Predicting Students' Engagement in Online Courses Using Machine Learning

Jawaher Alsirhani† and Khalaf Alsalem††,
Jawaheerr1995@gmail.com kosalem@ju.edu.sa
Jouf University, Sakaka Aljouf, Saudi Arabia

Summary
No one denies the importance of online courses, which provide a very important alternative, especially for students who have jobs that prevent them from attending face-to-face in traditional classes; Engagement is one of the most important fundamental variables that indicate the course's success in achieving its objectives. Therefore, the current study aims to build a model using machine learning to predict student engagement in online courses. An online questionnaire was prepared and applied to the students of Jouf University in the Kingdom of Saudi Arabia, and data was obtained from the input variables in the questionnaire, which are: specialization, gender, academic year, skills, emotional aspects, participation, performance, and engagement in the online course as a dependent variable. Multiple regression was used to analyze the data using SPSS. Kegel was used to build the model as a machine learning technique. The results indicated that there is a positive correlation between the four variables (skills, emotional aspects, participation, and performance) and engagement in online courses. The model accuracy was very high 99.99%, This shows the model's ability to predict engagement in the light of the input variables.

Keywords:
Machine learning - engagement - prediction - emotional aspects – participation.

1. Introduction

The importance of online learning is becoming clearer every day, whether it is a complement to traditional learning or an alternative to it. Earlier, some education systems questioned the ability of online courses to develop academic achievement, acquire skills and attitudes, and achieve the joy of learning. Now, all universities in developed and developing countries alike are planning to make online education an essential part of their educational system with the aim of expanding access. and efficiency [1].

There are many studies that have been concerned with studying the impact of students' engagement in online education courses on improving achievement and student learning outcomes, especially with regard to their involvement in educational activities and tasks [2] [3] [4].

The problem of the research is to determine the factors affecting students' engagement in online courses and to predict the extent of students' engagement in online courses through the use of machine learning techniques, analysis and interpretation of these results to identify the reasons that encourage students to students' engagement in online courses or the reasons that lead to non- students' engagement, to has decision-making to ensure greater students' engagement in online courses in the future.

2. Literature Review

Students' engagement is defined as the effort the student makes in learning processes for specific course content [5]. Encouraging students to engagement and predicting the level of engagement definitely contributes to improving learning outcomes and directs strategies that enable the management of educational institutions to overcome the barriers that prevent students from continuing their learning through online courses. Accordingly, the researchers were interested in the students' engagement as one of the variables and reasons for the drop in final grades for students, leaving the course, or failing [6].

Reeve et al., [7] refers to the student's engagement as the student's emotionally active behaviors while performing the tasks within the online course. Manwaring et al. [8] have dealt with three levels of engagement so that they can be analyzed within the online course, which are the activity level, the course level, and the institutional level.

Many studies have been conducted that have applied predictive machine learning algorithms and models, where Husain et al., [9] study applied supervised machine learning algorithms to predict students' engagement and interaction within the online course in virtual learning environments, while Motz et al., [4] study used a logistic regression model to predict student engagement in the online course within the electronic content management system. Cocea and Weibelzahl [10] also developed a predictive model for students' engagement through behavioral factors within the electronic content management system. As for Sadeque et al., [11] study, it used logistic regression as a model to predict students' engagement through discussion forums, and to monitor the number of responses, time spent by students in the forum, and forms of interaction within discussion forums. Whereas, Sharma et al., [12] study relied on emotional

Manuscript received September 5, 2022
Manuscript revised September 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.9.24

aspects by observing facial signs to predict students' engagement in the online course. Furthermore, Calvo and D'Mello, [13] note that detection systems that integrate data from different agents have been widely endorsed but rarely implemented.

as a term, is not well defined. Kuh [14] sees engagement as “the time and energy students devote to educationally sound activities” (p. 25). Handelsman, Briggs, Sullivan, and Towler’s [15] measure of traditional classroom student engagement. They found four factors illustrating how students devote time and energy in the classroom: skills engagement (keeping up with readings, putting forth effort); emotional engagement (making the course interesting, applying it to their own lives); participation/interaction engagement (having fun, participating actively in small group discussions); and performance engagement (doing well on tests, getting a good grade) [15]. They see student engagement as containing both affective and behavioral components.

