V.Narasimha Swamy, 2026, 14:1
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Efficient And Accurate Cloud-Assisted Medical Pre-Diagnosis With Privacy Preservation

[1]V.Narasimha Swamy, [2]M.Bhuvaneswari , [3]M.Pavan Kumar Reddy , [4]S.Sai Sindhuri
[1]Assistant Professor [2,3,4] UG students
1,2,3,4 Department of Computer Science and Engineering, Sai Rajeswari Institute of Technology

Abstract- Cloud-based healthcare services help doctors deliver quick assessments and early diagnoses, even when patients are far from hospitals. However, sending medical data and diagnostic models to the cloud raises serious privacy concerns because sensitive patient information may not always be safe. The NAIAD framework addresses this by using encrypted kNN and secure Mahalanobis Distance calculations, allowing the cloud to process medical queries without ever seeing the actual data. It also speeds up the search process using a hierarchical encrypted index tree. While NAIAD provides good privacy and accuracy, it still has gaps in data optimization, fine-grained user control, and transparency of the results.The proposed system enhances NAIAD by focusing on smarter data preparation, stronger protection, and better verification. It filters and encrypts only meaningful medical features before outsourcing them, reducing computation and improving performance. Enhanced access control ensures that patients, doctors, and administrators can securely interact with the system while keeping sensitive records fully protected. The system also adds a result-verification mechanism so patients can confirm that the cloud processed their data honestly without modification or tampering. Additionally, the framework introduces better data organization, reduced redundancy, and improved communication efficiency—giving quicker responses during diagnosis. Security layers are strengthened to withstand modern cyber-attacks, ensuring long term trust in the system

Keywords: Cloud-based healthcare, privacy preservation, encrypted kNN, secure Mahalanobis distance, hierarchical encrypted index tree, NAIAD framework, secure medical data outsourcing, smart data filtering, fine-grained access control, result verification, data integrity, secure cloud diagnosis, optimized data organization, reduced redundancy, improved communication efficiency, cyber-attack resistance.

I. INTRODUCTION

The rapid growth of cloud computing has transformed the healthcare sector by enabling remote medical services, efficient data storage, and fast access to diagnostic tools. Cloud-assisted medical diagnosis systems allow patients and doctors to perform preliminary diagnosis without physical visits. However, outsourcing sensitive medical data to cloud servers introduces serious concerns related to privacy, data security, and trust.
Medical information such as disease history, laboratory results, and personal health records is highly confidential. Many existing cloud-based diagnosis systems store or process this data in plaintext or apply insufficient security measures, making them vulnerable to data breaches, unauthorized access, and insider threats. Furthermore, the lack of result verification mechanisms allows untrusted cloud servers to modify diagnosis results, affecting reliability.
To address these challenges, this project proposes a secure and privacy-preserving cloud-assisted medical pre-diagnosis system that ensures data confidentiality, accuracy, and integrity while maintaining computational efficiency.

Objective

To Develop a Secure Cloud-Assisted Medical Pre-Diagnosis System

To design and implement a medical pre-diagnosis framework that securely utilizes cloud computing

resources while protecting sensitive patient information.

To Ensure Privacy Protection of Patient Data
To encrypt and securely outsource selected medical features to the cloud, preventing unauthorized access or data leakage.

To Perform Relevant Medical Feature Selection
To identify and select only essential medical attributes required for diagnosis, thereby reducing computational complexity and enhancing efficiency.

