

Global Warming Prediction in India using Machine Learning



D. Deva Hema, Anirban Pal, Vineet Loyer, Rajeev Gaurav

Abstract: Long term global warming prediction can be of major importance in various sectors like climate related studies, agricultural, energy, medical and many more. This paper evaluates the performance of several Machine Learning algorithm (Linear Regression, Multi-Regression tree, Support Vector Regression (SVR), lasso) in problem of annual global warming prediction, from previous measured values over India. The first challenge dwells on creating a reliable, efficient statistical reliable data model on large data set and accurately capture relationship between average annual temperature and potential factors such as concentration of carbon dioxide, methane, nitrous oxide. The data is predicted and forecasted by linear regression because it is obtaining the highest accuracy for greenhouse gases and temperature among all the technologies which can be used. It was also found that CO₂ is the plays the role of major contributor temperature change, followed by CH₄, then by N₂O. After seeing the analysed and predicted data of the greenhouse gases and temperature, the global warming can be reduced comparatively within few years. The reduction of global temperature can help the whole world because not only human but also different animals are suffering from the global temperature.

Index Terms: Global Warming, Temperature prediction, Greenhouse gases prediction, Linear Regression.

I. INTRODUCTION

Global Warming is the rise in the average for long-term temperature of the Earth's climate system. This climate change is occurred due to some physical factors like CO₂, N₂O, CH₄ etc. These factors are increasingly simultaneously in the weather. For this, the temperature of Earth's temperature is also increasing. The temperature for this reason was increasing since 1990 caused by emission of greenhouse gases from the modern industry, vehicle, burning of fossil fuels etc. The average temperature of earth is 1 degree Celsius higher than 100 years ago. Now, many scientists are expecting that within the next 200 years the average temperature will increase by approx. 6 degree Celsius than recent temperature. Global Warming occurs when the greenhouse gases absorb sunlight and solar radiation have reflected from the earth's surface.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

D. Deva Hema*, Assistant Professor in SRM Institute of Science and Technology in Ramapuram, Chennai, (Tamil Nadu), India.
Anirban Pal, Department of Computer Science and Engineering, from SRM Institute of Science and Technology, Chennai, (Tamil Nadu), India.
Vineet Loyer, Department of Computer Science and Engineering, from SRM Institute of Science and Technology, Chennai, (Tamil Nadu), India.
Rajeev Gaurav, Department of Computer Science and Engineering, from SRM Institute of Science and Technology, Chennai, (Tamil Nadu), India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

For that the atmosphere of the earth is increasing and ultimately the overall temperature of earth is also increasing. The result of the global warming is lethal. For the global warming the glacier is melting day by day. For that the average height of the sea level is also increasing and this is cause for flooding in many coastal areas. Not only that, but also the underground water of the earth is decreasing. The ocean's temperature is also increasing for this reason. For that many storms are taking place which has fatal result. Many diseases like allergies, asthma and many other skin problems is also facing by the human for his reason. Many scientists are thinking that, if the temperature will increase like this then many cities will be go under the ocean or submerged with the ocean. A paper [16] describes about global warming. They explain the causes, effects and probable solutions for global warming. There is a paper [7], which gives also some idea about the temperature prediction. This paper helps us to predict the air temperature, which is the main problem of global warming. "Prediction of the likely impact of climate change on monthly mean maximum and minimum temperature in the Chaliyar river basin, India, using ANN- based models" [10], "Big Data and Climate Change" [17], "Atmospheric Temperature Prediction using Support Vector Machines" [18], "Artificial Neural Networks (The Multilayer Perception) – A review of applications in the atmospheric sciences" [21], "Air quality prediction using optimal neural networks with stochastic variables" [23], "The forecasting research of early warning systems for atmospheric pollutants: A case in Yangtze River Delta region" [24] are some reference paper of temperature, air condition etc. analysis and prediction. This papers also provided some basic information about temperature, global warming etc. To get relief from the destructive result of Global Warming, many people are reducing the fossil fuel burning to reduce the Global Warming. Not only that but also, they are trying to aware the country people to emit less greenhouse gases. Due to the sea level rise many cities like London, New York, Virginia, Sydney, Charleston, Mumbai etc and sea ports are affected heavily. Due to Global Warming many forests are reducing, many trees are dying for the overheated temperature. For that the ratio of every gases is not perfect in atmosphere in many places. Though the ratio of this effect is lesser comparatively than other. For the high temperature the crops cannot be grown properly which affects the animal food cycle. That means we cannot explain the effects of Global Warming in a single sentence.

