



OPEN ACCESS

EURASIA Journal of Mathematics Science and Technology Education
ISSN: 1305-8223 (online) 1305-8215 (print)
2017 13(8):5499-5517
DOI 10.12973/eurasia.2017.00939a



E-Assessment and Computer-Aided Prediction Methodology for Student Admission Test Score

Muhammad Usman

University of Engineering and Technology, Taxila, Pakistan

Muhammad Munwar Iqbal

University of Engineering and Technology, Taxila, Pakistan

Zeshan Iqbal

University of Engineering and Technology, Taxila, Pakistan

Muhammad Umar Chaudhry

Sungkyunkwan University, Suwon, South Korea

Muhammad Farhan

COMSATS Institute of Information Technology, Sahiwal, Pakistan

Muhammad Ashraf

Sungkyunkwan University, South Korea

Received 17 November 2016 • Revised 30 December 2016 • Accepted 16 April 2017

ABSTRACT

Machine Learning is a scientific discipline that addresses learning in context is not learning by heart but recognizing complex patterns and makes intelligent decisions based on data. Currently, students have to face the problem of selecting the best suitable university for admission in engineering. There is no predictor system that recommends the students to select the specific category which is best to its academic career. Students have to first appear in the entry test and can't predict whether he/she can pass the entry test to get admitted in University. To tackle this problem the field of Machine Learning develops algorithms that discover knowledge from specific data and experience, based on sound statistical and computational principles. After going through the entry test students have to face problems for selecting the preferences among different categories due to the lack of knowledge of intake merits of preceding years. Another problem arises when students are waiting for admission in specific university, meanwhile, other universities finish their admission processes and select the students, but some students can't take admission in any university due to no prediction system for admission in universities. In this work, we would like to develop an E-Assessment and Computer-Aided Prediction online system that enables the student to predict the entry test numbers by giving the Metric and Intermediate marks and other academic numbers. The suggested scheme has been demonstrated to perform at the maximum speed under MATLAB setup.

Keywords: multiple regression, computer aided prediction, admission test score, machine learning

© **Authors.** Terms and conditions of Creative Commons Attribution 4.0 International (CC BY 4.0) apply.

Correspondence: Muhammad Munwar Iqbal, *University of Engineering and Technology, Pakistan.*

Phone: +92-321-4502399

✉ E-mail: munwariq@gmail.com

State of the literature

- The authors have been conducted by researchers to predict recommendations on the basis of 'Web of Trust' with accurate suggestions (Ben-Shimon et al., 2007; Ziegler & Lausen, 2004).
- The researchers Group Recommender System is also known as e-group activity recommender system, and suggest preferences in various domains for example music, movies, web links, events and vacation plans (Brusilovski, Kobsa, & Nejdli, 2007; Garcia & Sebastia, 2014).
- The work of discusses a G2B system 'Smart Trade Exhibition Finder' (STEF) which predicts on the basis of semantic similarity mechanism. Recommendations based on fuzzy relations are used to reflect the graded or uncertain information in the G2B techniques (Ghazanfar & Prügell-Bennett, 2014; Guo & Lu, 2007).
- Zaiane developed an e-learning agent which predicts the information by working on association rule mining. This recommender system suggests the courses, learning the material and interesting subjects to the students of e-learning systems (Klašnja-Milićević, Ivanović, & Nanopoulos, 2015).

Contribution of this paper to the literature

- The research focuses on the accurate usage of semantically relevant data of old students for the mark prediction.
- The first contribution of this research is a new realistic feature regression technique that supports understanding of machine learning.
- The second contribution is finding the best predictors variables that fit the best regression line.
- The third contribution is the presentation of results of prediction online system that shows the result on the basis of old data of students.

INTRODUCTION

In today's technology driven era it would be difficult to come across any field not yet affected by wonderful advancements in computer technology. Specifically, the domain of machine learning techniques enables the computer to learn without openly programmed on the basis of previous data sets. By using the Machine Learning techniques and Statistical tools it would be easy to enhance the predictive power. These techniques require the ability and skill on data analysis, data scientists on raw data by refining it for prediction. For data analysis, Machine Learning is a very powerful tool to get the required results. By data sources, computing power can be gain during the process. Predictions can be making easily by going straight to the data. It is one of the most significant ways for prediction. Traditional predicting from the old data is unavailable and retrieval approaches from old data have proven to be hopeless, time-consuming and burdensome. The incompetence of traditional prediction searches can be dealt with Machine Learning based predictor System Easily. Machine learning tools give the opportunity for a decision support system (DSS). Decision support system are computer application programs that go to the analyses of business data and present the data in an efficient way that will be easy for the user to make its own business decision very easy (Chen, Chiang, & Storey, 2012). Decision support system is basically informational applications that work in normal business operations.

Machine Learning is a scientific discipline that enables the user to make a program for a system that automatically learns and to improve with experience. Learning in this context is not learning by heart but recognizing complex patterns and make intelligent decisions based on data. Currently, students have to face the problem of selecting the best suitable University for admission in Engineering. There is no suitable predictor system that recommends the students to select the specific category which is best to its academic career. Students have to first appear in the entry test and can't predict whether a student can pass the entry test to get admitted in University (Koljatic & Silva, 2013). To tackle this problem the field of Machine Learning develops algorithms that discover knowledge from specific data and experience, based on sound statistical and computational principles. The presented system 'BizSeeker' provides a list of suggested partners by calculating semantic similarities. This system is further extended in and named as Smart BizSeeker based on a hybrid fuzzy semantic recommendation (HFSR) (Lu, Shambour, Xu, Lin, & Zhang, 2013).

