Analysing the learning style of an individual and suggesting field of study using Machine Learning techniques

Conference Paper · July 2019

DOI: 10.1109/ICCES45898.2019.9002051

5 authors, including:

Lekshmi S Nair
Amrita Vishwa Vidyapeetham Kollam
18 PUBLICATIONS **53** CITATIONS

SEE PROFILE

Proceedings of the Fourth International Conference on Communication and Electronics Systems (ICCES 2019)
IEEE Conference Record # 45898; IEEE Xplore ISBN: 978-1-7281-1261-9

Analysing the learning style of an individual and suggesting field of study using Machine Learning techniques

Ananthu S Kuttattu[1], Gokul G S[2], Hari Prasad[3], Jijith Murali[4] Lekshmi S Nair[5]
Department of Computer Science and Engineering
Amrita Vishwa Vidyapeetham
Amritapuri, India
[1]anandhus.kuttattu@gmail.com, [2]gokulg35@gmail.com, [3]prasadhari905@gmail.com,
[4]jijith331@gmail.com, [5]lekshmiprabha.s@gmail.com

***Abstract*—The learning style can be referred to as the way a student prefers to acquire, process and retain information. The prominent learning style classification model is the VAK model. [1] According to this model visual, kinesthetic and auditory are the three major kinds of learning styles. Many research have shown that people prefer more than one way of learning, hence categorizing a person to just one of the above types as done in traditional methods is not accurate.[7] A method to identify our learning styles more accurately is required. Machine learning can be applied in this field to achieve our aim in the most efficient way. Once we have an accurate information about learning styles, we can use it to suggest career options. This research aims to predict the learning style combinations of students and suggest field of study using algorithms like k-means, SVM and decision tree.**

***Index Terms*—Machine Learning, auditory, visual, kinesthetic, Rule based system, k-means, SVM, Decision Tree**

I. Introduction

Choosing an appropriate field for higher studies is a complicated task. Most students require guidance in making the perfect choice. The domain of study of a student needs to suite not only the interest but also the learning style which will impart a way for career advancement. Even though there are many options that help in shortlisting career options, not many of them use learning style as one of the criteria for coming up with the choice.

Most of the traditional ways to identify the learning styles are predefined mathematical equations and patterns of questions that comes to a conclusion. It is not accurate because students prefer more than one mode of learning. The most widely accepted categories of learning styles are:

- Auditory
- Visual
- Kinesthetic

A student is said to be an auditory learner if he/she learns by listening. They tend to understand information in a better way when they hear it. These types of learners are comfortable with lectures, discussions and reading aloud. They get the idea after being told a couple of times since they are good listeners. They have a tendency to read aloud, this helps them in understanding the information in a better manner.

The main tool for a visual learner are their eyes, i.e they learn through vision. Visual learners benefit more from a live demonstration of how things are done. They require diagrams, flowcharts, presentations, etc for a better understanding of the information presented to them.

Kinesthetic learners are those who require hands-on experience to understand the concept. They like to be more involved in the processes that require their attention when compared to the other two types of learners. These learners require active learning methods like role-playing, hands-on workshops, practicals, etc for processing the information in a better manner.

Categorizing a student based on the above mentioned learning styles is not an easy task, because most students prefer more than one way of learning. For example, a visual learner might be interested in hands-on workshops. In such case a student can be considered as a kinesthetic learner too. In these type of situations, the traditional methods for identifying the learning style which are based on predefined equations and question patterns are less accurate. Thus it demands the possibility of the existence of combinations in learning styles. A student may exhibit more than one learning style. In order to find this combination of learning styles, we apply machine learning. We will be using the classification and the clustering to fit a learner into an appropriate class of learning style.

Clustering is an unsupervised machine learning technique. We used the k-Means clustering algorithm as it is an iterative algorithm, that tries to partition the dataset into k predefined distinct non overlapping subgroups where each data point belongs to only one group and inter cluster data points can be made. We do the classification which groups new data into known classes while clustering relies on underlying patterns to come up with the clusters. We have used two classification algorithms, Support vector machines (SVM) and Decision tree. We have used SVM to predict the individual learning style and decision tree to predict the learning style combination. All these algorithms were chosen in order to maximise the

accuracy and efficiency of our program.

