Mobile Machine Learning Models for Emotion and Sarcasm Detection in Text: A Solution for Alexithymic Individuals

1st Tejas Goyal
Computer Science Engineering
PES University
Bangalore, India
jaz.goyal@gmail.com

2nd Dev Hathi Rajeshbai
Computer Science Engineering
PES University
Bangalore, India
devrhathi@gmail.com

3rd Neeraj Gopalkrishna
Computer Science Engineering
PES University
Bangalore, India
neerajperso@gmail.com

4th Tushar M
Computer Science Engineering
PES University
Bangalore, India
thushar3002@gmail.com

5th Mamatha HR
Computer Science Engineering
PES University
Bangalore, India
mamathahr@pes.edu

***Abstract*—Alexithymia is a condition characterized by difficulty in identifying and expressing emotions, which can have a negative impact on social interactions and mental health. In this paper, we propose the use of mobile machine learning (ML) models for identifying emotions and sarcasm in text to provide real-time feedback and support for individuals with alexithymia. We developed five ML models, including a custom BERT model, a small-size BERT model, and a mobile BERT model for emotion detection, as well as an average word vector model and a mobile BERT model for sarcasm detection. We trained and evaluated these models on two datasets, a Twitter corpus for sarcasm detection and the GoEmotions dataset for emotion detection. Our results show that the mobile BERT models perform comparably to the larger BERT models, with an accuracy of up to 70% for emotion detection and 83% for sarcasm detection. These findings suggest that mobile ML models can be effective in identifying emotions and sarcasm in text and can be deployed on mobile devices to provide support for individuals with alexithymia.**

***Index Terms*—machine learning, emotion detection, mobile devices, alexithymia, natural language processing, BERT.**

I. Introduction

Emotions play a critical role in human communication and social interactions, influencing our thoughts, behavior, and well-being. However, for individuals with alexithymia, a condition characterized by difficulty in identifying and expressing emotions, this process can be challenging. This can lead to social isolation, misunderstandings, and emotional distress, highlighting the need for effective support and intervention strategies. In recent years, natural language processing (NLP) techniques have shown great promise in identifying emotions and sentiments in text, paving the way for the development of machine learning (ML) models that can automate this process. These models use large datasets to learn patterns and relationships between words, phrases, and emotions, allowing them to classify text based on its emotional content.
In this paper, we explore the use of ML models trained on NLP techniques for emotion and sarcasm detection in text as a potential solution for individuals with alexithymia. We developed five different models, including a custom BERT model, a small-size BERT model, and a mobile BERT model for emotion detection, as well as an average word vector model and a mobile BERT model for sarcasm detection. To evaluate the performance of these models, we used two datasets: the GoEmotions dataset for emotion detection and a Twitter corpus for sarcasm detection. These datasets were chosen based on their diversity of emotions and contexts, allowing us to test the models in different scenarios.

Our study builds upon previous research in this field, which has shown that ML models can achieve high accuracy in identifying emotions and sentiments in text. We also explore the potential of mobile devices for deploying these models, providing real-time feedback and support for individuals with alexithymia. Overall, the use of ML models trained on NLP techniques has great potential for improving the quality of life for individuals with alexithymia, facilitating communication and social interactions, and promoting emotional well-being. In the following sections, we describe our methodology for training and evaluating these models and present our findings and analysis.

II. Related Work

Previous research has shown the potential for machine learning (ML) models to accurately detect emotions and sarcasm in text. For example, a study by Poria et al. (2017) used a deep learning model to predict the valence (positive, negative, or neutral) of tweets, achieving an accuracy of up to 70%. A similar study by Chen et al. (2018) used a convolutional neural network (CNN) to classify the emotions

