iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VII **Month of publication:** July 2023

DOI: https://doi.org/10.22214/ijraset.2023.54975

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VII Jul 2023- Available at www.ijraset.com

Future Prediction of Brain Stroke Using Machine Learning

Raavi Anusha[1], Syda Nahida[2]
[1]*Student, Department of CSE, NRI Institute of Technology, Pothavarappadu, Vijayawada, India*
[2]*Associate Professor, Department of CSE, NRI Institute of Technology, Pothavarappadu, Vijayawada, India*

Abstract: ***Ischemic stroke is a condition in which brain stops working due to lack of blood supply resulting in death of brain cells. Image processing is widely used to detect Ischemic Strokes in brain. Ischemic stroke is a leading cause of disability and death for which no acute treatments exist beyond recanalization. The pre-processing is performed including filtering on the raw data collected. According to WHO, if these brain strokes are continued, we were going to see massive deaths all over the world. This work presents about prediction of brain stroke using Machine Learning. It shows the outcome of numerical analysis that is meant as a preliminary step towards experimentation. For this, regular information is collected in real time and is computed. we present synthetic results that are obtained by experimentation. the strokes can be detected in early stages. According to previous study, we have Computed Tomography for identifying Brain Stroke. Our work shows the usage of Machine Learning tools using machine learning for early prediction of Brain Stroke.***
Keywords: Analysis, investigation, research

I. INTRODUCTION

Everyone's life is viewed as being centred around their health hence it is necessary to have a system in place for keeping track of disease statistics and their connections.

In clinic medical records, patient case summaries, and other manually preserved documents, the majority of information relevant to diseases can be discovered. These might be read using various text mining and machine learning approaches to decode the sentences (ML). Machine learning is a tool that can be used to spread content as part of information retrieval, where the semantic and syntactic components of the content are given precedence. In order to extract features and classify them, various ML and text mining approaches are presented and put into practice.

In today's scenario brain strokes are one of the leading diseases that is affecting the lifespan of a particular person Generally people of age limit 35 - 50 are being much affected by this brain stroke that is diminishing their lifespan. This brain stroke is leading to major side effect such as paralysis According to a survey every 4 out of 10 were facing the problem of paralysis. A person who is having a problem of diabetes have 2x times more effect than a general brain stroke treated person. So, in this project we aim to predict the brain stroke for a particular person based on their habits, medical conditions using machine learning.

II. LITERATURE SURVEY

In the research conducted by Manisha Sirsat, Eduardo Ferme, Joana Camara, the main aim of the research was to classify state-of-arts on ML techniques for brain stroke into 4 categories based on their functionalities or similarity, and then review studies of each category systematically. The study further discusses the outcomes and accuracies obtained by using different Machine Learning models using text and image-based datasets. In this study, the authors discussed many stroke related problems from the state-of-art. The reviewed studies were grouped in several categories based on their similarities. The study notes that it is difficult to compare studies as they employed different performance metrics for different tasks, considering different datasets, techniques, and tuning parameters. Hence, it only mentions the research areas which were targeted in more than one study and the studies which report highest classification accuracy in each section.

The paper finally concludes by discussing how Machine learning applications are expanding in the medical field for diagnostic and therapeutic purposes, and the rapidly expanding and increasingly neuro-imaging reliant field of AIS is proving to be fertile ground. There is a particular need for ML solutions in this field, which is faced with the challenge of increasingly complex data, with limited human expert resources. Future directions in ML for AIS may require collaborative approaches across multiple institutions to build a robust dataset for efficient training of ML networks.

III. EXISTING SYSTEM

The existing system was based on MRI and CT scans for classifying early strokes. Previously work is done for how to treat the brain stroke within three months of attack of a brain stroke from the time patient is admitted in the hospital based on properties of EEG. the recent work presents approaches such as LSTM, biLSTM, GRU, and FFNN but GRU performs the best with 95.6% accuracy, whereas biLSTM gives 91% accuracy and LSTM gives 87% accuracy and FFNN gives 83% accuracy.In the last five years, several scores such as the ASTRAL, DRAGON and THRIVE have been proposed as tools to help physicians predict the patient functional outcome after a stroke. Diffusion-weighted magnetic resonance imaging (DWI) is sensitive to acute ischemic stroke and is a common diagnostic method for the stroke.

