

## Evaluation of Student Performance based on Bridge Course

Geetha N<sup>1\*</sup>, Piyush Kumar Pareek<sup>2</sup>, Suhas G K<sup>3</sup>, Sandhya Soman<sup>4</sup>

<sup>1</sup>Dept. of Computer Science and Engineering, SOET, CMR University, Bangalore, India

<sup>2</sup>Dept. of Computer Science and Engineering, EWIT, Magadi Road Bangalore, India

<sup>3</sup>Dept. of Computer Science and Engineering, PES College, Bangalore, India

<sup>4</sup>Dept. of Computer Science and Engineering, KristuJayanti College, Bangalore, India

DOI: <https://doi.org/10.26438/ijcse/v7si15.248253> | Available online at: [www.ijcseonline.org](http://www.ijcseonline.org)

**Abstract**—Performance of the student is evaluated and estimated using various evaluation methods and parameters. Modern evaluation methods can have a tremendous impact on the student performance in their curricula. Some courses in the University curriculum has some prerequisites for particular courses and one such course in the University is Data Structures of computer science stream. Students haven't studied C programming as a prerequisite for this course and the test has been conducted. The results of this test are not satisfactory and hence a bridge course is introduced to overcome the problem of prerequisite and also the pre-test for C programming is also taken for future analysis. The bridge course is conducted for 30hrs in a laboratory since C is a programming course and post-test is conducted for both the courses. The improvement in results is identified and the performance of students is calculated. This research has been conducted on 58 students in the University, the null hypothesis is used and performed t-Test distribution to analyze the performance of students. This paper tells how a bridge course is useful for the students to perform better and suggests the best suited methods for capturing and analyzing data by choosing the right metrics and performance indicators.

**Keywords**— Education Data Mining, Bridge course, Prerequisite, t-Test, Null Hypothesis, Pre-test, Post-test.

### I. INTRODUCTION

The growth of the country is directly proportional to the student's education system. The innovative ideas in the education system for evaluation and assessing the students plays an important role. The various perspectives like an individual, personal, social and mental will be valuable to gauge the student performance. This may prompt find the understudies who are in hazard and it helps the administration to make the opportune move. The excellence of academic technical institutions is derived by the success percentage of the students and therefore the skill set of the academic institutes/university. To improve the productivity of the students it is essential to design and teach appropriate courses with the required prerequisite and have to address some issues and concerns, majorly in Engineering, master of computer science or other technical education programs. Some courses in the education system have some pitfalls in the designed curricula for students from different backgrounds like biology, computers, electronics, and others.

A survey has been conducted for 58 students who scored very fewer marks in Data Structures (DS) and the unified opinion has been taken from the students as well as faculty. The students from biology or non-computer background or without the knowledge of computer programming are

struggled in understanding the course better. So, it is decided to bring a bridge course into the picture.

**Bridge Course:** For a better understanding of the courses in a particular study program the Bridge courses are offered to students. University organizes the bridge course for new students to prepare them for university education. Every newly admitted student must have to take bridge course classes prior to starting the usual classes.

Computer (C) programming bridge course has been conducted by 2 faculties exclusively for 58 students and it went for 30 hours in the programming laboratory. Before starting of this bridge course the pre-test for C is conducted. After conducting the training program second internals for data structures is conducted and then the post-test for C is carried out. The data is collected and cleaned using data mining tools and an analysis has been made on the cleaned data using null-Hypothesis, t-Test distribution. The improvement in both courses (C and DS) is identified.

### II. RELATED WORK

Sharayu, in her paper she mentioned the broad view of analysis and useful literature survey, techniques of student data mining and also mentioned the challenges in educational data mining and the various perspectives like an individual,

personal, social and mental will be valuable to gauge the student performance[2].

A significant research is done by Bhardwaj and Pal [5] famous and useful classification method, the Naïve, on a group Bachelor of Computer Application students in Dr. R. M. L. Awadh University, Faizabad, India, who appeared for the final examination in 2010.

Abeer and Elaraby [6] processed previously enrolled students' data and analysed in a specific manner in course program for more than 5 years with significant multiple attributes that are collected from the database in university which results to predict the students' final grades in the particular course program, also "help the student's to improve their performance and suggesting them in various aspects includes special attention in particular domain to reduce failing ration and taking appropriate actions in proper time" [6].

R.Jindal, in his paper he introduced Academic Objectives and provided summary of required tools for according to their function/features [3].

