Celal Bayar University Journal of Science

Voting Combinations Based Ensemble: A Hybrid Approach

Abdul Ahad Abro[1]*, Mir Sajjad Hussain Talpur[2], Awais Khan Jumani[3], Waqas Ahmed Siddique[3], Erkan Yaşar[1]

[1]Department of Computer Engineering, Ege University, Izmir, Türkiye
[2]Information Technology Centre, Sindh Agriculture University, Pakistan
[3]Department of Computer Science, Ilma University, Pakistan
*abdulahadabro1@gmail.com
*0000-0002-3591-9231

Received: 2 November 2021
Accepted: 12 September 2022
DOI: 10.18466/cbayarfbe.1014724

Abstract

In the field of Artificial Intelligence (AI), Machine Learning (ML) is a well-known and actively researched concept that assists to strengthen the accomplishment of classification results. The primary goal of this study is to categories and analyze ML and Ensemble Learning (EL) techniques. Six algorithms Bagging, C4.5 (J48), Stacking, Support Vector Machine (SVM), Naive Bayes (NB), and Boosting as well as the five UCI Datasets of ML Repository are being used to support this notion. These algorithms show the robustness and effectiveness of numerous approaches. To improve the performance, a voting-based ensemble classifier has been developed in this research along with two base learners (namely, Random Forest and Rotation Forest). Whereas important parameters have been taken into account for analytical processes, including: F-measure values, recall, precision, Area under Curve (Auc), and accuracy values. As a result, the main goal of this research is to improve binary classification and values by enhancing ML and EL approaches. We illustrate the experimental results that demonstrate the superiority of our model approach over well-known competing strategies. Image recognition and ML challenges, such as binary classification, can be solved using this method.

Keywords: Artificial Intelligence, Data Mining, Machine Learning, Pattern Recognition.

1. Introduction

To discover; the relation and patterns in enormous datasets, sophisticated data analysis tools are being adopted and utilized for the extraction of data mining techniques [1]. Numerous theoretical and empirical research that demonstrate the benefits of the combination paradigm over separate classifier models have been published [2-5]. In recent years, ML has gained significant traction in a number of areas, including remote sensing, image categorization, and pattern identification.

These resources are interdisciplinary research fields including mathematical algorithms, statistical models, ML techniques, and intelligent information systems, etc [6].

A C4.5 decision is produced by the clear and easy-to-use algorithm J48 [7]. The classification process is modelled using a binary tree. It is a successor to the ID3 algorithm. Recursively choosing the test feature with the highest knowledge gain frequency as the test feature in [8] is an effective assessment model that eventually yields acceptable results.

One-dimensional convolutional neural network (1D-CNN), stacking-based ensemble deep learning model to carry out a multiclass classification on the five most prevalent kinds of cancer based on RNASeq data. The results of the single 1D-CNN, support vector machines with radial basis function, linear, and polynomial kernels, artificial neural networks, k-nearest neighbors, and bagging trees with the results of the novel suggested model with and without LASSO. The findings demonstrate that the suggested model, both with and without LASSO, outperforms competing classifiers. Additionally, the results demonstrate that under sampling improves performance compared to oversampling for the machine learning algorithms Support Vector Machine (SVM)-R, SVM-L, SVM-P, Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), and bagging trees.[9].

Scientific classifications bagging and Neural Networks among the best representations available for the other models. As ML models are used to assess the risk of the most common deadly diseases with low occurrence, it produces considerable presentation. ML outperforms traditional regression for illness forecast modelling when the likelihood of disease occurrence is low. [10].

The conceivable outcome is aggregated by the NB classifier with the Bayes paradigm in decision rules. The learning framework for large-scale computational value and multi-domain platform classification [11].

Multiple continuous and categorical variables can be handled by the robust and adaptable SVM method. In addition, the overall results and comparisons are provided, highlighting the BER drops considerable non-linear explanation. The SVM multi-in-phase classifier's and quadrature components are largely reliant on the in-phase and quadrature components, which are comparably perfect when considering the impacts of intention and storage [12].

The choice of accessible base classifiers and combiner techniques are two of the primary obstacles in creating an ensemble [13]. The stack of ensemble (SoE) is an ensemble classifier that uses parallel architecture to merge three separate ensemble learners—Random Forest, Gradient Boosting Machine, and Extreme Gradient Boosting Machine in a consistent way. According to their Matthews correlation coefficients, accuracies, false positive rates, and area under ROC curve metrics that are satisfactory in terms of the analyzed parameters, classification algorithms performance importance is statistically examined.

