



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: III Month of publication: March 2021

DOI: <https://doi.org/10.22214/ijraset.2021.33520>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Data Analysis of Customer Complaints of Loan Approval and Financing in Banking Industry using SVM Classification

A. P. Gopu¹, S. Manikandan², M. Nandhagopal³, V. Boopathi⁴

¹Assistant Professor, ^{2,3,4}UG Students – Final Year, Department of Computer Science and Engineering, Nandha College of Technology, Perundurai, Tamilnadu, India.

Abstract: This paper focus on discovering and analysing Customer Finance Complaints data, to find how many similar complaints are there in relation to the same bank or service or product. These datasets fall under the complaints of Credit reporting, Hypothecation, Dues Collection, Customer Loan and Finance Accounting. By using data mining techniques, cluster analysis as well as predictive modelling is applied to obtain valuable information about complaints in certain areas of the Country. The banks that are receiving customer complaints filed against them will analyse the complaint data to provide results on where the most complaints are being filed, what products services are producing the most complaints and other useful data. Our model will contribution banks in identifying the location and types of mistakes for resolution, important to increased customer satisfaction to drive income and profitability.

Keywords: Customer, Complaints, analysis, clustering, predictive.

I. INTRODUCTION

As we are awake that in today's modern world peoples are more into business, so receiving a complaint from a customer happens almost every day. A customer's complaints present bank or reporting agency with a prospect to detect and resolve specific problems with their current product and service. Service complaints management is a critical part of business management. A good complaint-management strategy will result in best customer relationship with least human-resource investment and so hope to find a connection between bank, and customers to refine bank applications to better accommodate customer needs. Progressively banks are identifying the value of a customer complaint in that it is feedback on their experience, and an opportunity to not only resolve a problem for that particular customer but feasibly also for a much larger number of customers and that leads to unavoidable amounts of data that has to be analysed and specific functions are used to aggregate the analysis results. A customer's complaints present bank or reporting agency with an opportunity to identify and rectify specific problems with their current product or service. The banks that are receiving customer complaints filed against them will analyse the complaint data to provide results on where the most complaints are being filed, what products services are producing the most complaints and other useful data. This project assists banks in identifying the location and types of errors for resolution, leading to increased customer satisfaction to drive income and profitability. This project finds a connection between complaints, banks and customers to refine bank applications to better accommodate customer needs using a hybrid approach of hierarchal and k-means clustering.

In addition, using SVM classification, the complaints sentiment values are analysed and classified into positive or negative reviews. The project is designed using R Studio 1.0 as front end. The coding language used is R version 3.4.4

II. LITERATURE REVIEW

The number of studies has been conducted regarding the services to customers and their awareness. As such, we have reviewed some of them.

- 1) Venkata Rao Maddumala(2020) determined accuracy of loan approval based on machine learning technique successfully.
- 2) Soni P M, Varghese Pau l(2019) predicted a loan credibility system using data mining concept successfully.
- 3) Kamakodi (2007) concluded that modern day generation is influenced by the computation features used by banks and so the banks study about factors influencing their preferences. Residence relocation, salary fluctuation and unavailability banking based services are reasons enough to change bank.
- 4) Uppal and Kau r(2007) determined how consumer's awareness of web domains used by banks and gave some measures to make these applications more successful. They concluded that the limitation about today's web domain application is spreading the awareness about the varied features offered.
- 5) Mishra and Jain (2007) took up dimensions of consumer satisfaction in national and private banks. The study talks about how satisfaction is the foremost asset to the organization, which provides unmatched competitive edge that helps achieving loyalty of a customer.

