International Journal of Modern Communication Technologies & Research (IJMCTR)
ISSN: 2321-0850, Volume-4, Issue-5, May 2016

IRS: Intelligent Recommender System as an Academic Advisor for Student Coordination

Mohamed Ezz, Hany Harb, Abeer Sayed

***Abstract*— The enrollment process of students in the colleges of Al-Azhar university (coordination) is mainly based on student final grades in Al Azhar high school Diploma (AHSD). The coordination does not consider the student qualifications and their talent in certain aspects. This paper proposes a system which acts as an academic advisor helping the students to check their qualifications to study commerce studies in general and specially accounting. This system can be adapted to any other fields of study in Al-Azhar university. The proposed model considers the student skills and knowledge and then decides how much the student is suitable for his/her chosen study field. We measure the student skills and knowledge earned during his/her study in the AHSD. The student attitude evaluation in different subjects is based on his/her earned degrees in the final AHSD. The model may take into consideration exams required by the faculty/department, constraints imposed by the faculty/department, the faculty/department capacity, the location of the faculty/department, the location of the student high school, and the student gender. The system works on a data set of the last five years of AHSD and the results of the students already enrolled in the faculty of Commerce/Accountant department. The research firstly apply the logistic regression model to check the suitability of the chosen major for a student then all the students who pass the suitability test are reevaluated to predict their performance in the major.**

***Index Terms*—Data Mining, Coordination System, Recommendation System, Machine learning, Logistic Regression**

I. INTRODUCTION

About 100,000 students are welcomed every year in the Al-Azhar University which is one of the oldest universities in the world. All of the applicants are accepted and then distributed to the different colleges of the university. There are about 76 college distributed all over Egypt. As the biggest Islamic school in the world, it offers its educational services to many students from different foreign countries. Some colleges are only reserved for girls and others for guys. The official coordination system is based on the student final score without regarding the student grades in the different subjects. The recommender system helps the student to check which is the most suitable specialization for his/her abilities so s/he can organize his/her choices properly. The already applied coordination system assigns the first student desire which satisfies the college requirements and based on the total student score. The case study for the proposed model has been applied in the faculty of Commerce/Accounting based on the data set collected over the last five years. This data set is a property of Al Azhar university.

Table1 presents a sample of students already enrolled in the faculty of commerce Al-Azhar University for each grade where P stands for pass, G for Good, VG for Very Good, and E for Excellent of AHSD certificate, during the last five years. Each high school student studies different subjects which are categorized such as Islamic, Arabic, and Scientific fields where each category contains more than one subjects. The score of subjects in each category are summed and its percentage (related to the maximum score of this category) is computed. For each student there exists the detailed student degrees in different subjects of the AHSD (percentage of sum of Arabic language subjects degrees, percentage of sum of Islamic studies subjects degrees, percentage of sum of cultural subjects degrees, percentage of sum of scientific subjects degrees and percentage of his/her final score), certificate type (literal/scientific), gender, age, the student study performance in his/her college/department during the last five years (if any), the student institute district (Cairo/North Egypt/South Egypt). There also exists the student performance history in his/her college/department. The student GPA for every year of college study is kept. Our case study considers the faculty of commerce where there are four years of study, so the student may have one record if he/she is in the second year of study, two records if he/she in the third year, and so one. One more record is added for every year the student has failed in. The first year and second year of study are common to all students. Starting from the third year, the student is assigned to one department (major): accounting, statistics, business administration, economy, statistics. There are two types of students: regular and irregular. We only consider regular students for this case study.

The correlation between the student final grade (as an independent variable) and his/her study status (as dependent variables) shows that there is no linear relation between the student performance in the faculty, and his/her final score in the AHSD. The result entails that there are other implicit factors that govern the success of students. The proposed model in this paper utilizes different machine learning algorithms for classification and prediction of student performance. Several machine learning algorithms have been applied during the research work, namely J48[1], Support Vector Machines (SVM)[2], Fuzzy Unordered Rule Induction Algorithm (FURIA) [3] algorithms.

The rest of the paper is organized as it follows. Section 2 summaries related works in predicting student model using data mining and classification techniques in educational environments. Section 3 describes the proposed enrollment advisory model and briefly reviews its components and methods. Section 4 gives a description of data set and features

Mohamed M. Ezz, Systems & Computers Engineering, Al Azhar University/ Faculty of Engineering, Cairo, Egypt, 00201001083969,

Hany M. Harb, Systems & Computers Engineering, Al Azhar University/ Faculty of Engineering, Cairo, Egypt.

Abeer Sayed, Al Azhar University/ Faculty of Commerce, Cairo, Egypt.

that have been used in this research and describes the preprocessing steps for data analysis. Section 5 presents the experimental results, and section 6 concludes the research.

II. Related Works

Many studies have been proposed, which attempt to predict the student performance using machine learning and data mining techniques where they work on different student features and different machine learning techniques and they have different accuracy measures. The accuracy measure is either the overall model accuracy (correctly classified instances), or the true positive (TP) rate of each classification in the model. Table 2 presents a comparison between some of these researches.

