient to satisfy the needs of end user, as one is interested in aspectual sentimental classification. For example, Indian government may be interested to know the change in the India before and after implementation demonetisation of currency. So, sentiment analysis system should be understand and identify the aspectual sentiments present in the text.

Aspect-level Sentiment Analysis intends to classify with respect to the particular aspects of entities. Initially, identify the entities and their aspects. The opinion possessor can give different opinions for different aspects of the same entity like this sentence “ The camera feature of iPhone is not good enough, but the battery life is very good’’. This survey deal with the first two kinds of Sentiment Analysis.

Adjectives are used to find emotions as they are good indicators of person’s emotions. In this we have to split the text into part-of-speech tags and then mark adjectives as emotions.

Emoticons were used in 37% of the posts; 22% of those contained more than one emoticon. Table illustrates the emotions that can be assigned to the used emoticons:

Table 2. Emoticons and expressed emotions

Emoticon	Emotion
:) :-) =) x) (:	Joy (smiling)
:D	Excitement
;) ;-) ;D ;P (;	Wink
xD =D ^^ ^.^	Happiness (laughing)
<3 Ɔ *_* *.* **	Love
:P	Playfulness
:O	Surprise
:S	Skepticism
(Y)	Support (thumbs up)
:(=(	Sadness
-.-	Annoyance

Table 3 : Interjections and their emotional meaning

Interjections	Emotion
Mmm	Pleasure
Hmm	Wondering
Mhmm	Confirmation
yeah, uee, juhu, jipi, wuhu, boah	Excitement
haha, hihi	Laughter
jumjum, njamnjam	Tasty
Wow	Surprise

Intentional misspelling and punctuations marks are interestingfrom the perspective of sentiment analysis as indication of emotion intensity. The recognized patterns include Capital letters,Repeating vocals andPunctuation marks, e.g.:"u r2 gud"

5.1.4 Source/Target

We can recognize emotions of a user, but what is user. Is it a reader, a writer or we can find emotions of any third entity or a target user.

FEATURE SELECTION AND EXTRACTION METHODS

Feature Selection methods works with annotating the terms on some bases, can be described into lexicon-based methods that require human annotation which begin with a small set of 'seed' words., and statistical methods which are based on automation that we frequently used.

The feature selection methods indulge the documents either as group of words aka Bag of Words, or as a string which retains the sequence of words in the document. Bag of words is used usually because of its ease for the classification.

5.2.1 Syntactic and Statistical techniques

Syntactic techniques can convey improved precision because they make use of the syntactic system of the language in order to identify the verbs, adjectives and nouns. Regrettably such techniques deeply depend on the language of the document and as a result, the classifiers can't be portable to different languages.

On the other hand statistical techniques have probabilistic environment and focus on the associations between the words and categories. Statistical techniques have two considerable benefits over the Syntactic ones: we can use them in further languages with less or no adaptations and we can use translation of machine language of the original dataset and still get fairly high-quality results.

Part-of-speech Tag

Part-of-Speech tagging is done to assign the speech to each word of the review so as to concentrate on the adjectives, verb and adverbs. These words of review are represented using n-grams. This representation is stored in a database for sentiment polarity calculation. The features from the database are retrieved and the sentiment polarity is calculated using sentiment analysis technique i.e. dictionary based technique.

N-Grams

N-grams are simply all combinations of adjacent words or letters of length n that you can find in your source text. For example, given the word cow, all 2-grams (or "bigrams") are co and ow. You may also count the word boundary – that would expand the list of 2-grams to #c, co, ow, and w#, where # denotes a word boundary.You can do the same on the word level. As an example, the hello, world! text contains the following word-level bigrams: # hello, hello world, world #.The basic point of n-grams is that they capture the language structure from the statistical point of view,like what letter or word is likely to follow the given one. The longer the n-gram (the higher the n), the more context you have to work with. Optimum length really depends on the application – if your n-grams are too short, you may fail to capture important differences. On the other hand, if they are toolong, you may fail to capture the "general knowledge" and only stick to particular cases.

