International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Machine Learning based IoT Flood Rediction Using Data Modeling and Decision Support System

Deivendran P[1], Selvakanmani S[2], Malathi G[3], Vinoth Kumar V[4]
[1] Department of Information Technology. Velammal Institute of Technology, Chennai, India
[2] Department of Computer Science and Engineering, RMK Engineering College, Chennai, India.
[3]School of Computer Science and Engineering,Vellore Institute of Technology,Chennai,India
[4] Department of Information Technology, Velammal Institute of Technology, Chennai, India.)
[1]deivendran1973p@gmail.com

Abstract - An essential step in supplying data for climate impact studies and evaluations of hydrological processes is rainfall prediction. However, rainfall events are complex phenomenon's that continue to be difficult to forecast. In this paper , we present unique hybrid models for the prediction of monthly precipitation that include Seasonal Artificial Neural Networks and Discrete wavelet transforms are two pre-processing methods, together with Artificial Neural Networks have two feed forward neural networks. The temporal series of observed monthly rainfall from Vietnam's Ca Mau hydrological station were decomposed into three subsets by seasonal decomposition and five sub signals and four levels by wavelet analysis. The methods for predicting rainfall that use feed forward artificial neural networks (ANN) and seasonal artificial neural network (SANN) were fed with the processed data. The classic genetic method and simulated annealing method backed by using an integrated moving average and autoregressive moving was contrasted with the predicted models for model evaluation. The results showed that non-stationary regarding issues with non-linear time series, such forecasting rainfall could be satisfactorily simulated. The SANN model was integrated with the wavelet transform and seasonal decomposition are both used. Techniques, however the wavelet transform method produced the most accurate monthly rainfall data, Predictions. Due to the effects of climate change, nations including the Japan, China, the United States of America, and Taiwan, etc., have recently experienced severe and devastating natural disasters. One of the biggest causes of the destruction in Asian nations like china, India, Bangladesh, Sri Lanka, etc. is the flood. The danger of fatality from these floods is increased by 78% as information technology advances; there is a demand for simple access to massive amounts of cloud storage and computing capacity.

Keywords - Prediction, Networks, IoT, Decomposition, Statical, geographical, framework

1. INTRODUCTION

Finding foods has taken more effort in previous years. Flood occur naturally catastrophe that has disturbing effects on commodities, products, services, and the lives of both animals and people. An early warning of such a catastrophe could be useful to lessen the effects and help people. Although river flooding cannot be prevented, its effects can be reduced and managed with the right planning and application of IoT and ML technologies. A recent Gartner analysis found that there were 6.5 billion linked things/objects discovered in 2021, an increase of almost 3% from 2020, and it is predicted that there will be 20.by the year 2022. Numerous sensors are among these items. that could be beneficial for enhancing the quality of data gathered in order to facilitate improved decision-making. IoT is a subject that is expanding and is commonly used for these things by the year 2022 Numerous sensors are among these items, and enables items to be detected or controlled remotely across various network contexts. For example, ML offers and enticing way to anticipate water levels in order to gather meaningful and productive data. An environmental protection, the great majority of environmental monitoring facilities have embraced IoT Sun and Scanlon report that the identification of early flood warnings utilizing potent deep learning algorithms has been greatly improved by the application of machine learning. Gait features can be utilized to measure flood levels, as demonstrated by panchal et al, 2021. The data analysis used support vector and random forest machines are two machine learning approaches. While furquim et al., propose an a flood detection system based on the wireless sensor networks of the internet of things (IoT), as well as fault tolerance to anticipate any communication failure. According to Belal et al., 2020 the difficulty of making early predictions and the need of knowledge regarding the importance of drinking water quality problems. For the analysis to be valid, the authors emphasized the significance of collecting trustworthy and high-quality data. In this study, we estimate river water levels using various sensor data collected from monitoring flood centre's spread throughout many nations. To do this, several cutting edge learning and prediction algorithms, including Long short term memory, support vector machine, and K-nearest neighbor classifier are all examples of machine learning algorithms, are all examples of artificial neural networks were developed. The goal is to analyze flood sensor log files, which exhibit nonlinearities and dynamic properties, using machine learning methods. Decision makers continue to be concerned about quality despite the quality raw data. For instance, the most frequent issues that affect the data quality include missing values, transfer distortions, and noise effects. This research uses a data science technique to analyze and extract features from sensory datasets that are unbalanced classes, noise, and missing values are characteristics. The data from the flood sensor is analyzed using a variety of ML approaches. ML classification models that were only used once utilized in the initial simulation studies did not deliver adequate performance and accuracy. Then, utilizing LSTM, ANN, and RF, a novel classifier that is based on ensemble learning was created. Statistics show that ensemble learning models outperform all single and individual ML classifiers. Statistical relationships between climate variables that are widely removed from one another are known as global teleconnections. Teleconnections, which join several into a unified global climate system, combining. The result of extensive interactions between the water and atmosphere are regionalclimates.Precipitation forecasting efforts have been done repeatedly utilizing data from several climatic indices that describe teleconnection patterns. Using monthly750-hpa teleconnection indexes, the southern oscillation of EI Nio indicators, and ANNs, silverma and Dracup predict the seven zones of California's total annual precipitation. They stressed the potential of employing large-scale climatic variables and ANNs to predict long term precipitation, EnSO, the Kumar et al., used the local ocean land temperature and the Indian Ocean oscillation contract as inputs,2022 development of ANN models for forecasting summer monsoon rainfall. For the usage of Pacific Ocean sea surface temperature anomalies and the delayed for august rainfall forecasts, use of pacific decadal oscillation index southern oscillation index indies, and north pacific index in Fukuoka, Japan, Iseri et al. 2021 developed ANN modes. Using a variety the SOI, the east Atlantic, and other climate indices the Scandinavian pattern, the polar Eurasia pattern, western Russia pattern, ad various indices created from data on sea surface temperatures and pressures at sea level, and snowfall. Hartman et al, 2022 The Yangtze River basin is expected to get summer rains. Yuvan et al, 2023 used ANN models using inputs from the West Pacific, ENSO, North Atlantic Oscillation, and trends to estimate summer precipitation in the Yellow River source basin in china.