After going through the entry test students have to face problems for selecting the preferences among different categories due to no knowledge of intake merits of preceding years. Another problem arises when students are waiting for admission in specific university, meanwhile, other universities finish their admission processes and select the students, but some students can't take admission in any university due to no prediction system for admission in Universities (Lu et al., 2013). In this work, we would like to develop an online decision support system that enables the student to predict the entry test numbers by giving the age, gender, Metric Maximum marks, Metric passing year, Metric marks obtained and as well as Intermediate data. The desired output can be achieved by effectively using the techniques of machine learning that are statistical tool regression. The regression line is used for predict the required results. The best-fitted line is calculated by using the high computational mathematical tool MATLAB (Thielicke & Stamhuis, 2014). MATLAB functions used on previous data of entry tests and applied the linear regression. Further is also enabling the student to select its choice of preference. The system supports the student whether a student can be eligible for that selected category. The system will manipulate from the old merit lists the best preference for the student (Hoxby & Avery, 2013). The focus of the presented dissertation is in more consistent ways to retrieve semantically correct result in an efficient way. The proposed technique resolves retrieval issues through machine learning techniques by improving the required results.

RESEARCH METHODOLOGY

The model used for the data collection and analysis is shown in Figure 1.

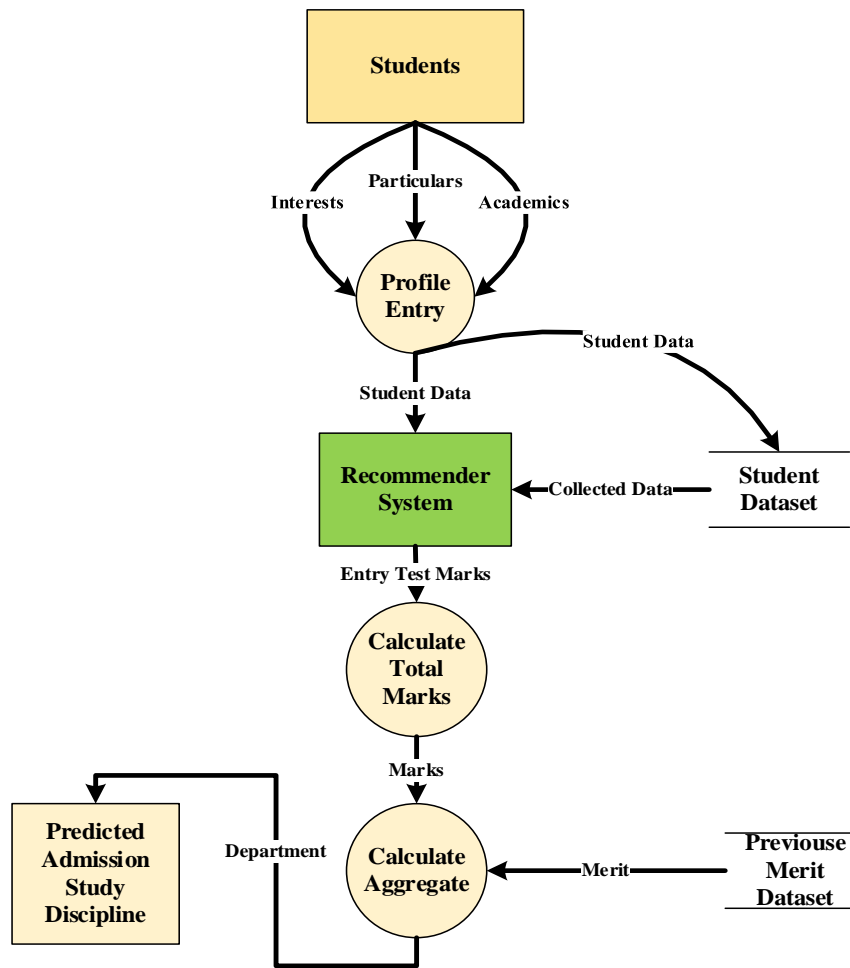


Figure 1. Model for the assessment of the Entry Test and Prediction Model

This model can be applied to any type of data problem on the basis of length and specific outcome of the dataset.

Linear Regression

This is most commonly and reliable estimation technique. This takes the actual values from the users and based on the constant variables it predicts the best-fitted line (Montgomery, Peck, & Vining, 2015). The example of linear regression is to know the estimated cost of the houses in the specific location, number of sales will be done in the specific region of the basis of old data, how much calls will be received in the specific time, what will be the rate of stock exchange with the real time up and downs. In the linear regression, the combination of only two variables will be used. The relationship between the dependent and independent variables has been established by using the best-fitted line. The best-fitted line is called regression line. Representation of lie is done by a linear equation. The linear equation is:

$$Y = a + b * X \quad (1)$$

In the above equation:

Y= Dependent variable
 a= Slope of the line
 X= Independent Variable
 b= intercept of the line

Where Y is the variable whose value we are going to be found or estimated for prediction. a is the slope of the line and will be calculated by using the prescribed formula. X value is the independent value and is given by the user. The intercept of the line is also calculated by the formula.

$$a = \frac{\sum_{i=1}^n XiYi - \sum_{i=1}^n Xi \sum_{i=1}^n Yi}{n \sum_{i=1}^n xi^2 - (\sum_{i=1}^n xi^2)} \quad (2)$$

$$b = \frac{1}{n} \sum_{i=1}^n Yi - a \sum_{i=1}^n Xi \quad (3)$$

The linear regression can be understood by the example of arranging the people in class on the basis of increasing order of their weight. But the weights of the people are unknown. The people will be arranged on relying the heights of the people which is to be known by analyzing the visually. The combination of the people is to be made on using the visible parameters. The heights are to be figured and correlation of height and weight is to be made by using the regression line given above for relationship (Faraway, 2016). The coefficients 'a' and 'b' are calculated by minimization the sum of squared difference of distance by using the data points and regression line. Suppose the best line of the data has been calculated for the linear equation is as:

$$a = 0.2811 \text{ and } b = 13.9$$

By putting the values into the regression straight line equation:

$$y = a * X + b$$

$$y = 0.2811 * X + 13.9$$

We have now calculated the linear regression line. By putting the height at the place of X, the weights of the people can be easily predicted.

