Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/120914

Using Markov Chains and Data Mining Techniques to Predict Students' Academic Performance

Saed Mallak[1], Mohammad Kanan[2,], Nidal Al-Ramahi[3], Aya Qedan[1], Hadi Khalilia[1], Ahmad Khassati[1], Rania Wannan[1], Mohammad Mara'beh[1], Samer Alsadi[1], and Abdalmuttaleb Al-Sartawi[4]*

[1]Electrical Engineering Department, Faculty of Engineering and Technology, Palestine Technical University, Kadoorie, Tulkarm, Palestine
[2]Industrial Engineering Department, Jeddah College of Engineering, University of Business and Technology, Jeddah 21448, Saudi Arabia
[3]Department of accounting, Business Faculty, Zarqa University, Zarqa 11831, Jordan
[4]Accounting Finance & Banking Department, College of Business & Finance, Ahlia University, Manama, Bahrain

Received: 2 Jul. 2023, Revised: 10 Aug. 2023, Accepted: 28 Aug. 2023.
Published online: 1 Sep. 2023.

Abstract: In this study, the academic performance of students from the E-Commerce department at Palestine Technical University – Kadoorie is predicted using a Markov chains model and educational data mining. Based on the complete data regarding the achievements of the students from the 2016 cohort of students obtained from the university's admissions and registration department, a Markov chain is built, in which the states are divided according to the semester average of the student, and the ratio of students in each state is calculated in the long run. The results obtained are compared with the data from the 2015 cohort, which demonstrates the efficiency of the Markov chains model. For educational data mining, the classification technique is applied, and the decision tree algorithm is used to predict the academic performance of the students, generalizing results with an accuracy of 41.67%.

Keywords: Prediction, Markov Chains, Academic Performance, Data Mining, Educational Data Mining, Decision Tree.

1. Introduction

Any educational institution that seeks to improve both its teaching and learning process and the development of its students and their achievements must first develop its ability to predict the academic performance of students. The traditional method of evaluating students focuses only on students' past achievements, but this lacks the ability to predict students' future development. Therefore, it is necessary to change the way in which grades for future academic performance are predicted; this is one of the most important things that all educational institutions must seek to strengthen and develop within their education administration [3].

Markov chains and fuzzy Markov chains using different approaches and different fuzzy numbers have proven to be effective tools in prediction, and many authors have used them in different areas, including [7], [9], [12–19], and [24–26].

Data mining has been used in the educational sector by extracting characteristics from databases that are most related to each other and have the greatest impact on a student's level or by predicting a student's academic performance based on their academic data or data from former students. It is also possible to sort students into groups according to their level in such a way that an instructor can determine a way to deal with each group according to its level, as in [8] and [10]. Many authors have used data mining in prediction in different areas, including [1–5], [6], [11], [20–21] and [23].

In this paper, the academic performance of students from the E-Commerce department at Palestine Technical University – Kadoorie is predicted using Markov chains and educational data mining. This is made possible by the fact that the standards and conditions are almost identical from year to year.

2. Markov Chains

A discrete time Markov chain is a Markov process in which the state space $\boldsymbol{S}$ (the range of possible values for the random variables $\boldsymbol{X_t}$) is a finite or countable set and the time index set is T={0,1,2,…}[22].

The Markov property is:

$$\boldsymbol{p(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i)}$$

*Corresponding author e-mail: m.kanan@ubt.edu.sa

$$= p(X_{n+1} = j | X_n = i) = p_{ij}^{n,n+1}$$

(the one step transition probability).

When the one step transition probabilities are independent of time n, the Markov chain is said to have stationary transition probabilities or be homogeneous, i.e., $p_{ij}^{n,n+1} = p_{ij}$.

The Markov matrix or transition probability matrix $P = [p_{ij}]_{i,j\in S}$ of the process satisfies $p_{ij} \geq 0, \forall i, j \in S$ and $\sum_j p_{ij} = 1$.

A Markov process is completely defined once its matrix and initial state X_0 are specified $p(X_0 = i) = p_i$.

The classifications of the states can be seen in [22].