There are several sections to the paper. Section 2 provides a brief description of the literature review. Section 3 contains the proposed approach used for carrying out various tests. Section 4 includes performance evaluation, experimental analysis, and detailed datasets. Finally, Section 5 recommends further work based on the findings and draws a conclusion.

2. Literature Review

Recently, research efforts focused on bagging, C4.5, stacking, SVM, boosting, and classification have increased [14]. In this study, we employ supervised learning's binary classification method. In classification, the goal class is anticipated properly and suitably for each situation involving data. The model is creating the training process, and a classification algorithm incorporates the standards of the analysts and the objectives [15]. Many classification algorithms use different methods to find associations. These relationships are models that can be applied to diverse datasets where the class is unknown [16]. The model was trained with the combined prediction model utilizing a vote-based ensemble learning technique. It demonstrates that when the vote-based ensemble method is combined with an ANN, the results are more accurate than those produced by an ANN alone

When the C-C4.5 procedure's applications on noisy facts are compared to the C4.5 algorithm in C4.5 [17], it is found to be more reliable. The various locations have a big impact on how the C-C4.5 method is presented. The C-C4.5 trees with large standards yield the outcomes that are effective and accurate on average.

In [18] stacking strategy for creating ensembles of machine learning models is described. The cases for logistic regression and time series forecasting have been taken into consideration. The findings indicate the enhancement in the performance of prediction models in the scenarios under consideration by applying stacking techniques [19].

In [20], properties of bagging and NB are being investigated and a contrast is made between them. The hybrid bagging-NB prototyping approach, which strategically monitors the pattern of controlling the tradeoff between prototypical bias and prototypical variation, reduces the sum of inaccuracies. By enhancing fewer factors, the hybrid prototype offers an improvement that is authentic in terms of the training period.

The most well-known and distinctive data mining algorithms in are NB [21]. According to the empirical findings, the intended NB exhibits improved classification performance while preserving simplicity and flexibility.

In [22], presented the concepts of incorporating imprecise previous knowledge and sophisticated ML SVM-constructed procedures techniques. It utilizes the duality illustration in the framework of the minimax approach of decision making, which allows us to get straightforward extensions of SVMs, comprising supplementary limitations for optimization variables.

Boosting is a basic classification method that generates a single-level decision tree, as described. It has the capability to grip misplaced values and numeric features representing flexibility instead of the easiness. The Boosting procedure creates set of instruction and every characteristic in the training data, then captures the instruction with the least error rate.

3. Proposed Methodology

The pre-processing step of the data and the classification algorithms utilized in this study are described in the overview of the suggested technique provided in this part.

3.1. Proposed System

The proposed system is given in Figure. 1. It consists of numerous phases: datasets, base learners and comparative analysis of results. Besides, the generalization presentation of the system, 10-fold cross-validation is helpful intended for all classifier learners and datasets.

3.2. Data Pre-processing

The data from various ML datasets may have high range values. In this situation, specific features may have a considerable positive or negative impact on the classification accuracy of algorithms. Therefore, using the min-max normalization technique [23], data standards are restricted to the [0,1] range.

3.3. Classification of Algorithms

In this study, base learners, including bagging, C4.5, stacking, SVM, NB and boosting, are employed.

There are numerous phases of method related to datasets and classifiers focused on ML. In this work, six ML classifiers, along with five datasets, are experienced for binary classification.

Among all the methods, including NB, multilayer perceptron, C4.5 and random forest produces effective outcomes. This hybrid algorithm offers a classification accuracy of 75.625 percent. Then, the C4.5 method and random forest algorithm were integrated, yielding a classification accuracy of 76.4583%. Compared to individual classification algorithms, the hybrid classification algorithm is more accurate [24].

The most widely utilized fraud detection approaches are NB, SVM, and KNN. These methods can be used independently or in conjunction with one another to create classifiers utilizing ensemble or meta-learning methods. Ensemble learning techniques, however, stand out among the rest of the methods available not just for their ease of use but also for their extraordinary ability to predict outcomes in real-world situations [25]. Due to its independence from attribute values, the bagging classifier based on decision three performs well with this kind of data.

In [26], NB and random forest overlap the implementation, and both ML techniques outperform a number of algorithms. ML techniques such as bagging, NB, and random forest can identify persistence at the population level. Even though all methods would have resulted in the same results in reality, it is preferable to pick the most appropriate course of action for every situation.

