



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 6, June 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 [ijirset@gmail.com](mailto:ijirset@gmail.com)

 [www.ijirset.com](http://www.ijirset.com)

# Campus Placements Prediction & Analysis Using Machine Learning

Mr.R.KARTHIKEYAN.,<sup>1</sup> Mr.S.NANDHAKUMAR.,<sup>2</sup>

Assistant Professor, Department of Master of Computer Application, Gnanamani College of Technology, Namakkal  
Tamilnadu, India<sup>1</sup>

PG Scholar, Department of Master of Computer Application, Gnanamani College of Technology, Namakkal,  
Tamilnadu, India<sup>2</sup>

**ABSTRACT:** Placement of students is one of the most important objectives of an educational institution. Reputation and yearly admissions of an institution invariably depend on the placements it provides its students with. That is why all the institutions, arduously, strive to strengthen their placement department so as to improve their institution on a whole. Any assistance in this particular area will have a positive impact on an institution's ability to place its students. This will always be helpful to both the students, as well as the institution. In this study, the objective is to analyze previous year's student's data and use it to predict the placement chance of the current students. This model is proposed with an algorithm to predict the same. Data pertaining to the study were collected from the same institution for which the placement prediction is done and also suitable data pre-processing methods were applied. This proposed model is also compared with other traditional classification algorithms such as Decision tree and Random forest with respect to accuracy, precision and recall. From the results obtained it is found that the proposed algorithm performs significantly better in comparison with the other algorithms mentioned.

**KEYWORDS:** Naïve Bayes classifier, Classification algorithms, ID3, Bayes Net, RBF network, J48, algorithms.

## I. INTRODUCTION

Placements are considered to be very important for each and every college. The basic success of the college is measured by the campus placement of the students. Every student takes admission to the colleges by seeing the percentage of placements in the college. Hence, in this regard the approach is about the prediction and analyses for the placement necessity in the colleges that helps to build the colleges as well as students to improve their placements.

In Placement Prediction system predicts the probability of a undergrad students getting placed in a company by applying classification algorithms such as Decision tree and Random forest. The main objective of this model is to predict whether the student he/she gets placed or not in campus recruitment. For this the data consider is the academic history of student like overall percentage, backlogs, and credits. The algorithms are applied on the previous year's data of the students.

## II. LITERATURE SURVEY

In this paper [2], used Logistic regression technique on their college placement dataset which got 83.33% of accuracy. Authors in [4] ID3, J48, REP Tree, NB Tree, MLP, Decision Table Classification techniques on the placement dataset collected from their college. The results had shown that ID3 predicted well among them with an accuracy of 82.1% . In this paper [5], used C4.5 classification technique on the placement dataset which was collected from their college which got 80% of accuracy.

## III. EXITING SYSTEM

NaiveBayes, OneR classification techniques on the data gathered from their high school. The results had shown that J48 and Simple Cart predicted well among them with an accuracy of 79.61% [5]. Ajay Kumar Pal and Saurabh Pal collected the data for the study and analysis of the student's educational performance basically for training and

placement. The authors used different classification algorithm and used WEKA data mining tool. They concluded that naive Bayes classification model is the better algorithm based on the placement data with found accuracy of 86.15% and overall time taken to build the model is at 0 sec. As compared with others Naïve Bayes classifier had lowest average error i.e. 0.28.

Built the prediction model to improve the placement of the students . They used WEKA as the data mining tool to build the model using random tree algorithm. They also used ID3, Bayes Net, RBF network, J48, algorithms on the student data set. They resolved that the RT (Random Tree) algorithm is more accurate with 73% for the classification/prediction of the model. The accuracy using ID3 and J48 is 71%. Bayes Net is 70%.

#### Disadvantages of existing system

- ❖ The system is not implemented Attribute selection which is not relevant to each other.
- ❖ The system is not implemented cleaning missing values.

### IV. PROPOSED SYSTEM

In Placement Prediction system predicts the probability of a undergrad students getting placed in a company by applying classification algorithms such as Decision tree and Random forest. The main objective of this model is to predict whether the student he/she gets placed or not in campus recruitment. For this the data consider is the academic history of student like overall percentage, backlogs, credits. The algorithms are applied on the previous years data of the students.

