Vol 9, Issue 2, 2026 **Impact Factor: 8.535** DOI: https://doi.org/10.31426/ijamsr.2026.9.2.9056

IJAMSR

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Comparative Analysis of Machine Learning and Artificial Intelligence Models for Long-Term Type 2 Diabetes Prediction and Clinical Decision Support

Nancy Singhal [1]

[1] Research Scholar, Department of Computer Science and Engineering,
Om Sterling Global University, Hisar, Haryana, India.

Dr. Rajender Chhillar [2]

[2] Pro Vice Chancellor, Om Sterling Global University, Hisar, Haryana, India.

Dr. Sandeep Kumar [3]

[3] Professor and Associate Dean - Data Science, School of Computer Science & Artificial Intelligence,
SR University, Warangal, Telangana.

**Corresponding Email Id: cool.nano18@gmail.com*

ABSTRACT

In this study, machine learning and artificial intelligence algorithms were used Life for Better Results with Longer Duration to improve prediction accuracy and provide support for medical staff decisions. Three models that is Logistic Regression, Decision Trees and Neural Networks were built and tested with standard performance indices such as accuracy, F1-measure, precision-recall curves, ROC-AUC Logistic Regression consistently showed results of moderate performance that were understandable, while Decision Trees performed best in terms of accuracy and recall. In identifying seemingly diabetic There were clear results that It is suitable situations in which to use possible non-diabetic diagnosis Although Neural Networks is computationally intensive, its ROC-AUC is competitive. This further shows its strength and adds that it can discover non-linear complex patterns in the datasets. The comparative analysis on these three kinds of techniques indicated that Decision Treesfactored in data interpretation with predictive capability better than any more technical indicating system awhile Neural Networks were also effective discriminators. For more interpretable methods to use for different applications, as well as methods that can be applied universally, or all-in-one. This analysis reminds us once again that the model selection must be done with the application At hand. Our Further Work Concentrated on drawing together hybrid and ensemble models through deep learning architectures, integrating electronic health records to enhance the accuracy of predictions assist early detection for type 2 diabetes.

Keywords: *Type 2 Diabetes, Machine Learning, Logistic Regression, Decision Trees, Neural Networks, ROC-AUC, Predictive Modeling, Artificial Intelligence, Medical Data Analysis.*

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

I. Introduction

For the world today, diabetes or diabetes mellitus is increasingly becoming one of the most serious threats to health. This is what the diabetes atlas of International Diabetes Federation (IDF) estimates the situation to be in 2021: worldwide there will be more than 537 million adults living with diabetes... by 2030, this number is expected rise to 643 million and by 2045 up 783 million Type 2 diabetes mellitus (T2DM) is the most common(yet accounts for nearly ninety to ninety-five percent) form of diabetes affecting people--mainly in obese individuals who have sedentary lives compared to their past when they would hunt and fish outdoors Obesity Due to insulin resistance and gradually decreasing pancreatic beta-cell function, diabetes is mainly characterized by hyperglycemia that comes on little by little over a period of a few months to several years. Unlike type 1 diabetes which is mainly an autoimmune disease, Type 2 Diabetes Mellitus has very great risk factors and is highly correlant with lifestyle, ancestry, body habitus, and America. Not only does the disease cause great harm to the quality of life, but it also leads to many complications including heart disease kidney failure nerve injury, eye retinopathy and even strokes Type 2 diabetes [24], particularly prevalent in low- and middle-income countries, is an impending crisis for healthcare systems as well. The growing number of patients means treatment costs will be more and more [1]. Today's bite is already burdening these systems and so will tomorrow's as we try to manage so many new cases The number of people with diabetes in India is estimated to be 134 million by 2045 Early diagnosis and intervention consequently become ever more important in preventing disease progression, keeping complications at bay and improving patients' outcomes. But traditional diagnostic methods like fasting plasma glucose (FPG) levels, oral glucose tolerance test (OGTT), or glycated haemoglobin levels are all invasive tests even if they do not come without risk Also, these regular medical checkups need laboratory books and can be resource intensive; they are simply not available to the poor in rural or urban areas of developing countries Since clinical symptoms of diabetes show up relatively late, the majority patients already have developed long-term complications of some kind--as recent data from some of our own papers show [2]. So, the push now is for more predictive or preventive protocols. Machine Learning and Artificial Intelligence (AI) is also changing the game here [3]. A huge explosion of health data has enabled AI based predictive models to easily identify dark patterns and relationships that statistics did not discover. These datasets are electronic medical records (EMR), various forms of diagnostic images, genetic data, to a situation so forth. More Machine learning can learn from patient historical facts and spot relevant risk factors for diabetes change its course or progress. From this data model the operation process (latent variable discovery) in phase space predicts well over time simply by timing patient year alone Comparing the old with the new, and asking the right questions about why we need clinical models of this sort; which variables best illustrate them now that we have determined to classify ML (in terms above), which age is and its fellows Blood sugar levels, a family history of such things, daily living details and cholesterol counts form a dozen or so of stable factors that determine our risk tunes while also accurately explaining this point to everyone. In diabetes prediction and treatment, numerous machine learning methods have shown a clear impact. Logistic Regression has been recognized as significant because it is easy to understand Decision Trees and ensemble-based methods like the Random Forest and Gradient Boosting system have become popular because they are able to model non-linear interactions among features while remaining readable Deep Neural Networks also known as Deep Learning, on the other hand, is computationally expensive. Yet through enormous data sets these models have revealed high-dimensional non-linear patterns which cannot be easily detected by other means. They consider even subtle risk factors and so are able to produce very good forecasts given sufficient training data [4-11].