In [27], a set of rules is suggested to improve the feature subclasses of models, and integrate the constraints of the SVM to use the sorting in a proper manner. The experimental outcomes depict that the procedure has a satisfactory consequence on the classification of adequate instant messaging evidence of the Internet of things big data and has a virtuous impact and applied application value.

Boosting classification algorithm produces for each analyst in the data population. In [28], a procedure for ML is a general evident unexpectedly useful on the ordinary datasets generally used for evaluating the results. It takes as input a set of incidents, each with various features and a category like other learning methods. The boosting algorithm selects the most revealing single feature and bases the idea on this feature alone. However, the result is not satisfactory with continuous-esteemed features and handling the hidden values.

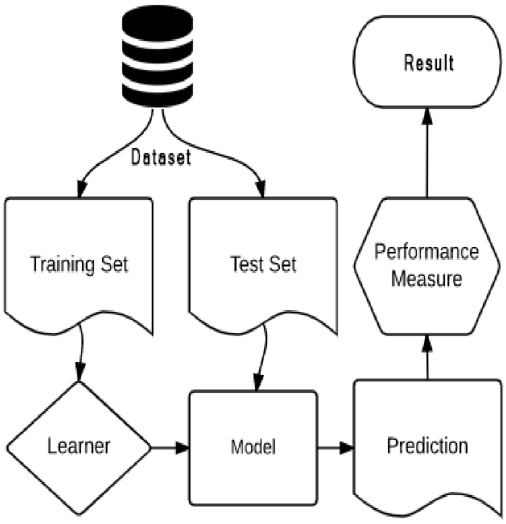

Figure 1. Proposed Layout.

4. Experiment and Analysis

In these subsections, we describe and present the experimental process, evaluation measures and experimental outcomes.

4.1 Experimental Process

The datasets utilized in the experiment extracted from the UCI ML Repository [29].

All studies rely on a total of 6 ML and EL classifiers thanks to the use of the WEKA (Waikato Environment for Knowledge Analysis) ML tools and the Java programming language. For all of WEKA's classifiers, we used the default parameter values [30].

To produce accurate findings, we use 10-fold cross-validation to all datasets. The initial dataset is subjected to the 10-fold cross-validation after being randomly divided into ten sets of equal size, one of which is utilized for test validation and the rest for testing. Ten times of the technique are done, and the averages of the results are calculated.

The attributes and the overall number of instances are considered when evaluating dataset properties. Usually, these datasets are utilized to address ML-related issues. Table 1 shows various numerical properties, instances, and class descriptions. The datasets were picked out of the UCI ML Repository based on the unique attributes that were being used for binary classification issues.

Table 1. Datasets detail.

Datasets	Instances	Attributes	Classes
Arrhythmia	452	279	2
Balance Scale	625	4	2
Car Evaluation	1728	6	4
Iris	150	4	3
Spambase	4601	8	6

The datasets used in this work have been considered suitable for classification, and various supervised ML techniques have been used. The performance measurements, however, are determined using confusion matrices to solve binary classification problems.

4.2 Assessment of Measures

This section describes the five performance evaluation measures of the proposed method, consisting of accuracy, Auc, precision, recall and F-measure.

Accuracy reflects how close an agreed number is to a measurement. It is specified further in Equation (1).

$$\text{Acc} = \left(\frac{TP+TN}{TP+FP+FN+TN}\right) \quad \textbf{(1)}$$

In equation 1, TN, FN, FP and TP show the number of True Negatives, False Negatives, False Positives and True Positives. The Auc represents the area under the ROC Curve. It procedures the whole two-dimensional region under the entire ROC curve from (0,0) to (1,1).

Precision is a positive analytical value [31]. Precision defines how reliable measurements are, although they are farther from the accepted value. The equation of precision is shown in Equation (2).

$$\text{Precision} = \left(\frac{TP}{TP+FP}\right) \quad \textbf{(2)}$$

The recall is the hit rate [32]. The recall is the reverse of precision; it calculates false negatives against true positives. The equation is illustrated in Equation (3).

$$\text{Recall} \quad \left(\frac{TP}{TP+FN}\right) \quad \textbf{(3)}$$

F-measure can be defined as the weighted average [32] of precision and recall [33]. This rating considers both false positives and false negatives. The equation is illustrated in Equation (4).

$$F = 2\,x\,\frac{Precision * Recall}{Precision + Recall} \quad \textbf{(4)}$$

These criteria are adjusted proportionally in the data by the reference class prevalence in the weighting operation.

