

Compare Different Prediction Algorithm

Devendra Kumar¹, Dr. Anil Kumar Dubey²

¹UG Scholar, ²Associate Professor, Poornima Institute of Engineering & Technology, Jaipur, India
¹2014pietcsdevendra@poornima.org, ^{2*} anil.dubey@poornima.org

Abstract: Now a days everywhere people are using intelligence devises for automation of work, advancement in technology and reducing the human effort. In this reference artificial intelligence provide the diverse applications with machine learning and deeper analysis. Deep Learning analysis is a part of artificial intelligence which inspired by the human brain called artificial neural networks. A Lot of researchers have already proposed several techniques and approaches regarding the predictive analysis of deep learning, but no one proposed the comparative analysis of given approaches to identify which technique are approach is suited for concern problem. To resolve such a problem we proposed a comparative analysis of deep learning techniques to support the future researcher to select the appropriate techniques as per the problem environment. Here we have compared 8 advanced predictive analysis techniques of deep learning through Weka3.0 tool of Open source.

Keywords: Artificial Intelligence; Machine Learning; Deep Learning; Neural Network; Data Mining.

I. INTRODUCTION

In recent years, Deep learning are used for deeper analysis and it is able to predict the more accurate results, which cannot be done by other technology. These prediction techniques and deep learning, are an area of interest, in which researchers can continue to discover knowledge and opinions that are useful to businesses from the large volumes of customer information data accumulated by these businesses using these techniques. Among these, the deep learning approach used in this study is a relatively new method of machine learning in which we can convert deep prediction results in neural network. The deep learning is well rooted in the classical neural network (NN) literature.

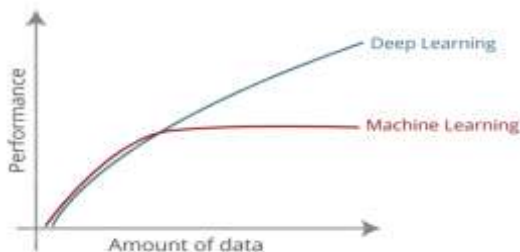


Fig. 1. Deep Learning

II. DEEP LEARNING

Deep learning is able to predict the deeper results in businesses data, health data, education data etc. We can call this prediction as static predictions. There are many regression algorithms which are responsible for predict the better results.

There is different examples like:

- Stock market data(benefits of company's)
- Health Disease (brain tumour, Diabetes, heart attacks etc).
- Education Data (we can verify how many students appear in exams)
- Company turnover (on the basis of previous data we predict)

We can predict the different results on different data. As we can see, with help Neural Networks we predict the results.. The deep learning moving toward a world of smarter that combine neural networks with other algorithms like reinforcement learning to attain goals. Deep learning is work on human neural networks; It has multiple hidden layers that predict deeper results. Deep learning has the multiple nodes and multiple inputs. Deep learning neural network worked on human neurons, which occurs when it encounters stimuli. Here's a diagram of what deep learning layers.

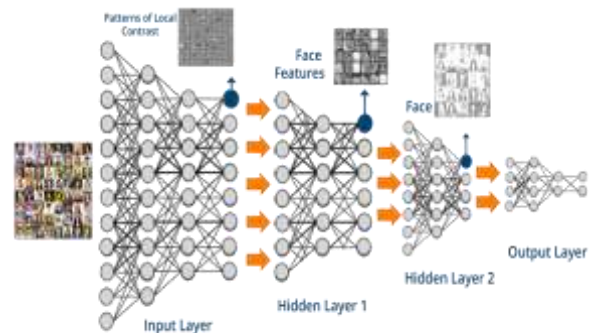


Fig. 2. Deep Learning Neural Network

There is a node layer which switches that turn on or off as the input. Every layer's output is simultaneously the layer's input, starting from an initial input layer receiving your data.

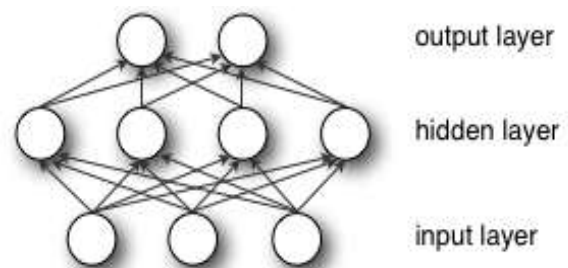


Fig. 3. Neural Network Layers

In deep learning we have basically three layers which are input layer, output layer and hidden layers. There are two or more than two hidden layer are present which provides the deeper results.

