








Research Article

Secure and Efficient Data Storage Operations by Using Intelligent Classification Technique and RSA Algorithm in IoT-Based Cloud Computing

Faiqa Sajid ¹, Muhammad Abul Hassan ², Ayaz Ali Khan ³, Muhammad Rizwan ^{1,4},
Natalia Kryvinska ⁵, Karovič Vincent ⁵, and Inam Ullah Khan ⁶

¹Department of Computer Science, Kinnaird College for Women, Lahore 54000, Pakistan

²Department of Computing and Technology, Abasyn University, Peshawar 25000, Pakistan

³Department of Computer Science, University of Lakki Marwat, Lakki Marwat, Pakistan

⁴Secure Cyber Systems Research Group, WMG, University of Warwick, Coventry CV4 7AL, UK

⁵Information Systems Department, Faculty of Management Comenius University in Bratislava, Odbojárov 10, Bratislava 82005, Slovakia

⁶Department of Electronic Engineering School of Engineering & Applied Sciences (SEAS), Isra University, Islamabad Campus, Islamabad, Pakistan

Correspondence should be addressed to Muhammad Abul Hassan; abulhassan900@gmail.com

Received 7 February 2022; Revised 7 March 2022; Accepted 16 March 2022; Published 14 April 2022

Academic Editor: Muhammad Zakarya

Copyright © 2022 Faiqa Sajid et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In mobile cloud services, smartphones may depend on IoT-based cloud infrastructure and information storage tools to conduct technical errands, such as quest, information processing, and combined networks. In addition to traditional finding institutions, the smart IoT-cloud often upgrades the normal impromptu structure by treating mobile devices as corporate hubs, e.g., by identifying institutions. This has many benefits from the start, with several significant problems to be overcome in order to enhance the unwavering consistency of the cloud environment while Internet of things connects and improves decision support system of the entire network. In fact, similar issues apply to monitor loading, resistance, and other security risks in the cloud state. Right now, we are looking at changed arrangement procedures in MATLAB utilizing cardiovascular failure information and afterward protecting that information with the assistance of RSA calculation in mobile cloud. The calculations tried are SVM, RF, DT, NB, and KNN. In the outcome, the order strategies that have the best exactness result to test respiratory failure information will be recommended for use for enormous scope information. Instead, the collected data will be transferred to the mobile cloud for preservation using the RSA encryption algorithm.

1. Introduction

Advancements in smart world have upgraded every industry. Intercommunication is the problem where desired results need to be obtained from implementing line models. Today's modern industry is comprised of complex system models which intend to improve over-all process. Sustained advancement in such industry required reliability and efficiency to Decision Support Systems (DSSs) [1].

The Internet is regarded as the most important invention since it allows users to make inter- and intra-connections among conventional and portable devices using various protocols [2]. Internet of things has newly emerged in last few decades. A large number of devices generate data in IoT environment and forward it for further processing. Main problem faced by devices attached with IoT environment is security which is not properly considered during manufacturing. Moreover, these devices have limited

memory, and data generated by each device are doubled of its memory size. IoT-based cloud computing is the integrated to process huge amount of data. Figure 1 briefly describes the merger of IoT and cloud computing to enhance the decision support system.

IoT-based cloud computing technology is providing the best possible practices which are focused on market research [3].

Cloud computing usually provides tools as entities on call, snappy transmission, and charge as a need to enhance user service. This technology offers utilities which include Platform as a Service (PaaS), Infrastructure as a service (IaaS), and Software as a service (SaaS). It is most continent way to provide computation services and can be accessed remotely [4]. In addition, the primary focal points are (i) offering on-demand solutions, (ii) compensation as invoice, and (iii) administration limits [5]. IoT-based cloud computing has a variety of applications such as web-based management. However, cloud storage requires specific databases to rely on the client's need; if the client wishes to hold their knowledge in free space, cloud computing exposes the server and then the knowledge (data) needs to be checked by Decision Support System for timely and accurate measurements.

Encryption techniques are being used to ensure privacy of data in Internet of Things (IoT) cloud and ensure confidentiality. Machine Learning (ML) Algorithms are employed including Decision Tree (DT), Support Vector Machine (SVM), Naive Bayes (NB), K -nearest Neighbors (KNN), and Random Forest (RF) [6]. Each of the layout strategies displays a characteristic suitability and specificity that depends on the type of dataset [7]. Moreover, there are different assessment measurements for looking at the characterization techniques that every one of them could be helpful relying upon the issue. Fundamental problems for cloud computing are (i) digital unloading, (ii) centralized implementation, and (iii) dynamics which have been extensively explored in previous compositions. A few further changes have been introduced to overhaul protection for Clouds themselves, such as protection structure [8] based on TCG [9], stable redistribution [10–13], cloud network defence [14, 15], board tool and detachment [16, 17], and insurance [18, 19]. Main concern to IoT-cloud professional organizations is to ensure the confidentiality of records created and managed by mobile gadgets or cloud server [20]. Data/report confidentiality is very critical to the data/record holder because it may include some secret data.

Main contribution of this research study is as follows:

- (i) Five different Machine Learning Algorithms are applied to classify the dataset
- (ii) RSA algorithm is utilized for the data security
- (iii) Fast Encryption and Decryption

Rest of the paper is organized as follows: Section 2 provides background study and related problem. Section 3 provides proposed methodology. Section 4 provides result and discussion, and conclusion is provided in Section 5.

2. Background Study

Machine Learning algorithms are being used for training to explain how the input variables apply to the category [5]. There are various identification approaches for data such as DT, SVM, NB, KNN, RF, LogR, LR, NB, ANN, LC, decision tree, and so on [6, 21]. The association of classifiers and usage of classifier is important [22]. Support Vector Machine (SVM) finds out hyperplane (N -characteristics) in N -dimensional space which distinguishes the data points. [23]. SVM has a few favourable circumstances in which it performs very well when the differentiation between bunches is self-evident and increasingly compelling for high dimensional spaces.. Because the number of tests is more than the number of measurements, it is persuasive.

There exist several drawbacks in SVM which do not deal with massive data sets. Every data point needs to access the sum of training dataset where SVM is considered the optimal choice in comparison with K -nearest neighbor (KNN). However, training tuples are linked to n -quality which usually refers to a stage in the domain of n -measures. Right now, the training tuples are put in a space of n -dimensional progression. The KNN classifier quickly determines sample space for the k preparation of tuples that are the most similar to dark tuples. The K tuples are the nearest neighbors of the darkest tuple [24, 25]. "Closeness" is known as the separation metric, e.g., the Euclidean separation. The distinction between Euclidean and two tuples or tuples $X_1 = (x_{11}, x_{12}, \dots, x_{1n})$ and $X_2 = (x_{21}, x_{22}, \dots, x_{2n})$ is derived from the following:

$$\text{dist}(X_1, X_2) = \sqrt{\sum_{i=1}^n (x_{1i} - x_{2i})^2}. \quad (1)$$

DT is a tree-like flowchart layout where each inner hub is a feature check and other branch is the outcome [26]. Inside a tree the top hub is the root hub.

During development, choice measures for credits are utilized to pick the best possible asset while three basic determination measures for characteristics are the knowledge, benefit ratio, and Gain index. Information gain is depicted as follows:

$$\text{Gain}(E, Y) = \text{Entropy}(E) - \text{Entropy}(E, Y), \quad (2)$$

where E is the target variable and Y is the feature to be split on.

Entropy (E, Y) describes the highlighted information. Tree pruning expects to perceive and take out these branches, with the objective of improving the precision of grouping on inconspicuous information. Random Forest (RF) can be utilized for selection process. Expectations across all trees were grouped to render the last forecast; class approach is utilized to collect information and indicate relapse expectations. The suggested techniques can be used for the progression of settle tests on ultimate conclusion.

Increase in base learner numbers (k) will decrease variance. However, k is lower due to that variance increases. However, for the entire cycle, bias remains constant. The

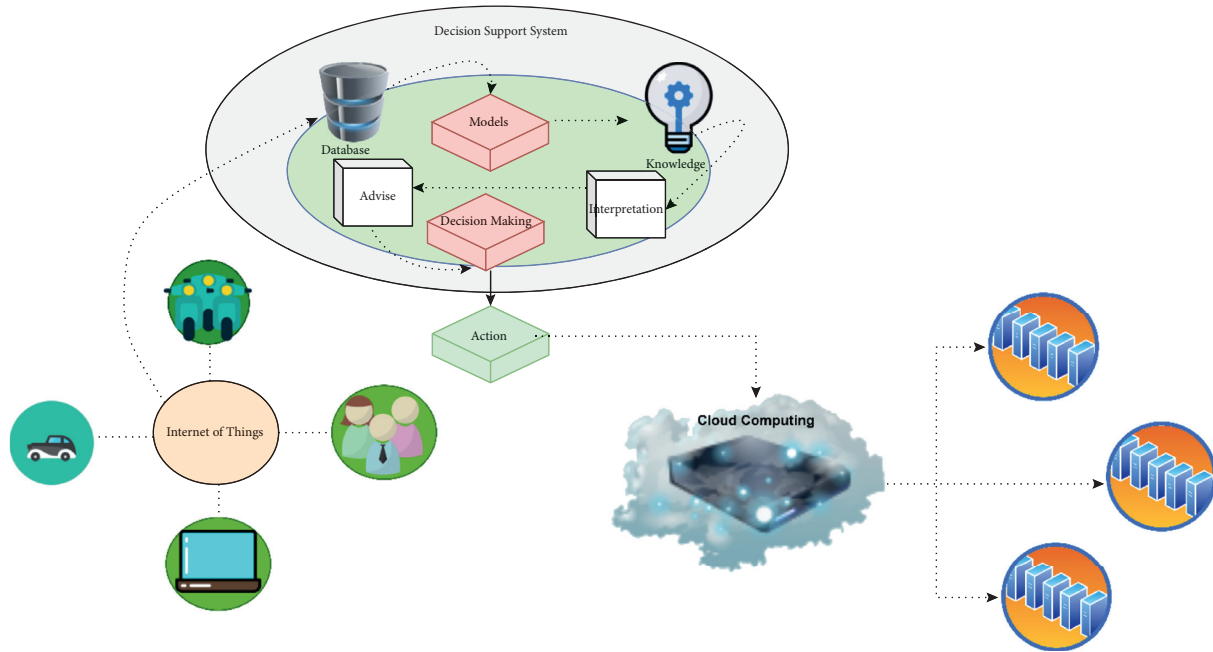


FIGURE 1: IoT-based cloud environment for the optimal decision support system.

cross-validation is used to find k as shown in the following equation:

$$\begin{aligned} \text{Random forest} = & \text{DT (base learner)} + \text{bagging (Row sampling with replacement)} + \text{feature bagging (column sampling)} \\ & + \text{aggregation} \left(\frac{\text{mean}}{\text{median}}, \text{majority vote} \right). \end{aligned} \tag{3}$$

Mobile computing has reduced the privacy and security concerns which usually mimic resource management [27]. Secure data organization is based on ethnicity-based intermediary which uses encryption while integrated phone applications have secured mobile data in better way. Users who store consumer data/records in IoT-based clouds use encryption and trustworthy computations. However, Diffie–Hellman key sharing, bilinear mapping, and the merkle hash tree compel the use of a stable knowledge ownership mechanism [28–31]. Table 1 describes multiple security risks which are elaborated in detail.

Provable Data Ownership (PDP) plot guarantees the clustering, confidentiality, and consistency of mobile consumer data processed on the IoT-based cloud. In Diffie–Hellman symmetric keys, key sharing is required where G_1 and G_2 are cyclical multiplicative social activities with main recruitment q and g as generators of G_1 . The bilinear solution is e : G_1 and G_2 above G_1 , where G_1 and G_2 are cyclical multiplicative social activities with main recruitment q and g as generators of G_1 . Merkle Hash Tree (MHT) is utilized as an analogous tree, with the leaves representing the actual results’ hash calculation. The establishment of a tree must be validated by the verifier. Hyrax, a Hadoop method that facilitates cloud computing on smartphones, was created

by Arockiam and Monikandan [32]. This provides a discerning approach to data exchange and suffers from a middle point of entry. Eugene in like manner executed a scattered media search and data sharing philosophy. Xu et al. [33] suggested methodology that promotes the enhancement of conventional encryption strategies by combining replacement ciphers and conversion ciphers. Both the replacement and the transposing methods used the alphabet for the cipher letter. In their proposed algorithm, plain text is initially translated to the appropriate ASCII code type of each alphabet. Rehman and Manickam [34] suggested an agent-aid paradigm by incorporating multiagent network and choice-making methodology to function load balancing problems in vast clouds. While the mentioned research, manage data sharing which boosts cloud computing efficiency through space-intensive applications, such as distributed data mining. In addition, nature inspired evolutionary computational technique (E-AntHocNet) and Fisheye State routing (FSR) can be considered for further improvements [35–37].

3. Methodology

The proposed model contains two stages. In the initial step, we pass our information to the five classification models. In

TABLE 1: Security risks with details.

Attacks	Explanation
Data revelation	Secure details of the owner shall be exposed to any unwanted user.
Tampering	At the point when any unapproved individual does a few changes in other client's information.
Rejection	At the point when an individual rejected in the wake of communicating something specific that he did not send it.
Virus	Those are particularly well-known threats. Those are symbols that will infect the show of a certain program.
Personality tricking	In this assault, a person mimics himself as someone who owns the details.

the second step, we can pass the highest precision model to the RSA algorithm which we proposed.

Figure 2 depicts the progression of the research that led to the development of the model. The proposed model consists of five techniques where data of patient or disease are sent directly to various classification techniques. Here, classification techniques that we use are SVM, NV, DT, RF, and KNN. SVM is accepted as the AI model utilized in characterization. Algorithm 1 shows the RSA after passing by all these techniques.

A support vector machine's objective is to achieve the most raised edge isolating hyperplane between two classes [24]. Hyperplane is never close to knowledge that focuses on different classes for greater supposition. Hyperplane has distant from the information targeting classification. Moreover, closest to the edge (classifier) is guidance for the vectors [38].

Naive Bayes utilized Bayes theorem and characterized those data, which is not known and trivial to each other. It further provides information of specific element and does not affect another in a class. In addition, likelihood of creating classification of clusters is very restrictive. For order of data, ground breaking calculation is used [25].

Random forests are an ensemble learning method for classification that involves creating a large number of option trees in a short amount of time and delivering a class. For individual trees, this strategy of classes (order) or mean forecast (relapse) is applied. Overfitting to their preparation set is caused by arbitrary tree propensity choices.

Decision Tree is based on supervised learning technique and utilized for the classification of data. Primary aim of DT is the possibility of target class. Parent node and its child nodes are used for clustering of data. DT checks for those nodes which is more gain among others [39].

These systems characterize data and generate outcomes. We consider the outcome in terms of precision. The most exact strategy is considered to be used in the future. Significances of these strategies will be examined in Result Section.

When one outcome is chosen from five possible outcomes, the information must be saved somewhere. As a result, we are sending our data to the cloud for the purpose of data compression. However, we require a secure method of sending data to the cloud. Because security is a top priority, we have implemented the RSA calculation [40] on the cloud to ensure that data are kept safe and secure.

Client data are encoded first and then stored in the mobile Cloud using the RSA technique. When necessary, the client submits a request for information to the Cloud supplier, who validates the customer and transmits the

information. RSA is an encryption that prevents each message from being mapped to a number. RSA is made up of two parts: public and private keys.

3.1. Environment Setup. This research study focuses on the classification of dataset of heart patients using five different techniques. Tools which we are using for classification are SVM, NB, RF, DT, and KNN. These techniques have more roots of them, e.g., SVM have 6 types: Linear, Cubic, Quadratic, Fine Gaussian, Medium Gaussian, and Coarse Gaussian. We are using fine Gaussian SVM technique here because its accuracy is highest as compared with types. Decision Tree has 3 types further: complex tree, simple tree, and medium Tree. Among these types, complex tree has more accuracy so we are using fine tree for classification of data. The third technique which we are using is KNN. It has 6 types further: Weighted KNN, Cubic KNN, Cosine KNN, Coarse KNN, Fine KNN, and Medium KNN. So, we are using Fine KNN from k-nearest neighbor because it has more accuracy than other types of KNN. The fourth technique which we are using is Naïve Bayes. It has 2 types further: Gaussian Naïve Bayes and Kernel Naïve Bayes. So, we are using Kernel Naïve Bayes because it has more accuracy than other types of Naïve Bayes. At the last, we are using bagged trees in Random Forest because it has highest accuracy than others.

According to a recent medical assessment, if factors such as smoking, cholesterol, and diabetes are addressed in a country, people suffering from heart disease can reduce by about 15%. This Cleveland data base was compiled by Robert Detrano and contains 303 occurrences of 76 attributes. Cleveland heart disease is the most commonly used dataset by data mining researchers in the UCI machine learning repository, with 76 variables. Only 11 variables have been employed by researchers to predict and analyse cardiac disorders. Table 2 shows that we have considered heart-attack dataset throughout the research. Total of 1311 instances are extracted from medical dataset. It has 11 numeric value attributes. One class has "0" value considered as negative and other class has "1" considered a positive for heart patients. The dataset originally had 76 features or qualities from 303 patients; however, published research only chose 11 features to be relevant in predicting heart disease. As a result, we will be working with a dataset of 303 patients and 11 characteristics.