IV. PROPOSED SYSTEM

The proposed system is an application that is based on machine learning algorithms. In this work we will predict the brain stroke in early stages than before based on the person's gender, age, previous medical history, person's food habits, the place where the person is staying, his/her marital status, smoking status, average glucose level, body mass index values, the kind of work (or) work environment, any previous surgeries happened, what are the symptoms the person is facing in the current situation. This is compared to the dataset that is already given through Random Forest Algorithm and using that comparison we will determine what the disease is whether it a brain stroke or any other else. If it is a brain stroke, then the type of brain stroke will be detected and can be forwarded for treatment.

V. METHODOLOGY

In this process, first the dataset is taken and performed data pre-processing which helps to clean the dataset. Then the data is divided into training data and test data respectively. Next, the Random Forest model will be run followed by training and testing. Finally, the output is shown through the application. If the outcome is positive then, the patient is advised to consult a doctor either wise if the result is negative, then the patient will be given with a prediction that is the person is having a chance of getting a stroke based on the daily routine, food habits, smoking status, age, work type, marital status, previous health history etc.

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. Random Forest is a classifier that contains a number of decision trees on various subsets of given dataset and takes the average to improve the predictive accuracy of that dataset. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

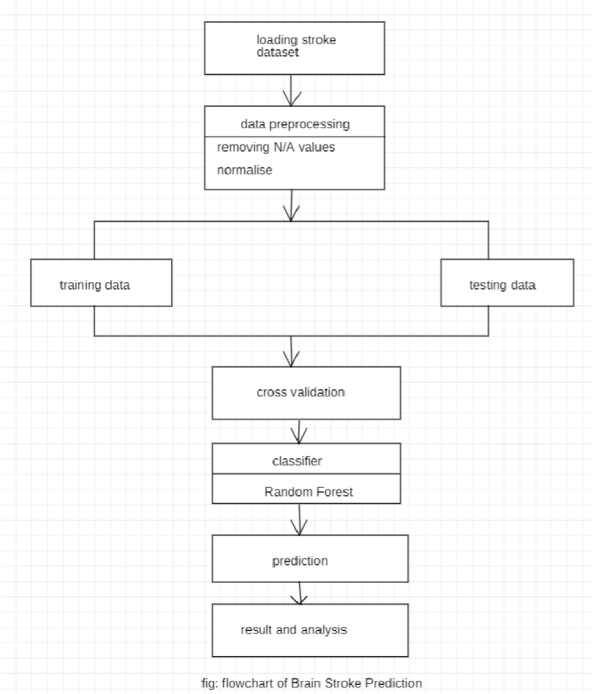


fig: flowchart of Brain Stroke Prediction

VI. IMPLEMENTATION & CONCLUSION

A. Data Collection

Data is the main source to get the result, to implement the model whether the brain stroke is present or not. Random Forest is one of the top and leading techniques of machine learning to solve any type of problem without any regard.

B. Pre-processing of Data

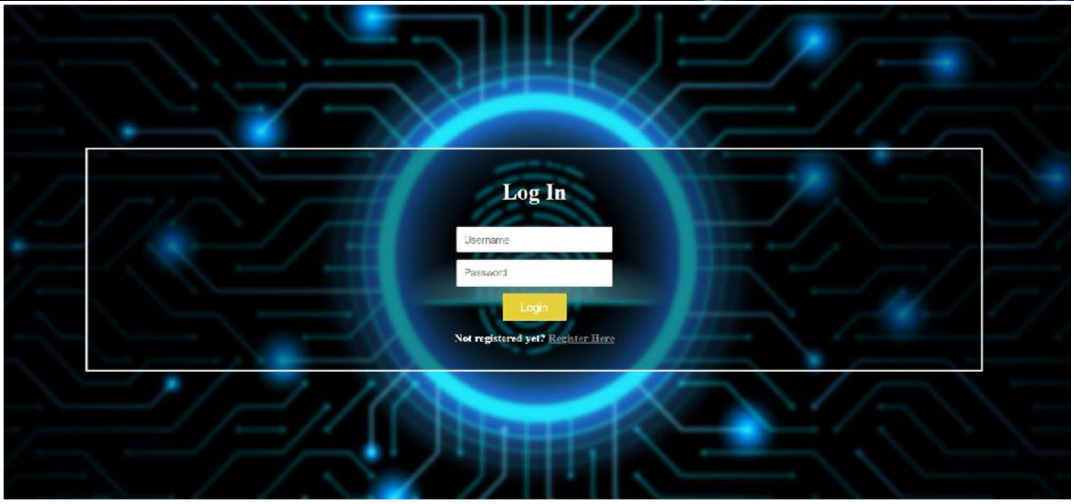
Data pre-processing is another important thing to be remembered while training a model. Before training any model the data that we are giving to the model should be proper and appropriate. For this purpose, we collected raw data, and we will remove all the improper values that can't be accepted while training the model. However our pre trained model has completed this process and hence our data is clean and accurate.