III. Logistic Regression verification

The main question of interest about our case study: which of the explanatory variables are predictive of the response (the major is recommended or not?). The considered explanatory variables are high school Arabic subjects score (HS-Arabic), high school Islamic subjects score (HS-Islamic), high school scientific subjects score (HS-scientific), high school final score (HS-total), High school certificate type (HS-Type) (either literal or scientific), High school area (area) (Cairo District, lower Egypt, upper Egypt), Faculty grade 1 score (grade1), Faculty grade 2 score (grade2), Faculty grade 3 score (grade3), Faculty grade 4 score (grade4).

Table 1: Student Study History for Commerce/statistics

AHSD Year/Certificate		**P**	**G**	**VG**	**E**	**Grand Total**
2011		7	73	254	66	400
	Literal	7	73	75	5	160
	Scientific			179	61	240
2012			19	184	107	310
	Literal		19	34	1	54
	Scientific			150	106	256
Grand Total		7	92	438	173	710

A. *The Dataset*

The training dataset has been collected from two databases (which are joined through student names): the AHSD DB and Commerce/Accounting major DB as a case study. The DB of any other major can be substituted. A consolidation process is performed as a preliminary step of the training phase as a pre-processing module. This dataset consists of 710 students who were enrolled in the Commerce/Accounting major and each instance consists of 10 attributes from the AHSD database. The student accumulated grades in the years of studying in the faculty of Commerce/Accounting major are added to the dataset as 5 attributes from the Commerce/Accounting major DB representing the student ID and the grades in the four studying years. Some irrelevant attributes have been removed manually such as AHSD student ID, student name, and college Student ID. The left 12 selected attributes are represented in Table 3.

Table 2: Related work summary

Research	**Objective of Research**	**Machine Learning algorithm**	**Dataset & Features used**	**Type of Classification**	**Accuracy Measure (Overall accuracy or individual class accuracy)**
[4] Q. a Al-Radaideh, E. M. Al-Shawakfa	Create student model for measure the student's performance in a specific courses (C++ language)	ID3, ,C4.5, and Naive Bayes used and compared	Undergraduate students took the C++ courses. 12 attributes collected using a questionnaire	Four classes as Course grade : A, B, C, D	ID3 (38% Overall accuracy) C4.5 (35%Overall accuracy) Naive Bayes (33%Overall accuracy)
[5] R. Kabra and R. Bichkar	Knowing the reasons of failure of student in Engineering faculty to help to take necessary actions to improve the success percentage	J48 decision tree algorithms	346 Engineering student. 16 attributes selected from high school final results and some subject results	Two classes : Promoted, and Failed	69.94 %Overall accuracy
[6] S. Sembiring, M. Zarlis	Predict the final grades of students based on behavioral (psychometric factors) of students	Smooth Support Vector Machine (SSVM) algorithm	1000 students from faculty of computer system and software. 5 attributes from behavioral variables used are Interest, Study Behavior, Engage Time, Believe, and Family	Five classes for final grade: Excellent, Very Good, Good, Average,	Excellent (92% TP rate), Very Good(75% TP rate), Good(61% TP rate), Average(69% TP rate), Poor(93% TP rate)

			Support.	Poor	
[7] Q. a Al-radaideh, A. Al Ananbeh	Predict the suitable track for the students in high school based on previous result in basic school	J48 decision tree algorithm	248 students from basic schools. Three attributes used; the average grade of the last year class (N), the average grade of classes (N, N-1,N-2), the minimum grade acceptable for each track.	Four classes: Science, Management , Academic, Profession,	Science(54% TP rate), Management(90% TP rate), Academic(100% TP rate), Profession(98% TP rate)
[8] J. Econ. Bus	Predict the final grades of students based on socio-demographic, high school final result, and study attitudes of students	C4.5, Multilayer Perceptron , Naive Bayes	270 students from Faculty of Economics. 11 attributes selected from students' socio-demographic, high school final result, and study attitudes	Two classes : Pass, Fail	C4.5 (73.93% Overall accuracy), Multilayer Perceptron (71.20% Overall accuracy), Naive Bayes(76.65% Overall accuracy)
[9] S. K. Yadav and S. Pal	Predict the performance of students in Engineering faculties to identifying the students that are most likely to fail to improve their performance	C4.5, ID3 and CART decision tree algorithms	90 students from faculty of Engineering. 16 attributes from student demographic data, plus student grade in high school and senior secondary schoo;d	Three classes: Pass, Fail, promoted	ID3 (62% Overall accuracy), C4.5(67% Overall accuracy), CART (62% Overall accuracy)
[10] D. Kabakchieva	Studying the data mining techniques for predicting student performance	J48, NaiveBayes, BayesNet, OneR, JRip	10330 students from 9 faculties, 13 attributes from student personal data such as gender and age and grade of high school in addition to some characteristics of high school.	Bad Average Good Very Good Excellent	J48(66% Overall accuracy), NaiveBayes(59% Overall accuracy), BayesNet(59% Overall accuracy), OneR(54% Overall accuracy), JRip(63% Overall accuracy)

Table 4 presents a training dataset which consists of 710 students from Commerce/Accounting major during the academic year 2014/2015 distributed over different accumulated GPA classes. The Excellent class is equivalent to 85% or more and is equivalent to GP of 4, Very Good is equivalent to 75% up to 85% and equivalent of GP of 3.0, Good is equivalent to 65% up to 75% and equivalent to GP of 2 and Pass is equivalent to 50% up to 65% and equivalent to GP of 1.0 as shown in Table 5.

