Predicting students' learning styles using regression techniques

Ahmad Mousa Altamimi[1], Mohammad Azzeh[2], Mahmoud Albashayreh[1]
[1]Department of Computer Science, Applied Science Private University, Amman, Jordan
[2]Department of Data Science, Princess Sumaya University for Technology, Amman, Jordan

Article Info

ABSTRACT

Traditional learning systems have responded quickly to the COVID pandemic and moved to online or distance learning. Online learning requires a personalization method because the interaction between learners and instructors is minimal, and learners have a specific learning method that works best for them. One of the personalization methods is detecting the learners' learning style. To detect learning styles, several works have been proposed using classification techniques. However, the current detection models become ineffective when learners have no dominant style or a mix of learning styles. Thus, the objective of this study is twofold. Firstly, constructing a prediction model based on regression analysis provides a probabilistic approach for inferring the preferred learning style. Secondly, comparing regression models and classification models for detecting learning style. To ground our conceptual model, a set of machine learning algorithms have been implemented based on a dataset collected from a sample of 72 students using visual, auditory, reading/writing, and kinesthetic (VARK's) inventory questionnaire. Results show that regression techniques are more accurate and representative for real-world scenarios than classification algorithms, where students might have multiple learning styles but with different probabilities. We believe that this research will help educational institutes to engage learning styles in the teaching process.

Corresponding Author:

Ahmad Mousa Altamimi
Department of Computer Science, Applied Science Private University
21 Al Arab st., Amman 11931, Jordan
Email: a_altamimi@asu.edu.jo

1. INTRODUCTION

Due to the COVID pandemic, the largest disruption to education in history has been recorded, which has had a nearly universal impact on learners and educators worldwide [1]. As a result, learning systems have successfully undergone critical changes and used new models (i.e., online learning or distance learning) supported by information and communication technologies [2]. Online learning enriches conventional learning by offering flexibility and self-paced learning with an efficient way to deliver knowledge through virtual communication and collaboration [3]. However, it requires a personalization method as learners have different backgrounds, knowledge, and various learning environments [4]. One of the personalization methods is detecting the learners' learning style as it influences individual academic achievement [5]. The concept of learning style was coined during the mid of 70's and is formally defined as: "an individual's mode of gaining knowledge" [6]. It is the best method a person uses to learn.

Many learning styles' theories have been introduced in the field of education and widespread recognition in education theory and learning strategies [7]. These theories include ones that classify people according to their own distinguishing features that differentiate one from others. The widespread theories are motivated by the fact that knowing a learner's learning style can enable instructors to maximize learners'

learning by using adapted teaching methods and allowing them to recognize their learning styles to find what study methods and activities help them learn best [8]. Thus, the awareness of the learning styles roles in the education process is very important for both learners and researchers [9].

Inventories are used to recognize individuals' learning styles, typically take the form of a questionnaire assessment, where a series of questions are asked and then scored the results to illustrate the dominant learning styles. There are many popular learning style inventories proposed in the literature, such as fleming's visual, auditory, reading/writing, and kinesthetic (VARK) learning style questionnaire [10], Kolb's learning style inventory (LSI) [11], Jackson's learning styles profiler (LSP) [12], and other. Each of these proposed a set of questions to identify the learners' different styles. For example, according to VARK, learners are categorized into four different types: visual, auditory, reading/writing, and kinesthetic [10]. On the other hand, Kolb's is also one of the widely used inventories identifying four learning styles [11].

More recently, considerable research has been devoted to automatically detecting the learning style [13]-[15]. In fact, educational data mining is the leading approach concerned with applying machine learning to the collected information from educational settings. Here, both classification and clustering algorithms have been applied. While the classification technique is applied for discrete variables, the regression technique is applied for continuous variables. Classification algorithms were the dominate into two approaches: clustering and classification. For example, Aissaoui et al. [15] utilized the K-modes clustering algorithm to improve the e-learning system. The model was implemented based on the Felder and Silverman learning style model using a dataset extracted from an e-learning system's log file.

Other classification algorithms have also been used. The decision tree (DT) was used in [16] to detect the learners' learning styles from students' weblogs. Pantho [17] also used the decision tree C4.5 algorithm to identify the learning styles. Here, the sample was collected from 1,205 students using the VARK questionnaire. Other algorithms have also been utilized. The neural networks was employed in [18], where Felder-Silverman's model was used to identify four dimensions of learning styles. These dimensions are sensing or intuitive, active or reflective, visual or verbal, and sequential or global. Felder-Silverman's model was also used in [14], but the fuzzy C-means was employed as a clustering algorithm to detect learners' styles based on their data stored in the log files.