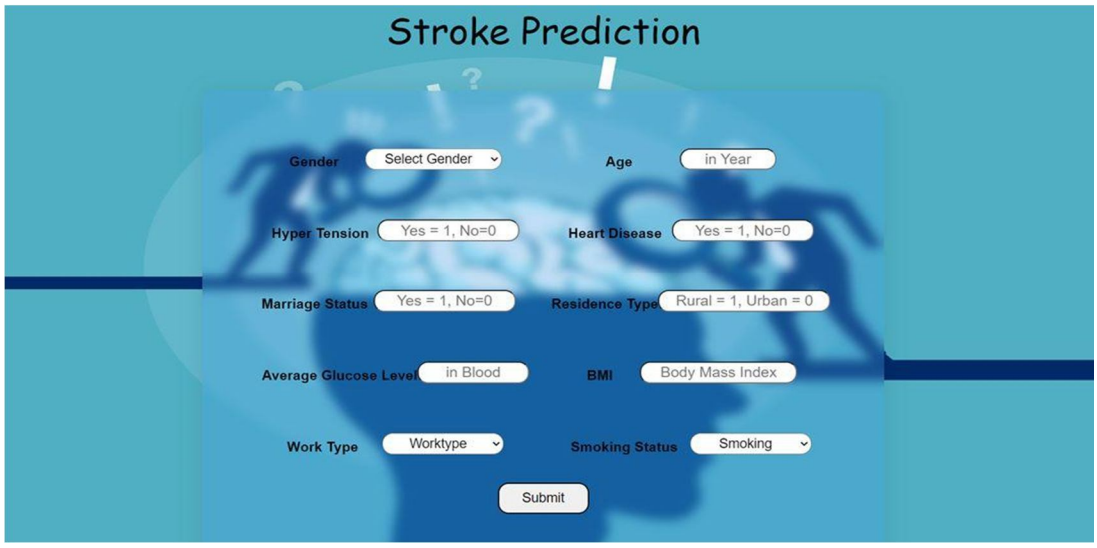
C. Workflow

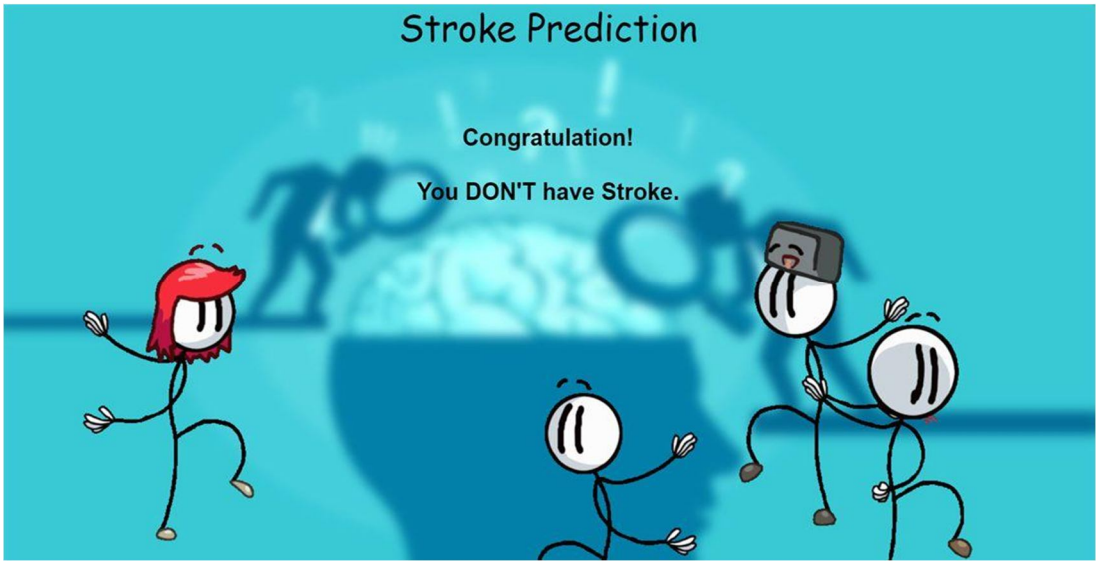
1) Collect data from outsource.
2) Login into the software.
3) Define the data.
4) Train the model
5) Test the model.

