

Performance of Classification Techniques on Medical Datasets

[Murat KOKLU, Kemal TUTUNCU]

Abstract— The definition of the data mining can be told as to extract information or knowledge from large volumes of data. One of the main challenging area of data mining is classification. There are so many different classification algorithm in literature ranging from statistical based to artificial intelligence based. This study make use of Waikato Environment for Knowledge Analysis or in short, WEKA to compare the different classification techniques on different medical datasets. 23 different classification techniques were applied to three different medical datasets namely EEG Eye State, Fertility and Thoracic Surgery Medical Datasets that were taken from UCI Machine Learning Repository. The results showed that Multilayer Perceptron (MLP) had highest accuracy for Fertility Dataset (90%), three different techniques namely Bagging, Dagging and Grading had highest and same accuracies for Thoracic Surgery Data Set (85.1064%) and finally Kstar had highest accuracy for EEG Eye State Dataset (96.7757%).

Keywords— Data Mining, Multilayer Perceptron, Grading, Kstar, Classification, EEG Eye State Dataset, Thoracic Surgery Data Set, Fertility Dataset

I. Introduction

In Classification, training examples are used to learn a model that can classify the data samples into known classes. The Classification process involves following steps:

Create training data set, Identify class attribute and classes. Identify useful attributes for classification, Relevance analysis, Learn a model using training examples in Training set, Use the model to classify the unknown data [1].

In this study, 23 classification techniques were applied to the three Medical Datasets from University of California Irvine Machine Learning Repository (UCI) namely EEG Eye State [2], Fertility [3] and Thoracic Surgery Datasets [4-5] using WEKA [6]. The obtained results were listed and compared to each other to determine the best classification technique for each medical dataset.

The remainder of this paper is organized as follows: In Section 2 previous literature studies on the same medical datasets are presented. In Section 3 materials and methods used in this study are described.

Murat Koklu
Department of Computer Engineering, Selcuk University
Konya, TURKEY

Kemal Tutuncu
Department of Electrical – Electronics Engineering, Selcuk University
Konya, TURKEY

In Section 4, 23 different classification technique are applied to three medical datasets. The results are compared to each other to determine the best classification technique for each medical dataset. And finally, in Section 5 the conclusions are stated.

II. Literature Survey

A. *The Classification Studies on EEG Eye State Medical Dataset*

Fukuda et al. make use of a log-linearized Gaussian mixture neural network for EEG eye state classification [7]. Yeo et al. successfully used support vector machines (SVMs) to detect drowsiness during car driving by eye blink [8]. Moreover, a hybrid system based on decision tree classifier and fast Fourier transform was applied to the detection of epileptic seizure by Polat and Güneş [9]. Sulaiman et all. also make use of K-nearest neighbour (K-NN) for stress features identification [10]. In addition, Hall et all. formed 117-second EEG eye state corpus and employed 42 different machine learnings and statistical approaches based on Weka [11] to predict eye states. It has been found that KStar is the best approach among these different methods [12]. Eye state corpus of them is now a benchmark problem saved by Machine Learning Repository, University of California, Irvine (UCI) [13].

Wang et all., (2014) applied a promising technique that uses incremental attribute learning (IAL) based on neural networks. IAL is a novel machine learning strategy which gradually imports and trains features one by one. IAL exhibited better classification performance in terms of classification error rates in comparison with conventional and some other approaches [14]. Roesler and Suendermann (2013) applied 42 different machine learning algorithms to the related dataset to predict the eye state after training with the corpus. The best-performing classifier was found as KStar [15].

Sahu et all., (2015) used binary classification for finding feature subset selection named as Incremental Feature Reordering (IFR), it gives most non dominant feature (MND) for Electroencephalography (EEG) signal corpus and create reorder set. They found that the removal of MND gives optimal subset feature and it increases the classifier accuracy and efficiency [16].

Roesler et all., [2014] investigated whether the price of an EEG device is directly correlated with the quality of the obtained data when applied to a simple classification task. The data of three different devices (one medical and two consumer) was used to determine the eye state (open or closed). For classification, 83 machine learning algorithms were used on the raw EEG data. While the cheapest device performed extremely poor with only one classifier better than the majority vote the other two devices achieved high

accuracy. The lowest error rate for a more expensive consumer EEG was 1.38% and produced by KStar [17].