RiasyahNovita, In his research he has tested a various depth of tree and threshold (alpha\_split) values with different numbers of various study programs and the result has been taken for that and the result shows that depth of tree can improve the accuracy but not significant with 85% average accuracy. [4]

It has also been predicted that end year exam performance through student activity with online tutors [8] and to predict item response outcome of the students [9].

The research is heavily dependent on the availability/accessibility of accurate and reliable data [7] so it is required to take reliable and correct data for analysis.

### III. METHODOLOGY

The marks of the first internals of data structures were collected and before doing the bridge course the pre-test for C programming is conducted. The pre-test checks the programming knowledge of the students irrespective of their biology or computer science background. The bridge course is conducted rigorously in the laboratory for 30 hours (2018-19, ODD semester) that includes all the required concepts for programming and DS. After the completion of bridge course the second internals marks of DS are collected and the post-test for C programming is conducted. A survey has been made on this collected data using suitable indicators.

#### A. Experiment settings for DS.

The prelims or first internals for this core subject is taken in a regular way along with other courses in University.

1. *Conduction:* In conduction of the internals general/regular question paper is set and the marks

has been allotted for the test. The obtained results are not satisfactory.

2. *Survey:* A survey had been made on students for the bad performance in the course and identified the problems faced by students.
3. *Biology background and C background:* According to the survey, students had a perception that they are from non-programming background like biology, electronics (other than computer language) and this has been addressed through the proper guidance to the students and decided to conduct the bridge course for 30 hrs which covers all the prerequisites for DS course and cover the programming concepts.

#### B. Pre-Test

Before starting the bridge course the knowledge of the C programming has been taken by conducting the Pre-Test. The test tests the programming skills and the marks are tabulated.

#### C. Bridge Course

For a better understanding of the courses in a particular study program the bridge courses are offered to students. Based on the opinion of the students and the faculty, University has conducted a bridge course of computer programming with the help of 2 faculty. The rigorous training has been conducted for 30 hours.

#### D. DS Post Internals

After the bridge course training the performance of the dependence course DS is tested in the second internals and the results are tabulated for analysis. Analysis gives the positive improvement in the performance than in the first internals. This analysis is conducted using t-Test and hypothesis and by setting standard alpha value to 0.05 and the results has been taken.

#### E. Post-Test

The post-test is conducted after the completion of the bridge course to test the knowledge of the students in C programming and evaluated. The results are significant than the previous test and the significance test has been performed using T-test and hypothesis and by setting standard alpha value as 0.05 and the results has been taken to check the significant dependency.

#### F. Data Collection And cleaning.

Data is collected and it is cleaned using data different mining techniques. Data cleaning in mining is the sequence process

of detecting, removing corrupt or inaccurate records from a record set, table or database and also inserting the data if necessary.

Some data cleaning methods:-

- Ignoring the tuple. This can be done when there is a missing class label. This method is not very effective, unless the tuple contains several attributes with missing values.
- Manual filling of the missing data: This method is more effect when there is a less data set. .
- Replacing with global constant: Replacing all missing attribute values with global constant, such as a label like “NA” or “Unknown” or minus infinity.
- Using the attribute mean to fill in the missing value.
- Filling with probable value.

**Clustering:** The group of similar values are clusters and values outside of this are outliers. These outliers may be detected by clustering. Values that are outside of the set of clusters may be considered outliers.

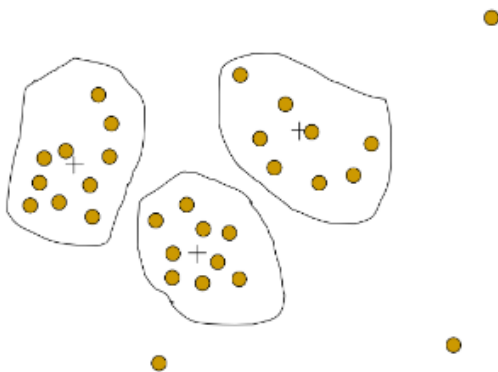


Figure 1: Clustering

**Regression:** Data can be smoothed by fitting the collected data into a regression functions.

**Binning:** Binning methods sorted data value by consulting its “neighbor- hood,” that is, the values around it.

#### G. Analysis

After data cleaning process the data is analysed using suitable mining methods and these are illustrated below.

#### 1. t-Test and Null-Hypothesis

The t-Test is performed by setting the alpha value to 0.05 and found the p-value for both DS and C. One sample t-Test is a statistical procedure that is used to determine whether a sample of observations could have been generated by a process with a specific generated mean.

Given the sample data, the main purpose of the one sample t-test is to determine if the null hypothesis should be rejected.

The null hypothesis remains the same for each type of one sample t-Test.