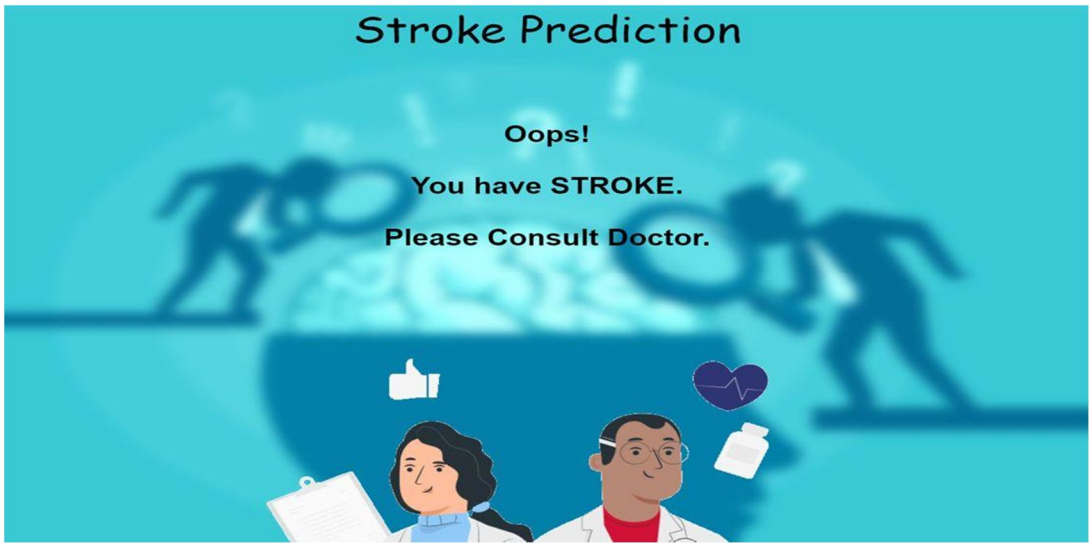
VII. RESULTS

A. Sample Screens




Stroke Prediction
Gender
Select Gender
Age
in Year
Hyper Tension
Yes = 1, No=0
Heart Disease
Yes = 1, No=0
Marriage Status
Yes = 1, No=0
Residence Type
Rural = 1, Urban = 0
Average Glucose Level
in Blood
BMI
Body Mass Index
Work Type
Worktype
Smoking Status
Smoking
Submit


Stroke Prediction
Congratulation!
You DON'T have Stroke.

