INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Adaptive Learning for Data Mining Classification

Mahendra Divekar[1], Ajay Patil[2], Pooja Patil[3], Anita Saini[4], Prof. Avinash Devare[5]

[1,2,3,4]*B.E. Student, Department of Computer Engineering, Jaywantrao Sawant College of Engineering Hadapsar, (Pune).*

[5]*Associate Professor, Department of Computer Engineering, Jaywantrao Sawant College of Engineering Hadapsar,(Pune).*

*Abstract***:** *The classification of algorithm for dataset Recommendation should be appropriate as it is very important and beneficial task but it goes through some difficulties. As per No-Free-Lunch theorem, for different classification problem there is no best classifier to decide which learning algorithm suits best for which type data and domain is difficult. A type of recommending classification algorithm is affirmed, this paper is based on meta-features promising classifier one recommended by Meta learning. As per the characteristics of data the problem algorithm is selected by adaptive learning. In knowledge discovery process the selection of algorithm will be advantageous by adaptive learning. In these theory measures of recommendation parameter and knowledge base architecture of adaptive learning is discussed. Different classification algorithms are applied on the different dataset and problems. And also algorithms are recommended for the classifications.*
*Keywords**: Supervised Learning, Dynamic Algorithm Selection, Classification, Data Characteristic, Ensemble Learning, No-Free-Lunch (NFL), Knowledge Base, K-Nearest Neighbour (KNN).*

I.   INTRODUCTION

The basic idea is to recommend the most and best suitable algorithm which is based on accuracy and measure. The best algorithm selection for the particular problem or large dataset is very big problem in different data analysis and real time application. The Adaptive learning is mainly adapted to solve these types of problems. This included the recommending the best classifier or suitable algorithm for the given problem and large dataset. Data mining is the process of extract pattern from large dataset. Data mining also included the clustering, classification, web mining, text mining. Different problems of algorithm selection are solved by using these techniques. The main approach of this is to recommend the more appropriate algorithm based on the data characteristics,   and choosing the suitable classifier for given problem to achieve the good results. The meta learning is used to learn about learning algorithm themselves to predict how learning classifier will perform on a given dataset,   this is based on extracting meta features. These Meta features are used to describing the dataset itself.

II.   PROPOSED SYSTEM ARCHITECTURE

The Adaptive learning classification is based on these components.
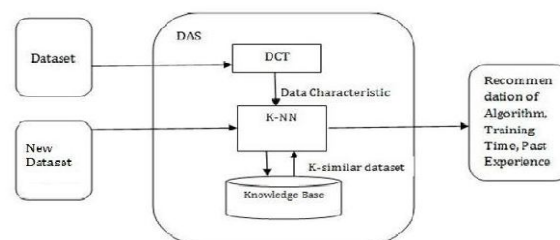New dataset
Historical data
K-NN DCT Knowledge base

Fig1.Adaptive learning architecture

The dataset is used to represent the training data as well as testing data. In the first stage the dataset are transferred to the data

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

characteristics (DCT) tool. DCT is the module this DCT module is used to calculation of different data characteristics such as accuracy, MIN, MAX, standard Deviation etc. Knowledge base contains of the experience of known applications. It also included the knowledge about the execution of special algorithm on specific dataset. The Nearest-neighbour classifier are based on learning by analogy, i.e. by comparing a given test tuple with training tuple that are similar to it. K-Nearest neighbour (KNN) classifier searches the pattern space for K training tuple that are closest to unknown tuple.

For solving the complication of algorithm selection different perspectives are given different classifiers are put in an application on each of them for given different dataset. By given some parameters such as performance of each classifier can be analysed so as to mention the given dataset. For deciding which classifier will operate best for a given dataset, there are two options in the first based on perception and practical knowledge is to put all trust in expert opinion. The second is to run through every possible classifier that could work on the dataset. Distinguish logically one which performs best, the lather alternatives while being the most through, would take time and request a significant amount of resources especially with large dataset .These are in such practical.

III. RELATED WORK

The term adaptive learning is very useful approach. There are different types of reasons why we are using the adaptive learning approach. By using ensemble learning it gives very different types benefits of using adaptive learning system. Following are the some approaches of algorithm selection ensemble based system is more better than the single classifier. The individual classifier are combined with different approaches which having the many advantages of algorithm selection user gets recommendation of algorithm.

IV. EXISTING SYSTEM

The gathering supervision or description is affected by first one. When there is massive volume of input data is available then it is exclusively high. If it is not possible to label all of them, for example it is not worth mentioning task to description of the large number of images for image classification of the exclusive description of everything in the existing system is not possible. In the administration or description there is unpredictability and Equivocation. for example two abstraction hot and cold is not defined for isolating by majority. The applications of the supervised learning paradigm in several number of framework may terminate the struggles to succeed in dealing with these restriction in implementation , other learning paradigm such as unsupervised learning , semi supervised learning , reinforcement learning, active learning some mixed learning approaches can be considered.

V. CONCLUSION

For dataset of all types there is no any single algorithm which acts better as per the No-free-lunch theorem. There are various ways to recommend the algorithm. The best classifier is recommended by adaptive learning approach which recommends the appropriate and best classifier. The recommendation is done on the basis of accuracy and performance measure whose main aim is to selecting the algorithm and this selection is assisted by non-expert system. Simple information theoretic and statistical are three types of meta features. These three categories are used and evaluated comparatively. Experimentation on selecting features which are essential task like number of attribute number of class labels , number of instances , class entropy for classifiers selection ,maximum class probability.

REFERENCES

[1] Q. Song, G,. Wang, C. Wang, "Automatic Recommendation of Classification Algorithms Based on Dataset Characteristic", Pattern Recognition 45(2012) 2672-2689

[2] Guido Lindner and Rudi Studer,"AST: Support for algorithm selection with a CBR Approach".

[3] Nikita Bhatt, Amit Thakkar, Amit Ganatra, " A survey and current research challenges in meta learning Approaches based on dataset characteristic", International Journal of soft computing and Engineering, ISSN: 2231-2307, Volume-2, Issue-1, March-2012

[4] Lior Rokach, Department of Information System Engineering, Ben-Gurion University of the Negev, Israel, Computational Statistics and Data Analysis 53 (2009), "Taxonomy for characterizing ensemble methods in classification tasks: A review and annotated bibliography"2001

[5] Geoffrey I. Webb (author for correspondence), School of Computer Science and Software Engineering Monash University, Clayton, Victoria, 3800, Australia," Multi strategy Ensemble Learning: Reducing Error by Combining Ensemble Learning Techniques" To appear in IEEE Transactions on Knowledge and Data Engineering.

[6] Nikolaj Tatti," Distance between dataset based on summary statistic", Journal of machine learning research 8(2007) 131-15

[7] Robi Polikar, "Ensemble Based system in decision making" IEEE circuit and System Magazine 2006.

[8] C. IGEL and M.TOUSSAINT,"A No-Free-Lunch Theorem for Non-Uniform Distributions of Target Functions", Journal of Mathematical Modeling and Algorithms 3: 313–322, 2004. © 2004 Kluwer Academic Publishers.

[9] R. King, C. Feng, A. Suterland,"Stat-log: Comparison of classification algorithm on large real-world problem".

[10] Richard O. Duda, Peter E. Hart, David G. Stork, A book on " Pattern Classification" ,II nd edition.

[11] R. King, C. Feng, A. Suterland, "Stat-log: Comparison of classification algorithm on large real-world problem".

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

[12] L.I. Kuncheva," Combining Pattern Classifiers Methods and Algorithms". New York, NY: Wiley Inter science, 2005.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089   (24*7 Support on Whatsapp)