JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DEVELOPMENT OF AN ARTIFICIAL INTELLIGENCE SUPPORTED HYBRID DATA MANAGEMENT PLATFORM FOR MONITORING DEPRESSION AND ANXIETY SYMPTOMS IN THE PERINATAL PERIOD

[1]Goranti Ramya , [2]Padmawar Srija , [3]Vollala Viha , [4]Mr. CH. Gopi

[1,2,3]UG Scholars , [4]Asst. Professor

[1,2,3,4]Department of Computer Science Engineering (Internet of Things)

Guru Nanak Institutions Technical Campus (Autonomous), Hyderabad, India

Abstract : Machine learning has emerged as a driving force in both scientific research and industrial applications, particularly in the context of handling vast amounts of data known as Big Data. With the availability of extensive healthcare datasets and continuous advancements in machine learning techniques, computers have become adept at accurately diagnosing various medical conditions. This study aims to address the pressing issues surrounding anxiety and depression in pregnant women by employing performance-optimized algorithms to extract pertinent features. The objective is to expedite the process and minimize the number of inquiries required to yield meaningful results. Building upon this, the research endeavors to develop an instantaneous remote health status prediction system tailored for assessing depression and anxiety in pregnant women. Leveraging the Apache Spark Big Data processing engine, which specializes in handling streaming Big Data, the scalable system acquires data from pregnant individuals to forecast their health condition. Subsequently, the system applies the Naïve Bayes machine learning algorithm, which has demonstrated superior performance on the dataset, achieving an accuracy of 90.8% and a precision of 81.71%. Through the integration of this Big Data platform, the traditionally time-consuming procedure of detecting anxiety and depression in pregnant women can be replaced by a computer-based technique that operates swiftly while maintaining a respectable level of accuracy.

Index Terms **:** Big Data Analytics Framework , Perinatal Mental Health , Machine Learning Techniques , Depression and Anxiety Disorders , Feature Selection , Hybrid Machine Learning , Scalable Big Data Platform , Rapid Disease Diagnosis

I. INTRODUCTION

The contemporary era, particularly the last two decades, has been characterized as the "age of big data," witnessing the increasing importance of digital data across various domains, including research, technology, society, and healthcare. However, the sheer volume and complexity of these datasets pose significant challenges to computing resources and analytics frameworks, hindering the efficient extraction of valuable insights. Consequently, there is a pressing need to develop effective big data analytics frameworks to address these challenges. Several research initiatives have focused on creating such frameworks, leveraging technologies like Apache Spark, Apache Hadoop, Apache Storm, and Apache Kafka to tackle healthcare-related issues. Mental health disorders, in particular, represent a significant healthcare concern due to their profound impact on individuals' lives, especially during the perinatal period encompassing pregnancy. Failure to timely detect psychological disorders during this critical phase can have detrimental effects on both mothers and infants, underscoring the importance of addressing mental health issues in society. Machine learning emerges as a powerful tool in informing illness models, drug discovery, and prevention strategies in psychiatry. Leveraging machine learning algorithms within big data analytics platforms holds promise in advancing the understanding and treatment of mental health disorders. However, existing techniques for recognizing brain and mental disorders have often yielded incomplete or inaccurate representations, highlighting the need for advancements in big data analytics to better address psychological disorders. The perinatal period, in particular, is characterized by a high prevalence of depression and anxiety disorders, with studies indicating global prevalence rates ranging from 10% to 20% for depression and 10% to 24% for perinatal anxiety disorders. Despite their high prevalence, these

