Applied Mathematics and Computational Intelligence
Volume 14, No. 4, 2025 [144 - 153]

Identifying Postpartum Depression Symptoms on Social Media Using Machine Learning Techniques

Zaber Al Hassan Ayon[1], Nur Hafieza Ismail[2*], Hoh Jia Xuan[3], Nur Shazwani Kamarudin[4]
Muhammad Zulfahmi Toh Abdullah[5]

[1,2,3,4,5]Faculty of Computing, University Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia

[*] Corresponding author: hafieza@umpsa.edu.my

Received: 15 August 2024
Revised: 11 November 2024
Accepted: 24 September 2025

ABSTRACT

Postpartum depression (PPD) is one of the most common maternal morbidities after delivery. Most new mothers are at risk for PPD not only after birth, but also during pregnancy. There is no single cause of PPD, but physical changes, emotional distress, and genetics may play a key role in this issue. The symptoms of PPD can be strong negative feelings faced by mothers, such as never-ending anxiety, sadness, fatigue, and mood swings. In this work, we proposed a framework for identifying symptoms of PPD based on linguistic characteristics in their textual posts on social media. Today, most people are active on social media platforms to keep in touch with family and friends. Social media allows users to have conversations, share information and feelings through the posting feature available. Thus, this has opened up an opportunity to explore the text content on social media posted by PPD sufferers. We crawled the data using Twitter API (currently known as X) and preprocessed it to remove noises. In the experiment, the Support Vector Machine (SVM) presented the highest accuracy of 87.5% compared to other algorithms. The results indicate that we can utilize the extracted model to gain a deeper understanding of this group.

Keywords: machine learning, NLP, PDD, sentiment analysis, social media

1 INTRODUCTION

The birth of a baby is a big and life-changing experience. Being a parent can trigger mixed feelings of excitement and anxiety at the same time. Those feelings are normal, especially for a first-time parent. Typically, mothers will experience the 'baby blues' phase within the first three days after delivery and may last up to two weeks. In the worst cases, the symptoms can last for months or years. Therefore, if the negative feelings, such as constant mood swings, severe sadness, and crying spells, become more extreme over time, the person may suffer from PPD. Moreover, PPD does not have a set duration. PPD is a mental health disorder that typically arises following childbirth. PPD can affect not only the new mother but also the father and adoptive parents. The new parent is not only experiencing changes in physical, hormonal, and emotional health but also in social patterns, lifestyles, and financial moves. These alterations may lead to PPD symptoms.

PPD affects about 10–15% of mothers globally after pregnancy, with a high prevalence in developing countries [1]. Pregnancy and childbirth are some of the major life events for women, which affect their physical and mental states. Depression after pregnancy may have a negative affect not only on the mother but also on the child, spouse, and other family members that live together [2]. The untreated mother with PPD will feel stressed and exhausted almost every day. If left untreated, the mother will have a detrimental influence on her ability to care for herself and the newborn [3]. Thus, it is crucial to manage the symptoms immediately so the mother can fulfill her new role as caretaker and the baby can grow in a healthy and positive environment.

Unfortunately, blood tests and body scans do not detect PPD. Normally, the doctor will conduct an interview to learn about the conditions and recent feelings, as well as detect any PPD mood-related symptoms that patients may be experiencing. The standard PPD screening tests used by medical experts are the Edinburgh Postnatal Depression Scale (EPDS), the 2-Question Patient Health Questionnaire (PHQ-2), and the 9-Question Patient Health Questionnaire (PHQ-9). However, individuals who suffer from PPD may be unable to identify or admit that they are depressed. In most cases, sufferers may be unaware of PPD's symptoms and indications. Furthermore, the limited time allocated for the routine check-up of the new mother and baby makes it very difficult for the doctor to recognize the symptoms in patients. As a result, it is critical to have alternative approaches or methods for uncovering PPD symptoms in new parents.