Null hypothesis:  $\mu_1 - \mu_2 = 0$

The hypothesis are formally defined below:

- The null hypothesis (H0) assumes that the difference between the true mean ( $\mu$ ) and the comparison value ( $m_0$ ) is equal to zero.
- The two-tailed alternative hypothesis (H1) assumes that the difference between the true mean ( $\mu$ ) and the comparison value ( $m_0$ ) is not equal to zero.

The mathematical representations of the null and alternative hypotheses are defined below:

- H0:  $\mu = m_0$
- H1:  $\mu \neq m_0$  (two-tailed)

The four steps are:

a. To calculate the sample mean.

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n}$$

b. To calculate the sample standard deviation.

$$\hat{\sigma} = \sqrt{\frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \dots + (y_n - \bar{y})^2}{n - 1}}$$

c. To calculate the test statistic.

$$t = \frac{\bar{y} - m_0}{\hat{\sigma} / \sqrt{n}}$$

d. To calculate the probability of observing the statistic test under the null hypothesis. The value is obtained by comparing  $t$  to a  $t$ -distribution with  $(n - 1)$  degrees of freedom.

$$p = 2 \cdot Pr(T > |t|) \text{ (two-tailed)}$$

$$p = Pr(T > t) \text{ (upper-tailed)}$$

$$p = Pr(T < t) \text{ (lower-tailed)}$$

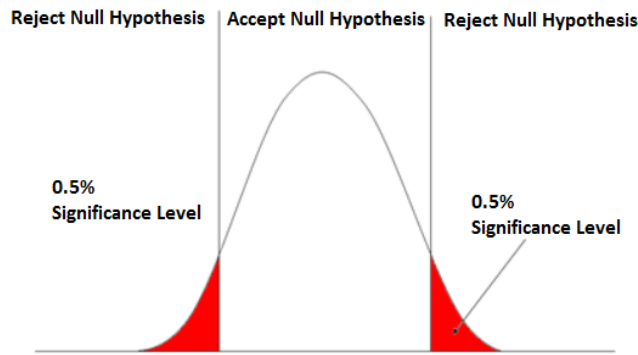


Figure 2: Two Tailed Hypothesis testing

If the statistical analysis results in significance level is below the cut-off or alpha value we have set, we can reject the null hypothesis and accept the alternative hypothesis and in our experiment the alpha value is set to 0.05 [1]. Statistical hypothesis testing was used to validate the model with a 5% level of significance.[7]

#### IV. RESULTS AND DISCUSSION

After the data cleaning the data is tabulated in excel and the data analysis is made t-Test for paired two sample for means is tested and by setting the hypothesized mean difference to zero and alpha or threshold value to 0.05 and the results are obtained.

Table 1: Marks of 58 students in DS(1<sup>st</sup> and 2<sup>nd</sup> IA) and C(pre and post)

SL. No	DS 1st IA	DS 2nd IA	C Pre Test	C Post Test
1.	19	6	6.5	11

2.	22.2	26.4	5	9
3.	19.8	25.8	5	13
4.	11.4	25.8	8	12
5.	16.8	22.2	12	14
6.	7.2	15	3	10
7.	15	15	11	11
8.	25.2	30	7.5	14
9.	8.4	24.6	15.5	13
10.	28.8	30	12.5	12
11.	27.6	28.2	9	13
12.	29	24	6.5	11
13.	8.4	18	3	13
14.	23.4	26.4	8	12
15.	25.8	27	7.5	14
16.	24	27	12	15
17.	25.8	29.4	7.5	12
18.	15	17.4	2	9
19.	27	30	7.5	13
20.	9.6	9.6	7	13
21.	9	19.2	4	12
22.	26.4	27.6	6.5	15
23.	16.8	20.4	4.5	15
24.	25.2	26.4	7	15
25.	13.2	13.8	6	10
26.	15	18.6	8	14
27.	25.8	27	10	13
28.	15	18.6	5	12
29.	21.6	25.8	5.5	14
30.	20.4	22.8	7.5	12
31.	27.6	29.4	7	14
32.	27.6	29	10	15
33.	21.6	22.2	6	13
34.	19.8	22.8	5	11
35.	16.8	26.4	5.5	14
36.	10.2	26.4	12.5	13
37.	4.2	12.6	7	13
38.	23.4	25.8	4	9
39.	15.6	22.6	5	11
40.	15	24	2	11
41.	28.8	28.8	7.5	11
42.	10.2	20	1.5	11
43.	10.2	9.6	3	13
44.	17.4	27	3	13
45.	21	22.2	3	12
46.	26.4	27	8	14

47.	5.4	24	1	11
48.	6.6	12.6	0	14
49.	11.4	12.6	2	10
50.	16.8	10.2	5	13
51.	10.8	8.4	1	11
52.	17.4	22.8	1	13
53.	15	15	2	10
54.	9.6	20.4	3	13
55.	15	15.6	1	13
56.	15	25	4	11
57.	16.2	18	2	11
58.	24	27.6	11	14

Table 2: Analysis for dependence course DS.