When the state space is finite and $\exists \mathbf{n}, \mathbf{p}_{ij}^{(n)} > 0, \forall \mathbf{i}, \mathbf{j} \in \mathbf{S}$, then all states are positive recurrent and aperiodic, i.e., it is an Ergodic Markov chain.

With Ergodic Markov chains, there exists a probability vector $\pi = (\pi_j)_{j\in S}$ in which:

1) $\pi_j > 0, \forall j$ 2) $\sum_{j\in S} \pi_j = 1$ 3) $\pi_j = \sum_{i\in S} \pi_i p_{ij}$ 4) $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$.

The convergence means that in the long run ($n \to \infty$), the probability of finding the Markov chain in state j is approximately π_j, regardless of the state in which it began at time zero This is the stability rule used in the practical part of this study.

Forecasting is an essential axis in the decision-making process, and it is one of the most important goals of the Markov model, for which it was designed. The prediction step depends mainly on the transitional probability matrix, and for the matrix to be able to predict, it must be Ergodic.

3. The Practical Part of Markov Chains (The Model)

3.1 Data

The sample for this study was the 2016 cohort of students from the department of business administration and electronic commerce. The results obtained by these students during the eight semesters from 2016/1 to 2019/2 were examined. However, the eighth semester was later excluded in response to the COVID-19 pandemic and its effects on the education process in universities; thus, the results of students that were analyzed were limited to those obtained during seven semesters.

The group of students was divided into five categories based on the average GPA of the student in each semester as follows: A (90–100), B (80–89), C (70–79), D (65–69), and E (less than 65).

The matrix of transitional probabilities of a Markov chain with states of these five categories was formed, where the transition probability p_{ij} represents the probability of reaching state j from state i in one step (one semester). The predicted values based on the limits of the probability transition matrix were obtained.

The results were then compared with the relevant data from the 2015 batch from the same department.

3.2 Modulation the Transition Matrix

The transition matrix is 5*5; the group of students from each department was divided into five categories based on the assessment of the students' average in each semester according to what is accredited at Kadoorie University, as shown in Table 1.

Table 1: Categories Based on the Assessment of the Semester Average

Number	Class GPA range	Category	GPA assessment
1	90≤GPA≤100	A	Excellent
2	80≤GPA<90	B	Very Good
3	70≤GPA<80	C	Good
4	65≤GPA<70	D	Satisfactory
5	GPA<65	E	Failed/Not Normal

From this, the transition probability matrix is:

$$P = \begin{bmatrix} P_{AA} & P_{AB} & P_{AC} & P_{AD} & P_{AE} \\ P_{BA} & P_{BB} & P_{BC} & P_{BD} & P_{BE} \\ P_{CA} & P_{CB} & P_{CC} & P_{CD} & P_{CE} \\ P_{DA} & P_{DB} & P_{DC} & P_{DD} & P_{DE} \\ P_{EA} & P_{EB} & P_{EC} & P_{ED} & P_{EE} \end{bmatrix}$$

To determine the transition probability matrix P related to the transition of student's semester GPA from one assessment category to another for the 2016 batch (during the seven semesters from 2016/1 to 2019/1), the students' data were analyzed and sorted. The transition of the semester GPA from one assessment category to another during these levels (the semesters) was traced, and these moves were assembled in a raw matrix called N.

The elements of each row were then divided by the total sum of that row, producing the transition probability matrix.

Finally, the limiting probability distribution π was calculated after verifying that the transition probability matrix was regular (Ergodic), which allowed the number of graduates in each class to be predicted.

3.3 Application and Practical Implementation

After sorting and analyzing the results of the students from the applied mathematics department in the 2016 batch and tracking the transfer of the semester GPA for all students through the seven levels (semesters), it was found that there were:

- 10 transitions from A to A, 10 transitions from A to B, and no transitions from A to any of C, D and E.
- 12 transitions from B to A, 55 transitions from B to B, 25 transitions from B to C, 1 transition from B to D, and 4 transitions from B to E.
- No transitions from C to A, 21 transitions from C to B, 85 transitions from C to C, 46 transitions from C to D, and 29 transitions from C to E.
- No transitions from D to A, 3 transitions from D to B, 33 transitions from D to C, 34 transitions from D to D, and 33 transitions from D to E.
- No transitions from E to A, no transitions from E to B, 22 transitions from E to C, 33 transitions from E to D, and 62 transitions from E to E.