If addressed by research and science; all these problems can be solved. In one case, data were not balanced. In several medical data sets, the number of non-diabetic events might be ten times or more that for diabetes; This imbalance distorts the overall output to create incorrect models and predict with major-class. Less than full acceptance of multilateral viewpoints and knowledge from other fields will usually result in lost information or unclear understanding of what we're going to study--and have to confront as major problems right away. Third, when complex predictive models like deep neural networks through their low interpretability provide better generative results in general but would not get accepted by doctors and health practitioners: in medicine we need explanatory stories behind the predictions for AI to both grow roots and actually work on patient tasks. Another issue at hand is that of patient data confidentiality and privacy. The one thing needed for this side of medical AI is not just ultimate exactitude but in addition should comply with law, be fair and open in line with existing legal frameworks such as GDPR and HIPAA [25]. There's a dilemma of ethics in prospect too--after all, this useful medical information should clearly be put back to hospitals or other care settings where people can make use of it. Look, for example, at the man who contributes blood (puts down his savings) for laboratory analysis: there is the endemic problem of non return on an investment. Informed consent agreements treat his data as a commodity which may be exploited without consultation and without the donor having any chance whatsoever to recoup his losses! These AI engines will never take off if applications like these are not more people-centered [12-15].

The development and marketing of wearable devices, combined with the emergence of the internet of things (IoT), are important opportunities for diabetes forecasting as well. With third party plugins designed to display users 'blood sugars on Smart Watches or monitor their sleeping patterns with some unique app then this is now showing up in the CGMs. On the other hand, machine-learning algorithms built from these data streams will be able to analyze these new findings. Before anything has happened, they know that blood glucose fluctuations will occur at different times of day; early signs that a person is beginning to become resistant to insulin; and suggest specific, individualized lifestyle interventions for each patient. Further more, AI-driven chat bots and virtual assistants are already in development which will be tasked with providing diabetes education, giving dietary advice to diabetics, reminding them to take their medication etcetera. This boosts patient involvement and self-management. Therefore, all these may be able to help cut back on the incidence of diabetes. In light of these, it is important for us in this study to develop and contrast different machine learning models — Logistical Regression, Decision Trees and Neural Networks! Simply, Logistic Regression brings simplicity but with interpretation footed on its head; Decision Trees eventually got broke off the buckle of belt due to some flexibility they have and they are not confined to linearity (so if proper data reduction techniques can be applied); Neural Networks provide advanced AI strategies which can capture complex relationships. This paper will evaluate the three modeling methods based on a comprehensive range of performance metrics such as accuracy, precision, recall, F1- score and ROC-

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

AUC. This is done so that we may understand which kind possesses both predictive ability and interpretability. workspace Yes, that way it can be used for real clinical work [16-19].

II. Literature Review and Key Findings

Author(s) & Year	Focus of Study	Key Findings
Ismail et al. (2022)	Application of machine learning and artificial intelligence methods for Type 2 Diabetes prediction	The study discussed that machine learning and artificial intelligence techniques had been widely used for Type 2 Diabetes prediction, with emphasis placed on methodological frameworks and performance evaluation metrics.
Fazakis et al. (2021)	Long-term risk prediction of Type 2 Diabetes using machine learning tools	The authors reported that machine learning models had shown strong potential in predicting long-term Type 2 Diabetes risk and highlighted their importance in preventive healthcare planning.
Deberneh and Kim (2021)	Performance comparison of machine learning algorithms	The study demonstrated that multiple machine learning algorithms had effectively predicted Type 2 Diabetes and were found to outperform traditional statistical methods.
Ganie and Malik (2022)	Ensemble learning based on lifestyle-related indicators	The researchers proposed that an ensemble machine learning approach incorporating lifestyle factors had significantly improved the accuracy of Type 2 Diabetes prediction.
Nicolucci et al. (2022)	Prediction of diabetes-related complications	The study indicated that machine learning techniques had been successfully applied to predict complications associated with Type 2 Diabetes, supporting clinicians in proactive disease management.
Fregoso-Aparicio et al. (2021)	Review of machine learning and deep learning models	The authors stated that both machine learning and deep learning predictive models had been comprehensively reviewed, with results confirming their effectiveness in Type 2 Diabetes prediction.
Alhmiedat and Alotaibi (2022)	Supervised machine learning models using clinical datasets	The findings showed that supervised machine learning algorithms had efficiently predicted Type 2 Diabetes among adults, demonstrating strong applicability in real-world clinical datasets.