#### Advantages of proposed system:-

Some of the attributes in the initial dataset that was not pertinent (relevant) to the experiment goal were ignored. The attributes name, roll no, credits, backlogs, whether placed or not, b.tech %, gender are not used.

The random forest algorithm can also be thought of as an ensemble method in machine learning. The input to a random forest algorithm is a dataset consisting of records, with attributes. Random subsets of the input are created.

### V. SYSTEM METHODOLOGY

**A. Data Acquisition:** The campus placement dataset is collected from Kagglewebsite. Here is the link for the dataset The dataset consists of various attributes such as Serial Number, Gender, SSC percentage, SSC Board - Central/ Others, HSC percentage, HSC Board, HSC Specialization, Degree Percentage, UG Degree Stream, Work Experience, E -test Percentage, Degree Specialization, Degree Percentage, Placement Status & Salary. The size of dataset is 19.71 KB & it has total 215 records.

Handling missing values: In our dataset missing values are present only in the salary column as these values correspond to the students who didn't get placed in any placement drive. So it is assumed that the missing values in Salary Column are Zero & replaced them by zero using `fillna(0,inplace=True)` function in Python. Handling categorical data: Since we cannot deal with categorical values directly, mapping is done for attributes having categorical values.

Gender attribute has values M (Male) & Female (F). Here, M is replaced by 0 & F is replaced by 1. SSC & HSC Board attributes has values 'Central' & 'Other.' Here, Central is replaced by 1 & Other is replaced by 0. Work Experience attribute has values 'Yes' & 'No'. Here, 'Yes' is replaced by 1 and 'No' is replaced by 0. Degree specialization attribute has values 'Marketing & Finance' & 'Marketing & HR'. Here, 'Marketing & Finance' is replaced by 1 and 'Marketing & HR' is replaced by 0. Status attribute has values 'Placed' and 'Not Placed'. Here, 'Placed' is replaced by 1 and 'Not Placed' is replaced by 0. This is achieved through map function in Python.

For e.g.,

- `df['gender']=df['gender'].map({'M':0,'F':1})`
- `df['ssc_b']=df['ssc_b'].map({'Central':1,'Others':0})`
- `df['workex']=df['workex'].map({'Yes':1,'No':0})`