5.2.2 Linguistic Resources

As we recognize emotions from social networking sites, the major step was to develop lexicons.There are many researcher who develop their lexicon based dictionaries that help to identify the emotion. Lexicon Resources are created to acquire the knowledge about emotions.In this regard, Philip Stones for the first time, identify emotions as compared to manual databases consist of emotion words. Then Fellbaum et al. developed WordNet, the lexicon dictionary consist of all English words. In 1999, Bradley et al. develop ANEW lexicon which consists of all affective words, then LIWC lexicon is proposed by Pennebaker et al. that works on frequency counts. In 2004 C. Strapparava et al. presents WordNet-Affect that emphasis on affective words. Wilson et al. develop MPQA subjectivity lexicon that calculates words valence and arousal level. In 2006, A. Esuli et al. develop SentiWordNet that extends the quality of WordNet. In 2007, C. Strapparava et al. develop Semeval that has numerous tasks. In this, emotions are fond to be categorized as positive, negative, and neutral and marked the text with affective words. Chaumartin et al. develop UPAR7, a rule based lexicon system using three defined lexicons sources, WordNet, WordNet-Affect and SentiWordNet. In 2010 Thelwall develop SentiStrength to find emotions from unstructured online post. Yassine et al. develop emoticons, acronyms and foreign lexicons. Neviarouskaya et al. develop EmoHeart, rule base lexicon that also helps to visualize emotions. Mohammed et al. develop NRC-Emotion lexicon that consists of part-of –speech tags for different set of emotions. In 2011 Cambria et al. give the concept of Concept-Net lexicon that recognize emotions based on common sense and Sentic-Net in 2012 that is based on Concept-Net. In 2014, Poria at al. presents EmoSenticSpace that is based on Concept Net and some machine learning algorithms. In 2015, Petra Kralj Novak et al. presents Emoji lexicons to analyse emotions from tet containing emoticons. In 2016, Hassan Saif et al. presents SentiCircles lexicon that dynamically enhance strength of emotions based on the meaning present behind the context.

5.2.3 Multilingual

To recognize emotions in cross-culture, many papers published uses the concept of translation to an intermediate state and then to a target language. Emoticons are one of the aspect to recognize emotions as emoticons, a graphical symbol is free from language and culture boundations and convey a same emotion throughout this world.

5.2.4 Feature Selection

In learning based techniques, before training the classifier, you must select the words/features that you will use on your model. You can't just use all the words that the tokenization algorithm returned simply because there are several irrelevant words within them.

Two commonly used feature selection algorithms in Text Classification are the Mutual Information and the Chi-square test. Each algorithm evaluates the keywords in a different way and thus leads to different selections. Also each algorithm requires different configuration such as the level of statistical significance, the number of selected features etc. Again you must use Trial and error to find the configuration that works better in your project.

Point-wise Mutual Information (PMI)

The mutual information determination provides a prescribed way to represent the mutual information between the features and the classes. This determination is predicted from the theory of information. The point-wise mutual information (PMI) among the word and the class is defined on the basis of the level of co-occurrence between the class and term .When PMI function returns the value greater than zero, the term is positively correlated to the class otherwise, the term is negatively correlated to the class.

PMI(term, class)=log[p(term, class)p(term)p(class)]

Many applications uses PMI, co-occurrence strength is only the consideration of it, so many changes are applied to it. [Yu and Wu] have enhance the basic PMI by developing a contextual entropy model to inflate a set of seed words generated from a small corpus of stock market news articles. Their contextual entropy model measures the similarity between two terms by comparing their contextual distributions using an entropy measure, allowing for the discovery of words similar to the seed words. Once the seed words have been expanded, both the seed words and expanded words are used to categorize the sentiment of the news articles. The results showed that their method can determine more useful emotion words, and its corresponding intensity improves the classification performance. This process outperformed the PMI-based expansion methods as they consider both co-occurrence strength and contextual distribution, thus acquiring more useful emotion words and fewer noisy words.

Chi-square (X2)

Chi-Square is one of the statistical approach like PMI approach that can assess the goodness of fit between a set of experimental values and those that are predictable hypothetically.

The test statistic for the chi-squared test of independence is

$$\chi^2 = \sum_{i=1}^{r} \sum_{k=1}^{c} [O_{ik} - E_{ik}]^2 \div E_{ik},$$

where

- r is the number of terms.
- c is the number of correlated term.
- O_{ik} is the observed count of the cell in the ith row and the kth column.
- E_{ik} is the expected count of the cell in the ith row and the kth column.

5.3.1 Dictionary based approach

The lexicon based methods use key Spotting method where they rely on emotion lexicons i.e., pre-built dictionary of words and their related sentiment orientation such as WordNet, LIWC lexicon, MPQA subjectivity lexicon, SentiStrength and SentiWordNet.