Table 3: The student Dataset attributes

Attribute Category	**Attribute Name**	**Attribute Label**	**Attribute type**
Personal data (from AHSD certificate)	DOB	Date Of Birth	Numeric attributes
	Area	The area of leaving	Nominal attribute (Cairo, North Egypt, South Egypt)
AHSD related info			
	HS-Arabic	AHSD Arabic total degree	Numeric attribute
	HS-Scientific	AHSD Scientific total degree	Numeric attribute
	HS-Islamic	AHSD Islamic total degree	Numeric attribute
	HS-Total	AHSD total degree	Numeric attribute
	Year	AHSD Year of graduation	Numeric attribute
	HS-Cert-Type	AHSD certificate type	Nominal attribute (Scientific, literary)

Faculty related info			
	Grade-1	The faculty 1st class grade	Nominal attributes (Excellent, Very Good, Good, Pass)
	Grade-2	The faculty 2nd class grade	Nominal attribute (Excellent, Very Good, Good, Pass)
	Grade-3	The faculty 3rd class grade	Nominal attribute (Excellent, Very Good, Good, Pass)
	Grade -4	The faculty 4th class grade	Nominal attribute (Excellent, Very Good, Good, Pass)

B. The Dataset Pre-processing

The pre-processing of training phase consolidates the AHSD DB and Commerce/Accounting DB using student name to join records from both databases. Next, irrelevant attributes are removed such as student ID in AHSD, student ID in faculty which do not affect the classification process. As another part of the pre-processing phase, in AHSD student record, all literature subjects of Arabic, all Islamic subjects, and all Scientific subjects degrees are summed to be as three fields only (HS-Arabic, HS-Islamic, and HS-Scientific respectively). Different weights can be given to literal Arabic, Islamic and Scientific subjects sum.

Table 4: The Commerce/Accounting Training dataset Sample

Accumulated Grade	**Count**
Excellent	13
Very Good	356
Good	329
Pass	12
Grand total	710

Percentage is computed for these three fields and turned into the equivalent grade class (E/VG/G/P). The HS-Total is the percentage of the final High school score.
For each student faculty performance data, the student GPA is computed as it follows. Excellent is given a weight of 4, very good is given a weight of 3, good is given 2, pass is given 1, pass with one or two subjects is given weight 0.5 and Failed is given weight 0 as shown in Table 5. The accumulated GPA is computed as the sum of the GPA for all years of study divided by the total number of study years.

All the score are turned into grade codes (E/VG/G/P/F) where E (excellent) for score >= 85%, VG (Very Good) for score >= 75% and less than 85%, G (Good) for score >= 65% and less than 75%, P(pass) for score >50% and less than 65%, and F (Fail) for score < 50%. The student GPA is calculated by giving E weight 4, giving VG weight 3, giving G weight 2, giving P weight 1, and giving F weight 0 and then compute the sum of all the faculty grades divided by all the number of years attended by the student. The outcome (response) variable (suitability) determines either the faculty/major is suitable for the student (value of '1') or not suitable (value of '0'). The binary outcome will be 1 if the student has GPA >= 1.5 otherwise it will be 0. There are two main reasons for applying logistic regression rather than other simple regression models.

C. The Logistic Regression Model

The outcome variable is binary rather than continuous, so logistic regression is considered rather than any other regression model which assumes that the outcome variable is continuous and it has a normal distribution with constant variance rather than binomial distribution [11]. Our analyses establish a relationship between the recommendation variable ("1" indicating recommend yes and "0" recommend no) and the above mentioned attributes which may have effect on the success in the faculty/major. The recommendation process proceeds in three steps: assessment of relationships between dependent and each independent attribute, adjustment of relationships for attributes in group, and then study the interaction effects between variables. Descriptive tools provide initial insights into the structure of the data and the associations between categorical independent variables (such as certificate type and area) and the recommendation variable. Cross-tabulations of recommendation by each of the categorical predictor variables are shown in the table

Table 5: Mapping of Commerce/Accounting Faculty Final Grade to classes

Grade	**Weight**
Excellent (85%>)	4
Very Good (75%-85%)	3
Good (65%-75%)	2
Pass ({50%-65%)	1
Pass with one or two subjects	0.5
Failed	0

attached in the Appendix I. The results show that in our sample of 710 students the recommendation proportion were less for literal certificate students. A scatter plot (especially when it is enhanced by a Lowes curve.) may be helpful for examining the association between certificate type and recommendation (suitability for the major). The percentage of all students who are suitable for the major is 81.6% and this is equivalent to saying that the probability of suitability in our sample is 0.816.
The logistic function transforms log odds into a proportion or a probably p: Log [p/(1-p)] = a + bx. The reverse of the log to both sides of the equation, eliminating the log on the left hand side, the formula can be rearranged to solve for the value p: p = Exp(a+bx) / [1 + Exp(a+bx)]. The Deviance -2 log-likelihood (-2LL) is the basic statistic measure of the model accuracy where the higher the statistic value means the less model accuracy. It sums the differences between the actual and predicted outcomes for each case as a measure of the total model error. This deviance depends on the number of model parameters, the sample size, and the fit goodness, so a

