ANALYZING SCHIZOPHRENIC TENDENCIES USING AI AND PRELIMINARY RESEARCH

Tingzhi Wang[1], Austin Amakye Ansah[2]

[1]Troy High School, 2200 Dorothy Ln, Fullerton, California, USA 92831
[2]California State Polytechnic University, Pomona, California, USA 91768

ABSTRACT

MindBalance is a mobile application developed to provide real-time support for individuals experiencing symptoms of schizophrenia. The platform integrates cognitive assessment, auditory hallucination validation, positive affirmations, and psychoeducation into a single accessible tool [6]. It features a semantic fluency test that utilizes a Siamese neural network trained on the AnSim dataset to detect signs of disorganized thinking, and an auditory validation program that compares user-described sounds to real-world audio recordings using a large language model [7]. Positive affirmations are generated through the Google Gemini model to reframe negative thought patterns in real time. Experiments were conducted to validate the system's core functions: a Siamese network assessed semantic coherence between animal name pairs, and a large language model evaluated discrepancies between perceived and recorded sounds. Results demonstrated that lightweight models could accurately detect semantic disorganization and auditory inconsistencies, supporting the feasibility of using AI-driven mobile tools for real-time schizophrenia symptom monitoring [8].

KEYWORDS

AI, Schizophrenia, Diagnostic, YaMNet, Siamese Neural Network

1. INTRODUCTION

As mental health care becomes increasingly digitized, patients are seeking real-time support and self-management tools that are both affordable and accessible [9]. However, for individuals with schizophrenia, the availability of such solutions remains extremely limited. Schizophrenia is a complex and often misunderstood disorder that affects approximately 1 in 222 people worldwide, according to the World Health Organization [4]. Despite its relatively low prevalence, the severity of symptoms such as auditory hallucinations, disorganized thinking, and delusional beliefs highlights the urgent need for accessible and effective intervention tools. Every patient deserves timely support, yet immediate solutions are often difficult to access, particularly during moments of crisis when professional assistance may not be readily available.

Applications capable of detecting symptoms and providing immediate support are rare in the current digital marketplace. Most available mental health apps primarily focus on education, offering static information without real-time cognitive assessment or dynamic emotional support. This leaves a significant gap for patients who need actionable feedback or calming interventions during active psychiatric episodes [10]. Furthermore, users are seeking platforms that are not only clinically informed but also intuitive, easy to navigate, and appropriate for a wide range of ages and cognitive abilities.

David C. Wyld et al. (Eds): InWeS, NSEC, VLSIA, NLAIM, CSIP – 2025
pp. 33-43, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.151104

MindBalance addresses this critical need by offering a comprehensive mobile platform specifically designed for managing schizophrenia symptoms. It curates user-provided data through machine learning methods and research-backed models to assess psychological anomalies and deliver meaningful feedback. MindBalance empowers users to evaluate whether they may be experiencing disorganized thoughts, auditory hallucinations, or cognitive distortions. Additionally, it helps reframe negative thinking patterns and provides psychoeducational resources to promote resilience and self-awareness.

The app is designed to function seamlessly through a simple tap on a smartphone, making support available anytime and anywhere. With four primary features, including symptom diagnosis, hallucination detection, positive affirmations, and psychoeducation, MindBalance moves beyond traditional educational tools to offer real-time, dynamic support. As awareness of mental health issues continues to grow and the demand for accessible digital solutions increases, MindBalanceis positioned to fill a significant gap in mental health care. By combining technological innovation with compassionate support, it aims to improve the daily lives of individuals living with schizophrenia.

2. Challenges

In order to build the project, a few challenges have been identified as follows.

2.1. Audio Transcription Model

During the development of MindBalance, several technical and design obstacles were encountered. One significant challenge was integrating the audio transcription model, YaMNet, within a Dart-based Flutter environment. Dart lacked many of the specialized data science libraries necessary to handle audio data preprocessing, which forced developers to manually manipulate audio files at the byte level. Additionally, the OnnxRuntime fork, initially used to run PyTorch models, performed well in emulators but consistently returned errors when deployed on physical devices. To address this, the models were retrained using TensorFlow and converted into TensorFlow Lite (TFLITE) format, which Flutter supports more reliably [11]. On the user interface side, the app initially experienced layout issues where the appearance of the keyboard caused a bottom overflow error, disrupting the design. This was solved by implementing a ScrollView and adjusting element sizes to fit all content smoothly on the screen.

