International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05682
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

A Machine Learning-Based System for Early Diabetes Prediction using Medical Data

Eshwar Palle, Tushar Bhandurge, Vikrant Satav

eshwarpalle123@gmail.com , tusharbhandurge8@gmail.com , vikrantsatav4@grmail.com,

Department of Computer Science, MIT ADT University, Pune, India

Abstract

Diabetes is one of the most common chronic diseases affecting millions of people worldwide. Early detection and prediction of diabetes can significantly reduce the risk of severe complications such as heart disease, kidney failure, and nerve damage. Traditional diagnostic methods rely on medical tests and physician analysis, which may delay early identification. Machine Learning (ML) techniques provide powerful tools for analysing medical datasets and predicting disease occurrence based on patterns and patient health parameters. This research presents a machine learning-based framework for predicting diabetes using clinical and lifestyle-related data. The study evaluates several supervised learning algorithms including Logistic Regression, Decision Trees, Random Forest, and Support Vector Machines. The proposed system aims to improve prediction accuracy and support healthcare professionals in early diagnosis and preventive treatment planning.

Keywords—Machine Learning, Diabetes Prediction, Healthcare Analytics, Data Mining, Artificial Intelligence, Disease Prediction

1. INTRODUCTION

Diabetes mellitus is a chronic metabolic disorder characterized by high blood glucose levels due to insufficient insulin production or improper insulin utilization by the body. According to the World Health Organization (WHO), diabetes is one of the leading causes of death globally and continues to increase rapidly due to lifestyle changes and unhealthy dietary habits.

The major types of diabetes include:

- Type 1 Diabetes – caused by the immune system attacking insulin-producing cells.
- Type 2 Diabetes – caused by insulin resistance and improper insulin production.
- Gestational Diabetes – occurs during pregnancy.

Early detection of diabetes is essential to prevent severe health complications such as cardiovascular diseases, kidney failure, nerve damage, and vision loss. Traditional diagnostic methods require laboratory testing and medical examination, which can sometimes delay diagnosis.

With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), healthcare systems are increasingly adopting data-driven techniques to analyse medical records and predict diseases. Machine learning algorithms can identify hidden patterns in medical datasets and assist doctors in predicting the likelihood of diabetes in patients.

The goal of this research is to develop a machine learning-based diabetes prediction system that analyses patient health parameters and predicts the probability of diabetes occurrence.

2. PROBLEM STATEMENT

Diabetes is often diagnosed only after the patient develops noticeable symptoms, which may occur in the later stages of the disease. Many individuals remain unaware of their condition during the early stages, which increases the risk of severe complications.

Traditional diagnostic methods have several limitations:

- Late detection due to lack of continuous monitoring
- Dependence on manual medical analysis
- Limited ability to analyse large healthcare datasets
- Difficulty in identifying hidden patterns in patient health records

Therefore, there is a need for an automated and intelligent prediction system that can analyse patient medical data and provide early prediction of diabetes risk.

Machine learning techniques can be used to develop predictive models capable of identifying high-risk individuals and assisting healthcare professionals in making timely medical decisions.

3. LITERATURE REVIEW

Several research studies have explored the use of machine learning techniques for predicting diabetes.

Early research focused on statistical methods such as logistic regression for analysing patient health parameters and predicting diabetes risk. Although these methods are simple and interpretable, they often fail to capture complex relationships between multiple health variables.

Recent studies have demonstrated the effectiveness of machine learning algorithms such as Random Forest, Decision Trees, Support Vector Machines (SVM), and Artificial Neural Networks (ANN) in improving prediction accuracy.

For example, research using the Pima Indians Diabetes Dataset has shown that Random Forest models can achieve high accuracy in diabetes prediction due to their ability to handle complex nonlinear relationships and multiple input features.

Deep learning models such as Artificial Neural Networks have also been applied in medical prediction systems. These models can learn complex patterns in healthcare datasets and provide accurate predictions when large datasets are available.

Despite these advancements, several challenges still exist, including:

- Limited availability of high-quality medical datasets
- Model interpretability issues in deep learning methods
- Data imbalance in medical datasets

This research aims to address these challenges by evaluating multiple machine learning algorithms and identifying the most effective model for diabetes prediction.