Age, sex, cerebral palsy, resting blood pressure, cholesterol, fasting blood sugar test, electrocardiogram, maximum heart rate achieved, exhang, oldpeak, and outcome are attributes used in this paper shown in Table 3.

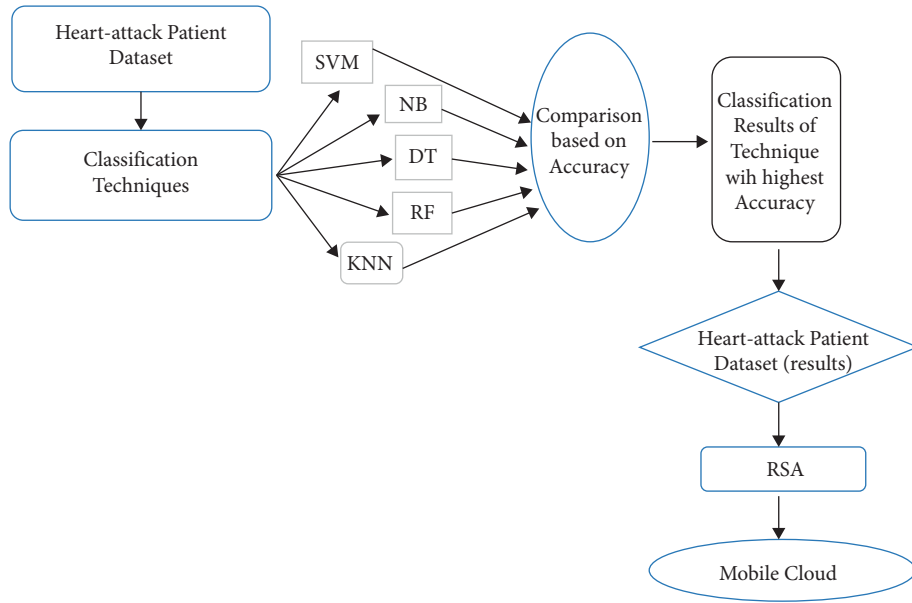


FIGURE 2: Model diagram of proposed technique.

```

    (1) Start
    (2) Choose P, Q where Prime = {P, Q}, P ≠ Q
    (3) Calculate S {S = P * Q}
    (4) Calculate Φ (S), {Φ (S) = (P - 1) (Q - 1)}
    (5) Choose Random W {gcd (Φ (n), W = 1)} and 1 < W < Φ (S)
    (6) Calculate D, {d.e ≡ 1 mod φ (S)}
    (7) Public key {W, n}
    (8) Private Key {D, n}
    (9) End
  
```

ALGORITHM 1: Pseudo code of RSA.

TABLE 2: Dataset with details.

Dataset	No. of rows	No. of columns	Total data
Heart attack	1311	11	14421

TABLE 3: Attributes of dataset.

Attributes	Abbreviation
Age	—
Sex	—
Cerebral palsy	Cp
Resting blood pressure	Trestbps
Cholesterol	Chol
Fasting blood sugar test	Fbs
Electrocardiogram	Restecg
Maximum heart rate achieved	Thalach
Exhang	—
Oldpeak	—
Outcome	Num

For Table 4, arrangement, accuracy, scatter plot, confusion matrix, and ROC metrics are considered.

Client information is first encrypted and then stored in the Cloud via RSA. When the client requests information from the Cloud supplier, the Cloud supplier verifies client and then sends the information to the client.

In RSA, each message is assigned a whole integer, which is a square figure. RSA is made up of two parts: public and private keys. In our Cloud situation, everyone has access to the Public Key, but only the client with the first access to the data has access to the Private Key.

3.2. Encryption and Authentication of Data. To ensure data security, all data packets are encrypted and decoded using a private key, with asymmetric cryptography providing authentication. An upgraded RSA cryptographic method is presented as a solution.

3.2.1. Key Generation. We choose two distinct prime numbers: $a = 61$ and $b = 53$.

Process $n = a * b$, in this manner $n = 61 * 53 = 3233$.

Process Euler's totient work, $\Phi(n) = (a - 1) * (b - 1)$.

Thus, $\Phi(n) = (61 - 1) * (53 - 1) = 60 * 52 = 3120$.

Picked any whole number e , with the end goal that $1 < e < 3120$ that is coprime to 3120. Here, we picked $e = 17$.

Process d , $d = e - 1 \pmod{\Phi(n)}$; therefore, $d = 17 - 1 \pmod{3120} = 2753$.

Public Key is $(e, n) = (17, 3233)$ and Private Key is $(d, n) = (2753, 3233)$.

3.2.2. Encryption. Cloud co-op provides the client with the Public Key (17, 3233) for storing the information. Let suppose $m = 65$ for Client mapped information. This information is further encoded to the specialist co-op by utilizing Public Key $C = 65^{17} \pmod{3233} = 2790$ which is shared among both parties. The Cloud specialist co-op is currently storing these encoded data, i.e., figure content.

3.2.3. Decryption. When the customer requests the information, the Cloud specialist cooperative verifies the client and sends the encrypted data (if the client is legitimate).

At that time, the cloud client computes to decode the information, $m = Cd \pmod{n} = 2790 * 2753 \pmod{3233} = 65$.

When the m esteem is gotten, client will get back the first information. Figure 3 describes the flow chart of RSA.

3.3. The RSA Algorithm for RSA Public and Private Key Pair Generation. Both key exchange and digital signatures can be done with the RSA technique. The mathematics behind RSA is extremely simple despite the fact that it is used with integers with hundreds of digits.

The steps below can be used to establish an RSA public and private key pair:

- (i) P and q are two prime numbers. The modulus, $n = pq$, can be calculated using these numbers.

TABLE 4: Accuracy measures.