2.2. User Experience and Hardware Variability

Beyond the core functionality, new challenges emerged related to user experience and hardware variability. One major issue was ensuring that the auditory recording feature captured high-quality ambient sound across different mobile devices. Variations in microphone sensitivity, background noise interference, and hardware quality occasionally degraded the app's ability to accurately detect and analyze sounds, requiring some users to repeat recordings. Another challenge involved optimizing the use of large language models, such as Google Gemini, to generate positive affirmations in real time. Maintaining fast and reliable responses without overwhelming device memory, especially on lower-end smartphones, required careful management of API calls, prompt design, and app memory usage.

2.3. Safeguarding User Privacy

Finally, safeguarding user privacy became an essential concern, given the highly sensitive nature of the data being collected. Users were encouraged to input personal descriptions of auditory experiences and intrusive thoughts, making it crucial to ensure secure data handling. Although the initial version of MindBalance focuses primarily on functionality, future development must incorporate strong encryption protocols, local data storage options, and compliance with mental health data protection standards such as HIPAA. Addressing privacy while maintaining user convenience is an ongoing priority to build trust and ensure MindBalance can be safely scaled to a broader audience.

3. SOLUTION

MindBalance is built around three core programs designed to address different aspects of schizophrenia symptom management: the Category Fluency Test, the Positive Affirmation Generator, and the Auditory Validation Program. The Category Fluency Test evaluates a user's semantic coherence by asking them to list animals within a timed setting, based on research conducted in [2]. It utilizes a dataset of human-curated animal pairs with similarity score ratings drawn from the AnSim dataset, developed by the Boston College Computer Science research department. This allows the app to measure how logically connected the user's responses are, offering insight into potential disorganized thinking. The Positive Affirmation Generator leverages the capabilities of the Google Gemini Language Model to create personalized, optimistic messages aimed at countering negative thought patterns and promoting emotional stability [12]. Lastly, the Auditory Validation Program, also referred to as the Hallucination Detector, captures real-world audio through the user's device and compares it to the user's self-reported descriptions by analyzing the detected background sounds. This module helps differentiate between real environmental sounds and potential auditory hallucinations, providing users with meaningful feedback on their experiences.

