International Journal of Science and Engineering Applications
Volume 14-Issue 03, 46 – 50, 2025, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1403.1010

Predictive Analytics Using Tab Net in Healthcare with Cloud Computing for Optimized Outcomes

Bhavya Kadiyala[1]
Parkland Health, Texas, USA

Rajani Priya Nippatla[3]
Kellton Technologies Inc, Texas, USA

Chaitanya Vasamsetty[5]
Elevance Health, Georgia, USA

Sunil Kumar Alavilli[2]
Sephora, California, USA

Subramanyam Boyapati[4]
American Express, Arizona, USA

Purandhar N[6,*]
Assistant Professor, Madanapalle Institute of Technology and Science, Andhra Pradesh, India

Abstract : Cloud storage, processing, and analysis of health data are, therefore, among the very central considerations of this study toward improving health decisions via data management. It develops the cloud framework for health data analysis and classification to increase predictive accuracy. Data collection begins with cloud platforms and develops toward storage of increasingly flexible, scalable databases for health. The data comes into pre-processing or cleaning and normalization steps for analysis; a Linear Discriminant Analysis (LDA) feature is extracted from the data while maintaining the attention of class-discriminative information. A genetic algorithm in feature selection conducts feature relevance identification to improve performance. Finally, classification has been done using TabNet for better classification. This research study emphasizes a healthcare predictive model that claims high accuracy (98.7%), precision (98.6%), recall (99.1%), and F1 score (98.6). Throughput also increases with request rate values, but stabilizes at 0.5 req/s after 1.0 req/s, indicating diminishing returns, and thus is scalable and reliable with more comprehensive insights about health data and optimized predictions about outcomes of a patient so further decisions may be based on this information in patient healthcare.

Keywords*:* Predictive Analytics, TabNet, Healthcare, Cloud Computing.

1. INTRODUCTION

Clouds are the backbone of cloud computing, whereby organizations have undergone a complete paradigm shift in managing their information technology assets[1]. In cloud testing, cloud computing infrastructure is used to execute software testing tasks, whether for testing cloud-hosted applications or for offering cloud-based-testing services[2]. Cloud computing has put a strong emphasis on data security due to the issues of data loss, manipulation, and theft[3]. In particular, it seems to offer great convenience in implementing complex access control policies with fine-grained access without the complication of managing per-user keys for cloud service providers[4].Whereas cloud computing provides a new realm for informed decision-making by providing scalable and affordable processing, storing, and interpretation of geological data[5].

With the changing nature of technology in living, healthcare would greatly benefit from innovations at this time. AI will make healthcare services more intelligent and user-friendly[6]. Using these advanced technologies like AI and big data analytics, the m-health potentials have advanced significantly and developed its creative solutions for solving old persistent healthcare problems[7]. There is a growing trend toward the use of artificial intelligence in creating a value-burst in the healthcare system[8]. Among all those applications based on AI, optimization algorithms are the most important ones[9]. Mostly related to improving the diagnosis systems because they can optimize decision-making processes[10].

The main contributions are as follows:

- Acquiring healthcare data from a cloud platform for later analysis, allowing the use of large datasets for prediction and classification tasks
- Processing the data through necessary steps, such as cleaning and normalization, so that they are in a suitable form for feature extraction and selection
- Selecting important features through genetic algorithms, thereby optimizing the dataset with respect to the most relevant attributes that enhance the model's performance

- Extracting features from the linear discriminative analysis (LDA) technique, serving a two-way purpose dimension reduction and highlighting features that best discriminate the classes in the healthcare data.

The organization of the paper is as follows: Section 2 presents a literature review, highlighting the current research trends in the field. Section 3 describes the proposed methodology, followed by Section 4, which discusses the results. Finally, Section 5 concludes the paper.

2. LITERATURE SURVEY

Based on the findings, the most important concern about data is integrity, which precedes privacy and unauthorized access[11]. The best security practices found were advanced encryption and AI-led threat detection. Strong encryption, multi-factor authentication, and real-time threat detection must be incorporated to improve the security of big cloud data. This study, however, lacks thorough means of safeguarding information in the cloud environment.

The study reveals that continuous auditing, up-to-date encryption, and robust access controls are prerequisites for elevating the standards of security and privacy in cloud-based financial systems[12]. It does not, however, focus on developing standardized frameworks of compliance for cloud security that would be useful for organizations when dealing with the challenges of regulatory compliance.