Recent studies have also explored the use of artificial intelligence techniques for predictive analysis in different domains. Nalavade, Auti, and Singh proposed a predictive system that uses blockchain and artificial intelligence-based programming to forecast Bitcoin prices by analysing historical data patterns and market trends [27]. Although their research focuses on financial prediction, the methodology demonstrates the effectiveness of machine learning algorithms in analysing large datasets and identifying hidden patterns. Similar predictive techniques can be applied in the healthcare domain for diseases such as diabetes. By analysing patient health parameters such as glucose level, body mass index (BMI), blood pressure, insulin level, and age, machine learning models can detect patterns that indicate the likelihood of diabetes. Such data-driven predictive systems can support healthcare professionals in early diagnosis and preventive treatment planning.

4. PROPOSED SYSTEM ARCHITECTURE

The proposed diabetes prediction system is designed as a data-driven healthcare decision support system consisting of multiple modules that process patient data and generate predictive outcomes.

The architecture consists of the following components:

1. Data Collection
2. Data Preprocessing
3. Feature Selection
4. Machine Learning Model Training
5. Prediction and Decision Support

4.1 Data Collection Layer

The system collects patient health information from medical datasets and healthcare databases.

Typical patient attributes used for diabetes prediction include:

- Age
- Body Mass Index (BMI)
- Blood Glucose Level
- Blood Pressure
- Insulin Level
- Skin Thickness
- Family History of Diabetes
- Number of Pregnancies

The dataset can be represented as:

$X = \{x_1, x_2, x_3, \dots, x_n\}$

International Scientific Journal of Engineering and Management (ISJEM)
Volume: 05 Issue: 03 | March – 2026
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI:10.55041/ISJEM05682

where each x_i represents patient medical attributes.

4.2 Data Preprocessing Module

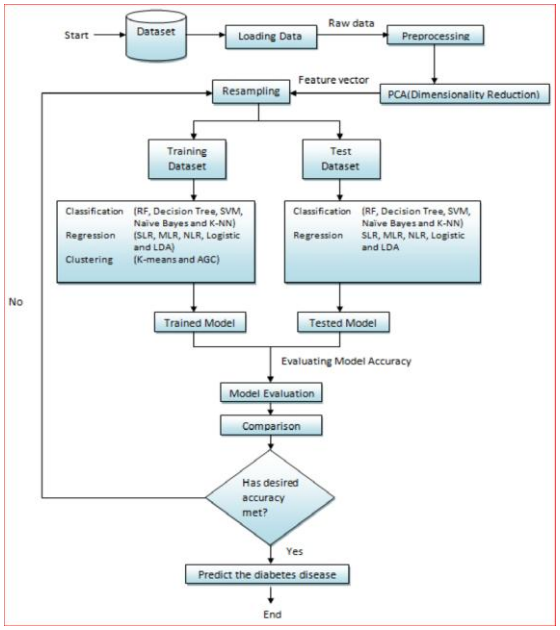
Medical datasets often contain missing values, inconsistent data, and noise. Therefore, preprocessing is required to ensure reliable model performance.

Preprocessing operations include:

- Handling missing values using mean or median imputation
- Removing duplicate records
- Normalizing feature values
- Converting categorical data into numerical format

The processed dataset is represented as:

$X' = f\_preprocess(X)$



4.3 Feature Selection

Feature selection improves prediction accuracy by selecting only the most relevant attributes from the dataset.

Common feature selection techniques include:

- Correlation analysis
- Recursive Feature Elimination (RFE)
- Information Gain

The optimized feature set is represented as:

$X^* = f\_feature(X')$

4.4 Machine Learning Model Training

The machine learning model learns patterns from the training dataset.

The prediction model can be represented as:

$\hat{Y} = f\_model(X^*)$

Where:

$\hat{Y}$ = Predicted diabetes outcome.