2. RELATED WORKS

From this the years 2019-2022 are taken into account. With a focus on data mining approaches, this study will help academics examine the most recent work on rainfall prediction and give a baseline for future et al shanmuganathan[29]. One of the challenging operational responsibilities handled by meteorological organizations globally is rainfall forecaster et al deivendran[30]. The phrase roadmap to success refers to the idea that failure to plan is equivalent to planning for failure et al Mafarja[32]. Future event information is essential for efficient and effective planning. Effective preparation is the only way to avoid the effects of natural disasters like flooding and drought [5]. The many methods include decision trees, clustering K-means, and fuzzy logic. Given the temperature Leaf sickness categorization has tested the efficacy of different pretrained models with the usage of several hyperparameter. They came to the conclusion, to identify centella asiatica leaf disease, use GoogleNet, alexNet[17], Resnet, and DenseNet, system due to their higher presentation compared to other models. Some of these experiments have also examined how different hyperparameter choices, including optimizers, batch diameters, the number of leaves, and fine-tuning the replica from numerous sources lowest points, affect performance. In the field of smart farming, there has been a lot of recent study on the application of machine learning techniques. The author examines the issue of predicting drought with deep learning algorithms in [12]. It suggests utilizing two constrained Boltzmann machines in a deep belief network and lagged standardized stream flow index values as inputs to predict drought over the long run. The study compares the effectiveness of the suggested strategy to that of traditional approaches for forecasting the various time scales of drought conditions, such as support vector regression and multilayer perception. The suggested method outperformed the traditional approaches, according to the results[4] The internet of Things (IoT) technology was used in [15] to the management of environmental challenges in crop fields. It is made up of three parts: a web application, physical components, and software. Using internet of Things applications and machine learning techniques,[12] provided an evaluation of numerous IoT based machine learning models to their unique designs and contributions in a variety of disciplines, including agriculture, the environment, and energy management.[16]. In the mysense system, this is a proposal to coordinate data-gathering activities to handle typical PA/PV problems. The system is built on four layers, like front end application support, crop field and sensor networks, sensor and sensor node networks, and cloud services, a sophisticated monitoring system for agriculture that measures soil humidity and temperature was presented in[32]. Without human interaction, the system processes the detected depending on the observations of soil temperature and humidity, and performs the necessary action various machine learning techniques are presented by the author in[22] the results are split into three groups, such as typical, a typical, and high risk floods, for the purpose of projecting flood risk. Numerous studies indicate that applying AI algorithms to the preprocessing of flood data can improve the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

final product. These methods have contributed to improved classification technique accuracy. In many applications, neural network engineering generally yields good results, but our investigations have demonstrated that random forest work yields the best outcomes when compared to benchmark models.

3. PROPOSED METHODS

Machine learning based method to improve accuracy and efficiency in order to get around the drawback in the current system. We used hadoop to store and retrieve data from the distributed hadoop file system in order to manage large amounts of data. Data can be loaded into a cluster using Hadoop. The random forest algorithm will be used to predict rainfall. Larger cloud service providers like Amazon Web Services, Microsoft, and Rack space already have risk reduction strategies and technology in place. DDOS attackers are drawn in by the cloud services. It's reasonable to predict that DDOS assaults on cloud services will increase in frequency as more people utilize them. In comparison to conventional composite materials and production processes, the hybrid SMC technology has many advantages in the realm of aircraft applications, especially for interior components. Cost effective techniques and materials, the ability to create parts with complicated 3D shapes, high repeatability, customized mechanical qualities, and quick lead times are the key advantages of SMC

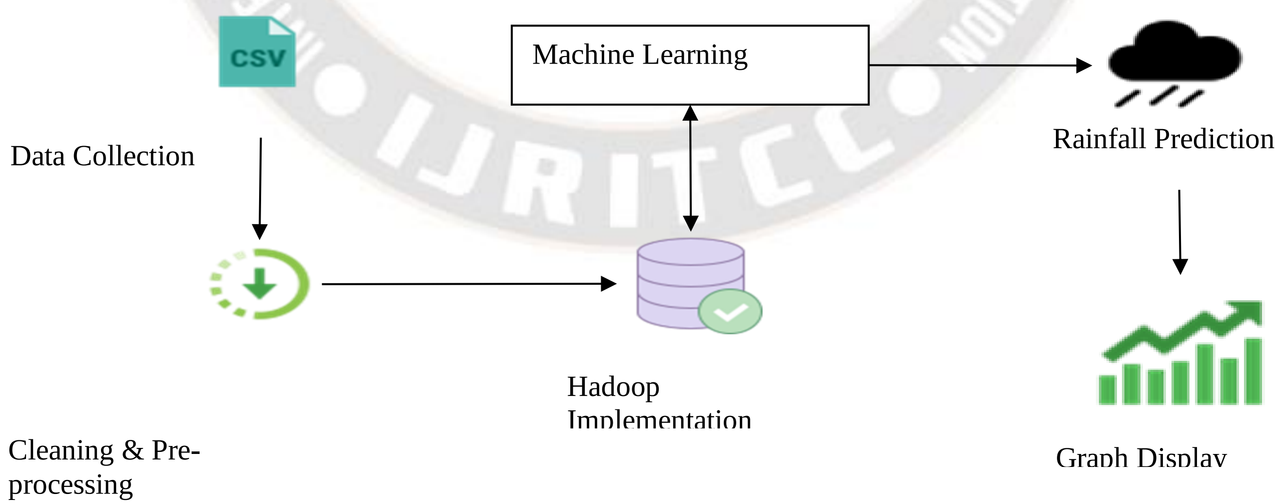
technology. A good strategy to lessen flood damage is to issue flood warnings. Forecasting uncertainties for floods using climate change and its effects on several economic sectors have become apparent.

The use of farming choice is contemporary information and communication technology in agriculture, such as machine learning algorithms as well as the rationalization of resource utilization in a capitalist system are all examples of smart farming. It also refers to the employment of cutting-edge technology in food production in sustainable and healthy ways. Many different applications of cutting-edge technologies are helping the majority of people around the planet. At this time, Big data analytics, data science, and the internet of things have all come to play an important role in people's daily lives, and expanding their capacity to modify their surroundings. In general, IoTs and data analysis are used in the agro industrial and environmental sectors to regulate smart agricultural systems and diagnose problems while also supplying crucial.

The process of changing raw data into something data preparation is the process of preparing data for use by a machine learning model. It is the initial and most important stage in the development of a machine learning system, mode. When working on machine learning projects, we almost always discover clean and prepared data. The dataset needed to develop a machine learning models, for example, will be different than the dataset needed for a patient with liver disease. Furthermore, several dataset formats may be utilized for varied aims; as a result, each dataset is distinct from the others. We usually convert datasets into CSV files before using them in our apps. Fig. 1.1. Shown the architecture of preprocessing architecture diagram is used to collect all the data set and stored in CSV files.