disorders are often underdiagnosed, leading to adverse consequences for both mothers and babies. Various risk factors, including unplanned pregnancies, low socioeconomic status, and exposure to violence, further complicate the diagnosis and treatment of these disorders. Given the potential reach and impact of addressing mental health issues during the perinatal period, there is a need for scalable and efficient big data platforms to facilitate analysis and accelerate the diagnostic process. In our study, we address this challenge by employing hybrid machine learning techniques that integrate feature selection methods and classification algorithms. The resulting big data architecture enables scalable and efficient analytics, allowing for the rapid identification of pregnant women at risk of anxiety and depression. Our approach involves preprocessing patient data, applying feature selection algorithms, and utilizing artificial intelligence techniques to develop models with acceptable sensitivity. We employ a 10-Fold cross-validation technique to validate model performance and identify the machine learning algorithm with the highest performance for deployment on the big data platform. The resulting system enables instant disease diagnosis and serves as a scalable infrastructure for handling larger datasets, thus minimizing the potential harm caused by depression and anxiety in women during the perinatal period.

II. LITERATURE SURVEY

In recent years, heightened attention has been directed towards perinatal mental health, prompting investments in specialized mental health services, including inpatient psychiatric mother and baby units, particularly in affluent nations. Howard and Khalifeh (2023) critically examine the epidemiology and impact of perinatal mental disorders, emphasizing their increased prevalence among young pregnant women. Such disorders rank among the most common morbidities during pregnancy, significantly contributing to maternal mortality and adverse outcomes for neonates, infants, and children. The review delves into current evidence-based interventions, encompassing individual-level and public health approaches, alongside various service delivery models. While randomized controlled trials demonstrate the effectiveness of psychological and psychosocial interventions at the individual level, determining the necessity for additional support for parenting remains uncertain for women with perinatal mental disorders. Research on psychotropic use during pregnancy primarily relies on observational studies.

Abdel-Fattah, Othman, and Goh (2022) highlight chronic kidney disease (CKD) as a prevalent ailment posing serious risks, such as cardiovascular disease and end-stage renal disease. Early detection and treatment of individuals at risk of CKD can potentially mitigate these risks. Machine learning algorithms integrated with Big Data platforms, such as Apache Spark, offer significant support to accurately diagnose CKD at its early stages.

Kanakasabapathy et al. (2022) investigate the potential of deep learning techniques for blastocyst selection in in vitro fertilization (IVF), aiming to democratize access to care by utilizing assistive tools compatible with conventional microscopes. Their study evaluates the efficacy of a deep convolutional neural network (CNN) trained on single time-point images of embryos, achieving an impressive accuracy rate in identifying the highest quality embryo for transfer. Moreover, a CNN trained to assess an embryo's implantation potential surpasses the performance of trained embryologists across different fertility centers.

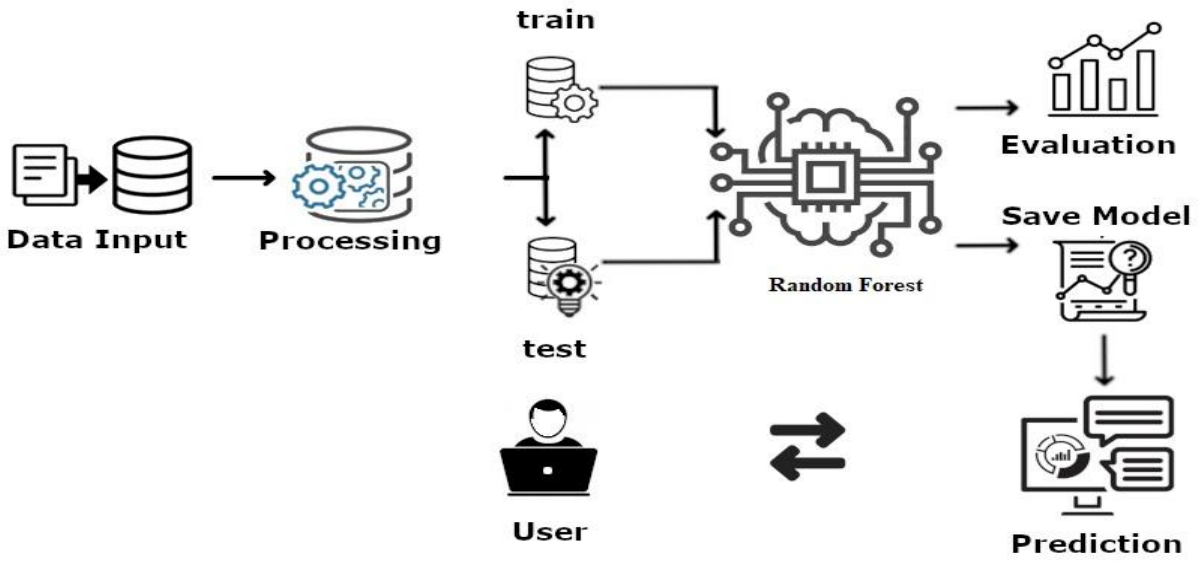
Chen et al. (2023) employ various machine learning algorithms to develop models for classifying and predicting breast cancer, aiming for early diagnosis. Their research contributes to advancing breast cancer diagnosis by offering insights into model performance and the potential for early detection and intervention.

