ISSN -0258-2724

DOI：10. 3969/j. issn. 0258-2724.2018.002

西南交通大学学报

JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY

第 _ 卷 第 _ 期
2018 年 6 月

Vol. 53 No.5
October 2018

ANALYZING STUDENTS' ANSWERS USING ASSOCIATION RULE MINING BASED ON FEATURE SELECTION

Ali Salah Hashim[a], Alaa Khalaf Hamoud [b], Wid Akeel Awadh[c]

[a]Computer Science Department,
College of Computer Science and Information Technology, University of Basrah
[b]Computer Information System,
College of Computer Science and Information Technology, University of Basrah
[c]Computer Information System,
College of Computer Science and Information Technology, University of Basrah

Abstract

Educational Institutions tend to find ways to improve academic performance by implementing different analytical tools and techniques. The explosive size of educational databases has many advantages and one of them is they can be used to enhance the performance of all academic staff and the institutions in the result. Educational Data Mining (EDM) as a new trend of data mining has faced a huge number of researches and tested almost all data mining techniques and algorithms which result in emerging enhancements in all educational fields. One of the most important functions in data mining is association rules mining. The association rules mining finds the association between dataset features and produces the association as rules. The resulting rules describe many hidden patterns inside the dataset and this feature is very helpful in EDM. Since all educational databases have many features which make the resulting association rules too many, so there is a need for using Feature Selection (FS) algorithm to reduce the features and use important features only. FS has become increasingly important as the size and dimensionality of educational datasets increase. FS is one of most important data mining research areas. It aims to identify several features that describe the dataset better than the original set of features. This objective can be achieved by removing redundant and less correlated features according to importance criteria in FS. Various selection algorithms have been proposed for filter-based FS. ReliefF is one of the most important algorithms that have been successfully implemented in many FS applications. This study presents a new model that can be utilized by any university that seeks to improve the quality of education by analyzing data and identifying factors that affect academic results to increase students' chances of success. We focus on Relief-based algorithms (RBAs) according to the technique of mining association rules. The main concept of the proposed algorithm is to find features that are closely correlative with the class attribute through the mining of association rules. Experimental results based on several real datasets confirm that the proposed model can obtain a smaller and promising feature subset compared with other models and techniques. The outcomes of this research can help improve policies in higher education.

Keywords: Educational Data Mining (EDM), Association Rules, Apriori, Feature Selection, ReliefF, Weka.

摘要

教育機構傾向於通過實施不同的分析工具和技術來找到提高學業成績的方法。教育數據庫的爆炸性規模有許多優點，其中之一是它們可以用來提高所有學術人員和機構的績效。教育數據挖掘（EDM）作為一種新的數據挖掘趨勢，面臨著大量的研究和測試，幾乎所有的數據挖掘技術和算法都導致了所有教育領域的新

興增強。數據挖掘中最重要的功能之一是關聯規則挖掘。關聯規則挖掘查找數據集要素之間的關聯，並將關聯生成為規則。生成的規則描述了數據集中的許多隱藏模式，此功能在 EDM 中非常有用。由於所有教育數據庫都具有許多使得結果關聯規則太多的特徵，因此需要使用特徵選擇（FS）算法來減少特徵並僅使用重要特徵。隨著教育數據集的規模和維度的增加，FS 變得越來越重要。 FS 是最重要的數據挖掘研究領域之一。它旨在識別比原始特徵集更好地描述數據集的幾個特徵。該目標可以通過根據 FS 中的重要性標準去除冗餘和較少相關的特徵來實現。已經為基於濾波器的 FS 提出了各種選擇算法。浮雕 F 是在許多 FS 應用程序中成功實現的最重要的算法之一。本研究提出了一種新模式，任何大學都可以利用這種模式，通過分析數據和確定影響學業成果的因素來提高教育質量，從而提高學生的成功機會。我們根據挖掘關聯規則的技術，專注於基於浮雕的算法（RBAs）。該算法的主要概念是通過挖掘關聯規則來找到與類屬性密切相關的特徵。基於幾個真實數據集的實驗結果證實，與其他模型和技術相比，所提出的模型可以獲得更小且有前途的特徵子集。這項研究的成果有助於改善高等教育政策。

关键词: 教育數據挖掘（EDM），關聯規則，先驗的，特徵選擇，浮雕 F，Weka

I. INTRODUCTION

Firstly, since data mining algorithms techniques differ and cover different paths such as clustering, classification, regression, prediction, and finding association rules, mining association rules is an important part of data mining discipline. Association rules mining can be used to find the hidden knowledge which can be support decisions. Association rules provide an easy form of knowledge like if then form which can be easy to read and interpret. The most used algorithm in association rule mining is Apriori which used in different fields [1, 2, 3]. Education data mining differs from other discipline of data mining in process and results. Mining educational association rules is an emerging domain since the result rules are critical to academic institutions success. A lot of researches based on association rule mining are applied to help lecturers and students to find the most knowledge to increase success and reduce failure [4, 5]

In the other hand, Feature Selection (FS) is one of the important common techniques which used in different paths of data pre-processing such as machine learning, pattern recognition, and data mining (DM). The goal of FS is to find the suitable variables (features) that better describing the dataset and to get more compact and essential representation of feature information. The results subset should be small and very informative to be useful for the model or application. These characteristics can be obtained by removing less correlated and redundant attributes based on most important criterion in FS. FS significantly reduces time of processing and produces accurate and appropriate results when used in data mining applications. FS is a best choice in data mining when the process of data collection is difficult and costly [6, 7]. FS is very important in different fields of machine learning such as computer vision, data mining, text categorization and information retrieval. The need for FS mechanism appears because the increasing size of datasets and dimensions due to development in storage and information acquisition. FS algorithms differ and vary in dealing with data types and produce accurate and efficient results. Recently FS takes a lot of attentions in academic path and used widely in machine learning due to its role in information reduction and then giving a high quality results [8, 9]. FS selection falls into three classes: wrapper, filter and embedded methods [10].

Filter methods include five algorithms (gain ratio-based feature selection, information gain-based feature selection, chi-squared ($\chi 2$) feature selection, symmetrical uncertainty-based feature selection and ReliefF Algorithm). Compare with all filter algorithms, ReliefF algorithm is simple, effective and widely used in feature selection and feature weight estimation [8]. Recently, academic performance has been extensively studied to enhance student's achievement and improve the performance of academic institutions. The use of DM techniques in the educational DM (EDM) field has gained the attention of researchers due to the different functionalities of these techniques, such as prediction, classification and analysis [11, 12, 13].

In this study, we proposed a feature subset selection technique that applies the ReliefF algorithm with association rule mining to select the most valuable features of a dataset (students' responses to a questionnaire). We aim to analyze collective student information through a questionnaire based on Lime Survey and Google Form and classify collected data to predict and categorize student performance. We also seek to identify features that are closely correlated to the class attribute by using the association rule mining method. The experimental results of the datasets demonstrate that the attribute evaluator

(ReliefF) is a good algorithm for FS because the weights of the result attributes affect the total number of association rules obtained after removing the less correlated attributes. Moreover, ReliefF does not affect the confidence of the result association rules when removing the less correlated questions.

II. Related Works

In [14], A. K. Hamoud proposed a model for classifying students' answers by using four clustering algorithms based on a suggested feature selection method, namely, principal component analysis (PCA). PCA is used as an FS method for attribute reduction to find the correlation between the goal class attribute and other attributes. The final step involves a comparison among clustering algorithms based on specific performance criteria to identify the optimal algorithm for clustering. The dataset consists of 62 questions that cover fields that are most related to the study topic, such as health, social activities, relationships and academic performance. Google Forms and LimeSurvey questionnaires (open source applications) are used to prepare the questionnaire, which comprises 161 answers from students. The answers are collected from the students of two departments (Computer Science and Computer Information Systems) in the College of Computer Science and Information Technology, University of Basrah, Iraq.