II. RELATED WORKS

There are so many ideas and project about weather prediction, rainfall prediction, temperature prediction. Some of those ideas are taken for the reference purpose.

In this project the temperature and the greenhouses gases are predicting for next few years. For that the weather prediction-based ideas and projects are focused mainly.

“A new approach for simulating and forecasting the rainfall-runoff process within the next two months” by M.J Alizadeh, M.R. Kavianpour, Ozgur Kisi, Vahid Nourani [1] is one of the good concepts for predicting the rainfall. In this ideas Artificial Neural Network, SVR etc techniques are used. This concept is focussing only rainfall. It doesn't provide any information about temperature or greenhouse gases.

“Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms” by S. Salcedo-Sanz, R. C. Deo, L. Carro-Calvo, B. Saavedra Moreno [2] is also a prediction-based idea. In this idea the temperature is only focused. The physical factors are not focused which are responsible for the Global Warming. In this project SVR and multilayer-perceptron methods are used.

“Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes” by F. Mekanik, M.A. Imteaz, S. Gato-Trinidad, A. Elmahdi [3] is also another prediction-based idea. In this project regression and artificial neural network are used to predict the rainfall. This idea is focussing on only the rainfall. It is not focussing on temperature nor greenhouse gases.

“Development and Analysis of ANN Models for Rainfall Prediction by Using Time-Series Data” by Neelam Mishra, Hemant Kumar Soni, Sanjiv Sharma, AK Upadhyay [4] is also used as a reference. In this project regression, mean square error and MRE are used. This idea also focused only on rainfall not on temperature nor greenhouse gases.

“Application of Artificial Neural Networks to Rainfall Forecasting in Queensland, Australia” by John Abbot and Jennifer Marohasy [5] is also taken for reference. In this project Artificial Neural Networks is used to observe and forecast rainfall. This idea also doesn't give any explanation about global warming.

“Analysis of Global Warming Using Machine Learning” by Harvey Zheng [8] is also used as a reference. In this project SVM, lasso, random forest is used. The idea has focused on the global warming but the explanation of this idea is more complex and not so clear.

“An Integrated Approach for Weather Forecasting based on Data Mining and Forecasting Analysis” by G.Vamsi Krishna[14] is also used as a reference for this paper. The explanation of the paper is very good but it is focused the overall weather, not a particular thing.

“An interactive predictive system for weather forecasting” by Ayham Omary, Ahmad Wedyan, Ahmed Zghoul, Ahmad Banihai, Izzat Alsmadi[12] is also used as a reference in our paper. The explanation is also based on the overall weather not in a particular field.

“Research on weather forecast based on neural networks” by Yuan Quan and Lu Yuchang[11] papers gives the concept of weather prediction using neural networks. They also don't focus on global warming.

“Weather Prediction Using Data Mining” [13] is also a weather prediction-based paper which don't focus on global warming.

“Machine Learning Applied to Weather Forecasting” [15], “Localized Precipitation Forecasts from a Numeric Weather Prediction Model Using Artificial Neural Networks” [19], “Neural Network Local Forecasting with Weather Ensemble Predictions” [20], “Multistage Artificial Neural Network Short-Term Load Forecasting Engine with Front-End Weather Forecast” [22], “Neural network based short-term load forecasting using weather compensation” [25] are also used as reference paper. They also don't focus on global warming.

“Comparison of neural network configuration in the long-range forecast of southwest monsoon rainfall over India” [6] is a paper which focused on rainfall prediction. They also don't focus on global warming. “A hybrid Double Feedforward Neural Network for Suspended Sediment Load Estimation” [9] is also a prediction-based paper but they don't focus on the temperature or global warming.

III. PROPOSED SYSTEM

The main objective of this study is to analyse global warming (temperature and greenhouse gases like- CO₂, N₂O, Methane) in India for approx. 100 – 150 years on the basis data recorded at different place like Kaggle, Data.gov.in etc. The specific objectives are:

1. To predict the temperature and greenhouse gases concentration for next 10 years.
2. To make a graphical interface based on the prediction to ease of understand.