Priya, Garg, and Tigga (2022) focus on predicting anxiety, depression, and stress using machine learning algorithms in today's fast-paced society. They address the imbalance in classes within the confusion matrix and identify the Random Forest classifier as the optimal choice for predicting these psychological conditions.

III. PROPOSED SYSTEM

We present a robust framework for detecting anxiety in pregnant women, utilizing supervised learning techniques, with a particular emphasis on Random Forest algorithms. Our primary evaluation metric is the handover failure rate, which comprehensively assesses mobility concerns, encompassing both premature and delayed handovers. Through this framework, we effectively identify instances of depression and anxiety among pregnant women. Random Forest, functioning as an ensemble learning approach, excels in capturing non-linear relationships and intricate dependencies among various health indicators. In our project aimed at predicting the health status of pregnant women, the versatility of Random Forest in constructing multiple decision trees and aggregating their predictions leads to the development of a more precise and resilient model. This algorithm proves adept at handling the complexities and subtleties inherent in health-related datasets, thus serving as a dependable tool for the creation of an instantaneous remote health status prediction system. Proficient at capturing non-linear associations and analyzing intricate datasets effortlessly. Capable of prioritizing critical factors during assessments and interventions effectively. Establishes an ensemble approach to augment robustness and reliability.

IV. SYSTEM ARCHITECTURE



V. HARDWARE REQUIREMENTS

The hardware requirements serve as a comprehensive specification for the system implementation and are essential for initiating the system design process. They outline what the system should encompass rather than dictating how it should be executed.

- Processor: Dual-core Intel Core 2 Duo
- RAM: 4GB DDR RAM
- Hard Disk: 250GB Storage Capacity

VI. SOFTWARE REQUIREMENTS

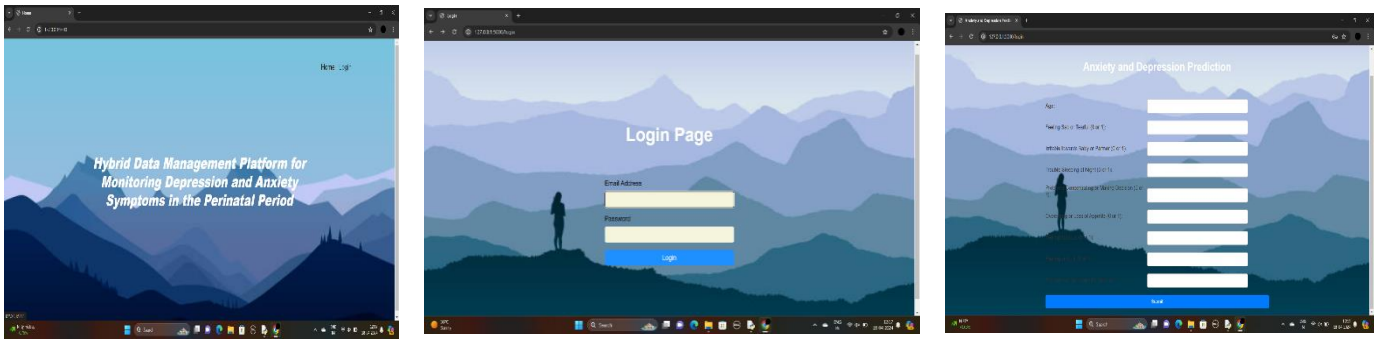
The software requirements document serves as a specification for the system, outlining what the system should accomplish rather than detailing the methods for achieving it. It provides a foundation for creating the software requirements specification and is instrumental in estimating costs, planning team activities, performing tasks, and tracking progress throughout the development process.