Multiple Linear Regressions

Multiple regression is an extension of simple linear regression technique. Simple linear regression deals with the bi-variant samples, which deals only with two variables the one is dependent variable Y and the another one is independent variable X. Multiple regression deals with three or more variables, in which the one is dependent variable and other are

multiple independent variables (Covrig & McConaughy, 2015). For this we will include new terms for every entrance of each independent variable, the equation will come as:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k \quad (4)$$

The β 's are coefficients for the independent variables in the true or population equation and the x 's are the values of the independent variables for the member of the population (Nath Das & Mukhopadhyay, 2016). The intercept α or β_0 is the value where regression plane cross on the Y axis. The value of the β_1 predicts Y per unit X_1 . The slope for variable X_2 (β_2) predicted the alteration in Y per unit X_2 having X_1 constant.

Significance Test

Testing the level of significance is done by F and t-test. When the simple linear regression line is fitted then the t and F test gives the answer. While in the multiple regression F and t tests provides a different conclusion.

F-test: In this test, we try to examine whether the significant relationship has been existing between the set of dependent variables and all independent variables. F test is also called the test of all the variables.

T-Test: F test checks the overall level of significance among the variables while the t-test determines the significance among all the independent variables separately. For each independent variable, the separate t-test will be determined. The t-test is also known as a test for individual significance.

Testing of F test significance

The hypothesis was made for checking the significance.

$$H_0 = \beta_0 = \beta_1 = \beta_2 = \dots = \beta_p = 0$$

H_1 = One of more of the parameters is not equal to 0

Testing of t test for significance

Hypothesis can be checked for t test is by

$$H_0: \beta_i = 0$$

$$H_a: \beta_i \neq 0$$

RESULTS AND DISCUSSION

The execution evaluation of any structure can give information about the aftereffects of trademark estimations. The method fundamentally examines the framework to upgrade computation's layout and demonstrates the results procure by realizing and differentiating the estimation. The execution setup of experimentation consolidates Matlab 2015 on

windows 7 working structure presented on the corei7 machine. The proposed arrangement is contemplated against standard counts similarly as precision, the exactness and correctness of results, and the audit estimations of the system.

Student Datasets

The dataset composed of 5042 students has been collected from the admission section of University of Engineering and Technology, (UET) Taxila. The dataset is collected from the admission department of the UET, Taxila. This dataset is gathered more than 20 years records of the students, those got admission in UET Taxila in different Disciplines.

Gender disparity in this study is very natural because the collected dataset is of UET and Engineering disciplines. Male students have more tendency towards engineering as compared to female students in Pakistan. Female students have more interest in Management Sciences, Arts / Humanities, Social Sciences, and other sciences related disciplines in Pakistan. Training data is not affected by gender differences because the features set for training data is not biased towards it. Moreover, research has also negated the use of convenience sampling.

This data is used for estimation of the results of students in entry test. 11 attributes have been taken that has the direct impact on the result of the entry test marks. The data is first organized as per requirements. Machine Learning techniques are used for predicting the best results. Multiple Regression tools are used for prediction purpose. This tool is implemented by using the MATLAB 2015.

EXPERIMENTAL SETUP

The selected data set is scanned and a regression line has been drawn to fit the dataset values. On getting multiple regression lines a straight line has been drawn. The linear regression line is then used for prediction purpose. The portal has been made that takes the student academic record from the user. This record has been used for predicting the marks a student can get in entry test. In machine learning techniques, textual data is first converted into numerical data, so when there is textual data e.g. Male and Female, it is converted to a numeric value. In this case, we have assigned the value of 1 to male students and 2 to the female students. We have 4452 records of the male student and 590 records of female students. Gender disparity in this study is very natural because the collected data is of engineering disciplines. Male students have more tendency towards engineering as compared to female students in Pakistan. Female students have more interest in Arts / Humanities, Social Sciences, and other sciences related disciplines in Pakistan. Training data is not affected with gender differences because the features set for training data is not biased towards it. The dataset also contains the information of passing students in matriculation exams of different years from 1993 to 2015. This information contains 30 different boards of examination which conduct matriculation exams. For the matriculation boards, different numbers are assigned to the boards for changing to quantitative data. The number of students is also mentioned below for information. The value has been assigned to the different boards as Table.

Table 1. Values allocated to boards

Serial No.	Board of Intermediate and Secondary Education (BISE)	Value	No. of Students
1	ABBOTABAD	1	19
2	AGHA KHAN, KARACHI	2	13
3	BAHAWALPUR	3	253
4	BANU	4	05
5	D.G. KHAN	5	369
6	FAISALABAD	6	396
7	GUJRANWALA	7	566
8	IBCC ¹ ISLAMABAD	8	49
9	IBCC LAHORE	9	45
10	ISLAMABAD	10	1026
11	KOHAT	23	12
12	LAHORE (PBTE) ²	11	718
13	MALAKAND	29	02
14	MIRPUR A.J.K	22	03
15	MARDAN	24	03
16	MULTAN	12	544
17	PESHAWAR	13	12
18	QUETTA	25	35
19	RAWALPINDI	14	537
20	SAHIWAL	15	104
21	SARGODHA	16	284
22	SUKKUR	17	03
23	MIRPUR KHAS SIND	18	01
24	KARAKORAM UNIVERSITY	28	01
25	HYDERABAD SINDH	19	09
26	LARKANA SIND	27	05
27	KARACHI	20	14

The student has to enter the maximum marks of metric. The numbers depend on the specific board that may be 850 or 1050. The board wise distribution of the students is shown as in Figure 2.