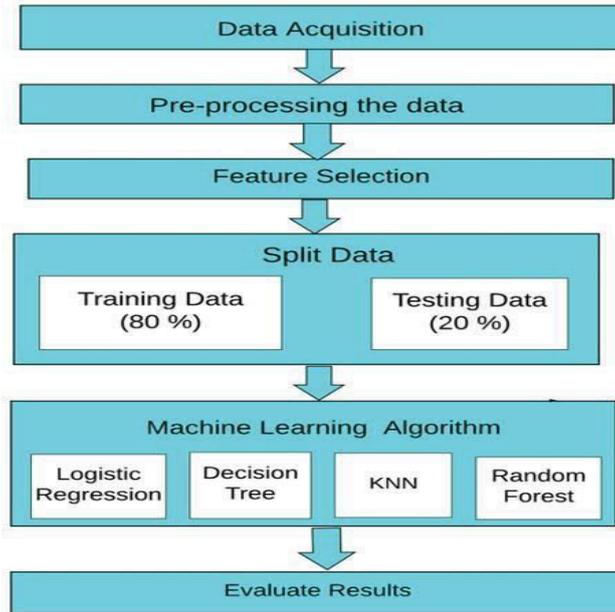


Fig. 5.1. Architecture Diagram

**Feature Selection:** Here, various features are visualized to understand their correlation with the target feature.

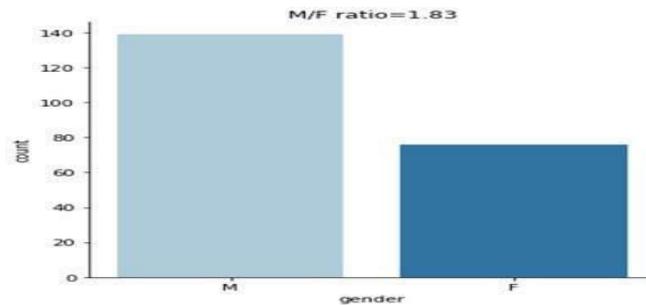
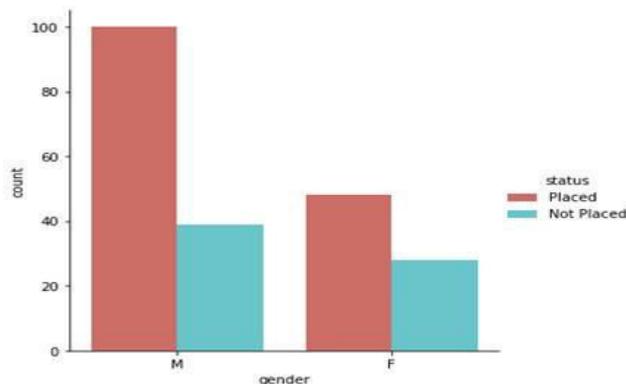


Fig.5.2. M/F ratio

Here, male : female ratio for one batch of students is approximately equal to 2. It means that there are 2 male candidates appearing for placement drives for every 1 female candidate.



From the above graph it can be concluded that the count of placed male candidates in a batch is higher as compared to

female candidates & the placement count is dependent on gender.

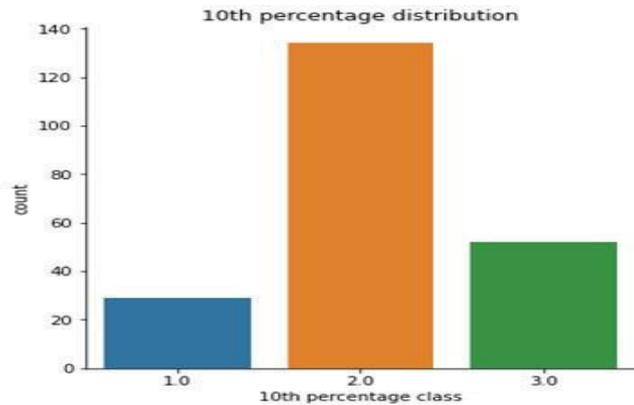


Fig.5.3. 10th standard percentage distribution

In the above graph, class 1 represents students having scores between 80-100%, class 2 represents students having scores between 60-80% and class 3 represents students having less than 60% score in 10th standard.

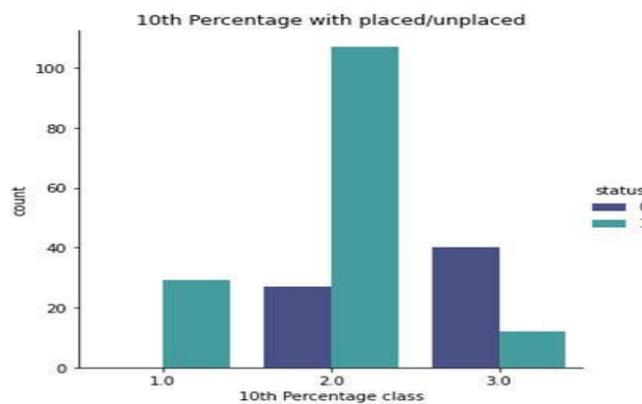


Fig. 5.4 Placement count vs. 10th percentage

From the above graph, it's observed that all the students having scores between 80-100% in 10th standard got placed. Very few students having scores between 60-80% in 10th standard couldn't get placed. Whereas, most of the students having below 60% score in 10th standard couldn't get placed.