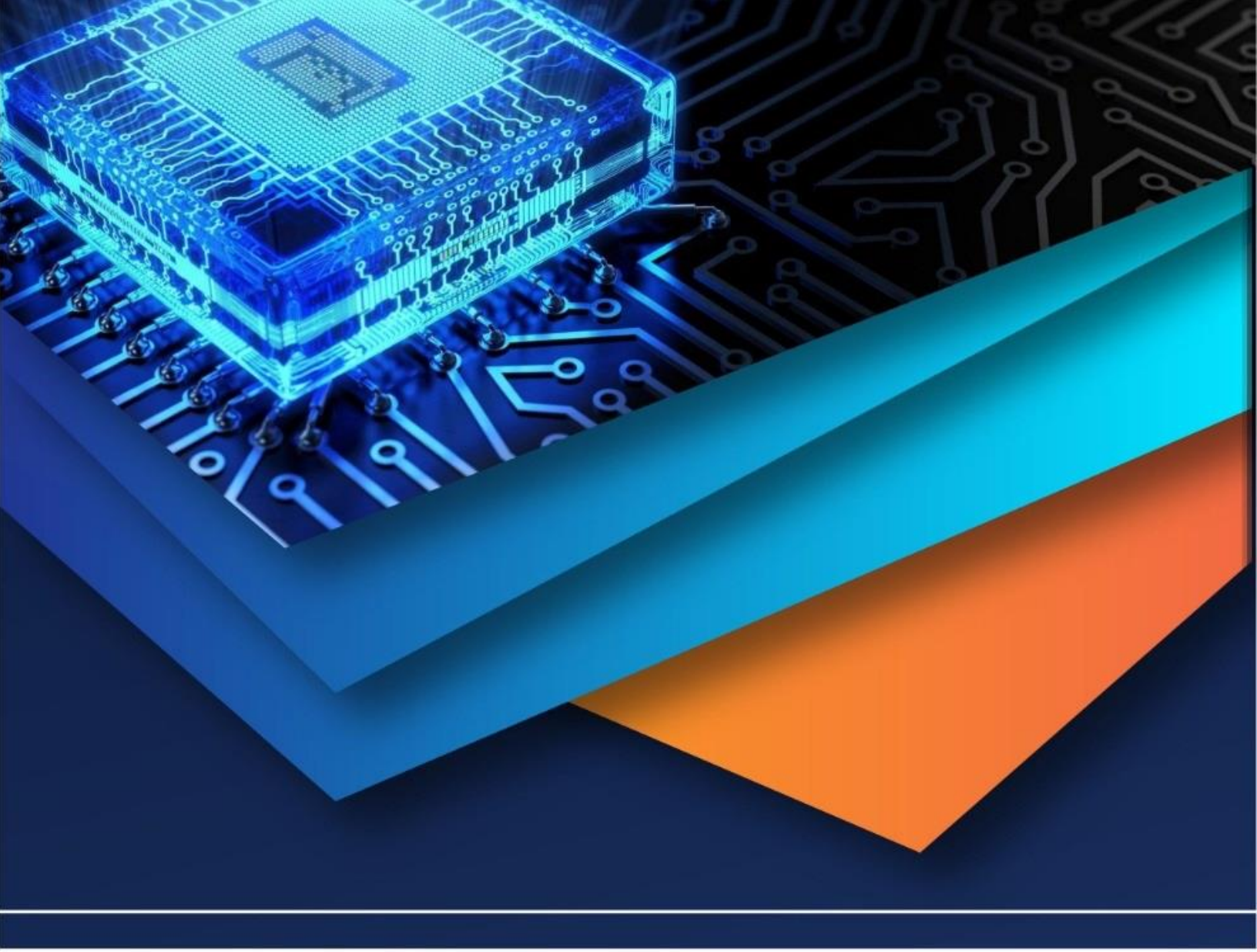

Stroke Prediction
Oops!
You have STROKE.
Please Consult Doctor.

VIII. CONCLUSION

We conclude that machine learning techniques can be effectively used to predict the functional outcome of a brain stroke. So, we wish to improve the prediction by using machine learning techniques with the help of Random Forest algorithm. To get this job done we created a software in which technician will login into those websites. This software is deployed only to the medical teams to keep the data confidential. So, this entire work is controlled by the technician. Firstly, the technician will register into the software. After logging in that software, the stroke page will be opened. Before that the technician will train the patient data with the help of random forest algorithm. So that the patient data is stored in the software which is called as trained data. After training the data, the technician will provide inputs to the software which is called test data. So, with the help of our software, it predicts whether brain stroke is present or not. If the result obtained is a positive one, then it will suggest the person to consult a doctor immediately.