III. RESULTS AND DISCUSSION

Here we are comparing the algorithm as per the prediction results. Initially we are selecting one data-set of university students in this we are analysing with three categories which are L (Low), M (Medium) and H (High).

This dataset has 16 attributes and 1 class and it has 480 instances. We apply this dataset on the Weka3.0 tool. On This dataset we predict the correctly classified instances result and compare with different algorithm. J48 Tree algorithm provides the better results of the prediction. J48 Tree algorithm has high value of the Kappa statics and the mean absolute error has low value so J48 provides the better results. Decision Stump algorithm provides the worst prediction results. It has the low value of the Kappa statics and the mean absolute error has high value so J48 provides the worst results.

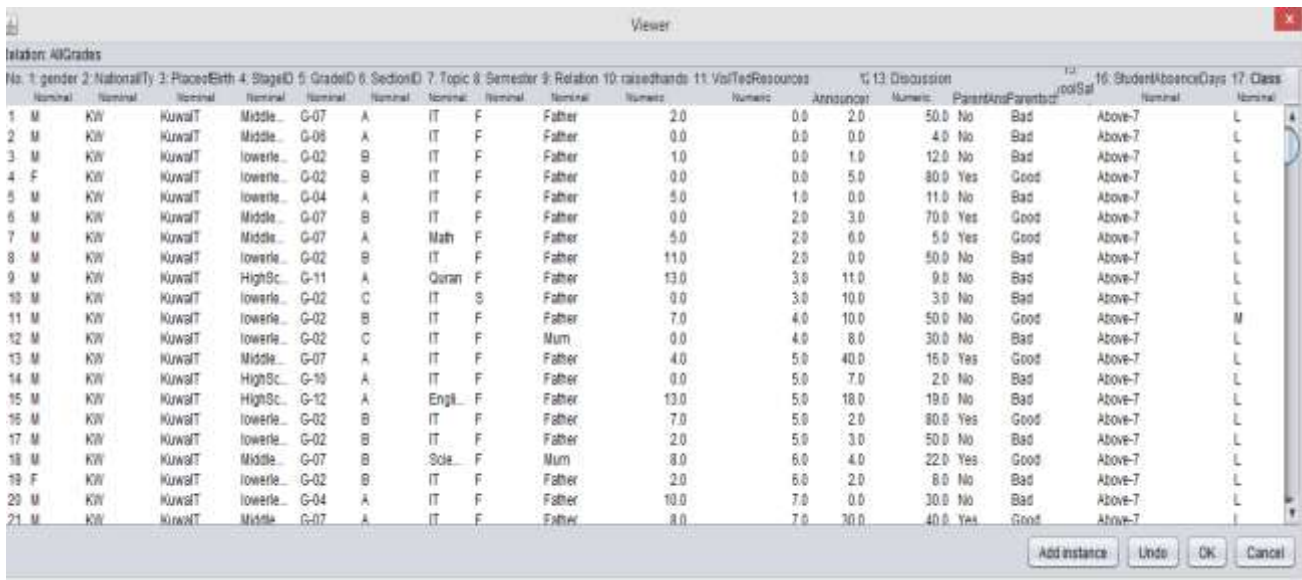


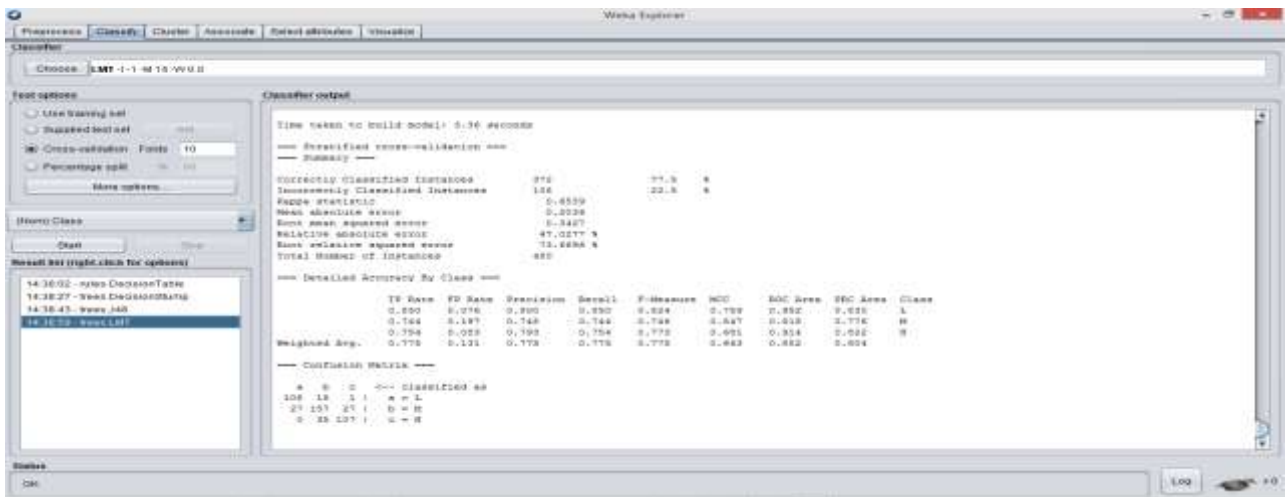
Table 1. Comparing Different Prediction Algorithm

