



ORIGINAL

A Text-based Intelligently driven Emotion Recognition Framework

Un marco de reconocimiento emocional impulsado por la inteligencia basado en texto

Xiaoping Wu¹ , Hanyu Lu¹ 

¹International College, Krirk University. Bangkok, 10220, Thailand.

Cite as: Wu X, Lu H. A Text-based Intelligently driven Emotion Recognition Framework. Salud, Ciencia y Tecnología. 2024; 4:.988. <https://doi.org/10.56294/saludcyt2024.988>

Submitted: 26-02-2024

Revised: 16-06-2024

Accepted: 09-10-2024

Published: 10-10-2024

Editor: Dr. William Castillo-González 

ABSTRACT

Introduction: emotion recognition from text has gained considerable interest due to its applications in human-computer interaction, emotion analysis, and psychiatric research. Traditional methods have struggled with emotional ambiguity, cultural nuances, and the dynamic nature of language, which affect the reliability of emotion recognition. This paper presents a novel emotion recognition framework named Artificial Gorilla Troops driven Modified Gaussian Support Vector Machine (AGT-MGSVM).

Method: we gathered a publicly available ISEAR dataset containing various textual emotional expressions and applied natural language processing (NLP) techniques for text pre-processing. The suggested AGT-MGSVM approach combines the resilience of the Gaussian assist vector gadget (GSVM) with the ability of AGT, a bio-stimulated optimization method. AGT improves the MGSVM with the aid of dynamically regulating its parameters based on the evolutionary conduct of gorilla troops, optimizing the version to enhance emotion popularity.

Results: we examine the performance of the proposed technique against traditional emotion reputation methods using standard metrics inclusive of recall (89,2 %), precision (89,5 %), F1-score (89,4 %), and accuracy (89,9 %).

Conclusion: the counseled AGT-MGSVM method is a promising improvement in intelligence-driven emotion reputation from the text.

Keywords: Emotion Recognition; Textual Content; Natural Language Processing (NLP); Activated Gorilla Troops Were Moved Using an Alternating Gaussian Support Vector Machine (AGT-MGSVM).

RESUMEN

Introducción: el reconocimiento de emociones a partir de texto ha ganado considerable interés debido a sus aplicaciones en la interacción humano-computadora, análisis de emociones e investigación psiquiátrica. Los métodos tradicionales han luchado con la ambigüedad emocional, los matices culturales y la naturaleza dinámica del lenguaje, que afectan la fiabilidad del reconocimiento emocional. Este artículo presenta un novemarco de reconocimiento de emociones llamado Artificial Gorilla Troops driven Modified Gaussian Support Vector Machine (AGT-MGSVM).

Método: reunimos un conjunto de datos ISEAR disponible públicamente que contiene varias expresiones textuales emocionales y técnicas aplicadas de procesamiento del lenguaje natural (PNL) para el pre-procesamiento del texto. El enfoque AGT-MGSVM sugerido combina la resistencia del Gaussian assist vector gadget (GSVM) con la capacidad de AGT, un método de optimización bio-estimulado. AGT mejora el MGSVM con la ayuda de la regulación dinámica de sus parámetros basados en la conducta evolutiva de las tropas gori, la optimización de la versión para mejorar la popularidad de la emoción.

Resultados: se examina el desempeño de la técnica propuesta frente a los métodos tradicionales de reputación emocional utilizando métricas estándar incluyendo memoria (89,2 %), precisión (89,5 %), F1-score (89,4 %) y precisión (89,9 %).

Conclusión: el método AGT-MGSVM aconseja una mejora promete en la reputación de la emoción impulsada por la inteligencia del texto.

Palabras clave: Reconocimiento de Emociones; Contenido Textual; Procesamiento del Lenguaje Natural (PNL); Activación de las Tropas Gorilas Movidas Usando una Máquina Alternativa de Vectores de Soporte Gaussiano (AGT-MGSVM).

INTRODUCTION

Emotion recognition from text has emerged as a key area of research in artificial intelligence (AI) to comprehend human emotions as they are expressed via written expression. This capability has important implications for various domains, including personalized user interactions in digital environments, customer service, and behavioral monitoring. Researchers and programmers are exploring sophisticated algorithms that can recognize emotions in text by using natural language processing (NLP) techniques.⁽¹⁾ The use of data-based emotional intelligence represents a major advance in AI-powered emotional intelligence, not least in recognizing basic emotions such as happiness, sadness, fear, anger, and emotion.

