Enhancing text using emotion detected from EEG signals

Akash Gupta · Harsh Sahu · Nihal Nanecha · Pradeep Kumar · Partha Pratim Roy · Victor Chang ·

Abstract Often people might not be able to express themselves properly on social media, like not being able to think of appropriate words representative of their emotional state. In this paper, we propose an end to end system which aims to enhance user-input sentence according to his/her current emotional state. It works by a) detecting the emotion of the user and b) enhancing the input sentence by inserting emotive words to make the sentence more representative of the emotional state of the user. The emotional state of the user is recognized by analyzing the Electroencephalogram (EEG) signals from the brain. For text enhancement, we modify the words corresponding to the detected emotion using correlation finder scheme. Next, the verification of the sentence correctness has been performed using Long Short Term Memory (LSTM) Networks based Language Modeling framework. An accuracy of 74.95% has been recorded for the classification of five emotional states in a dataset of 25 participants using EEG signals. Similarly, promising results have been obtained for the task text enhancement and overall end-to-end system. To the best of our knowledge, this work is the first of its kind to enhance text according to the emotional state detected by EEG brainwaves. The system also releases an individual from thinking and typing words, which might be a complicated procedure sometimes.

Akash Gupta
E-mail: akii.gupta16@gmail.com
Harsh Sahu
E-mail: harshsahu23@ymail.com and
Nihal Nanecha
E-mail: nihal07.iitr@gmail.com
Pradeep Kumar
E-mail: pradeep.iitr7@gmail.com
Partha Pratim Roy
E-mail: proy.fcs@iitr.ac.in
Department of Computer Science & Engineering, IIT Roorkee, India

Victor Chang
E-mail: victorchang.research@gmail.com
International Business School Suzhou, Xi'an Jiaotong-Liverpool University, China.

1 Introduction

We use language not just to convey facts, but also our emotions. In the recent years, there has been a huge inflow of data onto social media platforms [48]. People use these platforms to share their daily activities and communicate with the world. Most of the communication on social media platforms is done through text. A number of knowledge discovery systems have been developed by analyzing this data [52]. While expressing himself on social media, not every person is able to express his emotions in the text he writes. For example, when one reads the sentence, *I am lonely*, the person gets to know that the writer of the sentence was feeling sad when wrote this sentence. However, when someone reads the sentence, *I am watching a movie*, the reader might not be able to know what the writer of the sentence is feeling. He might be feeling happy, sad, surprised, or any other emotion. While expressing himself over social media, the writer might not be able to find the proper words to describe his emotional state. This paper proposes a novel system to help users to enhance their text input to social media through emotion analysis.

To discover knowledge through emotion analysis, researchers usually explore three components of emotions, namely the psychological component, the subjective component and the expressive component [9]. The psychological component can be understood by our body's reaction to emotions, like text [53], facial expressions [23], speech [41] or a combination of them. Also, internal signals like heart beat [21], pulse, and brain waves can also be used to recognize the emotional state of a person. There has been a lot of research in the field of mining of social media data to develop various knowledge discovery systems. For example, sentiment analysis is an important and growing field of research where the sentiment (positive and negative) is derived from the text posted on social media [1]. Emotional analysis of social media data takes it one step forward by differentiating the sentiment into different emotions [51]. All these works focus on obtaining the emotions/sentiment given a sample text.

Electroencephalogram (EEG) has been widely used by researchers to capture the brainwave signals for determining the emotional state of the user [29]. EEG based devices have also been used for Brain-Computer-Interfaces (BCI) like cursor control and thought driven robots [3]. They have also been applied to serve as a user identification method in [20, 24]. They also have several applications in the medical field [45]. Emotion recognition also has many applications in different domains such as education, commerce, security [7] and others [49]. Similarly, Long Short Term Memory (LSTM) networks have been used extensively in time series prediction [40], speech recognition [15], grammar learning [14], protein homology detection [18], handwriting recognition [16], etc.

In this paper, we propose a novel system which recognizes the emotional state of the user using the EEG signals captured and next understanding the emotional state of the user, the input-text of the user is enhanced to better depict the emotional state of the user. Here, the transformed sentence is an enhanced version of the original sentence, containing more emotive words. This releases the individual

of thinking and typing words, which might be a complicated procedure sometimes. The sentence provided by the user is next enhanced by text transformation module. The words corresponding to the detected emotion are first modified using correlation finder between words and emotion. Next, the correctness of the sentence is verified with LSTM networks based Language modeling. The proposed emotional-aware text enhancement system is more representative of the emotional state of the user. This is helpful in many ways as it relieves the user of thinking complicated words. It also helps the organizations and companies to better know their users as the opinion of the users is better represented. Also, the target audience of the user is better aware of the emotional state of the user while reading the enhanced text. The main contributions of the paper are as follows:

- Our first contribution is the recognition of user's emotional state while watching videos with the help of EEG signals.
- Our second contribution is text transformation module where the input-text of the user is enhanced using correlation finder between words and detected emotion.
- Finally, the correctness of the sentence is verified with LSTM networks based Language modeling framework.

The rest of the paper is organized as follows. Next, section presents the previous work done in various fields related to our paper. Then Section 3 presents the methodology of the system proposed in the paper. Dataset is described in Section 4. We conducted experiments of our modules individually as well as overall system, which are presented in Section 5.1 - 5.3 and discussion is presented in Section 5.4. Finally, Section 6 concludes with future work.

2 Related Work

In this section, we present a brief description of the work done in fields related to our work. In the following sub-sections, the work done in the field of emotion recognition using EEG brainwaves and text transformation are discussed.

2.1 EEG Emotion Recognition

Initial works include [25] where only 2 emotions, i.e., happiness and sadness were classified, which were evoked by showing subjects pictures of smiling and crying facial expressions. They used Common Spatial Patterns (CSP) and linear Support Vector Machine (SVM) to classify these 2 emotions. Emotions were further increased to 4 (joy, anger, sadness, pleasure) by [27]. The authors used SVM to classify emotions but used music listening approach to collect the EEG-data. A combination of EEG responses and eye gaze data was used in [44] where ground truth was based on average arousal and valence scores given to clips by subjects using online questionnaire. Valence Arousal Dominance (VAD) model was used in [6] to classify the video segments of different emotions. VAD values were calculated from alpha beta power spectral density in different parts of the brain. Fractal dimension method based on Higuchi algorithm [17] was applied to calculate valence arousal values on the data collected from users listening to different audio

clips from the Quran [12]. An EEG-based web-enabled music player that display music according to the user's emotional state was designed in [30]. The authors implemented emotion induction experiments with music and sound stimuli.