Several supervised learning algorithms are used:

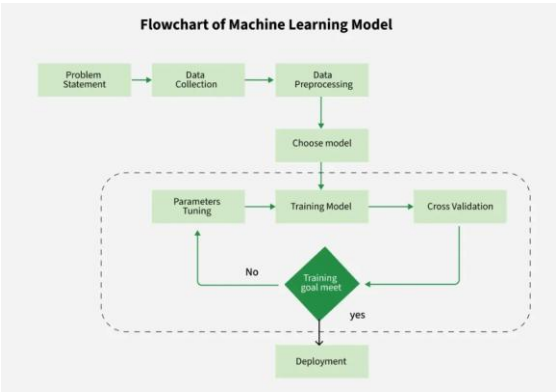
- Logistic Regression
- Decision Tree
- Random Forest
- Support Vector Machine

The dataset is divided into:

- 70% Training data
- 30% Testing data

5. METHODOLOGY

The proposed methodology follows a standard machine learning workflow.



5.1 Dataset Description

The study uses the Pima Indians Diabetes Dataset, which contains medical data of female patients.

Dataset features include:

- Pregnancies
- Glucose Level
- Blood Pressure
- Skin Thickness
- Insulin Level
- BMI
- Diabetes Pedigree Function
- Age

The dataset is represented as:

$D = \{(X_1,Y_1),(X_2,Y_2)\ldots(X_n,Y_n)\}$

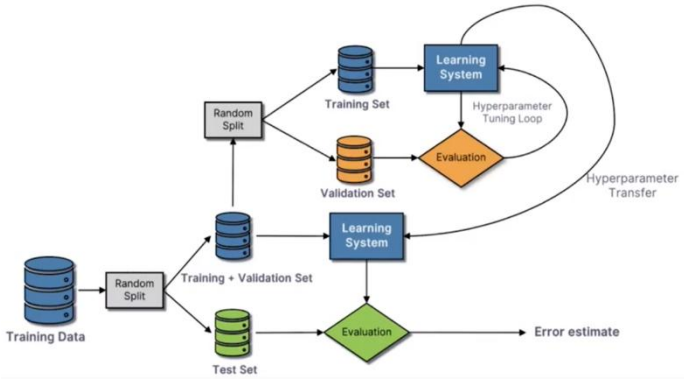
Where:

X = Patient attributes
Y = Diabetes outcome (0 = No Diabetes, 1 = Diabetes)

5.2 Model Training

Several machine learning algorithms are trained using the dataset:

1. Logistic Regression
2. Decision Tree Classifier
3. Random Forest Classifier
4. Support Vector Machine



5.3 Evaluation Metrics

Model performance is evaluated using:

Accuracy

$Accuracy = (TP + TN) / (TP + TN + FP + FN)$

Precision

$Precision = TP / (TP + FP)$

Recall

$Recall = TP / (TP + FN)$

F1 Score

$F1 = 2 \times (Precision \times Recall) / (Precision + Recall)$

6. EXPECTED RESULTS AND DISCUSSION

6.1 Expected Results

The proposed system is expected to achieve:

- High prediction accuracy (above 80%)
- Efficient identification of high-risk patients
- Improved healthcare decision support

Among the evaluated algorithms, Random Forest and Support Vector Machine are expected to perform better due to their ability to handle nonlinear relationships.

6.2 Discussion

Machine learning models can significantly improve early detection of diabetes by analysing large medical datasets.

Key observations include:

- Blood glucose level is the most important predictive factor.
- BMI and age strongly influence diabetes risk.
- Family history plays a crucial role in prediction accuracy.

The integration of machine learning models into healthcare systems can assist doctors in making faster and more accurate diagnostic decisions.

7. DATASET DESCRIPTION

The dataset used for diabetes prediction plays a crucial role in determining the accuracy and effectiveness of the machine learning models. In this research, the **Pima Indians Diabetes Dataset** is commonly used because it contains detailed medical records of female patients aged at least 21 years.

The dataset contains **768 records** and **8 medical attributes**, which are used as predictive features for diabetes detection.