5.3.2 Corpus based approach

Corpus based lexicon requires annotating corpus with labels that consist of semantic information which is highly associated as features[Pang,2008] and [I.Titov et al.2008] described a new statistical representation "MultiAspect Sentiment Model" consisting of two things, first is to process generating topics and second to find its associating emotion. The topics is usually an entity for which we find emotions in different linguistics. There are different methods to find relevant emotions from corpus i.e. AAA, stands for Annotation, Abstraction and Analysis. Annotation includes part-of-speech tagging, parsing of text. Abstraction consists of translating of one text to other applying the linguistic rules and Analysis consists of statistically probing, evaluating, manipulating using rule base, and generalizing from the given dataset. SentiWordNet3.0 is more useful dictionary in this case. There are various limitations of lexicon-based methods as they are limited by pre-built dictionaries and they are fully reliant of presence of words or syntactical features that can echo emotions. Although this limitation is improved by Hassan Saif et al. "Contextual Semantics for Sentiment Analysis of Twitter". So, we move to new approach called machine learning approach.

5.4 MACHINE LEARNING APPROACH

Machine learning techniques can be understood by its four categories supervised, semi-supervised, unsupervised and hybrid.

5.4.1 Supervised approach

Supervised learning approach requires training of data to learn emotion classifiers. Initially manual seed-words are use to classify the sentiments of a text whether its polarity is positive, negative or neutral. Each domain should have different classifiers as we have different set of features for different domains and at different levels. For example, A positive review of one product is a negative review for other product. Classifying the emotions has different approaches, polarity, subjectivity classification, feature selection at different levels.

There are many learning algorithms based on supervised approach. Supervised learning approaches include support vector machine, neural network, naive bayes, Bayesian network and maximum entropy classification.

Support Vector Machine(SVM) a emotion trained classifiers used to analyze the data and data patterns that can be used for classification, regression analysis, clustering of opinions of same type of emotions. SVM performs best for tri-gram model Jiang,2011 use SVM classifier to study the classification of target dependent emotions. So, this is also helpful in context aware environment.

Conditional Random Field(CRF) is a classifier used for sequential emotions using structural prediction. It predict a label taking into account a neighboring emotions.

Naïve Bayes Classifier is based on naïve baye's theorem and uses the concept of maximum likelihood and Bayesian probability. It is used in emotion-term model in Sem-Eval 2007 provides a technique to calculate term-emotion associations using their co-occurrence counts. This classifier give highest accuracy with storyline documents or articles.

By using naïve bayes, CRF, and SVM classifier, one can found an emotion in binary output i.e, positive or negative.But a user can be interested to identify an emotion at aspect level so move to next method, Maximum Entropy classifier.

Maximum entropy is a probability distribution estimation technique widely used for language modeling, part-of – speech tagging and text segmentation. It prefer uniform models that satisfy some constraints.

5.4.2 Semi-supervised approach

The limitations of supervised approach is that it needs to train classifier and its dependence on domain cost much so to overcome this limitation Go et al. propose distant supervision approach that makes use of automatic generated training data set where emoticons are use to tag tweets as positive or negative.

5.4.3 Unsupervised approach

Point-wise Mutual Information (PMI) Algorithm

For the first time Turney et al. used Point-wise mutual information, an unsupervised classifier to automate a system that can find consecutive words and their semantic polarity using emoticons "thumbs up" to represent positive and "thumbs down" to represent negative opinions.

Latent Dirichlet Allocation Algorithm

A domain-independent lexicon based on Latent Dirichlet Allocation for sentiment analysis is constructed. It's a way of automatically discovering topics that sentences contain. LDA is a probabilistic model to construct a lexicon. The lexicon constructed is highly related to the dataset. Precision of this lexicon is more than the Liu's lexicon, MPQA and GI. This method is better than trivial methods in all aspects as trivial approach builds the lexicon based on calculating the words appearing number of occurrences in positive and negative reviews.

Random Walk Algorithm

Random Walk Algorithm, an automatic construction of domain-oriented sentiment lexicon. However, most of the attempts rely on only the relationship between sentiment words, failing to uncover the mutual relationship between the words and the documents, as well as ignoring the useful knowledge of some existed domains (or "old domain").The approach simulates a random walk on the graphs that reflect four kinds of relationships (the relationship between words, the relationship from words to documents, the relationship between documents, the relationship from documents to words) between documents and words.

5.5 Hybrid Approach

In lexicon based, there is a problem of low recall and in machine learning technique, problem is to domain independence. To avoid these two limitations, hybrid approach is used to determine emotion.

