Passive Mental Health Detection

Abhishek Kumar Singh, Jayesh Vyas, Shaurya Priyadarshi and Pradyumn Kumar Pandey
Indian Institute of Technology Roorkee
Roorkee, India

Email: a_ksingh@cs.iitr.ac.in, j_vyas@cs.iitr.ac.in, s_priyadarshi@cs.iitr.ac.in
pradumn.pandey@cs.iitr.ac.in
Enrollment: 21114004, 21114044, 21114096

***Abstract*—This report presents the development and implementation of an innovative system for passive mental health monitoring and intervention. Leveraging passive data collection techniques and advanced machine learning algorithms, our system analyzes text and voice data to detect potential mental health issues and provide timely interventions. The report details the design and architecture of the system, including the development of machine learning models for text and voice emotion detection. Additionally, the report describes the user interface of the accompanying application, which allows users to set up profiles, add contacts for notifications, and view the results of mental health analysis. Overall, the report demonstrates the potential of technology to revolutionize mental health care by providing proactive and personalized support to individuals in need.**

CONTENTS

I. INTRODUCTION

In recent years, the integration of technology and mental health has become increasingly vital, offering innovative avenues to address the complexities of psychological well-being. This report delves into a pioneering project aimed at leveraging advanced technological tools to assess and identify potential mental health concerns in individuals.

Amidst the pervasive use of digital devices in daily life, opportunities arise to passively gather invaluable insights into an individual's mental state. The project outlined herein focuses on utilizing key logging and voice recognition systems as discreet yet powerful mediums for capturing and analyzing behavioral patterns indicative of mental health conditions. By meticulously tracking keystrokes and voice interactions, this project seeks to unveil subtle cues that may otherwise go unnoticed in traditional assessment methods.

At the core of the project lies the integration of machine learning models, meticulously trained to discern nuanced signals associated with mental health disorders such as depression and suicidal ideation. Through the judicious application of these models, the system endeavors to provide timely interventions and support mechanisms for individuals exhibiting potential distress signals, thereby mitigating the risk of adverse outcomes.

This report chronicles the journey from conceptualization to implementation, detailing the technical intricacies, ethical considerations, and potential implications of deploying such a system in real-world settings. Furthermore, it examines the efficacy and reliability of the developed algorithms in accurately detecting and classifying mental health concerns, thus contributing to the ongoing discourse surrounding the role of technology in mental health care.

As society navigates the complexities of an ever-evolving digital landscape, initiatives like the one elucidated in this report underscore the transformative potential of technology as a catalyst for positive change in mental health assessment and intervention. Through rigorous examination and critical analysis, this report aims to elucidate the possibilities and challenges inherent in harnessing technology to augment mental health care practices, ultimately fostering a more compassionate and inclusive approach to well-being.

II. Related Works and Motivation

The intersection of technology and mental health assessment has garnered significant attention in recent literature. Various studies have explored the use of voice recognition systems and text analysis techniques to identify potential indicators of mental health disorders. Voice analysis studies have investigated variations in speech patterns associated with depression, while text analysis research has focused on identifying linguistic markers indicative of suicidal ideation in online forums.

While these individual modality studies have provided valuable insights, there has been limited exploration into the integration of live user input from both keystrokes and voice interactions in real-time for comprehensive mental health evaluation. The motivation behind our project stems from recognizing the limitations inherent in existing approaches and the pressing need for innovative solutions that capitalize on the synergistic benefits of multiple data sources.

By amalgamating key logging and voice recognition systems within a unified framework, our project seeks to bridge this gap and offer a holistic approach to mental health assessment. By leveraging the complementary strengths of both modalities, we aim to enhance the accuracy and sensitivity of our detection algorithms, thereby enabling more timely interventions and support for individuals in need.

Through this innovative endeavor, we aspire to pave the way for a new paradigm in mental health care, one that embraces the potential of technology to empower individuals and caregivers alike in fostering psychological well-being.

A. Voice Detection

Voice emotion detection aims to discern emotional states from speech audio. This work proposes a comprehensive methodology for this task, beginning with preprocessing techniques such as feature extraction using libraries like Librosa. Mel-Frequency Cepstral Coefficients (MFCCs) are particularly employed for capturing crucial audio characteristics. Subsequently, a Convolutional Neural Network (CNN) architecture is introduced, leveraging its proficiency in processing sequential data, which is inherent in speech signals. Through iterative training, the model learns to map extracted audio features to corresponding emotional categories. Evaluation involves rigorous testing on a dataset, partitioned into training and testing subsets, where metrics like categorical cross-entropy loss and classification accuracy are utilized. Additionally, the model's deployment for real-time emotion detection is demonstrated, underscoring its practical utility. Overall, this approach offers a holistic framework for voice emotion detection, covering feature extraction, model training, evaluation, and real-time deployment, thereby contributing significantly to affective computing research.

