STM JOURNALS
Journal of Artificial Intelligence Research & Advances
ISSN: 2395-6720 (Online)
Volume 5, Issue 2
www.stmjournals.com

Emotion Evaluation from Text: A Survey

Aabiroo Bader**, Amjad Husain Bhat
Department of Information Technology, School of Engineering and Technology, Central University of Kashmir, Srinagar, Jammu and Kashmir, India

Abstract

Information extraction is the process of retrieving structured information from unstructured or semi-structured text. It has a wide range of applications in domains such as biomedical literature mining, business intelligence, sentiment analysis, and emotion analysis. This survey describes recent works in the field of emotion detection for text data, being a part of the broader area of affective computing. It has been inspired by the well-known fact that, despite there is much work done in this field of research, the clock of research is still ticking ON. The increment of this work is due to a significant amount of emotional data available on social web. Analyzing this data over the internet means we are spanning across the world, going through all the cultures and communities. This paper summarizes the previous works done in, by the various emotional models and computational approaches employed.

Keywords: *Sentiment analysis, emotion analysis, machine learning, distant supervision, bootstrapping*

****Author for Correspondence*** E-mail: baderabiro@gmail.com

INTRODUCTION

Emotion analysis is a local field of research that is intently interrelated to sentiment analysis. The objective of sentiment analysis is to reveal the positive, negative or neutral feelings from the text. However, the objective of emotion analysis is to spot and distinguish types of feelings from the expression of text, for instance, anger, happiness, trust, anticipation, disgust, fear, sadness, and surprise. The emotion analysis forms cardinal part of the Affective computing. Affective computing is everything that helps us to design systems and devices that can interpret, process, recognize and simulate the human affects [1, 2]. Emotion analysis and detection have been extensively studied in psychology, neuroscience, and behavioral science, as they are vital elements of human nature. A text-based emotion prediction procedure would benefit from recognizing the emotional affinity of data.

Sentiment and emotion analysis is a subtopic of NLP (natural language processing) and text mining that deals with the automated detection and mining of knowledge about people's sentiments, evaluation, and opinions from textual data such as articles, personal blogs, reviews and feedback forms. In recent years, sentiment analysis has changed from studying online product reviews to social media texts like Facebook, Twitter etc. We have selected to work with Twitter since we feel it is a better approximation of public sentiment in contrast to conventional internet articles and web blogs. The reason is that the amount of relevant data is much more significant for Twitter, as compared to conventional blogging sites. Evaluating the sentiments and emotions of different textual data on the internet possess its worth, viz. measuring the wellbeing of people, suicide prevention [3], target marketing, e-learning [4] and also it can be helpful for political parties to identify whether public supports their program or not.

This paper is grounded on previous works done in the arena of emotion analysis through text which is an evolving field with various applications in the physical world. Already much work has been completed in this field of research since long-ago, and the clock of research is still ticking ON notably, using the Tweets [5–10]. In this survey, we classify the most appropriate emotion recognition works by the emotional models and the approaches applied. Conversely, analysis of text emotions also hosts several challenges in work in the sense that the emotions and the methods to convey these emotions are all subjective.

Emotion analysis employs the NLP, text analysis and numerous computational procedures to reveal the emotions covered in a specific text. The analysis can be performed at several levels: document level [11, 12], sentence level [13], entity level [14], and aspect level [15].

1. Document-level: Document level analysis classifies the entire document text into different sentiments like positive, negative or neutral.
2. Sentence level: Sentence level analysis determines whether each sentence expresses a positive, negative or neutral opinion.
3. Entity and Aspect: Aspect level is the opinion mining and summarization based on the feature.

The classification is concerned about recognizing and extracting features from the source data.

The rest of the paper is organized into following sections: Next part presents the emotion models, followed by emotion lexicon, emotion corpora, computational approaches in emotion analysis and finally the conclusion.

EMOTION MODELS

When emotional recognition systems are analyzed, it is vital to focus our attention on defining and clarifying how the emotion models are formed, as they are the basis of these systems. According to research in psychology, there are a number of theories around how to connote emotions. This paper describes three significant emotion models: categorical, dimensional and extended.

Categorical Model

Emotional categorical approaches are persistent on model emotions based on clear emotion labels or classes. The categorical model follows the approach of dividing the emotions into several groups and assigns the emotion into a proper discrete category [16].

Categories which are used in the study of textual emotion detection are based on the idea of discrete emotion theory. This theory claims that there is only a limited set of emotions which can be experienced.

The most prominent categorical models are defined as under:

- The Ekman's basic emotion model is under this approach. Ekman's emotions are the most popular set of basic emotions and classify anger, disgust, fear, joy, sadness and surprise as the set of basic emotions [17, 18].
- Parrot's hierarchy model has categorized emotions into six discrete classes of fear, anger, joy, love, sadness, and surprise [19].
- Frijda in his model divided emotions into six discrete classes' viz. desire, happiness, interest, surprise, wonder, and sorrow [19].
- Plutchik's model is a superset of Ekman's basic emotion model with two additional classes: trust and anticipation (Figure 1) [20].
- Izard in his model divides emotions into 12 classes viz. contempt, self-hostility, fear, shame, interest, joy, surprise, sadness, anger, disgust, shyness and guilt [21].
- Extended Ekman's model divided emotions into 18 discrete classes of anger, disgust, fear, joy, sadness, surprise, amusement, relief, satisfaction, sensory pleasure, contempt, contentment, embarrassment, excitement, guilt, pride in achievement, and shame [22].

Dimensional Model

The dimensional emotion model follows the approach of representing the emotion classes in a dimensional form: either 2D or multi-dimension, with each emotion occupying a distinct position in space. Thus emotions are not a subject of assignment to a single category, but to several variables. Consider Ekman's basic emotions, the categorical model would allocate one single emotion label to a specific word. Extending the categorical emotion classification by adding a scale to each of the emotions, the result is more flexible emotion vector. By adding emotion scales for each category, the final classification forms a vector, where each item resembles specific emotion, and each value corresponds to emotional intensity.