Table 2. Results of ML methods for Arrhythmia dataset.

Arrhythmia					
Methods	**Acc (%)**	**Auc**	**Precision**	**Recall**	**F-Measure**
J48	85.1705	0.875	0.845	0.852	0.845
Stacking	79.8718	0.722	0.798	0.799	0.798
Bagging	**85.6988**	0.911	0.851	0.857	0.852
NB	82.3779	0.902	0.846	0.824	0.831
SVM	85.4367	0.764	0.848	0.854	0.848
Boosting	79.6569	0.592	0.797	0.797	0.747

Table 3. Results of ML methods for Balance Scale dataset.

Balance Scale					
Methods	**Acc (%)**	**Auc**	**Precision**	**Recall**	**F-Measure**
J48	**75.5245**	0.584	0.752	0.755	0.713
Stacking	72.3776	0.628	0.699	0.724	0.697
Bagging	68.8811	0.646	0.668	0.689	0.675
NB	71.6783	0.701	0.704	0.717	0.708
SVM	69.5804	0.590	0.671	0.696	0.677
Boosting	65.7343	0.542	0.624	0.657	0.635

Table 4. Results of ML methods for car evaluation dataset.

Car Evaluation					
Methods	**Acc (%)**	**Auc**	**Precision**	**Recall**	**F-Measure**
J48	92.3611	0.976	0.924	0.924	0.924
Stacking	93.5185	0.997	0.940	0.935	0.925
Bagging	93.1134	0.990	0.932	0.931	0.931
NB	85.5324	0.976	0.852	0.855	0.847
SVM	**93.7500**	0.953	0.939	0.938	0.938
Boosting	70.0231	0.500	0.700	0.700	0.824

Table 5. Results of ML methods for iris dataset.

Iris					
Methods	**Acc (%)**	**Auc**	**Precision**	**Recall**	**F-Measure**
J48	96.0000	0.968	0.960	0.960	0.960
Stacking	95.3333	0.966	0.953	0.953	0.953
Bagging	**96.1000**	0.981	0.960	0.960	0.960
NB	96.0000	0.994	0.960	0.960	0.960
SVM	96.0000	0.978	0.962	0.960	0.960
Boosting	92.0000	0.940	0.920	0.920	0.920

Table 6. Results of ML methods for Spambase dataset.

Spambase					
Methods	**Acc (%)**	**Auc**	**Precision**	**Recall**	**F-Measure**
J48	55.9299	0.733	0.549	0.559	0.553
Stacking	52.2911	0.685	0.524	0.523	0.522
Bagging	**58.6253**	0.825	0.585	0.586	0.577
NB	57.6146	0.816	0.585	0.576	0.566
SVM	57.0755	0.781	0.478	0.571	0.596
Boosting	53.2911	0.585	0.404	0.400	0.517

Table 7. Proposed Voting-Based Hybrid Apporach.

Hybrid Approach Voting-Based Random and Rotation Forest						
Datasets	**Acc (%)**	**Impr. (%)**	**Auc**	**Precision**	**Recall**	**F-Measure**
Arrhythmia	**85.7222**	0.0234	0.902	0.840	*0.857	**0.856**
Balance Scale	**75.5247**	**0.0002**	**0.651**	**0.707**	0.731	0.691
Car Evaluation	**97.8009**	**4.0509**	**0.999**	**0.979**	**0.978**	**0.978**
Iris	***96.1000**	**0**	**0.995**	0.947	0.947	0.947
Spambase	**62.4663**	**3.841**	**0.851**	**0.620**	**0.625**	**0.618**

- * Indicates the similar performance results concerning base learner.
- High Acc, Auc, Precision, Recall and F- measure is shown in Bold, while the greyed shows insufficient results.
- Impr. represents improvement according to best results of Tables 2-6.

4.3 Experimental Results

There are several algorithms for classification of which the most well-known and widely applicable dataset.

Tables 2-6 for all datasets present accuracy, Auc, precision, recall and F-measurement values of ML algorithms. In Table 2-6, high Acc, Auc, Precision, Recall and F-measure are shown in Bold, while the greyed shows insufficient results.

To sum up, Tables 2-6, has been designed in terms of different specifications according to the multiple datasets relating to the numerous approaches to ML. In Table 2, bagging has better outcomes, which provides 85.6988% Acc in comparison to others. Probably, in Table 3, J48 indicates 75.5245% Acc adequate consequences. Similarly, in Table 4, the SVM presents 93.7500% Acc effective results. Likewise, in Table 5, the bagging illustrates the 96.1000% Acc productive outcomes. However, in the end, bagging shows a 58.6253% Acc result in Table 6. Moreover, it is analyzed that bagging in Arrhythmia dataset Table 2, provides positive findings. Likely, J48 in the Balance Scale dataset concerning Table 3, indicates the progressive result.