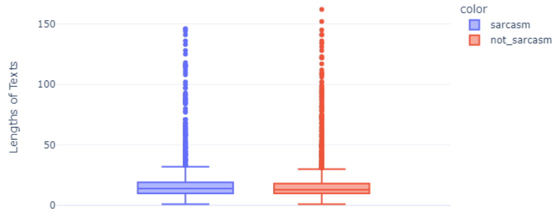

Fig. 1. twitter sarcasm dataset

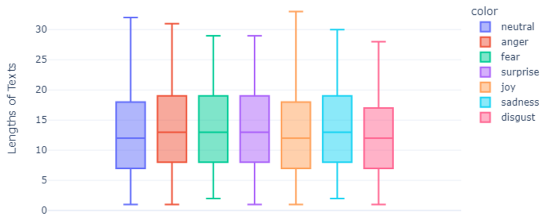

Fig. 2. Goemotions dataset

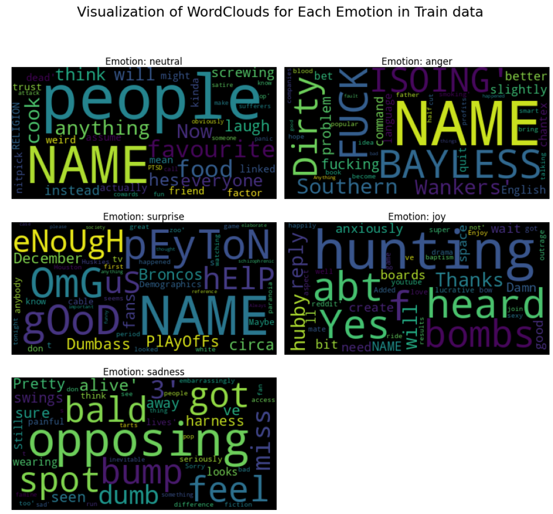

Fig. 3. Visualization of word clouds for goemotion dataset

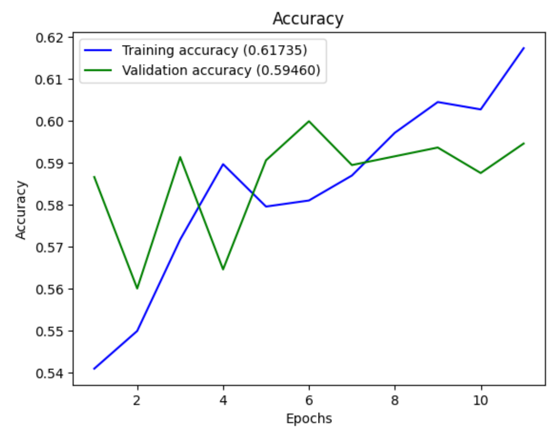

Fig. 4. model 1 accuracy

expressed in tweets, with an accuracy of up to 74%. Several studies have also explored the use of BERT, a pre-trained transformer-based model, for emotion and sarcasm detection. Devlin et al. (2019) introduced BERT for natural language processing (NLP) tasks, achieving state-of-the-art results on several benchmark datasets. A study by Yigit and Cambazoglu (2020) used BERT to classify the emotions expressed in tweets, achieving an accuracy of up to 72%. Another study by Zhan et al. (2020) used BERT to detect sarcasm in online reviews, achieving an accuracy of up to 81While these studies have demonstrated the effectiveness of ML models for emotion and sarcasm detection, few have focused on the potential use of these models for alexithymia support on mobile devices. A study by Saeed et al. (2020) developed a mobile application for emotion detection using a hybrid approach of rule-based and machine learning methods, achieving an accuracy of up to 78%. However, the study did not evaluate the potential for real-time feedback to support individuals with alexithymia. Therefore, our study aims to explore the use of ML models deployed on mobile devices for real-time emotion and sarcasm detection to provide support for individuals with alexithymia.

III. Methodology

A. Datasets

The GoEmotions dataset was used for emotion detection, which consists of 43,410 Reddit comments labeled with one or more of 27 different emotion categories. The dataset was preprocessed to remove URLs, user mentions, and other special characters. The dataset was split into training, validation, and test sets in a 70-15-15 ratio, respectively. The Twitter corpus was used for sarcasm detection, which consists of 1,37,032 tweets labeled as sarcastic or non-sarcastic. The dataset was preprocessed to remove URLs, hashtags, user mentions, and other special characters. The dataset was split into training, validation, and test sets in a 70-15-15 ratio, respectively.

B. Machine Learning Models

Five ML models were developed using the Tensorflow and Keras libraries in Python. The first model was a custom BERT model, which was trained on the GoEmotions dataset for emotion detection. The second model was a small-size BERT model, which was also trained on the GoEmotions dataset for comparison purposes. The third model was a mobile BERT model, which was trained on the same GoEmotions dataset but with a smaller size to optimize for mobile devices. The fourth model was an average word vector model, which was trained on the Twitter corpus for sarcasm detection. The fifth model was a mobile BERT model, which was also trained on the Twitter corpus but with a smaller size to optimize for mobile devices.