Recently, emotion recognition using a combination of facial expressions features and EEG signal has also been studied in [31]. The authors used a standard dataset, DEAP [22] where subjects watched emotion-evoking videos. Higher Order Crossings (HOC) based feature extraction scheme is discussed in [35] for emotion classification with the help of four different classifiers.Almehmadi and Bourque [2] presented a model to determine the emotional state of the user posting on social media. The emotional state as well as the geographical information were appended at the end of the tweet to add additional context to the tweet which might have been missed in the original tweet itself. Our work builds upon that context, as it aims to enhance the tweet itself according to the emotional state rather than just appending the emotion at the end of the tweet. Wang et al. [50] have proposed a system for assessing the memory workload levels using EEG signals. The authors have used the n-back task to classify the different memory workload levels for nine participants. Similarly, the authors in [38] have used EEG signals in conjunction with offline signature to develop a multimodal biometric system. The brain signals were acquired simultaneously during signature and the identification process was performed using Hidden Markov Model (HMM) classifier.
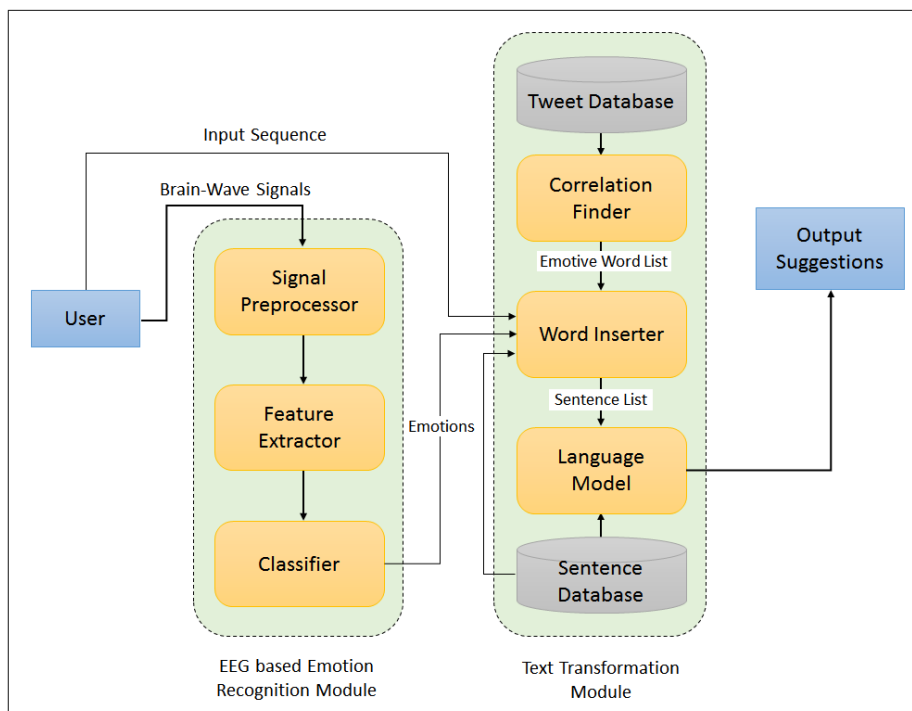
2.2 Text Transformation

There exist a number of work to discover knowledge using text transform at structural level, like transforming linear text documents into hypertext [13] or transforming XML documents into OWL ontologies [46]. It focused on determining relation in the text by examining the formatting of the input document. Text normalization [28] aims to convert words to standard dictionary forms from non standard forms (such as sms). Text normalization only aims at correcting existing words and does not provide for inserting new words. Text summarization is another area of research which extracts important details from a paragraph of sentences, but it uses the original sentences and words to compose the summary. Text simplification [10] is used to automatically convert a sentence for easier understanding. It is done by guessing meanings for the low frequency words. Research in the area of actually adding words to sentences is scarce. In linguistic personalization [37], skeleton sentences were enriched with adjectives and adverbs based on simple models using basic Point-wise Mutual Information. The Text transformation module of our proposed work differs from the linguistic personalization work because it involves more granular emotions rather than positive and negative sentiments and improves on the suggestions generated by using advanced techniques including LSTM language models. This makes the system more robust and capable of handling social media text for knowledge discovery.

Becker et al. [4] have proposed an emotion classification system in texts in multilingual context using a corpora written in two languages. Two different strategies of combining translated texts, i.e., the combination of features from multiple languages and the stacking of monolingual classifiers have also been used to boost the system performance. A relevance based language modeling framework for developing a recommendation system is proposed in [34]. The authors have proposed a

probabilistic clustering technique for performing the neighbor selection to achieve an approximation of the set of relevant items. Likewise, the authors in [36] have used language modeling techniques to characterize the syntactic and structural complexity of web search queries with natural language.

3 Proposed Methodology

As discussed earlier the proposed system is made up of two subsystems, namely EEG based emotion recognition and text transformation. Fig. 1 describes the proposed system via data flow diagram. The EEG based emotion recognition module detects the emotion of the user through EEG brainwaves and the Text transformation module enhances the text input by the user according to the emotion detected by the emotion recognition module. Each module of the system is described in details in the following sub-sections.


Fig. 1 Data flow Diagram: The brainwave signals are captured from the user to determine the emotional state which is fed to the text transformation module next to enhance the user-given text.

3.1 EEG based emotion recognition module

The EEG based emotion recognition module is used to identify the emotional state of the user by capturing the brainwave signals. The works by capturing the brainwaves and then preprocessing the signal. Features are extracted from the preprocessed signal which are used to determine the emotional class of the signal.

3.1.1 EEG System Description

Brainwave activities from a user wearing EEG device are recorded by the brain sensor device Emotiv EPOC. This device has 14 sensors, shown in Fig. 2, **which are placed over the scalp of participant as per international 10-20 system which is an internationally recognized method to describe and apply the location of scalp electrodes in the context of an EEG exam. The system is based on the relationship between the location of an electrode and the underlying area of the brain, specifically the cerebral cortex.** The number after the letters represent whether that sensor is on left side or right side of the brain e.g. odd number denotes the left side of the brain and even number represent right side of the brain. The device sends the raw EEG signals via bluetooth and is received on the computer through Software Development Kit (SDK). The device records the signals at a sampling rate of 128Hz.
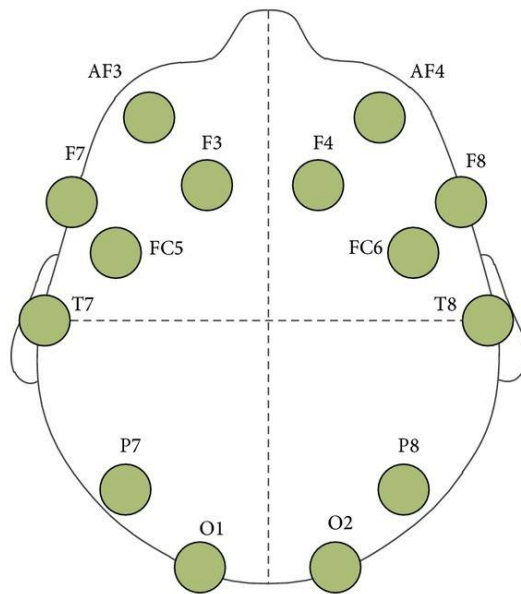
Fig. 2 Electrode map of EEG Emotiv EPOC+ device. The electrodes position are governed by international 10-20 system.