The application of NLP models to textual emotion interpretation has increased significantly in the last several years. These models can now recognize complicated emotional micro-contexts expressed in textual communication in addition to fundamental emotions. To improve sentiment analysis and context-aware identification of emotions, researchers are incorporating emotional intelligence into AI.⁽²⁾

According to social media and market research agencies, emotion identification in text is becoming increasingly popular for consumer sentiment monitoring, with the worldwide AI industry for emotion identification and recognition expected to reach \$38,6 billion by 2028. NLP-based models are used in a variety of industries, including marketing and healthcare, to assess how people feel about certain goods, services, or social events by analyzing huge amounts of textual data. User satisfaction has increased in the United States attributed to a 15 % improvement in accuracy observed in AI-driven customer support platforms that incorporate emotion identification.⁽³⁾

The efficacy of text-based emotion identification techniques remains limited despite these advancements due to several issues. It is challenging for AI to reliably examine emotions without a great deal of training and validation since emotional representation in natural language is frequently subtle, unclear, and context-dependent. Large, diversified datasets are also a source of ethical challenges for the gathering and processing of which user privacy and authorization are prominent among them. To guarantee that AI systems operate effectively and maintain moral principles, these problems must be resolved.⁽⁴⁾

Addressing those challenges requires collaboration among linguists, psychologists, records scientists, and ethicists to make sure technological accuracy and moral responsibility in the use of such technology will be used inside the process.⁽⁵⁾ The future of intelligent sensing lies in tailoring models to handle multiple input languages, integrating contextual information for increased accuracy, and explicit systems have been developed that prioritize user consent and data privacy similarly to the likes of AI and NLP. As advances continue, the potential of these systems to transform human-computer interaction and the social understanding of emotion rises, promising a sympathetic and efficient digital landscape.⁽⁶⁾ This research aims to create and assess a novel framework for emotion recognition by utilizing the Artificial Gorilla Troops driven Modified Gaussian Support Vector Machine (AGT-MGSVM). By using cutting-edge machine learning algorithms and investigating solutions for the problems of emotional uncertainty, dataset diversity, and ethical data management in AI applications, this project attempts to improve the accuracy of emotion recognition in text.

The following divisions are found in the final part of this study: part 2 contains a related study; part 3 contains the methodology; part 4 contains the analysis of the research's findings and part 5 has the research's conclusion.

METHOD

The TF-IDF and Bagging algorithms were chosen in the research to construct a categorization model. The outcomes demonstrated that the two algorithms outperformed other techniques in the Spanish poetry categorization domain in terms of precision and objectivity.⁽⁷⁾

The proposed method by Author utilized AI to enhance learning via customized interactions and intelligent instruction.⁽⁸⁾ The technique focused on mobile robots' vision-based target detection, which directly affects their navigation and location. When tested using a public database, the suggested model performed better at target recognition than other cutting-edge techniques.

A text-based emotion identification and forecasting system was developed in the research.⁽⁹⁾ Four machine learning algorithms namely MNB, SVM, DT Classifier, and KNN Classifier were compared for classifier efficiency

in their research. The greatest efficiency was achieved by the multinomial Naïve Bayes classifier, which had an average accuracy of 64,08 %.

Using the GA-SVM algorithm for audio signal extraction and emotional feature parameters, paper introduced a novel speech recognition-based Chinese pronunciation system.⁽¹⁰⁾ A PNN-HMM model was used for emotional temperament recognition in broadcast speech. The outcomes indicated that the PNN-HMM model had a high identification rate and a robust resistance to interference.

A basic but efficient text recognition model built on an encoder-decoder architecture was provided in research.⁽¹¹⁾ The suggested system could differentiate between regular and irregular text. An encoder-decoder design for irregular word identification in natural photographs was suggested that used TPS transformation, SAM, Bi-LSTM network, and DSAM. The suggested approach performed better on regular and irregular text identification techniques than the traditional techniques.

The goal of the semantic-enhanced encoder-decoder framework proposed in article for computer vision was to accurately recognize scene texts that were of low quality.⁽¹²⁾ By integrating the cutting-edge ASTER approach, the suggested approach performed better on datasets used as benchmarks and was more resilient to poor-quality text images.

The approach to text line identification, which employed two ANN combined by a dynamic programming methodology, was presented in article.⁽¹³⁾ The suggested structure demonstrated the effective uses of ANNs that were incredibly lightweight for camera-captured image identification and FCNs for text line classification.

A GSRM was presented in the research to record global semantic contexts through multi-way parallel exchange, and the result was a unique end-to-end trainable architecture termed SRN for effective scene text identification.⁽¹⁴⁾ The efficacy and reliability of the suggested approach were confirmed by existing findings on seven public benchmarks, involving regular text, irregular text, and non-Latin lengthy text.

The CRNN model for text identification was improved in the article.⁽¹⁵⁾ It performed poorly in identifying irregular text, had relatively low accuracy, and only took into account receiving text sequence data from a single element, leading to insufficient information collection. The findings demonstrated that the modified approach surpassed its initial version in recognition efficiency and that the approach operated well in the majority of benchmark assessments.

A CNN and RNN-based Quranic OCR method was presented in the research.⁽¹⁶⁾ The result demonstrated that the suggested method achieved a 98 % accuracy rate on the validation information and a 95 % WRR and 99 % CRR on the test dataset.

The data collection involved emotional expressions from the ISEAR dataset. Natural language processing (NLP) techniques were implemented for text pre-processing, such as tokenization and sentiment evaluation. Then, AGT-MGSVM is used to identify the emotion based on text. Figure 1 shows the overall flow.

