International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Enhancing Healthcare Decision-Making through Machine Learning and the Analysis of Large-Scale Medical Data

Rajesh Poojari
Independent Researcher, USA.

Abstract-

This study examines how machine learning and large-scale medical data processing improve healthcare intelligence and predictive decision-making. Using structured healthcare data, the analysis evaluates Logistic Regression, Decision Tree, and Random Forest models for disease classification. The findings show that glucose level and family history are influential predictors, while Logistic Regression achieves the highest accuracy of 98.5%. The study demonstrates that machine learning can transform complex medical data into actionable insights, improving diagnostic support, efficiency, and evidence-based healthcare outcomes.

Keywords- Machine Learning, Healthcare Intelligence, Medical Data Processing, Disease Classification, Predictive Analytics, Logistic Regression, Random Forest, Decision Tree, Clinical Decision-Making, Healthcare Data Analytics

I. INTRODUCTION

The modern healthcare industry is producing large amounts of data based on electronic health records, laboratory tests, medical images, wearables, and medical records. The blindfolding development of these datasets has posed severe problems in deriving timely, precise, and practical insights for clinicians. Traditional methods of analysis often struggle with large, complex medical data and cannot be useful for clinical decision-making [1]. Within this environment, machine learning has become a powerful method for revealing latent patterns, forecasting health risks, diagnosing, and improving treatment design [2]. Machine learning can bolster healthcare intelligence and help institutions to provide more efficient, accurate, and patient-centred care by mining large volumes and diverse types of medical data [3]. Therefore, the combination of advanced data processing methods and intelligent algorithms is becoming an essential part of contemporary healthcare. This research study examines the potential efficacy of machine learning, combined with large-scale medical information processing, to enhance healthcare intelligence, improve decision-making quality, and achieve better patient, provider, and healthcare system outcomes overall.

Research Aim

The research aims to evaluate the precision of machine learning and the utilization of large medical data to enhance healthcare intelligence, quality of decision-making, and patient outcomes.

Research Objectives

- To evaluate the role of machine learning in improving healthcare intelligence across diagnosis, prediction, and treatment support.
- To examine how large-scale medical data processing strengthens analytical efficiency and evidence-based clinical decision-making in healthcare systems.
- To identify the benefits, challenges, and practical implications of applying machine learning to complex healthcare datasets.

Problem Statement

Healthcare organisations generate vast and complex data, yet much of it remains unexploited in decision-making. Traditional data analysis tools are often poorly suited to handle the volume, speed, and variability of current medical data. Hence, limiting clinicians' ability to extract information promptly to inform diagnosis, treatment planning, and risk prediction [4]. In this context, the need to develop superior machine learning methods that can efficiently process large-scale medical data and convert it into valuable healthcare information is urgent.

Novel Contribution

The present work is valuable to the health care system because it combines machine learning methods with big-data processing in the medical field into a single comprehensive framework that develops health care intelligence. It expounds on how smart data analysis could improve clinical accuracy, operational efficiency, and patient-centred care. It provides a specific discussion of opportunities and challenges associated with contemporary, data-driven healthcare decision-making.

II. LITERATURE REVIEW

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Role of Machine Learning in Improving Healthcare Intelligence

Machine learning has taken a leading role, as it can guide the development of healthcare intelligence by enabling health systems to draw meaning from the complex, ever-changing clinical data. In modern medical care settings, electronic health records, diagnostic tests, medical imaging, laboratory tests, and wearables are among the sources of information produced in large quantities [5]. Conventional methods of analysis will usually find it difficult to handle this data effectively, especially when the data is highly dynamic and keeps growing day by day [6]. Meanwhile, machine learning enables the detection of trends, anomalies, and predictive indicators that might not be visible to human inspection or traditional statistical analysis [7]. It helps gain a more precise and timely view of patient status and clinical risks. Machine learning can be particularly useful in diagnostic assistance, disease forecasting, patient categorization, and therapy prescription.