3.1.2 Signal Preprocessing

The system was set up in a calm environment and subjects were told not to make unnecessary movements. However, several unwanted noise occurred due to hair, and muscular activities, which needs to filtered out. A number of filters e.g. IIR or FIR filter, forward-backward filter and Savitzky-Golay filter were tested to filter the noise. From these filters, we obtained best results with Savitzky-Golay filter [39]. This filter minimizes the noise present in the signal by increasing signal-to-noise ratio, also without distorting the signal by any considerable amount. The coefficients corresponding to that polynomial are $(C_{-n}, C_{-(n-1)}...C_{n-1}, C_n)$ which are called as convolution integers for fitting data. Smoothed data points $(S_k)_g$ can be calculated for the raw signal $(X_1, X_2...X_n)$ as given in Eq. (1).

$$(S_k)_g = \sum_{i=-n}^{n} C_i * X_{k+i} \tag{1}$$

After applying the filter, the signal is broken into smaller duration at regular time interval equal to 1 second. This technique is applied because analysis of smaller segments is much more efficient compared to larger segments.

3.1.3 Feature Extraction

There are 22 features which are extracted from processed signal via statistical values and wavelet transform. A feature vector FV[1...22] is constructed for each channel. Since there are 14 channels in EEG device, there are total $14 * 22 = 308$ features. Complete description of feature vector(FV) is described below:

– **FV[1]:** This feature vector corresponds to the mean of the raw signal and computed using (2).

$$\mu_X = \frac{1}{N} \sum_{n=1}^{N} X_n, \tag{2}$$

where X is input time series signal and N is size of the signal.
– **FV[2]:** In this feature, we have computed standard deviation of the raw signal with the help of (3).

$$\sigma_X = \sqrt[2]{\frac{1}{N} \sum_{n=1}^{N} (X_i - \mu_X)^2}, \tag{3}$$

where X is input time series signal and N is size of the signal.
– **FV[3]:** This feature corresponds to the variance of the raw signal calculated using (4), where X is input time series signal and N is size of the signal.

$$Var(X) = \sigma_X^2 = \frac{1}{N} \sum_{n=1}^{N} (X_i - \mu_X)^2, \tag{4}$$

– **FV[4]:** This feature is equivalent to the mean of the absolute values of the first differences of the raw signal.

$$\delta_X = \frac{1}{N-1} \sum_{n=1}^{N-1} \left| X(n+1) - X(n) \right|, \tag{5}$$

where X is input time series signal and N is size of the signal.

- **FV[5]:** In this feature, the mean of the absolute values of the first differences is calculated using (6).

$$\overline{\delta_X} = \frac{1}{N-1} \sum_{n=1}^{N-1} \left| \overline{X}(n+1) - \overline{X}(n) \right| = \frac{\delta_X}{\sigma_X},$$ (6)

where X is input time series signal and N is size of the signal.

- **FV[6]:** Here, the mean of the absolute values is calculated with the help of Eq. (7), of the second differences of the raw signal.

$$\gamma_X = \frac{1}{N-2} \sum_{n=1}^{N-2} \left| X(n+2) - X(n) \right|$$ (7)

where X is input time series signal and N is size of the signal.

- **FV[7]:** This feature is calculated using (8) and is equivalent to the mean of the absolute values of the second differences of the standardized signal.

$$\overline{\gamma_X} = \frac{1}{N-2} \sum_{n=1}^{N-2} \left| \overline{X}(n+2) - \overline{X}(n) \right| = \frac{\gamma_X}{\sigma_X}$$ (8)

where X is input time series signal and N is size of the signal.

- **FV[8]:** Mobility of the raw signal calculated using (9), where X is time series signal, $D1$ is first derivative of signal X and σ_X is defined in Eq. (3).

$$Mobility_X = \frac{\sigma_{D1}}{\sigma_X}$$ (9)

- **FV[9..12]:** *Power in different frequency bin:* Theta(4-8Hz), Alpha(8-12Hz), Beta(12-30Hz), Gamma(30-128Hz) is used. Power for i^{th} bin is calculated as:

$$Power_i = \sum_{n=Bs_i}^{Be_i} C(n) * C(n)$$ (10)

where $Power_i$ is spectral power corresponding to i^{th} frequency bin, Bs_i and Be_i are:

$$Bs_i = \frac{S_i * len(C)}{Fs}$$ (11)

$$Be_i = \frac{E_i * len(C)}{Fs}$$ (12)

where S_i and E_i are starting and ending frequency of ith bin, $len(C)$ is length of C and C is Fourier transform of time series signal X. Fs is sampling rate in physical frequency.

- **FV[13..16]:** *Ratio power in bins:* It is calculated for i_{th} band using (13), where $Power\_Ratio_i$ is the fraction of spectral power in i^{th} bin and $Power_i$, is defined in Eq. (10).

$$Power\_Ratio_i = \frac{Power_i}{\sum_{x=1}^{4} Power_x}$$ (13)

– **FV[17]:** Petrosian Fractal Dimension of raw signal can be defined using (14), where n is length of time series signal X, N_{delta} is number of times first derivative ($D1$) changes sign i.e. $D1(i) * D1(i+1) < 0$.

$$PFD_X = \frac{\log_{10} n}{\log_{10} n + \log_{10} \frac{n}{n+0.4*N_{delta}}} \quad (14)$$

– **FV[18,19]:** Hjorth Mobility and Complexity of raw signal are calculated using (15) - (17).

$$TP_X = \sum_{n=1}^{N} X(i)^2, \quad (15)$$

where $X(i)$ is the ith component of the input time series signal X and N is size of the signal.

$$Hjorth\ Mobility_X = \sqrt[2]{\frac{TP_{D1}}{TP_X}}, \quad (16)$$

$$Hjorth\ Complexity_X = \sqrt[2]{\frac{TP_X * TP_{D2}}{TP_{D1} * TP_{D1}}} \quad (17)$$

where X is input time series signal, $D1$ is first derivative of signal X, $D2$ is second derivative of signal X.

– **FV[20]:** Spectral Entropy calculated using average band power from Eq. (10) and using (18), where N_b is number of bands and $Power\_Ratio_i$ is defined in Eq. (13).

$$Spectral\ Entropy_X = \frac{-1}{\log N_b} * \sum_{i=1}^{N_b} Power\_Ratio_i * \log\left(Power\_Ratio_i\right) \quad (18)$$

– **FV[21]:** Singular Value Decomposition (SVD) Entropy of the raw signal can be calculated using (19), where X is input time series signal, τ and d_E are the delay and the embedding dimension, respectively. The embedding space is then constructed using (20).

$$y(i) = [X(i), X(i+\tau), ..., X(i+(d_E-1)\tau)] \quad (19)$$

$$Y = [y(1), y(2), ..., y(N(d_E-1)\tau)]^T \quad (20)$$

The SVD is then calculated using matrix Y to get M singular values (e.g. $\sigma_1, ..., \sigma_M$) defined in Eq. (3) known as the singular spectrum. The SVD entropy is defined using (21), where M and $\sigma_1, ..., \sigma_m$ are the number of singular values and normalized singular values by the term $\overline{\sigma_i} = \frac{\sigma_i}{\sum_{j=1}^{M} \sigma_j}$, respectively.

$$SVD\ Entropy_X = -\sum_{i=1}^{M} \overline{\sigma_i} * \log_2 \overline{\sigma_i} \quad (21)$$

– **FV[22]:** Fisher info calculated using (22) from standard deviation of raw signal, where M and $\sigma_1, ..., \sigma_M$ are used from Eq. (21).

$$FI_X = \sum_{i=1}^{M-1} \frac{(\overline{\sigma_{i+1}} - \overline{\sigma_i})^2}{\overline{\sigma_i}} \quad (22)$$

The summary of all the features described above is given in Table 1.

