IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

Student Exam Result Prediction and Analysis

Sathyanarayana K B[1], Pradeep H K[2], Girish Mantha[3]

Assistant Professor, ISE, JNNCE, Shivamogga, India[1]

Associate Professor, ISE, JNNCE, Shivamogga, India[2]

Assistant Professor, ISE, JNNCE, Shivamogga, India[3]

Abstract: Result varies from student to student, institution to institution, year over year, due to various reasons. This system is developed for prediction & analysis of the upcoming university results by considering certain dependent parameters. This project is based on data mining and analysis of the extracted data from the result sheets. The project outputs a prediction value which is near to the expected value with variance.

Keywords: Regression, Multiple Linear Regression**(**MLR), Data Mining, Hypothesis, Exam Result Analysis.

I. INTRODUCTION

Data mining is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems. Data mining is an interdisciplinary subfield of computer science and statistics with an overall goal to extract information from a data set and transform the information into a comprehensible structure for further use.

Analytics is the discovery, interpretation, and communication of meaningful patterns in data; and the process of applying those patterns towards effective decision making. In other words, analytics can be understood as the connective tissue between data and effective decision making, within an organization. Especially valuable in areas rich with recorded information, analytics relies on the simultaneous application of statistics, computer programming and operations research to quantify performance.

Students' result prediction & analysis is a web-based application designed and engineered for colleges that need to manage results across that need to track, manage, report and predict results. At a time we can see all the years result in a single sheet via graphical representation.

II. PROBLEM SPECIFICATION

A. PROBLEM STATEMENT:

The existing systems did not predict the upcoming over all academic results of institution which did not help institution to overcome their drawbacks and improve the quality of education that they provide.

B. OBJECTIVES:

Following are the objectives of the study:

1. To extract the data from result sheets and to analyse the same considering some of the parameters that are considered to be affecting the results of a student.
2. To provide a graphical representation of pass/fail, exam difficulty level feedback by student in percentages.
3. To provide the predicted result that the institution can get in the upcoming year.

C. SCOPE OF THE PROJECT:

- The predicted results are very near to the real values: When cross checked with the results of the previous years , the project is observed to provide the prediction value which is very near to the result obtained in the very next year.
- The project is able to predict the results for multiple subjects: The project can be helpful in predicting the results when multiple number of subjects are under consideration.

III. DESIGN AND ANALYSIS

A. EXISTING SYTEM:

In the existing systems, data was mined and was stored. The data was never predicted. No prediction of future results was carried out.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

PRACTICES:

- Storing and retrieving the unorganized data.
- Segregating the data based on a marks obtained in external and internal.
- Further segregating the data subject wise.
- Providing the result based on different attributes considered(standard pattern- internal marks, external marks, total marks, student feedback)

B. PROPOSED SYTEM

In the proposed system application organize the data, display all previous year's result in the form of a graph. It analyzes the underlying pattern and predicts the result.

MODULE DISCRIPTION :

- Admin:

Admin can add and delete students' details, is responsible to confirm the authenticated faculty registration. Without the confirmation from Admin, faculty registration cannot be done.

Functionality: Admin can add student details and authenticate faculty registration.

