# Recognition of Emotion from Emoticon with Text in Microblog Using LSTM

Juyana Islam[1], M. A. H. Akhand[1,*], Md. Ahsan Habib[1], Md Abdus Samad Kamal[2], Nazmul Siddique[3]

[1]*Dept. of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna 9203, Bangladesh*

[2]*Graduate School of Science and Technology, Gunma University, Kiryu 376-8515, Japan*

[3]*School of Computing, Engineering and Intelligent Systems, Ulster University, Northern Ireland, UK*

| ARTICLE INFO | ABSTRACT |
|---|---|
| | *With the advent of internet technology and social media, patterns of social communication in daily lives have changed whereby people use different social networking platforms. Microblog is a new platform for sharing opinions by means of emblematic expressions, which has become a resource for research on emotion analysis. Recognition of emotion from microblogs (REM) is an emerging research area in machine learning as the graphical emotional icons, known as emoticons, are becoming widespread with texts in microblogs. Studies hitherto have ignored emoticons for REM, which led to the current study where emoticons are translated into relevant emotional words and a REM method is proposed preserving the semantic relationship between texts and emoticons. The recognition is implemented using a Long-Short-Term Memory (LSTM) for the classification of emotions. The proposed REM method is verified on Twitter data and the recognition performances are compared with existing methods. The higher recognition accuracy unveils the potential of the emoticon-based REM for Microblogs applications.* |

## 1. Introduction

Expressions of emotion are fundamental features of intelligent beings, especially humans, that play important roles in social communication [1, 2]. In simple words, emotion represents a person's state of mind exposing into an expression such as happiness, sadness, anger, disgust, and fear. Humans express their emotions in verbal and nonverbal modes, such as speech [3], facial expression [4], body language [5], and expression using text [6]. Emotion has a strong correlation with mental health measured by positive and negative affect and plays a vital role in human social and personal life. With the advent of internet technology, people use various social networking platforms (e.g., Facebook, Twitter, Whatsapp) for social communication. People share their thoughts, feelings, and emotions on different socio-economic, politico-cultural issues using social media. Social media contents appear to emerge as a potential resource for research in human emotional and social behaviors.

Microblog is a common platform to share opinions; hence, it is an important source of emotion analysis of individuals. Microblog makes communication more convenient in our daily life. The most popular microblogs include Twitter [7], Facebook, Instagram,

LinkedIn, and Tumblr. A microblog-post can be reached a vast number of audiences within a short time through these platforms. Moreover, a post may reflect one's emotion or sentiment. Thus, sentiment analysis and emotion recognition are two critical tasks from microblog data [8–13]. Depression is the world's fourth major disease, which is deeply related to emotions [14] and often leads to suicidal tendencies. In the United States, suicide is the 10th major cause of death [15]. Social communications and microblog messages may reflect one's mental state. Thus, a potential application of microblog analysis is to take quick necessary actions against deeply depressed people (who might commit suicide) based on his microblog comments.

Analysis of social media contents, especially microblogs, has become very important in different prospects in the present internet era. Tracking and analyzing social media contents are advantageous for understanding public sentiment on any current socio, cultural, and political issue. Researchers have explored different techniques of extracting information from social media data, which have a direct impact on customer services, market research, public issues, and politics. Customer review analysis through such techniques plays an important role in improving the quality of the products and services for retaining customers and attract more [16]. The developed techniques are expected to play

*Corresponding Author: M. A. H. Akhand, Email: akhand@cse.kuet.ac.bd, and Tel.: +8801714087205

an essential role in the study of patients' psychology. Furthermore, emotion analysis is being considered as an emerging research domain for the assessment of mass opinion [17].

Automatic recognition of emotion from microblogs (REM) is a challenging task in the machine learning and computational intelligence domain. There are two main approaches to recognize emotions from microblogs: the knowledge-based approach and the machine learning (ML) approach [18]. In the knowledge-based approach, the task is to develop a set of rules analyzing the given data and then detect emotion using the rules [19]. In the ML approach, a dataset, which consists of patterns based on the features generated from the microblog data, is used to train an ML model, and then the model is used to predict the emotion for unseen data [20]. Typically, ML-based approaches are expected to perform better than knowledge-based approaches [29], [31]. Recently, deep learning (DL)-based methods, which work on the preprocessed data and do not require explicit features, are investigated for REM and found to be promising results [21], [22].

The existing REM methods ignored emoticons and other signs or symbols in the microblogs. These researches only considered texts for the recognition of emotion [23–25]. Nowadays, emoticons, the pictorial representations of facial expressions using characters and related symbols, are commonly used on social media sites. It is found that emoticons are becoming the most important features of online textual languages [26]. Among the few studies that dealt with emoticons is [22] using Convolutional Neural Network (CNN). In the study, words and emoticons from microblogs are processed separately in two different vectors and projected into the emotional space to classify using CNN. Emoticon consideration independent of the text seems not appropriate as emoticons embedded within the text fabricates a semantic or contextual meaning, which is important in emotion analysis. Placement of the emoticon within the text is also important as the different arrangement of emoticon within text may change the meaning. However, emoticon-based REM development is the motivation behind the present study.

This study aims to develop an improved REM method to keep the semantic links between emoticons and the relevant texts. Acknowledging emoticons as particular expressions of emotions, they are represented by suitable emotional words. The original sequence of emoticons in the microblog is unchanged since their sequence may have a vital role in expressing the appropriate emotion. With the necessary prepossesing of microblog data, a machine learning model suitable for examining the sequential or time-series information, known as the Long Short-Term Memory (LSTM), is employed to classify emotions. The recognition performances are compared with the existing method that uses only text expressions (i.e., ignores emoticons) in the recognition process. An initial version of the LSTM-based REM considering emoticons has been presented in a conference [1]; and, the present study is an extended version. The current REM presents the detailed theoretical analysis and experimental results. The higher recognition accuracy of the proposed REM justifies its use in emerging microblog applications.

