Nat. Volatiles & Essent. Oils, 2021; 8(5): 4550 – 4559

Ensemble Based Learning Style Identification using VARK

Ch. Sasidhar Rao [1] A. S. Arunachalam [2]

[1]Research Scholar, Department of Computer Science, Vels Institute of Science, Technology & Advanced Studies, Chennai - 600117, sdrteachings@gmail.com

[2]Associate Professor, Department of Computer Science, Vels Institute of Science, Technology & Advanced Studies, Chennai - 600117, arunachalam1976@gmail.com

ABSTRACT

In the current situation the need for e –learning grown larger in both education and training industry. The scope of e-learning paves way to adapt Learning Management System (LMS) an integrated web based learning environment and tool for instructional purpose to provide learning available at any time anywhere to the learner becomes essential tool for learners. Learning is considered as legacy process that differs for every individual. Each individual has their own Learning style in adapting new and concrete information. Each learning style got it own individual method in understanding learners learning style, among these VARK model developed by Fleming is widely accepted to enhance its functionality with the recent technologies. In this work we used ensemble learning a machine learning Meta approach to gain better predictive performance by aggregating the predictions from multiple models. Our main objective of adapting ensemble approach in learning prediction using VARK model has been carried out using classifiers such as J48, SVM, Naive Bayes and Random forest as initial step towards our objective. The bagging ensemble approach has been utilized under hard majority voting to improve more accuracy in learning style identification and to identify various attributes that influence in personalizing LMS. Thus, the efficient personalization of learning management system by understanding learners learning style helps to provide sophisticated Learning Environment to the Learner. It can also be extended in various levels of training in the corporate industry for the effective management of employee training and learning process can be made easier to the learner.

Keywords - Learning Management System, Ensemble learning, Classification Techniques, Learning Styles, Learning Environment.

1. INTRODUCTION

In the current scenario education and training industry understood the need of adapting e-learning platform called Learning Management system as a solution for their future growth. Learning Management system (LMS) is used for e-learning practices that provide instructor to create and deliver content, monitor student participation and to evaluate student performance through a web domain based technology. There are various factors that influence LMS in real time, as Learning Management system tools are user centric, it continuously tries to adapt and customize learning environment based on learner's preferences to provide sophisticated learning experience to learner through personalizing LMS. In this Juncture there arises need for Learning Management system to understand the Learning preferences of the Learner to provide more personalized and customizable Learning Management system. In this scenario the personalizing learning environment through prediction of learning style [7] supports us to provide more personalized Learning Environment. The learning style is viewed as individual perception on acquiring information and converting as knowledge by various experiences from day –to-day life [11]. Each individual have their own way of learning things. Learner style was visualized in Kolb's model (1984) as four Accommodator, Converger, Diverger, and Assimilator. Each learning style have individual approach in understanding learners learning style, among these VARK model developed by Fleming is widely accepted to enhance its functionality with the recent technologies[13]. Thus, the efficient personalization of learning management system by understanding learners learning style helps to provide sophisticated Learning Environment to the Learner.

The rest of the paper is arranged as follows section 2 discusses about the literature review and about roles of techniques utilized in adapting ensemble learning in VARK learning style identification then the Section 3 explains the proposed model for personalizing Learning Management System using ensemble learning, Section 4 Results and Discussion and Section 5 Concludes the motive of work.

2. LITERATURE REVIEW

2.1 Role of Learning Management System.

In recent years researchers understood the vital role of e –learning in the process of learning. The learners feels sophisticated to learn through e-learning tools so that learner can learn anywhere at any time. The advent need of e-learning systems has taken shape into an integrated web based learning environment and tool for instructional purpose called Learning Management systems (LMS). Learning Management systems (LMS) gives way to create and deliver content, monitor student behavior and to identify student performance [25]. LMS consists of various interactive features in learning through discussions, video conferencing and discussion forums. The Learning Management System consists of components (i) Course Management System (CMS), (ii) Learning Content Management System (LCMS), (iii) Managed Learning Environment(MLE),Learning Support Systems and Learning Platform. The LMS tries to adapt and personalize learning environment based on learner's preferences in learning [15].

