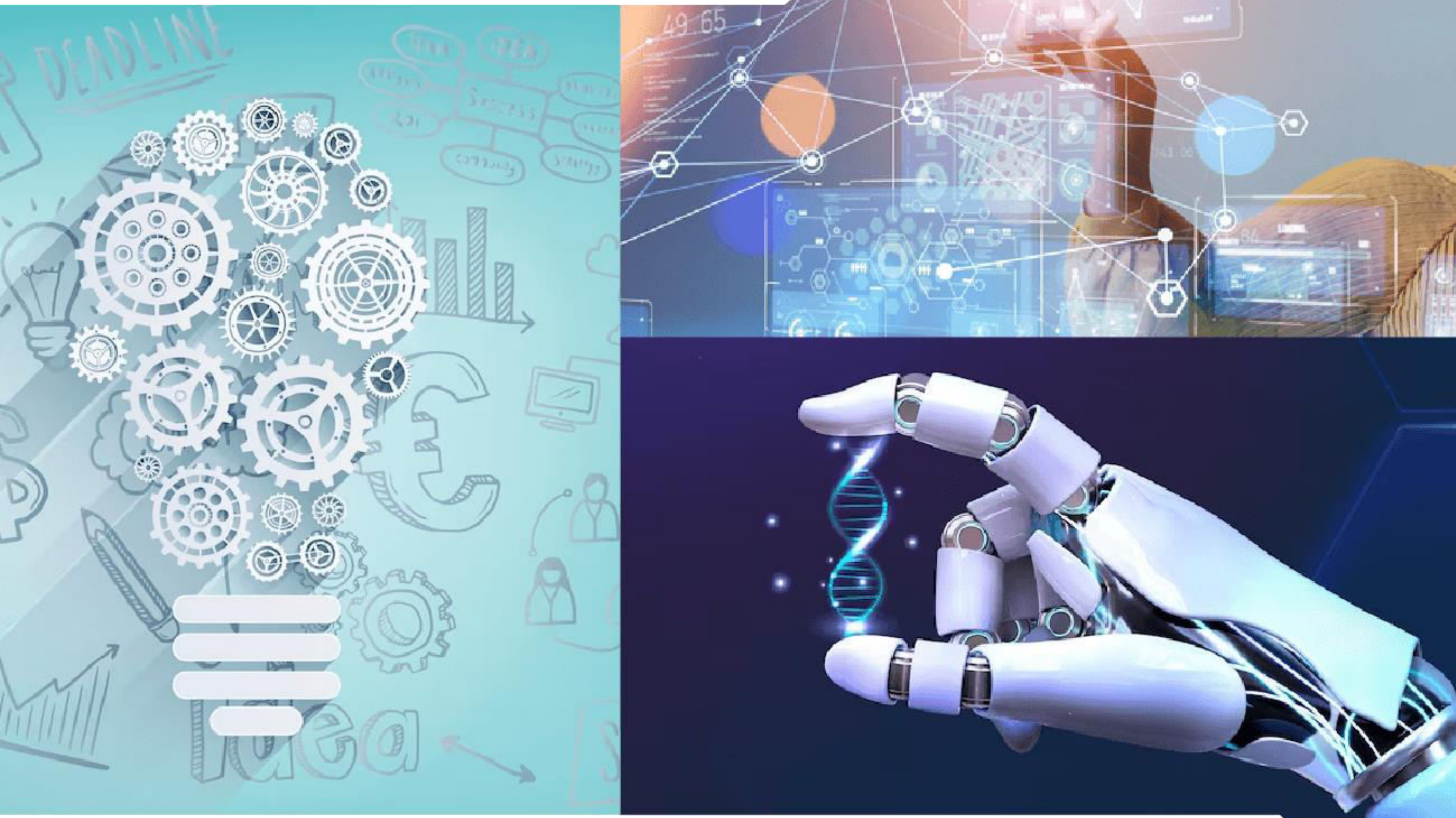
IJARETY



International Journal of Advanced Research in Education and TechnologY *(IJARETY)*

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 10, Issue 4, July-August 2023 ||

Heart Disease Detection using LSTM with Attention Mechanism: A Cloud-Based Classification Approach

Sabine Sane[1]

Ludwig Maximilian University of Munich, Munich, Germany[1]

ABSTRACT: Indeed, heart disease happens to be one of the major health problems of concern at the global level, and there arises the need for the development of effective, cost-efficient and non-invasive diagnostic methods for diagnosing such diseases. This paper discusses a cloud-based heart disease detection system using a hybrid Long Short-Term Memory (LSTM) model in combination with an Attention Mechanism. By utilizing cloud technologies, any unit of data may process real time as well as scalability, which would case facilitate the storage and use of very large medical data. The same model of LSTM captures the time dependencies in the medical history of the patients while the Attention Mechanism focuses on the key features necessary for the accurate diagnosis. Performance evaluation of the system was done based on evaluation metrics such as accuracy, precision, sensitivity, specificity, and F-measure, where the metrics revealed outstanding performance results of 98.9% accuracy, 99.0% precision, and 98.8% sensitivity, among other high scores. These findings show that the proposed system is timely, cheap, and scalable, thereby offering a good solution to heart disease detection and diagnosis early enough.

KEYWORDS: Cloud Computing, Health Care, Long Short-Term Memory (LSTM)

I. INTRODUCTION

The global health concern of heart disease continues to accrue enormous responsibility for a significant share of global mortality. Since healthcare has progressed, several diagnostic tools have been formed, but many of these methods under use are often invasive, costly, and protracted [1]. Hence, the pressing need for more efficient and non-invasive methods has spurred interest in machine learning and deep learning technologies. Given that these technologies, especially those that are cloud-based, can analyze vast medical data in a shorter period and with high accuracy, they hold great potential for early detection and intervention in heart disease management [2].

The factors in the development of heart disease are genes, lifestyle, and environmental influences. Hypertension, irregular cholesterol levels, smoking, poor diet, and inactivity are some of the most publicized risk factors [3]. Diabetes and obesity are in chronic form, and they associate closely with cardiovascular health. Recognizing such factors in people who are at risk for heart disease is significant in timely intervention. Better above-all predictive models for identification of the health factors will improve early detection, allowing intervention through prevention and direct treatment to decline the negative consequence of heart disease [4].

There are still a lot of challenges in heart disease diagnosis besides becoming much better in techniques for diagnosis. Diagnostic methods traditionally are often long and cumbersome which waste a lot of time to arrive at diagnosis [5]. In addition to that, the medical record data are often inconsistent, incomplete, and imbalanced among the different healthcare facilities which makes prediction model development equally challenging. Further complicating this is the fact that most existing machine learning models have poor interpretability, making it difficult for healthcare professionals to accept their predictions as being credible [6]. As a result of these factors, any diagnosis may be too late or faulty and seriously reduces the overall effectiveness of heart disease management.

