CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Predicting Students Performance is an essential Part. To make college affordable, it is thus crucial to ensure that many more students graduate on time through early interventions on students whose performance will be unlikely to meet the graduation criteria of the degree program on time. A critical step towards Effective Intervention is to build a system that can continuously keep track of students' academic Performance and accurately predict their future performance. Such as predicting one's subject marks based on his previous marks. Although predicting student performance has been extensively studied in the context of solving problems. With the wide usage of computers and internet, there has recently been a huge increase in publicly available data that can be analyzed.

1.2 RELEVANCE

The Evaluation is important to maintain student Performance based on Final exam score and Behavioral pattern . Universities today are operating in very complex and highly competitive environment. The main challenge for universities is to measure their performance and to build the strategy for further development and future action. Predicting students' performance helps students to know their weaker side; Teachers get to know overall performance of a Class.

1.3 PROJECT UNDERTAKEN

An algorithm for predicting marks of current semester based on previous semester record, previous 2 years record for same time and behavioral analysis is proposed. As per performance of student in written exam of all previous semesters, gradient of all semesters is calculated. This is then saved in database for further evaluation. Some fixed no. of questions is solved by student for behavioral analysis of student. Average of this questions leads to calculation of impact factor. The impact factor based marks prediction algorithm is use for calculating impact from average values of obtained marks. Previous 2 years data of students studying in same class is obtained from database. An exponential moving average and line regression is calculated and fed to marks prediction. Gradient, impact factor, moving average and regression together predicts a result of marks for current semester. Moving averages smooth the data to form a trend following indicator. They do not predict direction, but rather define the current direction. The exponential moving average is a type of moving average that gives more weight to recent values in an attempt to make it more responsive to new information.

1.4 ORGANIZATION OF REPORT

Our project “Students Performance Prediction Using Machine Learning” is thoroughly explained in all the chapters in this report. Planning and organization of this subject has been done with curiosity and as per the given deadline. So this project gives the entire overview of this subject.

The report is divided into various chapters to understand each aspect of the subject technically and separately.

Chapter1: Gives brief introduction of this project. This consists of background, relevance of project to wellbeing of human being, brief information on project undertaken and organization of report of this project.

Chapter 2: Gives brief review of the related Literature and present scenario of proposed system.

Chapter3: Describes Specification of proposed system which includes hardware and software specifications, problem definition, and hardware and software requirement.

Chapter4: Describes the design of proposed system, DFD and UMLs of same.

Chapter5: Describes the implementation of proposed system, this includes algorithm and installation.

Chapter6: Describes results and evaluation.

Chapter 7: Describes conclusion and future work of proposed system

CHAPTER 2

LITERATURE SURVEY/ BACKGROUND

J. Xu and et al build up a novel AI strategy for anticipating understudy execution in degree programs that can address these key difficulties. The proposed strategy has two noteworthy highlights. Initial, a bilayered structure involving various base indicators and a course of gathering indicators is created for making expectations dependent on understudies' advancing execution states. Second, information driven methodology dependent on inactive factor models and probabilistic network factorization is proposed to find course significance, which is imperative for developing effective base indicators. Through broad recreations on an undergrad understudy dataset gathered more than three years at UCLA. [1]

The mean square blunder of expectation is determined for an exponentially weighted moving normal (e.w.m.a.), when the arrangement anticipated is a Markov arrangement, or a Markov arrangement with superimposed error. The best decision of damping consistent is given; the decision isn't basic. There is an estimation of the Markov relationship to underneath which it is difficult to anticipate, with an E.W.M.A, the neighborhood varieties of the arrangement. The mean square error of an E.W.M.A is contrasted and the base conceivable esteem, to be specific that for the best direct indicator (Wiener). An adjusted E.W.M.A is developed having a mean square error moving toward that of the Wiener indicator. This modification will be of esteem if the Markov connection parameter is negative, and perhaps at the same time when the Markov parameter is ne [2]

Diverse strategies and procedures of information mining were thought about amid the expectation of understudies' prosperity, applying the information gathered from the reviews directed amid the late spring semester at the University of Tuzla, the Faculty of Economics, scholastic year 2010-2011, among first year understudies and the information taken amid the enlistment. The achievement was assessed with the passing evaluation at the test. The effect of understudies' socio-statistic factors, accomplished outcomes from secondary school and from the selection test, and frames of mind towards contemplating which can have an effect on progress, were altogether explored. In future examinations, with distinguishing and evaluating factors related with procedure of considering, and with the example increment, it is conceivable to deliver a model which would remain as an establishment for the advancement of choice emotionally supportive network in advanced education. [3]