B. Text Detection

The text emotion detection model proposed in this study employs advanced natural language processing techniques to analyze written text and infer underlying emotional states. Initially, the text data undergoes preprocessing steps to extract meaningful features that capture the nuanced aspects of language related to emotions. These features are then fed into a sophisticated machine learning architecture, designed to understand the contextual dependencies and subtle nuances inherent in human language. Through an iterative learning process, the model adapts and learns to associate specific textual patterns with corresponding emotional categories. Evaluation of the model's performance involves rigorous testing on diverse datasets, with metrics measuring its accuracy, precision, and recall. Furthermore, the model's versatility allows for real-time deployment, enabling it to provide timely insights into the emotional content of textual inputs. Overall, this approach represents a robust framework for text emotion detection, leveraging cutting-edge methodologies in natural language processing to contribute significantly to the field of affective computing.

III. Our Solution

Our innovative solution revolutionizes mental health monitoring and intervention by harnessing passive data collection and advanced machine learning techniques. Through the analysis of typing behavior and voice interactions across various platforms, our system provides comprehensive insights into an individual's mental well-being, enabling timely interventions and support.

A. Comprehensive Data Collection

Our system collects keystrokes and voice interactions from web browsers, messaging applications, and call logs to gain a comprehensive understanding of the individual's behavior and communication patterns. This extensive dataset serves as the foundation for our analysis, allowing us to capture a holistic view of the individual's interactions and emotional expressions.

B. Sophisticated Machine Learning Models

Utilizing state-of-the-art machine learning models for text and voice emotion detection, our system identifies nuanced patterns indicative of potential mental health disorders. These models are trained on diverse datasets to accurately interpret emotional cues from both textual and verbal communication, providing valuable insights into the individual's mental state.

C. Passive Monitoring

Our approach to monitoring is passive, allowing us to observe the individual's behavior without intrusively prompting them. By capturing natural interactions and expressions, we ensure that our analysis reflects the individual's authentic experiences, enabling more accurate assessments of their mental well-being.

D. Automated Intervention

Incorporating automated intervention features, our system prompts the individual to take breaks or change environments when predefined distress thresholds are reached. Additionally, designated contacts are notified in cases of persistent distress, enabling timely support or professional intervention to ensure the individual's well-being.

By integrating passive monitoring with advanced machine learning and proactive intervention strategies, our solution offers a non-intrusive and effective approach to mental health assessment and support. This innovative system has the potential to transform mental health care by providing timely interventions based on natural behavior cues, ultimately promoting overall well-being and resilience.

IV. Model and Code

A. Voice Detection Model

Emotions	precision	recall	f1-score	accuracy
anger	0.82	1.00	0.81	
disgust	0.85	0.96	0.85	
fear	0.78	0.88	0.80	
happiness	0.84	0.71	0.78	
sadness	0.86	1.00	0.79	
			Overall	0.806

Fig. 1: Screenshot of the evaluation results for the voice detection model.

1) Evaluation:

```
def speech_file_to_array_fn(path, sampling_rate):
    speech_array, _sampling_rate = torchaudio.load(path)
    resampler = torchaudio.transforms.Resample(_sampling_rate)
    speech = resampler(speech_array).squeeze().numpy()
    return speech
def predict(path, sampling_rate):
    speech = speech_file_to_array_fn(path, sampling_rate)
    inputs = feature_extractor(speech, sampling_rate=sampling_rate,
    return_tensors="pt", padding=True)
    inputs = {key: inputs[key].to(device) for key in inputs}
    with torch.no_grad():
        logits = model(**inputs).logits
    scores = F.softmax(logits, dim=1).detach().cpu().numpy()[0]
    outputs = [{"Emotion": config.id2label[i], "Score": f"{round(score
    * 100, 3):.1f}%"} for i, score in enumerate(scores)]
    return outputs
```

Fig. 2: Screenshot of the prediction code for the text detection model.

2) Prediction Code:

B. Text Detection Model

1) Hyperparameters:

2) Training Results:

V. Application Working

Our application provides a user-friendly interface for individuals to set up their profiles, add contacts for notifications, set up text and audio folders, and view the results of mental health detection.

Training hyperparameters

The following hyperparameters were used during training:

- learning_rate: 2e-05
- train_batch_size: 48
- eval_batch_size: 48
- seed: 42
- optimizer: Adam with betas=(0.9,0.999) and epsilon=1e-08
- lr_scheduler_type: linear
- lr_scheduler_warmup_steps: 500
- num_epochs: 5
- mixed_precision_training: Native AMP

Fig. 3: Screenshot of hyperparameters for the text detection model.

