

Frame Semantics Evolutionary Model for Emotion Detection

Mohamed H. Haggag¹

¹ Computer Science Department, Faculty of Computers and Information, Helwan University, Egypt

Correspondence: Mohamed H. Haggag, Computer Science Department, Faculty of Computers and Information, Helwan University, Egypt. E-mail: mohamed.haggag@fci.helwan.edu.eg

Received: December 27, 2013 Accepted: January 16, 2014 Online Published: January 18, 2014

doi:10.5539/cis.v7n1.p136

URL: <http://dx.doi.org/10.5539/cis.v7n1.p136>

Abstract

Emotions play a significant role in identifying attitude, state, condition or mode of a particular circumstance. Textual data, in particular, involves emotional state and affective communication beside its informative contents. Emotion extraction from text has been potentially studied to stimulate and elicit articulation features. In this study, a machine learning emotion detection model is proposed for textual emotion recognition. A frame semantics approach is identified to extract knowledge from the text in an evolutionary process that improves the detection capabilities. Emotion detection process is controlled by a rule base; each of its entries is generated by pre-invoking event, action and resulting emotion state. Frame entities semantically collaborated to evaluate the frame emotion. Individual entities may arbitrary substituted by their synonyms or opposites if a candidate frame doesn't match any of the knowledge set. The proposed model proves considerable capability of recognizing emotions by referencing their semantic relations. Results showed better detection accuracy for the proposed model compared with variety of emotion approaches including keyword spotting, knowledge-based ANN and supervised machine learning models. Experiments indicated encouraging results over both binary emotion and multiple labels classifiers.

Keywords: emotion detection, frame semantics, machine learning, knowledge based artificial neural network, genetic algorithms, semantic analysis, wordnet, evolutionary model

1. Introduction

Early emotion extraction algorithms were similar in their recognition to that of text categorization and topic detection in natural language applications. Emotion detection process has two main steps; preprocessing and emotion extraction. Preprocessing is concerned with textual data processing to produce output that is used as input to a later stage. Examples of preprocessing tasks are filtering, tokenizing, stemming, and stop words removal. Emotion extraction objective is to define the emotion that best fit to the context. Emotion detection approaches are classified into keyword based approaches and machine learning approaches.

Keyword-based approach is concerned with extracting the emotional keywords that mostly recognize user's emotions for the input text. This approach is based on initial identification of emotion categories and storing their corresponding words in a dictionary. The challenge with keyword based approaches is defining all emotional keywords including ambiguous ones. A major drawback to this approach is the limitation to keywords predefined in the manually designed lexicon. This will limit emotion categories to those fixed in the lexicon. Moreover, if the text contains keywords that do not appear in the dictionary, then its emotion would not be defined. An example here is "she saw a tiger". None of the sentence words have the fear sense. As a result, keyword based approach cannot detect any emotional keyword and accordingly no emotion could be detected.

The use of keyword spotting features is applied by Cheongjae Lee and Gary Geunbae Lee (2007) to improve the performance of emotion recognition from the natural language dialogues for entertainment. The model is a hybrid classification approach of statistical feature-based and keyword spotting classifications. Linguistic and pragmatic features are considered in addition to keyword spotting features. This approach supports both binary classification and multiple labels classification. Binary emotion classification recognizes "Neutral" and "Emotional" classes, which can be significantly improved using the keyword spotting features model. Multiple labels classification is to identify, out of multiple, the best emotion.

Lexical affinity approaches depend on semantic parsing. The emotion of a linguistic unit is classified based on the affinity of the linguistic unit and an affective keyword. For example, if a phrase appears closely much more

often with the word “happy” than “sad”, it is reasonable to believe that the emotion associated with this phrase is happy. Such approach requires large scale corpus and a sophisticated search engine to get statistically significant results. The search engine needs not only to retrieve all documents with the phrases but also needs to determine the distance between the two phrases. Without the distance information, the approach can be trapped easily. In this regard, the machine learning algorithm proposed by Lu, Hsu, Peng, Chung and Ho (2010) is based on sentence level emotion classification. A corpus of affectively annotated texts is used for training. An adaptive learning algorithm has been presented by Jagadeesh, Bose and Nagaraja (2004) to improve the training times and the performance of KBANN by using regularization methods. Semantic role labeling is associated with web mining in Lu, Hong and Cruz-Lara (2006) to detect basic emotions using common sense knowledgebase, ConceptNet.

Machine learning approach overcomes keyword based approach limitations through textual features learning techniques. Although some emotion recognition systems used feature-based statistical classification, a set of features have been incorporated for emotion detection through the use of machine learning methods. Learning classifiers are trained by features extracted from the text. Emotion tags could be identified by relevant features, where features extraction approach is imposed to the emotion detection process. Although some of the machine learning models resulted in less accuracy than keyword based approaches, they provide powerful decisions where keyword approaches fail to recognize the proper emotion. A hybrid model that uses keyword-based and machine learning methods is proposed by Seol, Dong-Joo Kim and Han-Woo Kim (2008). EmotiNet knowledge base Balahur, Hermida, Montoyo and Muñoz (2011) was identified to detect emotions based on commonsense knowledge on concept, interactions and affective consequences. Data set is built from self-reported affective situations and extended with external sources of commonsense knowledge on emotion-triggering concepts.

Machine learning emotion detection approaches are categorized into supervised and unsupervised techniques. Supervised learning methods use large set of emotions-labeled to train the emotion classifier. The problem with supervised learning is the unavailability of large training emotional data sets. Accordingly, training provided on one domain could not serve other domains. For example, a supervised learning used in Alm, Roth and Sproat (2005) for emotion prediction with the SNoW learning architecture. The model aims to classify the emotional affinity of sentences in a limited domain of children’s stories. Specific features have been extracted resulting in a typical feature vector size with around 30 features. Emotion prediction is achieved according to the training features set, which is limited to children stories and not covering other document categories. Polynomial kernel method, one of the supervised machine learning methods, is proposed by Quan (2010) to recognize eight emotions (Expect, Joy, Love, Surprise, Anxiety, Sorrow, Angry and Hate) for a Chinese emotion corpus. A SVM classifier was trained by Pak, Bernhard, Paroubek and Grouin (2012) to perform multi-label classification. Each emotion is trained independently. The provided system used two emotion classification approaches; machine learning and rule-based approaches.

On the other hand, unsupervised learning emotion detection approaches detection does not rely on the size of training data. Instead, such approaches are not restricted to a fixed number of emotion categories. The model proposed by Agrawal and An (2012) involved a novel unsupervised context-based approach to detect emotion from text at the sentence level. The proposed methodology does not depend on any existing manually crafted affect lexicons such as WordNet- Affect. The model is flexible enough to classify sentences and compute an emotion vector for each potential affect bearing word. Such classification is based on semantic relatedness between words and various emotion concepts. The scores are then fine-tuned using the syntactic dependencies within the sentence structure.

Specification of concepts has been developed by Shivhare and Khethawat (2012) using emotion ontology. Ontologies are defined by entities, domain vocabulary, attributes and interrelationships. Ontologies provide an understanding of a particular domain, which facilitates communication between different stakeholders like persons, institutions, and application systems.

Building emotion Corpus based on emotional expression analysis is provided by Quan and Ren (2009) for Chinese emotional expression. The corpus is then used to explore the emotion expressions in documents. The annotation frame includes three levels document, paragraph, and sentence. Sentence level is the basic level for emotion annotation. A corpus of emotion-annotated data, provided by Aman (2007), included fine-grained distinctions of emotion categories, emotion intensity levels and emotion indicators in text. This corpus enables building emotion lexicon by selecting a primary set of emotion words and deciding on the similarity score of semantic relatedness between words. An evaluation of knowledge-based and corpus-based methods for emotion identification is proposed by Strapparava and Mihalcea (2008), through a construction of a large data set annotated for six basic emotions; anger, disgust, fear, joy, sadness and surprise.

According to the text length, emotion detection from text is categorized on the basis of word, sentence or document detection model. In word level emotion extraction, lexicon including its word lists is used for emotional word tagging. Accordingly, emotion corpus is constructed for emotional expressions contextual analysis. Many studies have been presented at this level. A tagging system is provided to extract emotions from expressive real time Internet communications (Picard, 1997). Word level approaches have been extended to include semantic relations for sentences. Emotion estimation net (ESiN) model (Tao, 2004) is on top of this by using new lexicon structure which defines different word types like Content Word, Metaphor Word, Modifier Word and Emotion Keyword. Tokuhisa et al. (2008) saved pre invoking event to the emotion instead of emotion keywords. They used corpus from the web with a reasonable accuracy, where its examples can be used to build a semantic content-based model for fine-grained emotion classification.

Sentence level approaches have been developed to identify emotion tags within each sentence. Early researches in sentence level were concerned with detecting the appearance of emotional keywords. Word level emotion tags could be utilized to provide semantic analysis at the sentence level. Following approaches included syntactic and semantic basis for emotional information, rather than only word level emotions. Emotional keywords and emotion modification words are defined by Chuang and Wu (2004) to provide Support Vector Machine (SVM) based emotion recognition. Other researchers used a machine learning approach or feature extraction approach to get out sentence level emotions. Document level emotion detection approaches identify the emotion tags expressing the document context. Word level and sentence level approaches could interchangeably contribute in this aspect.

Recently, some researches like Shivhare and Khethawat (2012) have been contributed to solve the problems of keyword based approach using machine learning. Combining keyword based and machine learning, higher accuracy is resulted compared to each of the individual approaches. Initially, the model validates whether input text has emotional keywords or not. If one or more emotional keywords exist, using emotional keyword dictionary, keyword-based approach will be followed. If the input text has no emotional keywords, machine learning emotion classifier will be used for emotion recognition.