Thus the raw matrix N is:

$$N = \begin{bmatrix} 10 & 10 & 0 & 0 & 0 \\ 12 & 55 & 25 & 1 & 4 \\ 0 & 21 & 85 & 46 & 29 \\ 0 & 3 & 33 & 34 & 33 \\ 0 & 0 & 22 & 33 & 62 \end{bmatrix}$$

The elements of each row were then divided by the total sum of that row, producing the transition probability matrix:

$$P = \begin{bmatrix} \frac{10}{20} & \frac{10}{20} & 0 & 0 & 0 \\ \frac{12}{97} & \frac{55}{97} & \frac{25}{97} & \frac{1}{97} & \frac{4}{97} \\ 0 & \frac{21}{181} & \frac{85}{181} & \frac{46}{181} & \frac{29}{181} \\ 0 & \frac{3}{103} & \frac{33}{103} & \frac{34}{103} & \frac{33}{103} \\ 0 & 0 & \frac{22}{117} & \frac{33}{117} & \frac{62}{117} \end{bmatrix}$$

$$= \begin{bmatrix} 0.5000 & 0.5000 & 0 & 0 & 0 \\ 0.1237 & 0.5670 & 0.2577 & 0.0103 & 0.0413 \\ 0 & 0.1160 & 0.4696 & 0.2541 & 0.1603 \\ 0 & 0.0291 & 0.3204 & 0.3301 & 0.3204 \\ 0 & 0 & 0.1880 & 0.2821 & 0.5299 \end{bmatrix}$$

Note that the transition matrix is Ergodic. Hence: $\lim_{n\to\infty} P^n = \pi \ (*)$.

To reach the stability of the matrix, the equation (*) was applied, and the matrix was found to become stable at n=50. Matlab was used to reach the following result:

$$P^{43}=\begin{bmatrix} 0.0343 & 0.1388 & 0.3104 & 0.2378 & 0.2787 \\ 0.0343 & 0.1388 & 0.3104 & 0.2378 & 0.2787 \\ 0.0343 & 0.1388 & 0.3104 & 0.2378 & 0.2787 \\ 0.0343 & 0.1388 & 0.3104 & 0.2378 & 0.2787 \\ 0.0343 & 0.1388 & 0.3104 & 0.2378 & 0.2787 \end{bmatrix}$$

From the above matrix, it should be noted that a steady state is reached after 43 semesters,

The limiting probability distribution is as follows: π=(0.0343, 0.1388, 0.3104, 0.2378, 0.2787)

- $\pi_1 = 0.0343 = 3.43\%$, the rate of graduates with GPA A.
- π_2 =0.1388= 13.88%, the rate of graduates with GPA B.
- π_3 =0.3104=31.04%, the rate of graduates with GPA C.
- π_4 =0.2378=23.78%, the rate of graduates with GPA D.
- π_5 =0.2787=27.87%, the rate of students who will not graduate (GPA E).

Upon closer inspection, it should also be noted that the matrix P^8 indicates the probability of a student's GPA transferring from one assessment category to another after eight semesters.

$$P^8 = \begin{bmatrix} 0.0781 & 0.2314 & 0.2919 & 0.1858 & 0.2131 \\ 0.0573 & 0.1876 & 0.3007 & 0.2103 & 0.2445 \\ 0.0326 & 0.1351 & 0.3108 & 0.2395 & 0.2823 \\ 0.0277 & 0.1246 & 0.3127 & 0.2453 & 0.2898 \\ 0.0250 & 0.1188 & 0.3137 & 0.2486 & 0.2941 \end{bmatrix}$$

This transition probability matrix indicates the probability of a student obtaining a bachelor's degree in applied mathematics at the end of the normal period for obtaining a degree, which is four academic years (equivalent to eight semesters).

For the probability vector π_0, the initial probability distribution, which is obtained by dividing the GPA of the first semester for students of the 2016 batch by the total, was found to be: π_8 =(0.0361 0.1425 0.3093 0.2354 0.2767).

