Emotion Analysis from Text: A Survey

Vaibhav Tripathi[1] **Aditya Joshi**[1,2,3] **Pushpak Bhattacharyya**[1]
[1]IIT Bombay, India
[2]Monash University, Australia
[3]IITB-Monash Research Academy, India
{vtripathi,adityaj,pb}@cse.iitb.ac.in

Abstract

Computational analysis of emotions has been considered a challenging and interesting task. Researchers rely on various cues such as physiological sensors and facial expressions to identify human emotions. However, there are few prior works who work with textual input to analyse these emotions. This survey attempts to summarize these diverse approaches, datasets and resources that have been reported for emotional analysis from text. We feel that there is an essential need to have a collective understanding of the research in this area. Therefore, we report trends in emotion analysis research. We also present a research matrix that summarizes past work, and list pointers to future work.

1 Introduction

The idea of enabling computers with emotions was first discussed by Picard (1997). This paper titled **"Affective Computing"** introduces foundational ideas in incorporation of affect for computers - from perspectives of both generation and detection. Initial research in building affective computers focused on human studies that investigate cognition, psychology and behaviors of humans. Such an analysis of emotions from humans often involves use of passive sensors which capture data about the user's physical state or behavior. The data gathered is analogous to the cues humans use to perceive emotions in others. Picard et al. (2001; Nasoz et al. (2004; Schubert (1999) For example, a video camera can be used to capture facial expressions, body posture and gestures, while a microphone captures speech. Other sensors detect emotional cues by directly measuring physiological data, such as skin temperature.

Building affective computers requires multi-modal processing since there are a wide variety of cues to be captured. However, in the current scenario, the above-mentioned approaches are impeded by their requirement of such multi-modal data which may not be readily available. On the contrary, text-based approaches to emotion analysis have become popular. In this survey paper, we describe such emotion analysis approaches. The survey paper is aimed at researchers aiming to begin their exploration in emotion analysis.

Emotion analysis (EA) from text is the task of predicting emotion in a piece of text. The proliferation of emotion analysis approaches has been motivated by the rise of Web 2.0. Due to popularity of social media, people express emotions on the web these days. In addition, weblogs, discussion forums and comments are easily accessible. All these forms of text are available to people interested in emotion analysis research, who can apply variable natural language processing algorithms on such text data.

2 Emotive Potential of Text

Schwarz-Friesel (2015) state that text can capture but only a portion of emotions expressed by a human. Even in linguistic studies of emotions, emphasis has been given on intonation and accentuation of speech for emotion analysis. While these cues are absent from text, there are only two modes possible for the expression of emotion. They state these two modes as follows:

- **Emotive vocabulary:** The simplest way to express emotions through text is to use words from emotive vocabulary for example, *love, hate, good etc.* These words refer directly to emotions as symbols. Consider the example '*Excited to watch 'Finding Dory' this week-*

end' contains the word 'excited' which is a direct indicator of emotion. However, much attention must be given here to the sense associations of these words in multiple contexts. This also includes idiomatic phrases like '*a cup of cake*'.

- **Affective items:** Sometimes the expression of emotion does not include words from the emotive vocabulary. Interjections such as *ugh, uh-oh* and expletives can be useful indicators of emotion, especially in modern social media domains and microblogs. This is seen from the fact that many past approaches use online dictionaries of colloqial language like `www.urbandictionary.com`. Affective items may also be expressed in action words in text. For example, use of brackets or asterisks are commonly used as action indicators. For example, '*(rolls eyes) Yeah right!*'.

It must also be noted that expression of emotion in text takes place in all linguistic levels in text and discourse:

1. Morphological:
2. Lexical (As stated above)
3. Syntactic (as in the case of '*It was boring initially but later it picked up so well!*')
4. Figurative (as in the case of sarcasm, etc.)

As we move up the layers, automatic prediction of emotions becomes more challenging.

3 Basic emotions

In order to understand emotions, we look at literature from psychology that describes different properties of emotions. In the first subsection, we look at properties of emotions while in the second subsection, we look at different ways of representing emotions.

3.1 Properties of Emotions

(**?**) defined four key properties of emotions:

1. **Antecedent**: The antecedent is the event or situation that causes a given emotion. It acts as a trigger to the emotion. For example, in case of sadness, the antecedent is, say, the death of one's pet.
2. **Signal**: The signal is the physiological method that a human uses to express an emotion. A signal is generated when a person expresses a specific emotion. For example, in case of sadness, tears are one of the signals.
3. **Response**: The response is the expected, conventional reaction to an emotion. A response is generated when a person understands a given emotion in another person. In case of sadness, if a person sees another sad, one likely response is that the person consoles the sad person.
4. **Coherence**: Coherence indicates that a given emotion has similar antecedents, signals and responses across different living beings. Simply put, similar things are likely to make a dog and a human sad (antecedents), both dog and human are likely to express sadness in similar ways (signals), etc.

Similarly, if were to consider the emotion '*surprise*, an unexpected gift would be an antecedent, having one's mouth wide open a signal, while feeling happy or amused would be a response. The fact that most humans and animals experience and express surprise in a similar manner indicates the coherence of the emotion.

3.2 Representations of Basic Emotions

There are three ways in which emotions have been represented: (a) A simple list of basic emotions where two emotions are not related to each other, (b) Emotion dimensions where emotions are arranged as points in a n-dimensional space, and (c) More complex structures like the Plutchik wheel of emotions.

3.2.1 Basic Emotions

The binary basic emotions are happiness and sadness. (**?**) indicates six basic emotions, namely anger, disgust, fear, sadness, happiness, and surprise. (**?**) indicates eight basic emotions, namely anger, disgust, fear, sadness, happiness, surprise, anticipation, and trust.