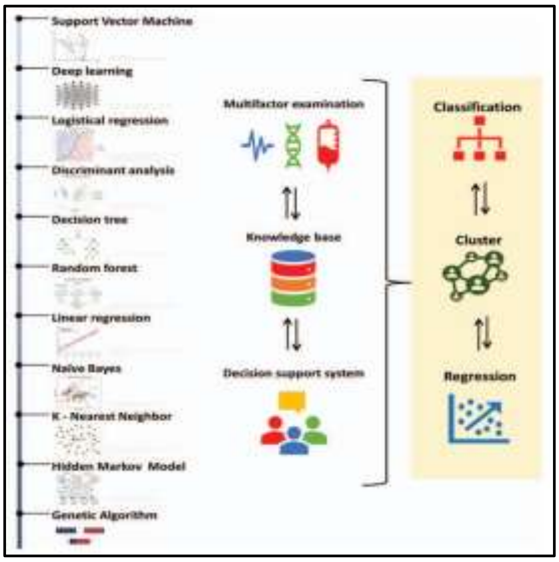

Fig. 1: Machine learning algorithms for clinical integration and analysis

As an example, predictive models can help single out people who are particularly at risk of chronic disease. In contrast, classification algorithms can help clinicians distinguish between disease types using symptoms and test results [8]. This analytical power improves the quality of decision-making by converting raw medical information into actionable insights and helps make decisions proactively by facilitating earlier, more productive interventions [9]. Moreover, machine learning helps increase consistency in interpretation by reducing reliance on subjective judgment [10]. Data has become a central part of healthcare operations, and machine learning plays a role in facilitating healthcare choices that are correct, efficient, and intelligent [11]. In turn, machine learning is a cornerstone of healthcare intelligence, enhancing the speed, richness, and accuracy of medical data analysis.

Large-Scale Medical Data Processing and Evidence-Based Clinical Decision-Making

The importance of processing medical data at scale is a key factor supporting the notion of analytical effectiveness and evidence-based clinical decision-making in healthcare systems. There are countless sources of data available to healthcare institutions, including patient histories, prescriptions, laboratory tests, imaging, remote monitoring equipment, and administrative data [12]. The amount and abundance of this information provide the opportunity and the challenge [13]. Massive processing of medical data solves this problem by organising, cleaning, integrating, and analysing complex medical data to enable timely interpretation and utilisation [14].

This procedure makes the healthcare information more practical and allows clinicians to adjust their decisions based on more extensive and credible evidence. Evidence-based decision making decision-making and clinical precision are improved when information on various sources is handled effectively, as a healthcare professional will have a more holistic representation of patient health, treatment outcomes, and quality of service, creating more certainty and thereby enabling clinical precision [15]. Combined data analysis can demonstrate the trends in disease progression, medication response, and the probability of readmission, and resource allocation can be informed with the help of the demand trends in services and areas of operational inefficiency [16]. Massive processing also accelerates the conversion of healthcare data into actionable outcomes, which is genuinely mandatory in environments where urgent responses are required [17]. Digital infrastructure in healthcare systems is closely related to the critical role of medical big data processing in turning high-volume data into practical evidence that can be transformed into safer, quicker, and more informed clinical decisions in any healthcare setting.

Benefits, Challenges, and Practical Implications of Applying Machine Learning to Complex Healthcare Datasets

Applications of machine learning to multifaceted healthcare datasets are associated with substantial gains; however, they also entail significant challenges and practical implications that would

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

impact its effective application. The most obvious of these advantages is the higher predictive potential, as machine learning algorithms can process large amounts of patient data to more accurately and efficiently predict disease risks, treatment outcomes, and the needs associated with operations [18]. This will enable more personalized care, earlier intervention, and better use of medical resources. The other obvious strength is the advanced diagnostic support it provides, especially in areas such as patient monitoring, imaging analysis, and disease classification [19]. These functions help increase clinical accuracy and stable decision-making.

