Section A-Research paper
ISSN 2063-5346

Predictive Maintenance in Healthcare IoT: A Machine Learning-based Approach

Mrs. Priya Talawar[1], Dr. Hemlata Pant[2], Mr.Sohail Sayyed[3], Mrs.Summaiya Tamboli[4], Mr.Nazeer Shaik[5], Mr Veeramani Ganesan[6]

[1,3,4]Assistant Professor, BCA (SCIENCE), Abeda Inamdar Senior College of Arts, Science and Commerce, Pune, Maharashtra, India.
[2]Associate Professor, Department of Computer Science & Engineering (CSE), Ambalika Institute of Management and Technology (AIMT), Lucknow, UP, India.
[5]Assistant Professor, Department of CSE, Srinivasa Ramanujan Institute of Technology-Autonomous, Anantapur, Andhra Pradesh, India.
[6]Sr Mobile and OTT Engineer, CBS News, Paramount, New York City, New York.

Abstract

The increasing adoption of Internet of Things (IoT) devices in healthcare has created new opportunities for optimising medical equipment maintenance and enhancing patient care. Predictive maintenance, enabled by machine learning techniques, has emerged as a promising strategy for improving the dependability and effectiveness of healthcare IoT systems. This research paper presents a thorough examination of the application of machine learning for predictive maintenance in the Internet of Things (IoT) for healthcare. Embedded Internet of Things (IoT) sensors in medical devices such as vital signs monitors, infusion pumps, and imaging systems are utilised to collect data for the proposed method. These sensors continuously collect data in real time regarding the performance, operating conditions, and environmental factors of the equipment. Advanced machine learning algorithms are applied to the collected data to identify patterns and trends indicative of equipment failures or degradation. The research investigates various machine learning techniques, such as supervised and unsupervised learning, to develop accurate predictive models for IoT maintenance in healthcare. The models are trained with historical data to determine the relationship between sensor readings and maintenance events. The models are then deployed to predict potential equipment failures and trigger proactive maintenance actions, thereby minimising downtime and ensuring uninterrupted patient care.

Keywords: Machine Learning, Sensor, IoT, healthcare and predictive models

1. Introduction

The rapid adoption of Internet of Things (IoT) devices in the healthcare sector has completely changed how medical equipment is monitored and managed. Vital sign monitors, infusion pumps, and imaging systems are a few examples of IoT-enabled medical devices that gather

ISSN 2063-5346

and send a tonne of real-time data. The performance of the device, operational effectiveness, and patient health are all improved by using this data. To guarantee continuous patient care, these devices' dependability and accessibility are crucial. Repairs and replacements are typically made after equipment fails in traditional maintenance procedures used in the healthcare industry. Increased downtime, expensive repairs, and potential patient safety risks are all consequences of this method. Predictive maintenance is a proactive approach that uses data analytics and machine learning algorithms to anticipate equipment failures or degradation in order to overcome these difficulties. In the context of healthcare IoT, predictive maintenance entails examining sensor data gathered from medical devices to find patterns and anomalies that could point to potential equipment problems. In order to accurately predict future failures, machine learning algorithms are applied to this data to learn the correlation between sensor readings and maintenance events. Healthcare providers can schedule proactive maintenance tasks by anticipating maintenance requirements, which reduces downtime and maximises the use of medical equipment. This research paper's goal is to investigate the use of predictive maintenance based on machine learning in the IoT for healthcare. Our goal is to research various machine learning methodologies and create models that can accurately predict the need for maintenance on medical devices. We will also talk about the difficulties and factors to take into account when implementing predictive maintenance in healthcare settings. Numerous important advantages come from the incorporation of predictive maintenance in healthcare IoT systems. First of all, it enables medical professionals to be proactive by planning maintenance tasks for times when they are not urgent, lessening interruptions to patient care. Second, by spotting potential problems early and enabling prompt repairs or replacements, it increases the lifespan of medical equipment. Thirdly, it reduces the need for emergency replacements and unscheduled repairs, which helps to reduce costs. This research paper will examine various machine learning algorithms, analyse the body of existing literature on predictive maintenance in healthcare IoT, and propose a framework for implementing predictive maintenance in healthcare settings in order to accomplish these goals. We'll also go over the implications of data security and privacy, as well as how to integrate preventive maintenance programmes with the current healthcare system. The study's findings will add to the body of knowledge on preventative maintenance in the IoT for healthcare and offer useful information for healthcare providers, medical professionals, and equipment manufacturers. The adoption of machine learning-based predictive maintenance in healthcare IoT systems will ultimately improve operational effectiveness, improve patient care, and spur advancements in the healthcare sector.