The most important on which the system depends is the metric and higher secondary school marks. The greatest influence is dependent on the regression is these two variables. Multiple regression lines are used for the prediction of entry test marks. As illustrated above, the line can be represented as:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

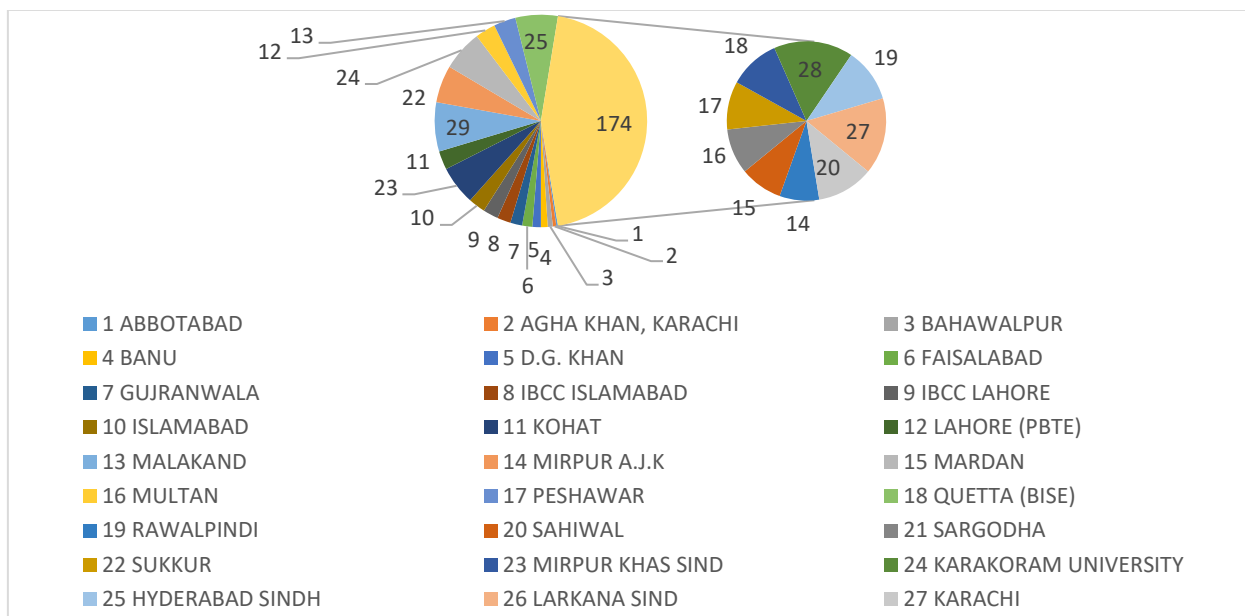


Figure 2. Board wise distribution for admitted Students

As the manual calculation for the intercept β_0 and coefficients are difficult, all the data has to be given to the MATLAB for getting the straight-line coefficients. On putting the data to the MATLAB, the outcome of coefficients comes as shown in Table 1.

Table 1. Coefficient with Values

Coefficient	Value
Intercept	24
Gender	-7.40445
Age	0.765236
Secondary School Certificate (SSC) Passing Board	0.552024
Secondary School Certificate (SSC) Maximum Marks	0.472080
Secondary School Certificate (SSC) Passing Year	0.0061357
Secondary School Certificate (SSC) Marks Obtained	0.0201989
Higher Secondary School Certificate (HSSC) Passing Board	0.8305830
Higher Secondary School Certificate (HSSC) Maximum Marks	-0.1398000
Higher Secondary School Certificate (HSSC) Passing Year	0.0167444
Higher Secondary School Certificate (HSSC) Marks Obtained	0.3449640

The above coefficients are then put into the Multiple Linear Regression Straight Line. The line is used to take the inputs from the user and by manipulation, on the inputs, the value of the dependent variable will be calculated. For evaluating the result of the prediction system record of different boards has been given to the system. Comparison of actual and predicted result has been compared and a graph of the result of prediction has been represented.

Ensemble learning method in classification, regression can be used for random forests or random decision forests that work by making multitude decision trees in training time and producing the class that is the way of classes or mean prediction of the individual trees. The correction of decision trees produced by random decision forests that over fitted to their training set. Random forest of 100 trees, and each constructed by taking 4 random technical features. Error results from out of bag error is 29.84 that is computed from dataset as shown in Table 3.

Table 3. Attribute Values generated by the Random Forest and Random Tree

Measures	Random Forests	Random Trees
Correlation coefficient	0.9657	0.9987
Mean absolute error	10.9768	0.232
Root mean squared error	14.042	2.3639
Relative absolute error	29.8459 %	0.6308 %
Root relative squared error	30.5709 %	5.1465 %

The dataset of 20 records for federal boards has been taken and their graph of actual and predicted marks by the Online Entry Test Marks Prediction System (OETMPS). Their dataset in the shown as in Table 4.

In the Figure 3 above graph, it can be seen that actual marks and predicted marks are close to each other. The difference of the predicted and actual entry test marks are represented and the median of the differences is 12.965. Another part of dataset has taken for the 20 students of Rawalpindi Board and their data is represented in tabular as shown in Table 5.