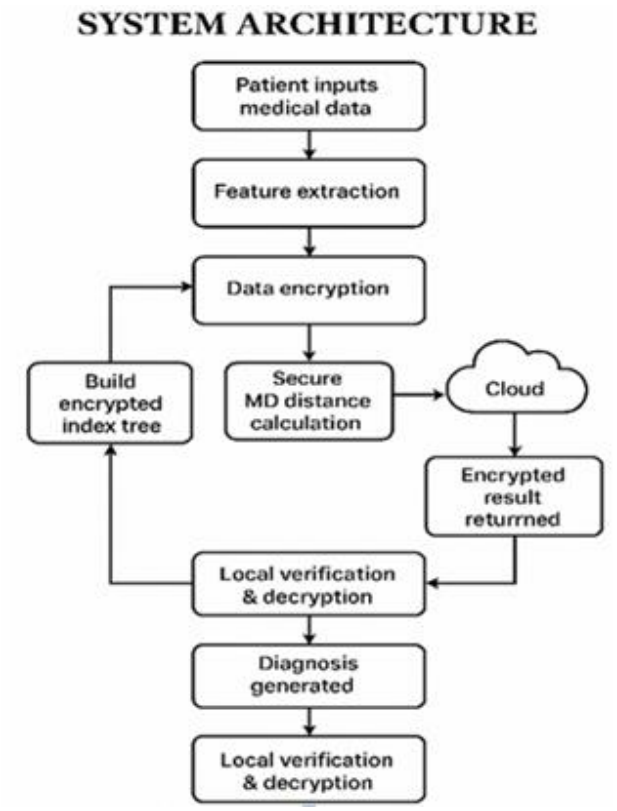
To Implement Distance-Based Classification for Diagnosis
To apply a distance-based classification algorithm for accurate medical pre-diagnosis without revealing raw patient data to cloud servers.

To Incorporate a Lightweight Verification Mechanism
To design a verification method that ensures the correctness and integrity of diagnostic results returned by the cloud.

II. METHODOLOGY / DESIGN

The methodology of the proposed Efficient and Accurate Cloud-Assisted Medical Pre-Diagnosis with Privacy Preservation system is designed to securely process sensitive medical data while maintaining computational efficiency and diagnostic accuracy. The system adopts a structured and modular approach to ensure that patient information is protected throughout data collection, storage, and diagnosis phases.

The proposed design emphasizes privacy preservation by encrypting medical data before cloud outsourcing and ensures reliability through local verification mechanisms. The system operates as a pre-diagnosis tool to assist healthcare professionals rather than replace clinical decision-making.



User Layer (Patient / Doctor)
The user layer represents patients and doctors who interact with the system through a secure web-based interface. Patients provide medical information such as age, blood pressure, glucose level, and body mass index, while doctors access the generated pre-diagnosis results. This layer enforces role-based access control to ensure that only authorized users can access specific system functionalities. Feature selection is performed locally at this stage to retain only relevant medical attributes. Before any data is transmitted to the cloud, it is encrypted on the client side, ensuring that raw medical data never leaves the user system in plaintext form.

Encryption and Feature Selection Module
The encryption and feature selection module operates at the client side prior to cloud interaction. This module is responsible for identifying and selecting only medically significant attributes while eliminating redundant or irrelevant features. After feature selection, the selected medical data is encrypted using symmetric encryption techniques. By minimizing the amount of data being encrypted and outsourced, this module significantly reduces

computational overhead, improves system efficiency, and enhances overall privacy protection.

Cloud Server Layer

The cloud server layer acts as an untrusted storage and processing environment within the system architecture. It is responsible for storing encrypted medical data and responding to authorized data retrieval requests. The cloud server does not possess encryption keys and therefore cannot access or interpret any plaintext medical information. The cloud is assumed to be honest-but-curious, meaning it performs requested operations correctly but may attempt to infer sensitive information. Due to the encrypted nature of stored data, the cloud is unable to extract any patient-related details.

Diagnosis Engine

The diagnosis engine operates in a trusted environment and is responsible for generating medical pre-diagnosis results. It securely retrieves encrypted medical data from the cloud and performs decryption locally. After decryption, a distance-based classification approach, such as k-Nearest Neighbor (kNN), is applied to analyze patient medical parameters and generate preliminary diagnostic insights. The diagnosis engine serves as a decision-support tool that assists healthcare professionals and does not replace expert clinical judgment

Result Verification Module

The result verification module ensures the integrity and trustworthiness of the diagnosis results generated by the system. This module verifies whether the diagnosis output has been altered during cloud storage or transmission. Hash-based verification techniques are used to detect any unauthorized modification of results. If tampering is detected, the system alerts the user; otherwise, verified results are presented to the doctor or patient. This module guarantees result integrity and prevents cloud-side manipulation of diagnosis outcomes.

III. IMPLEMENTATION

This section presents the partial implementation of the proposed Efficient and Accurate Cloud-Assisted Medical Pre-Diagnosis with Privacy Preservation system. The implementation focuses on key modules completed so far, including feature selection, data encryption, distance-based classification, and result verification. These modules demonstrate the core working of the system while full integration will be completed in later stages.