On the other hand, Genetic algorithms were employed to describe learning styles. Yannibelli et al. [19] define a group of chromosomes and assign the learner's action to each gene. Then used these genes generate new populations of chromosomes that describe learning styles. In the same vein, the work of [20] classified learners based on their learning styles by combining genetic algorithms with k-nearest neighbors (K-NN). In this work, the learners' behaviours are represented in an n-dimensional space. Learners are then considered to have the same learning style if they have a shorter distance to others. Lwande et al. [21] combined both felder-silverman learning style model and cognitive trait model to estimate learning styles from learning management system (LMS). Results showed a possible estimation for the learning styles. Another study [22] have conducted that used different machine learning models to predict learning outcome. The study read records from e-learning platform to get the relevant features.

Recently, educational data mining has been extensively considered in the literature. The educational data mining community defines it as an emerging discipline concerned with developing methods for exploring unique educational data types to understand students' learning settings better [23]. The spreading of educational data mining is due to the emergence of numerous public data mining tools such as R, waikato environment for knowledge analysis (WEKA), RapidMiner, and konstanz information miner (KNIME) [24]. Wahbeh et al. [25] demonstrated a comparison between these tools, and it concludes that each of these tools has its advantages and disadvantages. Educational data mining was also used to predict students' performance using classification and regression techniques [26]. While the classification technique is applied for discrete variables, the regression technique is applied for continuous variables [27]. Lincke et al. [22] employed the artificial neural network with a sample of 316 undergraduate students to predict academic performance. Results showed that students' performance in the course is improved when considering their learning styles. Aissaoui et al. [28] utilized multiple linear regression (MLR) to build a student' performance prediction model. The obtained results show that the model outperforms the other constructed models.

As one can be noticed, almost all of the proposed works were designed to detect the learners' learning styles and identify a single learning style for each learner. However, in practice, learners might have a single or multiple learning styles, where one can equally prefer both visual and auditory learning styles. For learners with a mix of learning styles (with probability) or with no dominant style of learning, detecting a single learning style is becoming ineffective. This was supported by Azzi et al. [29], where the researchers proved that learners have different learning styles, and thus, no single system can serve well with all learners. Therefore, the current approaches do not support this trend, and thus a new learning style detection system has been proposed to solve this issue based on regression analysis. To the best of our knowledge, no work

has provided a probabilistic approach to infer the learner's style. Hence, this work utilizes regression analysis to provide a probabilistic approach for inferring the preferred learning styles.

To this end, a dataset was collected using the VARK's inventory questionnaire from a sample of 72 students from applied science university. To easily collect the students' responses, we developed an online version of the questionnaire of 16 different questions using Microsoft Forms (part of Office 365). Here, we divided the whole dataset into an array of four matrices (A, V, K, R), a matrix for each learning style. Each matrix consists of 16 columns demonstrating the presence or absence of a learning style and 5 output columns. 4 columns of them represent the learning styles' probabilities, which served as output for regression models, and the last column represents the selected learning style label which served as output for classification models.

After preparing the dataset, multiple prediction models are developed using five machine learning algorithms (multi-layers perceptron neural network (NN), support vector machine (SVM), decision tree (DT), random forest (RF), and K-NN). The constructed model attempts to predict the probability of each learning style to identify the most favoured styles. In this case, the output of prediction would be in this format: <A=0.3, V=0.22, K=0.08, R=0.4>. Then a threshold can be specified to select the most favoured learning styles. To accomplish that, we compute the distance between the top learning style and the remaining learning styles. Any learning style that falls within the distance given by the threshold is selected as the nominated learning style. In the example above, if the threshold equals 0.2, then the selected learning style is {R, A, V}. We recommend the threshold value to be not very small, ignoring some interesting learning styles or too large that involve all learning styles.

The remainder of this paper is organized as shown in; The research methodology is given in section 2. The experimental work, along with the evaluation measures, and the discussion about our results, is presented in section 3. Finally, section 4 presents the conclusion and directions for future research.

