iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** VI **Month of publication:** June 2017

DOI:

www.ijraset.com

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

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

An Error Reduction Hybrid for Achieving Near to Exact Classification Using KSTAR and IBK Classifier

Prasanta Kumar Nanda[1], Praveen Chouksey[2]
[1,2]*Dr. C.V. Raman University, Bilaspur, Chhattisgarh, India*

Abstract: ***The data mining which is the source of the all the discovery of knowledge, classification always play a vital role and it really means, so without proper classification it is impossible to find the proper accuracy, so this paper gives the maximum error reduction by using the mention data mining classifier. The combination of both classifier exhibits the high rate of error elimination.***
Keywords: ***Weka, Data mining, Weka, Kapa statistic, classification, KSTAR, IBK***

I. INTRODUCTION

Data mining also known as the discovery of knowledge in databases in the field of computer science engineering, the process of discovering new pattern of useful data relationship in very large volume of data and classification plays a major role in the field of Data mining to help and analyse and obtained the required data very accurately by predictions. Sometimes it happens that all our works would ruin at the end of the result so when applying an algorithm in any of the field we must be knowing the liability of it. KStar algorithm is a case-based learner that makes an effort to improve its consistency of performance for dealing with misplaced values, efficiency problems and both real and abstract valued attributes and IBK is a k-nearest-neighbour classifier that applies the same distance metric. The number of nearest neighbours can be defined explicitly in the object editor or determined automatically.

II. LITERATURE RIVIEW

According to[1] P.Rajeswari et al (2010) focused on analyse a disordered in liver using weka tools. In this process of their experiment they have selected naive Bayes,KStar and FT tree. During test condition, they have selected 10 folds cross validation. After the completion of experiment they have found that out of three algorithm, FT Tree algorithm is the best algorithm among others for lever dataset on the basis of accuracy and time taken.

According to[2] Lokesh S.Katore et all(2015) worked on comparison of various classification of algorithm in weka tools for the experimental purpose. They have chosen J48, Naive based simple cart and KStar in terms of test condition. They have applied 10 fold cross validation on a training data set. This data set contains 7 attributes of 258 instances. After completion of experiment, they have found that J48 is performed better after applying a filtered technique name as SMOTE and reassemble during their experiment.They have used multiple parameters like MAE,RMSE,RAE,RRSE accuracy etc.

According to [3] Dorina Kabakchiva(2012) worked on performance improvement system using effective data mining classification algorithm ,the researcher used data mining algorithm to get maximum accuracy which increases the high probability of finding prediction.The data mining algorithm used in their experiment were OneR Rule Learner,Decision Tree, Neural Network and K-Nearest Neighbour.

According to [4] P.S Anoop et all(2016) worked on extract various features from vibration signal of a vehicle, for this purpose the researchers have used MEMS Sensor and their various sample. After this all the related data work evaluated using J48 for features selection and KStar future classification. During this experiment they have used multiple parameters as MAE,RMSE,RAE,RRSE and Kappa statistics etc.

According to [5] Abdel Hamid Salih Mohamed Salih et all(2015) focused on to give an approach for health care sector so that discussion support system with facility of automation can be implemented. During their experiment they have used multiple classifiers on test options like 5 fold cross validation and 10 fold cross validation. The algorithm which were used for evaluation were Begging, Random committee ,IBK, Part and J48 etc. The parameters during experiment was recall precision MAE,RMSE,RAE & RRSE. After completion of experiment the researchers concluded that Random Tree model is emerged as a winner for best algorithm for health care support system.

©IJRASET: All Rights are Reserved

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

III.METHODOLOGY

The following steps used in our methodology

A. Select the Dataset from the online data storage.
B. Install the weka tool.
C. Select the explorer option from the weka GUI Chooser then select autos.arff dataset.
D. Select attribute selection filter from the supervised filter.
E. Select the percentage split test option and sets it value to 87% which is comes under classify option.
F. Chose the KSTAR and IBK algorithm individually from lazy classifier and select the readings.
G. Choose vote classifier from meta classifier and choose to add both algorithms KSTAR & IBK.
H. Then run the process by clicking start option.

IV. RESULT AND DISCUSSION

The result of our experiment is described below using following tables:

TABLE I
List of attributes after applying filtered during pre-processing.

Attributes
normalized-losses
num–of-doors
wheel-base
length
height
symboling

TABLE II
Result after applying KStar

Kappa statistic	Mean absolute error	Root mean squared error	Relative absolute error	Root relative square error
0.8269	0.0444	0.1725	20.104	51.966

TABLE III
Result after applying IBK

Kappa statistics	Mean absolute error	Root mean squared error	Relative absolute error	Root relative square error
0.8902	0.0268	0.144	12.0826	43.2551

TABLE IV
Result after applying KSTAR + IBK

Kappa statistics	Mean absolute error	Root mean squared error	Relative absolute error	Root relative square error
0.9446	0.0254	0.1097	11.4605	32.9622

©IJRASET: All Rights are Reserved

TABLE V
Comparative analysis with previous work

	According to [2] for KSTAR	According to [4] for KSTAR	Our approach KSTAR+ IBK
MAE	0.1127	0.0897	0.0254
RMSE	0.2668	0.2395	0.1097
RAE	31.61	20.17	11.4605
RRSE	63.22	50.80	32.9622

