International Journal of Computational Intelligence Systems (2023)16:107
https://doi.org/10.1007/s44196-023-00234-5

RESEARCH ARTICLE

On the Effect of Emotion Identification from Limited Translated Text Samples Using Computational Intelligence

Madiha Tahir[1,5] · Zahid Halim[1] · Muhmmad Waqas[3,4] · Shanshan Tu[2]

Received: 10 September 2021 / Accepted: 2 April 2023

Abstract
Emotion identification from text data has recently gained focus of the research community. This has multiple utilities in an assortment of domains. Many times, the original text is written in a different language and the end-user translates it to her native language using online utilities. Therefore, this paper presents a framework to detect emotions on translated text data in four different languages. The source language is English, whereas the four target languages include Chinese, French, German, and Spanish. Computational intelligence (CI) techniques are applied to extract features, dimensionality reduction, and classification of data into five basic classes of emotions. Results show that when English text is translated to French, classification accuracy is higher than others, i.e., 99.04%. Whereas, when the same is translated to Chinese language, its detection rate is lowest among target languages. It is concluded that emotions remain preserved after translation to some extent. Framework consists of TFIDF features. PCA and Discriminant Analysis perform good to detect emotions from translated data.

Keywords Emotion identification · Dynamic · Machine analysis · Translated text · Data mining

Abbreviation

API	Application programming interface
CI	Computational intelligence
DA	Discriminant analysis
fMRI	Functional magnetic resonance imaging
IDF	Inverse document frequency
Lid	Language identification
MMC	Matthew's correlation coefficient
MARS	Multivariate adaptive regression splines
PPV	Positive predictive value
PCA	Principal component analysis
RSA	Royal Spanish Academy
SNS	Social networking service
SVM	Support vector machine
TF	Term frequency
TFIDF	Term frequency-inverse document frequency

✉ Muhmmad Waqas
engr.waqas2079@gmail.com

Madiha Tahir
madiha.tahir@giki.edu.pk

Zahid Halim
zahid.halim@giki.edu.pk

Shanshan Tu
sstu@bjut.edu.cn

1 The Machine Intelligence Research Group (MInG), Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi 23460, Pakistan

2 Beijing University of Technology, Beijing, China

3 School of Engineering, Edith Cowan University, Joondalup, WA 6027, Australia

4 Computer Engineering Department, College of Information Technology, University of Bahrain, Sakheer 32038, Bahrain

5 Institute of Space Technology, Islamabad 45730, Pakistan

1 Introduction

Emotion is a biological state linked with the human nervous system. It brings changes associated with one's feelings, thoughts, and attitude. Emotions play important role in the daily life of human beings. An individual, directly or indirectly, expresses her emotions in almost every activity, e.g., while listening to the music, handing over an item to another person or while chatting with friends and relatives. Generally, there are two aspects of human emotions, the first one is that there are no human experiences that are free of human emotion, and the second one is that all our memories

Published online: 23 June 2023

of the past events preserve emotions that are experienced when those memories get recalled [23]. According to the knowledge of psychology, emotion is a complex state of feelings that results in physical and psychological changes that influence thoughts and behavior. Emotions can be expressed through verbal communication, gestures, actions, and written expressions. In addition, to the intentional expression of human emotion, at times, one tends to hide these as well. However, through various cues the dominant emotion hidden in a communication can be identified. This can be conveniently identified in a face-to-face communication. However, for the communication that is made from a long distance, it is a challenging task. Such communication usually involves the text-based modes of message passing, like emails and letters.

Owing to the recent advancements in the domains of information, communication, and technology, exchange of views on current trends/events irrespective of geographical borders has increased exponentially [7]. This has been achieved through various online social networking websites (like, Twitter and Facebook) and email services (for example, Gmail and Yahoo). Additionally, communication between individuals having different native languages has also evolved through free online translation services [2]. All this not only involves sharing data and information, but also includes sharing of feelings. To share emotions in different languages (through text) appropriate and relevant words needs to be selected the get one's perspective across. Emotions also play a vital role in the lives of monolinguals and multilinguals. For example, when bilingual parents use the verbal communication mode to express their emotions to their children, it does not have to be their first language, but the language in which they are to express their emotions. Radoš in [19] gives example of an English–French bilingual who moved to France in early adulthood. She used French to express her emotions instead of English. Her believe is that she had discovered what love meant in the French language. Past studies show that people acquire second language in a naturalistic way or in an instructed manner [20]. This usually happens when individuals migrate from one place to another. Therefore, they end up learning a language in a natural way to survive in the new location/country and enable communicate with other individuals. Another way of learning second language may be because of personal interest. In the instructed way of learning, emotion is the basis of any learning or absence of the same. Positive and negative emotions influence a learning situation. On one hand, if stimulus for learning is positive then the learner will put efforts to learn it. On the other hand, if the stimulus is negative then less effort and attention will be devoted to it [19].

1.1 Research Gap and Problem Statement

In the field of emotion detection, we found following research gaps:

- A dataset consists of text data collected in non-acted environment and follow theory of basic human emotions.
- Datasets of low resource languages are less available online.
- Study of the effect on emotion detection rate after translation on a non-acted dataset.
- Study on the effect of emotions when original text data is translated into another language.