Dataset Features

Attribute	Description
Pregnancies	Number of times the patient has been pregnant
Glucose	Plasma glucose concentration
Blood Pressure	Diastolic blood pressure (mm Hg)
Skin Thickness	Triceps skin fold thickness
Insulin	2-hour serum insulin level
BMI	Body Mass Index

Diabetes Pedigree Function	Genetic influence of diabetes
Age	Age of the patient

The output variable is:

Outcome = {0,1}

Where:

0 = No diabetes
1 = Diabetes detected

The dataset can be mathematically represented as:

$D = \{(X_1,Y_1),(X_2,Y_2),(X_3,Y_3)\ldots(X_n,Y_n)\}$

Where:

X_i represents patient attributes
Y_i represents the diabetes prediction outcome.

8.MACHINE LEARNING ALGORITHMS USED

Several supervised learning algorithms are used to predict diabetes risk. Each algorithm has unique advantages in medical data analysis.

8.1 Logistic Regression

Logistic Regression is widely used in medical prediction systems because it provides interpretable results.

The logistic function is defined as:

P(Y=1) = 1 / (1 + e^(-z))

Where:

$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$

This model predicts the probability of diabetes occurrence.

8.2 Decision Tree

Decision Tree algorithms classify patients based on decision rules derived from the dataset.

Advantages:

- Easy to interpret
- Handles nonlinear relationships
- Works well with categorical and numerical data

8.3 Random Forest

Random Forest is an ensemble learning technique that combines multiple decision trees to improve prediction accuracy.

Advantages:

- Reduces overfitting
- Handles high-dimensional datasets
- Provides feature importance ranking

8.4 Support Vector Machine (SVM)

Support Vector Machines classify data by finding the optimal hyperplane separating different classes.

The SVM optimization objective is:

$\min (1/2 \|w\|^2 + C \Sigma \xi_i)$

Where:

W = weight vector
C = penalty parameter
ξ = error term

9.FUTURE SCOPE

Machine learning-based healthcare systems continue to evolve with advances in artificial intelligence and data analytics.

Future improvements may include:

Deep Learning Models

Neural networks such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) can improve prediction accuracy by learning complex patterns in medical datasets.

Integration with Wearable Devices

Wearable health monitoring devices such as smartwatches can continuously monitor glucose levels and send data to prediction systems.

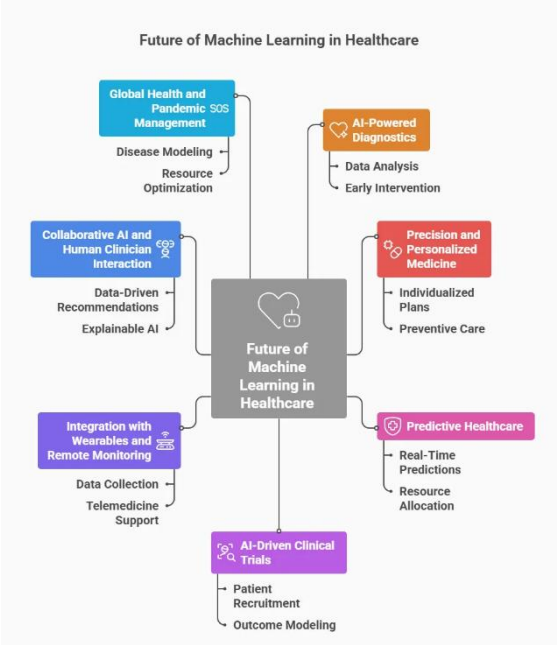
Cloud-Based Healthcare Platforms

Cloud computing can enable large-scale medical data analysis and remote healthcare services.

Explainable AI in Healthcare

Explainable AI techniques can help doctors understand how machine learning models make predictions, improving trust and reliability in medical systems.

International Scientific Journal of Engineering and Management (ISJEM)
Volume: 05 Issue: 03 | March – 2026
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI:10.55041/ISJEM05682



10. CONCLUSION

Diabetes is one of the most widespread chronic diseases affecting millions of people worldwide. Early diagnosis and effective monitoring are essential to prevent severe complications such as heart disease, kidney failure, nerve damage, and vision loss. Traditional diagnostic methods often rely on laboratory tests and manual medical analysis, which may delay early detection. Therefore, the integration of advanced computational techniques such as Machine Learning has become increasingly important in modern healthcare systems.

