

Prediction of Student Performance Based on Behavior using E-Learning During the Covid-19 Pandemic using Support Vector Machine

Agung Eka Widarta^{1)*}, Ahmad Luthfi²⁾, Chandra Kusuma Dewa³⁾

^{1,2,3)}Islamic University of Indonesia

¹⁾20917001@students.uui.ac.id, ²⁾ahmad.luthfi@uui.ac.id, ³⁾chandra.kusuma@uui.ac.id

Submitted : Aug 8, 2023 | Accepted : Oct 16, 2023 | Published : Jan 1, 2024

Abstract: The COVID-19 crisis has profoundly impacted many sectors globally, including education, necessitating the shift from traditional in-person learning to independent or online learning through various digital platforms. The integrity of e-learning can be ensured by leveraging e-learning behavioral data. The objective of this research is to develop a novel data model to navigate the educational challenges of the COVID-19 era. Previous studies employed the Support Vector Machine (SVM) technique to predict student performance in an e-learning setting, yet they failed to contrast different SVM kernels and their outcomes. In contrast, this study uses SVM and compares three types of kernels: Radial, Polynomial, and Linear. The dataset used for this research was procured from X-API-Edu-Data. The SVM technique was utilized in a unique way to process the data, which comprised 17 variables and 40 observations. Notably, all 17 variables were character variables, with only four being numeric. Two variables, Raisedhands and Discussion, were selected for analysis due to their key role in effective learning and their association with student performance in an e-learning environment. The evaluation of the model was performed using the Topic variable, which represents the subjects in the dataset. The research findings revealed a marked improvement in accuracy compared to earlier studies. Among the three SVM kernels tested - Radial, Polynomial, and Linear, the Polynomial kernel demonstrated superior accuracy with a score of 0.9979. Therefore, the Polynomial model was deemed most appropriate for analyzing the Topic variable. In conclusion, the study indicates that the application of the e-learning method, specifically during the COVID-19 pandemic, proved highly effective in forecasting student performance.

Keywords

Student performance, Covid-19 Pandemic, E-Learning, Support Vector Machine (SVM)

INTRODUCTION

Education is very much needed in the continuity and welfare of a person's life even in the welfare of a nation. With education one will avoid ignorance and poverty, because with the knowledge and skills acquired through the educational process students are able to overcome the various life problems they face (Sagala, 2003). Effective learning is expected to be well accepted and practiced by students (Kyriacou, 2009). This is in accordance with the provisions of PP No. 47 of 2008 Article 7 paragraphs 4 and 5 concerning the Legal Basis for 12 Year Compulsory Education. Educators and students are mutually successful in carrying out teaching and learning activities so that good and smooth education is created (Maswan and Muslimin, 2017).

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

However, there are always obstacles in terms of learning in the process of imparting knowledge between educators and students, one of which is the situation of the Covid -19 Pandemic which hit all parts of the country some time ago. (Anggun Wulandari et al, 2020). This condition then has implications for changes in the pattern of delivery of learning material that was previously face-to-face, then must be done remotely or online. This happened because the Covid-19 Pandemic is a deadly epidemic and one of the transmissions is social interaction.

In the process of student performance based on behavior using e-learning began to be studied to predict the way students learn during the Covid-19 pandemic (Sahfitri, 2015). In relation to student performance, it is very important to predict student performance during a pandemic which includes several things such as the main role of the teacher in successful learning using e-learning and analyzing student learning styles using e-learning during a pandemic.

Predicting student performance based on e-learning before the pandemic has been widely reported which only focuses on student performance models. However, the previous research did not describe deeper post-pandemic learning related to student performance while still focusing on the journal data set by adding a model to process the data set from the previous model, namely by using the Support Vector Machine (SVM), which is a classification algorithm for data. linear and non-linear. SVM uses non-linear *mapping to transform the initial training* data to a higher dimension (Han, 2012). This study aims to find high accuracy based on the data generated and analyzed Raisedhands and Discussion, these two variables were chosen because in my opinion they are the best activities in learning activities. While for testing the model is from the Topic which contains the subjects in the dataset.