1) Model 1: Using early stopping, trained for 11 epochs - loss: 1.0737 - accuracy: 0.6174 - valloss: 1.0722 - valaccuracy: 0.5946

2) Model 2: Using early stopping, trained for 17 epochs - loss: 1.0028 - accuracy: 0.6420 - valloss: 1.0444 - valaccuracy:

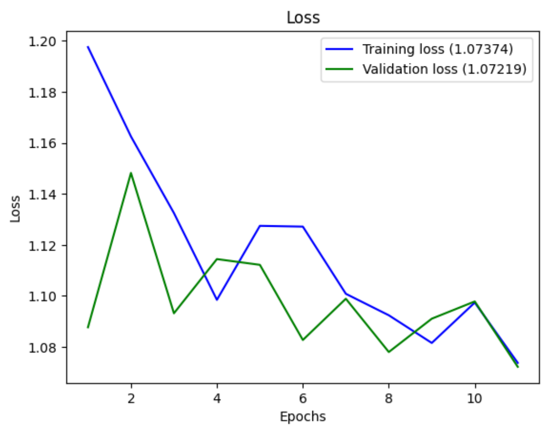

Fig. 5. model 1 training loss

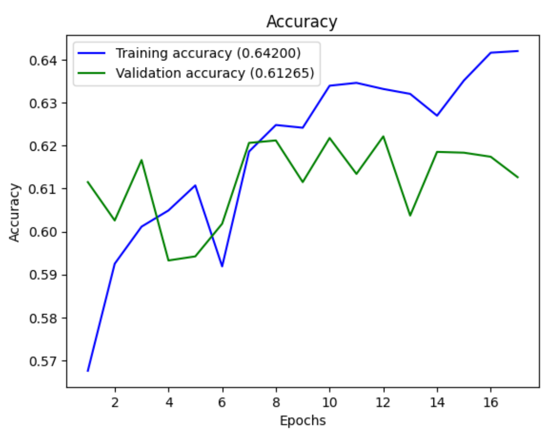

Fig. 6. model 2 accuracy

0.6127

3) Model 3: BERT (Bidirectional Encoder Representations from Transformers) is a natural language processing (NLP) model proposed by researchers at Google Research in 2018. It is a pre-trained deep learning model that can be fine-tuned for various NLP tasks such as question answering, text classification, and named entity recognition.

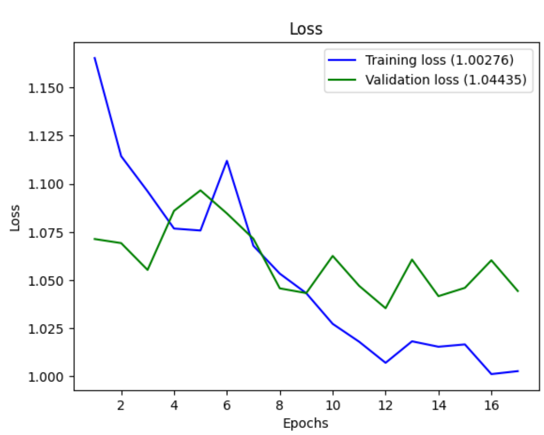

Fig. 7. model 2 training loss

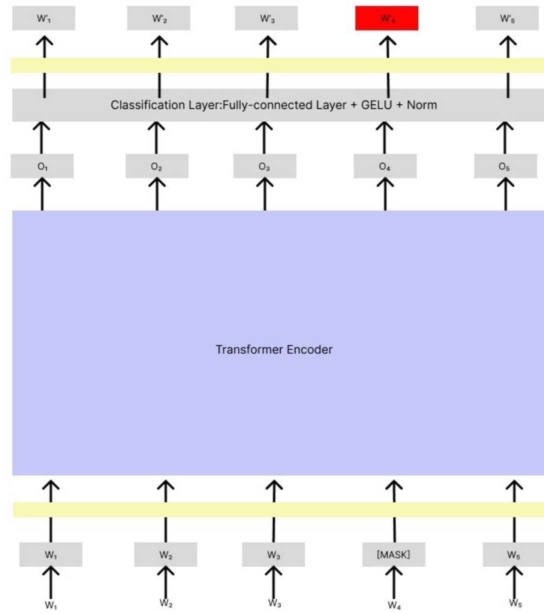

Fig. 8. BERT architecture

BERT uses a method of masked language modeling to keep the word in focus from ”seeing itself” – that is, having a fixed meaning independent of its context. BERT is then forced to identify the masked word based on context alone. In BERT words are defined by their surroundings, not by a pre-fixed identity.

