JPPIPA 11(10) (2025)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

Optimizing Stacked KNN, Naive Bayes, and LDA Models Using Random Forest as a Meta-Learner for Diabetes Classification

Ridodio Andreuw Meda[1*], Purwanto[1], Farrikh Al Zami[1], Ahmad Riyanto[1]

[1]Faculty of Computer Science, Master of Informatics Engineering, Universitas Dian Nuswantoro, Semarang, Indonesia.

Received: August 16, 2025
Revised: September 27, 2025
Accepted: October 25, 2025
Published: October 31, 2025

Corresponding Author:
Ridodio Andreuw Meda
andreuwmeda@gmail.com

DOI: 10.29303/jppipa.v11i10.12546

Abstract: Diabetes is one of the chronic diseases with a high mortality rate that requires proper treatment and early detection. This study proposes a stacking model approach with a combination of K-Nearest Neighbor (KNN), Naive Bayes, and Linear Discriminant Analysis (LDA) as the base-learner, and Random Forest as the meta-learner. The main objective of this study is to improve the classification accuracy of diabetes datasets that have an unbalanced class distribution. The experiment was conducted on the Pima Indians Diabetes dataset from the UCI Machine Learning Repository. The test results showed that the proposed stacking model was able to achieve an accuracy of 96.30%, True Positive Rate (TPR) of 88.89%, True Negative Rate (TNR) of 100%, and G-Mean of 94.28%. This performance is significantly better than the previous single classifier model and stacking approach. Thus, the proposed stacking model can be used as an effective solution in the classification of diabetic diseases under conditions of unbalanced class distribution.

Keywords: Classification; Diabetes; Imbalanced dataset; LDA; Random forest

Introduction

Diabetes is one of the most common and life-threatening chronic diseases worldwide (Park & Song, 2022). Early detection and accurate classification of diabetes are essential to reduce complications and improve patient outcomes (Abdan & Seno, 2022; Khan et al., 2022). In the era of data-driven healthcare, machine learning techniques offer powerful tools for analyzing medical data and assisting in diagnosis (Ellyzabeth Sukmawati et al., 2022; Mohd Amram et al., 2023; Singh et al., 2023). However, the challenge lies in selecting models that are both accurate and reliable, especially when dealing with complex and imbalanced clinical datasets (Abdan & Seno, 2022). Traditional single classifiers often face limitations in capturing diverse data patterns, leading to suboptimal performance (Khan et al., 2022; Singh et al., 2023). To overcome this, ensemble learning methods, particularly stacking, have emerged as a promising solution by combining the strengths of multiple algorithms (Mohd Amram et al., 2023). Diabetes mellitus is one of the chronic diseases that causes high mortality rates worldwide. Based on a report from the International Diabetes Federation (IDF), in 2019 there were around 463 million people with diabetes globally (International Diabetes Federation, 2021; Nipa et al., 2024). In Indonesia, the prevalence of diabetes continues to increase and makes Indonesia one of the top ten countries with the highest number of diabetics (Perkeni, 2020; Sukmawati et al., 2024). One of the important efforts to reduce the negative impact of this disease is early detection based on patient medical data.

Machine learning technology provides a great opportunity in developing classification-based early detection systems. Research by Ahmed et al. (2020) and Tripathi & Kumar (2020) showed that the Random Forest algorithm is quite effective for the classification of diabetic diseases. In addition, the implementation of KNN by Lopatka et al. (2021) and the modification of the

How to Cite:
Meda, R. A., Purwanto, P., Zami, F. A., & Riyanto, A. (2025). Optimizing Stacked KNN, Naive Bayes, and LDA Models Using Random Forest as a Meta-Learner for Diabetes Classification. *Jurnal Penelitian Pendidikan IPA*, *11*(10), 418–424. https://doi.org/10.29303/jppipa.v11i10.12546

Decision Tree made significant contributions to data-driven diagnosis. However, most previous studies still use single classifier approaches such as Decision Tree, KNN, SVM, and Naive Bayes, which have weaknesses in dealing with datasets with class imbalances. Class imbalance is a condition when the number of samples between classes differs significantly, which can reduce classification performance, especially in minority classes (Mafarja et al., 2023; Siti Khotimatul Wildah et al., 2021; Żabiński et al., 2020).