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	5.913793103	12.37931034
Variance	11.97489413	2.590441621
Observations	58	58
Pearson Correlation	0.379242016	
Hypothesized Mean Difference	0	
Df	57	
t Stat	-15.31222139	
P(T<=t) one-tail	3.70091E-22	
t Critical one-tail	1.672028888	
P(T<=t) two-tail	7.40181E-22	
t Critical two-tail	2.002465459	

Table 3: Analysis for bridge course C.

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	17.87586207	21.82758621
Variance	48.57098609	40.47641863
Observations	58	58
Pearson Correlation	0.668274137	
Hypothesized Mean Difference	0	
Df	57	
t Stat	-5.51438038	
P(T<=t) one-tail	4.42225E-07	
t Critical one-tail	1.672028888	
P(T<=t) two-tail	8.84451E-07	
t Critical two-tail	2.002465459	

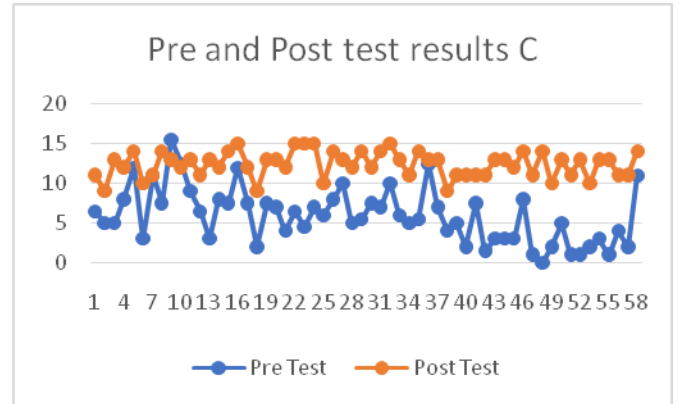


Figure 3: Comparison of Pre and Post test results for C

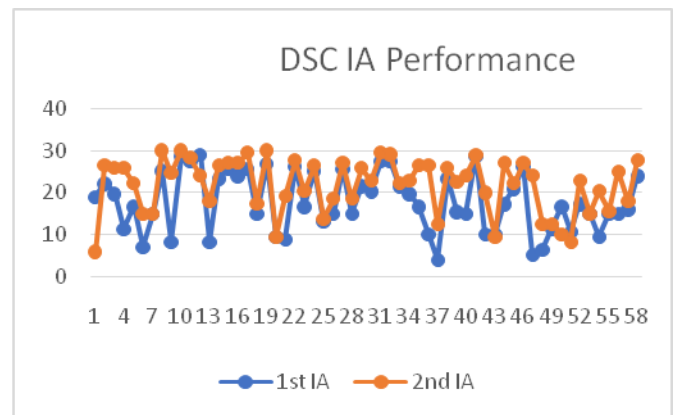


Figure 4: Comparison of first and second internal marks of DS

From the results one can conclude that the performance of the students is significantly improved both in C as well as in the dependent course DS and the p-value for DS and C are mentioned below:

The p-value for DS is 8.84451E-07

The p-value for C is 7.40181E-22

So the Null-Hypothesis is stating that “there is no significant difference between pre and post-test for C” and “no significance for first and second internals”, and hence the null-hypothesis is rejected and found the significant improvement in the student performance.

### V. CONCLUSION AND FUTURE SCOPE

The bridge course helps in improving the students’ performance in both the courses DS(dependent course) and the C(bridge course). The analysis is performed using the t-Test, null-hypothesis and some suitable indicators. The results tell us to reject the null-hypothesis that provides the significance improvement in the student performance in both courses (one is a prerequisite for other). But, this can be

enhanced to a large set of data so that the analysis will become more efficient and effective.