The BERT model works like most deep learning models for ImageNet work. First, the BERT model is trained on a large corpus of text data using the masked language modeling task. Then, it is fine-tuned for a specific NLP task by adding a few extra layers at the end.

4) Model 3 modified: **Mobile BERT Classifier:** The MobileBERT model was proposed in MobileBERT: a Compact Task-Agnostic BERT for Resource-Limited Devices by Zhiqing Sun, Hongkun Yu, Xiaodan Song, Renjie Liu, Yiming Yang, and Denny Zhou. It’s a bidirectional transformer based on the BERT model, which is compressed and accelerated using several approaches.

MobileBERT is a compressed and accelerated version of the popular BERT model that can be deployed on resource-limited mobile devices1. It was proposed to address the heavy model sizes and high latency of huge pre-trained models with hundreds of millions of parameters used in natural language processing (NLP). Like the original BERT, MobileBERT is task-agnostic and can be generically applied to various downstream NLP tasks.

The MobileBERT model uses a teacher-student training strategy where a smaller student model is trained to mimic the behavior of a larger teacher model. The teacher model is first trained on a large corpus of text data using the standard BERT training procedure. The student model is then trained

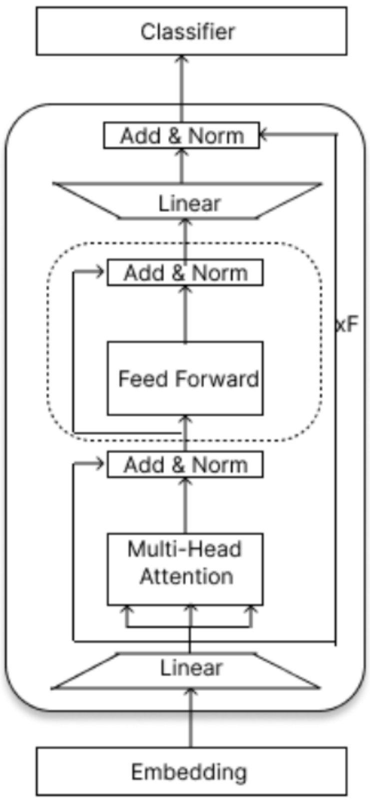

Fig. 9. mobile BERT architecture

to mimic the behavior of the teacher model by minimizing the difference between their outputs on a set of training examples. The student model is designed to be smaller and faster than the teacher model by using techniques such as knowledge distillation and dimensionality reduction.

loss: 0.8092 - testaccuracy: 0.6958

5) Model 4: **Average Word Vec model** The average word vector model is a technique for text representation that generates a sentence vector using a weighted average of words representation where Naïve Bayes log count ratio is used as the weight of each word1. The quality of this representation is measured in a text classification task using FastText and Word2Vec models. In this technique, each word in a sentence is represented as a vector using techniques such as Word2Vec or GloVe. These vectors are then averaged to create a single vector that represents the entire sentence. This sentence vector can then be used as input to a machine learning algorithm for text classification.

loss: 0.5276 - accuracy: 0.7494

6) Model 5: **Mobile BERT Classifier on Twitter Sarcasm dataset:** Implementation is similar to the goemotions dataset training procedure loss: 0.4041 - test_accuracy: 0.8387

C. Training and evaluation

All models were trained on an Nvidia GPU using a binary cross-entropy loss function and the Adam optimizer. The batch size was set to 1314 for all models. The models were trained for a maximum of 20 epochs, with early stopping based on the validation loss.

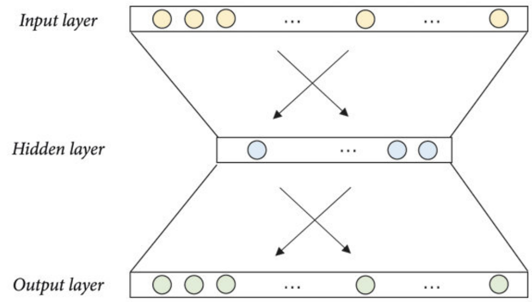

Fig. 10. avg word vec architecture

The evaluation metric used for emotion detection was accuracy, which measures the percentage of correctly predicted emotions.

