International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Emotion Detection from Text using Natural Language Processing and Neural Networks

[1]S. Arun Kumar S., [2]A. Geetha

Submitted: 08/12/2023 **Revised**: 19/01/2024 **Accepted**: 29/01/2024

Abstract: Emotion may be shown in a variety of ways, including voice, written texts, and facial expressions and movements. Emotion detection in text is essentially a content-based classification challenge that combines concepts from natural language processing and machine learning. This paper addresses textual data-based emotion identification algorithms and emotion detection.

Keywords: *emotion word ontology, human-computer interaction, textual emotion detection.*

1. Introduction

The majority of literary expressions come from interpreting the meaning of concepts and how those concepts interact with one another in addition to from the usage of emotive language. Analysing a person's emotional state through a text document they have written can be challenging, but it is also important in many circumstances. In the human-computer interface, it is crucial to recognise the text's emotional content [1]. Speech-based, facial-based, and text-based emotions, respectively, refer to the ways in which emotions can be represented by a human being via written language, spoken words, and facial expressions. Academics still need to pay attention to text-based emotion detection algorithms notwithstanding the fact that enough work has been done to recognise emotions in speech and on the face[2]. From an application standpoint, it is becoming more and more crucial in computational linguistics to identify human emotions in text.

Happy, sad, angry, surprised, fearful, and so on are emotions. Ekman [3] divides emotions into six separate categories, including joy, sadness, fear, surprise, rage, and disgust. Generally speaking, a person's voice, written words, gestures, and facial expressions all convey their mood and feelings [4]. Unlike facial expression and speech recognition, a written statement lacks taste and so loses its ability to identify itself. Because of the text's intricacy and ambiguity, identifying its emotions is a difficult task. It is hard to tell the mood of a given text since words might have many meanings and morphological forms.

One of the key benefits of human-machine contact is the ability for a nonliving object to sense or feel emotions similar to those of a human. Since tasteless text sentences lack any tone or expression, our suggested approach is able to identify emotions from them. A single dataset has been the subject of many studies. For the purpose of detecting emotions, however, we have worked with three datasets that include text in the form of comments, basic words, and dialogues. We can use any technology to implement our text-based emotion recognition model. This model's business potential includes the ability to identify feelings in consumer reviews of goods and services, as well as to give social media users security.

The remainder of the text is structured as follows: Section 2 presents the overview of the literature on emotion detection, and Section 3 provides details on the suggested method. In Section 4, findings and an analysis of the recommended tasks are provided. Section 5 provides the article's concluding recommendations.

2. Existing System

A crucial aspect of human existence is emotion. These feelings affect how people make decisions and improve how we express ourselves to others. The process of determining a person's various feelings or emotions (such joy, sorrow, or fury); also known as emotion detection or recognition. Researchers have worked hard in the last few years to automate the recognition of emotions. However, several bodily behaviours, hand trembling and voice pitch, can also be used to infer an individual's emotional state [5]. However, it can be challenging to discern emotions in text. The difficulty of detecting emotions in text is further increased by the numerous ambiguities and new lingo or terminology that are being used every day. Additionally, emotion recognition tends to go beyond simply detecting the basic psychological states (happiness, sadness, and anger).

[1] Research Scholar, Department Of Computer and Information Science, Annamalai University, Annamalai Nagar, Tamil Nadu, India
ORCID ID: 0000-0002-8222-6040
[2]Professor, Department Of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Tamil Nadu, India
ORCID ID:0000-0001-6670-0374
** Corresponding Author Email: arunkumar@apec.edu.in*