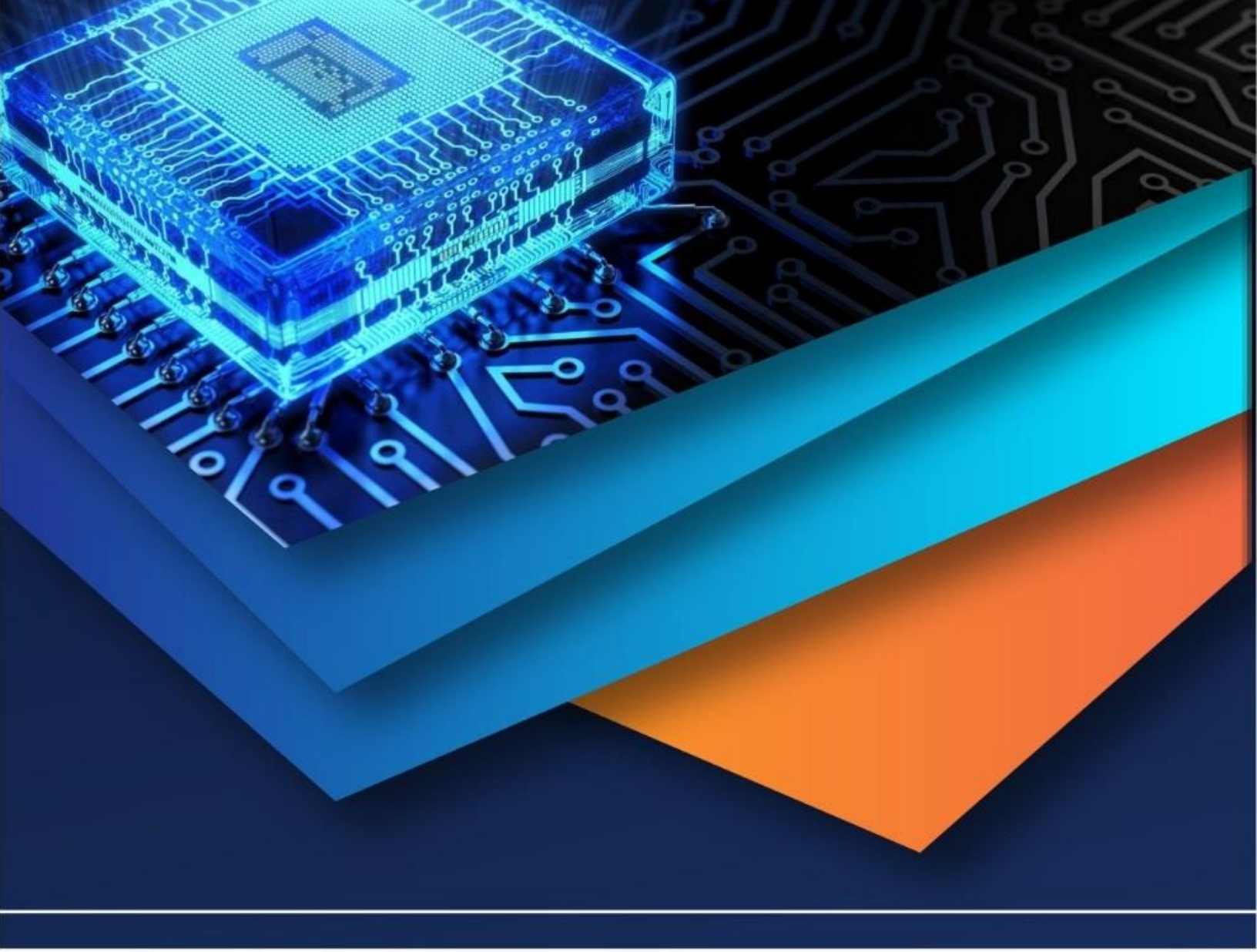
V. CONCLUSIONS

In this paper, experimentally analyzed and find out that the error rate has been minimized using the combination of two useful algorithm i.e KStar and IBK.. Simultaneously these algorithm has been applied with different conditions of percentage split. After completion of this project we can say that our project is successful so that we can achieve errorless classification to the desirable amount.

REFERENCES

[1] P.Rajeswari, G.Sophia Reena,"Analysis of Liver Disorder Using Data mining Algorithm",Global Journal of Computer Science and Technology Issue 14, November 2010, pp. 48-52
[2] Lokesh S. Katore,J.S.Umale,"Comparative Study of Recommendation Algorithms and Systems using WEKA",International Journal of Computer Applications",Volume 110,No. 3,January 2015,pp.14-17.
[3] Dorina Kabakchieva,"Student Performance Prediction by Using Data Mining Classification Algorithms",International Journal of Computer Science and Management Research,Vol 1, Issue 4, November 2012,ISSN 2278-733X,pp.686-690
[4] P. S. Anoop*, V. Sugumaran and Hemanth Mithun Praveen,"Implementing K-Star Algorithm to Monitor Tyre Pressure using Extracted Statistical Features from Vertical Wheel Hub Vibrations ",Indian Journal of Science and Technology, Vol 9(47),December 2016,ISSN (Print) : 0974-6846,ISSN (Online) : 0974-5645,PP. 1-7
[5] Abdel Hamid Salih Mohamed Salih,Ajith Abraham,"Computational Intelligence Data Analysis for Decision Support and Health Care Monitoring System",Journal of Network and Innovative Computing,Volume 3,ISSN 2160-2174,pp. 088-104.
[6] https://algorithmia.com/algorithms/weka/KStar.
[7] http://www.cs.waikato.ac.nz/ml/publications/1995/Cleary95-KStar.ps.gz.
[8] http://www.sciencedirect.com/science/article/pii/S0957417413009056
[9] http://machinelearningmastery.com/how-to-tune-a-machine-learning-algorithm-in-weka/

©IJRASET: All Rights are Reserved



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)