VI. CONCLUSION

Emotions are one of the major aspect of human life that are very useful in various applications as discussed in Table 1. So there is a need of something that recognizes emotions. Ekman, Pluchik and many other researchers define the group of emotions where we can classify our emotions. In this field we try our best to read as many papers as we can, summarise all papers, discuss some classifiers which we can use. To recognize emotions, a first step is to remove any content that can not be beneficial to recognize emotions like hashtagged content, url, email, etc. Then convert mis-spelled words like acronyms used, informal content used in messages etc. then using classifier, find the group which it belongs to, then find polarity, valence etc. to recognize emotions. Negative words, modals, adjectives, emoticons are good source to recognize emotions. We also identified that a emotion is not recognized only by its words, their co-occurrence but their semantic behind every context.

REFERENCES

[1] Osgood, Charles Egerton, William H. May, and Murray S. Miron. Cross-cultural universals of affective meaning.University of Illinois Press, 1975.

[2] Russell, James A. "Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies." Psychological bulletin115, no. 1 (1994): 102.

[3] Fahlman, Scott E., Jeff Baird, and Mike Jones. "Original Bboard Thread in which:-) was proposed." (1982).

[4] Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller. "Introduction to wordnet: An on-line lexical database*."International journal of lexicography 3, no. 4 (1990): 235-244.

[5] Ortony, Andrew, Gerald L. Clore, and Allan Collins. The cognitive structure of emotions.Cambridge university press, 1990.

[6] Shaver, Phillip, Judith Schwartz, Donald Kirson, and Cary O'connor. "Emotion knowledge: further exploration of a prototype approach." Journal of personality and social psychology 52, no. 6 (1987): 1061.

[7] Pluchik, R. "A general psychoevolutionary theory of emotions." Emotion: Theory, research and experience 1: 3-33.

[8] Hearst, Marti A. "Automatic acquisition of hyponyms from large text corpora." In *Proceedings of the 14th conference on Computational linguistics*-Volume 2, pp. 539-545.Association for Computational Linguistics, 1992.

[9] Ekman, Paul. "An argument for basic emotions." Cognition & emotion 6, no. 3-4 (1992): 169-200.

[10] Brill, Eric. "Some advances in transformation-based part of speech tagging."arXiv preprint cmp-lg/9406010 (1994).

[11] Hatzivassiloglou, Vasileios, and Kathleen R. McKeown. "Predicting the semantic orientation of adjectives." In *Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics*, pp. 174-181. Association for Computational Linguistics, 1997.

[12] Bradley, Margaret M., and Peter J. Lang. Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical report C-1, the center for research in psychophysiology, University of Florida, 1999.

[13] Hatzivassiloglou, Vasileios, and Janyce M. Wiebe. "Effects of adjective orientation and gradability on sentence subjectivity."In *Proceedings of the 18th conference on Computational linguistics*-Volume 1,

pp. 299-305. Association for Computational Linguistics, 2000

[14] Hatzivassiloglou, Vasileios, and Janyce M. Wiebe. "Effects of adjective orientation and gradability on sentence subjectivity."In *Proceedings of the 18th conference on Computational linguistics*-Volume 1, pp. 299-305. Association for Computational Linguistics, 2000

[15] Das, Sanjiv, and Mike Chen. "Yahoo! for Amazon: Extracting market sentiment from stock message boards." In *Proceedings of the Asia Pacific finance association annual conference (APFA)*, vol. 35, p. 43. 2001.

[16] Corney, Malcolm, Olivier De Vel, Alison Anderson, and George Mohay. "Gender-preferential text mining of e-mail discourse."In *Computer Security Applications Conference, 2002.Proceedings. 18th Annual*, pp. 282-289. IEEE, 2002.

[17] Turney, Peter D. "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews." In *Proceedings of the 40th annual meeting on association for computational linguistics*, pp. 417-424.Association for Computational Linguistics, 2002.

[18] Morinaga, Satoshi, Kenji Yamanishi, Kenji Tateishi, and Toshikazu Fukushima. "Mining product reputations on the web."*In Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 341-349.ACM, 2002.

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

[20] Pennebaker, James W., Matthias R. Mehl, and Kate G. Niederhoffer. "Psychological aspects of natural language use: Our words, our selves."Annual review of psychology 54, no. 1 (2003): 547-577.

[21] Yi, Jeonghee, Tetsuya Nasukawa, RazvanBunescu, and Wayne Niblack. "Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques." In Data Mining, 2003.ICDM 2003. *Third IEEE International Conference* on, pp. 427-434. IEEE, 2003.

[22] Yu, Hong, and VasileiosHatzivassiloglou. "Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences." In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pp. 129-136. Association for Computational Linguistics, 2003

[23] Dave, Kushal, Steve Lawrence, and David M. Pennock. "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews."In *Proceedings of the 12th international conference on World Wide Web*, pp. 519-528.ACM, 2003.