As mentioned earlier, the communication between speakers of different native languages requires utilization of text translation from one language to another. This is made convenient through various free online translation tools, especially via Google Translate.[1] However, this may compromise the expression of emotion which is present in the source language. There are a few languages which have very large vocabulary, for example, English having around 171,476 words. Whereas there are a few other languages with limited vocabulary, thus providing limited opportunities for the express of emotion. Currently, there is a need of intelligent methods that can enable to identify the hidden emotion in a text. Additionally, the ability of the same needs to be investigated when the actual text is translated into another language. This will not only enable to evaluate the robustness of the methods in identification of emotions but will ably assist in gauging the ability of the target language in expressing the emotion appropriately (as good as it was represented in the source language originally). Through translation of text to another language, the identification of feelings is a challenge. To the best of our knowledge there is no work reported in the past that investigates performance of the machine learning methods in identifying the emotion hidden in short, translated text. There are a few past works like [9–11, 17] that enable identification of emotion in limited text data; however, they are restricted to one language only.

1.2 Contribution

This work contributes towards the emotion detection from low resource languages. The present work aims to address the abovementioned gap in the past literature. Contributions of this work are highlighted below:

- Emotion detection from low resource languages.

[1] https://translate.google.com/.

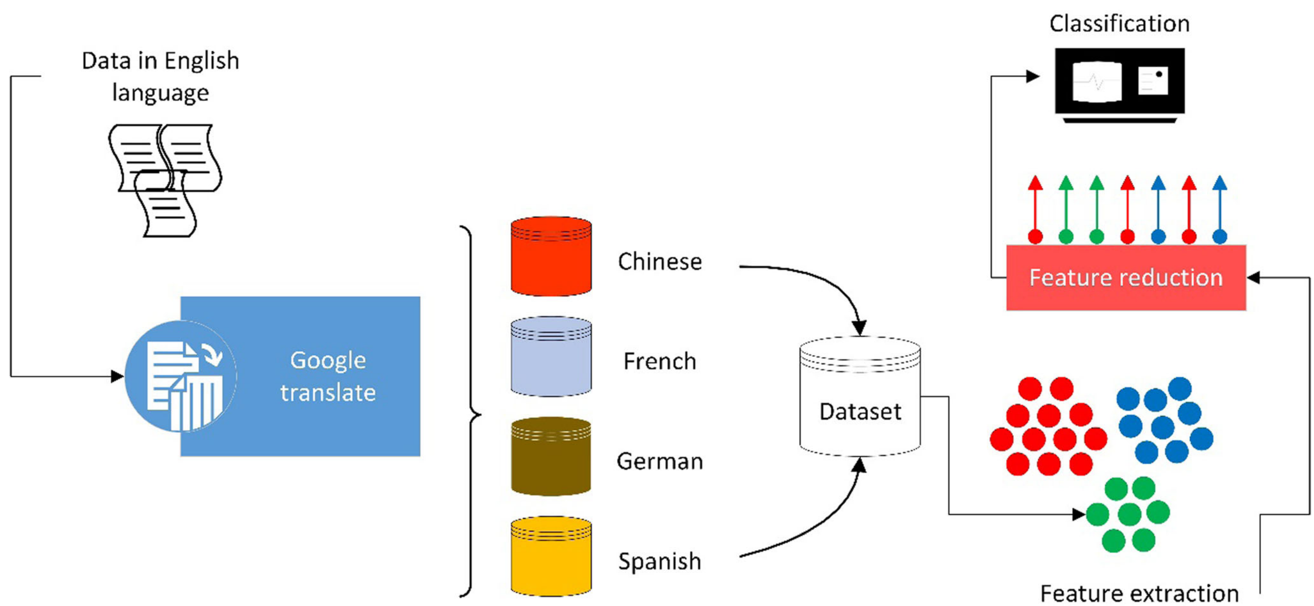

Fig. 1 Working of the overall solution

- Analysis of emotion detection rate from translated data.
- Study on a novel dataset.

For this purpose, the present work develops a machine learning-based framework to identify emotion that is hidden in limited text data. Any machine learning task requires the data at its base layer. Same goes for the present work. This work engages a set of volunteers to create a labeled dataset consisting of five basic human emotion using the English language. Later, the actual dataset is translated into four low resource languages, namely, Chinese, French, German, and Spanish utilizing Google Translate. Afterwards, the machine learning framework is executed on the five datasets to evaluate performance of the proposed framework. Overall framework is shown Fig. 1. This works also investigates the relation of text with its translated form and how much emotions of a person preserved by translating her written feelings. To analyze distribution of emotions, theory of Paul Ekman is followed. The current work applies machine learning techniques in this work. The five languages considered in this work are chosen because of their usage frequency and number of speakers worldwide. These five languages are among the most spoken languages across the globe. Therefore, the data and experiments performed in this work are based on these five languages.

The rest of the paper is organized as follows. Section 2 contains literature review; Sect. 3 describes process of dataset formation. Section 4 describes the experiments performed on the data. Section 5 consists of results and detailed analysis of this work and finally Sect. 6 concludes this paper with a few future directions mentioned.