The most prominent dimensional models are defined as under:

- Rusell's Circumplex Model [23]: This model suggests that emotions are distributed in a 2D circular space where the vertical axis represents the arousal and horizontal axis represents the valence and the origin represents neutral as shown in Figure 2. The valence dimension specifies how much PLEASANT and UNPLEASANT an emotion is. The arousal dimension distinguishes ACTIVATION and DEACTIVATION.

- Vector model [24]: This model consists of two vectors pointing in two directions assuming the existence of an underlying arousal dimension with valence dimension, vector determining the direction in which a particular emotion lies.
- PANA (positive activation-negative activation) model [25]: This model divides the system into positive and negative effect, with the vertical axis representing the positive effect and the horizontal axis representing the negative effect.
- Plutchik' model [20]: Plutchik contributed a hybrid model arranging the emotions into concentric circles where inner circle represents the basic and the outer circle represents more complex emotions.
- PAD (pleasure, arousal and dominance) model [26]: This model is a superset of circumplex model with addition of dominance dimension, which indicates whether the subject feels in control of the situation or not.
- Lovheim cube model [27]: It presents signal substances forming the axis of the coordinate system, and the eight basic Ekman's emotions are placed at eight corners of the cube (Figure 3).

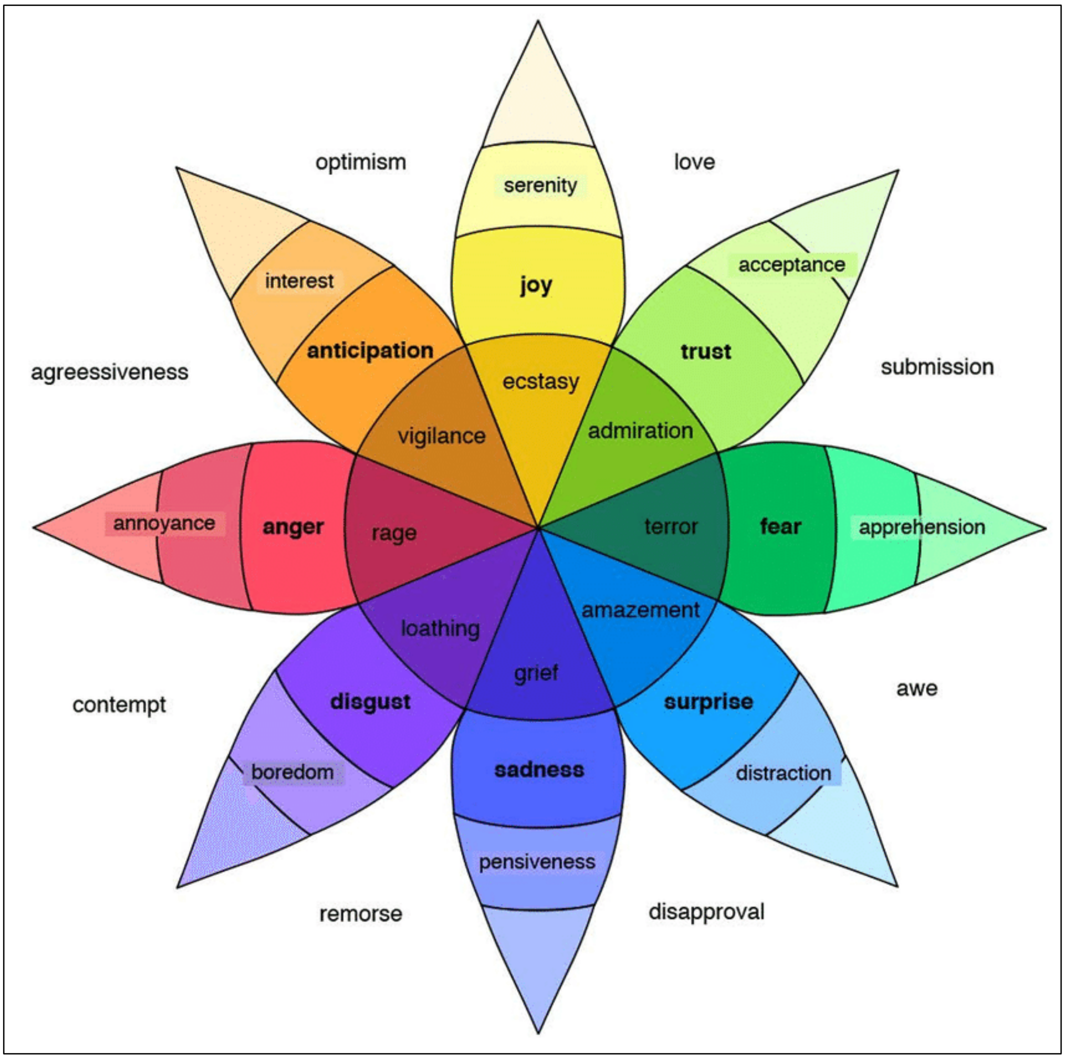

Fig. 1: *Plutchik's Model.*

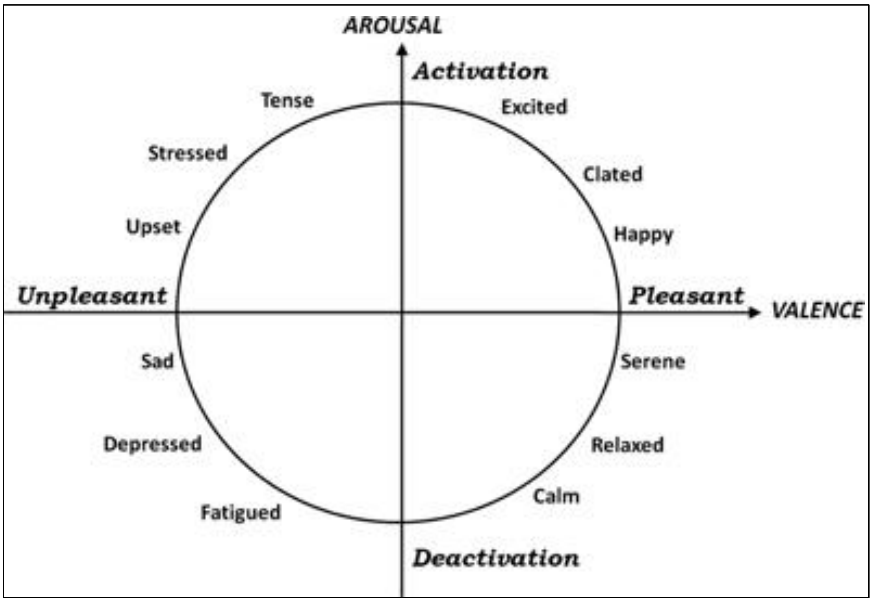

Fig. 2: *Circular Representation of Circumplex Model.*

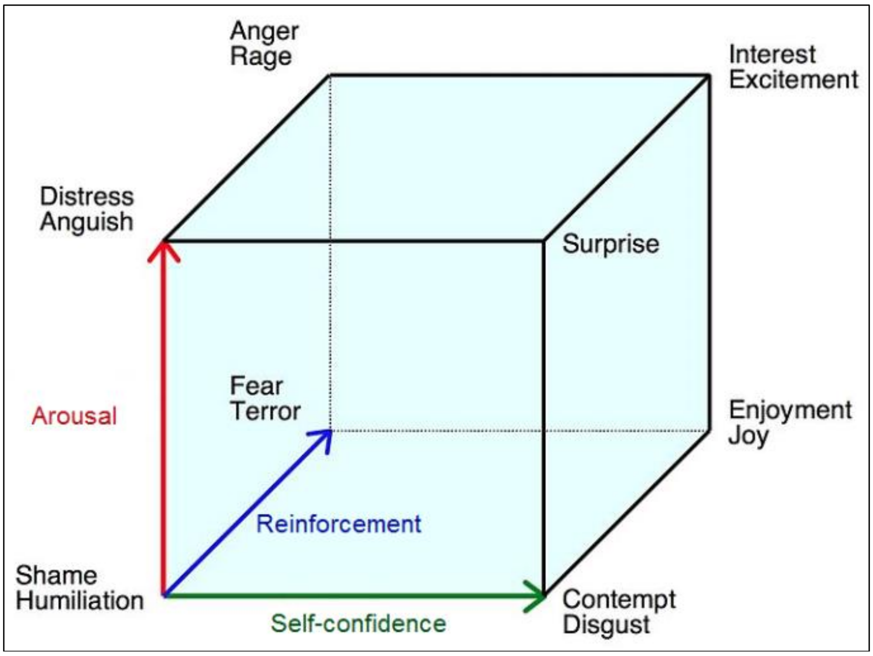

Fig. 3: *Lovheim Cube Model.*

Extended Models

In addition to models mentioned above, there are other few models considered in the field of textual emotion detection. Examples are Ortony, Clore and Collins model and Five Factors Model. Both take a little bit different approach than preceding ones. Firstly, these models are used in research, where the personality, targets, and desires of the presenting or receiving the text are taken into account. Like, in the work of Kao *et al.* FFM model is used and in the paper from Li *et al.* OCC integrated with FFM is used [28–30]. OCC's chief features are a decision tree to determine emotion, contemplating the event triggering the emotion, and assessing the relationship between the feeling and the situation.

EMOTION LEXICON

Emotion lexicons can be defined as a set of words characterized according to their emotional connotation. The label can be an emotional category when the categorical model is utilized or can be a value of the asset of a given emotion dimension in a concrete dimensional model. Consequently, in literature, we can attain lexicons labeled with one or another model.

