Pune Institute of Computer Technology
Dhankawadi, Pune

A SEMINAR REPORT
ON

DETERMINING THE RELATIONSHIP BETWEEN LEARNING STYLE AND ACADEMIC PERFORMANCE USING A MACHINE LEARNING APPROACH

SUBMITTED BY

Name : Isha Abhijit Joshi
Roll No. : 31231
Class: TE-2

Under the guidance of
Prof. Sheetal Sonawane

DEPARTMENT OF COMPUTER ENGINEERING
Academic Year 2021-22

DEPARTMENT OF COMPUTER ENGINEERING

Pune Institute of Computer Technology

Dhankawadi, Pune-43

CERTIFICATE

This is to certify that the Seminar report entitled

DETERMINING THE RELATIONSHIP BETWEEN LEARNING STYLE AND ACADEMIC PERFORMANCE USING A MACHINE LEARNING APPROACH

Submitted by

Isha Abhijit Joshi Roll No. : 31231

has satisfactorily completed a seminar report under the guidance of Prof. Sheetal Sonawane in the Third Year Computer Engineering Semester I, Academic Year 2021-22 of Savitribai Phule Pune University.

Prof. Sheetal Sonawane
Internal Guide

Prof. M.S.Takalikar
Head
Department of Computer Engineering

Place:Pune
Date: 10/10/2021

ACKNOWLEDGEMENT

It is with great delight that I present my report on "Determining the Relationship Between Learning Style and Academic Performance Using a Machine Learning Approach". I am immensely thankful towards our Seminar Coordinator Prof. D.D.Kadam, Head of Department Prof. M.S.Takalikar and Principal R.Sreemathy for their encouragement and support.

I am extremely grateful towards my respected seminar guide Prof. Sheetal Sonawane, Department of Computer Engineering, for her constant guidance and support. She constantly encouraged me and aided me patiently throughout the course of my seminar report. Her guidance was crucial in the success of my seminar completion.

Lastly, I would like to thank the faculty and my parents, along with my friends, for their essential support.

Contents

List of Figures

Abstract

Understanding the impact of a student's learning style on his or her academic performance can help educational institutions design personalized education programs for students who have a preference for a particular learning pattern. For the purpose of learning information, students use four sensory modalities (Visual, Aural, Read/Write and Kinesthetic).

The impact of preferred learning modalities on academic success has been analyzed through a machine learning approach. Firstly, the learning style of each student was determined through a V.A.R.K.(Visual, Aural, Read/Write and Kinesthetic) questionnaire. Secondly, different machine learning models were used to classify students into binary classes on the basis of their preferred learning style. The results were explored and studied to determine the relationship between academics and learning pattern in order to conclude if a particular learning style resulted in better academic achievement.

Keywords

Machine learning, Logistic regression, Support vector machines, K nearest neighbors

1 INTRODUCTION

Students use different patterns in order to learn information. Many learning models have been developed in order to determine this learning pattern. One widely used model is the V.A.R.K model. The model determines the student's scores for each of the four learning modalities - Visual, Aural, Read/Write and Kinesthetic. Students who have a preference for visual learning, learn through diagrams, pictures etc. whereas those who prefer auditory learning, learn better through discussions and lectures. While some students prefer reading and writing in order to learn, others prefer kinesthetic learning,i.e. they learn information through experiments and hands-on activities.

A few studies have been carried out to determine the relationship between learning patterns and academic performance. It is important to understand the impact of learning patterns on academic success so as to ascertain whether students who do not obtain information according to their preferred modality are unable to perform well, academically speaking. This information can be used to reform various educational programs as well. The resources provided to students in order to learn information can be tailor made in accordance with their learning style. This will give students with different learning patterns a chance to better their academic achievement.

In order to determine the relationship between learning style and academic achievement, one can perform classification of students into categories which determine their performance as "good" or "bad". This can be done by classifying students based on range of academic scores. In order to perform this classification, we require data. We must first determine the students' V.A.R.K scores in order to ascertain their learning style. After this, it is important to collect data related to the academic performance of these students, whether it is their percentage/ grade point average. These scores will have to be manually classified into categories. This data can be used to train a machine learning model to predict the category of the student based on their V.A.R.K scores. The data must be for students from the same discipline.

The proposed methodology in this report suggests that we perform binary classification of the students based on their score. A threshold value has been used to determine if a student has a "good" or "bad" academic score. Algorithms such as logistic regression, support vector machines and k-nearest neighbors can be used to build a model to perform binary classification on the data set. The algorithms can be compared on the basis of their F1 score, precision and recall. With the help of the above parameters, the best algorithm to describe the relationship between V.A.R.K scores and academic success can be chosen.