Table 4. Students marks of Islamabad Board

Sr#	SSc Marks	HSSC Marks	OETMPS Results	Actual Entry Test Marks	Difference
1	850	740	124.56	132	7.44
2	883	792	149.81	136	13.81
3	962	1000	225.65	221	4.65
4	972	922	198.94	187	11.94
5	716	887	177.84	192	14.16
6	946	919	195.72	205	9.28
7	977	935	201.04	205	3.96
8	817	836	160.33	175	14.67
9	923	905	190.40	181	9.40
10	830	916	186.60	190	3.40
11	937	881	176.62	162	14.62
12	820	891	177.70	132	45.70
13	837	915	192.14	179	13.14
14	928	892	181.97	197	15.03
15	974	968	207.38	178	29.38
16	960	908	192.21	204	11.79
17	947	936	195.79	183	12.79
18	915	836	160.65	189	28.35
19	767	877	179.27	171	8.27
20	943	928	197.10	174	23.10

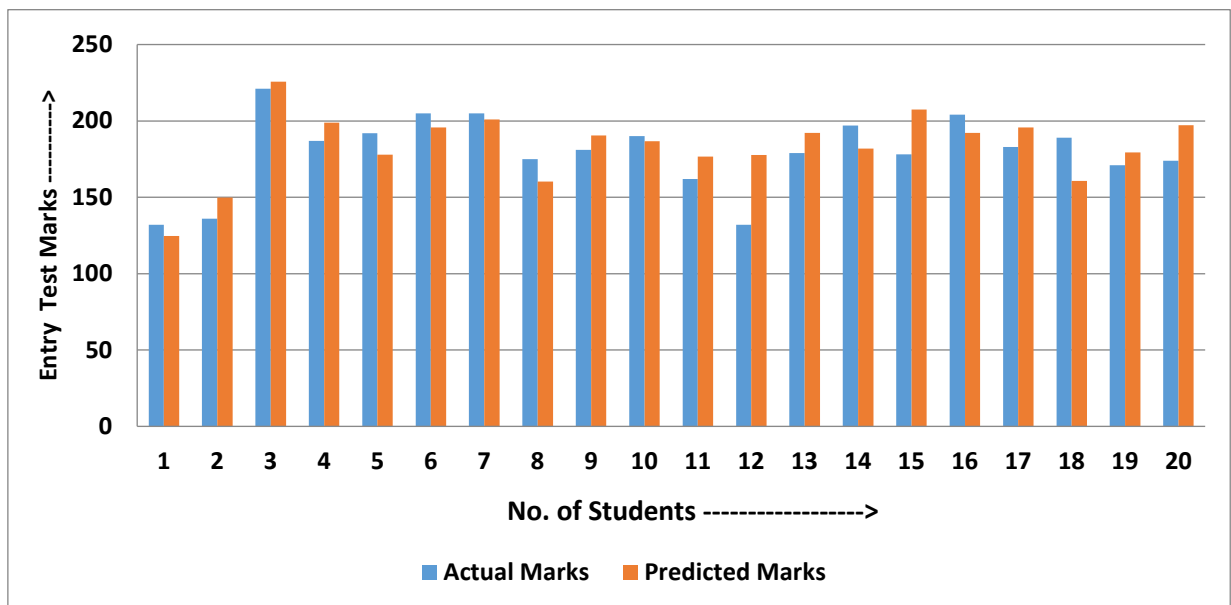


Figure 3. Analysis of Predicted and Actual Marks for the Islamabad Board

Table 5. Students marks of Rawalpindi Board

Sr#	SSC Marks	HSSC Marks	OETMPS Results	Actual Entry Test Marks	Difference
1	873	803	148.88	140	8.88
2	860	854	172.86	187	14.14
3	893	891	185.54	174	11.54
4	748	925	200.58	295	94.42
5	681	874	174.63	169	5.63
6	937	846	164.68	174	9.32
7	761	818	155.04	140	15.04
8	947	914	191.16	176	15.16
9	973	823	161.05	150	11.05
10	804	904	178.17	182	3.83
11	878	948	205.63	190	15.63
12	987	1002	222.41	201	21.41
13	851	864	162.08	150	12.08
14	900	900	191.27	187	4.27
15	956	916	195.36	182	13.36
16	618	678	106.50	89	17.5
17	950	934	197.97	187	10.97
18	873	742	135.16	129	6.16
19	891	845	156.18	144	12.18
20	711	785	143.78	155	11.22