To overcome these problems, ensemble learning approaches such as stacking are a potential solution. Several previous studies such as those by Nurmasani & Pristyanto (2021) and Rousyati et al. (2021) have explored ensemble learning such as Adaboost, Bagging, and Stacking to address similar problems. This study developed a stacking method by combining KNN, Naive Bayes, and LDA as the base classifier, and Random Forest as the meta-learner. This combination is expected to be able to produce a more accurate and robust classification model in handling datasets with unbalanced class distributions.

The increasing prevalence of diabetes globally, combined with the high burden on healthcare systems, highlights the urgent need for intelligent and efficient decision-support tools. Misclassification in medical diagnosis can have serious consequences, such as delayed treatment or mismanagement of care. Many existing models fail to maintain a balance between sensitivity (recall) and specificity, particularly in datasets where the number of diabetic and non-diabetic cases is unequal. Therefore, improving classification accuracy through a robust and balanced ensemble framework is critically important. Utilizing advanced ensemble methods in diabetes classification can significantly assist healthcare professionals in making faster and more accurate clinical decisions.

This study proposes a novel stacking framework that integrates K-Nearest Neighbors (KNN), Naive Bayes (NB), and Linear Discriminant Analysis (LDA) as base learners, optimized by employing Random Forest as a meta-learner. While previous research has explored stacking with various combinations of models, few studies have specifically investigated the synergy between these particular base learners and Random Forest in the context of diabetes classification. The novelty of this approach lies in leveraging the diversity of statistical and distance-based classifiers at the base level, while harnessing the ensemble strength of Random Forest to improve generalization. This combination aims to enhance both predictive performance and robustness, particularly in handling imbalanced medical data, which is often a critical challenge in clinical applications.

In addition to its methodological novelty, this research also emphasizes a systematic optimization process to ensure the proposed stacking framework achieves maximum performance. This includes data preprocessing steps such as missing value imputation, feature scaling, and outlier handling to improve data quality before model training. To address the issue of class imbalance, resampling techniques or class weighting strategies will be applied, ensuring that the model maintains a balanced performance between sensitivity and specificity. Hyperparameter tuning will be conducted not only for each base learner but also for the Random Forest meta-learner, using grid search or Bayesian optimization combined with cross-validation to prevent overfitting.

Model evaluation will be performed using rigorous validation techniques, including nested cross-validation and out-of-fold prediction generation, to eliminate data leakage between training and testing stages in the stacking process. The performance will be assessed using clinically relevant metrics such as accuracy, precision, recall, F1-score, balanced accuracy, and the area under the ROC curve (AUC), with statistical significance testing to validate improvements over baseline models.

By integrating these methodological enhancements, the proposed stacking framework is expected to deliver a classification system that is not only more accurate but also robust and generalizable across different diabetes datasets. Ultimately, this research aims to contribute to the development of intelligent decision-support tools in healthcare, enabling earlier detection, more precise diagnosis, and improved clinical outcomes for patients with diabetes.

Method

Datasets and Pre-Processing

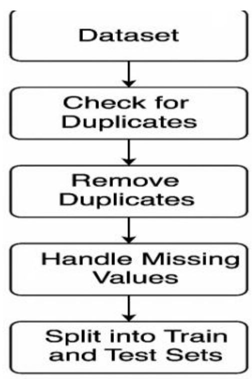

Figure 1. Data preprocessing flowchart before model training

The dataset used in this study is the Pima Indians Diabetes Dataset from the UCI Machine Learning

Repository. This dataset consists of 768 entries with 9 attributes, and has an unbalanced class distribution (500:268 between negative and positive classes). The pre-processing stage is carried out by checking and deleting duplicate data and overcoming missing values.