2 Literature Review

2.1 Emotion Analysis from Languages

This section presents state of the artwork in the field of emotion analysis from different languages. Shalunts et al. in [22] present a novel sentiment analysis method which apply trilingual sentiment classification problem in the domain of general news. They specifically cover the natural disasters in general news. The languages examined in their work are English [14], German, and Russian. Their solution is named SentiSAIL and it employs the methodology of SentiStrength and covers the domain of general news in English, German and Russian languages. A work on the analysis of emotion in code switching text is presented in [25], work by Wang et al. Data in their work is collected from Weibo.com, a popular Social Networking Service (SNS) website in China. They use encoding representation for each character in the post to identify the code-switching post. Noise and advertisements have been removed from the posts and the remaining 4195 code switching posts are used for emotion detection. Five basic emotions are annotated, namely sadness, happiness, surprise, fear, and anger. A method is proposed in [13] for emotion identification in Chinese song lyrics based on affective lexicon. The emotion of a sentence is calculated from its emotion units. Grouping of lyric's sentences according to their emotions is done by a fuzzy clustering method. Weights and confidence of sentences of a cluster are used to weight a cluster and singing speed of the sentences are used to adjust the cluster weights. A multilingual database is created in [28] for the detection of natural and stress emotion. The data was collected in their work from university students to study the stress expression differences in a variety of languages. The work in [21] evaluate six emotional speech corpora from three families of languages namely Germanic, Romance, and Sino-Tibetan for enhancing multilingual recognition of emotion in speech by language identification. Emotion classes considered in their work are positive/negative and arousal/valence. It is observed that automatic Language Identification (LID) for selecting training corpora is superior in using all available corpora when the spoken language is not known. The work in [17] model adaptive non-linear interaction effects of appraisal factors on the individuals' motivation, their expression, and physiology simultaneously. For this purpose, they use principal component analysis (PCA) for dimensionality reduction of the data and multivariate adaptive regression splines (MARS) for automatic interaction identification. The data utilized in their work was from 27 countries and represented semantic profiles of component information in 24 common emotion words. A.

2.2 Factors Involved in Emotions in Different Cultures

Emotional valence is considered as key emotional process. To disentangle emotional and non-emotional processes in healthy persons and patients, a parametric functional Magnetic Resonance Imaging (fMRI) is applied in [12]. Thirteen healthy volunteers participated in giving images to cover the entire range of emotional valences. It has been observed that emotional valence exerts its effects predominately via modulation of signal decrease. Their work concluded that emotional valence may be related to neural processing in cortical midline regions. Gender and culture differences in emotion are discussed in [5]. A study on the relationship between color perceptual attributes, color emotions, and the influence of different cultural backgrounds is taken in [6] Gao et al. In their work, 214 color samples were evaluated on 12 emotional variables. The data were collected from subjects from seven different regions. It was concluded that three factors were sufficient to represent 80 "region-emotion" variables. These factors mainly relate to Chroma, lightness, and hue. Chroma and lightness were the most important factors in color emotions, whereas hue and cultural background were the least important factors. Gorini et al. in [8] analyzed the role of the cultural and technological backgrounds of the users in the emotional responses to virtual reality. They used the core effect model by Russell to explore the factors. They concluded that civilized participants belonging to Mexico City were able to report a significant reduction in the self-reported anxiety while inhabitants of a rural area EI Tepeyac showed a reduction in their physiological reactions but not in their anxiety. The work in [1] applied principals of emotion expressions across different cultures. They concluded that different cultures have different rates of emotion expressions. Lind et al. in [15] study stress as a predisposing factor for somatoform disorder. Analysis was conducted on data collected from 24 semi-structured subjects by taking their life history interviews. They concluded that generally psychosocial stress effect during childhood/youth, but it can vary. Their main theme was emotion avoidance culture, further three related subthemes were identified namely, (a) experiencing difficulties in communicating concerns, problems, expressing complex feelings in close relations; (b) suppressing their needs, vulnerability, sadness and anger that were not recognized by significant adults; and (c) disconnecting their stress reaction awareness from stressful bodily sensations using avoidant behaviors, e.g., by being highly active. The work in [26] evaluate the performance of machine learning methods in detecting the within-person fluctuations in one emotional state. This is done utilizing the acoustic analysis. They find that it is not yet possible to automatically assess fluctuations in one emotional state.

3 Data

A novel dataset is created for emotion detection consists of text and keystrokes data. The solution proposed in this work is evaluated on the DEKT-345×2 data which we formed during a research work conducted at the research group's laboratory. The dataset consists of text and keystrokes information based on five basic classes of emotions. However, in the present work, only the text data are utilized because the problem at hand is based on translated text data. The text data of DEKT-345×2 are in English language. Details about the dataset are given in the following sections.