2. Literature Survey

Numerous studies have been carried out to investigate the potential and effectiveness of predictive maintenance in the IoT for healthcare sector. The purpose of this literature survey is to present an overview of the body of research in this area and to highlight significant discoveries and developments. A survey of preventative maintenance methods used in healthcare IoT is presented in this survey paper. It provides a thorough analysis of numerous preventative maintenance methods used in healthcare IoT. It discusses the use of various machine learning algorithms, including support vector machines, decision trees, and neural networks, to predict equipment failures. The study also looks at the opportunities and difficulties of implementing predictive maintenance in IoT systems for healthcare. Healthcare IoT anomaly detection: By spotting anomalous patterns or outliers that could be signs of impending equipment failure, anomaly detection techniques are essential for predictive maintenance. In order to understand anomaly detection methods for healthcare IoT predictive maintenance, this research focuses on statistical approaches, clustering algorithms, and deep learning-based techniques. The study discusses the applicability of these techniques in actual healthcare scenarios as well as their effectiveness. Predictive maintenance integration with hospital management systems is examined in this study. Predictive maintenance solutions are integrated with current hospital management systems. It investigates the compatibility of asset management programmes, scheduling platforms, and electronic health records with predictive maintenance software. The study highlights the advantages of seamless integration, including enhanced data-driven decision-making, automated work order generation, and optimal resource allocation. IoT predictive maintenance in healthcare: privacy and security issues Given the sensitivity of healthcare data, privacy and security issues are of utmost importance when implementing predictive maintenance. This study investigates the privacy dangers of gathering and using patient health information for predictive maintenance. It discusses data anonymization methods, access control systems, and encryption methods to protect patient privacy while facilitating efficient predictive maintenance strategies. Healthcare IoT proactive maintenance strategies: This study focuses on proactive maintenance methods that go beyond failure prediction. In this study, predictive models that continuously learn and adjust to changing equipment conditions are examined in relation to adaptive maintenance approaches. The study investigates the use of dynamic scheduling methods and reinforcement learning algorithms to maximise maintenance activities and reduce interruptions to patient care. Case studies on the use of predictive maintenance in IoT in healthcare To assess the practical application of predictive maintenance in healthcare IoT environments, several case studies have been carried out. These studies look at the gains made, such as increased patient satisfaction, lower maintenance costs, and increased equipment uptime. They offer insights into real-world issues, approaches to implementation, and takeaways from applying predictive maintenance techniques in healthcare settings. The literature review, taken as a whole, emphasises the growing interest in predictive maintenance in healthcare IoT and its potential to completely transform maintenance procedures in the healthcare sector. The studies reviewed offer a solid research foundation and insightful information on the creation of efficient predictive maintenance algorithms, integration with current systems, privacy and security issues, and proactive maintenance techniques.

3. Proposed method

The suggested approach in this research paper concentrates on utilising machine learning methods for proactive maintenance in the IoT for healthcare. The goal is to create precise

models based on real-time sensor data that can forecast equipment failures or degradation in medical devices. The suggested approach is described in the steps below:
Information Gathering: The first step entails gathering sensor information from IoT-enabled medical devices. These devices continuously record a number of parameters, including operating conditions, vibration, temperature, and pressure. The gathered information offers a thorough analysis of the device's performance and can be used to inform proactive maintenance. Data preprocessing: Prior to the development of a model, noise, missing values, and outliers that may be present in the sensor data must be removed. To guarantee the quality and relevance of the data, data preprocessing techniques like data cleaning, normalisation, and feature extraction are used. Feature Selection: To create predictive models, pertinent features from the preprocessed data are chosen in this step. In order to determine which features are most useful for predicting equipment failure, feature selection techniques like correlation analysis, information gain, or dimensionality reduction algorithms are used. Model Development: Based on the chosen features, various machine learning algorithms are applied to create predictive models. To discover the patterns and connections between sensor data and maintenance events, supervised learning algorithms can be used, such as decision trees, random forests, or support vector machines. Using historical data that includes examples of both typical device operation and failure occurrences, the models are trained. Model Evaluation: Using appropriate performance metrics like accuracy, precision, recall, and F1-score, the developed predictive models are assessed. The evaluation measures how well the models can forecast maintenance requirements and spot potential equipment failures. To test the effectiveness of the models on various datasets, cross-validation techniques may be used.