Frame semantics is the study that is concerned with knowledge of language. It is a set of features and interactions of linguistic concepts like words, phrases and grammatical patterns that provide the frame in a cognitive structure. Missing of the knowledge associated with any of these concepts results in incomplete knowledge and interpretation of the frame as a whole. A semantic frame identifies these concepts through labels with specific linguistic expression, which are called frame elements.

In this study, a new semantic frame approach for textual emotion detection is proposed. The model is based on machine learning where semantic and syntactic information for frames elements is generically stored in a structured knowledge base. Learning is provided to train knowledge based artificial neural network (KBANN) to enhance the classification decision of the textual emotions. Emotions are detected through a matching process against the knowledge base entries. Alternative matches of frames are constructed for individual sentence by substituting each entity with all of its synonyms and opposites. A trade of performance versus accuracy is controlled by three types of matching algorithm; first matching, best matching, best opposite matching and average matching.

2. Proposed Model

The proposed emotion extraction model is based on new frame semantics approach. Labels are identified for all linguistic elements of the sentence. These labels are linked in a syntactic structure forming the semantic frame corresponding to the sentence. Frames are constructed for the input text and will be the linguistic base for the emotion detection process. The frame entities are constructed with the use of link parser and word sense disambiguation at the semantic analysis stage. Link parser and word sense disambiguation are supporting techniques at the semantic analysis stage. Link parser processes defines the syntactic elements (noun, verb, proposition, ... etc.) of the sentence and how they are syntactically linked. Word sense disambiguation is provided using WordNet to define all possible senses for each word in the sentence. These senses are used to update the emotions knowledge. Semantic frame is evolved from such information in a way that expresses both syntactic and semantic details of the sentence. In other words, the frame contains all emotional information referenced within the context.

In addition to semantic parsing and the emotions knowledge representation, the proposed model utilizes machine learning process that lead to better emotions detection over the processed texts. The proposed machine learning is achieved through two main components; knowledge base artificial neural network (KBANN) and evolutionary algorithm to train the KBANN.

The proposed emotion detection problem is viewed as multi classification problem. Let “D”, “P” and “S” represent text document, paragraph and sentence respectively, where:

- $D = \langle P_1, P_2, \dots, P_n \rangle$, n is the number of paragraphs
- $P_i = \langle S_1, S_2, \dots, S_m \rangle$, i ranges from 1 to m and m is the number of sentences in P_i
- $S_j = \langle E_1, E_2, \dots, E_k \rangle$, j ranges from 1 to k and k is the number of frame entities in S_j

The objective of the proposed model is to get out the emotion vector for the object (D, P, S) within the text. This vector represented as:

$$EV = \{(e_1, score_1), (e_2, score_2), \dots, (e_q, score_q)\}$$

where q is the number of emotion classes, $e_q \in E$ emotion set and $score \in \{0 \rightarrow 1\}$.

2.1 Learning Process

On top of the proposed neural network architecture, the learning process is essential to assure accurate emotion detection over the subsequent experiments. The proposed evolutionary model relies on extensive KBANN learning of empirical knowledge. Using evolutionary algorithm improves detection results against those resulted from the back propagation algorithm.

One ANN is provided for each emotion class being considered in the detection performance. For example, if the model is designed to evaluate Happy and Anger emotions relatedness of an input text, two ANNs will be required. Figure 1 shows a schematic consequence of training the KBANN.

The proposed neural network uses logistic function as activation function. The EA starts with an initial population size of 100 chromosomes for training the KBANN. A set of experiments has been conducted to evaluate variation of chromosome length and its effect to the training time. A length is 1350 unit has been concluded. Figure 2 shows the evolutionary process for training the neural network.

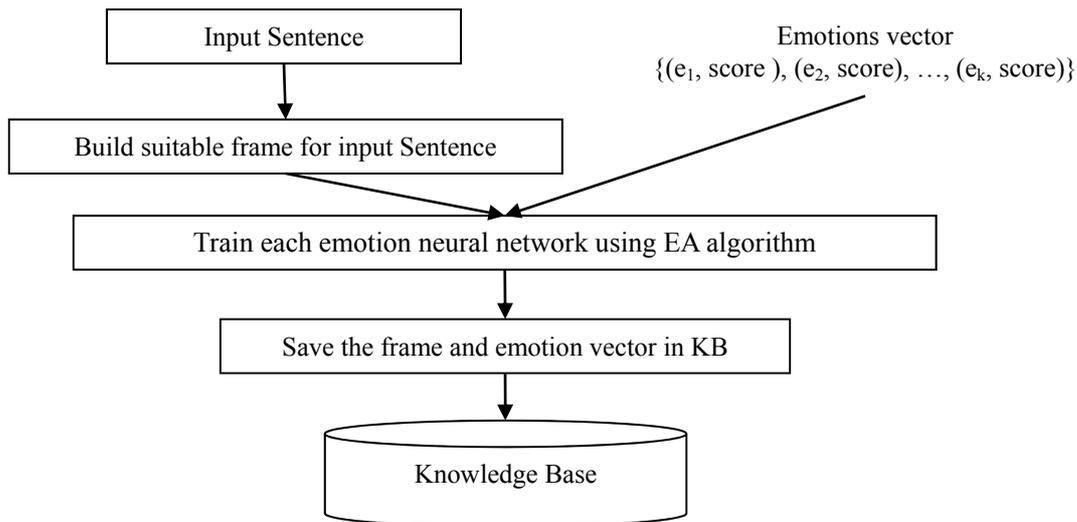


Figure 1. KBANN learning process

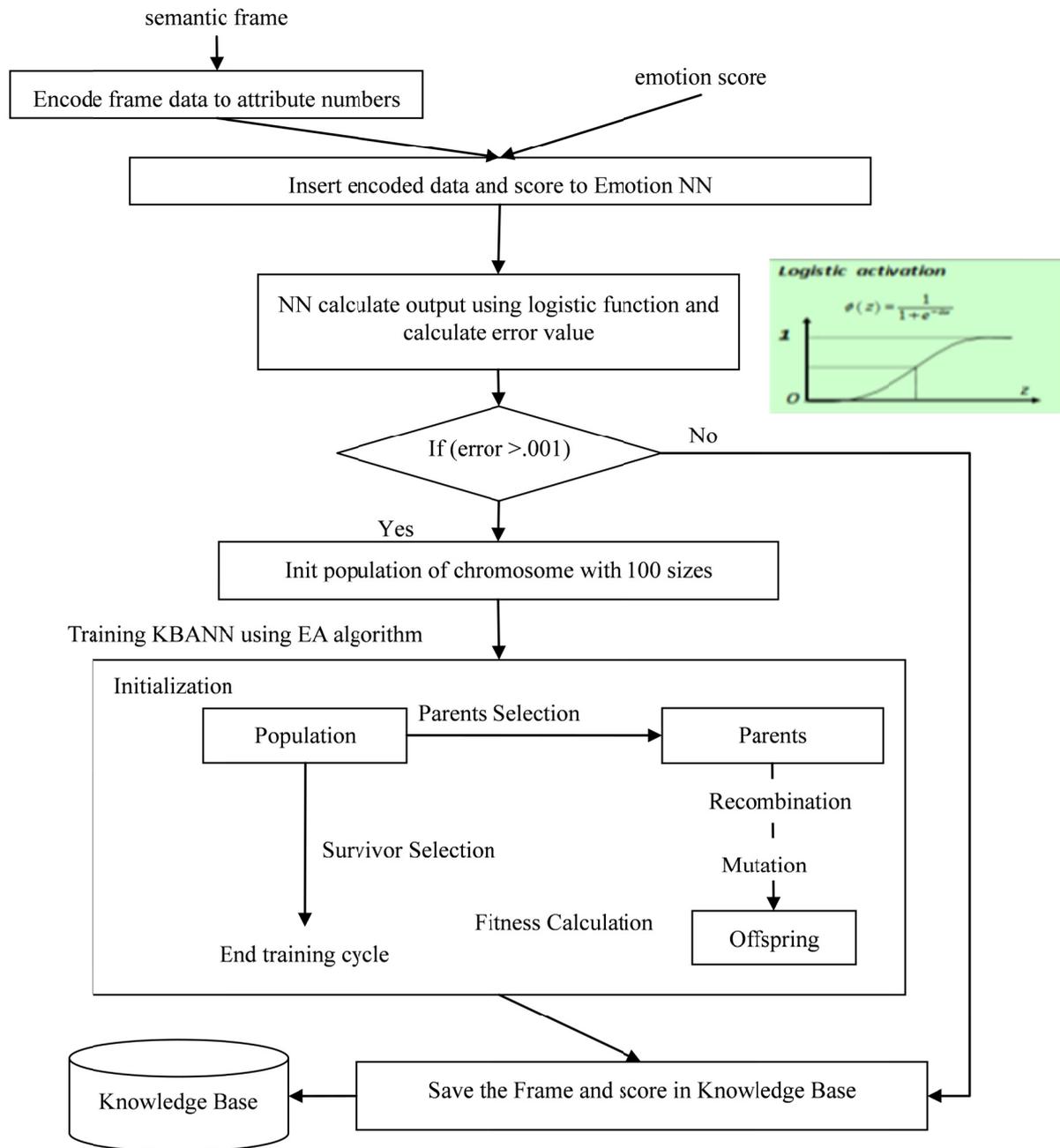


Figure 2. Evolutionary process

One ANN is provided for each emotion class being considered in the detection performance. For example, if the model is designed to evaluate Happy and Anger emotions relatedness of an input text, two ANNs will be required. Figure 1 shows a schematic consequence of training the KBANN.

The proposed neural network uses logistic function as activation function. The EA starts with an initial population size of 100 chromosomes for training the KBANN. A set of experiments has been conducted to evaluate variation of chromosome length and its effect to the training time. A length is 1350 unit has been concluded. Figure 2 shows the evolutionary process for training the neural network.