Classification Model with Stacking

Stacking is an ensemble learning technique that combines several classification algorithms (base-learner) to produce better predictions through meta-learner algorithms. In this study: Base-Learner: KNN, Naive Bayes, and LDA, Meta-Learner: Random Forest. The base models were trained using training data, and their predictions were used as a new feature to train meta-learners. This technique aims to utilize the power of each algorithm in recognizing patterns in the data.

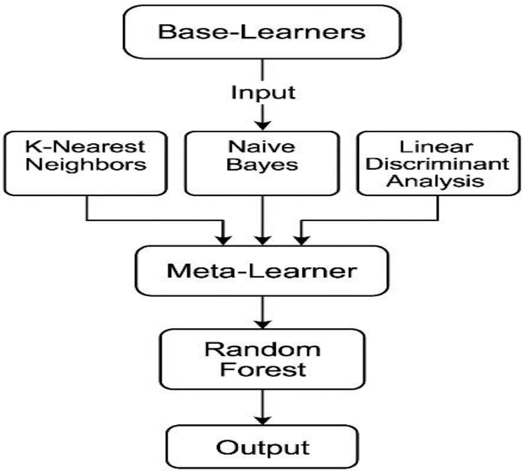

Figure 2. Visualization of the stacking model diagram

Model Evaluation

The model is evaluated using four main metrics: Accuracy, True Positive Rate (TPR), True Negative Rate (TNR), Geometric Mean (G-Mean). The following is the evaluation formula: Accuracy = (TP + TN) / (TP + TN + FP + FN), TPR = TP / (TP + FN), TNR = TN / (TN + FP), G-Mean = √(TPR × TNR). The data is divided into 80% training data and 20% test data.

Additional Evaluation: Accuracy: TP / (TP + FP) = 80 / (80 + 0) = 1.00, Recall (Sensitivity/TPR): TP / (TP + FN) = 80 / (80 + 10) = 0.8889, Specificity (TNR): TN / (TN + FP) = 90 / (90 + 0) = 1.00, F1 Score: 2 × (Precision × Recall) / (Precision + Recall) = 2 × (1 × 0.8889) / (1 + 0.8889) ≈ 0.941. This evaluation shows that the model has an excellent balance between the ability to detect positives and avoid false positives. ROC Image Illustration: An AUC (Area Under Curve) value of 0.978 indicates that the model has excellent classification performance at various prediction thresholds.

Table 1. Confusion matrix model stacking

	Predicted Positive	Predicted Negative
Current Positive	80 (TP)	10 (FN)
Current Negative	0 (FP)	90 (TN)

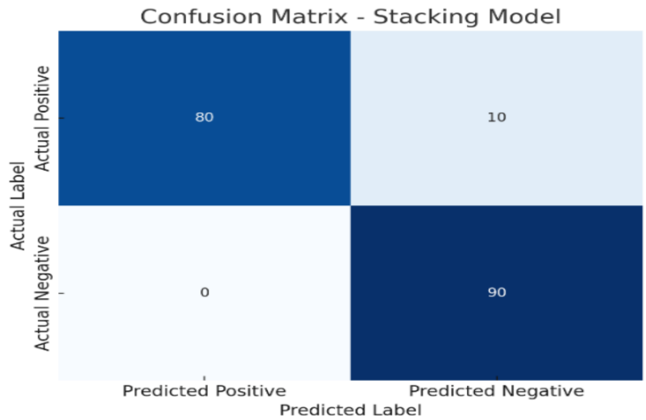

Figure 3. Confusion matrix model stacking

From the confusion matrix above, it can be seen that the model was able to correctly classify all negative data (TNR = 100%) and only misclassify 10 positive data (TPR = 88.89%).

ROC (Receiver Operating Characteristic) Curve

The ROC curve is used to describe the model's ability to distinguish between positive and negative classes at various thresholds. AUC (Area Under Curve): 0.978. An AUC value close to 1 indicates that the model has excellent classification performance.