2.   RESEARCH METHOD

The methodology employed in this study requires a clear understanding of the tradeoffs inherent in this domain. In fact, the key challenge is to classify learners according to their distinguishing features (learning styles), taking into account that learners may have a mix of learning styles with the probability of having no dominant learning style. As such, our general approach builds upon the regression analysis to provide a probabilistic approach for inferring the preferred learning styles. Because of this domain's maturity in general, it is important to ground our approach with a robust experimental evaluation. To this end, we constructed multiple prediction models using five machine learning algorithms. The dataset used in these experimental tests was collected using the VARK's inventory questionnaire from a sample of 72 students. We develop classification models for inferring the learning style label using the same machine learning algorithms to demonstrate our approach. The models are evaluated using recall, precision, accuracy, F1-score, and area under curve (AUC). Based on the obtained results, one can conclude that regression algorithms are accurate and representative for predicting learning styles' probabilities

2.1. Data collection

To conduct our study, a sample of 72 students was randomly selected from Applied Science University. The sample data was collected using VARK's inventory questionnaire, where four different learning styles are identified: visual (V), auditory (A), reading/writing (R), and kinesthetic (K). The questionnaire consists of 16 different questions that deal with the way(s) in which students like to learn or prefer to deliver. The questions are based on situations where there are choices and decisions about how those might happen. To easily collect the students' responses, we developed an online version using Microsoft Forms. The responses are then imported as an Excel file where each answer is represented as a vector of binary values denoted as <A, V, K, R>. The data is then preprocessed to be eligible for the employed machine learning algorithms. Here, we divided the whole dataset into an array of four matrices, a matrix for each learning style. Each matrix consists of 16 columns demonstrating the presence or absence of a learning style and 5 output columns (4 columns represent the learning styles' probabilities, and the last column represents the selected learning style label).

Figure 1 shows the probability distribution for each learning style using boxplots. The horizontal line inside the boxplot represents the median of probabilities, while the x symbol inside each box represents the mean of probabilities. We can observe that all learning styles have a relatively similar distribution. For instance, learning style (A) has the largest mean and median values with narrower distribution than other learning styles, suggesting that all students have the same learning style probabilities (A). Also, we can notice that most students favored learning style (A). Based on visual observation, we can see that the mean of probabilities for learning style (A) is significantly different from that of (V) and (R) learning styles.
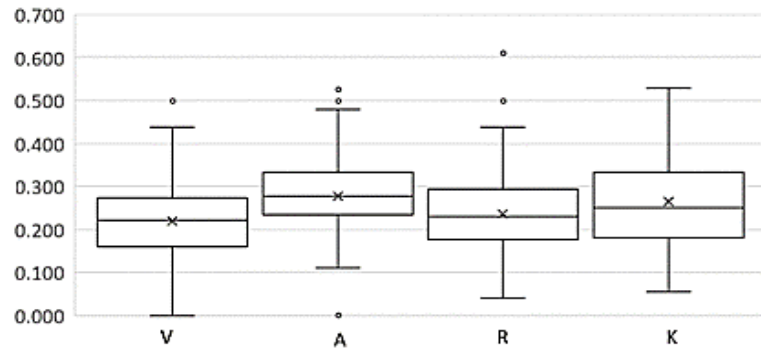
Figure 1. Boxplots of probabilities for all learning styles

2.2. Data preprocessing

The collected dataset has been first preprocessed to be eligible for the used machine learning algorithms. To this end, the original dataset is described by multiple rows and columns, where each row represents student responses, and the columns represent questions. Each response consists of a list of one or more styles selected from the complete response labels {A, V, K, R}. In other words, each student might provide multiple styles for the response to each question. We represent each answer as a vector of binary values denoted as <A, V, K, R> to facilitate processing the data. For example, the vector <1, 0, 1, 0> means that the student responded with A and K learning styles. The processed data has been manipulated to look like the dataset shown in Table 1, where the last five columns are considered output. Four of them are considered numeric, denoted by 'Prob of', the probability of each learning style that will serve as output for regression models. In contrast, the last column is the selected learning style label, given based on the maximal probability, which served as output for classification models.

Table 1. Processed dataset

ID	Q1	Q2	Q3	…	Q16	Prob of A	Prob of V	Prob of K	Prob of R	Learning Style
S1	<1,1,0,1>	<1,0,1,1>	<1,0,0,0>	…	<0,0,1,1>	0.36	0.24	0.25	0.15	A
S2	<0,0,0,1>	<0,1,0,1>	<1,0,0,1>	…	<0,1,0,0>	0.18	0.44	0.21	0.17	V
S3	<0,0,0,1>	<0,1,0,0>	<0,1,1,1>	…	<1,1,0,0>	0.14	0.25	0.19	0.42	R
				…						
N	<0,1,0,0>	<0,0,0,1>	<0,0,0,1>	…	<0,1,0,1>	0.42	0.49	0.07	0.02	V

Since each cell contains a binary vector representing student response, we divided the whole dataset into an array of four matrices, a matrix for each learning style. Specifically, each matrix has the same number of rows and columns, but each question cell represents the corresponding value for that learning style. For example, the matrix for learning style R will look as shown in Table 2. All matrices share the same set of numeric and label outputs.