standard is needed to evaluate the relative size. The value for our model is compared against a baseline model and the -2LL statistic variance test whether the model is getting more accurate. Initially, the baseline (step 0 where all independent variables are absent) is just guessing as the category with the largest number of cases. In suitability, 98.3% of student achieve suitability while 1.7% do not, so the probability of picking at random a student who achieves the suitability threshold is therefore slightly higher than the probability of picking a student who does not. This baseline model is the one we can test our later models against. Our later models are getting by adding independent variables either one by one or all at once. The model improvement can be computed as: X2= [-2LL (baseline)] - [-2LL (new)] where degrees of freedom= kbaseline- knew, and k is the number of parameters in each model. If a later model explains the data better than the baseline (or previous) model, then there should be a significant reduction in the deviance (-2LL) which can be tested against the chi-square distribution to give a p value (as shown in the Appendix I tables). The less deviance (-2LL) with each set of independent variables added to the model means this set is more significant and the model is more accurately predicting the outcome.

SPSS calculates and reports the Wald statistic (which tests whether the variable is making a significant contribution to the prediction.) and importantly the associated probability (p-value). The B coefficient indicates the increase in the log odds of the outcome for a one unit increase in the independent variable. Taking the exponent of the log odds allows interpretation of the coefficients in terms of Odds Ratios (OR) which are substantive to interpret. SPSS gives this OR for the explanatory variable labeled as Exp(B). The accuracy improves up to a satisfactory level (92.2%) by stepwise adding six variables which are HS-Arabic, HS-Islamic, HS-Scientific, HS-Total, Certificate Type, Area.

D. Setting up the logistic regression model

The logistic regression model is created with all six independent explanatory variables. The logistic regression pop-up box is appeared allowing us to input the variables. Our outcome measure is whether or not the student achieves suitability. This variable is labeled suitable and should be moved in to the Dependent box. Any independent variable is placed in the covariates box. If the explanatory variable is continuous (GPA), it can be directly placed but we have to set up dummy variables for categorical variables based on a specific baseline category. Move Area, Certificate Type, HS-Arabic, HS-Islamic, HS-Scientific, HS-total into the covariates box. Define them as categorical variables by clicking the button marked “Categorical” to open a submenu to move all of the categorical variables from the left hand list (Covariates) to the right hand window.

The reference (or baseline) category for each variable is decided by clicking on each in turn and using the controls which are marked “Change Contrast”. We tell SPSS whether the first or last category should be used as the reference and then click “Change” to finalize the setting. For our Area variable the first category is “Cairo” can be used as the reference category. Change the selection to “First” and click “Change”. For the Certificate Type variable we only have two categories and could use either literal or scientific as the reference, we have used scientific as the reference (change the selection to “First” and click “Change”). For all other variables, we use E as the reference. Click “Change”, then the selections have appeared in brackets next to each variable and then click “Continue” to close the submenu. All explanatory variables are entered together as one block by leaving “Enter” as it is. Only the suitable students are passed to The IRS system.

IV. THE IRS ARCHITECTURE

The IRS can be applied to any major where it takes the student data and recommends (or does not recommend) this major for the student. The system consists of three phases: pre-processing, features extraction, and the prediction model and it acts in two modes: training mode and runtime/production mode. During the training mode, the system is trained to generate the prediction model. It takes the high school student scores (final score and detailed grades in different subjects), and the student performance history during his/her major study as input and generates the prediction model. The Fig. 1-A shows the IRS model architecture while Fig. 1-B shows the system phases where the training algorithm may be altered to test different training methodology. Once the system is trained, it can run in the production phase.

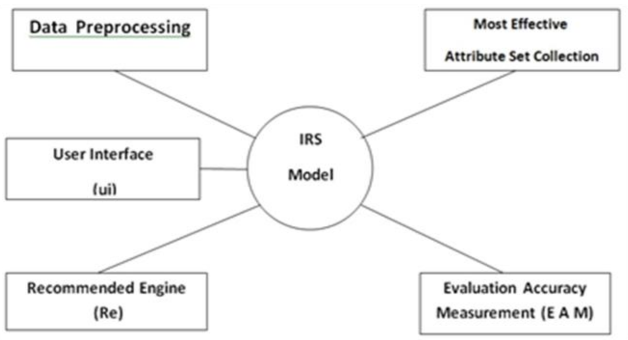

Fig. 1-A : The IRS Model Architecture

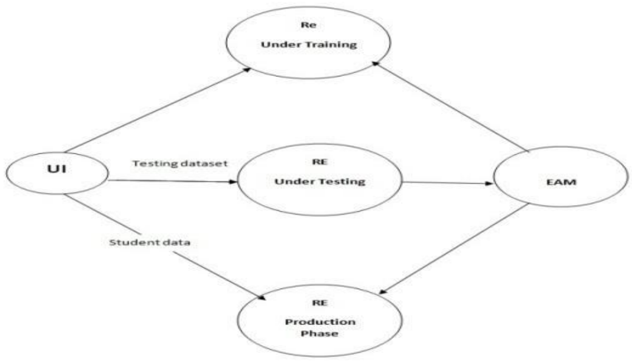

Fig. 1-B: The IRS Phases

A. The Features Extraction

The features extraction component filters the most relevant (effective) attributes for learning phase, by measuring the rank of each attribute. After filtration, only 9 relevant attributes (as shown in Table 6) from 11 attributes have been selected. The ranking algorithm measures the most effective attributes using Gain Ratio Attribute Evaluation [12] which evaluates the worth of an attribute by measuring the gain ratio with respect to the class.