Paper investigates the utility of bunching in lessening error in different expectation errands. Past work has alluded to the improvement in forecast exactness credited to bunching calculations whenever used to pre-process the information. In this work we all the more profoundly explore the immediate utility of utilizing bunching to improve expectation exactness and give clarifications to why this might be so. We take a gander at various datasets, run k-implies at various scales and for each scale we train indicators. This produces k sets of expectations. These forecasts are then consolidated by a gullible group. We saw that this utilization of an indicator related to grouping improved the expectation exactness in many datasets. We trust this demonstrates the prescient utility of abusing structure in the information and the information pressure gave over by grouping. We additionally discovered that utilizing this technique enhances the expectation of even a Random Forests indicator which recommends this strategy is giving a novel and valuable wellspring of change in the forecast procedure. [4]

As of late Educational Data Mining (EDM) has risen as another field of research because of the advancement of a few measurable ways to deal with investigate information in instructive setting. One such use of EDM is early forecast of understudy results. This is fundamental in advanced education for distinguishing the "feeble" understudies with the goal that some type of remediation might be sorted out for them. In this paper a lot of properties are first characterized for a gathering of understudies studying Computer Science in some undergrad schools in Kolkata. Since the quantities of characteristics are sensibly high, highlight determination calculations are connected on the informational index to diminish the quantity of highlights. Five classes of Machine Learning Algorithm (MLA) are then connected on this informational collection and it was discovered that the best outcomes were acquired with the choice tree class of calculations. It was likewise discovered that the forecast outcomes got with this model are tantamount with other recently created models. [5]

Lately Educational Data Mining (EDM) has risen as another field of research because of the advancement of a few measurable ways to deal with investigation information in instructive setting. One such use of EDM is early forecast of understudy results. This is vital in advanced education for distinguishing the "frail" understudies with the goal that some type of remediation might be sorted out for them. Since the quantities of qualities are sensibly high, highlight determination calculations are connected on the informational collection to lessen the quantity of highlights. Five classes of Machine Learning Algorithm (MLA) are then connected on this informational index and it was discovered that the best outcomes were acquired with the choice tree class of calculations. It was additionally discovered that the forecast outcomes acquired with this model are equivalent with other recently created models. [6]

In this period of computerization, training has additionally patched up itself and isn't restricted to old address technique. The customary journey is on to discover better approaches to make it increasingly successful and effective for understudies. These days, bunches of information is gathered in instructive databases, yet it remains unutilized. So as to get required advantages from such major information, incredible assets are required. Information mining is a rising integral asset for examination and expectation. It is effectively connected in the zone of misrepresentation location, promoting, advertising, advance evaluation and forecast. Be that as it may, it is in beginning stage in the field of instruction. Significant measure of work is done toward this path, yet at the same time there are numerous immaculate territories. Also, there is no brought together methodology among these looks into. [7]

The understudy's execution forecast is a vital research theme since it can enable educators to keep understudies from dropping out before last tests of the year and recognize understudies that need extra help. The target of this investigation is to anticipate the troubles that understudies will experience in an ensuing computerized structure course session. We broke down the information logged by an Technology Enhanced learning (TEL) framework called advanced gadgets instruction and structure suite (DEEDS) utilizing AI calculations. The AI calculations incorporated a counterfeit neural systems (ANNs), bolster vector machines (SVMs), strategic relapse, Naïve bayes classifiers and choice trees. The DEEDS framework enables understudies to settle computerized configuration practices with various dimensions of trouble while logging input information. The info factors of the present examination were normal time, all out number of exercises, normal inert time, normal number of keystrokes and absolute related action for each activity amid individual sessions in the advanced structure course; the yield factors were the student(s) grades for every session. We at that point prepared AI calculations on the information from the past session and tried the calculations on the information from the up and coming session. We performed k-crease cross-approval and registered the recipient working trademark and root mean square blunder measurements to assess the models' exhibitions. The outcomes demonstrate that ANNs and SVMs accomplish higher exactness than do different calculations. ANNs and SVMs can without much of a stretch be incorporated into the TEL framework; in this manner, we would anticipate that educators should report improved understudy's execution amid the consequent session. [8]

In this paper, we tended to the evaluation challenge in the assessment framework, which is an online coaching framework that fills in as an e-learning and e-appraisal condition. We - 19 - concentrated on the appraisal capacity of the framework and assessed it by mining our log information and contrasting and state sanctioned test outcomes. Some proof was displayed that the online evaluation framework completed a superior employment of anticipating understudy learning by having the capacity to null over what amount mentoring help was required, how quick an understudy tackles an issue and what number of endeavors were expected to complete an issue. [9]

To build adequacy in conventional homeroom courses just as in Massive Open Online Courses (MOOCs), computerized frameworks supporting the educator are required. One vital issue is to consequently distinguish understudies that will do inadequately in a course sufficiently early to have the capacity to take therapeutic activities. This paper proposes a calculation that predicts the last grade of every understudy in a class. It issues an expectation for every understudy exclusively, when the normal exactness of the forecast is adequate. The calculation realizes online what is the ideal forecast and time to issue an expectation dependent on previous history of understudies' execution in a course. We infer exhibit the execution of our calculation on a dataset acquired dependent on the execution of around 700 college understudies who have taken a starting computerized flag preparing in the course of recent years. Utilizing information got from a pilot course, our philosophy recommends that it is powerful to perform ahead of schedule in-class evaluations, for example, tests, which result in convenient execution forecast for every understudy, along these lines empowering auspicious intercessions by the teacher (at the understudy or class level) when vital. [10]