Social media platforms provide their users with the opportunity to create and share information, thoughts, feelings, ideas, and other forms of expression. Globally, there are 4.8 billion social media users, accounting for 59.9% of the world population and 92.7% of all internet users. This statistic is from Statista.com. Through the creation of blogs, podcasts, films, and game websites, people use social media to create memories, discover new things, market themselves, make friends, and expand ideas. Twitter, Instagram, TikTok, Facebook, and other social media platforms are particularly well-liked by users. These applications, due to the popularity of social media, can be an alternative platform for identifying PPD symptoms in new parents. Social media postings can provide important information or indications of PPD. For example, Twitter users often express their feelings and thoughts in writing [4]. Machine learning (ML) algorithms can analyze the posting content on Twitter to gain an overview of the individual's PPD status.

Numerous studies have explored PPD. The implementation of ML approaches and social media could improve the diagnosis by identifying the best possible PPD predictors. Previous studies trained ML models like logistic regression (LR), random forest (RF), and extreme gradient boosting (XGBoost) on PPD health data from the IBM MarketScan Medicaid Database. This dataset contains 573,634 female respondents from various backgrounds. The studies conclude that each model's performance varied depending on the technique of bias reduction and outcome label [5]. Another similar study crawled and trained ML models on user-generated text postings from Reddit to classify PPD content and non-PPD content. Multilayer perceptron (MLP), a type of deep learning (DL), did better than SVM and LR models in the experiment, correctly labeling PPD content 86.9% of the time [6].

Thus, the goal of this work is to identify the PPD symptoms in the linguistic features shown by the new parent through their text posting on Twitter. We manually crawled the dataset using PPD-related keywords and preprocessed it by removing unrelated information. We then analyze the cleaned dataset using several ML and DL algorithms, compare the results, and use the model with the best performance to identify PPD symptoms on social media. Specifically, we will answer the following research questions (RQs). RQ1: How do you identify tweets containing PPD symptoms (ground truth data collection)? RQ2: Which algorithm performed best on our crawled dataset? Following these RQs, we presented a framework that can identify PPD symptoms from text postings.

2 PREVIOUS WORK

In recent years, there has been a burgeoning interest in employing ML and DL techniques to tackle the challenge of PPD detection and prediction. Previous studies in this field have utilized electronic health records (EHRs) and ML algorithms to predict PPD [7]. It uses a vast amount of comprehensive patient data from EHRs. The EHRs are from Weill Cornel Medicine and New York-Presbyterian Hospital. They use six ML models, including SVM, RF, XGBoost, naïve bayes (NB), decision tree (DT), and L2-regularized logistic regression. Among these six ML algorithms, SVM has the highest specificity and the best performing model. The results demonstrated the ML approach's feasibility for PPD risk prediction based on information collected during prenatal care in an EHR. Additionally, they discovered several disease diagnoses and medications taken during pregnancy that could potentially contribute to the prediction of PPD.

Moreover, there are studies that have utilized ML algorithms to predict PPD based on a variety of factors, including the PHQ2 and EPDS questionnaires, ranging from demographic information to psychosocial variables. These endeavors have shown promising accuracy rates, offering insights into identifying women at risk for PPD and facilitating early intervention strategies [8]. Another study that used DL to find PPD discovered that the best way to look at a lot of data from various sources is with a hybrid model that combines transfer learning (TL), bi-directional long short-term memory (BLSTM) models, and convolutional neural networks (CNNs) [9]. In addition, researchers have shown that ML techniques such as SVM can accurately detect mental health symptoms from social media text, further highlighting the potential of ML for early mental health screening [10, 11].

Further advancements in DL techniques have been demonstrated by leveraging the long short-term memory (LSTM) with CNNs networks to predict PPD by capturing temporal patterns in mood fluctuations. This method worked better than traditional ML algorithms, showing that DL has the potential to capture the complex temporal dynamics that are part of PPD progression. The work aims to provide accurate and timely identification of PPD symptoms to enable early intervention and support for new mothers [12]. Another similar work applies the Relief algorithm to feature selection, enhancing the interpretability and accuracy of ML models in PPD classification. The suggested method aimed to improve prediction models for a more accurate risk assessment screening tool by picking out the most important features from a large group of predictors, such as lifestyle and demographic information about the mother [13].

