Contents lists available at GrowingScience

International Journal of Data and Network Science

homepage: www.GrowingScience.com/ijds

Examining the usability of mobile applications among undergraduate students using SUS and data mining techniques

Mohammed Afif[a*]

[a]*Department of Management Information Systems, College of Business Administration, Prince Sattam Bin Abdulaziz University, Saudi Arabia*

C H R O N I C L E

Article history:
Received: November 29, 2023
Received in revised format: January 16, 2024
Accepted: February 12, 2024
Available online: February 12, 2024

Keywords:
Usability
Mobile Application
Opinions mining
System Usability Scale
Higher Education
Classification
Clustering

A B S T R A C T

Mobile Applications offer a new style to service sectors, for instance, in higher education, mobile applications are utilized to provide access to academic resources and academic services. Despite the wealth of mobile applications, they encounter various challenges that have attracted the interest of academia and software developers. The usability issues of mobile applications may cause performance degradation, resulting in the company's loss in terms of cost. This study aims to investigate the usability of the Prince Sattam bin Abdulaziz University (PSAU) mobile application by adopting data mining as a descriptive and predictive process. The first step was gathering data of the usability of the PSAU mobile application using the system usability scale. Afterwards, data was preprocessed into a suitable format to apply data mining methods. Specifically, the explanatory model has been employed to describe and investigate insights related to the usability factors and features of the PSAU mobile application. Furthermore, this study adopted the Four Clustering methods to segment the usability levels of the PSAU mobile application into homogenous groups based on user behavior. Additionally, the predictive model was used to build models for predicting the usability level and Grade and five classification algorithms were employed to predict the usability level and Grade. Most algorithms have given positive results in all performance indicators, where the accuracy rate achieved is 98% to 95% for most methods. The results revealed that the PSAU mobile application has an acceptable usability level, and the data mining methods helped to discover hidden patterns. Furthermore, the findings will help the developers and policymakers understand users' and stakeholders' behavior to find the most common usability problems for each group, and customize the PSAU mobile application.

1. Introduction

The growth of mobile applications (mobile Apps) is astounding. It helps people directly access content, search for content, and perform transactions. Mobile applications are extensive in numerous segments, involving banking, healthcare, e-commerce, tourism, and education (Weichbroth, 2020; Lumor et al., 2020). The development of mobile applications offers a new style to service sectors. For instance, in higher education, mobile apps provide access to academic resources and services (Lumor et al., 2020). Mobile Apps allow teaching staff, students, and stakeholders to carry out their duties wherever they are. Despite the wealth of mobile apps, they still encounter various challenges that have drawn the interest of academia and software developers. Mobile Apps usability is the most critical challenge that has gained the interest of scientists and software program developers (Weichbroth, 2020; Lumor et al., 2020; Parsazadeh et al., 2018). Usability is defined by ISO 9241-11 as "the extent to which specified users can use a product to achieve specific goals with effectiveness, efficiency, and satisfaction in a specified context of use" (Weichbroth, 2020; Iso & ISO, 2018; Moumane et al., 2016). In software engineering, usability

* Corresponding author.
E-mail address: mohaafif@gmail.com (M. Afif)

ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print)

doi: 10.5267/j.ijdns.2024.2.008

is a critical software quality characteristic that appraises how easy the user interface is to use (Weichbroth, 2020). Thus, any application programs or products that have usability challenges will lead to the failure of software and product (Moumane, Idri, & Abran, 2016; Nielsen & Molich, 1990; Nielsen, 1994; Brooke et al., 1996; Kumar & Goundar, 2019). Conversely, good usability software will enhance productivity and revenues (Nielsen & Molich, 1990).

As the mobile app development industry evolves rapidly, Organizations must keep up with the latest trends by continuously updating and developing mobile applications to stay informed and adaptable to new developments. Therefore, it is critical to discover the significant issues facing these mobile applications' stakeholders and get feedback about their usability. Numerous approaches are offered for estimating mobile application usability (Weichbroth, 2020; Parsazadeh et al., 2018; Moumane, Idri, & Abran, 2016; Kumar & Goundar, 2019; Kortum & Sorber, 2015; Silfee et al., 2017; Fabil et al., 2015; Harrison et al., 2013; Az-zahra et al., 2019; Al-Mayyan & Al-Refai, 2020; Kaya et al., 2019; Hoehle & Venkatesh, 2015; Afif, 2021; Afif, 2023), where considerable models and methods were applied. In addition, most studies employed quantitative and qualitative-based statistical methodologies. The outcomes of these studies highlighted the most critical issues which are related to mobile application usability. Furthermore, giving feedback helps empower the mobile applications' functionality. Moreover, appraisal is essential in addressing mobile applications' usability problems. These intentions encourage us to concentrate on exploring the usability of the PSAU mobile application to highlight the usability issues that face its users by applying data mining as a descriptive and predictive process. Prince Sattam bin Abdulaziz University has developed a mobile application which allows faculty members, students, and other stakeholders to access various academic services and resources and perform transactions.

Having said that, this current study attempts to bridge the gap in the literature using data mining techniques to evaluate the usability of mobile applications and highlight mobile application usability difficulties at higher education institutions. In addition, the study attempts to give mobile application designers and stakeholders in higher education insights that may help them enhance mobile application usability.

The main objective of the present study is to examine the usability of mobile applications among university students of PSAU. To achieve this, knowledge discovery techniques will be employed to examine the present status of PSAU Apps, identify issues, develop models for predicting the degree of usability and grades for the PSAU mobile application.

The paper's organization is as follows: related works are presented in section 2; the methodology is explained in detail in section 3, and section 4 presents the results and discussions. Finally, the study concludes with section 5.

2. Related Work

To our best of knowledge, only some methods utilize data mining techniques to evaluate mobile applications' usability. So, this part listed the related works that utilized data mining to extract hidden knowledge from datasets related to users interacting with websites. Lin et al. (2022) performed a systematic literature review on opinion mining for software development activities. The survey involved more than 180 articles, and one of the review's main points was related to employing opinion mining to measure software product quality. Furthermore, the researchers investigated articles that employed opinion mining in software engineering in gathering informative app reviews to recognize how designers can upgrade their products and alter their release plans. Besides, opinion mining has been used to evaluate the quality of software products. Al-Omar (2018) used data mining techniques to extract usability problems encountered by users of the learning management system (LMS) at King Abdulaziz University. The study aimed to combine similar usability difficulties into one set. The findings assisted universities in identifying faculty members' and students' behaviors, detecting the most common issues for each set, and adapting the LMS for each set.

Sagar and Saha (2016, 2017) applied data mining techniques to assess the usability of academic websites in combination with traditional usability testing approaches. The study aimed to discover an everyday usability problematic pattern among the top 50 academic websites. The study uses the ISO9241-151 guidelines under 16 categories to gather data from hundreds of participants. Association rules and decision trees were used to extract hidden relationships and patterns. The findings indicated that most of these issues were found in search and social media categories, where 50.53% of website guidelines are fully functional, while 49.46% of characteristics have usability issues.