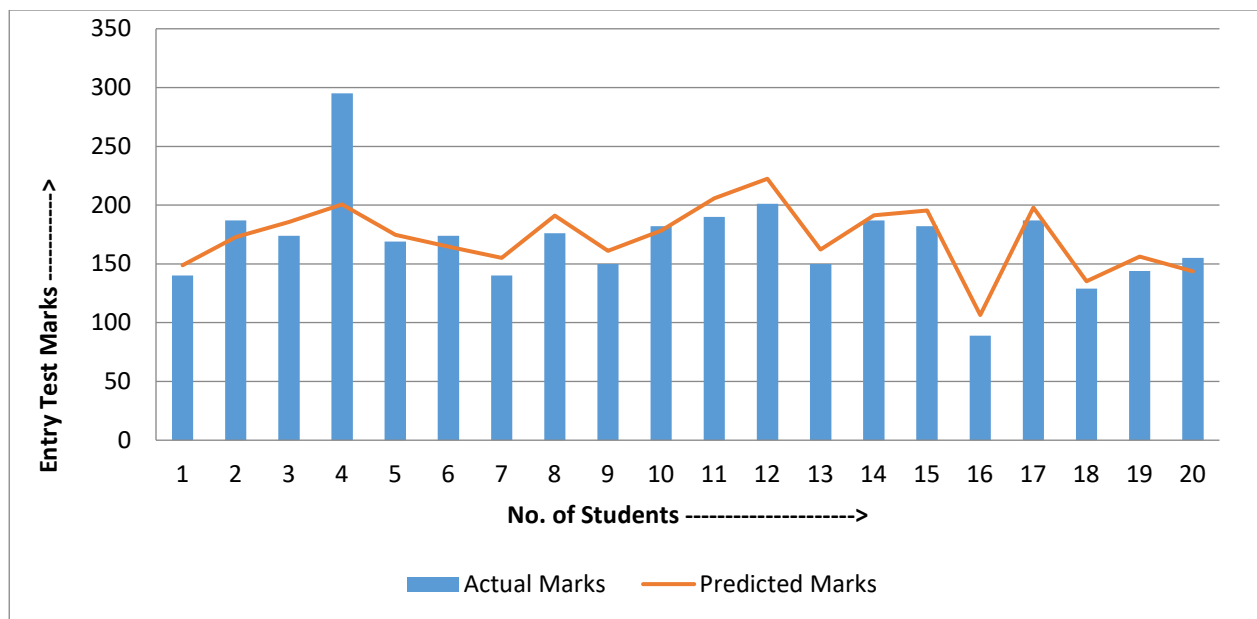


Figure 4. Analysis of Predicted and Actual Marks for the Rawalpindi Board

In the Figure 4 above graph, it can be seen that actual marks and predicted marks are close to each other. The difference of the predicted and actual entry test marks are represented and the median of the differences is 11.81. Another dataset of 20 students of Faisalabad Board has been selected from the entire dataset of students. Their dataset is shown as under in Table 6.

The Figure 5 shows the comparisons of the marks predicted by the system with actual entry test marks obtained by the student of BISE Faisalabad. The difference of the predicted and actual entry test marks are represented and the median of the differences is 9.735.

Table 6. Dataset of Faisalabad Board Students

Sr#	SSC Marks	HSSC Marks	OETMPS Results	Actual Entry Test Marks	Difference
1	892	706	114.64	128	13.36
2	971	880	180.41	174	6.41
3	983	888	179.26	212	32.74
4	853	747	127.99	115	12.99
5	946	934	183.65	176	7.65
6	791	887	171.71	165	6.71
7	953	848	167.35	175	7.65
8	899	784	141.69	130	11.69
9	798	721	118.06	66	52.06
10	964	913	186.67	187	0.33
11	929	803	141.52	150	8.48
12	876	862	164.81	154	10.81
13	920	786	135.40	145	9.6
14	917	823	155.50	115	40.5
15	925	900	181.40	117	64.4
16	893	871	174.07	168	6.07
17	955	883	169.56	172	2.44
18	947	903	187.03	194	6.97
19	939	882	168.90	184	15.1
20	726	850	168.87	159	9.87

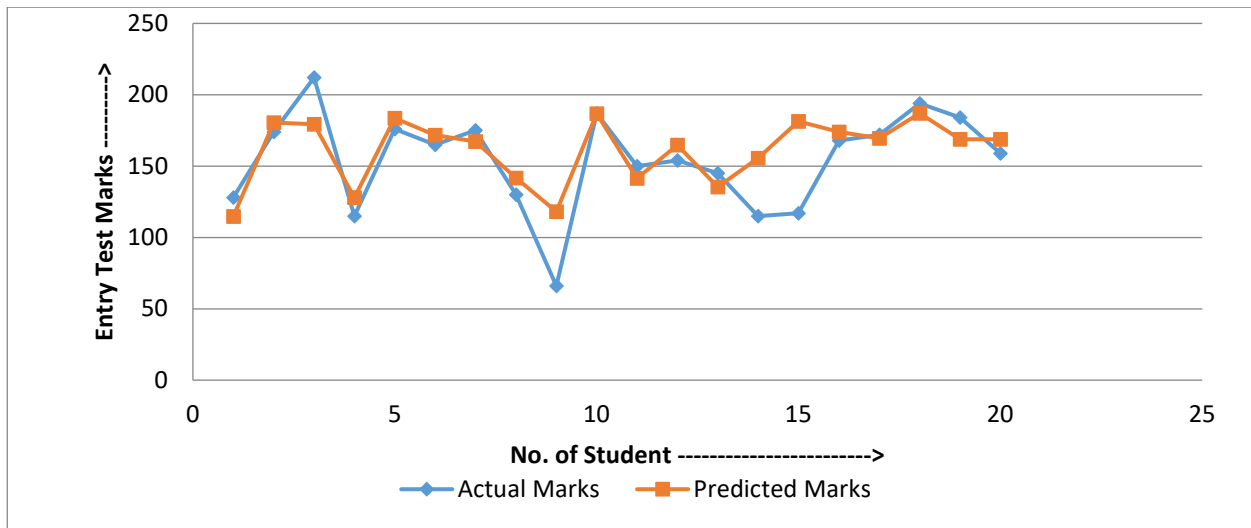


Figure 2. Analysis of Predicted and Actual Marks for Faisalabad Board

14 number of the records dataset has been selected from the list of records of the students for the Karachi Board. Their dataset is shown as under in Table 7. The Figure 6 shows the comparisons of the marks predicted by the system with actual entry test marks obtained by the student of BISE Karachi.