Moreover, recent research endeavors have explored the fusion of multimodal data sources to further enhance PPD prediction accuracy. This work investigated multimodal fusion strategies, integrating written materials and audio data using DL architectures equipped with attention mechanisms. This approach enabled capturing intricate interactions chronological interpersonal relationships among different modalities, leading to improved prediction outcomes [14]. In the experimental setup, researchers develop an integrated model that combines deep reinforcement learning (DRL) and differential evolution (DE) algorithms. Through interaction with a simulated environment, an agent acquires the ability to recognize PPD symptoms. The agent receives feedback and adjusts its strategy over time to improve its detection capabilities. Simultaneously, the DE algorithm optimizes the hyperparameters and structural configurations of the DRL model. This hybrid approach allows for the fine-tuning of the model's learning process and ensures that it can effectively handle the complexities and variability of PPD symptoms. The study's results show that using both DRL and DE algorithms together makes PPD identification much more accurate than using only one or the other [15].

Significant prior research has focused on PPD. Various studies have explored different forms of treatment utilizing DL and ML techniques, while others have examined the incidence and risk factors associated with PPD. Notably, researchers have employed ML algorithms to predict the likelihood of PPD based on demographic and clinical data. Researchers have used DL approaches, like CNNs, to analyze text data from social media to identify symptoms of PPD. These advanced methodologies aim to enhance early detection and intervention, potentially improving outcomes for affected mothers. Plus, the widespread use of social media can become a platform for testing theories about user behavior and mental health.

3 THE PROPOSED FRAMEWORK

This section explains in detail the proposed framework for identifying PPD symptoms from Twitter data using the SVM algorithm. Figure 1 depicts the main phases involved in this work. The five main phases are data collection, data preprocessing, experiment setup, and result evaluation.

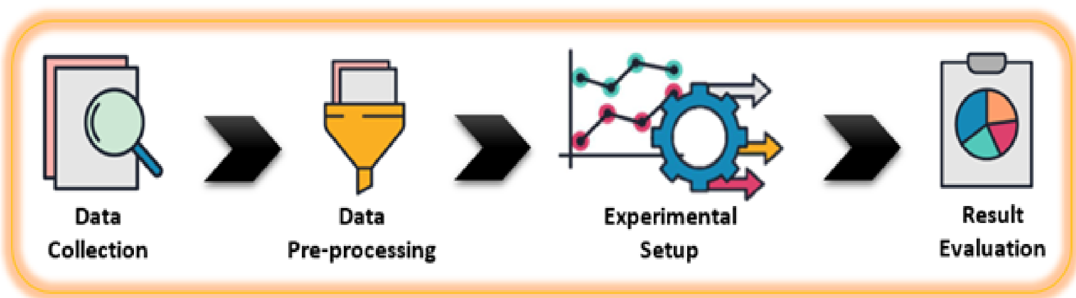

Figure 1 : PPD prediction framework

3.1 Data Collection

Data collection is the initial phase of the PPD prediction framework. During this phase, the Twitter API in Python crawls the text dataset using a list of PPD-related keywords. The keywords used are postpartum depression, depression, and PPD. During this process, we exclude the retweet feature to prevent repetitive tweets. On platforms like Twitter, the retweet feature allows users to share someone else's tweet with their own followers. This can result in the platform repeating a single tweet multiple times. This repetition can make it challenging to analyze or classify the content effectively because the repeated tweets can skew the results. Therefore, eliminating the retweet feature reduces the impact of repetitive content and ensures that word frequency accurately reflects a word's importance in the dataset. This leads to better and more reliable classification results. The processes involved in this phase have answered our RQ1. Figure 2 presents the sample raw tweets retrieved from Twitter. We crawled a total of 500 data points over several periods in 2020–2022.