- **WordNet Affect (WNA) [31]**

WordNet-Affect is a branch of WordNet Domains that incorporates a subset of synsets appropriate to signify affective concepts associated with affective words. The affective concepts indicating emotional states are individuated by synsets prominent with the a-label emotion. There are also additional a-labels for those concepts indicating moods, emotional responses or situations eliciting emotions.

- **ANEW [32]**

Affective norms for English words can be defined as a set of normative emotional ratings for an enormous quantity of words in the English language (1,034 English words comprise verbs, nouns, and adjectives). Each word is evaluated from 1 to 9 regarding valence, arousal, and dominance.

- **NRC Word-Emotion Association Lexicon (Emolex) [33]**

Also called Emolex is a dataset of general domain comprising around 14,000 English unigrams (words) associated with the Plutchik's eight basic emotions (anger, joy, trust, surprise, sadness, anticipation, disgust and fear) and the two sentiments (negative and positive), composed through manual annotation.

- **NRC Hashtag Emotion Lexicon [34]**

NRC Hashtag Emotion Lexicon can be well-defined as a dataset of Twitter domain entailing almost 16,000 words related with the Plutchik's eight basic emotions by automatic annotation from tweets with emotion word hash-tags.

- **DAL [35]**

The dictionary of affective language holds 8,742 words which have been assessed by people for their activation, evaluation, and imagery.

- **EmoSenticNet (ESN) [36]**

EmoSenticNet can be demarcated as a lexical resource consisting of 13,171 words that allocate qualitative emotion label and quantitative polarity values to SenticNet concepts. WordNet Affect emotion labels (Ekman's emotion labels) are the set of emotions used for labeling the concepts.

- **DepecheMood [37]**

DepecheMood can be described as a lexicon of roughly 37,000 terms related with emotion scores of eight emotion categories viz. inspired, happy, amused, angry, afraid, annoyed, sad and do not care). It was developed by exploiting the crowd-sourced emotional annotation implicitly offered by readers of news articles from Rapper website.

- **LIWC**

Linguistic Inquiry and Word Count (LIWC) system is an accessible text processing system. The core dictionary of LIWC consists of 4500 words, and word stems organized into four categories, affective processes, cognitive processes, pronouns-negation (no, never) and quantifiers (few, many).

Table 1 below shows the emotion lexicons.