Once we have acquired a student's learning style combination, we can compare it with the learning style combinations prevalent in existing field of study. Based on the degree of similarity we can suggest an area of study to that person. Career selection is a subjective process and depends on the interests of the individual too. With this work, we hope to suggest valid career options for prospective learners based on their learning style combination.

II. Background and related work

Much study was done in the field of identification or prediction of learning styles. For example, a survey was conducted in 2016 on 175 students studying MS programming in engineering. For this survey they have used the questions in Index of Learning Styles by Felder and Soloman [15]. According to this survey, there is a huge diversity in the learning styles among engineering students studying in different branches.[6]

A similar survey was conducted among medical students [3]. A questionnaire consisting thirty nine questions were used. Two-hundred and ninety students participated in the survey. Based on the results seventy three percent of the students who participated in the survey prefer only one learning style, twenty two percent prefer two learning styles and the rest prefer three learning styles. Most of the students who prefer single learning style prefer visual learning style followed by auditory and then kinesthetic. Among the students who prefer two learning style fifty four percent prefer visual and auditory, twenty six percent prefer visual and kinesthetic and the rest prefer auditory and kinesthetic.

An Indonesian translated version of Joy Reids (1984) Perceptual Learning Style Preference Questionnaire was used to conduct a survey among 3rd semester English Department students of Faculty of Cultural Studies at Universitas Brawijaya [2]. A total of hundred students participated in the survey, out of which sixty six were females and the rest were males. The study revealed that majority of the male students who took the survey prefers kinesthetic mode of learning while majority of female students preferred either kinesthetic or group study.

VARK questionnaire developed by Flemming and Miles (1992) was used to assess the learning style preferences of the engineering students at Atlm University [4]. The questionnaire was answered by 107 students out of which sixty one were from computer science, twenty two were from manufacturing engineering and twenty four were from mechatronics engineering. The study showed that around 25.2% of the students prefer more than one mode of learning style. 27.9% of the computer science students who took the survey prefers multiple modes of learning styles and among the students who prefer single learning style, the most preferred styles are auditory, kinesthetic and reading/writing. The most popular learning style among the manufacturing engineering students was found out to be auditory. 29.2% of mechatronics students prefer multiple learning style and among the rest of the students who prefer single learning style, the most preferred style was auditory.

One instance where machine learning was used to identify the learning style is the website 'First aid for you'. Here, machine learning was used to observe how the user interacts with the website and identify patterns in those interactions Those patters are then used to identify the best type of content for the next page. Here, they have used Felder & Solomon Index of Learning Style. Nave Bayes algorithm was used to identify the preferred learning style of the user and this information was used to customize the learning environment.[11]

In [5], the use of k-means in multi label classification is being discussed. The rule mining algorithm Apriori is a popular choice for rule generation, but when it comes to numerical data. With proper pre-processing, a lot of intricate rules can be derived form those numerical datasets. K-means algorithm can be used for this purpose. This helps in finding hidden patterns in huge numerical datasets.

A detailed study of the advantages and disadvantages of both traditional and modern techniques for learning style identification methods was done in [7]. Based on the result it was identified that most of the traditional methods which uses questionnaire and patters in questions lacks authenticity because they follow a general rule which cannot be applied in all individual cases. Modern techniques using machine learning are computationally expensive compared to the traditional ones but they tend to be more accurate.

Support vector machine (SVM) algorithm is widely used in classification and regression analysis. [8] does a detailed analysis on the implementation of SVM using both linear and Map-Reduce fashion. The performance of thus implemented algorithm is then compared.

A fuzzy inference system was used for predicting preferred learner style based on the features used for the VARK questionnaire. The system was tested on a group of students. The input data for the system was collected from a custom questionnaire that was designed specially to get the inputs for the fuzzy inference system. On comparing the output of the fuzzy system to that of the VARK system it was observed that 48% of the test cases were similarly classified.[10]

K-Means clustering encounters issues when it deals with text data. [12] focuses on eliminating this drawback by converting text data into a numerical value which results in more distinct clusters and accurate running times.

The use of tree augmented Bayes network was observed to have better classification accuracy when compared to naive Bayesian network. [13] Online learning environment was used as the platform for automatically detecting students learning style using tree augmented Bayesian network. The experimental results proved that the tree augmented naive Bayes network is capable of achieving higher detection accuracy when compared to the Bayesian network.