2.2 Role of VARK learning model:

VARK learning model that was introduced by Neil Fleming suggest four modulator methods for identifying learning style of individuals as Visual Learning, Auditory Learning, Physical Learning, and Social Learning. Daoruang, Beesuda [6] identified the impact of learning style prediction using VARK based on user

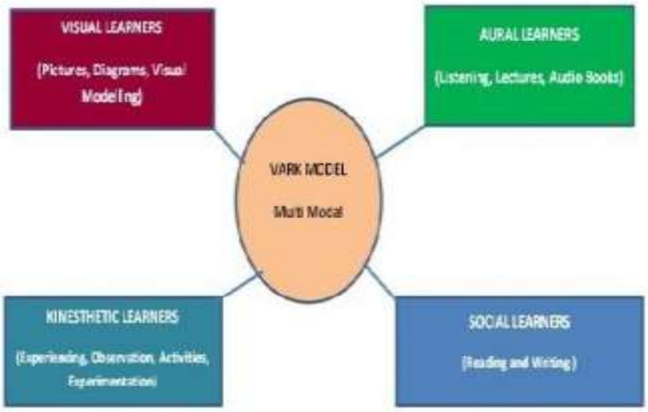


Figure 1: VARK Learning Style Model

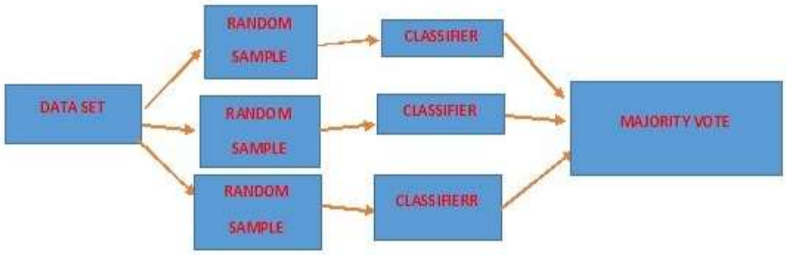
characteristics.

- Visual Learners learn things using real time visual tools such as graphs, charts, diagrams, symbols.
- The Auditory Learners prefer to understand through listening such as lectures, discussions, tapes.
- The Tactile/kinesthetic Learners tries to learn using real time experiencing such as project work.
- The Social Learners prefers to learn using
- social Skills like Reading and Writing

2.3 Role of Ensemble Learning:

The ensemble learning gives us a way to produce projection on data that has weak predictive features by utilizing voting mechanisms for better performance. The ensemble learning works as many variants, but three methods are dominantly used in ensemble learning [14].

2.3.1 Bagging



In 2007, Guohua Liang [15] defines the use of bagging also known as Bootstrap aggregation learning method is used to train data from a diverse group of ensemble members. The bagging ensemble fits the weak learner for each of these samples and finally aggregates their outputs models. For a classification problem, the probabilities of each classes returned by all the models, average these probabilities and keep the class with the highest average probability, this method is termed as **soft-voting** in Majority vote.

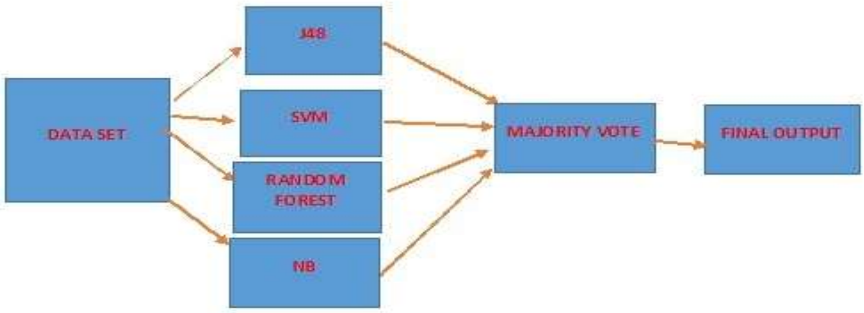


Figure 2: Bagging Ensemble Approach

2.3.2 Stacking

Stacking is an ensemble method using a model to combine predictions from diverse group of members that fits on training data. The stacking ensemble learning technique uses prediction from multiple models to build a new model. Each train set into n parts and fitted in n+ 1 part. The prediction of each train is built as train set. Each base model undergoes the same method and final test model for each classifier. Saso dzeroski et al [16] merits stacking approach has better performance one among various methods.

2.3.3 Boosting

In 1999, Robert E. Schapire [17] sources the advantage of Boosting ensemble that changes the training data based on the previous models from the dataset. The Boosting approach combines the weak learners to form a strong learner as the weak learner classifies data with less correlation among the actual classification.