- 6) Jain and Jain (2006) demonstrated that the banking sector, both private and public have suffered radical as well as revolutionary changes due to the liberalization act of 1991. Retail banking is the consumer preferred choice which articulates itself responses received from 200 customers of HDFC bank, ICICI bank and some other banks in the city of Varanasi, Uttar Pradesh and he looked upon the schemes offered by the banks, quantized satisfaction in different types of services, expectations about these schemes and the height of segmentation among the services offered.
- 7) Singh (2006) discusses CRM approaches in various banks. He emphasized on how the management targets customers in order to gain insight and gives out value added services and products. Web as provided a smooth user experience, giving access to the various features used by the customers thereby achieving customer satisfaction. Management has to strive to ensure end to end delivery and ensure customer satisfaction which is essential to the banks in terms of maintaining high regards and loyalty obtained from customers.
- 8) Bhaskar (2004) computed that expansion of banking is directly proportional to the quality of services provided by the banks and satisfaction is regarded highly as customer's feedback is the only thing to lean on, when it comes to the highly competitive banking industry. Arguably, India's banking industry is highly thriving and depends heavily on customer morale and loyalty.
- 9) Furthermore, Hasanbanu (2004) stated how the rural India is unaware about various schemes and benefits offered by the banks in order to ensure financial welfare. The majority of rural population is inaccessible to the web domain services of the banks and continue to prefer local moneylenders charging light interest rates. The study was conclusive and based on the data provided by the RBI, however, it is based on the questionnaire and surveys.
- 10) Although Singh (2004) spurred about the reality of banks in terms of providing customer support and found out that the customers are influenced by the banks location and the minutest detail of the banking details including the banking interest rates as well as attitudes and customer support provided by the personnel.

III. METHODOLOGY AND PROCEDURE

The following modules are present in the project.

A. K-Means Clustering

In this module, K-Means clustering of bank customer data are given as input to analyse and group the customers. The customer data set is taken with Customer Id, Annual Income, Loan Type, Interest Rate, Duration in Months, Loan Amount Requested and Loan Status columns of which Interest Rate and Duration in Months are taken as X and Y Axis for K- Means clustering. By default 3 is given for K, but we can give any number to cluster the data.

1. Initialize cluster centroids $\mu_1, \mu_2, \dots, \mu_k \in \mathbb{R}^n$ randomly.
2. Repeat until convergence: {

For every i, set $C^{(i)} := \operatorname{argmin}_j \|x^{(i)} - \mu_j\|^2$

For each j, set $\mu_j := \frac{\sum_{i=1}^m 1\{c^{(i)}=j\} x^{(i)}}{\sum_{i=1}^m 1\{c^{(i)}=j\}}$

B. SVM Classification

“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well.

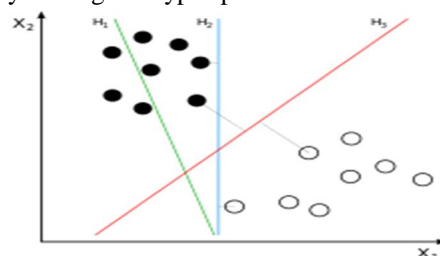


Figure 1 H_1 does not separate the classes. H_2 does, but only with a small margin. H_3 separates them with the maximal margin.

C. Neural Network Algorithm

A neural network is a series of algorithms that endeavours to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature.

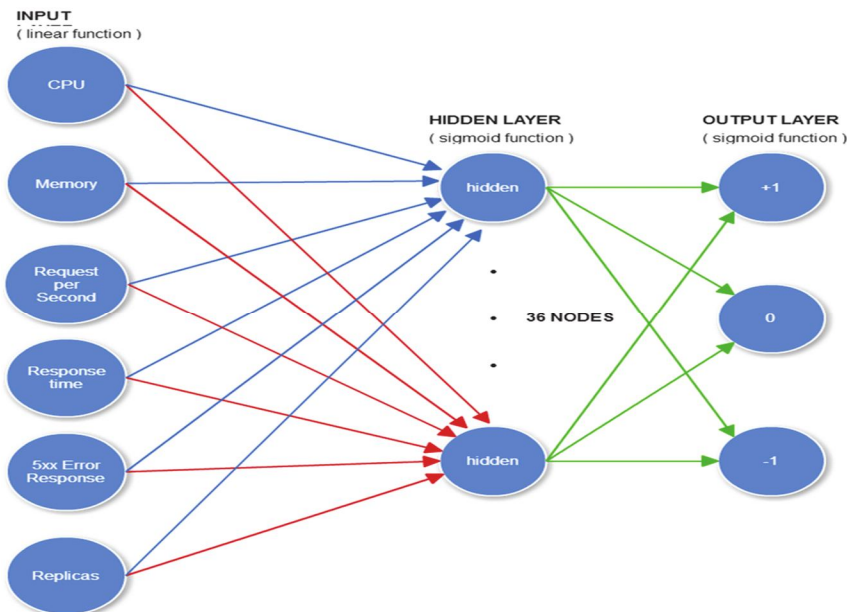


Figure2 The artificial neural network that is used in our scaling algorithm.

