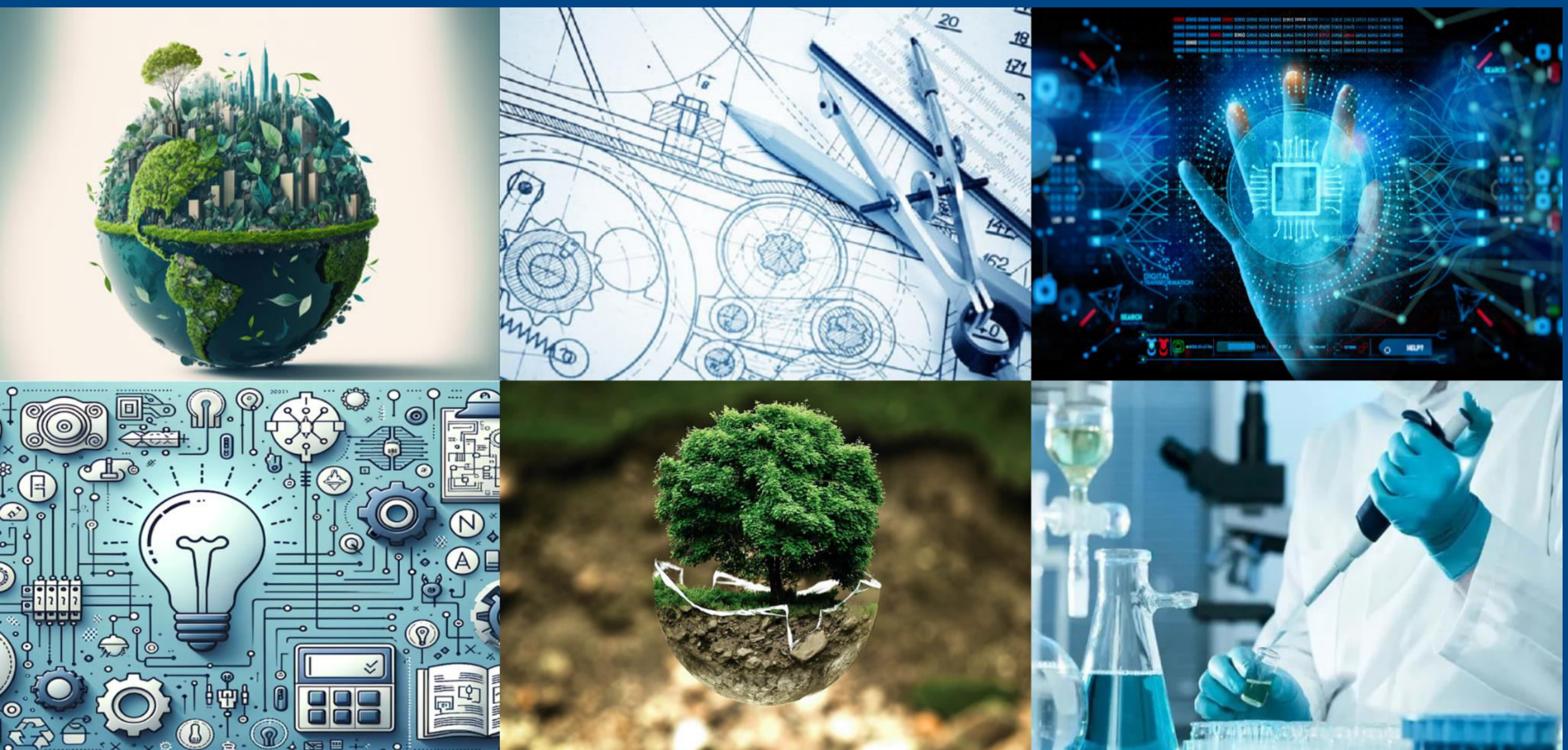
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 3, March 2025

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803038

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Heart Disease Prediction using Machine Learning Strategies

Dr.T.Subba Reddy[1], Y.Dora Babu Naidu[2], P.Gopi[3], SK.Sameer[4], P.Mahesh[5]

Associate Professor, Dept. of CSE-Data Science, KKR & KSR Institute of Technology and Sciences, Guntur, A.P, India[1]

Student (B. Tech), Dept. of CSE-Data Science, KKR & KSR Institute of Technology and Sciences, Guntur, A.P, India[2,3,4,5]