3. Proposed System

3 different systems are going to use in the proposed system. They are:

i) Emotion text identified with NRCLex method

ii) Emotion text identified with Neural Network method

iii) Emtoion text identified with NLP method

3.1 Emotion Text Identified With NRCLex Method

Social networking sites have become a vital tool for communicating feelings to people all over the world as a result of the Internet's phenomenal growth. Many people use writing, pictures, music, and videos to voice their viewpoints or ideas. In this study, we used NRCLex to analyse social media whatsapp status in terms of positive, negative, and neutral sentiment and other emotional categories like happiness, rage, melancholy. Stop words, unnecessary punctuation, and other distractions are removed during the pre-processing of the text data. The NRClex lexicon is then used to map the remaining words to their corresponding emotional and sentiment ratings. The analysis' results can be utilised to discover more about the opinions and preferences of social media users. Nevertheless, by providing insightful analyses of the sentiment and emotions reflected in text data, sentiment and emotion analysis using NRCLex may aid in guiding decision-making across a range of industries. In order to fit specific use cases, NRCLex additionally offers techniques for altering the emotion categories and score system. A strong and adaptable toolbox for the semantic and emotional analysis of text data is NRCLex. It is especially useful for applications like social media monitoring, market research, and customer feedback analysis are examples of text analysis.

The NRCLex has been applied as the baseline technique to calculate emotion scores [6]. It creates a list of English words together with their correlations with two negative and two good sentiments, as well as eight fundamental emotions (anger, fear, trust, surprise, sadness, joy, and disgust). In this approach, crowd sourcing was used to manually annotate words. With almost 27,000 entries, this vocabulary is derived from the National Research Council of Canada (NRC) impact lexicon and the Word Net synonym sets from the NLTK library. The scores for a specific word w and emotion e range from 0 to 1. The maximum level of emotion is represented by a score of 1, while the lowest level of emotion is represented by a score of 0. Sentiment-bearing words are identified in a phrase and given ratings for each emotion category using a 4-tuple (best to worst) scaling method. The resulting scores are then normalised between 0 and 1.0. Four emotional scores—fear, anger, sadness, and surprise—were computed for this investigation. Please see [7,8] for more information on the NRCLex in full.

A well-liked tool for text emotion recognition is the NRC Emotion Lexicon (NRC-LEX). Eight fundamental emotions are listed, including happiness, sadness, anger, fear, trust, disgust, surprise, and anticipation.

3.2 Emotion Text Identified with Neural Network Method

Social networking is become a part of everyday life. Internet use increases as social media usage grows because more people are participating simultaneously [9,10]. Human expressions are greatly influenced by emotion. Emotions can be detected in a person's face, speech, and written words.

The limited attempts to use deep learning for emotion recognition are discussed in the sections that follow, but we discover that there are few real performance reviews. Aggression is predicted using CNN with a sliding window and subsequent max-pooling.

Emotional expression was achieved using NLP [11]. Since the network is designed to handle only one dimension, it is unclear how well this approach can generalise to challenges involving regression or even multi-class predictions, which are common in dimensional emotion models. This approach so has several shortcomings. Although using a deep network, the method's network architecture limits the amount of texts it can process, like typical machine learning. This is where it varies from recurrent networks, which can handle texts of any size since they move over sequences.

For the purpose of recognising emotions, numerous recurrent neural network (RNN) techniques have been developed. Many supervised classification algorithms for emotions have been developed utilising data collected from microblogs like Twitter, using hashtags or emoticons as the emotion label for the data. This is because there aren't many emotion-labelled datasets available. The most advanced techniques for identifying fine-grained emotions are shown on the Gated Recurrent Unit (GRU) network [12]. The researchers extended the categorization to eight important emotional elements discovered in the psychology theory of emotion after first creating a large dataset for automatic emotion identification from Twitter.

Other combination architectures have recently been put up for tasks involving text-based emotion recognition. One example is the introduction of the mixed recurrent convolutional neural network (RCNN) approach [14] and has shown results that are competitive with well-tuned contextual and emotional word embeddings [15], [16], and [17].