GainR(Class, Attribute) = (H(Class) - H(Class | Attribute)) / H(Attribute).

As shown in Fig. 2 the true positive (TP) accuracy against the number of attributes. The maximum value is given when the number of attributes is equal to 9.

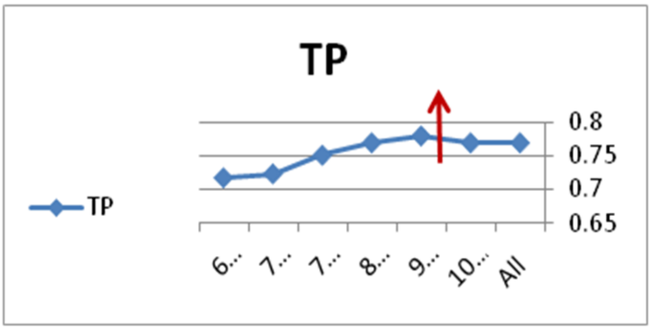

Fig. 2: The Selected attributes

The system evaluate attributes by using all attributes and the TP rate is measured (J48 algorithm utilized in this stage), then the attribute with the lowest effect factor is removed one by one until the best performance model is reached TP. At this time, the optimum and effective attributes are selected to be used in the learning phase and applied in the machine learning algorithms.

B. The Training Process

Machine learning algorithms[13] operate by building a system model from training dataset and then use the generated model to make predictions or decisions. The main component of the advisory system is the training component that employs different machine learning algorithms after the training student dataset is consolidate and optimized. After the student model is built, it acts an advisory system that outputs one of four classes (Excellent, Very Good, Good, Passed). In this paper, three types of machine learning algorithms are employed to be contrasted to pick the best performance one.

Table 6: The ranked attributes

Attribute #	Attribute Name	Rank
1	Grade-2	0.13006
2	HS-Islamic	0.1209
3	HS-Total	0.10896
4	Year	0.09923
5	Grade-1	0.09041
6	HS-Cert-Type	0.07907
7	HS-Scientific	0.07693
8	Grade-3	0.0706
9	HS-Arabic	0.0631
10	DOB	0.05253
11	Area	0.00588

1) The (J48) algorithm

Decision trees are powerful and popular tools for classification. A decision tree is a tree-like structure, which starts from root node (attribute), and ends with leaf nodes. Generally, a decision tree has several branches consisting of different attributes. The leaf node on each branch represents a class or a kind of class distribution. Decision tree algorithms describe the relationships among attributes and the relative importance of attributes. This algorithm uses decision trees with multiple linear regression models at the leaf nodes, and additive regression using forward stage-wise modeling is applied to grow the tree. The advantages of decision trees are that they represent rules which could easily be understood and interpreted by users. The WEKA J48[1] classification is applied on the training dataset during the experimental study. The result tree from the J48 algorithms is shown in Fig. 3 where the goal leaf appears for each class. The generated tree size is 97 (total number of nodes) and Leaves are 72 (number of predictor nodes).

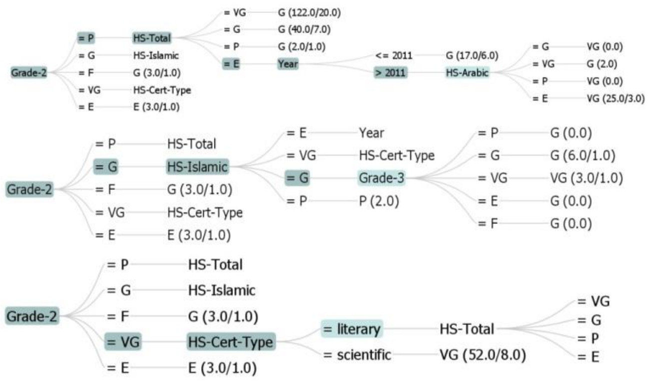

Fig. 3: Part of J48 model generated tree

2) The Support Vector Machines (SVM) Algorithm

The second algorithm used in this research is the SVM algorithm [14] which builds a model that assigns new examples to one class or the other, making it a non-probabilistic binary linear classifier. The model represents the examples as points in space, so that the examples of the separate classes are divided by a clear gap that is as wide as possible. New examples are then mapped to the same space and predicted to belong to a class based on which side of the gap they fall on. The SVM polynomial kernel version of SVM is a kernel function which is commonly used with support vector machines and other kernelized models, that represent the similarity of training samples in a feature space over polynomials of the original variables, allowing learning of non-linear models. The polynomial kernel looks not only at the given features of input samples to determine their similarity, but also at combinations of these. In the context of regression analysis, such combinations are known as interaction features. The feature space of a polynomial kernel is equivalent to that of polynomial regression, but without the combinatorial blowup in the number of parameters to be learnt. The WEKA SMO classification is applied on the dataset during the experimental study as shown in Table 7.