Table 1 Summary of features used in EEG based emotion classification, where X is raw signal and $\overline{X}$ is standardized signal

FV[1]:	Mean of X	**FV[12]:**	Power of Gamma band of X
FV[2]:	Standard Deviation of X	**FV[13]:**	Power ration of Theta band to overall power
FV[3]:	Variance of X	**FV[14]:**	Power ration of Alpha band to overall power
FV[4]:	Mean of first different of X	**FV[15]:**	Power ration of Beta band to overall power
FV[5]:	Mean of absolute values of first difference of X	**FV[16]:**	Power ration of Gamma band to overall power
FV[6]:	Mean of absolute values of second difference of X	**FV[17]:**	Petrosian Fractal Dimension of X
FV[7]:	Mean of absolute values of second difference of $\overline{X}$	**FV[18]:**	Hjorth mobility of X
FV[8]:	Mobility of X	**FV[19]:**	Hjorth complexity of X
FV[9]:	Power of Theta band of X	**FV[20]:**	Spectral Entropy
FV[10]:	Power of Alpha band of X	**FV[21]:**	Singular value Decomposition Entropy
FV[11]:	Power of Beta band of X	**FV[22]:**	Fisher Info

3.1.4 Emotion Classification using Random Forest Classifier

A total of 308 features were obtained after feature extraction. Data was trained using a Random Forest classifier having 100 trees. Random Forest classifier [26] has been successfully used to develop various knowledge discovery systems like recommendation system, sentiment analysis, etc. The classifier is an ensemble of decision trees where the vote of the trees decides the final class for a sample. Each decision tree is build on a bootstrap sample of the whole data and acts as an independent estimator. A bootstrap sample is obtained by selecting a random sample from the whole data with replacement, i.e., each sample is obtained from the whole data rather than the depleted data. A decision tree tries to predict the value of the target variable using simple decision rules based on the features. Tree based methods have an advantage over other methods like SVM and Logistic regression because they scale well to large number of data points, handle noise well, handle categorical values well and are quick to train. An ensemble of trees is helpful in preventing over fitting. The Algorithm for training a Random Forest classifier is summarized as in Algorithm 1. To determine the final class of a sample point x, it is passed through all the T trees of the forest until it reaches the leaf. The final class is given by Eq. (23), where $p(c|x)$ is the probability the given point belongs to class c considering all the trees, T is the number of trees, $p_t(c|x)$ is the probability that the given point belongs to class c by considering only the tree T.

$$p(c|v) = \frac{1}{T} \sum_{t=1}^{T} p_t(c|x) \tag{23}$$

3.2 Text transformation module

This subsection describes the various components of the Text transformation module. It receives the detected emotion from the EEG based emotion recognition

---

Algorithm 1 Training a Random Forest Classifier

---
1. Draw bootstrap samples from the whole data;
2. Construct a decision tree for each bootstrap sample by recursively splitting by a randomly selected subset of features;
3. Repeat the above steps until a predefined number of tress are grown;

---

module and the text input by the user. It outputs the suggestions generated by enhancing the text using the detected emotion. It contains 3 major components, namely-Correlation finder, Word inserter and Language modeling. The Correlation finder is pre-trained to find the lists of emotive adjectives and adverbs. The Word inserter generates a list of suggestions by inserting words into the original text and the Language model is used to rank the generated suggestions and prune the less probable ones.

3.2.1 Correlation Finder: Finding Correlation between words and emotions

Tweets tagged with corresponding text labels are used as training instances to a logistic regression model. Since, only the adjectives and the adverbs usually describe an emotion, all other words from the training corpus are removed for this task. The features extracted for the model were unigrams of the adjectives and adverbs used. The Logistic Regression model finds a linear combination of the features which describes the output variable the best (minimizes the error between actual output and the predicted output). The equation of the Logit function can be given as:

$$log_e(\frac{P}{1-P}) = \beta_0 + \sum_{i=1}^{k} \beta_i * X_i, \tag{24}$$

where X_i denotes a vocabulary word in a sentence, k is the vocabulary size of the dataset, P is the probability that output variable is equal to the given emotion and β_1, β_2.. are the coefficients of the features. Here, the coefficients of each feature (adjective or adverb) indicate how much that feature is correlated with the emotion. For each emotion class, the words having a correlation coefficient above a threshold T are kept. The threshold is found out by empirical testing. From experiment results, it was noted that T=0.8 provides good results. The list obtained by this method is referred as L1.

Strength of Association: In order to expand the vocabulary, another model named Strength of Association [32] was trained to get more adjectives and adverbs strongly associated with an emotion class. It is used as a measure to compute the association between a word and an emotion class.

The Strength of Association (SoA) for a given dataset of sentences (w) and their emotion labels (e) is computed using (25), where PMI is the point-wise mutual information calculated using (26),

$$SoA(w,e) = PMI(w,e) - PMI(w,\neg e) \tag{25}$$

$$PMI(w,e) = log_2 \frac{freq(w,e) * N}{freq(w) * freq(e)} \tag{26}$$

where $freq(w,e)$ denotes the number of times w is encountered in a sentence with label e. The terms $freq(w)$, $freq(e)$ are the frequencies of w, e in the labeled corpus and N represents the total words in the dataset.

$$PMI(w, \neg e) = log_2 \frac{freq(w, \neg e) * N}{freq(w) * freq(\neg e)} \qquad (27)$$

where $freq(w, \neg e)$ denotes the times w occurs in a sentence that does not have the label e. Thus, Eq. (25) is simplified to (28).

$$SoA(w, e) = log_2 \frac{freq(w, e) * freq(\neg e)}{freq(e) * freq(w, \neg e)} \qquad (28)$$

A word is said to be strongly associated with a label if it has a SoA score greater than 0. We made a list of each emotion class and the strongly associated words for that class. Let this list be referred as L2. As strength of association is not a good measure for words occurring with low frequencies, the words having frequencies less than 50 are removed from the training corpus.

Next, both these lists L1 and L2 are expanded by adding all other forms of adjectives and adverbs and also the possible spelling corrections. As there are some noisy words captured in L1 and L2, these lists are filtered by using the word-emotion lexicon [33]. The words not corresponding to an emotion in the word-emotion lexicon are removed from the lists of L1 and L2. Finally, lists L1 and L2 are merged to form a combined list L3 which contains the final list of adjectives and adverbs used by other modules.

3.2.2 Word Inserter: Inserting Adjectives for Nouns and Adverbs for Verbs

Next, we process a pair of sentences and given emotion one at a time. The nouns and verbs in the given sentence are sorted by their inverse document frequencies with respect to the training corpus. For each noun having an inverse document frequency above a certain threshold, a list of adjectives corresponding to the given emotion is fetched. The list is then pruned by removing the adjectives which do not have any co-occurrence with the noun in the training corpus or have the point-wise mutual information (PMI) lower than a threshold T1. For a pair of words (a and b),

$$PMI(a, b) = \log_2 \frac{p(a, b)}{p(a) * p(b)} \qquad (29)$$

where p(a) and p(b) refer to the occurrence probabilities of a and b, and the joint probability p(a, b) is given p(a)p(b|a).