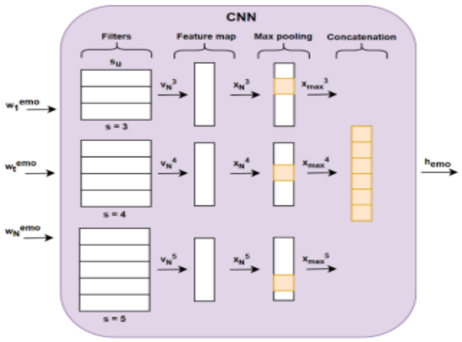

Fig 1. CNN Architecture

3.3 Emotion Text Identified With NLP Method

Natural language processing (NLP) consists of two parts: producing and understanding human language. The former is more difficult since natural language contains ambiguity. The former is more challenging since natural language contains uncertainty.

Speech recognition, machine translation, question answering, document summarization, speech synthesis, and many other uses of natural language processing (NLP) are available [18]. The two primary components of natural language processing are sentiment analysis and emotion identification. These two names are not the same, even though they are sometimes used synonymously. You may ascertain if a piece of sentiment analysis is used, data may be positive, negative, or neutral.

Either positive or negative data exist. On another hand, a technique for recognising certain human emotion types is called emotion detection including happiness, sadness, or anger. On sometimes, the phrases "emotion detection," "affective computing," the terms "emotion identification" and "emotion analysis" are interchangeable [19]. Individuals are using social media to communicate their emotions. as a result of improved Internet access. Individuals freely discuss their disagreements, ideas, and opinions on a variety of topics on social media platforms that are either neutral or active[20], [21], and [22].

Context extraction from text is one of the most amazing advances in natural language processing (NLP). Originally, context extraction was meant to reveal the sentiment polarity in text, but the globe advanced and began to recognise sentiment as feelings a few years ago. Both of these ideas are pretty dissimilar [23], [24] and [25]. Sentiment can be good, negative, or neutral, whereas emotions are fall between positive and negative. An mood that is cheerful, joyous, exciting, or even funny can be characterised as having a positive sentiment. Similar to how sadness, disgust, and fury make people feel bad. A number of machine learning techniques used to train emotion detection models, and studies a few years ago employed a lexical approach to emotion recognition based on lexicons as collections of emotional phrases common for diverse emotions[26]. All of these methods are rapidly getting obsolete due to the recent advancements in deep learning detection models, which are capable of performing an extremely accurate automatic interpretation of emotions from a text.

It appears difficult to infer a person's emotional state from their textual communication. In human-computer interaction (HCI), recognising the text's emotions is crucial. Speech, text-based, and facial emotions are the methods in which an individual can express their emotions verbally. Because enough research has been done on spoken and facial emotion recognition, a text-based emotional identification system also needs to draw in academics [27].

From an application standpoint, it is becoming increasingly important to understand how to detect human emotions in text using computational linguistics. Text emotion detection scans the writer's input text for emotions and assesses it. The premise behind this is that upbeat individuals would say encouraging things. These expressions may refer to the underlying unpleasant emotions of someone who is anxious, irritated, or unhappy.

It is essential that the text be able to identify human emotions since it is the primary means of communication between people and computers in chat rooms, forums, websites, blogs, product evaluations, and other social media sites like YouTube, Twitter, and Facebook. It uses the importance of the data it produces for sentiment analysis, user profiling, and a subfield of natural language processing, has benefited. Specifically, sentiment analysis is to classify relevant aspect of the statement according to three opposites: positive, neutral, or negative[28].

Word embeddings are frequently utilised in NLP applications because the vector representations of words enable deep learning techniques to extract useful semantic components and linguistic relationship between words [29].

4. Implementation

4.1 Software

Installing Python 3.5, Jupyter Notebook v0.27.0, and the Anaconda library set up the deep learning environment. The Python 3.5 interpreter has been successfully coupled with the OpenCV (version 3.3.0) library. Libraries like tensorflow, keras, pandas, stopwords, and seaborn are imported with the anaconda install command.

4.2 Dataset Used

The emotion_dataset_raw1.csv is used for the training process and testing process in this proposed work. The Dataset for training and testing constitute the Comments from Website, Sentences, words and Whatsapp status lines.