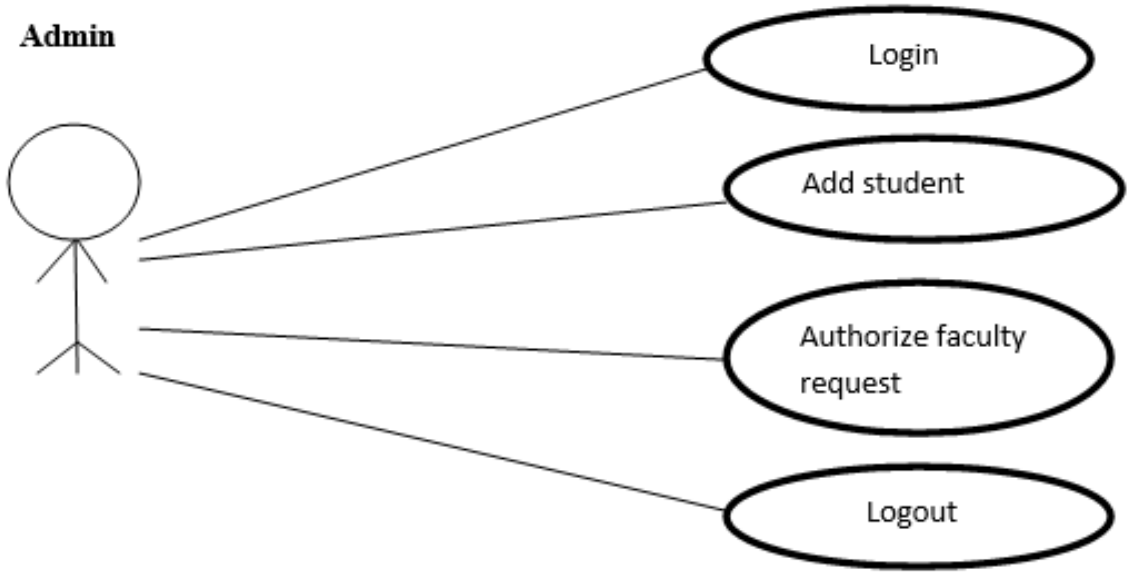

Fig 1: Use case diagram of admin

Student:

Student must be an authenticated user of the institution to access this application. The student can view all the internal assessment marks and it's aggregate. The student can give the feedback on difficulty level of the question paper.

Functionality: Student have an option to view their internal assessment marks. For result analysis and prediction students can give their feedback on difficulty level of the question paper.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

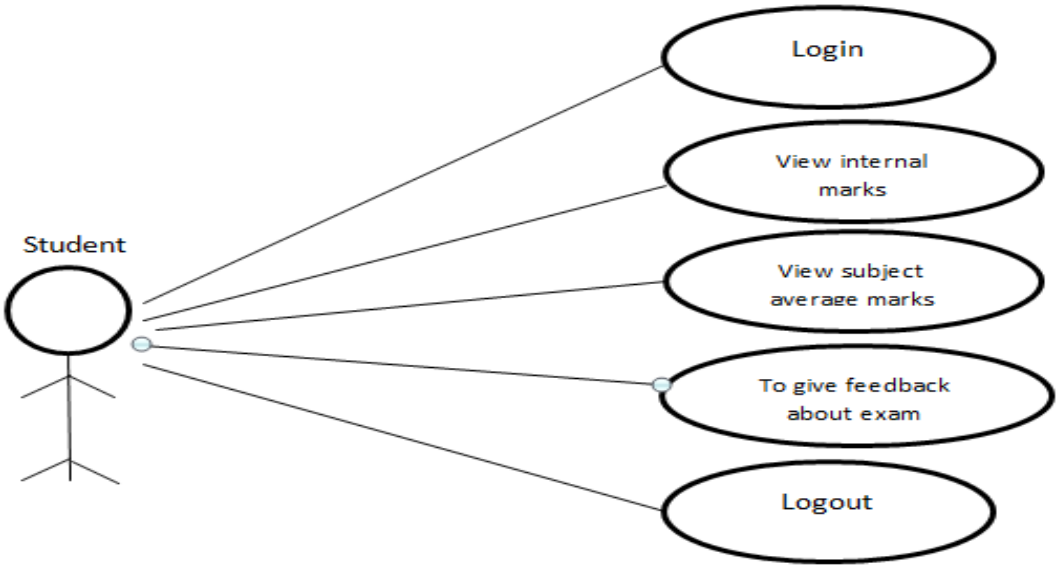

Fig 2: Use case diagram of student

Faculty:
Faculty must also be authorized users of the college. Faculty can view the overall class result, statistics report and students feedback. They can also view the individual student prediction.