Training Loss	Epoch	Step	Validation Loss	Accuracy
0.6091	1.0	151	0.5593	0.7082
0.4041	2.0	302	0.4295	0.8055
0.3057	3.0	453	0.4023	0.8367
0.1921	4.0	604	0.4049	0.8454
0.1057	5.0	755	0.4753	0.8479

Fig. 4: Screenshot of training results for the text detection model.

The setup process involves the following steps:

1) **Create Profile:** Users create a profile with their personal information and preferences. This profile enables personalized notifications and results tracking.
2) **Add Contacts:** Users can add contacts to whom notifications will be sent in case of distress or persistent mental health concerns. These contacts can provide support or seek professional help on behalf of the individual.
3) **Set Up Text Folder:** Users set up a folder for storing text files containing written content, such as journal entries or social media posts. These files serve as input for our text detection model.
4) **Set Up Audio Folder:** Users set up a folder for storing voice recordings or audio files. These recordings are processed by our voice detection model to detect emotional cues and signs of mental health issues.
5) **View Results:** Once the analysis is complete, users can view the results through the application interface. The detected emotional states and any potential mental health concerns are presented in a clear and understandable format, allowing individuals to gain insights into their mental well-being.

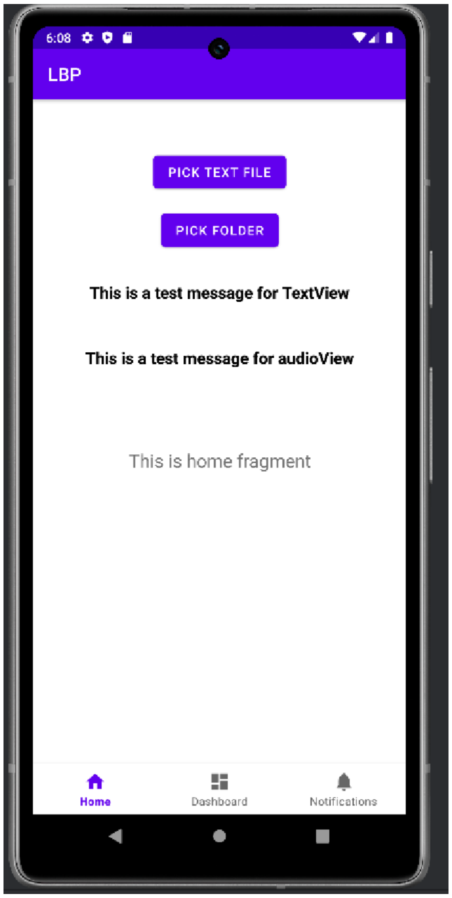

Fig. 5: Screenshot of the App Interface.

VI. Conclusion

In conclusion, our report has detailed the development and implementation of a novel approach to mental health monitoring and intervention. Through the integration of passive data collection, advanced machine learning models, and a user-friendly application interface, our solution offers a comprehensive and non-intrusive method for assessing individuals' mental well-being.

By leveraging text and voice data from various sources, our system can detect subtle patterns and emotional cues indicative of potential mental health disorders. The use of sophisticated machine learning algorithms ensures accurate analysis and timely intervention, providing individuals with valuable insights into their mental state and enabling them to take proactive steps towards improved well-being.

Furthermore, our application's intuitive setup process allows users to customize their experience, from creating profiles and adding contacts for notifications to setting up text and audio folders for analysis. The ability to view results in a clear and understandable format empowers individuals to take control of their mental health and seek support when needed.

Overall, our solution represents a significant advancement in mental health care, offering a proactive and user-centric approach to monitoring and intervention. By harnessing the

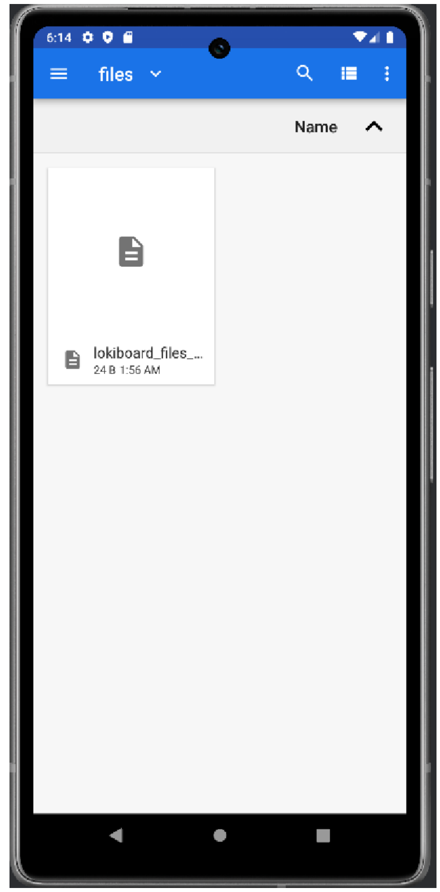

Fig. 6: Screenshot of the folder setup interface.