All two data sets—text sentences, and tweets are integrated in this study. More than 27000 words used in the pooled dataset. According to their syntactic and semantic polarity, the six types of emotions that can be associated with a text sentence are: happiness, disgust, fear, surprise, angry, and sadness. The content is written in English and includes some extra punctuation and emojis. Only text sentences and the emotions they indicate are included in the dataset. Each dataset is split into two categories: training and testing, with an 80:20 ratio between the two.

The steps that would generally take to do emotion detection using the NRC-LEX approach are as follows:

4.2.1 Text Preparation

The text data should be cleaned. Stop words, punctuation, and other distracting elements might need to be eliminated. The data should be broken text into distinct words or tokens by tokenizing it.

4.2.2 Emotional Rating

Check the NRC-LEX's associations for each word in your text with each emotion. Each word for each emotion in the lexicon is given a binary score (1 for presence, 0 for absence).

4.2.3 Aggregation

Calculate the overall emotion score for the entire text by adding the ratings for each emotion.

Setting a threshold will help you identify the main feeling. Consider classifying the text as expressing joy if the "joy" score is much higher than the others.

4.3 Training Phase

There are five categories for the sentences: neutral, happy, sad, angry, disgusting, and surprised. In NLP, around 24000 training examples are utilised, including neutral (4000), happy (4000), sad (4000), angry (4000), disgust (4000), and surprised (4000).

4.4 Testing Phase

Neutral 1000, happy 1000, sadness 1000, furious 1000, disgust 1000, and fear 1000 are approximately 6000 samples that are verified and evaluated.

5. Experimental Results

3 different systems are proposed in this work. They are:

i) Emotion text identified with NRCLex method

ii) Emotion text identified with Neural Network method

iii) Emotion text identified with NLP method

The Performance evaluation of the proposed work is shown in the Table 1.

Table 1. Performance Evaluation

Sl.No.	Method	Performance Accuracy (%)
1.	NRCLex	64.44
2.	Neural Network	99.56
3.	NLP	83.36

Out of these 3 proposed systems we are considering Emotion text identified with Neural Network method because its performance accuracy is 99.56 whereas NRCLex is 64.44 and NLP is 83.36 only.

6. Screen Shots

The screen shots are the propsed work are shown in figures 2 through 5.