Combining social constructivist notions of learning, the CoI model, and previous incarnations of engagement in the traditional classroom, a description of online student engagement emerges: Engagement involves students using time and energy to learn materials and skills, demonstrating that learning, interacting in a meaningful way with others in the class (enough so that those people become “real”), and becoming at least somewhat emotionally involved with their learning (i.e., getting excited about an idea, enjoying the learning and/or interaction). Engagement is composed of individual attitudes, thoughts, and behaviors as well as communication with others. Student engagement is about students putting time, energy, thought, effort, and, to some extent, feelings into their learning. Therefore, the OSE attempts to measure what students do (actively and in their thought processes) as well as how they feel about their learning and the connections they are making with the content, the instructor, and other students in terms of skills, participation, performance, and emotion [16].

The current study benefit from Online Student Engagement Scale (OSE) used in (Dixson, [16]). The OSE attempts to measure what students do (actively and in their thought processes) as well as how they feel about their learning and the connections they are making with the content, the instructor, and other students in terms of skills, participation, performance, and emotion.

3. Methodology

A predictive model of students’ engagement in online courses will be developed using machine learning outcomes based on a set of input variables, namely skills, emotional aspects, participation in activities and content, student performance in the web-based course, in addition to the variables of educational level, gender, and specialization.

As stated earlier, while there are strong theoretical foundations and a very useful model for engagement, student engagement,

The study tools will be applied to a sample of students at the Jouf University campus in Saudi Arabia.

The population in the current research will be all students at the Jouf University in Saudi Arabia. The sample size consist of 263 students who response to the online questionnaire.

The study used online questionnaire, the questionnaire was divided into two parts. The first is the demographic data, which is the name (optional), gender, college, major, and academic year; The second part relates to four axes dealing with the basic variables, namely skills (7 statements), emotional aspects (7 statements), participation (7 statements), and performance (6 statements). the response of the target population (Jouf University students) to the statements will be according to the Pentagonal Likert Scale, to determine the degree of their agreement with it gradually: always, often, sometimes, rarely, never.

The Kaggel program was used to build a machine learning model, where the data obtained from applying the questionnaire to the students was divided into two parts: training, which represents 80% of the data, and the testing, which represents 20% of the data.

The researcher performed a set of procedures to build the model, which are as follows:

3.1 Import libraries

```
import math
import seaborn as sns
sns.set(style="whitegrid", color_codes=True)

from wordcloud import WordCloud, STOPWORDS

import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import numpy as np # linear algebra
import matplotlib
import matplotlib.pyplot as plt
import sklearn
%matplotlib inline
import matplotlib.pyplot as plt
plt.rcParams["figure.figsize"] = [16, 12]
# Input data files are available in the "../input/" directory.
# For example, running this (by clicking run or pressing Shift+Enter) will list the files in the input directory

from subprocess import check_output
print(check_output(["ls", "../input/"]).decode("utf8"))

# Any results you write to the current directory are saved as output.
filenames = check_output(["ls", "../input/"]).decode("utf8").strip()
# helpful character encoding module
import chardet

# set seed for reproducibility
np.random.seed(0)
```

Fig. 1 Import libraries for predictive model.

3.2 Read the dataset

```
[5]: data = pd.read_excel('../input/dataset5/dataset5.xls')
```

```
data.head()
```

[6]:

	Gender	Specialization	Academic year	Skills	Emotional aspects	Participation	Performance	Engagement
0	2	2	4	4.14	4.14	4.29	4.67	4.31
1	2	1	3	4.29	3.00	3.57	4.50	3.84
2	1	1	1	5.00	5.00	5.00	5.00	5.00
3	2	1	2	5.00	4.86	3.86	5.00	4.68
4	2	2	4	4.71	5.00	5.00	5.00	4.93

Fig. 2 Read the dataset from excel file.

3.3 Data processing

- Check the null values in each column and review the data type for each column
- If categorical values are changed to numeric because regression deals with numeric values

```
[7]: data.isnull().sum()
```

```
[7]: Gender               0
     Specialization       0
     Academic year        0
     Skills               0
     Emotional aspects    0
     Participation        0
     Performance          0
     Engagement           0
     dtype: int64
```

```
data.dtypes
# no categorical values. all values are numeric
```

```
[8]: Gender                 int64
     Specialization         int64
     Academic year          int64
     Skills               float64
     Emotional aspects    float64
     Participation        float64
     Performance          float64
     Engagement           float64
     dtype: object
```

Fig. 3 Data processing.