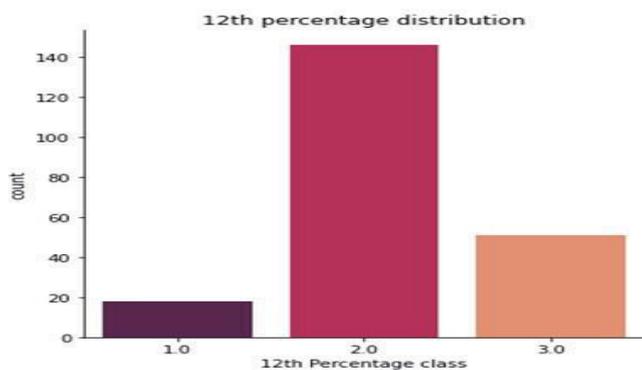


Fig. 5.5. 12th standard percentage distribution

In the above graph, class 1 represents students having scores between 80-100% , class 2 represents students having scores between 60-80% and class 3 represents students having less than 60% score in 12th standard.

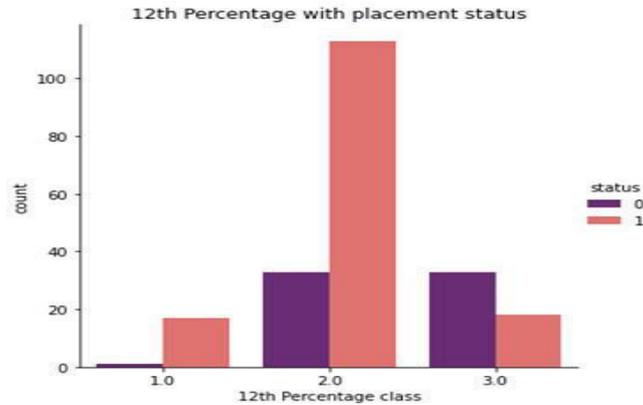


Fig. 5.6. Placement count vs. 12th percentage

From the above graph, it's observed that all the students having scores between 80-100% in 12th standard got placed. Very few students having scores between 60-80% in 12th standard couldn't get placed. Whereas, most of the students having below 60% score in 12th standard couldn't get placed.

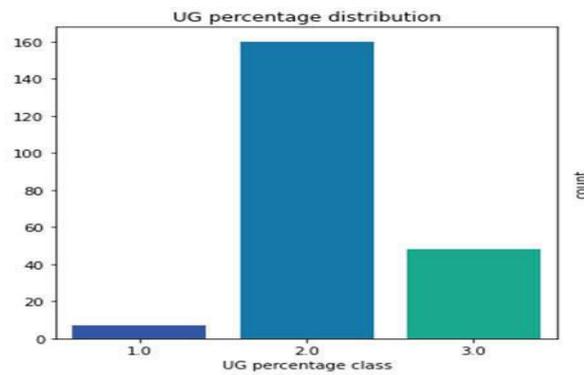


Fig.5.7. UG percentage distribution

In the above graph, class 1 represents students having scores between 80-100%, class 2 represents students having scores between 60-80% and class 3 represents students having less than 60% score in UG degree.

From the above graph, it's observed that most of the students having scores between 80-100% in UG got placed. Very few students having scores between 60-80% in UG couldn't get placed. Whereas, most of the students having below 60% score in UG couldn't get placed.

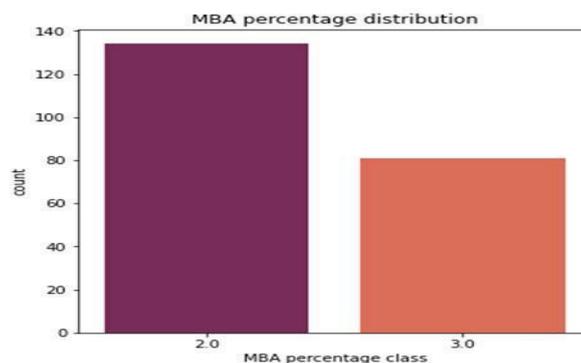


Fig.5.8. MBA percentage distribution

After studying MBA percentage data it is observed that no student has secured more than 80% marks. So the class 1 data isn't available for percentage of MBA.

## VI. SYSTEM IMPLEMENTATION

Hence, it is clear that placement count of the students is dependent on various features such as Gender, SSC percentage, SSC Board - Central/ Others, HSC percentage, HSC Board, HSC Specialization, Degree Percentage, UG Degree Stream, Work Experience, E-test Percentage, Degree Specialization, Degree Percentage.