Gradient Boosting:

Figure 3: Stacking Ensemble Approach

The gradient approach tr[illegible]entially to form a new model that gradually minimizes the loss function using the gradient descent method. The gradient tree Boosting algorithm takes decision trees as the weak learners by selecting the best split among the trees that are not similar and the decision is taken by calculating error rate by the previous tree. Yanka Aleksandrova1 et al[18] evidences to utilize this approach in classifying student log data for improving LMS performance.

AdaBoost:

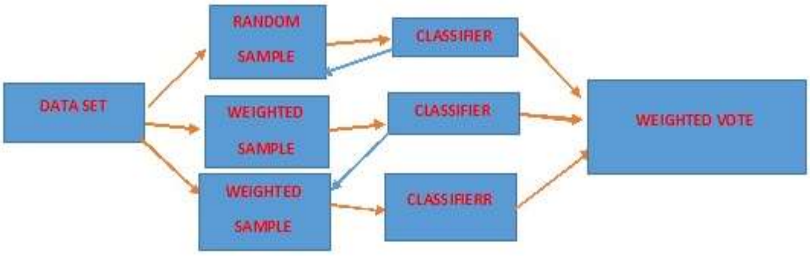


Figure 4: Boosting Ensemble Approach

In the AdaBoost algorithm approach all the weights are re-assigned so that the incorrectly classified models can fit into the sequence of weak learners on different weights. If the prediction is made based on the first learner considered as incorrect then the algorithm allocates the higher importance to the incorrectly predicted statement as an repetitive process, then new learners will be added until the threshold limit is reached in the model.

XGBoost:

Extreme gradient boosting consists of a collection of predictions from multiple models to provide better prediction accuracy. This technique calculates the errors identified by previous models and tries to rectify it through succeeding models by adding some weights to the models. Tianqi Chen et al [20] states the advent use of xg boosting approach in various types of classifying data using machine learning.

2.4 Role of Classification Algorithms

Classification is a classic data mining technique based on machine learning. Classification is used to classify each item in a set of data into one of a predefined set of classes or groups. Classification from large chunks of data provide be our first priority in utilizing data mining for personalizin sed for classification of them are J48 Decision trees, Naïve Bayes, Support Vector machines and Random Forest .

2.4.1 J48-Decision Tree

J48 is a successor of C4.5 decision algorithm and seen as extension of ID3. The algorithm features with missing values, decision trees pruning on derivation of rules and continuous attribute value. The J48 is a predictive learning model that calculates the learning value based on attribute values. The internal node denotes attributes, the branches represents the possible values in observed samples the terminal nodes identifies the final classification value. Amal Alhassan et a l[21] identified the use of machine learning algorithm in his learning style prediction.

2.4.2 Naive-Bayes

The Naive Bayesian works on statistical classification in predicting class members by probabilities belongs to a particular class. The Computational efficiency and simplicity makes the real world applications to widely use Naive Bayes and Bayesian networks for classifying data. It classifies data based on presence or absence of attribute value in the class .The naive Bayes works with small amount of training to identify required knowledge. This classification helps in identifying dissimilarities in Learner's data in LMS Yun-Fu Liu et al[22] estimated the use of naive bayes to improve the performance of Learning management system.

2.4.3 Support Vector Machine (SVM)

SVM is used for knowledge discovery through classification, regression and outlier's detection Ashish Dutt [24] considered SVM advantageous and efficient in its high dimensional spaces, efficient memory management but if the sample sizes become greater the performance becomes poor SVM will not support for probability estimates.

2.4.4 Random forest

Random forest uses bagging ensemble method by using multiple decision trees and separates classes into a class or a group .The random forest selection of the predictor variables to identify less correlation among the trees that has low error rate. The class with the maximum number of votes is identified as new instance. Dejan Ljubobratović [23] identifies the use of Random forest algorithm as high among various classifiers in prediction of learning style of the learner.

