Prediction of postpartum depression in Zacatecas Mexico using a machine learning approach

Lopez-Veyna J. Ivan[1], Ortiz-Garcia Mariana[1], Diaz-Diaz Alvaro Moises[1], Bermejo-Sabbagh Carlos[2]
[1]Division of Graduate Studies and Research, National Technological Institute of Mexico Campus Zacatecas, Zacatecas, Mexico
[2]Department of Systems and Computer, National Technological Institute of Mexico Campus Merida, Merida, Mexico

Article Info

Article history:

Received Nov 5, 2024
Revised Aug 13, 2025
Accepted Sep 1, 2025

Keywords:

Classification
Machine learning
Postpartum depression
Prediction
Pregnancy

ABSTRACT

Postpartum depression (PPD) is a silent disorder, difficult to detect by the mother who suffers from it. In this research project, we propose a classification model of PPD using machine learning (ML) techniques, following a supervised learning approach. This is model allows the prediction of PPD using sociodemographic and medical data through a dataset of 100 Zacatecan mothers previously classified with the result of Edinburgh Test. We use eight ML algorithms such as adaptative boosting classified (ABC), principal component analysis (PCA) boosting, decision trees (DT), k-nearest neighbors (KNN), support vector machines (SVM), random forests (RF), and boosting. Our results show that the proposed ML model based on ABC algorithm can outperform other classifiers yielding a precision of 90%, a recall of 90%, a F1-score of 78% and 74% for area under curve (AUC), illustrating a correct capability in the prediction of this disorder.

Corresponding Author:

Ortiz-Garcia Mariana
Division of Graduate Studies and Research, National Technological Institute of Mexico Campus Zacatecas
Carr. Panamericana Km. 7 La Escondida, Zacatecas, Mexico
Email: mariana.ortiz@zacatecas.tecnm.mx

1. INTRODUCTION

According to the World Health Organization, depression affects more than 300 million people in the world [1]. By 2030 it will be the second cause of the decrease in years of healthy life worldwide and the first in developed countries. The results of the National Mental Health Survey in 2018 revealed that 18% of the urban population of productive age (15-64 years of age) suffer from some mood disorder such as anxiety, depression or phobia [2] conforming to the DSM- IV, perinatal depression or postpartum depression (PPD) is like an episode of major depression [3]. In Mexico it is observed that between 12 and 14% of pregnant women have depression [4]. PPD is a disorder that affects the physical and mental state of the mother. Women with PPD have feelings of sadness, anxiety and fatigue, which makes it difficult for them to carry out the daily activities of caring for themselves and for other people [5]. This disorder is a silent disorder, difficult to detect by the mother who suffers it. A mother who has suffered PPD has between 10% and 50% higher risk to have again this disorder during a new pregnancy [6].

In the study "Maternal depressive symptomatology in Mexico: national prevalence, care and population risk profiles", headed by Dr. Filipa de Castro, researcher at the Center for Population Health Research (CISP) of the National Institute of Public Health (INSP) [7], reported that, the distribution of PPD in the state of Zacatecas has reached 20%. Also, Mendez-Cerezo in [8] reports that in Mexico, exists an incidence of 13% of PPD, and this percentage is incremented until 22.5% in the fourth weeks of puerperal.

The problem of PPD has been approached in a traditionally way from the psychological point of view using the Edinburgh postnatal depression scale (EPDS), suggested in 1987 by Cox *et al.* [9] EPDS is the most widely used screening questionnaire for PPD. It incorporates a 10-item self-report questionnaire, women are asked to rate how they have felt in the previous 7 days. The score of each one of the questions is 0–3 (resulting range 0–30) and responding takes around 5 min. EDPS are the widely accepted scale to classify 'possible' and 'probable' depression respectively.

Currently, PPD is a silent disorder, difficult to detect by the mother who suffers from it and is often not diagnosed. Furthermore, even if there exist some efforts from different research using a technological solution from a proper Diagnosis of PPD, currently, there does not exist a technological tool for a proper diagnosis. However, thanks to the technological advances on artificial intelligence (AI) and machine learning (ML) techniques, currently it is possible to offer to the woman's that suffer some type of PPD a technological alternative for a prediction and classification of this psychological disorder. Specifically in the City of Zacatecas, Mexico where there not exists any research project to predict and prevent PPD.