Boza et al. (2014) employed data mining techniques to extract relationships among Nielsen usability components. They aim to discover hidden relationships with features and components and discover website usability challenges. Experiments were applied using association rules and decision trees on a dataset containing evaluation reports of different websites. The outcomes indicated that the suggested method was promising for discovering exciting relations from this data type. The discovered patterns and relations are helpful for Website designers and give them insights related to usability issues that must be considered. El-Halees (2014) used opinion mining as an automatic technique to determine subjective usability and the approach developed a model for discovering knowledge from the opinion to increase subjective software usability. The presented methodology concentrated on three usability quality factors: effectiveness, efficiency, and satisfaction. Four software had been used to estimate the proposed model as well. Testing was performed on a dataset encompassing 565 reviews divided into 345 positive and the remaining negative. The outcome is 85% for accuracy.

The significant difference between the current study and the previous studies presented by Afif (2021, 2023) is that the current study uses a data mining technique for investigating the usability of the PSAU mobile application. Furthermore, the proposed approach utilizes the data mining as a descriptive and predictive process for extracting the hidden knowledge from users' opinions about the usability of PSAU mobile application.

After reviewing the literature, to the researcher's best knowledge, no studies have been conducted at Prince Sattam bin Abdulaziz University applying a data mining approach to extract hidden knowledge for evaluating the usability of mobile applications.

3. Materials and Method

The methodology involves various phases, starting with gathering data and then the preprocessing phase. After that, data is split into training and testing. The training data were used to build models for prediction using data mining algorithms, while the testing datasets were utilized to evaluate and assess the models. The last phase was the knowledge discovery phase and interpretation of the outcomes. Fig. 1 shows the methodology flowchart.

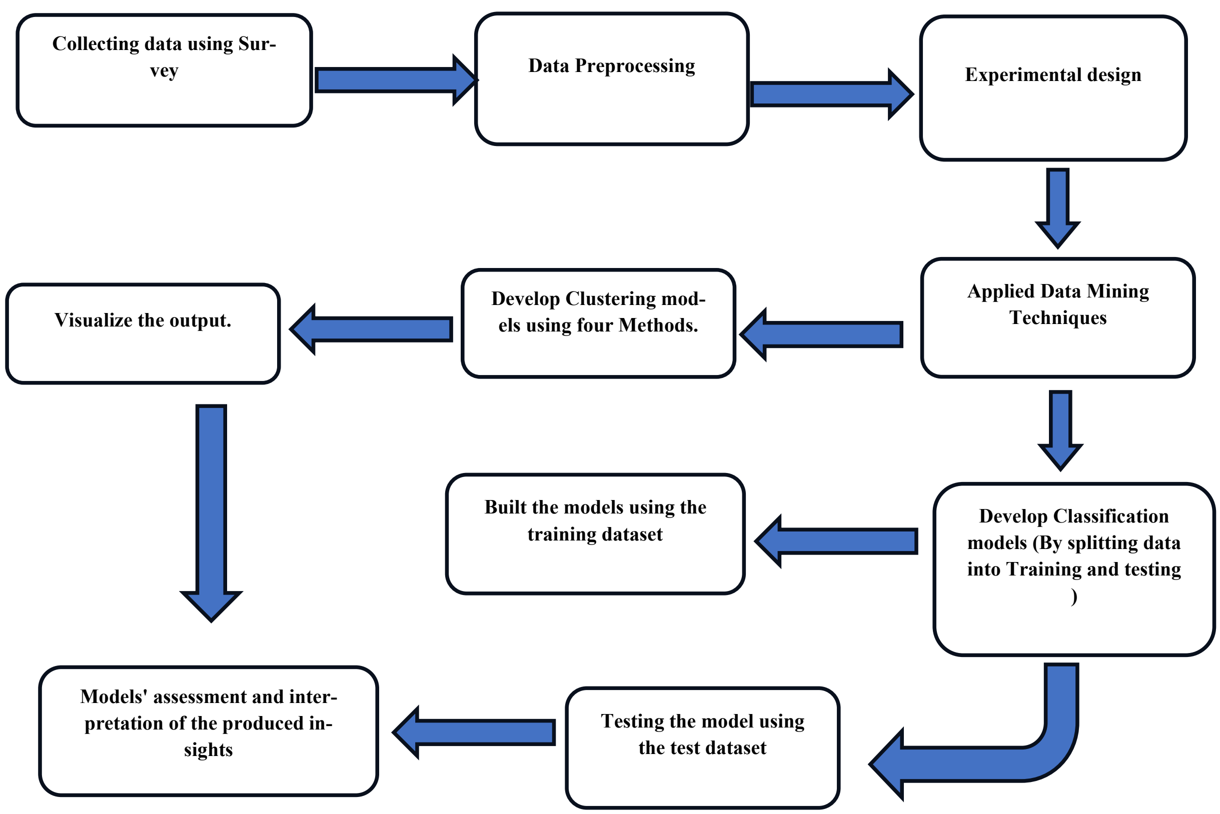


Fig. 1. Flowchart of the research process

In the first phase, a survey instrument was utilized to gather students' opinions about the usability of the PSAU mobile application. The survey contains two parts. The first of which contained four statements designed to collect data from the respondents: the study level, gender, type of smartphone, and smartphone usage. The second portion encompasses system usability scale (SUS) sentences suggested by Brooke (1996) and mainly exploited to gather the opinion of students about the usage of PSUA mobile application. The SUS is a popular measure for usability assessments and used by several studies (Brooke, 2013; Bangor et al., 2009; Al-Omar, 2018; Bangor et al., 2008; Harrison et al., 2013; Kaya et al., 2019; Vlachogianni & Tselios, 2022). A survey was published through Google form and sent to various students from different colleges at Prince Sattam bin Abdulaziz University, during the period from February 2022 until December 2022. The participants were selected randomly, without any criteria. The collected dataset contains 18 features and 523 records. Table one lists details about each attribute and includes a description of the dataset's features and frequency.

The next stage of methodology is preprocessing. The collected data are prepared by adopting many data preprocessing tasks such as selecting, cleaning, transforming, and constructing new attributes. The MS Excel 365 software has been employed during this phase. The collected data was first cleaned by removing incomplete data, missing values, and unsuitable data. Then, the encoding process is invoked by transforming the nominal and categorical data into numerical values. Next, scaling is applied by normalizing the values of features to a range between (0-1), not only this but, new attributes were constructed to store the values of derived features, SUS score, Grade, and usability levels. Specifying values for newly derived attributes depends on some rules that have been proposed (Bangor et al., 2008, 2009; Brooke, 2013; Vlachogianni & Tselios, 2022). The target attributes are usability level and Grade.