2.2 Rules Base Generation

KBANN uses approximate domain knowledge as a useful property for the emotion recognition. It uses two training sources, rules as the theory and data as the examples. The proposed KBANN is trained by the two common algorithms; Rules-to-Network” and “Neural Learning”. The first is used to feed approximately-correct,

symbolic rules into the network as the initial knowledge. The second algorithm refines the network using the back propagation learning algorithm, as will be described within the proposed approach. Six emotions classes have been identified while validating the proposed model. Each class has its own semantic rules that are merged with other classes' rules to compose the knowledge base of the neural network.

In this section, the rules entries of the actions that fire each emotion are identified. Some of these actions are subjective and can vary from one person to another. Accordingly, the rules are generic and open to accommodate for multiple entry to same action. Table 1 lists some actions that lead to the happy emotion state. Further actions could be identified by experts and added to the emotion identifier actions.

Table 1. Sample actions igniting happy emotions

Action	Dependent component
<i>Reaching to his goal</i>	<i>Something that can be a goal</i>
<i>Avoiding something harmful</i>	<i>Something harmful</i>
<i>Good description</i>	<i>Descriptor verb & good adjectives</i>
<i>Satisfaction</i>	<i>Satisfied action</i>
<i>Finishing task</i>	<i>Something can be task or job</i>
.....	

An emotion is addressed in the knowledge base through a set of rules. Each rule is identified by pre-invoking event, action and the resulting emotion state.

Example: "I am happy that I finished that annoying task".

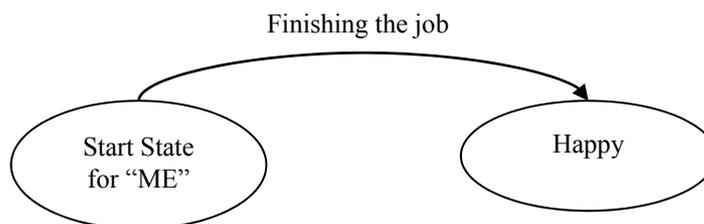


Figure 3. Emotional states transition

The state diagram indicated in Figure 3 shows the following:

- Pre-invoking event: finishing the *annoying task*
- Action: reaching that goal
- Resulting state: happy

Pre-invoking events are the events that are fired prior to the action. Such events are linguistically represented in either simple or composite events. Composite event is a set of simple or composite events that are correlated to each other in a sequential or parallel order. Simple event is identified as a relation between an action, state and set of attributes describing the event parameters. The following rules summarized different relations governing event structure.

$preinvoking\ event \leftarrow Composite\ event \cup simple\ event$
 $simple\ event \leftarrow action \cap state \cap attributes$
 $action \leftarrow sense \cap type \cap score$
 $state \leftarrow description \cap state\ type$
 $composite\ event \leftarrow sequence\ of\ events \cup parallel\ events$
 $sequence\ of\ events \leftarrow event_1 \cap event_2 \cap event_3 \cap event_4 \cap \dots$
 $parallel\ events \leftarrow event_1 \cup event_2 \cup event_3 \cup event_4 \cup \dots$

The following example shows a happy emotion sample rules structure.

$happy \leftarrow achieving\ goal \cup Satisfaction \cup avoiding\ harmful\ action$
 $achieving\ goal \leftarrow finishing\ a\ job \cup success\ in\ the\ exam$
 $finishing\ a\ job \leftarrow finishing\ a\ job \cup success\ in\ the\ exam$
 $finishing\ a\ job\ simple\ event$
 $\leftarrow (actor\ (Me) \cup actor\ (some\ one)) \cap (Action\ (Finish) \cup Action\ (Do))$
 $\cap (State\ (Finished\ job) \cup Action\ (do\ homework))$

The following sentence shows an example of composite event:

“While I was walking in the woods, I saw a lion attacking a Gazelle“

The main event of this sentence is composed of two consequence events. The actor of the sentence passed through one state to reach to the final emotion state as indicated in Figure 4.

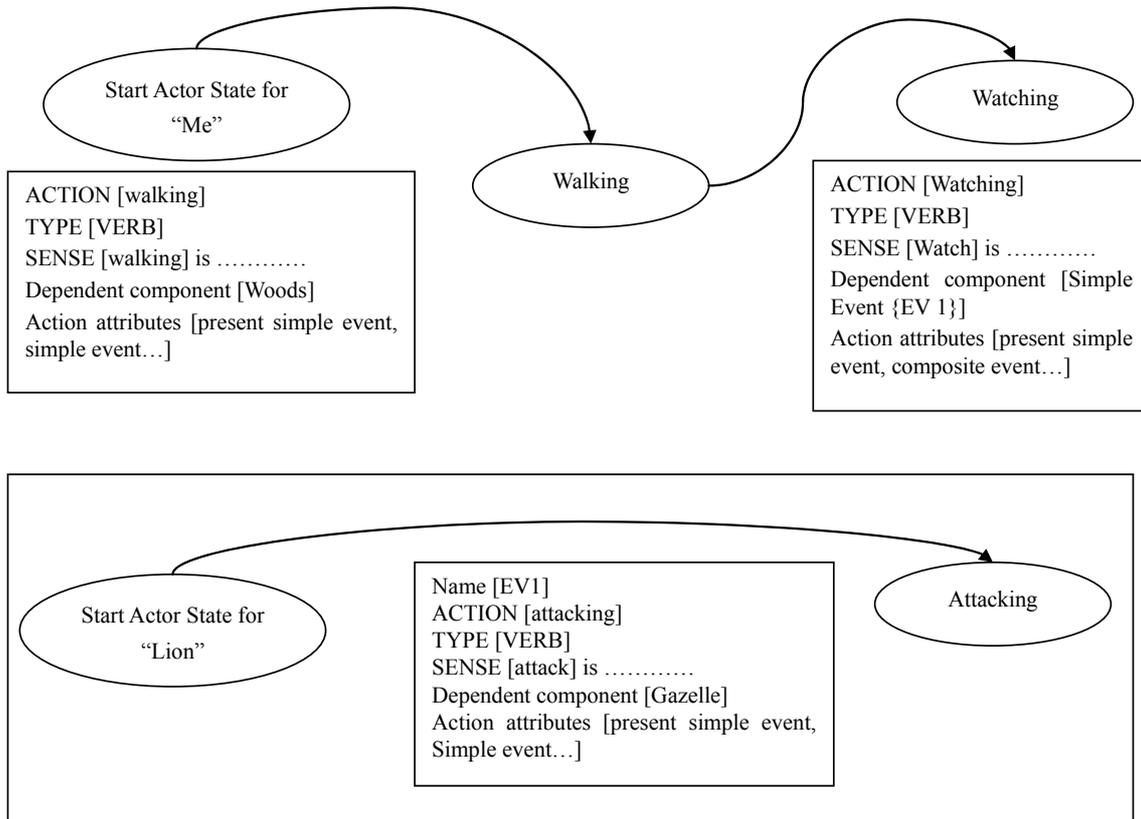


Figure 4. Event, action, state relationship

2.3 Frame Structure

A sentence being processed by the proposed algorithm is mapped into semantic frame. The frame elements are structured as in Figure 5.

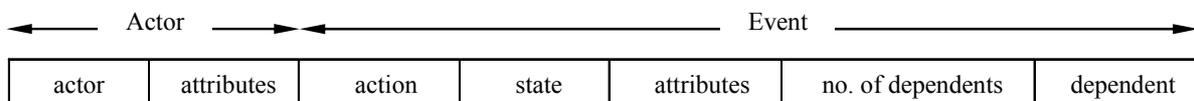


Figure 5. Proposed frame structure

The actor is the event causer. It might be a person, a something or another frame in case of complex sentence. The actor is associated with its attributes. Actor attributes are like type, sense and score.

The event is the coherent behavior of the frame. Generated rules base is used to identify the frame event. An event may be simple or composite. The simple event has three elements; action, state and attributes that describe the action characteristics. Event action can be verb or frame, in case of complex sentence. Event attributes include tense and type. Action tense indicates the time of event occurrence. Event type is either simple, sequence of events or composite of sequence and parallel events. Table 2 lists details of the event elements.

Table 2. Elements of the frame events

<i>Action</i>	<i>State</i>	<i>Attributes</i>
<i>Action name</i>	<i>State name</i>	<i>Action tense</i>
<i>Type</i>	<i>State type</i>	<i>Event type</i>
<i>Sense</i>		
<i>Dependent</i>		

The dependent includes the objects that complete sentence meaning. It may be an object like “woods”, an event like “*a lion attacking a Gazelle*”, time or place.

The following sample is the frame structure for this sentence “my cat died yesterday”

Frame item	Data
Actor	101, 2, 1 {sense, type, score}
Event action	1001, 2, 1 {sense, type, has dependent component}
Dependent component	150, 3, 1 {sense, type, score}
Action tense	Past simple 1
Event type	Simple event

Semantic information of the frame elements including sense and score are extracted by referencing the knowledge base and word sense disambiguation process while building the frame data. The actor is “my cat”, with type “my thing” decoded as “2”. The event is verb event decoded as “2”. This event depends on time object which is “yesterday”. After decoding the entire frame, it will be stored in the knowledge base for future processing.

An emotional item is constructed and added to the knowledge base by processing the sentence in three main steps. The first step is to extract the actor and event frame elements including their inner elements like actor name, action, state, dependent and other existing attributes. The second step is to get the relevant sense that corresponds to each of these elements. Word sense disambiguation task is fitting here. The last step is to evaluate the overall emotion class for the entire sentence based on the correlation between individual elements’ senses and the concerned emotions. The sentence along with the emotion class is finally stored in the knowledge base as an emotional item. The items in the knowledge base are then used to extract the rules needed for training the neural network and consequently build the KBANN.

