

# An Empirical and comparative Analysis on Meta Classification Algorithms

Putta Sandeep<sup>1</sup>, Dr. G.V.Ramesh Babu<sup>2</sup>

<sup>1</sup>PG Student, Department of Computer Science, Sri Venkateswara University, Tirupati

<sup>2</sup>Assistant Professor, Department of Computer Science, Sri Venkateswara University, Tirupati

**Abstract**— With the huge boom within side the statistics being gathered due to ubiquitous records accumulating devices, and the elevated want for doing statistics mining and analyses, there's a want for scaling up and enhancing the overall performance of conventional statistics mining and gaining knowledge of algorithms. Two associated fields of disburged statistics mining and ensemble gaining knowledge of goal to deal with this scaling issue. Ensemble gaining knowledge of strategies goal to create a meta-classifier via way of means of combining numerous classifiers created at the identical statistics and enhance their overall performance. This paper specializes in the category strategies primarily based totally on Bagging, Dagging and Stacking the usage of Page-blocks statistics set became used for the category with 5473 times with eleven attributes as unbiased variable and one as established variable for the analysis. The outcomes display that Bagging observed to be the set of rules with maximum precision (97.186%) and accuracy while as compared to Dagging and Stacking set of rules.

## I. INTRODUCTION

Data mining is a technology that offers extracting or discovering new relations, hidden knowledge and important patterns from such data. It is also known as Knowledge Discovery in Databases (KDD). Data mining technique is important for analysis purpose. Data mining supports different techniques such as classification, clustering, association rule mining, outlier analysis etc [1][2]. Data Mining(DM) discovers hidden relationships in data, in fact it is a part of wider process called “knowledge discovery”. Knowledge discovery describes the phases which must be done to ensure reaching meaningful results through research. The objective of DM process is to obtain information out of a dataset and converts it into a comprehensible outline. An understanding of different situations to reach quality solutions. Therefore, data mining is a research field where many advances are being done to accommodate and solves emerging problems [4]. For present study purpose classification technique is investigated.

## II. CLASSIFICATION

Classification plays an important role in data mining and machine learning. The purpose of classification algorithm is to construct a classifier, and then analyzes the characteristics of the unknown data to get an accurate model. The performance of the classifier is measured by its classification accuracy. Building effective classification systems is one of the central tasks of data mining. The main purpose of supervised learning is to build a simple and unambiguous model of the allocation of class labels in terms of predictor features [3][4]. The classifiers are then used to classify class labels of the testing instances where the values of the predictor features are known, to the value of the class label which is unknown [4][8]. Classification of this tremendous amount of data is time consuming and utilizes excessive computational effort, which may not be appropriate for many applications.

There are various Meta classification algorithms such as Bagging, Adaboost and Dagging, and so on. Bagging, boosting and dagging are well known re-sampling ensemble methods that generate and combine a diversity of classifiers using the same learning algorithm for the base-classifiers. Boosting algorithms are considered stronger than bagging and dagging on noise-free data. However, there are strong empirical indications that bagging and dagging are much more robust than boosting in noisy settings. For this reason, in this work we built an ensemble using a voting methodology of bagging, boosting and dagging ensembles with 8 sub-classifiers in each one. We performed a comparison with simple bagging, boosting and dagging ensembles with 25 sub-classifiers, as well as other well known combining methods, on standard benchmark datasets and the proposed technique had better accuracy in most cases.

## III. METHODOLOGY

Meta classifier includes a wide range of classifier. When the attributes have a large number of values because the time and space complexities depend not only on the number of attributes, but also on the number of values for each attribute. Many different types of Meta classification techniques have been proposed in literature that includes Bagging, boosting and dagging

etc. In this paper, we evaluate the performance of the meta classifiers algorithms on Page-blocks data set was used for the classification compared with the Bagging, Dagging and Stacking algorithm.

### 3.1 Bagging

Bootstrap Aggregating or Bagging generates multiple bootstrap training sets from the original training set using sampling with replacement and uses each of them to generate a classifier for inclusion in the ensemble. Given a set,  $D$ , of tuples, bagging works as follows. For iteration  $I$  ( $i = 1, 2, \dots, k$ ), a training set,  $D_i$ , of  $d$  tuples is sampled with replacement from the original set of tuples,  $D$ . Note that the term bagging stands for bootstrap aggregation. Each training set is a bootstrap sample. Because sampling with replacement is used, some of the original tuples of  $D$  may not be included in  $D_i$ , whereas others may occur more than once. A classifier model,  $M_i$ , is learned for each training set,  $D_i$ . To classify an unknown tuple,  $X$ , each classifier,  $M_i$ , returns its class prediction, which counts as one vote [5][7]. The bagging can be applied to the prediction of continuous values by taking the average value of each prediction for a give test tuple. The bagged classifier often has significantly greater accuracy than a single classifier derived from  $D$ , the original training data. It will not be considerably worse and is more robust to the effects of noisy data. The increased accuracy occurs because the composite model reduces the variance of the individual classifiers. For prediction, it was theoretically proven that a bagged predictor will always have improved accuracy over a single predictor derived from.

