INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 8, Issue 1 - V8I1-1248)
Available online at: https://www.ijariit.com

Heart Disease Prediction

Medasani Hari Kumar
harikumar.medasani@gmail.com
Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh

Kottamasu Sai Anila
k.saianila2000@gmail.com
Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh

Doppalapudi Sriram
sriramdoppalapudi7@gmail.com
Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh

Rizwan Patan
rizwan@vrsiddhartha.ac.in
Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh

ABSTRACT

Heart Disease prediction is one of the most complicated tasks in medical field. Day by day the cases of heart diseases are increasing at a rapid rate and it's very Important and concerning to predict any such diseases beforehand. This diagnosis is a difficult task i.e. it should be performed precisely and efficiently. We prepared a heart disease prediction system to predict whether the patient is likely to be diagnosed with a heart disease or not using the medical history of the patient. Traditional approaches that involve collection of data from devices into one centralized repository for further analysis are not always applicable due to the large amount of collected data,the use of communication channels with limited bandwidth, security and privacy requirements, etc. Federated learning (FL) is an emerging approach that allows one to analyse data directly on data sources and to federate the results of each analysis to yield a result as traditional centralized data processing. Available electronic medical records of patients quantify symptoms, body features, and clinical laboratory test values, which can be used in predicting heart disease.

Keywords: *Federated Learning, Machine Learning, Privacy, Prediction, Heart Disease.*

1. PROBLEM STATEMENT

Medical services is one of the greatest space, which needs combined learning for getting and handling information. Customary methodologies that include assortment of information from gadgets into one concentrated vault for additional investigation are not appropriate due all the time to the enormous measure of gathered information, the utilization of correspondence channels with restricted data transfer capacity, security and protection necessities. Herewe are utilizing AI calculation to accomplish protection from digital attacks.We proposed, involving AI calculation which give exactness in anticipating coronary illness

2. LITERATURE SURVEY

Federated Learning of predictive models from federated Electronic Health Records [1]: The focus of this paper is to develop a distributed (federated) method to predict hospitalizations during a targetyear for patients with heart diseases, based on their medical history as described in their Electronic Health Records (EHRs).In an era of "big data," computationally efficient and privacy-aware solutions for large-scale machine learning problems become crucial, especially in the healthcare domain, where large amounts of data are stored indifferent locations and owned by different entities. Past research has been focused on centralized algorithms, which assume the existence of a central data repository (database) which stores and can process the data from allparticipants. Such an architecture, however, can be impractical when data are not centrally located, it does not scale well to very large datasets, and introduces single-point of failure risks which could compromise the integrityand privacy of the data. Given scores of data widely spread across hospitals/individuals, a decentralized computationally scalable methodology is very much in need.They focus on the soft-margin l1-regularized sparseSupport Vector Machine (sSVM) classifier. They develop an iterative cluster Primal Dual Splitting (cPDS) algorithm for solving the large-scale sSVM problem in a decentralized fashion.

Reliable customer analysis using federated learning and exploring deep-attention edge intelligence [2]: The Internet of Things (IoT) and smart cities are flourishing with distributed systems in mobile wireless networks.As a result, an enormous amount of data are being generated for devices at the network edge. This results inprivacy concerns, sensor data management issues, and data utilization issues. In this research, we propose acollaborative clustering method where the exchange of raw data is not required. The attention-based model usedwith a federated learning framework. The edge devices compute the model updates using local

data and sendthem to the server for aggregation. Repetition is performed in multiple rounds until a convergence point reached.They propose an attention-based federated learning approach for customer analysis at the edge of mobilenetworks. For transaction data, the model learns low dimensional representation (embedding) in a fullyunsupervised manner.

A Review of application in federated learning[3]: Li Li Yuxi Fan Mike Tse Kuo-Yi Lin has given a study which review Federated learning and explores the main evolution path for issues exist in FL development process to advance the understanding of FL. They did it with an aim to review prevailing application in industrial engineering to guide for the future landing application. Thisstudy also identifies six research fronts to address FL literature and help advance our understanding of FL for future optimization. This study contributes to conclude application in industrial engineering and computer scienceand summarize a review of applications in FL.

A contemplative perspective on federated machine learning: Taxonomy, threats & vulnerability assessment and challenges[4]: Divya Jatain Vikram Singh Naveen Dahiya has given a study that Federated Learning's advantages over the traditional methods and/or its classification. They have also intended to address the totality of federated learningwith a complete vulnerability assessment. During the study of the literature, it is found that security being promised as one of the key advantages of federated learning can still not be guaranteed because of some issues inherently present, and this can lead to poisoning, inference attacks and insertion of backdoors, etc. This paper intends to provide a complete picture by giving an in-depth and comprehensive analysis of Federated Learning and its taxonomy. It also provides a detailed vulnerability assessment and highlights the challenges faced in the current setting and future research directions to make federated learning a more functional, robust and secure method to train machine learning models.