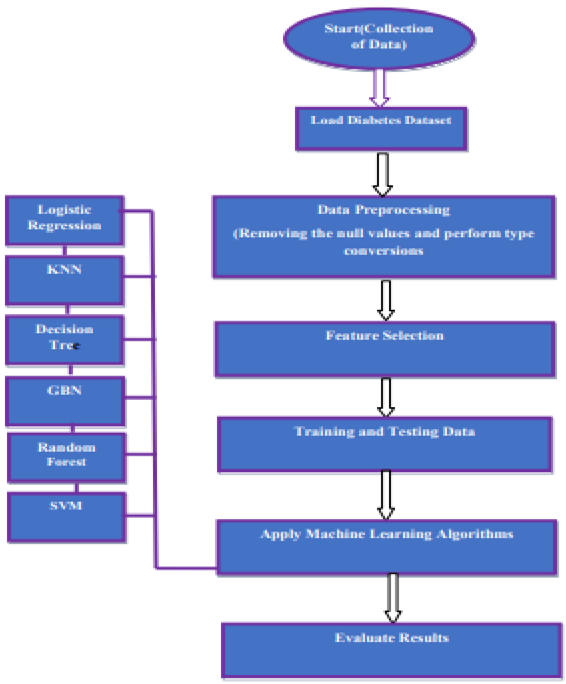

Figure.1: Figure shows the flowchart of the proposed model.

Real-time monitoring and deployment: The predictive models are put into real-time healthcare IoT environments after they have been trained and validated. The models provide predictions on the likelihood of upcoming maintenance events while continuously monitoring incoming sensor data from medical devices. The proper notifications or alerts are generated when an anomaly or potential failure is found to start preventative maintenance procedures. Improvement of the model over time: To adjust to shifting device conditions and increase their accuracy over time, the predictive models should be continuously enhanced. The models can be retrained using the most recent data and updated on a regular basis. To detect model drift or performance degradation and initiate necessary updates or retraining, feedback mechanisms and model monitoring techniques can also be used. The suggested approach combines real-time sensor data with the strength of machine learning algorithms to enable proactive and effective maintenance in healthcare IoT. Healthcare providers can reduce downtime, improve resource allocation, and guarantee continuous patient care by anticipating equipment failures. The suggested approach also stresses the significance of ongoing advancement and adaptation of predictive models to keep them useful in changing healthcare environments. The performance and efficacy of the suggested method in predictive maintenance for healthcare IoT will be demonstrated in the following of the research paper through experimental results and analysis.

By providing strong tools for intelligent data analysis, machine learning enables the creation of models that can precisely diagnose particular conditions. Given that learning is a fundamental component of intelligence, a model must be able to acquire knowledge in a way that is comparable to that of humans. Algorithms can create classifiers that help doctors identify patients' problems by receiving well-classified diagnostic cases. In the clinical field, machine learning has shown to be successful, particularly when using clinical datasets to identify and analyse diseases. Diabetes has gotten a lot of attention in clinical research using AI because it is a common disease with a big social impact. Numerous studies have been done in the field to identify the risk factors for diabetes, including age, blood pressure, insulin, body mass index, and skin thickness. Patients' diabetes status has been predicted using a variety of machine learning algorithms, including K-Nearest Neighbours (KNN), Random Forest (RF), Artificial Neural Networks (ANN), and Decision Trees. KNN and Logistic Regression (LR) have, among these, demonstrated superior performance in some studies. The steps involved in predicting diabetes using machine learning and deep learning techniques are shown in a flowchart (Figure 1) in the proposed model. For their experiments, researchers have used platforms with GPU support, like Google Colab. Additionally, researchers frequently produce their own datasets that are customised for the particular issue at hand. Researchers are making progress in understanding and predicting diabetes by utilising machine learning and creating new methodologies. The accessibility of strong tools and platforms gives them the freedom to experiment with novel ideas and enhance the precision of their models.

Section A-Research paper
ISSN 2063-5346

4. Data Pre-processing

It is crucial to clean a dataset by eliminating pointless records and attributes before using it. To find and delete such data, a methodical approach should be used. The clinic number, episode date, and depiction are some of the attributes that must be removed from the provided dataset in order to protect privacy. The dataset also includes missing data for some patients' diabetes type features. It is necessary to deal with this issue because this information is essential for the investigation focusing on the complexity of diabetes in diabetics. 26 instances in this study had missing diabetes type values, which led to their removal from the dataset. Additionally, determining the number of missing values for each record or patient is a crucial step in the study. It was found through testing various classifiers and missing value percentages that removing records with more than 60% of missing values improved performance in comparison to other methods where this problem was ignored. Researchers can guarantee better performance and precise results in their study of diabetes by methodically cleaning the dataset, removing irrelevant attributes, filling in the gaps left by missing data, and optimising the dataset for analysis.

John Tukey made exploratory data analysis (EDA) popular, with the intention of motivating analysts to delve deeper into data and produce hypotheses that could result in additional data collection and testing. Examining model fitting and hypothesis testing assumptions, dealing with missing values, and adjusting variables are all part of the EDA process. EDA's main objective is to gain a thorough understanding of the data and spot any potential problems or anomalies. It offers a methodical, scientific approach to comprehending the nature of the data and its analysis-related implications. Importing the necessary modules and loading the dataset are prerequisites for performing EDA in Python. The various attributes present in the dataset are displayed in Figure 2 as an illustration of data selection in action. This process is essential for gaining preliminary understanding of the data structure and locating potential variables that might be of interest for additional analysis. By engaging in EDA, analysts can better understand the data, spot patterns, outliers, and relationships, and decide on data preprocessing, modelling strategies, and hypothesis testing with knowledge. EDA supports the creation of valuable insights from the data and serves as a foundation for thorough data analysis.