3.2.2 Emotion Dimensions

Emotions may be organized as points on multiple dimensions. (**?**) gives two dimensions along which emotions may be arranged. These dimensions are: **Valence** and **Arousal**. Valence indicates how pleasant or unpleasant an emotion is, while arousal indicates whether or not the person experiencing an emotion feels in control of themselves. Valence divides emotions into pleasant and unpleasant emotions. For example, emotions such as '*excitement, happiness, contentment*' are positive emotions while emotions like '*stress, sadness, boredom, fatigue*' are negative emotions.

Arousal is not related to the magnitude of the emotion, but to the level of control. Consider the following example. A person who is excited and jumping out of excitement is not necessarily more happy than a person who is calmly meditating with eyes closed. This example shows that arousal is merely linked with an emotional state of excitement or calmness. The emotion dimensions allow us to observe similarities (along both the axes) and differences between emotions, instead of looking at them as a simple list of basic emotions.

3.2.3 Wheel of Emotions

(**?**) presented a wheel of emotions as given in Figure **??**. There are three components of the wheel:

1. **Basic emotion pairs**: Consider the emotions on the second ring from inside. These emotions are: joy, trust, fear, surprise, sadness, disgust, anger and anticipation. They are arranged in a ring indicating antonymy between pairs of emotions. Thus, the wheel states that joy and sadness are opposite emotions, as are anger and fear, anticipation and surprise and trust and disgust.

2. **Emotion Lobes**: The wheel consists of four lobes, each corresponding to a basic emotion pair. Consider the vertical lobe that consists of serenity, joy, ecstacy, grief, sadness and pensiveness. Serenity is a form of deactivated joy whereas ecstacy is a form of activated joy. Similarly, pensiveness is a form of deactivated sadness while grief is a form of activated sadness. Thus, the central circle of emotions rage, vigilance, ecstacy, admiration, terror, amazement, grief and loathing are the '*activated*' basic emotions, whereas the emotions on the outer ring, namely serenity, acceptance, apprehension, distraction, pensiveness, boredom, annoyance and interest are the '*deactivated*' basic emotions.

3. **Combination of Emotions**: The wheel of emotions also shows how basic emotions combine in order to form more complicated emotions. These emotions are listed outside each of the lobes on the outer ring. These combined emotions are love, submission, awe, disapproval, remorse, contempt, aggressiveness and optimism. Consider the emotion '*love*'. The wheel represents love as a combination of serenity and acceptance, two basic emotions.

The Plutchik wheel of emotions allows us to understand relationships between different basic and complex emotions in terms of a complex structure as a wheel. One peculiar problem that arises in the task of emotion analysis is that cognitive psychologists do not seem to agree on the number of basic emotions in humans. The idea of 'basic' emotions is similar to the concept of primary colors in color theory that is, once we decide on the set of basic emotions, all other emotions can then be considered as combinations of these basic emotions. Unfortunately, there is no consensus regarding the basic emotions yet. The following theories exist, however:

Ekman identified six basic emotions on the basis of facial expressions - *Anger, Happiness, Sadness, Surprise, Disgust, Fear*. Ekmans six basic emotions are the most widely used categorization in literature even today.

3.2.4 Other models of emotion

In addition to these models, others that have been reported are:

- **Parrott's hierarchy of emotions:** Parrott created a tree-structured list of more than 100 emotions, organising them into three levels of primary, secondary and tertiary emotions.

- **Rachel E. Jack's theory:** Rachel E. Jack analyzed the 42 facial muscles which shape emotions in the face and concluded that there cannot be more than four basic emotions.

It is interesting that most researchers believe that the number of basic emotions should be determined by the number of different emotions that a human face is able to express. In this paper, the datasets used in prior work also has varied number of emotions, with Ekman's six emotions being the most popular.

4 Datasets

In this section, we discuss datasets that have been introduced for computational studies in emotion analysis. We divide them into two: short text and long text.

4.1 Short Text

Most of the prior work uses short text as its target for emotion analysis research. This may be because longer text brings in nuanced forms of emotion expression which may be difficult to detect, while short text is direct and pointed in its emotion expression. Consider the following: *"I'm extremely happy for you, she said and quickly turned around, walking away slowly, so that he shouldn't notice her tears."* The most common short text used in emotion analysis research is news headlines Strapparava and Mihalcea (2007; Bellegarda (2010), followed by microblogs Aman and Szpakowicz (2007a; Chaffar and Inkpen (2011). Hashtag-based supervision is a popular technique for microblogs where hashtags indicated by authors of a tweet are used as its labels. For example, a tweet containing '#happy' is considered as a happy tweet.

4.2 Long Text

There have been few works that analyze long text for emotions. The most notable among these is Liu et al. (2003), who work with emails. Alm (2008) work on children's stories, but instead of treating one story as a large data sample, they break it down into sentences. The annotations is done for each sentence separately.

A few researchers such as Kang and Ren (2016) have also tried to explore datasets in languages other than English. This is important since emotive expressions vary greatly across languages.

5 Lexicons for EA

An emotion lexicon is a knowledge repository containing textual units annotated with emotion labels. In its simplest form, it may exist as a set of word lists. For example, we may have a word list corresponding to happiness. Such a word list would include words 'happy', 'happiness', 'excited', 'pleased', etc.

Lexicons are useful as a knowledge base to understand emotion in a language. They can be used directly in a simple rule-based emotion analysis system.

In this section, we describe six popular emotion lexicons, namely: (a) LIWC, (b) EmoLexi, (c) Wordnet-Affect, (d) ANEW, (e) ANEW for Spanish, and (f) Chinese emotion lexicon. For each lexicon, we describe what the lexicon contains, and the process that was used to create the lexicon. Following this, we then describe some common trends in emotion lexicon generation in terms of approaches used.

5.1 LIWC

Linguistic Inquiry and Word Count (LIWC) system by (**?**) is a popular text processing system. It provides a software to understand topic, sentiment, etc. in a given piece of text. LIWC is widely used as an off-the-shelf system by emotion-allied applications in order to predict sentiment/emotion in a given piece of text. In terms of research, LIWC is often cited as a baseline result.