3.4 Data cleansing

```
[9]: data.dropna(inplace = True)
```

```
data.describe()
```

[10]:

	Gender	Specialization	Academic year	Skills	Emotional aspects	Participation	Performance	Engagement
count	263.000000	263.000000	263.000000	263.000000	263.000000	263.000000	263.000000	263.000000
mean	1.711027	1.338403	3.041825	4.446388	4.223992	4.263954	4.489125	4.355779
std	0.454150	0.474069	1.092200	0.643121	0.775956	0.742254	0.601860	0.593151
min	1.000000	1.000000	1.000000	1.290000	1.860000	1.430000	1.170000	1.760000
25%	1.000000	1.000000	2.000000	4.140000	3.710000	3.860000	4.000000	4.015000
50%	2.000000	1.000000	3.000000	4.570000	4.430000	4.430000	4.670000	4.520000
75%	2.000000	2.000000	4.000000	5.000000	5.000000	5.000000	5.000000	4.825000
max	2.000000	2.000000	4.000000	5.000000	5.000000	5.000000	5.000000	5.000000

Fig. 4 Delete empty values and display a description of the data.

3.5 Determine the input and output of the variables

- X
- Iv= Gender', 'Specialization', 'Academic year', 'Skills',
- 'Emotional aspects', 'Participation', 'Performance'
- Y= engagement

```
[16]: X.shape
[16]: (263, 7)

[17]: y.shape
[17]: (263, 1)

[18]: X.columns
[18]: Index(['Gender', 'Specialization', 'Academic year', 'Skills',
             'Emotional aspects', 'Participation', 'Performance'],
            dtype='object')

y.columns
```

Fig. 5 Determine the input and output of the variables.

3.6 Import regression libraries

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error
import seaborn as sns
```

Fig. 6 Import regression libraries.

3.7 The separated data for training and testing

```
# splitting the data
x_train, x_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)
#shapes of splitted data
print("X_train:",x_train.shape)
print("X_test:",x_test.shape)
print("Y_train:",y_train.shape)
print("Y_test:",y_test.shape)
```

```
X_train: (210, 7)
X_test: (53, 7)
Y_train: (210, 1)
Y_test: (53, 1)
```

Fig. 7 Separated data for training and testing.

```
# creating an object of LinearRegression class
LR = LinearRegression()
# fitting the training data
LR.fit(x_train,y_train)