Federated Learning for Malware Detection in IoT Devices[5]: Valerian Rey and Martin Jaggi investigated the possibilities enabled by federated learning concerning IoT malware detection and studied security issues inherent to this new learning paradigm. They proposed a framework that uses federated learning to detect malware affecting IoT devices. It was N-BaIoT, a dataset modeling network traffic of several real IoT devices while affected by malware, has been used to evaluate the proposed framework. This Framework uses FL to detect, in a privacy preserving fashion, cyberattacks affecting IoT devices. The proposed framework covers both anomaly detection and classification approaches using multi-layer perceptron and auto encoder architectures. Traditional AI techniques have been widely applied in the literature to detect cybersecurity issues in IoT scenarios.

The future of digital health with federated learning[6]: Nicola Rieke and Micah Sheller has given a study which review implementation of Federated Learning could holda significant potential for enabling precision medicine at large-scale, leading to models that yield unbiased decisions,optimally reflect an individual's physiology, and are sensitive to rare diseases while respecting governance and privacyconcerns. However,They said that Federated Learning still requires rigorous technical consideration to ensure that the algorithm is proceeding optimally without compromising safety or patient privacy. Nevertheless, it has the potential to overcome the limitations of approaches that require a single pool of centralised data. They Shared their view of providing context and details for the community regarding the benefits and impact of Federated Learning formedical applications, as well as key considerations and challenges of implementing Federated Learning for digital health.

3. METHODOLOGY

The Methodology of the project disease prediction using machine learning consist of all the various aspects a normal flow diagram requires. This Lifecycle diagram shows how from the model starts its flows from one step to another, at first the dataset has been collected and to which some data pre-processing steps have been performed. Then the model is build according to the algorithms used and have checked the results, for better accuracy the model is again optimized for better results, Now the finalized model is sent to the server. Software design is a process to transform user requirements into some suitable form, which helps the programmer in software coding and implementation. A project is short-term for the reason that it has an identified commencingand finalizing over time, and so consequently outlined extent and solutions. Likewise, a project workforce frequently involves professionals who do not generally perform collectively occasionally from distinct agencies and throughout diverse geographies. Scrum is a verified and as well, broadly implemented approach for attainingapplication agility. Here, whenever the project terminates Scrum assures the majority of important work is finished**.**

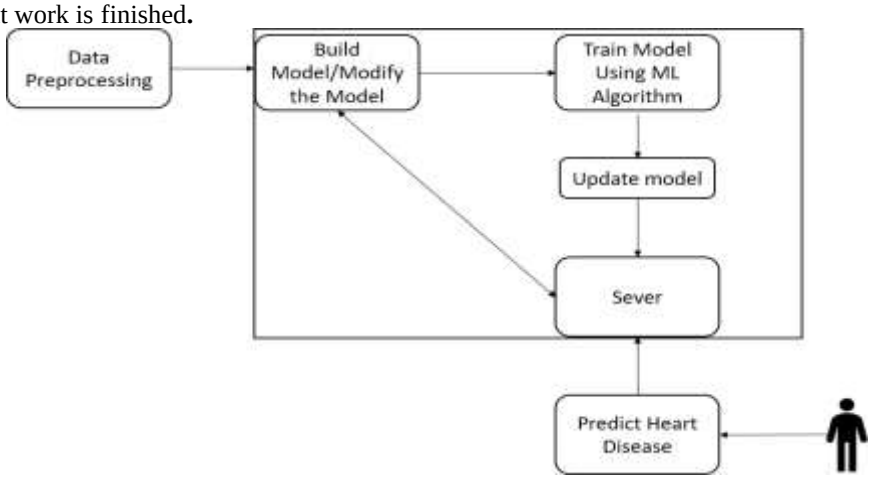

Chart -1 : Methodology

Process Flow Diagram

The proposed work predicts heart disease by exploring the classification algorithms and does performance analysis. The objective

of this study is to effectively predict if the patient suffers from heart disease. The health professional enters the input values from the patient's health report. The data is fed into model which predicts the probability of having heart disease. Fig. 2 shows the entire life cycle process involved.