LITERATURE REVIEW

It has been reported that the use of e-learning in an effort to predict based on student behavior to determine success in using e-learning as a learning medium has been done before and previous research that was studied related to student performance (Amrieh, 2015) used three types of research models, namely Artificial Neural Network (ANN), Decision Tree (DT), and Naïve Bayes (NB) results show that each classification algorithm introduces two classification results: (1) classification results with behavioral features (BF) and (2) classification results without behavioral features (WBF) and the classification is obtained in Table 1:

Table 1. Classification Method Results with Behavioral Features (BF) and Results without Behavioral Features (WBF)

Evaluation Measures	DT		ANN		NB	
	BF	WBF	BF	WBF	BF	WBF
Accuracy	61.3	55.6	73.8	45.8	72.5	50.4
Recall	61.3	55.6	73.8	45.9	72.5	50.4
Precision	60.9	56.2	73.9	45.2	72.7	49.6
F-Measure	60.1	53.4	73.2	44.8	71.9	49.4

It can be seen that the ANN model outperforms the others with Accuracy 73.8 with behavioral features (BF) and 55.6 without behavioral features (WBF) , 73.8 means 98 out of 133 students are classified correctly to the right class label (High, Medium and Low). And Recall gave results of 73.8 with behavioral features (BF) and 45.9 without behavioral features (WBF), 73.8 meaning that 98 students were correctly classified into the total number of cases that were not classified and the cases that were classified were correct. Precision gives results of 73.9 with behavioral features (BF) and 45.2 without behavioral features (WBF) behavioral features, 73.9 means 98 out of 133 students are classified correctly and 35 students are misclassified. While the F-Measure results are 73.2 behavioral features (BF) and 44.8 without behavioral features (WBF) (Amrieh, 2015).

While in other journals that are also related to student performance processes (Amrieh, 2016) have also reported student performance models with new feature categories, called behavioral features. The educational dataset was collected from a learning management system (LMS) called Kalboard 360. This model uses several data mining techniques to evaluate the impact of student behavior features on

*name of corresponding author



student academic performance. By using three types of research models, namely Artificial Neural Network (ANN), Decision Tree (DT), and Naïve Bayes (NB), the results obtained for each classification algorithm introduce two classification results: (1) classification results with behavioral features (BF) and (2) the results of the classification without behavioral features (WBF) and the classification are obtained in Table 2:

Table 2. Classification Method Results with Behavioral Features (BF) and Results without Behavioral Features (WBF)

Evaluation Measures	DT		ANN		NB	
	BF	WBF	BF	WBF	BF	WBF
Behavioral Features Existence						
Accuracy	75.8	55.6	79.1	57.0	67.7	46.4
Recall	75.8	55.6	79.2	57.1	67.7	46.5
Precision	76.0	56.0	79.1	57.2	67.5	46.8
F-Measure	75.9	55.7	79.1	57.1	67.1	46.4

As shown in the table, it can be seen that the ANN model outperforms other data mining techniques. The ANN model achieves an accuracy of 79.1 with behavioral features (BF) and 57.0 without behavioral features (WBF). Where Accuracy 79.1 means that 380 out of 480 students are classified correctly to the appropriate class labels (High, Medium and Low) and 100 students are classified incorrectly. For Recall, the result is 79.2 with behavioral features (BF) and 57.1 without behavioral features (WBF). Recall 79.2 means that 380 students were classified correctly to the total number of cases that were not classified and classified correctly. For Precision, the results were 79.1 with behavioral features (BF) and 57.2 without behavioral features (WBF). Precision of 79.1 means that 380 out of 480 students are classified correctly and 100 students are classified incorrectly. For the F-Measure, the results are 79.1 with behavioral features (BF) and 57.1 without behavioral features (WBF). The experimental results prove the strong influence of learner behavior on student academic achievement (Amrieh, 2016).

The conclusion from these two studies is that the ANN model outperforms other models such as DT and NB. If you look at the resulting accuracy, it is also quite high and for this reason, the latest research using Support Vector Machine (SVM) is expected to be able to produce higher classification or accuracy values as a comparison to predict student learning performance during the Covid-19 Pandemic.

METHODS

This research method uses quantitative methods. Quantitative research is research that uses data in the form of numbers as a tool to find out what you want to know (Samsu, 2017). Quantitative methods are methods that use the theory of measurement, observation, and testing (Emzir, 2010). This study uses explanatory research, which is research that tests a theory or hypothesis that aims to explain the relationship or influence of one variable on other variables (Hartono, 2018). This study will conduct research to determine predictive student learning outcomes based on their behavior in using bold learning during the Covid-19 pandemic.