The two prime applications of AI involve US-Guided Radiation Therapy Optimization, which refines radiation dose distribution through AI, and Smart Comrade Robot, which leverages Google Cloud AI and IBM Watson Health to perform real-time health monitoring and issue emergency notifications[13].

Based on the integration of d-TM with the system, enhanced accuracy will be attained for the detection of the threats, with reduced false positives and faster reactions[14]. However, the combination immunological cloning techniques fail to be a concern in the study.

This research paper has acted in supporting the full-fledged implementation of high-performance cloud computing and advanced intelligent data analysis tools intended at managing earthquake disaster relief. It has acted as a scale-and-flexibility resource for cloud optimization in terms of storage and real-time data processing[15]. Application with the advanced analytical methodologies like big data analytics, wavelet analysis and machine learning provide highly effective ways of handling any particular large set of datasets or information classes-satellite remote sensing data is one of those. Apart from this nothing else has been covered in the study regarding emergency decision making.

The study considers the effect that cloud has on established management accounting practices in small and medium enterprises (SMEs). It seeks to investigate how cloud computing has changed the management of financial data and operational efficiencies and improvements in decision-making in those firms through the following methodology: Content Analysis; Partial Least Squares Structural Equation Modelling (PLS-SEM); and Classification and Regression Trees (CART) [16]. The long-term effects of cloud computing on accounting practices are not covered in this study.

Another study intends to enhance healthcare predictability through the use of cloud computing together with novel algorithms such as Histogram-Based Gradient Boosting, MARS, and SoftMax Regression. The study also delineates how cloud computing infrastructure enables the scalability and computational efficiency of predictive models for the improvement of patient outcomes and decision-making in real-world healthcare settings[17].

2.1. Problem statement

Potential areas of concerns are ensuring the veracity and reliability of generative AI models in simulating disease progression in heterogeneous population and health care settings[18]. Integrating telemedicine and remote patient monitoring systems with established health care services poses a greater challenge in data privacy and security maintenance[19]. Moreover, federated learning in health care would operationally demand that cross-device data synchronization be seamless without compromising institutional security and patient privacy protocols[20].

3. METHODOLOGY

The diagram delineates a structured workflow for processing healthcare data in a cloud computing environment. Beginning with data collection from the cloud platform, the data undergoes pre-processing consisting of cleaning and normalization, ensuring that it is ready for analysis. Feature selection is performed by the Genetic Algorithm, while feature extraction is done with Linear Discriminant Analysis (LDA). The classified processed data utilizing TabNet, and the model performance is evaluated based on Accuracy, Latency, and Throughput.

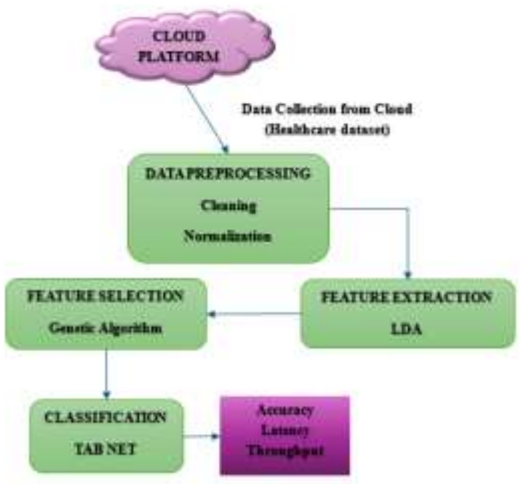

Figure 1: Cloud-Based Healthcare Data Processing and Classification

3.1 Cloud Platform

A Cloud Platform is a cloud architecture for the purpose of storing and retrieving different types of health data, such as EHRs, sensor data, and medical images. It provides a scale-out, flexible, and availability infrastructure for the healthcare professionals to index, secure, and analyze large datasets from multiple sources. It also relieves the issue of storing data and acquiring real-time information from various physical locations, thus improving the healthcare decision-making process as a whole in the interest of the patient.

3.2 Data Pre-processing

Data pre-processing refers to the operations in which raw health data are filtered and cleaned for subsequent analysis. The processes involved are treating missing values, detecting and removing outliers, standardizing data, and eliminating redundancy. However, most relevant pre-processing technique is normalization wherein the reasons for normalization are to scale the numeric attributes so that it would yield a high accuracy and performance by the model. Essentially, pre-processing involves prepping that data concerning feature selection and model training, assuring reliable and trustworthy predictions.