CHAPTER 3

SPECIFICATIONS

3.1 PROBLEM DEFINITION

Current systems use only academic data to analyze the result. The academic results are not only based on academic performance but individual behavior,that also affects the performance. Hence for better prediction behavior needs to be considered. The challenges faced by existing systems are for the accurate prediction the dependency between subjects must be identified, students evolving progress needs to be incorporated into the prediction. Hence to overcome all the drawbacks of current system, new system is being developed and is defined as 'Student Performance Prediction based on academic performance history and analyzing behavioral pattern to improve results.'

3.2 SYSTEM SPECIFICATION

Get survey of students for their daily routine, study timings and planning by using a questionnaire. Get the academic performance of students. Analyze the behavior and academic performance of the student and suggest improvements for his/her studies.

USERS: - 1. Students

2. Teachers

Analyze previous performance and forecast result, Find interdependency of subjects

3.2.1 Hardware Specifications

- Processor - Intel i3 core
- Speed - 1.1 GHz
- RAM - 2GB
- Hard Disk - 50 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

3.2.2 Software Specifications

- **Languages :-**

1. Python

- **Operating System :**

1. Windows
2. Linux

- **Database:**

1. MySQL

- **Software:**

1. Python IDLE

3.3 SOFTWARE REQUIREMENT

3.3.1 PYTHON

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

- **Python is Interpreted** − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.
- **Python is Interactive** − You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.
- **Python is Object-Oriented** − Python supports Object-Oriented style or technique of programming that encapsulates code within objects.
- **Python is a Beginner's Language** − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

Python Features

Python's features include −

- **Easy-to-learn** − Python has few keywords, simple structure, and a clearly defined syntax. This allows the student to pick up the language quickly.

- **Easy-to-read** − Python code is more clearly defined and visible to the eyes.
- **Easy-to-maintain** − Python's source code is fairly easy-to-maintain.
- **A broad standard library** − Python's bulk of the library is very portable and cross-platform compatible on UNIX, Windows, and Macintosh.
- **Interactive Mode** − Python has support for an interactive mode which allows interactive testing and debugging of snippets of code.
- **Portable** − Python can run on a wide variety of hardware platforms and has the same interface on all platforms.
- **Extendable** − You can add low-level modules to the Python interpreter. These modules enable programmers to add to or customize their tools to be more efficient.
- **Databases** − Python provides interfaces to all major commercial databases.
- **GUI Programming** − Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as Windows MFC, Macintosh, and the X Window system of Unix.
- **Scalable** − Python provides a better structure and support for large programs than shell scripting.

Apart from the above-mentioned features, Python has a big list of good features, few are listed below −

- It supports functional and structured programming methods as well as OOP.
- It can be used as a scripting language or can be compiled to byte-code for building large applications.
- It provides very high-level dynamic data types and supports dynamic type checking.
- It supports automatic garbage collection.
- It can be easily integrated with C, C++, COM, ActiveX, CORBA, and Java.

Python is available on a wide variety of platforms including Linux and Mac OS X. Let's understand how to set up our Python environment.

3.3.2 MYSQL

MySQL is the most popular Open Source Relational SQL Database Management System. MySQL is one of the best RDBMS being used for developing various web-based software applications. MySQL is developed, marketed and supported by MySQL AB, which is a Swedish company. This tutorial will give you a quick start to MySQL and make you comfortable with MySQL programming. A database is a separate application that stores a collection of data. Each database has one or more distinct APIs for creating, accessing, managing, searching and replicating the data it holds.

Other kinds of data stores can also be used, such as files on the file system or large hash tables in memory but data fetching and writing would not be so fast and easy with those type of systems.

Nowadays, we use relational database management systems (RDBMS) to store and manage huge volume of data. This is called relational database because all the data is stored into different tables and relations are established using primary keys or other keys known as **Foreign Keys**.

MySQL is a fast, easy-to-use RDBMS being used for many small and big businesses. MySQL is developed, marketed and supported by MySQL AB, which is a Swedish company. MySQL is becoming so popular because of many good reasons −

- MySQL is released under an open-source license. So you have nothing to pay to use it.
- MySQL is a very powerful program in its own right. It handles a large subset of the functionality of the most expensive and powerful database packages.
- MySQL uses a standard form of the well-known SQL data language.

- MySQL works on many operating systems and with many languages including PHP, PERL, C, C++, JAVA, etc.
- MySQL works very quickly and works well even with large data sets.
- MySQL is very friendly to PHP, the most appreciated language for web development.
- MySQL supports large databases, up to 50 million rows or more in a table. The default file size limit for a table is 4GB, but you can increase this (if your operating system can handle it) to a theoretical limit of 8 million terabytes (TB).
- MySQL is customizable. The open-source GPL license allows programmers to modify the MySQL software to fit their own specific environments.