S. N.	Algorithm	Correctly Classified Instances	Incorrectly Classified Instances	Kappa Statistic	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Root Relative Squared Error
1	J48	75.8333 %	24.1667 %	0.6289	0.2155	0.3632	49.7768 %	78.0737 %
2	Naïve Bayes	67.7083 %	32.2917 %	0.513	0.2251	0.397	52.0046 %	85.3312 %
3	Decision Table	71.0417 %	28.9583 %	0.5585	0.2901	0.367	67.0138 %	78.8864 %
4	Random Forest	77.0833 %	22.9167 %	0.6459	0.2422	0.3324	55.9451 %	71.4604 %
5	Random Tree	66.875 %	33.125 %	0.4886	0.2178	0.4565	50.3176 %	98.1204 %
6	REP Tree	64.5833 %	35.4167 %	0.4477	0.2727	0.4117	63.0007 %	88.5017 %
7	LMT	68.6916 %	31.3084 %	0.5702	0.099	0.2737	46.7534 %	84.3316 %
8	Decision Stump	52.5000 %	47.5 %	0.2581	0.3461	0.4164	79.9546 %	89.4966 %



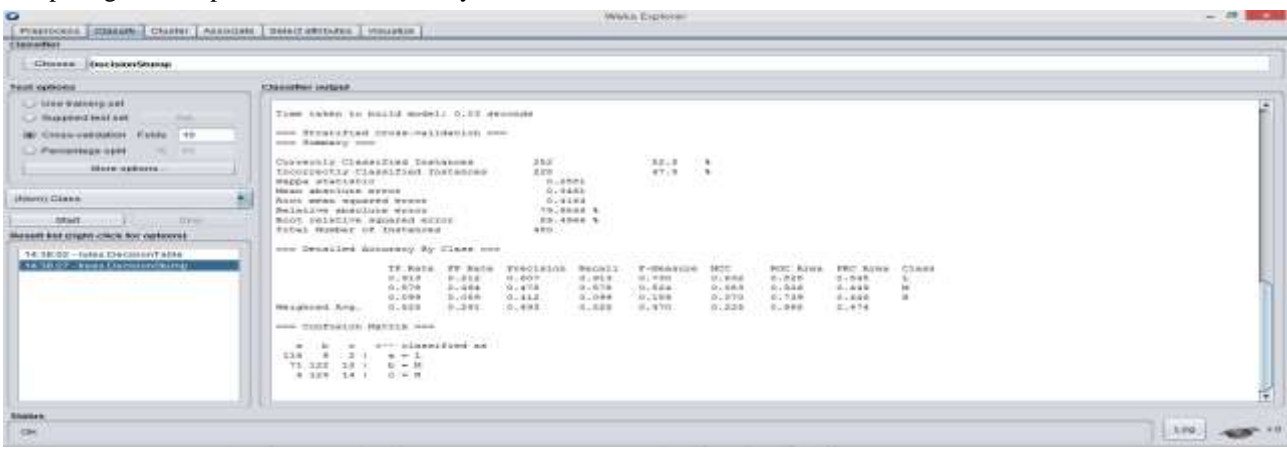
LMT (Logistic Model Tree) algorithm: A LMT is a classification model with supervised algorithm that combines LR and decision tree learning [1][2]. LMT provides the correctly classified instances results is

31.3084 % and incorrectly classified instances result is 31.3084 % which provides the moderate results for maximum dataset. It has the Kappa statistic value is 0.5702 which also good level.