5.1.1 Schematic Description

The core dictionary of LIWC consists of 4500 words and word stems organized in 4 categories. A word stem in this case corresponds to a regular expression. For example, if the word stem '*happ**' is known to be related to the emotion happiness, all forms of happ get included in the emotion. This includes the adjective, adverb, noun form, etc. The four categories of words in LIWC are:

1. **Non-sentiment categories**: The first three categories of words are non-sentiment-related. They are, namely:
   (a) **Linguistic processes**: These are pronouns, prepositions, conjunctions and other function words.
   (b) **Speaking processes**: These are words related to speech. They include words that arise from speech such as fillers (e.g. 'hmm') or interjections (e.g. 'wow!').
   (c) **Personal concerns**: These are words that define the topic of a given piece of text. These include word lists related to work, home, etc.
2. **Sentiment categories**: The fourth category of word consists of psychological processes, and incldues words dealing with affect and opinion. This category is divided into the following sub-categories:

Prior Work	Number of emotion labels	Domain of dataset	Approach Summary	Features Used/Lexicons
Olveres et al. (1998)	6	-	Rule Based (Natural Language Parsing)	-
Liu et al. (2003)	6	Emails	Real-world knowledge concept models (Commom Sense Affect)	-
Strapparava and Mihalcea (2007)	6	News headlines	Rule-based; Unsupervised Classification	Word-Net Affect
Yang et al. (2007)	10	Blogs	Supervised classification (SVM, CRF)	Custom-made lexicon
Aman and Szpakowicz (2007a)	7	Blogs	Supervised classification (Naive Bayes, Support Vector Machines)	Word-Net Affect, General Inquirer, Features such as presence of punctuations (?,! etc.)
Alm (2008)	8	Stories	Supervised Classification	Bag of words features; Roget's thesaurus
Strapparava and Mihalcea (2008)	6	News	Supervised (Naive Bayes) & Unsupervised (LSA) Classification	Lexical features, Word-Net Affect
Bellegarda (2010)	6	News	Supervised (NB, SVM, Decision Trees); New approach based on LSM	Bag of words features
Ghazi et al. (2010)	6	Blogs	Hierarchical Classifiers	Bag of words features
Chaffar and Inkpen (2011)	6	News, blogs, health, diary posts	Supervised Classification (Naive Bayes, SVM, Decision Trees)	Bag of words features; Word Net Affect
Kang and Ren (2016)	8	Chinese blogs	Hierarchical Bayesian Models	-

Table 1: Research matrix showing prior work in emotion analysis from text

(a) **Cognitive Processes**: This includes categories of words that express subjectivity or cognition. The cognitive process word lists correspond to possibility (e.g. 'possible'), certainty, (e.g. 'definitely') or inhibition (e.g. 'prevented'). The possibility and certainty words indicate emotion and also signify the probability of the truth value of the statement that follows. The inhibition words, on the other hand, indicate a reversal in the emotion expressed in the statement that follows.

(b) **Affective Processes**: This includes categories of words that express emotion/affect. The different sub-categories within the set of affective processes correspond to anxiety, anger, sadness, positive emotion and negative emotion.

LIWC consists of a total of 713 words corresponding to cognitive processes and 915 words corresponding to affective processes. As seen, the words are arranged in a hierarchy of categories.

5.1.2 Generation

LIWC is a manually created emotion lexicon. The steps of creating LIWC are:

1. **Define Category Scales**: The first step is to determine the hierarchy of categories that forms the framework of the lexicon. The categories were first determined. Then, they were grouped into a hierarchy by establishing commonalities between different categories.

2. **Populate Manually**: For each of the category determined, the lexicographers then populate it with a list of words. For every word to be added to a category, three judges manually determine which category it must be added to. For each word that must possibly be added to a given category, the authors determine if the word should be placed specifically to this category or moved higher in the hierarchy.

5.2 ANEW

A New English Wordlist (ANEW) is a lexicon by (**?**). It is a dictionary that consists of about 1000 words, each of which is annotated in terms of a 3-tuple: pleasure, arousal and activation.

5.2.1 Schematic Description

The attribute '*pleasure*' indicates whether the given word indicates a positive or a negative emotion. The attribute '*arousal*' indicates the intensity of the pleasure expressed using the word. Finally, the attribute '*activation*' indicates whether the given word gives an experience of being in control or not. For example, along this 3-tuple definition, the word '*afraid*' will be indicated by the tuple: Negative, 3, not. This means that the word '*afraid*' evokes a negative emotion, has a magnitude of 3 while the user using this word is not in control of their own emotions.

5.2.2 Generation

Like LIWC, the ANEW was also created manually. However, while LIWC was created top-down (i.e. the structure determined first, and then populated), ANEW generation proceeds in a bottom-up manner. This means that the annotators begin with a list of words for which they need to provide annotations along each of the three axes. The process of generation of ANEW involved a total of 25 annotators, and was as follows:

- Each experiment was conducted in runs of 100/150 words.

- For each word, an annotator uses a ScanSAM sheet as shown in Figure **??**. The ScanSAM sheet requires that an annotator mark whether the emotion is pleasant or not, aroused or not, and in control or not.

5.3 ANEW for Spanish

ANEW has been adapted to multiple languages, one of them being the Spanish version presented by (**?**). The ANEW-Spanish corpus, like ANEW, consists of words annotated with 3-tuple: pleasure, arousal and control. The ANEW for Spanish lexicon uses the English lexicon as a basis.

5.3.1 Generation

- **Bootstrap lexicon**: All 1302 words in the ANEW lexicon are translated into Spanish using automatic translation. The assumptions here are that the emotion of a word remains the same across languages, and that the translation is able to correctly determine the right sense of the word and hence, select the appropriate word in the target language.

- **Evaluatory annotation**: The annotators then merely '*correct*' the values determined by this translated version of ANEW.

In order to establish the quality of the new corpus ANEW-Spanish, the authors show the correlation between ANEW and ANEW-Spanish for the three attributes. They observe a correlation of 0.916 on valence, 0.746 on arousal and 0.720 on dominance. This indicates that while a positive emotion word in English is likely to remain a positive emotion word in Spanish, its magnitude or activation may differ.

5.4 Emo-Lexicon

Emo-Lexicon was proposed by (**?**). It was a lexicon created through crowd-sourcing by leveraging online crowd-sourcing portals such as Amazon Mechanical Turk. In a crowd-sourcing portal, the authors set up a task of obtaining annotations for their lexicon and assign workers throughout the world to create the lexicon. The lexicon resulting from this task consists of 10000 terms.

6 Schematic Description

The Emo-Lexicon consists of word-senses annotated with a set of given emotions. The 'textitword-sense' portion is important because emotion is specifically associated with a sense of word as opposed to the word alone. (Consider the example of the word '*deadly*'. A '*deadly*' snake evokes a negative emotion, while a '*deadly*' spinner evokes a positive emotion). The '*a set of given emotions*' indicates that the annotator annotates whether each emotion is present or not. This structure allows that a word may belong to more than one emotions.

7 Generation

Since Emo-Lexicon uses a crowd-sourcing platform for generation, it is susceptible to quality flaws. Workers on the internet who take up the task may (a) not be aware of a given word, or (b) not perform the task with the attention that may be necessary. In order to get around these quality control issues of crowd-sourcing, the lexicon generation uses a modified process of obtaining annotation. The process is described as follows:

1. **Collect list of words**: The authors first collect words from a thesaurus. These words are then compared with General Inquirer and Wordnet to select word-sense pairs. This results in the primary dataset that requires word-senses to be annotated with each of the emotions.

2. **Quality Control**: This step is to ascertain that a worker is conversant with a given sense of the word. In order to do so, the task of annotation of a word starts with quality control step. The worker is shown a target word and four words are displayed along with it. The worker is then asked to select one that is closest to the target words. The words given as options may either be synonyms or related words. If the worker is able to correctly identify the target word that is expected, the worker gets to annotate a given word with emotion. If not, the word-sense pair is discarded.

3. **Obtain emotion annotation**: The worker is then displayed a list of questions, each corresponding to whether or not a given emotion is present in the given word. This allows a word to be annotated with more than one emotion.

7.1 Wordnet-Affect

Wordnet-Affect by (**?**) is an annotated version of the Wordnet, a linked lexical repository that uses a semantic concept at the core. The semantic concept of Wordnet is called a synset which consists of a set of synonymous words that represent the concept.

7.1.1 Schematic Description

The Wordnet-Affect consists of a portion of Wordnet annotated with affective-labels, indicated as a-labels. An a-label may correspond to an emotion-specific label as happy, sad, or a set of cognition-specific label such as possibility, certainty, impossibility, etc. The Wordnet-Affect consists of 2874 synsets annotated with a-labels.

7.1.2 Generation

Unlike the previous lexicons, Wordnet-Affect was created in a semi-automated manner. In other words, the algorithm of generation uses a combination of manual annotation and automatic annotation in order to efficiently generate a good-quality lexicon. The steps used for generation of Wordnet-Affect are as follows:

1. **Create a set of core synsets**: This core of synsets are the ones whose emotions can be accurately and definitively determined the annotators. This forms the manual annotation component of the lexcion.

2. **Annotate with a-labels**: The core synsets from step 1 are annotated with a-labels. Each synset is assigned at most one a-label.

3. **Projection**: Synsets whose emotion is known are used along with the Wordnet graph structure. The label of a synset is projected to other synsets using Wordnet relations. Thus, for synsets connected through relations like troponymy in Wordnet, the a-label remains the same.

4. **Manual evaluation**: The resultant expanded lexicon is then manually evaluated and corrected, wherever necessary.

7.2 Chinese Emotion Lexicon

The Chinese Emotion Lexicon created by (**?**) consists of words labeled with emotions. The lexicon was created using a semi-automated method, like Wordnet-Affect. However, Wordnet-Affect used the graphical structure of Wordnet and its relations. This was not feasible in the case of the Chinese language. The authors devise a unique method in order to project the emotion of a label to another.

7.2.1 Generation

The steps of creation of the Chinese emotion lexicon are as follows:

1. Select a core set of words. Obtain labels for these words. This forms the manual annotation component of the lexicon.
2. The core set of words is iteratively expanded using a similarity matrix. The iterations are stopped when the algorithm converges, i.e., when no new words can be appended to the existing lexicon.

The similarity matrix that forms the basis of expansion gives how close two words are. The similarity value (that is present in each cell of the matrix) between two words is determined by three kinds of similarity:

- **Syntagmatic similarity**: If two words cooccur frequently in a large text corpora, they are likely to be similar. Hence, their similarity value in the similarity matrix is changed appropriately.
- **Paradigmatic similarity**: If two words are related to each other in a semantic dictionary, they are likely to be similar to each other. Hence, in order to incorporate paradigmatic similarity, the authors use a semantic dictionary that allows synonyms and antonyms to be looked up.
- textbtLinguistic peculiarity: Since Chinese involves words corresponding to symbols and combining with other words to form new words, the authors use syllable overlap as a basis of determining similarity between two words. This is similar to the LIWC approach towards word stems associated with emotion labels. Consider the example of the word 'happy'. If we know that the word 'happy' corresponds to a positive emotion, the word '*happily*' also corresponds to positive emotion because of an overlap of four consonants in the word, specifically '*happ*'.

8 Approaches

The approaches reported in emotion analysis from text can be summarized into the following: keyword spotting, lexical affinity and statistical NLP.

8.1 Keyword spotting

The most naïve approach and probably also the most popular because of its accessibility and economy is spotting keywords from an emotive vocabulary. Text is classified into affect categories based on the presence of fairly unambiguous emotion words like *"distressed", "enraged," and "happy."* Due to this, the use of emotion lexicons like WordNet Affect (Strapparava et al., 2004) is quite popular in prior work. (Chaffar and Inkpen, 2011; Aman and Szpakowicz, 2007a; Strapparava and Mihalcea, 2008) Lexical resources like this and Roget's Thesaurus (Roget, 2008) act like emotive vocabularies here. The emotion analysis approach is based on rules that use these vocabularies/dictionaries, apply linguistic modifications (such as negations) in order to make prediction about emotion.