After preparing the dataset the data mining was invoked to achieve two functions. First, it is invoked as a descriptive model used to define patterns in data and investigate the characteristics of the data detected. The next one is used as a predictive model, which is exploited to build a model for predicting the datasets whose results are unknown (Han et al., 2022; Gupta & Chandra, 2020; Kantardzic, 2011; Abu Saa et al., 2019; Al-Hagery et al., 2020; Yağcı, 2022; LIN et al., 2022). This study applied descriptive and predictive models to discover correlations, patterns, relationships, and trends by searching a large amount of data. This step aims to find significant knowledge from the collected data. Different data mining techniques have been exploited, i.e., classification, clustering, and filtered attributes. The classification technique is applied to predict the level and grades of usability. Five popular prediction techniques were used to build models: Support Vector Machine, Random Forest, neural networks, Naive Bayes, and logistic regression. Clustering is applied as a descriptive and predictive data mining technique to classify observations into homogeneous groups based on distance, density, or connectivity criteria. Four clustering methods (Hierarchical, K-Mean, DBSCAN, and Louvain) were applied; clustering primarily aims to discover patterns, segments, outliers, and trends in data sets.

4. Data Explanatory

This activity aims to investigate the data and its features over summaries and charts. Furthermore, exploring the data serves to understand and gain information about data in more detail. Table 1 summarizes the data, including feature name, description, code, frequency and percentages for each feature.

Table 1
Description of the dataset

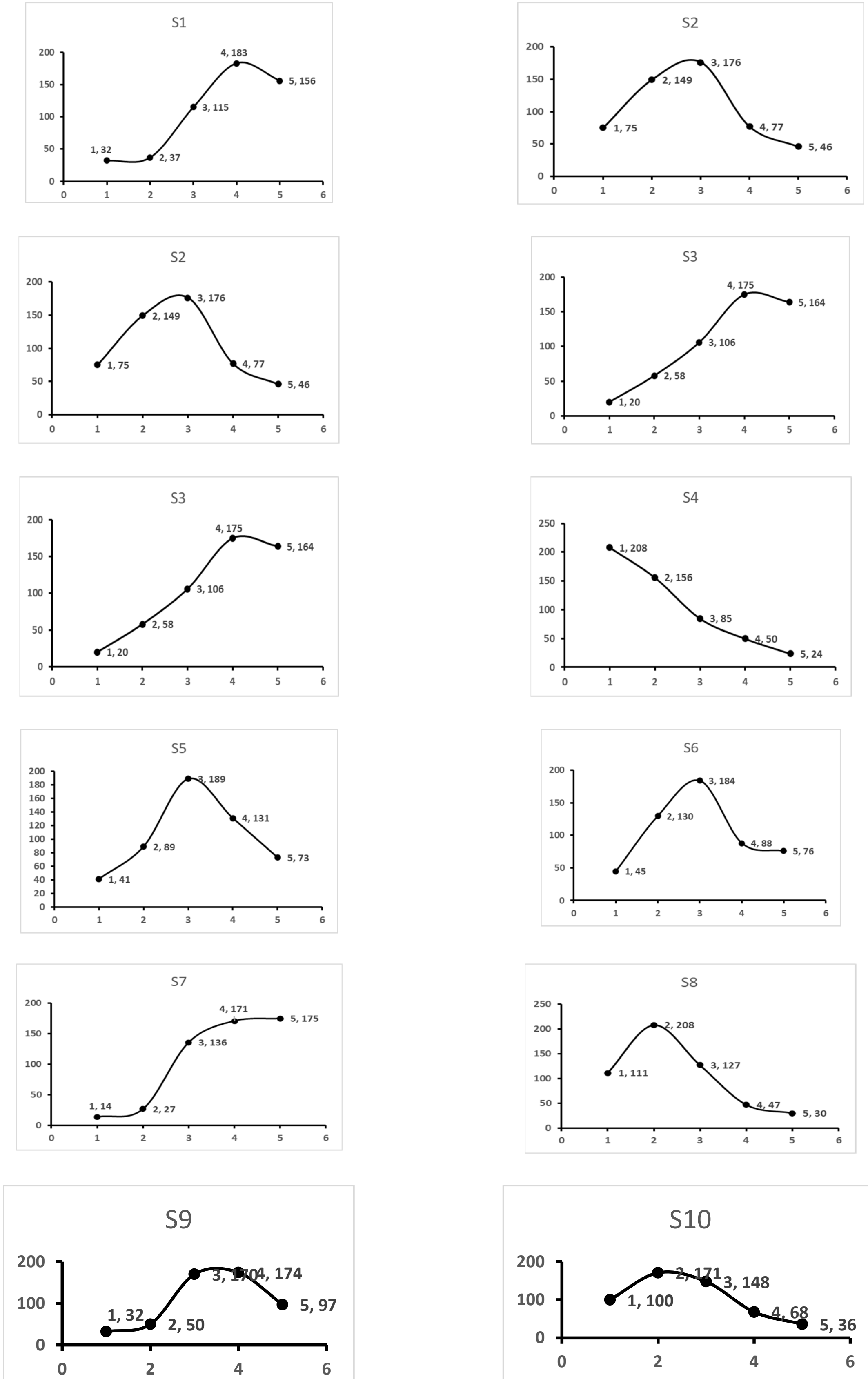
Name	Description	Code	Frequency	percents
F1	This feature represents the study level, encoded with 1,2,3 and 4. (1=First year and .etc.	1	184	35%
		2	90	17%
		3	127	24%
		4	122	23%
F2	This feature represents a gender, encoded by 1 for males and 0 for females.	1	380	73%
		0	143	27%
F3	Mobile device type website, 1==iPhone, 2= Samsung, 3= Huawei	1	462	88%
		2	39	8%
		3	22	4 %
F4	This attribute for usage experience of App 1= less than one year, 2= from one to 3 years, 3= From 4 and less than 6, 4= More than seven years	1	34	7%
		2	109	21%
		3	117	22%
		4	263	50%
F5-F14	The features from F5 to F14 represent the SUS scale. Table 2 lists descriptions of them.			
F15	This feature represents the value of the SUS score and calculated based on rules(Bangor et al., 2008; Brooke, 2013).			
F16	This feature represents the Grade of usability quality and has five letters (A, B, C, D, F)	A	80	15.3%
		B	125	23%
		C	0	0
		D	179	34%
		F	139	27%
F17	This feature represents the rating of usability quality, which has five values Excellent, Good, Okay, Poor, Awful)- Also, it represents the target feature for multiclass	Excellent	80	15.3%
		Good	125	23%
		Okay	0	0
		Poor	179	34%
		Awful	139	27%
F18	This feature is the target, and it represents the usability level and has two values (Acceptable encoded with one and Not acceptable encoded with -1) – For Binary classification.	Acceptable	384	73%
		Not acceptable	139	27%