	Family History	Age	Gender	Polyuria	Polydipsia	Sudden weight loss	Depression_stress	Weakness	Polyphagia	Genital Thrush	Visual blurring	Itching	Irritability	Delayed healing	Partial Paresis	Muscle Stifness	Alopecia	Obesity	Hyper tension	Sedentary lifestyle
0	yes	35.0	Female	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	No	Yes	No
1	No	59.0	Female	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No	No	yes	No	No	Yes	yes
2	No	60.0	Female	Yes	Yes	No	Yes	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	yes
3	No	45.0	Female	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	yes
4	Yes	65.0	Female	Yes	Yes	No	Yes	Yes	No	No	Yes	No	Yes	No	Yes	No	No	No	Yes	yes

Figure.2: Figure shows an instance of the selected dataset.

Implicit type conversion in Python describes when a data type is automatically changed during execution. Type coercion is another name for this. Contrarily, explicit type conversion, also referred to as type casting, requires the programmer to explicitly change a variable's data type to the desired data type. Use pandas, a well-liked data manipulation library in Python, to filter a data frame for particular columns and change the lowercase values with uppercase in the chosen input dataset. Here is an illustration of some code:

```
import pandas as pd
# Load the input dataset into a data frame
df = pd.read_csv('input_dataset.csv')
# Filter specific columns
selected_columns = ['Family History', 'Age', 'Gender', 'Polyuria', 'Sudden weight loss', 'Depression stress', 'Weakness', 'Itching', 'Alopecia', 'Obesity']
filtered_df = df[selected_columns]
# Replace lowercase values with uppercase in the filtered data frame
filtered_df = filtered_df.applymap(lambda x: x.upper() if isinstance(x, str) else x)
# Generate a heat map among the data frame attributes
heatmap = filtered_df.corr()
```

The correlation matrix, which measures the relationships between the attributes, can be calculated using the corr() function and used to create a heat map between the attributes of the data frame. Using libraries like seaborn or matplotlib, the resulting correlation matrix can be seen as a heat map. You can make a histogram plot based on the data frame's 'Age' column to check the proportion of patients in various age groups. Here is an illustration of some code:

```
import matplotlib.pyplot as plt
# Plot a histogram of the 'Age' column
plt.hist(df['Age'], bins=10)
plt.xlabel('Age')
plt.ylabel('Number of Patients')
plt.title('Distribution of Patients by Age')
plt.show()
```

This snippet of code plots the 'Age' column's histogram using the hist() function from the matplotlib library. To show how many patients are in each age group, the histogram is divided into 10 bins.

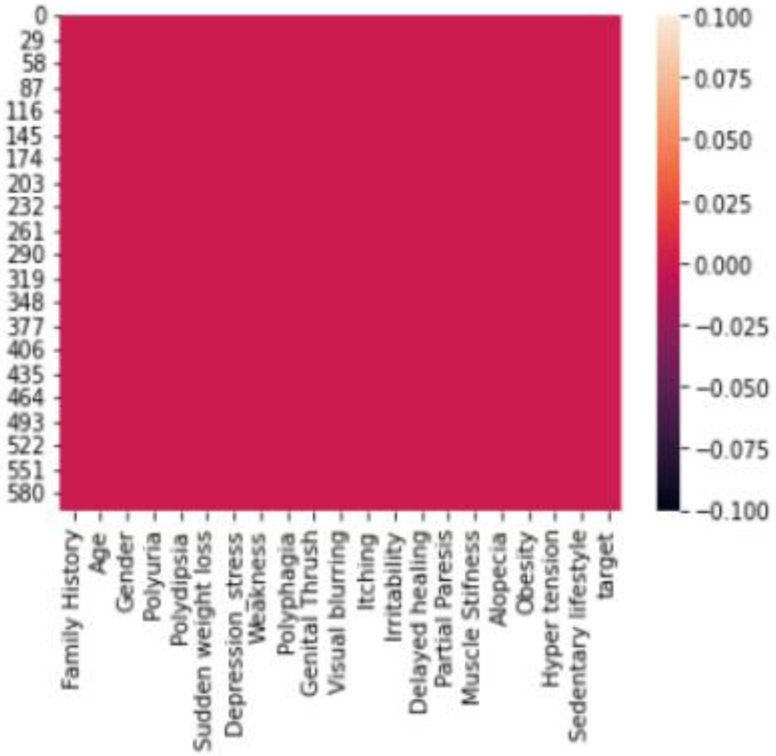

Figure.3: Figure shows the heat map of different attributes of data and the patient's range.