2.4 Emotion Detection Algorithm

The algorithm starts with parsing individual sentences of the input text. After identifying each sentence, semantic frame is constructed. This frame is used as a query to the knowledge base. If the frame matches any of the knowledge base items, the corresponding emotion vector is resulted.

In case the frame doesn’t exist in the knowledge base, a replacement process of frame elements, one by one, with all of their synonyms, one at a time. The possible number of frame synonyms is calculated as follows:

$$S = \prod_{i=1}^n (\text{synCount}_i + 1) - 1 \quad (1)$$

Where

S is the number of frame synonyms;

n is the number of frame entities;

synCount_i is the number of synonyms for the *ith* entity.

The set of synonym frames are used as new queries to the knowledge base. In case of any matches, the matched synonym frames will be handled as follows:

- 1) Individual synonym frames has the following synonymy relatedness (synonymy of the input frame to its constructed synonym):

$$\text{Frame Synonym Relatedness} = \sum_{i=1}^n \text{Syn}_j(E_i) * w_i \quad (2)$$

Where

n is the number of frame entities;

$\text{Syn}_j(E_i)$ is the synonymy score of the i^{th} entity with respect to its j^{th} synonym (j ranges from 1 to m_i);

w_i is the weight of the i^{th} entity within the frame.

Example: Suppose there is a frame composed of two entities {E1, E2} with equal weights.

- Word senses reveal two synonyms for each entity; synonyms for E1 are {E11, E12} and synonyms for E2 are {E21, E22} .
- Synonymy score for the different synonyms are $(E1, E11) \rightarrow 0.8$, $(E1, E12) \rightarrow 0.7$, $(E2, E21) \rightarrow 0.83$, $(E2, E22) \rightarrow 0.65$
- Frame synonym relatedness is calculated for all possible alternatives as follows:

$$\{E_{11}, E_2\} = (.8 * 1/2) + (1/2) = .9$$

$$\{E_{12}, E_2\} = (.7 * 1/2) + (1/2) = .85$$

$$\{E_{11}, E_{21}\} = (.8 * 1/2) + (.83 * 1/2) = .815$$

$$\{E_{11}, E_{22}\} = (.8 * 1/2) + (.65 * 1/2) = .725$$

$$\{E_{12}, E_{21}\} = (.7 * 1/2) + (.83 * 1/2) = .765$$

$$\{E_{12}, E_{22}\} = (.7 * 1/2) + (.65 * 1/2) = .675$$

$$\{E_1, E_{21}\} = (1/2) + (.83 * 1/2) = .915$$

$$\{E_1, E_{22}\} = (1/2) + (.65 * 1/2) = .825$$

- 2) The emotion of each match is weighted according to the synonym relatedness of the query frame.

Accordingly, the emotion score of the frame synonym is calculated as the following:

$$em_{syn} = (\sum_{i=1}^n \text{Syn}_j(E_i) * w_i) * em_{KB} \quad (3)$$

Where

em_{syn} is the emotion score of the frame synonym;

em_{KB} is the emotion score of the matched frame, as restored from the knowledge base.

- 3) This process will continue for all entities' synonym senses to get the emotions of matched frames.

If frame synonyms failed to match any of the knowledge base items, another round is processed with contrary senses of all frame entities, one by one. The possible number of frame opposites is calculated as follows:

$$O = \prod_{i=1}^n (\text{oppCount}_i + 1) - 1 \quad (4)$$

Where

O is the number of frame opposites;

n is the number of frame entities;

oppCount_i is the number of opposites for the i^{th} entity.

If a query of the opposite frames matches, the contrary or complement of the corresponding emotion score extracted from the knowledge base is then considered. The matched opposite frames will be handled as follows:

- 1) Individual opposite frames has the following contrariety relatedness (contrariety of the input frame to its constructed opposite):

$$\text{Frame Opposite Relatedness} = \sum_{i=1}^n \text{Opp}_k(E_i) * w_i \quad (5)$$

Where

n is the number of frame entities;

$Opp_k(E_i)$ is the contrariety score of the i^{th} entity with respect to its k^{th} opposite (j ranges from 1 to m_j);
 w_i is the weight of the i^{th} entity within the frame.

- 2) The emotion of each match is weighted according to the opposite relatedness of the query frame.

Accordingly, the emotion score of the frame opposite is calculated as the following:

$$em_{opp} = (\sum_{i=1}^n Opp_k(E_i) * w_i) * (1 - em_{KB}) \quad (6)$$

Where

em_{opp} is the emotion score of the frame opposite;

em_{KB} is the emotion score of the matched frame, as restored from the knowledge base;

- 3) This process will continue for all entities' opposite senses to get the emotions of matched frames.

If frame opposites failed again to match any of the knowledge base items, a third round is processed with a mix of synonymy and contrary senses of all frame entities, one by one. The possible number of frame mixed synonyms and opposites is calculated as follows:

$$M = \prod_{i=1}^n (synCount_i + oppCount_i + 1) - \prod_{i=1}^n (synCount_i + 1) - \prod_{i=1}^n (oppCount_i + 1) - 1 \quad (7)$$

Where

M is the number of frame mixed senses;

n is the number of frame entities;

$synCount_i$ is the number of synonyms for the i^{th} entity;

$oppCount_i$ is the number of opposites for the i^{th} entity.

If a query of the mixed frames matches, the matched frames will be handled as follows:

- 1) Individual opposite frames has the following mixed synonymy and contrariety relatedness (contrariety of the input frame to its constructed opposite):

$$Mixed\ Sense\ Frame\ Relatedness = \sum_{i=1}^l Syn_j(E_i) * w_i + \sum_{i=l+1}^n Opp_k(E_i) * w_i \quad (8)$$

Where

n is the number of frame entities;

l is the number of frame entities substituted by their synonyms;

$n-l$ is the number of frame entities substituted by their opposites;

$Syn_j(E_i)$ is the synonymy score of the i^{th} entity with respect to its j^{th} synonym (j ranges from 1 to $synCount_i$);

$Opp_k(E_i)$ is the contrariety score of the i^{th} entity with respect to its k^{th} opposite (k ranges from 1 to $oppCount_i$);

w_i is the weight of the i^{th} entity within the frame.

- 2) The emotion of each match is weighted according to the mixed relatedness of the query frame.

Accordingly, the emotion score of the frame of mixed synonym and opposite senses is calculated as the following:

$$em_{mix} = (\sum_{i=1}^l Syn_j(E_i) * w_i) * em_{KB} + (\sum_{i=l+1}^n Opp_k(E_i) * w_i) * (1 - em_{KB}) \quad (9)$$

Where

em_{mix} is the emotion score of the frame opposite;

em_{KB} is the emotion score of the matched frame, as restored from the knowledge base.

- 3) This process will continue for all entities' synonyms and opposite to get the emotions of matched frames.

Finally, if this round fails to match with the knowledge base again, an evolutionary process is triggered so that the neural network can resolve and get out the fitting emotion. The algorithm processing steps are illustrated in Figure 6.

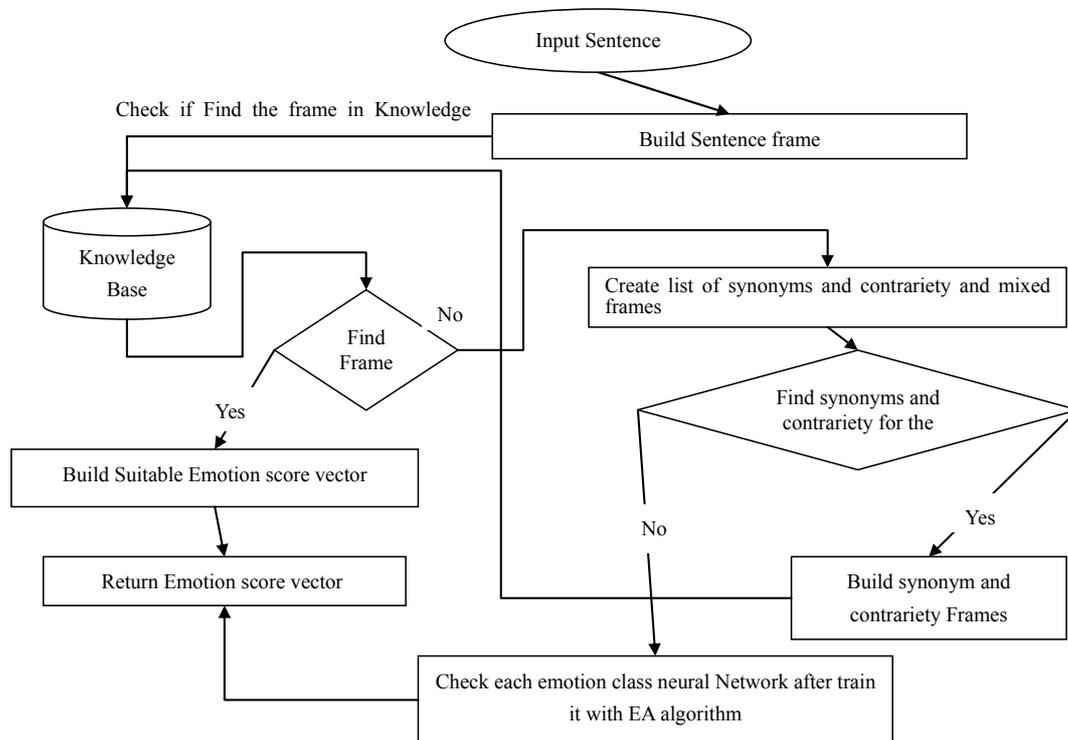


Figure 6. Emotion detection proposed model

2.5 Frame Matching Process

Matching a frame against the knowledge based frame set is provided through four different approaches. Each approach varies in the search time and matching quality. Such flexible approaches of matching enables control of the performance versus relatedness strength of the matched synonym or opposite.