D. Similarity Percent Finding

To find the nearest matching record values among the given training record set, hot encoding vector comparison is also made so that similarity percent of the test record in the overall training dataset records are also found out. The test record is classified into maximum similarity matching record.

IV. RESULTS

By performing K- Means, support vector machine (SVM) for positive negative analysis and neural network algorithm to find similarity percentage, we got a better insight by having 3 clusters. Figure1 shows clusters in having the highest density while the rest of the have a lower density in comparison. Figure2 plots below show us the results for our clustering model.

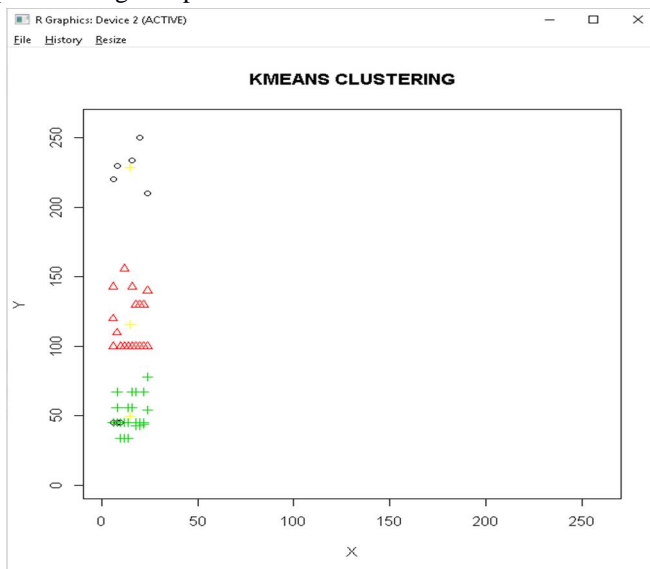


Figure3K-Means clustering

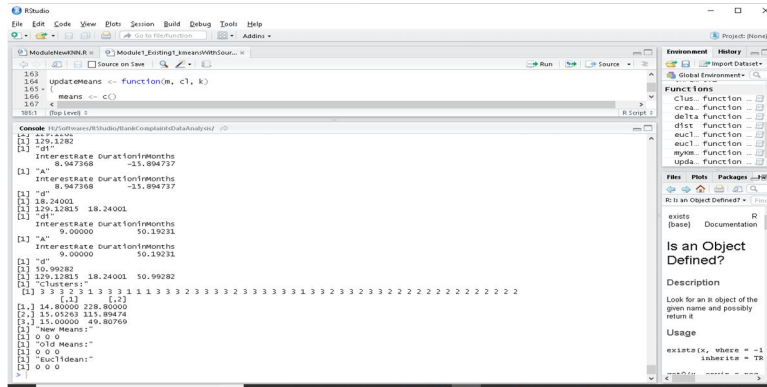