Content	Aut
Postpartum depression is not a sign of weakness. It's a condition that can be treated.	@a
Your mental health is important. Take the time you need to heal from postpartum depression.	@N
Postpartum depression has a way of making everything feel hopeless.	@E
Even simple joys seem distant when struggling with postpartum depression.	@t
Postpartum depression feels like a heavy fog that I can't seem to lift.	@A
The deep sadness of postpartum depression often feels like it's taking over my life.	@_
The isolation from postpartum depression often makes me feel completely alone in my struggles.	@l
Celebrate each small victory on your journey through postpartum depression. You are making progress.	@v
Believe in the power of your resilience. Each day brings new possibilities for healing and growth	@j
There is hope and healing on the horizon. Trust the process and be kind to yourself.	@c
Taking the first step towards recovery from postpartum depression can be daunting, but it's worth it.	@S
the father is a piece of shit for filming this obvious mental breakdown, likely due to postpartum depression or psychosis, instead of taking the baby somewhere safe and helping the mother get the mental health resources she needs. he doesn't care about either of them.	@J
the postpartum depression survey flagged me as depressed so I'm being strongly encouraged to see a therapist, but all the therapists I can find cost hundreds of dollars per week so basically I'm being strongly encouraged to lie going forward	@p
Postpartum depression is a medical condition, not a personal failure.	@N
Support from family and friends can be a lifeline during postpartum depression	@c
It's hard to see past the fog of postpartum depression.	@p
Postpartum depression casts a long shadow over every part of my life.	@n
The emotional exhaustion from postpartum depression is profound and consuming.	
Postpartum depression often feels like an invisible force pushing me down.	@_
Postpartum depression makes even the simplest tasks feel monumental.	@j
Postpartum depression may be tough, but your resilience and hope are powerful forces for change.	@c
Celebrate each small victory on your journey towards recovery. You are making progress every day	@b
Surround yourself with love and support. You are on the path to recovery.	@v
You are important, and your mental health matters. Seek help for postpartum depression.	@F
she's a teenage girl who went through childbirth and is taking on the stress of caring for a baby and the person recording	

Figure 2: Raw tweets from Twitter

3.2 Data Preprocessing

In the preprocessing phase, we exclusively utilized the textual content of the tweets, filtering out irrelevant entries, resulting in a final dataset of 497 tweets. To facilitate sentiment analysis, we employed a two-step annotation process. We first fine-tune a large language model (LLM) using a few-shot learning approach. This involved training the LLM on a small, curated dataset of tweets labeled for positive and negative sentiment, specifically related to postpartum depression. We then tasked the LLM with annotating the 497 tweets, assigning a label of "1.0" for negative sentiment (indicating PPD) and "2.0" for positive sentiment. To ensure annotation accuracy, all LLM-generated labels underwent human verification and correction. The goal of this rigorous process is to enhance the reliability of the dataset. The final dataset, comprising tweets and their corresponding sentiment annotations, served as the foundation for training the models.

In contrast to conventional natural language processing pipelines, the stop words and embedded links within the tweets are retained. This decision was motivated by the potential for these elements to carry valuable sentiment information, particularly in the context of social media language use. The tokenization technique is applied to prepare the textual data for model ingestion, segmenting each tweet into individual words. Next, the term frequency-inverse document frequency (TF-IDF) vectorization is used to figure out how important each word in a tweet is compared to the whole corpus. This gives us a numerical representation of the text that accurately shows what it means.

Furthermore, the results of our exploratory data analysis (EDA) on the dataset are presented. The EDA assists in understanding the distribution of labels and the characteristics of tweet lengths, which are critical for the subsequent modeling phase. The first analysis focused on the distribution of labels in the dataset as shown in Figure. 3. The bar chart demonstrates the balance of the dataset, showing approximately 200 samples for each label in Figure 4. This balance guarantees impartiality in the model training process and equal representation of both classes. To achieve reliable and generalizable performance in classification tasks, a balanced dataset is crucial because it keeps the model from being biased towards any particular class.