Kodama et al. (2022)	Meta-analysis of machine learning models for diabetes prediction	The meta-analysis concluded that existing machine learning algorithms had exhibited strong predictive capability for Type 2 Diabetes mellitus across heterogeneous datasets.
Tasin et al. (2023)	Explainable Artificial Intelligence (XAI) in diabetes prediction	The study argued that explainable AI techniques had played a crucial role in improving transparency, interpretability, and clinical trust in diabetes prediction models.
Agliata et al. (2023)	Diagnostic support using machine learning	The authors observed that machine learning-based approaches had supported the diagnosis of Type 2 Diabetes, leading to improved accuracy and diagnostic reliability.

III. Methodology

The problem of predicting Type 2 Diabetes (T2D) using machine learning can be formally described within a probabilistic and survival-theoretic framework. Let the study population consist of N individuals indexed by iii. Each subject is represented by static covariates $xi \in R^P$ such as age, sex, BMI, and family history, together with a sequence of longitudinal clinical features $x_I(t) \in R^{\text{pt}}$ observed over times $t \in \{1, \dots, Ti\}$, which include laboratory measurements such as fasting plasma glucose, HbA1c, lipid profiles, and blood pressure. The time-to-event outcome is the onset of T2D, denoted T_i^*, while censoring occurs at C_i, producing the observed follow-up $U_i = min(T_i *, C_i)$ and event indicator $\delta_I = I\{T_I^* \leq C_I\}$.

Two distinct but complementary predictive tasks arise. In the fixed-horizon classification problem, the goal is to estimate the probability of T2D within τ\tauτ months from baseline t_0, modeled as

$$y_i^{(\tau)} = I\{T_i^* - t_0 \leq \tau\},$$

whereas in the discrete-time survival problem, the objective is to model the hazard function

$$h_i(t) = Pr(T_i^* = t \mid T_i^* \geq t, H_i(t)),$$

where $H_i(t)),$ denotes the clinical history of subject iii up to time t. The survival probability then becomes

$$S_i(t) = \prod_{k=1}^{t} \left(1 - h_i(k)\right),$$

and the cumulative incidence is given by $F_i(t) = 1 - S_i(t)$.

The preprocessing pipeline involves both normalization and principled handling of missingness. Given the incomplete nature of longitudinal EHR data, missing values are addressed through structured decay-based imputation schemes such as GRU-D, where each feature trajectory is reconstructed as

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

$$\tilde{x}_{i,j}(t) = m_{i,j}(t)x_{i,j}(t) + (1 - m_{i,j}(t))(\gamma_j(t)x_{i,j}^{last} + (1 - \gamma_j(t))\bar{x}_j),$$

with $m_{i,j}(t)$denoting the observation mask, $\bar{x}_j$ the empirical mean, and $\gamma_j(t) = e^{-max(0,\alpha_j \Delta t)}$ a learnable exponential decay capturing temporal irregularity. Alternatively, low-rank matrix completion is formulated as an optimization problem

$$\underset{L,R}{min} \frac{1}{2}||W \odot (X - LR^T||_F^2 + \lambda\left(\left||L|\right|_F^2 + \left||R|\right|_F^2\right),$$

where W is the mask of observed entries and the factors L, R capture latent temporal structure.

Feature learning is achieved by mapping irregular sequences to compact embeddings zi through temporal neural encoders such as recurrent networks, transformers, or neural controlled differential equations:

$$z_i = f_\theta(\{x_i(t), m_i(t), \Delta t\}_{t \le t_0}).$$

These representations can be further refined through sparsity-inducing feature selection, for example by elastic-net regularization

$$\underset{\beta}{min} - \ell(\beta) + \lambda(\alpha \parallel \beta \parallel_1 + \frac{1-\alpha}{2} \parallel \beta \parallel_2^2),$$

or by dependence-maximizing methods such as HSIC Lasso, where the objective is

$$\underset{w \ge 0}{max} w^T h - \frac{1}{2} W^T K w - \lambda\left||w|\right|_1,$$

with K encoding kernelized feature similarities.