Similarly, Table 4, SVM presents effective results in the car evaluation dataset. Likewise, in Table 5, the iris dataset bagging provides a more accurate outcome and It indicates adequate consequences in Table 6, Spambase dataset. Finally, high Acc, Auc, precision, recall and F-measure is shown in Bold, while the greyed shows insufficient results.

Table 7, demonstrates the comparison of all datasets results, with respect to our proposed voting-based hybrid approach meta-ensemble method. As it is clearly shown in Table 7, a meta-ensemble classifier, voting with two base learners (namely, Random Forest and Rotation Forest) provide highly accurate outcomes as compare to others.

5. Conclusıon and Future Work

The outcomes of supervised ML algorithms such as bagging, C4.5, Stacking, SVM, NB and boosting to classify numerous datasets. Algorithm effectiveness is further broken down into recall/sensitivity, accuracy, precision, and F-score categories. A retrospective study that analysed the different sizes of training and test sets can have a significant impact on the sensitivity and specificity of the same algorithm. This study suggests a hybrid voting-based technique. With this strategy, we may produce more beneficial and successful results by utilising the advantages of these algorithms. Other data mining approaches, such as clustering and association, can benefit from this research.

We intend to enhance our research on classification models in the future by using a hybrid framework of an intelligent ML system to a large number of real-world datasets.

Author's Contributions

Abdul Ahad ABRO: Drafted and wrote the manuscript, performed the experiment and result analysis.

Mir Sajjad Hussain TALPUR: As the project consultant, supervised the works and helped prepare the manuscript.
Awais Khan JUMANI: Assisted in analytical analysis on the structure, supervised the experiment's progress, result interpretation and helped in manuscript preparation.
Waqas Ahmed SİDDİQUE: Assisted in analytical analysis on the structure, supervised the experiment's progress, result interpretation and helped in manuscript preparation.
Erkan YAŞAR: Searched the literature and helped in manuscript preparation.

Ethics

There are no ethical issues after the publication of this manuscript.

References

[1]. Accorsi R, Manzini R, Pascarella P, Patella M, Sassi S. "Data Mining and Machine Learning for Condition-based Maintenance". *Procedia manufacturing*, 11,1153–1161, 2017.

[2]. Shao Y, Liu Y, Ye X, Zhang S. "A Machine Learning based global simulation data mining approach for efficient design changes". *Advances in Engineering Software,*124, 22–41, 2018.

[3]. Hüllermeier E. "Fuzzy sets in Machine Learning and data mining". *Applied Soft Computing,* 11(2). 1493–1505, 2011.

[4]. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I." Machine Learning and data mining methods in diabetes research". *Computational and structural biotechnology journal*, *15*, 104-116, 2017.

[5]. Shafiq M, Tian Z, Bashir AK, Jolfaei A, Yu X." Data mining and Machine Learning methods for sustainable smart cities traffic classification: a survey". *Sustainable Cities and Society*, *60*, 102177, 2020

[6]. Deepajothi S, Selvarajan S. "A comparative study of classification techniques on adult data set". *International Journal of Engineering Research & Technology (IJERT)*, *1,* 2012.

[7]. Bansal D, Chhikara R, Khanna K, Gupta P. "Comparative analysis of various Machine Learning algorithms for detecting dementia". *Procedia computer science*, *132*, 1497-1502, 2018.

[8]. Wang X, Zhou C, Xu X. "Application of C4. 5 decision tree for scholarship evaluations". *Procedia Computer Science*, *151*, 179-184,2019.

[9]. Mohammed M, Mwambi H, Mboya, IB, Elbashir MK, & Omolo B. "A stacking ensemble deep learning approach to cancer type classification based on TCGA data. *Scientific reports*", 11(1), 1-22, 2021.

[10]. Nusinovici S, Tham YC, Yan MYC , Ting DSW, Li J, Sabanayagam C, Cheng CY. "Logistic regression was as good as Machine Learning for predicting major chronic diseases" *Journal of clinical epidemiology*, *122*, 56-69, 2020.

[11]. Xu F, Pan Z, Xia R. "E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework". *Information Processing & Management*, *57*(5), 102221,2020.