Figure 4 Records Grouped In Clusters

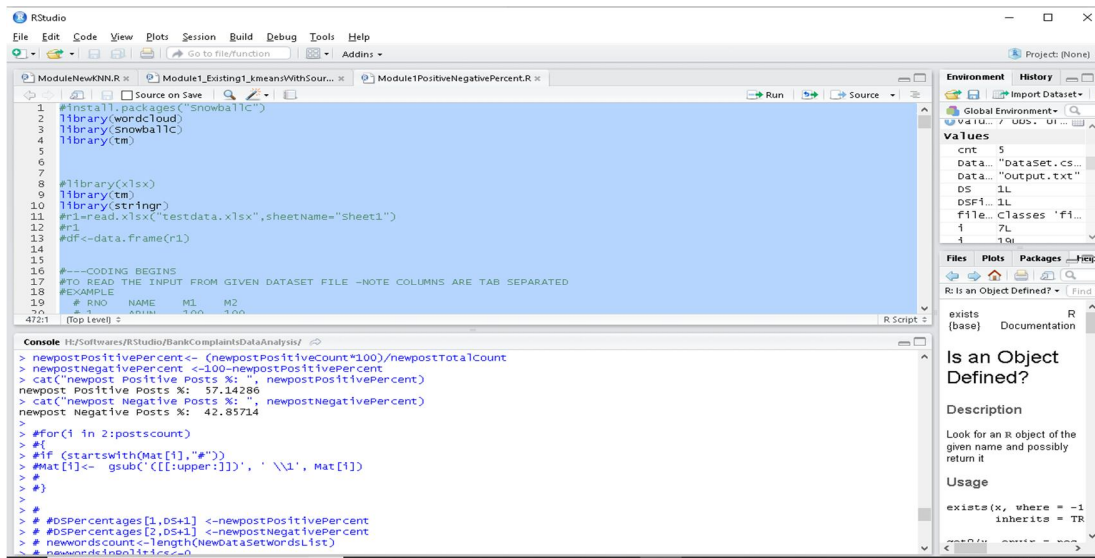


Figure 5 Positive/Negative Posts Percentage

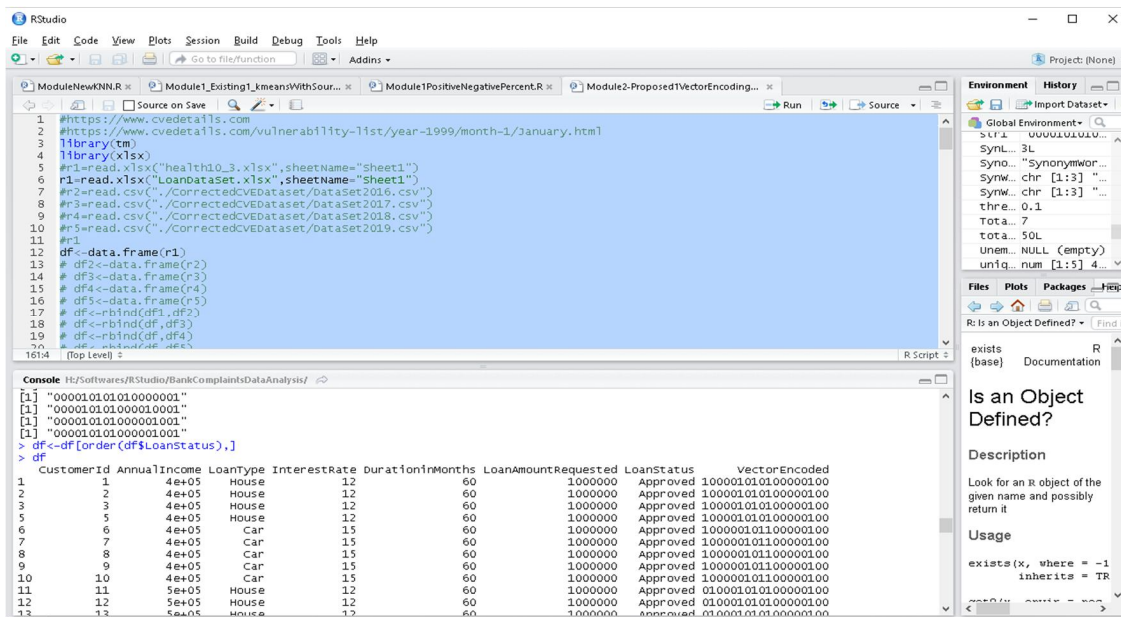


Figure 6 Vector Encoded Values

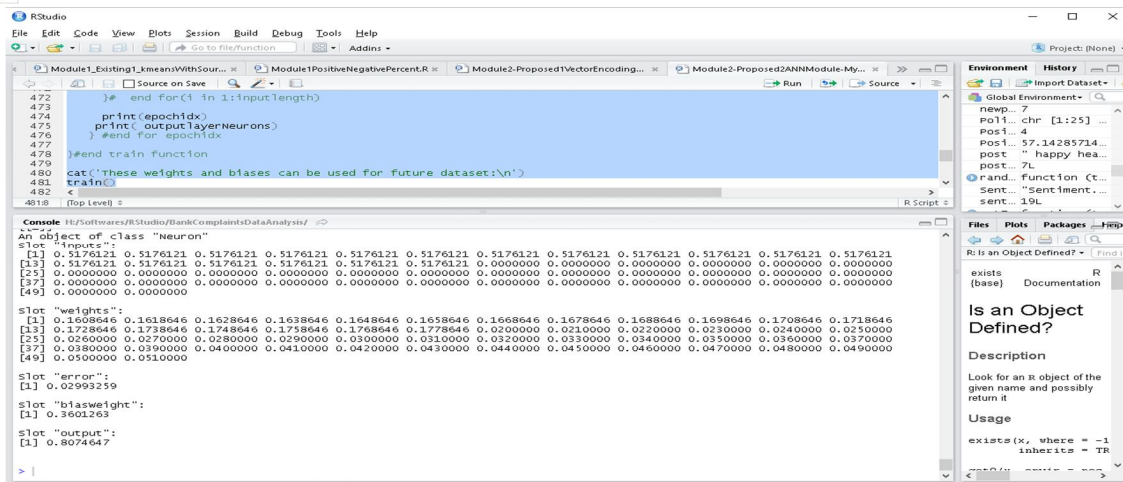


Figure 7 Neural Network Model [Weights and Biases]

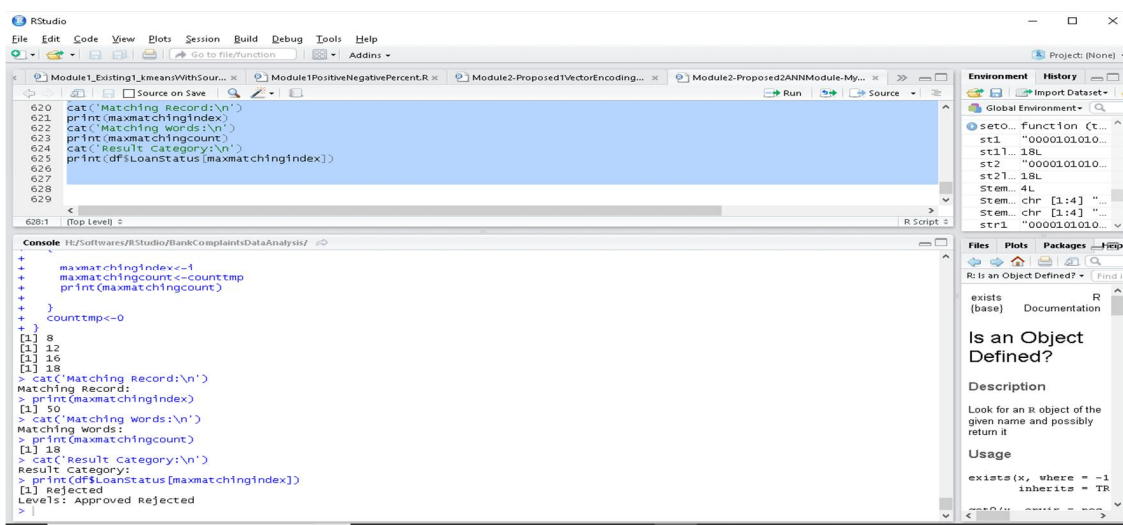


Figure 8 Matching Record in the Training Data for the Given Test Data.