3. Proposed work

In this work we utilized ensemble learning approach to improve accuracy in learning style prediction using classification algorithm and to identify attribute features that influence in personalizing LMS. The treatment of dataset under classification gives us insight to analyze the learning behavior of learners that affects personalization of LMS. Aldahwan, N.[1] highlights the need and advent use of artificial intelligence in personalizing LMS and also suggested various methods that paves the way to utilize machine learning techniques in personalizing LMS. The main objective is to improve accuracy in learning style prediction using ensemble learning approach, for adapting ensemble learning approach we devised an intermediate module named as Intelligent Learning Style Identifier The recommendation of intelligent learning style identifier has sourced by many researchers namely Beesuda Daourang [1] recommends an adaptive expert module for learning style prediction using VARK to enhance personal learning, Bindhia. K .Francis et al [9] recommends hybrid data mining model for improving prediction accuracy, Christos Trousaas [3] et al utilized ensemble learning to improve accuracy in learning style identification using FSLSM learning model. Ensemble learning a Meta approach to machine learning that seeks better predictive performance by combining the predictions from multiple models. Our work utilized ensemble classification method among classifiers J48 Decision trees, Naïve Bayes, Support Vector machines and Random Forest are combined based on the majority voting method that yields more accuracy in learning style identification using ensemble learning to improve accuracy in learning style identification that acts as source for personalization of LMS . The efficient prediction of learning style affects in personalizing Learning Management System (LMS) [15].

3.1 Improving classification accuracy in learning style identification Using VARK

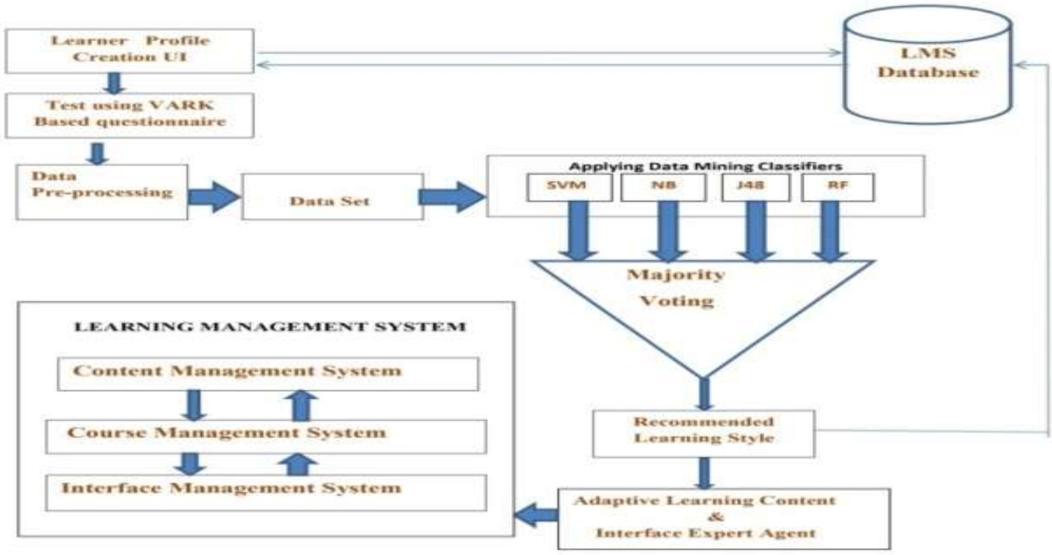
The recommendation of learning style identifier has sourced by many researchers namely Beesuda Daourang [1] recommends an adaptive expert module for learning style prediction using VARK to enhance personal learning, Hasibuan [3] adapted prior knowledge using VARK for obtaining more accuracy in learning style prediction . Arunachalam, A. Set al [9] recommends hybrid data mining model for improving prediction accuracy, Christos Trousaas [4] et al utilized ensemble learning to improve accuracy in learning style identification using FSLSM learning model. Da Silva [2] utilized ensemble approach with various classifiers on sentimental analysis. Rasheed, Fareeha [7] visualizes learning style identification in e-learning system helps in personalizing LMS environment. Hasibuan [10] focuses on improving accuracy in learning style prediction using prior knowledge .Thus Ensemble learning a Meta approach to machine learning that seeks better predictive performance by combining the predictions from multiple models. The classifier algorithms were selected after detailed literature review and from the results from our initial work [13] in learning style

Figure 5: Proposed model for LMS personalization Using Ensemble Learning

identification. Our work utilized ensemble classification method among three classifiers namely J48, NB,SVM and Random Forest are combined based on the majority voting method that yields more accuracy in learning style identification Ensemble learning to improve accuracy in learning style identification using emersions learning.