- Operating System: Windows 7/8/10
- Platform: Spyder3
- Programming Language: Python
- Front End: Spyder3

VII. FUTURE ENCHANCEMENT

The developed system has undergone testing for the immediate identification of anxiety and depression in expectant mothers. This study underscores the significance of employing hybrid big data platforms with machine learning technology for detecting mental health disorders. The proposed architecture shows promise for real-time big data processing applications, as it integrates scalable and high-performance data analytics tools such as Apache Kafka and Spark. By leveraging this architecture, the time-intensive analysis of anxiety and depression could be substituted with an automated computer-based approach, offering a reasonable level of accuracy instantaneously.

VIII. SNAPSHOTS



IX. CONCLUSION

In various stages of life, individuals may encounter situations that adversely affect their psychological well-being. Among these, the perinatal period emerges as a critical phase significantly impacting mental health. Without timely intervention during this period, it is evident that adverse consequences will manifest in both the child and the mother, subsequently affecting societal mental health. Recognizing the urgent need for a system facilitating faster and more accessible diagnosis and treatment, our study aims to bridge

this gap by leveraging a hybrid big data analytics platform. Focusing on hybrid big data platforms employing machine learning techniques for mental health condition detection, our approach aims to streamline the process for pregnant women, minimizing the burden of questionnaire-based assessments. We employed feature selection algorithms, including Sequential Forward (SFS), Sequential Backward (SBS), and Optimized feature selection techniques, to identify key data attributes crucial for detecting anxiety and depression in pregnant women. Notably, our findings revealed that the optimized selection technique incorporating genetic algorithms demonstrated superior performance. Furthermore, we evaluated six machine learning classifiers, namely Decision Tree, Naive Bayes, K-NN, Random Forest, GBT, Logistic Regression, and DFFNN, optimizing their parameters using cross-validation. Utilizing six evaluation methods, including accuracy, sensitivity, specificity, and precision, we validated the results with testing data. Our results indicate that the Naive Bayes Classifier, integrated with selected features, achieved outstanding performance with 90.80% accuracy. This high accuracy rate underscores the effectiveness of our developed model for diagnosis purposes. Notably, the Naive Bayes-based machine learning model was implemented on the Apache Spark platform, facilitating seamless interaction with streaming data transmitted via Apache Kafka to Apache Spark. These findings highlight the potential of our model to significantly enhance the efficiency and efficacy of mental health diagnosis and treatment during the perinatal period.

X. REFERENCES

[1] L. R. Nair, S. D. Shetty, and S. D. Shetty, ‘‘Utilizing Spark-Based Machine Learning Model on Streaming Big Data for Health Status Prediction,’’ in Computational and Electrical Engineering, vol. 65, pp. 393–399, Jan. 2018, doi: 10.1016/j.compeleceng.2017.03.009.

[2] Apache Spark. (2018). Apache SparkTM—Unified Analytics Engine for Big Data. [Online].

[3] Apache Hadoop. Accessed: Oct. 26, 2022. [Online].

[4] Apache Storm. Accessed: Oct. 26, 2022. [Online].

[5] Apache Kafka. Accessed: Oct. 27, 2022. [Online].