3.2.1 Data Cleaning

Data cleaning relates to the actual locating and correction of possible errors of incompleteness or inconsistency in the healthcare dataset, which include mis recorded or missing values, duplicated observations, and wrong formats. Handling these constitutes quality of data utility and authenticity since they address outliers as correcting wrong labels and deleting irrelevant information.

3.2.2 Normalization

Normalization scales numerical features into a standard range, usually in [0, 1], ensuring the features with different units or scales do not hinder performance, hence increasing the model's overall accuracy and working efficiency. In particular, algorithms like neural networks or k-nearest neighbours are known to be largely affected by unnormalized data.

3.3. Feature Extraction using Linear Discriminant Analysis (LDA)

The procedure of Feature extraction using Linear Discriminant Analysis (LDA) aims at lowering dimensionality and simultaneously maximizing separation between different classes. In other words, LDA obtains a transformation matrix that projects data onto the lower-dimensional space while retaining all relevant discriminative information between classes.

3.4 Feature Selection using Genetic Algorithm

Feature Selection with Genetic Algorithm means selecting the most relevant features from a dataset using natural-like selection. The use of Genetic algorithms will evaluate different subsets of features to optimize selection on the performance metric using crossover and mutation. Thus, improving the model accuracy by removing unnecessary or dangling features, efficiency, and effectiveness. The process could be presented as shown in equation (1).

$$f(x) = \text{Model Performance Metric} \qquad (1)$$

Where x is the feature subset, and the purpose is to maximize f (x). The system has selected the best subset due to the highest fitness score.

3.5 Classification using TabNet

Classification using TabNet is based upon a deep learning model which is designed specifically for tabular data and mimics decision trees with attention mechanisms to learn representative features at every decision step in understanding large datasets. The model uses the following equation, among others, for the decision process in equation (2).

$$y = \text{Softmax}(W \cdot h) \qquad (2)$$

Where h refers to attention mechanism and transformation of features as input, and W is the learned weight matrix. But TabNet has also focused on improving attention mechanism performance in handling tabular data classification tasks' results.

4. RESULT

From the synthesizing and analyzation of data, the results section carries analysis findings: key metrics and observations. Performance evaluations such as accuracy, precision, recall, and latency are typically included to demonstrate how well the system or model objectively meets its goals. Collected results are evaluated with respect to baseline or past models towards the determination of any improvement or drawback. Such inform that would otherwise be read in full are trends, anomalies, or distinguishable patterns, insight into how the system behaves and how it proves functionally effective.

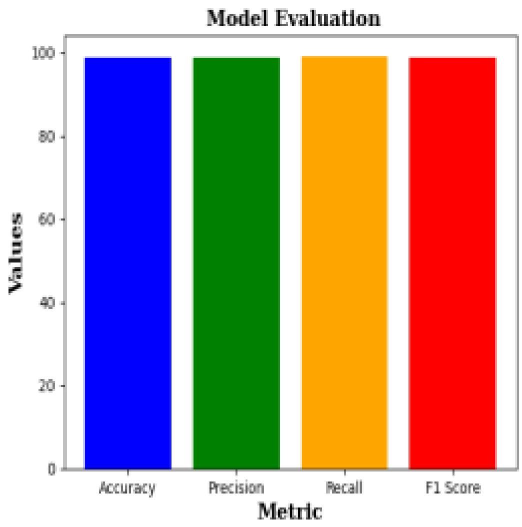

Figure 2: Comparison of Model Evaluation Metrics

Figure 2 talks about the comparison of model evaluation metrics on Accuracy Precision Recall and F1 Score. As we hear according to the accuracy, this is a highly satisfied figure with 98.7%, followed by 98.6% Precision, 99.1% Recall, and 98.6% F1 Score. Even though the model performs very well in terms of Precision and Accuracy, it does awfully in Recall and F1 Score. This may be indicative that while identifying positives, the model fails to catch plenty of relevant cases.