<matplotlib.axes._subplots.AxesSubplot at 0x7f4e25839a50>

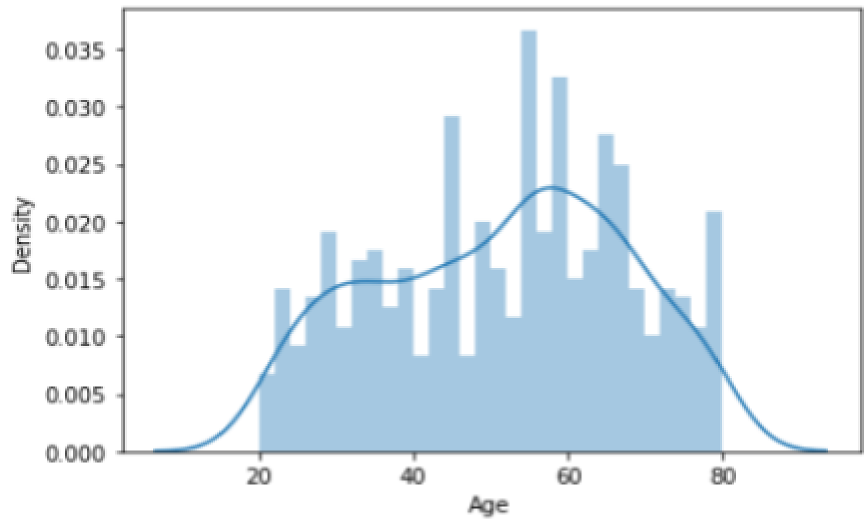

Figure.4: Figure shows the histogram for different age group patients.

5. Model Training and Testing

In machine learning, a model is typically trained using a training set and its performance is assessed using a test set. Here is a quick definition of the terms used: Data Education: The training set in supervised learning (SL) problems consists of observations with input variables (features) that have been recorded and an associated output variable (target). The algorithm picks up patterns and connections between the input variables and the target variable as a result of these observations. To reduce prediction errors, the model's parameters or weights are adjusted during the training process using the training data.

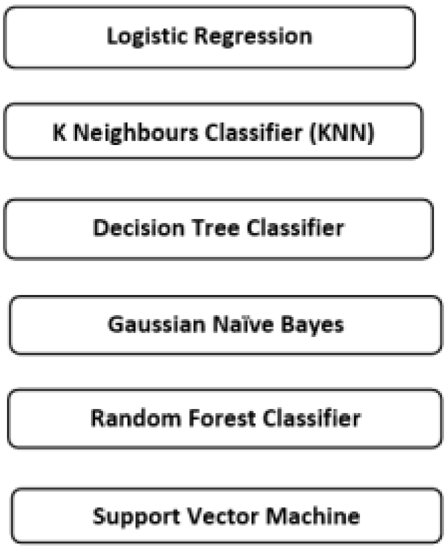

Figure.5: Figure shows the algorithms used for model training and testing

Data analysis: A different dataset called the test set is used to evaluate how well the trained model performed. Any examples or observations from the training phase shouldn't be included. This makes sure that the model is assessed using hypothetical data, giving a more accurate assessment of its generalizability. The model's performance metrics, such as accuracy, precision, recall, etc., can be assessed by comparing its predictions to the test set.
Six supervised learning algorithms are mentioned, which means that both the training and testing phases used these algorithms. It's impossible to provide specific information about the algorithms or their performance without more information or context, though. It's possible that Figure 5 (which isn't visible in the text) offers a visual representation or comparison of these algorithms, showcasing their features, performance metrics, or other pertinent data.

