DOI: 10.15849/IJASCA.240330.09

Int. J. Advance Soft Compu. Appl, Vol. 16, No. 1, March 2024
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Creating Visual Knowledge Representation Based on Data Mining in Educational Jordanian Databases

Muzhir Al-Ani[1], Qeethara Al-Shayea[2], Abdul Razzak Alshehadeh[3], Bintang Annisa Bagustari[4], Haneen A. Al-khawaja [5,6]

[1]Department of Information Technology, College of Science and Technology, University of Human Development, Sulaymaniyah, Iraq, Email: muzhir.al-ani@uhd.edu.iq
[2]MIS Department, Faculty of Business, Al-Zaytoonah University of Jordan, Amman, Jordan, Email: drqeethara@zuj.edu.jo
[3]Faculty of Business, Al-Zaytoonah University of Jordan, Amman, Jordan, Email: abdulrazzaqsh@zuj.edu.jo
[4]Department of Computer Science, Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia. Email: bintangnisa@ui.ac.id
[5]Department of Financial Technology and Banking, Faculty of Business, Ajloun National University, Jordan.
[6]Applied Science Research Center, Applied Science Private University, Jordan

Abstract

Extracting insights from educational data offers valuable business intelligence and web discovery solutions. Predicting student trajectories presents a significant challenge for higher education institutions today. They aim to determine which universities have the highest employment rates for their graduates and to understand if certain students are more likely to secure jobs than others. Accurate statistics are essential for planning, trend forecasting, and optimizing business processes. To tackle these challenges, we suggest an approach that employs historical data to classify the most suitable major for each student. This is achieved using the decision tree method and the Iterative Dichotomiser3 (ID3) algorithm. We've integrated a dashboard tailored specifically for Jordanian higher education institutions to visualize the knowledge derived from educational databases. Our research aims to assist higher education and training institutes in navigating the intricate shifts in the labor market. This paper emphasizes using the ID3 decision tree algorithm, a classification technique, to identify categories of workers. Such an approach aids companies in pinpointing the student groups with the highest likelihood of securing employment.

Keywords: *data mining, ID3, decision support system, classification, decision tree.*

1 Introduction

Jordan's education sector faces a significant challenge in terms of growth, while business and government leaders seek to identify the country's competitive advantage compared to other regional and global players. To better understand the importance of having a recommender system that helps students choose their majors, institutions need to focus on

Received 4 December 2022; Accepted 29 February 2024

marketing and promotional efforts and develop strategies to expand and diversify Jordan's higher education offerings.

Our study aims to help higher education and training institutes cope with the increasing complexity and pace of change in labor markets by assisting students in selecting the most suitable major. The interaction between knowledge discovery and visualization will result in a decision support system for various end users, from students to decision-makers in different institutions.

We use the Decision Tree algorithm in our approach, as it has been successfully used in expert systems to capture knowledge. The main task is to use inductive methods to determine the appropriate classification of an unknown object based on its attribute values according to decision tree rules.

Our study seeks to extract useful and important information about majors, universities, and faculties, tailored to each user. This includes a real-life decision support system that provides genuine advice to students, and a dashboard system to assist executives in making decisions based on historical data. To achieve our goals, we collect and clean data from higher education databases from 1999 to 2009 and apply the ID3 decision tree algorithm to examine the results of past decisions regarding which major to study. The inferred information should increase our knowledge about each group of students. Finally, the collected records will serve as the source of data to test the expected results, with implicit preprocessing modules employed to extract knowledge.

2 LITERATURE REVIEW

Over the past few years, the education sector in Jordan has undergone significant development. However, there is still a pressing need for a wider range of academic university majors to meet the demands of the job market. The absence of any form of student academic advisory system exacerbates the difficulties faced by many students when selecting the appropriate major. Moreover, education plays a crucial role in a country's progress and development as it contributes to the refinement and organization of its people. Thus, this study aims to use data mining techniques to extract knowledge from educational databases to analyze student trends and behaviors toward education. By doing so, it may be possible to address the lack of understanding and knowledge of the higher education system, which hinders management systems from achieving quality objectives. Ultimately, data mining technology has the potential to bridge the gaps in higher education systems [1-3].

In recent times, data mining techniques have been utilized to improve and evaluate higher education tasks. The educational sector offers great potential for various research trends to be explored. Data mining techniques provide many opportunities to examine student performance and study patterns [4, 5].

A significant advancement in data mining is the ability to create decision support systems that extract knowledge from various historical records and apply data mining tools and techniques to develop models that can help students determine their educational paths. This development is crucial for supporting the decisions made by decision-makers in different educational institutions [6]. Data mining is widely used in various fields such as banking, telecommunications, and fraud detection [7, 8].

Statistics play a vital role in the everyday life of an average business person. It enables them to make better decisions by allowing them to calculate risk and uncertainty when all the facts are unknown or cannot be collected. Although business decisions are still informed guesses, having more and better data, along with a better understanding of statistics, leads to better decision-making [9].

Search engines play a significant role in today's computing environment, allowing users to extract hidden facts and figures that are processed through well-defined decision support systems (DSS). Data mining relies on scalable statistics, artificial intelligence, and machine learning, while decision support systems utilize available information and data mining technologies to provide a decision-making tool that usually relies on human-computer interaction. DSS and DMT represent the spectrum of analytical information technologies (AIT) and provide a unified platform for an optimal combination of data-driven and human-driven analytics [10, 11].

The decision-tree approach is a popular logic method used to create classifiers from data efficiently. A multitude of decision-tree induction algorithms are described in machine learning and applied statistics literature. The decision tree method is known for its robustness and learning efficiency, with a learning time complexity of O(nlog2n), where n represents the number of rows in the training set. Quinlan's ID3 is a well-known tree-growing algorithm that generates decision trees based on univariate splits. Its extended versions, C4.5 and C5.0, are also popular. These algorithms typically use greedy search methods to grow and prune decision-tree structures, allowing them to explore the vast space of possible models and remove redundant preconditions. They employ a divide-and-conquer strategy to build the tree. Despite their limitations, such as those criticized by these references [12, 13]. Decision trees are widely used due to their speed and ease of use [14, 15].

3 MATERIALS AND METHODS

A well-scripted methods section lays the foundation for your research by outlining the different methods you used to derive your results. The methods used to achieve the objectives must be described precisely and in sufficient detail, so as to allow a competent reader to repeat the work done by the author.

The objective of this study is to create a decision support system (DSS) that can use historical records to predict students' educational paths, and also provide visual representations of the results to aid decision-makers. The system utilizes a decision tree method, specifically the ID3 algorithm, to extract knowledge from a student database. The process involves building a data warehouse by identifying source data systems, designing the warehouse, and performing the ETL process. The KDD module is then applied, which involves selecting a mining algorithm based on the required functionality, in this case, classification using the ID3 algorithm. The model is then trained, patterns of interest are extracted, and the model is applied to test data. The patterns are evaluated using the Lift Chart, and the extracted knowledge is visualized using dashboards to help decision-makers examine data volumes and detect patterns visually. Ultimately, the system aims to predict the best major for a student to study based on discovered knowledge.

3.1 The Implementation of Visual Classification System

This subsection outlines the different processes involved in the system and its modules. These include the data preparation module, building the classification model module, assessing the decision tree, and viewing a statistical report. The system flow diagram model is shown in Fig. 1.

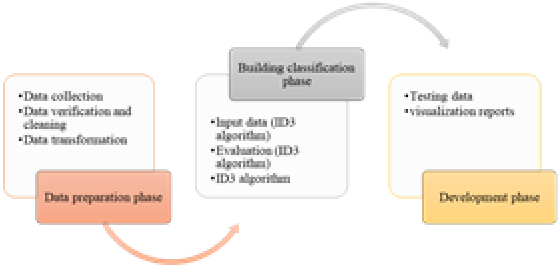

Figure 1: Proposed system model

Data comes to the system from different resources such as the Ministry of Higher Education, universities, colleges, social cooperation's, and information technology centers. The system was created using SQL Server Business Intelligence Development Studio and utilizes the ID3 Decision tree algorithm to build seven models that predict the percentage of students who are likely to find employment. The system comprises three main modules, each containing sub-modules. The Data Preparation Module includes data collection, data verification and cleaning, and data transformation processes. The Building Classification Module includes three main processes: ID3 Algorithm Input Data, ID3 Algorithm Evaluation, and ID3 Algorithm Application. Finally, the system applies the testing data and extracts reports from the Deployment Module for visualization.

3.2 The System Modules

3.2.1 Data Preparation Module

This module is responsible for preparing the database, which involves either the training data or the testing data, to be used for building the mining model, as explained in the preceding section.

The process of collecting data began in 2008, and it entailed the acquisition of pertinent unprocessed information concerning students from Jordan. This information comprises data on students who completed their studies between 1999 and 2009.

The data cleaning process involves a series of steps intended to purify the data, which are as follows [16-18]:

- Completing missing values.
- Removing errors from noisy data.
- Eliminating or identifying outliers.
- Resolving inconsistencies.
- Removing diacritics.
- Removing extra spaces.
- Standardizing words ending with characters that may cause confusion, such as "University of Jordan," which may be written differently by different users.
- Reducing character extenders, such as "Elmaddeh" in Arabic, which can be used to extend characters.
- Standardizing the writing of compound names and the use of quotation marks.
- Standardizing the "Nationality" field, which some users filled in as "Jordanian" while others used "Jordan."

- Improving the reliability of the "National Number" field data.
- Standardizing the different formats used for the "Gender" field.
- Standardizing the different formats used for the "Graduation Semester" field.

The Data Transformation component converts data into a format that is appropriate for data mining by obtaining relevant information. To transform the database, several steps were undertaken, as follows:

- Creating a data source by linking to a relational database to construct the data warehouse.
- Creating a data source view by establishing it on a data source view and defining a portion of the data that populates the data warehouse. Data source views permit the selection of tables, the establishment of relationships between tables, and the inclusion of calculated columns and named views without altering the original data source.

3.2.2 Building Classification Module

In this subsection, the ID3 algorithm is utilized by the researcher on the pre-processed data obtained from the prior section. The section comprises three major components: introducing the dataset, assessing the effectiveness of the ID3 algorithm, and ultimately implementing the algorithm to generate the resulting decision tree.

The decision tree algorithm processes input data based on their data type. Identifying the data type of a column gives the algorithm important information about the type of data in the columns of the mining model, enabling it to determine how to process the data. The content type of a column is related to the behavior of its content. For example, a continuous content type may be assigned to a column containing scalable measurements, such as temperatures.

The attributes used in this research consist of both discrete and continuous data types, each with unique characteristics. Discrete columns contain a fixed number of values with no range between values, and cannot include ordering even if the values are numeric. For example, gender can be represented as a discrete attribute with two possible values: male or female. Continuous columns, however, contain values that represent a scale of numeric data, allowing for serial values, such as temperature readings.

As the ID3 algorithm does not support continuous data types, the researcher made an extension to the algorithm to support continuous data types, particularly in the case of the "GPA" attribute. The researcher created new discrete-valued attributes that partition the continuous-valued attribute into symbolic attributes.

In this subsection, the researcher describes how the data is split into training and testing sets to evaluate the constructed structure. By separating the data in this way, the accuracy and error rates of the mining models can be assessed.

The training set is used to build the mining model, while the testing set is used to measure its accuracy and error rate. The testing set comprises approximately 30% of the total data, and the maximum number of cases in the testing set is limited to 1000. To test the accuracy, the following equation is used:

$$Accuracy = \frac{number\ of\ successful\ predictions}{number\ of\ all\ predictions} * 100\% \qquad (1)$$

$$Error\ Rate = \frac{number\ of\ failed\ predictions}{number\ of\ all\ predictions} * 100\% \qquad (2)$$

3.2.3 ID3 Algorithm

A decision tree is a graphical representation consisting of decision-making nodes, branches, and leaves, designed for decision-making. It is an effective classification algorithm for predictive modeling. Each internal node represents a test on an attribute, with the branches representing the different values of that attribute. Each leaf represents a type or distribution. The top node of the tree is called the root.

To classify a specific data project, a path is traced from the root to the appropriate leaf along a branch, which ultimately leads to a decision. Each path from the root to a leaf is a classification rule, and decision trees can be easily transformed into classification rules, making it a very intuitive classification model. The ID3 algorithm, introduced by Quinlan in 1986, is currently the most widely used and mature decision tree algorithm. Fig. 2 provides a conceptual representation of the decision tree structure.

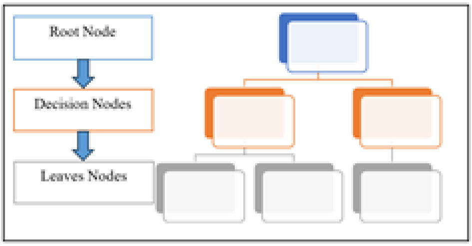

Figure 2: Conceptual model for the decision tree structure

The decision tree is used not only as a classifier, but also as a feature selection tool. It is a practical and efficient method due to its speed and the ease with which it can be converted into simple classification rules. The decision tree uses the information gain metric, which is based on the entropy measure, to determine the most useful attribute. In the ID3 algorithm, the attribute with the highest information gain value is chosen as the test attribute [17, 19].

The ID3 algorithm is based on the following key concepts:

- Each non-leaf node in the decision tree corresponds to an input attribute, and each branch corresponds to a possible value of that attribute.
- A leaf node represents the expected value of the output attribute when the input attributes are described by the path from the root node to that leaf node.

Fig. 3 provides a comprehensive conceptual figure of the decision tree concept.

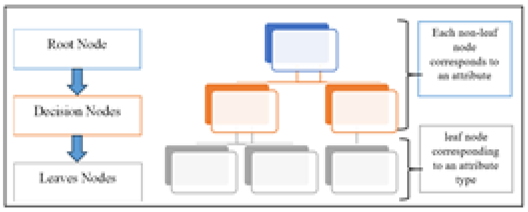

Figure 3: Full conceptual model for the decision tree structure

The decision tree should have each non-leaf node correspond to the input attribute that is the most informative about the output attribute among all the input attributes not yet considered in the path from the root node to that node so that the output attribute can be predicted using the smallest possible number of questions. The decision tree growing algorithm focuses on selecting which attributes to test at each node in the tree. Entropy, which is a measure of uncertainty, is used to determine how informative a particular input attribute is about the output attribute for a subset of the training data. The gain ratio is used to rank attributes based on their level of information gain, with the attribute having the highest gain ratio being considered as the root node of the decision tree. The process is repeated for the remaining attributes to reach the next level until they reach the end of the tree. To minimize the decision tree depth, the user needs to select the optimal attribute for splitting the tree node, which can be indicated by the attribute with the most entropy reduction. The information gain is defined as the expected reduction of entropy to that specific attribute. The formula for information gain is given for a collection S of c outcomes.

$$\boldsymbol{Entropy(S) = \sum -p(I)\log_2 p(I)} \qquad \textbf{(3)}$$

where $p(I)$ – the proportion of S belonging to class I. **Σ** – over c. Log2 – log base 2.

The information gain of an example set S on attribute A is denoted by Gain (S, A), and it is calculated as follows: The entropy of S is subtracted from the weighted sum of entropies of each subset created by partitioning S according to attribute A. The weighting factor is the proportion of examples in each subset relative to the total number of examples in S. In other words, Gain (S, A) measures the expected reduction in entropy achieved by partitioning S according to attribute A.

$$Gain(S, A) = Entropy(S) - \sum((\frac{|S_v|}{|S|} * \; Entropy(S_v)) \qquad (4)$$

where:

Σ is each value v of all possible values of attribute A

Sv = subset of S for which attribute A has value v

|Sv| = number of elements in Sv

|S| = number of elements in S

The ID3 algorithm's pseudo code is presented in Fig. 4.

To create a decision tree from a data set S, the following algorithm can be used:

Input: A data set, S

Output: A decision tree

- Check if all instances have the same value for the target attribute. If yes, return a decision tree with that value.
- If not, do the following:
  a. Compute the Gain value for each attribute and select the one with the highest value. Create a node for that attribute.
  b. For each value of the selected attribute, create a branch from the node.
  c. Assign all possible values of the attribute to the branches.
  d. Follow each branch by partitioning the dataset to include only instances with the corresponding value of the attribute, then go back to step 1.

Figure 4: ID3 algorithm's pseudo code

Once a decision tree is built, a set of classification rules can be generated by following each unique path in the decision tree. These paths represent a set of conditions that must be satisfied to reach a specific outcome or classification. Therefore, each path in the decision tree corresponds to a classification rule that can be easily interpreted and understood by humans.

3.2.4 Data Visualization Based on Dashboard Module

In this subsection, the process of creating visualizations to support interactive decision tree construction is discussed. Dashboards are utilized to create a platform for data sharing and usage, allowing decision-makers to make more informed choices. The collection of data and visualization in a clear and concise format can be a powerful tool for management in the higher education sector, which is the focus of this domain [20-27].

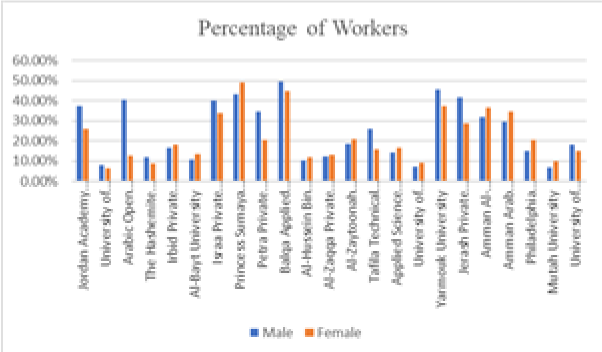

Figure 5: Graduated student based on their gender and percentage of workers.

4 RESULT AND DISCUSSION

In this section, we will discuss the results obtained from our classification system experiments. We will begin by examining the classification structure models, analyzing statistical reports, and evaluating the performance and accuracy rates of these models. The first section will involve a discussion of the structuring models, followed by a section on statistical reports. The final section will provide a closer look at the overall performance, specifically the accuracy of the structural models. We will also analyze the factors affecting the extracted results and make recommendations regarding the system. After conducting a wide-ranging search of available resources, we found that no knowledge extraction system in Jordan predicts the probability of students finding employment in specific majors or universities. While previous research has been done to forecast student performance based on their grades in certain subjects, the variation in systems and their objectives for different stakeholders plays a significant role in determining the value of our addition to scientific research. The results indicate that our system's overall performance is satisfactory.

4.1 Classification models

This section will focus on one of the implemented models (Predicting if students will work based on GPA and University Name attributes – Model1). One of the implemented models (model 1) focuses on predicting whether a student will work based on two input fields - the "GPA" field and the "University Name" field. The "Is_worker" field is used as the predictable field. The "GPA" field has a continuous data type, while the "University Name" field has a discrete data type. A portion of the graph view of the mining model 1 is shown in Fig. 6.

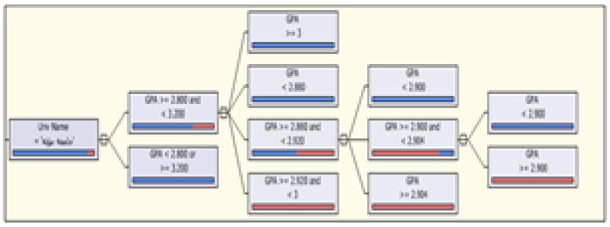

Fig. 6. Part of the graph view of mining model 1

The implemented model, called model 1, allows users to predict a student's likelihood of working based on two input fields: the "GPA" field and the "University Name" field. The "Is worker" field is the predictable field. The data type for "GPA" is continuous, and for "University Name" is discrete. The model consists of a decision tree with six levels and a root, with each node representing an input field. Users can navigate the tree by expanding or collapsing it using the (+) and (-) symbols, respectively. The background color of each node indicates the likelihood of a group of students working, with darker colors indicating a higher probability. The value "1" indicates that a student will work, while "0" indicates that they will not work. Additional information appears in a tooltip when the user hovers over a node, as depicted in Fig. 7.

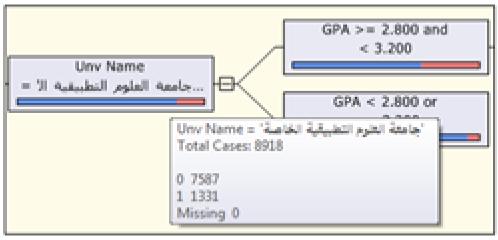

Fig. 7. Tooltip over a specific node in model 1

The mining legend contains various statistics and can be accessed by clicking on a node in the decision tree. It provides information on the number of cases studied, the number of students who have studied at a particular university, and the number of students who have worked. In Fig. 7, for instance, the tooltip shows that all 8918 students have studied at "Applied Science Private University" and 7587 students have not worked while 1331 students have worked. The mining legend also includes a scale that compares the number of students who have worked to those who have not. Fig. 8 shows an example of this mining legend. It provides detailed statistics such as the number of cases in the node, the probability of a student working given a certain GPA range, and the number of students falling into each category.

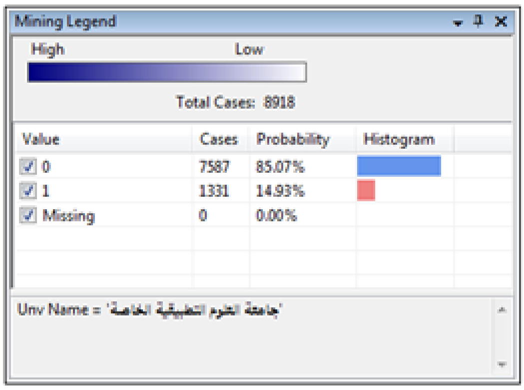

Fig. 8. Mining legend of a node in model 1

The mining legend provides the following information:

- The node represented by the legend tests the attribute "University Name" with the value "جامعة العلوم التطبيقية".
- The total number of cases studied is approximately 8918.
- For cases where the target attribute "Is worker" is 0 (students who are not working), there are 7587 cases, which represents a percentage of 85.07%. This is represented by a blue histogram.
- For cases where the target attribute "Is worker" is 1 (students who are working), there are 1331 cases, which represents a percentage of 14.93%. This is represented by a red histogram.
- The training data does not contain any missing values.

4.2 Accuracy using lift chart

To ensure the accuracy of the structural models and their predictions, the data is divided into training and testing sets. The training set is used to build the model, while the testing set is used to evaluate the accuracy of the mining model. The results of the model are plotted on a chart called the "Lift Chart" against known values for the predictable column in the dataset, along with a representation of the results that an ideal model would produce and the results of random guessing. The more lift the model demonstrates, the more effective it is considered. Only mining models with discrete predictable attributes can be compared on a lift chart.

For the current model, which predicts if students will work based on the attributes of GPA and University Name, the lift chart is shown in Fig. 9. The model's accuracy can be validated by testing it against the testing set to see how well it performs against real data before it is deployed into a production environment. The best-performing model will be used by the MOHE to identify students who will get a job.

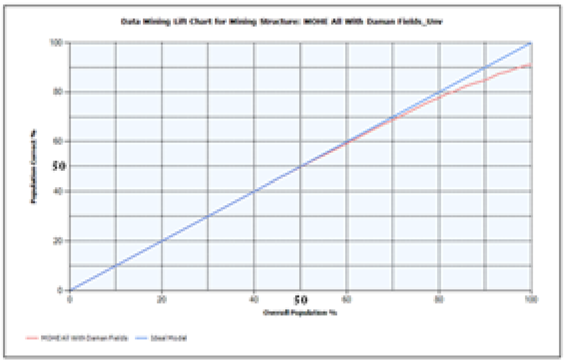

Fig. 9. Lift chart for mining model 1

To evaluate how well the model performs, the data is divided into a training set and a testing set. The testing set is used to assess the accuracy of the model's predictions. The Lift Chart is a chart that compares the model's predicted results with the actual results in the testing set. It shows how well the model predicts both students who are likely to get a job and those who are not. The x-axis of the chart represents the percentage of the testing data used to compare the predictions, while the y-axis represents the percentage of predictions that are correct. The ideal line is the diagonal line, which represents the maximum expected accuracy of 50%. The Lift Chart can be produced with or without target data, and it has a Mining Legend that displays the percentage of cases overall and the percentage of cases that were predicted correctly. The Mining Legend is connected to a gray slide bar that changes the value of the Mining Legend according to its position.