Table 7. Dataset of Karachi Board Students

Sr#	SSC Marks	HSSC Marks	OEMTPS Marks	Actual Entry Test Marks	Difference
1	997	949	203.41	206	2.59
2	796	864	174.25	164	10.25
3	913	930	204.37	241	36.63
4	978	971	214.01	202	12.01
5	897	937	205.55	199	6.55
6	781	782	141.51	137	4.51
7	838	862	177.50	169	8.5
8	834	886	181.70	171	10.7
9	987	967	210.17	197	13.17
10	919	796	156.53	160	3.47
11	863	871	171.37	93	78.37
12	976	880	186.66	172	14.66
13	950	853	173.49	167	6.49
14	957	986	226.31	296	69.69

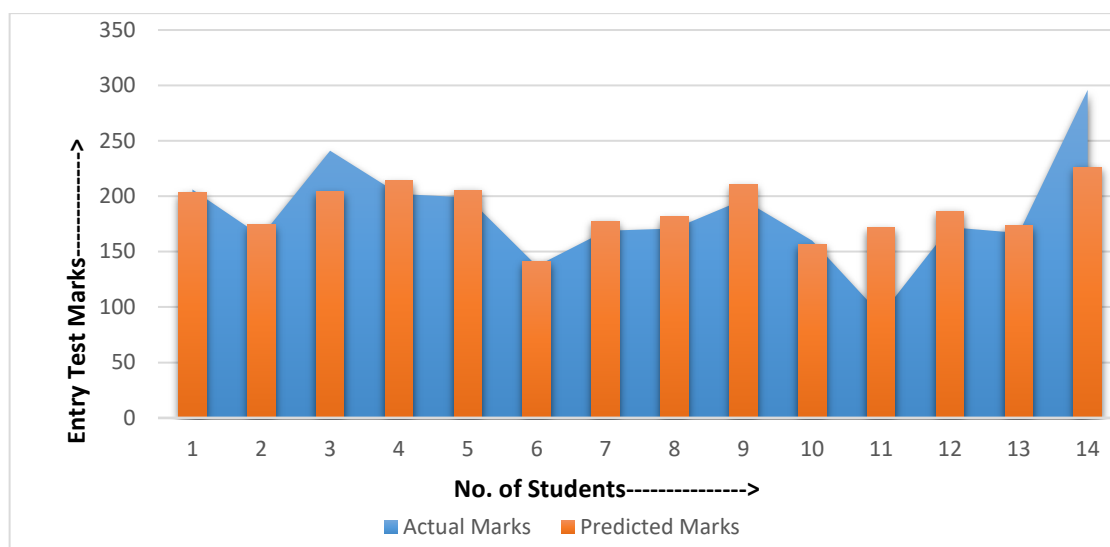


Figure 6. Comparison Analysis of Predicted and Actual Marks for Karachi Board

The 20 number of students has been selected from the Sargodha board and their respective marks are as under show in **Table 8**.

Table 8. Dataset of Sargodha Board Students

Sr#	S.Sc. Marks	HSSC Marks	OEMPTS Marks	Actual Entry Test Marks	Difference
1	814	835	162.40	154	8.4
2	848	874	176.54	181	4.46
3	882	882	179.99	169	10.99
4	959	953	205.21	208	2.79
5	810	845	166.35	178	11.65
6	690	745	129.58	98	31.58
7	888	884	180.63	173	7.63
8	906	795	143.74	129	14.74
9	948	938	201.32	191	10.32
10	866	872	171.99	179	7.01
11	948	868	169.77	172	2.23
12	823	887	183.77	190	6.23
13	955	945	193.23	187	6.23
14	866	948	209.00	199	10
15	830	696	115.53	71	44.53
16	920	947	200.69	201	0.31
17	969	972	212.87	205	7.87
18	741	741	129.33	119	10.33
19	904	666	105.69	98	7.69
20	911	852	167.58	171	3.42

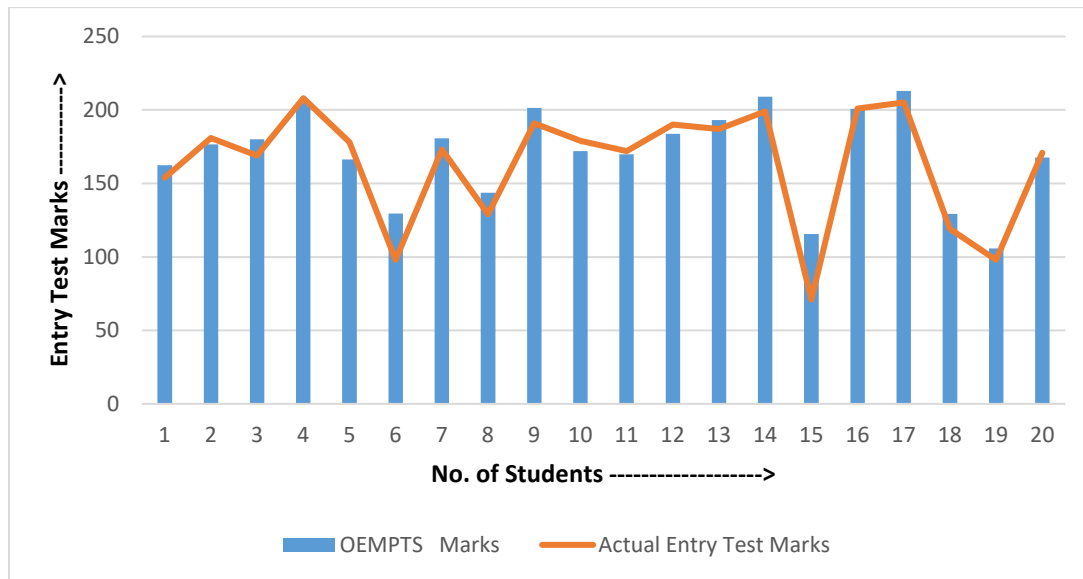


Figure 3. Analysis of Predicted and Actual Marks for Sargodha Board

The Figure 3 shows the comparisons of the marks predicted by the system with actual entry test marks obtained by the student of BISE Sargodha. The difference of the predicted and actual entry test marks are represented and the median of the differences is 10.475.

The system is developed and currently working online at the website of university of engineering and technology Taxila. Screen shots of the running system is shown as in Figures 8 and 9. The computer aided prediction system is named as Online Entry Test Marks Prediction System (OETMPS) used for calculating the student entry test marks.

It also aggregates percentage that will be possible according to the university admission rule and regulations as shown in Figure 9. This aggregate percentage result is further used for the identification of the discipline in which possible student take admission.