8.2 Lexical Affinity

This approach is slightly more sophisticated than keyword spotting. Detecting more than just obvious affect words, the approach assigns arbitrary words a probabilistic "affinity" for a particular emotion. For example, "accident" might be assigned a 75% probability of being indicating a negative affect, as in "car accident." These probabilities are usually trained from linguistic corpora. Liu et al. (2003) Though often outperforming pure keyword spotting, we see two problems with the approach. First, lexical affinity, operating solely on the word-level, can easily be tricked by negations and different word senses. Second, lexical affinity probabilities are often biased toward text of a particular genre, dictated by the source of the linguistic corpora. This makes it difficult to develop a reusable, domain-independent model.

8.3 Statistical Natural Language Processing

Statistical approaches are arguably the the most common class of approaches in emotion analysis from text. By feeding a machine learning algorithm a large training corpus of annotated texts, it is possible for the system to not only learn the relationship between lexical entities and their valence, but also the role of paradigmatic features like punctuations (?,!,“haha”) in text Aman and Szpakowicz (2007b). Statistical methods such as latent semantic analysis (LSA) have been popular for affect classification of texts, and have been used in Strapparava and Mihalcea (2008)

However, statistical methods are generally semantically weak, meaning that, with the exception of obvious affect keywords, other lexical or co-occurrence elements in a statistical model have little predictive value individually. Also, they have little use in modern social networks where language lacks subtlety. Just like other natural processing subdomains, emotion analysis from text has also seen a rising affinity towards full machine-learning systems.

Traditional machine learning algorithms like Naive Bayes and Support Vector Machines seem ubiquitous in emotion research. Yang et al. (2007; Aman and Szpakowicz (2007b; Alm (2008; Bellegarda (2010; Chaffar and Inkpen (2011). Few prior works have also tried to move beyond traditional classification by making subtle improvisations. Yang et al. (2007) were the first to consider conditional random fields (CRFs) Lafferty et al. (2001) for emotion analysis. This basically means that they thought about taking advantage of the sequential nature in his data to capture context transfer from sentence to sentence. Another important idea is by Ghazi et al. (2010) who identified hierarchical structure in emotion labels and used hierarchical classifiers to classify weblogs into one of Ekman’s emotion classes.

9 Results observed in prior work

Strapparava and Mihalcea (2007) presented detailed results of three systems that participated in the SemEval 2007 Affective Text task. Out of these, one in rule based, and the other two are unsupervised and supervised systems respectively. They observed that the rule-based system performed the best in 4 out of 6 emotion classes, while the supervised system performed the best in other two. Aman and Szpakowicz (2007b) and Bellegarda (2010) show that support vector machines are better in the task of emotion analysis when compared to Naïve Bayes systems. They use datasets consisting of blogs and news respectively. The values are shown in Table 2.

Yang et al. (2007) showed that sequence labelers can outperform traditional classifier (SVM) on a dataset of blogs, increasing the accuracy from **32.88** to **43.35**.

	Naive Bayes	**Support Vector machines**
Aman and Szpakowicz (2007b)	72.08	73.89
Bellegarda (2010)	59.72	71.69

Table 2: Prior work showing the advantage of using SVMs to Naive Bayes (F-Scores)

An interesting result was presented by Ghazi et al. (2010), who showed that by identifying hierarchical relationships between emotion labels and then using a hierarchical classification for prediction can prove beneficial for the dataset given by Aman and Szpakowicz (2007b). For a two-level hierarchical classification, the accuracy increased from **61.67** to **65.5**.

10 EA Applications

In this section, we look at different applications of Emotion Analysis. Specifically, we discuss: (a) Application of EA to chat clients, (b) Application of EA to intelligent tutoring systems, and (c) Application of EA to mental health monitoring.

10.1 EA & shat clients

Emotion analysis has been used effectively in several chat applications. Chat-based clients have the property of being real-time. Hence, the response of these applications is also critical. There has been some work in integrating emotion analysis-aware mechanisms in chat systems. This section describes two of these works in emotion analysis for chat applications.

10.1.1 For real-time internet communication

This section describes work by (**?**). The authors propose a real-time internet communication mechanism that performs emotion analysis and presents them to the user in the form of “*expression images*” i.e., images that indicate a certain emotion. The emotion labels that the authors consider are angry, disgust, aversion, happy, sad, fear and surprise. **Input/Output**
The input to the system is a sentence while the output is an expression image. The expression image depends on both the emotion expressed in the

sentence and the intensity of emotion expressed therein.
The system is a three-stage process: i) input analysis system, ii) tagging system, iii) parser. The role of these is explained in the next subsection.
Approach
The emotion extraction system introduced in this work uses a rule-based approach in order to predict emotion in a sentence. The approach works as follows:

- A word is looked up in the dictionary.
- If the word is not present, the word is stemmed and then looked up for a possible match.
- If either of the above steps returns positive, the emotion is appropriately assigned.
- While words indicate sentiment, this sentiment may be intensified or diminished based on the words around them. Towards this, the authors use a parser to understand the syntactic structure of the sentence. Based on the syntactic structure, the emotion of the sentence is updated:
    - The authors disregard negation i.e., if a "*not*" is present before an emotion word, the emotion is assumed to be absent. However, the presence a "*not happy*" does not amount to sad, as per the implementation of this paper.
    - The authors use intensifiers in order to understand intensity of words. This is done by increasing the emotion intensity score of a word.

Once the emotion of a sentence has been identified, an appropriate expression image that resonates with the emotion of the sentence is chosen to be displayed in a chat engine.
Evaluation
This subsection discusses evaluation mechanisms for the chat application implemented in their paper. The authors use two kinds of evaluations: a static evaluation and a dynamic evaluation.