3.3.3 PYTHON IDEL

- IDLE is integrated development environment (IDE) for editing and running Python 2.x or Python 3 programs.
- The IDLE GUI (graphical user interface) is automatically installed with the Python interpreter. IDLE was designed specifically for use with Python.
- IDLE has a number of features to help you develop your Python programs including powerful syntax highlighting.
- IDLE is integrated development environment (IDE) for editing and running Python 2.x or Python 3 programs.
- The IDLE GUI (graphical user interface) is automatically installed with the Python interpreter. IDLE was designed specifically for use with Python.

- IDLE has a number of features to help you develop your Python programs including powerful syntax highlighting.
- IDLE can be used to execute a single statement just like Python Shell and also to create, modify and execute Python scripts. IDLE provides a fully-featured text editor to create Python scripts that includes features like syntax highlighting, autocompletion and smart indent. It also has a debugger with stepping and breakpoints features.

CHAPTER 4

DESIGN

4.1 PROPOSED SYSTEM

In this paper, an algorithm for predicting marks of current semester based on previous semester record, previous 2 years record for same time and behavioral analysis is proposed. As per performance of student in written exam of all previous semesters, gradient of all semesters is calculated. This is then saved in database for further evaluation. Some fixed no. of questions is solved by student for behavioral analysis of student.

Average of this questions leads to calculation of impact factor. The impact factor based marks prediction algorithm is use for calculating impact from average values of obtained marks. Previous 2 years data of students studying in same class is obtained from database. An exponential moving average and line regression is calculated and fed to marks prediction.

Gradient, impact factor, moving average and regression together predicts result of marks for current semester.

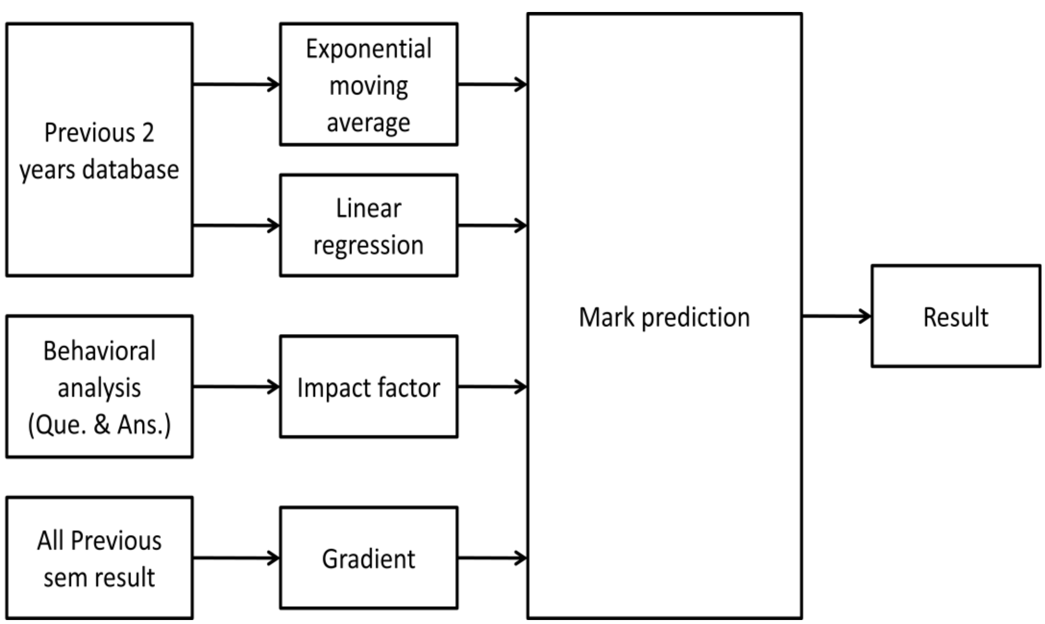

Figure 4.1 proposed algorithm

Moving averages smooth the data to form a trend following indicator. They do not predict direction, but rather define the current direction. The exponential moving average is a type of moving average that gives more weight to recent values in an attempt to make it more responsive to new information.

$$EMA = (P * \alpha) + \left(previous\ EMA * (1 - \alpha)\right) \qquad .... (1)$$

P=current price

α=smoothing factor= $\frac{2}{1+N}$

N=no. of time periods

Linear regression endeavors to demonstrate the connection between two factors by fitting a direct condition to watched information. One variable is viewed as an informative variable, and the other is viewed as a reliant variable. A linear regression line is given as

$$Y = a + bX \qquad .... (2)$$

Where X = explanatory variable

Y = dependent variable.

b= slope of the line is b

a=the intercept (the value of y when x = 0).