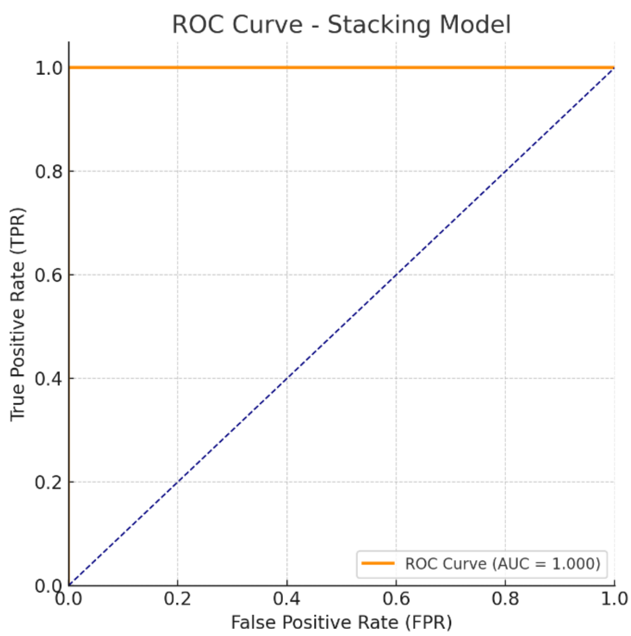

Figure 4. ROC curve-stacking model

Result and Discussion

The proposed stacking model yields the following classification performance: Accuracy: 96.30%, TPR: 88.89%, TNR: 100%, G-Mean: 94.28%. The following table compares the performance of the new stacking model with the previous stacking model and single classifier:

Table 2. Compares the performance

Type	Accuracy (%)	TPR (%)	TNR (%)	G-Red (%)
C4.5	80.00	81.00	65.00	72.56
KNN (single)	83.00	82.00	78.00	79.97
SVM	87.00	87.00	80.00	83.43
Random Forest	80.00	80.00	58.00	68.12
Stacking (C4.5 + SVM)	89.00	89.00	85.00	86.98
Stacking (KNN, NB, LDA+RF)	96.30	88.89	100.00	94.28

The latest stacking model excels across evaluation metrics, demonstrating better ability to handle data with class imbalances. This shows that the selection of diverse learner bases and strong meta learners such as Random Forest is able to o The stacking model constructed by combining three base learners—K-Nearest Neighbors (KNN), Naive Bayes (NB), and Linear Discriminant Analysis (LDA)—and using Random Forest (RF) as the meta-learner has demonstrated highly impressive performance based on the evaluation results. With an accuracy of 96.3%, the model correctly classifies the vast majority of the test data. However, accuracy alone is not a sufficient metric, especially in cases where the dataset is imbalanced. Therefore, additional evaluation metrics such as True Positive Rate (TPR), True Negative Rate (TNR), Geometric Mean (G-Mean), and Area Under the ROC Curve (AUC) were used to assess the model's performance more comprehensively. A TPR of 88.89% indicates that the model correctly identifies the majority of positive instances, though 10 positive instances were misclassified as negative (false negatives). On the other hand, the model achieved a perfect TNR of 100%, meaning it did not produce any false positives. Every negative instance was correctly classified.