➤ **Split data:** Here, data is divided into two parts i.e. training data & testing data. Where 80 % data is taken for training our machine learning algorithm and remaining 20 % data is used for testing whether our trained machine learning model is working correctly or not.

### Machine Learning Algorithm:

➤ **Logistic Regression:** Logistic regression is a statistical method used to determine the outcome of a dependent variable (y) based on the values of independent variable (x). In our problem dependent variable is placement status and independent variables are the features selected by us in the previous step. This algorithm is mostly used for the problems of binary classification.

➤ **Decision Tree:** A decision tree is a graph like a tree where nodes represent the position where we select the feature and ask a question, edges represent the answers of the question; and the leaves represent the final output or label of the class.

➤ **KNN:** K-NN stores all the training data into different classes based on the class labels and classifies new data by checking its similarity with data in the available classes.

➤ **Random Forest:** Random Forest classifier consists of a number of decision trees which apply on different subsets of our dataset and the average of outputs of all the decision trees is taken to improve the accuracy of output prediction.

**Evaluate results:** Accuracy is calculated by following formula,  $Accuracy = (TP + TN) / (TP + FP + TN + FN)$

Where,

TP: True Positive (the number of cases correctly identified as placed)

TN: True Negative (the number of cases correctly identified as unplaced).

FP: False Positive (the number of cases incorrectly identified as placed)

FN: False Negative (the number of cases incorrectly identified as unplaced)