- The **static** evaluation consists of a set of paragraphs from 7 novels that are given as input to the system. The authors report that they are able to correctly extract emotion in 98% of the sentences.
- The **dynamic** evaluation involves simulation of a chat setup where annotators are indulged in a chat conversation. The annotators are also expected to give their response in terms of an emotion label. The authors discover that their system is able to identify the correct emotion in 90% of the sentences.

The authors observe that such a system can be used for real-time communication interfaces where response of a chatbot can be tweaked based on the emotion undergoing in a conversation.

10.1.2 Emotion Analysis & Internet Chat

This section is based on work by (**?**). The authors present an emotion analysis system for instant messages in chat clients. While majority of work in sentiment analysis involves larger text such as reviews in different domains, short text has received lower focus. However, since chat applications are popular and commonly used, it is essential and useful to perform emotion analysis of chat messages.
Dataset creation
The authors describe their technique of dataset creation in detail:

- **Data gathering**: They use two corpora for their experiments: (a) a Naval Post Graduate School chat corpus (NPS corpus) consisting of 7933 sentences, and (b) an Internet Relay chat (IRC) corpus consisting of 2980 sentences. All sentences are in XML format where each word in the sentence is also annotated with its POS tag.
- **Data normalization**: In chat applications, it is important that data is normalized accurately since users tend to misspell words and also use colloqial/slang words. Towards this, the authors used two internet-based translators for removing slang words and normalizing sentences. These translators are called noslang and transl8it. After translating sentences, they are POS-tagged again to remove errors if any because of shortened forms of words in the original sentences.
- **Data annotation**: The authors implemented a web interface to annotate these sentences. The sentences were annotated by one person and hence, interannotator agreement has not been reported. The output labels considered are: anger, happiness, disgust, fear,

surprise, sadness and neutral. The neutral label corresponds to the situation where no emotion was discovered. The distribution of sentences among emotion labels shows that about 47.8% of sentences in NPS corpus are neutral while 48.4% of sentences in the IRC corpus are neutral.

Feature Engineering
In order to predict emotions in chat messages, the authors propose a feature set of around 71 features for the corpora. The features are broadly classified into following categories:

- **Features obtained from data itself**
  - Part of speech counts for different POS
  - Frequency of different emoticons
  - Frequency of extended words to indicate pragmatic intensity of emotion
  - Frequency of punctuations
  - Counts corresponding to words such as length of longest and shortest word, etc.
- **External features**
  - **Affective keywords**: The authors use General Inquirer corpus that contains a list of affective keywords corresponding to three sentiment labels: positive, negative and neutral.
  - **Slang words**: The authors elaborately normalize slang words and take into considerateion the presence and frequency of such words.
  - **Similarity of words**: The authors report similarity of emotion words in a sentence with key emotion words corresponding to each of the emotion labels.

10.1.3 Evaluation

The authors report ten-fold cross-validation results for four sets of experiments: subjectivity detection (neutral v/s not neutral), sentiment detection (positive v/s negative), pure emotion detection (only emotion labels) and general emotion classification (emotion labels + neutral label). The authors train a set of classifiers such as Naive Bayes, J48 and weka to perform each of these tasks. The results of each of the classifiers are reported in the paper.

10.1.4 Trends in EA-based chat applications

Chat applications are interesting for emotion analysis because of their real-time nature. Emotion in a chat conversation is also a time series of sorts. This is because a chat conversation may tread through multiple emotions. Hence, emotion analysis of chat applications is important for temporal sentiment analysis. We also observe a considerable stress on normalization of data since chats tend to be ill-formed in terms of misspellings and slangs.

10.2 EA & intelligent tutors

Several automatic tutoring systems have been deployed for different educational applications. For these tutoring systems to be effective teachers, it is important that they possess skills of a traditional teacher. One of these skills is to understand how well the class is following you. (**?**) implement a tutoring system that is aware of the emotional state of its students. The key idea of the paper is that a person's mood/affective state can be understood from how coherent he/she is, in a dialogue.

10.2.1 Approach

In order to understand the coherence of a student in a dialogue, the paper implements a set of cohesion-based features to determine emotional state of a student. The goal here is to predict the affective state of the learner, as one among, boredom, flow, confusion and frustration. 'Flow' is a positive emotion label while others are negative emotion labels. It is important to note how the authors did not rely on basic emotion list as the list of output label but chose task-specific output labels instead. The dataset consists of textual dialogue transcripts of 28 undergraduate students using an automatic tutoring system called AutoTutor. The audio transcripts are converted to text using a speech-to-text converter.

Annotation
Each video in a lecture along with the transcript is presented to an annotator. Each video is split into intervals of 20 seconds, indicating that a student is likely to have the same emotion in a 20-second interval. With this granularity, an annotator is expected to determine the emotion label of each 20-second interal as neutral, confusion, boredom, frustration, delight, flow and surprise. The annotation is validated in three levels: a student annotates his/her own video with one of the output labels after the session, a student annotates another student's session with the same label, and a third expert annotator annotates the session. The resultant label is considered as an average across

annotations provided at these three levels.

Feature Engineering

The system used to predict emotional state of the learner consists of six kinds of features:

1. **Co-referential cohesion**: This includes refering to entities that the tutor used. This could also mean an overlap between nouns that the tutor and the learner used. For example, if the tutor says, “Now you click here” and the learner says, “Oh, do you mean I click here?”, it is likely to indicate that the learner is satisfactorily engaged.

2. **Pronoun referential cohesion**: This inclues refering to entities that the tutor used with appropriate entities. This means that the learner has understood what the tutor just said. For example, if the tutor says, “Do you see why Jim did it?” and the learner says, “Maybe because he saw gravity at work?”, it is likely to indicate that because the learner correctly identified the pronoun and referred to it, the learner is satisfactorily engaged.

3. **Connectives**: Conjunctions and connectives like ‘because’, ‘however’ also indicate cohesion in a conversation.