Measures	Explanation
Accuracy	Accuracy decides the precision of the calculation in foreseeing occasions
ROC curve	The ROC (receiver operating curve) bends are used to assess the ease of testing
Scatter plot	Scatter (x, y) makes a disperse plot with hovers at the areas determined by the vectors x and y
Confusion matrix	The confusion matrix shows the total number of perceptions in every cell

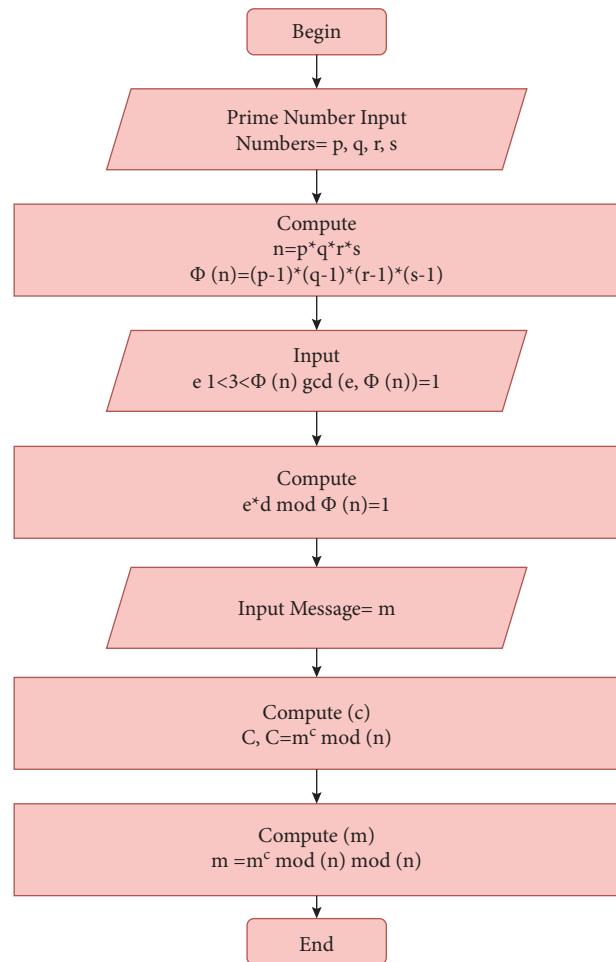


FIGURE 3: Flow chart depicted the RSA decryption algorithm.

- (ii) Choose a third number, r , that is substantially prime to the product $(p - 1)(q - 1)$; r is the public exponent.
- (iii) Use the quotient $((rs - 1)) / ((p - 1)(q - 1))$ to find an integer s . The private exponent is represented by the integer s .
- (iv) The number pair (n, r) is the public key. Although these data are public, determining r from n and s is computationally impossible unless p and q are large enough.
- (v) To encrypt a message, m generates the cipher-text C using equation 4 and the public key. $C = m^r \pmod{n}$ (4)

- (vi) The receiver then uses the private key to decrypt the cipher-text using the following:

$$m = c^5 \text{ mod } (n). \tag{4}$$

4. Results and Discussion

As previously stated, the proposed philosophy lashed back at the heart of informational collection understanding. Linear, Cubic, Quadratic, Fine Gaussian, Medium Gaussian, and Coarse Gaussian classifiers are processed by the SVM classifier with simplified parameters in order to find the ideal hyperplane that divides the pursuit space into two classes. SVM structured information into two classes using a reaction set for classifier in which 0 and 1 attributes are set on which assumption. Figure 4 shows the feature selection that was taken into account in the study.

Figure 4 depicts the feature selection which is age, sex, cerebral palsy, resting blood pressure, cholesterol, fasting blood sugar test, electrocardiogram, maximum heart rate achieved, exchang, and oldpeak considered in the study.

As we discussed in proposed methodology, total six techniques are used. First of all, data of disease are sent towards different classification techniques to classify data into two classes.

In SVM, there are further six types: Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, and Coarse Gaussian SVM. We classify our data through all these six SVMs, and we found the best result from Fine Gaussian SVM as it has maximum accuracy among all SVMs.

In Naive Bayes, there are further two types: Gaussian Naive Bayes and Kernel Naive Bayes. We classify our data with these two Naive Bayes, and we found the best result from Kernel Naive Bayes as it has highest accuracy compared with Gaussian Naive Bayes.

In Decision Tree, there are further three types: Fine Tree, Medium Tree, and Coarse Tree. We classify our data with these three Decision Trees, and we found the best result from Fine Tree as it has highest accuracy compared with other two. We also classify our data through Random Forest. In KNN, there are further six types: Fine KNN, Medium KNN, Coarse KNN, Weighted KNN, Cubic KNN, and Cosine KNN. We classify our data with all these six KNN, and we found the best result from Fine KNN as it has highest accuracy among all KNN.

Table 5 determines classifier’s performance on the basis of classified. Performance of SVM, NV, DT, RF, and KNN is calculated. The affectability, explicitness, exactness, and f-measure are calculated, and accuracy is checked, in order to analyze the classifier’s performance. FP and FN are the proportions of negative instances incorrectly categorised as positive and positive cases incorrectly classed as negative, respectively.

On the basis of classified cases, Table 6 determines the performance of the classifier. Accuracy is determined and analysed based on these classified cases. Out of a total number of examples, the performance of SVM is measured in terms of Correctly Classified Instances and Incorrectly

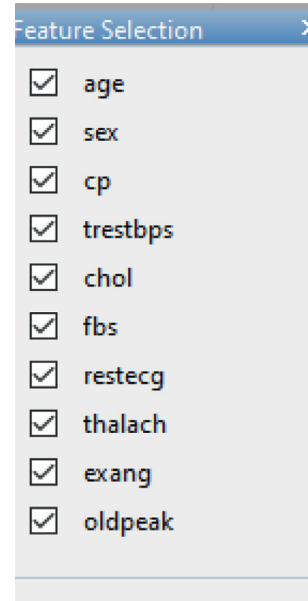


FIGURE 4: Feature selection.