This research thus proposes a solution that leverages cloud computing and LSTM networks combined with Attention Mechanisms in Long Term Memory Call. The LSTM model is very specific in the nature of the handling of sequential data making it a good fit for retrospectives of patients and checking on how the medical history of patients changes over time [7], [8]. The Attention Mechanisms augment the functionality of the model further by assisting in fetching factors that play an important role in the detection of heart diseases, thus improving both accuracy and interpretability. Providing a solution over cloud storage and computing resources makes the model scalable and breaks the limitations on space and processing of data [9]. The proposed approach is envisioned to provide a resourceful, reliable, and timely method of detecting the disease. Hence it can progress healthcare standards by making health care outcomes better through early and accurate detection.

1.1 PROBLEM STATEMENT

A slow, unreliable, and inefficient method of old has been used for heart disease diagnosis, often involving long waiting periods, irregular medical data, and poor interpretability of findings. These hurdles delay diagnosis and treatment with adverse impacts on the patient's health. With this in mind, the proposed work primarily addresses these issues by building a cloud-based heart disease detection system that uses LSTM in tandem with the attention mechanism. This causes prediction to be faster and more accurate, as it takes into account the patient medical history over time, looks at the most relevant features toward diagnosis, and provides a cloud-based infrastructure able to overcome limitations on data management. The further interpretability of the model provides more confidence to healthcare professionals in its predictions and thus aids in establishing better decisions and earlier intervention. This method is bound to make a huge impact toward timely and accurate detection of heart diseases.

1.2 OBJECTIVES

- **Define** the effectiveness of the proposed cloud-based heart disease detection system.
- **Examine** how the integration of LSTM and Attention Mechanism enhances heart disease detection accuracy.
- **Evaluate** the scalability and efficiency of the proposed system in handling large medical datasets.
- **Analyze** the potential of cloud technologies in improving processing, data management, and decision-making for healthcare applications.
- **Investigate** the interpretability of the model and its impact on healthcare professionals' confidence in the system's predictions.

II. LITERATURE SURVEY

Tunnel engineering is a very dangerous job, one that requires long construction periods and lots of money. TBMs make a reasonable impact on efficiency and safety through the collection of vast amounts of monitoring data [10]. emphasizes a hybrid data mining approach meant to achieve automation of real-time TBM data processing, which makes extensive use of association rule mining, decision tree classification, and neural network modeling. The system can analyze TBM parameters, recognize anomalies, categorize geological formations, and predict rates of penetration (ROP). The implementation of this process in one of the tunnel projects in China helped increase the operational efficiency and safety management of the project through the analysis of TBM data with high reliability and effectiveness [11].

Accurate risk prediction models are needed for managing rheumatoid arthritis (RA) patients, a population predisposed to CVD. examines the feasibility of using long-term blood samples in prediction of cardiovascular risk within RA populations by investigating biomarker stability over a period of 10 to 20 years [12]. Our intention is the assessment of serum quality and biomarker stability, including lipid profiles, inflammatory markers, and traditional risk factors, together with RA-specific markers such as disease activity. We will use longitudinal data analysis to evaluate cardiovascular outcomes together with disease activity along their timeline. Wearables and telemedicine platforms are also being coupled with omics, creating an opportunity for enhancing risk assessment and patient monitoring, ultimately aiding cardiovascular risk prediction and improving the end patient outcome through targeted therapies [13].

Cloud computing has transformed several industries by providing scaled and cost-effective solutions for storing and managing data. Health services face increased security considerations due to the very nature of patient data and stringent regulatory requirements [14], therefore, presents an integrated security management framework to deal with these challenges, which comprises risk assessment, security enforcement, continuous monitoring, and compliance management. The framework provides various means to secure data, such as authentication, encryption, and intrusion detection, with advanced security features based on new technologies like blockchain and multi-factor authentication. Application case studies of major healthcare organizations are demonstrative of the use of cloud computing solutions in the areas of data security and compliance, leaving healthcare organizations to focus more on enhancing patient care and operational efficiency while creating an avenue for the protection of sensitive information. Ensuring the security and privacy of medical data is paramount in healthcare applications. Insights from Ayyadurai [15], emphasizing threat detection via cloud-based analytics, validate the system's use of encrypted storage and secure access protocols. This integration significantly boosts trustworthiness and compliance in handling patient health records.

Djenna, A., et al [16] denotes that the primary requirement for cybersecurity risk management is an organization contending to be relevant in today's fast-paced and ever-evolving world. Cyber threats have, indeed, become so frequent and complex that it is now completely impossible to deal with any risks without the aid of technology. While previous literature and studies have dwelt significantly on risk assessment processes, analysis without the intelligence features or with emerging trends in attacks became irrelevant in contextual data, more so in terms of compliance to cybersecurity safety management. Now it offers the merged Cyber Security Risk Management (m-CSRM)-a decision support system based on fuzzy set mechanism to risk assessment and type of claims predicated on machine learning. In the end, am-Cyber Security Risk Management (m-CSRM) method has been automated and processed through an interesting tool and achieved an impressive 82.13 percent success rate with greater efficiency than currents methods.

Bagaa, M., et al [17] research work is an effort towards resource allocation in intelligent cloud computing systems workload prediction using back-propagation neural network algorithm along with game-theoretic principles. This method is done to augment the services offered by users and service providers availing the services through proper alignment. It applies the concept of Nash equilibrium in developing beneficial SLAs (service level agreement) among the above stakeholders. Experimentation on real datasets can prove the technical efficacy with speed of improving cloud computing operations with strategic alignment in use. Prioritize scalability against security and usability: this dimension has high potential for improving the situations in cloud resource management for many industries.

Kaissis, G., et al [18] seeks to solve collaborative computing systems' issues regarding the dissimilarity of attacks and the main thrust to protect the most important data privacy. It is based on cutting-edge innovations such as federated learning coupled with cloud-edge collaborative computing systems. A multinational validation framework has been built by the study, both against and without attacks, with the application of an "End-to-End Privacy-Preserving Deep Learning for Attack Classification (E2EPPDL)" method in every episode, for classifying attacks into some taxonomy. This central research approach is evaluated using some parameters like Time, Node Count, Routing Count, and Data Delivery Ratio regarding its effectiveness.

Duan, B., et al [19] which deals with the problem of integrating Internet of Things technologies to manufacturing systems have been focused mainly on the area of integrated inventory cost control and job-shop scheduling (JSP). The study proposes a new approach that combines Heterogeneous Genetic Algorithm (HGA) with Hybrid Particle Swarm Optimization (HPSO) to optimize production in the presence of complicated and dynamic constraints. HGA incorporates immune mechanisms to overcome limitations found in conventional Genetic Algorithm in terms of premature convergence, while HPSO blends strengths of PSO with genetic operators to improve job sequencing and minimize production time. Hybridization is more efficient than using either global or local search; thus, it outperforms all conventional approaches. Significant improvements in costs and schedule efficiencies resulted from empirical studies conducted by the approach.