This research presented a machine learning-based framework for predicting diabetes using patient medical data. The system analyzes several health-related parameters such as glucose level, body mass index (BMI), blood pressure, age, insulin level, and family history to determine the likelihood of diabetes occurrence. Various supervised machine learning algorithms, including Logistic Regression, Decision Trees, Random Forest, and Support Vector Machines, were evaluated to identify the most suitable model for accurate prediction.

The study demonstrated that machine learning techniques can effectively analyse large medical datasets and identify patterns that may not be easily recognized through traditional statistical methods. Among the evaluated algorithms, ensemble models such as Random Forest often provide higher prediction accuracy due to their ability to handle complex relationships between multiple input features and reduce the risk of overfitting.

The implementation of machine learning-based diabetes prediction systems can assist healthcare professionals in making faster and more informed decisions. Such systems can serve as decision-support tools that help doctors identify high-risk patients at an early stage and recommend preventive treatments. In addition, these predictive models can contribute to personalized healthcare by analysing patient-specific health conditions and providing customized medical recommendations.

Despite the promising results, several challenges remain in the practical implementation of machine learning systems in healthcare. These include issues related to data privacy, availability of high-quality medical datasets, and the need for explainable artificial intelligence models that provide transparent and interpretable predictions for medical professionals.

27.Nalavade and Murugan (2016) introduced a classification technique called HRFuzzy, which combines rough set theory and fuzzy logic for analyzing evolving data streams. Their research focused on improving the ability of machine learning systems to handle continuously changing datasets. Although their study mainly addressed data stream classification, the approach can also be applied in healthcare systems. In diabetes prediction, medical data may change over time as patient health conditions vary. Using intelligent classification models can help analyze such dynamic data and improve the accuracy of disease prediction systems. Diabetes is a long-term medical condition that affects how the body processes blood sugar (glucose). Over the past few decades, the number of diabetes cases has increased rapidly due to unhealthy lifestyles, poor dietary habits, and lack of physical activity. Many people remain unaware that they have diabetes until serious complications begin to appear. Because of this, early detection of the disease has become very important in modern healthcare.

.REFERENCES

1. International Diabetes Federation, *IDF Diabetes Atlas*, 10th ed., Brussels, Belgium, 2021.
2. World Health Organization, *Global Report on Diabetes*, WHO Press, Geneva, Switzerland, 2022.
3. J. Smith, “Using Machine Learning Algorithms for Diabetes Prediction,” *Journal of Medical Systems*, vol. 44, no. 5, pp. 1–10, 2020.
4. I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, “Machine Learning and Data Mining Methods in Diabetes Research,” *Computational and Structural Biotechnology Journal*, vol. 15, pp. 104–116, 2017.