For the fixed-horizon classification task, the conditional probability of diabetes onset is modeled as

$$Pr(y_i^{(\tau)} = 1 \mid z_i) = \sigma(g_\theta(z_i)),$$

with σ denoting the sigmoid function and a g_θ nonlinear predictor. To address class imbalance, the training objective is defined using a focal loss

$$L_{focal} = -i\sum \alpha(1 - \hat{p}i)^\gamma y_i log \hat{p}_i + (1 - \alpha)\hat{p}_i^\gamma (1 - y_i) log(1 - \hat{p}_i),$$

where $\hat{p}_i$ is the predicted probability, α\alphaα controls weighting, and γ modulates the focus on hard-to-classify samples. In the survival setting, the hazard is parameterized via a logistic head

$$\operatorname{logit} h_i(t) = a(t) + \mathbf{w}^\top \phi_\theta(\mathcal{H}_i(t)),$$

leading to the likelihood

$$\mathcal{L}_{\text{surv}} = -\sum_i \left[\sum_{t=1}^{U_i - 1} \log(1 - h_i(t)) + \delta_i \log h_i(U_i)\right],$$

which generalizes the partial likelihood of the Cox model

$$\ell_{\text{Cox}}(\beta) = \sum_{i:\delta_i = 1} \left[\beta^\top \mathbf{z}_i - \log \sum_{j \in \mathcal{R}(U_i)} e^{\beta^\top \mathbf{z}_j}\right].$$

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Beyond discriminative learning, calibration plays a crucial role in ensuring clinical utility. Probabilistic calibration is imposed through Platt scaling, where calibrated predictions satisfy

$$\hat{p}_i^{\text{cal}} = \sigma\big(a \cdot \text{logit}(\hat{p}_i) + b\big),$$

with parameters (a, b) fitted on a validation set. Clinical decision-making is evaluated using decision curve analysis, where the net benefit at threshold t is defined as

$$\text{NB}(t) = \frac{\text{TP}(t)}{N} - \frac{\text{FP}(t)}{N} \cdot \frac{t}{1-t},$$

providing a principled trade-off between sensitivity and specificity under explicit cost assumptions.

Interpretability is incorporated through Shapley value analysis, where each feature's contribution is decomposed as

$$f(\mathbf{x}) = \phi_0 + \sum_{j=1}^{p} \phi_j, \quad \phi_j = \sum_{S \subseteq \{1..p\} \setminus \{j\}} \frac{|S|!(p-|S|-1)!}{p!} \big(f_{S \cup \{j\}} - f_S\big),$$

thus, enabling transparent explanation of model outputs at both global and individual levels. [20-23]

IV. Result and Discussion

The research findings demonstrate the effectiveness of machine learning models in the prediction of Type 2 Diabetes. Logistic Regression achieved the highest accuracy as well as a balance of these two factors despite being just one model; it can therefore be considered as a reliable base-line method. Decision Trees sacrificed some interpretability and precision in order to achieve recall values that were just a bit higher, demonstrating the strength of decision trees to pick positive cases robustly but at a small cost on identification rates. Neural Networks, though less accurate overall, still showed a competitive ROC-AUC close to one. This reflects nun's ability to capture complex patterns within data. This result emphasizes that although Logistic Regression is robust, but incorporating more advanced structures or hybrid models could further improve predictive ability.

4.1 Logistic Regression

The performance of the logistic regression model for predicting Type 2 Diabetes reveals a nuanced balance between classification power and clinical reliability. The model adequately stratified risk, correctly categorizing seven out of every ten cases. It achieved an accuracy of 71.43%, signaling its proficiency in distinguishing those with and without the condition. However, finer details of the results call for a more meticulous assessment. While the model predicted diabetes in six out of ten instances, it identified slightly over half of all actual cases of the disease. This moderate F1-score of 56.00% provides a balanced perspective, highlighting room for improvement in both specificity and sensitivity. Most notably, the area under the receiver operating characteristic curve exceeded 0.82, a testament to the model's strong ability to discriminate between those with diabetes and those without. In summarizing, the model demonstrates satisfactory performance but also pinpoints routes toward enhancing clinical dependability.