Different approaches using ML techniques has been proved to predict PPD, among which we can mention, logistic regression (LR) and support vector machines (SVM) in [10] using social network post. Naïve Bayes (NB), decision tree (DT), LR, adaptative boosting classifier (ABC) and bagging to predict PPD were used in [11]. Adaptative neuro fuzzy was used in [12]. Gradient boosted tree (GBT) was used in [13]. SVM, random forest (RF), DT, XGBoost (XGB), and multilayer perceptron (MLP) where used in [14] using Twitter (tweets) and Instagram (comments). RF, stochastic gradient boosting (SGB), SVM, recursive partitional and regression trees (RT), NB, k-nearest neighbor (KNN), LR, and neural network (NN) were used in [15] using data from the pregnancy risk assessment monitoring system. RF and RT were used in [16] to explore physical, psychological and social factors to predict postpartum pain and depression. Convolutional neural networks (CNN) were used in [17], [18]. Extremely randomized trees (XRT) algorithm was used in [19]. Fazraningtyas *et al.* [20] presented an updated version of ML approaches for DPP, finding that RF, SVM, LR, XGB, and ABC show the highest predictive accuracy. DT, XGB, regularized LR, MLP were used in [21] with electronic health records to predict PPD. LR, SVM and MLP was used with social media text to predict PPD in [22]. LR, SVM, DT, NB, XGB, and RF were used in [23] to predict PPD using information related to demographics data, diagnoses and medication data.

According to previous research in this paper we propose a prediction and classification model for PPD using ML techniques, following a supervised learning approach. This proposal will contribute to predict and classify PPD, because in our geographical and cultural context there not exist a model that consider some specific characteristics of the state of Zacatecas, México, and can be used to a point of comparison with other models developed in the region, other states of Mexico, or in other countries. Also, this model could be an additional support for EPDS because this scale is based on the certainty of the mother's responses, and can yield false positive (FP) or false negative (FN) results, however, with an appropriate ML model and the appropriate data set we are allowed to help in the prediction of DPP with good results.

On the other hand, the dataset created with this research project could be an important repository for Secretary of Salud del Estado de Zacatecas, because currently and not only in the state of Zacatecas, but there are also few records and statistics about PPD in Mexico. With this proposal, we can generate statistics about PPD that could be analyzed and considering to know the real state of PPD in the City of Zacatecas and propose some strategies to reduce the apparition of this disorder.

2. METHOD

The scope of this research project is to propose a new approach to predict and classify PPD using ML as a supervised classification problem. For this purpose, we use the dataset obtained from Parga-Anaya and Garcia-Sanchez [24]. We create a model based on supervised learning to classify PPD. Our methodology includes two main activities: i) identify the supervised learning algorithm which generates the highest accuracy in the classification of PPD and ii) evaluate the performance of our model and compare the efficiency with other research projects. We follow a ML approach starting with data acquisition, preprocessing, classification and evaluation the performance of the model.

2.1. Data acquisition

In this study, we used the dataset from Parga-Anaya and Garcia-Sanchez [24] from 2014-2015, which include data from 100 women of the Zacatecas, Mexico state. These data were obtained from a questionary interview applied in the Clinic 1 of IMSS of Zacatecas City. This questionary contains clinical information and socio demographic data. The first criteria to consider is that all women were in a puerperal stage, this means that all of them have a baby born in the last one month, and until one year old. With respect to gynecological-obstetric clinical features we consider the type of birth (cesarean, natural birth), lactation, use of anesthesia, pregnancy planned, pregnancy wanted, primigravida, multigravida, number of children, abortions, neonate illness, and institution where the birth was attended. In relation with socio demographic data, the features such as age, marital status, schooling, occupation, monthly income for the family, partner support. With respect to PPD, this feature was measured with Edinburgh questionnaire [9] that allows to classify the women as healthy or depression.