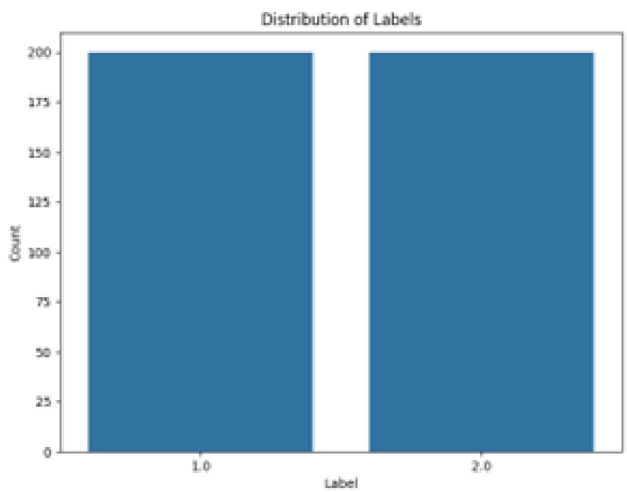

Figure 3: Distribution of label

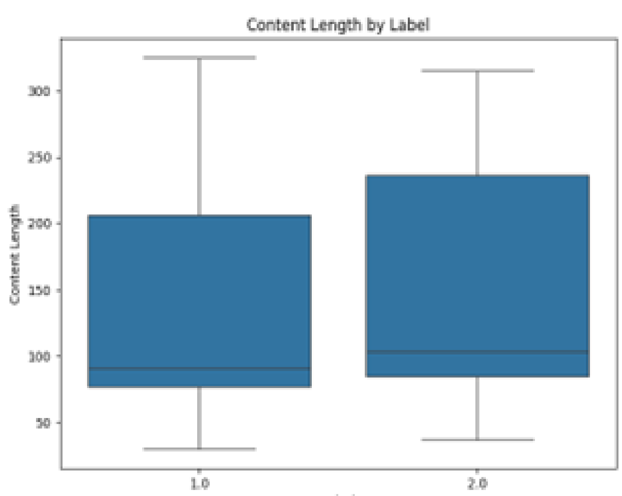

Figure 4: Tweet length by label

Next, using a box plot, the distribution of tweet lengths by label is examined. For label 1.0, the median tweet length is around 100 characters, with the interquartile range (IQR) spanning from approximately 50 to 200 characters. There are some outliers extending up to around 350 characters. Similarly, for label 2.0, the median tweet length is also around 100 characters, but the IQR is slightly broader, extending from approximately 50 to 250 characters, with outliers reaching up to 350 characters. This similarity in the median and IQR for both labels indicates that the tweet length distribution does not significantly differ between the two classes. The presence of outliers in both classes suggests that some posts are significantly longer than the majority, which might influence the text processing steps.

Finally, we analyzed the overall distribution of tweet lengths using a histogram augmented with a kernel density estimate (KDE). The histogram shows a concentration of most tweet lengths between 50 and 150 characters, with a prominent peak around 80 characters, suggesting that this is the most common length for tweets in the dataset as depicted in Figure 5. There are smaller peaks around 200 and 250 characters, suggesting that there are a few tweet clusters of these lengths. The KDE line helps visualize the distribution's smoothness and confirms the multimodal nature of tweet lengths, with a primary mode around 80 characters and secondary modes around 200 and 250 characters.