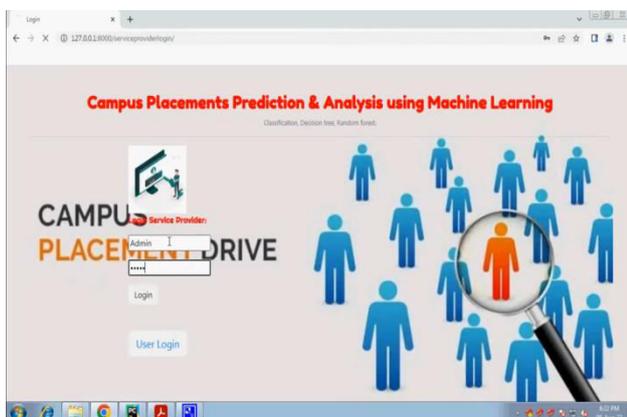


Fig: 6.1 Login Page

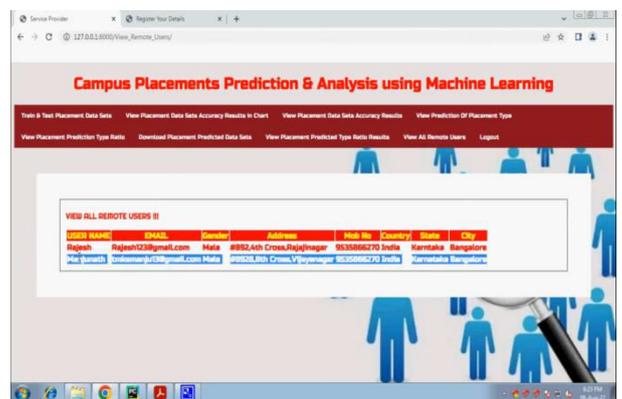


Fig: 6.2 View All Remote Users

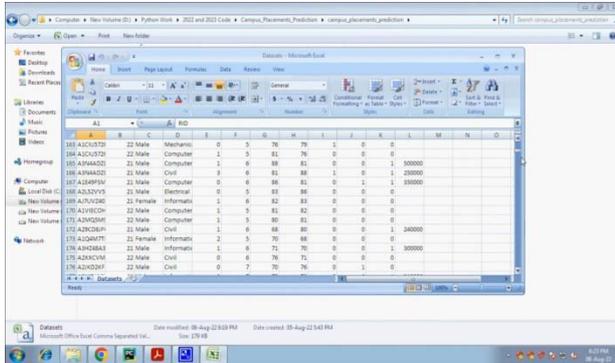


Fig: 6.3 Excel Sheet Deliver System

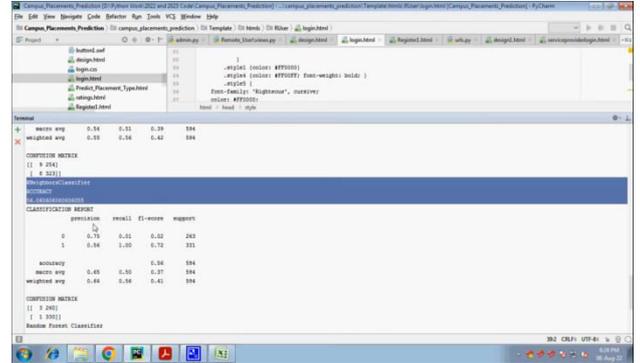


Fig: 6.3 Terminal

### VII. FUTURE ENHANCEMENT

Accuracy may further increase by application of more advanced techniques such as deep learning & experimenting with different activation functions of neural networks such as linear, sigmoid, tan h & ReLU. We can also experiment with different cross validation techniques such as 3 Fold, 5 Fold, 10 Fold, 15 Fold cross validation in order to analyze the change in accuracy.

### VIII. CONCLUSION

The campus placement activity is incredibly a lot of vital as institution point of view as well as student point of view. In this regard to improve the student’s performance, a work has been analyzed and predicted using the classification algorithms Decision Tree and the Random forest algorithm to validate the approaches. The algorithms are applied on the data set and attributes used to build the model. The accuracy obtained after analysis for Decision tree is 84% and for the Random Forest is 86%. Hence, from the above said analysis and prediction it’s better if the Random Forest algorithm is used to predict the placement results.

### REFERENCES

1. R.Karthikeyan, & et all "Biometric for Mobile Security" in the international journal of Engineering Science & Computing, Volume7,Issue6, June 2017, ISSN(0):2361-3361,PP No.:13552-13555.
2. R.Karthikeyan, & et all "Data Mining on Parallel Database Systems" in the international journal of Engineering Science & Computing, Volume7,Issue7, July 2017, ISSN(0):2361-3361,PP No.:13922-13927.
3. R.Karthikeyan, & et all "Ant Colony System for Graph Coloring Problem" in the international journal of Engineering Science & Computing, Volume7,Issue7, July 2017, ISSN(0):2361-3361,PP No.:14120-14125.
4. R.Karthikeyan, & et all "Classification of Peer –To- Peer Architectures and Applications" in the international journal of Engineering Science & Computing, Volume7,Issue8, Aug 2017, ISSN(0):2361-3361,PP No.:14394-14397.
5. R.Karthikeyan, & et all "Mobile Banking Services" in the international journal of Engineering Science & Computing, Volume7,Issue7, July 2017, ISSN(0):2361-3361,PP No.:14357-14361.
6. R.Karthikeyan, & et all "Neural Networks for Shortest Path Computation and Routing in Computer Networks" in the international journal of Engineering and Techniques, Volume 3 Issue 4, Aug 2017, ISSN:2395-1303,PP No.:86-91.
7. R.Karthikeyan, & et all "An Sight into Virtual Techniques Private Networks & IP Tunneling" in the international journal of Engineering and Techniques, Volume 3 Issue 4, Aug 2017, ISSN:2395-1303,PP No.:129-133.
8. R.Karthikeyan, & et all "Routing Approaches in Mobile Ad-hoc Networks" in the International Journal of Research in Engineering Technology, Volume 2 Issue 5, Aug 2017, ISSN:2455-1341, Pg No.:1-7.
9. R.Karthikeyan, & et all "Big data Analytics Using Support Vector Machine Algorithm" in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 6 Issue 9, Aug 2018, ISSN:2320 - 9798, Pg No.:7589 - 7594.
10. R.Karthikeyan, & et all "Data Security of Network Communication Using Distributed Firewall in WSN " in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 6 Issue 7, July 2018, ISSN:2320 - 9798, Pg No.:6733 - 6737.