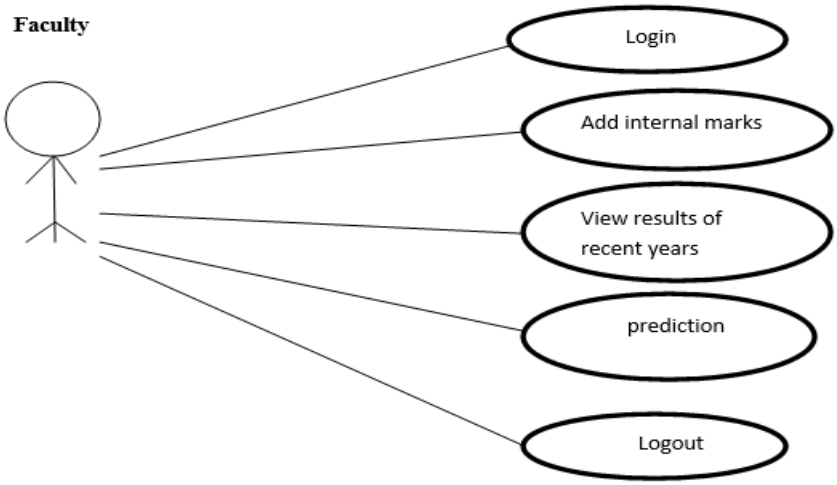

Fig 3: Use case Diagram of Faculty

Functionality: Updates the IA marks of the student. View the feedback provided by the students . View the results of the students enables the faculty to view the results of previous years in a graphical format. Views the prediction of the result.

IV. MATHEMATICAL MODEL

A. Parameter hypothesis

1. Parameter 1: Internal marks.

By assuming that, the students who score 18 out of 25 in internal test will utmost pass in the external examination. From this assumption, we calculate the pass percentage from internals marks from previous years' dataset for every year which happened to be precise with the original pass percentage.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

2. **Parameter 2: Students' Feedback on examination difficulty level.**

From previous years' dataset of marks, If marks scored in Internal is greater that 15(out of 25) & in external is less that 40(out of 100) , then we marked it as difficult. If marks scored in Internal is greater that 15(out of 25) & in external is more than 41 and less than 65(out of 100) , then we marked it as medium. If marks scored in Internal is greater that 15(out of 25) & in external is more than 66(out of 100) , then we marked it as Easy.

For the upcoming year result prediction, Students are given privilege to give feedback on examination difficulty level and we assume that every student gives the feedback. Having considered the three options of question paper level in the student's feedback system. Namely, 1.Easy 2.Medium 3.Difficult . For the mathematical prediction calculation, Assuming each of them as, Easy →3, Medium→2, Difficult→1, and taking mean of it, we get a value varying from 1 to 3 for every year. Subjecting the values calculated from both parameters to prediction Algorithm

B. Prediction Algorithm

Linear regression compares the response of a dependent variable given a change in some explanatory variable. However, it is rare that a dependent variable is explained by only one variable. In this case, an analyst uses multiple regression, which attempts to explain a dependent variable using more than one independent variable. Multiple regressions can be linear and nonlinear.Multiple regressions are based on the assumption that there is a linear relationship between both the dependent and independent variables. It also assumes no major correlation between the independent variables.

Multiple linear regression

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MLR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable. In essence, multiple regression is the extension of ordinary least-squares (OLS) regression that involves more than one explanatory variable.

The Formula for Multiple Linear Regression Is

$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots\ldots + \beta_p x_{ip} + €$ ---------- Equation 1

Where, for i = n observations:

Y_i = dependent variable

x_i = explanatory variables

β_0 = y- intercept (constant term)

β_p = slope coefficients for each explanatory variable

€ = the model's error term (also known as the residuals)

A simple linear regression is a function that allows an analyst or statistician to make predictions about one variable based on the information that is known about another variable. Linear regression can only be used when one has two continuous variables—an independent variable and a dependent variable. The independent variable is the parameter that is used to calculate the dependent variable or outcome. A multiple regression model extends to several explanatory variables.The multiple regression model is based on the following assumptions:There is a linear relationship between the dependent variables and the independent variables.The independent variables are not too highly correlated with each other.

Y_i observations are selected independently and randomly from the population.Residuals should be normally distributed with a mean of 0 and variance σ.The coefficient of determination (R-squared) is a statistical metric that is used to measure how much of the variation in outcome can be explained by the variation in the independent variables. R2 always increases as more predictors are added to the MLR model even though the predictors may not be related to the outcome variable.

R2 by itself can't thus be used to identify which predictors should be included in a model and which should be excluded. R2 can only be between 0 and 1, where 0 indicates that the outcome cannot be predicted by any of the independent variables and 1 indicates that the outcome can be predicted without error from the independent variables.When interpreting the results of a multiple regression, beta coefficients are valid while holding all other variables constant ("all else equal"). The output from a multiple regression can be displayed horizontally as an equation, or vertically in table form.