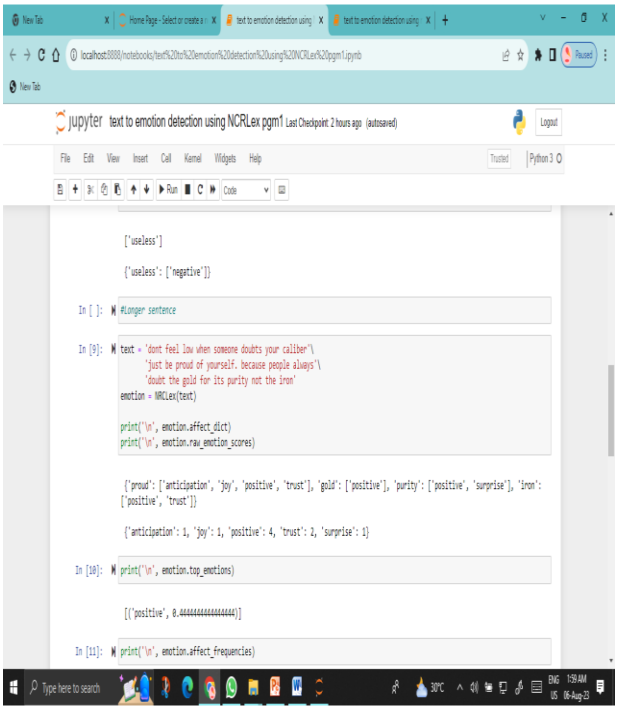

Fig 2. Accuracy of the emotional detection using NRCLex

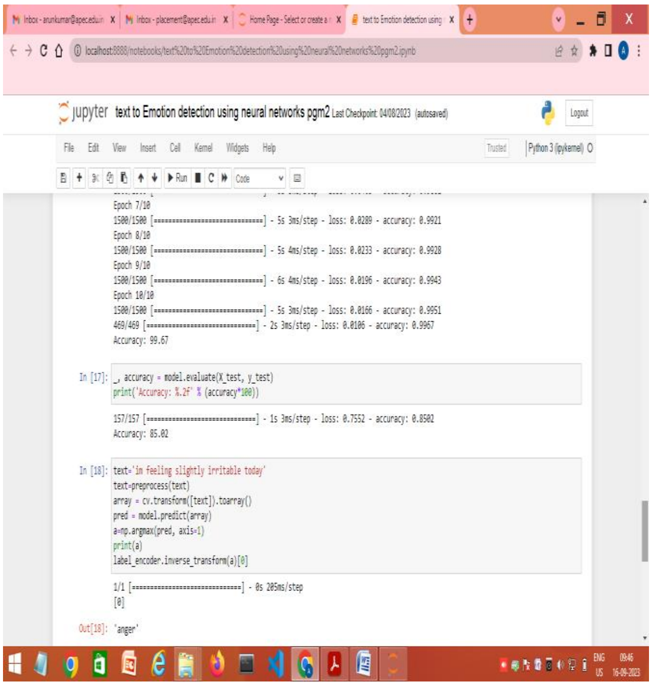

Fig 3. Accuracy of the emotional detection using neural network

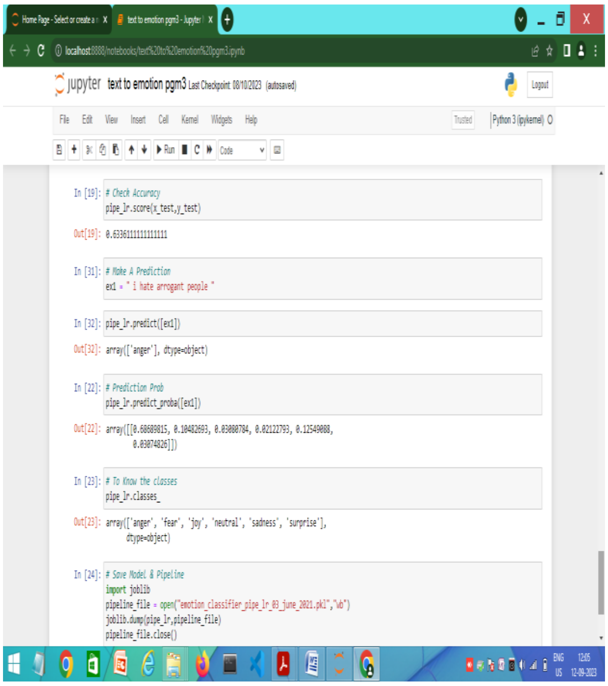

Fig 4. Result of the testing sentence or word classifications

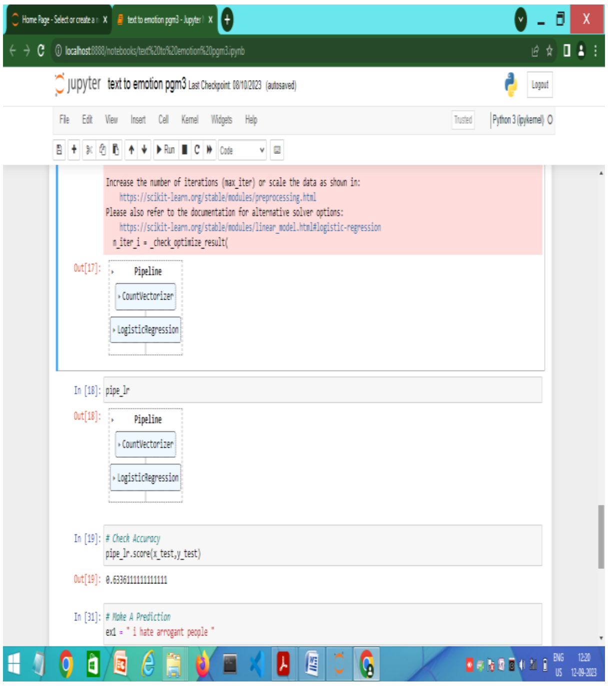

Fig 5. Accuracy of the emotional detection using NLP

7. Conclusion

In this proposed work, the emotion is detected from the input sentence. The suggested work is able to perform competitively in terms of accuracy and efficiency, according to experimental data. Emotion detection in text using NRCLEX shows the accuracy as 64.0%, using NLP getting accuracy is 83.0% and using neural network with an accuracy of 99.0%. According to the results, it is concluded that Neural Network is a promising method for emotion predictions from text. The proposed work may be extended with other deep learning models.

References

[1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, “Emotion Recognition In Human-Computer Interaction,” in IEEE Signal Processing Magazine, vol. 18(1), Jan. 2001, pp. 32-80, doi: 10.1109/79.911197.

[2] Nicu Sebea, Ira Cohenb, Theo Geversa, and Thomas S. Huangc “Multimodal Approaches For Emotion Recognition: A Survey”, USA.

[3] P. Ekman, “Basic Emotions,” Handbook of cognition and emotion, vol. 98, no. 45-60, p. 16, 1999.

[4] R. Plutchik, “E Nature Of Emotions,” American Scientist, vol. 89, no. 4, p. 344, 2001.

[5] Kratzwald B, Ilić S, Kraus M, Feuerriegel S, Prendinger H (2018) “Deep Learning For Affective Computing: Text-Based Emotion Recognition In Decision Support”,. Decis Support Syst 115:24–35

[6] Mohammad S, Turney P. “Crowdsourcing A Word-

Emotion Association Lexicon” Comput Intell. 2013;29(3):436–65.