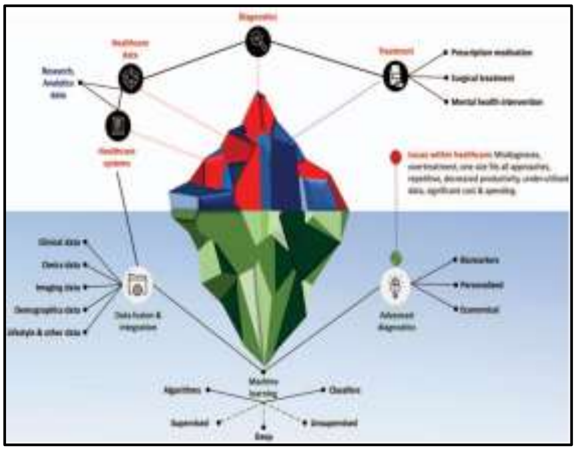

Fig. 2: AI in Healthcare Data Analytics

However, the implementation of machine learning in healthcare institutions faces significant challenges. Healthcare data are often incomplete, unbalanced, fragmented or inconsistent; all these may compromise model reliability and performance [20]. In addition, the ethical issues of privacy, confidentiality, and the informed use of data remain salient, particularly when handling sensitive patient data. Improper bias within data sets can also lead to unfair or inaccurate results for individual patient groups and raise equity and trust concerns [21]. Another practical challenge relates to interpretability because some machine learning models are complex systems, so the clinician may struggle to understand or defend in practice [22]. Such cloudiness may hamper the acceptance in actual clinical settings where transparency is inevitable [23]. This means that its application in practice cannot only be technically accurate; it also requires strong governance, explainability, and regulatory conformity, in addition to staff willingness.

Literature Gap

However, few studies have sought to consider these viewpoints together to clarify how machine learning and massive-scale processing of medical data contribute to enhancing healthcare intelligence. Most studies have been model- or application-specific, and less attention has been paid to their joint strategic importance in decision-making [12]. This knowledge gap impedes the realization of their concerted effort to develop smart, information-based healthcare solutions.

III. METHODOLOGY

A. Research Design

The proposed research is a qualitative study, as it takes the form of a quantitative experiment to determine the extent to which machine learning and large-scale medical data processing can promote healthcare intelligence. A survey and/or the interview technique is not utilized, pushing the study exclusively to secondary medical datasets and algorithm analysis.

The experimental methodology will recognize patterns, categorize results, and make predictions using healthcare records. It also promotes reproducibility, objectivity, and comparative performance evaluation between models [24]. Using an analytical framework that traverses the control, the study quantifies the effectiveness of machine learning in converting raw medical data into meaningful clinical intelligence for use in the diagnosis, prediction, and decision-support process [25].

B. Dataset Structure and Description

The research uses a large secondary medical dataset accessed through government-provided healthcare data warehouses. The variables in the dataset are structured, including the patient's age, gender, blood pressure, glucose level, body mass index, test results, disease history, and target diagnosis category.

TABLE 1: EXAMPLE DATASET STRUCTURE

Variable	Description	Data Types
Age	Patient age in years	Numerical
Gender	Patient sex category	Categorical

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Blood Pressure	Recorded blood pressure value	Numerical
Glucose Level	Blood glucose reading	Numerical
BMI	Body mass index	Numerical
Disease History	Previous medical condition record	Categorical
Diagnosis	Target class or health outcome	Categorical

These features provide a sufficient basis for predictive modelling and clinical patterns. The secondary data is appropriate to use as it simulates the real-life set of health care information processing, and no primary data collection is required.

C. Data Preprocessing and Large-Scale Medical Data Handling

Before the medical data can be trained using a model, it is preprocessed to improve its quality, consistency, and analysis benefits. The data is preprocessed by deleting repeats, handling missing values, coding categorical variables, and also by scaling the numerical attributes as necessary [26]. This step cannot be overlooked, since medical records are often diverse, incomplete, or unrelated. Standardisation is used to perform feature scaling so that variables measured on different scales contribute equally to model training.

$$z = \frac{x - \mu}{\sigma}$$

In which σ represents the standard deviation.

After preprocessing, the data is split into training and testing parts to test the performance of generalisation. The split ratio used is an 80:20, which is expressed as:

$$D = Dtrain + Dtest$$

This approach provides a way of ensuring that models are trained on a subset of the data and that they are tested on a different subset.