Table 2. Learning style (R) matrix

ID	Q1	Q2	Q3	……	Q16	Prob of A	Prob of V	Prob of K	Prob of R	Learning Style
S1	1	1	0	……	1	0.36	0.24	0.25	0.15	A
S2	1	1	1	……	0	0.18	0.44	0.21	0.17	V
S3	1	0	1	……	0	0.14	0.25	0.19	0.42	R
				……						
N	0	1	1	……	1	0.42	0.49	0.07	0.02	V

2.3. Research models

In this research, four machine learning algorithms have been used to build classification models to predict the learning style. These algorithms are DT, SVM, multi-layer perceptron NN, and k-NN. The NN is a feed-forward neural network algorithm with one input layer, at least one hidden layer, and one output layer. Each neuron of the input layer represents an input vector. The NN uses a nonlinear activation function in the neurons of the hidden layer. In contrast, a linear activation function is usually used in the output layer. The

number of neurons in the output layer depends on the problem type. If the problem type is classification, as in our case, then the number of neurons equals several labels, and the output is the probability of each label. Each output neuron represents a class label, and the one with a significant probability is chosen as output. The number of neurons in the hidden layer varies based on the number of input neurons and the type of training algorithm used. The standard training algorithms are the backpropagation algorithm and conjugate gradient algorithm. In this study, we used the backpropagation algorithm because of its advantages over the conjugate gradient algorithm. The number of neurons for each layer has been carefully chosen after multiple trials. The number of input neurons is four which equals the number of input features, the number of hidden neurons in the hidden layer is ten, and finally, the number of output neurons is 2. The activation function used in this research is the sigmoid function.

The K-NN uses the notion of retrieving by similarity and voting to classify data. The K-NN retrieves the closest k similar cases for the new one; then voting is applied to derive the final output. Choosing the value of k has a significant effect on the accuracy of K-NN; for instance, if we choose small k, other valuable cases might be ignored, thus reducing accuracy, whereas the enormous value of k is time and resource-consuming. There are several ways to choose the appropriate value of k; for instance, the most common way is to calculate the square root of the total number of data points. In this paper, we choose k=5 because it is a reasonable value that allows us to select the best closest cases without affecting the cost of the resources.

SVM is used to build an optimal hyperplane that can separate data with maximum margin. The margin is defined as the maximal width of the slab parallel to the hyperplane with no interior data points. The optimal hyperplane generation depends on kernel functions such as Gaussian, polynomial, and radial basis function. Both Gaussian and radial basis function kernels can benefit hyperplane generation because they support the locality of training data, which means that the data can be efficiently separated. In this study, we used a radial basis kernel.

To build the regression model, five machine learning algorithms have been used, which are multi-layers perceptron NN, SVM, DT, RF, and k-NN. The probabilities are used as output, where results are aggregated for each learning style as a prediction. Each matrix has applied these algorithms to predict each learning style's probability as a regression problem. We record the mean of absolute errors (MAE), median of absolute errors (MdAE), root mean of squared errors (RMSE). Then, we aggregate MAE, MdAE, and RMSE using the average aggregation method for each learning style probability.

The constructed model attempts to predict the probability of each learning style to identify the most favoured styles. In this case the output of prediction would be in this format: <A=0.3, V=0.22, K=0.08, R=0.4>. Then a threshold can be specified to select the most favoured learning styles. To accomplish that, we compute the distance between the top learning style and the remaining learning styles. Any learning style that falls within the distance given by the threshold is selected as the nominated learning style. In the example above, if the threshold equals 0.2, then the selected learning style is {R, A, V}. If the threshold is 0.1, then the selected learning style set is {R, A}. We recommend that the threshold value be not very small, ignoring some important learning styles or too large involving all learning styles.

3.   RESULTS AND DISCUSSION

For the regression task, we used the probabilities of each learning style as output. As mentioned in the methodology section, we have constructed an array of four binary matrices, where each matrix represents the presence or absence of a learning style, and the output is the probabilities of all learning styles. We applied five popular machine learning algorithms for each matrix as regression to predict each learning style's probability: NN, SVM, DT, RF, and K-NN. Then we aggregate results for each learning style from the four matrices. The predicted probabilities are compared to the actual probabilities using four performance metrics, as shown in Table 3. The values with boldface and underline represent the most accurate results. This Table shows that RF is almost the superior one for predicting all learning styles across all performance metrics. For learning style (A), we found that RF and SVM work well for predicting the probabilities of (A) for all students. For learning (V) and (R), we found that RF is the dominant model across all performance metrics, as shown in Figure 2. From the obtained results, we found that using regression algorithms to predict the learning styles probabilities are more accurate and representative for real-world scenarios where students might choose multiple learning styles but with different probabilities. Moreover, to examine the effectiveness of our approach, we used the same set of machine learning algorithms to develop classification models for predicting the learning style label. We aimed to compare our findings with the classification-based approach and show which served better.