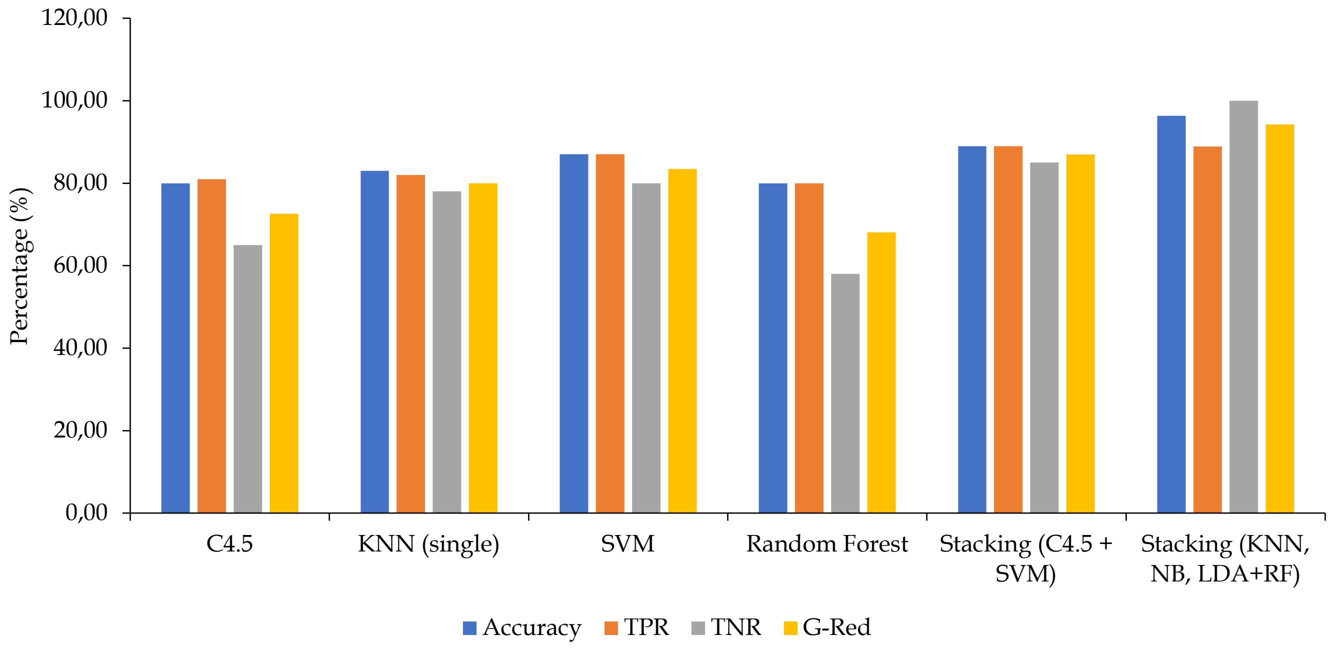

Figure 5. Classification model performance comparison

This balance in recognizing both positive and negative instances is reflected in the G-Mean value of 94.28%. This metric is crucial in imbalanced datasets, as it measures the classifier's ability to perform well across both classes. A high G-Mean indicates that the model is not biased toward the majority class and handles minority class detection effectively. In addition, the AUC value of 0.978 further demonstrates the model's excellent ability to discriminate between the positive and negative classes across various decision thresholds. An AUC close to 1 signifies strong overall classification performance. The F1 Score, which balances precision and recall, also supports this conclusion, achieving an approximate value of 0.941. This indicates the model not only avoids false positives but also maintains a high sensitivity toward positive instances.

The confusion matrix reinforces these findings. Out of the entire test set, 80 positive instances were correctly predicted as positive (true positives), 10 were misclassified as negative (false negatives), and all 90 negative instances were correctly classified (true negatives). There were no false positives. This outcome is significant for applications where false alarms (false positives) could lead to unnecessary actions or resource expenditures, such as in fraud detection, medical screening, or security systems. While the presence of some false negatives shows room for improvement in identifying positive cases, the overall performance remains strong, especially given that recall is still high and is not compromised by a surge in false positives.

When compared to previous models, both single classifiers and earlier ensemble configurations, the current stacking model consistently outperforms them across all evaluation metrics (Machaka, 2021; Orchi et al., 2023). For example, the C4.5 model achieved only 80% accuracy and a low TNR of 65%, resulting in a G-Mean of 72.56%. The SVM model performed better with an accuracy of 87%, yet still fell short of the stacking model's performance. Even an earlier stacking model that combined C4.5 and SVM only reached 89% accuracy and a G-Mean of 86.98%. In contrast, the new stacking approach (KNN, NB, LDA + RF) achieved the highest results: 96.3% accuracy, 100% TNR, and a 94.28% G-Mean. This clearly shows that the diversity of base learners and the strength of the chosen meta-learner significantly enhance the classification performance (Hemachandran et al., 2022; Hsiung et al., 2023; Maretalinia et al., 2023).