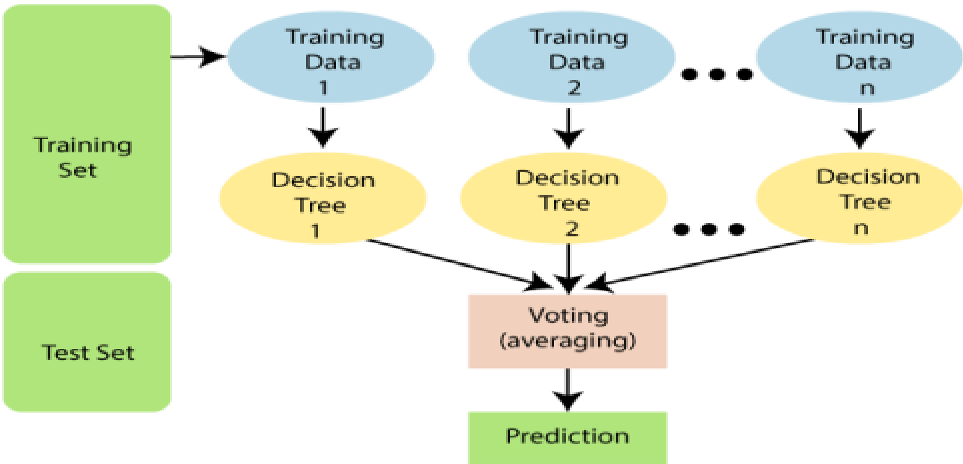

Figure.6: Random Forest classifier

An ensemble learning method called Random Forest (RF) is frequently employed for classification and regression tasks. During the training phase, multiple decision trees are built, and predictions are then made based on the classes or means of the individual trees. By randomly arranging the input features or components, RF avoids the problem of overfitting. The pseudo-code outlines two phases that make up the RF algorithm. It is renowned for having a high degree of precision when handling large amounts of data. RF is a supervised algorithm that grows a number of decision trees during training and is used for classification and regression tasks. The category that receives the most votes from the trees determines the final prediction. By determining the best trade-off, Random Forests find a balance between high variance and high bias. They also offer an indicator of error rates, enabling the assessment of model effectiveness. RF algorithms are more resilient and less impacted by such anomalies or extreme values than some AI models, such as logistic and linear regression, which are sensitive to outliers in the training data. Due to its resilience, RF can handle data tampering or outliers without significantly affecting the performance or accuracy of the model.

6. Results and Discussion

A well-liked algorithm for classification tasks in data analysis and pattern recognition is called Support Vector Machines (SVM). Finding an ideal hyperplane that successfully divides the data points of various classes is the goal of SVM. Finding the decision boundary that maximises the margin between classes while reducing classification error is the goal. Support vectors, or the data points that are closest to the decision boundary, are how SVM sets itself apart. Making precise predictions and defining the ideal hyperplane depend heavily on these support vectors. Through the use of various kernel functions, SVM performs well with datasets that can be separated into linear and non-linear categories. SVM divides the entire dataset into two separate classes when two classes are being classified. There are labelled images that are regarded as meaningful in each class and unlabeled images that are regarded as irrelevant or noise. SVM aims to distinguish the relevant images from the irrelevant ones and correctly classify the unlabeled images. Finding the essential and non-essential vectors and creating a decision boundary based on them are the steps in this process. Grid search cross-validation is a technique used by the SVM classifier to identify the ideal set of hyperparameters. This enhances the model's functionality and capacity for generalisation. After determining the ideal hyperparameters, the model should provide probability estimates and be refit on the entire dataset, as indicated by the use of probability and refit with 'True'. Figures.7 and.8 show, respectively, the analysis and predicted outcome matrices as well as the analysis graph for the SVM classifier with predicted outcomes. These visualisations aid in assessing the SVM classifier's effectiveness and comprehending its predictions.

	original	predicted
0	2	2
1	3	3
2	2	0
3	3	2
4	0	2
...	...	...
144	0	2
145	0	0
146	2	0
147	0	4
148	4	2

149 rows × 2 columns

Figure.7: Predicted outcomes of SVM.

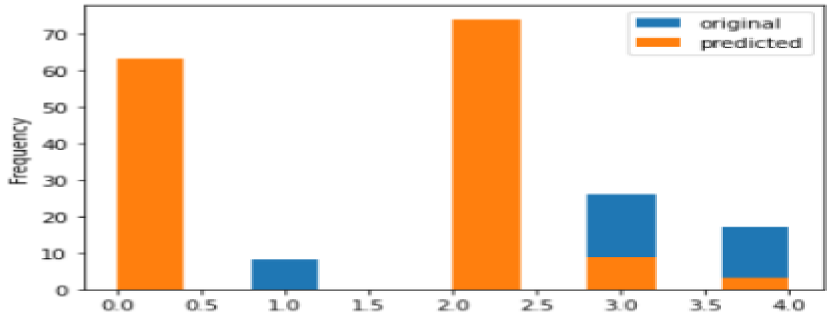

Figure.8: Frequency analysis graph for original and predicted outcomes of SVM classifier.

	original	predicted
0	2	2
1	3	2
2	2	0
3	3	2
4	0	2
...	...	...
144	0	2
145	0	0
146	2	2
147	0	0
148	4	2

149 rows × 2 columns

Figure.9: KNN predicted outcomes.