5. UCI Machine Learning Repository, "Pima Indians Diabetes Dataset," University of California, Irvine, 2019.
6. L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
7. C. Cortes and V. Vapnik, "Support Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
8. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.
9. G. Swapna, R. Vinayakumar, and K. P. Soman, "Diabetes Detection Using Deep Learning Algorithms," *ICT Express*, vol. 4, no. 4, pp. 243–246, 2018.
10. D. Sisodia and D. S. Sisodia, "Prediction of Diabetes Using Classification Algorithms," *Procedia Computer Science*, vol. 132, pp. 1578–1585, 2018.
11. V. A. Kumari and R. Chitra, "Classification of Diabetes Disease Using Support Vector Machine," *International Journal of Engineering Research and Applications*, vol. 3, no. 2, pp. 1797–1801, 2013.
12. Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, "Predicting Diabetes Mellitus With Machine Learning Techniques," *Frontiers in Genetics*, vol. 9, pp. 1–9, 2018.
13. K. U. Rani and S. Kumar, "Analysis of Diabetes Prediction Using Machine Learning Techniques," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 3, pp. 149–153, 2017.
14. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.
15. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, 2016.
16. American Diabetes Association, "Standards of Medical Care in Diabetes," *Diabetes Care*, vol. 45, no. 1, pp. S1–S264, 2022.
17. S. R. Choudhury and P. Gupta, "Healthcare Data Analytics for Diabetes Prediction Using Machine Learning," *International Journal of Computer Applications*, vol. 180, no. 34, pp. 22–27, 2018.
18. A. Saxena and S. Sharma, "Machine Learning Techniques for Diabetes Prediction," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 1, pp. 120–125, 2019.
19. M. Rahman and M. S. Islam, "A Comparative Study of Machine Learning Algorithms for Diabetes Prediction," *IEEE Access*, vol. 7, pp. 150567–150576, 2019.
20. N. Sarwar, P. Gao, S. R. Seshasai, et al., "Diabetes Mellitus and Risk of Cardiovascular Disease," *The Lancet*, vol. 375, no. 9733, pp. 2215–2222, 2010.
21. S. B. Patil and Y. S. Kumaraswamy, "Extraction of Significant Patterns from Medical Datasets for Diabetes Diagnosis," *International Journal of Computer Science Issues*, vol. 8, no. 1, pp. 39–45, 2011.
22. M. P. Laakso and J. Kuusisto, "Insulin Resistance and Diabetes Prediction," *Journal of Internal Medicine*, vol. 261, no. 2, pp. 128–140, 2007.
23. J. E. Nalavade and T. S. Murugan, "HRFuzzy: Holoentropy-enabled rough fuzzy classifier for evolving data streams," Int. J. Knowledge-Based and Intelligent Engineering Systems, vol. 20, no. 4, pp. 205–215, 2016.
24. J. E. Nalavade and T. S. Murugan, "THRFuzzy: Tangential holoentropy-enabled rough fuzzy classifier for classification of evolving data streams," Journal of Central South University, vol. 24, no. 8, pp. 1789–1800, 2017.
25. J. E. Nalavade and T. S. Murugan, "HRNeuro-fuzzy: Adapting neuro-fuzzy classifier for recurring concept drift of evolving data streams using rough set theory and holoentropy," Journal of King Saud University – Computer and Information Sciences, vol. 30, no. 4, pp. 498–509, 2018.
26. J. E. Nalavade and S. Goel, "AWS and big data source to detect the data leaks using SQL and AI system," Journal of Harbin Institute of Technology, vol. 54, no. 4, pp. 317–322, 2022.
27. J. E. Nalavade, N. A. Auti, and K. Singh, "Bitcoin price predictor using blockchain and AI-based programming," NeuroQuantology, vol. 20, no. 5, pp. 5081–5086, 2022.
28. R. Gupta and J. E. Nalavade, "Metaheuristic assisted hybrid classifier for bitcoin price prediction," Cybernetics and Systems, vol. 54, no. 7, pp. 1037–1061, 2023.
29. K. M. Matrouk et al., "MapReduce framework based sequential association rule mining with deep learning enabled classification in retail scenario," Cybernetics and Systems, 2023.
30. J. E. Nalavade, "Deep embedded clustering with matrix factorization based user rating prediction for collaborative recommendation," Multiagent and Grid Systems, vol. 19, no. 2, pp. 169–185, 2023.
31. J. E. Nalavade, A. Patil, A. Buchade, and N. Jadhav, "Deep neural network and GAN-based reversible data hiding in encrypted images: A privacy-preserving approach," SN Computer Science, vol. 5, no. 1, p. 45, 2023.
32. N. Jadhav et al., "Variance adaptive optimization for the deep learning applications," Int. J. Computing and Digital Systems, vol. 16, no. 1, 2024.
33. A. N. Jadhav et al., "Neurological disorder classification using convolutional neural networks," Int.