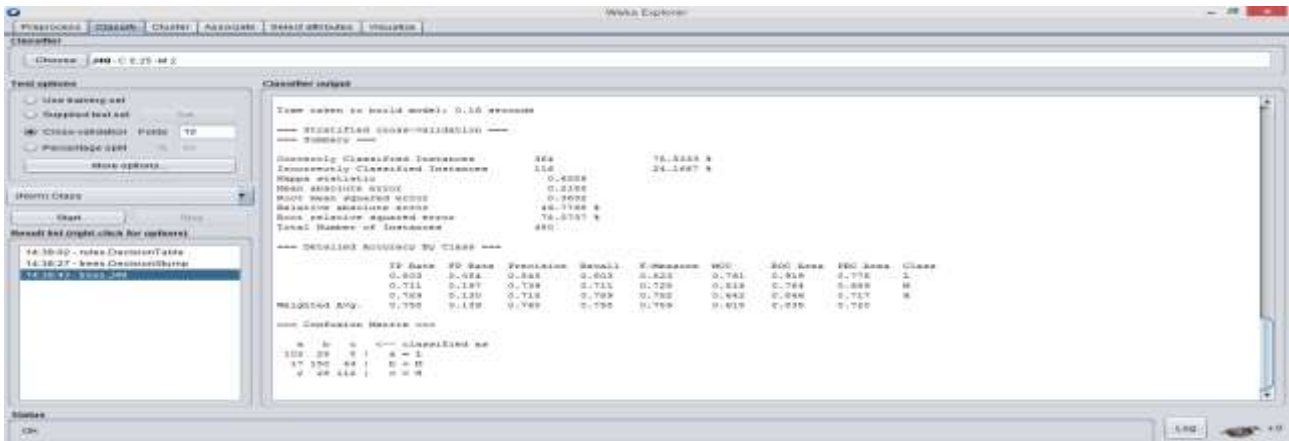
Decision Stump Algorithm: A decision stump is model consisting of a one-level decision tree [3]. That is, it is a decision tree with one internal or the root node which is immediately connected to the terminal nodes. Decision Stump algorithm provides the correctly classified

instances 52.5000 % and incorrectly classified instances 47.5 % which provides less result for maximum dataset. It has the Kappa statistic value is 0.2581 which also low.



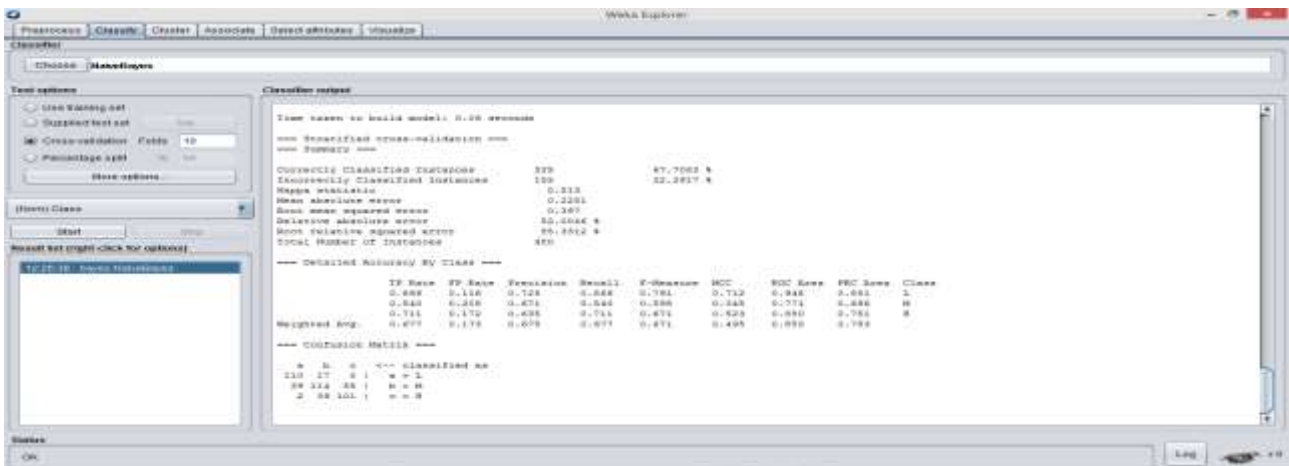
J48 Tree Algorithm: J48 is an extension of ID3. J48 tree algorithm provides the correctly classified instances result is 75.8333 % and incorrectly classified instances

result are 24.1667 % which provides better result for maximum dataset. It has the Kappa statistic value is 0.6289 which also very good.