4. FLOOD FORECASTS PREPROCESSING IN NETWORKS CLASSIFICATION ARCHITECTURE

Fig 1.1 Leaf Disease Prediction Classification Architecture



International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

The gathering of data for the ML model's training is the core step in the machine learning pipeline. The quality of the training data determines how accurately ML systems can predict the future. The following are few problems that could arise while gathering data. The information acquired might not be pertinent to the problem statement, and there might not be enough data or sub data. This may appear as blank values in columns of unbalanced data or missing images for a specific class of prediction. For some classes or categories in the data, there may be an abnormally large or low number of associated samples.

They run the risk of not being adequately represented in the model. Data skew is depending on the selection of the data, subjects, and labels. Real world raw data and images are typically incomplete, inaccurate, and devoid of any distinct patterns or trends. They most likely contain numerous errors as well. In order to preprocess as well, in order to preprocess them into a model compatible machine learning data format method. They are collected and then preprocessed. Data cleaning is one of many techniques and steps used in preprocessing.

These manual and automatic techniques remove data imputations that were incorrectly inserted or categorized. For balancing or adding missing the bulk of machine learning frameworks contain tools and APIs in addition to data. The mean, median, and K-NN of the data in the required fields, as well as the standard deviation, are typically used to impute missing values. Organizations are able to gather enormous amounts of data using Hadoop, which can then be utilized to derive insights with enormous economic value for a variety of use cases, such as user segmentation, sentiment analysis, churn analysis, risk assessment, predictive maintenance, and fraud detection. However, setting up Hadoop may be incredibly difficult as well as time consuming it challenging to gather the insights. A variety of specialist IT and analytics skills are needed to use the ecosystem of technologies and open-source projects known as Hadoop for storage and processing organizations spend their time on the architecture rather than concentrating on creating business value since it might be challenging and time consuming to integrate all of these different Hadoop technologies. The majority of their time is spent by data scientists mastering the numerous skills needed to. The Hadoop distributed file system stores unstructured image data across several nodes. The model makes use of this data to train and make predictions. Given that Hadoop offers real time data analysis and processing power required to input data into the model is greatly decreased. The Hadoop MapReduce engine is essential to the project as a whole. Additionally, because stored data is replicated, the idea of data corruption is diminished. System failure is improbable because backup nodes are available in the case of a node failure.

5. METHODOLOGY

Based on sensor data, ML methods to forecast the magnitude of flood disaster have been used. In this paper, we use data from IoT flood sensors to compare the performance of popular machine learning approaches, and we after a flood, offer an ensemble strategy for a separate flood severity classification system occurs anywhere, the system makes use of data from local sensors. The purpose is to categorize the data collected by flood sensors into three groups of classes. Floods are classified as normal, abnormal, or high risk. The proposed system consists of a number of creations and trading using process data, and model evaluation using testing, that hasn't been seen before prediction of the value**Table 1.1 Sample data set**

6. Leaf Classification Techniques

Based on sensor data, different ML approaches have been employed to forecast how severe flooding catastrophes will be. This study compares the effectiveness of common machine learning (ML) techniques using they use data from IoT flood sensors to offer an collection strategy for a certain the severity of the flood. Assessment system for a given that floods can occur anywhere, the system makes use of local sensor data. The purpose is to categorize the flood sensors collected data; there are three types, such as typical, unusual, and high-risk floods. IoT sensor data collection, preprocessing representation of feature space data, configuration and training of classification models using processed data, and evaluation of the models employing, a test dataset that hasn't been seen before, are all the successive components of the proposed system. A description of the processing and analysis in succession pipeline used to classify flood sensor data is shown in the Fig.1.2 diagram.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

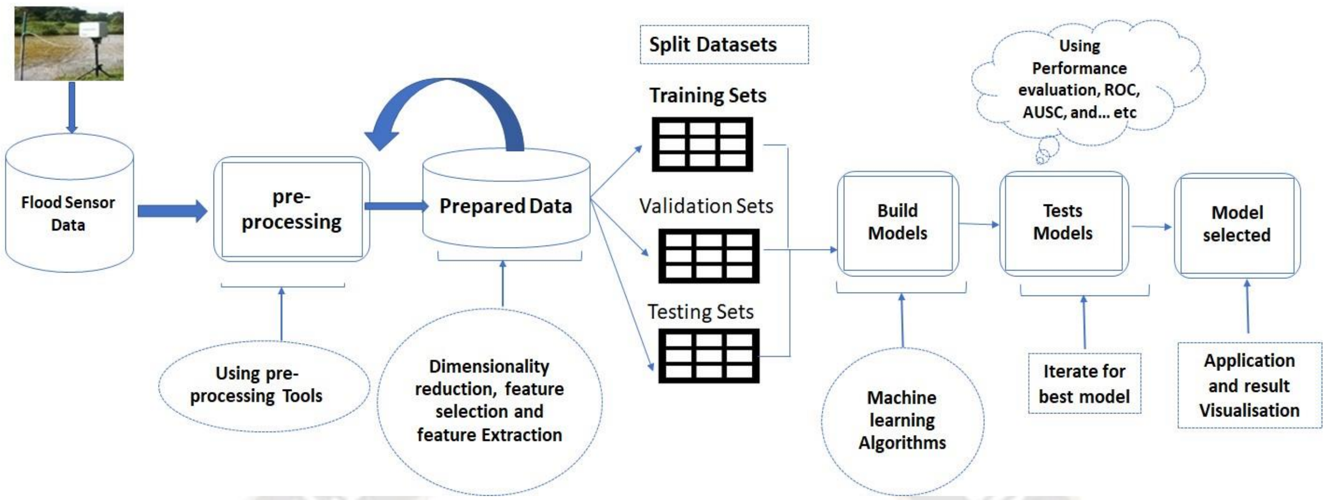


Fig.1.2 Overview of the Data Pre-Processing Diagram

For the purposes of training, validating, and testing ML models, the three partitions of the processed flood sensor dataset are as follows. Eleven features are used as inputs to three output classes of machine learning algorithms. Our system starts with the data collecting stage, where information on flood sensory data will be collected gathered. The preprocessing stage will be used to clean and normalize the collected data before it is sent on for analysis. Monsoonal rain, day length, number of fatalities, the number of displaced, the amount of snowmelt and ice jams, the size, the centroid of heavy rain, the entire area impacted, and the tropical storm are only a few of the 11 features that are gathered in this case. To categorize the severity of the flooding, trained machine learning algorithms are applied with the data that was cleaned and standardized according to the necessary standards.