The rest of the paper is organized as follows. Section 2 presents a brief survey of existing REM methods. Section 3 explains the proposed REM method. Section 4 provides detailed experimental results and analysis. Finally, the conclusion is presented in Section 5.

## 2. Related Works

Microblog analysis for REM is explored with the rapid growth of social media communication. Several studies were conducted in the last decade for REM from microblogs employing different ML methods, including Naive Bayes (NB) and Support Vector Machine (SVM). The DL-based techniques have also emerged remarkably in the recent REM studies.

Pre-processing of blog tests and distinguishable feature extraction with appropriate techniques are the two important tasks to apply any ML method for REM. Chaffar and Inkpen [18] extracted features from diary-like blog posts (called Aman's Dataset) using bags of words and N-grams. They used decision trees, NB, and SVM to recognize the six fundamental emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) using the features. The SVM is found best among the other classifiers. Silva and Haddela [27] also used Aman's data set and applied the SVM for REM purposes. But they investigated a concept called term weighting to enhance the conventional Term Frequency Inverse Document Frequency (TF-IDF) for feature extraction. Chirawichitchai [28] studied a feature selection technique by information gain and REM by SVM on Thai language blog texts from various social networking sites (e.g., Facebook).

In [29], the authors examined semi-supervised learning with SVM, called distant supervision, for REM from the Chinese tweets in Weibo using a large corpus with 1,027,853 Weibo statuses with emotion labels. Their proposed system predicted happiness emotion most accurately (90% accuracy rate) and worked well for anger. However, the system was less effective for detecting other emotions, e.g., fear, sadness, disgust, and surprise.

In [30], the authors used emoticons in their proposed REM method called the emoticon spaced model (ESM). The ESM learns a sentiment representation of words with the help of emoticons using a heuristic. Words with similar sentiments have similar coordinates in the emoticon space. The coordinates of words are fed into Multinomial naive Bayes (MNB) and SVM for classification. They applied their method on the Chinese microblog benchmark corpus NLP&CC2013 with 14,000 posts with the four most common emotion types (happiness, like, sadness, and disgust).

In [31], the authors performed REM from Twitter's data using NB; in preprocessing stage, they removed URL, special characters, stop-words, and few other things. In [32], the authors extracted features using different methods (e.g., Unigram, Bigram) on the collected 1200 Twitter emotional data and classified emotions using MNB. The large number of features combining Unigram and Bigram is shown to outperformed others with an accuracy of 95.3%.

In [14], the authors adapted emotional-related Chinese microblog (Sina Weibo) data for depression recognition adding "depression" as a new class and excluding the "surprise" class. They developed an emotion feature dictionary with seven types of emotions, namely depression, good, happiness, fear, sadness, disgust, and anger, for depression recognition using 1381 emotional words or phrases. In their study, Multi-kernel SVM is found better than KNN, NB, and standard SVM for depression

recognition from the combination of features from the user profile and user behavior and the features from blog texts.

Among different DL methods for REM, CNN and LSTM are the most well-known ones found in prominent studies recently. In [22], the authors proposed a CNN-based REM, called enhanced CNN (ECNN), that examines both texts and emoticons. Specifically, by placing the emoticons and words in two different vectors and projecting them into one emotional space, CNN is employed to classify emotion. They viewed emoticons as independent of the text, i.e., ignored the emoticon's order in the description. Such consideration might be misleading because emoticon placement or sequence in the text may have a specific meaning. ECNN applied on the Chinese Sina Weibo, NLPCC2013, and Twitter datasets (SEMEVAL). The experimental results on Chinese Sina Weibo, NLPCC2013, and Twitter microblog datasets showed that ECNN outperformed other methods, including SVM, bidirectional LSTM (BiLSTM).

On the other hand, in [21], the authors proposed a hybrid DL model, called Semantic-Emotion Neural Network (SENN), with BiLSTM and CNN for REM. BiLSTM is used to capture contextual information and focuses on the semantic relationship, and CNN is used to extract emotional features and focuses on the emotional connection between words. SENN was applied on Twitter and other social media data, but the use of emoticon is not clear in the decision process.

## 3. Recognition of Emotion from Microblog (REM) Managing Emoticon with Text

Recently, social media has become a dominant and popular platform for expressing and sharing emotion [3] using microblogs, photos, and videos. Remarkably, the microblog is the hot favorite choice, where one directly writes personal thoughts (e.g., own status, reactions to others, and opinions). Facebook and Twitter are examples of the most popular social media for expressing and communicating such personal thoughts in microblogs. Microblogs contain words, emoticons, hashtags, and various signs with distinct meanings. Since emoticons have become more popular elements besides the text than ever, they should be given proper attention in any microblog-based scheme of emotion recognition.

In this study, emphasis is given to emoticons and their association with texts in microblogs, considering that both are equally valuable to identify proper emotion. Some studies excluded emoticons in the preprocessing step considering those as noisy inputs [33]. But, in this study, emoticons are altered with emotional words and fused with texts for emotion recognition.