Now, for each noun, there is a list of adjectives having good chances of co-occurrence with that noun. A similar procedure is repeated to find the list of adverbs for each verb having inverse document frequency above a threshold. New sentences are created by inserting each word from both lists (adjectives and adverbs). This list of new sentences is checked for consistency. This is done by POS tagging (Part of Speech Tagging) each sentence and checking if the new word inserted is actually being used as an adjective for a noun or adverb for a verb. The inconsistent sentences are next removed from the new sentences list and this modified list is called L4.

3.2.3 Language Model

From the list of new sentences L4, the sentences which are most probable are chosen. For this purpose, a language model is used to find the probability of each sentence. A language model is defined as a probability distribution over sequences of words. The model assigns a probabilities $P(w_1, \ldots, w_m)$ to a sequence of length m. Thus, the probability of a sentence can be calculated using (30).

$$p(w_1, ..., w_N) = \sum_{i=1}^{N} p(w_i | w_1, ..., w_{i-1})$$
(30)

The probability of sentence is normalized by sentence length to prevent the model penalizing longer sentences.

The sentences are next sorted by the their normalized log probabilities and the sentences with normalized log probabilities below a threshold are removed. The final list of sentences is provided to the user as a set of suggestions.

In this work, for Language Model (LM) we used a BIG LSTM+CNN language model described in [19]. This model performs better compared to parametric (e.g., log-linear models) and nonparametric approaches (e.g., count-based LMs). The Long Short Term Memory (LSTM) is a Recurrent Neural Network (RNN) which is especially useful as a sequence learning model. A representation of an unrolled RNN network is shown in Fig. 3. The X_i is the input at time i while A is a RNN cell and h_i is output at time i. The RNN cell is used as a memory unit as well as an activation unit.
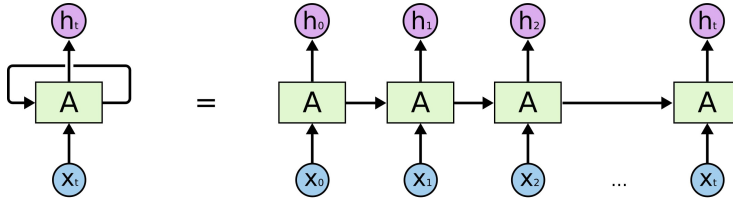

Fig. 3 An unrolling of a Recurrent Neural Network.

An LSTM network is made up of LSTM units which specialize in remembering values for either a short or long period of time. An LSTM does not have an activation function in its recurrent components which prevent the remembered things to be squashed over time. LSTM networks help retain the dependencies between words far away in the sentence. LSTM blocks are usually trained with Backpropagation through time. A CNN (Convolutional Neural Network) Softmax layer is used because character-level features help to make a smoother and compact parametrization of the word embeddings. The CNN Softmax layer is used to predict the probability of a word by computing its logit $z_w = h^T * e_w$ where h is a context vector and e_w the word embedding. The probability of a word is given by

$$p(w) = \frac{exp(z_w)}{\sum_{w' \epsilon V} exp(z_{w'})}$$
(31)

where z_w is logit of the word w and V is the vocabulary. The LSTM is a word level model which is combined with a character level model to improve the training time because word level models tend to take large amounts of time especially when the vocabulary size is very large (as in our case).

4 Dataset

In this section, we discuss the details of the datasets, and the preprocessing steps that have been used to remove noise from text.

4.1 Tagged Tweets Dataset

For the correlation finder, we use a dataset of tweets and tag them by their hashtag describing an emotion. The tweets were collected from the tweet ids given in [51]. For all access to Twitter API, we used python library Tweepy, which used the Twitter stream API and Twitter search API internally.

Five basic emotion categories were defined in existing psychology literature [42]. In this work, the emotion categories were joy, sadness, anger, surprise and fear. Each emotion category contained a list of subcategories under it. We checked the ending hashtag of each tweet to match it with corresponding emotion category. The tweet was assigned to an emotion category if the lemmatized version of its ending hashtag was contained in that emotion category. The number of tweets in emotion category is described in Table 2.

Table 2 Tweet Count for each emotion in Tagged Tweets Dataset

Emotions	Number of tweets
Sadness	336953
Joy	327810
Anger	304466
Fear	59220
Surprise	12943

Preprocessing:

- **Lowercase:** All the tweets are converted to lowercase.
- **Web Links:** Web Links are the URL addresses sometimes included in tweets. They do not depict any emotion and hence can be removed.
- **Usernames:** Sometimes user mentions like @nihal, @sahu are included in tweets. These words do not depict any emotion and hence are not required for the our task. So all user mentions are made anonymous to @user
- **White-spaces:** Multiple white-spaces and carriage returns were combined to to form a single white-space separator.
- **Numbers:** As any particular number is not associated with an emotion, all the occurrences of numbers are replaced with the string 'num'. Note that only the numbers preceded and succeeded by a white-space were replaced.

- **Stop-words:** The common words like a, an, the, is, are, etc in the text are known as stop-words. These were removed as they are usually present in most of the sentences.
- **Ending Hashtags:** The hashtagged words at the end of the tweets were removed because they were earlier used to tag the tweet with the emotion.
- **Hashtag and Special Symbols:** The hashtags in the middle of tweets and other special symbols like %,;*, etc. were also removed.

4.2 Language Modeling and PMI info Dataset

For the language modeling and for calculating the PMI information to insert words, a larger dataset was required that does not need to be tagged with emotion class. For this purpose, the One Billion Word Benchmark Dataset [8] was used in this work. This dataset contains mostly news text so most of the sentences are well structured making our language model learn the structure of English language well. The dataset contains about 0.8 billion words (829250940 words). A vocabulary (793471 words) was constructed by removing the words occurring below the count 3 and replacing them by <UNK>token. The sentence order was randomized making the data non contextual.

5 Experiments and Discussion

We conduct 3 experiments to test the individual modules and end-to-end system. First experiment (Section 5.1) is aimed at testing the suggestions generated by the text transformation module. Second experiment (Section 5.2) is aimed to test the performance of the emotion classification module and third experiment (Section 5.3) aims to test the prowess of the end to end system which is a combination of the emotion classification module and the text transformation module. Section 5.4 presents the discussion for the experiments.

5.1 Transforming sentences given an emotion

In this experiment, we test the performance of our Text transformation module. To implement the Text transformation module, the Stanford POS Tagger was used to tag the tweets with their corresponding parts of speech [47]. The pre-trained model used by the tagger was given by [11]. The NLTK library [5] was used to tokenize tweets and as a python wrapper for POS tagger. Pattern.en [43] library was used for lemmatization and getting all forms of adjectives and adverbs. We used pre-trained LSTM language model on the One Billion Word Benchmark given by [19].