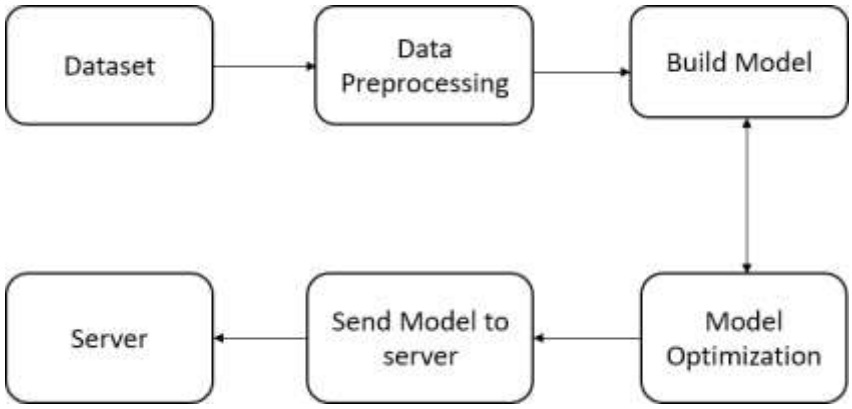

Chart 2: Process flow diagram

4. ALGORITHM

Logistic regression is named for the function used at the core of the method, the logistic function. The logistic function, also called the sigmoid function was developed by statisticians to describe properties of population growth in ecology, rising quickly and maxing out at the carrying capacity of the environment. It's an S-shaped curve that can take any real-valued number and map it into a value between 0 and 1, but never exactly at those limits.

1 / (1 + e^-value)

Where e is the base of the natural logarithms (Euler's number or the EXP() function in your spreadsheet) and value isthe actual numerical value that you want to transform.

5. CONCLUSION AND RESULTS

```
[71] LR_model = LogisticRegression() # Since Basic accuracy outcome gives the best model accuracy results, we will implement it
     LR_model.fit(X_train_scaled, y_train)
     y_pred = LR_model.predict(X_test_scaled)
     y_train_pred = LR_model.predict(X_train_scaled)

     log_f1 = f1_score(y_test, y_pred)
     log_acc = accuracy_score(y_test, y_pred)
     log_recall = recall_score(y_test, y_pred)
     log_auc = roc_auc_score(y_test, y_pred)

     print(confusion_matrix(y_test, y_pred))
     print("\033[1m--------------------------------------------------------\033[0m")
     print(classification_report(y_test, y_pred))
     print("\033[1m--------------------------------------------------------\033[0m")

     plot_confusion_matrix(LR_model, X_test_scaled, y_test)

     train_val(y_train, y_train_pred, y_test, y_pred)

     [[51 11]
      [ 6 70]]
```

Chart 3: Description of Logistic Regression accuracy

	train_set	test_set
Accuracy	0.862	0.877
Precision	0.867	0.864
Recall	0.887	0.921
f1	0.876	0.892

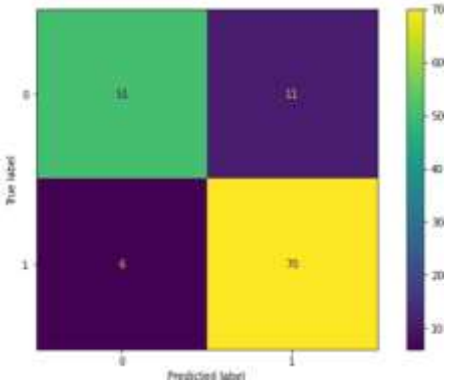

Chart 4*:* Description about Accuracy

Sex_M	ChestPainType_ATA	ChestPainType_NAP	ChestPainType_TA	RestingECG_Normal	RestingECG_ST	ExerciseAngina_Y	ST_Slope_Flat	ST_Slope_Up	HeartDisease	pred	pred_proba
1	0	0	0	0	1	1	0	0	1	1	0.896
0	0	0	0	0	0	0	1	0	0	1	0.571
1	0	0	0	1	0	1	1	0	1	1	0.969
1	0	1	0	0	1	0	1	0	0	1	0.909
1	1	0	0	1	0	0	0	1	0	0	0.049
1	0	1	0	1	0	0	1	0	1	1	0.647
0	0	0	1	1	0	0	1	0	0	0	0.289
1	0	0	0	1	0	0	1	0	1	1	0.861
1	0	0	0	0	0	1	1	0	1	1	0.949
1	0	0	0	0	1	0	0	1	1	1	0.746

Chart 5: Predicted Output Of Logistic Regression

This project helps people to predict the possibility of being diagnosed by heart disease. We performed preprocessingsteps on the dataset and applied logistic regression to get an accuracy of 87.7%. Previous models like ANN and Decision Tree having an accuracy of 87% and 84.8% respectively. Probing deeper, the results in this project also provide a strong foundation for future work in creating awareness. The future course of this research can be performedwith diverse mixtures of machine learning techniques to better prediction techniques. Furthermore, new feature selection methods can be developed to get a broader perception of the significant features to increase the performanceof heart disease prediction. The future of this project is to use a decentralised approach which allows devices to learna shared prediction model collaboratively while maintaining the training data on the computer and uploading and storing it on a central server to increase the accuracy and reliability to better validate the results and predict the best outcome possible.