Table 3. Accuracy results of all regression models

	RMSE	MdAE	MAE	Metric
A	0.1102	0.0671	0.0849	NN
	0.0835	0.0405	0.0614	SVM
	0.0864	0.0405	0.0640	kNN
	0.0853	0.0476	0.0653	DT
	0.0843	0.0451	0.0612	RF
V	0.1184	0.0882	0.0962	NN
	0.0786	0.0523	0.0622	SVM
	0.0897	0.0573	0.0713	kNN
	0.0816	0.0580	0.0653	DT
	0.0773	0.0520	0.0614	RF
K	0.1115	0.0812	0.0923	NN
	0.0930	0.0615	0.0737	SVM
	0.1043	0.0608	0.0793	kNN
	0.0941	0.0564	0.0731	DT
	0.0912	0.0594	0.0730	RF
R	0.1038	0.0732	0.0831	NN
	0.0882	0.0560	0.0688	SVM
	0.0909	0.0595	0.0701	kNN
	0.0936	0.0541	0.0728	DT
	0.0867	0.0561	0.0685	RF
All	0.0762	0.0686	0.0713	NN
	0.0594	0.0519	0.0532	SVM
	0.0631	0.0497	0.0569	kNN
	0.0602	0.0517	0.0553	DT
	0.0580	0.0507	0.0528	RF
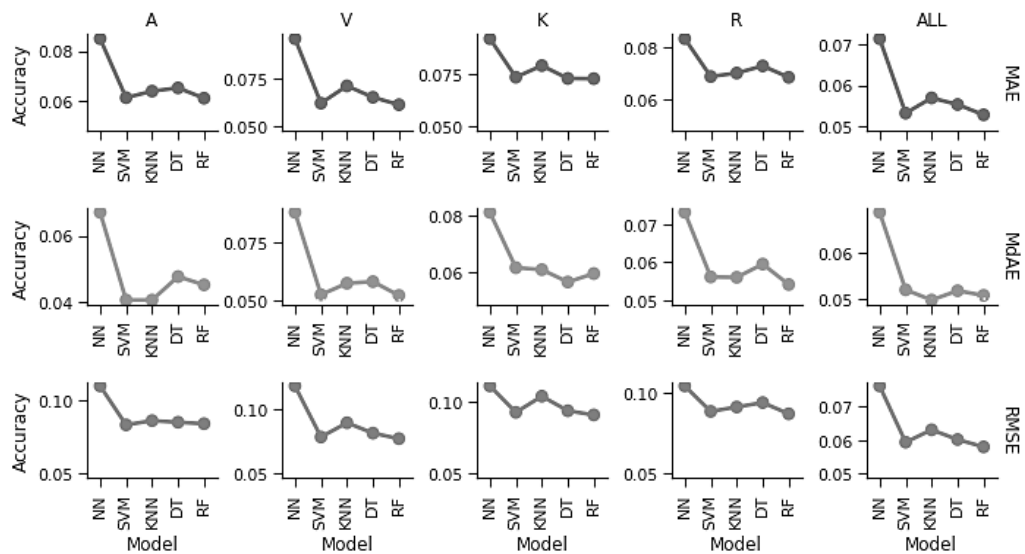
Figure 2. Accuracy results for all prediction models

The Wilcoxon significance tests between each pair of models based on absolute errors over each learning style are presented in Table 4. The results show that predictions produced by NN are almost different from those generated by other prediction models. However, the accuracy of NN in Table 4 was significantly different from other models but not necessarily superior. On the other hand, we did not have any significant difference between each pair of models, which means that all prediction models, except NN, produce relatively similar predictions. From these results, we can conclude that NN is the only model that can generate different predictions than others, while all remaining models behave similarly overall learning styles. This is because we have used a small sample size. As future work, we are planning to collect more student's responses and conduct more critical analyses. Interval plots confirm these results in Figure 3.