The authors of [15, 16] used the academic data from four courses, namely, Introductory Programming (IP), Algorithm and Data Structures (ADS) and Change Management (CM) in the undergraduate program, and Creative Leadership (CL) in the master's degree program, as a case study of the Management Study Program at the Faculty of Economics in Maranatha Christian University, and [16] introduced a study that explored educational data from a learning management system. The main goal of the current study is to provide a feedback for the learning process by using a learning management system, thereby enhancing students' success and achievement. The proposed DM algorithms are association rule mining and decision tree J48. Association rule mining produces two sets of interesting rules for CM and IP courses and three sets of rules for CL and ADS courses. Suggestions have been proposed to enhance the learning management system and to encourage student involvement in blended learning.

In [17], the authors proposed an improved apriori algorithm and presented the evaluation of its results after testing the application on ideological and political courses in colleges and universities. The study requirement for college students was obtained after analyzing the correlation found through the network questionnaire. The result of this application can be used to improve the effectiveness of teaching ideological and political courses.

In [18], an application based on different DM techniques, such as Apriori association rules and the k-mean clustering algorithm, was proposed. These authors used a dataset of four government schools from the Vellore District, Tamil Nadu for 100 students and 45 factors. The factors used for classifying student performance varied from personal characteristics and academic data to learning methodology and family history. Some suggestions were obtained from this research, such as the number of teachers per students, providing parents with daily updates of student performance and monthly meetings between staff and parents.

III. Educational Data Mining (EDM)

EDM is defined as new emerging discipline which focuses on exploring and finding new methods and techniques to utilize data collected from educational institutes for enhancing performance of students and exploring new tools for this purpose [19]. EDM uses and analyses the data collected from educational organizations using different educational systems. The main goal is developing models to enhance learning system and improve the effectiveness of educational institutes [20].

The main components of EDM models are three stages: collecting data, archiving data and analyzing that educational data. The first stage fall into tools used for collecting educational data such as answers of online students quizzes, events of educational intelligent systems and all relevant information. The second stage is storing, browsing and archiving educational data. The last stage is analyzing educational data using various tools such as machine learning to get full understanding of the educational data, exploring the relationship among data attributes and developing a model to get generally deep quantitative understanding of cognitive processes [21].

EDM utilizes different techniques and algorithms in different functions such as decision tree algorithms, association rules algorithms, clustering algorithms, k-nearest neighbor algorithms, neural networks algorithms, support vector machine algorithms and naïve Bayes algorithms [22]. Many kinds of knowledge can

be gained can be discovered by using data mining algorithms and functions which can be used for prediction students' success, syllabus organization, students' enrolment is specific course and enhancing students' performance by discovering the most factors that affect the students' pass rate [23]. EDM includes many systems, such as (not limit to) visual data analytics, domain driven data mining tools, information retrieval systems, recommending systems, social networks analysis systems, psychopedagogy, cognitive psychology and psychometrics. Figure (1) shows that the main three parts components of EDM: computer science, education and statistics. These parts are intersecting and form some other areas which are closely related to EDM such as machine learning systems, learning analytical systems, and computer based educational systems [24].

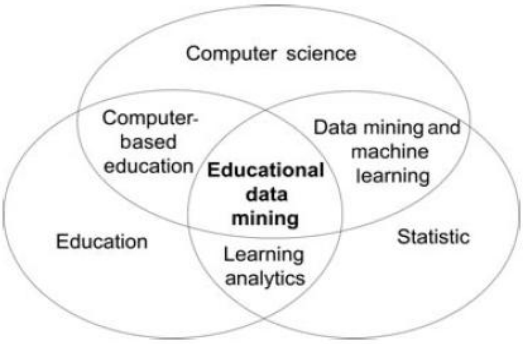

Figure 1. Educational Data Mining

IV. Association Rule Mining (Apriori Algorithm)

Association rules can be used in different disciplines to determine the relationship among dataset attributes and to generate recommendations to support decisions [25]. Association rules mining can be utilized to predict the values of single or a combination of attributes in dataset. Association rules algorithms are differ and the strength point of them are they can express the regularities in a specific dataset and predict the values of dataset attributes. Many association rules can be derived from large and even small dataset, and thus the accuracy of reasonably rules plays the vital role in determining the preferable rules. Two concepts are used to determine the most important rules if association rules mining used, support and confidence. The first concept is support which is the coverage of association rules, the number of instances (records) with can be used to make correct predictions, while the second one is confidence (the accuracy), the number of instances (records) that it predicts correctly. The confidence is the proportion of all records to which it applies [26].

An association rule takes the form of [27]

$A \rightarrow C$,

where A is an item set called the antecedent, and C is an item set called the consequent. A and C have no common items, i.e. $A \cap C = \emptyset$ (an empty set). The relationship between A and C in the association rule indicates that the presence of item set A in a data record implies the presence of item set C in the data record. That is, item set C is associated with item set A.

The Apriori algorithm which was proposed by Agrawal and Srikant in 1994 provides an efficient procedure for generating frequent item sets by considering that an item set can be a frequent item set only if all of its subsets are frequent item sets. Table (1) provides the steps of the apriori algorithm for a given dataset D [28].

Table 1.Apriori Algorithm Steps

Step	Description of the Step
1	F_1 = {frequent one-item sets}
2	$i = 1$
3	while $F_i \neq \emptyset$
4	$i = i + 1$
5	$C_i = \{\{x_1, \ldots, x_{i-2}, x_{i-1}, x_i\} \mid \{x_1, \ldots, x_{i-2}, x_{i-1}\} \in F_{i-1}$ and $\{x_1, \ldots, x_{i-2}, x_i\} \in F_{i-1}\}$
6	for all data records $S \in D$
7	for all candidate sets $C \in C_i$
8	if $S \supseteq C$
9	C.count = C.count + 1
10	$F_i = \{C \mid C \in C_i$ and C.count $\geq$ minimum support$\}$
11	return all F_j, $j = 1, \ldots, i-1$

The Apriori algorithm can be more effective in the process of candidate generation. Apriori algorithm uses pruning techniques to guarantee the completeness of item sets and avoid measuring particular item sets. Apriori algorithm can ensure that these item sets will not be considerable. However, the apriori algorithm has two bottlenecks. The first is the complex candidate generation process that utilizes most of the time, space and memory. The second is the multiple scanning of the database. Many new algorithms were designed based on the apriori algorithm with certain modifications or improvements [28, 29].

V. Feature Selection Methods (ReliefF Algorithm)

FS is an emerging area in the machine learning and DM fields. There is a need for reducing efforts required for processing data sets and get high accurate results. FS can ensure performing this goal by removing redundant and uncorrelated features [30]. There are many characterizations and objectives for FS:

1. Idealized Subset: find the minimal-sized feature subset that is necessary and sufficient for describing the target concept.
2. Target Feature Count: select a subset of m features from a total set of n features, m < n, such that the value of a criterion function is optimized over all subsets of size m.
3. Prediction Accuracy Improvement: choose a subset of features that optimally increases prediction accuracy or decreases model complexity without significantly decreasing prediction accuracy.
4. Approximate Original Class Prediction Probability Distribution: for classification problems, select a feature subset that yields a class prediction probability distribution that is as close as possible to the class prediction probability distribution considering all features. In contrast to prediction accuracy, this perspective seeks to preserve additional information regarding the probabilities of class predictions.
5. Rank and Define Cutoff: rank all features using some surrogate measure of feature 'value', and then define the feature subset by applying an ad-hoc cutoff. This cutoff may be determined via the statistical or subjective likelihood of relevance or by simply using a desired number of features in the subset.