IV. THE PROPOSED GLOBAL WARMING PREDICTION SYSTEM

There are so many technologies to predict data like svm, linear regression, lasso etc. We have tried many of the algorithm to get the highest accuracy. All the methods have different working procedure. The working procedure of some of those predicting techniques are discussed below:

1. Linear Regression

Linear Regression is a method which give a relationship between a dependent variable or scalar variable and an independent variable or explanatory variable. In this method the relationships are modelled using linear predictor function. Here the data is trained by this method. Linear predictor function is used to make an object of that function and used it for prediction. After creation of the object, the data is forecasted for future.

2. Multiple Regression

Multiple Regression is a method which give a relationship between a dependent variable and one or more independent variable. The dependent variable is modelled as a function of the different independent variable. Here the basic difference from the linear regression is here the independent variable may or may not be more than one, remaining all other things are same.

Here also linear predictor function is used to make an object of that function and used it for further prediction.

3. Support Vector Regression

Support Vector Regression is one part of Support Vector Machine. SVR follows the same principle which is followed by SVM. For support vector regression, the prediction method is difficult comparative to other methods. The algorithm is more complicated.

Among all the technologies more accuracy is observed in Linear Regression. For that in this paper Linear Regression is used. The complexity on Linear Regression is also comparatively much lesser than the other technologies.

There are some modules which is required to develop the Global Warming Prediction System. Those modules are briefly explained below:

A. Data Collection

In this module the raw is collected data from different data set. Then the data set is changed as per need. This raw data cannot be predicted directly. So, it is needed to clean and pre-process.

B. Data Pre-processing

In this module the data is cleaned. After cleaning of the data, the data is grouped as per requirement. This grouping of data is known as data clustering. Then check if there is any missing value in the data set or not. If there is some missing value then change it by any default value. After that if any data need to change its format, it is done. That total process before the prediction is known as data pre-processing. After that the data is used for the prediction and forecasting step.

C. Data Prediction and forecasting:

In this step, the pre-processed data is taken for the prediction. This prediction can be done in any process which are mentioned above. But the Linear Regression algorithm score more prediction accuracy than the other algorithm. So, in this project the linear regression method is used for the prediction. For that, the pre-processed data is splitted for the train and test purpose. Then a predictive object is created to predict the test value which is trained by the trained value. Then the object is used to forecast data for next few years.

D. Visualization:

In this step, the predicted and forecasted data is used to provide a graphical interface separately. At first the predicted data is plotted in a graph separately with the help of matplotlib library. Then the forecasted data of temperature is plotted in graph with proper scale. Then the greenhouse gases forecasted data are plotted in a single graph with proper scale.

V. PROPOSED SYSTEM ARCHITECTURE

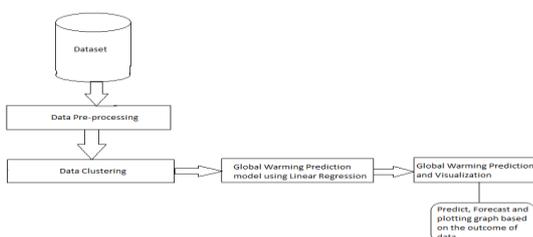


Fig: 1 (Proposed System Architecture)

VI. DATA SET USED

A. Temperature

The temperature data set from 'Kaggle' is taken. After taking the data, it is modified and reduced as per need. In the raw data, the temperature of each month. Those data are converted into average temperature for better and ease usage. In the temperature data the average temperature of India is given. So that we can easily predict the average temperature of India by trained those data.

B. Carbon di-oxide:

The data set for the carbon di-oxide is not directly taken data. At first, the data from 'data.gov.in' is taken and then the data set is modified as per need. In the raw data set, there are only 20 years data. After that, by referring some other website the data are entered for 70 to 80 years. Then the data is used for further process.

C. Methane

For the methane data set, 'methanelevels.org' is referred. In the raw data set of methane, the 80 percent of data is given. Remaining data are filled by referring some other sources.

D. Nitrous oxide

For the N2O data set, 'n2olevels.org' is referred. In the raw data set of nitrous oxide, the 90 percent of data is given. Remaining data are filled by referring some other sources. After analysing all the data, 90% data is selected to train. After training the data, an object is created for the linear model. Then, remaining data is predicted and checked the prediction score with the help of that object. Then the temperature and greenhouse gases data are forecasted for the next ten years