3) The Fuzzy Unordered Rule Induction Algorithm (FURIA) Algorithm

The third algorithm used in this research is rule based algorithm which is named the Fuzzy Unordered Rule Induction Algorithm (FURIA) algorithm [3].The algorithm extends the well-known RIPPER algorithm (kind of rule learner), while preserving its advantages such as simplicity and its comprehensible rule sets. In addition, it includes a number of modifications and extensions. The FURIA learns fuzzy rules instead of conventional rules and unordered rule sets instead of rule lists. Moreover, to deal with uncovered examples, it makes use of an efficient rule stretching method. The advantages of rule based are that they represent rules which could easily be understood and interpreted by users. The WEKA FURIA classification is applied on the dataset during the experimental study. The resultant rules from the FURIA algorithms is shown in Table 8.

4) Enhancing classification results by Voting

The final enhancement performed on the model using one of the ensemble learning method is called "Voting" [15]. The Voting algorithm combines the above mentioned classifications algorithms (J48, SVM, and FURIA) to combine several classifiers together in order to achieve improved recognition performance .

V. The Results and Analysis

The TP for final grade (Excellent, Very Good, Good, Pass) are the percentages of students who are truly classified as correct grade and predicated by the system relative to the total number of students who are incorrect grade. These two measures are computed for different learning algorithms, namely J48, SVM, FURIA, and Voting three algorithms as shown in Table 9 and Fig. 4. The overall accuracy is computed as the percentage of the correctly predicated students relative to the total number of students of the training set. Also the details of evaluation measure for voting three algorithms presented in Table 10.

Table 7: The sample from generated support vectors

Classifier for classes:					
P, E	**VG, E**	**VG, P**	**G, E**	**G, P**	**G, VG**
-0.915 *Grade -2=P	-0.2019 *Grade-2=P	-0.6751 *Grade-2=P	-0.383 *Grade-2=P	0.0691 *Grade-2=P	-0.8183 *Grade-2=P
0.4966 *Grade -2=G	-0.3511 *Grade-2=G	-0.6743 *Grade-2=G	-0.555 *Grade-2=G	-0.2226 *Grade-2=G	-0.2418 *Grade-2=G
0.4184 *Grade -2=VG	-0.2434 *Grade-2=VG	- 0.3795 *Grade-2=F	-0.4001 *Grade-2=VG	-0.2192 *Grade-2=F	-0.5764 *Grade-2=VG
0.2293 *HS -Islamic=E	-0.7964 *Grade-2=E	-0.3804 *Grade-2=VG	-0.5379 *Grade-2=E	-0.0658 *Grade-2=VG	-0.2521 *HS-Islamic=E
0 *HS -Islamic=VG	-0.1238 *HS-Islamic=E	-0.3118 *HS-Islamic=E	-0.4961 *HS-Islamic=E	-0.298 *HS-Islamic=E	-0.5955 *HS-Islamic=VG
-0.2293 *HS -Islamic=P	-0.1238 *HS-Islamic=VG	-0.2287 *HS-Islamic=VG	-0.4693 *HS-Islamic=VG	-0.5793 *HS-Islamic=VG	-0.3434 *HS-Islamic=G
-0.1588 *HS -Total=VG	-0.1625 *HS-Total=VG	- 0.5404 *HS-Islamic=P	-0.0268 *HS-Islamic=G	-0.4282 *HS-Islamic=G	-0.2203 *HS-Total=VG
0 *HS -Total=G	-0 *HS-Total=G	- 0.3364 *HS-Total=VG	-0.0223 *HS-Total=VG	-1.3056 *HS-Islamic=P	-0.1281 *HS-Total=G
-0.2293 *HS -Total=P	-0.1625 *HS-Total=E	-0.3101 *HS-Total=G	-0.0268 *HS-Total=G	-0.1467 *HS-Total=VG	-0.6882 *HS-Total=P
0.3882 *HS -Total=E	-0.3281 *Year	- 0.5404 *HS-Total=P	-0.0045 *HS-Total=E	-0.0931 *HS-Total=G	-0.3398 *HS-Total=E
0.1241 *Year	-0.3634 *Grade-1=P	-0.5668 *HS-Total=E	-0.1736 *Year	-0.3056 *HS-Total=P	-2.8203 *Year
-0.2293 *Grade -1=P	-0.3791 *Grade-1=G	-0.1573 *Year	-0.427 *Grade-1=P	-0.0658 *HS-Total=E	-0.1905 *Grade-1=P
-0.0616 *Grade -1=G	-0.0539 *Grade-1=VG	- 0.4394 *Grade-1=P	-0.4145 *Grade-1=G	-0.3559 *Year	-0.7483 *Grade-1=G
0.291 *Grade -1=VG	-0.7964 *Grade-1=E	- 0.1028 *Grade-1=G	-0.3132 *Grade-1=F	-0.1606 *Grade-1=P	-1 *Grade-1=F