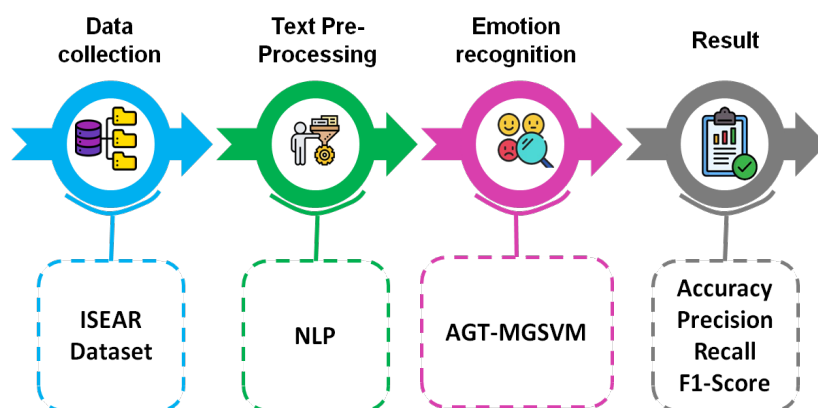


Figure 1. Processing of Emotion Recognition Framework based on Text

Dataset

The seven main emotions joy, fear, anger, sadness, disgust, humiliation, and guilt were asked to be reported by student responses, psychologists, and non-psychologists alike. The questions in each scenario focused on their assessment of the circumstances and their response. Approximately 3 000 respondents from 37 countries across five continents contributed reports on seven different emotions to the final data collection.⁽¹⁷⁾

Preprocessing

The pre-processing stage involves gathering the ISEAR dataset and applying natural language processing (NLP) methods such as tokenization, stemming, removing stop words, normalization, and lemmatization to clean and prepare the text data for effective emotion recognition by the AGT-MGSVM model.

Removing Stop words

Stop words are terms that are frequently found in datasets, such as papers, predication, adjectives, etc., but do not provide context or provide syntactic, semantic, or sentiment meaning.

Natural language processing must be used to eliminate these words that stop.

Tokenization

The process of turning sentences into words is known as tokenization. Tokenization also involves breaking the language up into separate phrases and adding each one to the list.

Normalization

To ensure consistency in the preparation of every text, standardization is employed. Many activities are carried out simultaneously, such as changing all text to upper or lower case and numbers to the corresponding words. Words that have the same meaning as “girl” and “GIRL” constitute non-identical words, therefore, normalization is crucial. All material in this publication has been changed to lowercase letters.

Stemming

Stemming is a method of returning various word tenses to their original form. Stemming offers the assistance required to remove terms from the list that are not intended to be computed. For instance, stemming techniques will transform the words “lose,” “losing,” and “lost” into “lose.”

Lemmatization

Lemmatization involves a method of combining two or more words into a single word by employing synonyms to reduce the words to a word that already exists in the language. Alternatives for every word are combined into one in this stage.

Table 1. Preprocessing result

Index	Emotion	Text	Removing Stop words	Tokenization	Normalization	Stemming	Lemmatization	Detokenized
31406	Joy	“Your future is bright”	“Future bright”	[“future”, “bright”]	[“future”, “bright”]	[“future”, “bright”]	[“future”, “bright”]	future bright
41124	Sadness	“I can’t pull myself out of depression.”	“can’t pull depression.”	[“can’t”, “pull”, “depression”]	[“can’t”, “pull”, “depression”]	[“can’t”, “pull”, “depress”]	[“can’t”, “pull”, “depression”]	can’t pull depression
21518	Fear	“I’m nervous about this test rip.”	“Nervous test rip”	[“nervous”, “test”, “rip”]	[“nervous”, “test”, “rip”]	[“nervous”, “test”, “rip”]	[“nervous”, “test”, “rip”]	nervous test rip
11009	Anger	“I really wish my road wasn’t so damn bad lmaoooo”	“Really wish the road wasn’t damn bad lmaoooo”	[“really”, “wish”, “road”, “wasn’t”, “damn”, “bad”, “lmaoooo”]	[“really”, “wish”, “road”, “wasn’t”, “damn”, “bad”, “lmaoooo”]	[“really”, “wish”, “road”, “wasn’t”, “damn”, “bad”, “lmaoooo”]	[“really”, “wish”, “road”, “wasn’t”, “damn”, “bad”, “lmaoooo”]	really wish the road wasn’t damn bad lmaoooo
7662	Disgust	“I felt sick seeing how animals are mistreated”	“Felt sick seeing animals mistreated”	[“felt”, “sick”, “seeing”, “animals”, “mistreated”]	[“felt”, “sick”, “seeing”, “animal”, “mistreated”]	[“felt”, “sick”, “see”, “animal”, “mistreated”]	[“felt”, “sick”, “see”, “animal”, “mistreated”]	felt sick see animal mistreated

Emotion Recognition Framework Based on Text

The purpose of AGT (Artificial Gorilla Troops) is to dynamically optimize parameters based on evolutionary ideas, enhancing adaptability in emotion recognition. MGSVM (Modified Gaussian Support Vector Machine) to offer robustness in dealing with textual statistics, enhancing accuracy and reliability in emotion recognition frameworks.

Artificial Gorilla Troops (AGT)

A troop consisting of five to thirty individuals is led by each silverback gorilla. For the optimizing process, the AGT algorithm employs various techniques. Three distinct mechanisms migration to a known location, migration to an unknown location, and approaching other gorillas are employed in the exploration phase of gorilla behavior. These three pathways are selected using a broad methodology. p is a parameter that determines which migratory method to utilize at an unknown location. When $\text{rand} < p$, the first mechanism is selected. The motion technique toward other gorillas is chosen if $\text{rand} \geq 0.5$. The migration strategy to a known location is chosen if rand is less than 0.5. The AGT algorithm benefits greatly from each process based on the mechanisms in use. The method can efficiently control the entire problem space to the first process, its heuristic efficiency is improved by the second mechanism, and its ability to escape from local excellent places is strengthened by the third technique. The three processes that were employed during the investigation phase are simulated using equation (1).

$$\begin{aligned} HW(s+1) &= \{(VA - KA) \times q_1 + KA \\ &\text{if } \text{rand} < 0 (q_2 - D) \times W_q(s) + K \times G \\ &\text{if } \text{rand} \geq 0.5 W(j) - K \times [K \times (W(s) - HW_q(s)) + q_3 \times (W(s) - HW_q(s))]\} \text{ if } \text{rand} < 0.5 \end{aligned} \quad (1)$$

At every iteration, the gorilla's location component $s+1$ is denoted by $HW(s+1)$. The gorilla's present position vector is denoted by $W(s)$. Moreover, each iteration updates each of the random numbers rand , q_1 , q_2 , and q_3 in the interval $[0,1]$. Prior to the optimization process, the value of the parameter 0 needs to be supplied within the interval $[0,1]$. The likelihood of selecting the migration technique to a foreign location is determined by this variable. The maximum and minimum choice variables are denoted by VA and KA , accordingly. HW_{qls} one of the randomly chosen gorilla candidate position vectors including the most recent locations at each stage, and W_q is a single of the randomly chosen components of the gorilla troop representing the total populations. In the end D , K , and G are determined using equations (2), (3), and (4), respectively.

$$D = E \times \left(1 - \frac{Js}{\text{MaxJs}}\right) \quad (2)$$

$$E = \text{Cos}(2 \times q_4) + 1 \quad (3)$$

$$K = D \times k \quad (4)$$

The total number of the present variation is Max in equation (2). The value of E is derived from equation (3) and represents the total number of iterations required to complete the process of optimizing. Cos is the cosine function in this formula, while q_4 is a random number that is updated during each iteration and falls between 0 and 1. The optimized values in equation (2), when first formed, are created with abrupt changes in a big interval but, in subsequent iterations, this interval is reduced. Equation (4), where l is any value that occurs in the interval $[1, 1]$, is used to find K . To model management in silverback gorillas, utilize equation (4). In the real world, a silverback gorilla's lack of experience in early management stages may cause it to make poor decisions when searching for food or managing the troop. But it does acquire sufficient experience to take the lead. Furthermore, Y is computed using equation (6), which has a random value in the problem's parameters and is described in the interval $[-D,D]$, in equation (5), and G is determined using equation (5).

$$G = Y \times W(s) \quad (5)$$

$$Y = [-D, D] \quad (6)$$

When the exploration stage concludes, all HW solutions have their cost function emotion recognition based on text evaluated. If the objective function is $HW(s) < Y(s)$, then the value of $HW(s)$ is used instead of $W(s)$. As a result, the stage's best solution is sometimes referred to as a silverback. Using the value of D in equation (2), one of the two mechanisms in the exploitation stage is the option to follow the silverback gorilla and compete with adult female gorillas. If $D < X$, the competition for mature female gorillas is over; if $D > X$, the process of following the silverback rhino is selected. It is necessary to modify the value of parameter X before beginning the optimizing process. Relation (7) is employed to model the mechanism of following the silverback.

$$HW(s+1)K \times N \times [W(s) - W_{\text{silverback}}] + W(s) \quad (7)$$

$$N = \left[\left| \frac{1}{M} \sum_{j=1}^M HW_j(s) \right|^h \right]^{\frac{1}{h}} \quad (8)$$

$$h = 2^K \quad (9)$$

Equation (7) is used to mimic the Silverback's tracking system. The gorilla's movement vector in equation (7), denoted by $W(s)$, and the silverback gorilla position vector (best solution), denoted by $W_{silverback}$. Furthermore, N is computed by equation (8), and K by equation (4). The position vector of each gorilla candidate in the repetition s is represented by $HW_j(s)$ in equation (8). Here M is the entire population of gorillas in the world. Moreover, h is computed utilizing equation (9). In the exploitative phase, the strategy of competition for adult female gorillas is selected if $D < X$. Whenever the juvenile gorillas evenly reach adolescence, they engage in violent rivalry involving other males to choose the adult female gorilla, which spreads across their group. The above scenario is simulated using equation (10).