For this experiment a group of 10 people was asked to note a set of sentences which one would want to post to a social network and also inform the mental state they were in when stating that sentence. A total of 50 sentences were received and after filtering out similar sentences, we were left with a set of 32 sentences. We then applied our NLP module to those sentences and listed out all the suggestions it made for those sentences.
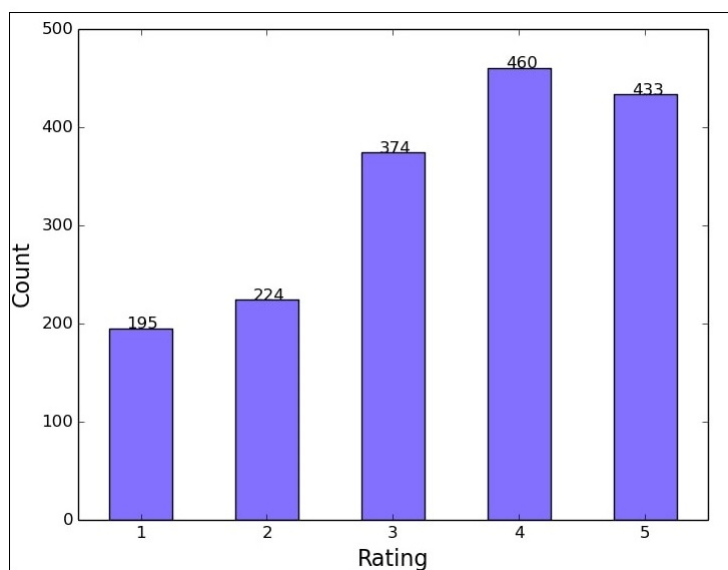
Fig. 4 Number of each Rating for Text Transformation.

Finally, a survey was created and 54 people (excluding those who were involved in posting sentences) were asked to rate the quality of those suggestions on a scale from 1 to 5 (1 means they are very less likely to use a suggestion and 5 means very much likely to use a suggestion). We got an average score of 3.39 for the suggestions. This means on an average, each participant is likely to use a suggestion suggested by our module. The detailed analysis of the responses is shown in Fig. 4. Some sample suggestions generated are shown in Table 3.

Table 3 Sentence Transformation Examples for each emotion

Emotion	Original Sentence	Suggestions
Anger	I am tired with my job.	1. I am tired with my ridiculous job. 2. I am tired with my awful job. 3. I am tired with my damn job.
Sadness	I don't like this world.	1. I don't like this bad world. 2. I don't like this miserable world. 3. I don't like this terrible world.
Surprise	That was awesome.	1. That suddenly was awesome. 2. That was unexpectedly awesome. 3. That astonishingly was awesome.
Joy	I love this world.	1. I love this wonderful world. 2. I love this lovely world
Fear	I have an infection.	1. I have an bad infection. 2. I have an unsafe infection. 3. I have an fatal infection

5.2 Detecting the emotional state of the user

This experiment was conducted to collect training data for the EEG emotion recognition module as well as to test the performance of the emotion recognition module.

5.2.1 Subject and Stimuli

A total of 25 subjects participated for the experiment. They were students of IIT Roorkee, age varying from 18 to 30 years. None of the subjects had any history of mental illness nor had performed any BCI activities before. It was ensured that subjects were not under stress during the experiment. A set of videos clubbed together sequentially was used as the stimuli for different emotions. The length of the each video ranged from 30 seconds to 2 minutes. There was a relaxation time of 10 sec between different videos. Each video contains certain emotion associated with it and the participant did not have prior knowledge of what emotion is associated with each video. Videos corresponding to different emotions are taken such that it can have strong stimuli for that emotion. The detailed description of the videos can be found in Table 4. **The criteria for selecting these videos is as follows. Initially, a set of 20 videos was proposed with 4 videos for each emotion. A group of 10 people with age between 22-40 with average of 26, were asked to select the best video for each emotion. Then, the video with the maximum votes has been chosen for every emotion.**

Table 4 Description of the videos shown to subjects to evoke emotions

Youtube Link	Emotion	Start Time (mm:ss)	End Time (mm:ss)	Duration (seconds)
https://www.youtube.com/watch?v=HuHVqxRpyqc	joy	00:10	01:30	80
https://www.youtube.com/watch?v=gOKyTtJg1O	surprise	01:36	03:36	120
https://www.youtube.com/watch?v=BPcipZxZoGI	anger	01:26	03:36	130
https://www.youtube.com/watch?v=GePiE3mh5gQ	sad	00:10	01:40	90
https://www.youtube.com/watch?v=JhMWopjJiI8	fear	00:30	02:40	130
https://www.youtube.com/watch?v=1OX1hBwNCJc	sad	00:50	01:42	52
https://www.youtube.com/watch?v=jXmluLg2XP0	sad	00:00	01:00	60
https://www.youtube.com/watch?v=qJ9AIz7OLBE	anger	00:00	00:36	36
https://www.youtube.com/watch?v=c9QeTThOxNg	joy	00:00	00:28	28

5.2.2 EEG Data Collection

Before beginning of the experiment, subjects were informed about the procedure, and were asked to behave normally as they would behave while watching any other video. Subjects were also asked about their will to participate as videos could trigger extreme emotions. They were also instructed not to make unnecessary movements as it could cause noise in the signal. Thereafter, EPOC device was mounted along the scalp of the subject. Next, the video was showed after the subjects felt comfortable. An illustration of a subject watching the video is shown in Fig. 5.

Fig. 5 A subject watching a video while wearing EPOC Emotive Device (EEG Device setup).

5.2.3 Data Processing and Feature Extraction

Once, the EEG signal is recorded through Emotiv EPOC device, data is split into different files where each file corresponds to the sub-part of the video. Here, each sub-part of the video contains a particular emotion associated with it. After this partitioning, each file has an emotion associated with it. Data samples thus obtained are used for preprocessing and feature extraction as explained in the methodology (Section 3.1.2 and 3.1.3).

5.2.4 Classification and Validation

The classification results were validated using K-fold cross validation. In K-fold cross validation, the data is divided into K subsamples. Out of which, 1 subsample is retained as the testing data and the remaining $(K-1)$ subsamples are used for training. This process is repeated $(K-1)$ more times where the other $(K-1)$ subsamples are kept as testing data once. The results are averaged across the K iterations to obtain final accuracy, average precision, recall and f1 scores.

In this work, we have used $K = 10$, i.e. 10-fold cross validation to validate the performance of our method. The overall accuracy was found out to be 74.95%. The class wise performance indicators are displayed in Table 5.
The confusion matrix (averaged across all folds) is displayed in Table 6.

Table 5 Class wise Performance Indicators for the task of emotion classification

Emotions	Joy	Surprise	Anger	Sadness	Fear
Precision	75.58%	78.30%	74.78%	74.41%	76.36%
Recall	89.32%	80.79%	62.04%	63.31%	79.28%
F1 score	0.81	0.79	0.67	0.68	0.77

Table 6 Confusion Matrix, where each row describes the true emotion and each column describes the % values classified as each emotion

Input	Output (%)				
	Joy	Surprise	Anger	Sadness	Fear
Joy	**89.32**	6.97	0.58	1.4	1.73
Surprise	9.2	**80.8**	4.7	1.95	3.36
Anger	5.88	10.73	**62.05**	11.32	10.03
Sadness	5.93	7.92	7.69	**63.32**	15.15
Fear	7.08	4.8	2.55	6.29	**79.28**

5.3 Detecting the emotional state of the user and using it to enhance the input sentence

In this experiment, we aim to test the end-to-end performance of our proposed system. For this, we prepared a set of 5 videos, 1 for each emotion (joy, sadness, surprise, anger and fear). 4 videos were shown to a set of 20 people who were asked to write a sentence after each video. The EEG signals were recorded for each participant while watching the videos. The setup was similar to what was done in experiment described in Section 5.2. The data collected in the previous experiment was used as training data for this experiment. The features were constructed for the last 5 seconds of each video and were passed into the classifier for prediction. Thus the emotional state of each participant was determined for the last 5 seconds of each video. A majority voting of the predictions for each second of 5 seconds provides the final emotional state. The emotional state determined and the sentence written for each video were given as an input to the Text transformation module which gave out a list of suggestions. Each participant was asked to rate the list of suggestions for each video on a scale of 1 to 5 (1 means very less likely to use a suggestion and 5 means very much likely to use a suggestion). A total of 20 * 4 = 80 responses were obtained. The analysis of the ratings is shown in Fig. 6. The average rating was found out to be 3.95. It shows that most of the users are much likely to use at least one of the suggestions generated by our system.