Gradient in python is given by numpy.gradient to get an array with the numerical derivative for every variable. The gradient is defined as

$$\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x} \qquad .. .. (3)$$

4.2 UML DIAGRAM

4.2.1 Use Case Diagram

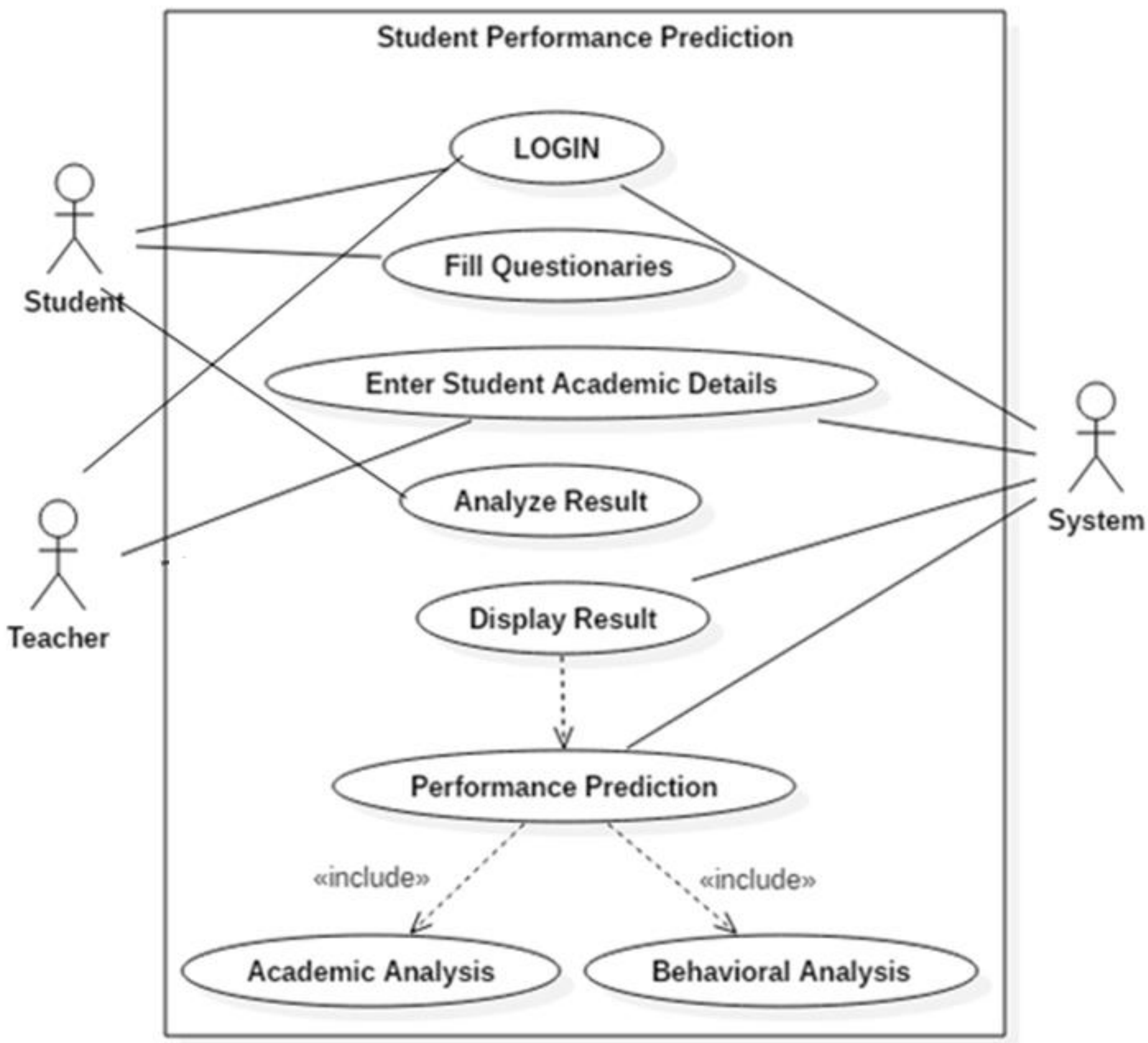

Figure 4.2.1. Use case diagram

- Student as well as Teachers will Login into the System.
- Students will have to Fill the Behavioral Based Questionnaires' .
- Teachers have to Enter Students Academic Records
- System considers both the inputs (Questionnaires' and Academic Records)
- Machine Learning Algorithm is applied on to the data and Prediction Results are displayed.