[7] Mohammad S, Turney P. “Emotions Evoked By Common Words And Phrases: Using Mechanical Turk To Create An Emotion Lexicon”, In-Proc. NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, LA, California; 2010. p. 26–34.

[8] Mohammad S. “Word Afect Intensities” In: Proc. of the 11th International Conference on Language Resources and Evaluation (LREC-2018); 2018. p. 174–183.

[9] D. Gordeev, ‘‘Detecting State Of Aggression In Sentences Using CNN,’’ in Proc. Int. Conf. Speech Comput. Cham, Switzerland: Springer, 2016, pp. 240–245.

[10] M. Abdul-Mageed and L. Ungar, ‘‘Emonet: Fine-Grained Emotion Detection With Gated Recurrent Neural Networks’’ in Proc. 55th Annu. Meeting Assoc. Comput. Linguistics, vol. 4, Jul. 2017, pp. 718–728.

[11] Kim, Y., ‘‘Convolutional Neural Networks For Sentence Classification,’’ Aug. 2014, arXiv:1408.5882.

[12] Mittal P, Garg S, Yadav S. “Social Network Analysis using Interest Mining: A Critical Review, Indian Journal of Science and Technology. 2016”, Apr; 9(16):1−8. Crossref

[13] Vivekananthamoorthy N, Naganathan ER, Rajkumar R. “Determinant Factors On Student Empowerment And Role Of Social Media And Ewom Communication: Multivariate Analysis On Linkedin Usage”, Indian Journal of Science and Technology. 2016 Jul; 9(25):1−12. Crossref

[14] P. Zhong and C. Miao, ‘‘Ntuer At Semeval-2019 Task 3: Emotion Classification With Word And Sentence Representations In RCNN’’ Feb. 2019.

[15] B. Felbo, A. Mislove, A. Søgaard, I. Rahwan, and S. Lehmann, ‘‘Using Millions Of Emoji Occurrences To Learn Any-Domain Representations For Detecting Sentiment, Emotion And Sarcasm’’ , Aug. 2017.

[16] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, ‘‘Deep Contextualized Word Representations,’’ Feb. 2018.