---

**Algorithm 1:** Recognition of emotion from microblogs (REM)

**Input**: Microblog $M$ with $W$ Words
**Output**: Emotion Category $EC \in$ (Happy, Sad, Angry, Love)

// Process 1: Emoticon alteration with textual meaning
**For** all $w \in W$ **do**
    **If** $M[w]$ is emoticon **then**
        $M[w] = Emoticon.meaning (M[w])$
    **End If**
**End For**

//Process 2: Tokenization with integer encoding
**For** all $w \in W$ **do**
    $IW[w] = Tokenizer (M[w])$
**End For**

// Process 3: Padding with 0 to render fixed $S$ size
**For** all $w \in [S - W]$ **do**
    $P[w] = 0$    // Consider 0 padding
**End For**
**For** all $w \in [S - W + 1 : S]$ **do**
    $P[w] = IW[w]$ // Copy the rest values
**End For**

// Process 4: Emotion recognition using LSTM
//Embedding integer to 2D vector
**For** all $w \in S$ **do**
    $V[x, y] = Embedding (P[w])$
**End For**
$EC = LSTM (V)$

---

These interpreted emotional words and other texts presented in the proposed REM help the model perform improved emotion classification.

Figure 1 illustrates the framework of REM proposed in this study for a sample microblog containing an emoticon in the text. The REM consists of four sequential processes. In Process 1, emoticons are converted into relevant emotional words according to a predefined lookup table. The words are transformed into a sequence of integer numbers in Process 2. In Process 3, padding is conducted to form a vector containing the sequence of words with equal length. Finally, in Process 4, the LSTM is employed for classification of emotions into Happy, Sad, Angry, or Love.

Algorithm 1 shows the proposed REM where individual processes are marked. It takes microblog $M$ with $W$ words as input and provide emotion category $EC$. The whole method is broadly divided into two major parts: processing microblog data using processes 1, 2, and 3, and recognition with the LSTM network. The processes are briefly described in the following subsections.
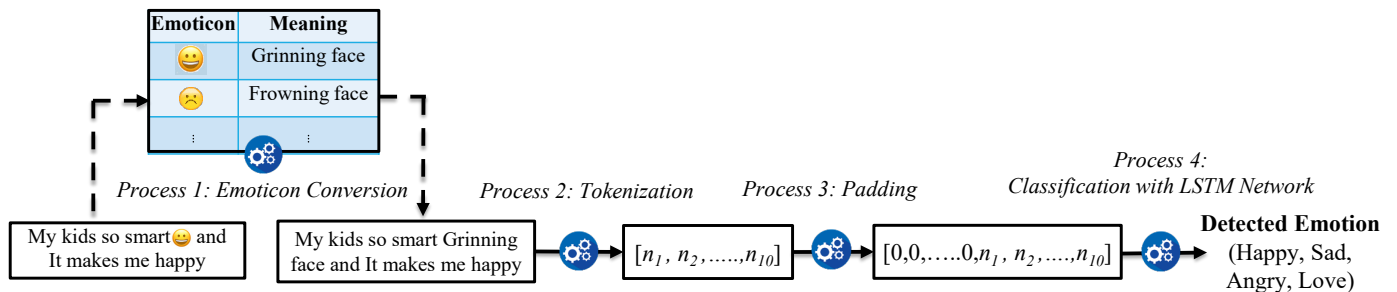


Figure 1: The proposed REM framework illustrating different processes for a sample microblog.
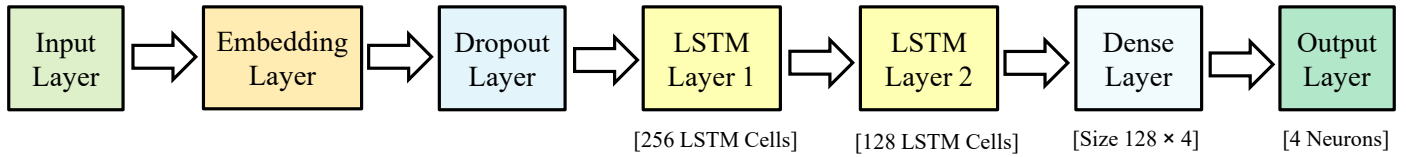
| Input Layer | Embedding Layer | Dropout Layer | LSTM Layer 1 | LSTM Layer 2 | Dense Layer | Output Layer |
|---|---|---|---|---|---|---|
| | | | [256 LSTM Cells] | [128 LSTM Cells] | [Size 128 × 4] | [4 Neurons] |

Figure 2: LSTM network architecture used in the proposed REM.

### 3.1. Microblogs Processing

Twitter is a popular microblog platform and it allows emoticons with texts. Thus, Twitter microblogs are collected and processed for REM in this study. Social media data contains noisy information that needs to be cleaned up to use in the system. Then the clean microblogs with emoticons and texts go through the Processes 1 to 3 (Fig. 1). In Process 1, each emoticon is replaced with corresponding text (i.e., equivalent word for the emoticon) using a function, called *Emoticon.meaning()*, with the help of a lookup table with equivalent words and emoticons. Process 2 is the Tokenization step: it removes unnecessary information and then, generates an integer vector sequence of words (*IW)* through integer encoding. Finally, Process 3 transforms *IW* to a defined fixed length size (say *S*) of words with zero initial paddings. If *IW* contains *W* integer values, the padding outcome *P* vector will contain zeros (i.e., 0) in initial *S-L* positions and the rest are *L* values from *IW*.

### 3.2. Emotion Recognition Using LSTM

The proposed structure of the LSTM network for the REM is shown in Figure 2. The network consists of an input layer, an embedding layer, a dropout layer, two LSTM layers, a dense layer, and finally, the output layer. The input in the LSTM network comprises a sequence of an integer number (defined fixed length *S*) with zero paddings in initial positions. The embedding layer simply transforms each integer word into a particular embedding vector. In the proposed architecture, the sizes of input integer words and embedding vectors are 78 and 128, respectively. Therefore, the output of the embedding for a microblog text is a 78×128 sized 2D vector. A dropout layer is placed just after the input layer, which randomly selects input features during training. The purpose of the dropout layer is to reduce overfitting and improve the generalization of the system.