No missing values were detected, and the primary indicators highlighted first regarding gender: 73% were males, and 27% were females, while the second was the study level, where 35% were from the first level and 24%,23%, and 17% from the third, fourth, and second levels correspondingly. The next indicator was related to the smartphone type: 88% use an iPhone, and the next was 11% for Samsung and Huawei. Regarding smartphone usage, it indicates that 50% of students have used smartphones for more than five years, while 22% have used them for less than five years. The rest of the students have used smartphones for less than three years. As stated in the previous section, the SUS was used to collect opinions about the usage of the PSAU mobile application. The SUS helps measure a combination of usability factors, including efficiency, intuitiveness, ease, and satisfaction (Bangor et al., 2009; Al-Omar, 2018; Goel, 2018; Brooke, 1996; Bangor et al., 2008; Brooke, 2013; Vlachogianni & Tselios, 2022). Also, the SUS contains ten statements. Table 2 displays descriptive pointers of the SUS elements, including the mean and median, whose values are 2.1 - 3.77 and 2—4, respectively. The dispersion was in the range (of 0.21 and 0.55). The dispersion reflects the coefficient of variation of the values of the features, whose values were in the range of 0.21 and 0.55. This measure indicates the variation of the dataset is relative to the mean.

Table 2
Description of SUS features

Name	Description	Mean	Median	Dispersion	Stdev	Min	Max	Frequency		
								Code	**Freq**	**Percent**
S1	I think that I would like to use this mobile App frequently	**3.75**	4	0.30	1.14	1	5	1	**32**	6.1%
								2	**37**	7%
								3	**115**	22%
								4	**183**	35%
								5	**156**	30%
S2	found this mobile App unnecessarily complex.	2.75	3	0.41	1.14	1	5	1	75	14%
								2	149	28.5%
								3	**176**	34%
								4	77	15%
								5	46	9%
S3	This mobile App was easy to use.	**3.77**	4	0.30	1.1	1	5	1	20	4%
								2	58	11%
								3	106	20%
								4	**175**	33.5%
								5	**164**	31%
S4	I would need the support of a technical person to be able to use this mobile App.	2.1	2	0.55	1.2	1	5	1	**208**	40%
								2	**156**	30%
								3	85	16%
								4	50	10%
								5	24	5%
S5	I found the various functions in this mobile App were well-integrated	3.2	3	0.35	1.11	1	5	1	41	8%
								2	89	17%
								3	**189**	36%
								4	**131**	25%
								5	73	14%
S6	I thought there was too much inconsistency in this mobile App."	3.04	3	0.38	1.2	1	5	1	45	9%
								2	130	25%
								3	**184**	35%
								4	88	17%
								5	76	15%
S7	Most people would learn to use this mobile App very quickly.	**3.9**	4	0.26	1	1	5	1	14	3%
								2	27	5%
								3	136	26%
								4	**171**	33%
								5	**175**	33%
S8	found this mobile App very cumbersome to use.	**2.4**	2	0.46	1.1	1	5	1	111	21%
								2	**208**	40%
								3	**127**	24%
								4	47	9%
								5	30	6
S9	felt very confident using this mobile App	**3.5**	4	0.31	1.1	1	5	1	32	6%
								2	50	10%
								3	**170**	33%
								4	**174**	33%
								5	97	19%
S10	I needed to learn many things before I could get going with this mobile App.	2.56	2	0.45	1.1	1	5	1	100	19%
								2	**171**	33%
								3	148	28%
								4	68	13%
								5	36	7%

The standard deviation values were in the range (1- 1.2). As reported in Table 2, the features S1, S3, S7, and S9 obtain the highest means values with rates (3.9, 3.8, 3.8, and 3.5), while the means for features S2, S4, S5, 6, 8, and S10 were in the range between (3.01 to 2.1). Figures from 2 to 10 demonstrate the distribution of each statement, including the scale and its frequencies. The significant indicators extracted regarding the usability of the PSAU mobile application are summarized as follows: the PSAU mobile application is easier to learn, easy to use, and less complex. In addition, the effectiveness and efficiency of the PSAU mobile application are acceptable. Furthermore, the PSAU mobile application requires more attention to enhance its functionality by adding more functions to increase stakeholder satisfaction and empower usability. Fig. 11 summarizes the complete PSAU mobile application usability indicators that encompass scale and frequencies for each choice.



Figs. 2-10. Distribution of Scale and frequencies for S1 to S10

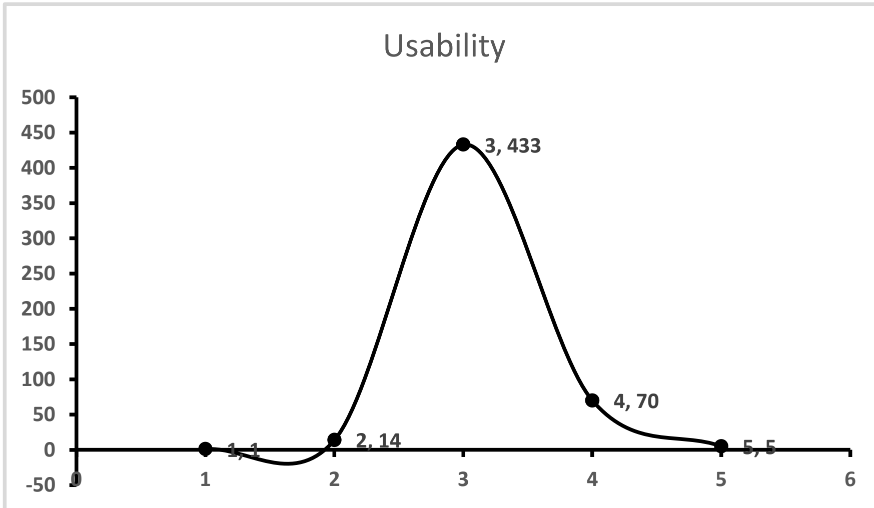


Fig. 11. Overall usability scales and frequencies

5. Results and Discussion.

Orange machine learning software has been utilized because it is an open-source software, has a user-friendly interface, is simple in developing models, and involves many methods. All experiments were performed on a machine with a core i7 CPU, 16 GB of RAM, and Windows 10 OS. Fig. 12 and Fig. 13 illustrate the workflow diagram for the proposed methods using the Orang platform. The workflow is determined by the steps implemented to achieve the study objectives. 10-fold cross-validation and holdout methods split the dataset into Training and testing. The training part was employed in order to develop the models while testing and assessing them. Five widespread prediction techniques were used to build models: Support Vector Machine, Random Forest, Neural Networks, Naive Bayes, and logistic regression.