- First Matching, where the algorithm stops at the first synonym frame matched in the knowledge base. In this approach, the search time is the optimal and is useful when the performance is being concerned.
- Best Matching, where the highest synonymy related frame is selected to result the maximum emotion score. Although the processing is not the optimal, the match quality is better and accordingly the resulted emotion is the more accurate.
- Best Opposite Matching, where the highest opposite relatedness is selected to result the maximum emotion score. This approach takes longer time than first matching however match quality is better as well.
- Average Matching, where the emotions score is evaluated as the average of all derived synonym and opposite frames.

3. Case Studies

In this section, case studies are presented to illustrate the procedures of two main functionalities in the proposed model; the evolutionary process for training an ANN and the emotion detection process.

3.1 ANN Evolutionary Training

The first illustration is a case study that shows the ANN evolutionary process used to train the model. In the happy emotion example considered in Figure 7, the ANN is assumed to have the following characteristics:

- One neuron.
- Three inputs to this neuron representing the frame with 3 entities.
- Weight of inputs is one of the set $\{1, 0, -1\}$, where (1) represents full positive effect of the input to the output, 0 represents no effect of the input to the output and (-1) represents full negative effect of the input to the output.
- The activation function $f(x)$ is threshold activation function defined as:

$$f(x) = \begin{cases} 0, & Y < 0 \\ 1, & Y \geq 0 \end{cases}$$

Where y is calculated as

$$Y = \sum_i^n W_i * E_i$$

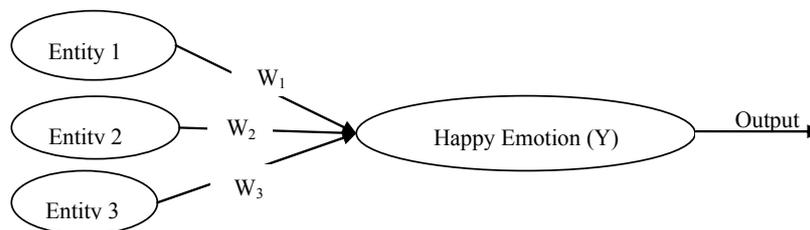


Figure 7. Happy emotion activation function

Training data set for the input entities and the corresponding desired output are identified in Table 3.

Table 3. Input weights-desired output data set

Entity ₁	Entity ₂	Entity ₃	Desired Output
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

The fitness function for the case example is defined as the count of matching results of actual versus desired results.

$$fitness(x) = \sum_{i=1}^8 (Actual_i == Desired_i)$$

Actual	Desired	Matching
1	1	Yes
1	0	No
1	1	Yes
1	0	No
0	1	No
1	1	Yes
0	1	No
1	1	Yes

Applying genetic algorithm, the following steps will proceed.

Step 0: Population initialization. Four random weights sets used as an initial population. The corresponding fitness value for each chromosome is indicated as follows.

Chromosome Number	W_1	W_2	W_3	$fitness(x)$
1	-1	0	1	4
2	1	0	0	2 (Low will discarded)
3	0	1	-1	6
4	1	-1	1	5

Step 1: Solutions are randomly generated according to the fitness probability distribution. Accordingly, the second solution in the above step is the lowest, where fitness is 2. It will be dropped and the third solution with fitness = 6 is generated twice.

Chromosome Number	W_1	W_2	W_3	$fitness(x)$
1	-1	0	1	4
2	0	1	-1	6
3	0	1	-1	6
4	1	-1	1	5

Step 2: Iteration 1; the first two and last two solutions are randomly mated. Then their crossing sites are randomly chosen. We assume that there is no mutation.

Chromosome Number	W_1	W_2	W_3
1	-1	0	1
2	0	1	-1
3	0	1	-1
4	1	-1	1

Iteration 2, new population

After the crossover breeding mentioned in the previous iteration, the new population along with the fitness values are:

Chromosome Number	W_1	W_2	W_3	$fitness(x)$
1	-1	0	-1	4
2	0	1	1	6
3	0	-1	1	4
4	1	1	-1	7

Based on the fitness probability distribution, the solutions are randomly generated. In this specific run, above solutions are regenerated in a different order.

Chromosome Number	W_1	W_2	W_3
1	0	1	1
2	0	-1	1
3	-1	0	-1
4	1	-1	-1

Randomly mating the first two and last two solutions, their crossing sites are randomly chosen as highlighted. It has been assumed that there is no mutation.

Iteration 3, new population

Upon crossover breeding, the new population along with fitness values will be:

Chromosome Number	W_1	W_2	W_3	<i>fitness(x)</i>
1	0	1	1	6
2	0	-1	1	4
3	0	-1	1	4
4	1	0	-1	8

The last solution ($w_1 = 1, w_2 = 0, w_3 = -1$) matches all desired outputs corresponding to actual inputs as listed in Table 4.

Table 4. Desired output vs actual output

Entity ₁	Entity ₂	Entity ₃	Desired output	Actual output	<i>Matching</i>
0	0	0	1	1	Yes
0	0	1	0	0	Yes
0	1	0	1	1	Yes
0	1	1	0	0	Yes
1	0	0	1	1	Yes
1	0	1	1	1	Yes
1	1	0	1	1	Yes
1	1	1	1	1	Yes

3.2 Detecting Text Emotion Using Proposed Algorithm

The second case study shows the emotion detection algorithm processing for a text segment extracted from <http://www.studymode.com/essays/September-11-Essay-911420.html>. The first paragraph will be processed.

“September 11, 2001. A day no American will ever forget. It was a day that this nation realized that we as Americans are not invincible. We can be hurt. Planes flew into both the north and south towers of the world trade center. The towers collapsed killing thousands of people. This event has been not only terrifying, but it has also been insightful. In my English IV class, we read multiple articles that were published after 9/11. These articles gave us insight by looking into the historical significance as well as the memories of that terrible day. “

The processing output will be as follows:

3.2.1 Semantic Parsing

Input text is parsed using one of the semantic parser frameworks like OpenNLP or link parser. The text is semantically parsed resulting in the parsed text listed in Table 5.

Table 5. Sentences semantic parsing

#	Sentence	Semantic parser output
S1	<i>September 11, 2001. A day no American will ever forget</i>	(TOP (S (NP (DT A) (NN day)) (NP (DT no) (NNP American)) (VP (MD will) (ADVP (RB ever)) (VP (VB forget))) (. .)))
S2	<i>It was a day that this nation realized that we as Americans are not invincible.</i>	(TOP (S (NP (PRP It)) (VP (VBD was) (NP (DT a) (NN day)) (SBAR (IN that) (S (NP (DT this) (NN nation)) (VP (VBD realized) (SBAR (IN that) (S (NP (NP (PRP we)) (PP (IN as) (NP (NNPS Americans)))))) (VP (VBP are) (RB not) (ADJP (JJ invincible)))))))) (. .)))
S3	<i>We can be hurt</i>	(TOP (S (NP (PRP We)) (VP (MD can) (VP (VB be) (VP (VBN hurt)))) (. .)))
S4	<i>Planes flew into both the north and south towers of the world trade center.</i>	(TOP (S (NP (NNS Planes)) (VP (VBD flew) (PP (IN into) (NP (DT both) (NP (DT the) (NN north)) (CC and) (NP (NP (JJ south) (NNS towers)) (PP (IN of) (NP (DT the) (NN world) (NN trade) (NN center)))))) (. .)))
S5	<i>The towers collapsed killing thousands of people</i>	(TOP (S (NP (DT The) (NNS towers)) (VP (VBD collapsed) (S (VP (VBG killing) (NP (NP (NNS thousands)) (PP (IN of) (NP (NNS people)))))) (. .)))
S6	<i>This event has been not only terrifying, but it has also been insightful.</i>	(TOP (S (S (NP (DT This) (NN event)) (VP (VBZ has) (VP (VBN been) (CONJP (RB not) (RB only)) (ADJP (JJ terrifying)))) (. ,) (CC but) (S (NP (PRP it)) (VP (VBZ has) (ADVP (RB also)) (VP (VBN been) (ADJP (JJ insightful)))) (. .)))
S7	<i>. In my English IV class, we read multiple articles that were published after 9/11.</i>	(TOP (S (PP (IN In) (NP (PRP\$ my) (NNP English) (NNP IV) (NN class))) (. ,) (NP (PRP we)) (VP (VBD read) (NP (NP (JJ multiple) (NNS articles)) (SBAR (WHNP (WDT that)) (S (VP (VBD were) (VP (VBN published) (PP (IN after) (NP (CD 9/1) (CD 1)))))) (. .)))
S8	<i>These articles gave us insight by looking into the historical significance as well as the memories of that terrible day.</i>	(TOP (S (NP (DT These) (NNS articles)) (VP (VBD gave) (NP (PRP us)) (NP (NN insight)) (PP (IN by) (S (VP (VBG looking) (PP (IN into) (NP (DT the) (JJ historical) (NN significance))) (ADVP (ADVP (RB as) (RB well)) (PP (IN as) (NP (NP (DT the) (NNS memories)) (PP (IN of) (NP (DT that) (JJ terrible) (NN day)))))) (. .)))

Link parser results show eight sentences, leading to eight corresponding frames for the paragraph.

3.2.2 Frame Structure Representation

A semantic frame is to be generated for each sentence. At this stage, senses corresponding to each word or phrase will be substituted. Table 6 listed sample senses that have been referenced throughout the frames generation.