This study summarizes some of the research phases that will be carried out in each phase. A diagram of the research steps is shown in Figure 1, source taken from Gopinath, 2015 (https://www.researchgate.net/publication/281521947_Research_Proposal-Flow_Chart-New_Research_Scholars-Reg.)

Research Proposal Flow Chart

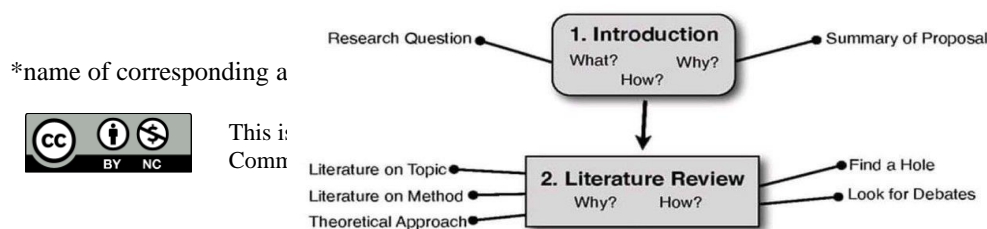


Figure 1. Research Stages (Gopinath (2015))

The following are the research steps along with explanations for each research step carried out:

a. Identification of problems

Problem identification is a number of successfully identified problems describing the context of the problem or the problem placement to be investigated and the scope of the issue in general. (Sugiyono, 2017).

The first step in the research is the identification of problems, at this stage the researchers elaborate backgrounds related to prediction of student performance based on behavior using E-Learning during the Covid-19 pandemic.

b. Literature review

Library assessment is an important thing in a research because with a library survey, the researcher will know more about the research that has been done, which is relevant and relevant to the research to be conducted, because So researchers will be thorough researchers and be able to understand and know previous research problems. (Sugiyono, 2016).

The next step of the research was library assessment, in which researchers collected and studied research references related to behavior-based prediction of student learning outcomes. students learning online during the Covid-19 pandemic. Preparation of Research Model

The next research stage is the proposed research model and research hypothesis which refers to predicting student performance based on behavior using E-Learning during the Covid-19 Pandemic.

c. Research Model

Research methodology is essentially the scientific method of collecting data for specific purposes and uses. There are four keywords to consider, which are scientific method, data, purpose, and utility. (Sugiyono, 2018).

The next step of the research is to propose a research model and research hypothesis to predict student learning outcomes based on online learning usage behavior during the Covid-19 pandemic.

d. Data collection

According to (Sugiyono, 2018), there are two main factors that influence it: the quality of the research data, i.e. the quality of the research instruments and the quality of data collection. Data collection can be done in different contexts, from different sources, and in different ways. The next

*name of corresponding author



step is to collect the data from the previous research by updating it with the same dataset but describing the data using different techniques than before using the Support Vector Machine (SVM) method, Kernel usage: Radial, polynomial and linear.

e. Machine Learning Model Training

The model training step uses the Support Vector Machine (SVM) method, using the kernel: Radius, Polynomial and Linear to find the highest accuracy of three kernels with variable classification Discussions and Raisedhands on the Topic as reference to predict student learning outcomes based on online learning usage behavior.

f. Testing and Data Analysis

Data analysis is the activity after collecting data from all respondents or other data sources. (Sugiyono, 2019).

The next step is to test the data that was performed using the Support Vector Machine (SVM) method to measure the performance of the kernel model: Radial, Polynomial and Linear are used to achieve higher accuracy than previous research.

g. Results and Conclusions

The conclusion is that qualitative research results may answer a pre-conceived problem formulation, but this may not be the case, because as suggested, problem formulation and problem formulation In research, quality is still temporary. and will develop upon completion of fieldwork. (Sugiyono, 2018).

The final step is to explain the results of the research that has been conducted and draw conclusions from the analyzed data to predict student learning outcomes based on online learning usage behavior. during the Covid -19 pandemic. Based on the conclusions of the study results, the researchers then make suggestions for improvements to stakeholders and suggest further research.

RESULTS

For this dataset, it is a collection of educational data collected from a learning management system "LMS". called Kalboard 360. Calboard 360 is a multi-agency LMS, designed to facilitate learning through the use of innovative technology. The system provides users with synchronized access to educational resources from any device with an Internet connection ("Kalboard 360 e-learning system", 2000) also accessible via <https://elearningindustry.com/directory/elearning-software/kalboard360>.