To check these results, they were compared with data from the 2015 batch obtained from the university's admission and registration department. There were 91 students in the 2015 batch, including 74 graduates. Tables 2 and 3 show the distribution of these students according to their graduation GPA.

Table 2: Graduates from 2015 Distributed According to their Graduation GPA

Graduation GPA	A	B	C	D	E
Number of graduates	2	17	45	10	17

Table 3: Expected Number of Graduates from 2015 in Each Category at the End of the Normal Period of Study

Graduation GPA	A	B	C	D	E
Graduates 2015	2	17	45	10	17
Π*n	3.1031	12.5853	28.2282	21.5943	25.4891
E_1	1.1031	4.4147	16.7718	11.5943	8.4891
π_8*n	3.2851	12.9675	28.1463	21.4214	25.207
E_2	1.2851	4.0325	16.8537	11.4214	8.207

From these results, it is clear that at the end of the normal period of study, the ratios of graduates are very close to each other.

4. Decision Tree Algorithm

Data mining can be defined in different ways. Educational data mining (EDM) is an emerging discipline that focuses on the application of data mining tools and techniques to educationally related data. The discipline focuses on the analysis of educational data to develop models for improving learning experiences and improving institutional effectiveness [10].

There are many data mining techniques used in the field of education to extract knowledge from educational data and present it to the decision maker in order to improve the education process. The decision tree is one of the easiest and most popular classification algorithms used in understanding and interpreting data. It can also be used for classification and prediction.

A decision tree is a flowchart-like tree structure where the inner node represents a feature (or attribute), the branch represents a decision rule, and each leaf node represents the outcome. The top node in a decision tree is known as the root node. It learns to segment based on the attribute value, and the tree is split in a recursive way known as recursive partitioning. This flowchart-like structure helps in the decision-making process; its visualization like a flowchart stimulates the level of human thinking necessary to understand and interpret data efficiently [11]. Decision trees are inexpensive and easy to build and read, and they work to split and segment data. The decision tree is formed by setting the appropriate and logical question that divides the data into two parts: the first part of the data applies to the question and the second part of the data does not apply to the question. It should, however, be noted that in cases where the sheet that contains data consists of one record (which means that it cannot be divided) or contains identical records (which means that the division of data will not lead to new and useful information), it is stopped from the beginning.

The attribute selection scale is a guide for defining a segmentation criterion that splits data in the best possible way. It is also known as partitioning rules because it helps the determination of breakpoints of clusters on a particular node. ASM provides a rank for each feature (or trait) by explaining the specific data set. The best degree attribute will be selected as the split (source) attribute. In the case of an attribute with a continuous value, the split points of the branches must also be defined. In tree construction, attribute selection is based on two measures: information gain and the Gini index [8] and [11].

The Gini index is used in the Classification and Regression Tree algorithm. The impurity of D, a data partition or set of training tuples, can be measured by the Gini index, as follows:

$$Gini\ (D) = 1 - \sum_{i=1}^{m} {P_i}^2$$

Where P_i is the probability that a tuple in D belongs to class C_i.

The Gini index considers a binary split for each attribute; thus, a weighted sum of the impurity of each partition can be computed. If a binary split on attribute A partitions data D into D1 and D2, the Gini index of D is:

$$Gini_A\,(\mathrm{D}) = \frac{|D_1|}{|D|}\ Gini(D_1) + \frac{|D_2|}{|D|}\ Gini(D_2)$$

In the case of a discrete-valued attribute, the subset that gives the minimum Gini index for that chosen is selected as a splitting attribute. However, in the case of continuous-valued attributes, each pair of adjacent values is selected as a possible split-point and the point with a smaller Gini index chosen as the splitting point.

The impurity reduction of a binary split on a discrete- or continuous-valued attribute A is shown in the following equation:

$$\Delta Gini(A) = Gini(D) - \ Gini_A(D)$$

The attribute with the minimum Gini index is chosen as the splitting attribute [8].

5. Experiment

This study has made use of data regarding the 2016 cohort of students from the department of E-Commerce. The semester GPA for each student during eight consecutive semesters was examined, and the students were divided into groups depending on whether the GPA increased, decreased, or remained stable. The training data used by the classification technique in this thesis was created and a decision tree was formed. A prediction model was then built using the training data set, and the model and its accuracy was tested using data from the 2015 cohort of students; these data were given to the model to predict the results, and the predicted results were compared with the real results to reach generalizations and determine the accuracy.

5.1 Data Set

The data set used in this study was obtained from the admission and registration department at the Palestine Technical University – Kadoorie. The sample was formed from the group of graduating students from the 2016 cohort from two departments within each of the following three faculties:

1) Faculty of Applied Sciences: Department of Applied Mathematics and Department of Applied Computing.

2) Faculty of Business and Economic: Department of Business Administration and Electronic Commerce and Department of Industrial Management.

3) Faculty of Engineering and Technology: Department of Industrial Automation Engineering and Department of Computer Engineering.

Students who graduated after eight semesters were selected only if their data was complete; students whose data was lacking, such as students who changed major or postponed one of the semesters, were excluded.

The results of students were followed during eight semesters from 2016/1 to 2019/2, and the group of students from each department was divided into five categories, based on the assessment of the semester average, as follows: A (90–100), B (80–89), C (70–79), D (65–69), and E (less than 65).

In the next step, a nominalization for the cases in which the semester GPA of the student changed category from one semester to the next was made as follows:

- An increase in the assessment of a student's semester GPA is symbolized with (1).
- A decrease in the assessment of a student's semester GPA is symbolized with (-1).
- Where a student's semester GPA remains the same, it has been symbolized with (0).

A Microsoft Excel spreadsheet was used to save and arrange the data for easy export to PyCharm later.

The E-Commerce department was chosen to present the idea of classification technique in the EDM, as the number of records in the Excel sheet for this major was the largest. The other departments were excluded because the number of records in their Excel sheets was insufficient to allow a strong decision tree to be formed. Thirty-one students from the 2016 cohort of E-Commerce majors graduated in the second semester of 2019; Table 4 shows the distribution of these students according to the assessment of their graduation GPA.

Table 4: Graduates of the 2016 Cohort Distributed According to Their Graduation GPA

Graduation GPA	A	B	C	D
Number of graduates	1	6	21	3

The semester GPA for each student during eight consecutive semesters was followed from 2016/1 to 2019/2; this was stored in Table 5, which contains seven columns (representing features). The last column represents the decision or result and is represented by an assessment of the student's graduation GPA (which is predicted later).

5.2 Experiment Setup

The results for the 2016 cohort of students during eight semesters from 2016/1 to 2019/2 were followed, and the students were split into groups based on whether their semester GPA increased, decreased, or remained stable. An increase in the assessment of a student's semester GPA is symbolized by (1); a decrease in the assessment of a student's semester GPA is symbolized by (-1); and in the case where a student's semester GPA remained the same, it is symbolized with (0).

Microsoft Excel was used to save and arrange the data for easy export to PyCharm. The strength of the Python language lies in its ease of learning and programming and its speed of work, and it has proven its effectiveness and efficiency in several areas. In addition, it also uses many important and diverse packages and libraries.

5.3 Methodology

A decision tree that goes through two steps was used as the classification algorithm. In the first stage, the training data was given to the algorithm to build a model that could later be used in prediction. In the second stage, the model obtained was tested by giving it a set of tested data; this enabled the accuracy of the results given by the model to be determined.

Table 5: Status of Assessment of the Semester GPA for Each Student During Eight Consecutive Semesters from 2016/1 to 2019/2