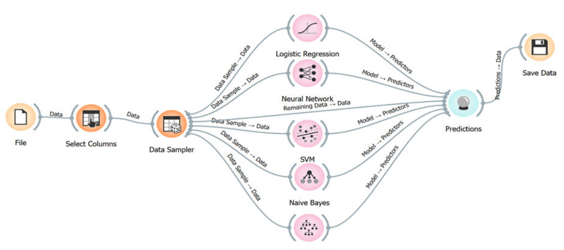


Fig. 12. The workflow diagram for classification Models- Multiclass

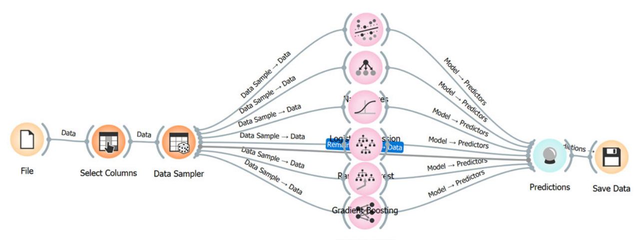


Fig. 13. The workflow diagram for classification Models- Binary Class

Furthermore, experiments are performed to discover the patterns and hidden relationships among features and predicate the usability level and Grade. Two types of prediction models were developed. The first one was to predicate the usability grade for the PSAU mobile application as A, B, C, D, or F. Grades A, B, C, and D mean that the usability ranges from excellent to acceptable. In contrast, an F grade means the usability level is unacceptable and denotes usability factors such as effectiveness, efficiency, and satisfaction that need more attention.

Table 3
Results of usability grade as Multiclass problem using 10-fold CV

Model	AUC	CA	F1	Precision	Recall
SVM	99	**97**	**97**	**97.3**	**97**
Random Forest	99	93.3	93.3	93.7	93.3
Neural Network	98.9	95.6	95.6	96.1	95.6
Naive Bayes	98.6	92.4	92.5	92.7	92.4
Logistic Regression	100	**99**	**99**	**99**	**99**

Table 4
Results of usability grade as Multiclass problem using Hold-out method

Model	AUC	CA	F1	Precision	Recall
SVM	**99.3**	**95.2**	**95.2**	**95.4**	**95.2**
Random Forest	98.7	93.8	93.8	93.9	93.8
Neural Network	98.3	92.1	92.1	92.6	92.2
Naive Bayes	98.2	88.3	88.4	88.7	88.3
Logistic Regression	100	**98.7**	**98.7**	**98.7**	**98.7**

Table 3 and Table 4 report the outcomes for predicting the usability grade using the 10-fold cross-validation and hold-out method. The Tables contain Area Under the Curve (AUC), classification accuracy (CA), F1, precision, and recall in the first column. All these are the performance indicators. Logistic regression and SVM give high values for all performance indicators, where it gives 100%, 99%, 99%,99%, and 99%, while SVM got 99%, 97%, 97%, 97%, and 97%.

Figs. (15-17) display visual presentations of all algorithm performance pointers and show the best-obtained results by Logistic regression and SVM using two validation methods.