Pelekoudas-Oikonomou, F., et al [20] proposes the advanced model of heart disease monitoring system by using IoMT devices and blockchain technology, addressing the existing system limitations and those with the ECG & PCG data with arrhythmias' consequences. The doctor and patient registration and login include key generation for sharing authentication. The sensed data is uploaded to IPFS and hash codes stored on the blockchain for secure data verification. The model developed with OA-CNN achieves an impressive classification accuracy of 98.32% against the existing systems by using an in-depth heart disease classification system trained on various preprocessing techniques with feature extraction techniques. Results show that the proposed methodology can accurately predict heart disease. The prediction model is trained using historical heart disease data, applying classification algorithms to learn risk patterns. Pulakhandam et al. [21] demonstrated how SNNs, CMA-ES, and HESN models improve scalability and adaptive learning. Adoption of these AI strategies enhances model tuning and improves robustness, ultimately boosting the system’s ability to respond effectively to new patient data.

Badru, L. O., et al [22] denotes that present invention relates to a new smart education management platform that supports intelligent automation and personalized learning through cloud computing and artificial intelligence (AI) technology. It is intended to perform services for such education. It "SOA, an architecture service oriented} platform deployed on a MongoDB implemented Hadoop server cluster," allowing scalable data management processing and resource-heavy processing. It also supports huge data access and high concurrency, thus facilitating easier management of educational resources and quality distance learning. Intelligent applications such as recommendation engines and predictive analytics introduce flexibility, thus creating a user-centered learning environment. Stress tests show that the platform performs very heavy loads but still produces reliable information, thus indicating a possible transformation in the educational service delivery platform.

This growth in cloud computing has made it increasingly difficult to manage vast amounts of data and resources, specifically in cases with highly complex scheduling. Zheng, J., & Wang, Y. [23] presents a combination of Resource Allocation and Task Scheduling using the Improved Bat Optimization Algorithm (IBOA) and Modified Social Group Optimization (MSGO). IBOA adopts dynamic weights and updates the speed using a formula that enhances its random searchability, while the MSGO alters the acquiring phase for better performance. We also ran simulations to compare the performance of the proposed methods against existing strategies; the results indicated that IBOA and MSGO optimize energy consumption immensely, accommodating 32.5 watts for 100 tasks, which closely surpassed the grey wolf optimization method, Multi-Objective Task Scheduling Grey Wolf Optimization.

The emergence of Robotic Process Automation (RPA) has been a game-changer in the field of finance by leading to the automation of repetitive tasks, reduction of errors, and giving the opportunity for real-time reporting. For this work to come to fruition, a comparative assessment of how RPA is integrated into cost accounting and financial systems in terms of quality and scalability will be practiced. It was a holistic procedure and this included the process identification, creation of a workflow, RPA, and finally performance evaluation. The findings indicated that RPA shortened the processing time by 95%, led to the cost allocation accuracy by 99.5%, and finally, it also resulted in a reduction of errors by 95. Qiu, Y. L., & Xiao, G. F. [24] concludes that by integrating RPA into the failed process and enhancing the operation efficiencies and financial controls, cost reductions can be achieved and the optimization of financial processes can be performed efficiently.

AI systems driven by machine learning are the new driving force in the financial industry. IoT-based environments (IoT) are the way by which devices could coordinate with each other: IoT data is analyzed so that fraud patterns can be easily and quickly singled out. In addition to the periodic transaction, the functionality of these AI types includes the dynamic allocation of the types of improvements such as the selection of data mining techniques. Moreover, it is also important to mention that AI makes use of different between supervised and unsupervised learning techniques of transaction data but also, the models described. For a more detailed and reproducible methodology, the datasets include many original attributes; every statement of fraudulent behavior is associated with those attributes. DeZoort, F. T., & Harrison, P. D. [25] we focus on the content that is very necessary for the development of the detection models, which include methodological datasets and evaluation metrics. The study compares and analyzes the different methods, datasets, and assessment metrics needed for fraud detection model development. This study will give for every statement of fraudulent behavior, the attributes that are associated with fraudulent approach and, thus, you can be sure that you know how to respond.

Qiu, M., et al [26] Denotes that P2DS is a new-age protector for the financial data in mobile cloud environment. The ever-widening security concerns encountered by the financial institution are resolved with the help of new cutting-edge techniques such as Attribute-Based Encryption (ABE), Attribute-Based Semantic Access Control, and the Proactive Determinative Access algorithm. The system performs exceptionally well in accurate access control, fast threat detection and response, and high encryption efficiency. With such capabilities, P2DS gives a strong promise of being a reliable solution to safeguard sensitive financial data in the fast-evolving digital landscape. The system collects health data via secure endpoints, processes it, and stores it in the cloud for further analysis. Inspired by Radhakrishnan et al. [27], implementing TLS and Salsa20 encryption in a similar framework, this approach benefits from enhanced transmission security and performance. This strengthens the integrity and confidentiality of sensitive patient data across the entire pipeline.

The purpose of this article is to elaborate on how the confluence of big data analytics, cloud computing, and ABE may be used to improve the security of financial data in the digital age. As with the other industries in banking, attackers become much more complex concerning cyber-attacks, which can only be addressed by innovative approaches in safeguarding sensitive data. Fugkeaw, S. [28] presents an integrated model of ABE and big data analytics as a cloud system while demonstrating that ABE fits into fine-grained access control over encrypted data, scalability, and confidentiality. It describes how big data analytics assist greatly in fraud detection, anomaly detection, risk management, and legal compliance. The interests of case studies applied in these technologies for financial institutions shown to enhance data security and legal standards against increasing numbers of cyber threats against modern banking.

Agrawal, N. [29] proposes a hybrid Edge-Fog-Cloud architecture to improve the latency, scalability, and processing efficiency of IoT systems in smart cities. The dynamic orchestration algorithms integrated in the architecture assign tasks based on their complexity and the available resources so as to optimize performance among the Edge, Fog, and Cloud layers. The method herein proposed works much better than the already established ones like SVM, NSGA-III, and MEC, thereby eliciting vast increases in accuracy, efficiency, and latency. The architecture ensures real-time data analytics and decision-making while reducing processing time and power consumption. The Edge-Fog-Cloud architecture is suitable for large-scale but resource-constrained IoT systems, which means that it can provide even better in terms of scalability and energy efficiency required by future IoT applications.