VII. GRAPHICAL OUTPUT

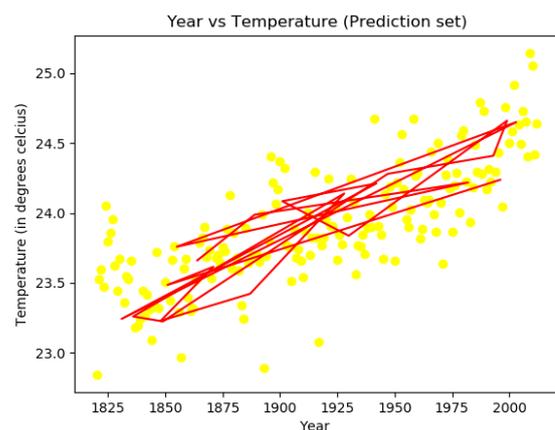


Fig: 2 (Year VS Temp prediction)

Global Warming Prediction in India using Machine Learning

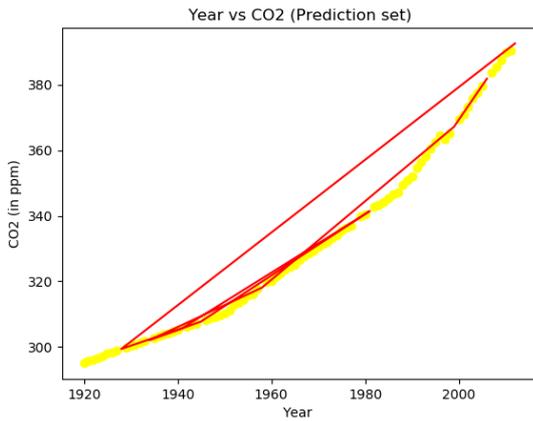


Fig: 3 (Year VS CO2 prediction)

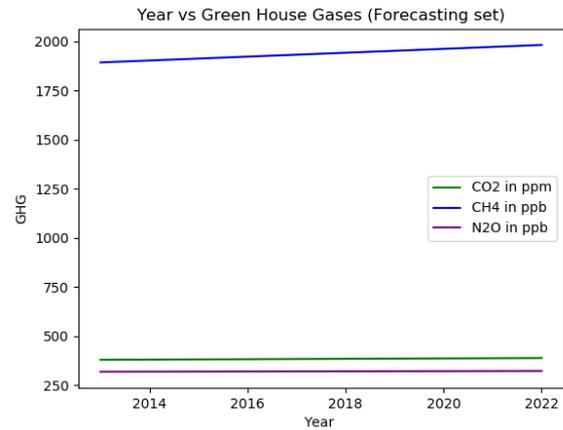


Fig: 6 (Year VS GHG forecasting)

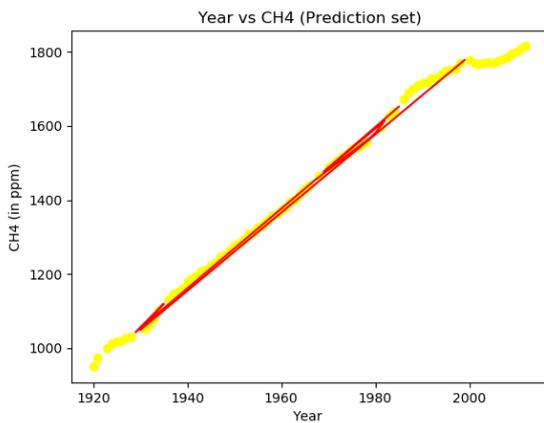


Fig: 4 (Year VS CH4 prediction)

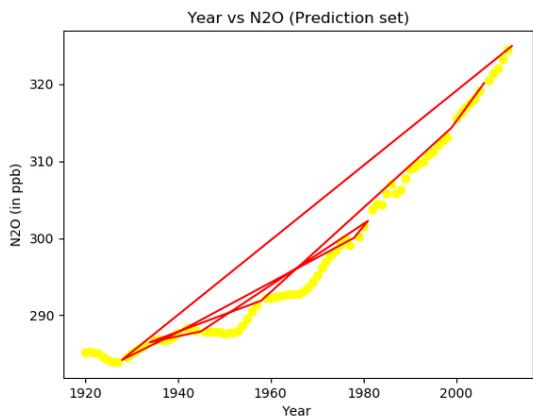


Fig: 5 (Year VS N2O prediction)