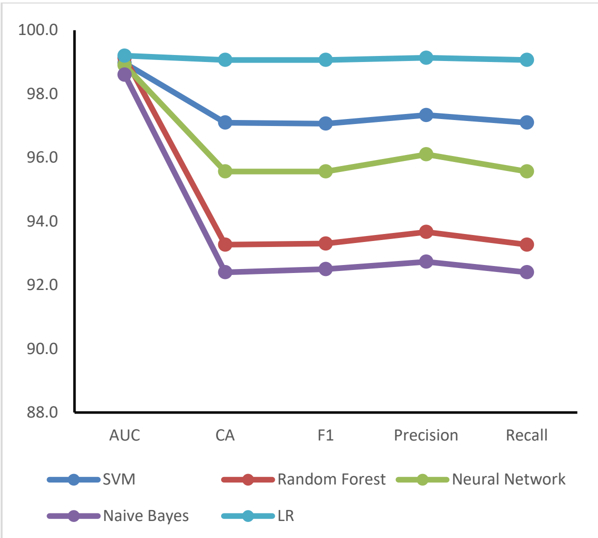


Fig. 15. Results for Usability Grade Prediction

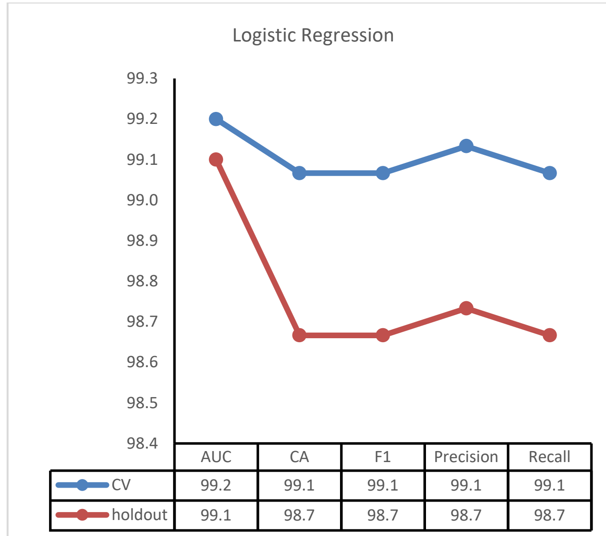


Fig. 16. Results for Usability Grade Prediction

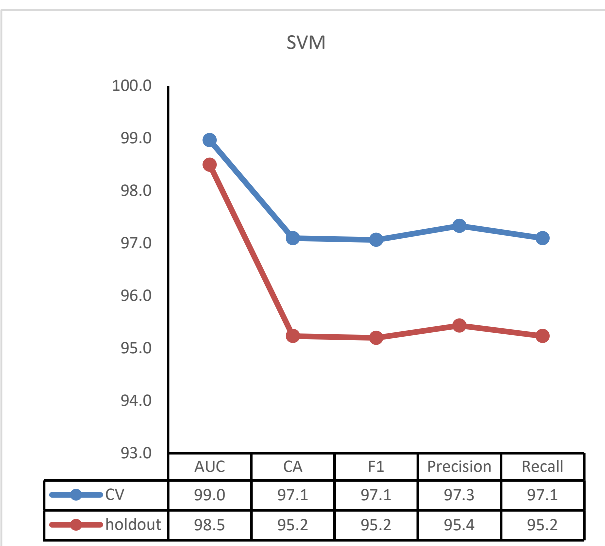


Fig. 17. Results for Usability Grade Prediction

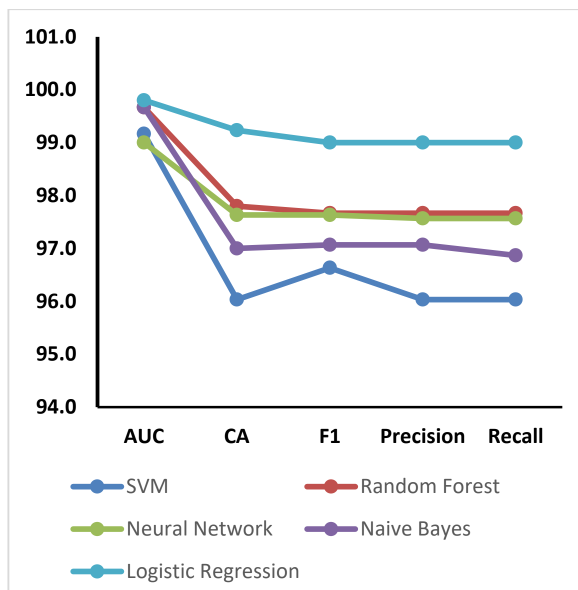


Fig. 18. Usability Level prediction of all methods

The second model has been for forecasting the usability level of the PSAU mobile application as acceptable or not acceptable. In this part, the model is developed to forecast the usability level (Acceptable or not acceptable level). An acceptable level means the usability is good, while an unacceptable level means some issues need improvement. The model was built using the same algorithms based on the training data and then tested using tested data. The cross-fold and holdout approaches were utilized for splitting the dataset.

Tables five and six report the obtained results for predicting the usability level using the same algorithms used in the first models. The logistic regression and random forest techniques give the best results for all performance pointers, achieving 99.9%,99.8%, 99%, 99%, and 99%, while the Random Forest method got 99.8%, 98.6%, 98.6%, 98.6%, and 98.6% for all performance indicators. Figs. (15-18) display a visual presentation of all performance indicators from different points of view to summarize the different aspects of the obtained results.

Table 5
Results of usability Binary class using 10-fold CV method

Model	AUC	CA	F1	Precision	Recall
SVM	99.4	95.4	95.4	95.5	95.4
Random Forest	99.8	98.6	98.6	98.6	98.6
Neural Network	99.4	96.7	96.7	96.8	96.7
Naive Bayes	98.9	96.5	96.5	96.8	96.5
Logistic Regression	99.9	99.8	99	99	99

Table 6
Results of usability Binary class using holdout method

Model	AUC	CA	F1	Precision	Recall
SVM	99.2	96.0	96.6	96.0	96.0
Random Forest	99.7	97.8	97.7	97.7	97.7
Neural Network	99.0	97.6	97.6	97.6	97.6
Naive Bayes	99.7	97.0	97.1	97.1	96.9
Logistic Regression	99.8	99.2	99.0	99.0	99.0

The clustering technique has been applied in the next part of the experiments. Clustering is a data mining approach applied as a descriptive and predictive technique. It helps to discover patterns, segments, outliers, and trends in the dataset. Furthermore, in this section, clustering methods are utilized to classify users of the PSAU mobile application based on their usage behavior. Figure 19 illustrates the workflow steps that were implemented to achieve the Clustering. Hierarchical, K-means, DBSCAN, and Louvain clustering techniques were exploited to classify the dataset into homogeneous groups based on distance, density, or connectivity criteria. Each method gives different numbers of clusters based on the Grade of usability. The grades are (A, C, D, and F), where A is excellent, C is good, and D is poor. F is Awful (Bangor et al., 2008, 2009; Kaya et al., 2019).

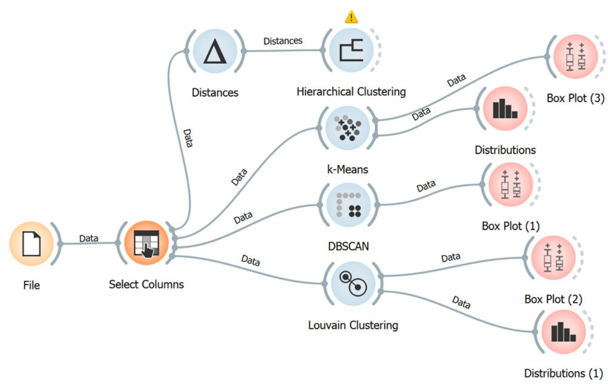


Fig. 19. The workflow diagram for Clustering

Table 7 reports results, including in the first column: the name of the clustering method, then the number of clusters, afterwards, the grade label, and then the frequency and percentage for each label in each cluster. As it is illustrated in the table, the Louvain method produces five clusters. Cluster one contains 157 data elements with a ratio of 30% and includes 20 and 137 elements from the D and F grades. Cluster two identified 139 data points distributed among 109, 29, and 1 for D, B, and A grades, respectively, with a ratio of 27%. Cluster three represented 18% of all segments and contained 92 elements. It has been distributed as 3, 43, 44, and two data points for grades A, B, D, and F, respectively. The last two clusters, 4 and 5, obtain the same value of elements with a percentage of 12.5% of the total. Figure 21 displays a visual representation of obtained clusters with their distribution. The main insights extracted from this clustering method may be summarized as the method segments the data points into five segments based on their similarities with percentages 30%, 27%, 18%, 12.5%, and 12.5%. The first cluster, which represented 30% of data elements, has a lower level of usability with grades F and D. While the other clusters representing 70% classify the usability level as high-grade, including grades A, B, and D. The outcomes provide an indicator of the effectiveness of Clustering in giving a descriptive view of a dataset by classifying the dataset into segments of users based on their usage behavior and similarities of features.

The following method, K-means, classifies the dataset into four clusters, as illustrated in Fig. 22. The first cluster contains 128 elements from only grade F, With a ratio of 24.5%. The next cluster includes 122 elements distributed among three grades, A, B, and D, with 23% of the total. The remaining clusters comprises of 144 and 129 elements with 27.5% and 24.5% ratios. The grades are distributed among A, B, D, and F. Segments two and four contain the highest grades from A, B, and D. Segments one and three contain data points from only grade F in cluster one. Cluster three contains data points distributed among 11 from grade F and 133 from the other grades B and D. This reflects the method's success in grouping the data points according to their similarities of features of the dataset as well as the user's usage behavior. The Silhouette Polt was used to measure the clustering quality, as illustrated in Fig. 24. As in the figure, the ratio for each cluster was C1=0.456, C2=0.47, C3= 0.38 and C4= 0.426. Figure 24 shows the Silhouette analysis for k-means methods to measure clustering quality.

The third method is DBSCAN, which produces two clusters. The first one contains 495 elements distributed among A, B, D, and F grades. In comparison, the second cluster contains only 28 elements from grade D. Figure 23 visualizes the produced clusters using the DBSCAN method.

The last method for Clustering is Hierarchical which creates a hierarchy of clusters by recursively trimming them based on their similarities. In this method, the division starts with all data points in a single cluster and divides them into smaller ones. As in Fig. 20, the method produces five segments. Clusters one and two represent the dataset that has F and D. Clusters three, four, and five contain a dataset that has good Grades.