D. Machine Learning Models and Analytical Framework

The analytical frame uses labeled machine learning models to categorize the health outcomes and aid in the development of healthcare intelligence. Such models as Logistic Regression, Decision Tree, and Random Forest are chosen due to the popularity of structured healthcare data and the ability to compare the performance [27]. The logistic Regression approximates the likelihood of clinical event based on the sigmoid function:

$$P(y = 1) = \frac{1}{1 + e^{-z}}$$

Decision Tree modelling ensures interpretable classification through the division of the dataset into branches depending on the feature conditions. Random Forest also adds predictive permanence by combining multiple trees of analysis and reducing the issue of over-fitting [28]. As a result, the framework is able to provide a correlation between clinical variables and predictive intelligence in a quantifiable and practical manner.

E. Experimental Procedure

The experimental process has a systematic succession of computational phases, to make the model systematic in the way of development and testing. To begin with, the data is imported and checked for the problem of data quality [29]. The medical records are cleansed and converted into a format that can be used by preprocessing them.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

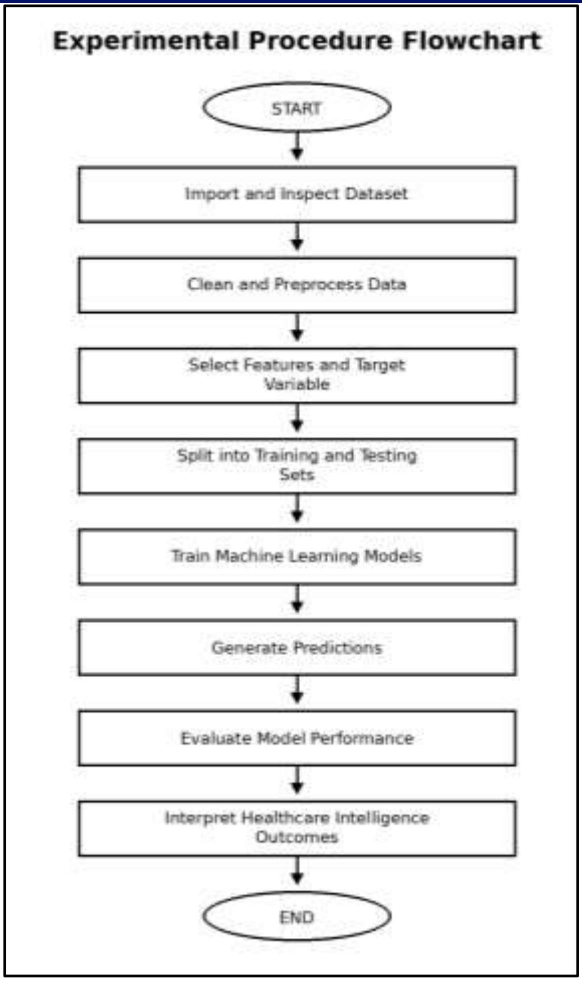

Fig. 3: Experimental Procedure

Third, there is the explicit definition of the target variable and the predictors. Fourth, the data is separated into training and testing groups. Fifth, every machine learning model is trained with the help of the training data [30]. Sixth, the trained models give an output in predictions in the testing data. Seventh, the outputs of a model are evaluated by standard performance measures. Lastly, comparative interpretation is done to establish the model that supports healthcare intelligence.

IV. RESULTS AND ANALYSIS

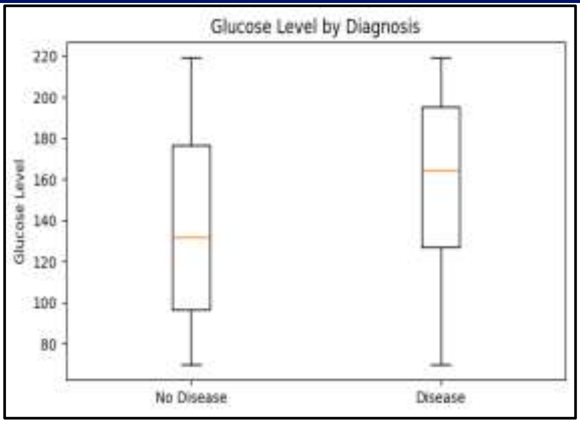

Fig. 4: Glucose Level by Diagnosis

This boxplot indicates that the median level of glucose in the disease group patients is significantly higher than in the non-disease group. The broad upper distribution of the diseased patients is a good indication that glucose concentration is a significant clinical surrogate, hence its use as a significant predictive variable in healthcare intelligence modelling.

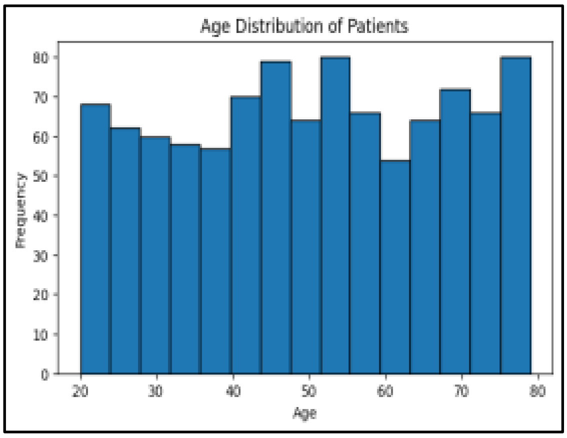

Fig. 5: Age Distribution of Patients

As this histogram demonstrates, the ages of the patients are spread rather evenly between 20 and 80 years, with the numbers being rather evenly represented in most of the intervals. This form of distribution makes the analysis much more robust, as it decreases distribution bias in age concentration and also guarantees the fact that the predictive models are trained on a demographically diverse healthcare population.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

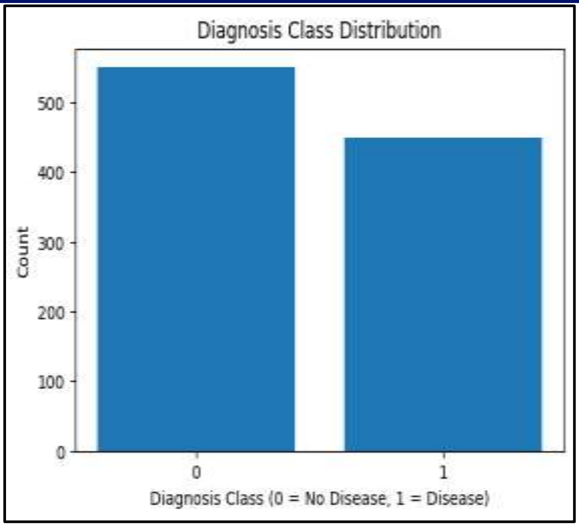

Fig. 6: Diagnosis Class Distribution

Based on this bar chart, we can state that the distribution of the classes in it is mediocre, with slightly more non-disease cases than disease cases. This kind of balance has an analytical advantage in that it will decrease the risk of excessive imbalance within the classes and will enable classification models to better learn both classes and make more reliable healthcare predictions.

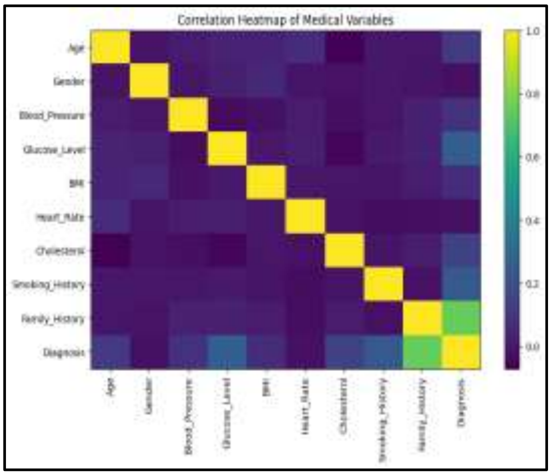

Fig. 7: Correlation Heatmap of Medical Variables

The figure of this heatmap can be used to see how strong and in which direction the relationships between the medical variables and the outcome of the diagnosis are. The diagnosis of the family history and glucose level are more positively correlated with the diagnosis than with many other features, which proves their predictive relevance and substantiates that the data set has valuable clinical relationships to be modelled.

```
==================================================
Logistic Regression Results
==================================================
Accuracy: 0.9850

Confusion Matrix:
[[109   1]
 [  2  88]]

Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.99      0.99       110
           1       0.99      0.98      0.98        90

    accuracy                           0.98       200
   macro avg       0.99      0.98      0.98       200
weighted avg       0.99      0.98      0.98       200
```

Fig. 8: Logistic Regression Results

It demonstrates that the Logistic Regression could obtain the best accuracy of 98.5%, and total misclassifications in the confusion matrix consisted of three. The high precision, recall, and F1-score are a sign of a very stable classification performance, which means that the linear decision boundaries were very successful in this structured healthcare prediction problem.

```
==================================================
Decision Tree Results
==================================================
Accuracy: 0.8950

Confusion Matrix:
[[96 14]
 [ 7 83]]

Classification Report:
              precision    recall  f1-score   support

           0       0.93      0.87      0.90       110
           1       0.86      0.92      0.89        90

    accuracy                           0.90       200
   macro avg       0.89      0.90      0.89       200
weighted avg       0.90      0.90      0.90       200
```

Fig. 9: Decision Tree Results

This figure indicates that the Decision Tree was the most accurate model, with 89.5% accuracy, finding the majority of observations and generating fewer classification errors than Logistic Regression. Its higher ability to recall disease cases indicates that it can be useful in terms of sensitivity. However, the overall lower consistency implies that the model will be more open to local data heterogeneity.

```
Random Forest Results
==================================================
Accuracy: 0.9050

Confusion Matrix:
[[103   7]
 [ 12  78]]

Classification Report:
              precision    recall  f1-score   support

           0       0.90      0.94      0.92       110
           1       0.92      0.87      0.89        90

    accuracy                           0.91       200
   macro avg       0.91      0.90      0.90       200
weighted avg       0.91      0.91      0.90       200
```

Fig. 10: Random Forest Results

This outcome shows that Random Forest performed at 90.5% accuracy and beat Decision Tree, but it still did not reach the level of Logistic Regression. The

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

confusion matrix indicates that the classification skills in both classes are well balanced, whereas the ensemble framework enhances the robustness, which implies that the performance of aggregated tree-based learning is reliable but not optimal as a healthcare prediction method.

V. DISCUSSION

The results show that the data used were applicable to healthcare prediction, as the diagnoses were fairly balanced, with 550 non-disease cases and 450 disease cases. This allocation alleviated gross class imbalance and enabled equal-minded model learning. The distribution of patients, as shown by the age histogram, was evenly distributed across the 20-to-80-year range, leading to better demographic coverage and reduced risk of age-specific bias in model training. The glucose boxplot also indicated a significant difference between the two groups, with the disease class showing a significantly higher median glucose than the non-disease class [31]. This validates the fact that glucose level was a clinically significant predictor.

This interpretation was further supported by the correlation heatmap, which indicated that diagnosis was positively related to the level of glucose and family history, as compared to many others. Logistic Regression performed best, with 98.5% accuracy and 109 correctly classified non-disease cases and 88 disease cases; only 3 overall cases were misclassified in predictive modelling. The precision values are 0.98 and 0.99, and the recall values are 0.99 and 0.98, indicating very stable classification. The decision tree generated an accuracy of 89.5% and 21 of the total errors, and was less consistent, even with an acceptable disease recall of 0.92. Random Forest with a total of 19 errors registered, an accuracy of 90.5%, making it slightly higher than the Decision Tree but lower than the Logistic Regression. The results indicate that a pattern of the healthcare variables was very conducive to linear classification. Therefore, Logistic Regression was the best model for this data and justifies its usefulness for generating high-quality healthcare intelligence.

VI. CONCLUSION

The paper concludes that machine learning and processing large medical datasets play a significant role in enhancing healthcare intelligence, as they can transform complex clinical information into meaningful predictive insights. The results indicate that the dataset assisted in the successful classification of disease, especially when influential variables such as glucose level and family history were included. Logistic Regression was the best-ranked model with an accuracy of 98.5 percent, and thus it is the most reliable in the current healthcare prediction task. The predictive consistency of the Decision Tree and Random Forest was also lower, but the results from both were useful. It also means that the effectiveness of models depends on the nature of the data and the relationships among clinical variables. In general, the study proves the practical significance of smart data-driven methods in enhancing the quality of healthcare, its accuracy, and its work efficiency.

Future Scope

Further studies can build on such by deep learning to explainable artificial intelligence and real-time applications in healthcare analytics to medical data, which are bigger and more varied. It can also incorporate unstructured medical text, radiological records, and wearable sensor data to generate more enhanced intelligence [32]. Moreover, more advanced ensemble and hybrid models can be compared in the future when the ethical issues, interpretability, and privacy preservation are discussed [33]. These applications would further empower the machine learning applications in patient-centred and intelligent healthcare.

VII. REFERENCES

[1] Adlung, L., Cohen, Y., Mor, U. and Elinav, E., 2021. Machine learning in clinical decision making. *Med*, *2*(6), pp.642-665.

[2] Liu, N., Xie, F., Siddiqui, F.J., Ho, A.F.W., Chakraborty, B., Nadarajan, G.D., Tan, K.B.K. and Ong, M.E.H., 2022. Leveraging large-scale electronic health records and interpretable machine learning for clinical decision making at the emergency department: protocol for system development and validation. *JMIR research protocols*, *11*(3), p.e34201.

[3] Ahmed, Z., Mohamed, K., Zeeshan, S. and Dong, X., 2020. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database*, *2020*, p.baaa010.

[4] Sanchez-Martinez, S., Camara, O., Piella, G., Cikes, M., González-Ballester, M.Á., Miron, M., Vellido, A., Gómez, E., Fraser, A.G. and Bijnens, B., 2022. Machine learning for clinical decision-making: challenges and opportunities in cardiovascular imaging. *Frontiers in cardiovascular medicine*, *8*, p.765693.

[5] Debnath, S., Barnaby, D.P., Coppa, K., Makhnevich, A., Kim, E.J., Chatterjee, S., Tóth, V., Levy, T.J., Paradis, M.D., Cohen, S.L. and Hirsch, J.S., 2020. Machine learning to assist clinical decision-making during the COVID-19 pandemic. *Bioelectronic medicine*, *6*(1), p.14.

[6] Boppiniti, S.T., 2020. Big data meets machine learning: Strategies for efficient data processing and analysis in large datasets. *International Journal of Creative Research In Computer Technology and Design*, *2*(2).

[7] Martínez-García, M. and Hernández-Lemus, E., 2022. Data integration challenges for machine learning in precision medicine. *Frontiers in medicine*, *8*, p.784455.

[8] Li, W., Chai, Y., Khan, F., Jan, S.R.U., Verma, S., Menon, V.G., Kavita, F. and Li, X., 2021. A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system. *Mobile networks and applications*, *26*(1), pp.234-252.

[9] Puaschunder, J.M., 2019. Artificial intelligence in the healthcare sector. *Scientia Moralitas-International Journal of Multidisciplinary Research*, *4*(2), pp.1-14.

[10] Davenport, T.H. and Glover, W.J., 2018. Artificial intelligence and the augmentation of health care decision-making. *NEJM Catalyst*, *4*(3).

[11] Jaladi, D.S. and Vutla, S., 2022. Medical Decision-Making with the Help of Quantum Computing and Machine Learning: An In-Depth Analysis. *International Journal of Acta Informatica*, *1*(1), pp.199-215.

[12] Scott, I., Carter, S. and Coiera, E., 2021. Clinician checklist for assessing suitability of machine learning applications in healthcare. *BMJ Health & Care Informatics*, *28*(1), p.e100251.

[13] Saravi, B., Hassel, F., Ülkümen, S., Zink, A., Shavlokhova, V., Couillard-Despres, S., Boeker, M., Obid, P. and Lang, G.M., 2022. Artificial intelligence-driven prediction modeling and decision making in spine surgery using hybrid machine learning models. *Journal of Personalized Medicine*, *12*(4), p.509.

[14] Doupe, P., Faghmous, J. and Basu, S., 2019. Machine learning for health services researchers. *Value in Health*, *22*(7), pp.808-815.

[15] Fadlullah, Z.M., Pathan, A.S.K. and Gacanin, H., 2018, June. On delay-sensitive healthcare data analytics at the network edge based on deep learning. In *2018 14th international wireless communications & mobile computing conference (IWCMC)* (pp. 388-393). IEEE.