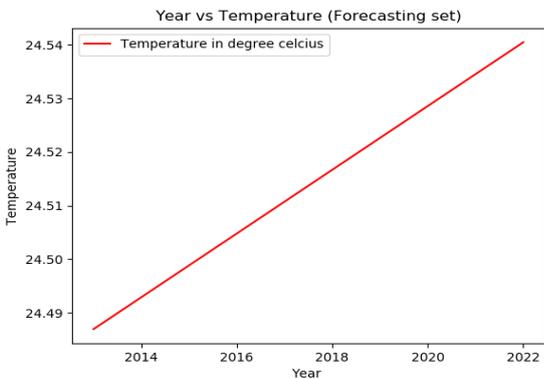


Fig: 6 (Year VS Temp forecasting)

VIII. RESULT AND DISCUSSION

After successful training of the temperature and the greenhouse gases data, more than satisfactory accuracy is done for the temperature prediction as well as the greenhouse gases prediction. After the successful prediction, forecasted data for the next 10 years and also graphical representation of those predicted and forecasted data with the help of matplotlib library is obtained. The graph (Fig: 2) is plotted with the help of the predicted data of year versus temperature. The graph (Fig: 3) is plotted between the predicted carbon di-oxide and year. The next graph (Fig: 4) is plotted between predicted data of methane and year. The last prediction graph (Fig: 5) is plotted between nitrous oxide and year. For the forecasting part, the temperature is plotted separately (Fig: 6) because the unit is different for the temperature from the others physical factors. And for the CO2, N2O and CH4 are plotted in a same graph (Fig: 7) where the data are compared between years and ppm or ppb (for only CO2). Because of using Linear Regression methodology, the graph is observed in a linear pattern (i.e.: $Y = MX + C$), where the X axis represents the years every time whereas the Y axis represents the temperature and other physical factors in different unit. After seeing the graph for the forecasting, it can be said that the temperature and physical factors which are responsible for global warming will increase day by day in a linear pattern and the result of this global warming will become more and more lethal.

IX. CONCLUSION

In this paper, the data (temperature and greenhouse gases) of 100-150 years is analysed. Linear Regression and Linear model are used to predict and forecast the temperature and greenhouse gases for the next 10 years in average. The matplotlib library is used to plot the predicted and the forecasted data. So, at last the following conclusion can be drawn – A model for forecasting data for next 10 years is trained and tested with different input variables like temperature, carbon di-oxide, methane, nitrous oxide by linear regression. Some graphs are plotted as a graphical interface for the predicted and forecasted data for all the inputs with the help of matplotlib library.

LIMITATION AND FUTURE WORK

There is some limitation in this study or this project. This model is predicted only the average temperature and greenhouse gases concentration of India. It doesn't explain each months and each state's data separately. The data set is used for this model for 100 – 150 years. But the prediction will be better if we go more than 200- or 250years data.

In future the prediction can be better for using a greater number of data. The other physical particle which is responsible for temperature increasing can be also predicted. The data can be predicted for each month as well as each state of India, which will make this more attractive and reasonable for each people of India.