Data is collected using a learning tracking tool, called the Experience API (xAPI). XAPI is a Training and Learning Architecture (TLA) component that tracks learning progress and learning actions like reading articles or watching training videos. The Experience API helps learning providers identify which students, activities, and objects describe the learning experience. The dataset includes 480 student records and 16 features. Features are classified into three main categories: (1) Demographic characteristics such as gender and nationality. (2) Characteristics of academic training such as educational level, grade level, and department. (3) Behavioral characteristics such as raising hands in class, opening resources, responding to parent surveys, and school satisfaction. The dataset includes 305 males and 175 females. The students come from many different backgrounds such as 179 students from Kuwait, 172 students from Jordan, 28 students from Palestine, 22 students from Iraq, 17 students from Lebanon, 12 students from from Tunisia, 11 students from Saudi Arabia, 9 students from Egypt, 7 students from Egypt. student. students from Syria, 6 students from the United States, Iran and Libya, 4 students from Morocco and one student from Venezuela. The dataset was collected over two semesters of research: 245 student records were collected in the first semester and 235 students registered in the second semester. The dataset also includes school attendance characteristics, as students are classified into two categories based on their number of days absent: 191 students were absent from school for more than 7 days and 289 students were absent from school for less than 7 days. This data group also includes a new feature category; This trait is part of parents in the educational process. The Parental Involvement feature has two sub-features: Parent feedback survey and parent satisfaction at school. There were 270 parents who responded to the survey and 210 parents who did not respond, 292 parents were satisfied with the school and 188 parents were dissatisfied (Amrieh, 2015).

*name of corresponding author



Data taken from X-Api Edu Data is processed using SVM. support-vector machine (SVM) is a classification algorithm for linear and non-linear data. SVM uses non-linear *mapping to transform the initial training* data to a higher dimension. In a way that has never been used before. The data consists of 17 variables and 40 observations, of which 17 of the 17 variables are character variables and 4 are numeric. While the summary of the 17 variables is detailed in Table 3:

Table 3. Classification of Data Variables and Data Types in X-Api-Edu-Data

No.	Data Variables	Data Type
1	Gender	Character
2	NationalITy	Character
3	PlaceofBirth	Character
4	StageID	Character
5	GradeID	Character
6	SectionID	Character
7	Topic	Character
8	Semester	Character
9	Relations	Character
10	Raisedhands	Numeric
11	VisITedResources	Numeric
12	AnnouncementsView	Numeric
13	Discussion	Numeric
14	ParentAnsweringSurvey	Character
15	Parentschool Satisfaction	Character
16	StudentAbsenceDays	Character
17	Class	Character

Next is to make variable plots from the data, the data I use for plots are *Raisedhands* and *Discussion*, these two variables were chosen because in my opinion they are the best activities in learning activities. While for testing the model is from *the Topic* which contains the subjects in the dataset, the details of the *Topic* are in Table 4:

Table 4. Topic Classification and Data Types in X-Api-Edu-Data

No.	Topic	Data Type
1	Arabic	Character

*name of corresponding author



No.	Topic	Data Type
2	Biology	Character
3	Chemistry	Character
4	English	Character
5	French	Character
6	Geology	Character
7	History	Character
8	IT	Character
9	Math	Character
10	Quran	Character
11	Science	Character
12	Spanish	Character

With the plot results get the mapping results in Figure 2:

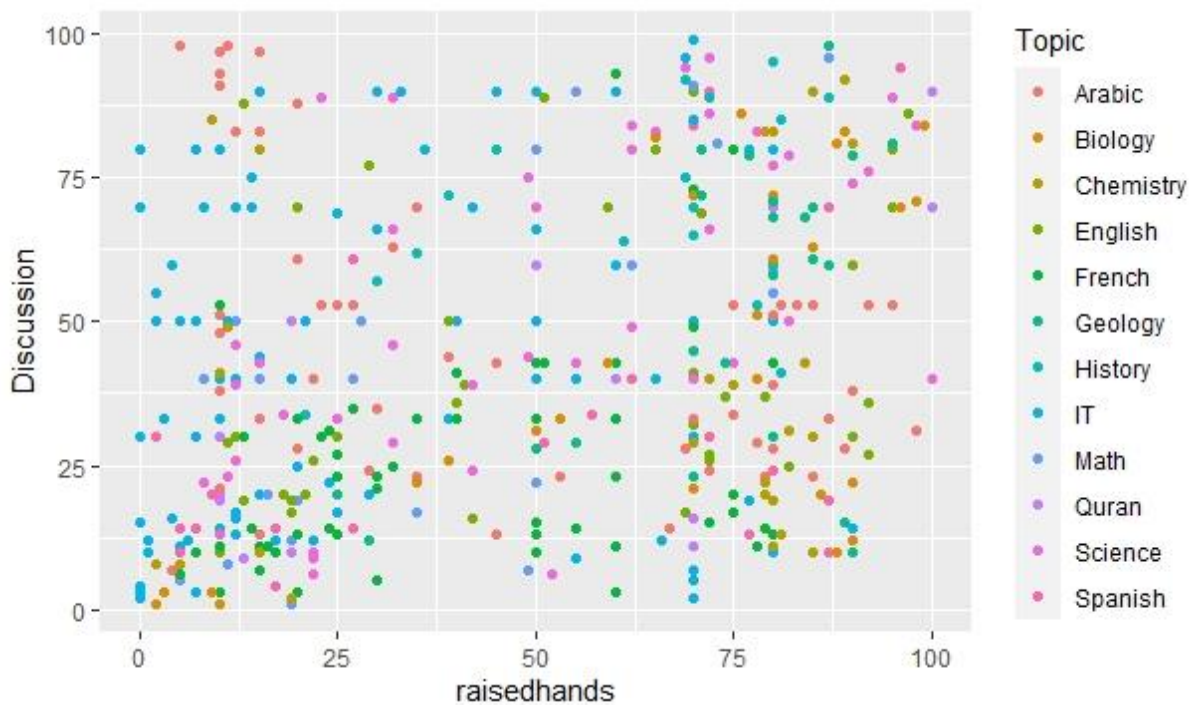


Figure 2. Discussion and Raisedhands Classification on Topic Variables on X-API-Edu-Data

From Figure 2 the plot with the Horizontal line is Raisedhands and the Vertical line is Discussion where the colored dots represent each subject in the Topic category. From the image the spread of each subject is based on the activities of Raisedhãnds and Discussion tends to be balanced, but the most dominant is on the IT subject, which is the subject that corresponds to the use of E-Learning in supporting learning activities. Next is a model test on the dataset in Figure 3 :

```
call:
svm(formula = Topic ~ ., data = df)
```

*name of corresponding author



<p>Parameters: SVM-Type: C-classification SVM-Kernel: radial costs: 1</p> <p>Number of Support Vectors: 453 (75 21 59 50 45 22 25 59 19 30 24 24)</p> <p>Number of Classes: 12</p> <p>Levels: Arabic Biology Chemistry English French Geology History IT Math Quran Science Spanish</p>

Figure 3. Classification of Discussion and Raisedhands Parameters on Topic Variables in X-API-Edu-Data

The next step is to make a confusion matrix or classification on the Topic variable, the observations can be explained in Table 5:

Table 5. Observation of Confusion Matrix and Classification on Topic Variables

Predictions	Arabic	Biology	Chemistry	English	French	Geology	History	IT	Math	Quran	Science	Spanish
Arabic	36	0	3	0	0	0	0	0	3	2	4	21
Biology	0	29	0	0	0	0	0	2	1	8	6	0
Chemistry	0	0	10	0	0	3	2	0	0	0	0	0
English	0	0	0	26	0	0	0	1	1	0	1	0
French	19	0	0	2	57	0	0	6	1	0	5	0
Geology	0	0	11	0	0	21	10	0	0	0	0	1
History	0	0	0	0	0	0	0	0	0	0	0	0
IT	4	0	0	16	4	0	3	83	13	8	6	0
Math	0	0	0	0	0	0	0	0	0	0	0	0
Quran	0	0	0	0	0	0	0	0	0	2	0	0
Science	0	1	0	1	4	0	4	3	2	2	29	0
Spanish	0	0	0	0	0	0	0	0	0	0	0	3

From the observation data of the confusion matrix and the classification of the topic variables, the accuracy values are obtained in Figure 4:

<p>Overall Statistics</p> <p>Accuracy : 0.6167</p>
--

*name of corresponding author



95% CI : (0.5715, 0.6604)
 No Information Rate : 0.1979
 P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.5604

McNemar's Test P-Value : NA

Figure 4. Accuracy Value of Observation Confusion Matrix and Classification of Topic Variables

DISCUSSIONS

In Figure 4, an accuracy score of 0.6167 is shown, which shows a fairly high score, but further research will use three Radial, Polynomial and Linear kernels to look for higher accuracy as a variation of data modeling results that have never been used in previous research.