Hossain, M. A., et al [30] investigates how cloud computing affects the accounting practices in SMEs. This has been done by employing a multi-method approach, which includes Partial Least Squares Structural Equation Modeling, Content Analysis and Classification and Regression Trees. In using the above-mentioned methods, the research explores how cloud computing enhances financial data management, operational efficiency, and making decisions. The results show that cloud-based accounting solutions provide access to real-time data that improves regulatory compliance and strategic decisions taken. Likewise, predictive analytics has taken over traditional methods in management accounting. In spite of various problems, including data security with privacy, employee training, etc., this

| ISSN: 2394-2975 | www.ijarety.in| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 10, Issue 4, July-August 2023 ||

study concludes that the merits of cloud computing would definitely accrue to SMEs through efficiency and strategic decision making. A robust design ensures that all components of the system comply with privacy laws while safeguarding sensitive health records. Drawing from Narla, S.'s [31] work on Continuous Data Protection and Data Obliviousness, the framework integrates secure data handling through encryption and backup. This enhances patient data confidentiality, ensuring compliance with GDPR and HIPAA standards.

Using this approach, Langmead, B., & Nellore, A. [32] delve into the different ways in which big data, hash graph, and cloud computing may be leveraged within the Kinetic methodology for the organization and analysis of large datasets. Cloud computing provides the elasticity required in terms of vast computational resources to process a massive amount of data rapidly and securely. Big data analytics extract valuable insights that improve decision-making through faster computation. Their fast and secure consensus process enables hash graph technology to deliver operational effectiveness and data integrity. The issues would include-the challenges of interoperability, scalability, and regulatory compliance-and show how all these technologies can help improve productivity, improve decision-making, and security of data.

Modern finance data is a mix of structured and unstructured types and calls for advanced methodologies of machine learning for effective analysis. Most traditional models are unable to adapt to high dimensions and complexities, which limit their scalability and accuracy. AlJame, M., et al [33] proposes a combination of Gradient Boosting Decision Trees (GBDT) and ALBERT along with optimization through the Firefly Algorithm all in the framework of a cloud-based solution. This enables real-time processing, scalability, and improved security in delivering a strong and effective solution for the handling of complex financial datasets with performance and accuracy improvements.

Khan, A. W., et al [34] deals with the significant security challenges that software vendor companies face when handling large volumes of data in cloud computing environments. Data management has revolutionized big data integrated with cloud computing those scales and deserves cost efficiency. It is high time that such cloud service providers now deal with major security issues like data integrity, unauthorized access, and data privacy. The research examines the security problems using an analytic hierarchy process, which enables identification, ranking and evaluation. Data privacy and unauthorized access have been determined as the main issues of priority. Advanced encryption, AI threat detection, and multifactorial authentication are the recommended measures for securing cloud data. It is also advised that the state-of-the-art technologies like artificial intelligence and quantum encryption should be looked upon for the additional data security level in the cloud environments.

III. METHODOLOGIES

Figure 1 consitutes the cloud-based detection system for heart diseases using the LSTM model combined with the Attention Mechanism. The whole system begins by importing data collected from the Clinic Heart Disease dataset and storing it in cloud storage for easy access. The next step is data preprocessing processes. The classification via the LSTM model with an attention mechanism improves the modeling so that it looks at significant features with improved precision. The system classifies heart disease instances into two categories: Detected and Not Detected performance metrics such as accuracy, precision, recall, and F1-score that are used to evaluate the performance of the developed model. The architecture designed for the heart disease prediction system consists of multiple layers, ensuring efficiency and modular integration of components. Referencing Kodadi's [35] work on cloud-based seismic systems, the architecture is developed by aligning with proven high-performance models that facilitate data processing and improve scalability, responsiveness to user demands, and resource efficiency attributes that become most pertinent when timely medical predictions need to be delivered.

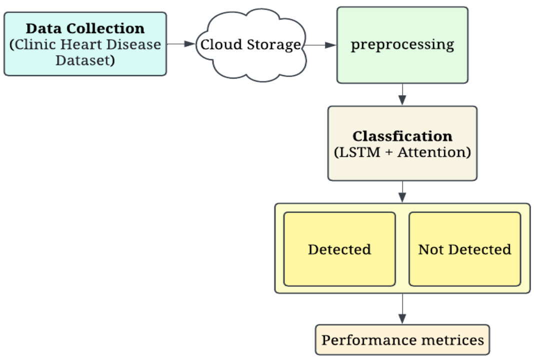


Figure 1: Cloud-Based Heart Disease Detection System

DATA COLLECTION

The Clinic Heart Disease Dataset [36] involved in this study is a complete patient data bank including major health indicators, demographics, medical histories, and so on. The variables include age, sex, cholesterol, smoking, and any medical conditions existing prior to heart disease [37], [38]. This database is stored in a cloud base healthcare database, providing excellent storage and access facilities for the voluminous data. Incorporating Alagarsundaram, P's [39] Deduplicable Proof of Storage (DPOS) model, which leverages symmetric key encryption and integrity auditing, enables the system to optimize storage by eliminating redundant data and validating integrity without decryption. This positively impacts the scalability and security of patient records as the dataset grows. Because of this facility, the dataset acquired enormous scaling possibilities so that it can grow with further collections of patient data. In conjunction with cloud infrastructure, it supports seamless access and processing of the data set—that is, efficient management and analysis [40], [41].

$$S_{t+1} = S_t + D_t - R_t \quad \ldots (1)$$

where S_t represents the total stored data size at time t, D_t is the new incoming patient data at time t, and R_t is the redundant data removed by DPOS at time t.

$$I_t = \text{Audit}\big(E(K, S_t)\big) \quad \ldots (2)$$

where I_t is the integrity status at time $t, E(K, S_t)$ is the symmetric key encryption of the data S_t using the key K, and $\text{Audit}(\cdot)$ is the integrity auditing function.

$$A = \frac{N_u \times R_d}{L_c} \quad \ldots (3)$$

where A denotes the data access throughput (requests per second), N_u is the number of simultaneous users, R_d is the average request data size per user, and L_c is the latency caused by the cloud infrastructure.

These mechanisms together ensure scalable, secure, and efficient management of the patient dataset as it grows.

CLOUD STORAGE

Cloud Storage is used to store and manage the Clinic Heart Disease Dataset securely [42]. By leveraging cloud infrastructure, the dataset is easily accessible from anywhere, enabling seamless retrieval and storage of large volumes of patient data. Cloud storage ensures that the data is scalable, allowing for continuous updates as new patient information is added. A distributed approach, such as the hybrid Edge-Fog-Cloud model proposed by Yalla et al. [43], significantly enhances data handling efficiency and minimizes latency. This directly strengthens the model's scalability and supports access to heart disease data, aligning with cloud-based infrastructure. Additionally, cloud solutions provide reliable backups, reducing the risk of data loss. The flexibility and scalability of cloud storage also support efficient processing and integration with machine learning models, facilitating smooth data handling and analysis [44], [45].

The total latency L in this distributed setup can be modeled as:

$$L = \alpha L_e + \beta L_f + \gamma L_c \quad \ldots (4)$$

where L_e, L_f, and L_c represent the latencies at the Edge, Fog, and Cloud layers respectively, and α, β, and γ are the corresponding workload proportions (with $\alpha + \beta + \gamma = 1$).