[6] A. M. Y. Tai, A. Albuquerque, N. E. Carmona, M. Subramanieapillai, D. S. Cha, M. Sheko, Y. Lee, R. Mansur, and R. S. McIntyre, ‘‘Machine Learning and Big Data: Implications for Disease Modeling and Therapeutic Discovery in Psychiatry,’’ in Artificial Intelligence in Medicine, vol. 99, Aug. 2019, Art. no. 101704, doi: 10.1016/j.artmed.2019.101704.

[7] D. Bzdok and A. Meyer-Lindenberg, ‘‘Machine Learning for Precision Psychiatry: Opportunities and Challenges,’’ in Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, vol. 3, no. 3, pp. 223–230, Mar. 2018, doi: 10.1016/j.bpsc.2017.11.007.

[8] N. B. Dorio, S. S. Fredrick, and M. K. Demaray, ‘‘School Engagement and the Role of Peer Victimization, Depressive Symptoms, and Rumination,’’ in Journal of Early Adolescence, vol. 39, no. 7, pp. 962–992, Aug. 2018, doi: 10.1177/0272431618797007.

[9] L. M. Howard and H. Khalifeh, ‘‘Perinatal Mental Health: A Review of Progress and Challenges,’’ in World Psychiatry, vol. 19, no. 3, pp. 313–327, Oct. 2020, doi: 10.1002/WPS.20769.

[10] C.-L. Dennis, K. Falah-Hassani, and R. Shiri, ‘‘Prevalence of Antenatal and Postnatal Anxiety: Systematic Review and Meta-Analysis,’’ in British Journal of Psychiatry, vol. 210, no. 5, pp. 315–323, May 2017, doi: 10.1192/BJP.BP.116.187179.

[11] C. Reck, K. Struben, M. Backenstrass, U. Stefenelli, K. Reinig, T. Fuchs, C. Sohn, and C. Mundt, ‘‘Prevalence, Onset and Comorbidity of Postpartum Anxiety and Depressive Disorders,’’ in Acta Psychiatrica Scandinavica, vol. 118, no. 6, pp. 459–468, Dec. 2008, doi: 10.1111/j.1600-0447.2008.01264.x.

[12] S. Misri, J. Abizadeh, S. Sanders, and E. Swift, ‘‘Perinatal Generalized Anxiety Disorder: Assessment and Treatment,’’ in Journal of Women's Health, vol. 24, no. 9, pp. 762–770, Sep. 2015, doi: 10.1089/JWH.2014.5150.

[13] D. Aslan and R. Edelmann, ‘‘Demographic and Offence Characteristics: A Comparison of Sex Offenders Convicted of Possessing Indecent Images of Children, Committing Contact Sex Offences or Both Offences,’’ in Journal of Forensic Psychiatry & Psychology, vol. 25, no. 2, pp. 121–134, Mar. 2014, doi: 10.1080/14789949.2014.884618.

[14] T. S. Kirkan, N. Aydin, E. Yazici, P. Akcali Aslan, H. Acemoglu, and A. G. Daloglu, ‘‘The Depression in Women in Pregnancy and Postpartum Period: A Follow-up Study,’’ in International Journal of Social Psychiatry, vol. 61, no. 4, pp. 343–349, Jun. 2015, doi: 10.1177/0020764014543713.

[15] D. Hubbeling, ‘‘Overtreatment: Is a Solution Possible?’’ in Journal of Evaluation in Clinical Practice, vol. 28, no. 5, pp. 821–827, Oct. 2022, doi: 10.1111/jep.13632.

[16] M. Furtado, C. H. T. Chow, S. Owais, B. N. Frey, and R. J. Van Lieshout, ‘‘Risk Factors of New Onset Anxiety and Anxiety Exacerbation in the Perinatal Period: A Systematic Review and Meta-Analysis,’’ in Journal of Affective Disorders, vol. 238, pp. 626–635, Oct. 2018, doi: 10.1016/J.JAD.2018.05.073.