Table 6. Sample referenced senses

Sense Description	REFERENCE ID
subjective Pronoun	SEN#1
twenty-four hours, twenty-four hour period, 24-hour interval, solar day, mean solar day	SEN#2
nonfictional prose forming an independent part of a publication	SEN#4
clear or deep perception of a situation	SEN#5
a native or inhabitant of the United States	SEN#6
dismiss from the mind; stop remembering	SEN#111
any physical damage to the body caused by violence or accident or fracture	SEN#112
being verb, copula, used with an adjective or a predicate noun	SEN#110
cause to have, in the abstract sense or physical sense	SEN#113
enjoying or showing or marked by joy or pleasure or good fortune	SEN#20
an emotion experienced in anticipation of some specific pain or danger	SEN#21
a feeling of understanding	#SEN53
the income or profit arising from such transactions as the sale of land or other property	#SEN116
try to forget; put out of one's memory or knowledge	#SEN115
feeling sorrow or unhappiness	#SEN22
Article	#SEN50
a set of questions or exercises evaluating skill or knowledge	#SEN51
find the solution	#SEN114
occupation, business, job, line of work	#SEN52

The semantically parsed text resulted from the previous step is processed to define objects of each sentence frame. Semantic labels of each sentence are structured in Actor-Event scheme according to the proposed frame representation. Frame items of sentences S1, S2, S3 and S8 are as follows:

S_1 : "September 11, 2001. A day no American will ever forget".

1. Actor object: - (American)

Sense :- SEN#6

Type: - person {nothing, person, something, frame}

Score :- 1 {1 except the type is frame}

2. Event object :- (will ever forget)

• Action

• Name :- Forget

• Type : verb {nothing, verb, connector, frame,}

• Sense :- SEN#111

• Dependent components (September 11, 2001. A day)

1. Type :- object {nothing, object, frame}

2. Sense SEN#2

3. Score 1

• State

• Name : not forgotten

• Type normal {nothing, normal, opposite}

• Event attributes

• Event tense:- future

• Direct or indirect : indirect

• Used modification words :- yes

• Event type:- simple

S₂: "It was a day that this nation realized that we as Americans are not invincible".

3. Actor object: - (it)

Sense: - SEN#1

Type: - something {nothing, person, something, frame}

Score: - 1 {1 except the type is frame }

4. Event object :- (was a day that this nation realized that we as Americans are not invincible.)

• Action

• Name :- is

• Type : verb {nothing, verb, connector, frame,}

• Sense :- SEN#110

• Dependent components

1. Object 1 (A day)

○ Type :- object { nothing , object , frame}

○ Sense SEN#2

○ Score 1

2. Object 2 (this nation realized that we as Americans are not invincible.)

○ Type :- frame {nothing , object , frame}

1. Actor : this nation

2. Action (realized)

3. Dependent components

• Object 1 (we as Americans are not invincible)

○ Type :- frame {nothing, object, frame}

○ Sense SEN#20

○ Score .85

○ Sense SEN#20

○ Score .85

• State

• Name :be realize

• Type normal {nothing , normal , opposite}

• Event attributes

• Event tense:- past

• Direct or indirect : Direct

• Used modification words :- yes

• Event type:- complex

S₃: "We can be hurt".

5. Actor object: - (we)

Sense: - SEN#1

Type: - person {nothing, person, something, frame}

Score: - 1 {1 except the type is frame}

6. Event object :- (can be hurt)

• Action

• Name :- hurt

• Type : verb {nothing, verb, connector, frame,}

• Sense :- SEN#112

• State

• Name : hurting

• Type normal {nothing, normal, opposite}

• Event attributes

• Event tense:- present

• Direct or indirect : indirect

- *Used modification words :- No*
- *Event type:- simple*

We complete parsing each sentence depend on the semantic parsing data from link parser the last senescence will be

S8: "These articles gave us insight by looking into the historical significance as well as the memories of that terrible day".

7. *Actor object: - (these articles)*

Sense: - SEN#4

Type: - something {nothing, person, something, frame}

Score: - 1 {1 except the type is frame}

8. *Event object :- (gave us)*

- *Action*

- *Name :- gave*

- *Type : verb {nothing, verb, connector, frame,}*

- *Sense :- SEN#113*

- *Dependent components*

1. *Object 1 (us)*

- *Type :- object {nothing, object, frame}*

- *Sense SEN#1*

- *Score 1*

2. *Object 1 (insight)*

- *Type :- object {nothing, object, frame}*

- *Sense SEN#5*

- *Score 1*

3. *Object 2 (by looking into the historical significance as well as the memories of that terrible day.)*

- *Type :- frame {nothing, object, frame, complement}*

- *Sense SEN#20*

- *Score. .7*

- *State*

- *Name :be giving*

- *Type normal {nothing, normal, opposite}*

- *Event attributes*

- *Event tense:- past*

- *Direct or indirect : Direct*

- *Used modification words :- No*

- *Event type:- complex*

3.2.3 Frame Generation

Frames are constructed from the semantic data excluded above. Senses of frames entities corresponding to S1, S2, S3 and S8 are shown in Table 7.

Table 7. Referenced senses for frame entities

#	Sentence	Frame items
S1	September 11, 2001. A day no American will ever forget	SEN#6,Person,1,SEN#111,Verb,1,1,SEN#2,Object,1,Future,INDIRECT,YES,SIMPLE
S2	It was a day that this nation realized that we as Americans are not invincible.	SEN#1,Something,1,SEN#110,Verb,1,2,SEN#2,Object,1,SEN#20,frame,.85,Past, Direct, Yes, COMPLEX
S3	We can be hurt	SEN#1,Person,1,SEN#112,Verb,1,0,Present,INDIRECT,NO,SIMPLE
S8	These articles gave us insight by looking into the historical significance as well as the memories of that terrible day.	SEN#4,something,1,SEN#113,Verb,1,3,SEN#1,Object,1,SEN#5,Object,1,SEN#21,frame,.6,Past,DIRECT,NO,COMPLEX

3.2.4 Emotion Recognition

Emotion detection processed according to the proposed algorithm detailed above. The model utilizes all trained frames to extract proper emotion. To illustrate the detection performance, sample of the frames knowledge needed for the emotion detection of sentences S1, S2, S3 and S8 are listed in Table 8.

Table 8. Sample knowledge base frames

ID	Frame	Emotion class	score
1(synonym)	SEN#6,Person,1,SEN#115,Verb,1,1,SEN#2,Object,1, FUTURE, INDIRECT, YES, SIMPLEX	sorrow	.63
2	SEN#1,Person,1,SEN#112,Verb,1,0,Present,INDIRECT,NO,SIMPLE	Fear	.71
3(opposite)	SEN#1,Something,1,SEN#110,Verb,1,2,SEN#2,Object,1,1,SEN#22,frame,.9, Past, DIRECT, YES, COMPLEX	sad	.9
4	SEN#4,Something,1,SEN#116,Verb,1,3,SEN#1,object,1,SEN#53,object,1, #SEN21,frame,.6, PAST,DIRECT,NO,COMPLEX	Exclamation	.74
5	SEN#51, Something,1,SEN#116,VERB,1,0, Present, INDIRECT, YES, SIMPLE	Sad	.74
6	SEN#52,Something,1,SEN#116,VERB,1,3,SEN#1,OBJECT,1,SEN#53, Object,1,#SEN21, frame,.6, Past, Direct, NO, Complex	anger	.81
7	SEN#1, Something,1, SEN#110,verb,1,2,SEN#3,Object,1,SEN#20,frame,.9, Past, DIRECT, YES,COMPLEX	happy	.82

The following are the mapping outcomes of sentences frames through the proposed algorithm:

- S1: Initial lookup for the frame in the knowledge base, results no match. Accordingly, synonym frame is constructed by substituting the word “forget” with its synonyms “unlearn”, $f(\text{syn}) = .8$ and weight = .25. The resulted frame is:
 “SEN#6,Person,1,SEN#115,Verb,1,1,SEN#2,Object,1, FUTURE, INDIRECT, YES, SIMPLEX”
 Lookup for this frame in knowledge base, it exists providing the emotion calculation as $(.81*.25*.63) + (.75*1*.63) = .5995$; the resulted emotion evaluation is (Sorrow, .5995).
- S2: Initial lookup for the frame in the knowledge base, results no match. Accordingly, synonym frame is constructed by substituting the frame “Americans are not invincible” with its opposite “Americans are invincible”. $f(\text{opp}) = 1$ and weight = .2. The resulted frame is:
 “SEN#1,Something,1,SEN#110,Verb,1,2,SEN#2,Object,1,1, SEN#22,frame,.9, Past, DIRECT, YES,

COMPLEX”

Lookup for this frame in knowledge base, it exists providing the emotion calculation as $(1*.2*)*(1-.9)=.02$; the resulted emotion evaluation is (sad, .02).

- S3: - Initial lookup for the frame in the knowledge base, results a match. The resulted emotion evaluation is (Fear, 71).
- S8: Initial lookup for the frame in the knowledge base, results no match. Accordingly, mixed frame is constructed by substituting the word "give" with its opposite, $f(\text{opp})=.88$ and weight = .17 and substituting the word "insight" with its synonyms "penetration", $f(\text{syn})=.75$ and weight = .17. The resulted frame is:

“SEN SEN#4, Something, 1, SEN#116, Verb, 1, 3, SEN#1, object, 1, SEN#53, object, 1, #SEN21, frame, .6, PAST, DIRECT, NO,COMPLEX”

Lookup for this frame in knowledge base, it exists providing the emotion calculation as $(.88*.17)*(1-.74) + (.17*.74*.75) + (4*.74*.17) = .034 + .09435 + .5032 = .63155$; the resulted emotion evaluation is (Exclamation, .63155).

4. Semantic Knowledge Base Representation

The proposed model is based on a knowledge base neural network (KBANN). The knowledge base is used to both train neural network and to detect the best emotion that suits the underlying text. The knowledge base is structured to compose a list of emotional items, each is constructed as a sentence or phrase, emotion class and emotion score.