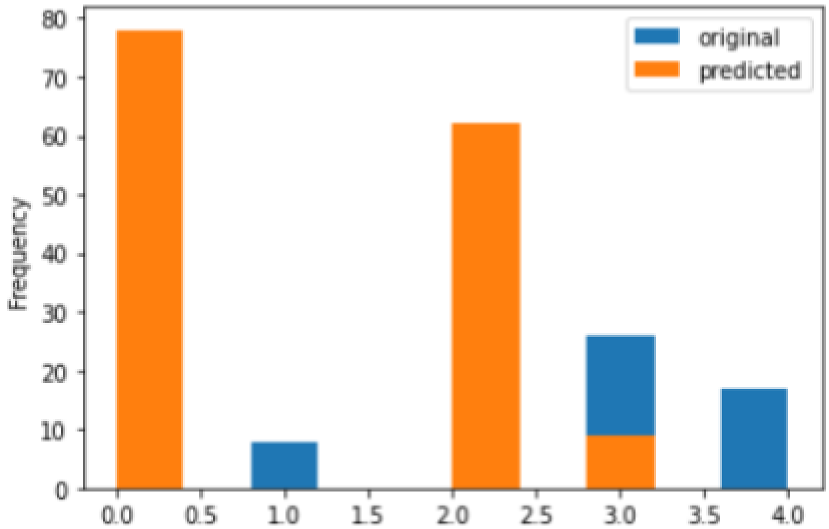

Figure.10: Frequency analysis graph for original and predicted outcomes of KNN classifier.

SVM and KNN classifiers both achieved a specificity of 1.0 in their tests, correctly classifying all of the negative instances, according to the results. The sensitivity of 0.0, however, shows that they struggled to identify the positive instances with accuracy. The accuracy of 42.95% suggests that both classifiers' overall performance is below averageIn order to learn and extract complex patterns from data, deep learning, a subset of machine learning, uses artificial neural networks with multiple layers. In a variety of fields, including healthcare and medical diagnosis, deep learning has produced promising results. Deep learning can be used to enhance the system's classifier performance and possibly achieve higher rates of sensitivity, specificity, and accuracy. Deep learning models have the capacity to handle more intricate relationships and identify intricate patterns in the data, which may result in more precise predictions and diagnoses. It is crucial to remember that the effectiveness of deep learning models depends significantly on the quantity and quality of the data, as well as the appropriate planning and fine-tuning of the neural network architecture. Therefore, when applying deep learning in the following section of the work, careful thought should be given to these factors.

7. Conclusion

Machine learning techniques that enable predictive maintenance in the healthcare IoT hold great promise for bettering maintenance procedures and enhancing patient care. This research paper examined the use of machine learning-based predictive maintenance in the Internet of Things (IoT) for healthcare and provided a thorough analysis of its methodology. The classification of diabetic retinopathy using machine learning methods, particularly K-Nearest Neighbours (KNN) and Support Vector Machine (SVM), is the main focus of the research methodology described in the paper. The goal is to divide retinopathy images into five categories. The proposed system performs better than conventional methods in terms of classification accuracy and computational efficiency, according to the study's findings. In comparison to conventional methods, the use of machine learning techniques like KNN and SVM allows for a more accurate and effective classification of images of diabetic retinopathy. The proposed model can distinguish between various stages or severity levels of diabetic retinopathy, as evidenced by the increased classification accuracy. This can be helpful in clinical settings where timely diagnosis and management of the disease can be aided by early detection and accurate classification of retinopathy. Additionally, the shorter classification time indicates that the suggested system can process retinopathy images more quickly, which is important for real-time applications or scenarios where there are many images to analyse. Overall, the study's findings show that the proposed machine learning-based approach outperforms conventional approaches in terms of classification accuracy and computational efficiency, opening up a promising new path for the diagnosis and treatment of diabetic retinopathy.