TABLE 5: Classifier’s performance on the basis of classified instances.

Classifiers	Correctly classified	Incorrectly classified
SVM	1308	2
KNN	1306	4
NB	1075	235
RF	1300	10
DT	1282	28

TABLE 6: True and negative values of classifiers.

Classifiers	TP	TN	FP	FN
SVM	742	566	2	0
KNN	742	564	4	0
NB	625	450	118	117
RF	742	558	10	0
DT	742	540	28	0

TABLE 7: Classifier’s performance.

Sensitivity	TP/TP + FN
Specificity	FP/FP + TN
Precision	TP/TP + FP
F-measure	2 (Sen * Pre)/(Sen + Pre)
Accuracy	TP + TN/TP + FP + TN + FN

Classified Instances. Four different classifiers are true positive, true negative, false positive, and false negative values. SVM displays the largest true positive values, while KNN displays the lowest false negative values, DT displays the lowest false positive values, and RF displays the highest false negative values.

Measures used for classification are shown in Table 7 where TP and TN address true positive and true negative which are the degree of positive and negative cases that were really perceived independently. FP and FN address false

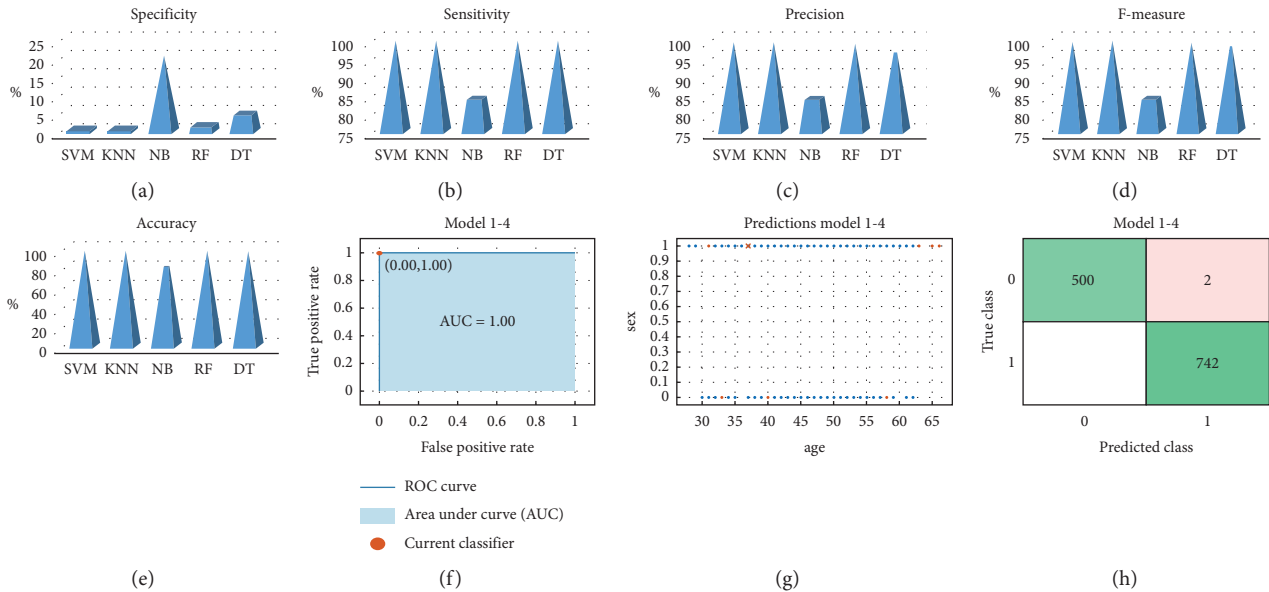


FIGURE 5: Performance measure comparison of five classification techniques. (a) Specificity. (b) Sensitivity. (c) Precision. (d) F -measure. (e) Accuracy. (f) ROC. (g) Scatter plot. (h) Confusion matrix.

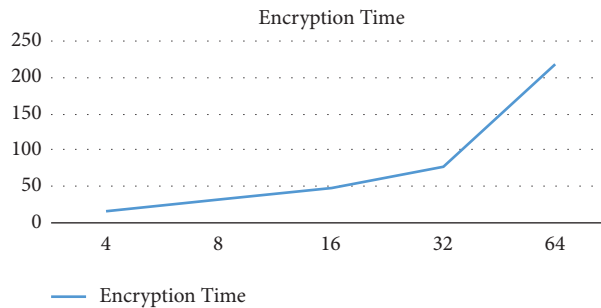


FIGURE 6: Encryption time using RSA algorithm.

TABLE 8: Prediction speed and training time of classifiers.

Classifiers	Prediction speed (obs/sec)	Training time (sec)
SVM	28000	23.01
KNN	16000	3.273
DT	12000	14.37
RF	8000	7.317
NB	14000	8.364

positive and false negative which are the degree of negative cases that were erroneously designated positive and the degree of positive cases that were erroneously named negative independently.

Figures 5(a)–5(d) and 6(e) illustrate the specificity, sensitivity, precision, f -measurement, and accuracy, which are used to calculate and estimate the effectiveness of the classifiers. Figure 5(a) shows that NB has the highest specificity. In Figures 5(b) and 5(c), NB has the lowest sensitivity and precision of 0.844 and 0.824, respectively. In Figure 5(c) and Figure 5(d), SVM achieves the highest precision and f -measure of 0.984 and 0.954, respectively. In Figure 5(e), the accuracy of several classifiers is shown in

percentages, with SVM having the greatest accuracy of 0.974. SVM has the highest accuracy of all the classifiers. The percentage of correctly expressed positive and negative examples is denoted by TP and TN, respectively. Figure 5(f) depicts the categorization model's performance. In Figure 5(g), a scatter plot with horizontal and vertical axes is displayed. As illustrated in Figure 5, FP and FN are the proportions of negative instances incorrectly identified as positive and positive cases incorrectly classified as negative, respectively. SVM has a TP rate of 0.980, while KNN has a rate of 0.963, DT has a rate of 0.933, RF has a rate of 0.982, and NB has a rate of 0.802. SVM has an FP rate of 0.043, KNN of 0.030, DT of 0.14, RF of 0.050, and NB of 0.330.