[24] Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." the Journal of machine Learning research 3 (2003): 993-1022.

[25] Liu, Hugo, Henry Lieberman, and Ted Selker. "A model of textual affect sensing using real-world knowledge." In *Proceedings of the 8th international conference on Intelligent user interfaces*, pp. 125-132. ACM, 2003.

[26] Wilson, Theresa, JanyceWiebe, and Rebecca Hwa. "Just how mad are you? Finding strong and weak opinion clauses."In aaai, vol. 4, pp. 761-769. 2004.

[27] Kamps, Jaap, M. J. Marx, Robert J. Mokken, and M. de Rijke. "Using wordnet to measure semantic orientations of adjectives." (2004): 1115-1118

[28] Hu, M. and Liu, B., 2004, August. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).ACM.

[29] Strapparava, Carlo, and Alessandro Valitutti. "WordNet Affect: an Affective Extension of WordNet." In LREC, vol. 4, pp. 1083-1086. 2004.

[30] Aue, Anthony, and Michael Gamon. "Customizing sentiment classifiers to new domains: A case study." In Proceedings of recent advances in natural language processing (RANLP), vol. 1, no. 3.1, pp. 2-1. 2005.

[31] Alm, Cecilia Ovesdotter, Dan Roth, and Richard Sproat. "Emotions from text: machine learning for text-based emotion prediction." In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pp. 579-586. Association for Computational Linguistics, 2005

[32] Read, Jonathon. "Using emoticons to reduce dependency in machine learning techniques for sentiment classification."In *Proceedings of the ACL student research workshop*, pp. 43-48.Association for Computational Linguistics, 2005.

[33] Wiebe, Janyce, and Ellen Riloff. "Creating subjective and objective sentence classifiers from unannotated texts." In Computational Linguistics and Intelligent Text Processing, pp. 486-497. Springer Berlin Heidelberg, 2005.

[34] Ma, Chunling, Helmut Prendinger, and Mitsuru Ishizuka. "Emotion estimation and reasoning based on affective textual interaction." In Affective computing and intelligent interaction, pp. 622-628.Springer Berlin Heidelberg, 2005.

[35] Wilson, Theresa, JanyceWiebe, and Paul Hoffmann. "Recognizing contextual polarity in phrase-level sentiment analysis."In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pp. 347-354.Association for Computational Linguistics, 2005.

[36] Gamon, Michael, Anthony Aue, Simon Corston-

Oliver, and Eric Ringger. "Pulse: Mining customer opinions from free text." In Advances in Intelligent Data Analysis VI, pp. 121-132.Springer Berlin Heidelberg, 2005.

[37] Niu, Yun, Xiaodan Zhu, Jianhua Li, and Graeme Hirst. "Analysis of polarity information in medical text."In AMIA. 2005.

[38] Zhao, Yan-Yan, Bing Qin, and Ting Liu. "Sentiment analysis." Journal of Software 21, no. 8 (2010): 1834-1848.

[39] Andreevskaia, Alina, and Sabine Bergler. "Sentiment tagging of adjectives at the meaning level."In Advances in Artificial Intelligence, pp. 336-346.Springer Berlin Heidelberg, 2006.

[40] Mao, Yi, and Guy Lebanon. "Isotonic conditional random fields and local sentiment flow." In Advances in neural information processing systems, pp. 961-968. 2006.

[41] Wiebe, Janyce, and RadaMihalcea. "Word sense and subjectivity." In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pp. 1065-1072. Association for Computational Linguistics, 2006.

[42] Wu, Chung-Hsien, Ze-Jing Chuang, and Yu-Chung Lin. "Emotion recognition from text using semantic labels and separable mixture models." *ACM transactions on Asian language information processing (TALIP)* 5, no. 2 (2006): 165-183.

[43] Wang, Xuerui, and Andrew McCallum. "Topics over time: a non-Markov continuous-time model of topical trends." In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 424-433.ACM, 2006.

[44] Esuli, Andrea, and FabrizioSebastiani. "Sentiwordnet: A publicly available lexical resource for opinion mining." In Proceedings of LREC, vol. 6, pp. 417-422. 2006.

[45] Bethard, Steven, and James H. Martin. "Identification of event mentions and their semantic class."In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pp. 146-154.Association for Computational Linguistics, 2006.

[46] Eguchi, Koji, and Victor Lavrenko. "Sentiment retrieval using generative models."In *Proceedings of the 2006 conference on empirical methods in natural language processing*, pp. 345-354.Association for Computational Linguistics, 2006.