ABSTRACT: Heart disease plays a crucial role in every living creature. Heart disease is one of the most significant causes of the mortality in the world. As we all know that predicting the cardio vascular diseases is not an easy challenge in the aspect of clinical data analysis. Machine learning (ML) algorithms are being extensively used for the prediction of heart disease based on different parameters and data sets of health. These algorithms try to detect patterns and enhance the correctness of diagnosis, thus eventually leading to early intervention and improved patient outcomes. This work examines machine learning methods for forecasting heart disease based on the evaluation of a wide range of clinical features. We used a range of algorithms such as k-nearest neighbors (KNN), support vector machines (SVM), and deep learning (DL) models to evaluate their respective predictive potential using the Framingham Heart Study dataset. Later, the work applied methods like hyper parameter tuning and cross-validation to enhance model performance.

Our observations indicate that the artificial neural network (ANN) model had the maximum accuracy of 90% and was effective in identifying heart disease risk individuals. This work highlights the value of ML in transforming cardiovascular risk assessment and encouraging proactive healthcare approaches.ML prediction of heart disease has become increasingly popular as a tool to improve clinical decision-making and patient outcomes. This work seeks to create a predictive model of heart disease risk based on a mix of demographic, clinical, and lifestyle information.

KEYWORDS: Heart Study Dataset, Prediction, Machine Learning, Patterns, Hyper parameter Tuning, Cross Validations.

I. INTRODUCTION

Conventional diagnostic procedures, though useful, tend to be dependent upon subjective judgments and are also subject to limitations of access to specialized medical professionals. Thus, increasing attention is being sought through the use of sophisticated technologies, with a focus on machine learning (ML), to improve heart disease risk prediction. With the recent emergence of machine learning (ML), there are new prospects for improving the predictive power of healthcare systems.

Heart disease is one of the foremost global health issues, causing about 32% of deaths every year worldwide. Early prediction and diagnosis of heart diseases are critical to safe and effective treatment and management. Even with the development of medical technology, misinterpretation, and misdiagnosis of test results by medical practitioners are still possible. Cardiovascular diseases (CVDs) are a major public health threat, and therefore there is an urgent need to develop sophisticated predictive analytics to improve early diagnosis.

ML, a subset of artificial intelligence (AI), is the practice of creating algorithms that can learn patterns from data, by recognizing the patterns, and make predictions without being specially programmed. Accurate heart disease prediction demands a sophisticated approach capable of deciphering the complex interplay among numerous risk factors, a task beyond the scope of traditional methods. ML algorithms reveal hidden connections within extremely big datasets comprising demographic, clinical, and lifestyle details that inform clinical decision-making.Numerous studies demonstrate the effectiveness of various ML algorithms in predicting heart disease. These studies have successfully

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

applied methods such as logistic regression (LR), decision trees (DT), random forests (RF), SVM, and neural networks (NN) to diverse datasets, including the Framingham Heart Study, the Cleveland Heart Disease dataset, and electronic health records.

Table 1: Clinical Parameters for Heart Disease Prediction

Clinical Parameters	
Age	Blood Pressure
Gender	Cholesterol
Sugar	Thalassemia

Looking to the future, the application of ML in clinical processes has both benefits and challenges. Cooperative collaboration between clinicians, policymakers, and data scientists is required to make machine learning models clinically applicable and ethical in addition to being accurate. Future studies ought to aim at creating strong models that are easily interpretable and can be readily incorporated into the current healthcare system, as well as the ethics of applying AI in patient treatment.

Many studies have shown the predictive power of numerous machine learning models for heart disease. Examples of decision trees and random forests are considered more interpretable, accommodating nonlinearities between variables, whereas logistic regression is being too widely used to solve binary classification problems.

SVM have proved to be promising in high-dimensional spaces, and deep learning methods, especially neural networks, have been popular due to their capability to learn sophisticated patterns in big data. The models have been tested using several performance measures such as accuracy, precision, recall, and area under the receiver operating characteristic curve (AUC-ROC), which have turned out to be better than classical statistical approaches. In spite of the encouraging progress, a number of challenges limit the extensive application of machine learning in clinical practice.