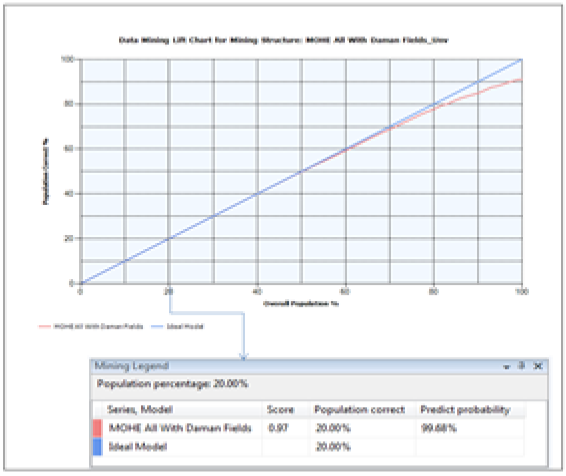

Fig. 10. Lift chart for mining model 1 for (20%) population percentage

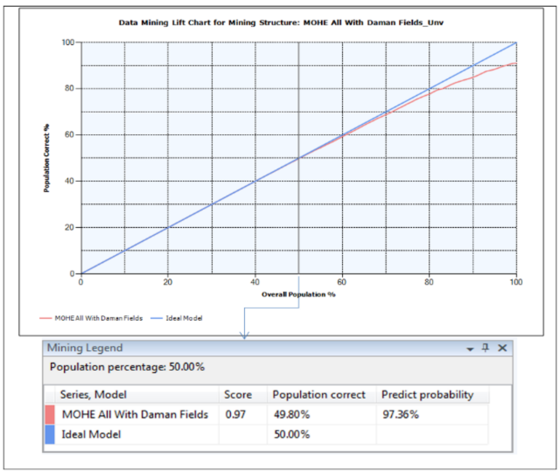

Fig. 11. Lift chart for mining model 1 for (50%) population percentage

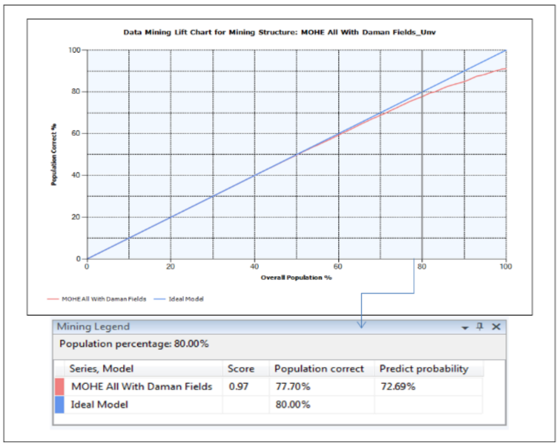

Fig. 12. Lift chart for mining model 1 for (80%) population percentage

Fig. 9, Fig. 10, and Fig.12 display a Lift Chart that does not have a target value. However, a Lift Chart with a target value can also be created, such as "Is worker=1". In this case, the x-axis represents the percentage of the test dataset used to compare predictions, while the y-axis represents the percentage of predicted values. An example of a Lift Chart with a target value and its corresponding mining legend can be seen in Fig. 13.

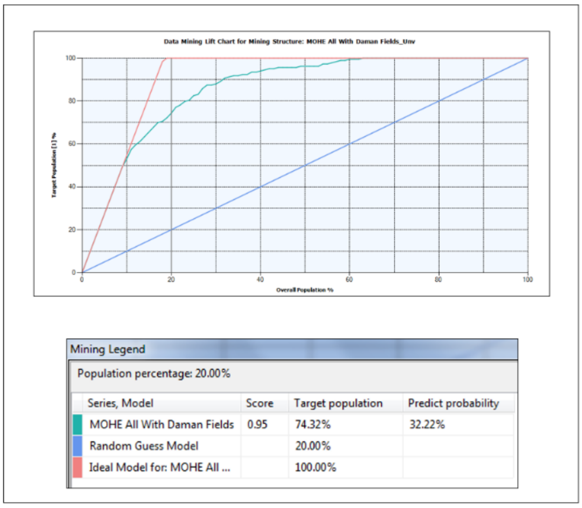

Fig. 13. Mining legend for model 1 lift chart

5 CONCLUSION

The purpose of this study is to use the ID3 decision tree algorithm to identify categories of employment. By using information such as majors, universities, faculties, and GPAs, we can classify future data samples to predict employment outcomes. This system will help companies select the most suitable candidates for employment. Data mining techniques have been widely used in recent decades to discover unknown relationships and patterns to predict actions and behaviors. However, further research is needed in educational domains, especially in different countries and aspects, and when incorporating various institutions into the study.

References

[1]National Information Technology Center. (2024). Retrieved from http://www.nitc.gov.jo/

[2]D. A. Keim, (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1), 1-8. http://doi.org/10.1109/2945.981847.

[3]K. Mierle, K. Laven, S. Roweis, & G. Wilson, (2005). Mining student CVS repositories for performance indicators. *In Proceedings of the 2005 International Workshop on Mining Software Repositories (MSR '05)* (pp. 1-5). ACM. https://doi.org/10.1145/1082983.1083077

[4]J. Han, M. Kamber, (2006). Data Mining: Concepts and Techniques. San Francisco: *Morgan Kaufmann Publishers*. https://doi.org/10.1016/B978-0-12-381479-1.X0001-0

[5]C. Kelleher, B. Tierney, (2018). Data science for business: What you need to know about data mining and data-analytic thinking. *O'Reilly Media*, Inc. http://dx.doi.org/10.5555/3220542

[6]E. R. Tufte, (2001). The visual display of quantitative information. *Graphics Press*. http://dx.doi.org/10.1109/MPER.1988.587534

[7]Alshehadeh, A.R., Ghaleb A. Elrefae, H.A., Al-Khawaja, H.A., Eletter, S.F., & Amer Q. (2022). The Impact of Data Mining Techniques on Information Quality: Insurance Companies as Case. *International Arab Conference on Information Technology (ACIT),* DOI: 10.1109/ACIT57182.2022.9994227.

[8]Musa, K., Alshehadeh, A.-R., Alqerem, R.,(2019), The Role of Data Mining Techniques in the Decision-Making Process in Jordanian Commercial Banks, *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT),* DOI: 10.1109/JEEIT.2019.8717461.