4. Results and Discussions

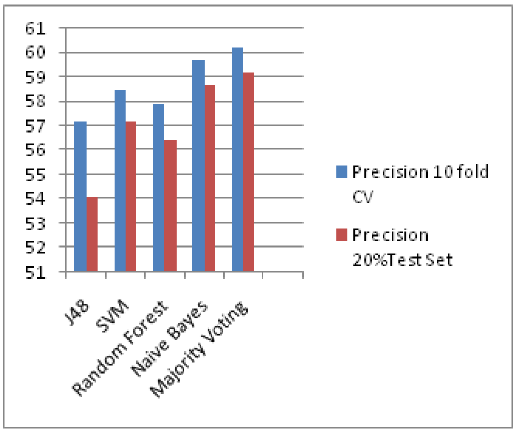
In this work our motive to adapt ensemble classification techniques in predicting VARK learning style give us insight to utilize classification algorithms J48, NB, SVM and Random Forest classifiers as an initial query. In our approach bagging method of ensemble for classification with hard majority voting for more accurate learning style identification as suggested Han,Bo[11] in classification using stacking. In this process after various levels of data pre-processing, we used heterogeneous classifiers to build new model using data mining Tool WEKA for investigating the utilization of ensemble approach in VARK learning style identification. Each individual classifier (J48, NB, SVM and Random Forest) trained to build the output model, and then the individual classifier is used to determine the final output under majority voting. The initial phase work we utilized bagging ensemble using Majority vote. In this approach dataset is loaded into each individual classifier and the performance is evaluated in terms of accuracy and precision is calculated separately, then majority

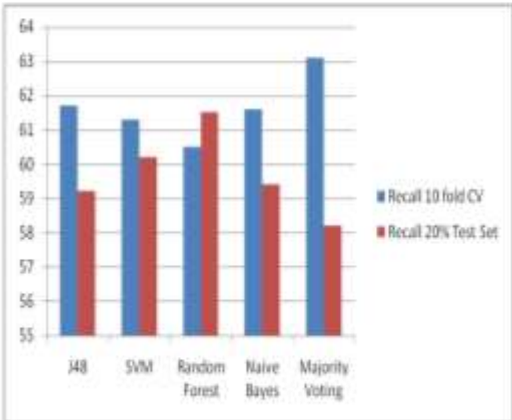


voting ensemble is build to find the usefulness of ensemble approach in classifying learning style.

The class label y^ is calculated using majority voting of each classifier Cj:

y^=mode{C1(x),C2(x),...,Cm(x)}





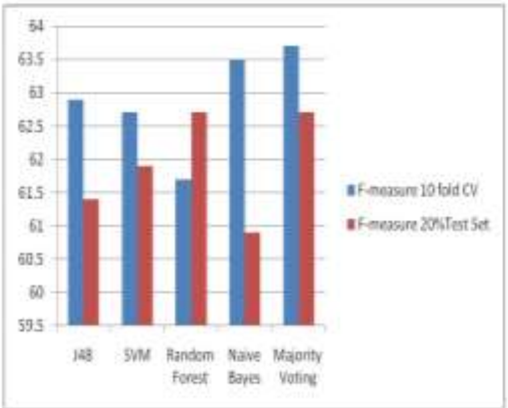
The training sample utilizes three samples for classification:
Classifier algorithm 1 -> class 0
Classifier algorithm 2 -> class 0,
Classifier algorithm 3 -> class 1
Y^=mode {0, 0, 1} =0, the majority voting algorithm will classify the result as "class 0."

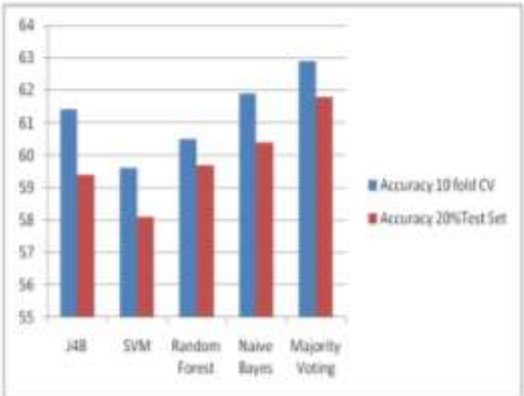
Table1: Results output of Classifier Algorithms

Experiment	Algorithm							Accuracy%	
		10 fold CV	20% Test Set	10 fold CV	20% Test Set	10 fold CV	20% Test Set	10 fold CV	20% Test Set
1	J48	57.2	54.1	61.7	59.2	62.9	61.4	61.4	59.4
2	SVM	58.5	57.2	61.3	60.2	62.7	61.9	59.6	58.1
3	Random Forest	57.9	56.4	60.5	61.5	61.7	62.7	60.5	59.7
4	Naive Bayes	59.7	58.7	61.6	59.4	62.5	61.9	61.9	60.4
5	Majority Voting	60.2	59.2	63.1	58.2	63.6	62.4	62.9	61.8