II. METHODOLOGY

The proposed method of our Heart Disease Risk Prediction System is based on some clinical parameters and previous medical reports and is thus aimed at detecting fatigue.

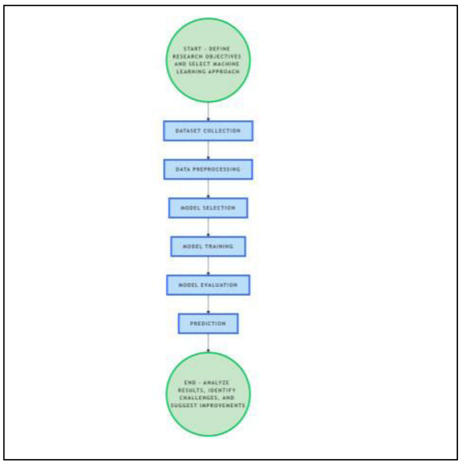


Figure 1: Flow of the Proposed Work

1. Heart Disease Prediction using clinical parameters
There are several clinical parameters as below- The sugar levels in the fasting aspect must be below the 100mg/dL and in the 2-Hour Postprandial aspect has to be below 140mg/dL. Coming to blood pressure levels, it should be below 120/80 mmHg. The total cholesterol has to be under 200 mg/dL, low-density lipoprotein (LDL) is less than 100 mg/dL and high-density lipoprotein (HDL) is must be 60 mg/dL or higher.

2. Random Forest Model for accuracy
To make it accurate in detecting the heart disease, the random forest algorithm combines multiple decision trees to improve predictive accuracy and control overfitting. It is mostly useful for the classification tasks, such as prediction of the heart disease, because it can handle the large datasets also not having more noise and outliers.

3. Real-time implementation
Electronic Health Records (EHR) will be integrated with the system to extract real-time patient data for tracking their health status and ensuring a smooth implementation for real-time healthcare. Mobile apps will be able to use health data to provide real-time risk assessments, thereby improving preventative treatment and early detection of diseases. Machine learning is key in modern healthcare since it enhances the accuracy and reliability of predictions by allowing constant update and retraining of the model with new patient data.

4. A Schematic Work flow of the System
For the purpose of machine learning-based prediction for heart disease, the research methodology comprises a workflow divided into structural steps. These consist of:

- Acquisition of Data - This constitutes gathering data about patients either from hospital records or from publicly available datasets.
- Preprocessing of the Data - Cleaning data, its handling of missing values, encoding of categorical variables and feature scaling.
- Exploratory Data Analysis (EDA) - Understanding trends, correlations and visualizing of relevant features.
- Feature Engineering - Selection and transformations of features that contribute most to prediction.
- Model Selection & Training - Using various ML algorithms on predictive models.
- Model Evaluation - Use various statistical values such as accuracy, precision, recall, and F1 score to compare and contrast which model appeared to perform the best.
- Deployment - Deploy the model into a real-time application or system of health-care.
- Monitoring and Improvement - Continuous Check and Retraining.

III. RESULTSANDDISCUSSIONS

1. Performance Evaluation
Primary performance metrics, including as accuracy, precision, recall, and F1-score, were used to evaluate the trained models. These criteria, which are mostly based on each model's accuracy level as a success metric, provide for an initial assessment of how well each model predicts cardiac disease.

2. Model Comparison
Different ML models exhibited varying levels of predictive accuracy. The results indicated that some models performed better than others in detecting heart disease. This section highlights the strengths and weaknesses of each model based on the evaluation results.

3. Challenges and Limitations
Yet, some challenges were faced during the study, including data imbalance, high computational cost, and overfitting. This section will explain how these challenges affected the study and some ideas for dealing with them going forward in future research.

4. Best Performing Model
Of all the trained models, this one stood out as the best with regard to precision and recall. This section examines the ways and causes for the outperformance of this model over others-the features of importance in deciding, hyper

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

parameter considerations, and attributes attributable to the specific dataset.

IV. THEORY AND CONCLUSIONS

Machine learning techniques for medical diagnosis typically follow a structured recipe involving data preprocessing, model selection, training, and evaluation. The primary philosophy behind these predictive models is recognition of the pattern in the medical databases: classifying patient records on the basis of risk factors. Supervised learning algorithms such as logistic regression and decision trees are commonly used in clinical predictions, as these allow clear decision-making processes.