Some features of the dataset were described in next paragraphs. For the type of birth 45% have a cesarean and 55 have a vaginal birth. A 3.3 month was the mean for lactation, with a maximum of 12 months and a minimum of 0 months or without lactation. In the case of the use of anesthesia during childbirth 65% of women have using some type of anesthesia in their labor. 54% have a pregnancy planned, while 89% have a pregnancy wanted. 31% of women were primigravida, while 69% were multigravida. 31% of women were their first child, while 67% have two or more children's. With respect to abortions only 13% of the sample have one abortion, while 10% have two. 8% have a neonate illness such as cardiopathy, lactose intolerance or tachycardia. With respect to institution where the birth was attended 51% were on the IMMSS, 29% on Hospital de la Mujer Zacatecana, 13% on private hospital and 7% on ISSSTE. With respect to socio demographic data the dataset includes a mean of 26.5 years, with a maximum of 43 and a minimum of 16 years. In the marital status 76% of women are married, 18% have a free union, 5% were single, and 1% were widow. In relation with level of schooling 43% have junior high school finished, 30% have high school, 19% have a university preparation, 5% have a master or a graduate degree, and only 3% have an elementary schooling. With regards to occupation 62% of interviewees are housewives, while 38% were employed. The monthly income for the family has a mean of 6,595 Mexican pesos, with a maximum of 30,000 pesos and a minimum of 1,000 pesos. 100% of women have a partner support, and only 94% have a family support. With respect to PPD, this feature was measured with Edinburgh questionnaire [9] that allows to classify the women as healthy or depression, 39% of women were classified with PDD, while 61% were classified as healthy.

2.2. Pre-processing

To pre-processing our dataset, we follow a series of previous steps, including data cleaning, removal observations with missing values, data extraction, normalize and standardize features. Feature extraction methods involve removing features with no-linear relationships between independent features or little-to-none-variance, with the aim to select a subset of prominent features that allow to identify or describe symptoms of depression. In the process of cleaning, we review the dataset with 25 features with the purpose to identify and establish some empty data or data out of range. Next, we apply a normalize and standardization processes proposed by Brownlee [25] who established that normalization scales each input variable individually to the range 0-1, is in this range were floating-point values have the highest precision. Standardization is the process of scaling each input variable individually by subtracting the mean (called centring) and dividing by the standard deviation so that the distribution has a mean of zero and a standard deviation of one. These two processes were generated using the sklearn preprocessing Python library [26]. Standardization of a dataset is a commonly task for several ML estimators because they might behave badly if the individual features do not more or less look like standard normally distributed data.

2.3. Classification

For our classification modeling of PPD, we used the pre-processed dataset, which was divided into two parts: the first part (70% of the data) will be used to create a classification model using different supervised learning algorithms (ABC, principal component analysis (PCA) bosting – PCB, DT, KNN, SVM, RF, NB, and LR). The second part (30% of the data) will be used to test the efficiency in the classification of the model and thus be able to choose which is the best algorithm that yields the highest accuracy by classifying PPD with the aim to generate a PPD prediction model. Once the model is determinate, it will be used within the mobile application. The model was generated used Python and its libraries, in a Dell processor: intel(R) Core (TM) i3-7130U CPU, @2.70 GHz, RAM 4GB laptop.

2.4. Evaluation

To evaluate our classification model, we consider some ML classification metrics used in the literature. This metrics shows the ability to predict PPD correctly. The confusion matrix is used to visualize the performance of an algorithm, verifying that our ML model is not confusing any class or prediction. In this case we have two classes: with PPD (positive) and without PPD (negative). The predicted results must be compared with real results, with the aim to provide the true positive (TP), which are the cases that our model predict PPD and match with the actual value classified positive previously. True negative (TN), which represent the items that do not present PPD and are correctly classified as negative. FP which represents a case that the model classified as positive, but the original classification was negative. FN represent cases in which the model classified negative, but the actual classification was positive. Seven common ML metrics are used to evaluate the performance of our model, to calculate accuracy, precision, recall, F1-score, Specificity and area under curve (AUC). These metrics are described.

Accuracy: is the proportion of correct prediction out of the total predictions. This is used to analyze the classification performance It is computed in (1).

$$Accuracy = \frac{TP+TN}{Total\ Predictions\ (TP+TN+FP+FN)} \tag{1}$$

Precision: is the proportion of TP out of all positive predictions, it is calculated in (2).

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Recall (or sensitivity): is the proportion of actual positives correctly (TP) predicted by the model. A low percentage of this metric indicates that the algorithm classifies less FN items. It is computed in (3).