4. **Causal cohesion**: The authors use presence of words with a high causal cohesion ratio in order to record cohesion. For example, the words ‘killed’ and ‘die’ have a high causal cohesion ratio. Although they are two different words, if the tutor says ‘*He was killed*’ and the learner asks, “In what year did he die?”, it indicates good engagement.

5. **Semantic cohesion**: Semantic cohesion between words using LSA are measured.

6. **Readability measures**: Features such as verbosity, percentage of out of vocabulary features, use of fillers and pauses, etc. are also used as features.

The above features are specific to the task. It is also important to note that most of them indicate engagement i.e. the positive label.

Evaluation

The authors present the standardized coefficient weights between each of the features and the emotion labels. The authors observe that a positive correlation of 0.5 between boredom and incidence of negations indicating that bored learners are likely to use negation words. Also, the authors observe a negative correlation of -0.526 between noun overlap between adjacent sentences and frustration. This indicates that frustrated learners tend to have little overlap between their nouns. There was also a high positive correlation between causal ratio and flow indicating that a high causal ratio is a positive indicator of favorable affective state of a learner.

10.3 EA for mental health monitoring

Mental health issues pose risks to lives and wellness of millions of people. It is important that these issues are monitored and addressed well. In times of social media, users often use media like Twitter in order to express their emotion. This opens up opportunities to understand emotional well-being of people. This has resulted in several applications of EA to mental health monitoring. The 1st Workshop on Computational Linguistics and Clinical Psychology was held collocated with ACL 2014.

It is important that everyone is susceptible to mental health risks. Thompson et al (2014) talks about suicide risks in military officials. The authors state that fear prevents courage to commit suicides. Since military culture nurtures resilience and fearlessness, they are as susceptible to mental health risks as others. The goal here is to understand if emotion analysis can be used to predict or assess mental health risks. The expected deliverable is a mental health monitor for mental health illness X. A labeled dataset is used to train a classifier that predicts health risk of illness X for a set of unlabeled textual units. The general architecture that most papers use consists of:

- Obtain labeled dataset
- Decide the goal
- Obtain inputs from clinical psychology
- Implement the desired classifier/topic model

Obtain labeled dataset

The following datasets have been used for various studies that combine EA and mental health monitoring:

1. **Medical Transcripts**: This includes sentences like “Doctor, I had a sever pain in my head when I woke up this morning...”. These transcripts could be audio transcripts.

Thompson et al (2014) use medical transcripts of medical officers talking to therapists as a part of the Durkheim project. The output labels considered are whether the official was typical, contemplating suicide or had attempted suicide already. These transcripts could also be chat transcripts to a web-based doctor, as in the case of Howes et al (2014).

2. **Experience Descriptions**: This includes sentences like "I used to be low on Friday evenings. That was strange..". These descriptions could be as described on discussion forums. Ji et al (2014) use data from Aspies, a discussion forum that is used by autism patients and their family memebrs and caretakers. These discussion forum posts may contains experiences of patients, approaches suggested for treatments and how effective these approaches were.

3. **Written communications**: This includes text written by mental health patients themselves. Glasgow et al (2014) use a dataset consisting of threat notes sent to judges by undertrials. The task then is to predict whether the undertrial is suffering from a mental illness or not. On the other hand, the most popular form of written communication is social media text. Coppersmith et al (2014) use tweets of people. In order to obtain gold annotations, the authors select users who have declared a medical diagnosis for a mental health issue. The authors believe that this is a reliable way of obtaining annotations for tweets.

Decide the goal
The goal of the system is then determined and the approach is selected appropriately.

- If the task is to predict the risk of an individual to a given mental illness, a classifier may be used. Homan et al (2014) proposes a classifier that predicts suicide risks in twitter, in terms of four distress levels.

- If the task is to analyze aspects of a given illness, a topic model may be used. Rouhizadeh et al (2014) measure similarity of token overlap to detect linguistic idiosyncrasies in children with autism.

Obtain inputs from clinical psychology
Since the application lies at the confluence of clinical psychology and computational linguistics, it is also important to analyze the typical traits of the mental health issue being considered. This is useful in order to design features on the basis of these traits. For example, Orimeye et al (2014) predict Alzheimer's disease using medial transcript data. Since such patients are likely to have trouble making inflected forms of words, morphemes are used as features. Caines et al (2014) aim to identify linguistic impairments related to disfluency. The classifier uses features based on repetitions, revisions and pauses.

Implement the desired classifier/topic model
In this subsection, we would like to discuss in detail two works:

1. A classifier that predicts linguistic impairments due to progressive aphasia

2. Assessment of discussion forums about autism using an author-topic model

1. Classifier that predicts progressive aphasia
Primary progressive aphasia (PPA) is characterized by linguistic impairment without any other notable impairment. There are two subtypes of PPA: Semantic dementia characterized by fluent byt spared grammar and syntax, and Progressive non-fluent aphasia that involves reduced syntactic complexity, word-finding difficulties, etc. The output labels considered for this task are prediction of SD, PNFA and typical development.

The dataset consists of 24 patients with PPA and 16 typical individuals. Given a topic, the subjects were asked to describe the story of Cinderella. The description was recorded and later transcripted using a speech-to-text converter. The classifier used consists of the following features:

- **POS features**: # adjectives, nouns, etc.

- **Complexity features**: Depth of parse tree, etc.

- **CFG features**: Average phrase length, etc.

- **Fluency features**: Indicators for umms, etc.

- **Psycholinguistic features**: Age of language acquisition, etc.

- **Acoustic features**: Jitters, pause, etc.

- **Vocabulary richness features**

The evaluation shows that the complete feature set is able to achieve an accuracy of 96.3% for the distinction between SD and typical. The accuracy, as expected, is lowest for the classifier to predict between SD and PNFA. The CFG features perform the best in this case, with an accuracy of 79.2%.