Table 4. Wilcoxon statistical significant test of absolute residuals between each pair of models

Model 1	Model 2	A	V	K	R	All
NN	SVM	0.036	0.002	0.06	0.02	0.003
NN	kNN	0.061	0.036	0.1	0.032	0.006
NN	DT	0.13	0.01	0.04	0.02	0.006
NN	RF	0.036	0.0034	0.05	0.02	0.0001
SVM	kNN	0.82	0.41	0.93	0.9	0.76
SVM	DT	0.54	0.71	0.84	0.96	0.80
SVM	RF	0.84	0.97	1.00	0.99	0.99
kNN	DT	0.66	0.6	0.74	0.85	0.89
kNN	RF	0.71	0.41	0.94	0.89	0.7
DT	RF	0.45	0.66	0.83	0.96	0.76

Figure 3 shows the interval plots of absolute errors for all prediction models over each learning style. It is clear that NN produces significantly different predictions than other models of overall learning styles, as shown in Table 4. Despite that, the NN produces bad results than other models. This is perhaps because of the small dataset used in this study. The remaining models behave similarly with no significant differences between their predictions, which means that any one of them can perform the job. Nevertheless, we recommend using the most accurate one, which is RF.
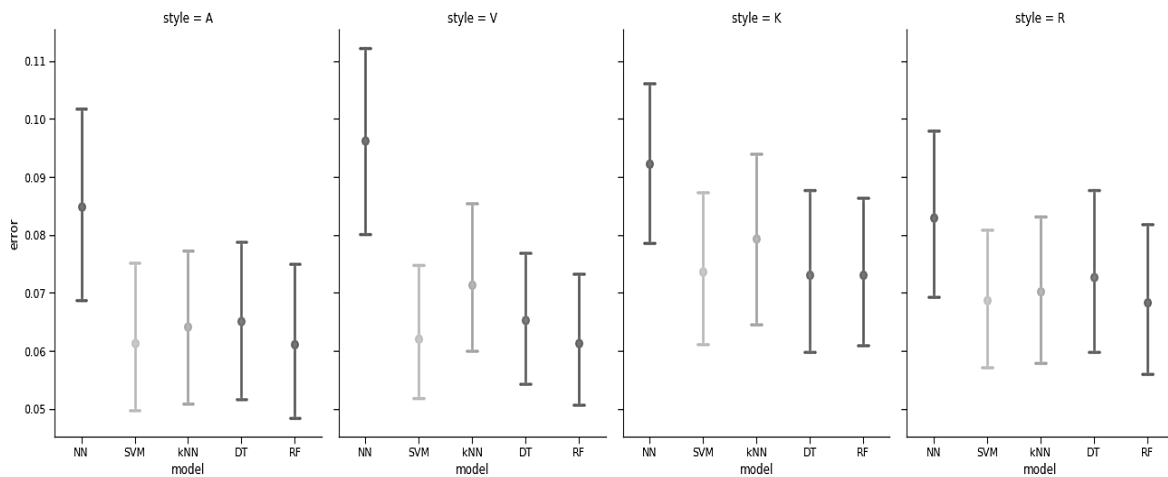
Figure 3. Comparison between prediction models for each learning style, using interval plots

From the above results (both regression and classification), we can conclude that predicting optimal learning style for students as classification is not accurate as predicting the learning style as probabilities. Therefore, the regression algorithms are more accurate. Finally, it is important to mention that several limitations are apparent in our study's last part. The size sample was relatively small, and they studied at the same university what might have influenced the study results. The results would be more precise if the sample size was larger and taken from different universities.

4.   CONCLUSION

Students might find that understanding their learning preferences are helpful. This is supported by recognizing the students' learning styles and approved by many studies that found the use of learning styles in conjunction with other learning methods enhances academic achievements or, at the very least, makes studying more enjoyable. This study is a mixed-method approach that aims to predict the learning styles for learners with mixed styles (with probability). To this end, theories and strategies have been investigated that identify the students' features according to their learning styles. Then the regression analysis was utilized to provide a probabilistic approach for predicting the preferred learning styles. Here, five machine learning algorithms were applied as regression to predict the probability of learning styles, which are multi-layers perceptron NN, SVM, DT, RF, and K-NN. A sample of 72 students was randomly selected to conduct our study. The sample data was collected using VARK's inventory questionnaire with 16 different questions to identify four different learning styles: VARK's. Results showed that the RF algorithm was the superior one

for predicting the probabilities of all learning styles. To examine the effectiveness of our approach, the same set of machine learning algorithms were used to develop classification models for predicting the learning style label. We aimed to compare the finding with the classification-based approach. We observed that the accuracies of all classification models are relatively low. The RF showed the best accuracy with 0.53 to predict learning style (A). Moreover, the overall results are not encouraging, suggesting that none of the models can produce highly accurate predictions. So, we conclude that regression algorithms are more accurate and representative for predicting learning styles' probabilities. As future work, we plan to apply different techniques to our dataset and collect more students' responses to conduct more critical analyses.