Table 7
Results of Clustering

Method	Clusters	Label	Frequency	Ratio
Louvain clustering	**Cluster 1**	D	20	13%
		F	137	87%
		C1	**157**	**30%**
	Cluster 2	A	1	1%
		B	29	21%
		D	109	78%
		C2	**139**	**27%**
	Cluster 3	A	3	3.3%
		B	43	47%
		D	44	48%
		F	2	2.2%
		C3	**92**	**18%**
	Cluster 4	A	67	100%
		C4	**67**	**12.5%**
	Cluster 5	A	8	12%
		B	53	79%
		D	6	9%
		C5	**67**	**12.5%**
K-means Clustering	**Cluster 1**	**F**	**128**	**24.5%**
	Cluster 2	A	7	6%
		B	66	54.1%
		D	49	40.2%
		C2	**122**	**23%**
	Cluster 3	B	3	2.1%
		D	130	90.3%
		F	11	7.6%
		C3	**144**	**27.5%**
	Cluster 4	A	73	56.6%
		B	56	43.4%
		C4	**129**	**24.6%**
DBSCAN	**Cluster 1**	A	80	16.2%
		B	124	25.1%
		D	154	31%
		F	137	27.7%
	Cluster 2	D	28	

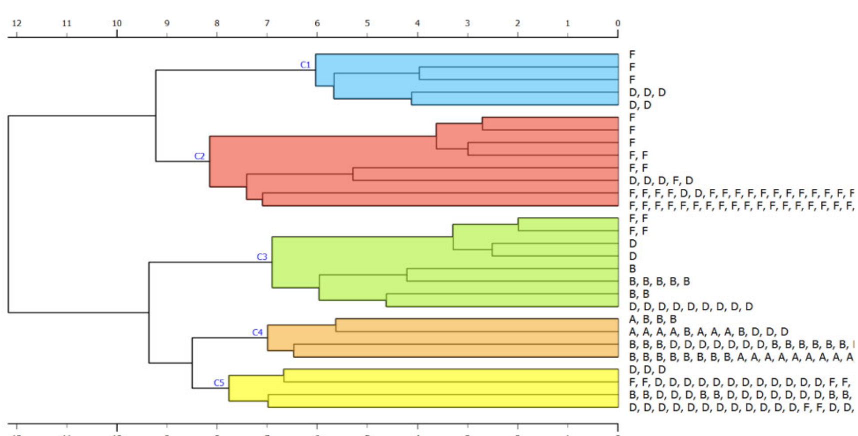


Fig. 20. Hierarchical Clustering

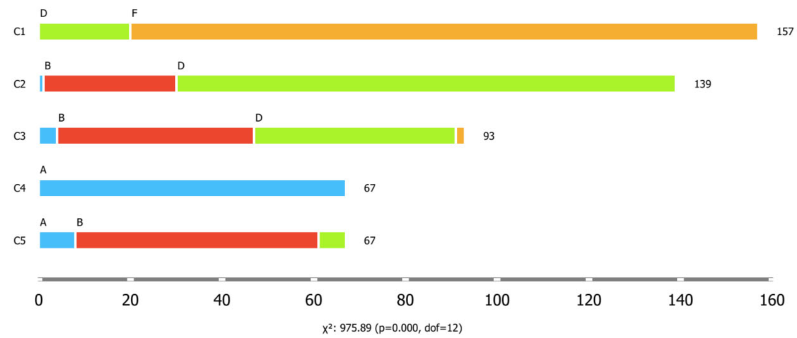


Fig. 21. Louvain Clustering

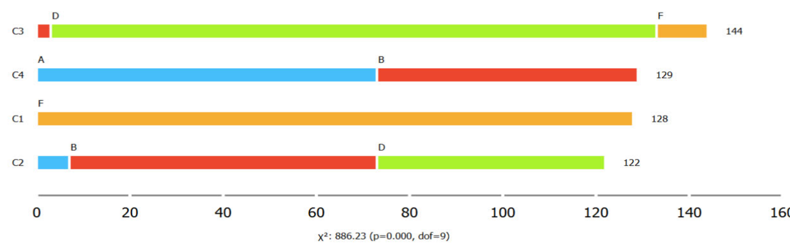


Fig. 22. K-means Clustering

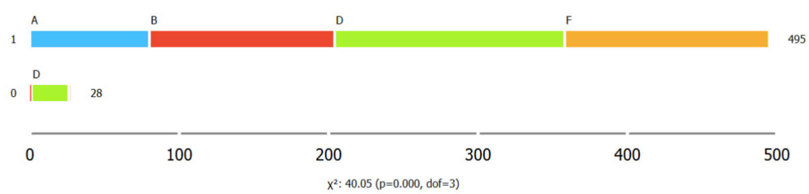


Fig. 23. DBSCAN Clustering

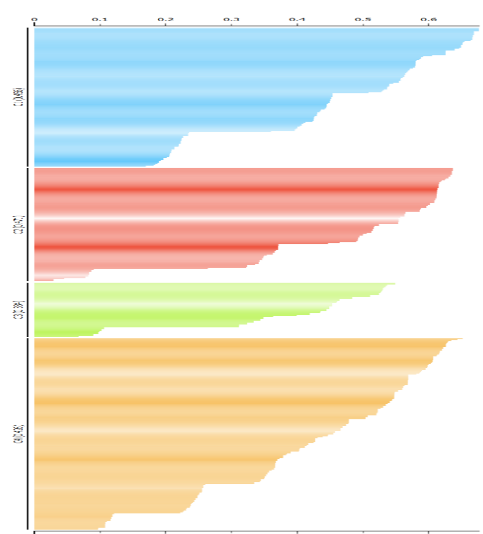


Fig. 24. Silhouette Polt for K-means Clustering

6. Conclusion and Future works

This study has presented examining the usability of mobile applications in higher education using data mining techniques. The system usability scale is used to collect users' data concerning the PSAU mobile application's usability. After that, the data mining methodology has been applied as descriptive and predictive methods. Explanatory models are utilized to examine the current situation, describe the PSAU mobile application's usability, understand and interpret its features, and explore the efficiency, simplicity, and satisfaction of the PSAU mobile application. Predictive models have been developed to forecast the usability of the PSAU mobile application at an acceptable or unacceptable level. The second model predicates usability grades for the PSAU mobile application as A, B, C, D, or F. Grades from A to D mean the usability is good.

In contrast, an F grade means the usability has some problems. Five standard classification algorithms and four clustering approaches are adopted. Most approaches give positive results in all performance indicators. The contribution of the proposed study can be summarized as follows: applied data mining approaches produce promising outcomes in discovering patterns and identifying hidden relationships among features for the dataset related to assessing the usability of the PSAU mobile application. The PSUA mobile application has acceptable usability for most factors, with a ratio rate of 70%. The proposed method consolidates related usability problems into one group with a ratio rate of 30%. The insights obtained will help policymakers, developers, and designers empower the functionality and customize the PSAU mobile application. In future studies, the researcher plans to expand this study by using other data mining techniques such as the association rule in order to discover hidden patterns.