2. Topic model that discovers topics in autism communities
Aspies Central Forum is a discussion forum where individuals with autism and their family, practitioners and caretakers write. The goal of this topic model is to discover topics that the users talk about on the forum.

A topic model based on LDA is discovered. Each post has words and author name as observed variables. Based on these, topic-author distributions and word-topic distributions are estimated. The following topics are discovered:

- weed marijuana pot smoking fishing
- empathy smells compassion emotions emotional
- relationship woman relationships sexual sexually
- classroom campus tag numbers exams
- yah supervisor behavior taboo phone
- depression believe christianity buddhism

A topic model, thus, helps in understanding different issues that are discussed on these forums.

11 Conclusion

In this paper, we presented a survey of the work done in the area of identifying emotions from textual data. We observeed various trends in research - There is a noticeable move from rule-based and unsupervised classifiers to supervised approaches. This must be because of the increased availability of emotion-annotated data. Very recently, the role of hierarchy is being explored in emotion analysis research.

Emotion analysis research has flourished significantly in the past few years, necessitating a look-back at the overall picture that these individual works have led to. Based on our survey, we present the following directions for the future:

11.1 Emotions from objective expressions

One of the primary issues with emotion analysis is that humans can express emotions in disguised ways, without using any emotive vocabulary. For example, *"My day job requires me to work 25 hours a day."* Of course, in this sentence, the speaker is expressing a hyperbolic negative sentiment, suggesting that he has to work too much for his job. [1] Such expressions can be captured only with the help of a real-world knowledge base. Work by Liu et al. (2003) was a good step in this direction. However, with new advancements in linguistic resources like FreeBase Bollacker et al. (2008), we can do a lot better.

11.2 Phrase-based emotion expression

Expressions arising from an interplay of multiple words for example, figurative expressions is not handled properly in prior work. Most of them use a bag of words feature model or n-grams which doesn't extend beyond a particular window size. Consider this: *"I am at an all time low in my life."* Keyword spotting, in such cases will not go a long way. We need sophisticated methods to capture such nuanced forms of sarcasm. With the ever-increasing power of computers, deep learning must come to the rescue.

11.3 Figurative expression of emotions

Sarcasm and other forms of figurative language are a challenge to emotion analysis. Sarcasm is a peculiar form of emotion expression in which the surface emotion is different (even opposite) than the implied emotion or sentiment. Consider this: *"I love you so much I want to shoot you."* A lot of modern systems pre-process a sentence with emotionality (or neutrality) detector before feeding it to a emotion analyzer. Likewise, there must be a figurative language detection pre-processing step.

References

Ebba Cecilia Ovesdotter Alm. 2008. *Affect in Text and Speech*. Ph.D. thesis, University of Illinois at Urbana-Champaign.

Saima Aman and Stan Szpakowicz, 2007a. *Identifying Expressions of Emotion in Text*, pages 196–205. Springer Berlin Heidelberg, Berlin, Heidelberg.

[1]The online classifier at http://text-processing.com/demo/sentiment/ gives it neutral.

Saima Aman and Stan Szpakowicz. 2007b. Identifying expressions of emotion in text. In *Text, Speech and Dialogue*, pages 196–205. Springer.

Jerome R Bellegarda. 2010. 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*, pages 1–9. Association for Computational Linguistics.

Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. Freebase: a collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 1247–1250. ACM.

Soumaya Chaffar and Diana Inkpen, 2011. *Using a Heterogeneous Dataset for Emotion Analysis in Text*, pages 62–67. Springer Berlin Heidelberg, Berlin, Heidelberg.

Diman Ghazi, Diana Inkpen, and Stan Szpakowicz. 2010. Hierarchical versus flat classification of emotions in text. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, CAAGET '10, pages 140–146, Stroudsburg, PA, USA. Association for Computational Linguistics.

Xin Kang and Fuji Ren. 2016. Understanding blog author's emotions with hierarchical bayesian models. In *2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC)*, pages 1–6. IEEE.

John Lafferty, Andrew McCallum, and Fernando CN Pereira. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data.

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

Fatma Nasoz, Kaye Alvarez, Christine L Lisetti, and Neal Finkelstein. 2004. Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cognition, Technology & Work*, 6(1):4–14.

Jimena Olveres, Mark Billinghurst, Jesus Savage, and Alistair Holden. 1998. Intelligent, expressive avatars. In *Proceedings of the First Workshop on Embodied Conversational Characters*, pages 47–55.

Rosalind W Picard, Elias Vyzas, and Jennifer Healey. 2001. Toward machine emotional intelligence: Analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(10):1175–1191.

R. Picard. 1997. *Affective Computing*. The MIT Press.

Peter Mark Roget. 2008. *Roget'S International Thesaurus, 3/E***. Oxford and IBH Publishing.

Emery Schubert. 1999. Measurement and time series analysis of emotion in music.

Monika Schwarz-Friesel. 2015. Language and emotion. *Emotion in Language: Theory–research–application*, 10:157.

Carlo Strapparava and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 70–74. Association for Computational Linguistics.

Carlo Strapparava and Rada Mihalcea. 2008. Learning to identify emotions in text. In *Proceedings of the 2008 ACM Symposium on Applied Computing*, SAC '08, pages 1556–1560, New York, NY, USA. ACM.

Carlo Strapparava, Alessandro Valitutti, et al. 2004. Wordnet affect:@inproceedingsbollacker2008freebase, title=Freebase: a collaboratively created graph database for structuring human knowledge, author=Bollacker, Kurt and Evans, Colin and Paritosh, Praveen and Sturge, Tim and Taylor, Jamie, booktitle=Proceedings of the 2008 ACM SIGMOD international conference on Management of data, pages=1247–1250, year=2008, organization=ACM an affective extension of wordnet. In *LREC*, volume 4, pages 1083–1086.

Changhua Yang, Kevin Hsin-Yih Lin, and Hsin-Hsi Chen. 2007. Emotion classification using web blog corpora. In *Web Intelligence, IEEE/WIC/ACM International Conference on*, pages 275–278. IEEE.