Next is to create a Radial kernel with the syntag results in Figure 5:

```
call:
svm(formula = Topic ~ ., data = df, type = "C-classification", kernel = "radial",
cost = 1, scale = FALSE)
```

```
Parameters:
SVM-Type: C-classification
SVM-Kernel: radial
costs: 1
```

Number of Support Vectors: 479

```
> summary (model _svm_radial)
```

```
call:
svm(formula = Topic ~ ., data = df, type = "C-classification", kernel = "radial",
cost = 1, scale = FALSE)
```

```
Parameters:
SVM-Type: C-classification
SVM-Kernel: radial
costs: 1
```

Number of Support Vectors: 479

```
( 95 21 58 51 45 22 25 65 19 30 24 24 )
```

Number of Classes: 12

```
Levels:
```

```
Arabic Biology Chemistry English French Geology History IT Math Quran Science Spanish
```

Figure 5. Data Processing Process with Kernel Radial Parameter Discussion and Raisedhands on Topic Variables in X-API-Edu-Data

And next is to look at the classification of each category on the Topic variable using the Radial kernel with the results in Table 6:

Table 6. Classification Results of Data Processing Processes with Kernel Radial Parameter Discussion and Raisedhands on Topic Variables in X-API-Edu-Data

*name of corresponding author



Predictions	Arabic	Biology	Chemis	English	French	Geolog	History	IT	Math	Quran	Science	Spanish
Arabic	57	1	1	0	0	0	0	0	0	0	0	0
Biology	0	28	1	0	0	0	0	0	0	0	0	0
Chemistry	0	0	19	0	0	0	0	0	0	0	0	0
English	0	0	0	42	0	0	0	0	0	0	0	0
French	1	1	0	1	65	0	0	0	1	1	0	0
Geology	0	0	0	0	0	24	0	0	0	0	0	0
History	0	0	0	0	0	0	17	0	0	0	0	0
IT	1	0	3	2	0	0	0	95	2	2	1	1
Math	0	0	0	0	0	0	0	0	18	0	0	0
Quran	0	0	0	0	0	0	0	0	0	18	0	0
Science	0	0	0	0	0	0	0	0	0	0	50	0
Spanish	0	0	0	0	0	0	0	0	0	0	0	24

From these data the accuracy value is obtained in Figure 6:

<p>Overall Statistics</p> <p>Accuracy : 0.9521 95% CI : (0.929, 0.9694) No Information Rate : 0.1979 P-Value [Acc > NIR] : < 2.2e-16</p> <p>Kappa : 0.9459</p> <p>Mcnemar's Test P-Value : NA</p>

Figure 6. Accuracy results using the Kernel Radial Parameter Discussion and Raisedhands on Topic Variables in X-Api-Edu-Data

Next is to create a Polynomial kernel with the syntag results in Figure 7:

<p>call: svm(formula = Topic ~ ., data = df, type = "C-classification", kernel = "polynomial", cost = 1, scale = FALSE)</p> <p>Parameters: SVM-Type: C-classification SVM-Kernel: polynomial costs: 1 degrees: 3 coef. 0: 0</p> <p>Number of Support Vectors: 376</p> <p>> summary (model _svm_radial) call: svm(formula = Topic ~ ., data = df, type = "C-classification", kernel = "polynomial", cost = 1, scale = FALSE)</p>
--

*name of corresponding author



<p>Parameters: SVM-Type: C-classification SVM-Kernel: polynomial costs: 1 degrees: 3 coef. 0: 0</p> <p>Number of Support Vectors: 376 (67 20 50 35 35 21 20 45 16 23 22 22)</p> <p>Number of Classes: 12</p> <p>Levels: Arabic Biology Chemistry English French Geology History IT Math Quran Science Spanish</p>

Figure 7. Data Processing Process with Kernel Polynomial Parameter Discussion and Raisedhands on Topic Variables in X-API-Edu-Data

And next is to look at the classification of each category on the Topic variable using the Polynomial kernel with the results in Table 7:

Table 7. Classification Results of Data Processing Processes with Kernel Polynomial Parameter Discussion and Raisedhands on Topic Variables in X-API-Edu-Data