Table 1: *Emotion Lexicons.*

Dictionary	Annotation	Size	Emotions
WordNet Affect (WNA) [31]	Manual	4,787 terms	Semantic levels, valence, arousal and affective labels
ANEW [32]	Manual	1,034 words	Valence, Arousal and Dominance
NRC Word-Emotion Association Lexicon (Emolex) [33]	Manual	14,182 words	Plutchick's emotions
NRC Hashtag Emotion Lexicon [34]	Automatic	16,862 words	Plutchick's emotions
DAL [35]	Manual	8,742 words	Activation, Evaluation and Imaginary
EmoSenticNet (ESN) [36]	Automatic	13,171 words	Ekman's Emotions
LIWC	Manual	4500 words	Four emotion categories
DepecheMood [37]	Automatic	37,216 words	Eight emotion categories

EMOTION CORPORA

An emotion corpus is a dataset of largely structured sentences where each sentence is tagged with one or more emotional labels. Corpora are a fundamental part of supervised learning approaches, as they depend on a labeled training data. Supervised learning algorithm evaluates the training data and infers a function, which it uses to label the unlabeled data. Most automatic emotion recognition systems are centered on a limited set of proposed basic emotions. Thus, the majority of corpora developed so far have been annotated with categorical emotion model.

Majority emotion annotation task has been approached with a manual process. This way, machine learning systems learn from human annotations that are usually more accurate. Amongst these resources, we can attain corpora labeled with the six basic emotions categories proposed by Ekman for instance: a sentence-level corpus of approximately 185 children stories annotated with emotion categories; blog posts collected directly from Web annotated with emotion categories and intensity [38].

Although there is corpus of mined 700 sentences from blog posts provided by BuzzMetrics [39] annotated with one of nine emotion categories (a subset of emotional states defined by Izard) and a consequent intensity value; corpus extracted 1,000 sentences from numerous stories [40], annotated with one of 14 categories of their annotation scheme (between them Izards emotions) and a subsequent score (the strength or intensity value). Emotiblog-corpus, that contains of a set of blog posts annotated by three annotation levels: document, sentence and element using a collection of 15 emotions [41]; or EmoTweet-28 corpus, that comprises of a group of tweets annotated with 28 emotion categories [42].

In the present era, there is a new kind of social media corpora which incorporate multi-layered manual annotations, amongst which the annotations for basic emotions are also included. For example; (a) the Twitter corpus, where a set of 2012 US presidential election tweets annotated by crowdsourcing for a number of attributes concerning to sentiment, emotion, purpose or intent behind the tweet, and style [43]; (b) Multi-View Sentiment Corpus (MVSC), manually labelled with various aspects of the natural language text: objective/subjective, polarity, implicit/explicit, emotion and irony, which includes 3,000 English micro-blogs posted associated with the movie domain [44].

The main obstacle of manual annotation is that the procedure is time-consuming and costly due to different factors viz. highly subjective, annotation scheme, the difficulty of the task or the training of the annotators.

Therefore with the aim of overcoming the short comes of manual annotation, some emotion resources have recently been developed which employ emotion word hash-tags to create automatic emotion corpus from Twitter. Mohammad describes how they generated a corpus from Twitter post (Twitter Emotion Corpus TEC) using this technique [45, 46].

Thus, in textual emotion analysis research community, the interest in developing amounts of emotion corpora has raised because that would allow refining supervised machine learning systems. For example, emotion intensity (EmoInt) proposed at WASSA 2017, which proposes a novel related challenge for emotion recognition. Therefore, although the use of emotion word hash-tags as a technique to label data is secure and efficient regarding time and cost, this method can be entirely useful to social networks and micro-blogging services.

Table 2 shows the emotion corpora.

COMPUTATIONAL APPROACHES

The computational approaches incorporate all the techniques that are used to design and implement an emotional classifier system (Figure 4). They are primarily classified into two main categories: lexicon-based approach and the machine learning approach. On the one hand, lexicon-based approach depends on lexical resources like lexicons, bags of words or ontologies. On the other hand, a machine learning approach (ML) employs ML algorithms grounded on linguistic features to train the system and infers a function for future classification of emotion.

Table 2: *Emotion Corpora.*

Corpus	Source	Size	Annotation	Emotions
Alm corpus [38]	Children stories	185 stories 1,200 sentences	Manual	Ekman's emotions
Nevioroskaya corpus (1) [40]	Stories	1,000 sentences	Manual	14 emotional categories (Izard's emotions)
Nevioroskaya corpus (2) [39]	Blog posts	700 sentences	Manual	Izard's emotions and polarity categories
Emotiblog-corpus [41]	Blog posts	1,950 sentences	Manual	15 emotions
EmoTweet-28 corpus [42]	Tweets	15,553 tweets	Manual	28 emotion categories, valence, arousal and emotion cues
2012 US Electoral Corpus [43]	Tweets	1,600 tweets	Manual	Sentiment, emotions (Plutchick's emotions), purpose, and style
Multi-View Sentiment Corpus [44]	Tweets	3,000 tweets	Manual	Subjectivity, Emotions (Plutchick's emotions), sentiment, explicit/implicit, and irony
Twitter Emotion Corpus (TEC) [45]	Tweets	21,000 tweets	Automatic	Ekman's emotions
Wang corpus [46]	Tweets	2.5 million Tweets	Automatic	7 Emotion categories