This class of models is often generalized with such advanced techniques as ensemble learning and deep learning to enhance generalization capacity and predictive accuracy by optimizing complex feature interactions made possible by large amounts of data.

Feature selection is significant in determining the performance of a model for the reason that irrelevant and redundant variables may decrease accuracy. Union scaling methods and one-hot encoding allow the models to learn optimally without biases from differing distributions of data. Additionally, methods for dealing with class imbalance are utilized, and one such is SMOTE, which enhances credibility by ensuring well-representation of minority class samples during the training phase. This methodological consideration produces a chance to make machine learning a useful system to adopt in medical diagnosis and decision-support systems.

Advanced AI integration: Integration with machine learning APIs and cloud-based data could improve specificity and personalization.

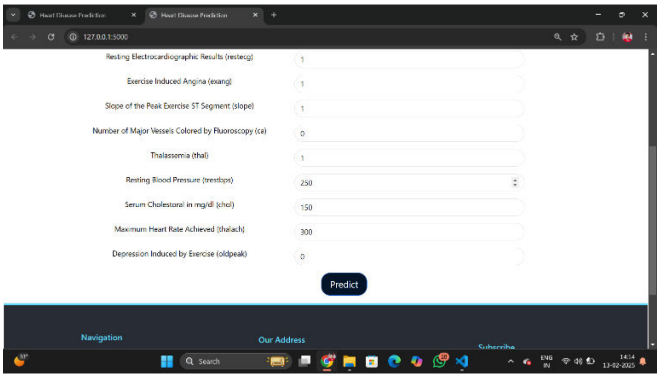


Figure 2: Input Screen

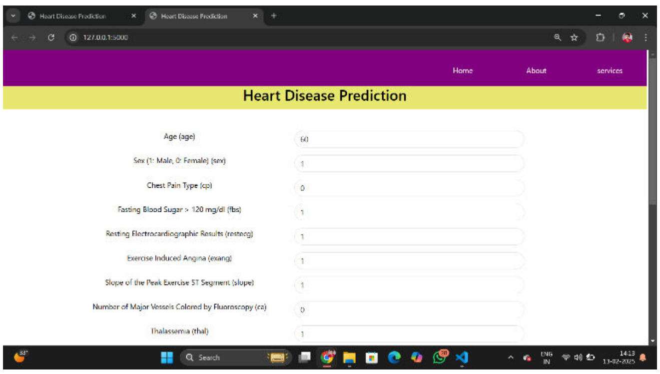


Figure 3: Output Screen

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Furthermore, the study indicates that ML-based heart disease prediction models may assist in the early diagnosis and thus preventative measures for healthcare practitioners.The conclusion is that decisions made based on data could lead to better care for patients and better management of hospital workflow. Insights from artificial intelligence concerning healthcare could allow hospitals to provide personalized treatment programs to patients and facilitate access to healthcare services.

V. CONCLUSIONS AND FUTUREWORK

The study has successfully implemented various predictive machine-learning models which help in predicting potential heart ailments. The results suggest that some models outperform others in terms of accuracy and performance attributes. In the near future, attempts will be made to apply deep learning approaches, conduct dynamic increase of datasets, and apply online patient data in real-time monitoring and prediction. In addition, placing these models in clinical applications will fast-track diagnoses and enhance patient care outcomes. For further research, feature engineering could be refined and explainable AI approaches could be added to strengthen interpretability for medical decision-making.

This study shows that machine learning can be used to predict heart disease, and the results show that some models are more accurate and reliable than others. Future research will involve integrating deep learning techniques, expanding the dataset, and incorporating real-time patient data for continuous monitoring and prediction. Deploying these models in clinical settings will improve patient care outcomes and early diagnosis. Additional research may concentrate on improving feature engineering techniques and integrating explainable AI approaches to improve interpretability in medical decision-making.

ACKNOWLEDGMENTS

We sincerely thank Dr.B.Bhanu Prakash Sir (Head of the Department),Dr.T.Subbareddy and all the trainers for their unwavering support, motivation, direction and invaluable during this endeavor. We would especially want to thank for their unwavering support, perceptive remarks, and inspirational contributions, all of which have been crucial in helping to frame our project. Their support and belief in our work have been essential to our success. We really appreciate their unwavering support, which was crucial to the accomplishment of this study.