There are two LSTM layers in the proposed architecture which are the main functional elements of the system. The first and second LSTM layers contain 256 and 128 hidden LSTM cells, respectively. Each LSTM cell in the first layer processes 128 sized embedding vectors and generates single output; therefore, the first LSTM layer produces 256 values which propagate to the input of each LSTM cell of the second layer. The second LSTM layer produces 128 values and the dense layer generates the emotional response from the values. Emotion recognition of this study is a multiclass (i.e., 4-class) classification problem to classify microblogs into four different emotion categories. Thus, it requires the dense layer to be of size 128×4. The output layer has to yield one of the four classes and therefore, the output layer comprises four neurons where each neuron represents a particular emotional state.

An LSTM cell is the heart of the LSTM network architecture illustrated in Fig. 3, which shows the basic building block of an
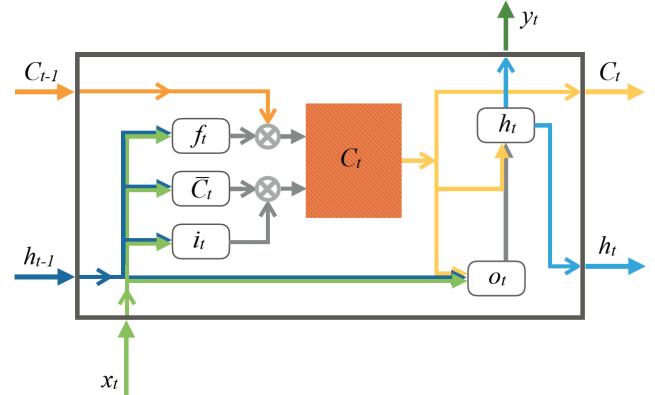


Figure 3: Basic building block of an LSTM cell.

LSTM cell. The LSTM cell consists of a forget gate (*f*), a memory cell (*C*), an input gate (*i*), and an output gate (*o*), At any state *t*, the memory block uses both the current input ($x_t$) and the previous hidden layer output ($h_{t-1}$) as inputs and generates new output ($h_t$) of the hidden layer. This memory block enables the LSTM network in forgetting and memorizing information as required. Hyperbolic tangent or *tanh (*symbol $\phi$) and *sigmoidal* (symbol $\sigma$) functions are used as the gates. The memory unit calculates the candidate memory $\overline{C_t}$, $f_t$ and input gate $i_t$ at state *t* according to Eqs. (1-3).

$$\overline{C_t} = \phi \ (W_C * h_{t-1} + U_C *x_t + b_C \ ) \tag{1}$$

$$f_t = \sigma \ (W_f * h_{t-1} + U_f *x_t + b_f \ ) \tag{2}$$

$$i_t = \sigma \ (W_i * h_{t-1} + U_i *x_t + b_i \ ) \tag{3}$$

Then, memory $C_t$ is calculated using Eq. (4).

$$C_t = f_t * C_{t-1} + i_t + \overline{C_t} \tag{4}$$

$h_t$ comes from $C_t$ through the output gate $o_t$ following Eqs. (5-6).

$$o_t = \sigma \ (W_o * h_{t-1} + U_o *x_t + b_o \ ) \tag{5}$$

$$h_t = o_i * \phi(C_t) \tag{6}$$

In Eqs. (5-6), *W* and *U* denote the respective shared weights, and *b* denotes the bias vector. Finally, the output of an LSTM cell comes from $h_t$ through the weight vector *V* defined as

$$y_t = V * h_t . \tag{7}$$

The LSTM is suitable for modeling complex time-series data since it can classify from any given sequence upon training. Therefore, the LSTM is chosen in the proposed REM to classify the processed microblogs. A detailed description of the LSTM is available in [34].

## 4. Experimental Studies

This section describes Twitter data preparation, experimental settings and experimental results of this study.

### 4.1. Dataset Preparation

To make the processed microblog data compatible with LSTM network, tokenization is performed using Keras open-source neural-network library [36] to convert the words to numerical values, and then padded with zeros for a fixed-sized vector. The

Table 1: Emoticons and corresponding word meanings

| Sl. | Emoticon | Unicode | Textual Meaning | Emotion |
|---|---|---|---|---|
| 1 | 😀 | U+1F600 | Grinning face | Happy |
| 2 | 😄 | U+1F604 | Grinning face with smiling eyes | |
| 3 | 😁 | U+1F601 | Beaming face with smiling eyes | |
| 4 | ☺ | U+263A | Smiling face | |
| 5 | 😭 | U+1F62D | Loudly crying face | Sad |
| 6 | 😢 | U+1F622 | Crying face | |
| 7 | 🥺 | U+1F97A | Pleading face | |
| 8 | ☹ | U+2639 | Frowning face | |
| 9 | 😠 | U+1F620 | Angry face | Angry |
| 10 | 😡 | U+1F621 | Pouting face | |
| 11 | 😤 | U+1F624 | Face with steam from nose | |
| 12 | 🤬 | U+1F62C | Face with symbols on mouth | |
| 13 | 😍 | U+1F60D | Smiling face with heart-eyes | Love |
| 14 | 🥰 | U+1F970 | Smiling face with hearts | |
| 15 | 😘 | U+1F618 | Face blowing a kiss | |
| 16 | 😗 | U+1F617 | Kissing face with closed eyes | |

Table 2: Samples Twitter microblogs and emotion calabel

| Microblog | Emotion Category |
|---|---|
| Yay! Very well deserved 😄 | Happy |
| I'm so sorry 😭 | Sad |
| 😭 I just want the pain to go away | Sad |
| Apologies to him right now! 😠 | Angry |
| 😡 go follow right now | Angry |
| 😍😍😍 Good night love | Love |

size of the padded numeral blog was 78, whereas, it was 33 while emoticons were discarded.