Table 8: The generated FURIA Rules

#	FURIA rules (23 Rules)
1	(Grade-2 = P) and (HS-Total = VG) => Grade-4(class)=G (CF = 0.83)
2	(HS-Total = G) => Grade-4(class)=G (CF = 0.8)
3	(Year in [-inf, -inf, 2011, 2012]) and (Grade-2 = P) => Grade-4(class)=G (CF = 0.79)
4	(HS-Cert-Type = literary) and (HS-Islamic = VG) => Grade-4(class)=G (CF = 0.82)
5	(Year in [-inf, -inf, 2011, 2012]) and (Grade-3 = P) => Grade-4(class)=G (CF = 0.85)
6	(Year in [-inf, -inf, 2011, 2012]) and (HS-Scientific = P) and (Grade-3 = G) => Grade-4(class)=G (CF = 0.82)
7	(Grade-1 = P) and (HS-Scientific = G) and (HS-Arabic = VG) and (Grade-3 = G) and (Grade-2 = G) => Grade-4(class)=G (CF = 0.95)
8	(Grade-3 = P) and (HS-Total = VG) and (Grade-1 = P) => Grade-4(class)=G (CF = 0.98)
9	(Year in [2011, 2012, inf, inf]) and (Grade-1 = G) => Grade-4(class)=VG (CF = 0.9)
10	(Year in [2011, 2012, inf, inf]) and (HS-Arabic = E) and (Grade-2 = G) and (HS-Total = E) => Grade-4(class)=VG (CF = 0.96)
11	(HS-Arabic = E) and (Year in [2011, 2012, inf, inf]) and (Grade-3 = VG) => Grade-4(class)=VG (CF = 0.95)
12	(HS-Arabic = E) and (Year in [2011, 2012, inf, inf]) and (Grade-3 = G) and (HS-Total = E) => Grade-4(class)=VG (CF = 0.98)
13	(HS-Scientific = VG) and (Grade-1 = P) and (Grade-3 = G) => Grade-4(class)=VG (CF = 0.84)
14	(HS-Arabic = E) and (Grade-3 = E) and (HS-Total = VG) => Grade-4(class)=VG (CF = 0.91)
15	(HS-Arabic = E) and (Grade-2 = G) and (HS-Total = E) and (HS-Scientific = G) => Grade-4(class)=VG (CF = 0.74)
16	(HS-Arabic = E) and (Grade-2 = G) and (Grade-3 = G) and (HS-Scientific = G) and (Grade-1 = P) and (HS-Cert-Type = scientific) => Grade-4(class)=VG (CF = 0.9)
17	(Grade-3 = VG) and (Grade-2 = G) and (HS-Total = VG) => Grade-4(class)=VG (CF = 0.68)
18	(Grade-2 = VG) and (HS-Scientific = P) and (HS-Islamic = E) => Grade-4(class)=VG (CF = 0.93)
19	(HS-Total = E) and (Grade-1 = P) and (Grade-3 = G) => Grade-4(class)=VG (CF = 0.88)
20	(HS-Scientific = P) and (Grade-3 = P) and (HS-Islamic = G) => Grade-4(class)=P (CF = 0.51)
21	(HS-Scientific = P) and (HS-Islamic = P) => Grade-4(class)=P (CF = 0.51)
22	(Grade-1 = VG) and (HS-Total = E) and (Grade-3 = E) and (HS-Scientific = VG) => Grade-4(class)=E (CF = 0.61)
23	(Grade-1 = VG) and (HS-Scientific = E) => Grade-4(class)=E (CF = 0.51)

Table 9: The accuracy of different machine learning algorithms applied for IIR model

	J48	SVM	FURIA	Voting
Overall Accuracy (Correctly Classified Instances)	77.9%	74.6%	75.1%	78.5%

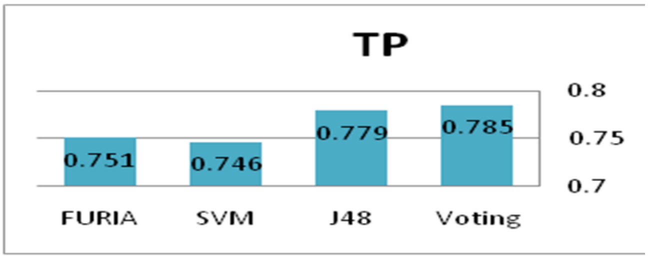

Fig. 4: The accuracy of different machine learning algorithms applied for IIR model

Table 10: The details accuracy of voting three algorithms

Class	Recall	Precision	FP Rate	TP Rate
P	0.167	0.4	0.004	0.167
G	0.754	0.803	0.16	0.754
VG	0.851	0.777	0.246	0.851
E	0.308	0.667	0.003	0.308
Weighted Avg.	0.785	0.78	0.198	0.785

The results show that J48 algorithm has the highest accuracy of 77.9% among all algorithms. The FURIA is the second best accuracy and finally the SVM algorithm comes the last. The results also reveal that the True Positive Rate is higher for the classes Very Good & Good (85, 75 %) compared to Excellent & Passed classes (30, 16 %). From the above results, it was concluded that J48 outperforms among other tested algorithms. To better enhance the obtained result, the three algorithms were stacked by the Voting technique and achieved 78.5% overall accuracy

VI. THE CONCLUSIONS

In this paper, an advisory model has been proposed for checking the suitability of a major for a student. The research firstly applies the logistic regression model to check the suitability of the chosen major for a student then all the students who pass the suitability test are reevaluated to predict their performance in the major. The evaluation is based on different learning criteria, namely, student grades, certificate, age, and gender. The proposed model is applied on a selective case study, namely the faculty of Commerce /accounting, Al-Azhar University. After regression model to test suitability has been applied, three machine learning algorithms were employed and compared. The J48 algorithm outperforms other algorithms, so it is considered as the best algorithm to employ in the proposed model. The overall accuracy of the model approaches 78.5%, which would later be enhanced by utilizing other learning techniques and adding more explanatory variables. The model may also be used by faculty staff to predict and identify weak students and then can take appropriate actions to help them.