This problem of determining the relationship between learning preference and academics is very important to explore. It can change the way learning occurs in educational institutes. It will certainly help students perform better and give them the ability to learn information more effectively and with ease. This challenge can be further explored with more data and advanced deep learning techniques, such as neural networks, can be used to further studies on this topic.

2 MOTIVATION

Educational programs are not tailor made for students based on their learning preferences. This causes them to lose confidence in themselves and believe that they do not have the ability to learn. They are not provided with the type of resources that suit their learning style and hence, are not able to learn as efficiently as they potentially can.

With the advancement in the field of machine learning, we can apply a machine learning approach to determine the relation between academics and learning modalities defined by various learning models like the Felder-Silverman and 4MAT models. By training different models to determine the same, we can find a conclusive relationship for further studies.

Thus, we can use this information so as to help educational institutes reform their programs. These programs can be designed as per the learning modalities of students. Resources which align with their learning pattern can be provided for a better learning experience. Further studies can be performed to observe the impact of these changes on academic performance. This will most definitely aid students in achieving success with respect to their academics.

3 LITERATURE SURVEY

In order to determine the relationship between academic success and learning style, it was important to take a look at the different approaches and performance metrics that were used in various research papers.

3.1 An Adaptive Learning System Based on Proportional VARK to Enhance Learning Achievement Concept [1]

This paper proposes tailoring resources for learners with different styles of learning to improve learner achievement.

The study uses adaptive models to tailor material. It determines the learner's style based on exercises and tests performed by the learner. It then generates the next few resources for learning on the basis of the learning pattern that is determined. These results are further analyzed and the paper concludes that adaptive learning helps learners achieve academic success.

3.2 Process of building a dataset and classification of vark learning styles with machine learning and predictive analytics models [2]

This paper aims to find a way to classify students into the four modalities defined by the V.A.R.K model. The algorithm used in this paper makes use of features that can influence one's learning style, for example, demographic or behavioural descriptors. Many algorithms were used to perform the classification and metrics such as accuracy, recall, precision, f1-score, cross validation score were used to determine the best performing algorithm. It was concluded that the voting classifier performed better with an accuracy of 0.7, recall score of 0.7, precision score of 0.73 and f1-score of 0.675.

3.3 A Comparative Study of Teaching Styles in Online Learning Environment [3]

This paper worked towards concluding if the teaching style of an instructor with respect to online learning made an impact on academic performance of students. Students were divided into two classes for an online course which was 12 weeks long. The first class was assigned to an instructor who was not very proactive. However, the second instructor monitored student discussions and gave timely feedback for all the coursework. The paper concluded that the second instructor's class performed better as compared to the first. It was discovered that a more responsive instructor who monitored the progress of students was likely to ensure a good academic performance.

3.4 The Exploration of a Machine Learning Approach for the Assessment of Learning Styles Changes [4]

In this paper, the learning style of undergraduate students who are learning English as a foreign language/second language was studied. It aims to assess the change in learning style of such students and to also analyze which year of their undergraduate studies is a key contributor with respect to this change. Firstly, the learning style of these students was determined for data collection. Then, a machine learning approach was used (Support Vector Machines) to determine the degree of separation between various years of the undergraduate course. It was found that there was a large degree of separation between first years and fourth year students. It was also discovered that the third year is the year in which there is a major change in learning style. The machine learning technique was found to be very effective in determining the change in learning style in ESL/EFL undergraduate students.

3.5 Does a Good Match of Trainees' Learning Styles to their Tutors' Instructional Strategies Contribute to Trainees' Academic Achievements? [5]

In this paper, the correlation between learning style of students and instructional strategies of teachers was determined. This paper made use of the Felder-Soloman index of learning styles to measure the learning style and instructional strategies. Tutors were paired with students with learning disabilities for this study. It was found that there was no consequential correlation between the two.

The above papers explore the many aspects of learning patterns. While some papers focus on the learning resources and their impact on different learning modes, some papers propose some methods to determine the learning style itself. However, the dominant learning modality of students has not been directly compared with the academic performance of students. It is important to understand the immediate impact of learning patterns on academics in a quantitative manner.

4 PROBLEM DEFINITION AND SCOPE

4.1 Problem Definition

To determine the relationship between learning style and academic performance using a machine learning approach

4.2 Scope

The V.A.R.K scores of students, determined through a questionnaire, are used as features (along with the gender of the student) in order to perform binary classification of students. Students were classified into those who were able to succeed academically and those who were not.