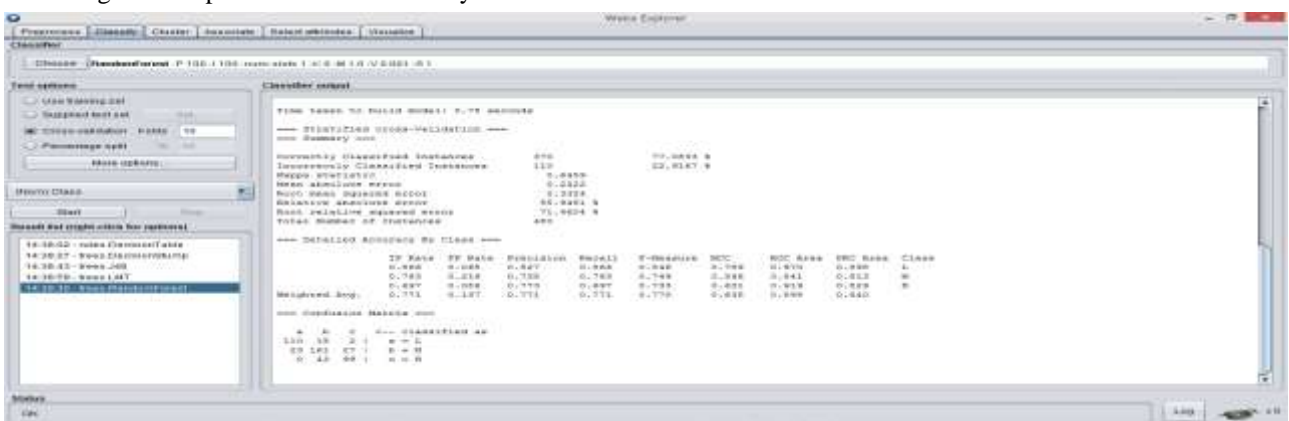
Naïve Bayes Algorithm: Naive Bayes algorithm provides the correctly classified instances result is 67.7083 % and incorrectly classified instances result are

32.2917 % which provides good result for maximum dataset. It has the Kappa statistic value is 0.513 which also good.



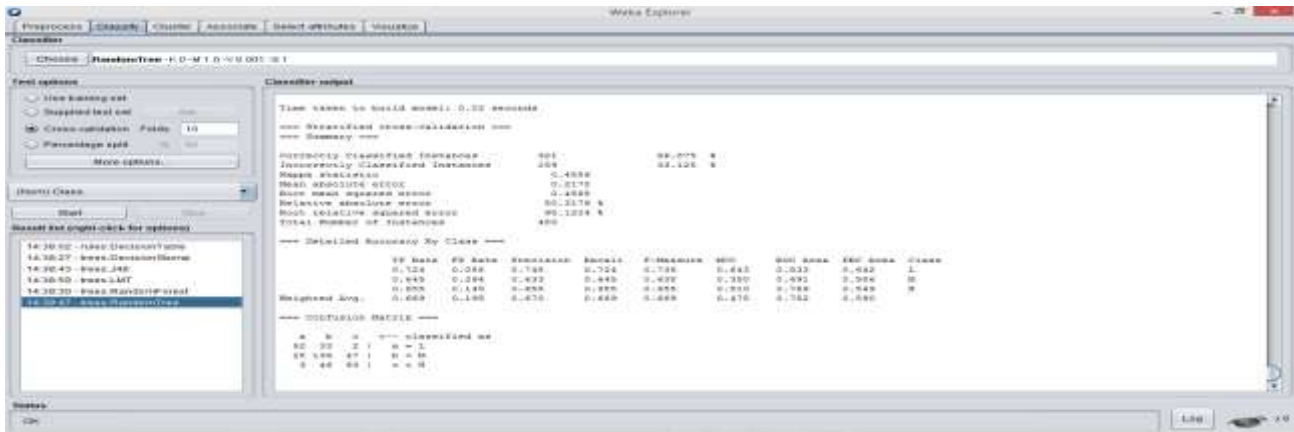
Random Forest Algorithm: Random forests algorithm are learning method for regression and other tasks, that constructing a multitude of decision trees [4]. Random forest algorithm provides the correctly classified

instances result is 77.0833 % and incorrectly classified instances result are 22.9167 % which provides better result for maximum dataset. It has the Kappa statistic value is 0.6459 which also very good.