7.DATA COLLECTION

The environment agency, which has a collection of datasets from different places throughout the world, is where the sensory data is gathered. As shown in the below Table 1.1, every dataset from the flood sensors sample comprises 11 variables that are considered crucial for estimating the severity. There are 4214 flood samples from the information from flood sensors used in this investigation are divided into their target classes, each of which indicates the intensity of the flood. Class one outputs had 1182 the class two had 308 data points for an abnormal water level, the class three had data points and the usual water level. had 457 data points, similarly for high risk water level. The aggregate function is used in statistical techniques combining data from several measures. Additionally, summary data replaces groupings of observations when data is aggregated.

1	Instances of Data	4214	Data for a 31-year period were gathered from the atmosphere agency's website.
2	Class of Variables	3	Course 1: Standard water level Course 2: Excessive water level Course 3: Unsafe water level
3	Features Attributes	11	Monsoonal rainstorm, the length of the day, the number of individuals who have died and the number of person who have been displaced. Somewhat, as well as Ice Jams Centroid, a lot of rain,. Total affected area Torrential Rain. The tropical storm.
4	Characteristics	5	correctness, understanding, accurateness and Specificity.
5	Revelation Techniques	5	Stochastic from the environment T-Distributed Neighborhood embedding. The curve agency's website includes receiver operating characteristics, principal component analysis of

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

			curves, and others, are all terms used in this context.
6	Assembly Classifier	2	ANN with RF,LSTM with RF, LSTM mode with RF, and LSTM with SVM.

Table 1.1 Table Sensor Dataset Description

8. EXPLORATORY ANALYSIS

In this crucial to properly prepare the flood sensor data that was previously provided using preprocessing method, such as purification and normalization, in order to produce reliable findings when using ML approaches. Better overall performance and prediction accuracy are obtained after using data normalization approaches, which rely on noise reduction and missing value management. In order to locate potential anomalies in pre-processed data, exploratory analysis was used. This is a crucial stage since it offers more information and improves how effectively models can use the data set to be learned from. A PCA based visualization of the flood sensor dataset serves as an illustration for this. PCA Fig.1.3 is frequently employed for dimensionality reduction and is very handy in a variety of applications. To prepare for projection into a lower dimensional space and remove potential correlations in feature space, it executes a linear mapping. The stochastic T-Scattered model neighborhood embed is another style of visualization, we used as seen in Fig.1.3 in actuality, this is a machine learning technique that, in contrast to PCA produces, a nonlinear projection of high dimensional data into a lower dimensional space, dimensional space. Dimensional subspace, often two or three dimensions. Patterns in the original feature space that resemble each other rare projected. Elsewhere. Adjacent in T-SNE, whereas different patterns are most likely at distant positions. Examining in Fig.1.4, we can see that patterns from various classes share a lot of similarities.

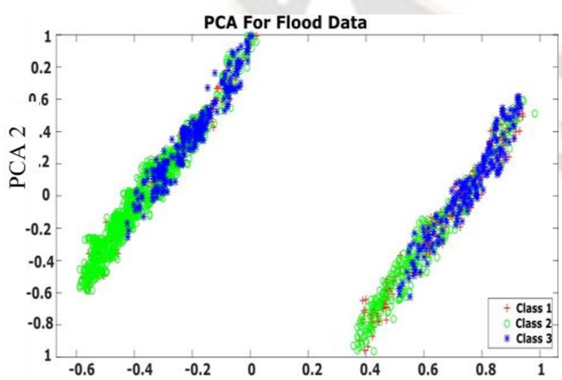


Fig 1.3 Projections of 3 Classes

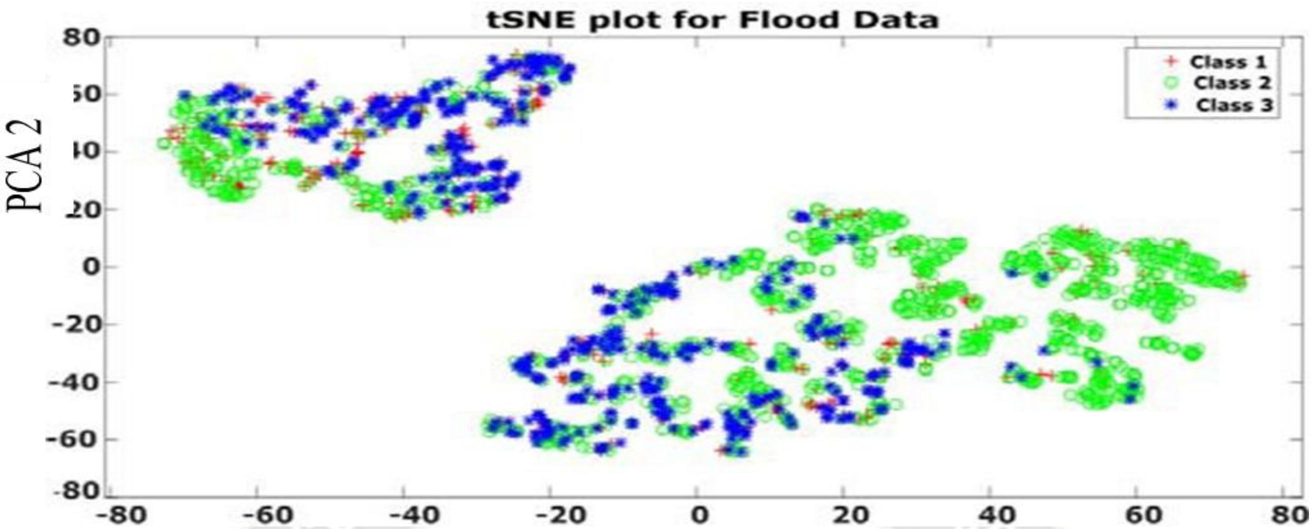
The analysis of standardized data requires the identification of features in order for our machine learning algorithms to train more quickly and avoid overfitting. In ML related problems, a variety of feature selection techniques have been utilized, together with key examples include component analysis, information gain, the chi-square test, and many other techniques other techniques others. The chi-square test ranks characteristics according to their statistical significance, implying that the current feature is affected by the target cl.