$$HW(j) = W_{silverback} - (W_{silverback} \times R - W(s) \times R) \times B \quad (10)$$

$$R = 2 \times q_5 - 1 \quad (11)$$

$$B = \beta \quad (12)$$

$$F = \{M_1 \text{ if } rand \geq 0.5, M_2 \text{ if } rand < 0.5\} \quad (13)$$

The optimal solution's location vector, $Y_{silverback}$, is found in equation (10), where $W(s)$ is the gorilla's current location vector. In order to replicate the impact force, R is computed using equation (11). The random values in the interval $[0, 1]$ represent q_5 in problem (11). Equation (12), which is used for calculating the parameter vector B , is used to measure the degree of violence in conflicts. E is defined from equation (13) and is used to mimic the effect of aggression on the physical dimensions of the alternatives. In equation (12), B is a parameter that should be determined before the optimization process operation. In the event where $rand$ is less than 0.5, then F is equivalent to a random value in the normal distributions; otherwise, it is equal to any value that occurs in both the problematic parameters and the normal distribution's value. Another randomly generated value in the interval $[0, 1]$ is $rand$. The collective creation function is carried out afterward the stage of exploitation. During this process, the expenses of each HW solution are estimated based on the cost function of $HW(s)$ remedies.

Modified Gaussian support vector machine (MGSVM)

The capacity to automatically translate non-linearly separable data points into an additional dimension where they are separated by linearity is support vector machines. Including cost in the mapping of the data points to higher dimensions. Greater dimensions translate into larger vector calculations, which increase memory needs and computation time. Fortunately, multidimensional matrices do not require explicit storage in MGSVM. They are merely utilized to store inner products after mapping the input data into the higher dimensions.

Various mappings are offered by various kernel functions. Inevitably, there isn't a perfect kernel to choose. There are benefits and drawbacks to every kernel. The accuracy of the SVM's classification of data points is significantly impacted by the kernel characteristic selection.

We adapt the current Gaussian foundation to suit our needs. Efficiency is improved over the original Gaussian kernel by using this improved kernel. Each observation in the input space Q is mapped into the characteristics space E by a linear SVM. An embedding of T into E is defined by its transformation as a curve sub manifold. In the highlighted space, indicate the transformed examples of S . The tiny vector dx has been assigned to:

$$\phi(dx) = \nabla \phi \cdot dx = \sum_j \frac{\partial}{\partial x^{(j)}} \phi(w) dx^{(j)} \quad (14) \quad ds^2 = |\phi(dx)|^2 = \sum_{j,i} h_{ji}(w) dx^{(j)} dx^{(i)} \quad (15)$$

$$h_{ji}(w) = \left(\frac{\partial}{\partial x^{(j)}} \phi(w) \right) \cdot \left(\frac{\partial}{\partial x^{(i)}} \phi(w) \right) \quad (16)$$

The dot represents the total across the index. The Riemann measurement vector that develops in is the positive-definite matrices.

$$h_{ji}(w) = \frac{\partial}{\partial x^{(j)}} \frac{\partial}{\partial x^{(i)}} L(w, w_j) \quad (17)$$

To enhance the SVM's effectiveness, we can raise the margin or the distances (ds) between classes. It enables to elevate the L Riemannian metrics tensor around the border and decrease it around additional samples. We can alter the kernel L so that it is big around the border in light of equation (17).

Kernel modification: according to the Riemannian geometric framework, let's say the operating system can be changed to:

$$L(w, w_j) = o(w) o(w_j) L(w, w_j) \quad (18)$$

Is known as a kernel's geodesic modification by value $o(w)$.

We consider the Gaussian kernel, which is the base function utilized in SVM, i.e.:

$$L(w, w_j) = \exp(-\|w - w_j\|^2 / \sigma^2) \quad (19)$$

The argument in this case is kernel width. The related Riemannian measured vector is shown to be transformed into the following categories:

$$h_{ji}(w) = \frac{1}{\sigma^2} h_{ji} \quad (20)$$

Following this Riemannian metric vector modification, it becomes:

$$\tilde{h}_{ji}(w) = o_j(w) o_i(w) + \sigma^2(w) h_{ji}(w) \quad (21)$$

To guarantee that $o(w)$ has a significant value surrounding a support vector (SV), its Gaussian kernel's geodesic modification:

$$o_j(w) = \frac{\partial o(w)}{\partial w_j} \quad (22)$$

For maximum $o(w)$ the value of $o_j(w) = 0$

To guarantee big numbers at the support vector's reference points for $o(w)$, it can be generated in a manner that depends on the information at hand, as follows:

$$o(w) = \sum_{j \in TV} \alpha_j \exp(-\|w - w_j\|^2 / 2\tau^2) \quad (23)$$

Where τ the summation is applied across all support vectors and $o_j(w)$ is a free parameter. It is obvious that when w is near the vectors of support, $o_j(w)$ and $o(w)$ are huge, while individuals are tiny when y is far from SVs. Consequently, the $h_{ji}(w)$ around vectors of support increases when y is near those vectors. As a result, the depth of coverage near the border is increased, and SVM's categorization power increases.

AGT-MGSVM

The "Artificial Gorilla Soldiers (AGT)" and "Modified Gaussian Support Vector Machine (MGSVM)" methods are integrated in the proposed "Text-Based Intelligence-Driven Emotion Recognition Framework". AGT dynamically modifies MGSVM parameters. This framework enhances emotion detection from textual content, promises to improve accuracy and performance in human-computer interaction, emotion analysis, and psychiatric diagnostic applications, and overcomes the limitations of traditional methods on cultural nuances and linguistic development.