FS or feature reduction used in different fields as a primary process or as a secondary process to get high accurate results. FS used in artificial intelligence, machine learning, image processing and data mining [31]. FS falls into two distinct directions: feature ranking and feature subset selection. There is a problem in the first direction (feature selection), how can the algorithm of learning select the relevant subset of attributes or features upon which to focus the attention of algorithm while ignoring the other features [32].

There are three major groups which FS can be classified into: filter, wrapper and embedded methods. In wrapper method, a concrete classifier is used as black box for estimating feature subsets. In this method, the problem occurs for high dimensional datasets where the computational cost of training the classifier becomes forbidden. The good generalization is the granular of wrapper method. In filter methods, the features are selected by using the general characterizations of features and without utilizing any classifier. The last category of methods is embedded methods where feature selection and learning parts cannot be separated. The class structure of the functions plays critical role in these methods [10]. The selected features with less computational expensive are relay on the machine learning system. One of the most simple and effective algorithm in FS methods is ReliefF, where it is considered as the widely used method for feature subset estimation. For feature ranking, ReliefF uses information ranking instead if viewing ranks and estimating the features with high correlation. The most important objective is to find the features with high importance and low similarity which ReliefF algorithm seeks by avoiding redundant feature selection by applying greedy search algorithm to ensure the optimization [33].

The original Relief algorithm [34] is considered as one of the most effective algorithm in feature weighting. However, Relief is rarely applied anymore and has been supplanted by ReliefF [35] which can be considered as a one of the simples, fast and effective algorithms for feature weighing and selection. The character 'F' in ReliefF refers to the sixth generation (A to F) of Relief. The central idea of relief algorithm based on evaluating features' quality through their capability to distinguish among instances from one feature or class to another class in a local neighborhood. The best attributes are selected based on contribution between instances, such as the best ones are those with high contribution to increase the distance between feature instances and low contribution to increase the distances between same feature instances. The weight in ReliefF Wi falls in the interval [-1,1], where the features with -1 weight can be considered as irrelevant and features with 1 considered as relevant [36]. ReliefF Algorithm [37] takes two vectors the input vector of feature value of each training instances in a dataset, while output vector W is the estimations of attributes qualities.

Table 2.ReliefF Algorithm Steps

Step	Description of the Step
1.	set all weights $W[A] := 0.0$;
2.	**for** i := 1 **to** m **do begin**
3.	randomly select an instance R_i;
4.	find k nearest hits H_j;
5.	**for** each class $C \neq class(R_i)$ **do**
6.	from class C find k nearest misses $M_j(C)$;
7.	**for** A := 1 **to** a **do**
8.	$W[A] := W[A] - \sum_{j=1}^{k} \text{diff}(A, R_i, H_j)/(m \cdot k) +$
9.	$\sum_{C \neq class(R_i)} [\frac{P(C)}{1-P(class(R_i))} \sum_{j=1}^{k} \text{diff}(A, R_i, M_j(C))]/(m \cdot k)$;
10.	**end**;

Similar to Relief, in table (2) (Line 3) shows that ReliefF selects an instance Ri randomly and searches for *k* of the nearest neighbors of the same attribute which called the nearest hits Hj in

(Line4). The nearest misses Mj(C) in (Line 5 and 6) represent the k of nearest neighbors from each of the different features.

In (Line 7,8 and 9) the value of W[6] (the quality estimation) is updated for all features depending on their values for Ri, misses Mj(C) and hits of Hj. ReliefF and Relief are similar in update formula (Line 5 and 6) but ReliefF calculate the average of hits and misses. The contribution of missed features is measured and weighted by using prior probability of that feature P(C) (the estimated value from training dataset). To ensure that the contribution values of misses and hits must be symmetric in each step and fall in interval [0,1], the sum weights of misses' probabilities must be equal to 1. Since the hits class is missing in sum, so the probability of weight is divided with factor 1-P [the class (Ri)] which is the sum of all probabilities for misses classes. This step is repeated for *m* times. The basic difference between Relif and ReliefF is the selection of k misses and hits. ReliefF ensures the robustness is greater with regard to noise. The parameter k is defined differently and this parameter controls the locally estimation values [37].

To handle complete data, the diff function is changed. The missing values of features are handled based on probability. The probability calculation is done based on concept of two instances have different values for a specific value over a feature value [30]:

$$diff(A, I_1, I_2) = \begin{cases} 0 & \text{if } value(A, I_1) = value(A, I_2), \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

$$diff(A, I_1, I_2) = \frac{|value(A, I_1) - value(A, I_2)|}{max(A) - min(A)} \quad (2)$$

$$diff(A, I_1, I_2) = \begin{cases} 0 & \text{if } d \le t_{eq}, \\ 1 & \text{if } d > t_{diff}, \\ \frac{d - t_{eq}}{t_{diff} - t_{eq}} & \text{if } t_{eq} < d \le t_{diff} \end{cases} \quad (3)$$

VI. MODEL

Figure 2 shows the model structure which passes through four steps, data pre-processing, Attribute selection using ReliefF (feature selection), association rules creation using Apriori algorithm and result evaluation. It should be noticed that questionnaire development is considered part of the data pre-processing step. In data preprocessing task, all operations of data evaluation, data cleaning, creating attributes abbreviations, converting data ranges in order to make them acceptable by the model, and creating the final derived column (class) based on specific question (Number of Failed Class) are part of first stage of the model. The derived (Failed attribute) which is the class attribute of the model is created based on the equation:

$$\text{Failed} = \begin{cases} F \; If \; FailedCourses > 0 \\ P \; If \; FailedCources < 0 \end{cases} \quad (4)$$

where F is abbreviation to Failed, P abbreviation to P.

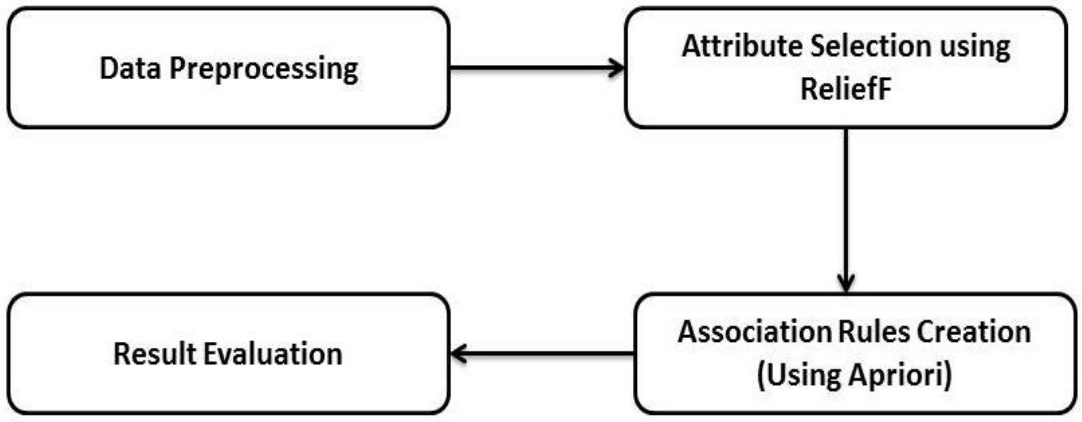

Figure 2. Model Construction Processes

A. Data Pre-processing

Data pre-processing involves two steps: data collection and ensuring reliability. For data collection step, two applications (online google form and open source Lime survey) are used to collect and manage students' answers of department of computer science and department of computer information system in college of computer science and information technology, university of Basrah. The dataset files are combined and one csv file in order to analyze the results using Weka application. The result file holds 161 students' answers which can be considered as a good start to build the model. The dataset with 161 records can be considered as acceptable sample for college of computer science and information technology with 10% percentage of data as errors for the study [38].