Decision trees are commonly used for representing data. In data mining, many different decision tree algorithms are used and each of those provide a unique decision tree from the input data. [14] focus on the comparison of different decision tree algorithms for data analysis.

A study was conducted among the dental students of King

Saud University to find their learning style pattern [16]. An arabic version of VARK questionnaire was used to conduct the survey. The results showed that most students (58.4%) preferred more than one learning style. Among the students who preferred single learning style, Kinesthetic and Auditory were the most prominent.

In [17] the authors have used supervised and unsupervised machine learning algorithms to predict the learners learning style. They have used the Felder and Silverman learning style model for categorizing learners. The unsupervised learning algorithm that they have used is K means and the supervised learning algorithm that they have used is Naive Bayes.

The learning style identification survey done among the architecture and interior design students reveal that there is no significant difference between the learning styles between students of architecture and interior design [18]. Here, the authors used Felder and Soloman's Index of Learning Styles questionnaire to conduct the survey.

III. Proposed Work

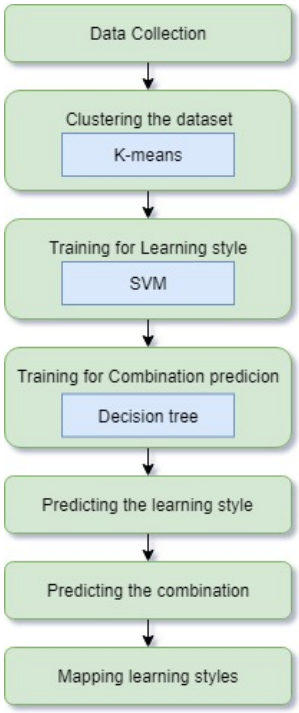

Fig. 1. Steps involved

A. Data Collection

A survey was conducted among the student community of Amrita Vishwa Vidyapeetam University. The survey consisted of fifteen questions. All the questions were multiple choice type and choices included "Strongly Disagree", "Disagree", "Neutral", "Agree" and "Strongly Agree". Each of the above value is mapped to a numerical value, with Strongly disagree as one and Strongly agree as five. The survey was conducted among people with age ranging from fifteen to twenty five. We were able to acquire one thousand one hundred and twenty three inputs. The obtained results was then stored in CSV format.

B. Data Clustering

Clustering of the data helps us to find underlying patterns and categorize them to different groups called clusters. Here we are using K-means clustering algorithm. k-means algorithm is computationally faster when compared to other clustering algorithms when the number of variables are more. According to this algorithm, the initial step is to assign k centroids randomly. Once we identify the centroids, Euclidean distance between each data point and the centroid is calculated. Each data points is then assigned to the centroid which is closest to it. The centroids are then re-positioned for optimization. These processes are iteratively repeated until the centroids doesn't change anymore i.e the centroids are at the center of the clusters that we wanted to create. Here, after applying the k-means algorithm we end up with fifty one clusters. Each and every data will have a cluster number associated with it. We add this cluster number as a feature in the dataset, so the total number of features will be sixteen after data clustering. Once the cluster number is also added to the dataset, we can divide the dataset into test set and train set. In this case we will be dividing the data set equally and one will be taken as test set and the other will be the train set.

C. Data Classification

Studies have shown that clustering the data before classification can increase the accuracy of the classification process. [9] Also the cluster number that we added after the clustering will play a vital role in classifying the data. Logistic regression is a simple and fast classification algorithm and we will be using it in our project. One of the simplest and highly preferred classification algorithm which produce significant accuracy with comparatively less computation power is the Support Vector Machines (SVM). Support Vector Machine is used for regression and classification. The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. We apply this algorithm on the training set and trains the algorithm. Once the training is done we apply the same on the test set to predict the learning style. We compare the predicted value of learning style to the existing value and calculates the accuracy of our prediction. We obtained an accuracy of 92.87%.

D. Combination Prediction

We are at a stage where we could predict the learning style of a person but we are limited to a single learning style per person. Our next aim is then to find a combination of learning styles per person. For this we have used a statistical technique called Rand index. Rand index is used for finding the similarity between clusters. Combinations are obtained from clusters that are similar. If a data point that is in a cluster

that is predominantly auditory but if that cluster is similar to another cluster that is predominantly visual, we can trace that data point to visual learning style too, in such a way we find the combinations for all the data points present in training set.