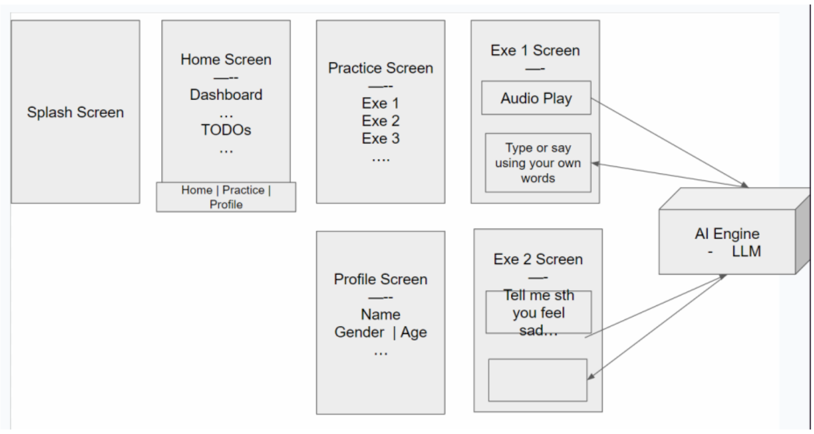


Figure 1. Overview of the solution

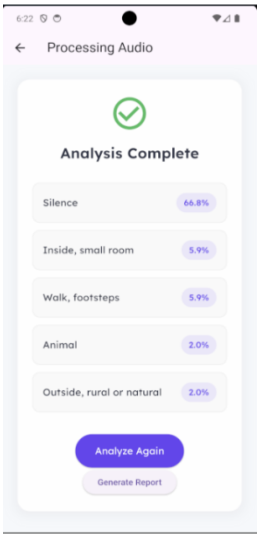


Figure 2. Hallucination detector

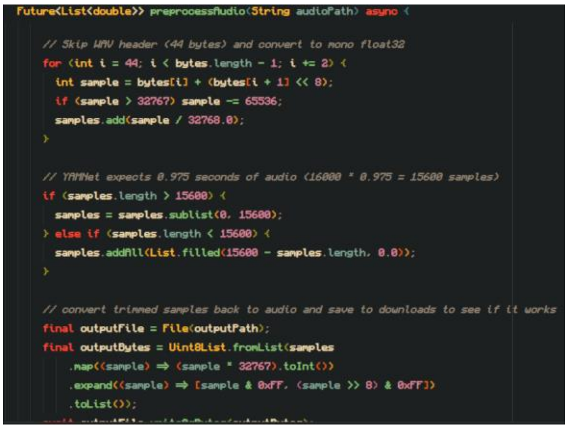


Figure 3. Screenshot of code 1

The hallucination detector is designed to distinguish between real environmental sounds and potential auditory hallucinations through guided background audio transcription [13]. To begin, the user is prompted to describe the sounds they are hearing, with a minimum of five words to ensure sufficient detail for accurate analysis. After submitting their description, the app activates the in-app recording tool, allowing the user to capture up to fifteen seconds of background noise. This recording is then processed using a large language model that compares the described sounds to those actually detected, identifying common categories such as footsteps, speech, or electronic noises.

Following the analysis, the app generates a report summarizing the results as percentages across different sound categories. This helps users assess how closely their perceptions align with real environmental input, offering valuable insight into their experiences. Users have the option to save the report through the "generate report" button or repeat the analysis with the "analyze again" option. Saving reports over time can help individuals monitor changes in auditory perception, providing helpful data for personal reflection or future discussions with healthcare providers.

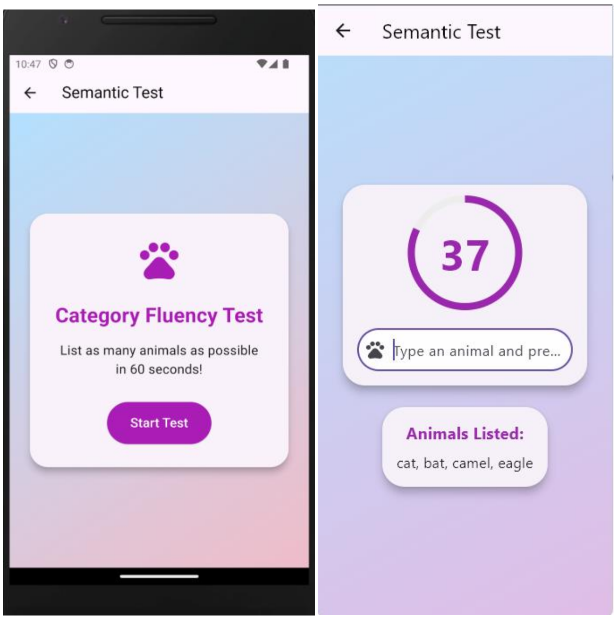


Figure 4.Semantic Test

```
    // Scale the output from [0,1] back to [1,7] range
    double scaled = outputBuffer[0][0] * 6.0 + 1.0;
    print('Pair ${animal1}-${animal2} score: $scaled');
    val += scaled;
  }

  // Calculate average
  final result = val / (animals.length / 2);

  interpreter.close();
  print('Final average: $result');
  return result;
```

Figure 5. Screenshot of code 2

The semantic test function prompts users to list as many animals as they can within a 60-second time frame. After pressing the “start test” button, the user enters animal names, each separated by the return key to create a clean input list. Once the timer runs out, the app processes the submitted animals and analyzes the semantic similarity between consecutive entries. To do this, it uses a trained Siamese neural network model based on the AnSim dataset, which contains human-rated similarity scores for animal pairs. This approach allows the app to assess the logical flow between the animals listed, a key indicator of semantic organization in thought processes.