The questionnaire includes 61 questions, and thus, shortening the questions became a necessity because questions' abbreviations are needed to understand the result association rules later. The questions' descriptions were abridged for use by both the attribute selection algorithm (ReliefF) and the association rule algorithm (apriori). Table (2) lists some of descriptions that will be used in the next table.

Table 3.Abbreviations and Descriptions of Questions

Question	Abbreviation	Description
Q1	Dep	Your department
Q2	Age	Your age
Q3	Stage	Your stage
Q4	Gender	Your gender
Q5	Address	Where do you live?
Q6	Status	Your status
Q7	Work	Are you working now?
Q8	LiveWithParent	Do you live with your parents?
Q9	ParentAlive	Are your parents alive?
Q10	FatherWork	What is your father's work scope?
Q11	MotherWork	What is your mother's work scope?
Q12	FCourses	Number of courses you failed in per semester
Q13	AbsenceDays	Absence days per semester
Q14	Credits	Number or registered credits per semester
Q15	GPA	GPA
Q16	ComCredits	Number of completed credits
Q17	YearsOfStudy	Number of academic years completed to date
Q18	ListImporPoints	Am I able to write down important points whilst reading a material?
Q19	WriteNotes	During lectures, am I able to write notes and use them for exam preparation?
Q20	PrepStudySchedule	Do I prepare my schedule for studying?
Q21	CalmDurExam	During exams, do I stay calm and coherent?
Q22	LDegNotMakeMeFail	Does getting low grades make me feel like a failure?

From the result of Table (3) where the questions were shortened, table (4) provides both the questions' description and the range of answers to the questionnaire. The descriptions of all the questions in table (4) were condensed based on table (3), such that it can be perused easily to facilitate understanding of the structure of association rules. The questions' ranges were also shortened and converted from numeric to nominal for ease of use and comprehension and for usage with the association rule algorithm (apriori).

Table 4. Description of Questions

Question	**Description**	**Range**	**Question**	**Description**	**Range**
Q1	Dep	IS,CS	Q32	I Hv Enrgy Enjoy	1,2,3,4,5
Q2	Age	1,2,3,4	Q33	Pract Regular	1,2,3,4,5
Q3	Stage	1,2,3,4	Q34	My Health Help	1,2,3,4,5
Q4	Gender	F,M	Q35	Fresh Food	1,2,3,4,5
Q5	Address	IN,OUT	Q36	Cn Use Laptp To Achv Succ	1,2,3,4,5
Q6	Status	M,S	Q37	Plan For Week	1,2,3,4,5
Q7	Work	YES,NO	Q38	Plan Daily	1,2,3,4,5
Q8	Live With Parent	YES,NO	Q39	Plan To Not Read Again	1,2,3,4,5
Q9	Parent Alive	0,1,2,3	Q40	Plan To Do Fun Thing	1,2,3,4,5
Q10	Father Work	0,1,2	Q41	Contrl My Budget	1,2,3,4,5
Q11	Mother Work	1,2	Q42	ClrIdea Abt My Budget	1,2,3,4,5
Q12	FCourses	0,1,2,3	Q43	Cn Work	1,2,3,4,5

Q13	Absence Days	0,1,2	Q44	My Edu Suppo My Goal	1,2,3,4,5
Q14	Credits	0,1,2	Q45	I Hv Sav Plan	1,2,3,4,5
Q15	GPA	1,2,3,4	Q46	EduIs Live Job	1,2,3,4,5
Q16	Com Credits	1,2,3,4	Q47	Clr Abot My Live Goal	1,2,3,4,5
Q17	Years Of Study	1,2,3,4	Q48	Respon Abt My Edu	1,2,3,4,5
Q18	List Impor Points	1,2,3,4,5	Q49	Respo Abt My Live Quality	1,2,3,4,5
Q19	Write Notes	1,2,3,4,5	Q50	Redy To Fac Challng	1,2,3,4,5
Q20	Prep Study Schedule	1,2,3,4,5	Q51	ClrIdea ABout Plans	1,2,3,4,5
Q21	Calm Dur Exam	1,2,3,4,5	Q52	Worked Recently	1,2,3,4,5
Q22	LDeg Not Make Me Fail	1,2,3,4,5	Q53	Knowled Wt Boss Expect Frm Me	1,2,3,4,5
Q23	Eas Can Chos Colg Study	1,2,3,4,5	Q54	EduChoicesToAchivGoal	1,2,3,4,5
Q24	Optim To Achv Goals	1,2,3,4,5	Q55	I Hv Enough Money	1,2,3,4,5
Q25	Cn Study EvU Impo Both Me	1,2,3,4,5	Q56	Relation With Others	1,2,3,4,5
Q26	Exi To Mater	1,2,3,4,5	Q57	Enough Budget	1,2,3,4,5
Q27	Clr Idea Abt Benifit	1,2,3,4,5	Q58	Reques to Help From Others	1,2,3,4,5
Q28	Dev Relation With Others	1,2,3,4,5	Q59	Try To Enhanc My Self	1,2,3,4,5
Q29	Contrl My Anger	1,2,3,4,5	Q60	I Hv Skill To Schv Acadm Work	1,2,3,4,5
Q30	Make Friends	1,2,3,4,5	Q61	IHv Skills For Self Feel	1,2,3,4,5
Q31	Open With Others	1,2,3,4,5			

The first step in data pre-processing involves converting the ranges of the questions' answers into nominal values as shown in the table (5). The details in the table are useful because they provide a basic understanding of the users of the model and allow for the subsequent application of the association rule algorithm. Rows with empty values are not ignored because the association rule algorithm (apriori) can handle missing values. A total of 161 rows are processed in the model.

Table 5. Details of the Answers' Ranges

Question	Range	Details	Question	Range	Details
Q1	IS	Information System Department	Q11	1	(H) Housewife
	CS	Computer Science Department		2	(E) Employee
Q2	1	18 Years	Q12	0	(N) None
	2	19 Years		1	(O) From 1–5 Days
	3	20 Years		2	(M) More than 5 Days

	4	>20 Years	Q13	0	(L) Less than 13 Credits
Q3	1	(F) First Stage		1	(M) 13–17 Credits
	2	(S) Second Stage		2	(N) 18–21 Credits
	3	(T) Third Stage	Q14	1	(C) <60
	4	(Fo) Fourth Stage		2	(C+) 60–70
Q4	F	Female		3	(B) 71–80
	M	Male		4	(A) >80
Q5	IN	Inside Basrah	Q15	1	(F) <36 Credits
	OUT	Outside Basrah		2	(S) 36–72 Credits
Q6	S	Single		3	(T) 72–107 Credits
	M	Married		4	(G) >107 Credits
Q7	Yes	Working	Q16	1	(O) One Year
	No	Not Working		2	(TW) Two Years
Q8	Yes	Living with Parent/s		3	(T) Three Years
	NO	Not Living with Parent/s		4	(F) Four Years
Q9	0	(N) No One Alive	Q17-Q61	1	(W) Weak
	1	(Y) Both Alive		2	(M) Moderate
	2	(FO) Father Only		3	(G) Good
	3	(MO) Mother Only		4	(VG) Very Good
Q10	0	(N) Not Working		5	(E) Excellent
	1	(Em) Employee			
	2	(W) Worker			

The second part of the first stage (reliability) is used for describing the overall measure's consistency. A measure is said to have a high reliability if it produces same results under consistent conditions. For example, the measurements of people's height and weight are frequently extremely reliable [39].