V. RESULTS AND ANALYSIS

A. RESULTS

This section contains snapshots of the system designed.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

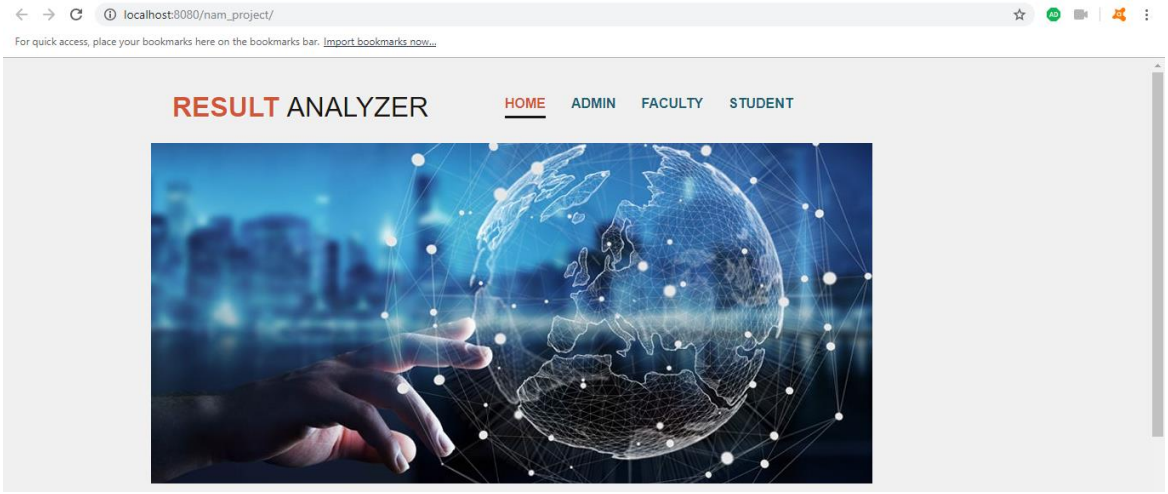

Fig 4 Home Page

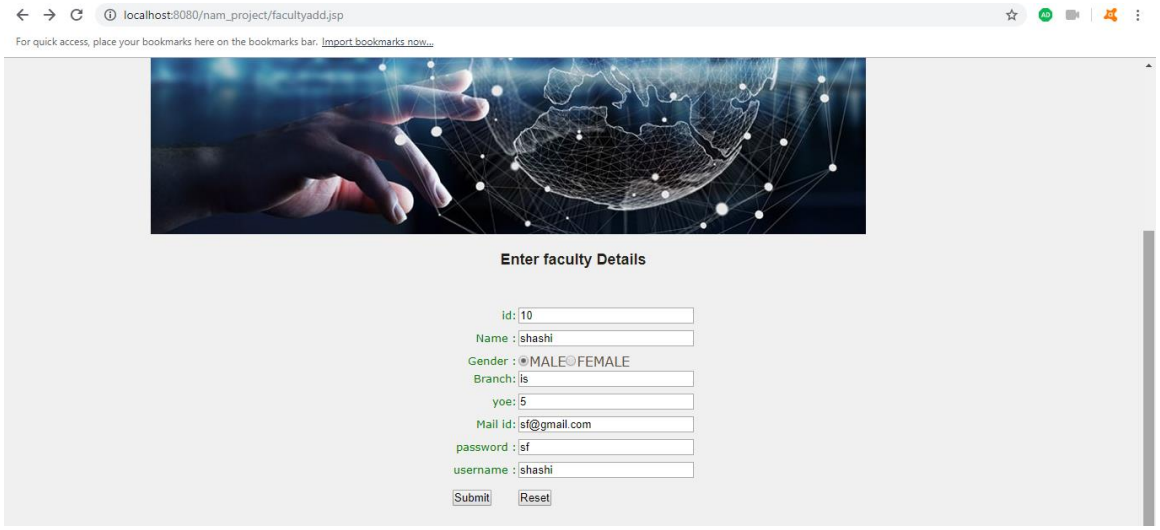

Fig 5 Faculty registration

Fig 4 Displays the landing page. The choice will be given for the user to choose the user role as Admin, Faculty or Student. Fig 5 denotes the registration of new faculty members.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