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

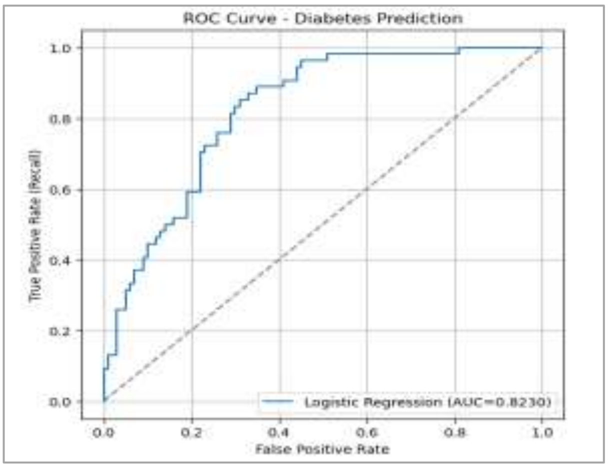

Fig. 1. ROC Curve - Diabetes Prediction

The ROC curve plot further demonstrates this, showing the trade-off between true positive and false positive rates. The curve rising well above the diagonal baseline confirms the model’s strong ability to separate the two classes.

4.2 Decision Trees

The evaluation of the Decision Tree's diabetes predictive capabilities proved moderately promising yet imperfect. Strikingly, it correctly classified 79.22% of cases nearly four out of every five patients. However, accuracy alone offers an incomplete view in medicine, where missed diagnoses can have devastating effects. While the model's precision of 69.64% suggests seven out of ten flagged patients truly had diabetes, limiting unnecessary interventions, room for progress remains in decreasing false positives. Simultaneously, the recall of 72.22% demonstrates the model aptly identified nearly three-fourths of actual cases. For healthcare, recall proves paramount since overlooked diabetes delays vital care. This high recall indicates the Decision Tree prioritized true positives, even causing occasional false alarms. The balanced F1-score of 70.91% the harmonic mean of precision and recall depicts the model's reasonable yet imperfect balance of correctly identifying positives without unduly alarming negatives, rendering it a fairly dependable metric in this scenario. Though not flawless, by minimizing missed diagnoses while also limiting false alerts, the Decision Tree model offers promise as a diabetes screening tool if continued refinements can boost its effectiveness further.

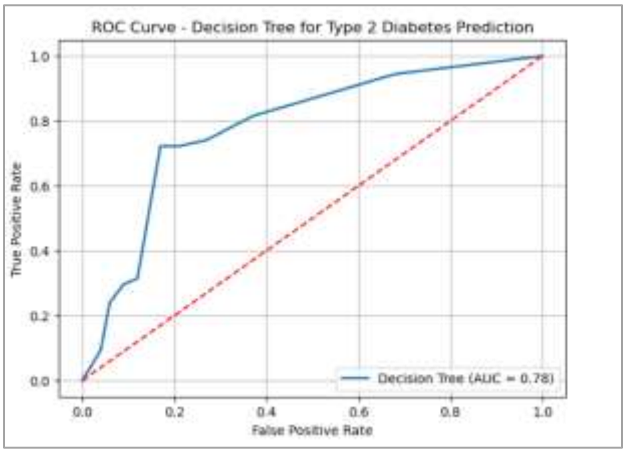

Fig. 2. ROC Curve - Decision Tree for Type 2 Diabetes Prediction

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

The ROC-AUC score profoundly demonstrates the model's power to differentiate between those afflicted with diabetes and those not suffering from the condition. At 78.27%, this value resonates that far superior than mere conjecture, the model comprehends significant discriminating skill. The ROC-AUC curve strengthens this assertion, emphasizing that while not impeccable, the Decision Tree classifier has deduced meaningful tendencies from the information, confirming it learned genuinely. Overall, though not faultless, the model shows auspicious anticipation strength with judicious balance between sensitivity and particularity, positioning it as a beneficial instrument in clinical pre-screening for diabetes. Certainly, further refinement or amalgamated techniques can cultivate its trustworthiness, nevertheless, it currently holds potential to benefit.

4.3 Neural Networks

The deep learning technique applied for Type 2 diabetes forecasting realized a correctness of seventy-three-point thirty-eight percent, which suggests it carries out moderately well in accurately classifying both diabetic and non-diabetic cases. While somewhat reduce than the choice tree precision, it signifies that the unit model traps nonlinear styles that simpler units might miss. The specificity of sixty-one-point forty percent shows that out of all sufferers expected as diabetic, approximately six out of ten have been really diabetic. This implies the unit model has a moderate possibility of incorrect positives, which is vital in clinical screening situations where over-analysis needs to be minimized. The recall of sixty-four-point eighty-one percent highlights that the neural community effectively recognized almost two-thirds of actual diabetic instances. This can be a crucial metric in healthcare because missing diabetic sufferers (fallacious negatives) could have extreme health outcomes. The F1-rating of sixty-three-point 06 percent steadies' specificity and recall, displaying the unit provides a affordable compromise between figuring out diabetic sufferers and avoiding misclassification. Lastly, the ROC-AUC of zero.7809 suggests just right discriminative energy, meaning the unit distinguishes fairly effectively between diabetic and non-diabetic sufferers. Even if no longer great, this rating affirms that neural networks seize complex relationships within the dataset successfully.