y_prediction =  LR.predict(x_test)
#y_prediction
```

+ Code   + Markdown

Fig. 8 Prediction model.

4. Results

The researcher will follow in presenting the results in two stages as stated in the research methodology, where the prediction was made through the application of the statistical method (linear regression) through the SPSS statistical analysis program in order to predict the impact of the seven factors, which are the input variables (specialization, gender, academic year, skills, aspects affectivity, participation, performance) on the dependent variable (engagement). Then compare these results with what was shown by the machine learning prediction model that also used the regression function and the training and testing data to predict students' engagement in courses via the web.

4.1 SPSS results

1. **Standard linear regression between a skill variable and online course engagement:**

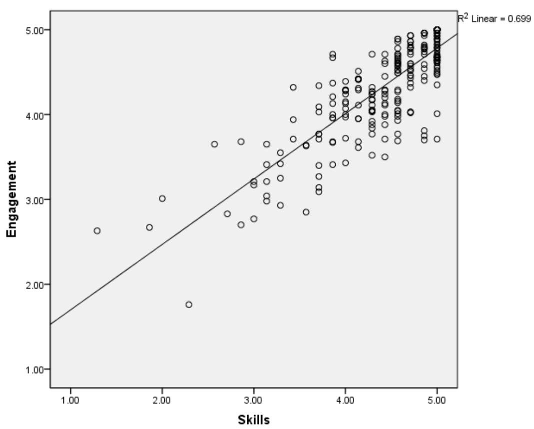


Fig. 9 Relation between the skills variable and online course engagement.

Fig. 9 illustrate a standard linear regression analysis was used, which declare through the previous graph that there is a strong positive correlation between skills and engagement, where the correlation coefficient (0.699) that is indicative of a significant correlation that predicts increased engagement in online courses whenever students perform skills in a questionnaire.

2. **Standard linear regression between emotional aspects variable and online course engagement:**

Fig. 10 illustrate a standard linear regression analysis was used, which declare through the previous graph that there is a strong positive correlation between emotional aspects and engagement, where the correlation coefficient (0.705) that is indicative of a significant correlation that predicts increased engagement in online courses whenever students has emotional aspects in a questionnaire.

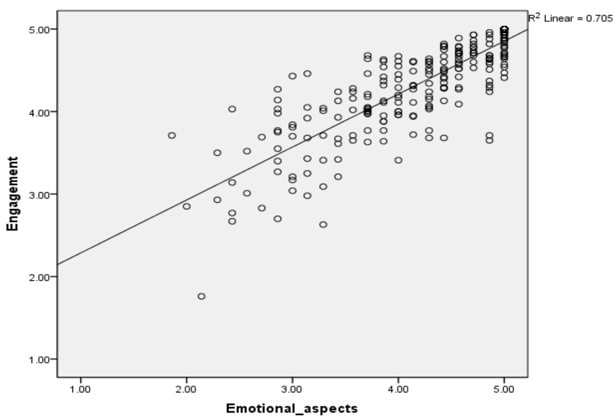


Fig. 10 Relation between the emotional aspects and online course engagement.

3. Standard linear regression between participation variable and online course engagement:

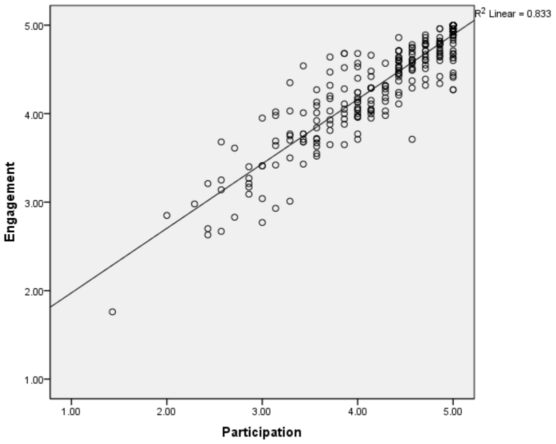


Fig. 11 Relation between the participation variable and online course engagement.

Fig. 11 illustrate a standard linear regression analysis was used, which declare through the previous graph that there is a strong positive correlation between participation and engagement, where the correlation coefficient (0.833) that is indicative of a significant correlation that predicts increased engagement in online courses whenever students perform participation statements in a questionnaire.

4. Standard linear regression between performance variable and online course engagement:

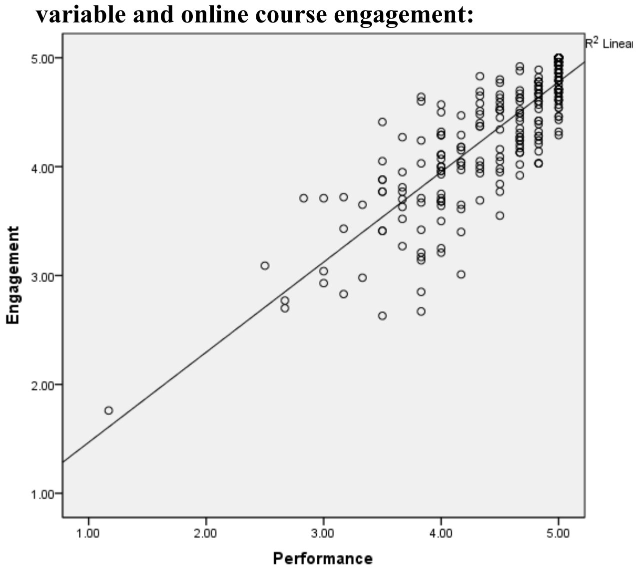


Fig. 12 Relation between the performance variable and online course engagement.

Fig. 12 illustrate a standard linear regression analysis was used, which declare through the previous graph that there is a strong positive correlation between performance and engagement, where the correlation coefficient (0.705) that is indicative of a significant correlation that predicts increased engagement in online courses whenever students perform performance statements in a questionnaire.

5. Standard linear regression between all variables (Gender – Specialization – Academic year – Skills – Emotional aspects – Participation - Performance) and online course engagement:

Table 1: Correlations

		Engagement	Gender	Specialization
Pearson Correlation	Engagement	1.000	-.040-	.106
	Gender	-.040-	1.000	-.040-
	Specialization	.106	-.040-	1.000
	Academic_year	-.146-	.155	.135
	Skills	.836	.042	.071
	Emotional_aspects	.840	-.114-	.148
	Participation	.913	-.077-	.064
	Performance	.840	.038	.073
Sig. (1-tailed)	Engagement	.	.258	.043
	Gender	.258	.	.257
	Specialization	.043	.257	.
	Academic_year	.009	.006	.014
	Skills	.000	.248	.125
	Emotional_aspects	.000	.032	.008
	Participation	.000	.106	.149
	Performance	.000	.271	.119
N	Engagement	263	263	263
	Gender	263	263	263
	Specialization	263	263	263
	Academic_year	263	263	263
	Skills	263	263	263
	Emotional_aspects	263	263	263
	Participation	263	263	263
	Performance	263	263	263