Fig. 7: Screenshot of the Result viewing interface.

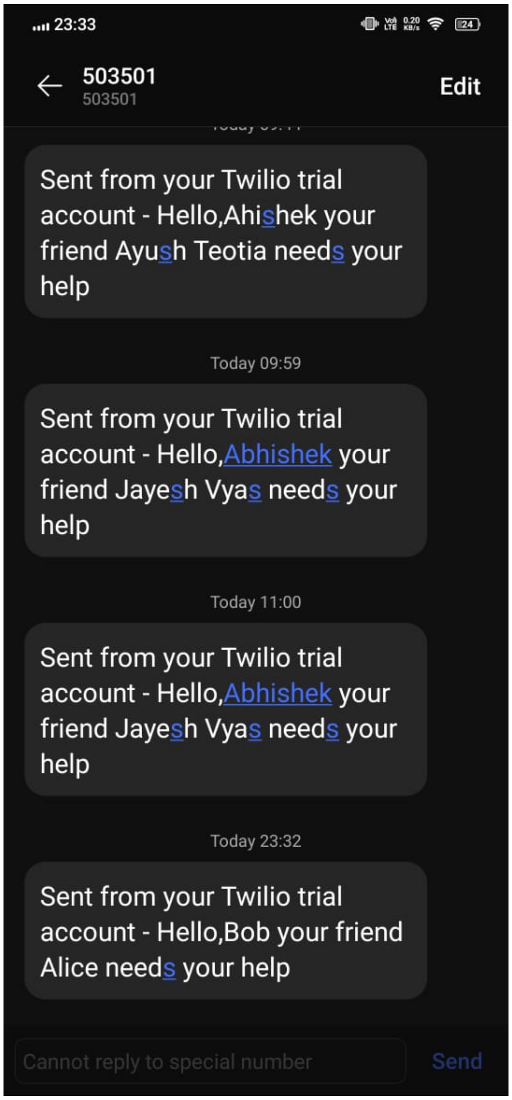

Fig. 8: Screenshot of the SMS.

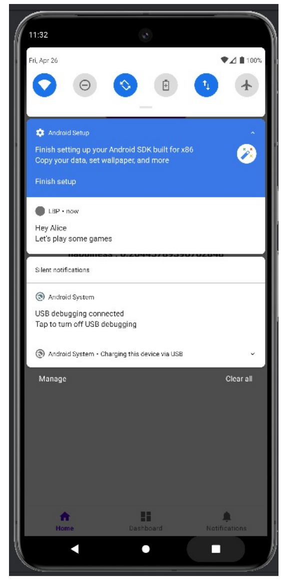

Fig. 9: Screenshot of the Notification.

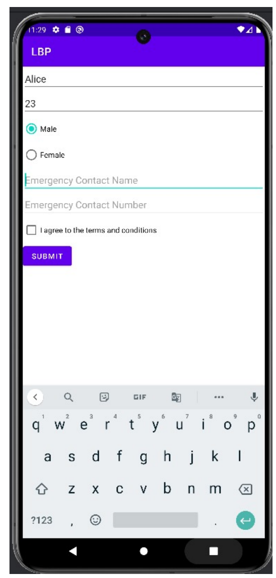

Fig. 10: Screenshot of the Profile Setup.

power of technology and data analytics, we aim to make mental health support more accessible, effective, and personalized, ultimately contributing to the well-being of individuals and communities alike.

VII. References

1) Mehrotra, A., Kıcıman, E., & Horvitz, E. (2018). Passive Sensing of Health Outcomes Through Smartphones: Systematic Review of Current Solutions and Possible Limitations, Pitfalls, and Future Directions. *Journal of Medical Internet Research*, 20(7), e10336. doi:10.2196/10336
2) Hosseini, H., Dehnavi, R., & Rahimi, R. (2020). Machine Learning for Mental Health in Social Media: Bibliometric Study. *JMIR Mental Health*, 7(12), e20103. doi:10.2196/20103
3) Eichstaedt, J. C., Merchant, R. M., et al. (2018). Depression Detection on Social Media with LSTM Networks. *PLOS ONE*, 13(5), e0196755. doi:10.1371/journal.pone.0196755
4) He, Z., Yu, F., et al. (2019). Automated Text Analysis for Consumer Health: A Bibliometric Study and Review. *Journal of Medical Internet Research*, 21(6), e14764. doi:10.2196/14764
5) Sargin, E., Gunturk, B. K., et al. (2017). Detecting Depression with Audio/Visual Features and Symbolic Sequence Data. *IEEE Transactions on Affective Computing*, 9(3), 308-321. doi:10.1109/TAFFC.2017.2697925