Factors Beyond System

Some of the factors that not considered for the implementation of the system. These factors directly or indirectly affect the performance of the students during the process of entry test. Some students take admission in academies for the preparation of entry test. They can take higher marks in the entry test as compared to the other students who don't take admission in any academy. The result their percentage of getting better marks before the entry test as compared to other students. One factor may be that a student reached at its test place after the start of entry test then it's impossible for students to solve the entire test in a limited time. The case may be that student has a better percentage in its background academic career but due to this, it's impossible to take the better entry test marks.

The screenshot shows the OETMPS (Online Entry Test Marks Prediction System) data entry form. It features a grey background with the title 'OETMPS' and subtitle '(Online Entry Test Marks Prediction System)' at the top. The form consists of several input fields, each with a blue icon of a person on the left. The fields contain the following data: '16', 'Male' (dropdown), '2014', 'AGHA KHAN, KARACHI' (dropdown), '1050', '1000', '2016', 'LAHORE (BISE)' (dropdown), '1100', '900', and 'Mechanical' (dropdown). A blue 'Calculate' button is located at the bottom right of the form.

Figure 8. Student data entry form used for the assessment of entry test marks, aggregate and admission

The screenshot shows the OETMPS results screen. It features a grey background with the title 'OETMPS' and subtitle '(Online Entry Test Marks Prediction System)' at the top. Below the title, the word 'Results' is displayed. The screen contains three lines of text: 'Dear Student Predicted Entry Test Marks = 181.41985344', 'Expected aggregate of marks = 70.879216280727', and 'The expected Admission will be in Electronics technology'.

Figure 9. Result screen for the support of student's admission

CONCLUSION AND FUTURE WORK

The focus of this research was to present a new entry test predictor system for appropriate understanding by using the statistical tools. The research activities not only show the entry test marks but also helps the student to choose the best of its carrier on the basis of previous results. This segment highlights the most vital result of this research. The explanation of this study is to beat the confinements and downside of existing methods. The presented technique offers a new procedure for test prediction results by using the statistical tools. Linear Regression Line presented in the proposed work provides the line that best fitted to the provided data. The main reason for the success of this research work is to facilitate the fresh graduates going to enter in the universities. The proposed features are classified through machine learning technique to improve the performance. In future, the proposed solution is more generic and collect the dataset of the most technical universities. This technique will be extended and can be implemented for the field of Medical Science, Basic Sciences. Other factors will be included that a student is mentally not prepared due to any reason for the test and cannot fully take attention to the test. The result marks of entry test will be affected. Some students reached their destination after the start of the test and they cannot fully give attention to their paper. Some students do not have material for their filling of test and as a result, their test has been canceled. There exists the valuable scope in the field of Machine Learning to an emphasis on the development of methods that would be able to improve the performance of the results of the entry test prediction marks.

NOTES

1. Inter Board Committee of Chairmen (IBCC)
2. Punjab Board of Technical Education (PBTE)

REFERENCES

- Ben-Shimon, D., Tsikinovsky, A., Rokach, L., Meisles, A., Shani, G., & Naamani, L. (2007). *Recommender system from personal social networks Advances in Intelligent Web Mastering* (pp. 47-55): Springer.
- Brusilovski, P., Kobsa, A., & Nejdl, W. (2007). *The adaptive web: methods and strategies of web personalization* (Vol. 4321): Springer Science & Business Media.
- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, 36(4), 1165-1188.
- Covrig, V., & McConaughy, D. L. (2015). Public versus Private Market Participants and the Prices Paid for Private Companies. *Journal of Business Valuation and Economic Loss Analysis*, 10(1), 77-97.
- Faraway, J. J. (2016). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models* (Vol. 124): CRC press.

- Garcia, I., & Sebastia, L. (2014). A negotiation framework for heterogeneous group recommendation. *Expert Systems with Applications*, 41(4), 1245-1261.
- Ghazanfar, M. A., & Prügél-Bennett, A. (2014). Leveraging clustering approaches to solve the gray-sheep users problem in recommender systems. *Expert Systems with Applications*, 41(7), 3261-3275.
- Guo, X., & Lu, J. (2007). Intelligent e-government services with personalized recommendation techniques. *International Journal of Intelligent Systems*, 22(5), 401-417.
- Hoxby, C., & Avery, C. (2013). The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity*, 2013(1), 1-65.
- Klašnja-Milićević, A., Ivanović, M., & Nanopoulos, A. (2015). Recommender systems in e-learning environments: a survey of the state-of-the-art and possible extensions. *Artificial Intelligence Review*, 44(4), 571-604.
- Koljatic, M., & Silva, M. (2013). Opening a side-gate: engaging the excluded in Chilean higher education through test-blind admission. *Studies in Higher Education*, 38(10), 1427-1441.
- Lu, J., Shambour, Q., Xu, Y., Lin, Q., & Zhang, G. (2013). a web-based personalized business partner recommendation system using fuzzy semantic techniques. *Computational Intelligence*, 29(1), 37-69.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2015). *Introduction to linear regression analysis*: John Wiley & Sons.
- Nath Das, R., & Mukhopadhyay, A. C. (2016). Correlated random effects regression analysis for a log-normally distributed variable. *Journal of Applied Statistics*, 1-19.
- Thielicke, W., & Stamhuis, E. (2014). PIVlab—towards user-friendly, affordable and accurate digital particle image velocimetry in MATLAB. *Journal of Open Research Software*, 2(1) pp. 1-10.
- Ziegler, C. N., & Lausen, G. (2004, March). *Analyzing correlation between trust and user similarity in online communities*. In International Conference on Trust Management (pp. 251-265). Springer Berlin Heidelberg. Chicago

<http://www.ejmste.com>