11. R.Karthikeyan, & et all “An Internet of Things Using Automation Detection with Wireless Sensor Network” in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 6 Issue 9, September 2018, ISSN:2320 - 9798, Pg No.:7595 - 7599.
12. R.Karthikeyan, & et all “Entrepreneurship and Modernization Mechanism in Internet of Things” in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 7 Issue 2, Feb 2019, ISSN:2320 - 9798, Pg No.:887 - 892.
13. R.Karthikeyan & et all “Efficient Methodology and Applications of Dynamic Heterogeneous Grid Computing” in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 7 Issue 2, Feb 2019, ISSN:2320 - 9798, Pg No.:1125 -1128.
14. R.Karthikeyan & et all “Entrepreneurship and Modernization Mechanism in Internet of Things” in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 7 Issue 2, Feb 2019, ISSN:2320 - 9798, Pg No.:887– 892.
15. R.Karthikeyan & et all “Efficient Methodology for Emerging and Trending of Big Data Based Applications” in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 7 Issue 2, Feb 2019, ISSN:2320 - 9798, Pg No.:1246– 1249.
16. R.Karthikeyan & et all “Importance of Green Computing In Digital World” in the International Journal of Innovative Research in Computer and Communication Engineering, Volume 8 Issue 2, Feb 2020, ISSN:2320 - 9798, Pg No.:14 – 19.
17. R.Karthikeyan & et all “Fifth Generation Wireless Technology” in the International Journal of Engineering and Technology, Volume 6 Issue 2, Feb 2020, ISSN:2395–1303.
18. R.Karthikeyan & et all “Incorporation of Edge Computing through Cloud Computing Technology” in the International Research Journal of Engineering and Technology, Volume 7 Issue 9, Sep 2020 ,p. ISSN:2395–0056, e. ISSN:2395–0072.
19. R.Karthikeyan & et all “Zigbee Based Technology Appliance In Wireless Network” in the International Journal of Advance Research and Innovative Ideas in Education, e.ISSN:2395 - 4396, Volume:6 Issue: 5 , Sep. 2020. Pg.No: 453 – 458, Paper Id:12695.
20. R.Karthikeyan & et all “Automatic Electric Metering System Using GSM” in the International Journal of Innovative Research in Management, Engineering and Technology, ISSN: 2456 - 0448, Volume:6 Issue: 3 , Mar. 2021. Pg.No: 07 – 13.
21. R.Karthikeyan & et all “Enhanced the Digital Divide Sensors on 5D Digitization” in the International Journal of Innovative Research in Computer and Communication Engineering, e-ISSN: 2320 – 9801, p-ISSN: 2320 - 9798, Volume:9 Issue: 4 , Apr. 2021. Pg.No: 1976 – 1981.
22. R.Karthikeyan & et all “Crop Yield Prediction Based On Indian Agriculture Using Machine Learning” in the International Journal Of Engineering and Techniques, ISSN: 2395-1303, Volume:8 Issue: 4 , July. 2022. Pg.No: 11 – 22.
23. R.Karthikeyan & et all “College Bus Transport Management Web Application” in the International Journal Of Multidisciplinary Research In Science, Engineering and Technology, ISSN: 2582-7219, Volume: 6 Issue: 6, June. 2023. Pg.No: 1619 – 1625.
24. R.Karthikeyan & et all “Face Recognition Based Attendance System” in the International Journal of Innovative Research in Computer and Communication Engineering, ISSN: e 2320-9801, Volume: 11 Issue: 6, June. 2023. Pg.No: 8710 – 8717.
25. R.Karthikeyan & et all “Cloud data Deduplication System using per File parity and File Name Interpreter” in the International Journal of Advanced Research in Arts, Science, Engineering and Management, ISSN: 2395-7852, Volume: 10 Issue: 3, May. 2023.
26. R.Karthikeyan & et all “Secure Photo Sharing Social Networks Using Coverless Image Steganography Techniques” in the International Journal of Research in Science, Engineering and Technology, e - ISSN: 2319-8753, Volume: 12 Issue: 6, June. 2023.
27. R.Karthikeyan & et all “Hotel Booking Mobile And Web Application” in the International Journal Of Multidisciplinary Research In Science, Engineering and Technology, ISSN: 2582-7219, Volume: 6 Issue: 6, June. 2023.



Impact Factor: 8.423



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details