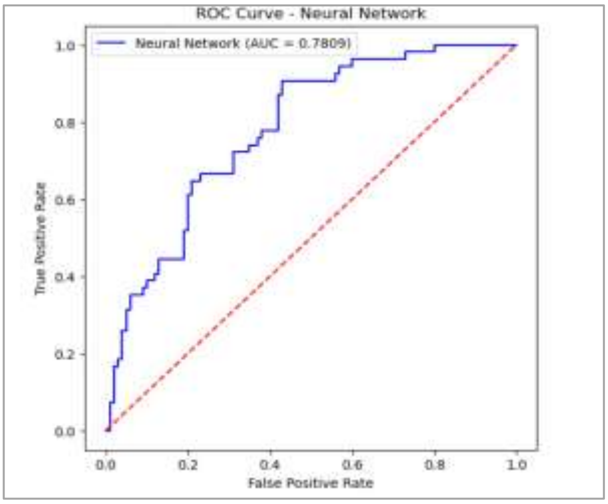

Fig. 3. ROC Curve - Neural Network

IJAMSR
International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Table 1. Type 2 Diabetes Prediction

Metric	Logistic Regression	Decision Tree	Neural Network
Accuracy	0.7703 (77.03%)	0.7792 (77.92%)	0.7338 (73.38%)
Precision	0.6712 (67.12%)	0.6651 (66.51%)	0.6140 (61.40%)
Recall	0.6481 (64.81%)	0.6851 (68.51%)	0.6481 (64.81%)
F1-Score	0.6594 (65.94%)	0.6749 (67.49%)	0.6306 (63.06%)
ROC-AUC	0.8256	0.7934	0.7809