REFERENCES

1. G. R. Finnie and G.E. Wittig, (1997). A Comparison of Software Effort Estimation Techniques: Using Function Points with Neural Networks, Case Based Reasoning and Regression Models. Journal of Systems and Software, vol.39, pp.281-289.
2. M. O. Elish, (2009). Improved Estimation of Software Project Effort using Multiple Additive Regression Tree. Expert Systems with Applications, vol.36, pp. 10774-10778.
3. K. Srinivasan and D. Fisher, (1995). Machine Learning Approaches to Estimating Software Development Effort. IEEE Transactions on Software Engineering, vol.21.
4. Parag C. Pendharkar, (2010). Probabilistic estimation of software size and effort. An International Journal of Expert Systems with Applications, vol. 37, pp.4435-4440.
5. B.L DeekshatuluaPriti Chandra “Classification of Heart Disease Using K- Nearest Neighbor and Genetic
6. Michael W.Berryet.al, Lecture notes in data mining, World Scientific(2006)
7. C.-L. Chang and C.-H. Chen, “Applying decision tree and neural network to increase quality of dermatologic
8. C.-L. Chang and C.-H. Chen, “Applying decision tree and neural network to increase quality of dermatologic
9. T. Azar and S. M. El-Metwally, “Decision tree classifiers for automated medical diagnosis,” Neural Comput.
10. B.L Deekshatulua Priti Chandra “Reader, PG Dept. Of Computer Application North Orissa University,
11. E. Hassanien and T. Kim, “Breast cancer MRI diagnosis approach using support vector machine and pulse
12. Dastjerdi, M., & Mollahosseini, A. (2020). Machine learning for heart disease prediction: A review. Journal of Healthcare Engineering, 2020.
13. Kourou, K., et al. (2015). Machine learning techniques for heart disease prediction: A review. Computational and Structural Biotechnology Journal, 13, 104-112.
14. Rani, R., & Rani, S. (2023). Predictive Analytics for Heart Disease Risk: A Deep Learning Approach. Journal of