ID	ST.1	ST.2	ST.3	ST.4	ST.5	ST.6	ST.7	GGPA
40	0	0	-1	1	0	-1	1	C
72	0	1	0	0	-1	0	1	D
81	-1	0	0	0	0	0	1	C
208	1	-1	-1	1	-1	1	1	C
267	0	-1	1	-1	-1	1	0	C
327	0	0	-1	1	-1	1	1	C
352	-1	0	0	1	-1	0	1	D
389	1	1	0	-1	1	1	-1	C
454	0	0	1	-1	-1	1	1	C
459	1	-1	0	1	0	-1	1	C
483	0	0	0	-1	0	1	-1	D
524	0	0	0	0	-1	1	1	C
684	-1	1	-1	0	1	-1	1	C
710	0	0	1	-1	1	0	0	C
731	0	0	0	0	0	1	0	C
851	0	1	0	-1	1	0	-1	B
904	1	-1	0	0	0	-1	1	C
924	1	-1	0	0	1	0	1	C
929	0	0	0	0	0	0	0	B
930	-1	0	1	0	0	1	0	C
951	0	0	0	0	-1	1	0	B
964	1	0	0	0	1	0	-1	B
972	-1	0	0	0	0	1	-1	B
1025	0	-1	0	0	0	-1	1	C
1036	1	-1	1	0	0	0	0	A
1289	1	-1	1	0	-1	0	1	C
1382	1	0	-1	0	0	-1	1	C
1384	0	-1	0	-1	0	0	1	C
1386	-1	-1	1	0	0	0	1	C
1419	1	0	0	-1	1	0	1	C
1564	-1	1	-1	0	1	0	0	B

5.3.1 Training Phase

The data was trained in order to build a model that could later be used for the purpose of prediction. The Excel file containing the training data was exported to PyCharm in order to train the classification algorithm; a code was applied to create the decision tree shown in Figure 1.

Fig. 1: The Decision Tree

5.3.2 Testing Phase

In the testing phase, the decision tree was tested to determine its accuracy. To achieve this, data regarding the 2015 cohort of E-Commerce students (provided by the admission and registration department at Palestine Technical University) was used. Thirty-six students from the 2015 cohort graduated in the second semester of 2018, after eight semesters; Table 6 shows their results, distributed according to the assessment of the graduation GPA.

Table 6: Graduates From the 2015 Cohort Distributed According to Their Graduation GPA

Graduation GPA	A	B	C	D
Number of graduates	1	9	24	2

The assessment status of the semester GPA for each student from this cohort was examined; Table 7 contains the results of this process.

Table 7: Assessment Status of the Semester GPA for Each Student During Eight Consecutive Semesters from 2015/1 to 2018/2

ID	ST.1	ST.2	ST.3	ST.4	ST.5	ST.6	ST.7	GGPA
5	0	0	0	0	1	-1	0	C
12	-1	0	-1	1	-1	1	0	C
15	0	0	1	-1	0	0	0	B
16	0	0	1	-1	-1	1	0	C
17	-1	1	0	0	0	0	0	B
20	-1	0	0	0	0	0	0	C
21	0	1	0	-1	0	1	0	B
28	0	1	-1	0	0	0	0	C
29	-1	-1	1	0	-1	0	1	D
32	-1	0	1	-1	1	-1	0	C
34	-1	1	1	-1	-1	1	0	D
38	-1	1	-1	0	0	-1	1	C
41	1	1	-1	0	0	1	-1	C
50	1	-1	1	0	-1	1	0	C
59	-1	1	-1	1	0	1	0	C
61	0	1	0	-1	0	0	0	C
300	-1	0	1	-1	0	0	0	C
453	0	1	-1	1	0	0	0	B
796	0	0	0	-1	1	0	0	B
1121	1	-1	0	0	-1	1	0	C
1172	-1	1	0	-1	0	0	1	B
1291	1	-1	1	0	0	1	0	C
1323	1	-1	1	1	-1	0	0	C
1339	1	-1	0	1	0	-1	0	B
1387	0	-1	1	-1	0	-1	1	C
1447	-1	0	1	0	0	-1	1	C
1455	0	0	0	0	0	0	0	A
1456	0	0	0	0	0	0	0	B
1475	0	0	1	-1	0	0	0	B
1674	-1	1	1	-1	1	-1	1	C
1781	0	1	-1	0	0	1	-1	C
1786	-1	1	-1	0	1	0	0	C
1916	1	1	0	-1	1	1	0	C
2091	-1	1	0	-1	-1	1	0	C
2145	-1	1	-1	1	-1	0	0	C
2172	0	0	0	-1	0	-1	1	C

Another code was then applied to obtain the following expected graduation GPA results:

- GPA A: 3
- GPA B: 17
- GPA C: 15
- GPA D: 1

This prediction has an accuracy of 0.4167. With an accuracy of 41.67%, it is possible to predict that the percentage of graduates with a GPA of A is 8.33%, the percentage of graduates with a GPA of B is 47.22%, the percentage of graduates with a GPA of C is 41.67%, and the percentage of graduates with a GPA of D is 2.78%.