Instead of calculating the combinations every time using adjusted rand index, we could use machine learning to predict the combinations. One of the most widely used predictive analytic algorithm is decision tree. Decision tree construct a tree of predictive decisions made based on actual values of attributes in the data. We create a training model using decision tree algorithm which takes the cluster value, the individual learning style and the combinations obtained using the rand index method as input. We pass the above said features of the training set as input to the training model. By doing so we are training the algorithm for prediction of combinations. Once we have trained the algorithm, we could pass the test set to the trained algorithm and obtain combinations as output.

E. Mapping

Now we are able to predict the learning style combination, in-order to suggest career options we require the information about the prevalent learning styles present in each career. The only way to acquire such information is through extensive survey. Once we have a database containing the most common career options and their prevalent learning style combinations, we could compare it with the learning style of a person and come up with valid suggestions. As of now, information in this regard is limited and based on some of the previous studies done in various parts of the world, we have acquired the following information:

- The prevalent learning styles present among **Computer Science** graduates are ***Auditory*** and ***Kinesthetic***.
- The prevalent learning style present among **Manufacturing Engineering** graduates is ***Auditory***.
- The prevalent learning style present among **Mechatronics Engineering** graduates is ***Auditory***.
- The prevalent learning style present among **Medical** graduates is ***Visual***.

Based on the learning style combination of a person we could suggest any one or more of the career options from the above list. We have used a rule based system for mapping the learning styles. The rules that we have created for the rule based system are as follows:

- ***Auditory***, ***Kinesthetic*** and ***Visual***: Computer Science, Manufacturing engineering, Mechatronics engineering and Medical
- ***Auditory*** and ***Kinesthetic***: Computer Science, Manufacturing engineering and Mechatronics engineering
- ***Kinesthetic*** and ***Visual***: Computer Science and Medical
- ***Auditory*** and ***Visual***: Computer Science, Manufacturing engineering, Mechatronics engineering and Medical
- ***Auditory***: Computer Science, Manufacturing engineering and Mechatronics engineering
- ***kinesthetic***: Computer Science
- ***Visual***: Medical

IV. RESULT

The initial distribution of learning styles in our database is as shown in Fig. 2.

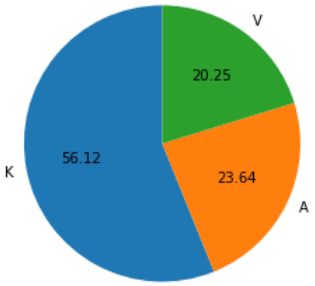

Fig. 2. Initial classification based on learning styles

Based on the rules that we have developed for mapping we can infer that fifty six percent of the students who took our survey can choose Computer Science related field as a career option and twenty three percent of the students can choose Computer science, Manufacturing engineering or Mechatronics engineering as a career option while the rest can choose medical field as their career choice. Using our proposed model we were able to efficiently predict the learning style combination of a student. After passing the same database to our program we obtained the result as depicted in Fig. 3.

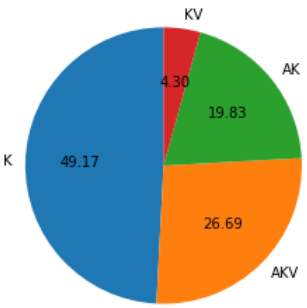

Fig. 3. classification based on learning styles after using our program

Here we observe the presence of combinations like Kinesthetic and Visual (KV), Auditory and Kinesthetic (AK) and Auditory, Kinesthetic and Visual (AKV) along with Kinesthetic. If we apply our mapping rules to the above result, we observe that around twenty six percent of the students who took our survey can choose any of the four career options, around twenty percent of the students can take Computer Science, Manufacturing engineering or Mechatronics engineering, nearly five percent of the students can take Computer Science or Medical as their choice while the rest can choose

Computer Science as their career option. Compared to the previous observation, learning style combinations have given us more accurate and specific results. Thus, identifying a student's learning style combination can help us in coming up with more precise and unique career suggestions.

V. Conclusion

We were able to learn in depth about student learning styles. Unlike the traditional methods, we are now able to predict the learning style combinations of a student using machine learning. The use of machine learning algorithms has provided a new perspective on the technique of finding learning styles. Using a rule-based system, we were able to map learning styles to certain fields of study and identifying the fields of study that suits a student's learning style can help the student in shortlisting career choices.