[12]. Wang C, Du J, Chen G, Wang H, Sun L, Xu K, He Z. "QAM classification methods by SVM Machine Learning for improved optical interconnection. " *Optics Communications*, *444*, 1-8,2019.
[**13**]. Tama BA, & Lim S. "Ensemble learning for intrusion detection systems: A systematic mapping study and cross-benchmark evaluation", *Computer Science Review*, 39, 100357, 2021.

[14]. Abro AA, Yimer MA, Bhatti Z. "Identifying the Machine Learning Techniques for Classification of Target Datasets". *Sukkur IBA Journal of Computing and Mathematical Sciences*, *4*(1), 45-52,2020.

[15]. Abro AA, Taşcı E, Ugur A. "A Stacking-based Ensemble Learning Method for Outlier Detection". *Balkan Journal of Electrical and Computer Engineering*, *8*(2), 181-185,2020

[16]. Abro AA. "Vote-Based: Ensemble Approach". *Sakarya University Journal of Science*, *25*(3), 871-879, 2021.

[17]. Mantas CJ, Abellán J, Castellano JG. "Analysis of Credal-C4. 5 for classification in noisy domains. *Expert Systems with Applications". 61*, 314-326, 2016.

[18]. Pavlyshenko B. "Using stacking approaches for machine learning models", I*EEE Second International Conference on Data Stream Mining & Processing*, 255-258, 2018.

[19]. Sikora R. "A modified stacking ensemble machine learning algorithm using genetic algorithms", *In Handbook of research on organizational transformations through big data analytics*, 43-53, 2015.

[20]. Tan Y,Shenoy PP. "A bias-variance based heuristic for constructing a hybrid logistic regression-naïve Bayes model for classification" *International Journal of Approximate Reasoning*, 117, 15-28, 2020.

[21]. Chen S, Webb GI, Liu L, Ma X. "A novel selective naïve Bayes algorithm". *Knowledge-Based Systems*, *192*, 105361, 2020.

[22]. Utkin LV. "An imprecise extension of SVM-based Machine Learning models". *Neurocomputing*, *331*, 18-32, 2019.

[23]. Singh BK, Verma K, Thoke AS. "Investigations on impact of feature normalization techniques on classifier's performance in breast tumor classification". *International Journal of Computer Applications*, *116*(19), 2017.

[24]. Kumar AD, Selvam RP, & Palanisamy V. "Hybrid classification algorithms for predicting student performance", *International Conference on Artificial Intelligence and Smart Systems*, 1074-1079, 2021.

[25]. Zareapoor M, & Shamsolmoali P. "Application of credit card fraud detection: Based on bagging ensemble classifier", *Procedia computer science*, 48(2015), 679-685, 2015..

[26]. Van der Heide EMM, Veerkamp RF, Van Pelt ML, Kamphuis, C, Athanasiadis I, Ducro BJ. "Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle". *Journal of dairy science*, *102*(10), 9409-9421, 2019.

[27]. Chen Y. "Mining of instant messaging data in the Internet of Things based on support vector machine" *Computer Communications*, *154*, 278-287., 2020., 2020.

[28]. Nevill-Manning CG, Holmes G, Witten IH. "The development of Holte's 1R classifier" In *Proceedings 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems,* 239-242,1995.

[29]. Dua D, Graff C. "UCI Machine Learning Repository". http://archive.ics.uci.edu/ml (9.07.2021).

[30]. Engel TA, Charão AS, Kirsch-Pinheiro M, Steffenel LA. "Performance improvement of data mining in Weka through GPU acceleration". Procedia Computer Science, 32, 93-100,2014.

[31]. Abro, A. A., Siddique, W. A., Talpur, M. S. H., Jumani, A. K., & Yaşar, E. "A combined approach of base and meta learners for hybrid system". *Turkish Journal of Engineering*, 7(1), 25-32, 2023.

[32]. Abro, A. A., Khan, A. A., Talpur, M. S. H., & Kayijuka, I. "Machine Learning Classifiers: A Brief Primer". *University of Sindh Journal of Information and Communication Technology, 5(2), 63-68, 2021.*

[33]. Chandio, J. A., Talpur, M. S. H., Abro, A. A., Bux, H., Khokhar, N. U. A. A., Shah, A. A., & Saima, M. "Study Of Customers Perception About Shopping Trend Involving E-Commerce: A Comparative Study". *Turkish Online Journal of Qualitative Inquiry,* 12(8), 5415-5424, 2021.