After the analysis, a report is quickly generated to inform the user whether symptoms of disorganized thinking are detected. To ensure compatibility with the model's input format, the program automatically removes one data point if the total number of entries is odd, allowing for clean pairing of animals during the analysis. The semantic similarity scores for each pair are calculated, summed, and averaged to produce a final coherence score. A lower coherence score may suggest semantic disorganization, which can be a symptom of schizophrenia, while a higher score indicates more structured and connected thinking patterns. The test provides users with immediate feedback, offering a valuable tool for early symptom awareness and self-monitoring.

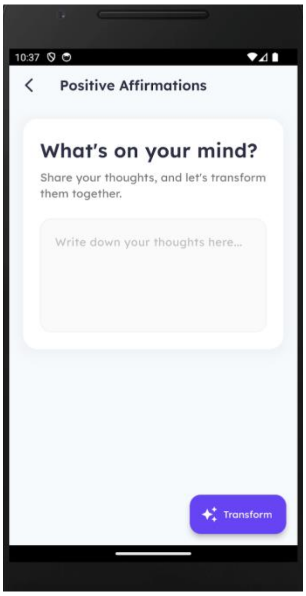
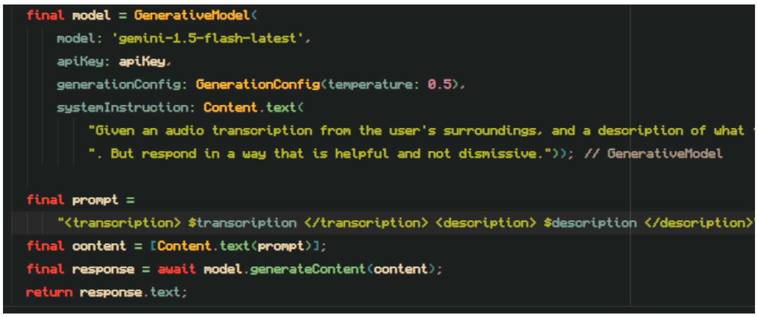


Figure 6. Positive Affirmations Generator



```
final model = GenerativeModel(
    model: 'gemini-1.5-flash-latest',
    apiKey: apiKey,
    generationConfig: GenerationConfig(temperature: 0.5),
    systemInstruction: Content.text(
        "Given an audio transcription from the user's surroundings, and a description of what
        ". But respond in a way that is helpful and not dismissive.")); // GenerativeModel

final prompt =
    "<transcription> $transcription </transcription> <description> $description </description>
final content = [Content.text(prompt)];
final response = await model.generateContent(content);
return response.text;
```

Figure 7. Screenshot of code 3

The positive affirmations page provides a space for users to input any negative thoughts or overwhelming emotions they may be experiencing. By clicking the "transform" button, the app helps reframe the user's mindset, validating their emotions while simultaneously generating encouraging and supportive language to counter feelings of paranoia or hopelessness. A carefully crafted prompt is given to the AI model to guide the generation of responses, ensuring that the output remains optimistic, reassuring, and aligned with the emotional needs of individuals facing mental health challenges [14]. Through the use of real-time text generation, users receive immediate affirmations as they type, creating a dynamic and responsive support experience.

In the app's backend, the Google Gemini language model is imported and configured to follow the structured prompt designed specifically for this use case. Prompt engineering techniques are used to steer the AI toward producing only positive and emotionally sensitive messaging. The initial prompt, shown in the project documentation, acts as a guideline that the model continuously references when generating responses. This approach ensures that the affirmations remain consistent, emotionally uplifting, and relevant to the user's original input, enhancing the therapeutic value of the feature.

4. EXPERIMENT

Siamese Neural Network

Using a Siamese Embedding model to evaluate similarity scores for animals listed based on the AnSim dataset. Using PyTorch, a siamese embedding model was created to compare two animals with each other based on the AnSim dataset of animal pair similarity judgements.