Table 6.Questionnaire Reliability

No. of items	**No. of respondents**	**% of respondents**	**Cronbach's alpha**
62	161	100%	0.85

In statistics, the coefficient alpha which is the most frequently used method for measuring internal consistency, is used as a measure of reliability for the dependent of the studies variables. The previous table (6) shows that the coefficient alpha is 0.85 for the scaled variables of the study, which include 62 questions (items) and 161 students' answers where a Cronbach's alpha of 0.7 indicates satisfactory internal consistency in reliability [40].

B. Attribute Selection Using ReliefF

The second step in the model is attribute selection based on a ReliefF classifier. Each of the 61 attributes in the questionnaire represents a question. Hence, identifying the questions with high correlation to the final class is crucial. In the model, a ReliefF attribute evaluator with a search method (Ranker) is proposed as an attribute selector to the final class (Failed or Pass) with an evaluation mode (full training set) to ensure highly accurate results. The application of the ReliefF attribute evaluator yields the weight, which, in turn, is the predictive value of each attribute to the final class. Weight indicates values between 1 and −1, and the more positive the values are, the more correlated they are to the final class attribute.

Table 7. Attribute Weights Using ReliefF

Seq	Attribute Number	Weight	Seq	Attribute Number	Weight
1	14	0.122188	31	29	0.007125
2	25	0.054	32	24	0.00525
3	52	0.0535	33	4	0.00375
4	51	0.050375	34	12	0.0025

5	50	0.049125	35	58	0.00225
6	17	0.048625	36	23	0.002125
7	13	0.043125	37	36	0.002
8	60	0.041875	38	18	0.001125
9	10	0.034375	39	5	−0.000625
10	53	0.034125	40	8	−0.00125
11	15	0.034062	41	28	−0.00475
12	22	0.034	42	41	−0.004875
13	47	0.031	43	37	−0.006125
14	42	0.030125	44	46	−0.006125
15	26	0.029625	45	55	−0.0065
16	1	0.028125	46	30	−0.008
17	3	0.026875	47	49	−0.009
18	43	0.0245	48	6	−0.009375
19	39	0.023875	49	33	−0.0105
20	35	0.023875	50	11	−0.010625
21	56	0.0235	51	32	−0.012375
22	2	0.019375	52	57	−0.01275
23	45	0.018875	53	34	−0.013
24	48	0.015375	54	7	−0.01375
25	20	0.013375	55	31	−0.01675
26	9	0.011875	56	27	−0.019125
27	44	0.011375	57	59	−0.0225
28	40	0.01075	58	21	−0.022875
29	16	0.009063	59	38	−0.032375
30	19	0.008	60	54	−0.035875

From table (7), note that certain questions (5, 8, 28, 41, 37, 46, 55, 30, 49, 6, 33, 11, 32, 57, 34, 7, 31, 27, 59, 21, 38 and 54) have lower weight values, which indicates that they are less correlated to the final class. The weight value can be utilized to observe the extent of its effect on the total number of rules obtained after applying the Apriori algorithm. These outcomes are depicted in Figures (3) and (4), respectively.

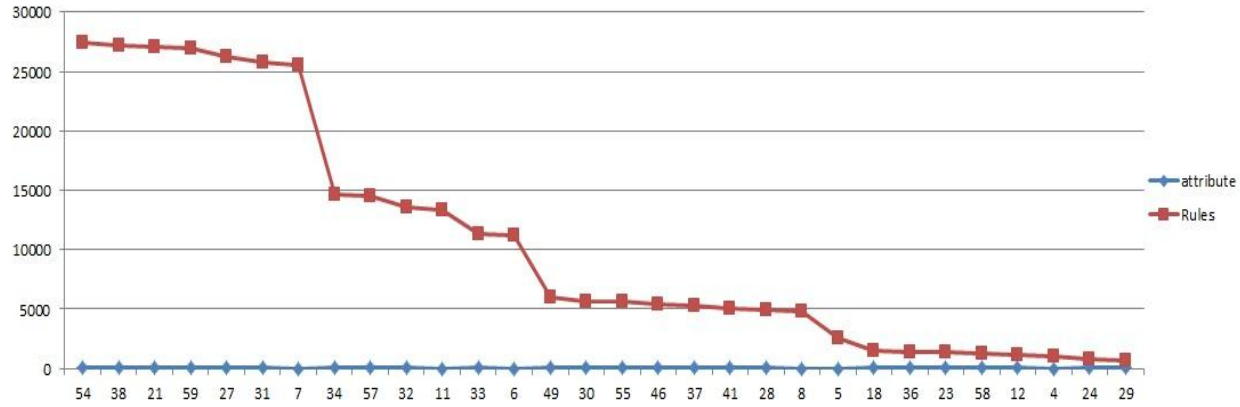

Figure 3.Number of Rules for the First 30 Attributes

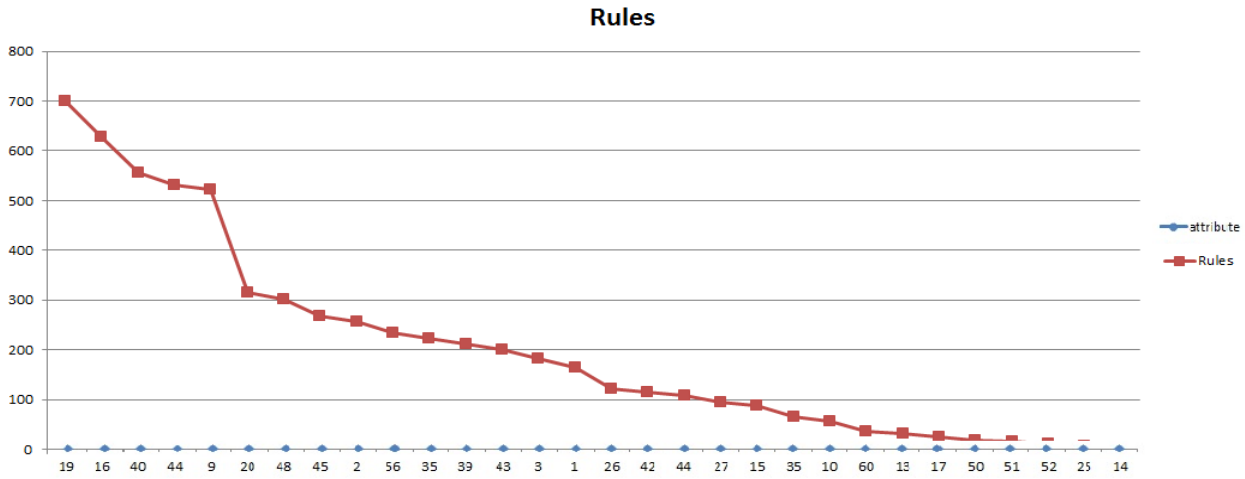

Figure 4. Number of Rules for Each Attribute for the Next 30 Attributes

In both figures, the X-axis represents the number of attributes, whereas the Y-axis illustrates the total number of rules. The red line represents the number of rules, whereas the blue line represents the number of attributes. The arrangement from left to right indicates the less correlated attributes to the more correlated attributes based on Table (7). In both figures, the number of rules per attribute is obtained after removing the attribute and applying the apriori algorithm. Car is set to 'true', and the class index is set to the number of the final class in the algorithm setting to determine only the correlated rules to the final class instead of the general association rules. The minimum confidence is set to 0.5 to filter the acceptable rules. Thus, rules with a confidence of <0.5 are ignored.