3.1 Data Collection

To identify the dominant emotion hidden in multilingual data, a dataset is created in English language based on five basic emotions introduced by Paul Ekman. The dataset is developed to train and test the designed modals. The novelty of the data is that it consists of text features representative of various human emotional states in the form of text and keystrokes. The complete dataset consists of 345 samples of short text and 345 samples of keystrokes corresponding to the same data. Thus, there are a total of 2× 345 samples in the dataset. Therefore, it is named as DEKT-

Table 1 Statistics of DEKT-345×2 database

Statistics related to videos	
Number of videos	5
Video duration	4–5 min
Selection and labeling method	Content, comments, no. of likes and dislikes, and no. of views
Language of videos	English
Statistics related to samples and participants	
No. of participants	69
Gender and age range	(16 females, 53 males), 19–25 years old
No. of samples	345

Table 2 Description of emotion stimuli used for data collection

Video number	Content	Label	URL
1	Last victory moments of final match of World Cup 1992	Happy	https://www.youtube.com/watch?v=rL4xKdqyMgY
2	How tourist enjoy tourism in Pakistan	Happy	https://www.youtube.com/watch?v=nXwijt1EUDo
3	The difficulties faces by immigrants during illegal immigration	Disgusting, Fear	https://www.youtube.com/watch?v=oyBRr8h4ZrI
4	Destruction of natural resources and how they are converting to industrial products	Disgusting	https://www.youtube.com/watch?v=WfGMYdalClU
5	How students are mentally effected due to education system	Fear, sad	https://www.youtube.com/watch?v=BE4oz2u6OHY
6	The summary of causes of Syria war	Anger	https://www.youtube.com/watch?v=K5H5w3_QTG0
7	The condition of affected children is shown in this video	Fear, Sad and Anger	https://www.youtube.com/watch?v=jdKHVnHTXkU
8	Mystery of Bermuda triangle	Surprise	https://www.youtube.com/watch?v=q_5n7URd2Gk
9	Passion of a scientist to observe lava lake closely	Surprise	https://www.youtube.com/watch?v=egEGaBXG3Kg

345×2, i.e., dataset for emotion detection using keystrokes and text. The procedure of dataset preparation was designed in a way that we can achieve the required data against each emotion. The main procedure of this data acquisition was through (a) induction of emotions in the participants using videos, (b) collection of short text which is written in the context of the induced emotion, and (c) tracking of keystrokes of text while the participants give their samples. This work focused on using only the short text of length between 5 and 150 words approximately. Data were collected in two different sessions; first session was conducted during the months of June and July of the year 2018 while the second session was conducted in the month of March 2019. In the dataset, the shortest text is of 4 words while the longest text is of 116 words. Table 1 lists the key features of the data.

3.2 Participants

The volunteers who participated in the data acquisition phase were the undergraduate and graduate students of the institute. A total of 48 individuals participated in the first session. Out of these 48 participants, 75% were males and the rest were females. In the second session, 21 undergraduate students participated, out of which 19% were female and the remaining were male.

3.3 Emotion Stimuli

Emotions were induced in the participants by showing them video clips. For this, nine different videos were initially selected from the YouTube against the five basic emotions. Finally, five videos were shown to participants, one against each emotion. Selection of the final five videos was made based on the number of views, content, topic, and comments on the video. The detail about each video is listed in Table 2. Once the emotion was induced, the participants were asked to express their view in a text file using a maximum of 150 words. Afterwards the text data of DEKT-345×2 dataset was translated into four different languages: namely, Chinese, French, German, and Spanish. Following is a text sample and its translated version in the five languages.

English: Very sad situation in Syria, the country is caught in a proxy war between multiple nations or groups aiming to further their own interests.

Chinese: 敘利亞非常悲慘, 該國陷入了旨在增進自己利益的多個國家或集團之間的代理戰爭.

French: situation très triste en syrie, le pays est pris dans une guerre par procuration entre plusieurs nations ou groupes visant à promouvoir leurs propres intérêts.

German: Sehr traurige Situation in Syrien. Das Land befindet sich in einem Stellvertreterkrieg zwischen mehreren Nationen oder Gruppen, um ihre eigenen Interessen zu fördern.

Spanish: Muy triste situación en Siria, el país está atrapado en una guerra de poder entre múltiples naciones o grupos con el objetivo de promover sus propios intereses.

4 Experiments

Experiments are performed on the text data using machine learning techniques and are evaluated using metrics of accuracy, error, precision, sensitivity, specificity, kappa, F1 score, and Matthews correlation coefficient. Details about experiments are given in the following.

4.1 Machine Translation

Machine translation or automated translation is a field of research since 1950s. Machine translation provides text translation by the computer without human interaction. This work utilized the Google Translator tool to translate the original English text data to five other languages. Google Translator has a free web interface, but it does not have free application programming interface (API). Google's free service instantly translates words, phrases, and web pages between English and 100 other languages.

4.2 Feature Extraction

Evaluation of the proposed framework is based machine learning techniques. Features from each set of data are extracted using scikit-learn software of machine learning used in Python programming language [31]. Term frequency (TF) and inverse document frequency (IDF) based features are extracted using Eq. (1).

$$\mathrm{IDF}(\mathrm{t},\mathrm{d}) = \log\left(\frac{\mathrm{N}}{|\{\mathrm{d}\in\mathrm{D}:\mathrm{t}\in\mathrm{d}\}|}\right), \tag{1}$$

where N is the number of documents (samples), d is a document belonging to the set of documents D and t is the term in the document d. The term IDF is inverse document frequency of term t in document d. It gives a value which shows how important a word (term) is to a document (sample). Term frequency is the count of the term that appears in a document and inverse document frequency is measure of information that the term provides.