4.2.2 CLASS DIAGRAM

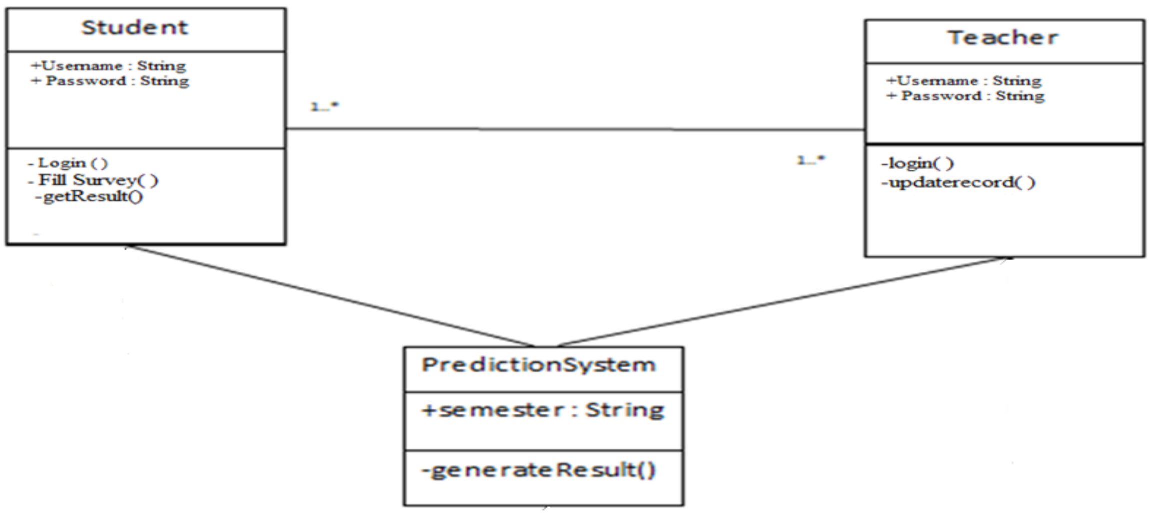

Figure 4.2.2 class diagram

4.2.3 Sequence Diagram

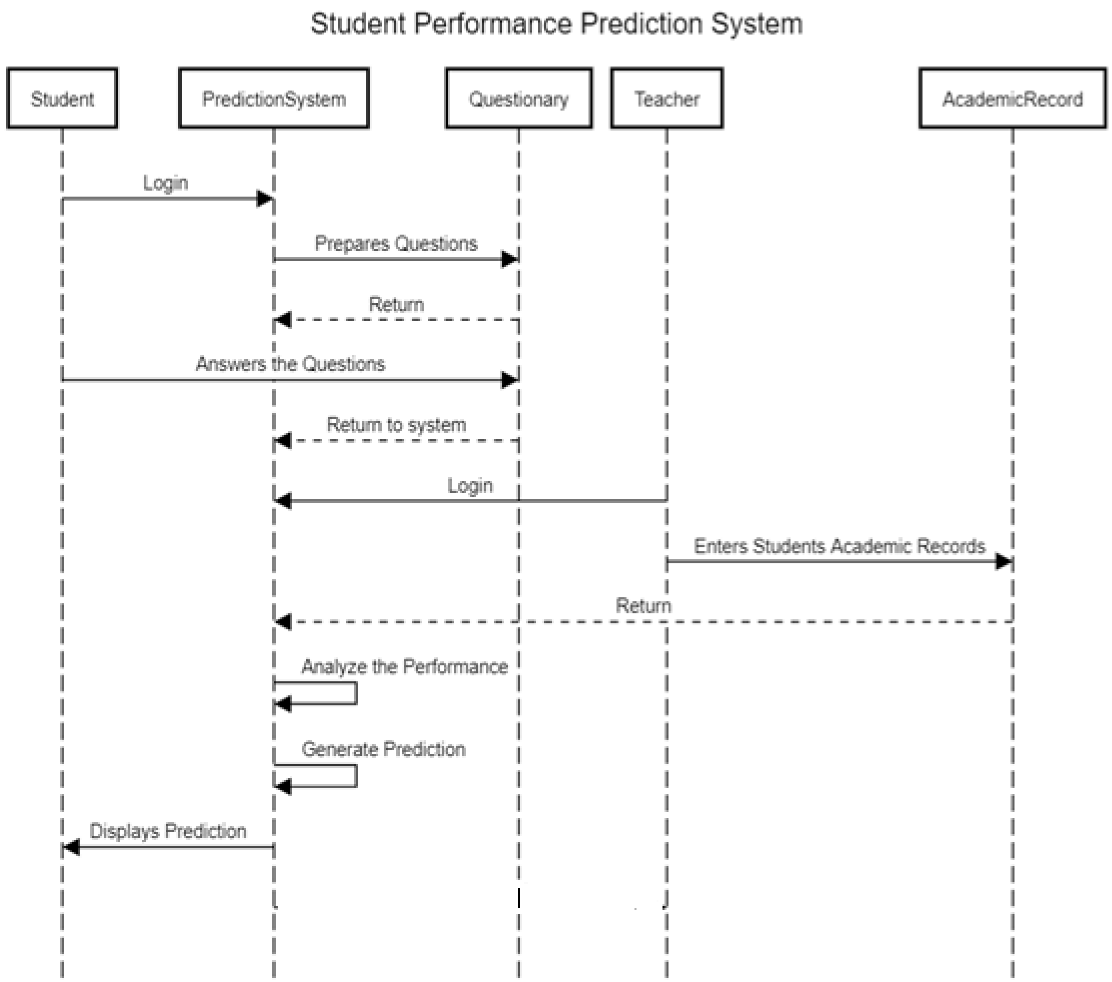

Figure4.2.3 sequence diagram

CHAPTER 5

IMPLEMENTATION

5.1 System Architecture

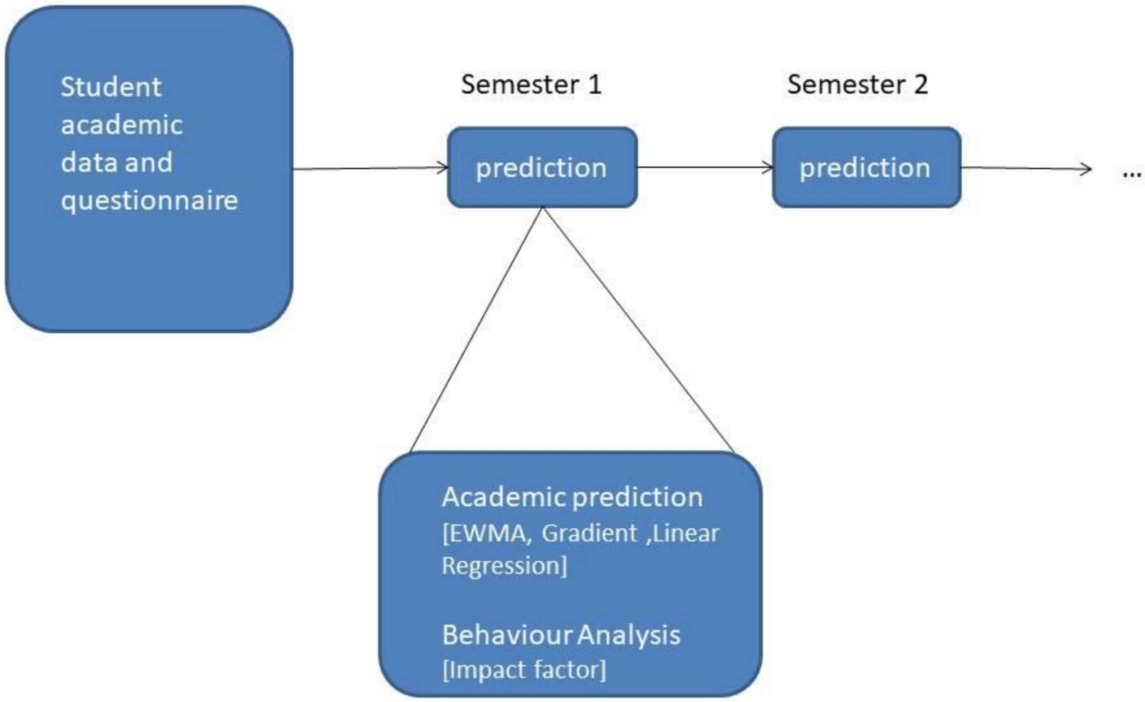

Fig. 5.1 System Architecture

The student's academic record which is semester wise marks of students is collected for teachers.Next each student can login into the student account to fill the questionnaire. This questionnaire includes questions to understand about the routine and preferences about study. Both of these records are used together to predict the next semester marks of the student. It uses the impact factor method for calculating the score of behavior. The linear regression, EWMA and gradient is used for calculating the prediction on academic records. These two results are combined together to predict students' performance.

5.1.1 EXPONENTIAL MOVING AVERAGE

Moving averages act as a technical indicator to show you how a security's price has moved, on average, over a certain period of time. Moving averages are often used to help highlight trends, spot trend reversals and provide trade signals. There are several different types of moving averages, but they all create a single smooth line that can help show you which direction a price is moving.

- **Simple Moving Average Calculation**

The simple moving average (SMA) calculates an average of the last **n** prices, where **n**represents the number of periods for which you want the average:

$$\boldsymbol{Simple\ moving\ average\ =\ (P1\ +\ P2\ +\ P3\ +\ P4 + \ldots + Pn)\ /\ n}$$

For example, a four-period SMA with prices of 1.2640, 1.2641, 1.2642, and 1.2641 gives a moving average of 1.2641 using the calculation [(1.2640 + 1.2641 + 1.2642 + 1.2641) / 4 = 1.2641].

While knowing how to calculate a simple average is a good skill to have, trading and chart platforms calculate this for you. Simply select the SMA indicator from the list of charting indicators, apply it to the chart, and adjust the number of periods you want to use.

You typically make adjustments to the indicators in the **Settings** menu section of a trading platform. On many platforms, you can locate the settings by double-clicking on the indicator itself.

The advantage of an SMA is that you know exactly what you are getting. The SMA value equals the average price for the number of periods in the SMA calculation.