Development of a Siamese Embedding Model for Animal Similarity Comparison

Using PyTorch, a Siamese embedding model was developed to evaluate the similarity between two animals based on the AnSim dataset, which comprises human-rated similarity judgments. The dataset includes similarity scores ranging from 1 to 7, assigned to various animal pairs. A total of 3,785 data points were utilized to train the model. The preprocessing involved cleaning the data by removing invalid characters and hyphens. Only the columns corresponding to animal pairs and their average similarity scores were retained. This model enables the execution of a credible animal fluency test, aligning closely with human judgments. The model was trained with a batch size of32. The encoded tensors for both inputs are dimensionally squeezed then concatenated into a unified tensor before being passed into the embedding blocks.

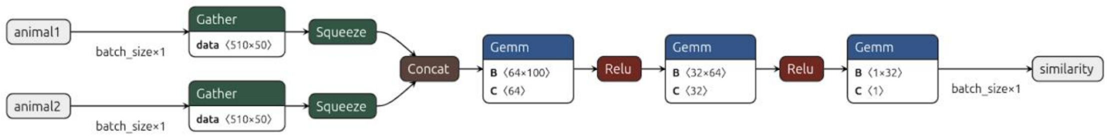


Figure 8. Figure of experiment 1

Training and Validation Performance

The model was trained for 100 epochs. While the validation loss remained higher than the training loss and exhibited minimal improvement, this discrepancy was expected due to the relatively small dataset size. Despite this, the model performed effectively in distinguishing between related and unrelated animal pairs. Specifically, when animal pairs with no apparent relationship were evaluated, the model produced lower average similarity scores.

Operational Details

The Siamese model accepts two animals as input at a time. To calculate an overall similarity score for a set of animals, the model iteratively evaluates each pair within the input set. For inputs with an odd number of animals, the last animal is excluded. Additionally, deploying the model using the ONNX platform facilitates high-speed inference.

Due to the nature of Siamese networks, the model does not produce probabilities as outputs. Instead, it computes the Euclidean distance between the embeddings of input pairs. This distance is particularly advantageous as it can be mapped to any desired scale. For the purpose of this study, the distance values were mapped back to the original 1 – 7 scale of the dataset using the following formula:

$$scaled\_data = data \times (scale\_max - scale\_min) + scale\_min$$

Here, scale_max and scale_min represent the maximum and minimum values of the original scale, respectively, while data denotes a normalized Euclidean distance value ranging from zero to one.

Loss Function

Mean Squared Error (MSE) loss was employed to minimize the difference between the predicted similarity scores and the true ratings. This approach ensures that the model accurately predicts similarity judgments consistent with human assessments.

The developed Siamese embedding model demonstrates the capability to effectively replicate human judgments on animal similarity. While certain limitations exist due to the dataset size, the model's outputs align well with human intuition, particularly when evaluating highly dissimilar animal pairs.