ACKNOWLEDGEMENTS

REFERENCES

[1]   M. S. Hasibuan, L. E. Nugroho, and P. I. Santosa, "Model detecting learning styles with artificial neural network," *Journal of Technology and Science Education*, vol. 9, no. 1, pp. 85–95, 2019, doi: 10.3926/jotse.540.

[2]   E. Goh and M. Sigala, "Integrating Information & Communication Technologies (ICT) into classroom instruction: teaching tips for hospitality educators from a diffusion of innovation approach," *Journal of Teaching in Travel and Tourism*, vol. 20, no. 2, pp. 156–165, Apr. 2020, doi: 10.1080/15313220.2020.1740636.

[3]   J. L. Moore, C. Dickson-Deane, and K. Galyen, "E-Learning, online learning, and distance learning environments: Are they the same?," *Internet and Higher Education*, vol. 14, no. 2, pp. 129–135, 2011, doi: 10.1016/j.iheduc.2010.10.001.

[4]   J. Bernard, T. W. Chang, E. Popescu, and S. Graf, "Learning style Identifier: Improving the precision of learning style identification through computational intelligence algorithms," *Expert Systems with Applications*, vol. 75, pp. 94–108, 2017, doi: 10.1016/j.eswa.2017.01.021.

[5]   G. Cheng and J. Chau, "Exploring the relationships between learning styles, online participation, learning achievement and course satisfaction: An empirical study of a blended learning course," *British Journal of Educational Technology*, vol. 47, no. 2, pp. 257–278, Mar. 2016, doi: 10.1111/bjet.12243.

[6]   M. Rais, F. Aryani, and A. S. Ahmar, "The influence of the inquiry learning model and learning style on the drawing technique of students," *Global Journal of Engineering Education*, vol. 20, no. 1, pp. 64–68, 2018, doi: 10.26858/gjeev20i1y2018p6468.

[7]   J. W. Keefe, *Learning style theory and practice*. 1987.

[8]   T. R. Frame, S. M. Cailor, R. J. Gryka, A. M. Chen, M. E. Kiersma, and L. Sheppard, "Student perceptions of team-based learning vs traditional lecture-based learning," *American Journal of Pharmaceutical Education*, vol. 79, no. 4, 2015, doi: 10.5688/ajpe79451.

[9]   A. Bhagat, R. Vyas, and T. Singh, "Students awareness of learning styles and their perceptions to a mixed method approach for learning," *International Journal of Applied and Basic Medical Research*, vol. 5, no. 4, p. 58, 2015, doi: 10.4103/2229-516x.162281.

[10]   N. D. Fleming and D. Baume, "Learning Styles Again: VARKing up the right tree!," *Educational Developments, SEDA Ltd*, vol. 7, no. 4, pp. 4–7, 2006.

[11]   D. C. Kayes, "Internal validity and reliability of Kolb's learning style inventory version 3 (1999)," *Journal of Business and Psychology*, vol. 20, no. 2, pp. 249–257, 2005, doi: 10.1007/s10869-005-8262-4.

[12]   F. Coffield, D. Moseley, E. Hall, and K. Ecclestone, "Learning styles and pedagogy in post-16 learning: A systematic and critical review. National Centre for Vocational Education Research (NCVER)," 2004.

[13]   F. Rasheed and A. Wahid, "Learning style detection in E-learning systems using machine learning techniques," *Expert Systems with Applications*, vol. 174, 2021, doi: 10.1016/j.eswa.2021.114774.

[14]   O. El Aissaoui, Y. El Alami El Madani, L. Oughdir, and Y. El Allioui, "A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments," *Education and Information Technologies*, vol. 24, no. 3, pp. 1943–1959, May 2019, doi: 10.1007/s10639-018-9820-5.

[15]   O. El Aissaoui, Y. E. A. El Madani, L. Oughdir, and Y. El Allioui, "Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles," in *Procedia Computer Science*, 2019, vol. 148, pp. 87–96, doi: 10.1016/j.procs.2019.01.012.

[16]   G. A. M. Kalhoro, A. Ahmed, and S. Rajper, "Detection of E-Learners' Learning Styles: An Automatic Approach using Decision Tree," *International Journal of Computer Science and Information Security*, vol. 14, no. 8, p. 420, 2016.

[17]   O. Pantho, "Using Decision Tree C4. 5 Algorithm to Predict VARK Learning Styles," *International Journal of the Computer, the Internet and Management*, vol. 24, no. 2, pp. 58–63, 2016.

[18]   B. Hmedna, A. El Mezouary, and O. Baz, "Identifying and tracking learning styles in MOOCs: A neural networks approach," *Advances in Intelligent Systems and Computing*, vol. 520, no. 2, pp. 125–134, 2017, doi: 10.1109/ICMCS.2016.7905606.