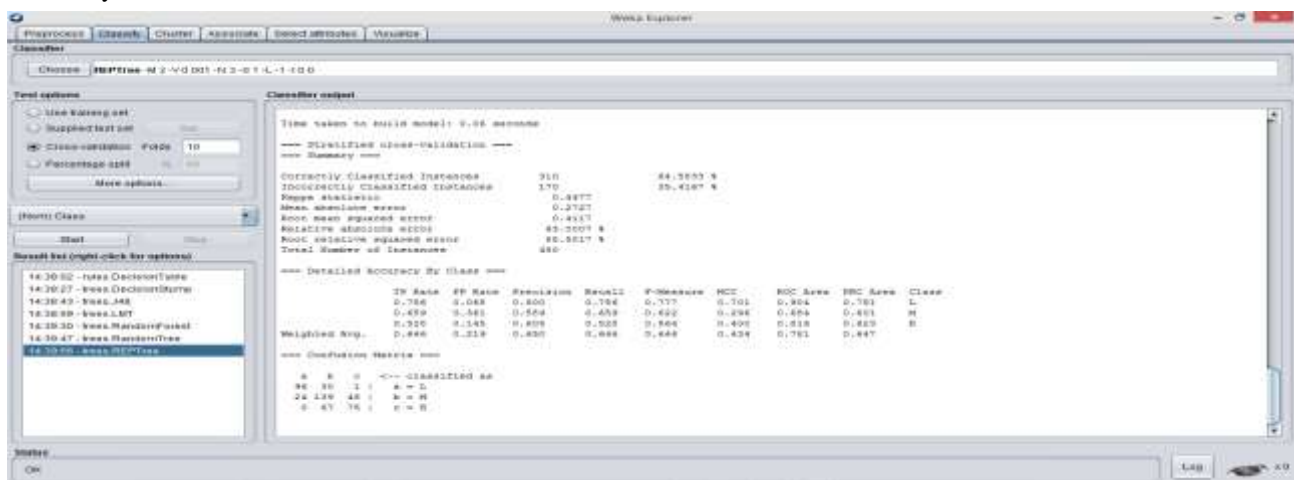
Random Tree Algorithm: Random tree algorithm provides the correctly classified instances result is 66.875 % and incorrectly classified instances result are

33.125 % which provides better result for maximum dataset. It has the Kappa statistic value is 0.4886 which also very good.



REP Tree Algorithm: REP tree algorithm provides the correctly classified instances result is 64.5833 % and incorrectly classified instances result are 35.4167 %

which provides good result for maximum dataset. It has the Kappa statistic value is 0.4477 which also good.



IV. CONCLUSION

Deep Learning Analysis Provides The Deeper Analysis Of The Data. Deep Learning Uses The Neural Network Which Inspired By Human Brain. In Artificial Intelligence Neural Network We Have Multiple Hidden Layer Which Provides The Highly Efficient Results. In This Paper, We Compared The Different Prediction Algorithm On The Basis Of Accuracy Of Prediction. We Take Different Attribute Through Which We Compare. These Are Correctly Classified Instances, Incorrectly Classified Instances, Kappa Statics, Mean Absolute Error, Root Mean Squared Error, Relative Absolute Error And Root Relative Squared Error. Here We Have Compared 8 Advanced Predictive Analysis Techniques Of Deep Learning Through Weka3.0 Tool Of Open Source. We Get That J48 Tree Algorithm And Random Forest Tree Algorithm Provides The Better Results And Decision Stump Tree Algorithm Provides Worst Results. In Future Any Researcher Wants To Select The Prediction Algorithm Then They Can Select Easily On Basis Of Compression Table. In Future We Plan To Extend Our Work With Different Environments And Different Parameters.

V. REFERENCES

- [1] Niels Landwehr, Mark Hall, and Eibe Frank (2003). Logistic model trees (PDF). ECML PKDD.
- [2] Landwehr, N.; Hall, M.; Frank, E. (2005). "Logistic Model Trees" (PDF). Machine Learning. 59: 161. doi10.1007/s10994-005-0466-3.
- [3] Iba, Wayne; and Langley, Pat (1992); Induction of One-Level Decision Trees, in ML92: Proceedings of the Ninth International Conference on Machine Learning, Aberdeen, Scotland, 1–3 July 1992, San Francisco, CA: Morgan Kaufmann, pp. 233–240.
- [4] Ho, Tin Kam (1995). Random Decision Forests (PDF). Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.
- [5] Ho, Tin Kam (1998). "The Random Subspace Method for Constructing Decision Forests" (PDF). IEEE Transactions on Pattern Analysis and Machine Intelligence. 20 (8): 832–844.