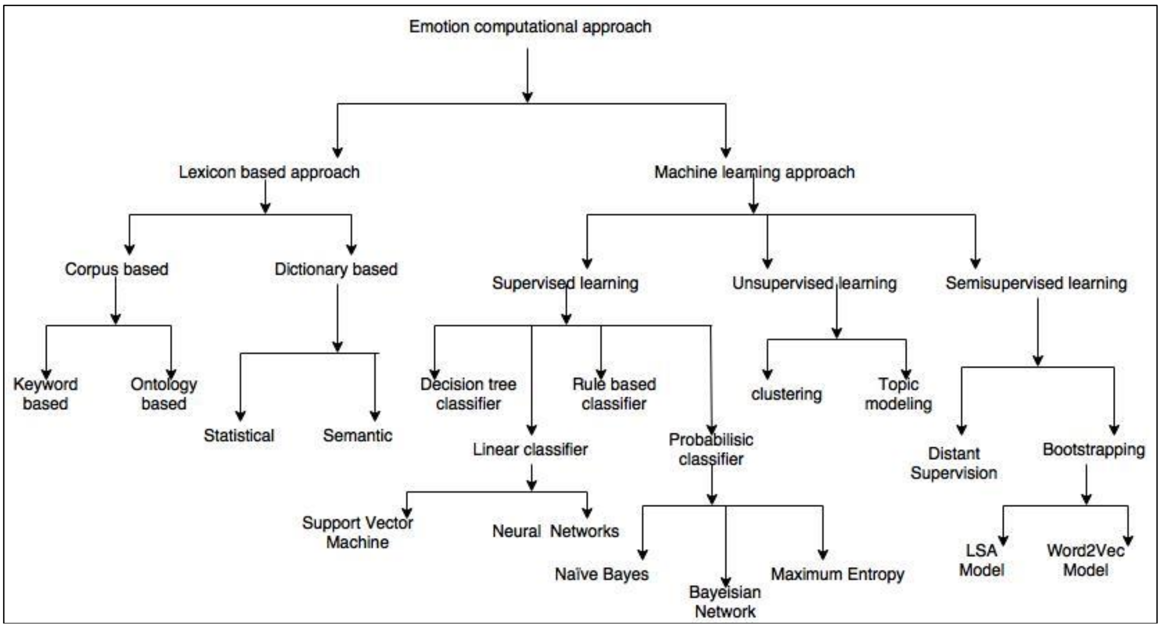

Fig. 4: *Computational Approaches.*

Lexicon Based Approach

Lexicon based approaches are approaches that merely use one or several lexical resources to identify emotions [47]. These are divided into two subtypes: dictionary based and corpus-based approach.

The corpus-based approach employs any conventional data for emotion analysis. Here the corpus (data) is first annotated by obeying a set of abstract rules, from a text for governing a natural language emotion analysis. Corpus-based approaches are further classified into two subtypes: Keyword-based and ontology-based approach.

Keyword-Based Approach

The keyword-based approaches are grounded on predetermining a set of terms to categorize the text into emotion classes. Strapparava and Mihalcea, in their study, as a baseline, employed a simple algorithm that examined the presence of affective words in the headlines and calculated a score that reflected the frequency of the words in the affective lexicon of the text [48]. They utilized WordNet-Affect.

Ontology-Based Approach

Ontology-based ones utilize EmotiNet, a resource for the recognition of emotion from text based on commonsense knowledge on concepts, their collaboration and their affective consequence to reveal emotion [49]; EmotiNet model situations as chains of actions and their subsequent emotional effect applying an ontological representation. Fine-grained emotion detection also follows this approach.

The dictionary-based approaches initiate with a specifically predefined dictionary of emotional words and then apply several measures, viz. term frequency, word count (so-called statistics) or word synonyms (called semantics), etc. to label the sentences. The dictionary-based approaches are further of two subtypes: statistical and semantic approach. Most statistical approaches employ Latent Semantic Analysis (LSA), and their variation has been used for studying the relationship between set of documents and the terms in these documents to generate meaningful patterns related to documents and terms [47]. They applied LSA and the Hyperspace Analogue to Language (HAL) to automatically compute the semantic similarity amongst the texts and emotions keywords.

Machine Learning Approach

Machine learning is a scientific discipline that deals with the study and implementation of algorithms that can learn from data by making usage of linguistic features of the text [48]. These algorithms operate by implementing a model based on inputs and applying these inputs to make predictions and decisions, instead of following only explicitly programmed instructions [49]. Especially in emotion recognition, machine learning algorithms are utilized to learn how to detect emotions. These approaches can be divided further into supervised, semi-supervised and unsupervised learning.

Supervised Machine Learning

Supervised learning approaches depend on a labeled training data, a set of training instances. The supervised learning algorithm analyses the training data and implies a function, which we employ for mapping future examples. A labeled corpus is a large and structured set of data that is necessarily annotated with emotional labels. Here, the annotation process is considered as one of the most critical drawbacks as it becomes an annoying and time-consuming task. However, there are recent works associated with emotion detection in Twitter tweets, where the training examples are automatically labeled using hashtags and emoticons contained, also methods for automatically labeling training data. Furthermore, it confirms that hashtags are great emotion labels. The SVM is a traditional approach in this regard. Few researchers have moved further than these traditional approaches to more efficient and reliable methods like CRF (conditional random field) [50, 51].

Regarding works that employ supervised learning algorithms, we can locate both the categorical and the dimensional approaches to base their emotional models. Categorical approaches are normally used in emotion detection, an empirical study of applying supervised machine learning with the SNoW learning architecture. They used an annotated corpus with an Extended Ekman basic emotion model.

Concerning works that employ supervised learning approach with dimensional emotion model, where an approach is applied for automatically classifying text messages of the individual to infer their emotional states. They used the Rusell's circumplex dimensional model of affect, and trained supervised classifiers to identify multiple emotions. Primarily, they have evaluated the accuracy of SVM, KNN, Decision Tree and Naïve Bayes for classifying Twitter tweets.

The supervised learning models are further of following types:

Decision Tree Classifiers: make use of the hierarchical recursive decomposition of training data based on the value of the attributes; till some leaf nodes are accessed that comprise values for the classification purpose [52].

Rule-Based Classifier: here, the classification is modeled on the basis of some set of rules, i.e., either manually defined or automatically learned. Each token in the text is characterized by a set of features. The conditions in disjunctive standard form in the features are represented by the LHS (left-hand side) whereas the class labels are represented by RHS (right-hand side) of the rule [53].

Linear Classifiers: It categorizes the emotions by making a decision based on the value of linear combination of characteristics of the input text, also known as feature values. These feature values are represented in a vector form called feature vectors. It consists of many models; among them, the most important are the support vector machines [54] and Neural Networks [55].

Probabilistic Classifier: assumes each class to be a component of the mixture which offers the probability of sampling a specific term for that component. It further entails following classifiers: Naïve Bayes classifier, which calculates the posterior probability based on words in a document [56]; Bayesian network, is an acyclic graph whose nodes signify random variable, and edges signify conditional dependencies; Maximum entropy, applies encoding to convert labeled features into vectors. These vectors then compute weights for each feature which are combined for the detection of the class for each feature set [57].