6. REFERENCES

[1] Brisimi, Theodora S., Ruidi Chen, Theofanie Mela, Alex Olshevsky, Ioannis Ch Paschalidis, and Wei Shi. "Federated learning of predictive models from federated electronic health records." International journal of medical informatics 112 (2018): 59-67.

[2] Ahmed, Usman, Gautam Srivastava, and Jerry Chun-Wei Lin. "Reliable customer analysis using federated learning and exploring deep-attention edge intelligence." Future Generation Computer Systems 127 (2022): 70-79

[3] Li, Li, Yuxi Fan, Mike Tse, and Kuo-Yi Lin. "A review of applications in federated learning." Computers &Industrial Engineering (2020): 106854.

[4] Jatain, Divya, Vikram Singh, and Naveen Dahiya. "A Contemplative Perspective on Federated Machine Learning: Taxonomy, Threats & Vulnerability Assessment and Challenges." Journal of King Saud University-Computer and Information Sciences (2021).

[5] Kholod, Ivan, Evgeny Yanaki, Dmitry Fomichev, Evgeniy Shalugin, Evgenia Novikova, Evgeny Filippov, andMats Nordlund. "Open-Source Federated Learning Frameworks for IoT: A Comparative Review and Analysis."

[6] Sensors 21, no. 1 (2021): 167

[7] Du, Zhaoyang, Celimuge Wu, Tsutomu Yoshinaga, Kok-Lim Alvin Yau, Yusheng Ji, and Jie Li. "Federatedlearning for vehicular internet of things: Recent advances and open issues." IEEE Open Journal of the ComputerSociety 1 (2020): 45-61

[8] Rey, Valerian, Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, and Martin Jaggi."Federated learning for malware detection in iot devices." arXiv preprint arXiv:2104.09994 (2021).

[9] Rieke, Nicola, Jonny Hancox, Wenqi Li, Fausto Milletari, Holger R. Roth, Shadi Albarqouni, Spyridon Bakas etal. "The future of digital health with federated learning." NPJ digital medicine 3, no. 1 (2020): 1-7.

[10] Xu, Jie, Benjamin S. Glicksberg, Chang Su, Peter Walker, Jiang Bian, and Fei Wang. "Federated learning forhealthcare informatics." Journal of Healthcare Informatics Research 5, no. 1 (2021): 1-19.

[11] Theodora S., Ruidi Chen, Theofanie Mela, Alex Olshevsky, Ioannis Ch Paschalidis, and Wei Shi. "Federated learning of predictive models from federated electronic health records." International journal of medical informatics 112 (2018): 59-67.

[12] Ahmed, Usman, Gautam Srivastava, and Jerry Chun-Wei Lin. "Reliable customer analysis using federated learningand exploring deep-attention edge intelligence." Future Generation Computer Systems 127 (2022): 70-79

[13] Li, Li, Yuxi Fan, Mike Tse, and Kuo-Yi Lin. "A review of applications in federated learning." Computers & Industrial Engineering (2020): 106854.

[14] Jatain, Divya, Vikram Singh, and Naveen Dahiya. "A Contemplative Perspective on Federated Machine Learning:Taxonomy, Threats & Vulnerability Assessment and Challenges." Journal of King Saud University-Computer and Information Sciences (2021).

[15] Kholod, Ivan, Evgeny Yanaki, Dmitry Fomichev, Evgeniy Shalugin, Evgenia Novikova, Evgeny Filippov, and Mats Nordlund. "Open-Source Federated Learning Frameworks for IoT: A Comparative Review and Analysis."

[16] Sensors 21, no. 1 (2021): 167

[17] Du, Zhaoyang, Celimuge Wu, Tsutomu Yoshinaga, Kok-Lim Alvin Yau, Yusheng Ji, and Jie Li. "Federated learning for vehicular internet of things: Recent advances and open issues." IEEE Open Journal of the Computer Society 1 (2020): 45-61

[18] Rey, Valerian, Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, and Martin Jaggi. "Federated learning for malware detection in iot devices." arXiv preprint arXiv:2104.09994 (2021).

[19] Rieke, Nicola, Jonny Hancox, Wenqi Li, Fausto Milletari, Holger R. Roth, Shadi Albarqouni, Spyridon Bakas et al. "The future of digital health with federated learning." NPJ digital medicine 3, no. 1 (2020): 1-7.

[20] Xu, Jie, Benjamin S. Glicksberg, Chang Su, Peter Walker, Jiang Bian, and Fei Wang. "Federated learning for healthcare informatics." Journal of Healthcare Informatics Research 5, no. 1 (2021): 1-19.