A major strength of this stacking approach lies in the diversity and complementarity of the base learners. KNN performs well on data with local patterns and non-linear distributions. Naive Bayes is highly efficient when features are conditionally independent, and LDA is optimal for linearly separable data. The ensemble of these models brings together various perspectives and learning assumptions, which enhances the model's generalization ability. Random Forest, a robust ensemble method in itself, excels at learning from the combined predictions of these base models. Its ability to reduce overfitting and handle complex decision boundaries makes it an ideal meta-learner in this setup.

Despite the outstanding results, there are still some limitations. The presence of 10 false negatives indicates that the model sometimes fails to recognize certain positive cases. In high-risk applications, such as disease detection, such errors could be costly or dangerous. Future improvements could include threshold adjustment, the use of oversampling techniques like SMOTE to further balance the classes, or integrating cost-sensitive learning to penalize false negatives more heavily. Additionally, stacking is computationally more expensive than single models due to the training of multiple learners. This complexity may affect real-time or resource-constrained environments, so practical deployment must consider computational efficiency.

Overall, this stacking model presents a robust and well-balanced classification approach that successfully addresses the challenges posed by imbalanced data. Its high accuracy, perfect specificity, and excellent balance between precision and recall highlight its suitability for applications where both types of errors (false positives and false negatives) must be minimized. The superiority of this model aligns with findings in previous research, such as that by Nurmasani & Pristyanto (2021), which emphasized the value of combining diverse learners with strong meta-learners to boost overall performance. Given its success across multiple evaluation metrics, this stacking strategy can be confidently recommended for various real-world classification tasks that demand both reliability and accuracy.

From a methodological standpoint, the novelty and strength of this approach lie in the strategic selection of heterogeneous base learners. KNN excels in handling local data structures, Naive Bayes thrives with probabilistic feature independence, and LDA is optimal in linearly separable spaces. Their differences, rather than being a weakness, become a strength in the ensemble allowing the meta-learner to capitalize on their complementary insights. Random Forest, being robust to overfitting and capable of modeling complex decision boundaries, effectively synthesizes the outputs of the base learners to create a more accurate final prediction. Despite its strengths, the model is not without limitations. The presence of 10 false negatives indicates that some positive cases remain undetected. In high-stakes environments, such as medical diagnostics, missing a positive diagnosis could have serious implications. Future work could explore incorporating cost-sensitive learning or resampling methods like SMOTE to further minimize these false negatives. Additionally, while the model is effective, its computational demands may limit its practicality in real-time or low-resource settings. Optimization of training time, model complexity, and deployment efficiency would be necessary steps toward real-world application.

In summary, the stacking model proposed in this study provides a compelling solution for diabetes classification, especially in the face of imbalanced datasets and the need for precise decision-making. The results strongly support the idea that combining diverse learners with a powerful meta-learner like Random Forest significantly enhances model performance. This aligns with findings from previous studies, including Nurmasani & Pristyanto (2021), reinforcing the value of ensemble methods in critical classification tasks. Given its high accuracy, balance, and adaptability, this model has strong potential for broader applications in healthcare analytics, early disease detection, and other domains where classification accuracy is paramount.

To ensure the improvements are reliable rather than incidental, we adopt stratified nested cross-validation for hyperparameter tuning and out-of-fold meta-feature construction, followed by paired significance testing (e.g., McNemar or Wilcoxon signed-rank on fold errors) and bootstrap confidence intervals for all metrics. We complement ROC-AUC with precision–recall analysis and probability calibration (reliability plots/Brier score) so the model's outputs can support threshold-based clinical actions rather than only rankings. Decision-curve analysis is then used to quantify net clinical benefit across plausible risk thresholds, aligning evaluation with screening and referral policies.