REFERENCES

1. "A new approach for simulating and forecasting the rainfall-runoff process within the next two months" by M.J Alizadeh, M.R. Kaviani-pour, Ozgur Kisi; Vahid Nourani, Journal of Hydrology (Vol – 548); May-2017
2. "Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms" by S. Salcedo-Sanz, R. C. Deo, L. Carro-Calvo, B. SaavedraMoreno; Theoretical and Applied Climatology (Vol – 125); July-2016
3. "Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes" by F. Mekanik, M.A. Imteaz, S. Gato-Trinidad, A. Elmahdi; Journal of Hydrology (Vol – 503); October-2013
4. "Development and Analysis of ANN Models for Rainfall Prediction by Using Time-Series Data" by Neelam Mishra, Hemant Kumar Soni, Sanjiv Sharma, AK Upadhyay; International Journal of Intelligent Systems and Applications; January-2018
5. "Application of Artificial Neural Networks to Rainfall Forecasting in Queensland, Australia" by John Abbot and Jennifer Marohasy, Advances in Atmospheric Sciences; July-2012
6. "Comparison of neural network configuration in the long-range forecast of southwest monsoon rainfall over India" by Snehasish Chakraverty and Pallavi Gupta; Neural Computing and Applications (Vol – 17); March2008
7. "Support vector regression with reduced training sets for air temperature prediction a comparison with artificial neural networks" by Robert F. C, Gerrit Hoogenboom, Ronald W. McClendon, J. A. Paz; Neural Computing and Applications (Vol – 20); February-2011
8. "Analysis of Global Warming Using Machine Learning" by HarveyZheng, Computational Water, Energy; and Environmental Engineering (Vol - 7); July2018
9. "A hybrid Double Feedforward Neural Network for Suspended Sediment Load Estimation" by Xiao Yun Chen and Kwok Wing Chau; Water Resources Management (Vol - 30); May-2016
10. "Prediction of the likely impact of climate change on monthly mean maximum and minimum temperature in the Chaliyar river basin, India, using ANN- based models" by N.R. Chithra, S.G. Thampi, S. Surapaneni, R. Nannapaneni, A.A.K. Reddy, J.D. Kumar; Theoretical and Applied Climatology (Vol – 121); August-2014
11. "Research on weather forecast based on neural networks" by Yuan Quan and Lu Yuchang; August-2002
12. "An interactive predictive system for weather forecasting" by Ayham Omary, Ahmad Wedyan, Ahmed Zghoul, Ahmad Banihai, Izzat Alsmadi; IEEE; June-2012
13. "Weather Prediction Using Data Mining" by Prashant Biradar, Sarfraz Ansari, Yashavant Paradkar, Savita Lohiya; IJERD (Vol – 5); 2017
14. "An Integrated Approach for Weather Forecasting based on Data Mining and Forecasting Analysis" by G.Vamsi Krishna; International Journal of Computer Application (Vol – 120); June-2015
15. "Machine Learning Applied to Weather Forecasting" by Mark Holmstrom, Dylan Liu, Christopher Vo; 2016
16. "Global Warming: Causes, Effects and Solutions" by Umair Shahzad; Durreesamin Journal (Vol – 1), August2015
17. "Big Data and Climate Change" by H. Hassani, Xu Huang and E. Silva; February-2019
18. "Atmospheric Temperature Prediction using Support Vector Machines" by Y.Radhika and M.Shashi; International Journal of Computer Theory and Engineering (Vol – 1); April-2009
19. "Localized Precipitation Forecasts from a Numeric Weather Prediction Model Using Artificial Neural Networks" by R. J.

20. "Neural Network Local Forecasting with Weather Ensemble Predictions" by J. W. Taylor and R. Buizza; IEEE Transactions and Power Systems (Vol – 17); August-2002
21. "Artificial Neural Networks (The Multilayer Perception) – A review of applications in the atmospheric sciences" by M. W. Gardner and S. R. Dorling; Atmospheric Environment (Vol – 32), June-1998
22. "Multistage Artificial Neural Network Short-Term Load Forecasting Engine with Front-End Weather Forecast" by K. Methaprayoon; IEEE Transactions on Industry Application (Vol – 43); December-2007
23. "Air quality prediction using optimal neural networks with stochastic variables" by A. Russo, F. Raischel, P. G. Lind; Atmospheric Environment; 2013
24. "The forecasting research of early warning systems for atmospheric pollutants: A case in Yangtze River Delta region" by Y. Song, S. Qin, J. Qu, Feng Liu; Atmospheric Environment; 201
25. "Neural network based short-term load forecasting using weather compensation" by T.W.S. Chow and C.T. Leung; IEEE Transactions on Power Systems (Vol – 11), Nov - 1996

AUTHORS PROFILE



D. Deva Hema received the M.E Degree in Anna University in 2007 and pursuing Ph.D. in Satyabhama Institute of Science and Technology. She is currently Assistant Professor in SRM Institute of Science and Technology in Ramapuram, Chennai, Tamil Nadu, India. Her research interests include the area of Artificial Intelligence and Machine Learning including crash prediction and occupant protection during vehicular crashes. She is currently focusing on crash prediction for intelligent transportation systems, injury mechanisms during rollover crashes, and the protection of pedestrians and drivers.



Anirban Pal is currently pursuing B. Tech degree in Computer Science and Engineering from SRM Institute of Science and Technology and will be graduated in 2021. He is learning Full Stack Web Development and Machine learning. His main ambition is to combine his knowledge on website and machine learning to make very portable, dynamic and attractive web page. Not only that but also, he wants to learn Data Science. He wants to be a professional Data Scientist in future and do Web Development for his passion.



Vineet Loyer is pursuing B. Tech degree in Computer Science and Engineering from SRM Institute of Science and Technology and will be graduated in 2021. He has strong interest in Machine Learning, Artificial Intelligence and Data Science. He wants to do his master degree with specialization in Machine Learning



Rajeep Gaurav is pursuing B. Tech degree in Computer Science and Engineering from SRM Institute of Science and Technology and will be graduated in 2021. He has strong interest in Machine Learning and Database Management System. He wants to do his master degree with specialization in Machine Learning.