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Medical Systems, 47(4), 1-13.
15. International Journal on "Wielding Neural Networks to Interpret Facial Emotions in Photographs with Fragmentary Occlusion", on American Scientific Publishing Group (ASPG) Fusion: Practice and Applications(FPA) ,Vol. 17, No. 01, August, 2024, pp. 146-158.
16. International Journal on "Prediction of novel malware using hybrid convolution neural network and long short-term memory approach", on International Journal of Electrical and Computer Engineering (IJECE),Vol. 14, No. 04, August, 2024, pp. 4508-4517.
17. International Journal on "Cross-Platform Malware Classification: Fusion of CNN and GRU Models",on International Journal of Safety and Security Engineering (IIETA),Vol. 14, No. 02, April, 2024, pp. 477-486
18. International Journal on "Enhanced Malware Family Classification via Image-Based Analysis Utilizing a Balance-Augmented VGG16 Model,onInternational information and Engineering Technology Association (IIETA),Vol. 40, No. 5, October, 2023, pp. 2169-2178
19. International Journal on "Android Malware Classification Using LSTM Model, International information and Engineering Technology Association (IIETA) Vol. 36, No. 5, (October, 2022), pp. 761 – 767. Android Malware Classification Using LSTM Model | IIETA.
20. International Journal on "Classification of Image spam Using Convolution Neural Network", Traitement du Signal, Vol. 39, No. 1, (February 2022), pp. 363-369 .
21. International Journal on "Medical Image Classification Using Deep Learning Based Hybrid Model with CNN and Encoder", International information and Engineering Technology Association (IIETA), Revue d'IntelligenceArtificiellVol. 34, No. 5, (October, 2020), pp. 645 – 652.
22. International Journal on "Prediction of Hospital Re-admission Using Firefly Based Multi-layer Perception, International information and Engineering Technology Association (IIETA) Vol. 24, No. 4, (sept, 2020), pp. 527 – 533.
23. International Journal on "Energy efficient intrusion detection using deep reinforcement learning approach",Journal of Green Engineering (JGE),Volume-11, Issue-1,January 2021.625-641.
24. International Journal on "Classification of High Dimensional Class Imbalance Data Streams Using Improved Genetic Algorithm Sampling", International Journal of Advanced Science and Technology, Vol. 29, No. 5, (2020), pp. 5717 – 5726.
25. Dr. M. Ayyappa Chakravarthi etal. published Springer paper "Machine Learning-Enhanced Self-Management for Energy-Effective and Secure Statistics Assortment in Unattended WSNs" in Springer Nature (Q1), Vol 6, Feb 4th 2025
26. Dr. M. Ayyappa Chakravarthi etal. published Springer paper "GeoAgriGuard AI-Driven Pest and Disease Management with Remote Sensing for Global Food Security" in Springer Nature (Q1), Jan 20th 2025.
27. Dr. M. Ayyappa Chakravarthi etal. presented and published IEEE paper "Machine Learning Algorithms for Automated Synthesis of Biocompatible Nanomaterials" , ISBN 979-8-3315-3995-5, Jan 2025.
28. Dr. M. Ayyappa Chakravarthi etal. presented and published IEEE paper "Evolutionary Algorithms for Deep Learning in Secure Network Environments" ISBN:979-8-3315-3995-5, Jan 2025.
29. Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Time Patient Monitoring Through Edge Computing: An IoT-Based Healthcare Architecture" in Frontiers in Health Informatics (FHI), Volume 13, Issue 3, ISSN-Online 2676-7104, 29th Nov 2024.
30. Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Amalgamate Approaches Can Aid in the Early Detection of Coronary heart Disease" in Journal of Theoretical and Applied Information Technology (JATIT) , Volume 102, Issue 19, ISSN 1992-8645, 2nd Oct 2024.
31. Dr. Ayyappa Chakravarthi M, etal, published Scopus paper "The BioShield Algorithm: Pioneering Real-Time Adaptive Security in IoT Networks through Nature-Inspired Machine Learning" in SSRG (Seventh Sense Research Group) -International Journal of Electrical and Electronics Engineering (IJEEE), Volume 11, Issue 9, ISSN 2348-8379, 28th Sept 2024.
32. Ayyappa Chakravarthi M, Dr M. Thillaikarasi, Dr Bhanu Prakash Battula, published SCI paper "Classification of Image Spam Using Convolution Neural Network" in International Information and Engineering Technology Association (IIETA) - "Traitement du Signal" Volume 39, No. 1
33. Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published Scopus paper "Classification of Social Media Text Spam Using VAE-CNN and LSTM Model" in International Information and Engineering Technology Association (IIETA) - Ingénierie des Systèmes d'Information (Free Scopus) Volume 25, No. 6.
34. Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published Scopus paper "Social Media

Text Data Classification using Enhanced TF_IDF based Feature Classification using Naive Bayesian Classifier" in International Journal of Advanced Science and Technology (IJAST) 2020

35. Ayyappa Chakravarthi M. presented and published IEEE paper on "The Etymology of Bigdata on Government Processes" with DOI 10.1109/ICICES.2017.8070712 and is Scopus Indexed online in IEEE digital Xplore with Electronic ISBN: 978-1-5090-6135-8, Print on Demand(PoD) ISBN:978-1-5090-6136-5, Feb'2017.
36. Subba Reddy Thumu & Geethanjali Nellore, Optimized Ensemble Support Vector Regression Models for Predicting Stock Prices with Multiple Kernels. Acta Informatica Pragensia, 13(1), x–x. 2024.
37. Subba Reddy Thumu, Prof. N. Geethanjali. (2024). "Improving Cryptocurrency Price Prediction Accuracy with Multi-Kernel Support Vector Regression Approach". International Research Journal of Multidisciplinary Technovation 6 (4):20-31.
38. Dr syamsundararaothalakola et.al. published Scopus paper "An Innovative Secure and Privacy-Preserving Federated Learning Based Hybrid Deep Learning Model for Intrusion Detection in Internet-Enabled Wireless Sensor Networks " in IEEE Transactions on Consumer Electronics 2024.
39. Dr syamsundararaothalakola et.al. published Scopus paper "Securing Digital Records: A Synerigistic Approach with IoT and Blockchain for Enhanced Trust and Transparency " in International Journal of Intelligent Systems and Applications in Engineering 2024.
40. Dr syamsundararaothalakola et.al. published Scopus paper "A Model for Safety Risk Evaluation of Connected Car Network " inReview of Computer Engineering Research2022.

INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com