### REFERENCES

- [1] N G Das, "Statistical Methods", McGraw Hill, Baltimore, India, 9780070083271, 2017
- [2] Miss. Sharayu N, "Survey on Evaluation of Student's Performance in Educational Data Mining", Proceedings of the 2nd International Conference ICICCT, 978-1-5386-1974-2, 2018
- [3] R.Jindal, M.D Borah, "A Survey on Educational Data Mining and Research trends", International Journal of Database Management System (IJDS), 5(3), pp-53-73, 2013
- [4] Riayah Novita, Mira Kania Sabariah, Veronikha Effendy, "Identifying Factors That Influence Student Failure Rate using Exhaustive CHAID (Chi-Square Automatic Interaction Detection)" IEEE International Conference on Information and Communication Technology (ICoICT), pp-482-487, 2015,
- [5] Bhardwaj, B.K. and Pal, S, "Data Mining: A prediction for performance improvement using classification", (IJSIS) International Journal of Computer Science and Information Security, Vol. 9, No. 4, April 2011.
- [6] Ahmed, A.B.E.D. and Elaraby, "Data Mining: A prediction for Student's Performance Using Classification Method", World Journal of Computer Application and Technology, I.S.2(2), pp.43-47, 2014
- [7] Oyerinde O. D, Chia P. A "Predicting Students' Academic Performances – A Learning Analytics Approach using Multiple Linear Regression", Volume 157– No 4, 0975 – 8887, January 2017 .
- [8] Ayers E., Junker B.W, "Do skills combine additively to predict task difficulty in eighth grade mathematics?", In AAAI Workshop on Educational Data Mining: Menlo Park, 14-20, 2006.
- [9] Desmarais, M.C., Gagnon, M., Meshkinfram, "P. Bayesian Student Models Based on Item to Item Knowledge Structures", In Conference on Technology Enhanced Learning, Crete, Greece, 1-10, 2006.

### Authors Profile

Ms. Geetha N received the B.E degree in Information Science and Engineering and M.Tech degree in Computer Network and Engineering from VTU, Belgaum in 2012 and 2014 respectively, and currently working toward Ph.D degree from VTU, Belgaum. Currently working as Assistant Professor in Department of Computer Science and Engineering, CMR University, SOET and total teaching experience is 5 Years and her areas of interest are Data Mining, Machine Learning, Artificial Intelligence, IoT, Python, Information and network Security, Cryptography. She has published papers in 2 National Conference and 6 International journals. Her major is in Student Data Mining.



Dr. Piyush Kumar Pareek is a Ph.D (2016) in Computer Science Engineering from Jain University, Bengaluru, M. Tech (2012) from Dayanand Sagar College of Engineering, Bengaluru & B.E (2010) from Basaveshwar Engineering College, Bagalkote, Author has over 8+ Years of Academic Experience, He is author or co-author of over



53 international journal papers or conference proceedings, He has FourBooks published in Lambert Publications, Germany. He is BOS Member – Alumni – ISE - of Basaveshwar Engineering College, Bagalkote, He is PhD Alumni Convener of Jain University, Bengaluru. He is Recognized Research Supervisor in Computer Science & Engineering (VTU0817401), Visvesvaraya Technological University, Balagavi. He is currently serving in department of Computer Science & Engineering at East West Group of Institutions, Bengaluru. His interested areas of Research are Software Engineering, Computer Networks & Machine Learning. He is Editorial Board Member & Reviewer of International Journal of Advanced Research in Computer Science & Technology (IJARCST), Brittany, FRANCE. Reviewer of International journal of Engineering Research and Technology ERSRA Publications (IJERT), INDIA, International journal of Computer Science Research and Technology (IJSRT), INDIA, Inter Science Research Network (IRNET), Bhubaneswar, INDIA, International Journal of Computer Science and Informatics (IJCSI), Bhubaneswar, INDIA, International Journal of Engineering, Sciences and Management (IJESM), U.P, INDIA., International Journal of Scientific & Technology Research (IJSTR), Paris, FRANCE, Journal of Theoretical and Applied Computer Science (JTACS), Szczecin, Poland, International Journal of Advanced Research in Computer Science & Technology (IJARCST), INDIA, Journal of Harmonized Research in Engineering, INDIA., IASTER's International Research Journals, INDIA, IJMAS Journals, Bhopal, INDIA, RR Journals, INDIA, Reviewer of IRD INDIA & STEM Journals, INDIA.

Mr. Suhas G received his B.E degree and M.Tech degree in from VTU, Belgaum, and also pursuing his Ph.D degree from VTU, Belgaum. Currently working as Assistant Professor at PES university in the Department of Computer Science and Engineering and his areas of interests are computer network security, big data and published one paper in international conference.



Ms. Sandhya Soman received Master's Degree in Computer Applications from Guru Gobind Singh Indraprastha University, New Delhi and is currently pursuing her PhD in Computer Science. She is currently working as an Assistant Professor in Kristu Jayanti College, Bengaluru. Her research areas include Agile and lean Software Development.