The above results in Fig 6 interpret the motive of work in utilizing ensemble approach in learning style identification using VARK learning model. The work utilized heterogeneous classifiers J48, Naive Bayes, Support Vector Machines and Random Forest

Figure 7a: Precision attained in classification using Ensemble Technique





The results are validated in terms of Precision, Recall, F-measure and Accuracy performed by classification algorithms and Majority voting approach. The above results in Fig 7a, 7b, 7c and 7d shows the performance evaluations of classifiers in terms of Precision, Recall, F-measure and Accuracy.

The J48 decision tree classifier algorithm produced 57.2 % precision level and 61.4% accuracy with 10 fold cross validation and 54.1 % precision level and 59.4 % accuracy in 20% Test set. The J48 attained 61.7 % Recall value and 62.9 % F-measure value with 10 fold cross validation and 59.2 % Recall value and 61.4% F-measure value with 20% Test set.

The SVM classifiers produced 58.5 % precision level and 59.6% accuracy with 10 fold cross validation and 57.2 % precision level and 58.1 % accuracy in 20% Test set. The SVM attained 61.3% Recall value and 62.7 % F-measure value with 10 fold cross validation and 60.2% Recall value and 61.9 % F-measure value with 20% Test set.

The Random Forest produced 57.9 % precision level and 60.5% accuracy with 10 fold cross validation and 56.4% precision level and 59.7% accuracy in 20% Test set. The Random Forest attained 60.5 % Recall value and 61.7 % F-measure value with 10 fold cross validation and **61.5** % Recall value and 62.7 % F-measure value with 20% Test set.

The Naïve Bayes produced 59.7% precision level and 62.3% accuracy with 10 fold cross validation and % precision level and 61.2% accuracy in 20% Test set. The Naive Bayes attained 61.6% Recall value and 62.5% F-measure value with 10 fold cross validation and 59.4 % Recall value and 62.5 % F-measure value with 20% Test set.

The Hard Majority voting produced 60.2% precision level and 62.9% accuracy with 10 fold cross validation and 59.2% precision level and 60.4% accuracy in 20% Test set. The Hard Majority voting attained 63.1% Recall value and 63.6 % F-measure value with 10 fold cross validation and 58.2% Recall value and 62.4 % F-measure value with 20% Test set.

The graphical representation and the discussion witnesses the work carried produced promising results to claim ensemble learning using majority voting technique improves the classification of learning style than individual classifiers, . This work gives us insight towards our motive to improve accuracy in learning style identification that helps to personalize Learning Management Systems (LMS).

5. Conclusion

In this paper, our motive towards personalizing Learning Management System by efficient learning style prediction has been taken the phase in utilizing ensemble approach in learning style prediction using VARK. This work gives us insight to our future approach to build hybrid data mining model towards personalizing LMS using efficient prediction in learning style. Initially, the work utilized bagging ensemble with majority voting for improving efficiency using classification algorithms. The work carried out shows promising results to claim ensemble learning technique utilized improves the classification of learning style than individual classifiers The solution can also be extended by adapting hybrid data model utilizing efficient ensemble technique to improve the performance in learning style prediction that helps to personalize Learning Management Systems (LMS) in future.

References

1. Aldahwan, N., and N. Alsaeed. "Use of artificial intelligent in Learning Management System (LMS): a systematic literature review." International Journal of Computer Applications 175, no. 13 (2020): 16-26.
2. Da Silva, Nadia FF, Eduardo R. Hruschka, and Estevam R. Hruschka Jr. "Tweet sentiment analysis with classifier ensembles." Decision Support Systems 66 (2014): 170-179.
3. Hasibuan, Muhammad Said, Lukito Edi Nugroho, and Paulus Insap Santosa. "Model Detecting Learning Styles with Artificial Neural Network." Journal of Technology and Science Education 9, no. 1 (2019): 85-95.
4. Troussas, Christos, Akrivi Krouska, Cleo Sgouropoulou, and Ioannis Voyiatzis. "Ensemble learning using fuzzy weights to improve learning style identification for adapted instructional routines." Entropy 22, no. 7 (2020): 735.
5. Daoruang, Beesuda, Krich Sintanakul, and Anirach Mingkhwan. "The study of learning achievement of learners classified VARK learning style in blended learning." In Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval, pp. 34-38. 2019.
6. Daoruang, Beesuda, Krich Sintanakul, and Anirach Mingkhwan. "The study of learning achievement of learners classified VARK learning style in blended learning." In Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval, pp. 34-38. 2019.
7. Rasheed, Fareeha, and Abdul Wahid. "Learning style detection in E-learning systems using machine learning techniques." Expert Systems with Applications 174 (2021): 114774.
8. Singh, Asmita, and R. Lakshmiganthan. "Impact of different data types on classifier performance of random forest, naive bayes, and k-nearest neighbors algorithms." (2018).