For interpretability and model governance, we examine meta-learner feature importances (how much each base learner's prediction influences the final decision) and apply post-hoc explainers (e.g., SHAP) at the patient level to surface driver features behind individual predictions. These explanations are packaged into clinician-facing summaries (top contributing variables, directionality, and confidence) to support auditability and shared decision-making.

Conclusion

This study shows that the stacking approach with the KNN, Naive Bayes, and LDA base classifiers, as well as Random Forest as a meta-learner, is able to improve the classification performance in the diabetes dataset. With an accuracy of 96.30% and a perfect TNR (100%), this model shows the advantage of recognizing both classes in a balanced manner. Compared to conventional single classifier and stacking models, this approach is superior in handling unbalanced data. This model is expected to be used as a basis for the development of decision support systems in the health sector, especially for early detection of diabetes. Further research can explore more variations of algorithms and datasets with more complex volumes and distributions. In addition, hybrid approaches such as a combination of stacking and sampling techniques (oversampling/undersampling) are also worth considering for further exploration.

Acknowledgments
I sincerely thank all those who have supported and assisted me throughout the course of my research.

Author Contributions
Conceptualization, methodology, software, formal analysis, investigation, resources, data curation, writing—original draft preparation, funding acquisition R.A.M. and P.; writing—review and editing, validation, visualization, supervision, project administration, F.A.Z. and A.R. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare no conflict of interest.

References

Abdan, M., & Seno, S. A. H. (2022). Machine Learning Methods for Intrusive Detection of Wormhole Attack in Mobile Ad Hoc Network (MANET). *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/2375702

Ahmed, H., Younis, E. M. G., & Ali, A. A. (2020). Predicting Diabetes using Distributed Machine Learning based on Apache Spark. *Proceedings of 2020 International Conference on Innovative Trends in Communication and Computer Engineering, ITCE 2020*, 44–49. https://doi.org/10.1109/ITCE48509.2020.9047795

Ellyzabeth Sukmawati, Iwan Adhicandra, & Nur Sucahyo. (2022). Information System Design of Online-Based Technology News Forum. *International Journal Of Artificial Intelligence Research*, *1.2*. https://doi.org/https://doi.org/10.29099/ijair.v6i1.2.593

Hemachandran, K., Verma, P., Pareek, P., Arora, N., Rajesh Kumar, K. V., Ahanger, T. A., Pise, A. A., & Ratna, R. (2022). Artificial Intelligence: A Universal Virtual Tool to Augment Tutoring in Higher Education. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/1410448

Hsiung, S. Y., Deng, S. X., Li, J., Huang, S. Y., Liaw, C. K., Huang, S. Y., Wang, C. C., & Hsieh, Y. S. Y. (2023). Machine learning-based monosaccharide profiling for tissue-specific classification of Wolfiporia extensa samples. *Carbohydrate Polymers*, *322*. https://doi.org/10.1016/j.carbpol.2023.121338

International Diabetes Federation. (2021). IDF Diabetes Atlas 2021 _ IDF Diabetes Atlas. *IDF Official Website*, 1–4.

Khan, A., Khan, A., Khan, M. M., Farid, K., Alam, M. M., & Su'ud, M. B. M. (2022). Cardiovascular and Diabetes Diseases Classification Using Ensemble Stacking Classifiers with SVM as a Meta Classifier. *Diagnostics*, *12*(11). https://doi.org/10.3390/diagnostics12112595

Lopatka, V., Meniailov, I., & Bazilevych, K. (2021). Classification and Prediction of Diabetes Disease

Using Modified k-neighbors Method. *2021 IEEE 12th International Conference on Electronics and Information Technologies, ELIT 2021 - Proceedings*, 46–50. https://doi.org/10.1109/ELIT53502.2021.9501151

Machaka, R. (2021). Machine learning-based prediction of phases in high-entropy alloys. *Computational Materials Science*, *188*. https://doi.org/10.1016/j.commatsci.2020.110244

Mafarja, M., Thaher, T., Al-Betar, M. A., Too, J., Awadallah, M. A., Abu Doush, I., & Turabieh, H. (2023). Classification framework for faulty-software using enhanced exploratory whale optimizer-based feature selection scheme and random forest ensemble learning. *Applied Intelligence*, *53*(15). https://doi.org/10.1007/s10489-022-04427-x