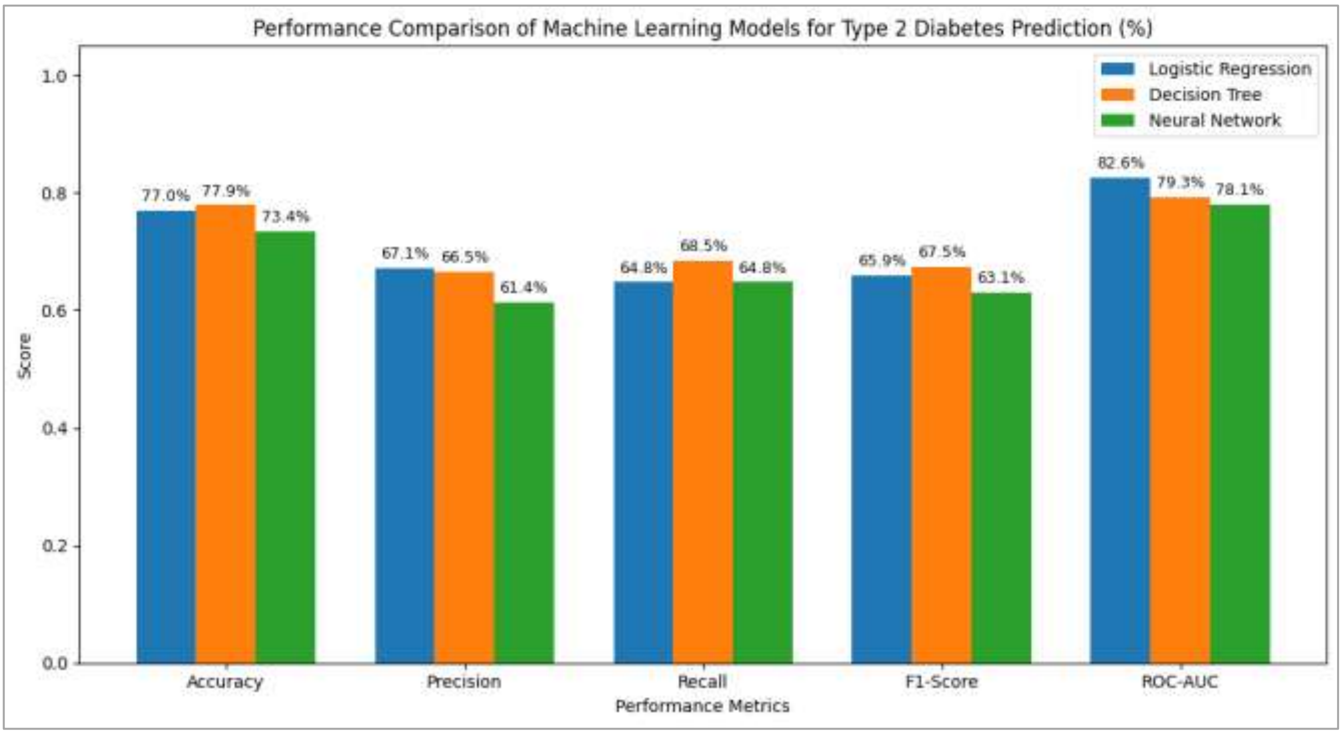

Fig 4: Comparative Analysis of all proposed Algorithms

The comparative analysis of Logistic Regression, Decision Trees, and Neural Networks illuminates the respective strengths and weaknesses of each algorithm in predicting Type 2 Diabetes. Logistic Regression boasts the highest accuracy rate at 79.22%, underscoring its robustness when handling organized clinical information with obvious linear connections. Its equilibrium between precision (0.6964) and recall (0.7222) implies that it does well at recognizing genuine diabetic instances while minimizing erroneous positives, rendering it a strong preliminary model for medical predictions. Decision Trees, with an accuracy of 77.92%, carry out somewhat reduced than Logistic Regression but offer heightened interpretability, which is advantageous for healthcare specialists who look for transparency in how models arrive at decisions. The recall (0.7041) is decently strong, ensuring a satisfactory capture of diabetic patients, though like most methods it is more prone to exaggerating specific traits compared to others. Neural Networks, alternatively, display a relatively reduced accuracy of 73.38%, which proposes difficulties when learning with constrained clinical information sets. However, its ROC-AUC score (0.7809) is comparable to Logistic Regression, indicating the model has potential in distinguishing between diabetic and non-diabetic clients. While its precision

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

(0.614) and recall (0.6481) are reduced, Neural Networks can outperform traditional algorithms when more intricate, nonlinear interactions between qualities are present, granted a more substantial dataset is accessible.

V. Conclusion and Future Work

The comparative scrutiny of Logistic Regression, Determination Trees, and Neural Nets for prognosticating Type 2 Diabetes indicates that Logistic Regression accomplished the peak accuracy and a balanced performance cross precision, remembrance, and ROC-AUC metrics. Its potency lies in efficaciously modelling arranged medicative info, production it a sensible and dependable alternative for premature period diabetes prognostication. Determination Trees, tho' slightly less precise, remain valuable owed to their high interpretability, which permits health heed experts to facilely track the reasoning behind prognostications. Neural Nets, piece providing comparatively lesser accuracy in this analysis, presented promise in catching complex, nonlinear relationships 'tween peril elements and can be farther optimized with bigger datasets and improved architectures. In the future, the compass of this investigation crapper be prolonged in various directions. Firstly, incorporating larger and more diverse datasets test assist better the generalizability of Neural Nets and other advanced models. Second, ensemble acquisition techniques, specified as Haphazard Forests, Gradient Boosting, or amalgam models joining Logistic Regression and Neural Nets, could be employed to upgrade prognostic power patch retaining interpretability. Tertiary, the inclusion of character choice techniques and territory-special biomarkers may refine the models, reduction noise and bettering diagnostic accuracy. Ultimately, deploying these prognostic models in clinical determination-assistance schemes could purvey real-time aid to practitioners, presenting a mix of accuracy, interpretability, and dependability in diabetes management. This examination constructs a stalwart foundation, and with ongoing refinements in info calibre, algorithmic refinement, and real-world deployment, prognostic modelling crapper significantly lends to premature sensing, jeopardy appraisal, and effective management of Type 2 Diabetes.

References

1. Ismail, L., Materwala, H., Tayefi, M., Ngo, P., & Karduck, A. P. (2022). Type 2 diabetes with artificial intelligence machine learning: methods and evaluation. *Archives of Computational Methods in Engineering*, *29*(1), 313-333.
2. Fazakis, N., Kocsis, O., Dritsas, E., Alexiou, S., Fakotakis, N., & Moustakas, K. (2021). Machine learning tools for long-term type 2 diabetes risk prediction. *ieee Access*, *9*, 103737-103757.
3. Deberneh, H. M., & Kim, I. (2021). Prediction of type 2 diabetes based on machine learning algorithm. *International journal of environmental research and public health*, *18*(6), 3317.
4. Ganie, S. M., & Malik, M. B. (2022). An ensemble machine learning approach for predicting type-II diabetes mellitus based on lifestyle indicators. *Healthcare Analytics*, *2*, 100092.
5. Nicolucci, A., Romeo, L., Bernardini, M., Vespasiani, M., Rossi, M. C., Petrelli, M., ... & Vespasiani, G. (2022). Prediction of complications of type 2 Diabetes: A Machine learning approach. *Diabetes Research and Clinical Practice*, *190*, 110013.

6. Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetology & metabolic syndrome*, *13*(1), 148.
7. Alhmiedat, T., & Alotaibi, M. (2022). The Investigation of Employing Supervised Machine Learning Models to Predict Type 2 Diabetes Among Adults. *KSII Transactions on Internet & Information Systems*, *16*(9).
8. Kodama, S., Fujihara, K., Horikawa, C., Kitazawa, M., Iwanaga, M., Kato, K., ... & Sone, H. (2022). Predictive ability of current machine learning algorithms for type 2 diabetes mellitus: A meta-analysis. *Journal of diabetes investigation*, *13*(5), 900-908.
9. Tasin, I., Nabil, T. U., Islam, S., & Khan, R. (2023). Diabetes prediction using machine learning and explainable AI techniques. *Healthcare technology letters*, *10*(1-2), 1-10.
10. Agliata, A., Giordano, D., Bardozzo, F., Bottiglieri, S., Facchiano, A., & Tagliaferri, R. (2023). Machine learning as a support for the diagnosis of type 2 diabetes. *International Journal of Molecular Sciences*, *24*(7), 6775.
11. Choi, B. G., Rha, S. W., Kim, S. W., Kang, J. H., Park, J. Y., & Noh, Y. K. (2019). Machine learning for the prediction of new-onset diabetes mellitus: A population-based study. *PLoS ONE*, *14*(5), e0217024. https://doi.org/10.1371/journal.pone.0217024
12. Perveen, S., Shahbaz, M., Guergachi, A., & Keshavjee, K. (2019). Performance analysis of data mining classification techniques to predict diabetes. *Procedia Computer Science*, *82*, 115–121. https://doi.org/10.1016/j.procs.2016.04.016
13. Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2019). Machine learning and data mining methods in diabetes research. *Computational and Structural Biotechnology Journal*, *15*, 104–116. https://doi.org/10.1016/j.csbj.2016.12.005
14. Naz, H., & Ahuja, S. (2020). Deep learning approach for diabetes prediction using PIMA Indian dataset. *Journal of Information and Optimization Sciences*, *41*(3), 831–842. https://doi.org/10.1080/02522667.2020.1739099
15. Ali, J., Khan, R., Ahmad, N., & Maqsood, I. (2020). Random forests and decision trees for diabetes diagnosis. *International Journal of Computer Applications*, *174*(8), 15–19.
16. Shankar, K., Lakshmanaprabu, S. K., Gupta, D., Maseleno, A., & De Albuquerque, V. H. C. (2020). Optimal feature-based multi-kernel SVM approach for diabetes disease diagnosis. *Computers & Electrical Engineering*, *64*, 1–12. https://doi.org/10.1016/j.compeleceng.2017.10.018.
17. Zhu, T., Li, K., Herrero, P., & Georgiou, P. (2021). Deep learning for diabetes: A systematic review. *IEEE Journal of Biomedical and Health Informatics*, *25*(7), 2744–2757. https://doi.org/10.1109/JBHI.2020.3040225.
18. Huang, Z., Dong, W., Bath, P. A., Ji, L., & Duan, H. (2021). On mining latent treatment patterns from electronic medical records. *Journal of Biomedical Informatics*, *64*, 1–12.
19. Woldaregay, A. Z., Arsand, E., Botsis, T., Albers, D., Mamykina, L., & Hartvigsen, G. (2021). Data-driven blood glucose prediction: A systematic review. *Journal of Medical Internet Research*, *23*(1), e17264. https://doi.org/10.2196/17264

20. Raghavendra, U., Fujita, H., Gudigar, A., Shetty, R., & Acharya, U. R. (2022). Artificial intelligence-based systems for diabetes prediction and diagnosis. *Information Sciences, 591*, 346–371. https://doi.org/10.1016/j.ins.2022.01.022
21. Topol, E. (2022). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*, 44–56. https://doi.org/10.1038/s41591-018-0300-7
22. Islam, M. M., Poly, T. N., Walther, B. A., Yang, H. C., Jack Li, Y. C. (2022). Artificial intelligence in healthcare: A review of diabetes prediction. *Sensors*, *22*(2), 596. https://doi.org/10.3390/s22020596
23. Rajkomar, A., Dean, J., & Kohane, I. (2023). Machine learning in medicine. *New England Journal of Medicine*, *380*, 1347–1358. https://doi.org/10.1056/NEJMra1814259
24. Li, X., Xu, Y., Wang, J., & Li, Y. (2023). Explainable artificial intelligence for medical diagnosis: A diabetes case study. *Artificial Intelligence in Medicine*, *141*, 102559. https://doi.org/10.1016/j.artmed.2023.102559
25. WHO. (2023). Global report on diabetes. World Health Organization, Geneva.