- The sentence or phrase is the set of co-related words constituting the sentence to provide an understanding of the subject-link-object structure. The sentence is semantically parsed into a set of semantic labels. All labels and other elements like dependent components will be considered in emotion evaluation. Each individual element will be identified by their sense, type and sense score. Sense is identified by its text description, type and code. Table 9 shows sample senses along with their types and the indicating words. Examples of subject and object types are “Me”, “My thing”, “Someone”, “Something” and “Sentence”. Examples of links types are “None”, “Noun”, “Event”, “Modification Word Sentence”, “Connector and Symbol”. Other sentence elements like conjunctions words, punctuations words, positive modification words and negative modification words are also considered as part of the knowledge details.
- The emotion class refers to the one of the emotions that describes the processed context like happy, sad, anger, fear, astonishment, natural..... etc.
- The score is a value ranging from 0 to 1 defining the relatedness of the sentence to the emotion class. A sentence might belong to multiple emotion classes with different scores.

Table 9. Sample senses’ identifications

Sense text	Sense Type	Sample Words	Sense Code
feeling sorrow or unhappiness	Frame		1
feeling or marked by joy or pleasure or good fortune	Frame		2
feeling concern or insecurity	Frame		3
any of various predatory carnivorous canine mammals	Noun	Wolf, Lion, tiger, cheetah	100
being begetter	Noun	father, mother	101
person who is a fellow member	Noun	brother, sister	102
small swift graceful antelope	Noun	gazelle, deer	103
eline mammal usually having thick soft fur	Noun	fox, cat	104
fail to keep	Verb	lose, lost	1000
go away	Verb	leave, die	1001
being verb, copula, used with an adjective or a predicate noun	Verb	is, am, are	SEN#110
subjective Pronoun	Noun	I, you, she, it	SEN#1
objective Pronoun	Noun	me, her, him	106

Sense text	Sense Type	Sample Words	Sense Code
Possessive adjective Pronoun	Noun	my, your, his	107
Possessive Pronoun	Noun	mine	108
have or possess	Verb	have, has	1003
human offspring	Noun	son, child	109
a person whose occupation is teaching	Noun	teacher	110
putting to death	Verb	Kill	1004
leave something undone	Verb	Fail	1005
Article	Noun	a , the, an	111
a set of questions or exercises evaluating skill or knowledge	Noun	test, exam	112
make something	Verb	do, make, prefer	1006
negation of a word or group of words	Noun	not	113
Reflexive Pronoun Pronoun	Noun	themselves, myself	114
isolated from others	Noun	Alone , lonely	115
affected by an impairment of normal physical or mental function	Noun	ill, sick	116
have a great affection or liking for	Verb	love , like	1006
Relative Pronoun	Noun	who, which, whom	117
perceive by sight or have the power to perceive by sight	Verb	see, saw	1007
occupation, business, job, line of work	Noun	job, task	118
modification Word	Noun	but	119
Scream	Verb	cry	1008
Preposition		of, in, at	5000
Timer Adverb	Adverb	yesterday, tomorrow	5001
the official currency issued by a government or national bank	Noun	money	120
fill with distaste	Emotion	sentence	4
feeling or showing anger	Emotion	sentence	5
a state of difficulty that needs to be resolved	Noun	problem	121
find the solution	Verb	solve	1009
undergo an emotional sensation	Verb	feel	1010
attractive especially by means of smallness or prettiness or quaintness	Noun	cute , smart	122
any physical damage to the body caused by violence or accident or fracture	Verb	hurt	SEN#112
dismiss from the mind; stop remembering	Verb	forget	SEN#111
nonfictional prose forming an independent part of a publication	Noun	article	SEN#4
cause to have, in the abstract sense or physical sense;	Verb	Give , gave	SEN#113
clear or deep perception of a situation	NOUN	insight	SEN#5
try to forget; put out of one's memory or knowledge	Verb	unlearn	1014
a feeling of understanding	NOUN	penetration	127
the income or profit arising from such transactions as the sale of land or other property	Verb	take	1015
twenty-four hours, twenty-four hour period, 24-hour interval, solar day, mean solar day	NOUN	day	SEN#2

5. Experimental Results

The proposed model has been evaluated through a data set whose items have been compiled from Oxford practice grammar book and different electronic linguistic magazines. Items were pre-classified into eight emotion categories; happy, sad, anger, disgust, fear, joy, astonishment and natural. A corpus of 1000 frame items for each emotion is used throughout the model evaluation. Out of these items, 600 are labeled as positive emotion and 400 labeled as negative emotion. Each item is either an individual sentence or a short paragraph of about 3-4 sentences. The knowledge base is open to extend for further learning, which will directly impact the detection accuracy.

The objective is to evaluate the behavior of both training and testing processes. In other words, efficient capability of the model to learn will lead to efficient detection of text emotions. Initial training set were labeled, through experts, with emotion state of each item. In the following section, the proposed emotion detection model is compared against other models.

5.1 Accuracy Evaluation

The corpus data set has been divided into two sub sets; one for training and the other for testing. Four experiments have been conducted to evaluate the proposed semantics frame model accuracy. Each of them differs in the ratio of the training set to testing set. Accuracy results for 80%-20%, 70%-30%, 60%-40% and 50%-50% of training-to-testing sets ration was evaluated for six emotions. The overall accuracy of emotion recognition is calculated as the average value for each emotion as indicated in Table 10.

Table 10. Emotion detection accuracy evaluation

Emotion Class	Accuracy
Happy	85%
Sad	85%
Fear	74%
Astonishment	72%
Anger	80%
Natural (No emotion)	88%

As the knowledge base increase by storing extra sets of semantic frames, the indicated accuracy will increase. This will be indicated in the following experiment.

5.2 Accuracy versus Frame Knowledge Size

In this experiment, the training frames are portioned in incremental samples to evaluate the accuracy variation according to the size of the training set. Sad emotion training set are examined in six increments of 50, 100, 200, 250, 400 and 600 frames respectively. Results presented by Figure 8 showed that starting training the sad emotion neural network with the first sample, the accuracy was 25%. Adding the second sample to the training set, the resulted accuracy increased to 40%. The experiment proved increased accuracy as the training set increase. Stepping up over incremental samples leads to 85% accuracy for 600 training samples. Training further samples ensures increased accuracy.

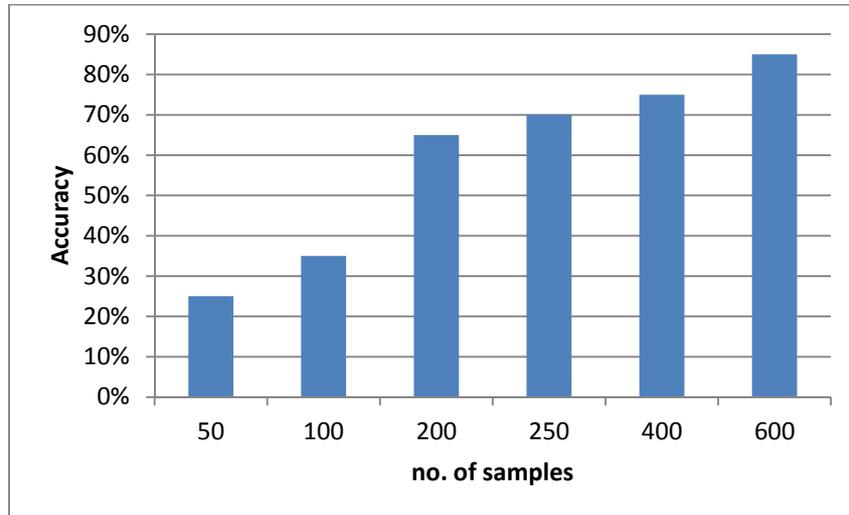


Figure 8. Accuracy vs sample size results

5.3 Learning Time versus Frame Knowledge Size

The evolutionary process varies in its learning time depending on the size of the training set. Figure 9 shows the learning time, represented by number of learning iterations for successive increments of training sets. It is clear from the figure that as the accumulated training sets increase, the learning time increases. At some limit comprising relatively high number of frames in the training set, the learning time will be almost stable with no further increase. The reason for that is that at some level of knowledge experience, all emotion detection inquiries could be accurately revealed. Extra frame knowledge will not add a lot to the detection accuracy.

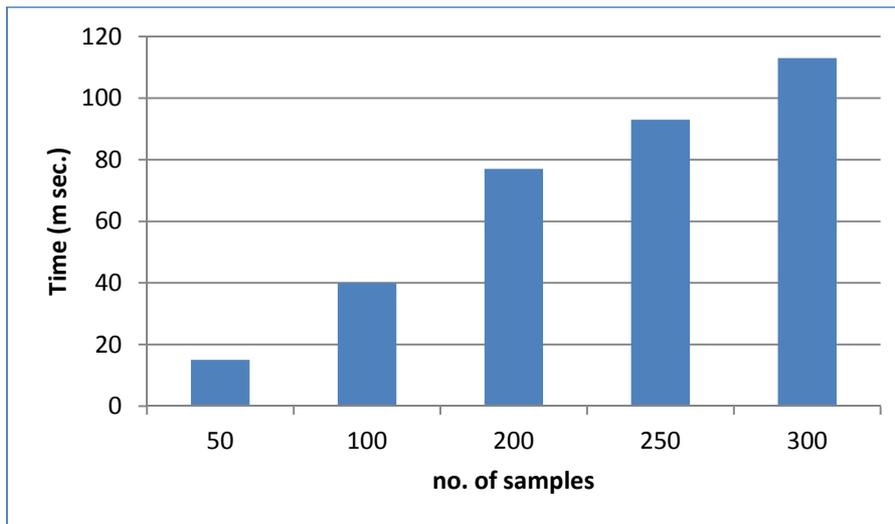


Figure 9. Learning time vs sample size results

5.4 Chromosome Length versus Training Iterations

One of the evolutionary algorithm characteristics that need to be tracked is the chromosome length. The impact of the chromosome length is directly affecting the KBANN training time. Results of the learning iterations size versus chromosome length is indicated in Figure 10.