Hassani and Lee (2014), proposed Incremental Quantum Particle Swarm Optimization to develop an Incremental Framework for Classification of EEG Signals. They compared the performance of IQPSO against ten other classifiers on two EEG datasets. The results suggested that IQPSO outperformed other classifiers in terms of classification accuracy, precision and recall [18].

Singla et al., (2011) compared SVM with the Artificial Neural Network (ANN) on EEG dataset. A one-against-all SVM and a multilayer ANN is trained to detect the eye events. They presented comparison result in related paper [19]. Razzaghi et al (2015), proposed a multilevel algorithmic framework for SVP that includes (a) construction of hierarchy of large-scale data coarse representations, and b) a local processing of updating the hyper plane throughout this hierarchy. Their multilevel framework substantially improved the computational time without losing the quality of classifiers. Experimental results were presented for balanced and imbalanced classification problems. Quality improvement on several imbalanced data sets had been observed [20].

B. *The Classification Studies on Fertility Medical Dataset*

Piles et al., (2013) categorized Fertility rate into five classes of equal length. Linear Regression (LR), Ordinal Logistic Regression (OLR), Support Vector Regression (SVR), Support Vector Ordinal Regression (SVO R), and Non-deterministic Ordinal Regression (NDOR) were compared in terms of their predictive ability with two base line algorithms: MEAN and MODE which always predict the mean and mode value of the classes observed in the data set, respectively. Predicting ability was measured in terms of rate of erroneous classifications, linear loss (average of the distance between the predicted and the observed classes), the number of predicted classes and the F1 statistic (which allows comparing procedures taking into account that they can predict different number of classes). For all methods, the reduced models showed almost an irrelevant decrease in their predictive abilities compared to the corresponding values obtained with the full models [21].

Gil et al.,(2012) compared three artificial intelligence techniques, decision trees, Multilayer Perceptron and Support Vector Machines, in order to evaluate their performance in the prediction of the seminal quality from the data of the environmental factors and lifestyle. From the studied methods, Multilayer Perceptron and Support Vector Machines were the most accurate in the prediction. Therefore these tools, together with the visual help that decision trees offer, were the suggested methods to be included in the evaluation of the infertile patient [22].

C. *The Classification Studies on Thoracic Surgery Medical Datasets*

Zieba et al.(2014) applied boosted SVM for medical application of predicting post-operative life expectancy in the lung cancer patients [23]. Harun et al., (2015) presented a guide of the application of data mining techniques for predicting outcome of thoracic surgery. A wide range of

data mining techniques was considered by using WEKA software package [24]. Shahian et al. (2001) [25] used artificial neural network for classification of thoracic surgery data. Sindhu et al.(2014) [26] used various classification techniques to analyse thoracic surgery data and they found that J48 gives better accuracy.

III. Material and Method

A. *Medical Dataset*

Thoracic Surgery Medical Datasets: The data is dedicated to classification problem related to the post-operative life expectancy in the lung cancer patients: class 1 - death within one year after surgery, class 2 – survival [27, 28].The dermatology dataset used in this study was taken from UCI Machine Learning Repository [28]. The aim of the dermatology data set is to diagnose one of six possible types of erythemato-squamous diseases (6 classes). Twelve clinical and 24 histopathological measurements of the patient are given (36 attributes). The data set contains 366 patterns, which are to be used for both classifier design and testing.

Fertility Medical Dataset: 100 volunteers provided a semen sample analysed according to the WHO 2010 criteria. Sperm concentration are related to socio-demographic data, environmental factors, health status, and life habits. There are 9 attributes and 1 output in the dataset [29].

EEG Eye State Medical Dataset: All data is from one continuous EEG measurement with the Emotiv EEG Neuroheadset. The duration of the measurement was 117 seconds. The eye state was detected via a camera during the EEG measurement and added later manually to the file after analysing the video frames. '1' indicates the eye-closed and '0' the eye-open state. All values are in chronological order with the first measured value at the top of the data. There are 14 attributes and 1 output in the dataset. Total number of the record is 14980 [30].