The growth of cloud storage capacity S_c over time t is described by the exponential model:

$$S_c = S_0 \times (1 + r)^t \quad \ldots (5)$$

where S_0 is the initial storage capacity, and r is the data growth rate due to continuous updates. To ensure data durability, the probability of successful backup across n distributed nodes is:

$$P_{\text{backup}} = 1 - \prod_{i=1}^{n} (1 - p_i) \quad \ldots (6)$$

where p_i denotes the probability that the i-th node successfully stores the backup. This distributed cloud architecture provides reliable backups, minimizes data loss risks, and supports smooth integration with machine learning models for further analysis.

PREPROCESSING

Preprocessing is a vital step to ensure a commonly used dataset is appropriate for the training of any model. This begins with a phase of Data Cleaning that identifies and treats missing or inconsistent values so that the dataset can be considered complete and reliable. Later, Normalization is applied to scale data values and bring all input features at the same level, thereby increasing model performance. Lastly, Encoding converts the categorical variables into meaningful number values using label encoding-discrete variables like smoking status and heart disease diagnosis for example [46]. Without such transformations, the algorithm has a hard time distinguishing and processing these features, so this step completes making the data ready for the behavior-train ML model. Based on Sareddy's [47] work in AI-enhanced hiring systems, the emphasis on blockchain for trust and AI for automation strengthens the model's data validation approach, positively impacting reliability and reducing manual error.

$$X_{\text{clean}} = X_{\text{raw}} + M \times \hat{X} \quad \dots (7)$$

where X_{clean} is the cleaned dataset, X_{raw} is the original dataset with missing entries, M is a mask matrix indicating missing data positions, and $\hat{X}$ contains the imputed estimates.

Normalization is then applied to scale features to a common range to enhance model performance:

$$X_{\text{norm}} = \frac{X - \mu}{\sigma} \quad \dots (8)$$

where X_{norm} is the normalized feature value, X is the original feature value, μ is the mean of the feature, and σ is the standard deviation.

Finally, categorical variables such as smoking status and heart disease diagnosis are transformed into numeric representations through label encoding:

$$X_{\text{encoded}} = \text{LabelEncode}\left(X_{\text{categorical}}\right) \quad \dots (9)$$

where X_{encoded} is the numeric encoding of the categorical feature $X_{\text{categorical}}$, and LabelEncode $(\cdot)$ is the encoding function mapping categories to integers.

These preprocessing steps ensure the data is complete, normalized, and suitably formatted, thus facilitating effective machine learning model training and reducing manual errors.

CLASSIFICATION

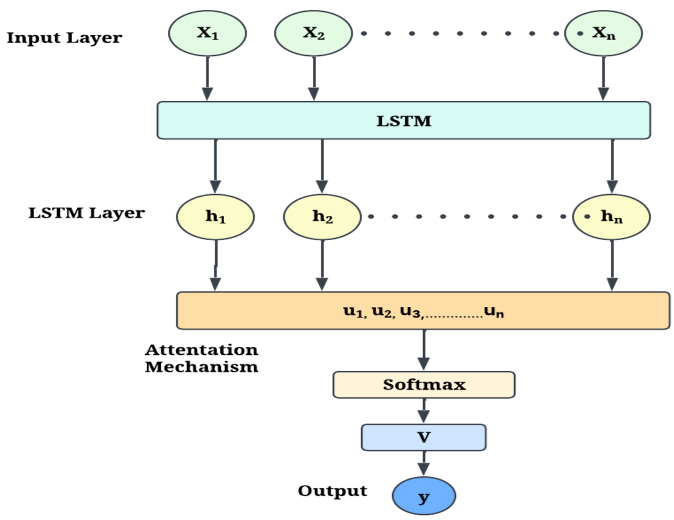


Figure 2: Architecture of LSTM & Attention Mechanism

At the center of this approach stands the Hybrid LSTM + Attention Model, which tries to incorporate features from both LSTM networks and the Attention Mechanism. This hybrid model uses sequential data and focuses on the most important features to detect heart disease. It gets trained on preprocessed data, where LSTM learns temporal dependencies and patterns over the dataset and the Attention Mechanism makes the model concentrate on the most important features at each time step to complement its accuracy and interpretability [48].

$$\hat{y} = \operatorname{softmax}(W_{\text{out}} \cdot c_t + b_{\text{out}}) \quad \dots (10)$$

Were, W_{out} is the weight matrix for the output layer, b_{out} is the bias term.

LSTM

LSTM is a class of recurrent neural network that is capable of learning and remembering long-term dependencies in sequential data [49]. Basically, it is needed in places where the system has to remember some information for long sequences, for example, in terms of time-series analysis or any medical history data. An LSTM accomplishes this by having special units called forget gates, input gates, and output gates, regulating the flow of information to retain only what is relevant for a longer period and to discard irrelevant information: LSTM learns sequentially along the data from changes in a patient's history, cholesterol, etc., to arrive at useful predictions for heart disease detection [50]. In the system proposed, real-time data input from the user is passed through the trained model to generate an instantaneous prediction corresponding to the heart disease risk. Gudivaka's [51] integration of LSTM/GRU with RPA showcases a highly efficient predictive analytics approach, achieving 0.89 accuracy. This reinforces the model's capability to deliver fast, reliable predictions and supports future automation of diagnostic workflows.

Input Gate

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad \dots (11)$$

Were, i_t is the input gate output, $W_i$is the weight matrix for the input gate, b_i is the bias term.

Output Gate

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad \dots (12)$$

Were, o_t is the output gate output, W_o is the weight matrix for the output gate, b_o is the bias term.

Attention Mechanism

An Attention Mechanism is any technique that allows the model to concentrate on certain portions of the input sequence that carry more weight for the purpose at hand. Thus, it differs from otherwise treating every feature as equal by assigning attention weights to each feature. This allows the model to consider one or a few features over others. All of this is equally important when addressing complex data set issues such as medical records. In heart disease detection, attention mechanism ensures that the model is able to consider very relevant features such as age, smoking status, or past medical history. Thus, improving both prediction and interpretation of the prediction. With both these aspects, the model focuses on really important features that result in more reliable and interpretable results, which helps bring further clarity to health-care decisions.

Attention Scores

$$e_{ij} = \operatorname{score}(h_i, h_j) = h_i^T W_a h_j \quad \dots (13)$$

Were, e_{ij} is the attention score between the h_i^T and h_j hidden states, W_a is the learnable attention weight matrix.

Attention Weights

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_k \exp(e_{ik})} \quad \dots (14)$$

Were, α_{ij} is the normalized attention weight between the $i-th$ and $j-th$ hidden states.