C. Association Rule Creation Using Apriori

In our model, the association rules are obtained using Weka 3.8, where the apriori algorithm can be performed easily and effectively. As mentioned earlier, the dataset of the questionnaire answers consists of 61 questions with 161 answers, where 6 rows contain empty answers for some questions. The large number of questions compelled us to select only the most correlated attributes and identify the association rule after removing the less correlated attributes. The previous stage shows the correlation between attributes and the final attribute (Class), such that the attributes with a negative weight are removed and the apriori algorithm is applied. A total of 22 attributes has a negative weight, as shown in table (7). Association rule mining provides the association among all items (questions) as rules in the dataset, and thus, only the association rules related to the final class (Class) should be found.

The class index is set to the index of the final class (Class), and (car) setting is set to TRUE to show only the rules with Class in the consequent part of the association rule. Minimum confidence is set to 0.5, and number of rules is set to 20 to show the sample of the association rules, as indicated in table (8). This table presents the association rules with metrics, i.e. confidence (conf.), lift, leverage (Lev.) and conviction (conv.).

Table 8. Result of Association Rules

Seq	Rule	Conf.	Lift	Lev.	Conv.
1	YearsOfStudy=TW ExiToMater=M ==> Class=F	1	1.65	0.04	7.09
2	ParentAlive=Y Credits=M YearsOfStudy=TW ==> Class=F	1	1.65	0.04	7.09
3	ParentAlive=Y YearsOfStudy=TW ExiToMater=M ==> Class=F	1	1.65	0.04	6.69
4	FatherWork=Em Credits=M YearsOfStudy=TW ==> Class=F	1	1.65	0.04	6.69
5	Credits=M YearsOfStudy=TW ==> Class=F	0.95	1.57	0.05	4.33
6	GPA=C YearsOfStudy=TW ==> Class=F	0.95	1.57	0.05	4.13
7	ParentAlive=Y GPA=C YearsOfStudy=TW ==> Class=F	0.95	1.57	0.04	3.94
8	GPA=C ComCredits=F ==> Class=F	0.95	1.56	0.04	3.74
9	ParentAlive=Y GPA=C ComCredits=F ==> Class=F	0.95	1.56	0.04	3.74
10	Credits=M ComCredits=S YearsOfStudy=TW ==> Class=F	0.94	1.56	0.04	3.54
11	GPA=C EduChoicesToAchivGoal=V ==> Class=F	0.94	1.55	0.04	3.35
12	Dep=CS FatherWork=EmYearsOfStudy=TW ==> Class=F	0.94	1.55	0.04	3.35
13	GPA=C ComCredits=S YearsOfStudy=TW ==> Class=F	0.94	1.55	0.04	3.35
14	Age=T WorkedRecently=W ==> Class=F	0.91	1.51	0.04	3.02
15	Age=T ParentAlive=Y WorkedRecently=W ==> Class=F	0.9	1.48	0.04	2.63
16	Credits=M GPA=C ==> Class=F	0.89	1.48	0.03	2.49
17	Age=T Gender=F WorkedRecently=W ==> Class=F	0.89	1.48	0.03	2.49

18	ParentAlive=Y Credits=M GPA=C ==> Class=F	0.89	1.48	0.03	2.49
19	FatherWork=Em GPA=C ==> Class=F	0.89	1.47	0.05	2.66
20	ParentAlive=Y FatherWork=Em GPA=C ==> Class=F	0.88	1.46	0.05	2.56

D. Result Evaluation

The results of the ReliefF attribute evaluator are depicted in Figures 3 and 4. The relationship between the numbers of attributes and result association rules is proportional, such that the more the attributes, the more rules are present. Measures (confidence, lift, leverage and conviction) are used to assess the frequency of the item set and how much they are associated. The confidence value shows the rule strength that can be used to evaluate rules of interested rules. The confidence is a number between 1 and 0, and the rule with the highest number is the most interested rule.

Lift has a value greater than 1, where 1 indicates that the right and left parts are independent. The higher the lift value, the higher the probability that the left and right parts of the rule will occur together. A higher value also represents the relationship between the left and right parts of the rules, and thus, all 20 result rules can be considered interested. Conviction is similar to lift except that the former has no upper bound and its value represents the probability of the right part when not being true. Leverage is the same as lift, but the former measures the occurrence of additional cases covered by the left and right parts of the rule above those expected if the left part and right part cases are independent of each other. Thus, for leverage, any value above 0 is desirable. Accordingly, all the result rules are dependable.

VII. Conclusion and Future Works

This study presents a model for applying an association rule mining algorithm (apriori) and an FS algorithm (ReliefF) to students' answers to a questionnaire. The results of the model confirm that the attribute evaluator (ReliefF) is a good algorithm for FS because the weights of the result attributes affect the total number of association rules obtained after removing the less correlated attributes. Moreover, ReliefF does not affect the confidence of the result association rules when removing the less correlated questions. Association rules can be applied to studying factors that affect the overall success and failure of undergraduate students. The result of the suggested model can assist stockholders (students, lecturers, managers and decision makers) in enhancing the progress of academic institutions. The model facilitates the understanding of how certain activities can influence student success and the overall academic performance of an institution. The association rules indicate that numerous factors can improve student performance. Finally, the findings of this model can be used to develop an application based on association rules to provide recommendations for improving the performance of academic institutions and stakeholders.

References

[1] KUOK, Ch.M., FU, A., and WONG, M.H. (1998) Mining fuzzy association rules in databases. *ACM Sigmod Record*, 27(1), pp. 41-46. doi: 10.1145/273244.273257

[2] NGAI, E.W.T., XIU, L., and CHAU, D.C.K. (2009) Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications 36*(2), pp. 2592-2602. doi: 10.1016/j.eswa.2008.02.021

[3] PASQUIER, N., BASTIDE, Y., and TAOUIL, R., LAKHAL, L. (1999) Efficient mining of association rules using closed itemset lattices. *Information Systems* 24(1), pp. 25-46. doi: 10.1016/S0306-4379(99)00003-4

[4] MERCERON, A., and YACEF, K. (2008) Interestingness measures for association rules in educational data. *Proceedings of the 1st International Conference on Educational Data Mining, Montreal, Québec, Canada, June 20-21, 2008.*, pp. 57-66.

[5] ZAILANI, A., HERAWAN, T., AHMAD, N., DERIS, M.M. (2011) Mining significant association rules from educational data using critical relative support approach. *Procedia-Social and Behavioral Sciences 28*, pp. 97-101. doi: 10.1016/j.sbspro.2011.11.020

[6] BASKAR, SS. (2014) Feature Selection Techniques to Enhance Classification Accuracy in Data Mining. PhD Dissertation, Department of Computer Science, St. Joseph's College, Tirchirappalli, India.

[7] BIJANZADEH, E., EMAM, Y., and EBRAHIMIE, E. (2010) Determining the most important features contributing to wheat grain yield using supervised feature selection model. *Australian Journal of Crop Science,* 4(6), pp. 402-407.