Table 2: Correlations

		Academic_year	Skills	Emotional_aspects
Pearson Correlation	Engagement	-.146-	.836	.840
	Gender	.155	.042	-.114-
	Specialization	.135	.071	.148
	Academic_year	1.000	-.100-	-.120-
	Skills	-.100-	1.000	.551
	Emotional_aspects	-.120-	.551	1.000
	Participation	-.137-	.694	.705
	Performance	-.142-	.659	.562
Sig. (1-tailed)	Engagement	.009	.000	.000
	Gender	.006	.248	.032
	Specialization	.014	.125	.008
	Academic_year	.	.052	.026
	Skills	.052	.	.000
	Emotional_aspects	.026	.000	.
	Participation	.013	.000	.000
	Performance	.010	.000	.000
N	Engagement	263	263	263
	Gender	263	263	263
	Specialization	263	263	263
	Academic_year	263	263	263
	Skills	263	263	263
	Emotional_aspects	263	263	263
	Participation	263	263	263
	Performance	263	263	263

Table 3: Correlations

		Participation	Performance
Pearson Correlation	Engagement	.913	.840
	Gender	-.077-	.038
	Specialization	.064	.073
	Academic_year	-.137-	-.142-
	Skills	.694	.659
	Emotional_aspects	.705	.562
	Participation	1.000	.714
	Performance	.714	1.000
Sig. (1-tailed)	Engagement	.000	.000
	Gender	.106	.271
	Specialization	.149	.119
	Academic_year	.013	.010
	Skills	.000	.000
	Emotional_aspects	.000	.000
	Participation	.	.000
	Performance	.000	.
N	Engagement	263	263
	Gender	263	263
	Specialization	263	263
	Academic_year	263	263
	Skills	263	263
	Emotional_aspects	263	263
	Participation	263	263
	Performance	263	263

The previous three tables (1,2,3) show the inter-relationships between each of the six variables (Gender – Specialization – Academic year – Skills – Emotional aspects – Participation - Performance) and the engagement variable,

The correlation between the variables (Skills – Emotional aspects – Participation - Performance) with the engagement variable are (0.836 – 0.840 - 0.913 - 0.840) which are positive and strong correlation coefficients that predict increased engagement.

While the correlation between the variables (gender, specialization, academic year) with the variable of engagement was (-0.040 – 0.106 - -0.146), is low correlation, which indicates that the three variables have no influence on the engagement.

Table 4: Variables entered/removed[b]

Model	Variables Entered	Variables Removed	Method
1	Performance, Gender, Specialization, Academic_year, Emotional_aspects, Skills, Participation	.	Enter
a. All requested variables entered. b. Dependent Variable: Engagement			

Table 5: Model summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	1.000[a]	1.000	1.000	.00300

Table 6: Model summary[b]

Model	Change Statistics				
	R Square Change	F Change	df1	df2	Sig. F Change
1	1.000	1461464.101	7	255	.000

a. Predictors: (Constant), Performance, Gender, Specialization, Academic_year, Emotional_aspects, Skills, Participation
b. Dependent Variable: Engagement

Table 7: ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	92.177	7	13.168	1461464.101	.000[a]
	Residual	.002	255	.000		
	Total	92.179	262			

a. Predictors: (Constant), Performance, Gender, Specialization, Academic_year, Emotional_aspects, Skills, Participation
b. Dependent Variable: Engagement

Table 8: Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Correlations
		Beta			Zero-order
1	(Constant)		-.828-	.409	
	Gender	.000	.783	.434	-.040-
	Specialization	.000	.018	.986	.106
	Academic_year	.000	-1.264-	.207	-.146-
	Skills	.271	585.336	.000	.836
	Emotional_aspects	.327	722.424	.000	.840
	Participation	.312	556.334	.000	.913
	Performance	.254	531.598	.000	.840

The previous table declare the correlation significant between the variables (gender, specialization, academic year) with the variable of engagement were (0.409 – 0.434 - 0.986), They are non-significant values at the significance level (0.05), which indicates that the three variables have no influence on the engagement.

While the correlation significant between the variables (Skills – Emotional aspects – Participation - Performance) with the engagement variable were (0.000 - 0.000 - 0.000 – 0.000) which are significant values at the significance level (0.05), which indicates that the four variables have high influence on the engagement and predict of it.

4.2 The Kaggle predict model results

1. Predicting the accuracy score:

```
# predicting the accuracy score
Accuracy=r2_score(y_test,y_prediction)*100
print(" Accuracy of the model is %.4f" %Accuracy)
print('Mean_Sqrd_Error is==',mean_squared_error(y_test,y_prediction))
print('Root_mean_squared error of is ', np.sqrt(mean_squared_error(y_test,y_prediction

Accuracy of the model is 99.9976
Mean_Sqrd_Error is== 8.461196325186892e-06
Root_mean_squared error of is  0.29088135597158665
```

Fig. 13 Predicting the accuracy testing model.

The following functions were used to verify the accuracy of the testing model: (r2_score, Mean_ squared _Error, Root_Mean_Squared Error). The accuracy of the testing model was 99.99%.

```
from sklearn.metrics import mean_absolute_error

y_train_predict =  LR.predict(x_train)
y_test_predict = LR.predict(x_test)

train_accuracy=r2_score(y_train,y_train_predict)*100

mae_train=mean_absolute_error(y_train, y_train_predict)
mae_test=mean_absolute_error(y_test, y_test_predict)

print(" Accuracy of training model is %.5f" %train_accuracy)
print(" mean_absolute_error of training model is %.5f" %mae_train)
print(" mean_absolute_error of testing model is %.5f" %mae_test)
```

Fig. 14 Predicting the accuracy training model.

The following functions were used to verify the accuracy of the training model: (r2_score, Mean_ squared _Error, Root_Mean_Squared Error). The accuracy of the testing model was 99.99%.

2. The results of predicting model:

OLS Regression Results

Dep. Variable:	Engagement	R-squared:	1.000
Model:	OLS	Adj. R-squared:	1.000
Method:	Least Squares	F-statistic:	1.461e+06
Date:	Sat, 28 May 2022	Prob (F-statistic):	0.00
Time:	01:22:56	Log-Likelihood:	1158.5
No. Observations:	263	AIC:	-2301.
Df Residuals:	255	BIC:	-2272.
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0015	0.002	-0.828	0.409	-0.005	0.002
Gender	0.0003	0.000	0.783	0.434	-0.001	0.001
Specialization	7.06e-06	0.000	0.018	0.986	-0.001	0.001
Academic year	-0.0002	0.000	-1.264	0.207	-0.001	0.000
Skills	0.2503	0.000	585.336	0.000	0.249	0.251
Emotional aspects	0.2501	0.000	722.424	0.000	0.249	0.251
Participation	0.2495	0.000	556.334	0.000	0.249	0.250
Performance	0.2505	0.000	531.598	0.000	0.250	0.251

Omnibus:	12.704	Durbin-Watson:	2.173
Prob(Omnibus):	0.002	Jarque-Bera (JB):	5.742
Skew:	-0.071	Prob(JB):	0.0566
Kurtosis:	2.290	Cond. No.	93.4

Fig. 15 OLS regression results.