REFERENCES

[1] S. Krishnan, P. Magalingam, and R. Ibrahim, "Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction," Int. J. Electr. Comput. Eng., vol. 11, no. 6, pp. 5467–5476, 2021, doi: 10.11591/ijece. v11i6.pp5467-5476.
[2] M. S. Phipps and C. A. Cronin, "Management of acute ischemic stroke," BMJ, vol. 368, 2020, doi: 10.1136/bmj. l6983.
[3] E. C. Djamal, R. I. Ramadhan, M. I. Mandasari, and D. Djajasasmita, "Identification of post-stroke eeg signal using wavelet and convolutional neural networks," Bull. Electr. Eng. Informatics, vol. 9, no. 5, pp. 1890–1898, 2020, doi: 10.11591/eei. v9i5.2005.
[4] P. Parmar, "Stroke: Classification and diagnosis," Clin. Pharm., vol. 10, no. 1, 2018, doi: 10.1211/CP.2018.20204150.
[5] V. K. Chaithanya Manam, V. Mahendran, and C. Siva Ram Murthy. " Performance Modeling of DTN Routing with Heterogeneous and Selfish Nodes." Wireless Networks, vol. 20, no. 1, pp. 25-40, January 2014.
[6] V. K. Chaithanya Manam, Gaurav Gurav, and C. Siva Ram Murthy." Performance Modeling of Message-Driven Based Energy-Efficient Routing in Delay-Tolerant Networks with Individual Node Selfishness." In COMSNETS'13: Proceedings of the 5th International Conference on Communication Systems and Networks, pp. 1-6, January 2013.
[7] V. K. Chaithanya Manam, V. Mahendran, and C. Siva Ram Murthy. "Message-Driven Based Energy-Efficient Routing in Heterogeneous Delay-Tolerant Networks." In MSWiM HP- MOSys'12: Proceedings of ACM MSWIM Workshop on High-Performance Mobile Opportunistic Systems, pp. 39-46, October 2012.
[8] V. K. Chaithanya Manam, V. Mahendran, and C. Siva Ram Murthy. "Performance Modeling of Routing in Delay-Tolerant Networks with Node Heterogeneity" In COMSNETS'12: Proceedings of the 4th International Conference on Communication Systems and Networks, pp. 1-10, January 2012.
[9] V. K. Chaithanya Manam, Dwarakanath Jampani, Mariam Zaim, Meng-Han Wu, and Alexander J. Quinn. "TaskMate: A Mechanism to Improve the Quality of Instructions in Crowdsourcing." In Companion Proceedings of The 2019 World Wide Web Conference (WWW '19). Association for Computing Machinery, New York, NY, USA, pp. 1121–1130, May 2019.
[10] V. K. Chaithanya Manam, and A. Quinn. " WingIt: Efficient Refinement of Unclear Task Instructions." In HCOMP'18: Proceedings of the 6th AAAI Conference on Human Computation and Crowdsourcing, pp.108-116, June 2018
[11] S. Nyamathulla , Dr. P. Ratnababu , Dr. G. Shobana , Dr. Y. Rokesh Kumar4 , K.B.V. Rama Narasimham A Fast, Dynamic method to identify attributes sets using Corelation-Guided Cluster analysis and Genetic algorithm Techniques" in Design Engineering ISSN: 0011-9342 | Year 2021 Issue: 7 | Pages: 5497-5510.
[12] Mrs.Shobana gorintla ,2 Mr.B.Anil Kumar ,3Mrs.B.Sai Chanadana ,4 Dr.N.Raghavendra Sai,5 Dr.G.Sai Chaitanya Kumar "Deep-Learning-Based Intelligent PotholeEye+ Detection Pavement Distress Detection System" in Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC 2022) IEEE Xplore Part Number: CFP22BC3-ART; ISBN: 978-1-6654-9710-
[13] G. Shobana, Dr Bhanu Prakash Battula" A Novel Imbalance Learning with Fusion Sampling using Diversified Distribution" in International Journal Of Research In Electronics And Computer Engineering -IJRECE VOL. 6 ISSUE 3 (JULY - SEPTEMBER 2018) ISSN: 2393- 9028 (PRINT) | ISSN: 2348-2281 (ONLINE).
[14] G. Shobana, Dr Bhanu Prakash Battula " A comparitive study of skewed data sources using fusion sampling Diversified Distributon" International Journal of Research in Advent Technology, Special Issue, March 2019 E-ISSN: 2321- 9637



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY