IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 9, May 2023

Analysing Health Care Data and Preserving Security using Decoy File and Fog Computing

Prof. Y. R. Chikane[1], Mr. Mahesh Patil[2], Ms. Shrushti Satpute[3], Mr. Suyog Bangar[4], Ms. Snehal Salve[5]
Professor, Department of Information Technology[1]
Student, Department of Information Technology[2,3,4,5]
Amrutvahini College of Engineering, Sangamner, India

Abstract: *Currently, humans face various diseases due to the current environmental condition and their living habits. Early identification and prediction of diseases are crucial to prevent their severity. However, doctors may find it challenging to accurately identify illnesses manually. This project aims to use predictive analysis to identify and predict patients with a higher risk of developing chronic diseases. By utilizing data analysis techniques, including machine learning and data mining, healthcare professionals can get early warnings of potential illnesses and take timely preventive measures, ultimately improving patients' overall health outcomes. Automating the diagnosis process can also help reduce healthcare costs, making healthcare more affordable and accessible. The use of data mining is crucial in predicting diseases, which is a difficult task. The proposed system employs machine learning algorithms such as the convolutional neural network (CNN) to provide a comprehensive disease prognosis based on a patient's symptoms. This system offers an automated feature extraction approach, enabling healthcare professionals to make more informed decisions. Therefore, it can contribute significantly to enhancing disease prediction and providing appropriate treatment. Storage and evaluation of big data is flexible and scalable on the cloud but at the same time there are issues in security such as data theft attacks. The problem of security can be resolved by using the technique called Fog computing. The data can be stored the security by using implementing decoy technique in fog. Decoy file are Honeypots and other false information which is used to secure the original data from an unauthorized user who is trying to access the private data. We are using decoy file to contain fake data which confusing the attacker.*

Keywords: Text detection, Morphological operations, painting, Connected component labelling

I. INTRODUCTION

Chronic conditions are persistent physical or mental health conditions that require regular treatment or monitoring, and limit daily activities. Such conditions are widespread and expensive, affecting nearly half of all individuals. Managing chronic diseases necessitates substantial medical attention. Managing chronic health conditions is critical to prevent complications and worsening of the condition. This involves following a healthy diet, engaging in regular exercise, and attending medical appointments. A healthy diet can help control blood sugar, blood pressure, and cholesterol levels. Being physically active can improve overall health and reduce the risk of further health problems. Regular medical appointments help identify changes in the condition and allow for timely intervention. By taking these necessary steps, individuals can better manage their chronic health conditions and live a healthier, more fulfilling life.Chronic diseases have impacted global health, quality of life, and healthcare costs. The Centres for Disease Control identifies them as a major driver of absenteeism and workforce patterns. These long-term health conditions, including diabetes, heart disease, and cancer, are responsible for 71% of deaths worldwide and often result from lifestyle choices, such as diet and exercise, and environmental factors. Preventive measures, like healthy lifestyle choices and access to healthcare, are crucial to reducing the prevalence of chronic illnesses and their negative impact on individuals, communities, and economies.

Big Data refers to a vast and increasingly expanding collection of data that exceeds the storage and processing capabilities of traditional data management tools. It is characterized by its large volume and complexity, making it

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 9, May 2023

difficult to manage, interpret, and derive value from. With the advancement of technology, Big Data has become an essential aspect of modern-day business and research, as it holds immense potential for generating valuable insights that can drive decision-making and innovation..Big data refers to large volumes of information generated through digital healthcare technologies, which are difficult to manage using traditional methods. This data is created through the collection of patients' records and management of hospital performance. Big data in healthcare is a critical tool used to manage and analyze massive amounts of information that would otherwise be too large and complex to handle through traditional technology. ig-style data in healthcare refers to the aggregation, analysis, and application of vast amounts of digitized information, with the goal of improving healthcare outcomes. Health data from populations as well as individuals can be leveraged to inform decisions, identify trends, and even predict disease incidences. The use of advanced analytical technologies can enable healthcare providers to provide more personalized care, improve patient outcomes, and optimize healthcare resource allocation. Overall, big-style data has the potential to revolutionize the healthcare industry by driving innovation and advancing scientific understanding.

Fog computing, also called edge computing, is intended for distributed computing where numerous "peripheral" devices connect to a cloud. (The word "fog" suggests a cloud's periphery or edge). Many of these devices will generate voluminous raw data (e.g., from sensors), and rather than forward all this data to cloud-based servers to be processed, the idea behind fog computing is to do as much processing as possible using computing units co-located with the data-generating devices, so that processed rather than raw data is forwarded, and bandwidth requirements are reduced. An additional benefit is that the processed data is most likely to be needed by the same devices that generated the data, so that by processing locally rather than remotely, the latency between input and response is minimized. This idea is not entirely new: in non-cloud-computing scenarios, special-purpose hardware (e.g., signal-processing chips performing Fast Fourier Transforms) has long been used to reduce latency and reduce the burden on a CPU. Fog networking consists of a control plane and a data plane.