```
sns.regplot(x=y_test,y=y_prediction,ci=None,color ='red');
```

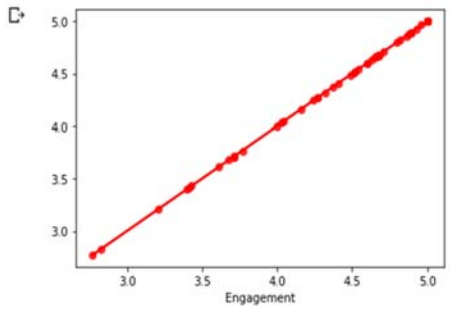


Fig. 16 regression chart

The results indicate the model’s ability to predict with high accuracy the students’ engagement in electronic courses, where the correlation was positive and statistically significant for the input variables, namely skills, emotional aspects, participation and performance, as these factors collectively and individually are considered influential in engaging in electronic courses and through which it can predict the future participation of students in Electronic courses if these variables are taken care of among students.

5. Discussion

The results that were conducted using the SPSS statistical analysis program and using the Kegel program through the machine learning model indicate that there is a positive relationship between the four input variables (skills, emotional aspects, participation and performance) and the dependent variable (engagement), which indicates their ability to influence students’ engagement in electronic courses; As for the other three variables, which are related to demographic data (specialization, gender, academic year), they were not significant, indicating their inability to influence enrollment in electronic courses.

The results of the current study agree with the study results of the Inder [17], which showed the ability of the skills variable to predict students' involvement in online courses at the university, and that students' possession of these skills makes them more active and positive during learning in the course and makes them more involved in the components of the course.

Dixson [16] stresses that students' possession of the skills of dealing with courses via the web, such as dealing with content, responses to course requirements, interacting with it, and holding meetings via the web are all skills that contribute to students' involvement in the online course if this is taken into account and training students before joining the courses to possess these skills because The complexity of students' skills will lead to students' reluctance to engage in the course.

The current study agrees with Redmond et al., [18] study, which indicated that emotional aspects are among the factors affecting students' engagement in online courses, and through which it can predict students' engagement if these aspects are taken care of among students, which benefits the faculty member in increasing students' motivation and attention to emotional aspects of the course, which encourages them to engage more in the courses throughout the learning period.

Also interested in the study of Altuwairqi et al. [19] about emotional aspects and their role in students' involvement in online courses, where it created a model to predict students' engagement in electronic courses by measuring emotional aspects through facial expressions via video clips, while the current study focused on monitoring emotional aspects through online questionnaire and measuring them according to textual expressions and responses the students themselves, in spite of that, the two studies agreed on the importance of the emotional aspects variable as one of the factors affecting students' engagement, and this was indicated by the results of the regression analysis in the two studies, which confirmed the ability of emotional factors to predict students' engagement in online courses.

The results of the current study agree with the study results of the Inder [17], which showed the ability of the emotional aspects variable to predict students' engagement in online courses at the university and that students' acceptance and satisfaction within the learning stages of online courses make students prefer learning through these courses and increases students' engagement in online courses.

Dixson [16] stresses that taking into account the emotional aspects of the course in terms of ease of navigation in the course, good design, content organization and content presentation methods are all things that make the electronic content more interesting and increase the students' motivation to make more effort to continue learning via the electronic course, which leads to an increase enroll students in online courses.

The results of the current study agree with the study results of the Inder [17], which showed the ability of the participation variable to predict students' involvement in online courses at the university, and that students' participation in activities, forums, and learning objects in the course is one of the most important indicators of predicting student engagement, as there is a positive correlation between the participation variable and the engagement variable.

Many studies [20] [16] [21] have agreed that students' participation in online courses, especially interactions with electronic content or interactions with students each other, and interactions between students and the course instructor, make the course more interesting and facilitate the learning process for students through the electronic course, and they are more eager to engage positively in the course.

The results of the current study agree with the results of the study of Wells et al., [22] which indicated that there is a positive correlation between students' performance and their participation in online courses, but the two studies differ in the type of data on which the model was built, as the current study relied on students' self-report through the questionnaire, while the Wells et al., [22] study on the performance data recorded in the electronic content management system, which monitors the performance of students during the learning period.

Many studies (Hussain et al., 2018; Bahati et al., 2017; Rodgers, 2008) also agree that there is a positive correlation between students' involvement in online courses and their performance in the course and their results in tests, and this is consistent with the results of the current study of a positive correlation between university students' performance in tests and assessment results and their engagement in the online course.

The results of the current study agree with the study results of the Inder [17], which showed the ability of the performance variable to predict students' engagement in online courses at the university, and that students' performance in tests and in self- and group assessment activities, and their progress in these assessments greatly contribute to students' involvement in online courses.