[9]M. Moussa, J. Y. Ruwanpura, G. Jergeas, (2004) "Decision tree module within decision support simulation system," *Simulation Conference*, 2004. *Proceedings of the 2004 Winter*, vol.2, no., pp. 1268- 1276 vol.2, 5-8 Dec. 2004. http://dx.doi.org/10.1109/WSC.2004.1371459

[10] J. R. Quinlan, C4.5: *Programs for Machine Learning*. Morgan Kaufmann Publishers, Inc., Los Altos CA, 1993.

[11] J. Han, M. Kamber, & J. Pei, (2011). Data mining: concepts and techniques. *Elsevier*. https://doi.org/10.1016/C2009-0-61819-5

[12] D. Squire, (2004). CSE5230 Tutorial: The ID3 Decision Tree Algorithm 3. Retrieved from https://people.eng.unimelb.edu.au/dsquire/ID3-tutorial.pdf (Accessed on May, 2023).

[13] J. Luan, (2004). Data Mining Application in Higher Education. *SPSS Executive Report. Chicago,* IL: SPSS.

[14] S. Few, Show me the numbers: Designing tables and graphs to enlighten, Analytics Press. , 2012.

[15] S. Thomassey, & A. Fiordaliso, (2006). A hybrid sales forecasting system based on clustering and decision trees. Decision Support Systems, 42(1), 307-319. https://doi.org/10.1016/j.dss.2005.01.008

[16] A. Adeyemo, & G. Kuye, (2006). Mining Students' Academic Performance Using Decision Tree Algorithms. Journal of Information Technology Impact, 6(3). Retrieved from http://www.jiti.com/v06/jiti.v6n3.161-170.pdf (Accessed on May 8, 2023).

[17] Q. Al-Radaideh, E. Shawakfa, & M. Al-Najjar, M. (2006). Mining student data using decision trees. *In Proceedings of the 2006 International Arab Conference on Information Technology* (ACIT'2006) (pp. 1-5).

[18] J. Liu, H. Zhao, & Y. Liu, (2022). Visualizing Knowledge Graphs: A Survey. IEEE Transactions on Visualization and Computer Graphics, 28(1), 708-727. https://doi.org/10.48550/arXiv.2002.00388.

[19] J. Heer, & B. Shneiderman, (2012). Interactive dynamics for visual analysis. Communications of the ACM, 55(4), 45-54. https://doi.org/10.1145/2133806.2133821.

[20] A. Dunoff, (2009). Are facility professionals using dashboards? Retrieved from http://mcmorrowreport.com/WPs/wpFMDashboards.pdf (Accessed on May, 2023).

[21] Al Houl, M. A. A., Alqudah, M. T. S., Almomani, M. A. A., & Eid, Q. M. A. (2023). The Risks of Financial Derivatives and Alternatives from the Viewpoint of Islamic Economics. International Journal of Professional Business Review, 8(4), e01213-e01213.

[22] B. Fry, Visualizing data: Exploring and explaining data with the processing environment, O'Reilly Media, Inc., 2008.

[23] J. Kehrer, & H. Hauser, (2013). Visualization and visual analysis of multifaceted scientific data: a survey. IEEE Transactions on Visualization and Computer Graphics, 19(3), 495-513, http://doi.org/10.1109/TVCG.2012.110.

[24] Ahmed Badran, T. (2023). Artificial Intelligence between Government and SelfRegulation Policies: A Theoretical Approach. Hikama, 4 (7), pp. 93-110. doi: 10.31430/IJZH4708

[25] Abdou Moussa El-Bermawy (2022). Government AI Readiness Index 2022. Hikama, 2 (3), pp. 214-221. doi: 10.31430/BZUJ6248

[26] Al-Zaqeba, M.A.A., Jarah, B.A.F., Al-Bazaiah, S.A.I., Malahim, S.S., Hamour, A.M.A., Alshehadeh, A.R., Almatarneh, Z., Al-Khawaja, H.A (2022). The effect of reverse factoring financial changes on supply chain, Uncertain Supply Chain Management, Vol10(4), pp:1331-1338.

[27] Alshehadeh, A.R., Mohammad A. Alia, Yousef Jaradat, Ehab Injadat, and Haneen Al-khawaja, (2023), Big data analytics techniques and their impacts on reducing information asymmetry: Evidence from Jordan, International Journal of Data and Network Science, 7, 1259–1266.