### 3.2 Dagging

This Meta classifier creates a number of disjoint, stratified folds out of the data and feeds each chunk of data to a copy of the supplied base classifier. Predictions are made via majority vote, since all the generated base classifiers are put into the Vote meta classifier. Useful for base classifiers that are quadratic or worse in time behavior, regarding number of instances in the training data[7] [9].

*Dagging (Disjoint samples aggregating)* is a variant of *Bagging* that creates a few disjoint groups with stratified data, each considered as a learning subset. The predictions are obtained from a majority vote for supervised classification problems. This method has shown its importance in recent works. We propose to use it and to study the behavior of our *CNC* classifier based on stratified sampling and analysis the difference compared to random sampling.

The principle of *DNC (Dagging Nominal Classifier)* is to take several disjoint and stratified samples  $\{O^1, \dots, O^{2b}\}$ . Our learning algorithm *CNC* is then built on each of them to get a collection of classifiers  $\{h_1, \dots, h_b\}$  that will be combined by majority voting. In the case of learning from stratified samples, the characteristic of this type of sample is that it allows the learning algorithm to generate classifiers from representative samples having the same composition as the initial training set.

### 3.3 Stacking

The process of classification is recursively testing the node value from root to leaf, according to the attribute value of the testing instance. Stacking is one of the ways of ensembling [8]. The main idea of the stacking is that the combiner algorithm is used on top of regular algorithms to make a final prediction using the predictions of the trained algorithms as a new input. The classical approach to the stacking in neural network ensembling is to train individually several neural networks and then use major voting principles for the calculation of the final result [10] [12].

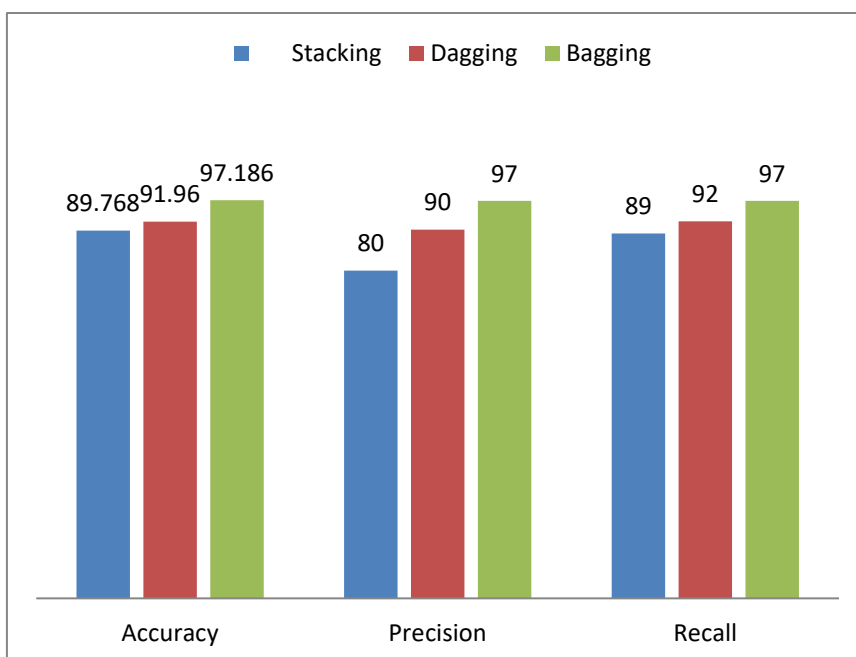
## IV. EXPERIMENTAL RESULTS

The analyses had been directed via way of means of making use of R programming Language. R is a complicated statistical software program package, which offers new methods to information mining. it's far an open-supply device for evaluation of information mining algorithms. The R Language is a package deal for facts characterization, grouping and representation. We have taken into consideration the Page-blocks dataset from the UCI Machine Learning Repository for assessing the productiveness and adequacy of Meta classifiers [11]. The hassle includes classifying all of the blocks of the web page format of a record that has been detected via way of means of a segmentation process. In this dataset 5473 times and eleven attributes. This is an vital step in record evaluation if you want to separate textual content from image areas, includes the 5 training are: textual content magnificence consists of 4913 times , horizontal line magnificence includes 329 statistics, photo magnificence include 28 times, verticalline magnificence consists of 88 statistics and image magnificence consists of a hundred and fifteen times. The Page-blocks feature information facts is consolidated in Table-1. The trendy dataset is parceled into units one for training (80%) and any other set for testing (20%).