Students' engagement is a complex and intertwined paradigm between a multi-faceted group where there is a correlation between students' possession of the skills that make them able to learn through electronic courses and the emotional aspects that push them to learn through these courses and their preference over traditional courses in the classroom, and there is also a correlational relationship between skills and emotional aspects and between students' performance in electronic tests and various assessments via the electronic course, and their participation in forums, meetings, chat rooms, and in various activities, and the correlation of all these factors in an interdependent relationship and between engagement, which was shown by the results of the analysis of the current study, which agrees with the results of studies [23] [24] [17].

6. Recommendation

In light of the study results, the researcher recommends the following:

1. Paying attention to the variables (skills, emotional aspects, participation, performance) when designing online courses because of their effective role in the positive engagement of students.
2. Building new models using machine learning to predict the enrollment of students in online courses

using algorithms other than regression and comparing their accuracy with the current study.

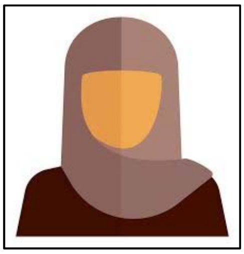
3. Building new models using machine learning to predict students' performance in light of using the input variables (engagement - emotional aspects - participation - skills).

Acknowledgment

The authors would like to thank the Deanship of Graduate Studies at Jouf University for funding and supporting this research through the initiative of DGS, Graduate Students Research Support (GSR) at Jouf University, Saudi Arabia.

References

[1] Ayouni S, Hajjej F, Maddeh M, Al-Otaibi S. A new ML-based approach to enhance student engagement in online environment. PLoS ONE 16(11): e0258788 (2021)
[2] Robinson C. C. and Hullinger H. New Benchmarks in Higher Education: Student Engagement in Online Learning. Journal of Education for Business, 84 (2) (2010)
[3] Christenson S. L., Reschly A. L. and Wylie C. Handbook of Research on Student Engagement, Springer (2012)
[4] Motz, B., Schroeder, N. Zook, J. and Gunkel, M. The validity and utility of activity logs as a measure of student engagement., in 9th International Conference on Learning Analytics & Knowledge (LAK19) (2019)
[5] Coates, H. Student Engagement in Campus-Based and Online Education, University Connections Routledge, London, UK (2006)
[6] Staikopoulos, A. O'Keeffe, I. Yousuf, B. et al. Enhancing student engagement through personalized motivations, in Proceedings of IEEE 15th International Conference on Advanced Learning Technologies, pp. 340–344, Hualien, Taiwan, July (2015)
[7] Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. Enhancing students' engagement by increasing teachers' autonomy support. Motivation and Emotion, 28(2), 147–169 (2004).
[8] Manwaring, K. C. Larsen, R. Graham, C. R. Henrie, C. R. and Halverson, L. R. Investigating student engagement in blended learning settings using experience sampling and structural equation modelling. Internet and Higher Education, 35, 21–33 (2017)
[9] Hussain, M.; Zhu,W.; Zhang,W.; Abidi, S.M.R. Student engagement predictions in an e-learning system and their impact on student course assessment scores. Comput. Intell. Neurosci (2018)
[10] Cocea, M., & Weibelzahl, S. Disengagement Detection in Online Learning: Validation Studies and Perspectives. IEEE Transactions on Learning Technologies, 4(2), (2011)
[11] Sadeque, F., Solorio, T., Pedersen, T., Shrestha, P., & Bethard, S. Predicting continued participation in online health forums. In Proceedings of the Sixth International Workshop on Health Text Mining and Information Analysis (pp. 12-20). doi:10.18653/v1/W15-2602 (2015).
[12] Sharma, P., Joshi, S., Gautam, S., Filipe, V., & Reis, M. J. Student engagement detection using emotion analysis, eye tracking and head movement with machine learning. CORR, 1909.12913. abs-1909-12913 (2019)
[13] Calvo, R. A., & D'Mello, S. Affect detection: An interdisciplinary review of models, methods, and their applications. IEEE Transactions on Affective Computing, 1(1), 18–37(2010)
[14] Kuh, G. D. What were learning about student engagement from nsse: Benchmarks for effective educational practices. Change: The Magazine of Higher Learning (2003)
[15] Handelsman, M. Briggs, W. Sullivan, N. and Towler, A. A measure of college student course engagement. The Journal of Educational Research, 98 (3), 184–192 (2005)
[16] Dixson, M. D. Measuring student engagement in the online course: The Online Student Engagement scale (OSE). Online Learning, 19(4), n4 (2015)
[17] Inder, S. Factors Influencing Student Engagement for Online Courses: A Confirmatory Factor Analysis. Contemporary Educational Technology, 14(1), (2022)
[18] Redmond. An Online Engagement Framework for Higher Education. Online Learning Journal, 22(1) (2018)
[19] Altuwairqi, K., Jarraya, S.K., Allinjawi, A., & Hammami, M. A new emotion–based affective model to detect student's engagement. Journal of King Saud University-Computer and Information Sciences, 15, pages1387–1395 (2018)
[20] Bigatel, P. M., Ragan, L. C., Kennan, S., May, J., & Redmond, B. F. The identification of competencies for online teaching success. Journal of Asynchronous Learning Networks, 16(1), 59-77 (2012)
[21] Ragusa, A. T., & Crampton, A. Sense of connection, identity and academic success in distance education: Sociologically exploring online learning environments. Rural Society, 27(2), 125-142 (2018)
[22] Wells, M. Wollenschlaeger, A. Lefevre, D. Magoulas, G. D. and Poulovassilis, A. Analysing engagement in an online management programme and implications for course design. Proceedings of Sixth International Conference on Learning Analytics & Knowledge, Edinburgh, UK, (2016)
[23] Trowler, V., & Trowler, P. Student engagement evidence summary. The Higher Education Academy (2010)
[24] Kahu, E. R. Framing student engagement in higher education. Studies in Higher Education, 38(5), 758-773 (2013)



Jawaher Alsirhani received the B.E. in information systems from Jouf University. She is currently a master's student in data science at the Collage of computer and Information Sciences at Jouf University, Saudi Arabia. She holds a training certificate in 2019 as a professor at the Higher Support Institute under the supervision of the Technical and Vocational Training Corporation. Her interests include information technology and data science, such as data entry, word processing, and applying algorithms using coding in several programs, the most famous of which is the kaggle program.

Khalaf Alsalem received the B.E. in computing from Aljouf Teachers' collage in 2006, M.Sc. in information science from Otago University in 2010 , and PhD degree in information systems from Claremont Graduate University in 2019. He is an assistant professor in information systems department at the Collage of computer and Information Sciences at Jouf University, Saudi Arabia.