V. CONCLUSION

The project has explored various concepts of data mining of bank customers' data set and the results show what problems customers are having with specific problems in particular loan provisions. This valuable information will show where companies will need to invest in to improve their overall performance in the view of their customers. This will lead to improved customer satisfaction. By maximizing customer satisfaction, the opportunity for repeat sales to customers can be increased. Customer satisfaction also helps to increase customer loyalty, reducing the need to allocate marketing budget to acquire new customers. Satisfied customers may also recommend your products or services to other potential customers, increasing the potential for additional revenue and profit. Future research can be done to collect more complaint data so that we can perform analysis for other products, services and companies. Since neural network is applied, the weight values and bias values are calculated which can be used for further records from customers in classifying the data.

REFERENCES

Book References

- [1] Krishna Rungta, "Learn R Programming in 1 Day: Complete Guide for Beginners". Kindle Edition.
- [2] SandipRakshit, "R for Beginners" Edition 1, 2017, McGraw-Hill Education, ISBN: 9789352604555, 9352604555
- [3] Mike. McGrath, "R for Data Analysis in easy steps: R Programming Essentials", Kindle Edition. Sold by: Amazon Asia-Pacific Holdings Private Limited.
- [4] Mark Gardner, "Beginning R: The Statistical Programming Language", Kindle Edition.
- [5] <http://www.rstudio.com>
- [6] <http://www.rstudio.com/online-learning>
- [7] <https://www.statmethods.net/r-tutorial/index.html>