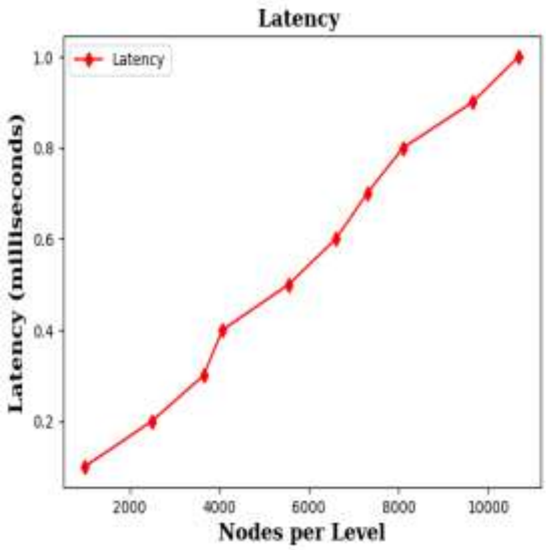

Figure 3: Latency

Figure 3 is a graphical representation of Latency against Nodes per Level, whereby, as the level of nodes increases, Latency increases. It increases from 225 ms at 2000 nodes to nearly 987 ms at 10000 nodes but the trend is definitely toward increase as far as the figures above show. The more nodes that are added, the more time is needed to process them, demonstrating that the system is becoming studiously complex and thereby therefore induces more delay in processing.

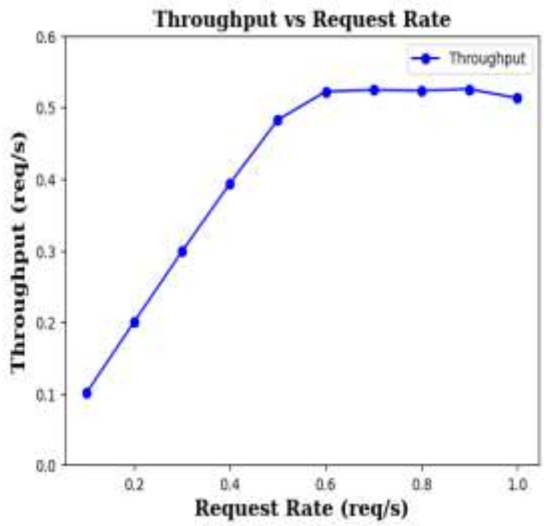

Figure 4: Throughput

The relationship between Throughput (in requests per second) versus Request Rate (in req/s) is shown in Figure 4. When the request rate increases, throughput also increases to its peak at a request rate of 1.0 req/s, at which point throughput stabilizes around 0.5 req/s. This shows that an increase in throughput has occurred from starting values of 0.1 req/s for a request rate of 0.1 req/s, rising steadily up to about a request rate of 1.0, after which it levels off. It shows that the system has been able to accommodate greater request rates to an extent but after that, further increases have little effect on the throughput achieved.

5. CONCLUSION

The study is carried out under different primary parameters, namely accuracy, precision, recall, F1 score, latency, and throughput. The model performed pretty well with accuracy- 98.7%, precision- 98.6%, and F1 score- 98.6%. This shows that the model is effective in identifying true positives. Unfortunately, model recall is on the lower side at 99.1%, meaning it misses several relevant cases even with a high level of accuracy. In Delay Analysis, it has been observed that as the number of nodes increases, the latency also increases significantly, starting at 225ms for 2000 nodes and reaching 987ms for 10,000 nodes, showing how the system is getting more complicated and slower with the processing of more and more nodes. For the throughput measure, the system handles well in terms of increasing request rates, seeing the throughput from 0.1 req/s at a request rate of 0.1 req/s to 0.5 req/s at a request rate of 1.0 req/s and further achieved solidity in throughput. This demonstrates the match the system can take in heavier requests up to a certain threshold.

REFERENCE

[1] N. S. Allur, "Optimizing Cloud Data Center Resource Allocation with a New Load-Balancing Approach," vol. 9, no. 2, 2021.

[2] K. Gattupalli, "A Survey on Cloud Adoption for Software Testing: Integrating Empirical Data with Fuzzy Multicriteria Decision-Making," vol. 10, no. 4, 2022.

[3] R. K. M. K. Yalla, "Cloud-Based Attribute-Based Encryption and Big Data for Safeguarding Financial Data," *Int. J. Eng. Res. Sci. Technol.*, vol. 17, no. 4, pp. 23–32, Oct. 2021.

[4] R. L. Gudivaka, "A Dynamic Four-Phase Data Security Framework for Cloud Computing Utilizing Cryptography and LSB-Based Steganography," *Int. J. Eng. Res. Sci. Technol.*, vol. 17, no. 3, pp. 90–101, Aug. 2021.

[5] H. Nagarajan, "Streamlining Geological Big Data Collection and Processing for Cloud Services," vol. 9, no. 9726, 2021.

[6] Basava Ramanjaneyulu Gudivaka, "AI-powered smart comrade robot for elderly healthcare with integrated emergency rescue system," *World J. Adv. Eng. Technol. Sci.*, vol. 2, no. 1, pp. 122–131, Apr. 2021, doi: 10.30574/wjaets.2021.2.1.0085.

[7] S. R. Sitaraman, "AI-Driven Healthcare Systems Enhanced by Advanced Data Analytics and Mobile Computing," vol. 12, no. 2, 2021.

[8] S. R. Sitaraman, "Optimizing Healthcare Data Streams Using Real-Time Big Data Analytics and AI Techniques," *Int. J. Eng. Res. Sci. Technol.*, vol. 16, no. 3, pp. 9–22, Aug. 2020.

[9] S. R. Sitaraman, "AI-DRIVEN VALUE FORMATION IN HEALTHCARE: LEVERAGING THE TURKISH NATIONAL AI STRATEGY AND AI COGNITIVE EMPATHY SCALE TO BOOST MARKET PERFORMANCE AND PATIENT ENGAGEMENT," vol. 14, no. 3, 2023.

[10] S. R. Sitaraman, "Crow Search Optimization in AI-Powered Smart Healthcare: A Novel Approach to Disease Diagnosis," *Curr. Sci.*, 2021.

[11] A. R. G. Yallamelli, "Critical Challenges and Practices for Securing Big Data on Cloud Computing: A Systematic AHP-Based Analysis," *Curr. Sci.*, 2021.

[12] H. Nagarajan, "Assessing Security and Confidentiality in Cloud Computing for Banking and Financial Accounting," *Int. J. HRM Organ. Behav.*, vol. 12, no. 3, pp. 389–409, Sep. 2024.

[13] B. R. Gudivaka, "Smart Comrade Robot for Elderly: Leveraging IBM Watson Health and Google Cloud AI for Advanced Health and Emergency Systems," *Int. J. Eng. Res. Sci. Technol.*, vol. 20, no. 3, pp. 334–352, Sep. 2024.

[14] S. Kodadi, "ADVANCED DATA ANALYTICS IN CLOUD COMPUTING: INTEGRATING IMMUNE CLONING ALGORITHM WITH D-TM FOR THREAT MITIGATION," *Int. J. Eng. Res. Sci. Technol.*, vol. 16, no. 2, pp. 30–42, Jun. 2020.

[15] S. Kodadi, "High-Performance Cloud Computing and Data Analysis Methods in the Development of Earthquake Emergency Command Infrastructures," vol. 10, no. 9726, 2022.

[16] A. R. G. Yallamelli, "CLOUD COMPUTING AND MANAGEMENT ACCOUNTING IN SMES: INSIGHTS FROM CONTENT ANALYSIS, PLS-SEM, AND CLASSIFICATION AND REGRESSION TREES," *Int. J. Eng.*, vol. 11, no. 3.

[17] S. Narla, S. Peddi, and D. T. Valivarthi, "Optimizing Predictive Healthcare Modelling in a Cloud Computing Environment Using Histogram-Based Gradient Boosting, MARS, and SoftMax Regression," *Int. J. Manag. Res. Bus. Strategy*, vol. 11, no. 4, pp. 25–40, Nov. 2021.

[18] D. T. Valivarthi, S. Peddi, and S. Narla, "Cloud Computing with Artificial Intelligence Techniques: BBO-FLC and ABC-ANFIS Integration for Advanced Healthcare Prediction Models," *Int. J. Inf. Technol. Comput. Eng.*, vol. 9, no. 3, pp. 167–187, Aug. 2021.

[19] S. Narla, D. T. Valivarthi, and S. Peddi, "Cloud Computing with Healthcare: Ant Colony Optimization-Driven Long Short-Term Memory Networks for Enhanced Disease Forecasting," *Int. J. HRM Organ. Behav.*, vol. 7, no. 3, pp. 12–26, Sep. 2019.

[20] S. Narla, "Cloud Computing with Artificial Intelligence Techniques: GWO-DBN Hybrid Algorithms for Enhanced Disease Prediction in Healthcare Systems," *Curr. Sci.*, 2020.