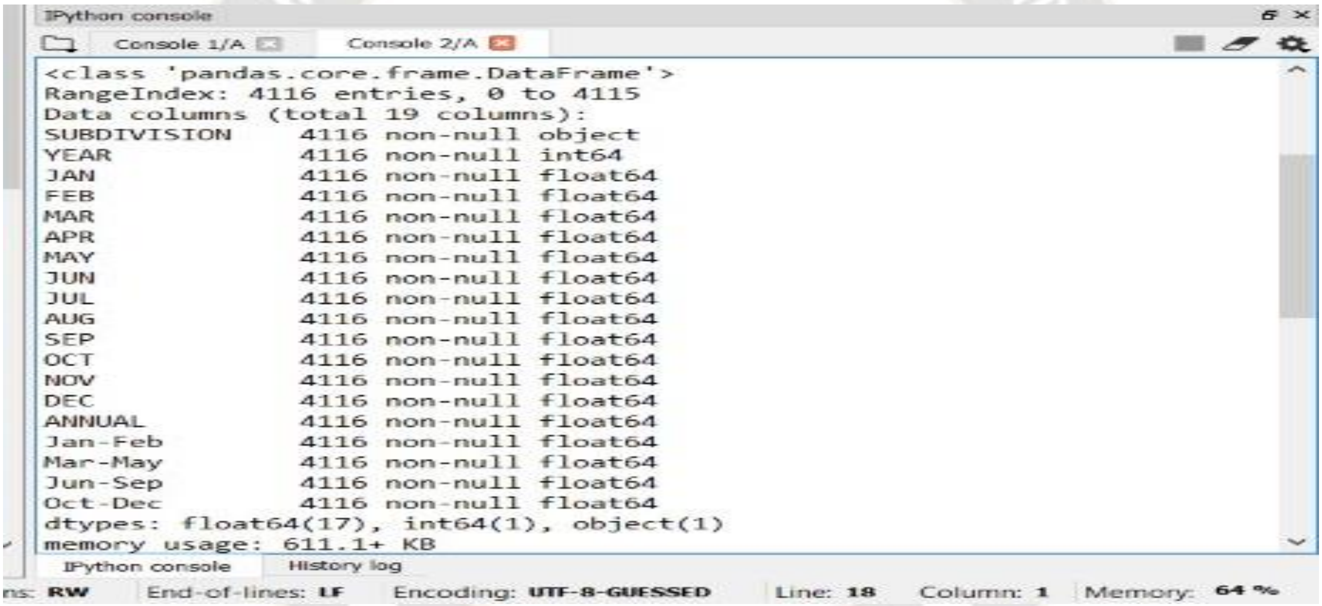
Table 1.2 Dataset value Preprocessing Techniques

Algorithm	**Dataset 1**	**Dataset 2**	**Dataset 3**	**Dataset 4**
RF	62.82	96.32	95.54	96.23
;R	6.86	85.76	89.4	95.49
LMM	49.3	76.57	88.71	81.63
AMM	37.56	89.34	95.38	87.59
NB	11.35	91.92	92.72	95.38
Proposed Algorithm	78.61	97.7	98.41	96.83

9.RESULT AND DISCUSSION

The latest three or four year's worth of rainfall data are compiled in a comma separated values (CSV) use, The month wise aggregate is included in the dataset. The dataset could include errors, negative numbers, or empty values. In the preprocessing, the dataset is cleaned. The preprocessing techniques involve eliminating incomplete records. We must get ready to send the clean dataset to the machine learning algorithm once it is available. Random forests or random decision forests generate a large number of decision trees during the training phase and output the class that is the mean of the classes or mean prediction of the classes or mean prediction of the individual trees as an ensemble learning technique for classification, regression, and other problems. The analysis of standardized data requires the identification of features in order for our machine learning algorithms to train more quickly and avoid overfitting. In ML related problems, a variety of feature selection techniques have been utilized, together with key examples include component analysis, information gain, the chi-square test, and many other techniques other techniques others.





```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4116 entries, 0 to 4115
Data columns (total 19 columns):
SUBDIVISION     4116 non-null object
YEAR            4116 non-null int64
JAN             4116 non-null float64
FEB             4116 non-null float64
MAR             4116 non-null float64
APR             4116 non-null float64
MAY             4116 non-null float64
JUN             4116 non-null float64
JUL             4116 non-null float64
AUG             4116 non-null float64
SEP             4116 non-null float64
OCT             4116 non-null float64
NOV             4116 non-null float64
DEC             4116 non-null float64
ANNUAL          4116 non-null float64
Jan-Feb         4116 non-null float64
Mar-May         4116 non-null float64
Jun-Sep         4116 non-null float64
Oct-Dec         4116 non-null float64
dtypes: float64(17), int64(1), object(1)
memory usage: 611.1+ KB
```


The above Fig 1.5 shows the average of rainfall in every month using the LST and random woodland are two examples, powerful data analysis classifiers that beat other models in terms of performance. Instead of using cross validation to improve these models used a model of decision trees-based training and testing results, out of bag technique Generally speaking ensemble classifier maintains the desirable qualities of decision trees, for example handling irrelevant or duplicate descriptors. In comparison to the ensemble classifier, out model was substantially faster during the training process. The ensemble classifier's ability to improve generalization leveraging the evidence from its member classifiers combined is a major factor in why it produced greater accuracy results. Despite the fact that certain samples found in the data sets have incorrect labels, the ensemble classifier is still able to predict. The Fig.1.6 shows the average month of the prediction values for June for the month of 1400 ranges to 3500 for that month alone highest rainfall occurs.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