Maretalinia, Rusmitasari, H., Supriatin, Amaliah, L., Sukmawati, E., & Suwarni, L. (2023). Factors influencing the utilization of the Modern Family Planning (MFP) method under the National Health Insurance in Indonesia: An analysis of the 2017 IDHS. *Public Health of Indonesia*, *9*(2). https://doi.org/10.36685/phi.v9i2.694

Mohd Amram, N. A. L., Keikhosrokiani, P., & Asl, M. P. (2023). Artificial intelligence approach for detection and classification of depression among refugees in selected diasporic novels. *Social Sciences and Humanities Open*, *8*(1). https://doi.org/10.1016/j.ssaho.2023.100558

Nipa, N., Riyad, M. H., Satu, S., Walliullah, Howlader, K. C., & Moni, M. A. (2024). Clinically adaptable machine learning model to identify early appreciable features of diabetes. *Intelligent Medicine*, *4*(1). https://doi.org/10.1016/j.imed.2023.01.003

Nurmasani, A., & Pristyanto, Y. (2021). Algoritme Stacking Untuk Klasifikasi Penyakit Jantung Pada Dataset Imbalanced Class. *Pseudocode*, *8*(1), 21–26. https://doi.org/10.33369/pseudocode.8.1.21-26

Orchi, H., Sadik, M., Khaldoun, M., & Sabir, E. (2023). Automation of Crop Disease Detection through Conventional Machine Learning and Deep Transfer Learning Approaches. *Agriculture (Switzerland)*, *13*(2). https://doi.org/10.3390/agriculture13020352

Park, K., & Song, Y. (2022). Multimodal Diabetes Empowerment for Older Adults with Diabetes. *International Journal of Environmental Research and Public Health*, *19*(18). https://doi.org/10.3390/ijerph191811299

Perkeni. (2020). Tetap Produktif, Cegah Dan Atasi Diabetes Mellitus. *Pusat Data Dan Informasi Kementrian Kesehatan RI*.

Rousyati, R., Rais, A. N., Rahmawati, E., & Amir, R. F. (2021). Prediksi Pima Indians Diabetes Database Dengan Ensemble Adaboost Dan Bagging. *EVOLUSI : Jurnal Sains Dan Manajemen*, *9*(2). https://doi.org/10.31294/evolusi.v9i2.11159

Singh, G., Pal, Y., & Dahiya, A. K. (2023). Classification of Power Quality Disturbances using Linear Discriminant Analysis. *Applied Soft Computing*, *138*. https://doi.org/10.1016/j.asoc.2023.110181

Siti Khotimatul Wildah, Agustiani, S., Ali Mustopa, Nanik Wuryani, Hendri Mahmud Nawawi, & Rizky Ade Safitri. (2021). Pengenalan Wajah Menggunakan Pembelajaran Mesin Berdasarkan Ekstraksi Fitur Pada Gambar Wajah Berkualitas Rendah. *INFOTECH : Jurnal Informatika & Teknologi*, *2*(2). https://doi.org/10.37373/infotech.v2i2.189

Sukmawati, E., Didik, N., & Imanah, N. (2024). *Motivation for Pregnant Women to Get Covid-19 Vaccination ARTICLE INFO*. https://doi.org/10.33860/jbc.v6i4.3978

Tripathi, G., & Kumar, R. (2020). Early Prediction of Diabetes Mellitus Using Machine Learning. *ICRITO 2020 - IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)*, 1009–1014. https://doi.org/10.1109/ICRITO48877.2020.9197832

Żabiński, G., Gramacki, J., Gramacki, A., Miśta-Jakubowska, E., Birch, T., & Disser, A. (2020). Multi-classifier majority voting analyses in provenance studies on iron artefacts. *Journal of Archaeological Science*, *113*. https://doi.org/10.1016/j.jas.2019.105055