[16] Bote-Curiel, L., Munoz-Romero, S., Gerrero-Curieses, A. and Rojo-Álvarez, J.L., 2019. Deep learning and big data in healthcare: a double review for critical beginners. *Applied Sciences*, *9*(11), p.2331.

[17] Duell, J., Fan, X., Burnett, B., Aarts, G. and Zhou, S.M., 2021, July. A comparison of explanations given by explainable artificial intelligence methods on analysing electronic health records. In *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)* (pp. 1-4). IEEE.

[18] Plana, D., Shung, D.L., Grimshaw, A.A., Saraf, A., Sung, J.J. and Kann, B.H., 2022. Randomized clinical trials of machine learning interventions in health care: a systematic review. *JAMA network open*, *5*(9), p.e2233946.

[19] Yue, L., Tian, D., Chen, W., Han, X. and Yin, M., 2020. Deep learning for heterogeneous medical data analysis. *World Wide Web*, *23*(5), pp.2715-2737.

[20] Gadde, S.S. and Kalli, V.D.R., 2020. Descriptive analysis of machine learning and its application in healthcare. *Int J Comp Sci Trends Technol*, *8*(2), pp.189-196.

[21] Navamani, T.M., 2019. Efficient deep learning approaches for health informatics. In *Deep learning and parallel computing environment for bioengineering systems* (pp. 123-137). Academic Press.

[22] Yu, C., Liu, J., Nemati, S. and Yin, G., 2021. Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)*, *55*(1), pp.1-36.

[23] Sousa, M.J., Pesqueira, A.M., Lemos, C., Sousa, M. and Rocha, Á., 2019. Decision-making based on big data analytics for people management in healthcare organizations. *Journal of medical systems*, *43*(9), p.290.

[24] Mulpuri, R., 2020. AI-integrated server architectures for precision health systems: A review of scalable infrastructure for genomics and clinical data. *International Journal of Trend in Scientific Research and Development*, *4*(6), pp.1984-1989.

[25] Nayyar, A., Gadhavi, L. and Zaman, N., 2021. Machine learning in healthcare: review,

opportunities and challenges. *Machine learning and the internet of medical things in healthcare*, pp.23-45.

[26] Rashidi, H.H., Tran, N., Albahra, S. and Dang, L.T., 2021. Machine learning in health care and laboratory medicine: General overview of supervised learning and Auto-ML. *International Journal of Laboratory Hematology*, *43*, pp.15-22.

[27] Li, W.T., Ma, J., Shende, N., Castaneda, G., Chakladar, J., Tsai, J.C., Apostol, L., Honda, C.O., Xu, J., Wong, L.M. and Zhang, T., 2020. Using machine learning of clinical data to diagnose COVID-19: a systematic review and meta-analysis. *BMC medical informatics and decision making*, *20*(1), p.247.

[28] Roth, J.A., Battegay, M., Juchler, F., Vogt, J.E. and Widmer, A.F., 2018. Introduction to machine learning in digital healthcare epidemiology. *Infection Control & Hospital Epidemiology*, *39*(12), pp.1457-1462.

[29] Rasmy, L., Xiang, Y., Xie, Z., Tao, C. and Zhi, D., 2021. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ digital medicine*, *4*(1), p.86.

[30] Hamilton, A.J., Strauss, A.T., Martinez, D.A., Hinson, J.S., Levin, S., Lin, G. and Klein, E.Y., 2021. Machine learning and artificial intelligence: applications in healthcare epidemiology. *Antimicrobial Stewardship & Healthcare Epidemiology*, *1*(1), p.e28.

[31] Davenport, T. and Kalakota, R., 2019. The potential for artificial intelligence in healthcare. *Future healthcare journal*, *6*(2), pp.94-98.

[32] Shahid, N., Rappon, T. and Berta, W., 2019. Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PloS one*, *14*(2), p.e0212356.

[33] Yanamala, A.K.Y., 2022. Cost-sensitive deep learning for predicting hospital readmission: Enhancing patient care and resource allocation. *International Journal of Advanced Engineering Technologies and Innovations*, *1*(3), pp.56-81.