In order to do that, the dataset was first normalized. Secondly, three classification algorithms were used (logistic regression, support vector machines and knn neighbors classifier) to determine which algorithm performs the binary classification with the most accuracy. The accuracy was determined with the help of some performance metrics.

5 DIFFERENT ALGORITHMS USED

The following algorithms have been used for determining the relationship between learning style and academics. Due to the fact that these algorithms are suitable for the current application with respect to the dataset size, only these three algorithms have been explored.

5.1 Logistic Regression

Logistic regression is a machine learning algorithm used for classification. It maps the input to a value between 0 and 1 with the help of a sigmoid or logistic function.

5.2 Support Vector Machines

Support vector machines are a powerful tool used in machine learning for classification problems. It learns a decision boundary to perform classification on the given data points.

5.3 K-nearest Neighbors Classifier

This machine learning algorithm can be used for both regression and classification problems. It makes use of 'k' number of nearest neighbors so as to classify a given data point. The distance between data points can be measured in many ways, for example, euclidean distance.

6 METHODOLOGY

6.1 Workflow

The methodology followed in this project has been highlighted in Figure 1 below. The workflow has been divided into 4 major steps.

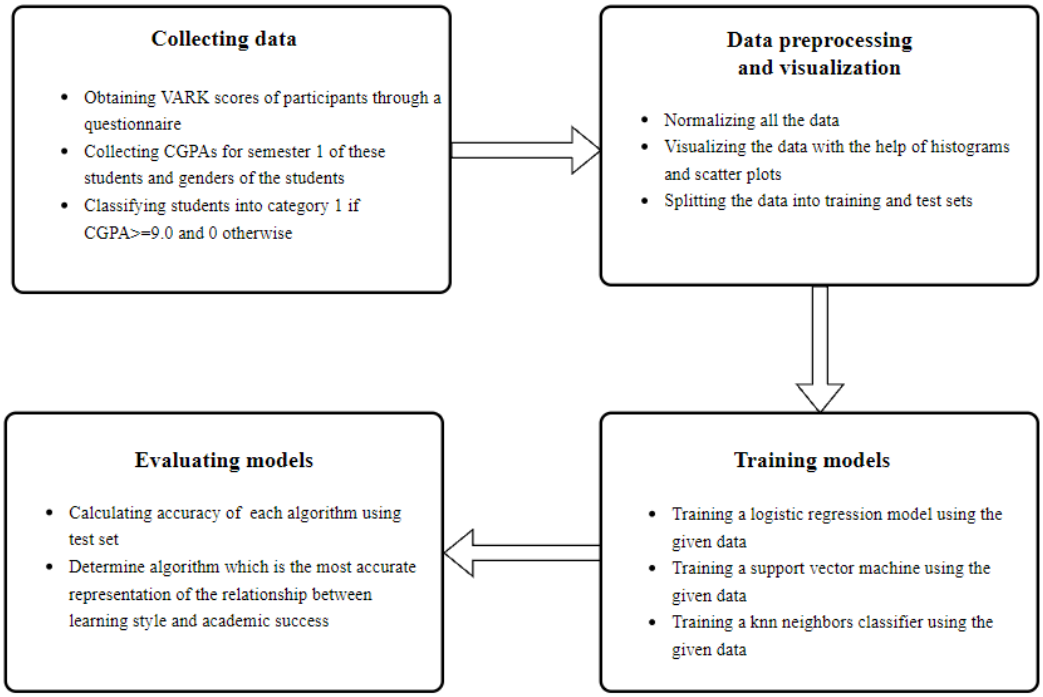

Figure 1: Relationship evaluation method: Workflow

The first step involves collecting the scores of the students for each learning modality. The cumulative grade point average (scored out of 10.0) along with the gender of the students was also collected. Students with a grade point average greater than or equal to 9 were placed in the academically successful category, whereas the rest were categorized into another category. In the second step, the scores were normalized. The required data visualizations were also performed to ensure that the data was balanced.

In the third step, three machine learning algorithms were used to build a model to classify students into those who performed well, academically and those who did not, based on their learning modality scores and other features. The algorithms were compared using the performance metrics mentioned in the figure.

6.2 Mathematical model

Sigmoid function for logistic regression

$$\sigma(x) = \frac{1}{1 + e^{-}x}$$

$$\sigma'(x) = \sigma(x).[\sigma(x) - 1]$$

$$\sigma(x) = \begin{cases} 1, & x \to \infty \\ 0, & x \to -\infty \end{cases}$$

The sigmoid function is used in machine learning to perform classification on the given data. In order to perform binary classification, a threshold is used to determine which category the data point belongs to.