B. *Software-WEKA*

Weka (Waikato Environment for Knowledge Analysis) written in Java, developed at the University of Waikato, New Zealand (6). Weka supports several standard data mining tasks, more specifically, data preprocessing, clustering, classification, regression, visualization, and feature selection. All techniques of Weka's software are predicated on the assumption that the data is available as a single flat file or relation, where each data point is described by a fixed number of attributes (normally, numeric or nominal attributes, but some other attribute types are also supported) [31].

C. *Methods*

BayesNet: It is probabilistic graphical model (a type of statistical model) that represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG). For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. [32].

Naïve Bayes: Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. [32].

MultilayerPerceptron: A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate outputs. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. [32].

SMO: Sequential Minimal Optimization (SMO) is a new algorithm for training Support Vector Machines (SVMs). Training a support vector machine requires the solution of a very large quadratic programming (QP) optimization problem. [32].

KStar: K-star or K^* is an instance-based classifier. The class of a test instance is based on the training instances similar to it, as determined by some similarity function. It differs from other instance-based learners in that it uses an entropy-based distance function. [33].

Locally Weight Learning (LWL): It has the standard form $y = f(x) + \epsilon$, where $x \in \mathbb{R}^n$ is a n -dimensional input vector, the noise term has mean zero, $E\{\epsilon\} = 0$, and the output is one-dimensional. The key concept is to approximate nonlinear functions by means of piecewise linear models [34],

Bagging: The Bagging (Bootstrap Aggregating) algorithm [1] uses bootstrapping (equiprobable selection with replacement) on the training set to create many varied but overlapping new sets. The base algorithm is used to create a different base model instance for each bootstrap sample, and the ensemble output is the average of all base model outputs for a given input.[35, 36]

ClassificationViaRegression: It is possible to use regression methods to solve classification tasks. In order to apply the continuous prediction technique of regression models to discrete classification problems, an approximation of the conditional class probability function can be considered. [37].

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. [38]

Decorate: Diverse Ensemble Creation by Oppositional, Relabeling of Artificial Training Examples (Decorate) is presented that uses a learner (one that provides high accuracy on the training data) to build a diverse committee. This is accomplished by adding different randomly constructed examples to the training set when building new committee members. [38].

DecisionTable: Decision Table is an accurate method for numeric prediction from decision trees and it is an ordered set of If-Then rules that have the potential to be more compact and therefore more understandable than the decision trees [39].

Decision Table/Naive Bayes (DTNB): Class for building and using a decision table/naive Bayes hybrid classifier. At each point in the search, the algorithm evaluates the merit of

dividing the attributes into two disjoint subsets: one for the decision table, the other for naive Bayes. [40].

Java Repeated Incremental Pruning (JRip): JRip is a propositional rule learner, i.e. Repeated Incremental Pruning to Produce Error Reduction (RIPPER). Initial rule set for each class is generated using IREP. [41].

OneR: Class for building and using a 1R classifier; in other words, uses the minimum-error attribute for prediction, discretization numeric attributes [42].

Partial C4.5 (PART): The PART technique avoids global optimization step used in C4.5 rules and RIPPER. It generates an unrestricted decision list using basic separate and-conquer procedure. It builds a partial decision tree to obtain a rule. [43].

J48: J48 algorithm is a modified version of c4.5 and ID3 algorithm which is used to construct the decision trees. The decision tree uses tree like graph and acts as decision support system. [44].

Logistic Model Tree (LMT): LMT is a classifier that combines logistic regression and decision tree learning. LMT uses cost-complexity pruning. This algorithm is significantly slower than the other algorithms [45].

Naive Bayes Tree (NBTree): The NBTree algorithm is a hybrid between decision-tree classifiers and Naive Bayes classifiers. It represents the learned knowledge in the form of a tree which is constructed recursively. [46].