Table 3: *Summarization of Prior Working in the Field.*

Prior Work	Year	Domain of Dataset	Emotion Labels Used	Approaches Used	Features/Lexicons Used	Performance
Alm *et al*. [16]	2005	Children Fairy Tales	Angry, Disgust, Fear, Happy, Sad, positively and negatively surprised	Supervised Learning (SNoW learning Architecture)	WordNet-Affect, conjunction of features, special punctuations	63.31%
Yang *et al.* [60]	2007	Web-blogs	Positive (happy, joy) and Negative (sad, angry)	Supervised Learning, SVM and CRF (Conditional Random field)	Custom made lexicon, Emotion Words and emoticon Features	SVM: 48.83% CRF: 56.0%
Aman and Szpakowicz [61]	2007	Blogs-posts	Happy, Sad, Angry, Disgust, Surprise, Fear, mixed and no emotions	Supervised Learning (NB, SVM)	Word Net-Affect, General Inquires	NB: 72.0% SVM: 73.89%
Bellegarda [50]	2010	News Headlines	Joy, Sad, Surprise, Anger, Disgust, Fear	Supervised (NB, SVM, and decision trees), Latent Semantic Mapping (LSM)	WordNet-Affect, WordNet synset, Bag of words	NB: 59.72% SVM: 71.69%
Chaffarand Inkpen [62]	2011	News, blogs, Fairy tales, diary like blog posts	Anger, Disgust, Fear, Happiness, Sad, Surprise	Supervised learning (NB, SVM and decision trees)	WordNet-Affect, n-grams and bag of words	NB: 59.72% DT: 64.70% SVM: 71.69%
Balabantaray *et al.* [6]	2012	Blog	Anger, Disgust, Fear, Joy, Sad, Surprise, Neutral	Supervised Learning (Multi-Class SVM)	n-grams, pos, WordNet Affect, emoticons, dependency parsing	73.24%
Sykora *et al.* [63]	2013	Blog	Anger, Disgust, Fear, Joy, Sad, Surprise, confusion, shame	Ontology Engineering approach, Custom NLP pipeline	Emotive ontology Lexicon, feature intensifiers, negators	85.0%
Wang and Zheng [51]	2013	ISEAR	Anger, Disgust, Fear, Joy, Sad, Surprise, guilt, shame	Improved LSA Algorithm	Scientific Library	80.6%
Suttles and Ide [64]	2013	Blog	Anger, Disgust, Fear, Happy, Sad, Surprise, Trust, Anticipation	Supervised Learning (NB, ME), Natural language toolkit	Hashtags, Emoticons, Emojis	82.9%
Hasan *et al.* [8]	2014	Blog	Happy active, happy inactive, unhappy active, unhappy inactive	Supervised learning	ANEW lexicon, LIWC dictionary, AFINN, emoticons, punctuations	87.0%
Ahmad *et al*.	2017	Social media	Anger, Disgust, Fear, Joy, Sad, Surprise, Anticipation, Trust	TF-IDF along with naive Bayes	WordNet Affect, Hashtags, Emoticons, Emojis	81.24%
Meo and Sulis [10]	2017	Blog	Anger, Disgust, Fear, Joy, Sad, Surprise, Anticipation, Trust	NB, SVM, Random Forest, Logistic Regression	Emotion lexicon, polarity lexicon, latent factor, 5 dictionaries	SVM: 75.7% NB: 89.5% RF: 93.5% LR: 93.5%

Semi-Supervised Machine Learning

Semi-supervised algorithms employ the idea of automatic labeling, most crucial semi-supervised technique for relation extraction is bootstrapping, which commences from a small set of seed relation instances and iteratively learns more relation instances and extraction patterns [58, 59]. An illustrative work on bootstrapping for relation extraction is the Snowball system proposed by Agichtein and Gravano [58], which enhanced an earlier system called DIPRE given by Brin [59]. More recently, another learning paradigm called distant supervision has been developed to get used of a large number of known relation instances in existing large knowledge bases to create training data, for example, Wikipedia and Freebase. Noisy training data is automatically generated for both bootstrapping and distant supervision, to increase efficiency, careful feature selection and pattern filtering need to be carried out.

Unsupervised Machine Learning

Unsupervised learning approaches do not necessitate any training data, thus can be applied to any text data without demanding any manual effort. These systems try to reveal the hidden structures in the input data and employing those structures map the unlabeled future data to emotion classes. The two critical unsupervised learning methods usually used are clustering and topic modeling. In clustering approach, a corpus of documents is segmented into partitions, each corresponding to a local cluster. The problems of clustering and topic modeling are correlated. In topic modeling, a probabilistic model is used to determine a soft clustering, in which each document has a membership probability of the cluster, as opposed to a hard segmentation of the documents. Thus, topic models can be thought as the process of clustering with a generative probabilistic model.

CONCLUSIONS

Social media networks, especially Twitter, provide a vital source of information regarding users and their emotions towards a particular product, service or event. In this paper, we presented a survey of the work done in the arena of identifying emotions from textual data. We observed several trends in research; there is a noticeable move from rule-based and unsupervised approaches to supervised approaches. This must be because of the increased accessibility to emotion-annotated data. Very recently, the role of hierarchy is being explored in emotion analysis research. In this paper, we have presented a framework of experimentation result of some of the works for various computational models with their overall system accuracy, shown in Table 3. We note that there has been a significant improvement in the system accuracy over the time with the improvement or modification of traditional computational approaches, the lexicons and the features generated. Based on our survey, in future, we can deal with further improvement or modification of traditional computational approaches, supervised learning augmented with deep learning.

ACKNOWLEDGEMENTS

I would like to express my gratitude towards my parents and guide (Mr. Amjad Husain Bhat) for their support and encouragement which motivated and helped me in completion of this study.