[47] Yang, Changhua, Kevin Hsin-Yih Lin, and Hsin-Hsi Chen. "Emotion classification using web blog corpora."In Web Intelligence, IEEE/WIC/ACM International Conference on, pp. 275-278.IEEE, 2007.

[48] Strapparava, Carlo, and RadaMihalcea. "Semeval-2007 task 14: Affective text." In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pp. 70-74.Association for Computational Linguistics, 2007.

[49] Hancock, Jeffrey T., Christopher Landrigan, and Courtney Silver. "Expressing emotion in text-based communication." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 929-932. ACM, 2007.

[50] Redondo, Jaime, Isabel Fraga, Isabel Padrón, and Montserrat Comesaña. "The Spanish adaptation of ANEW (affective norms for English words)."Behavior research methods 39, no. 3 (2007): 600-605

[51] Chaumartin, François-Régis. "UPAR7: A knowledge-based system for headline sentiment tagging." In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pp. 422-425.Association for Computational Linguistics, 2007.

[52] Mei, Qiaozhu, Xu Ling, Matthew Wondra, Hang Su, and ChengXiangZhai. "Topic sentiment mixture: modeling facets and opinions in weblogs." In *Proceedings of the 16th international conference on World Wide Web*, pp. 171-180.ACM, 2007.

[53] Titov, Ivan, and Ryan T. McDonald. "A Joint Model of Text and Aspect Ratings for Sentiment Summarization."In ACL, vol. 8, pp. 308-316. 2008.

[54] Strapparava, Carlo, and RadaMihalcea. "Learning to identify emotions in text." In Proceedings of the 2008 ACM symposium on Applied computing, pp. 1556-1560. ACM, 2008.

[55] Zhao, Jun, Kang Liu, and Gen Wang. "Adding redundant features for CRFs-based sentence sentiment classification."In Proceedings of the conference on empirical methods in natural language processing, pp. 117-126.Association for Computational Linguistics, 2008.

[56] Pang, Bo, and Lillian Lee. "Opinion mining and sentiment analysis."Foundations and trends in information retrieval 2, no. 1-2 (2008): 1-135.

[57] Ganesan, Kavita A., NeelakantanSundaresan, and HarshalDeo. "Mining tag clouds and emoticons behind community feedback." In *Proceedings of the 17th international conference on World Wide Web*, pp. 1181-1182.ACM, 2008.

[58] Jindal, Nitin, and Bing Liu. "Opinion spam and analysis."In *Proceedings of the 2008 International Conference on Web Search and Data Mining*, pp. 219-230.ACM, 2008.

[59] Bao, Shenghua, ShengliangXu, Li Zhang, Rong Yan, Zhong Su, Dingyi Han, and Yong Yu. "Joint emotion-topic modeling for social affective text mining."In Data Mining, 2009.ICDM'09. Ninth IEEE International Conference on, pp. 699-704. IEEE, 2009.

[60] Go, Alec, RichaBhayani, and Lei Huang. "Twitter sentiment classification using distant supervision." CS224N Project Report, Stanford 1 (2009): 12.