APPENDIX

a. Variable(s) entered on step 1: HSCertType, Area, HSIslamic, HSArabic, HSScientific, HSTotal.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1a	HSCertType(1)	.989	.493	4.023	1	.045	2.688
	Area(2)	-.713-	.338	4.437	1	.035	.490
	HSIslamic(3)	-1.268-	1.936	.429	1	.513	.281
	HSArabic(3)	-1.000-	.800	1.564	1	.211	.368
	HSScientific(3)	-.444-	.514	.746	1	.388	.641
	HSTotal(3)	.831	1.417	.344	1	.557	2.296
	Constant	1.600	.551	8.434	1	.004	4.953

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	420.823a	.056	.112

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	38.247	15	.001
	Block	38.247	15	.001
	Model	38.247	15	.001

Classification Table

	Observed		Predicted: Suitability 0	Predicted: Suitability 1	Percentage Correct
Step 1	suitability	0	0	73	.0
		1	0	587	100.0
	Overall predictive accuracy Percentage				88.9

REFERENCES

[1] S. L. Salzberg, "C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993," *Mach. Learn.*, vol. 16, no. 3, pp. 235–240, Sep. 1994.

[2] Y. Goldberg and M. Elhadad, "splitSVM: fast, space-efficient, non-heuristic, polynomial kernel computation for NLP applications," *HLT '08 Proc. 46th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol.*, no. June, pp. 237–240, 2008.

[3] J. Hühn and E. Hüllermeier, "FURIA: An algorithm for unordered fuzzy rule induction," *Data Min. Knowl. Discov.*, vol. 19, pp. 293–319, 2009.

[4] Q. a Al-Radaideh, E. M. Al-Shawakfa, and M. I. Al-Najjar, "Mining student data using decision trees," in *International Arab Conference on Information Technology (ACIT'2006).*, 2006.

[5] R. Kabra and R. Bichkar, "Performance Prediction of Engineering Students using Decision Trees.," *Int. J. Comput. Appl.*, vol. 36, no. 11, pp. 8–12, 2011.

[6] S. Sembiring, M. Zarlis, D. Hartama, S. Ramliana, and E. Wani, "Prediction of Student Academic Performance By an Application of Data Mining Techniques.," *Int. Conf. Manag. Artif. Intell.*, vol. 6, pp. 110–114, 2011.

[7] Q. a Al-radaideh, A. Al Ananbeh, and E. M. Al-shawakfa, "a Classification Model for Predicting the Suitable Study Track for School Students," *Ijrras*, vol. 8, no. August, pp. 247–252, 2011.

[8] E. Review, "Data mining approach for predicting student performance," *J. Econ. Bus.*, vol. X, no. 1, pp. 3–12, 2012.

[9] S. K. Yadav and S. Pal, "Data Mining : A Prediction for Performance Improvement of Engineering Students using Classification," *Wcsit*, vol. 2, no. 2, pp. 51–56, 2012.

[10] D. Kabakchieva, "Predicting student performance by using data mining methods for classification," *Cybern. Inf. Technol.*, vol. 13, no. 1, pp. 61–72, 2013.

[11] P. McCullagh and J.A. Nelder., *Generalized linear models*. 1989.

[12] "GainRatioAttributeEval." [Online]. Available: http://weka.sourceforge.net/doc.stable/weka/attributeSelection/GainRatioAttributeEval.html. [Accessed: 19-Mar-2016].

[13] R. Kovahi; and F. Provost, "Glossary of Terms," *Mach. Learn.*, vol. 30, no. 2–3, pp. 271–274, Feb. 1998.

[14] C. Chang and C. Lin, "LIBSVM : A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 1–39, 2011.

[15] *Combining Pattern Classifiers: Methods and Algorithms - Kuncheva - Wiley Online Library*. 2004.

Mohamed M. Ezz is associate professor in Faculty of Engineering Al Azhar University. He received his B.Sc., M.Sc. and Ph.D. in Systems & Computers Engineering from Faculty of Engineering, Al Azhar University. He is IEEE member. His area of interest includes application security, intrusion detection, semantic web, and pattern recognition. He has published 20 scientific papers in various national and international journals and conferences. He has contributed in more than 16 mega software projects in Electronic banking EBPP, EMV, mobile banking and e-commerce.

Hany M. Harb is a professor in Faculty of Engineering at Al-Azhar University. He is currently the Head of the Computers and Systems Engineering Department. He received his Doctor of philosophy (Ph.D.), Computer Science, Illinois Institute of Technology (IIT), Chicago , Illinois , USA, 1986. He received his Master of Science in Operations Research (MSOR), IIT, 1987. He received his Master of Science in Computers and Systems Engineering, Faculty of Engineering, Azhar University , Cairo , Egypt, 1981. He received his Bachelor of Science in Computers and Control Engineering, Faculty of Engineering, Ain Shams University, Cairo, Egypt, 1978. His area of interest includes object Oriented Software Engineering Modeling Using UML, Software engineering, Distributed Systems and Computer Networks.

Abeer Sayed is a assistant professor in the faculty of Commerce, Statistics Department, AlAzhar University. She got her Doctor of philosophy (Ph.D.), Master of Science (M.Sc.), and B.Sc. from the same department. Her area of interest includes statistics, data analytics, Big data, and data mining.