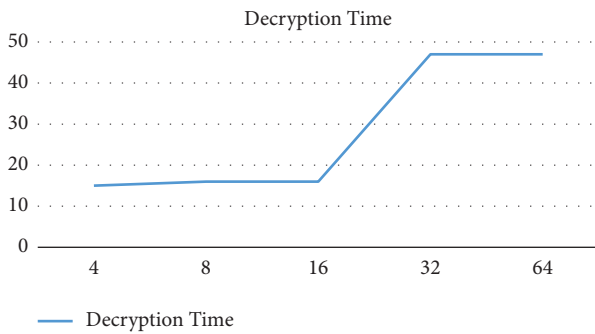


FIGURE 7: Decryption time using RSA algorithm.

The prediction speed and training time of each classifier are shown in Table 8, with SVM having the highest Prediction Speed observed per second (obs/sec) and KNN having the quickest training time.

Graph in Figure 6 represents a file of size 4 kb, 8 kb, 16 kb, 32 kb, and 64 kb taking 16 ms, 31 ms, 47 ms, 78 ms, and 218 ms encryption time using the RSA algorithm.

In RSA, exchange is the public key with the recipient is used for a decryption request. Hold it safe is a proprietary key that is never shared with any other person.

Graph in Figure 7 represents a file of size 4, 8, 16, 32, and 64 kb taking 15, 16, 16, 47, and 47 ms decryption time using the RSA algorithm.

If the key size is identical to the size of the packet to be sent across the network, an effective cryptosystem will produce the best results. Algorithm focused on parameters such as key length, block size, features, and functions. Because we realize data are held elsewhere in cloud storage, we need fast processing speed as well as fast protection. Here, the graph in Figure 6 indicates the success of the suggested case that how long it takes to encrypt the files.

5. Conclusions

An asset-restricted mobile PC saves data on an IoT-based cloud. However, constantly a significant inquiry regarding cloud specialist deals with the records appropriately while mobile cloud computing security is the fundamental concern. The proposed system gives a security system to make sure about the information in mobile cloud computing with the assistance of RSA calculation and other five different intelligent classification techniques.

A classification method based on the proposed methodology was presented to classify heart-attack data. First, dataset is generated for different parameters of heart attack. Second, based on these data, result is calculated in “0” and “1” form. Now, the generated data are sent straight to SVM, KNN, NB, RF, and DT for classification. The accuracies of the training dataset with classifiers are 99.80%, 99.6%, 82%, 99.2%, and 97.8% of SVM, KNN, NB, RF, and DT, respectively. SVM has high accuracy in comparison with other contemporary algorithms. In future, SVM could be utilized for grouping reason. The execution of the RSA calculation provides information assurance. Cloud security depends on reliable computing and encryption. Only the approved user

can access the data in the proposed work. However, if any intruder (unauthorized user) tries to attack the system to collect sensitive data, the original information will be secure and cannot be recovered.

Data Availability

The data used to support the findings of the study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors specially thank Muhammad Abul Hassan for his continuous support and help throughout this research study. This research was supported by the Faculty of Management of Comenius University in Bratislava, Slovakia.

References

- [1] M. L. Hoffmann Souza, C. A. da Costa, G. de Oliveira Ramos, and R. da Rosa Righi, “A survey on decision-making based on system reliability in the context of Industry 4.0,” *Journal of Manufacturing Systems*, vol. 56, pp. 133–156, 2020.
- [2] M. Abdel-Basset, G. Manogaran, A. Gamal, and V. Chang, “A novel intelligent medical decision support model based on soft computing and IoT,” *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4160–4170, 2019.
- [3] P. Mell and T. Grance, *The NIST Definition of Cloud Computing*, NIST, Maryland, United States, 2011.
- [4] M. Armbrust, A. Fox, R. Griffith et al., “A view of cloud computing,” *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.
- [5] M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, John Wiley & Sons Publishing, New Jersey, United States, 2003.
- [6] Y. Kim, “Comparison of the decision tree, artificial neural network, and linear regression methods based on the number and types of independent variables and sample size,” *Expert Systems with Applications*, vol. 34, no. 2, pp. 1227–1234, 2008.
- [7] N. Santos, K. Gummadi, and R. Rodrigues, “Towards trusted cloud computing,” *Proceedings of USENIX HotCloud*, 2009.
- [8] “Tcg specification architecture overview,” <https://www.trustedcomputinggroup.org>.
- [9] K. Bowers, A. Juels, and A. Oprea, “Proofs of retrievability: theory and implementation,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, p. 43, Chicago Illinois USA, November 2009.
- [10] W. Wang, Z. Li, R. Owens, and B. Bhargava, “Secure and efficient access to outsourced data,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, pp. 55–66, Chicago Illinois USA, November 2009.
- [11] A. Yun, C. Shi, and Y. Kim, “On protecting integrity and confidentiality of cryptographic file system for outsourced storage,” *P*, in *Proceedings of the ACM Workshop on Cloud Computing Security*, pp. 67–76, Chicago Illinois USA, November 2009.
- [12] R. Chow, P. Golle, M. Jakobsson et al., “Controlling data in the cloud: outsourcing computation without outsourcing control,”

- P, in *Proceedings of the ACM Workshop on Cloud Computing Security*, pp. 85–90, Chicago Illinois USA, November 2009.
- [13] P. Lam, E. Bursztein, and J. Mitchell, “TrackBack spam: abuse and prevention,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, Chicago Illinois USA, November 2009.
- [14] J. Sobey, T. Whalen, R. Biddle, P. V. Oorschot, and A. Patrick, “Browser interfaces and extended validation SSL certificates: an empirical study,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, Chicago Illinois USA, November 2009.
- [15] J. Wei, X. Zhang, G. Ammons, V. Bala, and P. Ning, “Managing security of virtual machine images in a cloud environment,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, pp. 91–96, Chicago Illinois USA, November 2009.
- [16] M. Christodorescu, R. Sailer, D. Schales, D. Sgandurra, and D. Zamboni, “Cloud security is not (just) virtualization security: a short paper,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, pp. 97–102, Chicago Illinois USA, November 2009.
- [17] M. Chase, K. Lauter, J. Benaloh, and E. Horvitz, “Patient Controlled Encryption: patient privacy in electronic medical records,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, p. 91, Chicago Illinois USA, November 2009.
- [18] M. Raykova, B. Vo, S. Bellovin, and T. Malkin, “Secure anonymous database search,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, Chicago Illinois USA, November 2009.
- [19] S. Liu, Yu Han, C. Miao, and C. Alex, “A fuzzy logic based reputation model against unfair ratings,” in *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-Agent Systems*, pp. 821–828, International Foundation for Autonomous Agents and Multiagent Systems, St. Paul MN USA, May 2013.
- [20] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*, Addison-Wesley Publishing, Massachusetts, United States, 2006.
- [21] I. U. Khan, A. Ryan, H. J. Alyamani et al., “RSSI-controlled long-range communication in secured IoT-enabled unmanned aerial vehicles,” *Mobile Information Systems*, vol. 2021, 2021.
- [22] D. Sisodia, Shrivastava, and Jain, “ISVM for face recognition,” in *Proceedings of the 2010 International Conference on Computational Intelligence and Communication Networks*, pp. 554–559, IEEE, Bhopal, India, 26–28 Nov. 2010.
- [23] I. Rish, “An empirical study of the naive Bayes classifier,” in *Proceedings of the IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, pp. 41–46, IBM, Vancouver, Canada, August 2001.
- [24] S. Ray, *6 Easy Steps to Learn Naive Bayes Algorithm (With Code in Python)*, 2017, <https://www.datasciencecentral.com/6-easy-steps-to-learn-naive-bayes-algorithm-with-code-in-python/>.
- [25] A. N Khan, S. U Mat Kiah, and S. A. Madani, “Towards secure mobile cloud computing: a survey,” *Future Generation Computer Systems*, vol. 29, no. 5, pp. 1278–1299, 2013.
- [26] W. Itani, A. Kayssi, and A. Chehab, “Energy-efficient incremental integrity for securing storage in mobile cloud computing,” in *Proceedings of the Int. Conference on Energy Aware Computing*, Dec. 2010.
- [27] W. Jia, H. Zhu, Z. Cao, L. Wei, and X. Lin, “SDSM: a secure data service mechanism in mobile cloud computing,” in *Proceedings of the IEEE Conference on Computer Communications Workshops*, Apr. 2011.
- [28] https://wiki.cloudsecurityalliance.org/guidance/index.php/Cloud_Computing_Architectural_Framework.
- [29] M. L. Mat Kiah, “Towards secure mobile cloud: a survey,” *Proceedings of Analyses paper*, vol. 29, p. 1278, 2012.
- [30] J. Yang, H. Wang, J. Wang, C. Tan, and D. Yu1, “Provable data possession of resource constrained mobile devices in cloud computing,” *Journal of Networks*, vol. 6, no. 7, pp. 1033–1040, 2011.
- [31] E. Eugene, *Hyrax: Cloud Computing on Mobile Devices*, *Dissertation of Thesis*, Carnegie Mellon University, Pittsburgh, 2009.
- [32] Dr. L. Arockiam and S. Monikandan, “Data security and privacy in cloud storage using hybrid symmetric encryption algorithm,” *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 8, 2013.
- [33] Y. Xu, L. Wu, L. Guo, Z. Chen, L. Yang, and Z. Shi, “An intelligent load balancing algorithm towards efficient cloud computing,” in *Proceedings of the AI for Data Center Management and Cloud Computing: Papers from the 2011 AAAI Workshop*, California, USA, August 7, 2011.
- [34] S. U. Rehman and S. Manickam, “Denial of service attack in IPv6 duplicate address detection process,” *International Journal of Advanced Computer Science and Applications*, vol. 7, pp. 232–238, 2016.
- [35] I. U. Khan, I. M. Qureshi, M. A. Aziz, T. A. Cheema, and S. B. H. Shah, “Smart IoT control-based nature inspired energy efficient routing protocol for flying ad hoc network (FANET),” *IEEE Access*, vol. 8, pp. 56371–56378, 2020.
- [36] M. Abul Hassan, S. Irfan Ullah, A. Salam, A. Wajid Ullah, M. Imad, and F. Ullah, “Energy efficient hierarchical based fish eye state routing protocol for flying ad-hoc networks,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 1, pp. 465–471, 2021.
- [37] M. A. Hassan, S. I. Ullah, I. U. Khan, S. B. Hussain Shah, A. Salam, and A. W. Ullah Khan, “Unmanned Aerial Vehicles Routing Formation using fisheye state routing for flying ad-hoc networks,” in *Proceedings of the 4th International Conference on Future Networks and Distributed Systems (ICFNDS)*, pp. 1–7, New York, NY, United States, November 2020.
- [38] D. Sisodia, L. Singh, and S. Sisodia, “Fast and accurate face recognition using SVM and DCT,” in *Advances in Intelligent Systems and Computing*, vol. 2012, pp. 1027–1038, Springer, 2014.
- [39] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Elsevier, Second Edition, 2006.
- [40] A. Ahmad and D. Ruelens, “Development of digital logic design teaching tool using MATLAB & SIMULINK,” *IEEE Technology and Engineering Education (ITEE)*, vol. 8, no. 1, pp. 7–11, 2013.