[17] J. H. Goodman, K. L. Chenausky, and M. P. Freeman, ‘‘Anxiety Disorders During Pregnancy: A Systematic Review,’’ in Journal of Clinical Psychiatry, vol. 75, no. 10, p. 1177, Oct. 2014, doi: 10.4088/JCP.14R09035.

[18] M. A. Abdel-Fattah, N. A. Othman, and N. Goher, ‘‘Predicting Chronic Kidney Disease Using Hybrid Machine Learning Based on Apache Spark,’’ in Computational Intelligence and Neuroscience, vol. 2022, pp. 1–12, Feb. 2022, doi: 10.1155/2022/9898831.

[19] C. L. Bormann et al., ‘‘Performance of a Deep Learning Based Neural Network in the Selection of Human Blastocysts for Implantation,’’ in eLife, vol. 9, pp. 1–14, Sep. 2020, doi: 10.7554/ELIFE.55301

[20] Y. Arai et al., ‘‘Prediction of Acute Graft-Versus-Host Disease Following Allogeneic Hematopoietic Stem Cell Transplantation Using a Machine Learning Algorithm,’’ in Blood, vol. 132, p. 68, Nov. 2018, doi: 10.1182/BLOOD-2018-99-114794.

[21] H. Chen et al., ‘‘Classification Prediction of Breast Cancer Based on Machine Learning,’’ in Computational Intelligence and Neuroscience, vol. 2023, Jan. 2023, Art. no. 6530719.

[22] Yaseen, G.-Y. Son, and S. Kwon, ‘‘Classification of Heart Sound Signal Using Multiple Features,’’ in Applied Sciences, vol. 8, no. 12, p. 2344, Nov. 2018, doi: 10.3390/APP8122344.

[23] G. Zini, ‘‘Artificial Intelligence in Hematology,’’ in Hematology, vol. 10, no. 5, pp. 393–400, Oct. 2005, doi: 10.1080/10245330410001727055.

[24] A. Subasi, ‘‘EEG Signal Classification Using Wavelet Feature Extraction and a Mixture of Expert Model,’’ in Expert Systems with Applications, vol. 32, no. 4, pp. 1084–1093, May 2007, doi: 10.1016/J.ESWA.2006.02.005.
[25] A. Priya, S. Garg, and N. P. Tigga, ‘‘Predicting Anxiety, Depression and Stress in Modern Life Using Machine Learning Algorithms,’’ in Proceedings of Computer Science, vol. 167, pp. 1258–1267, Jan. 2020, doi: 10.1016/J.PROCS.2020.03.442.
[26] A. Sau and I. Bhakta, ‘‘Screening of Anxiety and Depression Among the Seafarers Using Machine Learning Technology,’’ in Information in Medicine: Unlocked, vol. 16, Jan. 2019, Art. no. 100149, doi: 10.1016/J.IMU.2018.12.004.
[27] P. J. C. Suen et al., ‘‘Classification of Unipolar and Bipolar Depression Using Machine Learning Techniques,’’ in Psychiatry Research, vol. 295, Jan. 2021, Art. no. 113624, doi: 10.1016/J.PSYCHRES.2020.113624.
[28] Y. Luo et al., ‘‘Improvement of APACHE II Score System for Disease Severity Based on XGBoost Algorithm,’’ in BMC Medical Informatics and Decision Making, vol. 21, no. 1, pp. 1–12, Dec. 2021, doi: 10.1186/S12911-021-01591-X.
[29] S. Zhou et al., ‘‘Machine Learning to Predict Clinical Remission in Depressed Patients After Acute Phase Selective Serotonin Reuptake Inhibitor Treatment,’’ in Journal of Affective Disorders, vol. 287, pp. 372–379, May 2021, doi: 10.1016/J.JAD.2021.03.079.