9. Sarojini, T. K., and R. Gobinath. "Admiration of internal and external factors affecting teachers implementation of seminal evaluation to support learning." International Journal of Pure and Applied Mathematics 119.10: 209-213, 2018.
10. Arunachalam, A. S., and T. Velmurugan. "Measures for predicting success factors of ELearning in Educational institutions." International Journal of Pure and Applied Mathematics 118.18 (2018): 3673-3679.
11. Han, Bo, Paul Cook, and Timothy Baldwin. "A stacking-based approach to twitter user geolocation prediction." In Proceedings of the 51st annual meeting of the association for computational linguistics: system demonstrations, pp. 7-12. 2013.
12. Marakas, George M. Decision support systems in the 21st century. Vol. 134. Upper Saddle River, NJ: Prentice Hall, 2003.
13. Sasidhar Rao, Ch and Arunachalam, A.S "Personalization of Learning Management System using VARK." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12, no. 7 (2021): 1538-1545.
14. Dong, Xibin, Zhiwen Yu, Wenming Cao, Yifan Shi, and Qianli Ma. "A survey on ensemble learning." *Frontiers of Computer Science* 14, no. 2 (2020): 241-258.
15. Liang, Guohua, Xingquan Zhu, and Chengqi Zhang. "An empirical study of bagging predictors for different learning algorithms." In *Twenty-Fifth AAAI Conference on Artificial Intelligence*. 2011.
16. Džeroski, Saso, and Bernard Ženko. "Is combining classifiers with stacking better than selecting the best one?." *Machine learning* 54, no. 3 (2004): 255-273.
17. Schapire, Robert E. "A brief introduction to boosting." In *Ijcai*, vol. 99, pp. 1401-1406. 1999.
18. Aleksandrova, Yanka. "Predicting Students Performance in Moodle Platforms Using Machine Learning Algorithms." In *Conferences of the department Informatics*, no. 1, pp. 177-187. Publishing house Science and Economics Varna, 2019.
19. Kondo, Nobuhiko, Midori Okubo, and Toshiharu Hatanaka. "Early detection of at-risk students using machine learning based on LMS log data." In *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, pp. 198-201. IEEE, 2017.
20. Chen, Tianqi, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, and Hyunsu Cho. "Xgboost: extreme gradient boosting." *R package version 0.4-2* 1, no. 4 (2015): 1-4.
21. Alhassan, Amal, Bassam Zafar, and Ahmed Mueen. "Predict Students' Academic Performance based on their Assessment Grades and Online Activity Data." *International Journal of Advanced Computer Science and Applications (IJACSA)* 11, no. 4 (2020).
22. Liu, Yun-Fu, Jing-Ming Guo, and Jiann-Der Lee. "Halftone image classification using LMS algorithm and naive Bayes." *IEEE Transactions on Image Processing* 20, no. 10 (2011): 2837-2847.
23. Ljubobratović, Dejan, and Maja Matetić. "Using LMS Activity Logs to Predict Student Failure with Random Forest Algorithm." *The Future of Information Sciences* (2019): 113.
24. Dutt, Ashish, and Maizatul Akmar Ismail. "Can we predict student learning performance from LMS data? A classification approach." In *3rd International Conference on Current Issues in Education (ICCIE 2018)*, pp. 24-29. Atlantis Press, 2019.
25. Sarojini, T. K., and R. Gobinath. "Admiration of internal and external factors affecting teachers implementation of seminal evaluation to support learning." International Journal of Pure and Applied Mathematics 119.10: 209-213, 2018. W.Hong (2017) "A prediction and Early Alert Model using Learning Management System Data and Grouping in Learning Science Theory"ICEDM.