[61] Denecke, Kerstin. "Are SentiWordNet scores suited for multi-domain sentiment classification?." In Digital Information Management, 2009.ICDIM 2009. Fourth International Conference on, pp. 1-6. IEEE, 2009.
[62] Das, Dipankar, and SivajiBandyopadhyay. "Sentence-level emotion and valence tagging." Cognitive Computation 4, no. 4 (2012): 420-435.
[63] Mohammad, Saif M., and Peter D. Turney. "Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon." In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, pp. 26-34.Association for Computational Linguistics, 2010.
[64] Go, Alec, RichaBhayani, and Lei Huang. "Twitter sentiment classification using distant supervision." CS224N Project Report, Stanford 1 (2009): 12.
[65] Pak, Alexander, and Patrick Paroubek. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." In LREc, vol. 10, pp. 1320-1326. 2010.
[66] Thelwall, Mike, Kevan Buckley, Georgios Paltoglou, Di Cai, and ArvidKappas. "Sentiment strength detection in short informal text." Journal of the American Society for Information Science and Technology 61, no. 12 (2010): 2544-2558.
[67] Yassine, Mohamed, and Hazem Hajj. "A framework for emotion mining from text in online social networks."In Data Mining Workshops (ICDMW), 2010 IEEE International Conference on, pp. 1136-1142.IEEE, 2010.
[68] Neviarouskaya, Alena, Helmut Prendinger, and Mitsuru Ishizuka. "EmoHeart: conveying emotions in second life based on affect sensing from text."Advances in Human-Computer Interaction 2010 (2010): 1.
[69] Baccianella, Stefano, Andrea Esuli, and FabrizioSebastiani. "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining." In LREC, vol. 10, pp. 2200-2204. 2010.
[70] Batra, Siddharth, and Deepak Rao. "Entity based sentiment analysis on twitter." Science 9, no. 4 (2010): 1-12.
[71] Mohammad, Saif M., and Peter D. Turney. NRC Emotion Lexicon.NRC Technical Report, 2013.
[72] Davidov, Dmitry, Oren Tsur, and Ari Rappoport. "Enhanced sentiment learning using twitter hashtags and smileys." *In Proceedings of the 23rd international conference on computational linguistics: posters, pp. 241-249*. Association for Computational Linguistics, 2010.
[73] Boyd-Graber, Jordan, and Philip Resnik. "Holistic sentiment analysis across languages: Multilingual supervised latent Dirichlet allocation." *In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 45-55.Association for Computational Linguistics, 2010.
[74] Joshi, Aditya, A. R. Balamurali, and Pushpak Bhattacharyya. "A fall-back strategy for sentiment analysis in hindi: a case study." *Proceedings of the 8th ICON (2010).*
[75] Kouloumpis, Efthymios, Theresa Wilson, and Johanna D. Moore. "Twitter sentiment analysis: The good the bad and the omg!." Icwsm 11 (2011): 538-541.
[76] Taboada, Maite, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. "Lexicon-based methods for sentiment analysis." Computational linguistics 37, no. 2 (2011): 267-307.
[77] Zhang, Yudong, Zhengchao Dong, Lenan Wu, and Shuihua Wang. "A hybrid method for MRI brain image classification." Expert Systems with Applications 38, no. 8 (2011): 10049-10053.
[78] Das, Swagatam, and PonnuthuraiNagaratnamSuganthan. "Differential evolution: a survey of the state-of-the-art." Evolutionary Computation, IEEE Transactions on 15, no. 1 (2011): 4-31.
[79] Cambria, Erik, Thomas Mazzocco, Amir Hussain, and Tariq Durrani. "Switching Between Different Ways to Think."In Analysis of Verbal and Nonverbal Communication and Enactment. The Processing Issues, pp. 56-69. Springer Berlin Heidelberg, 2011.
[80] Hernández, Sergio, and Philip Sallis. "Sentiment-preserving reduction for social media analysis." In Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, pp. 409-416. Springer Berlin Heidelberg, 2011.
[81] Agarwal, Apoorv, BoyiXie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. "Sentiment analysis of twitter data."In *Proceedings of the workshop on languages in social media*, pp. 30-38.Association for Computational Linguistics, 2011.
[82] Wang, Xiaolong, Furu Wei, Xiaohua Liu, Ming Zhou, and Ming Zhang. "Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach." In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pp. 1031-1040.ACM, 2011.
[83] Cui, Anqi, Min Zhang, Yiqun Liu, and Shaoping Ma. "Emotion tokens: Bridging the gap among multilingual twitter sentiment analysis." InInformation retrieval technology, pp. 238-249.Springer Berlin Heidelberg, 2011.
[84] Das, Dipankar, Anup Kumar Kolya, AsifEkbal, and SivajiBandyopadhyay. "Temporal analysis of sentiment events–a visual realization and tracking."InComputational Linguistics and Intelligent Text Processing, pp. 417-428. Springer Berlin Heidelberg, 2011.
[85] Cambria, Erik, Marco Grassi, Amir Hussain, and

Catherine Havasi. "Sentic computing for social media marketing." Multimedia tools and applications 59, no. 2 (2012): 557-577.

[86] Bao, Shenghua, ShengliangXu, Li Zhang, Rong Yan, Zhong Su, Dingyi Han, and Yong Yu. "Mining social emotions from affective text." Knowledge and Data Engineering, IEEE Transactions on 24, no. 9 (2012): 1658-1670.

[87] Chamlertwat, Wilas, PattarasineeBhattarakosol, TippakornRungkasiri, and ChoochartHaruechaiyasak. "Discovering Consumer Insight from Twitter via Sentiment Analysis." J. UCS 18, no. 8 (2012): 973-992.

[88] Mohammad, Saif M., Svetlana Kiritchenko, and Xiaodan Zhu. "NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets." arXiv preprint arXiv:1308.6242 (2013).

[89] Cambria, Erik. "An introduction to concept-level sentiment analysis."InAdvances in Soft Computing and Its Applications, pp. 478-483. Springer Berlin Heidelberg, 2013.