4.3 Feature Reduction

Term frequency-inverse document frequency (TFIDF) yields the importance of each word to a document but in this way the obtained dimensionality of total features become very high. To select the most informative features and to reduce the dimension of features, this work used principal component Analysis (PCA) method in [27] and tree-based method proposed by Geurts, Pierre et al. [30]. The PCA extracts the dominant patterns in the input matrix. A covariance matrix of dimensions is generated from which eigenvectors and eigenvalues are generated. PCA choses the top k eigenvectors that correspond to largest eigenvalues. TB uses meta estimator that fits a number of randomized decision trees on various subsamples of data. TB method uses extra tree classifier to reduce feature dimensions. Extra tree classifier improves predictive accuracy and avoid overfitting.

4.4 Classification of Emotions

Classification of emotions from DEKT-345×2 dataset and translated datasets is performed by dividing the data into training, test, and validation sets. The 20% data is used for testing, and 80% data is used for training and validation. Classification is performed here using support vector machine (SVM) proposed by Chang et al., 2011 in [4] and discriminant analysis proposed by Mika, Sebastian, et al. In [30] SVM classifier utilized for classification of text and keystrokes features uses linear kernel as it performs well in case of high dimensionality of the features. Text and keystrokes are classified on one-against-one multiclass classification approach. The discriminant analysis classifier uses discrimination types of linear or diaglinear to classify the data.

Table 3 Results obtained using SVM and PCA on DEKT-345×2 dataset and its translated data

Language	Accuracy (%)	Error	Sensitivity	Specificity	Precision	F1 score	Matthews Correlation Coefficient	Kappa
English	92.75	0.07	0.93	0.98	0.92	0.93	0.91	0.77
Chinese	79.71	0.20	0.81	0.95	0.81	0.79	0.75	0.36
French	84.06	0.16	0.84	0.96	0.83	0.83	0.79	0.51
German	85.51	0.14	0.87	0.96	0.88	0.86	0.84	0.55
Spanish	89.86	0.11	0.91	0.97	0.91	0.89	0.87	0.68

Table 4 Results obtained using SVM and TB on DEKT-345×2 dataset and its translated data

Language	Accuracy (%)	Error	Sensitivity	Specificity	Precision	F1 score	Matthews correlation coefficient	Kappa
English	75.36	0.25	0.74	0.94	0.73	0.73	0.673	0.23
Chinese	35.4	0.65	0.29	0.82	0.71	0.244	0.24	0.51
French	67.6	0.33	0.67	0.91	0.69	0.67	0.59	0.04
German	68.12	0.32	0.69	0.92	0.73	0.67	0.62	0.003
Spanish	68.12	0.32	0.67	0.92	0.69	0.68	0.602	0.0036

4.5 Evaluation Metrics

The performance metrics utilized here for reporting results include accuracy, precision, recall, and F measure. Prerequisite to these metrics is the computation of the confusion matrix. Each row of the matrix shows the observations of the predicted class and each column visualizes the actual value. On the basis of event observation, the positive class represents the positive events and vice versa. The term true positive (TP) means that the observation is positive along with a positive predication made by the classification model. The term false negative (FN) signifies that the prediction is negative, but the observation still remains positive. The term false positive (FP) means that the observation is negative, but the model prediction is positive. The negative observation and prediction is indicated by the term true negative (TN).

Accuracy: Accuracy is the measure to evaluate a classifiers' performance. It is defined as a ratio of correctly identified observations to the total number of observations. The higher the accuracy, better are the results. Its value ranges from 0 (worst) to 1 (best). Equation (2) shows computation of the accuracy.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}. \tag{2}$$

Precision: It is also known as positive predictive value (PPV) and is defined as the ratio of total correct positive predictions and all positive predictions. The best value of specificity is 1 and its worst value is 0. Equation (3) shows the computation of precision.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{3}$$

F measure: The F score (Eq. 4) includes both recall and precision. It has 1 as the best value and 0 as the worst. The F measure is a harmonic mean therefore; it is inclined towards the smaller of the two elements. The F1 score will be small if either precision or recall is small.

$$F.\text{measure} = \frac{\left(1 + \beta^2\right)(\text{Precision} \times \text{Recall})}{\left(\beta^2 \times \text{Precision} + \text{Recall}\right)}, \tag{4}$$

where β is commonly set to a value of 0.5, 1, or 2.