Random Forest (RF): RF is an algorithm based on ensemble techniques for classification and regression analysis [47]. The main principle in RF is the combination of various decision trees using several bootstrap samples and selecting a subset of explanatory variables at every node. [48].

iv. Experimental Study

Following classifier techniques of WEKA have been applied to EEG Eye State, Fertility and Thoracic Surgery Medical Datasets: Bayes Net, Naïve Bayes, Logistic, Multilayer Perceptron, RBF Network, SMO, KStar, LWL, Bagging, Classification Via Regression, Dagging, Decorate, Grading, Decision Table, DTNB, JRip, OneR, PART, BFTree, J48, LMT, NBTree, Random Forest. The results obtained from related classification techniques were presented in Table 1 according to each dataset

v. Conclusions

In this study the performances of 23 different classification methods were evaluated in terms of classification accuracy on three different medical datasets namely EEG Eye State, Fertility and Thoracic Surgery Datasets. It has been seen that Multilayer Perceptron (MLP) had highest accuracy for Fertility Dataset (90%), three different techniques namely Bagging, Dagging and Grading had highest and same accuracies for Thoracic Surgery Data Set (85.1064%) and finally Kstar had highest accuracy for EEG Eye State Dataset (96.7757%). JRip had worst accuracy for Fertility Dataset (90%), Decorate had worst accuracy for Thoracic Surgery Data Set (82,3404%) and

finally OneR had worst accuracy for EEG Eye State Dataset (62,6035%).method.

TABLE I. RATIO OF EACH CLASSIFICATION TECHNIQUE ON EACH DATASET

Method	Accuracy (%)		
	Fertility	Thoracic Surgery	EEG Eye State
Bayes Net	88,0000%	81,9149%	68,3111%
Naive Bayes	88,0000%	78,5106%	46,7690%
Logistic	88,0000%	82,7660%	63,4513%
Multilayer Perceptron	90,0000%	80,2128%	54,8064%
RBF Network	87,0000%	83,1915%	55,8945%
SMO	88,0000%	84,8936%	55,1268%
KStar	84,0000%	81,4894%	96,7757%
LWL	86,0000%	84,4681%	59,2991%
Bagging	88,0000%	85,1064%	89,3858%
Classification Via Regression	86,0000%	84,0426%	86,2550%
Dagging	88,0000%	85,1064%	60,6943%
Decorate	88,0000%	82,3404%	91,5621%
Grading	88,0000%	85,1064%	55,1202%
Decision Table	88,0000%	84,4681%	73,1041%
DTNB	88,0000%	81,7021%	75,4673%
JRip	87,0000%	84,8936%	83,5714%
OneR	88,0000%	83,4043%	62,6035%
PART	88,0000%	79,5745%	83,6515%
BF Tree	87,0000%	84,6809%	84,3792%
J48	85,0000%	84,4681%	84,4993%
LMT	89,0000%	84,6809%	87,7704%
NB Tree	87,0000%	82,3404%	83,7850%
Random Forest	85,0000%	83,8298%	93,5381%

For the future work more classification algorithms should be applied to more medical datasets to see impacts of the different performance of algorithms on different medical datasets. Additionally one can also study one classification technique to adapt it to have best results on different medical datasets. This can be done by studying hard on the parameters to related classification technique to adapt them according to the structures of different medical datasets.

Acknowledgment

This study has been supported by Scientific Research Project of Selcuk University.

References

[1] Trilok Chand Sharma, Manoj Jain, WEKA Approach for Comparative Study of Classification Algorithm, International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 4, April 2013

[2] O. Rosler and D. Suendermann, "First step towards eye state prediction using EEG," in Proceedings of the International Conference on Applied Informatics for Health and Life Sciences (AIHLS '13), Istanbul, Turkey, 2013.

[3] David Gil, Jose Luis Girela, Joaquin De Juan, M. Jose Gomez-Torres, Magnus Johnsson, Predicting seminal quality with artificial intelligence methods, Expert Systems with Applications, Volume 39, Issue 16, 15 November 2012, Pages 12564-12573.

[4] UCI Machine Learning Repository, <https://archive.ics.uci.edu/ml/datasets/Thoracic+Surgery+Data#>, Last accessed date: September 2014.

[5] Zieba, M., Tomczak, J. M., Lubicz, M., & Swiatek, J. (2014). Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients, Applied Soft Computing, 14, pp. 99-108.

[6] Weka: <http://www.cs.waikato.ac.nz/~ml/weka/> Last access: 10.04.2015

[7] O. Fukuda, T. Tsuji, and M. Kaneko, "Pattern classification of EEG signals using a log-linearized Gaussian mixture neural network," in Proceedings of the IEEE International Conference on Neural Networks. Part 1 (of 6), pp. 2479-2484, Perth, Australia, December 1995.