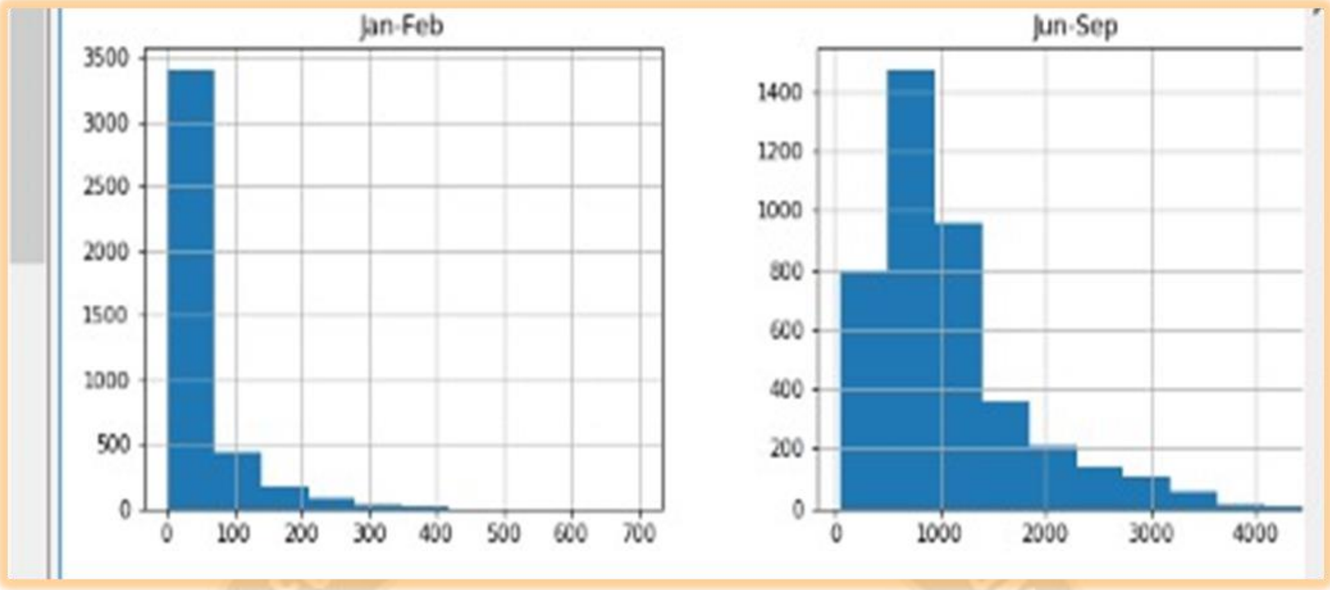


Fig 1.6 Average Rainfall using Bar graph

The above bar graph Fig 1.6 gives the monthly rainfall for the average and that, Individual classifiers fared well for the usual severity level class in terms of sensitivity and specificity, but did not perform well for the other two classes, such as average, for all severity level classes, ensemble classifiers produced excellent sensitivity and specificity, produced low sensitivity and specificity. This is brought on by the dataset's uneven representation of three classifications. In contrast, the ensemble model outperformed the individual models and demonstrated supremacy in the sensitivity specificity tradeoff for the above graph ranges from 200 to 3480 and its referred to as C2 and C3