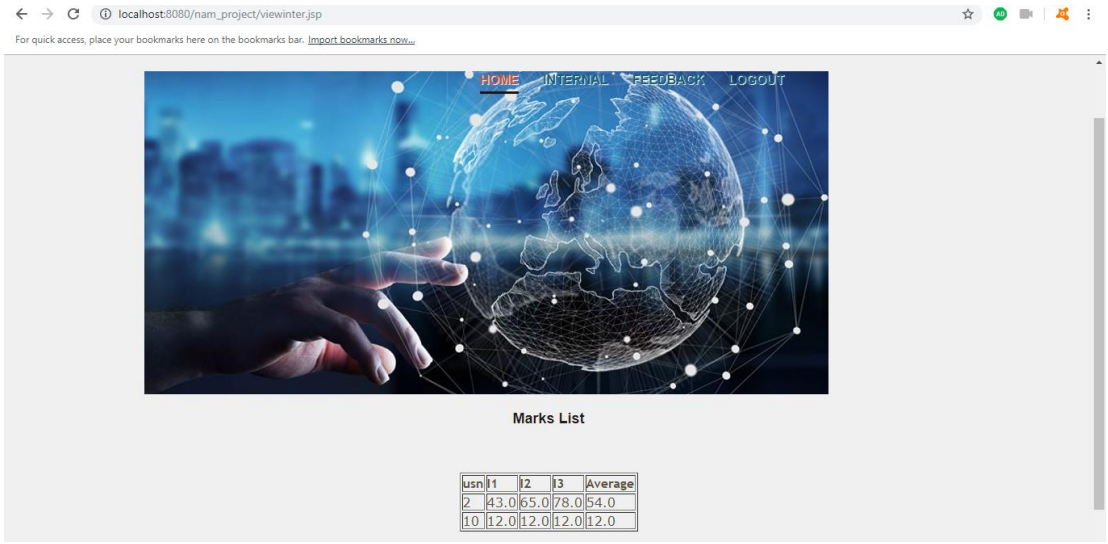

Fig 6 Student marks displayed

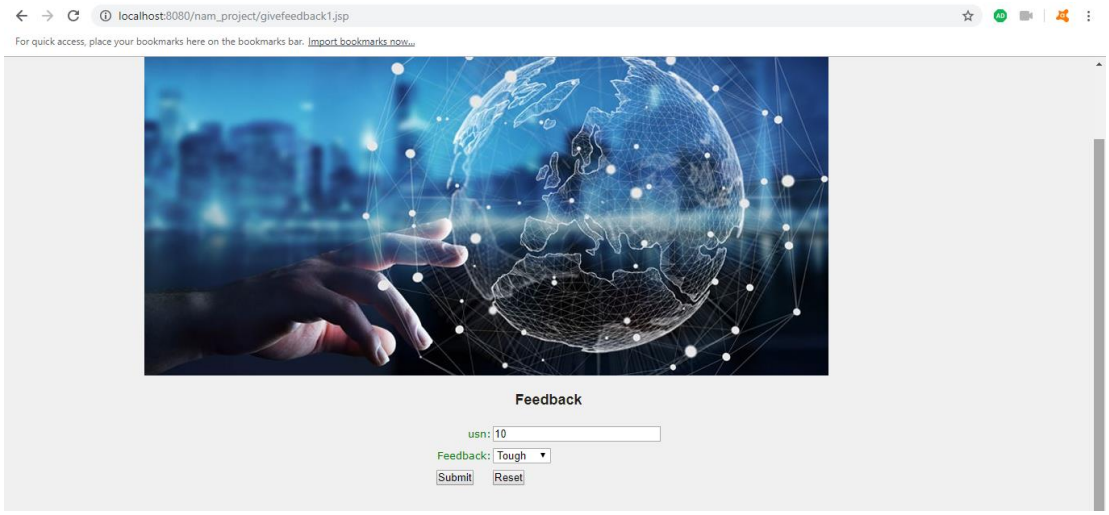

Fig 7 Student giving feedback

Fig 8 Feedback graph displayed to faculty

Fig 6 denotes the display of the marks secured by student in a course. Fig 7 student providing the difficulty level feedback about the question paper. Fig 8 displays the feedback graph with the level of difficulty about the question paper in various years.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