Error: Error in classification is calculated using (5), by subtracting accuracy from 1.

$$\text{Error} = 1 - \text{accuracy}. \tag{5}$$

Sensitivity: Sensitivity is calculated using (6), i.e. by dividing true positive samples with sum of true positive and false negative samples.

$$\text{Sensitivity} = \text{TP}/\text{TP} + \text{FN}. \tag{6}$$

Specificity: Specificity is calculated by dividing true negative with sum true-negative and false-positive samples, using (7).

$$\text{Specificity} = \text{TN}/\text{TN} + \text{FP}. \tag{7}$$

Matthews correlation coefficient (MCC): Accuracy is sensitive to class imbalance, precision, F1 score are

Table 5 Results obtained using discriminant analysis and PCA on DEKT-345×2 dataset and its translated data

Language	Accuracy (%)	Error	Sensitivity	Specificity	Precision	F1 score	Matthews correlation coefficient	Kappa
English	93.3	0.07	0.93	0.98	0.93	0.93	0.91	0.79
Chinese	90.3	0.09	0.91	0.98	0.91	0.91	0.88	0.69
French	99.04	0.009	0.991	0.997	0.99	0.99	0.99	0.97
German	93.27	0.07	0.93	0.98	0.93	0.93	0.91	0.79
Spanish	98.08	0.02	0.98	0.99	0.98	0.98	0.98	0.94

Table 6 Results obtained using discriminant analysis and TB on DEKT-345×2 dataset and its translated data

Language	Accuracy (%)	Error	Sensitivity	Specificity	Precision	F1 score	Matthews correlation coefficient	Kappa
English	65.38	0.35	0.66	0.91	0.71	0.66	0.59	0.07
Chinese	30.77	0.69	0.29	0.82	0.85	0.23	0.26	0.54
French	73.08	0.27	0.74	0.93	0.76	0.75	0.68	0.16
German	65.38	0.35	0.66	0.91	0.67	0.65	0.57	0.07
Spanish	70.19	0.29	0.69	0.92	0.74	0.71	0.64	0.07

symmetric to that, so MCC is introduced. The higher the correlation between true and predicted values, the better the prediction. MCC is computed using (8).

$$\mathrm{MCC} = \mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN} / \sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}. \tag{8}$$

Kappa: Kappa statistics is a measure that can handle both multiclass and imbalanced class problem.

$$k = p_o - \frac{p_e}{1} - p_e, \tag{9}$$

where p_o is the observed class and p_e is the expected class.

5 Results and Analysis

This section evaluates the difference between emotion detection rates using the ML model among different languages. Here, the dominant emotion is identified from the text data is translated into the five languages using machine learning techniques. Whereas the original labelled dataset in English language is utilized as the ground truth. For this purposes, five versions of the original dataset representing the earlier mentioned five emotions are created by translating the English text data of DEKT-345×2. Evaluation of each dataset is performed using the same set of machine learning techniques. The results are evaluated using accuracy, precision, recall, and F1 score. The obtained results by classifying the data into 4 classes of emotions on the DEKT-35×2 data and its translated data using SVM and PCA are listed in Table 3, using SVM and TB are listed in Table 4, using PCA and Discriminant Analysis are listed in Table 5, using TB and Discriminant Analysis are listed in Table 6.

Table 7 Results obtained using extra-trees on DEKT-345×2 data and its translated data

Dataset	Features	Accuracy (%)
English	Linguistic	37.68
	Statistical	72.46
French	Linguistic	43.48
	Statistical	60.87
German	Linguistic	30.43
	Statistical	72.46
Spanish	Linguistic	34.78
	Statistical	71.01

6 Results

It can be seen from the results that the accuracy achieved using SVM and PCA on the Chinese language is 79.71%, for the French language an accuracy of 84.06% is achieved, using the translated German language 85.51% accuracy is obtained, and Spanish language achieved 89.86% accuracy. For the translated Chinese language higher error rate, i.e., 0.20 is observed. Whereas, the Spanish language has a low error rate, i.e., a misclassification rate of 0.11. The Spanish language has the high sensitivity of 0.91 while Chinese language has the low sensitivity of 0.81. Similarly, Spanish language has the high specificity of 0.97 and Chinese language has the low specificity of 0.95. Spanish language has the high precision rate of 0.91 and Chinese language has the low precision value of 0.81. The Spanish language has a higher F1 score of 0.89 while Chinese language has the lowest F1 score of 0.79. Similarly, Chinese language has the high Matthews correlation coefficient value and kappa values of 0.87 and 0.68, respectively, whereas Spanish language's Matthew's correlation coefficient value and kappa values are 0.75 and 0.36, respectively. Results achieved by PCA and Discriminant Analysis (DA) are better than using SVM by yielding 99.04% accuracy on French data, 98.08% on Spanish data, 93.3% on both English and German and 90.3% on German data. Results achieved by SVM and TB also show that emotions remain preserved in the order of Spanish>German>French>Chinese. Results achieved by TB and DA show that emotion detection rate decreases in the order of French>Spanish>German>Chinese.

6.1 Analysis

From the results, it can be inferred that Spanish language is richer in emotions among the five languages considered in this work. It can be seen that for the Chinese language the ML methods yield the lowest emotion identification accuracy. The reason underlying these results is that there are many similarities between the English and Spanish languages. This is supported by the fact that around 30–40% of all words in English have a related word in Spanish [24]. Spanish and English language also use mostly the same letters. The Royal Spanish Academy (RSA) determined in 2010 that only *ñ* is a different letter in Spanish than the letters in English, making Spanish consist of 27 letters. Spanish and English have thousands of cognates. When compared to most other languages, the syntax of Spanish is observed to be very similar to that of English as well as discussed in [18]. Another similarity between the two is that both languages pluralize the words by adding "s" or "es" at the end of the words. Spanish also capitalizes many of the same words and use much of the same punctuation as the English language, but they just use less of them and in a different way. The rules which are same to English for capitalization are:

o The first word of a sentence.
p In use of proper noun.
q Titles, but in most cases just the first letter.