Decoy data can be created as bait for attackers trying to access sensitive information and to corrupt the stolen data. Decoy documents and honey pots are effective in confusing unauthorized access attempts and reducing risks of data breaches. The use of decoys is an innovative technique to detect and prevent cyber attacks in real time, deceiving hackers into thinking they have obtained valuable data when all they have is bogus information. The strategic deployment of decoys in a network can enhance security measures, providing an essential layer of protection from malicious attacks .. When unauthorized access to a cloud service is detected, the cloud may respond with decoy information that appears normal and legitimate. This tactic alerts the true owner of the information to the security breach. By using decoy information, the cloud provider can prevent sensitive data from being compromised by unauthorized users. The decoy information is designed to appear like the real data, but is essentially a trap meant to identify and stop unauthorized access. This proactive approach to cloud security can help protect valuable data stored in the cloud from potential threats The Cloud security system can prevent unauthorized access by providing attackers with fake information. Two additional security features can also be implemented, including the validation of data and limiting data access based on user authentication and authorization. By implementing these measures, users can ensure that their data is protected from potential breaches and unauthorized access attempts. It is important to prioritize security measures when utilizing cloud services to prevent potential threats to sensitive data.

II. RELATED WORK

KR Protector: Detection and Files Protection for IoT Devices on Android Without ROOT Against Ransomware Based on Decoys

The rise of cryptographic ransomware on Android poses a significant threat as cybercriminals encrypt private data on IoT devices and demand a ransom from users. The lack of file protection solutions in IoT devices without root access makes it challenging to stop such attacks. This highlights the need for robust security measures and proper encryption protocols to safeguard user data from potential cyber threats. As digitalization continues to expand its reach, addressing these security vulnerabilities is critical to prevent financial losses and damage to reputation.

KR Recover: An Auto-Recovery Tool for Hijacked Devices and Encrypted Files by Ransomwares on Android

Android ransomwares pose a significant threat by hijacking screen resources, locking devices, and encrypting files. With the evolution of ransomwares, they now even have the power to disable USB interfaces on mobile devices,

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 9, May 2023

making it difficult to recover data. This paper investigates the relationship between ransomwares and traditional malware via Symmetric Encryption and notes how defences in anti-malware can also protect against ransomwares. However, current anti-malware solutions for mobiles are under-equipped to deal with these attacks. All mobile users should remain cautious and employ best security practices while browsing their devices.The authors propose several strategies, including recovering hijacked resources and devices as well as decrypting encrypted files, to combat the effects of these ransomwares.

Research on Privacy Data Protection in Mobile Applications

The protection of user privacy data in mobile apps has become a challenging task as revealed in the investigation of 20 mobile apps where more than 100 settings related to privacy data were discovered. These settings are not limited to "privacy settings" only but are found in "message notification" and "general" settings also. It highlights the need for app developers to prioritize user privacy and security, and for users to be vigilant in managing their app permissions. Most people (over 84%) care about their personal privacy data on smart mobile applications, according to a recent survey. However, users often struggle to find ways to protect their data, for six reasons. These findings suggest that there is a strong desire for privacy protection among the public, but more guidance is necessary to empower users to safeguard their information.

Prediction of Chronic Kidney Disease - A Machine Learning Perspective

The article discusses the reliability of machine learning techniques in medical treatment, particularly in detecting diseases on time using a machine learning classifier algorithm. It uses Chronic Kidney Disease prediction as an example and highlights the importance of timely diagnosis using data sets. The article emphasizes the potential of machine learning in improving medical technology and healthcare overall. This research utilized seven classifier algorithms, including artificial neural network, C5.0, Chi-square Automatic interaction detector, logistic regression, linear support vector machine with penalty L1 & L2, and random tree. Additionally, an important feature selection technique was employed to optimize the results. This combination of techniques and methods allowed for a comprehensive analysis and interpretation of the dataVarious feature selection methods have been applied to five classifiers in this study. The classifiers include random forest, decision tree, k-nearest neighbor, support vector machine, and logistic regression. The feature selection methods used are correlation-based feature selection, Wrapper method feature selection, Least absolute shrinkage and selection operator regression, and synthetic minority oversampling technique with least absolute shrinkage. The results show that most classifiers have better performance with feature selection methods than with full features. SVM and logistic regression perform better with correlation-based feature selection, while random forest and decision tree prefer the Wrapper method. K-nearest neighbor performs better with synthetic minority oversampling technique combined with least absolute shrinkage.