RESULTS

In this section, we used ensemble methodologies to improve the accuracy of the results assessed with conventional techniques such as backpropagation neural (BPN), Bidirectional long-term short-term memory (BiLSTM)⁽¹⁸⁾, and Convolutional Neural Network (CNN) + Bidirectional Gated Recurrent Unit (Bi-GRU)- Support Vector Machine (SVM).⁽¹⁹⁾ A novel emotion recognition framework named Artificial Gorilla Troops driven Modified Gaussian support vector machine (AGT-MGSVM) method is a promising improvement in intelligence-driven emotion recognition from text. The study evaluated the suggested approach with other traditional techniques using several metrics, such as F1-score, Recall, Precision, and Accuracy. The results showed that the suggested strategy outperformed other conventional methods based on these factors.

Accuracy

The percentage of correctly estimated samples among the actual cases is measured by the accuracy metric. Figure 2 and table 2 show the accuracy performance of our suggested AGT-MGSVM approach in comparison to the conventional methodology. The accuracy level of existing methods BPN, BiLSTM, and CNN + Bi-GRU-SVM achieved 71,27 %, 87,66 %, and 80,11 %, respectively. The proposed method (AGT-MGSVM) achieved 89,9 % accuracy. The performance of our suggested method in textual emotion recognition is better than that of the traditional approaches.

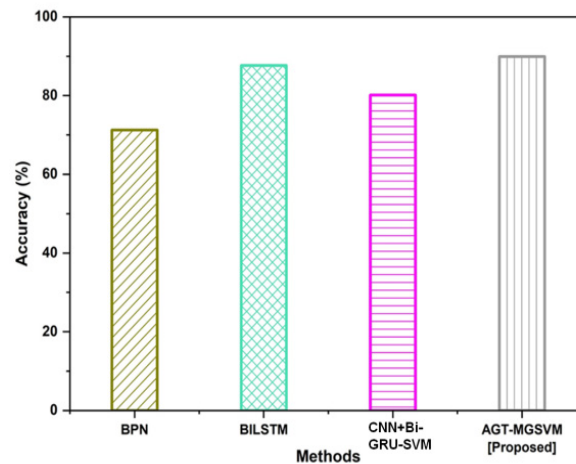


Figure 2. Output of Accuracy

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
BPN	71,27	67	66	67
BiLSTM	87,66	87,66	87,66	87,66
CNN+Bi-GRU-SVM	80,11	82,39	80,4	81,27
AGT-MGSVM [Proposed]	89,9	89,5	89,2	89,4

Precision

The precision and dependability with which text-based may be utilized to figure out known as the precision of emotion recognition followed by figure 3 and table 2 show the precision result. The precision of the traditional techniques, CNN+Bi-GRU-SVM, BiLSTM, and BPN, was 82,39 %, 87,66 %, and 67 %, respectively. The proposed AGT-MGSVM approach achieved an 89,5 % precision result when compared to the conventional methods. This shows our proposed method delivers outperformance in text-based emotion recognition.

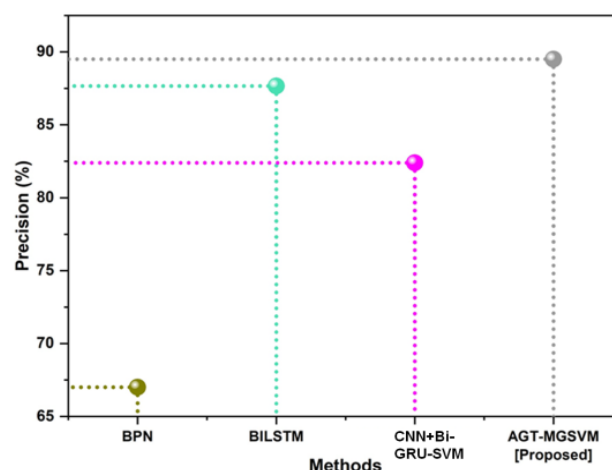


Figure 3. Performance of precision

Recall

The recall of a classification model is an efficiency metric that assesses how successfully the model discovers relevant samples inside a particular class. It is calculated as the ratio of positive forecasts to the sum of FN and TP. Table 2 and figure 4 show the recall procedure. The traditional techniques, CNN+Bi-GRU-SVM, BiLSTM, and BPN, all achieved recall levels of 80,40 %, 87,66 %, and 66 %, respectively. The suggested AGT-MGSVM approach produced an 89,2 % recall rate when compared to the conventional approaches. Our suggested strategy performs better in textual emotion recognition than the conventional approach.