[90] Montejo-Ráez, Arturo, Eugenio Martínez-Cámara, M. Teresa Martín-Valdivia, and L. Alfonso Ureña-López. "Ranked wordnet graph for sentiment polarity classification in twitter." Computer Speech & Language 28, no. 1 (2014): 93-107.

[91] Narayanan, Vivek, IshanArora, and Arjun Bhatia. "Fast and accurate sentiment classification using an enhanced Naive Bayes model."InIntelligent Data Engineering and Automated Learning–IDEAL 2013, pp. 194-201. Springer Berlin Heidelberg, 2013.

[92] Petz, Gerald, MichałKarpowicz, HaraldFürschuß, Andreas Auinger, VáclavStříteský, and Andreas Holzinger. "Opinion mining on the web 2.0–characteristics of user generated content and their impacts." In Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data, pp. 35-46. Springer Berlin Heidelberg, 2013.

[93] Ortega, Reynier, Adrian Fonseca, and Andrés Montoyo. "SSA-UO: Unsupervised Twitter sentiment analysis." In Second Joint Conference on Lexical and Computational Semantics (* SEM), vol. 2, pp. 501-507. 2013.

[94] Cui, Anqi, Haochen Zhang, Yiqun Liu, Min Zhang, and Shaoping Ma. "Lexicon-based sentiment analysis on topical chinesemicroblog messages." In Semantic Web and Web Science, pp. 333-344. Springer New York, 2013.

[95] Qadir, Ashequl, and Ellen Riloff. "Bootstrapped learning of emotion hashtags# hashtags4you."In Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp. 2-11. 2013.

[96] Mao, Xudong, YanghuiRao, and Qing Li. "Recipe popularity prediction based on the analysis of social reviews." In Awareness Science and Technology and Ubi-Media Computing (iCAST-UMEDIA), 2013 International Joint Conference on, pp. 568-573. IEEE, 2013.

[97] Kiritchenko, Svetlana, Xiaodan Zhu, and Saif M. Mohammad. "Sentiment analysis of short informal texts." Journal of Artificial Intelligence Research(2014): 723-762.

[98] Shaheen, Shadi, Wassim El-Hajj, Hazem Hajj, and Shady Elbassuoni. "Emotion Recognition from Text Based on Automatically Generated Rules." In Data Mining Workshop (ICDMW), 2014 IEEE International Conference on, pp. 383-392.IEEE, 2014.

[99] Poria, Soujanya, Erik Cambria, GregoireWinterstein, and Guang-Bin Huang. "Sentic patterns: Dependency-based rules for concept-level sentiment analysis." Knowledge-Based Systems 69 (2014): 45-63.

[100] Rao, Yanghui, Qing Li, Xudong Mao, and Liu Wenyin. "Sentiment topic models for social emotion mining." Information Sciences 266 (2014): 90-100.

[101] Fraisse, Amel, and Patrick Paroubek."Toward a unifying model for Opinion, Sentiment and Emotion information extraction."In LREC, pp. 3881-3886. 2014.

[102] Fraisse, Amel, and Patrick Paroubek. "Twitter as a comparable corpus to build multilingual affective lexicons." In The 7th Workshop on Building and Using Comparable Corpora, pp. 26-31. 2014.

[103] Khan, Farhan Hassan, Saba Bashir, and UsmanQamar. "TOM: Twitter opinion mining framework using hybrid classification scheme." Decision Support Systems 57 (2014): 245-257.

[104] Poria, Soujanya, Alexander Gelbukh, Erik Cambria, Amir Hussain, and Guang-Bin Huang. "EmoSenticSpace: A novel framework for affective common-sense reasoning." Knowledge-Based Systems 69 (2014): 108-123.

[105] Koto, Fajri, and MirnaAdriani."The Use of POS Sequence for Analyzing Sentence Pattern in Twitter Sentiment Analysis." In Advanced Information Networking and Applications Workshops (WAINA), 2015 IEEE 29th International Conference on, pp. 547-551. IEEE, 2015.

[106] Amalanathan, Anthoniraj, and S. Margret Anouncia."Social network user's content personalization based on emoticons." Indian Journal of Science and Technology 8, no. 23 (2015).

[107] Novak, Petra Kralj, JasminaSmailović, BorutSluban, and Igor Mozetič."Sentiment of emojis." PloS one 10, no. 12 (2015): e0144296.

[108] Saif, Hassan, Yulan He, Miriam Fernandez, and HarithAlani."Contextual semantics for sentiment analysis of Twitter." Information Processing & Management 52, no. 1 (2016): 5-19.