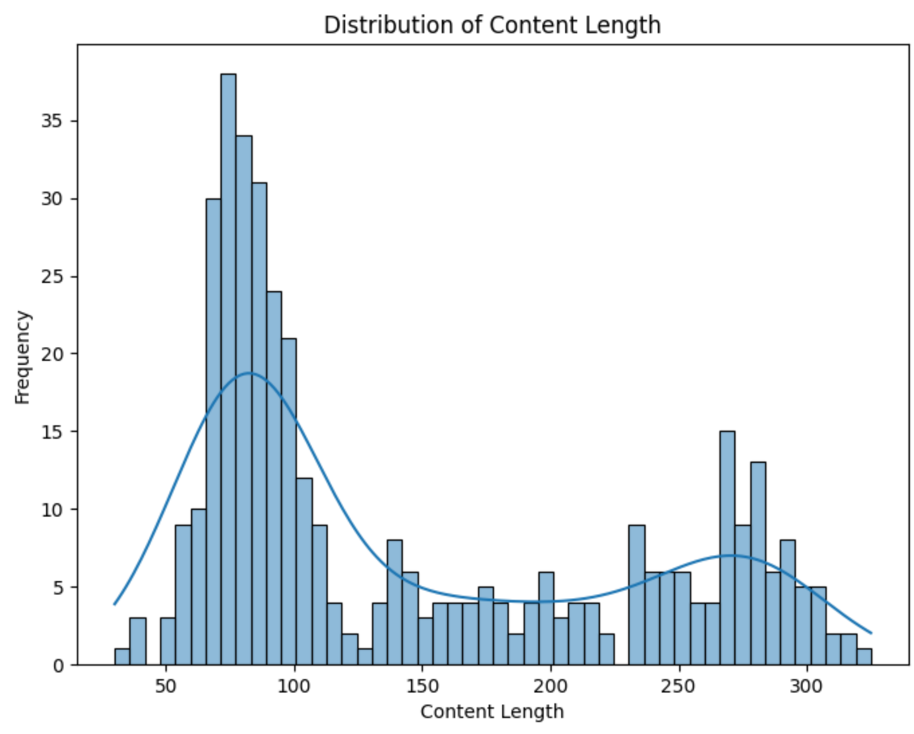

Figure 5: Tweet length distribution

In summary, the exploratory data analysis provided valuable insights into the dataset. The balanced label distribution ensures unbiased model training, while the tweet length analysis revealed similarities between the two classes and highlighted the presence of outliers. The overall tweet length distribution showed distinct peaks, indicating common tweet lengths within the dataset.

These findings are instrumental in guiding the preprocessing steps and informing the development of our classification model.

3.3 Experiment

In the next phase, we conducted an experiment using Python and the Scikit-Learn and Keras libraries. These libraries provide a range of functions for data analysis with SVM, KNN, and CNN. In addition to the Scikit-Learn library, the code also embedded the Pandas library to read the CSV file containing the dataset. Previous studies selected these three algorithms due to their outstanding performance in handling text analysis. While SVM divides groups based on binary categorization, KNN's core premise predicts a data point's label using the same or similar labels of its nearest neighbors. CNN streamlines the word embedding process for tweet categorization. The last phase results evaluation where the experiment output from selected algorithms were presented in accuracy, precision, recall, and F1-score values. The detailed output is discussed in the next section.

4 RESULT AND DISCUSSION

With data collection, data pre-processing, and experimentation out of the way, result evaluation is the last phase for the PPD prediction. Table 1 presents the experimental outputs of the selected algorithms applied in this case study. The SVM shows the highest accuracy of 87.5% compared to the other models in predicting tweets correctly. The goal of precision is to calculate the model to not label a positive as a negative, and SVM is also obtained the highest among the rest and this has answered our RQ2. While the recall can label all the positive data and CNN obtained the highest value in this regard. Lastly, the F1-Score was able to harmonic the mean value between precision and recall and CNN again obtained the highest score.

Based on the overall experiment results in Table 1, SVM has been shown to be most effective in detecting PPD in parents' written postings, where its ability to handle outliers is better than that of KNN and CNN. CNN presents the lowest, which might be due to the limited size of the dataset. Plus, text data is inherently sequential and contextual, whereas CNN is better suited for capturing spatial patterns. On the other hand, KNN measures the similarity between data points using distance metrics, such as the cosine or Euclidean distance. These distance measurements may lose their ability to discriminate between classes in high-dimensional text data since all points will often be equally spaced from one another, resulting in less precise categorization. SVM is particularly excellent at identifying PPD symptoms because it can analyze complex patterns in text data that may not be easily done by other methods. This is because SVM can identify the non-linear relationships that exist between variables.

Table 1 : The experimental results

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM	**87.50**	**87.73**	**87.50**	**87.46**
KNN	81.67	81.68	81.67	81.65
CNN	51.67	51.67	100.00	68.13

5 CONCLUSION

PPD is one of the psychiatric disorders that develop in some of the new parents after having a baby. Untreated PPD can cause permanent emotional damage due to long periods of negative feelings. Moreover, PPD could be difficult to recognize as the symptoms are not manifest in behaviors shown to the surrounding people. Thus, to identify this group, we proposed a model to identify PPD symptoms on social media text postings. From the experiment, SVM obtained the highest accuracy value compared to KNN and CNN. The model can identify the important and complex text features in data that are associated with PDD in written format. This extracted model can be a helping tool to alarm new parents who are at risk of developing PDD and suggest they seek a medical doctor for a further check on their mental health status.

There are several limitations in this study where we only focus on text postings. The contextual information of the social media user such as gender, geodemographic, etc., are excluded from the experiment. Moreover, the dataset is also very limited due to the time restriction and the social media users may use a wide range of terms and hashtags to describe their postpartum experiences. We believe that, the performance of the developed model can be improved if we have more datasets and amount of algorithms for the experiment. Future study should concentrate on refining and applying the techniques used in this study to more databases, maybe including Instagram, Facebook, TikTok and so on. Text classification can be used broadly across many other different topics and would be able to detect other mental illnesses like anxiety, depression, or even bipolar disorders. This study may pave the way for more, crucial studies that may aid those with mental illnesses who express their ideas and feelings on social media.