As far as the Chinese language is concerned, the ML has lesser emotion detection rate when text is translated from English to Chinese. This is mainly because the Chinese language has so many significant differences from English. This makes English learning difficult for Chinese native speakers. Chinese does not have alphabets but uses a logographic system for its written language. Some English phonemes do not exist in Chinese at all; stress and intonation patterns are different. Additionally, there are various differences in word order between Chinese and English languages [3]. All this makes it challenging for the ML to accurately identify the dominant emotion hidden in the Chinese language that is originally translated from English.

6.2 Comparison with State-of-the-Art Methods

Proposed work is compared with state of the artwork in affective computing field. Work presented by Jatinder kumar R. Saini et al. [29] is based on emotion detection from an annotated corpus Kavi which is based on Navrasa. It is about emotion detection from Punjabi poetry based in Indian concept of Navrasa. In this corpus, 948 poetries were classified into 9 emotion states namely karuna, shringar, hasya, raudra, veer, bhayanak, vibhata, adbhur, and shaanti. The corpus is manually annotated, and Kappa Flies index is used for inter-annotator agreement. Three type of features were extracted from data, linguistic, poetic, and statistical. Naïve Baes and Support Vector Machine classifiers are used for classification. It was concluded that SVM improves overall accuracy by yielding 70.02% accuracy, further poetic features outperform the linguistic features for emotion detection. We applied their framework on our data. Obtained results are shown in Table 7. It can be seen that statistical features yield better accuracy than linguistic features. Methodology of [29] gives 72.46% accuracy on German and English dataset while our features yield 92.75% maximum accuracy on English data while on German data it gives 85.51% accuracy using SVM which is better than comparison method's results.

7 Conclusion and Future Directions

The paper presented an analysis of emotion detection from text when it is translated from original language to other languages. We took the experiments by taking English as the root language and translated it to four international languages: Chinese, French, German, and Spanish. Google translator was used to translate text in English to other languages. Data were based on short text written by individuals who were induced with five different emotions separately. Participants had to write about their feelings they had after watching each video. Videos were selected based on five emotion categories while the emotion categories were selected based on Ekman's theory. These emotions included angry, disgusting, fear, happy, and surprise. Machine learning methods were used to classify the data into five classes of emotions. Term frequency-inverse document frequency (TFIDF)-based features were extracted from text data. Later, principal component analysis (PCA) was used for feature reduction as dimensions obtained from TFIDF were high. This work utilized support vector machine (SVM) classifier for prediction of an emotion in the text. The performance evaluation criteria for experiments were based on accuracy, error, precision, F1 score, sensitivity, specificity, Matthew's correlation coefficient, and kappa value. The obtained results suggested that emotions were detected with an average accuracy of 89.86% when translated from English to Spanish language. The reason to this is that Spanish and English languages are very similar to each other in many contexts. Detection rate of Chinese language was lower than other languages because Chinese is different in many aspects than English. For the translated Chinese text, the classifier achieved 79.71% accuracy. Results achieved by PCA and Discriminant Analysis are better than using SVM by yielding 99.04% accuracy on French data, 98.08% on Spanish data, 93.3% on both English and German and 90.3% on German data.

Proposed work can be applied in different real-world problems solve them, like in embassies, immigration process, socializing with international persons. The presented framework can be applied on data available on social media platforms in form of different languages. It can increase the growth of business of products either through electronic media or through person-to-person interaction.

There are some future directions regarding this work. Text data of affective computing domain in other language than English can be analyzed to check the emotion detection rate in that language and its translated versions. One can analyze the difference in emotion detection from text in native language and in translated form. Deep learning models can be applied to compare the performance of deep learning with machine learning techniques [16]. Another future direction is to experiment by extracting distinct features which influence the emotion detection accuracy when the text is translated.

Acknowledgements The authors would like to thank GIK Institute for providing research funding.

Author contributions MT: conceptualization, methodology, software, and writing—original draft. ZH: data curation, writing—review and editing, and software. MW: visualization, investigation, and software. ST: software, and writing—original draft.

Funding This work was sponsored by the GIK Institute graduate research fund under GA3 scheme.

Availability of Data and Material The dataset for the experimentation is available at https://www.minrg.org/toolsanddata.

Declarations

Conflict of interest The authors have no conflict of interest.

Ethical Approval and Consent to Participate All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Consent for Publication Informed consent was obtained from all individual participants included in the study.