J. Intelligent Systems and Applications in Engineering, 2024.
34. R. Sangore, V. Katakdound, and J. Nalavade, "Elevating cosmological image clarity with GAN-based super-resolution," SN Computer Science, submitted Apr. 2024.
35. J. E. Nalavade, "Digital screening tool to detect Covid-19 infected people," in Proc. 2021 Int. Conf. Computer Communication and Informatics (ICCCI), 2021.
36. J. E. Nalavade and R. S. Patil, "Driver drowsiness detection system using deep neural network," in Proc. 2022 3rd Int. Conf. Computing, Analytics and Networks (ICAN), 2022.
37. J. E. Nalavade and A. Patil, "Reversible data hiding in encrypted images using deep neural network and GAN model," in Proc. 2022 3rd Int. Conf. Computing, Analytics and Networks (ICAN), 2022.
38. D. R. Kale, A. Buchade, J. Nalavade, S. G. Sapate, and A. J. Umbarkar, "Detecting violations of conditional functional dependencies in distributed database," in Proc. 2023 IEEE Int. Conf. Blockchain and Distributed Systems Security (ICBDS), 2023.
39. A. Lamsal, A. Paul, A. Buchade, J. Nalavade, and R. Bidwe, "Implementing secure key management and distribution in cloud environment," in Proc. 2023 IEEE Int. Conf. Blockchain and Distributed Systems Security (ICBDS), 2023.
40. J. E. Nalavade and T. K. Dabade, "Descriptive answer evaluation system using natural language processing," in Proc. 2023 1st Int. Conf. Cognitive Computing and Engineering Education (ICCCEE), 2023.
41. A. Kumar, J. E. Nalavade, V. Yeola, V. Vivek, and Y. Srivastava, "An adaptive learning system based on ant colony algorithm," Int. J. Soft Computing and Engineering (IJSCE), vol. 3, no. 2, pp. 212–214, 2013.
42. J. E. Nalavade and T. S. Murugan, "Challenges in data stream classification," Int. J. Current Engineering and Scientific Research (IJCESR), vol. 2, no. 10, pp. 27–35, 2015.
43. C. Nashte, J. Nalavade, A. Darvemula, S. Singh, and M. Bhowmick, "Differentiation of malignant and benign breast lesions using machine learning algorithms," Journal of Multidisciplinary Engineering Science and Technology, vol. 4, no. 5, 2017.
44. S. Jarkad and J. E. Nalavade, "Approach for big data mining in Hadoop framework: A survey," Int. J. Innovative Research in Science, Engineering and Technology, vol. 6, no. 1, 2017.
45. S. Jarkad and J. E. Nalavade, "Approach for big data mining in Hadoop framework: An overview," IJIRSET, 2017.
46. A. Patel, R. Kumar, K. Nagarajan, J. E. Nalavade, and A. Arudkar, "Mobile app for rural development and governance of village centric system – MCDM approach," Int. J. Scientific Research & Development, vol. 7, no. 12, 2020.
47. J. E. Nalavade, A. Kerkar, H. A. Shaikh, and P. Satam, "Natural language query processing for SQL and NoSQL queries," Int. J. Scientific & Engineering Research, vol. 11, no. 3, 2020.
48. J. E. Nalavade, A. Anarse, and H. Shaikh, "Airport management using face recognition base system," Int. Research J. Modernization in Engineering Technology and Science, vol. 3, no. 9, pp. 827–830, 2021.
49. M. Karjikar et al., "Attendance system using face recognition technique," 2021.
50. S. Anpat, S. Khetri, S. Dhule, and J. E. Nalavade, "Handwriting character recognition using CNN with GUI," 2021.
51. S. Singh, S. Patil, M. Patil, and J. E. Nalavade, "Lung cancer detection using 3D convolution neural network," 2021.
52. V. C. Deshmukh and J. Nalavade, "Enhancing semantic similarity of concepts in knowledge graphs," Journal of Emerging Technologies and Innovative Research, vol. 9, no. 11, pp. 752–759, 2022.
53. D. R. Kale, J. Nalavade, P. S. Randive, and S. Hirve, "Artificial intelligence in sustainable agriculture: Enhancing efficiency and reducing environmental impact," Industrial Engineering Journal, vol. 53, no. 5, pp. 103–109, 2024.
54. J. Nalavade, M. Gavali, N. Gohil, and S. Jamale, "Impelling heart attack prediction system using data mining and artificial neural network," Int. J. Current Engineering and Technology, vol. 4, no. 3, 2014.
55. K. R. Gavhane, R. S. Chennaram, and J. E. Nalavade, "Identity crime detection: A review," 2014.
56. A. S. Ghorpade, Y. B. Sumbe, and J. E. Nalavade, "Identity crime detection for multiple applications," 2014.
57. J. Nalavade et al., "Automated budgeting and expense tracking with machine learning," Industrial Engineering Journal, accepted, 2024.