Predictions	Arabic	Biology	Chemistry	English	French	Geology	History	IT	Math	Quran	Science	Spanish
Arabic	59	0	0	0	0	0	0	0	0	0	0	0
Biology	0	30	0	0	0	0	0	0	0	0	0	0
Chemistry	0	0	24	0	0	0	0	0	0	0	0	0
English	0	0	0	45	0	0	0	0	0	0	0	0
French	0	0	0	0	65	0	0	0	0	0	0	0
Geology	0	0	0	0	0	24	0	0	0	0	0	0
History	0	0	0	0	0	0	19	0	0	0	0	0
IT	0	0	0	0	0	0	0	95	1	0	0	0
Math	0	0	0	0	0	0	0	0	20	0	0	0
Quran	0	0	0	0	0	0	0	0	0	22	0	0
Science	0	0	0	0	0	0	0	0	0	0	51	0
Spanish	0	0	0	0	0	0	0	0	0	0	0	25

From these data, the accuracy value is obtained in Figure 8:

*name of corresponding author



Overall Statistics
Accuracy : 0.9979
95% CI : (0.9884, 0.9999)
No Information Rate : 0.1979
P-Value [Acc > NIR] : < 2.2e-16
Kappa : 0.9977
Mcnemar's Test P-Value : NA

Figure 8. Accuracy results using Kernel Polynomial Parameter Discussion and Raisedhands on Topic Variables in X-API-Edu-Data

Next is to create a Linear kernel with the syntag results in Figure 9:

```
call:
svm(formula = Topic ~ ., data = df, type = "C-classification", kernel = "linear",
cost = 1, scale = FALSE)

Parameters:
SVM-Type: C-classification
SVM-Kernel: linear
costs: 1

Number of Support Vectors: 396

> summary(_svm_linear model)
call:
svm(formula = Topic ~ ., data = df, type = "C-classification", kernel = "linear",
cost = 1, scale = FALSE)

Parameters:
SVM-Type: C-classification
SVM-Kernel: linear
costs: 1

Number of Support Vectors: 396

( 55 21 54 48 39 22 19 44 19 29 23 23 )

Number of Classes: 12

Levels:
Arabic Biology Chemistry English French Geology History IT Math Quran Science Spanish
```

Figure 9. Data Processing Process with Kernel Linear Parameter Discussion and Raisedhands on Topic Variables in X-API-Edu-Data

And next is to look at the classification of each category on the Topic variable using the Linear kernel with the results in Table 8:

*name of corresponding author



Table 8. Classification Results of Data Processing Processes with Kernel Linear Parameter Discussion and Raisedhands on Topic Variables in X-API-Edu-Data

Predictions	Arabic	Biology	Chemistry	English	French	Geology	History	IT	Math	Quran	Science	Spanish
Arabic	49	0	0	0	2	0	0	0	2	0	1	6
Biology	0	30	0	0	0	0	0	1	0	1	0	0
Chemistry	0	0	19	0	0	0	2	0	0	0	0	0
English	0	0	0	41	0	0	0	3	1	1	0	0
French	9	0	0	0	62	0	0	1	0	0	0	0
Geology	0	0	3	0	0	22	1	0	0	0	0	0
History	0	0	2	0	0	2	14	1	1	0	1	0
IT	1	0	0	1	0	0	1	86	5	1	3	0
Math	0	0	0	2	0	0	0	1	11	2	0	0
Quran	0	0	0	0	0	0	0	1	0	17	1	0
Science	0	0	0	1	1	0	1	1	1	0	45	0
Spanish	0	0	0	0	0	0	0	0	0	0	0	19

From these data, the accuracy value is obtained in Figure 11:

<p>Overall Statistics</p> <p>Accuracy : 0.8646 95% CI : (0.8307, 0.8939) No Information Rate : 0.1979 P-Value [Acc > NIR] : < 2.2e-16</p> <p>Kappa : 0.8475</p> <p>Mcnemar's Test P-Value : NA</p>

Figure 11. Accuracy results using Kernel Linear Parameter Discussion and Raisedhands on Topic Variables on X-API-Edu-Data

From Table 5 to Table 8, using three tested kernels, namely the Radial, Polynomial and Linear Kernels, it appears that IT variables are more dominant than the other Topic variables, due to the use of a digital-based E-Learning learning model that students must master to support school learning activities as a form of modernization in receiving lessons and when learning cannot be done directly at school, then another alternative is to use E-learning so that the learning process can run well to predict student performance.