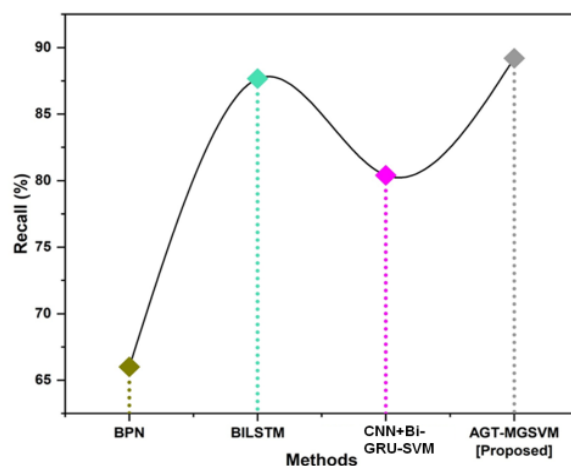


Figure 4. Output of Recall

F1 Score

The F1 score is a statistical metric that is commonly used to evaluate the efficacy of a text-based emotion identification framework or categorization model. The results of the F1 Score are presented in table 2 and figure 5. The traditional techniques, CNN +Bi-GRU-SVM, BiLSTM, and BPN, each obtained an F1 Score level of 81,27 %, 87,66 %, and 67 % respectively. Our suggested AGT-MGSVM approach achieved 89,4 % of the F1 Score outcome when compared to the conventional approaches.

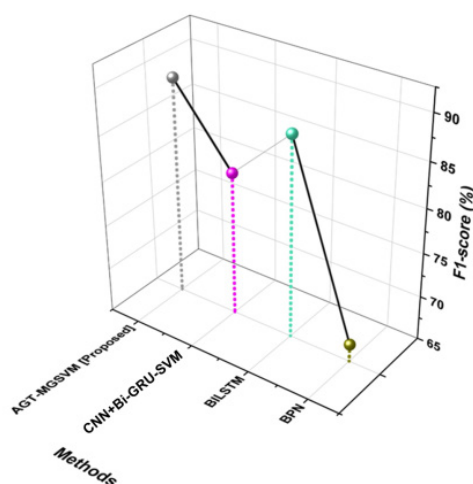


Figure 5. Outcome of F1-Score

DISCUSSION

The BPN for emotion recognition performed well in terms of precision (67 %), F1-score (67 %), recall (66 %), and accuracy (71,27 %). Long-range relationships in the text might be difficult for BPN to comprehend, resulting in less than ideal performance on tests involving the perception of emotions. Furthermore, because BPN does not have temporal information, it is less useful for sequential data, like text, where word relationships and order are essential for accurately identifying emotions.⁽¹⁷⁾

The findings demonstrate that, in comparison to the previous research, the Bi-LSTM model yielded better outcomes in terms of enhanced recall (87,66 %), precision (87,66 %), F1-score (87,66 %), and accuracy (87,66 %). Although BiLSTM is quite good at capturing contextual data in both directions, during training it can be computationally costly and slow, especially for large sequences. When analyzing lengthy texts, it may also be exploding problems with gradient disappearing or overstating. Moreover, BiLSTM's efficacy in practical applications for recognizing emotions may be impacted by its inconsistent ability to generalize well to new information.⁽¹⁸⁾

With a F1 score of 81,27, recall of 80,40, precision of 82,39, and accuracy of 80,11 %, the hybrid CNN+Bi-GRU-SVM algorithm has demonstrated its efficacy. Enhancing feature extraction and sequence estimation with CNNs in combination with Bi-GRU and SVM presents an intricate and computationally expensive hybrid

technique that requires longer training periods. Hyperparameter tuning may become more difficult as a result of an optimization procedure being more complex due to the combination of numerous models.⁽¹⁹⁾

In comparison with existing methods, our suggested AGT-MGSVM approach has several benefits. By using strong feature extraction and noise reduction techniques, it improves classification accuracy and optimizes the process by utilizing the combined intelligence of artificial gorilla troops. Effective emotional state distinction is made possible by the modified Gaussian kernel's improved ability to handle non-linear relationships in data. It is also appropriate for a variety of text-based applications due to its versatility. Through comparison with other traditional approaches, our proposed AGT-MGSVM demonstrate better performance in terms of recall (89,2 %), precision (89,5 %), F1-score (89,4 %), and accuracy (89,9 %).

CONCLUSIONS

The proposed emotion recognition framework, Artificial Gorilla Troops driven Modified Gaussian Support Vector Machine (AGT-MGSVM), offers a significant improvement in the field of textual emotion recognition. The suggested strategy improves the precision and dependability of emotion recognition from text by leveraging the ISEAR dataset and sophisticated natural language processing techniques. Our suggested model was implemented in Python software. The suggested AGT-MGSVM was a strong and reliable structure for text-based emotion recognition, improving its accuracy and dependability. The AGT-MGSVM framework for emotion recognition shows promise but faces limitations like dataset dependency, computational complexity, and handling cultural nuances. Future work could enhance dataset diversity, enable real-time detection, integrate multimodal data, and personalize models for improved accuracy and applicability across different contexts and domains.