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

F1-score: is the harmonic mean of the precision and recall, balancing both when you need a trade-off. This is used to assess the classification performance in one measure using the harmonic average. It is calculated in (4).

$$F1 = 2 * \frac{Precission*Recall}{Precision+Recall} \tag{4}$$

Specificity: is also known as true negative rate (TNR), is a used to evaluate the performance of a ML model, particulary in binary classification tasks. It measures the proportion of actual negative cases that are correctly identified by the model, it is computed in (5).

$$Specificity = \frac{TN}{TN+FP} \tag{5}$$

AUC: measures the trade-off between the TP rate and the FP rate at various thresholds. It provides a measure of the model’s ability to distinguish between classes and different thresholds. It is calculated in (6).

$$AUC = True\ Positive\ Rate\ (TPR)\ vs\ False\ Positive\ Rate\ (FPR) \tag{6}$$

3. RESULTS AND DISCUSSION

3.1. Results

Based on classification and regression models implemented with the aim to predict the binary class of PPD. Figure 1. Presents the performance comparison of various models on the prediction of PPD for our research.

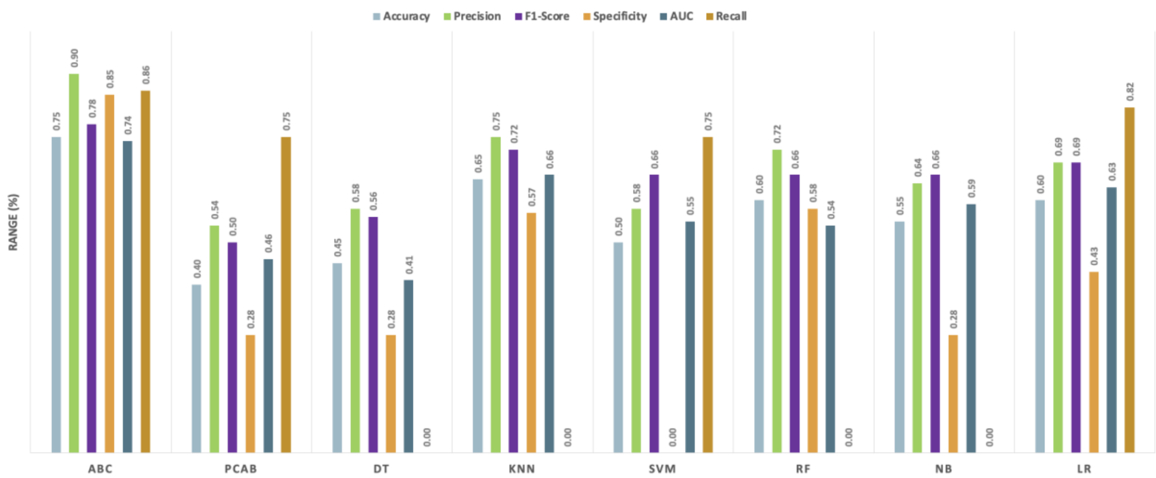

Figure 1. Performance comparison of various models of PPD

Table 1 presents the performance of the ML metrics with the different implemented algorithms. As we can see, for to the accuracy ABC outperforms the rest of classifiers with 75%, followed by KNN with 65%, and by RF and LR both with 60%. The precision for the ABC with 90% outperforms the other classifiers. This is followed by KNN with 75% and by RF with 72%. Recall for the ABC is 86% followed by LR with 82%, and PCAB together with SVM with 75%. The F1-score for ABC is 78% followed by KNN with 72% and LR with 69%. The specificity for ABC is 85% followed by RF with 58% and KNN with 57%. Finally, the AUC for ABC is 74%, which is the highest value, this is followed by KNN with 66% and LR with 63%. Our results reports that through metrics associated to the confusion matrix, and analyzing the model performance, our ABC model obtained excellent results to predict PPD.

Table 1. Performance of the ML metrics with different algorithms

Measure	ABC	PCAB	DT	KNN	SVM	RF	NB	LR
Accuracy	0.75	0.40	0.45	0.65	0.50	0.60	0.55	0.60
Precision	0.90	0.54	0.58	0.75	0.58	0.72	0.64	0.69
Recall	0.86	0.75	0.00	0.00	0.75	0.00	0.00	0.82
F1-score	0.78	0.50	0.56	0.72	0.66	0.66	0.66	0.69
Specificity	0.85	0.28	0.28	0.57	0.00	0.58	0.28	0.43
AUC	0.74	0.46	0.41	0.66	0.55	0.54	0.59	0.63

3.2. Discussion

The results of this research suggest that our model based on sociodemographic data is useful to identify PPD. There exist different proposals to predict PDD in the literature. Next paragraphs discuss about the obtained results with similar research articles that use same ML algorithms but with different datasets. Even it is not a fair comparison, because each research uses different computer resources, different size of data, and in some cases other ML algorithms, this comparison only pretends to show that our proposal has similar results to those reported in the state of the art.