From the three kernels, the Radial, the Polynomial and the Linear, we can conclude that the highest accuracy is 0.9979. Then the Polynomial model has been suited to the conclusion that the accuracy of Parameter Discussion and Raisedhands on Variable Topic on X-Fire-Edu-Data to illustrate student performance processes based on behavior using e-learning methods during the Covid-19 pandemic is effectively done to predict student performance.

*name of corresponding author



CONCLUSION

The COVID-19 pandemic has drastically reshaped numerous sectors, including education, compelling a shift from conventional face-to-face instruction to autonomous or online learning via diverse digital platforms to ensure continuity in education. A viable strategy to safeguard the quality of e-learning is to harness behavioral data derived from these e-learning platforms. The primary objective of this research is to devise a fresh data model suitable for the educational landscape reshaped by the pandemic. Previous research has delved into predicting student performance in e-learning contexts pre-pandemic, concentrating mainly on student performance models. However, prior work, such as that of Amrieh (2015), which employed three models, namely Artificial Neural Network (ANN), Decision Tree (DT), and Naïve Bayes (NB), doesn't sufficiently address post-pandemic learning intricacies related to student performance. This study thus proposes to augment the dataset of the earlier model with a new model that employs the Support Vector Machine (SVM) method, not previously utilized in this context. The dataset used for this study was sourced from X-API-Edu-Data. It contains 17 variables and 40 observations, all 17 of which are character variables, and four are numeric. The analyzed data includes variables such as Raisedhands and Discussion, which were selected due to their integral role in effective learning and association with e-learning-based student performance. The model was evaluated using the Topic variable, which represents the subjects within the dataset. The research findings indicate superior accuracy compared to previous studies. Among the three SVM kernels tested - Radial, Polynomial, and Linear, the Polynomial kernel demonstrated the highest accuracy, scoring an impressive 0.9979.

REFERENCES

- Amrieh, EA, Hamtini, T., & Aljarah, I. (2016). Mining Educational Data to Predict Student's academic Performance using Ensemble Methods. *International Journal of Database Theory and Application*, 9(8), 119-136.
- Amrieh, EA, Hamtini, T., & Aljarah, I. (2015, November). Preprocessing and analyzing educational data sets using X-API for improving student's performance. In *Applied Electrical Engineering and Computing Technologies (AEECT), 2015 IEEE Jordan Conference on* (pp. 1-5). IEEE.
- Emzir. 2010. *Educational Research Methodology: Quantitative and Qualitative*. Jakarta: Rajawali Press.
- Hartono, jogiyanto. 2018. *Methods of data collection and analysis techniques (ANDI, 2018)* page 31
- J, Han (2012). *Data mining: Data mining concepts and techniques*. Morgan Kaufmann.
- Kyriacou, C. (2009) *Effective Teaching in Schools: Theory and Practice*. Third Edition. Delta Place, Cheltenham, UK: Nelson Thornes Ltd
- Maswan & Khoiril Muslimin. (2017). *Educational Technology Application of Systematic Learning*. Yogyakarta: Student Libraries.
- R., Dr. Gopinath. (2015). *Research Proposal-Flow Chart-New Research Scholars-Reg*
- Sagala, Saiful. 2010. *The Concept and Meaning of Learning to Help Solve Learning and Teaching Problems*. Bandung: Alfabet
- Sahfitri, V., & Ulfa, M. (2015). Evaluation of the Usability of the E-Learning System as an Application Supporting the Learning Process in Higher Education Using the USE Questionnaire. *MATRIK Scientific Journal*, 17(1), 53-66.
- Samsu, Research University, Jambi: STS Press, 2011, p. 4.
- Sugiyono. (2016). *Quantitative, Qualitative Research Methods, R & D*. Bandung: CV Alfabet.
- Sugiyono. (2017). *Quantitative, Qualitative Research Methods, R & D*. Bandung: CV Alfabet.
- Sugiyono. (2018). *Quantitative Research Methods, Qualitative Intelligence, and R&D*, Alfabet Publisher, Contents
- Sugiyono. (2019). *Methodology of Quantitative and Qualitative Research and R&D*. Bandung: Alfabet.
- Wulandari A., et al. 2020. The Relationship between Individual Characteristics and Knowledge on Prevention of Coronavirus Disease 2019 in Communities in South Kalimantan. *Journal of Indonesian Public Health*, 15(1):42-46.

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.