Data mining techniques are very useful. More accurate results can be obtained, and thus a stronger model for prediction, when there is a larger volume of data; this is one of the primary reasons for the emergence of this science.

6. Conclusion

After examining the complete data regarding the achievements of the 2016 cohort of students from the E-Commerce department, the following conclusions can be made:

1. The academic performance of students stabilizes at (n=43), and it is expected that the percentage of graduates for this department will be 3.43%, 13.88%, 31.04% and 23.78% with a GPA of A, B, C, D, respectively, while the percentage of students who are not expected to graduate (because they are under academic probation and banned from graduating) will be 27.87%.
2. A classification model has been proposed related to data mining for predicting students' academic performance. A decision tree algorithm was selected for the construction of this model, which obtained an accuracy of 41.67%. The expected graduate rate is 8.33%, 47.22%, 41.67%, 2.78% with a GPA of A, B, C, D, respectively.
3. The two proposed models help the administration to identify strong and weak departments so that appropriate decisions can be taken to raise the level of student performance and improve the educational process in order to increase the number and quality of graduates at the university.

Table 8 shows a comparison of the two models presented in this study for the department of E-Commerce.

Table 8: Comparison of the Two Models

Graduation GPA	A	B	C	D
Percentage of graduates (Real data)	4.69%	23.44%	57.81%	14.06%
Predicted percentage of graduates using the Markov model	3.43%	13.88%	31.04%	23.78%
Predicted percentage of graduates by data mining model	8.33%	47.22%	41.67%	2.78%
E_1	1.26%	9.56%	26.77%	9.72%
E_2	3.64%	23.78%	16.14%	11.28%

The results of the Markov model are more accurate in this study.

Conflict of interest

The authors declare that there is no conflict regarding the publication of this paper.

References

[1] Amjad Abu Saa, "Educational Data Mining & Students' Performance Prediction" International Journal of Advanced Computer Science and Applications (IJACSA), 7(5), 2016. http://dx.doi.org/10.14569/IJACSA.2016.070531.

[2] Abeer Badr El Din Ahmed and Ibrahim Sayed Elaraby, "Data Mining: A prediction for Student's Performance Using Classification Method," World Journal of Computer Application and Technology (CEASE PUBLICATION), 2(2), pp. 43–47, 2014. DOI: 10.13189/wjcat.2014.020203.

[3] Amal Alhassan, Bassam Zafar and Ahmed Mueen, "Predict Students' Academic Performance based on their Assessment Grades and Online Activity Data," International Journal of Advanced Computer Science and Applications (IJACSA), 11(4), 2020. DOI: 10.14569/IJACSA.2020.0110425.

[4] Abdulmohsen Algarni, "Data Mining in Education," International Journal of Advanced Computer Science and Applications (IJACSA), 7(6), 2016. DOI: 10.14569/IJACSA.2016.070659.

[5] Mohd Maqsood Ali, "Role of Data Mining in Education Sector," International Journal of Computer Science and Mobile Computing, 2(4), pp. 374–383, 2013.

[6] Brijesh Kumar Baradwaj and Saurabh Pal, "Mining Educational Data to Analyze Students Performance," International Journal of Advanced Computer Science and Applications (IJACSA), 2(6), 2011. DOI: 10.14569/IJACSA.2011.020609.

[7] Alenka Brezavšček, Mirjana Pejić Bach and Alenka Baggia, "Markov Analysis of Students' Performance and

Academic Progress in Higher Education", Organizacija, 50(2), pp. 83–95, 2017. DOI: 10.1515/orga-2017-0006.

[8] Jiawei Han, Micheline Kamber and Jian Pei, *Data mining: Concepts and techniques*, 3rd edition, Elsevier, 2012.

[9] Robert Hlavatý and Ludmila Dömeová, "Students' Progress throughout Examination Process as a Markov Chain," International Educations Studies, 7(12), 2014. DOI: 10.5539/ies.v7n12p20.

[10] Richard A. Huebner, "A Survey of Educational Data-Mining Research," Research in Higher Education Journal, 19, 2013.