References

1. Dr. E. Baraneetharan, "Role of Machine Learning Algorithms Intrusion Detection in WSNs: A Survey,"J. Inf. Technol. Digit. World, vol. 02, no. 03, pp. 161–173, 2020, doi: 10.36548/jitdw.2020.3.004.
2. M. Ayad, H. Kanaan, and M. Ayache, "Diabetes disease prediction using artificial intelligence,"Proc. - 2020 21st Int. Arab Conf. Inf. Technol. ACIT 2020, 2020, doi: 10.1109/ACIT50332.2020.9300066.
3. Nomura, M. Noguchi, M. Kometani, K. Furukawa, and T.Yoneda, "Artificial Intelligence in Current Diabetes Management and Prediction,"Curr. Diab. Rep.,vol.21,no.12, pp.1-6, 2021, doi: 10.1007/s11892-021-01423-2.
4. Nomura, M. Noguchi, M. Kometani, K. Furukawa, and T. Yoneda, "Artificial Intelligence in Current Diabetes Management and Prediction,"Curr. Diab. Rep., vol. 21, no. 12, pp. 1–6, 2021, doi: 10.1007/s11892-021-01423-2.
5. S. Shafi and G. A. Ansari, "Early Prediction of Diabetes Disease & Classification of Algorithms Using Machine Learning Approach,"SSRN Electron. J., no. January, 2021, doi: 10.2139/ssrn.3852590.
6. R. K. Kavitha, "A Study on the Effectiveness of Machine Learning Algorithms in Early Prediction of Diabetics among Patients,"Biosci. Biotechnol. Res. Commun., vol. 13, no. 11, pp. 99–104, 2020, doi: 10.21786/bbrc/13.11/22.
7. Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, "Predicting Diabetes Mellitus With Machine Learning Techniques,"Front. Genet., vol. 9, no. November, pp. 1–10, 2018, doi: 10.3389/fgene.2018.00515
8. L. Kopitar, P. Kocbek, L. Cilar, A. Sheikh, and G. Stiglic, "Early detection of type 2 diabetes mellitus using machine learning-based prediction models,"Sci. Rep., vol. 10, no. 1, pp. 1–12, 2020, doi: 10.1038/s41598-020-68771-z.
9. J. T C, "A Study on Various Machine Learning Classification Algorithms for Diabetes Prediction,"Int. J. Eng. Res. Technol., vol. 10, no. 8, pp. 425–427, 2021, [Online]. Available: https://www.ijert.org/a-study-on-various-machine-learning-classification-algorithms-fordiabetes-prediction.
10. Viloria, Y. Herazo-Beltran, D. Cabrera, and O. B. Pineda, "Diabetes Diagnostic Prediction Using Vector Support Machines,"Procedia Comput. Sci., vol. 170, pp. 376–381, 2020, doi: 10.1016/j.procs.2020.03.065.

11. O. O. Oladimeji, A. Oladimeji, and O. Oladimeji, "Classification models for likelihood prediction of diabetes at early stage using feature selection,"Appl. Comput. Informatics, vol. ahead-of-p, no. ahead-of-print, 2021, doi: 10.1108/aci-01-2021-0022
12. U. M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, and H. H. R. Sherazi, "Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications,"J. Healthc. Eng., vol. 2021, 2021, doi: 10.1155/2021/9930985.
13. G. D. Kalyankar, S. R. Poojara, and N. V. Dharwadkar, "Predictive analysis of diabetic patient data using machine learning and Hadoop,"Proc. Int. Conf. IoT Soc. Mobile, Anal. Cloud, I-SMAC 2017, no. Dm, pp. 619–624, 2017, doi: 10.1109/I-SMAC.2017.8058253.
14. N. Sneha and T. Gangil, "Analysis of diabetes mellitus for early prediction using optimal features selection,"J. Big Data, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0175-6.
15. M. F. Faruque, Asaduzzaman, and I. H. Sarker, "Performance Analysis of Machine Learning Techniques to Predict Diabetes Mellitus,"2nd Int. Conf. Electr. Comput. Commun. Eng. ECCE 2019, pp. 1–4, 2019, doi: 10.1109/ECACE.2019.8679365.
16. M. K. Hasan, M. A. Alam, D. Das, E. Hossain, and M. Hasan, "Diabetes prediction using ensembling of different machine learning classifiers,"IEEE Access, vol. 8, pp. 76516–76531, 2020, doi: 10.1109/ACCESS.2020.2989857.
17. Mujumdar and V. Vaidehi, "Diabetes Prediction using Machine Learning Algorithms,"Procedia Comput. Sci., vol. 165, pp. 292–299, 2019, doi: 10.1016/j.procs.2020.01.047.
18. M. Rout and A. Kaur, "Prediction of Diabetes Risk based on Machine Learning Techniques,"Proc. Int. Conf. Intell. Eng. Manag. ICIEM 2020, pp. 246–251, 2020, doi: 10.1109/ICIEM48762.2020.9160276.
19. W. Chen, S. Chen, H. Zhang, and T. Wu, "A hybrid prediction model for type 2 diabetes using K-means and decision tree,"Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS, vol. 2017-Novem, no. 61272399, pp. 386–390, 2018, doi: 10.1109/ICSESS.2017.8342938.
20. Yahyaoui, A. Jamil, J. Rasheed, and M. Yesiltepe, "A Decision Support System for Diabetes Prediction Using Machine Learning and Deep Learning Techniques,"1st Int. Informatics Softw. Eng. Conf. Innov. Technol. Digit. Transform. IISEC 2019 - Proc., no. 2, pp. 1–4, 2019, doi: 10.1109/UBMYK48245.2019.8965556.
21. K. Bhatia, S. Arora, and R. Tomar, "Diagnosis of diabetic retinopathy using machine learning classification algorithm,"Proc. 2016 2nd Int. Conf. Next Gener. Comput. Technol. NGCT 2016, vol. 3, no. 2, pp. 347–351, 2017, doi: 10.1109/NGCT.2016.7877439.