References

1. Bhugra, D., McKenzie, K.: Expressed emotion across cultures. Adv. Psychiatr. Treat. **9**(5), 342–348 (2003)
2. Canfora, C., Ottmann, A.: Risks in neural machine translation. Transl. Spaces **9**(1), 58–77 (2020)
3. Cao, F., Perfetti, C.A.: Neural signatures of the reading-writing connection: greater involvement of writing in Chinese reading than English reading. PLoS ONE **11**(12), e0168414 (2016)
4. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 1–27 (2011)
5. Fischer, A.H., Rodriguez Mosquera, P.M., Van Vianen, A.E., Manstead, A.S.: Gender and culture differences in emotion. Emotion **4**(1), 87 (2004)
6. Gao, X.P., Xin, J.H., Sato, T., Hansuebsai, A., Scalzo, M., Kajiwara, K., Billger, M.: Analysis of cross-cultural color emotion. Color Res Appl **32**(3), 223–229 (2007)
7. Gelman, J., Wilson, S.L., Sanhueza Petrarca, C.: Mixing messages: how candidates vary in their use of Twitter. J Inform Technol Polit **18**, 1–15 (2020)
8. Gorini, A., Mosso, J.L., Mosso, D., Pineda, E., Ruíz, N.L., Ramíez, M., Riva, G.: Emotional response to virtual reality exposure across different cultures: the role of the attribution process. Cyberpsychol Behav **12**(6), 699–705 (2009)
9. Gupta, N., Gilbert, M., Fabbrizio, G.D.: Emotion detection in email customer care. Comput. Intell. **29**(3), 489–505 (2013)
10. Halim, Z., Rehan, M.: On identification of driving-induced stress using electroencephalogram signals: a framework based on wearable safety-critical scheme and machine learning. Inform Fus **53**, 66–79 (2020)
11. Halim, Z., Waqar, M., Tahir, M.: A machine learning-based investigation utilizing the in-text features for the identification of

dominant emotion in an email. Knowl.-Based Syst. **208**, 106443 (2020)

12. Heinzel, A., Bermpohl, F., Niese, R., Pfennig, A., Pascual-Leone, A., Schlaug, G., Northoff, G.: How do we modulate our emotions? Parametric fMRI reveals cortical midline structures as regions specifically involved in the processing of emotional valences. Cogn. Brain Res. **25**(1), 348–358 (2005)
13. Hu, Y., Chen, X., & Yang, D.: Lyric-based song emotion detection with affective lexicon and fuzzy clustering method. In ISMIR (pp. 123–128) (2009)
14. Lee, J.S.: Informal digital learning of English and second language vocabulary outcomes: can quantity conquer quality? Br. J. Edu. Technol. **50**(2), 767–778 (2019)
15. Lind, A.B., Delmar, C., Nielsen, K.: Struggling in an emotional avoidance culture: a qualitative study of stress as a predisposing factor for somatoform disorders. J. Psychosom. Res. **76**(2), 94–98 (2014)
16. Rostami, M., Oussalah, M., Farrahi, V.: A novel time-aware food recommender-system based on deep learning and graph clustering. IEEE Access **10**, 52508–52524 (2022)
17. Meuleman, B., Moors, A., Fontaine, J., Renaud, O., Scherer, K.: Interaction and threshold effects of appraisal on componential patterns of emotion: a study using cross-cultural semantic data. Emotion **19**(3), 425 (2019)
18. Nediger, W., Pires, A., Guijarro-Fuentes, P.: Integrating linguistic theory and experimentation in L2 acquisition: learning of Spanish differential object marking by Portuguese and English speakers. Appl. Linguist. Rev. **10**(2), 219–239 (2019)
19. Richards, D.: Exploring the effect of language of instruction on cross-linguistic influence in third language acquisition (Doctoral dissertation) (2020)
20. Sagha, H., Matejka, P., Gavryukova, M., Povolný, F., Marchi, E., Schuller, B.W.: Enhancing multilingual recognition of emotion in speech by language identification, pp. 2949–2953. Interspeech (2016)
21. Shalunts, G., Backfried, G., Prinz, P.: Sentiment analysis of German social media data for natural disasters. In ISCRAM (2014)
22. Stanley, R.O., Burrows, G.D.: Varieties and functions of human emotion. In: Emotions at work: theory, research and applications in management, pp. 3–19. Wiley (2001)
23. Vélez-Uribe, I., Rosselli, M.: The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals. Bilingualism Lang Cogn **22**(1), 30–46 (2019)
24. Wang, Z., Lee, S., Li, S., Zhou, G.: Emotion detection in code-switching texts via bilingual and sentimental information. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (VOLUME 2: Short Papers) (pp. 763–768) (2015)
25. Weidman, A.C., Sun, J., Vazire, S., Quoidbach, J., Ungar, L.H., Dunn, E.W.: (Not) hearing happiness: predicting fluctuations in happy mood from acoustic cues using machine learning. Emotion **20**, 642 (2019)
26. Wold, S., Esbensen, K., Geladi, P.: Principal component analysis. Chemom. Intell. Lab. Syst. **2**(1–3), 37–52 (1987)
27. Zuo, X., Lin, L., & Fung, P.: A multilingual database of natural stress emotion. In: Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12) (pp. 1174–1178) (2012)
28. Saini, J. R. et al.: Procedia Comput. Sci. **167**:1220–122 (2020)
29. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Mach. Learn. **63**(1), 3–42 (2006)
30. Sebastian, M., et al.: Fisher discriminant analysis with kernels. In: Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop (cat. no. 98th8468). Ieee, (1999)
31. Azadifar, S., Rostami, M., Berahmand, K., Moradi, P., Oussalah, M.: Graph-based relevancy-redundancy gene selection method for cancer diagnosis. Comput. Biol. Med. **147**, 105766 (2022)