IV. RESULT AND DISCUSSION

The system proposed for the detection of heart diseases from the cloud has utilized LSTM and Attention Mechanism and therefore has really exhibited outstanding results with an accuracy of 98.9%; a precision of 99.0%; specificity of 99.0%; and sensitivity of 98.8% with an F-measure of 98.9%. The lower number of false positives and false negatives indicates that the model is reliable and accurate for heart disease detection. The model performed well in Negative Predictive Value (NPV), indicating that it is correct 98.7% of the time when predicting non-disease cases. These results validate the effectiveness in terms of adding LSTM and attention mechanism, aiding the model in concentrating on the most relevant features in patient data, increasing both accuracy and interpretability of the biosensor. Therefore, this study is an indication of using machine learning solutions that are cloud-based for detection of heart disease in a scalable and reliable manner, enabling a strong instrument for early diagnosis and intervention. Future works may also optimize the model for real-time application and continuous monitoring.

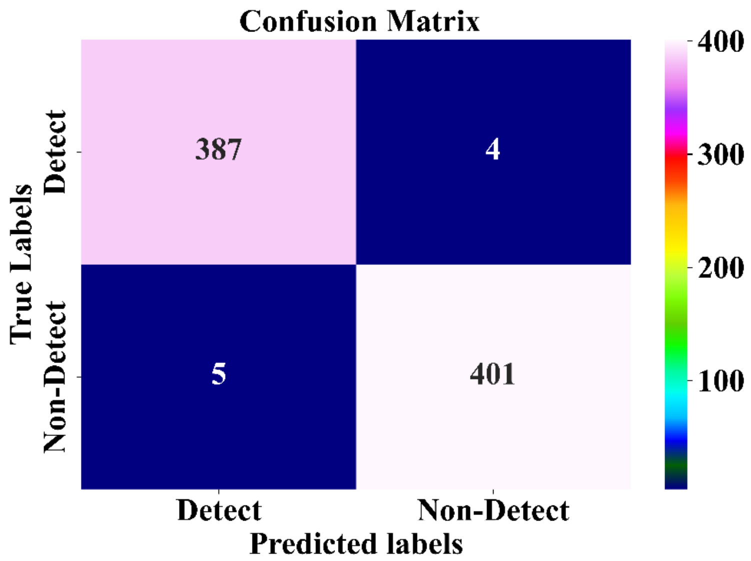


Figure 3: Confusion Matrix

In the matrix seen above is found a model confusion matrix for heart disease prediction. It consists of the following four quadrants: 387 true positives (Detection is Prophetic), 4 false positives (Detect is not really Detect), 5 false negatives (Not detected as non-detect) and 401 true negatives (the correct prediction as One Non-Detect). The matrix presented above demonstrates the excellent performance of the model: prediction ratios for both Detect and Non-Detects are very high, while the associated false positives, as well as negatives, are low. The color gradient depicts the frequencies of each category ranging upwards from lighter to darker values. From the information provided, an understanding of the model is demonstrated in detection, accuracy, and reliability regarding heart disease.

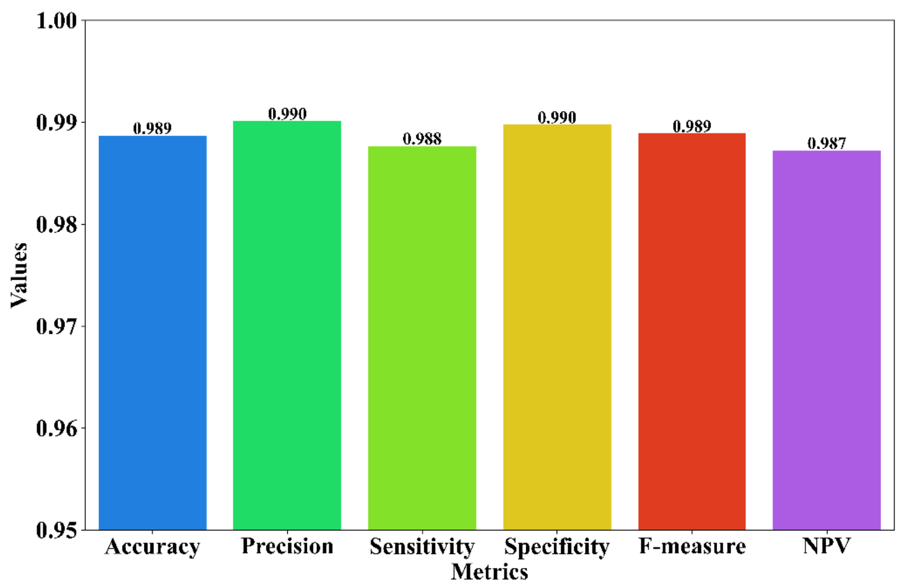


Figure 4: Metrics

The performance metric of heart disease detection model is presented in the form of bar charts of some important parameters. The model executed very high Accuracy (0.989), Precision (0.990) and Specificity (0.990) representing its strong ability to identify rightly positives and negatives. Sensitivity (0.988) and F-measure (0.989) also indicate the capability of the model in detecting true positives against precision and recall. NPV (0.987) proves the model's capability in correctly predicting negative cases. Overall, the model did exceptionally perform across each and every particular metric; thus, it is highly reliable and precise in heart disease detections.

V. CONCLUSIONS

This research shows that the integration of LSTM and the Attention Mechanism into cloud technology can lead to delivering high-performance and scalable solutions for automated diagnosis of heart diseases. Alarming accuracy, precision, and sensitivity suggest that such a system may improve early diagnosis and intervention effectiveness in healthcare. Traditional approaches usually have low processing speeds, data inconsistency, and poor interpretability. The incorporation of cloud computing into this approach turns out to be very beneficial in dealing with large medical datasets in real time. This will allow the improvement of heart disease detection effectiveness and efficiency while also adding credibility to decision-making among health practitioners. Future scopes of work can focus on improving the model for real-time applications and monitoring patients continuously to provide personalized and proactive prototype health solutions.

REFERENCES

[1] S. Hussain *et al.*, "Modern Diagnostic Imaging Technique Applications and Risk Factors in the Medical Field: A Review," *BioMed Res. Int.*, vol. 2022, no. 1, p. 5164970, 2022, doi: 10.1155/2022/5164970.

[2] A. A. Nancy, D. Ravindran, P. M. D. Raj Vincent, K. Srinivasan, and D. Gutierrez Reina, "IoT-Cloud-Based Smart Healthcare Monitoring System for Heart Disease Prediction via Deep Learning," Electronics, vol. 11, no. 15, Art. no. 15, Jan. 2022, doi: 10.3390/electronics11152292.

[3] J. Sharifi-Rad *et al.*, "Diet, Lifestyle and Cardiovascular Diseases: Linking Pathophysiology to Cardioprotective Effects of Natural Bioactive Compounds," *Int. J. Environ. Res. Public. Health*, vol. 17, no. 7, Art. no. 7, Jan. 2020, doi: 10.3390/ijerph17072326.

[4] S. P. Karunathilake and G. U. Ganegoda, "Secondary Prevention of Cardiovascular Diseases and Application of Technology for Early Diagnosis," *BioMed Res. Int.*, vol. 2018, no. 1, p. 5767864, 2018, doi: 10.1155/2018/5767864.