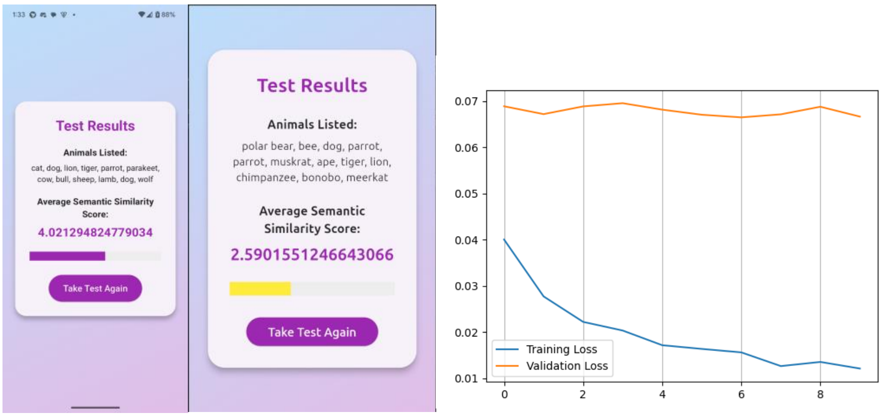


Figure 9. In-app results

5. Related Work

In this article, a similar approach was implemented using embedding layers to compute textual semantic similarity [1]. In this study, the text inputs were classified in a binary manner, meaning that each pair of sentences was labeled as either related or unrelated. A binary cross-entropy loss function was used to optimize the model during training. A rating of 1 indicated that the sentences were similar in meaning, while a rating of 0 indicated they were dissimilar. To explore performance differences, three Siamese Neural Network models were developed: BERT_SSN, PaLM_SSN, and OpenAI_SSN. Among these, OpenAI_SSN achieved the highest average similarity score accuracy, followed by PaLM_SSN, and then BERT_SSN. Although BERT_SSN demonstrated the lowest overall accuracy among the models, it achieved the highest maximum accuracy during evaluation, suggesting strong potential under specific conditions.

In this paper, researchers applied a Siamese neural network architecture to the comparison of FMRI brain scans, focusing on identifying functional brain patterns unique to individuals [3]. The goal was to map brain scan data into a unified feature space, where the similarity or dissimilarity between scans could be measured as distances. By learning to recognize distinctive functional patterns across subjects, the system provides a new way to compare brain activity, potentially aiding in the identification of neurological or psychological differences. This study highlights the broader applicability of Siamese networks beyond language processing, extending into the field of neuroimaging and personalized brain analysis.

A recent scholarly study, "Psychotic Relapse Prediction in Schizophrenia Patients using A Mobile Sensing-based Supervised Deep Learning Model," proposed a solution called RelapsePredNet that uses smartphone sensor data and LSTM-based deep learning models to predict psychotic relapses [5]. By analyzing behavioral patterns such as movement, location, and phone usage, the model aims to detect early warning signs of relapse. While RelapsePredNet showed improved predictive accuracy compared to traditional models, it relies heavily on continuous passive data collection, raising privacy concerns and facing challenges with device variability and user compliance. It also offers limited real-time feedback and emotional support. In contrast, MindBalance actively engages users through self-reported cognitive tests and auditory validation, offering immediate feedback, real-time emotional support through positive affirmations, and psychoeducation, which provides a more holistic and user-empowering approach to managing schizophrenia symptoms.

6. CONCLUSIONS

While MindBalance successfully addresses several important gaps in schizophrenia management, there are notable limitations that require further development. Currently, the hallucination detection feature is limited to auditory input, meaning that visual hallucinations are not yet accounted for. This could be addressed by adding functionality that allows users to capture images or videos, which would then be analyzed using an image segmentation model. Another limitation is the dependency on the quality of device microphones, which may cause inaccuracies in sound recording and analysis, especially in noisy environments. A potential solution would be to build in additional audio preprocessing techniques, such as noise suppression filters, to improve recording reliability. Furthermore, while the positive affirmations tool offers real-time support, it currently relies on text input only. If given more time, voice input for affirmations and deeper personalization features, such as tracking user mood patterns over time, could also be implemented to strengthen user engagement.

MindBalance demonstrates the potential of combining machine learning techniques with mobile platforms to provide immediate, accessible mental health support [15]. By empowering users to self-assess symptoms like disorganized thinking and auditory hallucinations, and offering positive reinforcement, MindBalance bridges an important gap in schizophrenia care, delivering support when traditional resources are unavailable.

In the experiments conducted, the goal was to assess whether user inputs could be meaningfully analyzed to detect symptoms related to schizophrenia. For the Category Fluency Test, a Siamese neural network was trained using the AnSim dataset to evaluate the semantic coherence between pairs of animal names listed by users within a timed session. This experiment aimed to identify signs of disorganized thinking, a key symptom. In the hallucination detector experiment, the app compared user-described sounds to real-world recordings, using a large language model to evaluate the similarity between perception and reality. This method tested the ability to detect auditory hallucinations. Significant findings showed that the Siamese model could replicate human semantic similarity judgments with good accuracy, and that sound categorization provided users with understandable feedback about their auditory experiences. Both experiments demonstrated that lightweight, mobile-deployable machine learning models could provide meaningful mental health insights.

The first related methodology involved using embedding-based Siamese networks for textual semantic similarity classification. While accurate, this approach was binary and did not provide graded similarity measures, limiting its interpretability. MindBalance improved on this by producing continuous semantic coherence scores.