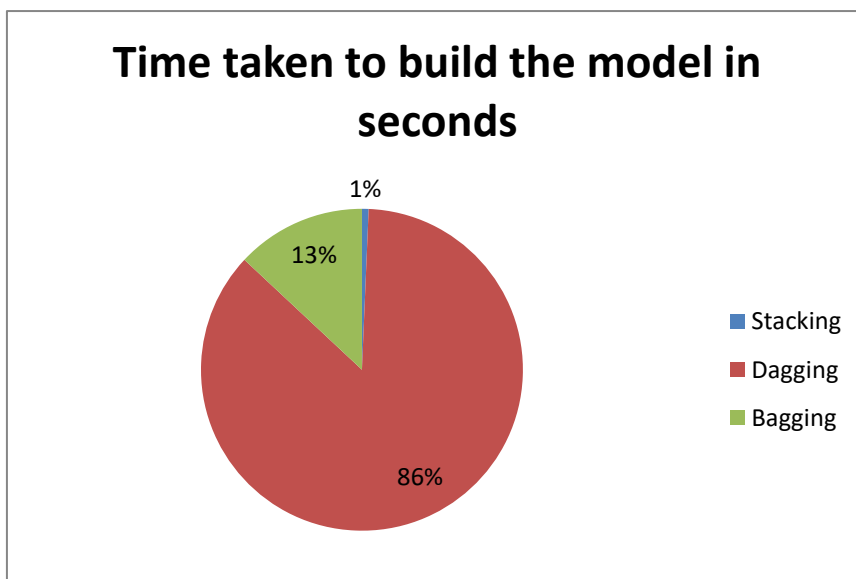
We survey our two models using assorted execution estimations like Accuracy, Precision and Recall, the Experimental results are showed up in the table-1 and same showed up in the Figure-1.

**TABLE 1**  
**PERFORMANCE OF CLASSIFIERS**

Algorithm	Accuracy	Precision	Recall
Stacking	89.768	80	89
Dagging	91.960	90	92
Bagging	97.186	97	97



**Figure-1: Experimental Results**



**Figure-2: Experimental Time**

We find in the Figure-1, the introduction of the Bagging estimation has accomplished 97.186% precision while Dagging and Stacking has achieved 91.960%, 81.768%. As the result from assessment among the three computations, we find that most

vital precision of Classification model is Bagging (97.186%). So, the Bagging algorithm have got highest accuracy, with a 5.226%, 15.418 difference when compared to Dagging, Stacking algorithms.

## V. CONCLUSION

Classification is a supervised manner that learns to categorise new times primarily based totally at the expertise learnt from a formerly labeled education set of times. In these studies, we've done the experiments for you to decide the category accuracy and mistakes price of 3 algorithms specifically Bagging, Dagging and stacking and assist of R Programming language. The intention of this exam paintings is deliberate to expose to categorise all of the blocks of the web page format of a file that has been detected via way of means of a segmentation technique from the Page-blocks dataset. So, the Bagging set of rules have were given maximum accuracy (97.18%), with a 5.226%, 15.418 distinction whilst in comparison to Dagging, Stacking algorithms.

## REFERENCES

- [1] D. Hand, H. Mannila, P. Smyth.: Principles of Data Mining. The MIT Press. (2001)
- [2] G Ravi Kumar, K Tirupathaiah and B Krishna Reddy, "Client Churn prediction of banking and fund industry utilizing machine learning techniques", IJCSE, Volume-7, Issue- 6, PP:842-846, 2019
- [3] G Ravi Kumar, K Venkata Sheshanna and G Anjan Babu, "Sentiment Analysis for airline tweets utilizing machine learning techniques", International Conference on Mobile Computing and Sustainable Informatics, Springer, Cham, PP:791-799, 2020
- [4] J. Han and M. Kamber, "Data Mining concepts and Techniques", the Morgan Kaufmann series in Data Management Systems, 2nd ed. San Mateo, CA; Morgan Kaufmann, 2006.
- [5] L. Breiman, "Bagging predictors," Machine learning, vol. 24, no. 2, pp. 123–140, 1996.
- [6] L. Mason, J. Baxter, P. L. Bartlett, and M. R. Freaun, "Boosting algorithms as gradient descent," in Advances in neural information processing systems, pp. 512–518, 2000.
- [7] Noor A. Samsudin, Andrew P. Bradley, "Group-based Meta-classification", School of Information Technology and Electrical Engineering, The University of Queensland, Australia.
- [8] P.-N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. Reading, MA: Addison-Wesley, 2005.
- [9] S. Dzeroski and B. Zenko, "Is combining classifiers with stacking better than selecting the best one?," Machine Learning, vol. 54, pp. 255–273, 03 2004.
- [10] Ting K. M. Witten H. I. (1997). Stacking Bagged and Dagged Models. In Proceedings of the Fourteenth International Conference on Machine Learning (ICML '97). Morgan Kaufmann Publishers Inc.
- [11] UCI machine learning repository. <http://archive.ics.uci.edu/ml/>
- [12] Z. Ma, P. Wang, Z. Gao, R. Wang, and K. Khalighi, "Ensemble of machine learning algorithms using the stacked generalization approach to estimate the warfarin dose," PLOS ONE, vol. 13, pp. 1–12, 10 2018.