REFERENCES

1. Rosalind Picard W. *Affective Computing.* Cambridge, MA, USA: MIT Press; 1997.
2. Calvo A (Senior Member). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Trans Affect Comput*. 2010; 1(1): 18–37p.
3. Desmet Bart, Hoste V´eronique. Emotion Detection in Suicide Notes. *Expert Syst Appl*. Nov 2013; 40(16): 6351–6358p.
4. Rodriguez Pilar, Ortigosa Alvaro, Carro Rosa M. Extracting Emotions from Texts in E-Learning Environments. *Complex, Intelligent and Software Intensive Systems (CISIS), 2012 6th International Conference on.* 2012.
5. Mohammad SM, Kiritchenko S. Using Hash Tags to Capture Fine Emotion Categories from Tweets. *Comput Intell.* 2015; 31(2): 301–326p.
6. Balabantaray RC, Mohammad Mudasir, Sharma Nibha. Multi-Class Twitter Emotion Classification: A New Approach. *International Journal of Applied Information Systems (IJAIS).* 2012; 4(1): 48–53p.

7. Wang W, Chen L, Thirunaraya K, *et al*. Harnessing Twitter "Big Data" for Automatic Emotion Identification. *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom).* 2012.
8. Hasan M, Agu E, Rundensteiner E. Using Hashtags as Labels for Supervised Learning of Emotions in Twitter Messages. *ACM SIGKDD*. 2014.
9. Roberts Kirk, Roach Michael A, Johnson Joseph, *et al*. EmpaTweet. Annotating and Detecting Emotions on Twitter. *LREC*. 2012; 3806–3813p.
10. Meo Rosa, Sulis Emilis. Processing Affect in Social Media: A Comparison of Methods to Distinguish Emotions in Tweets. *ACM Trans Internet Technol*. 2017; 17(1): 1–25p.
11. Moraes Rodrigo, Valiati João Francisco, Gavião Neto Wilson P. Document-Level Sentiment Classification: An Empirical Comparison between SVM and ANN. *Expert Syst Appl*. 2013; 40(2): 621–633p.
12. Ainur Yisong Y, Claire C. Multi-Level Structured Models for Document-Level Sentiment Classification. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, MIT, Massachusetts, Association for Computational Linguistics, USA. 2010; 1046–1056p.
13. Noura F, Elie C, Rawad AA, *et al*. Sentence-Level and Document-Level Sentiment Mining for Arabic Texts. *Proceeding IEEE International Conference on Data Mining Workshops*. 2010.
14. Nikos E, Angeliki L, Georgios P, *et al*. ELS: A Word-Level Method for Entity-Level Sentiment Analysis. *WIMS'11: Proceedings of the International Conference on Web Intelligence, Mining and Semantics*. 2011.
15. Haochen Z, Fei S. Aspect-Level Sentiment Analysis Based on a Generalized Probabilistic Topic and Syntax Model. *Proceedings of the 28th International Florida Artificial Intelligence Research Society Conference*, Association for the Advancement of Artificial Intelligence. 2015.
16. Cecilia Ovesdotter Alm, Dan Roth, Richard Sproat. Emotions from Text: Machine Learning for Text-Based Emotion Prediction. *Proc Conf Human Language Technology and Empirical Methods in Natural Language Processing*. 2005; 579–586p.
17. Ekman P, Friesen VW, O'Sullivan M, *et al*. Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion. *J Pers Soc Psychol*. 1987; 53(4): 712–27p.
18. Ekman P. Universals and Cultural Differences in Facial Expressions of Emotion. *Nebraska Symposium on Motivation*. 1972; 207–282p.
19. Hancock TJ, Landrigan C, Silver C. Expressing Emotion in Text-Based Communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'07*. 2007; 929–932p.
20. Plutchik R. Emotion: Theory, Research and Experience. In: *Theories of Emotion*. Academic Press; 1980; 11: 399p.
21. Izard EC, Libero ZD, Putnam P, *et al*. Stability of Emotion Experiences and Their Relations to Traits of Personality. *J Pers Soc Psychol*. 1993; 64(5): 847–860p.
22. Alm OC, Roth D, Sproat R. Emotions from Text: Machine Learning for Text-Based Emotion Prediction. *Proc Hum Lang Technol Conf Empir Methods Nat Lang Process*. Oct 2005; 579–586p.
23. Russell JA. A Circumplex Model of Affect. *J Pers Soc Psychol*. 1980; 39(6): 1161–1178p.
24. Bradley MM, Greenwald KM, Petry MC, *et al*. Remembering Pictures: Pleasure and Arousal in Memory. *J Exp Psychol Learn Mem Cogn*. 1992; 18(2): 379–390p.
25. Watson D, Tellegan A. Towards a Consensual Structure of Mood. *Psychol Bull*. 1985; 98(2): 219–235p.
26. Mehrabian A. Pleasure-Arousal-Dominance: A General Framework for Describing and Measuring Individual. *Current Psychology*. 1996; 15(4): 505–525p.
27. Lövheim H. A New Three-Dimensional Model for Emotions and Monoamine Neurotransmitters. *Med Hypotheses*. 2011; 78(2): 341–8p.