- [8] <https://www.tutorialspoint.com/r>
- [9] Goyal S, Thakur KS (2008). A Study of Customer Satisfaction Public and Private Sector Banks of India Punjab, J. Bus. Stud., 3(2): 121- 127.
- [10] Uppal RK (2007). Customer Service in Banks- An Empirical Study', Bankers Conference Proceedings, pp. 36-42.
- [11] Kamakodi N (2007). Customer Preferences on e-Banking Services Understanding through a Sample Survey of Customers of Present Day Banks in India Contributors, Banknet Publications, 4: 30-43.
- [12] Mishra JK, Jain M (2007). Constituent Dimensions of Customer Satisfaction: A Study of Nationalized and Private Banks Prajnan, 35(4): 390-398.
- [13] Jain AK, Jain P (2006). Customer Satisfaction in Retail Banking Services NICE, J. Bus. Stud., 1(2):95-102.
- [14] Singh SB (2006). Customer Management in Banks Vinimaya, 37(3): 31- 35.
- [15] Bhaskar PV (2004). Customer Service in Banks IBA Bulletin, 36(8): 9- 13.
- [16] Hasanbanu S (2004). Customer Service in Rural Banks: An Analytical Study of Attitude of Different types of Customers towards Banking Services IBA Bulletin, 36(8): 21-25.
- [17] Singh S (2004). An Appraisal of Customer Service of Public Sector Banks IBA Bulletin, 36(8): 30-33.
- [18] Shankar AG (2004). Customer Service in Banks IBA Bulletin, 36(8): 5- 7.
- [19] Ganesh C, Varghese ME (2003). Customer Service in Banks: An Empirical Study'. Vinimaya, 36(2): 14-26.
- [20] Nandagopal S, Arunachalam VP, KarthikS, "A Novel Approach for Mining Inter-Transaction Itemsets", European Scientific Journal, Vol.8, pp.14-22, 2012.
- [21] V.S. Suresh kumar "Frequent Pattern Complex query management using FIUT Approach", South Asian Journal of Engineering and Technology, pp: 300-304, issue 204, volume 202, 2018
- [22] Gokulraj P and Kiruthikadevi K, "Revocation and security based ownership deduplication of convergent key creating in cloud", International Journal of Innovative Research in Science, Engineering and technology. Vol. 3, Issue 10, ISSN: 2319-8753, October 2014.
- [23] Sureshkumar V S, Chandrasekar A," Fuzzy-GA Optimized Multi-Cloud Multi-Task Scheduler For Cloud Storage And Service Applications" International Journal of Scientific & Engineering Research, Vol.04, Issue.3,pp-1-7, 2013
- [24] E.Prabhakar, V.S.Sureshkumar, Dr.S.Nandagopal, C.R.Dhivyaa, Mining Better Advertisement Tool for Government Schemes Using Machine learning " , International Journal of Psychosocial Rehabilitation, Vol.23,Issue.4, pp. 1122-1135, 2019
- [25] Vijayakumar M, Prakash s, "An Improved Sensitive Association Rule Mining using Fuzzy Partition Algorithm", Asian Journal of Research in Social Sciences and Humanities, Vol.6,Issue.6, pp.969-981, 2016.
- [26] Prakash S, Vijayakumar M, " Risk assessment in cancer treatment using association rule mining techniques",Asian Journal of Research in Social Sciences and Humanities,Vol.6,Issue.10, pp.1031-1037, 2016.
- [27] Prabhakar E, " Enhanced AdaBoost algorithm with modified weighting scheme for imbalanced problems, The SIJ transaction on Computer science & its application, Vol.6,Issue.4, pp.22-26, 2018.
- [28] Suresh kumar V S ,Thiruvankatasamy S, Sudhakar R, "Optimized Multicloud Multitask Scheduler For Cloud Storage And Service By Genetic Algorithm And Rank Selection Method", Vol.3,Issue.2, pp.1-6, 2014
- [29] Nandagopal S, Malathi T, "Enhanced Slicing Technique for Improving Accuracy in Crowd Sourcing Database", International Journal of Innovative Research in Science, Engineering and Technology, Vol.3,Issue.1, pp.278-284, 2014
- [30] Prabhakar E, Santhosh M, Hari Krishnan A, Kumar T, Sudhakar R," Sentiment Analysis of US Airline Twitter Data using New Adaboost Approach" ,International Journal of Engineering Research & Technology (IJERT),Vol.7,Issue.1, pp.1-6, 2019
- [31] V.S. Suresh kumar "E-Farming by means of E-Mandi Process", International Journal of Research and Advanced Development (IJRAD), ISSN: 2581-4451, pp: 55-57, Issue 6, volume 2, 2019
- [32] Dr.C.R. Dhivyaa, R. Sudhakar, K. Nithya and E. Prabhakar "Performance Analysis of Convolutional NeuralNetwork for Retinal Image Classification", International Journal of Psychosocial Rehabilitation, Vol. 23, no.4, pp.1149-1159,November 2019.
- [33] S Nandagopal, S Karthik, VP Arunachalam," Mining of meteorological data using modified apriori algorithm", European Journal of Scientific Research , Vol. 47, no.2, pp. 295-308, 2010.
- [34] P Gokulraj, K Kiruthika-Devi," Revocation and security based ownership deduplication of convergent key creating in cloud", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, no.10, pp16527-16533, October 2014.
- [35] E Prabhakar, R Parkavi, N Sandhiya, M Ambika," Public Opinion Mining for Government Scheme Advertisement", International Journal of Information Research and Review, Vol. 3, no.4, pp2112-2114, February 2016.
- [36] E Prabhakar, G Pavithra, R Sangeetha, G Revathy," Mining Better Advertisement Tool For Government Schemes", International Journal For Technological Research In Engineering, Vol. 3, no.5, pp1023-1026, January 2016.
- [37] Karthik.S. Nandagopal.S, Arunachalam.V.P.," Mining of Datasets with Enhanced Apriori Algorithm", Journal of Computer Science, Vol. 8, no.4, pp599-605, 2012.
- [38] E. Prabhakar," Enhanced Adaboost Algorithm With Modified Weighting Scheme For Imbalanced Problems", The SIJ Transactions on Computer Science Engineering & its Applications (CSEA) , Vol. 6, no.4, pp22-26,July 2017.
- [39] Nandagopal.S. Malathi.T.," Enhanced Slicing Technique for Improving Accuracy in Crowd Sourcing Database", International Journal of Innovative Research in Science, Engineering and Technology) , Vol. 3, no.1, pp278-284,2014.
- [40] V Dharani S Thiruvankatasamy, P Akhila, V Arjitha, K Bhavadharani," A MD5 Algorithm Approach to Monitor Village Using Mobile Application", South Asian Journal of Engineering and Technology, Vol. 8, no.s1, 2019.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)