[19]   V. Yannibelli, D. Godoy, and A. Amandi, "A genetic algorithm approach to recognise students' learning styles," *Interactive Learning Environments*, vol. 14, no. 1, pp. 55–78, Apr. 2006, doi: 10.1080/10494820600733565.

[20]   Y. C. Chang, W. Y. Kao, C. P. Chu, and C. H. Chiu, "A learning style classification mechanism for e-learning," *Computers and Education*, vol. 53, no. 2, pp. 273–285, 2009, doi: 10.1016/j.compedu.2009.02.008.

[21]   C. Lwande, L. Muchemi, and R. Oboko, "Identifying learning styles and cognitive traits in a learning management system," *Heliyon*, vol. 7, no. 8, p. e07701, Aug. 2021, doi: 10.1016/j.heliyon.2021.e07701.

[22]   A. Lincke, M. Jansen, M. Milrad, and E. Berge, "The performance of some machine learning approaches and a rich context model in student answer prediction," *Research and Practice in Technology Enhanced Learning*, vol. 16, no. 1, pp. 1–16, Dec. 2021, doi: 10.1186/S41039-021-00159-7/TABLES/5.

[23]   C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, p. e1355, May 2020, doi: 10.1002/widm.1355.

[24]   A. Jović, K. Brkić, and N. Bogunović, "An overview of free software tools for general data mining," in *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2014 - Proceedings*, 2014, pp. 1112–1117, doi: 10.1109/MIPRO.2014.6859735.
[25]   A. H. Wahbeh, Q. A. Al-Radaideh, M. N. Al-Kabi, and E. M. Al-Shawakfa, "A Comparison Study between Data Mining Tools over some Classification Methods," *International Journal of Advanced Computer Science and Applications*, vol. 1, no. 3, 2011, doi: 10.14569/specialissue.2011.010304.
[26]   A. M. Shahiri, W. Husain, and N. A. Rashid, "A Review on Predicting Student's Performance Using Data Mining Techniques," in *Procedia Computer Science*, Jan. 2015, vol. 72, pp. 414–422, doi: 10.1016/j.procs.2015.12.157.
[27]   H. Aldowah, H. Al-Samarraie, and W. M. Fauzy, "Educational data mining and learning analytics for 21st century higher education: A review and synthesis," *Telematics and Informatics*, vol. 37, pp. 13–49, Apr. 2019, doi: 10.1016/J.TELE.2019.01.007.
[28]   O. El Aissaoui, Y. El Alami El Madani, L. Oughdir, A. Dakkak, and Y. El Allioui, "A Multiple Linear Regression-Based Approach to Predict Student Performance," in *Advances in Intelligent Systems and Computing*, Jul. 2020, vol. 1102 AISC, pp. 9–23, doi: 10.1007/978-3-030-36653-7_2.
[29]   I. Azzi, A. Jeghal, A. Radouane, A. Yahyaouy, and H. Tairi, "A robust classification to predict learning styles in adaptive E-learning systems," *Education and Information Technologies*, vol. 25, no. 1, pp. 437–448, Jan. 2020, doi: 10.1007/s10639-019-09956-6.

BIOGRAPHIES OF AUTHORS

Ahmad Mousa Altamimi ⓘ 🔍 SC Ⓟ is an associate professor of computer science at Applied Science Private University. He has been received his PhD degree in Computer Science from Concordia University-Montreal, Canada, in 2014. His research interests are primarily in machine learning, Cybersecurity, and online education. Dr Altamimi participated in the organization of many conferences. He has many publications in reputable journals and international conferences. He can be contacted at email: a_altamimi@asu.edu.jo.

Mohammad Azzeh ⓘ 🔍 SC Ⓟ is a professor of Computing at Princess Sumaya University for Technology. He holds PhD in computing from the University of Bradford, UK and MSc in Software Engineering from the University of the West of England, UK. His research interests focus on Data Science, Mining Software Repositories, Machine Learning for Software Engineering Problems, and Software Cost Estimation. Dr. Azzeh is an invited referee for high-quality journals and published over 50 research articles in reputable journals and conferences. He can be contacted at email: m.azzeh@psut.edu.jo.

Mahmood AlBashayreh ⓘ 🔍 SC Ⓟ is a PhD holder in Software Engineering since 2014. His research interests include reuse-based software engineering, context-aware computing, mobile patient monitoring systems, application frameworks, and conceptual modelling of information systems. Dr Mahmoud has recently focused on using machine learning and deep learning techniques for Natural Language Processing (NLP). He can be contacted at email: m_albashayreh@asu.edu.jo.