BIBLIOGRAPHIC REFERENCES

1. Erenel Z, Adegbeye OR, Kusetogullari H. A new feature selection scheme for emotion recognition from text. *Applied Sciences*. 2020 Aug 3;10(15):5351. <https://doi.org/10.3390/app10155351>
2. Hodel TM, Schoch DS, Schneider C, Purcell J. General models for handwritten text recognition: Feasibility and state-of-the art. *german kurrent as an example. Journal of open humanities data*. 2021 Jul 9;7(13):1-0. <http://dx.doi.org/10.5334/johd.46>
3. Vidal E, Toselli AH, Ríos-Vila A, Calvo-Zaragoza J. End-to-end page-level assessment of handwritten text recognition. *Pattern Recognition*. 2023 Oct 1;142:109695. <https://doi.org/10.1016/j.patcog.2023.109695>
4. Xue X, Feng J, Sun X. Semantic-enhanced sequential modeling for personality trait recognition from texts. *Applied Intelligence*. 2021 Nov 1:1-3. <https://doi.org/10.1007/s10489-021-02277-7>
5. Saxena A, Khanna A, Gupta D. Emotion recognition and detection methods: A comprehensive survey. *Journal of Artificial Intelligence and Systems*. 2020 Feb 7;2(1):53-79. <https://doi.org/10.33969/AIS.2020.21005>
6. Acheampong FA, Wenyu C, Nunoo-Mensah H. Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*. 2020 Jul;2(7):e12189. <https://doi.org/10.1002/eng2.12189>
7. Deng S, Wang G, Wang H, Chang F. An Artificial-Intelligence-Driven Spanish Poetry Classification Framework. *Big Data and Cognitive Computing*. 2023 Dec 14;7(4):183. <https://doi.org/10.3390/bdcc7040183>
8. Yin G. Intelligent framework for social robots based on artificial intelligence-driven mobile edge computing. *Computers & Electrical Engineering*. 2021 Dec 1;96:107616. <https://doi.org/10.1016/j.compeleceng.2021.107616>
9. Nasir AF, Nee ES, Choong CS, Ghani AS, Majeed AP, Adam A, Furqan M. Text-based emotion prediction system using machine learning approach. *InOP Conference Series: Materials Science and Engineering 2020 Feb 1 (Vol. 769, No. 1, p. 012022)*. IOP Publishing. 10.1088/1757-899X/769/1/012022
10. Yang H. Application of PNN-HMM model based on emotion-speech combination in broadcast intelligent communication analysis. *IEEE Access*. 2023 Aug 2. <https://doi.org/10.1109/ACCESS.2023.3301127>
11. Prabu S, Abraham Sundar KJ. Enhanced Attention-Based Encoder-Decoder Framework for Text Recognition. *Intelligent Automation & Soft Computing*. 2023 Feb 1;35(2). 10.32604/iasc.2023.029105
12. Qiao Z, Zhou Y, Yang D, Zhou Y, Wang W. Seed: Semantics enhanced encoder-decoder framework for scene text recognition. *InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2020 (pp. 13528-13537)*.

13. Chernyshova YS, Sheshkus AV, Arlazarov VV. Two-step CNN framework for text line recognition in camera-captured images. *IEEE Access*. 2020 Feb 14;8:32587-600. <https://doi.org/10.1109/ACCESS.2020.2974051>
14. Yu D, Li X, Zhang C, Liu T, Han J, Liu J, Ding E. Towards accurate scene text recognition with semantic reasoning networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2020* (pp. 12113-12122).
15. Yu W, Ibrayim M, Hamdulla A. Scene text recognition based on improved CRNN. *Information*. 2023 Jun 28;14(7):369. <https://doi.org/10.3390/info14070369>
16. Mohd M, Qamar F, Al-Sheikh I, Salah R. Quranic optical text recognition using deep learning models. *IEEE Access*. 2021 Mar 4;9:38318-30. <https://doi.org/10.1109/ACCESS.2021.3064019>
17. <https://paperswithcode.com/dataset/isear>
18. Asghar MZ, Lajis A, Alam MM, Rahmat MK, Nasir HM, Ahmad H, Al-Rakhami MS, Al-Amri A, Albogamy FR. A deep neural network model for the detection and classification of emotions from textual content. *Complexity*. 2022;2022(1):8221121. <https://doi.org/10.1155/2022/8221121>
19. Bharti SK, Varadhaganapathy S, Gupta RK, Shukla PK, Bouye M, Hingaa SK, Mahmoud A. Text-Based Emotion Recognition Using Deep Learning Approach. *Computational Intelligence and Neuroscience*. 2022;2022(1):2645381. <https://doi.org/10.1155/2022/2645381>

FINANCING

The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Methodology: Hanyu Lu.

Drafting - original draft: Xiaoping Wu.

Writing - proofreading and editing: Xiaoping Wu.

ANNEXES

GA-SVM	Genetic Algorithm-Support Vector Machine	DT	Decision Tree
SVM	Support Vector Machine	TPS	Thin-Plate Spline
TF-IDF	Term Frequency-Inverse Document Frequency	PNN-HMM	Probabilistic Neural Network with Hidden Markov Models
ASTER	Attentional Scene Text Recognizer	DSAM	Deformable Spatial Attention Module
SAM	Spatial Attention Module	KNN	k-nearest Neighbors
MNB	Multinomial Naïve Bayes	ANN	Artificial neural networks
CRNN	Convolutional Recurrent Neural Network	GSRM	Global semantic reasoning module
OCR	Quranic optical character recognition	CNN	Convolutional neural network
FCN	Fully convolutional networks	FN	False negatives
RNN	Recurrent neural network	SRN	Semantic reasoning network
TP	True positives	CRR	Character recognition rate
CRR	Word recognition rate (WRR)	AI	Artificial intelligence