[5] S. Roy, T. Meena, and S.-J. Lim, "Demystifying Supervised Learning in Healthcare 4.0: A New Reality of Transforming Diagnostic Medicine," *Diagnostics*, vol. 12, no. 10, Art. no. 10, Oct. 2022, doi: 10.3390/diagnostics12102549.

[6] R. ElShawi, Y. Sherif, M. Al-Mallah, and S. Sakr, "Interpretability in healthcare: A comparative study of local machine learning interpretability techniques," *Comput. Intell.*, vol. 37, no. 4, pp. 1633–1650, 2021, doi: 10.1111/coin.12410.

[7] B. K. Reddy and D. Delen, "Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology," *Comput. Biol. Med.*, vol. 101, pp. 199–209, Oct. 2018, doi: 10.1016/j.compbiomed.2018.08.029.

[8] E. Yang, H. Zhang, X. Guo, Z. Zang, Z. Liu, and Y. Liu, "A multivariate multi-step LSTM forecasting model for tuberculosis incidence with model explanation in Liaoning Province, China," *BMC Infect. Dis.*, vol. 22, no. 1, p. 490, May 2022, doi: 10.1186/s12879-022-07462-8.

[9] A. K. Sandhu, "Big data with cloud computing: Discussions and challenges," Big Data Min. Anal., vol. 5, no. 1, pp. 32–40, Mar. 2022, doi: 10.26599/BDMA.2021.9020016.

[10] M. Wei, Y. Song, X. Wang, and J. Peng, "Safety diagnosis of TBM for tunnel excavation and its effect on engineering," Neural Comput. Appl., vol. 33, no. 3, pp. 997–1005, Feb. 2021, doi: 10.1007/s00521-020-05371-y.

[11] Z. Chen, Y. Zhang, J. Li, X. Li, and L. Jing, "Diagnosing tunnel collapse sections based on TBM tunneling big data and deep learning: A case study on the Yinsong Project, China," Tunn. Undergr. Space Technol., vol. 108, p. 103700, Feb. 2021, doi: 10.1016/j.tust.2020.103700.

[12] A. R. Folsom et al., "Longitudinal increases in blood biomarkers of inflammation or cardiovascular disease and the incidence of venous thromboembolism," J. Thromb. Haemost., vol. 16, no. 10, pp. 1964–1972, Oct. 2018, doi: 10.1111/jth.14241.

[13] J.-D. Huang, J. Wang, E. Ramsey, G. Leavey, T. J. A. Chico, and J. Condell, "Applying Artificial Intelligence to Wearable Sensor Data to Diagnose and Predict Cardiovascular Disease: A Review," Sensors, vol. 22, no. 20, Art. no. 20, Jan. 2022, doi: 10.3390/s22208002.

[14] D. McGraw and K. D. Mandl, "Privacy protections to encourage use of health-relevant digital data in a learning health system," Npj Digit. Med., vol. 4, no. 1, pp. 1–11, Jan. 2021, doi: 10.1038/s41746-020-00362-8.

[15] R. Ayyadurai, "Transaction Security in E-Commerce: Big Data Analysis in Cloud Environments," Int. J. Inf. Technol. Comput. Eng., vol. 10, no. 4, pp. 176–186, Dec. 2022.

[16] A. Djenna, S. Harous, and D. E. Saidouni, "Internet of Things Meet Internet of Threats: New Concern Cyber Security Issues of Critical Cyber Infrastructure," Appl. Sci., vol. 11, no. 10, Art. no. 10, Jan. 2021, doi: 10.3390/app11104580.

[17] M. Bagaa, T. Taleb, J. Riekki, and J. Song, "Collaborative Cross System AI: Toward 5G System and Beyond," IEEE Netw., vol. 35, no. 4, pp. 286–294, Jul. 2021, doi: 10.1109/MNET.011.2000607.

[18] G. Kaissis et al., "End-to-end privacy preserving deep learning on multi-institutional medical imaging," Nat. Mach. Intell., vol. 3, no. 6, pp. 473–484, Jun. 2021, doi: 10.1038/s42256-021-00337-8.

[19] B. Duan, C. Guo, and H. Liu, "A hybrid genetic-particle swarm optimization algorithm for multi-constraint optimization problems," Soft Comput., vol. 26, no. 21, pp. 11695–11711, Nov. 2022, doi: 10.1007/s00500-022-07489-8.

[20] F. Pelekoudas-Oikonomou et al., "Blockchain-Based Security Mechanisms for IoMT Edge Networks in IoMT-Based Healthcare Monitoring Systems," Sensors, vol. 22, no. 7, Art. no. 7, Jan. 2022, doi: 10.3390/s22072449.

[21] W. Pulakhandam, V. R. Vallu, and V. K. Samudrala, "OPTIMIZING CLOUD, FINANCE, AND E-COMMERCE WITH AI: ADVANCING DECISION-MAKING USING SNNS, CMA-ES, AND HESN FOR SCALABLE AND ADAPTIVE SYSTEMS," Int. J. Inf. Technol. Comput. Eng., vol. 10, no. 4, pp. 257–277, Dec. 2022.

[22] L. O. Badru, V. Vasudevan, G. I. Lingam, and M. G. M. Khan, "MERN Stack Web-Based Education Management Information Systems for Pacific Island Countries," SN Comput. Sci., vol. 4, no. 1, p. 70, Nov. 2022, doi: 10.1007/s42979-022-01457-7.

[23] J. Zheng and Y. Wang, "A Hybrid Multi-Objective Bat Algorithm for Solving Cloud Computing Resource Scheduling Problems," Sustainability, vol. 13, no. 14, Art. no. 14, Jan. 2021, doi: 10.3390/su13147933.

[24] Y. L. Qiu and G. F. Xiao, "Research on Cost Management Optimization of Financial Sharing Center Based on RPA," Procedia Comput. Sci., vol. 166, pp. 115–119, Jan. 2020, doi: 10.1016/j.procs.2020.02.031.

[25] F. T. DeZoort and P. D. Harrison, "Understanding Auditors' Sense of Responsibility for Detecting Fraud Within Organizations," J. Bus. Ethics, vol. 149, no. 4, pp. 857–874, Jun. 2018, doi: 10.1007/s10551-016-3064-3.

[26] M. Qiu, K. Gai, B. Thuraisingham, L. Tao, and H. Zhao, "Proactive user-centric secure data scheme using attribute-based semantic access controls for mobile clouds in financial industry," Future Gener. Comput. Syst., vol. 80, pp. 421–429, Mar. 2018, doi: 10.1016/j.future.2016.01.006.