Feature Selection Implementation

Feature selection is implemented to retain only relevant medical attributes required for diagnosis. This reduces computational overhead and improves efficiency.

```
import pandas as pd

# Load dataset

dataset = pd.read_csv("medical_data.csv")

# Select relevant medical features

selected_features = ['Age', 'BP', 'Glucose', 'BMI', 'Cholesterol']

filtered_data = dataset[selected_features]
```

Data Encryption Implementation

To preserve privacy, selected medical data is encrypted at the client side before cloud storage. Symmetric encryption is used for efficiency.

```
from cryptography.fernet import Fernet

# Generate encryption key

key = Fernet.generate_key()

cipher = Fernet(key)

# Encrypt a sample medical value

encrypted_value = cipher.encrypt(b"120")
```

Cloud Data Storage (Simulation)

The encrypted data is stored in a simulated cloud environment. The cloud stores only encrypted data and does not have access to decryption keys.

```
# Simulated cloud storage

cloud_storage = []

cloud_storage.append(encrypted_value)
```

Distance-Based Classification (kNN)

A distance-based classification approach is used for medical pre-diagnosis. The k-Nearest Neighbor (kNN) algorithm is implemented as a baseline classifier.

```
from sklearn.neighbors import KNeighborsClassifier

# Training data

X_train = filtered_data.iloc[:, :-1]

y_train = dataset['Disease']

# Train kNN model

knn = KNeighborsClassifier(n_neighbors=3)

knn.fit(X_train, y_train)

# Prediction

prediction = knn.predict(X_train.iloc[:1])
```

Result Verification Implementation

To ensure result integrity, a hash-based verification mechanism is implemented. This helps detect any tampering of diagnosis results.

```
import hashlib

# Generate hash for verification

result_hash = hashlib.sha256(str(prediction).encode()).hexdigest()
```

IV. CONCLUSION

The proposed encrypted medical pre-diagnosis system presents a secure and efficient framework for handling sensitive healthcare data while enabling reliable disease prediction. In modern healthcare environments, protecting patient information is extremely important due to increasing cyber threats and strict privacy regulations. The developed system addresses these concerns by applying advanced encryption techniques that ensure patient records remain confidential even while being processed or stored in the cloud.

In conclusion, the proposed encrypted medical pre-diagnosis system successfully combines strong privacy protection, efficient encrypted computation, cloud transparency, and accurate diagnostic support. By addressing the limitations of the NAIAD approach and introducing optimized algorithms and verification mechanisms, the system provides a reliable, secure, and scalable solution for modern healthcare environments. With further development and real-world implementation, such systems have the potential to transform digital healthcare by enabling secure, privacy-preserving medical data analysis on a global scale.

Another important advantage of the proposed model is its scalability. As healthcare systems generate massive amounts of patient data, the system is designed to efficiently manage and process large datasets without compromising privacy or performance. This makes it suitable for integration with modern cloud-based healthcare infrastructures, telemedicine platforms, and digital health management systems.

Furthermore, the system contributes to improving early disease detection by enabling secure pre-diagnosis analysis. Doctors and healthcare professionals can use the insights generated by the system to make faster and more informed clinical decisions. At the same time, patients benefit from stronger privacy protection and improved trust in digital healthcare services.

REFERENCES

1. Wang, C., Ren, K., Lou, W., & Li, J. (2012).Toward Secure and Dependable Storage Services in Cloud Computing.IEEE Transactions on Services Computing. (Focuses on cloud security and data integrity verification.)
2. Li, M., Yu, S., Ren, K., & Lou, W. (2010). Securing Personal Health Records in Cloud Computing: Patient-Centric and Fine-Grained Data Access Control.IEEE Transactions on Parallel and Distributed Systems.(Privacy-preserving cloud-based healthcare systems.)
3. Zhang, R., & Liu, L. (2010).Security Models and Requirements for Healthcare Application Clouds.IEEE Cloud Computing. (Security challenges in healthcare cloud systems.)
4. Chen, Y., Ding, S., Xu, Z., Zheng, H., & Yang, S. (2018).Privacy-Preserving k-Nearest Neighbor Query Over Encrypted Data in Cloud Computing. Sensors (MDPI). (Distance-based classification over encrypted data.)
5. Liu, J., Huang, X., Li, J., & Chen, X. (2019).Secure and Privacy-Preserving Data Sharing in Cloud-

Based e-Health Systems. Future Generation Computer Systems.