5.4 Discussion

As we can see from the results, majority of the subjects were willing to use the suggestions generated by our system. The emotion detection system is fairly accurate in detecting the emotional state of the user. It has an average accuracy of 74.95%. As we observe in Table 5, the precision scores for all the classes are around 75% barring the classes sadness and anger. The confusion matrix presented in Table 6 further strengthens the evidence that the classification of emotions of anger and sadness has not been as good as the other 3 emotions. One of the reasons behind this could be the fact that the anger video could not evoke anger in the subject watching that video. Further research on evoking anger in test subjects and EEG based detection of anger emotion can be carried out. For the other emotions, our system performed well with around 80% or more samples lying in the correct classes.
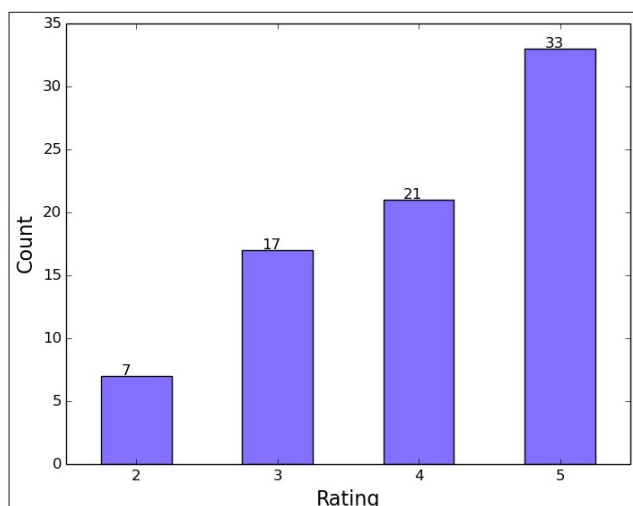
Fig. 6 Number of each Rating for end to end system.

The Text transformation module generates suggestions liked by most of the users. Out of the many suggestions, a few of the suggestions from the Text Transformation module were wrongly emotive or had a changed meaning from the original sentence, while the remaining suggestions retained the same meaning and were more emotive. This makes our system suggest at least one good emotive sentence in most of the cases.

Sometimes the words inserted, though grammatically correct and representative of the emotion, change the context of the sentence. For example *Original Sentence*-I am going to class, *Emotion*-Sadness, *Suggestion*-I am going to lowest class. Here, the human brain knows that going to lowest class has a very different meaning than simply going to class. But, for our system, these kind of suggestions are considered valid. But, applying LSTM Language model ensures that these kind of suggestions are ranked lower than the ones which retain the context.

6 Conclusion and Future Work

We have developed an end to end system which recognizes the emotional state of the user and enhances the text input by the user. Our system provides a transformation of a sentence by detecting the emotional state of the user and adding adjectives and adverbs and checking consistency by Language models. We have used some well established techniques including LSTM Language Models, POS tagging and Point-wise Mutual Information (PMI). From the results of the experiments, we have shown that the enhanced sentences are appreciated by most of the users.

Although appreciable results are obtained, in this work, we did not consider to change the semantic structure of the sentence. Research can be done to allow for modification of semantic structure of sentence which allows for changed placements of the words. While our system works well for simple sentences with one subject, it

struggles with complex sentences with more than one subject as it fails to recognize which part of the sentence to emote correctly. Example: *Original Sentence*-My daughter was killed by a gangster, *Emotion*-Sadness, *Suggestion*-My bad daughter was killed by a gangster. We noted that the bad word is being associated with the wrong subject which changes the complete meaning of the sentence. Research needs to be done to identify the correct subject related with the emotion.

References

1. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R.: Sentiment analysis of twitter data. In: Proceedings of the workshop on languages in social media, pp. 30–38. Association for Computational Linguistics (2011) 2
2. Almehmadi, A., Bourque, M., El-Khatib, K.: A tweet of the mind: automated emotion detection for social media using brain wave pattern analysis. In: International Conference on Social Computing, pp. 987–991 (2013) 4
3. Amarasinghe, K., Sivils, P., Manic, M.: Eeg feature selection for thought driven robots using evolutionary algorithms. In: 9th International Conference on Human System Interactions, pp. 355–361 (2016) 2
4. Becker, K., Moreira, V.P., dos Santos, A.G.: Multilingual emotion classification using supervised learning: Comparative experiments. Information Processing & Management **53**(3), 684–704 (2017) 4
5. Bird, S., Klein, E., Loper, E.: Natural language processing with Python: analyzing text with the natural language toolkit. " O'Reilly Media, Inc." (2009) 15
6. Blaiech, H., Neji, M., Wali, A., Alimi, A.M.: Emotion recognition by analysis of eeg signals. In: 13th International Conference on Hybrid Intelligent Systems, pp. 312–318 (2013) 3
7. Boldrini, E., Balahur Dobrescu, A., Martínez-Barco, P., Montoyo, A., et al.: Emotiblog: a fine-grained model for emotion detection in non-traditional textual genres (2009) 2
8. Chelba, C., Mikolov, T., Schuster, M., Ge, Q., Brants, T., Koehn, P., Robinson, T.: One billion word benchmark for measuring progress in statistical language modeling. arXiv preprint arXiv:1312.3005 (2013) 15
9. Cherry, K.: The Everything Psychology Book: Explore the human psyche and understand why we do the things we do. Simon and Schuster (2010) 2
10. De Belder, J., Moens, M.F.: Text simplification for children. In: Prroceedings of the SIGIR workshop on accessible search systems, pp. 19–26 (2010) 4
11. Derczynski, L., Ritter, A., Clark, S., Bontcheva, K.: Twitter part-of-speech tagging for all: Overcoming sparse and noisy data. In: RANLP, pp. 198–206 (2013) 15
12. Fattouh, A., Albidewi, I., Baterfi, B.: Eeg-based emotion recognition of quran listeners. In: 3rd International Conference on Computing for Sustainable Global Development, pp. 1338–1342 (2016) 4
13. Furuta, R., Plaisant, C., Shneiderman, B.: Automatically transforming regularly structured linear documents into hypertext. ELECTRON. PUBL. **2**(4), 211–229 (1989) 4
14. Gers, F.A., Schmidhuber, E.: Lstm recurrent networks learn simple context-free and context-sensitive languages. IEEE Transactions on Neural Networks **12**(6), 1333–1340 (2001) 2
15. Graves, A., Schmidhuber, J.: Framewise phoneme classification with bidirectional lstm and other neural network architectures. Neural Networks **18**(5), 602–610 (2005) 2
16. Graves, A., Schmidhuber, J.: Offline handwriting recognition with multidimensional recurrent neural networks. In: Advances in neural information processing systems, pp. 545–552 (2009) 2
17. Higuchi, T.: Approach to an irregular time series on the basis of the fractal theory. Physica D: Nonlinear Phenomena **31**(2), 277–283 (1988) 3
18. Hochreiter, S., Heusel, M., Obermayer, K.: Fast model-based protein homology detection without alignment. Bioinformatics **23**(14), 1728–1736 (2007) 2

19. Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N., Wu, Y.: Exploring the limits of language modeling. arXiv preprint arXiv:1602.02410 (2016) 13, 15
20. Kaur, B., Singh, D., Roy, P.P.: A novel framework of eeg-based user identification by analyzing music-listening behavior. Multimedia Tools and Applications pp. 1–22 (2016) 2
21. Kim, K.H., Bang, S.W., Kim, S.R.: Emotion recognition system using short-term monitoring of physiological signals. Medical and biological engineering and computing **42**(3), 419–427 (2004) 2
22. Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing **3**(1), 18–31 (2012) 4
23. Kumar, P., Roy, P.P., Dogra, D.P.: Independent bayesian classifier combination based sign language recognition using facial expression. Information Sciences **428**, 30–48 (2018) 2
24. Kumar, P., Saini, R., Roy, P.P., Dogra, D.P.: A bio-signal based framework to secure mobile devices. Journal of Network and Computer Applications **89**, 62–71 (2017) 2
25. Li, M., Lu, B.L.: Emotion classification based on gamma-band eeg. In: International Conference of the Engineering in Medicine and Biology Society, pp. 1223–1226 (2009) 3
26. Liaw, A., Wiener, M., et al.: Classification and regression by randomforest. R news **2**(3), 18–22 (2002) 10
27. Lin, Y.P., Wang, C.H., Jung, T.P., Wu, T.L., Jeng, S.K., Duann, J.R., Chen, J.H.: Eeg-based emotion recognition in music listening. IEEE Transactions on Biomedical Engineering **57**(7), 1798–1806 (2010) 3
28. Liu, F., Weng, F., Wang, B., Liu, Y.: Insertion, deletion, or substitution?: normalizing text messages without pre-categorization nor supervision. In: 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2, pp. 71–76. Association for Computational Linguistics (2011) 4
29. Liu, Y., Sourina, O., Nguyen, M.K.: Real-time eeg-based human emotion recognition and visualization. In: International Conference on Cyberworlds, pp. 262–269. IEEE (2010) 2
30. Liu, Y., Sourina, O., Nguyen, M.K.: Real-time eeg-based emotion recognition and its applications. In: Transactions on computational science XII, pp. 256–277. Springer (2011) 4
31. Matlovic, T., Gaspar, P., Moro, R., Simko, J., Bielikova, M.: Emotions detection using facial expressions recognition and eeg. In: 11th International Workshop on Semantic and Social Media Adaptation and Personalization, pp. 18–23 (2016) 4
32. Mohammad, S.M., Kiritchenko, S.: Using hashtags to capture fine emotion categories from tweets. Computational Intelligence **31**(2), 301–326 (2015) 11
33. Mohammad, S.M., Turney, P.D.: Crowdsourcing a word–emotion association lexicon. Computational Intelligence **29**(3), 436–465 (2013) 12
34. Parapar, J., Bellogín, A., Castells, P., Barreiro, Á.: Relevance-based language modelling for recommender systems. Information Processing & Management **49**(4), 966–980 (2013) 4
35. Petrantonakis, P.C., Hadjileontiadis, L.J.: Emotion recognition from eeg using higher order crossings. IEEE Transactions on Information Technology in Biomedicine **14**(2), 186–197 (2010) 4
36. Roy, R.S., Agarwal, S., Ganguly, N., Choudhury, M.: Syntactic complexity of web search queries through the lenses of language models, networks and users. Information processing & management **52**(5), 923–948 (2016) 5
37. Roy, R.S., Padmakumar, A., Jeganathan, G.P., Kumaraguru, P.: Automated linguistic personalization of targeted marketing messages mining user-generated text on social media. In: International Conference on Intelligent Text Processing and Computational Linguistics, pp. 203–224. Springer (2015) 4
38. Saini, R., Kaur, B., Singh, P., Kumar, P., Roy, P.P., Raman, B., Singh, D.: Dont just sign use brain too: A novel multimodal approach for user identification and verification. Information Sciences (2017) 4
39. Savitzky, A., Golay, M.J.: Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry **36**(8), 1627–1639 (1964) 7
40. Schmidhuber, J., Wierstra, D., Gomez, F.J.: Evolino: Hybrid neuroevolution/optimal linear search for sequence prediction. In: 19th International Joint Conferenceon Artificial Intelligence (2005) 2
41. Schuller, B., Reiter, S., Muller, R., Al-Hames, M., Lang, M., Rigoll, G.: Speaker independent speech emotion recognition by ensemble classification. In: International Conference on Multimedia and Expo, pp. 864–867. IEEE (2005) 2

42. Shaver, P., Schwartz, J., Kirson, D., O'connor, C.: Emotion knowledge: further exploration of a prototype approach. Journal of personality and social psychology **52**(6), 1061 (1987) 14

43. Smedt, T.D., Daelemans, W.: Pattern for python. Journal of Machine Learning Research **13**(Jun), 2063–2067 (2012) 15

44. Soleymani, M., Pantic, M., Pun, T.: Multimodal emotion recognition in response to videos. IEEE transactions on affective computing **3**(2), 211–223 (2012) 3

45. Teplan, M., et al.: Fundamentals of eeg measurement. Measurement science review **2**(2), 1–11 (2002) 2

46. Thuy, P.T.T., Lee, Y.K., Lee, S.: Dtd2owl: automatic transforming xml documents into owl ontology. In: 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, pp. 125–131 (2009) 4

47. Toutanova, K., Manning, C.D.: Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In: Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics, pp. 63–70 (2000) 15

48. Tsou, M.H.: Research challenges and opportunities in mapping social media and big data. Cartography and Geographic Information Science **42**(sup1), 70–74 (2015) 2

49. Vogt, T., André, E., Bee, N.: Emovoicea framework for online recognition of emotions from voice. Perception in multimodal dialogue systems pp. 188–199 (2008) 2

50. Wang, S., Gwizdka, J., Chaovalitwongse, W.A.: Using wireless eeg signals to assess memory workload in the n-back task. IEEE Transactions on Human-Machine Systems **46**(3), 424–435 (2016) 4

51. Wang, W., Chen, L., Thirunarayan, K., Sheth, A.P.: Harnessing twitter" big data" for automatic emotion identification. In: International Conference on Privacy, Security, Risk and Trust and International Confernece on Social Computing (SocialCom), pp. 587–592 (2012) 2, 14

52. Wu, X., Zhu, X., Wu, G.Q., Ding, W.: Data mining with big data. IEEE transactions on knowledge and data engineering **26**(1), 97–107 (2014) 2

53. Yang, C., Lin, K.H.Y., Chen, H.H.: Building emotion lexicon from weblog corpora. In: 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions, pp. 133–136 (2007) 2