Support vector machine

$$SVM = \begin{cases} w.u + b > 0, & positive\,class \\ w.u + b \leq 0, & negative\,class \end{cases}$$

A support vector machine is a robust algorithm used in classification and regression problems. It uses a hyper-plane to classify the given points into classes. The boundary learned determines the category of the data point.

Euclidean distance for k-nearest neighbors

$$d\left(p, q\right) = \sqrt{\sum_{i=1}^{n} \left(q_i - p_i\right)^2}$$

The k-nearest neighbors classifier makes use of data points and classifies an unknown point by using its k-nearest neighbors. The distance between the points is determined with the help of many distance formulae, like the Euclidean distance formula for example.

6.3 Key Terms

TP: True Positives
TN: True Negatives
FP: False Positives
FN: False Negatives
SVM: Support Vector Machine
RBF: Radial Basis Function
m: Number of Training Examples
t: Number of Training Examples
y-pred[ ]: Predicted Values for Training Examples
y-true[ ]: Actual Values for Training Examples

Calculations

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$f1 - score = \frac{2 * precision * recall}{precision + recall} \tag{3}$$

6.4 Algorithm

Number of data points := n
model-one := logistic regression model
model-two := SVM using an RBF kernel
model-three := K-nearest neighbors classifier

for k=1 to n do
    normalize visual, aural, read/write and kinesthetic scores
end for

for j=1 to m do
    train model-one using training data
    train model-two using training data
    train model-three using training data
end for

for every model do
    TP := TN := FP := FN := 0
    for i=1 to t do
        y-pred[t] := predicted value
        y-true[t] := true label for example
        if y-pred[t]= 1 and y-true[t]=1 then
            TP := TP + 1

```
            end if
            if y-pred[t]= 1 and y-true[t]=0 then
                  FP := FP + 1
            end if
            if y-pred[t]= 0 and y-true[t]=1 then
                  FN := FN + 1
            end if
            if y-pred[t]= 0 and y-true[t]=0 then
                  TN := TN + 1
            end if
      end for


precision := TP/(TP+FP)
recall := TP/(TP+FN)
f1-score:= 2 * precision * recall/ (precision + recall)
end for
```

7 CONCLUSION

Three classification algorithms were used to fit the given data, logistic regression, support vector machine (with an RBF kernel) and K-nearest neighbors classifier (with k=2). Performance metrics were used to measure the accuracy of all three models.

Upon comparing the accuracy of the models that were chosen, it was found that both the k-nearest neighbors classifier and logistic regressor gave a slightly better performance as compared to the support vector machine

References

[1] Yueer Wei, Qingxia Yang, Jiangping Chen and Jie Hu, "The Exploration of a Machine Learning Approach for the Assessment of Learning Styles Changes," Mechatronic Systems and Control, Vol. 46, No. 3,pp.121-126, 2018 " .

[2] Beesuda Daoruang, Suthida Chaichomchuen and Anirach Mingkhwan, "An Adaptive Learning System Based on Proportional VARK to Enhance LearningAchievement Concept," The 4th Joint Symposium on Computational Intelligence (JSCI 4), 2 February 2018, Bangkok, Thailand

[3] Titan, Ferdianto, G. G. Faniru Pakuning Desak, "A comparative study of teaching styles in online learning environment," International Conference on Information Management and Technology (ICIMTech), pp.25-30, 2017

[4] D. Krishnamoorthy, Duvvuru Lokesh, "Process of building a dataset and classification of vark learning styles with machine learning and predictive analytics models," Journal of Contemporary Issues in Business and Government Vol. 26, No. 2, pp.904-912, 2020

[5] Nissim Sabag, Hagit Krisher, "Does a Good Match of Trainees' Learning Styles to their Tutors' Instructional Strategies Contribute to Trainees' Academic Achievements?," IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE), pp.95-99, 2017

Curiginal

Document Information

Analyzed document	Report_Plagiarism_Check_31231.pdf (D118656804)
Submitted	2021-11-15 15:09:00
Submitted by	
Submitter email	sssonawane@pict.edu
Similarity	1%
Analysis address	sssonawane.pict@analysis.urkund.com

Sources included in the report

W URL: http://ir.uew.edu.gh/bitstream/handle/123456789/1421/Exploring%20the%20correlations%20between%20learning%20styles%20and%20academic%20performances%20of%20shs%20students%20in%20integrated%20science%20in%20gomoa%20east%20district.pdf?sequence=1&isAllowed=y
Fetched: 2021-11-15 15:36:00 1