[8] M. V. M. Yeo, X. Li, K. Shen, and E. P. V. Wilder-Smith, "Can SVM be used for automatic EEG detection of drowsiness during car driving?" Safety Science, vol. 47, no. 1, pp. 115-124, 2009.

[9] K. Polat and S. Güneş, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform," Applied Mathematics and Computation, vol. 187, no. 2, pp. 327-1026, 2007.

[10] N. Sulaiman, M. N. Taib, S. Lias, Z. H. Murat, S. A. M. Aris, and N. H. A. Hamid, "Novel methods for stress features identification using EEG signals," International Journal of Simulation: Systems, Science and Technology, vol. 12, no. 1, pp. 27-33, 2011.

[11] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," ACM SIGKDD Explorations Newsletter, vol. 11, pp. 10-18, 2009.

[12] O. Rösler and D. Suendermann, "First step towards eye state prediction using EEG," in Proceedings of the International Conference on Applied Informatics for Health and Life Sciences (AIHLS '13), Istanbul, Turkey, 2013.

[13] A. Frank and A. Asuncion, "UCI machine learning repository," <http://archive.ics.uci.edu/ml>. 2014.

[14] Ting Wang, Sheng-Uei Guan, Ka Lok Man, and T. O. Ting, EEG Eye State Identification Using Incremental Attribute Learning with Time-Series Classification, Hindawi Publishing Corporation, Mathematical Problems in Engineering, Volume 2014, Article ID 36532, 9 pages

[15] O. Roesler and D. Suendermann: A First Step towards Eye State Prediction Using EEG. In Proc. of the AIHLS 2013, International Conference on Applied Informatics for Health and Life Sciences, Istanbul, Turkey, September 2013

[16] Mridu Sahu, N. K. Nagwani., Shrish Verma, Saransh Shirke, An Incremental Feature Reordering (IFR) Algorithm to Classify Eye State Identification Using EEG, Information Systems Design and Intelligent Applications, Advances in Intelligent Systems and Computing Volume 339, 2015, pp 803-811

[17] O. Roesler, L. Bader, J. Forster, Y. Hayashi, S. Hessler, and D. Suendermann-Oeft, Comparison of EEG Devices for Eye State Classification. In Proc. of the AIHLS 2014, International Conference on Applied Informatics for Health and Life Sciences, Kusadasi, Turkey, October 2014.

[18] Kaveh Hassani and Won-Sook Lee, An Incremental Framework for Classification of EEG Signals Using Quantum Particle Swarm Optimization, Conference: 2014 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA) DOI: 10.1109/CIVEMSA.2014.6841436

[19] Rajesh Singla, Brijil Chambayil, Arun Khosla, Jayashree Santosh, Comparison of SVM and ANN for classification of eye events in EEG, J. Biomedical Science and Engineering, 2011, 4, 62-69, doi:10.4236/jbise.2011.41008.

[20] Razzaghi, T., Roderick, O., Safro, I., & Marko, N. (2015). Fast Imbalanced Classification of Healthcare Data with Missing Values. arXiv preprint arXiv:1503.06250.