For accuracy, our model obtains 75% for ABC, which is 5% less than 80% reported by Fatima *et al.* [22] for SVM, and 7% less than 82% reported by Natarajan *et al.* [11] for NB. However outperform the 67% reported by Shin *et al.* [15] for NB. This comparison shows that our proposal is close to the reported in state of the art.

In the case of precision, our model achieves 90% for ABC, which is the same result the reported by Fatima *et al.* [22] but with different ML algorithm SVM and LR. Besides, has superior performance that the reported by Natarajan *et al*. [11] and Shin *et al.* [15] for ABC with 73% and 81%. This comparison confirms that our research has good precision detecting PPD.

For recall, our model obtains 86% ABC, which is like reported by Zhang *et al.* [21] for DT. But is 4% less than 90% reported by Wang *et al.* [23] for SVM and LR, and the percentage is less than 92% reported by Shin *et al.* [15] for KNN. However, outperform the 79% reported by Natarajan *et al.* [11] for ABC. In the same sense as for precision, our proposal has good recall detecting PPD. Even is not the best performance for recall, our results are close to the reported by other proposals.

In the case of F1-score our research achieves 78% for ABC, which is similar to reported by Natarajan *et al.* [11] for ABC and Shin *et al.* [15] for SVM but is two and 4% less than 80% and 82% for LR and SVM reported by Natarajan. However, outperform 76% for ABC and 78% for SVM reported by Shin. Same as recall, even though our proposal does not have the best performance, our results are very close to those reported in the literature.

For specificity our model achieves 85% obtaining 1% more than 84% reported by by Zhang *et al.* [21] for the same algorithm ABC and DT. However, is 2% less than 87% reported for RF. Also, is 1% less than 86% reported by Shin *et al.* [15] for RF. Similar as the previous measures, our results obtain same performance to other research reported.

Finally for AUC our research obtains 74% which is 13% less than 87% reported by Shin *et al.* [15] for RF and eleven percent less than 85% for ABC. Also, this measure is 12% less than 86% reported by Zhang *et al.* [21]. Natarajan *et al.* [11] and Fatima *et al.* [22] both do not report results for this measure. Even though is not the best performance for AUC, our results are close to the reported by other proposals.

Thus, our ML model implemented with boosting adaptative classifier has demostrated better performance that other ML algorithms in some cases, but lower performance in others. However it shown that a proper classification of dataset along with correctly process of data collection, cleaning and preprocessing process, have improved the classification process with the aim to contribute to a correct classification and prediction of PDD, with similar results to those reported in literature.

4. CONCLUSION

In this paper we present a ML model to predict post partum depression for the state of Zacatecas, Mexico. The data collected has been classified previously with PPD and without PPD using socio demographic data and the result of Edimburgo Test. Our results shows that applying ML algorithms we can create a model that classify mothers at risk of PDD. We used eigth ML algorithms to create a predictive model for PPD. ABC, PCAB, DT, KNN, SVM, RF, NB, and LR. However, our ABC yielded 75% of accuracy, 90% of precision, 86% of recall, 78% of F1-score and 74% for AUC, ilustrating strong predictible capabilities to predict PDD correctly. As future work, the creation of new dataset with a large number of items, and a Mobile App for predict and classify PDD using our model can be carried on.

FUNDING INFORMATION

This research is funded by National Institute of Technology of Mexico. (Project ID M00/1582/2023).

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Lopez-Veyna J. Ivan	✓	✓		✓	✓	✓			✓	✓		✓	✓	✓
Ortiz-Garcia Mariana			✓			✓	✓	✓	✓	✓	✓		✓	✓
Diaz-Diaz Alvaro Moises	✓	✓	✓			✓		✓		✓	✓			
Bermejo-Sabbagh Carlos				✓	✓		✓			✓				

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

All authors declare no conflicts of interest financial, personal or professional.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, Lopez-Veyna J. Ivan, upon reasonable request.