[8] XIE, J., WU, J., and QIAN, Q. (2009) Feature selection algorithm based on association rules mining method. *Proceedings of the 8th*

IEEE/ACIS International Conference on Computer and Information Science, pp. 357-362. doi: 10.1109/ICIS.2009.103
[9] KIM, Y., STREET, W., and MENCZER, F. (2003) Feature selection in data mining, *Data Mining: Opportunities and Challenges. 3*(9), pp. 80-105. doi: 10.4018/978-1-59140-051-6.ch004
[10] MALDONADO, S., WEBER, R., and BASAK, J. (2011) Simultaneous feature selection and classification using kernel-penalized support vector machines. *Information Sciences*, 181(1), pp. 115-128. doi: 10.1016/j.ins.2010.08.047
[11] HAMOUD, A.K., HUMADI, A., HASHIM, A.S. and AWADH, W.A. (2017) Students' Success Prediction Based on Bayes Algorithms. *International Journal of Computer Applications, 178*(7), pp. 6-12. doi: 10.2139/ssrn.3080633
[12] HAMOUD, A.K., HASHIM, A.S. and AWADH, W.A. (2018) Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis. *International Journal of Interactive Multimedia and Artificial Intelligence, 5*(2), pp. 26-31. doi: 10.9781/ijimai.2018.02.004
[13] HAMOUD, A.K. (2016) Selection of Best Decision Tree Algorithm for Prediction and Classification of Students' Action. *American International Journal of Research in Science, Technology, Engineering & Mathematics, 16*(1), pp. 26-32.
[14] HAMOUD, A.K. (2018) Classifying students' answers using clustering algorithms based on principle component analysis. *Journal of Theoretical and Applied Information Technology, 96*(7), pp. 1813-1825. http://www.jatit.org/volumes/Vol96No7/6Vol96No7.pdf
[15] AYUB, M., TOBA, H., WIJANTO, M.C., and YONG, S. (2017) Modelling online assessment in management subjects through educational data mining. IEEE 2017 International Conference on Data and Software Engineering (ICoDSE), pp. 1-6, doi: 10.1109/ICODSE.2017.8285881
[16] AYUB, M., TOBA, H., YONG, S., and WIJANTO, M.C. (2017) Modelling students' activities in programming subjects through educational data mining. *Global Journal of Engineering Education* 19(3), pp. 249-255.
[17] MAO, Ch.-L., ZOU, S., and YIN, J. (2017) Educational Evaluation Based on Apriori-Gen Algorithm. *Eurasia Journal of Mathematics, Science and Technology Education 13*(10), pp. 6555-6564. doi: 10.12973/ejmste/78097
[18] GOWRI, G.Sh., THULASIRAM, R., and BABURAO, M. (2017) Educational Data Mining Application for Estimating Students Performance in Weka Environment. 14th ICSET-2017. *IOP Conference Series: Materials Science and Engineering*, 263(3), pp. 1-9. doi:10.1088/1757-899X/263/3/032002
[19] International Educational Data Mining Society. (2011) http://www.educationaldatamining.org/.
[20] DUTT, A., ISMAIL, M., and HERAWAN, T. (2017) A systematic review on educational data mining. *IEEE Access, 5*, pp. 15991-16005. DOI: 10.1109/ACCESS.2017.2654247
[21] de LAFAYETTE WINTERS, T. (2006) Educational data mining: collection and analysis of score matrices for outcomes-based assessment. University of California, Riverside. PhD Dissertation.
[22] EL-HALEES, A. (2009) Mining students data to analyze e-Learning behavior: A Case Study. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.592.4341&rep=rep1&type=pdf
[23] KUMAR, V., and CHADHA, A. (2011) An empirical study of the applications of data mining techniques in higher education. *International Journal of Advanced Computer Science and Applications 2*(3), pp. 80-84. doi: 10.14569/IJACSA.2011.020314
[24] ROMERO, C., and VENTURA, S. (2013) Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 3(1), pp. 12-27. doi: 10.1002/widm.1075
[25] HAMOUD, A.K. (2017) Applying Association Rules and Decision Tree Algorithms with Tumor Diagnosis Data. *International Research Journal of Engineering and Technology, 3*(8), pp. 27-31. doi: 10.2139/ssrn.3028893
[26] WITTEN, I.H., and Frank, E. (2016) Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann Publishers.
[27] YE, N. (2013) Data mining: theories, algorithms, and examples. CRC Press.
[28] KOTSIANTIS, S., and KANELLOPOULOS, D. (2006) Association rules mining: A recent overview. *GESTS International Transactions on Computer Science and Engineering 32*(1), pp. 71-82.
[29] AGRAWAL, R. and SRIKANT, R. (1994). Fast algorithms for mining association rules. Proceedings of the 20th International Conference on Very Large Data Bases, pp. 487-499.
[30] PALMA-MENDOZA, R.-J., RODRIGUEZ, D., and de-MARCOS, L. (2018) Distributed ReliefF-based feature selection in Spark. *Knowledge and Information Systems, 57*(1), pp. 1-20. doi: 10.1007/s10115-017-1145-y

[31] ROSARIO, S. F., and THANGADURAI, K. (2015) RELIEF: Feature Selection Approach. *International Journal of Innovative Research and Development* 4(11), pp. 218-224.
[32] CHAVES, R., RAMÍREZ, J., GÓRRIZ, J.M., and PUNTONET, C.G. (2012) Association rule-based feature selection method for Alzheimer's disease diagnosis. *Expert Systems with Applications 39*(14), pp. 11766-11774. doi: 10.1016/j.eswa.2012.04.075
[33] GENG, X., LIU, T.-Y., QIN, T., and LI, H. (2007) Feature selection for ranking. Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, pp. 407-414. doi: 10.1145/1277741.1277811
[34] KIRA, K., and RENDELL, L.A. (1992) A practical approach to feature selection. *Machine Learning Proceedings 1992*, pp. 249-256. doi: 10.1016/B978-1-55860-247-2.50037-1
[35] KONONENKO, I. (1994) Estimating attributes: analysis and extensions of RELIEF. In: Bergadano F., De Raedt L. (eds) Machine Learning: ECML-94. ECML 1994. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), vol 784. Springer, Berlin, Heidelberg, pp. 171-182. doi: 10.1007/3-540-57868-4_57
[36] AGRE, G., and DZHONDZHOROV, A. (2016) A Weighted Feature Selection Method for Instance-Based Classification. International Conference on Artificial Intelligence: Methodology, Systems, and Applications. Springer, Cham, LNAI 9883, pp.14-25. doi: 10.1007/978-3-319-44748-3_2
[37] ROBNIK-ŠIKONJA, M., and KONONENKO, I. (2003) Theoretical and empirical analysis of ReliefF and RReliefF. *Machine Learning*, *53*(1-2), pp. 23-69. doi: 10.1023/A:1025667309714
[38] ISRAEL, G.D. (2009) Determining Sample Size. (Publication No. PEOD6). University of Florida IFAS extension.
[39] CARSON, B. (2009) The Transformative Power of Action Learning. Chief Learning Officer. https://www.clomedia.com/2009/08/20/the-transformative-power-of-action-learning/
[40] SEKARAN, U. and BOUGIE, R. (2016) Research Methods for Business: A Skill Building Approach. John Wiley & Sons.

参考文:

[1] KUOK，Ch.M.，FU，A.和 WONG，M.H.（1998）挖掘数据库中的模糊关联规则. ACM Sigmod 记录，27（1），第 41-46 页. doi：10.1145/273244.273257
[2] NGAI，E.W.T.，XIU，L.和 CHAU，D.C.K.（2009）数据挖掘技术在客户关系管理中的应用：文献综述与分类. 应用专家系统 36（2），第 2 2592-2602 页. doi：10.1016 / j.eswa.2008.02.021
[3] PASQUIER，N.，BASTIDE，Y.和 TAOUIL，R.，LAKHAL，L.（1999）使用闭项集格的高效挖掘关联规则。信息系统 24（1），第 25-46 页. doi：10.1016/S0306-4379（99）00003-4
[4] MERCERON，A.和 YACEF，K.（2008）教育数据中关联规则的兴趣度量。第一届国际教育数据挖掘会议论文集，蒙特利尔，魁北克，加拿大，2008 年 6 月 20 日至 21 日，第 57-66 页.
[5] ZAILANI，A.，HERAWAN，T.，AHMAD，N.，DERIS，M.M.（2011）使用关键相对支持方法从教育数据中挖掘重要的关联规则. 程序 - 社会和行为科学 28，第 97-101 页. doi：10.1016/j.sbspro.2011.11.020
[6] BASKAR，S.S.（2014）提高数据挖掘分类准确性的特征选择技术. 博士论文，计算机科学系，圣约瑟夫学院，印度.
[7] BIJANZADEH，E.，EMAM，Y. 和 EBRAHIMIE，E.（2010）使用监督特征选择模型确定有助于小麦籽粒产量的最重要特征. 澳大利亚作物科学杂志，4（6），第 402-407 页.
[8] XIE，J.，WU，J.和 QIAN，Q.（2009）基于关联规则挖掘方法的特征选择算法。第 8 届 IEEE / ACIS 计算机与信息科学国际会议论文集，第 357-362 页. doi：10.1109/ICIS.2009.103
[9] KIM，Y.，STREET，W。和 MENCZER，F。（2003）数据挖掘中的特征选择，数据挖掘：机遇与挑战。 3（9），第 80-105 页. doi：10.4018/978-1-59140-051-6.ch004