[17] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, ‘‘Bert: Pre-Training Of Deep Bidirectional Transformers For Language Understanding,’’ Oct. 2018.

[18] Itani M, Roast C, Al-Khayatt S (2017), “Developing Resources For Sentiment Analysis Of Informal Arabic Text In Social Media”, Procedia Comput Sci 117:129–136.

[19] Munezero M, Montero CS, Sutinen E, Pajunen J (2014), “Are They Different? Affect, Feeling, Emotion, Sentiment, And Opinion Detection In Text”, IEEE Trans Afect Comput 5(2):101–111.

[20] Shiv Naresh Shivhare. "Emotion Detection from Text" , 'Academy and Industry Research Collaboration Center (AIRCC)' , 2012.

[21] Jia Guo. "Deep learning approach to text analysis for human emotion detection from big data" , Journal of Intelligent Systems, 2022.

[22] Erdenebileg Batbaatar, Meijing Li, Keun Ho Ryu. "Semantic-Emotion Neural Network for Emotion Recognition From Text" , IEEE Access, 2019

[23] Meenu Mathew, Jay Prakash. "Weighted Pooling RoBERTa for Effective Text Emotion Detection" , 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), 2023

[24] Darwin D, Sujitha Juliet. "Sentimental Analysis Based on User Emotions Using Machine Learning" , 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), 2023.

[25] Ravinder Ravinder, Naresh Vurukonda. "Cloud computing: an efficient load balancing and scheduling of task method using a hybrid optimization algorithm" , Indonesian Journal of Electrical Engineering and Computer Science, 2023.

[26] Md Khayrul Bashar. "A Hybrid Approach to Explore Public Sentiments on COVID-19" , SN Computer Science, 2022.

[27] "Progress in Advanced Computing and Intelligent Engineering" , Springer Science and Business Media LLC, 2021.

[28] Ghazanfar Latif, Sherif E. Abdelhamid, Khaled S Fawagreh, Ghassen Ben Brahim, Runna Alghazo. "Machine Learning in Higher Education: Students’ Performance Assessment considering Online Activity Logs" , IEEE Access, 2023.

[29] S Ibrahiem, K Bahnasy, M Morsey, M Aref. "FEATURE EXTRACTION ENCHANCEMENT IN USERS’ ATTITUDE DETECTION", International Journal of Intelligent Computing and Information Sciences, 2018

Authors

S. Arunkumar, M.C.A.,M.Phil.,

Research scholar in Department of Computer and Information Science at Annamalai University. He has been completed his Master of Computer Applications degree in the year 1998 at Annamalai University, Annamalai Nagar, Tamilnadu, India and completed his Master of Philosophy in Computer Science in the year of 2008 at Annamalai University, Chidambaram, Tamilnadu, India. He also published one International journal. His area of research is image processing.

Dr. A. Geetha, M.E., Ph.D.,

She obtained her B.E degree in Computer Science and Engineering from Periyar Maniammai University, Thanjavur in the year 1994. She received her M.E degree in Computer Science from Regional Engineering College, Trichy in the year 1997. She has been working in Annamalai University, since 1999. She was awarded Ph. D in Computer Science and Engineering at Annamalai University in the year 2014. Presently she works as a Professor in the same department. She has published around 40 papers in national & international journals and conferences. She has coordinated a workshop in the field of speech and image processing. An international conference was also organized by her to her credit. Also she was one of the editors for the International Journal of Applied Engineering Research which was published in the international conference. Two book chapters were published by her. She has won the best paper award in National Conference on Multimedia Signal Processing. She has delivered special lectures on several topics in her area, in workshops and staff development programme. Presently she is the coordinator of the Value Added Course, FEAT, and Annamalai University. She has visited and presented papers in international conferences conducted in Malaysia and Japan. Currently, she is a member of the Doctoral Committee in the Annamalai University and Research Advisory Committee at Anna University. Her areas of research interest include pattern classification, image & video processing and human-computer interaction. She has guided 2 PhD Scholars in the area of Human Computer Interaction.