A Machine Learning Analysis of Health Records of Patients with Chronic Kidney Disease at Risk of Cardiovascular Disease

Chronic kidney disease (CKD) affects millions of people worldwide, with a long-term decline in kidney function caused by various factors. It has a significant negative impact on patients and can further worsen when combined with cardiovascular disease (CVD). Patients with both conditions have an increased risk of adverse outcomes, including hospitalization, morbidity, and mortality. It is important to raise awareness and improve management strategies for CKD and its complications to decrease the burden on patients, healthcare systems, and society at large.Machine learning was applied to medical records of patients with CKD & CVD in order to provide physicians with insights that can help improve their decision-making about prognoses or therapies. The study found that machine learning can successfully predict outcomes such as hospitalization and mortality in these patients, indicating its potential value in clinical decision-making. Overall, the use of computational intelligence in electronic health records represents a promising avenue for improving patient care and outcomes in the field of medicine.

Comprehensive Performance Assessment of Deep Learning Models in Early Prediction and Risk Identification of Chronic Kidney Disease

In summary, this paper examined clinical features of CKD and used seven different deep learning algorithms for CKD prediction and classification. The algorithms included ANN, LSTM, GRU, Bidirectional LSTM, Bidirectional GRU, MLP, and Simple RNN. The proposed models were evaluated based on their accuracy and reliability for CKD prediction and classification. The study focused on assessing the performance of a prediction model by

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 9, May 2023

measuring accuracy, precision, recall, loss, and validation loss, as well as computation time, prediction ratio, and AUC. The researchers also examined statistical significance to compare the model with others. The results indicated that the model performed well on all fronts, with high accuracy and precision, low loss and validation loss, and short computation time. These findings suggest that the model is a reliable and efficient tool for predicting outcomes. Advanced machine learning methods have demonstrated their superiority over traditional data classification techniques in healthcare. These models have shown potential value in improving healthcare outcomes through more accurate diagnoses and treatment recommendations. The ability of machine learning models to analyze large amounts of data and identify patterns and correlations can greatly enhance the decision-making process for healthcare professionals. As the field of healthcare continues to evolve, the use of advanced machine learning methods is expected to become increasingly prevalent.

A Novel Fog Node Aggregation Approach for Users in Fog Computing Environment

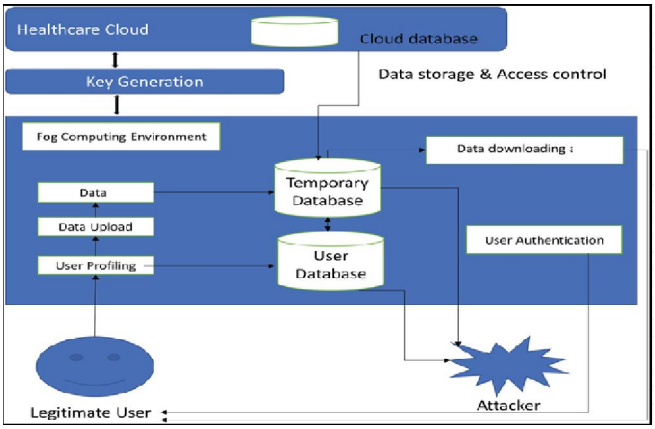
Fog computing uses geographically distributed fog nodes to offer accessible and efficient computing and services to end-users. However, a single fog node's resources are often insufficient for deploying services due to their inherent limitations. To tackle this issue, multiple fog nodes are used to pool their resources together and enable the deployment of advanced services. This approach ensures optimal service delivery and response times to the end-users The proposed paper introduces a new approach to aggregate resources of fog nodes. The approach involves multiple fog nodes being considered as a single resource node for hosting virtual machines and deploying services. The process is split into two phases: seeking suitable fog nodes and formulating cooperation strategies between them. By implementing this approach, fog computing can become more efficient and provide better services for users. We propose a two-stage approach to optimizing Fog computing resource allocation. Firstly, we consider resource allocation across different services, using a multi-objective evolutionary algorithm. Secondly, we propose a Fog node Selection algorithm based on Simulated Annealing to meet the resource requirements of an individual service while minimizing user overhead. These approaches can help to achieve efficient and sustainable utilization of Fog computing resources while ensuring that individual services are adequately supported.