28. Li H, Pang N, Guo S, *et al*. Research on Textual Emotion Recognition Incorporating Personality Factor. In *Robotics and Biometics, 2007. ROBIO 2007. IEEE International Conference on*. 2007; 2222–2227p.
29. Kao CCE, Liu C-C, Yang T-H, *et al.* Towards Text-based Emotion Detection A Survey and Possible Improvements. In *2009 International Conference on Information Management and Engineering*. 2009; 70–74p.
30. Ortony A, Turner JT. What's Basic about Basic Emotions? *Psychol Rev.* 1990; 97(3): 315–331p.
31. Strapparava C, Valitutti A. WordNet-Affect: An Affective Extension of WordNet. *Proceedings of the 4th International Conference on Language Resources and Evaluation*. 2004; 1083–1086p.
32. Bradley MM, Lang JP. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*. Psychology, Technical (C-1). The Center for Research in Psychophysiology, University of Florida. 1999; 1–49p.
33. Mohammad MS, Turney DP. Crowdsourcing a Word Emotion Association Lexicon. In *Comput Intell*. 2013; 29(3): 436–465p.
34. Mohammad MS, Kiritchenko S. Using Hashtagsto Capture Fine Emotion Categories from Tweets. *Comput Intell*. 2015; 31(2): 301–326p.
35. Whissell C, Fournier M, Pelland R, *et al*. A Dictionary of Affect in Language. *Reliability, Validity, and Applications, Perceptual and Motor Skills.* 1986; 62(3): 113–888p.
36. Poria S, Gelbukh A, Hussain A, *et al*. Enhanced SentiNet with Affective Labels for Concept-Based Opinion Mining. *IEEE Intell Syst*. 2013; 28(2): 2–9p.
37. Staiano J, Guerini M. Depechemood: A Lexicon for Emotion Analysis from Crowd-Annotated News. *CoRR.* 2014; (1405-1605).
38. Alm OC, Roth D, Sproat R. Emotions from Text: Machine Learning for Text-Based Emotion Prediction. In *Proceedings of the Conference on HLT-EMNLP*. 2005; 579–586p.
39. Neviarouskaya A, Prendinger H, Ishizuka M. Recognition of Affect, Judgment, and Appreciation in Text. Affect Analysis Model: Novel Rule-Based Approach to Affect Sensing from Text. *Nat Lang Eng*. 2011; 17(01): 95–135p.
40. Neviarouskaya A, Prendinger H, Ishizuka M. Recognition of Affect, Judgment, and Appreciation in Text. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*. 2010; 806–814p.
41. Boldrini E, Mart´ınez-Barco. *EMOTIBLOG: A Model to Learn Subjective Information Detection in the New Textual Genres of the Web 2.0-Multilingual and Multi-Genre Approach.* Ph.D. Dissertation, University of Alicante; 2012.
42. Liew YJS, Turtle RH, Liddy DE. EmoTweet-28: A Fine-Grained Emotion Corpus for Sentiment Analysis. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*. 2016.
43. Mohammad MS, Zhu X, Kiritchenko S, *et al*. Sentiment, Emotion, Purpose, and Style in Electoral Tweets. *Inf Process Manage*. Jul 2015; 51(4): 480–499p.
44. Nozza D, Fersini E, Messina E. A Multi-View Sentiment Corpus. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017)*. 2017; 272–280p.
45. Mohammad S. #Emotional tweets. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics, Vol. 1. Proceedings of the Main Conference and the Shared Task, and Vol. 2. Proceedings of the 6th International Workshop on Semantic Evaluation (SemEval 2012),* Montreal, Canada. Association for Computational Linguistics. 7–8 Jun 2012; 246–255p.
46. Wang W, Chen L, Thirunarayan K, *et al*. Harnessing Twitter "Big Data" for Automatic Emotion Identification. In *International Conference on Social Computing (SocialCom).* 2012.
47. Strapparava Carlo, Mihalcea Rada. Learning to Identify Emotions in Text. In *Proceedings of the 2008 ACM Symposium*

on Applied Computing, SAC'08, New York, NY, USA. 2008; 1556–1560p.
48. Strapparava C, Mihalcea R. Learning to Identify Emotions in Text. *Proceedings of the 2008 ACM Symposium on Applied Computing (SAC08)*. 2008.
49. Balahur A, Hermida MJ, Montoyo A. Detecting Emotions in Social Affective Situations Using the EmotiNet Knowledge Base. *Advances in Neural Networks Lecture Notes in Computer Science (LNCS) Vol. 6677*. 2011; 611–620p.
50. Jerome Bellegarda R. Emotion Analysis Using Latent Affective Folding and Embedding. In: *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Association for Computational Linguistics. 2010; 1–9p.
51. Wang Xuren, Zheng Qiuhui. Text Emotion Classification Research Based on Improved Latent Semantic Analysis Algorithm. In *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*, Paris, France. Atlantis Press; 2013; 210–213p.
52. Quinlan JR. Induction of Decision Trees. *Mach Learn*. 1986; 1(1): 81–106p.
53. Medhat W, Hassan A, Korashy H. Combined Algorithm for Data Mining Using Association Rules. *Ain Shams J Electr Eng*. 2008; 1(1).
54. Cortes C, Vapnik V. Support-Vector Networks. *Mach Learn*, Springer. 1995; 20(3): 273–297p.
55. Ruiz ME, Srinivasan P. Hierarchical Neural Networks for Text Categorization (Poster Abstract). *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.* 1999; 280–28p.
56. Kang H, Yoo SJ, Han D. Senti-Lexicon and Improved Naïve Bayes Algorithms for Sentiment Analysis of Restaurant Reviews. *Expert Syst Appl, Elsevier*. 2012; 39(5): 6000–6010p.
57. Kaufmann M, Max Align J. A Maximum Entropy Parallel Sentence Alignment Tool. *COLING'12 (Demos*). 2012; 277–288p.
58. Agichtein E, Gravano L. Snowball. *Proceedings of the 5th ACM Conference on Digital Libraries (DL00)*. 2000.
59. Brin S. Extracting Patterns and Relations from the World Wide Web. *Lecture Notes in Computer Science (LNCS). The World Wide Web and Databases*. 1999; 172–183p.
60. 62Changhua Yang, Hsin-Yih Lin Kevin, Hsin-His Chen. Emotion Classification Using Web Blog Corpora. In *Web Intelligence, IEEE/WIC/ACM International Conference on.* 2007; 275–278p.
61. Aman S, Szpakowicz S. Identifying Expressions of Emotion in Text. *International Conference on Text, Speech and Dialogue*, Springer. 2007; 196–205p.
62. 63Chaffar S, Inkpen D. Using a Heterogeneous Dataset for Emotion Analysis in Text. *Canadian Conference on Artificial Intelligence, Springer*. 2011.
63. 64Martin Sykora D, Thomas Jackson W, Elayan Suzanne. Emotive Ontology: Extracting Fine-Grained Emotions from Terse, Informal Messages. *IADIS International Journal on Computer Science and Information Systems (IJCSIS)*. 2013; 8(2): 106–118p.
64. Suttles Jared, Ide Nancy. Distant Supervision for Emotion Classification with Discrete Binary Values. In: Alexander Gelbukh, editor. *Computational Linguistics and Intelligent Text Processing*. Vol. 7817 of Lecture Notes in Computer Science (LNCS), Springer Berlin Heidelberg. 2013; 121–136p.

Cite this Article
Aabiroo Bader, Amjad Husain Bhat. Emotion Evaluation from Text: A Survey. *Journal of Artificial Intelligence Research & Advances*. 2018; 5(2): 8–20p.