REFERENCES

[1] World Health Organization, “147. Depression facts and figures,” who.int. [Online]. Available: https://www.who.int/es/news-room/fact-sheets/detail/depression. (Date accessed: Aug. 22, 2024).
[2] M. Sanchez, “Mental Health in Mexico,” www.foroconsultivo.org.mx. [Online]. Available: https://www.foroconsultivo.org.mx/INCyTU/documentos/Completa/INCYTU_18-007.pdf. (Date accessed: Aug. 22, 2024).
[3] R. Cooper, “Diagnostic and statistical manual of mental disorders (DSM),” *Knowledge Organization*, vol. 44, no. 8, pp. 668–676, 2017, doi: 10.5771/0943-7444-2017-8-668.
[4] M. E. Medina-Mora, E. Sarti, and T. Real, “Depression and other psychiatric disorders (in Mexico: *La depresión y otros transtornos psiquiátricos*),” 2015.
[5] M. Juarez-Romero, C. B. Nava-Cruz, P. Méndez-Hernández, J. Cuacenetl-Jiménez, E. Rojas-Lima, and M. Francisco-García, “Prevalence of postpartum depression at the Women’s Hospital in the state of Tlaxcala,” *South Mexico Medical Research Journal*, vol. 22, no. 3, pp. 132–135, 2016.
[6] K. Langdon, “Postpartum Depression Statistics,” postpartumdepression.org. [Online]. Available: https://www.postpartumdepression.org/resources/statistics/. (Date accessed: Aug. 22, 2024).
[7] F. De Castro, J. M. Place, A. Villalobos, and B. Allen-Leigh, “Maternal depressive symptomatology in México: National prevalence, care, and population risk profiles,” *Salud Publica de Mexico*, vol. 57, no. 2, pp. 144–154, Mar. 2015, doi: 10.21149/spm.v57i2.7410.
[8] Á. Méndez-Cerezo, “Postpartum depression and exercise,” *Perinatology and Human Reproductive Sciences*, vol. 28, no. 4, 2014.
[9] J. L. Cox, J. M. Holden, and R. Sagovsky, “Detection of Postnatal Depression: Development of the 10-item Edinburgh Postnatal

Depression scale," *British Journal of Psychiatry*, vol. 150, pp. 782–786, Jun. 1987, doi: 10.1192/bjp.150.6.782.

[10] S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, "Detecting depression and mental illness on social media: an integrative review," *Current Opinion in Behavioral Sciences*, vol. 18, pp. 43–49, Dec. 2017, doi: 10.1016/j.cobeha.2017.07.005.

[11] S. Natarajan, A. Prabhakar, N. Ramanan, A. Bagilone, K. Siek, and K. Connelly, "Boosting for Postpartum Depression Prediction," in *Proceedings - 2017 IEEE 2nd International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2017*, IEEE, Jul. 2017, pp. 232–240, doi: 10.1109/CHASE.2017.82.

[12] V. I. Osubor and A. O. Egwali, "A neuro fuzzy approach for the diagnosis of postpartum depression disorder," *Iran Journal of Computer Science*, vol. 1, no. 4, pp. 217–225, Dec. 2018, doi: 10.1007/s42044-018-0021-6.

[13] K. S. Betts, S. Kisely, and R. Alati, "Predicting common maternal postpartum complications: leveraging health administrative data and machine learning," *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 126, no. 6, pp. 702–709, May 2019, doi: 10.1111/1471-0528.15607.

[14] D. Suganthi and A. Geetha, "Predicting Postpartum Depression with Aid of Social Media Texts Using Optimized Machine Learning Model," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 3, pp. 417–427, Jun. 2024, doi: 10.22266/ijies2024.0630.33.

[15] D. Shin, K. J. Lee, T. Adeluwa, and J. Hur, "Machine learning-based predictive modeling of postpartum depression," *Journal of Clinical Medicine*, vol. 9, no. 9, pp. 1–14, Sep. 2020, doi: 10.3390/jcm9092899.

[16] W. Xu and M. C. Sampson, "Prenatal and Childbirth Risk Factors of Postpartum Pain and Depression: A Machine Learning Approach," *Maternal and Child Health Journal*, vol. 27, no. 2, pp. 286–296, Feb. 2023, doi: 10.1007/s10995-022-03532-0.