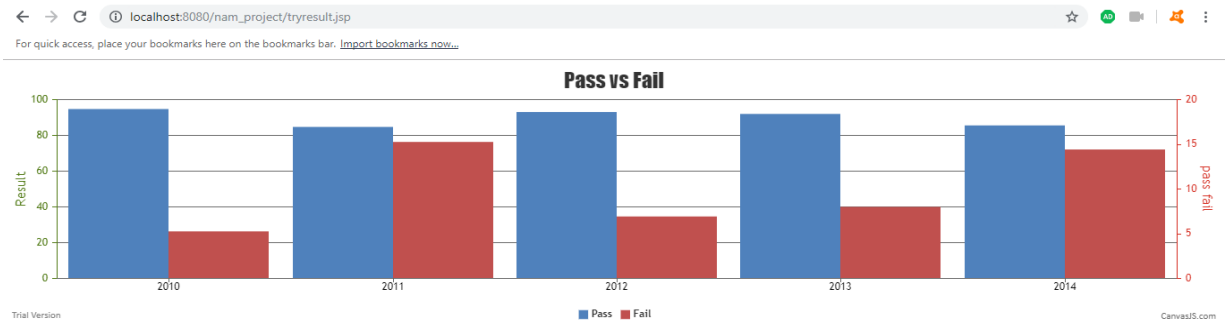

Fig 9 Pass fail percentage graph

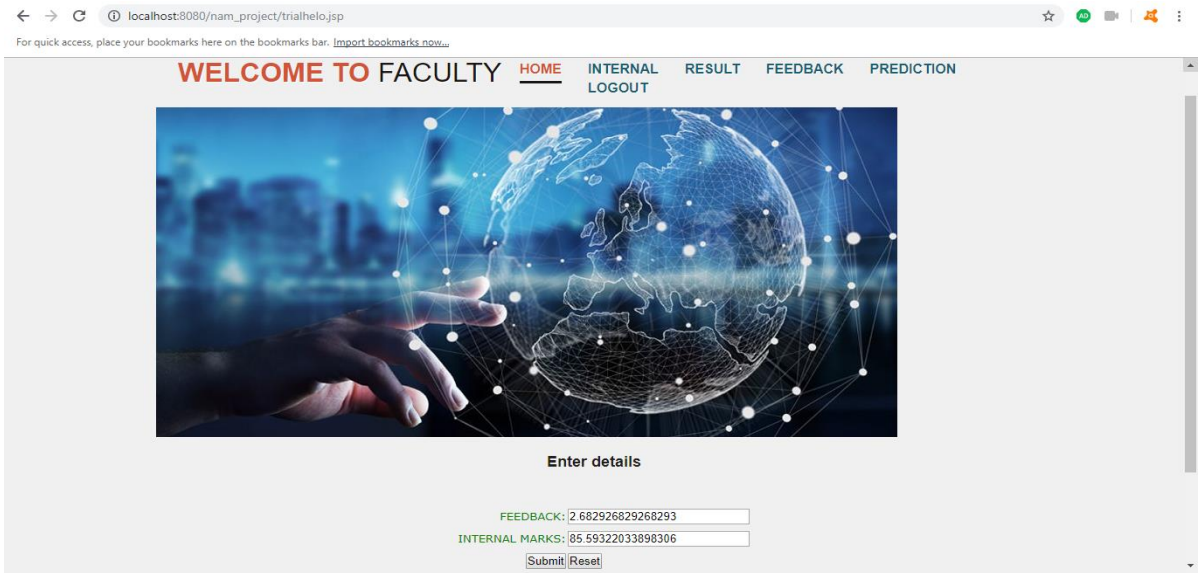

Fig 10 Data from parameters given to prediction

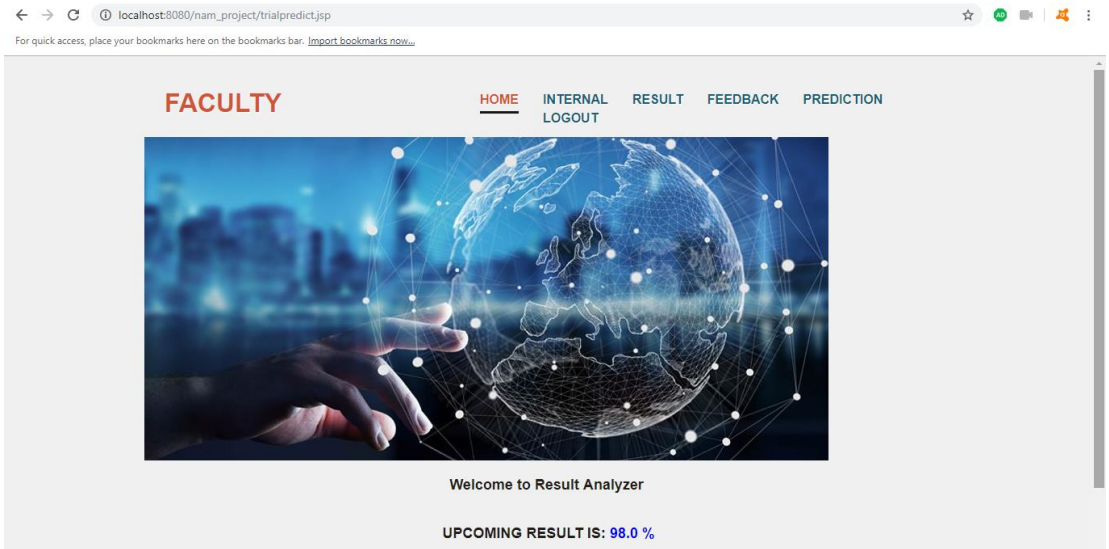

Fig 11 Prediction output

Fig 9 displays the graph of pass over fail percentage over the years. Fig 10 provides an interface to enter the feedback mean value and average marks scored by the students in the current year, which predicts the result for the current year, before the result announcement as shown in Fig 11.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11153

VI. CONCLUSION

The proposed system mines the data from database and analyzes it along with giving out the predicted results for the upcoming year. The predicted results meet the actual results as far the project is verified with the real results. There exists very little variance from the expected result , as any prediction system cannot be accurate. This can be highly beneficial for the institutions to work on their drawbacks and to get better results as the years pass by.

REFERENCES

[1]. Result management system, Ashwin mehta,Dept.of computer science,Volume 05 Issue 04,IRJET Journal,Apr 2018
[2]. Progress report and result analysis,Miss Ghadge Vrushali S, Volume 03 issue 02,ISO 9001:2008 Certified journal,IRJET.
[3]. Result analysis from PDF file, Akash B. Nikalje
[4]. Result analysis of choice based credit system proposed Miss. Koli Sheethal S,IEEE Journal for Data Mining,19(4),pp.1193-1208