ACKNOWLEDGEMENT

This research was fully funded by the UMPSA Research Grant Scheme under grant RDU220350, UMPSA.

REFERENCES

[1] S. Shorey, C. Y. I. Chee, E. D. Ng, Y. H. Chan, W. W. San Tam, and Y. S. Chong, “Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis,” *J Psychiatr Res*, vol. 104, pp. 235–248, 2018.

[2] N. Vliegen, S. Casalin, and P. Luyten, “The course of postpartum depression: a review of longitudinal studies,” *Harv Rev Psychiatry*, vol. 22, no. 1, pp. 1–22, 2014.

[3] A. Gopalakrishnan, R. Venkataraman, R. Gururajan, X. Zhou, and G. Zhu, “Predicting women with postpartum depression symptoms using machine learning techniques,” *Mathematics*, vol. 10, no. 23, p. 4570, 2022.

[4] S. R. Lim, N. S. Kamarudin, N. H. Ismail, N. A. H. Ismail, and N. A. M. Kamal, "Predicting Mental Health Disorder on Twitter Using Machine Learning Techniques," in *2023 IEEE 8th International Conference On Software Engineering and Computer Systems (ICSECS)*, 2023, pp. 23–27.

[5] Y. Park *et al.*, "Comparison of methods to reduce bias from clinical prediction models of postpartum depression," *JAMA Netw Open*, vol. 4, no. 4, pp. e213909–e213909, 2021.

[6] I. Fatima, B. U. D. Abbasi, S. Khan, M. Al-Saeed, H. F. Ahmad, and R. Mumtaz, "Prediction of postpartum depression using machine learning techniques from social media text," *Expert Syst*, vol. 36, no. 4, p. e12409, 2019.

[7] S. Wang, J. Pathak, and Y. Zhang, "Using electronic health records and machine learning to predict postpartum depression," in *MEDINFO 2019: Health and Wellbeing e-Networks for All*, IOS Press, 2019, pp. 888–892.

[8] J. F. Raisa, M. S. Kaiser, and M. Mahmud, "A machine learning approach for early detection of postpartum depression in Bangladesh," in *International Conference on Brain Informatics*, 2022, pp. 241–252.

[9] U. K. Lilhore *et al.*, "Unveiling the prevalence and risk factors of early stage postpartum depression: a hybrid deep learning approach," *Multimed Tools Appl*, pp. 1–35, 2024.

[10] Ibrahim, M. A., Ismail, N. H., Kamarudin, N. S., Nafis, N. S. M., & Nasir, A. F. A. (2023, August). Identifying PTSD Symptoms Using Machine Learning Techniques on Social Media. In *2023 IEEE 8th International Conference On Software Engineering and Computer Systems (ICSECS)* (pp. 392-395). IEEE.

[11] Ismail, N. H., Liu, N., Du, M., He, Z., & Hu, X. (2019). Using Deep Neural Network to Identify Cancer Survivors Living with Post-Traumatic Stress Disorder on Social Media. In *SEPDA@ ISWC* (pp. 48-52).

[12] P. Srivatsav and S. Nanthini, "Detecting Early Markers of Post-Partum Depression in New Mothers: An Efficient LSTM-CNN Approach Compared to Logistic Regression," in *2024 5th International Conference on Innovative Trends in Information Technology (ICITIIT)*, 2024, pp. 1–6.

[13] D. Shin, K. J. Lee, T. Adeluwa, and J. Hur, "Machine learning-based predictive modeling of postpartum depression," *J Clin Med*, vol. 9, no. 9, p. 2899, 2020.

[14] U. K. Lilhore *et al.*, "Prevalence and risk factors analysis of postpartum depression at early stage using hybrid deep learning model," *Sci Rep*, vol. 14, no. 1, p. 4533, 2024.

[15] S. Shen, S. Qi, and H. Luo, "Automatic Model for Postpartum Depression Identification using Deep Reinforcement Learning and Differential Evolution Algorithm.," *International Journal of Advanced Computer Science & Applications*, vol. 14, no. 11, 2023.