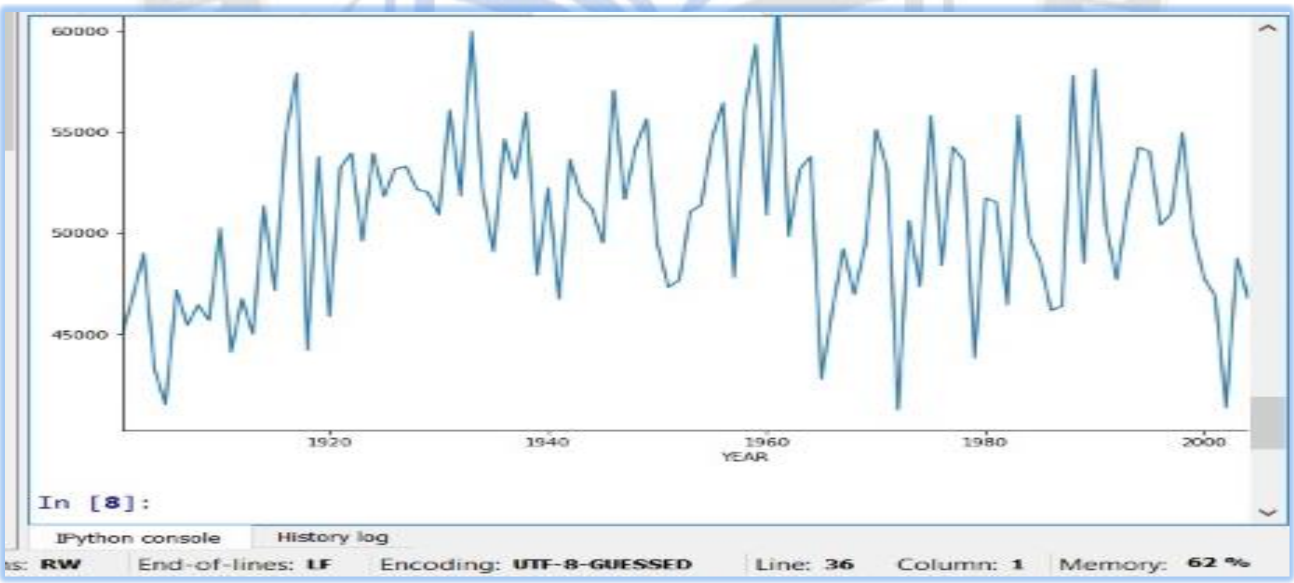


Fig. 1.7 Existing method for average rainfall using K-NN

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

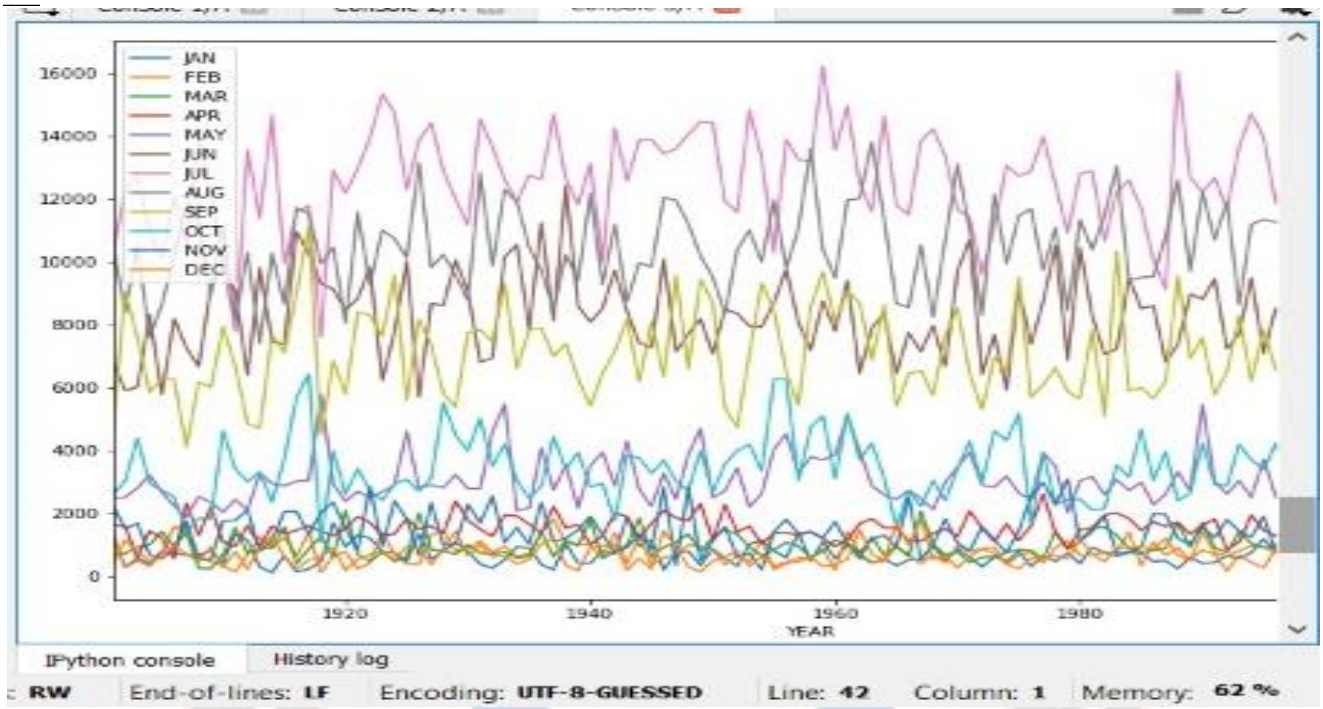


Fig. 1.7 Proposed method for average rainfall using Random Forest, CNN, Regression

In order to detect the severity of the water level, 11 parameters linked to flood sensor data from 4216 sensors records are analyzed in this study using an effective data science approach. The ensemble LSTM and RF classifier has an accuracy of 0.998 during the training and validation phases, but only 0.867 during 0.876 during the testing phase using fictional data, but only 0.874 during the training and validation stages. This classifiers ROC for training, and testing is displayed in the above graph 1.7 and 1.6 respectively. Additionally, during testing, the ensemble classifiers analyze data more quickly than single classifiers

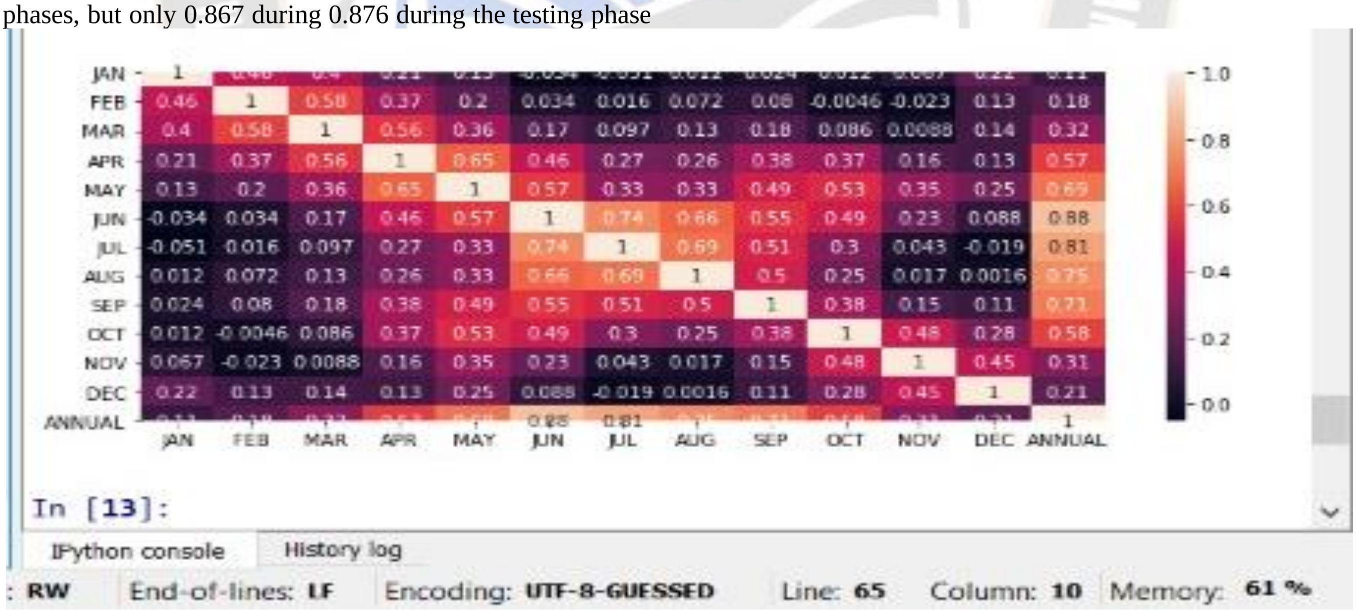


Fig. 1.8 Proposed method for average rainfall using LSTM Classifier

In the above Fig.1.8 is shows the several statistical in a precise classification and prediction test, various mathematical approaches are used to assess the effectiveness of a machine learning model, precision, recall ad accuracy, and f-score are all metrics, some of the performance evaluation criteria that are used to compare the proposed algorithms performance in comparisons to other algorithms. By taking into account True Negative (TN), True Positives (TP), False Positives (FP), and False Negatives (FN), the statements are accomplished. The percentage of the test set that the classifier successfully classifies indicates the model's classification accuracy on a given test. Several statistical and mathematical methods are utilized to evaluate some performance evaluation measures used to evaluate the performance of the proposed machine learning model include precision, recall, and the f-score, accurate classification and forecasting test, algorithm to that of other algorithms. By taking into account, the statements include true negatives; it is possible to achieve genuine positives, false positives, and false negatives. The percentage of the test set that the classifier successfully classifies indicates the classification model's accuracy on a given test.

10. CONCLUSION

In order to execute as inputs for data science methodologies for deterring river flood severity, the sensors and IoT platforms deployed rivers can be dangerous, leveraged. This paper's ensemble model produced good results for flood detection as well as tools for, flood warning in the future. This study takes into account three types of flood data, normal, abnormal, and dangerous water levels. The suggested ensemble machine learning solution is assessed using performance evaluation measures including sensitivity and specificity as well as visualization methods. The findings suggest that utilizing the right data science based on ensemble machine learning can provide early warning of flood severity approaches. Opportunity projects will utilize.