[27] P. Radhakrishnan, S. Ganesan, V. S. Musam, N. K. Musham, and Karthick M, "OPTIMIZING HEALTHCARE DATA COLLECTION AND SECURITY THROUGH CLOUD COMPUTING," Int. J. Mod. Electron. Commun. Eng., vol. 10, no. 1, 2022, [Online]. Available: https://ijmece.com/ijmeceadmin/upload/ijlbps_67e3db479042c.pdf

[28] S. Fugkeaw, "A Fine-Grained and Lightweight Data Access Control Model for Mobile Cloud Computing," IEEE Access, vol. 9, pp. 836–848, 2021, doi: 10.1109/ACCESS.2020.3046869.

[29] N. Agrawal, "Dynamic load balancing assisted optimized access control mechanism for Edge-Fog-Cloud network in Internet of Things environment," Concurr. Comput. Pract. Exp., vol. 33, no. 21, p. e6440, 2021, doi: 10.1002/cpe.6440.

[30] M. A. Hossain, R. Shams, Q. Mohammed, H. Elsie, and O. Abdus-Samad, "Factors Affecting Performance of Open Government Data Initiatives: A Multi-Method Approach Using Sem and FSQCA," J. Organ. Comput. Electron. Commer., vol. 31, no. 4, pp. 300–319, Oct. 2021, doi: 10.1080/10919392.2021.2018258.

[31] S. Narla, "BIG DATA PRIVACY AND SECURITY USING CONTINUOUS DATA PROTECTION DATA OBLIVIOUSNESS METHODOLOGIES," J. Sci. Technol., vol. 7, no. 2, Art. no. 2, Mar. 2022.

[32] B. Langmead and A. Nellore, "Cloud computing for genomic data analysis and collaboration," Nat. Rev. Genet., vol. 19, no. 4, pp. 208–219, Apr. 2018, doi: 10.1038/nrg.2017.113.

[33] M. AlJame, I. Ahmad, A. Imtiaz, and A. Mohammed, "Ensemble learning model for diagnosing COVID-19 from routine blood tests," Inform. Med. Unlocked, vol. 21, p. 100449, Jan. 2020, doi: 10.1016/j.imu.2020.100449.

[34] A. W. Khan et al., "Analyzing and Evaluating Critical Challenges and Practices for Software Vendor Organizations to Secure Big Data on Cloud Computing: An AHP-Based Systematic Approach," IEEE Access, vol. 9, pp. 107309–107332, 2021, doi: 10.1109/ACCESS.2021.3100287.

[35] S. Kodadi, "High-Performance Cloud Computing and Data Analysis Methods in the Development of Earthquake Emergency Command Infrastructures," J. Curr. Sci., vol. 10, no. 3, 2022.

[36] S. Ulianova, "cardiovascular disease dataset," Kaggle. [Online]. Available: https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset

[37] Y. Zhang et al., "High triglyceride-glucose index is associated with adverse cardiovascular outcomes in patients with acute myocardial infarction," Nutr. Metab. Cardiovasc. Dis., vol. 30, no. 12, pp. 2351–2362, Nov. 2020, doi: 10.1016/j.numecd.2020.07.041.

[38] S. Sze, P. Pellicori, J. Zhang, J. Weston, and A. L. Clark, "The impact of malnutrition on short-term morbidity and mortality in ambulatory patients with heart failure," Am. J. Clin. Nutr., vol. 113, no. 3, pp. 695–705, Mar. 2021, doi: 10.1093/ajcn/nqaa311.

[39] P. Alagarsundaram, "SYMMETRIC KEY-BASED DUPLICABLE STORAGE PROOF FOR ENCRYPTED DATA IN CLOUD STORAGE ENVIRONMENTS: SETTING UP AN INTEGRITY AUDITING HEARING," Int. J. Eng. Res. Sci. Technol., vol. 18, no. 4, pp. 128–136, 2022.

[40] N. L. Fitriyani, M. Syafrudin, G. Alfian, and J. Rhee, “HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System,” IEEE Access, vol. 8, pp. 133034–133050, 2020, doi: 10.1109/ACCESS.2020.3010511.

[41] D. Shah, S. Patel, and S. K. Bharti, “Heart Disease Prediction using Machine Learning Techniques,” SN Comput. Sci., vol. 1, no. 6, p. 345, Oct. 2020, doi: 10.1007/s42979-020-00365-y.

[42] P. G. Shynu, V. G. Menon, R. L. Kumar, S. Kadry, and Y. Nam, “Blockchain-Based Secure Healthcare Application for Diabetic-Cardio Disease Prediction in Fog Computing,” IEEE Access, vol. 9, pp. 45706–45720, 2021, doi: 10.1109/ACCESS.2021.3065440.

[43] R. K. M. K. Yalla, A. R. G. Yallamelli, and V. Mamidala, “A Distributed Computing Approach to IoT Data Processing: Edge, Fog, and Cloud Analytics Framework,” Int. J. Inf. Technol. Comput. Eng., vol. 10, no. 1, pp. 79–94, Jan. 2022.

[44] F. Desai et al., “HealthCloud: A system for monitoring health status of heart patients using machine learning and cloud computing,” Internet Things, vol. 17, p. 100485, Mar. 2022, doi: 10.1016/j.iot.2021.100485.

[45] A. S. Mohammad and M. R. Pradhan, “Machine learning with big data analytics for cloud security,” Comput. Electr. Eng., vol. 96, p. 107527, Dec. 2021, doi: 10.1016/j.compeleceng.2021.107527.

[46] P. Cerda, G. Varoquaux, and B. Kégl, “Similarity encoding for learning with dirty categorical variables,” Mach. Learn., vol. 107, no. 8, pp. 1477–1494, Sep. 2018, doi: 10.1007/s10994-018-5724-2.

[47] M. R. Sareddy, “Revolutionizing Recruitment: Integrating AI and Blockchain for Efficient Talent Acquisition,” IMPACT Int. J. Res. Bus. Manag. IJRBM, vol. 10, no. 8, pp. 33–44, Aug. 2022.

[48] C. Xu, C. Li, and X. Zhou, “Interpretable LSTM Based on Mixture Attention Mechanism for Multi-Step Residential Load Forecasting,” Electronics, vol. 11, no. 14, Art. no. 14, Jan. 2022, doi: 10.3390/electronics11142189.

[49] A. Sherstinsky, “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network,” Phys. Nonlinear Phenom., vol. 404, p. 132306, Mar. 2020, doi: 10.1016/j.physd.2019.132306.

[50] R. Huang et al., “Well performance prediction based on Long Short-Term Memory (LSTM) neural network,” J. Pet. Sci. Eng., vol. 208, p. 109686, Jan. 2022, doi: 10.1016/j.petrol.2021.109686.

[51] B. R. Gudivaka, “Real-Time Big Data Processing and Accurate Production Analysis in Smart Job Shops Using LSTM/GRU and RPA,” Int. J. Inf. Technol. Comput. Eng., vol. 10, no. 3, pp. 63–79, Sep. 2022.

IJARETY

INNO SPACE
SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर
NISCAIR

International Journal of Advanced Research in Education and TechnologY (IJARETY)

www.ijarety.in editor.ijarety@gmail.com