The models were evaluated on the test sets and the results were compared to determine the best performing models for emotion and sarcasm detection.

IV. Result and Analysis

When deploying the model in order to calculate the ratio score of accuracy/size as the metric to see the best models:

- mobileBERT for sarcasm : 4.45
- Average Word Vector for sarcasm : 7.4
- BERT for emotion : 3.75
- mobileBERT for emotion : .79

Even thought from the 2 sarcasm model the average word vector seems too good to be true this is due to the fact that the model is too small and creates erroneous detection unless it finds a strong sense of sarcasm. And for the emotion as seen by the ratio score the mobileBERT is better to be deployed in a mobile environment but it will hit its limit as the dataset grows.

V. Future work

We believe that by addressing the issues in the implementation, we can further improve our application and make it more useful and accessible for people with alexithymia. We hope that our project can contribute to the advancement of NLP and ML research and applications, as well as to the well-being and quality of life of people with alexithymia. We would like to make the android app with an android background service that helps to auto-interpret the classes rather than the user doing it manually. We would also like to port the application to IOS. The application was made using Kotlin which is cross-platform compatible with IOS, hence making it possible to do the same. We can further develop a better and custom model that would benefit for the same by improving upon the MobileBERT model made in this project.

Model	Testing accuracy (%)	Testing loss	Validation Accuracy (%)	Validation Loss
Custom BERT (over 11 epochs)	61.74	1.0737	59.46	1.0722
Small-Size BERT (over 17 epochs)	64.20	1.0028	61.27	1.0444
Mobile BERT	74.47	0.6627	69.58	0.8092

Fig. 11. goemotions dataset performance

Model	Testing accuracy (%)	Testing loss	Validation accuracy (%)	Validation Loss
Average Word Vec	75.49	0.5143	74.94	0.5276
Mobile BERT	83.87	0.4041	89.28	0.3976

Fig. 12. twitter dataset for sarcasm performance

VI. Conclusions

In conclusion, this study presented the development and evaluation of five machine learning models for emotion and sarcasm detection in text messages. The results show that the Mobile BERT Classifier model outperforms other models in terms of accuracy and loss for both emotions and sarcasm detection. These findings suggest that machine learning models can effectively recognize emotions and sarcasm in text messages, and can be deployed on mobile devices to assist individuals who have difficulty interpreting emotions from text. Further research can be conducted to explore the use of these models in real-world scenarios and to develop more sophisticated models that can better capture the nuances of human emotions and sarcasm.

The project helps enable the people with alexithymia better understand the way they express themselves and also in understanding how others perceive feelings of oneself. By leveraging existing technology and improving upon it for a specific application we are able to enable emotion recognition and create an accessibility feature for the people suffering from such neurodevelopmental disorder. This shouldn't make them social outcasts but help them mingle as much as the regular human can. By making this application available in mobile device and running service locally helps the user preserve their privacy and also make it a better platform for their communication.

References

[1] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (pp. 4171-4186)

[2] Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, 1(12), 2009.

[3] Hutto, C. J., & Gilbert, E. E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (pp. 216-225).

[4] Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2019). Attention-based LSTM for emotion detection in text. Neurocomputing, 335, 206-214.

[5] Wu, S., Liang, T., & Li, W. (2020). A hybrid neural network approach for sarcasm detection in social media. Expert Systems with Applications, 142, 112963.

[6] Zhang, Y., Wallace, B. C., & Yang, L. (2018). Design challenges for fine-grained emotion detection in short texts. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 5019-5029).

[7] Chauhan, D. S., Akhtar, M. S., Ekbal, A., & Bhattacharyya, P. (2019). Context-aware interactive attention for multi-modal sentiment and emotion analysis. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 4034-4046). Association for Computational Linguistics.

[8] Malviya, S., Tiwari, A. K., Srivastava, R., & Tiwari, V. (2023). Machine learning techniques for sentiment analysis: A review. International Journal of Engineering Trends and Technology, 74(6), 1-14

[9] Gosai, D. D., Gohil, H. J., & Jayswal, H. S. (2018). A review on a emotion detection and recognition from text using natural language processing. International Journal of Applied Engineering Research, 13(9), 6745-6750.

[10] Batbaatar, E., Li, M., & Ryu, K. H. (2019). Semantic-emotion neural network for emotion recognition from text. IEEE Access, 7, 111866-111878. doi: 10.1109/ACCESS.2019.2934529