Funding

This research was funded by Prince Sattam bin Abdulaziz University, project number PSAU/2023/02/ 25720.

Acknowledgment

The author extends his appreciation to Prince Sattam bin Abdulaziz University for funding this research work through project number (PSAU/2023/02/ 25720).

References

Abu Saa, A., Al-Emran, M., & Shaalan, K. (2019). Factors affecting students’ performance in higher education: a systematic review of predictive data mining techniques. *Technology, Knowledge and Learning*, pp. 567-598.

Afif, M. (2023). Assessing Students' Perceptions of Mobile Applications Usability using System Usability Scale. *Journal of Computer Science, 19*(1), 11--19. doi: https://doi.org/10.3844/jcssp.2023.11.19

Afif, M. H. (2021). Evaluating PSAU mobile application based on people at the center of mobile application development (PACMAD) usability model: empirical investigation. *Journal of Computer Science, 17*(3), 275--283.

Al-Hagery, M. A., Alzaid, M. A., Alharbi, T. S., & Alhanaya, M. A. (2020). Data mining methods for detecting the most significant factors affecting students’ performance. *International Journal of Information Technology and Computer Science (IJITCS)*, *12*(5), 1-13.

Al-mayyan, H., & Al-Refai, S. (2020). An Approach to Evaluate Software Usability of E-Communication Application. *International Journal of Computer Science and Information Security (IJCSIS), 18*(2), 17-22.

Al-Omar, K. (2018). Evaluating the Usability and Learnability of the" Blackboard" LMS Using SUS and Data Mining. *In 2018 Second International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 386--390). IEEE.

Az-zahra, H. M., Fauzi, N., & Kharisma, A. P. (2019). Evaluating E-marketplace mobile application based on people at the center of mobile application development (PACMAD) usability model. *2019 International Conference on Sustainable Information Engineering and Technology (SIET)* (pp. 72--77). IEEE.

Bangor et al. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies, 4*(3), 114--123.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Intl. Journal of Human--Computer Interaction, 24*(6), 574--594.

Boza, B. C., Schiaffino, S., Teyseyre, A., & Godoy, D. (2014). An Approach for Knowledge Discovery in a Web Usability Context. *In Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems.*

Brooke, J. (1996). SUS A quick and dirty usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester, *Usability Evaluation In Industry* (Vol. 189). CRC Press.

Brooke, J. (2013). SUS: a retrospective. *Journal of usability studies, 8*(2), 29-40.

El-Halees, A. M. (2014). Software Usability Evaluation Using Opinion. *Journal of software, 9*(2).

Fabil, N. B., Saleh, A., & Isamil, R. B. (2015). Extension of PACMAD model for usability evaluation metrics using Goal Question Metrics (GQM) approach. *Journal of Theoretical and Applied Information Technology*.

Goel, S., Nagpal, R., & Mehrotra., D. (2018). Mobile applications usability parameters: Taking an insight view. *Information and Communication Technology for Sustainable Development* (pp. 35--43). Singapore: Springer.

Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. *International Journal of Information Technology, 4*(12), 243-1257.

Han, J., Pei, J., & Kamber, M. (2022). *Data mining: concepts and techniques.* Morgan Kaufmann.

Harrison, R., Flood, D., & Duce, D. (2013). Usability of mobile applications: literature review and rationale for a new usability model. *Journal of Interaction Science*, *1,* 1-16.

Hoehle, H., & Venkatesh, V. (2015). Mobile Application Usability Conceptualization and instrument development. *MIS quarterly, 39*(2), 435-472.

Iso, & ISO. (2018). 9241-11 (2018) Ergonomics of human-system interaction—part 11: usability: definitions and concepts. *International Organization for Standardization. https://www. iso. org/obp/ui/\# iso: std: iso, 9241*(11).

Kantardzic, M. (2011). *Data mining: concepts, models, methods, and algorithms.* John Wiley & Sons.

Kaya, A., Ozturk, R., & Gumussoy, C. A. (2019). Usability measurement of mobile applications with system usability scale (SUS). In *Industrial engineering in the big data era* (pp. 389--400). Springer.

Kortum, P., & Sorber, M. (2015). Measuring the Usability of Mobile Applications for Phones and Tablets. *International Journal of Human–Computer Interaction, 31*(8), 518--529. doi:https://doi.org/10.1080/10447318.2015.1064658

Kumar, B. A., & Goundar, M. S. (2019). Usability heuristics for mobile learning applications. *Education and Information Technologies, 24*(2), 1819--1833.

LIN, B., Cassee, N., Serebrenik, A., Bavota, G., Novielli, N., & Lanza, M. (2022). Opinion mining for software development: a systematic literature review. *ACM Transactions on Software Engineering and Methodology (TOSEM), 31*(3), 1-41.

Lumor, T., Pulkkinen, M., & Hirvonen, A. (2020). The actual adoption and use of mobile apps: The case of a higher education context. *Americas Conference on Information Systems.* Association for Information Systems.

Moumane, K., Idri, A., & Abran, A. (2016). Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. *SpringerPlus, 5*(1), 1--15.

Nielsen, J. (1994). Usability inspection methods. *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 249-256). ACM. doi:https://doi.org/10.1145/97243.97281

Parsazadeh, N., Ali, R., Rezaei, M., & ZolfagharTehrani, S. (2018). The construction and validation of a usability evaluation survey for mobile learning environments. *Studies in Educational Evaluation, 58*, 97-111.

Sagar, K., & Saha, A. (2016). Enhancing Usability Inspection Through Data-Mining Techniques An Automated Approach for Detecting Usability Problem Patterns of Academic Websites. *In International Conference on Intelligent Human Computer Interaction, IHCI 2016. Lecture Notes in Computer Science. 10127*, pp. 229-247. Springer, Cham. doi:https://doi.org/10.1007/978-3-319-52503-7_19

Sagar, K., & Saha, A. (2017). Qualitative usability feature selection with ranking a novel approach for ranking the identified usability problematic attributes for academic websites using data-mining techniques. *Human-centric Computing and Information Sciences*, 1-24.

Silfee, V. J., Waring, M. E., Boudreaux, E. D., Sadasivam, R. S., Mullen, S. P., Carey, J. L., . . . Pagoto, S. L. (2017). Methods for Evaluating the Content, Usability, and Efficacy of Commercial Mobile Health Apps. *JMIR mHealth and uHealth, 5*(12), e8758.

Vlachogianni, P., & Tselios, N. (2022). Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review. *Journal of Research on Technology in Education, 54*(3), 392-409. doi:10.1080/15391523.2020.186793

Weichbroth, P. (2020). Usability of Mobile Applications: A Systematic Literature Study. *IEEE Access, 8*, 55563--55577.

Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, pp. 1-19.