[11] Hadi Khalilia, Thaer Sammar and Yazeed Sleet, "Predicting Students Performance Based on Their Academic Profile," Palestine Technical University Research Journal, 8(2), pp. 23–39, 2020.

[12] Saed Mallak and Deema Abdoh, "Predicting the Behavior of Gold Price Using Markov Chains and Markov Chains of the Fuzzy States," Mathematical Statistician and Engineering Applications, 71(4), pp. 2906–2920, 2022.

[13] Saed F. Mallak, Ghadeer A. Al-Mallak, Nasouh Sos and Saed Khayat, "Using Markov Chains Models to Predict Water Supply and Demand and their Behavior: Case Study from Tulkarm City," Palestine Journal of Mathematics, 12(1), pp. 608–619, 2023.

[14] Saed F. Mallak and Duha Bedo, "A Fuzzy Comparison Method for Particular Fuzzy Numbers," Journal of Mahani Mathematical Research Center (JMMRC), 2(1), pp. 1–14, 2013.

[15] Saed F. Mallak and Duha Bedo, "Particular Fuzzy Numbers and a Fuzzy Comparison Method between Them," International Journal of Fuzzy Mathematics and Systems (IJFMS), 3(2), pp. 113–123, 2013.

[16] Saed Mallak, Mohammad Kanan, Samer Alsadi, Ghadeer Sabbah, Siraj Zahran, Hani Attar and Allam Hamdan, "Predicting the Behavior of Solar Energy in Tulkarm City Using Markov Chains and Fuzzy Markov Chains," Appl. Math. Inf. Sci., 17(2), pp. 285–292 (2023).

[17] Saed F. Mallak, Mohammad Mara'Beh and Abdelhalim Zaiqan, "A Particular Class of Ergodic Finite Fuzzy Markov Chains," Advances in Fuzzy Mathematics, 6(2), pp. 253–268, 2011.

[18] Vivian R. Moody and Kanita K. DuCloux, "Application of Markov Chains to Analyze and Predict the Mathematical Achievement Gap between African American and White American Students," Journal of Applied & Computational Mathematics, 3(3), 2014. DOI: 10.4172/2168-9679.1000161.

[19] Murtala Adam Muhammad, Jamilu Yunusa Falgore and Usman Hashim Sani, "Analysis of Student' Academic Performance and Progression (Using Markov Chain Approach)," International Journal of Engineering Applied Sciences and Technology, 4(7), pp. 187–193, 2019.

[20] Ajay Kumar Pal and Saurabh Pal, "Analysis and Mining of Educational Data for Predicting the Performance of Students," International Journal of Electronics Communication and Computer Engineering, 4(5), pp. 1560–1565, 2013.

[21] Umesh Kumar Pandey and Saurabh Pal, "Data Mining: A Prediction of Performer or Underperformer Using Classification," International Journal of Computer Science and Information Technologies (IJCSIT), 2(2), pp. 686–690, 2011.

[22] Howard M. Taylor and Samuel Karlin, *An Introduction to Stochastic Modeling*, Library of Congress Cataloging-in-Publication Data. San Diego, California, 3rd ed., 1998.

[23] Surjeet Kumar Yadav, Brijesh Bharadwaj and Saurabh Pal, "Data Mining Applications: A Comparative Study for Predicting Student's Performance," International Journal of Innovative Technology & Creative Engineering, 1(12), 2012.

[24] Yehia Abd. Yehia, "Prediction of the Probabilities for Changing of the Agricultural Loans Using Markov Chain Model," Egyptian Journal of Agricultural Research, 96(3), pp. 1229–1257, 2018.

[25] Xu Zhang, Ruojuan Xue, Bin Liu and Yiqun Zhang, "Grade Prediction of Student Academic Performance with Multiple Classification Models," 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp. 1086–1090, 2018.

[26] Kanan M, Zerban A, Abunar S, El Harbi A, Weheba G, Ahmed RA, Abdultawwab M, Haddad T. Online Education and Managing Service Quality with the Challenges of COVID 19: The Case of University of Business and Technology (UBT) Saudi Arabia. Appl. Math. 2023;17(2):201-7.