[10] MALDONADO，S.，WEBER，R。和 BASAK，J。（2011）使用内核惩罚支持向量机的同时特征选择和分类。信息科学，181（1），第 115-128 页. doi：10.1016/j.ins.2010.08.047

[11] HAMOUD，A.K.，HUMADI，A.，HASHIM，A.S. 和 AWADH，W.A。（2017）基于贝叶斯算法的学生成功预测。国际计算机应用杂志，178（7），第 6-12 页. doi: 10.2139/ssrn.3080633

[12] HAMOUD，A.K.，HASHIM，A.S。和 AWADH，W.A。（2018）使用决策树分析预测高等教育机构的学生表现。国际互动多媒体和人工智能杂志，5（2），第 26-31 页。doi：10.9781/ijimai.2018.02.004

[13] HAMOUD，A.K.（2016）选择最佳决策树算法预测和分类学生行为。美国国际科学，技术，工程与数学研究期刊，16（1），第 26-32 页.

[14] HAMOUD，A.K。（2018）使用基于主成分分析的聚类算法对学生的答案进行分类. 理论与应用信息技术，96（7），第 1813-1825 页. http://www.jatit.org/volumes/Vol96No7/6Vol96No7.pdf

[15] AYUB，M.，TOBA，H.，WIJANTO，M.C。和 YONG，S。（2017）通过教育数据挖掘在管理主体中建模在线评估。 IEEE 2017 国际数据与软件工程会议（ICoDSE），第 1-6 页. doi：10.1109/ICODSE.2017.8285881

[16] AYUB，M.，TOBA，H.，YONG，S。和 WIJANTO，M.C.（2017）通过教育数据挖掘对学生在编程科目中的活动进行建模。全球工程教育杂志 19（3），第 249-255 页。

[17] MAO，Ch.-L.，ZOU，S。和 YIN，J。（2017）基于 Apriori-Gen 算法的教育评估。欧亚大学数学，科学和技术教育杂志, 13（10），第 6555-6564 页。 doi：10.12973/ejmste/78097

[18] GOWRI，G.Sh.，THULASIRAM，R., 和 BABURAO，M.（2017）用于评估 Weka 环境中学生表现的教育数据挖掘应用. ICSET-2017 第 14 届. IOP 会议系列：材料科学与工程，263（3），第 1-9 页. doi：10.1088/1757-89X/263/3/032002

[19]国际教育数据挖掘学会.（2011）http://www.educationaldatamining.org/。

[20] DUTT，A.，ISMAIL，M. 和 HERAWAN，T.（2017）教育数据挖掘的系统评价。IEEE Access，5，第 15991-16005 页. doi：10.1109/ACCESS.2017.2654247

[21] de LAFAYETTE WINTERS，T.（2006）教育数据挖掘：收集和分析基于结果的评估的评分矩阵。加州大学河滨分校。博士论文.

[22] EL-HALEES，A.（2009）挖掘学生数据以分析电子学习行为：案例研究. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.592.4341&rep=rep1&type=pdf

[23] KUMAR，V. 和 CHADHA，A.（2011）数据挖掘技术在高等教育中应用的实证研究. 国际高等计算机科学与应用杂志 2（3），第 80-84 页. doi：10.14569/IJACSA.2011.020314

[24] ROMERO，C。和 VENTURA，S.（2013）教育中的数据挖掘. Wiley 跨学科评论：数据挖掘和知识发现，3（1），第 12-27 页. doi：10.1002/widm.1075

[25] HAMOUD，A.K.（2017）将关联规则和决策树算法应用于肿瘤诊断数据。国际工程与技术研究期刊，3（8），第 27-31 页. doi：10.2139/ssrn.3028893

[26] WITTEN，I.H。和 Frank，E.（2016）数据挖掘：实用机器学习工具和技术。摩根考夫曼出版社。

[27] YE，N.（2013）数据挖掘：理论，算法和例子. CRC 出版社。

[28] KOTSIANTIS，S.和 KANELLOPOULOS，D.（2006）关联规则挖掘：最近的概述.

GESTS 计算机科学与工程国际交易 32（1），第 71-82 页.

[29] AGRAWAL，R。和 SRIKANT，R。（1994）。用于挖掘关联规则的快速算法。第 20 届超大型数据库国际会议论文集，第 487-499 页。

[30] PALMA-MENDOZA，R.-J.，RODRIGUEZ，D.和 de-MARCOS，L.（2018）Spark 中基于分布式 ReliefF 的特征选择。知识与信息系统，57（1），第 1-20 页. doi：10.1007/s10115-017-1145-y

[31] ROSARIO，S.F.和 THANGADURAI，K.（2015）RELIEF：特征选择方法。国际创新研究与发展杂志 4（11），第 218-224 页.

[32] CHAVES，R.，RAMÍREZ，J.，GÓRRIZ，J.M.和 PUNTONET，C.G.（2012）基于关联规则的阿尔茨海默病诊断特征选择方法. 应用专家系统, 39（14），第 11766-11774 页. doi：10.1016/j.eswa.2012.04.075

[33] GENG，X.，LIU，T.-Y.，QIN，T。和 LI，H。（2007）特征选择的排名。第 30 届国际 ACM SIGIR 信息检索研究与发展会议论文集. ACM，第 407-414 页. doi：10.1145/1277741.1277811

[34] KIRA，K。和 RENDELL，L.A.（1992）一种实用的特征选择方法. 机器学习程序 1992，第 249-256 页. doi：10.1016/B978-1-55860-247-2.50037-1

[35] KONONENKO，I.（1994）估计属性：RELIEF 的分析和扩展。在：Bergadano F.，De Raedt L.（编辑）机器学习：ECML-94。ECML 1994.计算机科学讲义（人工智能讲义），第 784 卷。斯普林格，柏林，海德堡，第 171-182 页。 doi.org/10.1007/3-540-57868-4_57

[36] AGRE，G.和 DZHONDZHOROV，A.（2016）基于实例的分类的加权特征选择方法. 国际人工智能会议：方法论，系统和应用. Springer，Cham，LNAI 9883，第 14-25 页. doi：10.1007/978-3-319-44748-3_2

[37] ROBNIK-ŠIKONJA，M.和 KONONENKO，I.（2003）ReliefF 和 RReliefF 的理论和实证分析。机器学习，53（1-2），第 23-69 页. doi：10.1023/A：1025667309714

[38]以色列，G.D.（2009）确定样本量.（公开号 PEOD6）.佛罗里达大学 IFAS 扩展.

[39] CARSON，B.（2009）行动学习的变革力量. 首席学习官. https://www.clomedia.com/2009/08/20/the-transformative-power-of-action-learning/

[40] SEKARAN，U. 和 BOUGIE，R.（2016）商业研究方法：技能构建方法. John Wiley & Sons.