VI. Future Scope

As per the current implementation, we are able to suggest only four field of study. This is because of the lack of data regarding the prevalent learning styles in other fields of study. Much research can be done in this area and once we have enough information, we can suggest more and more options to students.

VII. Acknowledgement

We would like to express our sincere gratitude to the students of Amrita University who took part in our survey.

References

[1] MR. SREENIDHI S K and MS. TAY CHINYI HELENA, "Styles of Learning Based on the Research of Fernald, Keller, Orton, Gillingham, Stillman , Montessori and Neil D Fleming", INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD ISSN 2455-0620 Volume - 3, Issue - 4, Apr - 2017

[2] Peptia Asrining Tyas, Mega Safitri, "Kinesthetic Learning Style Preferences: A Survey of Indonesian EFL Learners by Gender", Journal of English Educators Society, (April 29, 2017).

[3] Alina-Mihaela Buan, "Learning Styles of Medical Students - Implications in Education", US National Library of Medicine National Institutes of Health, (April 2014)

[4] Burcu Devrim Ictenbas, Hande Eryilmaz, "Determining Learning Styles of Engineering Students to Improve the Design of a Service Course ", Procedia - Social and Behavioral Sciences (Volume 28 2011)

[5] H Haripriya, Prathibhamol Cp, Yashwant RPai and M Sai Sandeep, "Multi Label Prediction Using Association Rule Generation and Simple k-Means", 2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)

[6] Tulsi PK; Poonia MP and Anu Priya, "Learning Styles of Engineering Students.", Journal of Engineering Education Transformations , Volume 30 , No. 2, October 2016

[7] Y. Helmy., A. Abdo. and R. Abdallah., "A Proposed Framework for Learning Style Prediction in Higher Education Environments.", International Journal of Advanced Research in Computer Science and Software Engineering (Volume 6, Issue 3, March 2016).

[8] M. V. Sukanya, Shiju Sathyadevan, and Sreeveni, U. B. Unmesha, Benchmarking support vector machines implementation using multiple techniques, in Advances in Intelligent Systems and Computing, 2015, vol. 320, pp. 227-238.

[9] Yaswanth Kumar Alapati, Korrapati Sindhu "Combining Clustering with Classification: A Technique to Improve Classification Accuracy", International Journal of Computer Science Engineering (IJCSE).

[10] Marwah Alian and Adnan Shaout, "Predicting learners styles based on fuzzy model", Springer Science + Business Media New York 2016

[11] Declan Kelly, Brendan Tangney, "First Aid for You: Getting to know your Learning Style using Machine Learning. In: Advanced Learning Technologies", 2005. ICALT 2005. Fifth IEEE International Conference (August 2005)

[12] Unnikrishnan, S.; Sreelakshmi, S.; Deepa, G.,"Enhancement of accuracy in K-means clustering" , International Journal of Control Theory and Applications Volume 9, Issue 15, 2016, Pages 7619-7626

[13] Ling Xiao Li and Siti Soraya Abdul Rahman, "Students learning style detection using tree augmented naive Bayes", Royal Society Open Science (June 21, 2018).

[14] Sathyadevan S., Nair R.R. (2015), "Comparative Analysis of Decision Tree Algorithms: ID3, C4.5 and Random Forest.", In: Jain L., Behera H., Mandal J., Mohapatra D. (eds) Computational Intelligence in Data Mining - Volume 1. Smart Innovation, Systems and Technologies, vol 31. Springer, New Delhi

[15] "Index of Learning Styles Questionnaire", https://www.webtools.ncsu.edu/learningstyles/

[16] Moshabab A. Asiry, "Learning styles of dental students", The Saudi Journal for Dental Research (2016) 7, 1317

[17] Ouafae EL AISSAOUI, Yasser EL ALAMI EL MADANI, Lahcen OUGHDIR and Youssouf EL ALLIOUI, "Combining supervised and unsupervised machine learning algorithms to predict the learners learning styles", Second International Conference on Intelligent Computing in Data Sciences (ICDS 2018)

[18] Wafa Labiba, Irene Pasinaa, Abdelhakim Abdelhadia, Goze Bayrama, Mohammad Nurunnabi, "Learning style preferences of architecture and interior design students in Saudi Arabia: A survey", MethodsX Volume 6, 2019, Pages 961-967