- [21] M. Piles, J. Díez, J.J. del Coz, E. Montañés, J.R. Quevedo, J. Ramon O. Rafel, M. López-Béjar, L. Tusell, Predicting fertility from seminal traits: Performance of several parametric and non-parametric procedures, *Livestock Science* 155 (2013) 137–147.
- [22] David Gil, Jose Luis Girela, Joaquin De Juan, M. Jose Gomez-Torres, Magnus Johnsson, Predicting seminal quality with artificial intelligence methods, *Expert Systems with Applications* 39 (2012) 12564–12573.
- [23] Maciej Zieba, Jakub M. Tomczak, Marek Lubicz, Jerzy Swiatek, Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients, *Applied Soft Computing* 14 (2014) 99–108.
- [24] Ahasan Uddin Harun and Nure Alam, Predicting Outcome of Thoracic Surgery by Data Mining Techniques, *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 5, Issue 1, January 2015, ISSN: 2277 128X.
- [25] D. M. Shahian, S. L. Norman, D.F. Torchiana, “Cardiac Surgery Report Cards: Comprehensive Review and Statistical Critique” , *Annals of Thoracic Surgery*, vol. 72, pp 2155-2168, 2001.
- [26] V. Sindhu, S. A. S. Prabha, S. Veni , and M. Hemalatha, “Thoracic surgery analysis using data mining techniques” , *International Journal of Computer Technology & Applications* , vol. 5 pp 578-586, May,2014.
- [27] UCI Machine Learning Repository, <https://archive.ics.uci.edu/ml/datasets/Thoracic+Surgery+Data#>, Last accessed date: September 2014.
- [28] Zieba, M., Tomczak, J. M., Lubicz, M., &Swiatek, J. (2014). Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients, *Applied Soft Computing*, 14, pp. 99-108.
- [29] David Gil, Jose Luis Girela, Joaquin De Juan, M. Jose Gomez-Torres, Magnus Johnsson, Predicting seminal quality with artificial intelligence methods, *Expert Systems with Applications*, Volume 39, Issue 16, 15 November 2012, Pages 12564-12573.
- [30] Oliver Roesler, [it12148 '@' lehre.dhbw-stuttgart.de](mailto:it12148@lehre.dhbw-stuttgart.de) , Baden-Wuerttemberg Cooperative State University (DHBW), Stuttgart, Germany.
- [31] Rohit Arora and Suman, Comparative Analysis of Classification Algorithms on Different Datasets using WEKA, *International Journal of Computer Applications* (0975 – 8887) Volume 54– No.13, September 2012
- [32] Weka: <http://www.cs.waikato.ac.nz/~ml/weka/> Last access: 10.04.2015.
- [33] John G. Cleary, Leonard E. Trigg: “K*: An Instance based Learner Using an Entropic Distance Measure”, 12th International Conference on Machine Learning, 108-114, 1995.
- [34] W. S. Cleveland, “Robust locally weighted regression and smoothing scatterplots,” *Journal of the American Statistical Association*, vol. 74, pp. 829-836, 1979.
- [35] Breiman, L., “Bagging Predictors”, *Machine Learning*, Vol. 24, No. 2, pp. 123–140, 1996.
- [36] Barutcuoglu Zafer and Alpaydin Ethem, A Comparison of Model Aggregation Methods for Regression, *ICANN/ICONIP 2003, LNCS 2714*, pp. 76–83, 2003.
- [37] Frank Eibe, Wang Yong, Inglis Stuart, Holmes Geoffrey, Witten Ian H., Using Model Trees for Classification, *Machine Learning*, Volume 32, Issue 1, pp 63-76. July 1998,
- [38] Bhoomi Trivedi, Neha Kapadia, INDUS institute of Eng. & Tech,TCET, Kandivali(E),Ahmedabad Modified Stacked generalization withSequential Learning. TCET 2012 on IJCA.
- [39] Kohavi R. (1995), “The Power of Decision Tables”, 8th European Conference on Machine Learning, pp. 174-189.
- [40] Mark H. and Eibe F., Combining Naive Bayes and Decision Tables, In: *Proceedings of the 21st Florida Artificial Intelligence Society Conference (FLAIRS)*, 2008.
- [41] Cohen William W. (1995) Fast Effective Rule Induction, Twelfth International Conference on Machine Learning, pp.115-123.
- [42] Holte R.C., (1993). Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11, pp.63-91.
- [43] Frank, E. and Witten, I. (1998). Generating Accurate Rule Sets Without Global Optimization, Shavlik, J., ed., *Machine Learning: Proceedings of the Fifteenth International Conference*, pp., 144-151, Madison, Wisconsin, Morgan Kaufmann, San Francisco.
- [44] Quinlan J.R., Improved use of continuous attributes in C4.5, *J Artif Res*, 4 (1996), pp. 77–90
- [45] Landwehr N., Hall M., Frank E., Logistic model trees, *European Conference on machine learning*, Springer-Verlag, Cavtat-Dubrovnik, Croatia (2003), pp. 241–252.
- [46] R. Kohavi. Scaling up the accuracy of Naive-Bayes classifiers: a decision-tree hybrid. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pages 202–207, 1996.
- [47] Breiman, L., *RandomForests*, *Machine Learning*, 45 (2001), pp. 5–32
- [48] Liaw A. and Wiener M., Classification and regression by randomForest, *R News*, 2 (3) (2002), pp. 18–22.