[17] 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*, IEEE, Mar. 2024, pp. 1–6, doi: 10.1109/ICITIIT61487.2024.10580321.

[18] U. K. Lilhore *et al.*, "Unveiling the prevalence and risk factors of early stage postpartum depression: a hybrid deep learning approach," *Multimedia Tools and Applications*, vol. 83, no. 26, pp. 68281–68315, Jan. 2024, doi: 10.1007/s11042-024-18182-3.

[19] 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, pp. 1–26, Dec. 2022, doi: 10.3390/math10234570.

[20] W. A. Fazraningtyas, B. Rahmatullah, D. D. Salmarini, S. A. Ariffin, and A. Ismail, "Recent advancements in postpartum depression prediction through machine learning approaches: a systematic review," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 4, pp. 2729–2737, Aug. 2024, doi: 10.11591/eei.v13i4.7185.

[21] Y. Zhang, S. Wang, A. Hermann, R. Joly, and J. Pathak, "Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women," *Journal of Affective Disorders*, vol. 279, pp. 1–8, Jan. 2021, doi: 10.1016/j.jad.2020.09.113.

[22] 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 Systems*, vol. 36, no. 4, pp. 1–13, Aug. 2019, doi: 10.1111/exsy.12409.

[23] S. Wang, J. Pathak, and Y. Zhang, "Using electronic health records and machine learning to predict postpartum depression," *Studies in Health Technology and Informatics*, vol. 264, pp. 888–892, 2019, doi: 10.3233/SHTI190351.

[24] M. Parga-Anaya and M. Garcia-Sanchez, "Caesarean section and its relationship with postpartum depression," Autonomous University of Zacatecas, Zacatecas, Mexico, 2015.

[25] J. Brownlee, *Data Preparation for Machine Learning: Data Cleaning, Feature Selection, and Data Transforms in Python*. Machine Learning Mastery, 2020.

[26] D. Cournapeau, "scikit-learn Machine Learning in Python," scikit-learn.org. [Online]. Available: https://scikit-learn.org/stable/about.html. (Date accessed: Aug. 25, 2024).

BIOGRAPHIES OF AUTHORS

Lopez-Veyna J. Ivan is a Professor at the Division of Graduate Studies and Research at Technological Institute of Zacatecas (ITZ), Mexico. He received the Engineer degree in Computer Systems from the same institute in 1997. He obtained a master's degree in computer sciences (MCC) from Technological Institute of Toluca (ITT), Mexico in 2001. And in 2014 completed his Ph.D. in Computer Sciences from the Information Technology Laboratory of the Center for Research and Advanced Studies of the National Polytechnic Institute, Mexico. He completed a postdoctoral stay at Barcelona Supercomputing Center in 2017. His research interests include machine learning, big data, large-scale storage, graph databases, and search engines. He can be contacted at email: ivanlopezveyna@zacatecas.tecnm.mx.

Ortiz-Garcia Mariana is a Professor at the Department of Systems and Computing at Technological Institute of Zacatecas (ITZ), Mexico, where she has been a faculty member since 1995. Mariana was graduate from Bachelor degree in Informatics. She received the master's degree in computer systems (MSC) from Technological Institute of Orizaba (ITT), Mexico in 2014. Her research interests are primarily in analysis, design and development of information systems, software engineering, agile methodologies, and web modeling. She can be contacted at email: mariana.ortiz@zacatecas.tecnm.mx.

Diaz-Diaz Alvaro Moises graduated as a Computer Systems Engineer in 2016 from the Technological Institute of Tapachula, Chiapas, Mexico. Currently, he is pursuing a master's degree in computer systems at the Technological Institute of Zacatecas, where he has developed a particular interest in machine learning, systems security, and mobile development. As part of his training, he has obtained various certifications in cybersecurity, including Fortinet certified fundamentals cybersecurity, Fortinet certified associate cybersecurity, and ethical hacking professional certification (CEHPC™). He can be contacted at email: alvaro.diaz@zacatecas.tecnm.mx.

Bermejo-Sabbagh Carlos is a professor at the Department of Systems and Computing at the Technological Institute of Merida (ITM), Mexico, since 1994. He is an electronic engineer and holds a Master of Science from the same Institute. His areas of interest are computer architecture and machine learning. He can be contacted at email: carlos.bs@merida.tecnm.mx.