[5]. Result management system,Dipali Meher, Dept. of Computer Science,Modern college of science,Pune,Maharastra,India.
[6]. Student Result analysis and report generator (general paper).

BIOGRAPHY

Sathyanarayana K B, Assistant Professor from JNN College of Engineering Shivamogga, Karnataka. Completed M.Tech in Networking and Internet Engineering from VTU Belagavi and BE in Computer Science and Engineering from VIT, Bengaluru, Karnataka. Research interests include Machine Learning, Image Processing, IoT. Guided many web application.

Pradeep H K, Associate Professor from JNN College of Engineering Shivamogga, Karnataka. Completed M.Tech in Computer Science from VTU Belagavi and BE in Electronics and Communication Engineering from SJCE, Mysore. Research interests include IoT, Security of IoT, Process automtion.

Girish Mantha, Assistant Professor from JNN College of Engineering Shivamogga, Karnataka. Completed M.Tech in Computer Science from VTU Belagavi and BE in Information Science from BVBCET Hubli, Karnataka. Research interests include IoT, Security of IoT applications using Block chain, Mobile Social Networks. Undertaken research project Funded by NewGen IEDC, Dept of DST, GOI, Applied and received funding from DST and GOI of about 20 Lakhs for various events. Awarded as Best Achiever award for Drafting 'JNNCE Green Policy' on the occasion of JNNCE Sneha Sammilana – 2019. Having interest in trekking, Trekked many Himalayan circuits and major south Indian peaks.