Edge server placement in mobile edge computing

Mobile edge computing is being used to decrease latency by diverting some workload from mobile devices to local edge servers with adequate Mobile edge computing is being used to decrease latency by diverting some workload from mobile devices to local edge servers with adequate computing capabilities. computing capabilities.In order to address the problem of edge server placement in mobile edge computing for smart cities, a multi-objective constraint optimization problem has been formulated. The goal of this approach is to effectively place edge servers in a way that meets multiple criteria and objectives. By doing so, it is possible to optimize the functionality and efficiency of smart city systems and ensure that they are able to provide the services and capabilities required to meet the needs of citizens and stakeholders The utilization of this approach may result in the development of optimized and efficient mobile edge computing solutions for smart cities.

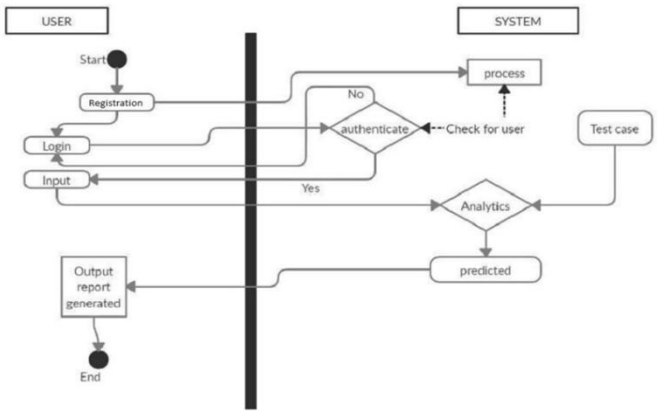
III. METHODOLOGY

System Architecture



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 9, May 2023

System Flow:



Methodology

for Prediction

Training data is crucial for machine learning algorithms as it helps them learn patterns and make predictions. The data scientist feeds the algorithm with a large amount of diverse and relevant data, and the algorithm extracts useful insights from it. During the training process, the algorithm learns to recognize these patterns and uses them to make predictions about unseen data. The quality and quantity of training data are essential for the success of machine learning models, and require careful consideration and construction. With appropriate training data, machine learning algorithms can deliver powerful and accurate results.

For Security

Decoy File: The decoy file is a resource that can be accessed locally or through a network, and may be monitored. This file can come in various forms, such as configurations, documents, executables, or other file formats, and serves a specific purpose. The goal of using a decoy file is to divert attackers away from sensitive data by tricking them into thinking they have found what they are looking for. By monitoring access to the decoy file, organizations can gain insight into potential threats and take appropriate action.

Algorithm

Algorithm SVM:

Input: D dataset, on-demand features, aggregation-based features,
Output: Classification of Application
for each application App-id in D do
Get on-demand features and stored on vector x for App-id
x.add (Get-Features(app-id));
end for
for each application in x vector do
Fetch first feature and stored in b, and other features in w.
hw,b (x) = g (z) here z= (wT x + b)
if (z 0)
assign g(z)=1;
else g(z)=-1;
end if
end

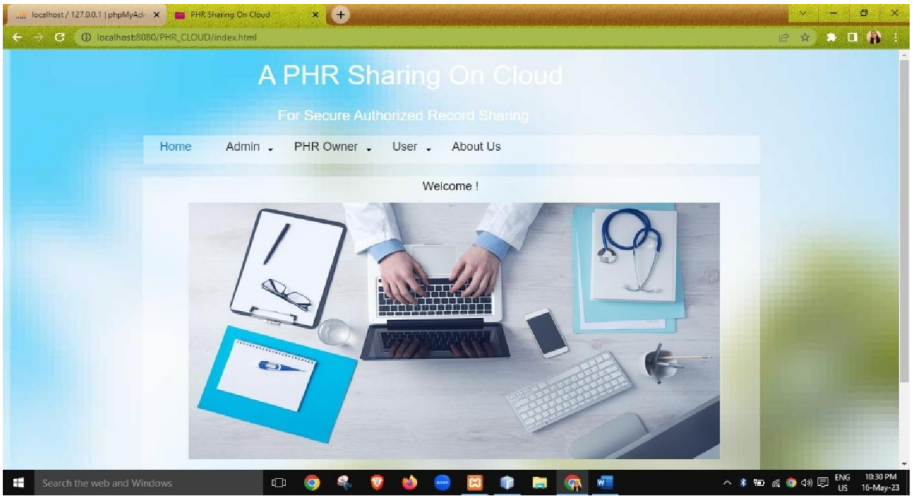
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 9, May 2023