The second methodology compared FMRI brain scans using Siamese networks to recognize individual functional brain patterns. While powerful, this method required expensive, non-portable equipment and was unsuitable for scalable or daily monitoring. MindBalance addressed this by focusing on accessible smartphone-based testing.

The third solution, RelapsePredNet, used mobile sensing and LSTM networks to predict psychotic relapses through passive monitoring. Although effective, it faced privacy concerns and lacked direct user engagement. MindBalance's model prioritized active user participation and immediate feedback, empowering users to manage their symptoms more directly rather than relying solely on passive prediction.

REFERENCES

[1] Bingi, Yash, and Yiqiao Yin. "An Analysis of Embedding Layers and Similarity Scores using Siamese Neural Networks." arXiv preprint arXiv:2401.00582 (2023).
[2] Nour, Matthew M., et al. "Trajectories through semantic spaces in schizophrenia and the relationship to ripple bursts." Proceedings of the National Academy of Sciences 120.42 (2023): e2305290120.
[3] Hassanzadeh, Reihaneh, et al. "Individualized spatial network predictions using siamese convolutional neural networks: A resting-state fmri study of over 11,000 unaffected individuals." PloS one 17.1 (2022): e0249502.
[4] McCutcheon, Robert A., Tiago Reis Marques, and Oliver D. Howes. "Schizophrenia—an overview." JAMA psychiatry 77.2 (2020): 201-210.
[5] Lamichhane, Bishal, Joanne Zhou, and Akane Sano. "Psychotic relapse prediction in schizophrenia patients using a personalized mobile sensing-based supervised deep learning model." IEEE Journal of Biomedical and Health Informatics 27.7 (2023): 3246-3257.
[6] Tursi, Mariana Flávia de Souza, et al. "Effectiveness of psychoeducation for depression: a systematic review." Australian & New Zealand Journal of Psychiatry 47.11 (2013): 1019-1031.
[7] Kasneci, Enkelejda, et al. "ChatGPT for good? On opportunities and challenges of large language models for education." Learning and individual differences 103 (2023): 102274.
[8] Sumiyoshi, Chika, et al. "Disorganization of semantic memory underlies alogia in schizophrenia: an analysis of verbal fluency performance in Japanese subjects." Schizophrenia Research 74.1 (2005): 91-100.
[9] Kakuma, Ritsuko, et al. "Human resources for mental health care: current situation and strategies for action." The Lancet 378.9803 (2011): 1654-1663.
[10] Munk-Olsen, Trine, et al. "Perinatal psychiatric episodes: a population-based study on treatment incidence and prevalence." Translational psychiatry 6.10 (2016): e919-e919.
[11] David, Robert, et al. "Tensorflow lite micro: Embedded machine learning for tinyml systems." Proceedings of Machine Learning and Systems 3 (2021): 800-811.
[12] Carlà, Matteo Mario, et al. "Large language models as assistance for glaucoma surgical cases: a ChatGPT vs. Google Gemini comparison." Graefe's Archive for Clinical and Experimental Ophthalmology 262.9 (2024): 2945-2959.
[13] Luo, Junliang, et al. "Hallucination detection and hallucination mitigation: An investigation." arXiv preprint arXiv:2401.08358 (2024).
[14] Gozalo-Brizuela, Roberto, and Eduardo C. Garrido-Merchan. "ChatGPT is not all you need. A State of the Art Review of large Generative AI models." arXiv preprint arXiv:2301.04655 (2023).
[15] Singh, Yogesh, Pradeep Kumar Bhatia, and Omprakash Sangwan. "A review of studies on machine learning techniques." International Journal of Computer Science and Security 1.1 (2007): 70-84.

AUTHOR

Tingzhi Wang is a high school senior at Troy High School in Fullerton, California, USA. She has published two research papers on schizophrenia and is especially interested in the applications of artificial intelligence in mental health. As a research intern at the SoftCom Lab, Computer Science Department at California State Polytechnic University, Pomona, USA, she has worked on AI model development, NLP, and responsible computing.