References

[1] Xiong, Lihua, and Kieran M. OConnor. "An empirical method to improve the prediction limits of the GLUE methodology in rainfall runoff modeling." Journal of Hydrology 349.1-2 (2008): 115-124.

[2] Deivendran.p Emotion Recognition for Challenged People Facial Appearance in Social using Neural Network. international Journal of Engineering Trends and Technology this link is disabled, 2022, 70(6), pp. 272–278

[3] Schmitz, G. H., and J. Cullmann. "PAI-OFF: A new proposal for online flood forecasting in flash flood prone catchments." Journal of hydrology 360.1-4 (2008): 1-14.

[4] Riordan, Denis, and Bjarne K. Hansen. "A fuzzy case-based system for weather prediction." Engineering Intelligent Systems for Electrical Engineering and Communications 10.3 (2002): 139-146

[5] Maruthi.R, A novel pixel shuffling and amalgamation scheme for visual secret sharing in cloud environment, International Journal of Engineering Trends and technology this link is disabled, 2021, 69(4), pp. 44–48

[6] Rathish Babu.T.K.S, An assessment of software defined networking approach in surveillance using sparse optimization algorithm, Computer Communications, 2020, 151, pp. 98–110.

[7] Selvakanmani.S, Analysis of taskable mobile IoT sensing systems for coverage and throughput, International Journal of System Assurance Engineering and Management this link is disabled, 2023.

[8] Guhathakurta, P. "Long-range monsoon rainfall prediction of 2005 for the districts and subdivision Kerala with artificial neural network." Current Science 90.6 (2006): 773-779.

[9] Deivendran.P, Scalability assurance process in replication and migration using cloud simulator, International Journal of Networking and Virtual Organizations link is disabled, 2019, 21(1), pp. 112–126.

[10] Selvakanmani.S, Deep learning approach to solve image retrieval issues associated with IOT sensors, Measurement: Sensors, 2022, 24, 100458.

[11] Pilgrim, D. H., T. G. Chapman, and D. G. Doran. Problems of rainfall-runoff modelling in arid and semiarid regions." Hydrological Sciences Journal 33.4 (1988): 379-400. 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence) 395 .

[12] Deivendran,P ,Scalability services in cloud computing using eyeos, Journal of Computer Science this link is disabled, 2015, 11(1), pp. 254–261

[13] Lee, Sunyoung, Sungzoon Cho, and Patrick M. Wong. "Rainfall prediction using artificial neural networks." Journal of geographic information and Decision Analysis 2.2 (1998): 233- 242.

[14] Charaniya, Nizar Ali, and Sanjay V. Dudul. "Committee of artificial neural networks for monthly rainfall prediction using wavelet transform." Business, Engineering and Industrial Applications (ICBEIA), 2011 International Conference on. IEEE, 2011. 64

[15] Shanmuganathan.C, Performance analysis of secure group key mechanism in Mobile Ad hoc networks, International Journal of Engineering and Technology (UAE), 2018, 7(4.10 Special Issue 10), pp. 344–348.

[16] Maruthi.R, Power management in server farms, International Journal of Innovative Technology and Exploring Engineering, 2019, 8(10), pp. 1135–1139.

[17] Ben-Bassat M (1982) Pattern recognition and reduction of dimensionality. Handbook of Statistics 2(1982):773–910

[18] Deivendran.P, Smart IoT based an Intelligent System for Needy People to Recognition Voice Detection of Obstacle, International Conference on Innovative Data Communication Technologies and Application, ICIDCA 2023 - Proceedings, 2023, pp. 814–818.

[19] Selvakanmani.S, Audio Segmentation Techniques and Applications Based on Deep Learning, Scientific Programming this link is disabled, 2022, 2022, 7994191.

[20] Rathish Babu.T.K.S, Improvement of the dependency structure of the software architecture model with risk estimation, Journal of Statistical Computation and Simulationthis link is disabled, 2016, 86(5), pp. 908–921.

[21] Hemdan, EED and Manjaiah, DH (2018). Springer, Cybercrimes investigation and intrusion detection in internet of things based on data science methods. In Cognitive Computing for Big Data Systems Over IoT (pp. 39–62).

[22] Shanmuganathan.C, enabling security in MANETs using an efficient cluster-based group key management with elliptical curve cryptography in consort with sail fish optimization algorithm, Transactions on Emerging Telecommunications Technologies, 2023, 34(3), e4717.

[23] Maruthi.R, Novel image blend scheme based visual cryptographic in cloud secret image sharing, international Journal of Scientific and Technology Research, 2020, 9(3), pp. 778–781.

[24] Rathish Babu.T.K.S, A new framework for software architecture generation and fault rectification in software engineering, International Journal of Applied Engineering Research, 2015, 10(16), pp. 37813–37819.

[25] Deivendran.P, Scalability and security requirements for the Internet of Things architecture, Artificial Intelligence for Internet of Things: Design Principle, Modernization, and Techniques, 2022, pp. 109–147.

[26]] Morais R, Silva N, Mendes J, Adão T, Pádua L, López-Riquelme JA, Pavón-Pulido N, Sousa JJ, Peres E (2019) My sense: A comprehensive data management environment to improve precision agriculture practices. Compute Electron Agric 162:882–894 .

[27] Deivendran.P, 2022, "Liver Infection Prediction Analysis Using Machine Learning to Evaluate Analytical Performance in Neural Networks by Optimization Techniques, International Journal of Engineering Trends and Technology, Volume 71 Issue 3,377-384, March2023,https://doi.org/10.14445/22315381/IJETT-V71I3P240.

[28] Deivendran.P, 2023 Liver Infection Prediction Analysis using Machine Learning to Evaluate Analytical Performance in Neural Networks by Optimization Techniques, International Journal of Engineering Trends and technology this link is disabled, 2023, 71(3), pp. 377–384.

[29] Shanmuganathan.C,Copy-paste forgery detection using deep learning with error level analysis, Multimedia Tools and Applicationsthis link is disabled, 2023.

[30] Deivendran.P, Data Analysis for Centella Asiatica Leaf Disease Prediction in Agriculture Using Machine Learning,, Journal of Coastal Life Medicine,JCLMM 2/11(2023) 406-418,ISSN2309-6152.

[31] Mafarja MM, Mirjalili S (2019) Hybrid binary ant lion optimizer with rough set and approximate entropy redacts for feature selection. Soft Compute 23(15):6249–6265.