For Decision Tree

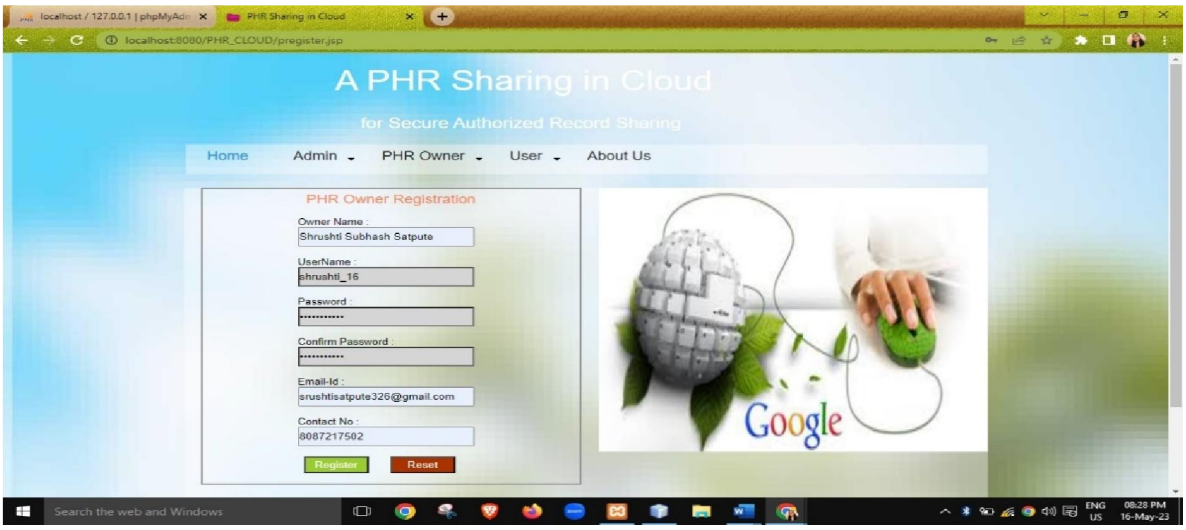
1. Check if algorithm satisfies termination criteria
2. Computer information-theoretic criteria for all attributes
3. Choose best attribute according to the information-theoretic criteria
4. Create a decision node based on the best attribute in step
5. Induce (i.e., split) the dataset based on newly created decision node in step 4
6. For all sub-dataset in step 5, call C4.5 algorithm to get a sub-tree (recursive call)
7. Attach the tree obtained in step 6 to the decision node in step 4
8. Return tree

IV. EXPERIMENTAL RESULTS

User Registration : This figure shows the registration window view of project. User can register on the system. While registering user must enter their name, username, mobile number, e-mail id, password. "Registration Successful" message is displayed on the screen after registration.



Admin Login : This figure shows the admin login window view of project. admin can login on the system. While login admin must enter correct user name and password.



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 9, May 2023

V. CONCLUSION

The decoy file is a resource that can be accessed locally or through a network, and may be monitored. This file can come in various forms, such as configurations, documents, executables, or other file formats, and serves a specific purpose. The goal of using a decoy file is to divert attackers away from sensitive data by tricking them into thinking they have found what they are looking for. By monitoring access to the decoy file, organizations can gain insight into potential threats and take appropriate action.

REFERENCES

[1]. S. Wang et al, "KRProtector: Detection and Files Protection for IoT Devices on Android Without ROOT Against Ransomware Based on Decoys", in IEEE Internet of Things Journal, vol. 9, no. 19, pp. 18251-18266, 1 Oct.1, 2022, doi: 10.1109/JIOT.2022.3156571.

[2]. Wang, Senmiao, Sujuan Qin, Nengqiang He, Tengfei Tu, Junjie Hou, Hua Zhang, and Yijie Shi. 2021. "KRRecover: An Autoecovery Tool for Hijacked Devices and Encrypted Files by Ransomwares on Android" Symmetry 13, no. 5: 861. https://doi.org/10.3390/sym13050861

[3]. H. Chen, Y. Gu, P. Wang, J. Dong and Y. Ren, "Research on Privacy Data Protection in Mobile Applications", 2021 33rd Chinese Control and Decision Conference (CCDC),2021, pp. 4912-4915, doi: 10.1109/CCDC52312.2021.9602169.

[4]. P. Chittora et al., "Prediction of Chronic Kidney Disease - A Machine Learning Perspective", in IEEE Access, vol. 9, pp. 17312-17334, 2021, doi: 10.1109/ACCESS.2021.3053763.

[5]. D. Chicco, C. A. Lovejoy and L. Oneto, "A Machine Learning Analysis of Health Records of Patients with Chronic Kidney Disease at Risk of Cardiovascular Disease", in IEEE Access, vol. 9, pp. 165132-165144, 2021,doi:10.1109/ACCESS.2021.3133700