### 4.2. Experimental Settings

Adam algorithm [37], a popular optimization algorithm in computer vision and natural language processing applications, is used to train LSTM. Softmax and categorical-cross entropy are considered as activation function and loss function, respectively. The dropout rate of the dropout layer is set to 0.3, and each LSTM layer contains 30% dropout and 20% recurrent dropout while training the model. Batch-wise training is common nowadays and LSTM training was performed for batch sizes 32, 64, and 128 which are commonly used in related studies. Among the collected 16012 tweets, 75% (i.e., 12009) were used to train LSTM, and the remaining 25% (i.e., 4003) were reserved for the test set to check the generalization ability of the system.

The proposed model was implemented in Python programming language. The experiments were performed using a jupyter notebook as well as Kaggle online environment. A PC with the following configuration is used for conducting the experiment model: HP ProBook 450 G4, processor: Intel(R) Core (TM) i5-5200U, CPU: 2.20 GHz, RAM: 8 GB, OS: Windows 10.

### 4.3. Experimental Results and Analysis

The emoticon consideration with text is the core significance of the proposed REM (with emoticons embedded in texts) from real-life Tweeter data. An experiment discarding emoticons (i.e., using texts only) is also carried out, it may be called REM without emoticon or text-only REM. The outcomes of text-only REM are compared with the proposed REM to observe the effect of emoticon.

In Figure 4, the accuracies of the LSTM for both the training and test sets are evaluated by varying the training epochs up to 200 for the different batch sizes (BSs). It is very clear from the graph that the accuracy of the proposed REM (as shown by the solid curves) is always better than the accuracy of the text-only REM (as shown by the dashed curve). Remarkably, the accuracy with the text-only case is compatible with the proposed REM while training. For example, at 100 epochs for BS=32 (in Fig. 4(a)), the achieved accuracies on the training set of the proposed REM and text-only REM are 0.994 and 0.957, respectively. However, regarding the test data, the accuracy of the proposed REM is much better than that of the method without emoticon (i.e., text-only REM). It is remarkable that the text-only REM test set accuracy is placed in a graph doubling its achieved value to make the graph better visualization. At a glance, the test set accuracy of the proposed REM is almost double that of the text-only REM with any BS values. As an example, at 100 epochs for BS=64 (in Fig. 4(b)), the achieved test set accuracy for proposed REM is 0.873; whereas, the value is only 0.424 for text-only REM. A similar observation is also visible for BS=128 in Fig. 4(c).
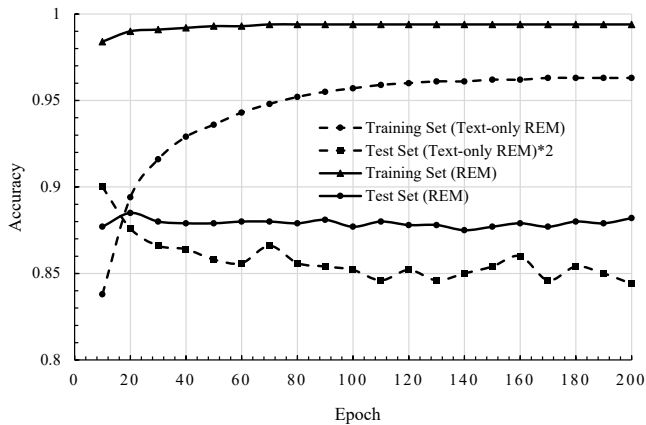
Figure 5 shows the best accuracies from individual experiments with different BS values. The training set accuracy for the text-only REM and the proposed REM are 0.963 and 0.994 for any BS value, respectively. Although performance on the
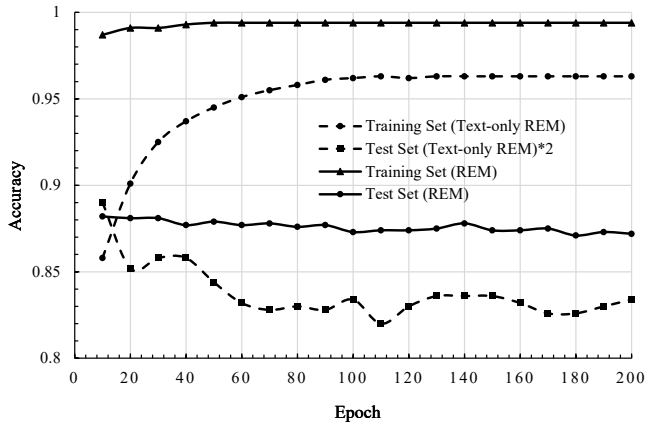
(a) BS= 32
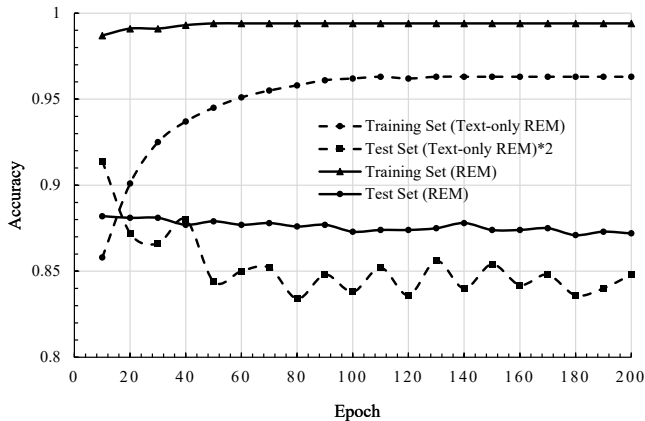


(b) BS= 64



(c) BS= 128

Figure 4: Recognition accuracies of proposed REM (using both text and emoticon) and text-only REM varying training epochs for three different batch sizes (BSs) 32, 64 and 128.

training set is competitive with and without emoticon, it is remarkable on test set accuracy with emoticons. The best test set accuracies for the text-only REM are 0.445, 0.457, and 0.45 at BS values 32, 64, and 128, respectively. On the other hand, considering both emoticon and text, the proposed REM achieved much better test set accuracies and the values are 0.884, 0.882, and 0.885 at BS values 32, 64, and 128, respectively. Notably, training set accuracy indicates the memorization ability and test set accuracy indicates the generalization ability of a system to
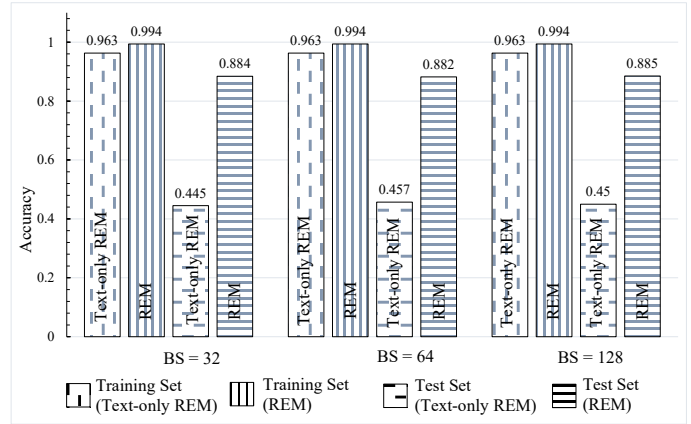
Figure 5: Best accuracies for training and test sets for proposed REM (text and emoticon) and text-only REM for training batch sizes (BSs) 32, 64 and 128.

work on unseen data. Test set accuracy is the key performance indicator for any machine learning system and it is a better score of proposed REM over text-only REM (i.e., without emoticon), which revealed that the use of emoticons enhances the ability of the proposed method in learning the emotion properly. However, the reason behind the worse performance with text-only REM is that the texts are limited in the selected tweet data and, in many cases, the text becomes meaningless without emotion, which has been explained in the data preparation section. Moreover, people do not care about meaning with text only when they use emoticon within it.

Table 3 and Table 4 show the emotion category-wise performance matrices of the proposed REM and text-only REM, respectively, for the best test accuracy cases shown in Fig. 4. The table shows the variation in actual and predicted emotions labeled for the individual emotion category. In the test set, 'Happy', 'Sad', 'Angry' and 'Love' emotion categories hold 1024, 956, 949, and 1074 tweet data consecutively. For the 'Happy' case in Table 3, for example, the proposed REM truly classified 914 cases, and the remaining 110 cases were misclassified into 'Sad', 'Angry' and 'Love' categories as 41, 33, and 36 cases, respectively. On the other hand, text-only REM truly classified only 442 cases as from Table 4. The proposed REM showed the best performance for the 'Angry' category by truly classifying 862 cases out of 949 cases

Table 3: Emotion category wise test set performance matrix for proposed REM method considering both emoticon and text.

| Emotion Category | Total Sample | Detected Category | | | |
|---|---|---|---|---|---|
| | | *Happy* | *Sad* | *Angry* | *Love* |
| Happy | 1024 | **914** | 41 | 33 | 36 |
| Sad | 956 | 21 | **857** | 42 | 36 |
| Angry | 949 | 29 | 31 | **862** | 27 |
| Love | 1074 | 43 | 68 | 53 | **910** |

Table 4. Emotion category wise test set performance matrix for text-only REM.

| Emotion Category | Total Sample | Detected Category | | | |
|---|---|---|---|---|---|
| | | *Happy* | *Sad* | *Angry* | *Love* |
| Happy | 1024 | **442** | 175 | 173 | 234 |
| Sad | 956 | 177 | **419** | 194 | 166 |
| Angry | 949 | 195 | 209 | **434** | 111 |
| Love | 1074 | 205 | 190 | 145 | **534** |

(i.e., 90.83% accuracy). On the other hand, the text-only method showed the best performance for the 'Love' category, truly classifying 534 cases out of 1074 cases (i.e., 49.72% accuracy). The performance comparison in the individual emotion category visualizes the proficiency of the proposed REM with emoticon and text.

On Twitter data, Table 5 compares the classification accuracy of the proposed method with other existing methods. The table also includes the methods used by various studies with a variety of dataset sizes. The existing methods considered Naive Bayes, CNN, and BiLSTM. The self-processed 16012 Twitter data used in the present study. The dataset used in [21] is larger than this study, but the authors did not mention how training and test sets are partitioned. Due to varying dataset sizes, the comparison with other methods may not be completely fair. However, the proposed method has outperformed any other methods showing a test set accuracy of 88.5%. The achieved accuracy is much better than the traditional machine learning with Naive Bayes [31] and deep learning methods with CNN and BiLSTM [21] [22]. It is already mentioned that study in [22] used emoticons but processed by separating them from the text. The main reason behind the outperforming ability of the proposed method is its emoticon management with text which is not appropriately handled in the existing methods. Finally, managing emoticons and texts simultaneously and classification with LSTM have been revealed as a promising emotion recognition method for microblogs.

Table 5: Performance comparison with existing methods on Twitter data.

| Sl. | Author [Work Ref], Year | Sample Size (Training + Test) | Method Used | Test Set Accuracy |
|---|---|---|---|---|
| 1 | Wikarsa and Thahir [31], 2015 | 268 (116+152) | Naive Bayes | 71.3% |
| 2 | Yang et al. [22], 2018 | 7200 (4600+2600) | CNN | 72.6% |
| 3 | Batbaatar et al. [21], 2019 | 19678 (Not Mentioned) | BiLSTM + CNN | 61.3% |
| 4 | Proposed Method | 16012 (12009 + 4003) | LSTM | 88.5% |

## 5. Conclusions

Nowadays, people are very active on social media and frequently express their emotions using both texts and emoticons in microblogs. Emotion recognition from social media microblogs (i.e., REM) emerges as a promising and challenging research issue. It is essential to consider all necessary microblog information for comprehensive REM. Unlike many existing methods that only view the textual expressions for simplicity, this study has investigated REM utilizing both emoticons and texts simultaneously. Using the underlying LSTM technique, the proposed REM could interpret the emoticons in the context of text expressions in Twitter data to precisely classify the user emotions and outperformed the existing methods. The proposed REM method is expected to be an effective tool in emerging emotion

recognition-based applications and play a vital role in social communication.

This study has revealed the proficiency of REM managing emoticons, and at the same time, several research directions are opened from its motivational outcomes and gaps. REM is developed collecting Twitter data for only 16 selected emoticons related to the four emotions (Happy, Sad, Angry, and Love); and system including other emotional states (e.g., Disgust, Surprise) and more emoticons might be interesting. Another thing, the texts were limited in the selected blogs, information degraded due to emoticon removal in text-only REM, and finally, recognition performance with LSTM was poor without emoticons. It might be interesting research to investigate text and emoticon trade-off effects on REM performance. In addition, instead of LSTM, any other deep learning method (e.g., CNN) might also be investigated owing to achieve better classification performance. We wish to work in such directions for developing a more comprehensive REM in the future study.

## References

[1] J. Islam, S. Ahmed, M.A.H. Akhand, N. Siddique, "Improved Emotion Recognition from Microblog Focusing on Both Emoticon and Text," in 2020 IEEE Region 10 Symposium (TENSYMP), IEEE: 778–782, 2020, doi:10.1109/TENSYMP50017.2020.9230725.

[2] L. Al-Shawaf, D. Conroy-Beam, K. Asao, D.M. Buss, "Human Emotions: An Evolutionary Psychological Perspective," Emotion Review, **8**(2), 173–186, 2016, doi:10.1177/1754073914565518.

[3] A.R. Avila, Z. Akhtar, J.F. Santos, D. OShaughnessy, T.H. Falk, "Feature Pooling of Modulation Spectrum Features for Improved Speech Emotion Recognition in the Wild," IEEE Transactions on Affective Computing, **12**(1), 177–188, 2021, doi:10.1109/TAFFC.2018.2858255.

[4] M.A.H. Akhand, S. Roy, N. Siddique, M.A.S. Kamal, T. Shimamura, "Facial Emotion Recognition Using Transfer Learning in the Deep CNN," Electronics, **10**(9), 1036, 2021, doi:10.3390/electronics10091036.

[5] G. Castellano, L. Kessous, G. Caridakis, Emotion Recognition through Multiple Modalities: Face, Body Gesture, Speech, Springer Berlin Heidelberg, Berlin, Heidelberg, Heidelberg: 92–103, 2008, doi:10.1007/978-3-540-85099-1_8.

[6] N. Alswaidan, M.E.B. Menai, "A survey of state-of-the-art approaches for emotion recognition in text," Knowledge and Information Systems, **62**(8), 2937–2987, 2020, doi:10.1007/s10115-020-01449-0.

[7] A. Java, X. Song, T. Finin, B. Tseng, "Why we twitter: understanding microblogging usage and communities," in Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis - WebKDD/SNA-KDD '07, ACM Press, New York, New York, USA: 56–65, 2007, doi:10.1145/1348549.1348556.

[8] A. Ortigosa, J.M. Martín, R.M. Carro, "Sentiment analysis in Facebook and its application to e-learning," Computers in Human Behavior, **31**(1), 527–541, 2014, doi:10.1016/j.chb.2013.05.024.

[9] S. Gangrade, N. Shrivastava, J. Gangrade, "Instagram Sentiment Analysis: Opinion Mining," SSRN Electronic Journal, 2019, doi:10.2139/ssrn.3372757.

[10] M. Nam, E. Lee, J. Shin, "A Method for User Sentiment Classification using Instagram Hashtags," **18**(11), 1391–1399, 2015.

[11] M.Z. Naf'an, A.A. Bimantara, A. Larasati, E.M. Risondang, N.A.S. Nugraha, "Sentiment Analysis of Cyberbullying on Instagram User Comments," Journal of Data Science and Its Applications, **2**(1), 88–98, 2019, doi:10.21108/jdsa.2019.2.20.

[12] L. Huang, S. Li, G. Zhou, Emotion Corpus Construction on Microblog Text, 204–212, 2015, doi:10.1007/978-3-319-27194-1_21.

[13] R. Poonguzhali, V. Waldiya, S. Vinothini, K. Livisha, "Sentiment Analysis on LinkedIn Comments," International Journal of Engineering Research & Technology (IJERT), **6**(07), 2018.

[14] Z. Peng, Q. Hu, J. Dang, "Multi-kernel SVM based depression recognition using social media data," International Journal of Machine Learning and Cybernetics, **10**(1), 43–57, 2019, doi:10.1007/s13042-017-0697-1.

[15] S.R. Braithwaite, C. Giraud-Carrier, J. West, M.D. Barnes, C.L. Hanson, "Validating Machine Learning Algorithms for Twitter Data Against

Established Measures of Suicidality," JMIR Mental Health, **3**(2), e21, 2016, doi:10.2196/mental.4822.

[16] D. Gräbner, M. Zanker, G. Fliedl, M. Fuchs, Classification of Customer Reviews based on Sentiment Analysis, Springer Vienna, Vienna: 460–470, 2012, doi:10.1007/978-3-7091-1142-0_40.

[17] S. Shayaa, N.I. Jaafar, S. Bahri, A. Sulaiman, P. Seuk Wai, Y. Wai Chung, A.Z. Piprani, M.A. Al-Garadi, "Sentiment Analysis of Big Data: Methods, Applications, and Open Challenges," IEEE Access, **6**, 37807–37827, 2018, doi:10.1109/ACCESS.2018.2851311.

[18] S. Chaffar, D. Inkpen, Using a Heterogeneous Dataset for Emotion Analysis in Text, 62–67, 2011, doi:10.1007/978-3-642-21043-3_8.

[19] S. Nirenburg, K. Mahesh, Knowledge-Based Systems for Natural Language Processing., 1997.

[20] S.B. Kotsiantis, I.D. Zaharakis, P.E. Pintelas, "Machine learning: a review of classification and combining techniques," Artificial Intelligence Review, **26**(3), 159–190, 2006, doi:10.1007/s10462-007-9052-3.

[21] E. Batbaatar, M. Li, K.H. Ryu, "Semantic-Emotion Neural Network for Emotion Recognition from Text," IEEE Access, **7**, 111866–111878, 2019, doi:10.1109/ACCESS.2019.2934529.

[22] G. Yang, H. He, Q. Chen, "Emotion-semantic-enhanced neural network," IEEE/ACM Transactions on Audio Speech and Language Processing, **27**(3), 531–543, 2019, doi:10.1109/TASLP.2018.2885775.

[23] N. Colneric, J. Demsar, "Emotion Recognition on Twitter: Comparative Study and Training a Unison Model," IEEE Transactions on Affective Computing, **11**(3), 433–446, 2020, doi:10.1109/TAFFC.2018.2807817.

[24] A. Yousaf, M. Umer, S. Sadiq, S. Ullah, S. Mirjalili, V. Rupapara, M. Nappi, "Emotion Recognition by Textual Tweets Classification Using Voting Classifier (LR-SGD)," IEEE Access, **9**, 6286–6295, 2021, doi:10.1109/ACCESS.2020.3047831.

[25] K. Sailunaz, R. Alhajj, "Emotion and sentiment analysis from Twitter text," Journal of Computational Science, **36**, 101003, 2019, doi:10.1016/j.jocs.2019.05.009.

[26] M.S. Schlichtkrull, "Learning affective projections for emoticons on Twitter," in 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), IEEE: 539–543, 2015, doi:10.1109/CogInfoCom.2015.7390651.

[27] J. De Silva, P.S. Haddela, "A term weighting method for identifying emotions from text content," in 2013 IEEE 8th International Conference on Industrial and Information Systems, IEEE: 381–386, 2013, doi:10.1109/ICIInfS.2013.6732014.

[28] N. Chirawichitchai, "Emotion classification of Thai text based using term weighting and machine learning techniques," in 2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE: 91–96, 2014, doi:10.1109/JCSSE.2014.6841848.

[29] Z. Yuan, M. Purver, "Predicting emotion labels for Chinese microblog texts," Studies in Computational Intelligence, **602**, 129–149, 2015, doi:10.1007/978-3-319-18458-6_7.

[30] F. Jiang, Y.-Q. Liu, H.-B. Luan, J.-S. Sun, X. Zhu, M. Zhang, S.-P. Ma, "Microblog Sentiment Analysis with Emoticon Space Model," Journal of Computer Science and Technology, **30**(5), 1120–1129, 2015, doi:10.1007/s11390-015-1587-1.

[31] L. Wikarsa, S.N. Thahir, "A text mining application of emotion classifications of Twitter's users using Naïve Bayes method," in 2015 1st International Conference on Wireless and Telematics (ICWT), IEEE: 1–6, 2015, doi:10.1109/ICWT.2015.7449218.

[32] J.K. Rout, K.K.R. Choo, A.K. Dash, S. Bakshi, S.K. Jena, K.L. Williams, "A model for sentiment and emotion analysis of unstructured social media text," Electronic Commerce Research, **18**(1), 181–199, 2018, doi:10.1007/s10660-017-9257-8.

[33] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. de Jong, U. Kaymak, "Exploiting emoticons in sentiment analysis," in Proceedings of the 28th Annual ACM Symposium on Applied Computing - SAC '13, ACM Press, New York, New York, USA: 703, 2013, doi:10.1145/2480362.2480498.

[34] S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural Computation, **9**(8), 1735–1780, 1997, doi:10.1162/neco.1997.9.8.1735.

[35] Full Emoji List, Available: https://unicode.org/emoji/charts/full-emoji-list.html, Accessed: March 31, 2021.

[36] Chollet François, Keras: The Python Deep Learning Library, 2015.

[37] D.P. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization," 2014.