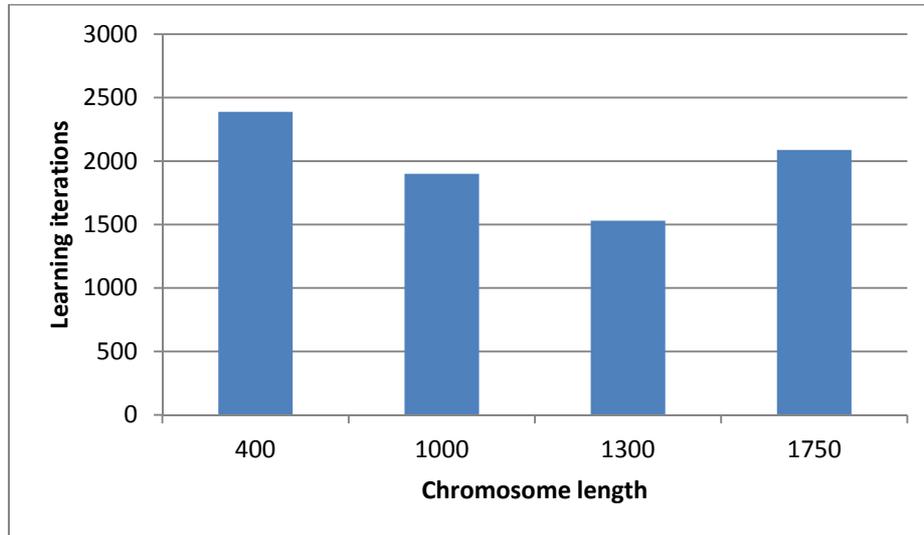


Figure 10. Chromosome length vs training time results

5.5 Proposed Model versus Binary Classifier Models

Binary classifier emotion detection models results are either “emotional” or “neutral”. A comparative experiment is conducted to evaluate the proposed model with three binary classifier models; Linguistic, pragmatic, and keyword spotting features model, emotion estimation net (ESiN) model (Tao, 2004) and supervised machine learning with SNoW learning architecture (Alm, Roth, & Sproat, 2005). The last two models have been examined under the same conditions of the proposed model. In such experiment, the 400 testing frames are divided into four subsets, each of 100 testing frames. The experiment is repeated four times, each with using one of the testing samples subsets. Accuracy is calculated based on the desired output matching as the average value of the four subsets emotion classification results.

Linguistic, pragmatic, and keyword spotting features model (C. Lee & G. Lee, 2007) accuracy result is based on a corpus consists of 2900 items from in 10 domains including “sports”, “love”, “weather”, “music” and others. 2122 items are natural items and 778 items are emotional item. Chat corpus of entertainment like the goal or task-oriented dialogs is used. It’s clear from the experiment that the majority of the testing items are natural, which will be supporting in the binary classification. On the other hand, emotional items set comprise relatively low number of items compared with the natural items. This will be tightly model related to the multi label classifier. Table 11 shows the accuracy evaluation for the proposed and the three mentioned binary classifier models.

Table 11. Binary classifiers accuracy evaluation

Model	Accuracy
Proposed Model_	91.73%
Linguistic, pragmatic, and keyword spotting features model	90.21%
Emotion estimation net (ESiN) Model	80.00%
Supervised machine learning with SNoW learning architecture	63.00%

Results show that the proposed model has higher accuracy than the other three models. The nearest one is the linguistic, pragmatic and keyword spotting features model, which is 1.52% less than the proposed model. However, this model is focused on the recognition of emotions in the natural language dialogs for entertainment rather than generic text as covered by the proposed model.

5.6 Proposed Model versus Multiple Labels Classifier Models

In multi label classification problem, the detected emotion is one of an emotion classes $E = \{em_1, em_2, \dots, em_k, \dots\}$, where em_k denotes a specific emotion. Table 12 shows the detection accuracy for the proposed model

compared with three multi label emotion classification models, linguistic, pragmatic, and keyword spotting features model, emotion estimation net (ESiN) model (Tao, 2004) and recognition from text using knowledge-based ANN. The observed results proved that accuracy improvements can be obtained through the proposed model. Similar to the binary classification experiment, linguistic, pragmatic, and keyword spotting features model has the nearest accuracy results with 2.76% less. Taking into consideration the limitation of dialogs specific nature, this model could not fit in other types of texts that the proposed model can process.

Table 12. multiple labels accuracy evaluation

Model	Accuracy
Proposed Model_	87.00%
Linguistic, pragmatic, and keyword spotting features	84.24 %
Emotion estimation net (ESiN) Model	70.00%
Recognition from Text Using knowledge-Based ANN	65.00%

6. Conclusion

This study proposed frame semantics based evolutionary model for emotion detection in text. The proposed model deals any textual information that maybe different text in domain, linguistic type, length, nature and sentences complexity. It has been realized that the frames knowledge set is significantly reduced. Synonyms and opposites of frame entities are semantically processed to generate frames that are not initially trained showing extensibility of other models if built using this method. The observed results show relatively good accuracy at low training frame set. The matching performance versus accuracy is controlled by variant search approaches; first matching, best matching, best opposite matching and average matching. Although domain free text, the proposed model assured better accuracy results compared with other models although less emotional frames knowledge set. In addition, the model fits for binary emotion detection and multi emotions classifications with enhanced accuracy results. Further enhancements to the evaluation results could be achieved by future training sets in both the semantic frames and emotion rule based knowledge. Despite the generic advantage of the proposed study, the application of the model for domain specific emotions classification could, as well, provide more accuracy results.

References

- Agrawal, A., & An, A. (2012). Unsupervised Emotion Detection from Text using Semantic and Syntactic Relations. *WI-IAT '12 Proceedings of the 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology, 1*, 346-353. <http://dx.doi.org/10.1109/WI-IAT.2012.170>
- Alm, C. O., Roth, D., & Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, Association for Computational Linguistics, Vancouver, Oct. (pp. 579-586). <http://dx.doi.org/10.3115/1220575.1220648>
- Aman, S. (2007). *Recognizing Emotions in Text*. M. Sc. Thesis, Ottawa-Carleton Institute for Computer Science, School of Information Technology and Engineering, University of Ottawa, Canada.
- Balahur, A., Hermida, J. M., Montoyo, A., & Muñoz, R. (2011). EmotiNet: A Knowledge Base for Emotion Detection in Text Built on the Appraisal Theories. *Natural Language Processing and Information Systems, Lecture Notes in Computer Science, 6716*, 27-39. http://dx.doi.org/10.1007/978-3-642-22327-3_4
- Chuang, Z. J., & Wu, C. H. (2004). Multi-Modal Emotion Recognition from Speech and Text. *The Association for Computational Linguistics and Chinese Language Processing, 9(2)*, 45-62.
- Jagadeesh, R. P., Bose, C., & Nagaraja, G. (2004). Performance Studies on KBANN (pp.198-203). *Fourth International Conference on Hybrid Intelligent Systems (HIS'04)*. <http://dx.doi.org/10.1109/ICHIS.2004.74>
- Lee, C., & Lee, G. G. (2007). Emotion Recognition for Affective User Interfaces using Natural Language Dialogs. *16th IEEE International Conference on Robot & Human Interactive Communication*, August 26-29, Korea.
- Lu, C. Y., Hong, J. S., & Cruz-Lara, S. (2006). Emotion Detection in Textual Information by Semantic Role Labeling and Web Mining Techniques. *Third Taiwanese-French Conference on Information Technology –*

TFIT.

- Lu, C. Y., Hsu, W. W. Y., Peng, H. T., Chung, J. M., & Ho, J. M. (2010). Emotion Sensing for Internet Chatting: A Web Mining Approach for Affective Categorization of Events (pp. 295-301). *13th IEEE International Conference on Computational Science and Engineering*.
- Pak, A., Bernhard, D., Paroubek, P., & Grouin, C. (2012). A Combined Approach to Emotion Detection in Suicide Notes. *Biomedical Informatics Insights*, 5(Suppl. 1), 105-114.
- Picard, R. W. (1997). *Affective Computing*. Cambridge: MIT Press.
- Quan, C., & Ren, F. (2009). Construction of a Blog Emotion Corpus for Chinese Emotional Expression Analysis. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 1446-1454).
- Quan, C. (2010). Sentence Emotion Analysis and Recognition Based on Emotion Words Using Ren-CECps. *International Journal of Advanced Intelligence*, 2(1), 105-117.
- Seol, Y. S., Kim, D. J., & Kim, H. W. (2008). Emotion Recognition from Text Using Knowledge-based ANN, The 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 2008 (pp. 1569-1572).
- Shivhare, S. N., & Khethawat, S. (2012). Emotion Detection from Text. *Computer Science & Information Technology*, 5, 371-377.
- Strapparava, C., & Mihalcea, R. (2008). Learning to Identify Emotions in Text. *SAC'08 Proceedings of the 2008 ACM symposium on Applied computing* (pp. 1556-1560). <http://dx.doi.org/10.1145/1363686.1364052>
- Tao, J. (2004). Context Based Emotion Detection from Text Input. *8th International Conference on Spoken Language Processing, ICSLP*, Oct. 4-8 (pp. 1337-1340).
- Tokuhisa, R., Inui, K., & Matsumoto, Y. (2008). Emotion Classification Using Massive Examples Extracted from the Web. *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, Manchester, Aug. 2008 (pp. 881-888).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).