Common SMA Values are 8, 20, 50, 100 and 200. For example, if using a 100-period SMA, the current value of the SMA on the chart is the average price over the last 100 periods or price bars.

Fig 5.1.1 shows a 50-period SMA, along with an exponential moving average (EMA) and a weighted moving average (WMA) on a one-minute stock chart. Due to their different calculations, the indicators appear at different price levels on the chart.

Fig 5.1.1 Chart showing moving average

- **Exponential Moving Average Calculation**

The exponential moving average **(**EMA) is a weighted average of the last **n** prices, where the weighting decreases exponentially with each previous price/period. In other words, the formula gives recent prices more weight than past prices.

$$\textbf{\textit{Exponential moving average}} = [\textbf{\textit{Close}} - \textbf{\textit{previous EMA}}] * (\mathbf{2} / \boldsymbol{n} + \mathbf{1}) + \textbf{\textit{previous EMA}}$$

For example. a four-period EMA with prices of 1.5554, 1.5555, 1.5558, and 1.5560, with the last value being the most recent, gives a current EMA value of 1.5558 using the calculation [(1.5560 - 1.5558) x (2/5) + 1.5558 = 1.55588].

As with the SMA, charting platforms do all the EMA calculations for you. Select the EMA from the indicator list on a charting platform and apply it to your chart. Go into the settings and adjust how many periods the indicator should calculate, for example, 15, 50 or 100 periods.

The EMA adapts more quickly to price changes than the SMA. For example, when a price reverses direction, the EMA will reverse direction quicker than the SMA. This takes place because the EMA formula gives more weight to recent prices, and less weight to prices that occurred in the past.

5.1.2 LINEAR REGRESSION

When we have a single input attribute (x) and we want to use linear regression, this is called simple linear regression.

If we had multiple inputs attributes (e.g. x1, x2, x3, etc.) This would be called multiple linear regressions. The procedure for linear regression is different and simpler than that for multiple linear regressions, so it is a good place to start.

With simple linear regression we want to model our data as follows:

$$y = B0 + B1 * x$$

This is a line where y is the output variable we want to predict, x is the input variable we know and B0 and B1 are coefficients that we need to estimate that move the line around. Technically, B0 is called the intercept because it determines where the line intercepts the y-axis. In machine learning we can call this the bias, because it is added to offset all predictions that we make. The B1 term is called the slope because it defines the slope of the line or how x translates into a y value before we add our bias.

The goal is to find the best estimates for the coefficients to minimize the errors in predicting y from x.

Simple regression is great, because rather than having to search for values by trial and error or calculate them analytically using more advanced linear algebra, we can estimate them directly from our data.

We can start off by estimating the value for B1 as:

$$B1 = sum((xi - mean(x)) * (yi - mean(y))) / sum((xi - mean(x))^2)$$

Where mean () is the average value for the variable in our dataset. The xi and yi refer to the fact that we need to repeat these calculations across all values in our dataset and i refers to the i'th value of x or y.

We can calculate B0 using B1 and some statistics from our dataset, as follows:

$$B0 = mean(y) - B1 * mean(x)$$

5.1.3 GRADIENT

The **gradient** of a function f, denoted as ∇f, is the collection of all its partial derivatives into a vector.

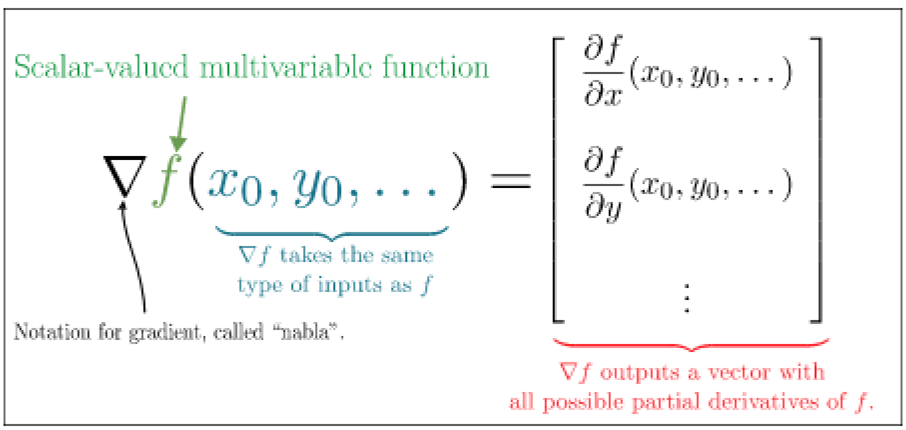

Figure.5.1.3 Gradient

Example: Gradient of $f(x, y) = x^2 - xy$, is as follows:

Notice, ∇f, f is a **vector-valued function**, specifically one with a two-dimensional input and a two-dimensional output. This means it can be nicely visualized with a vector field. That vector field lives in the input space of fff, which is the xyxyx, y-plane.

This vector field is often called the **gradient field** of f.

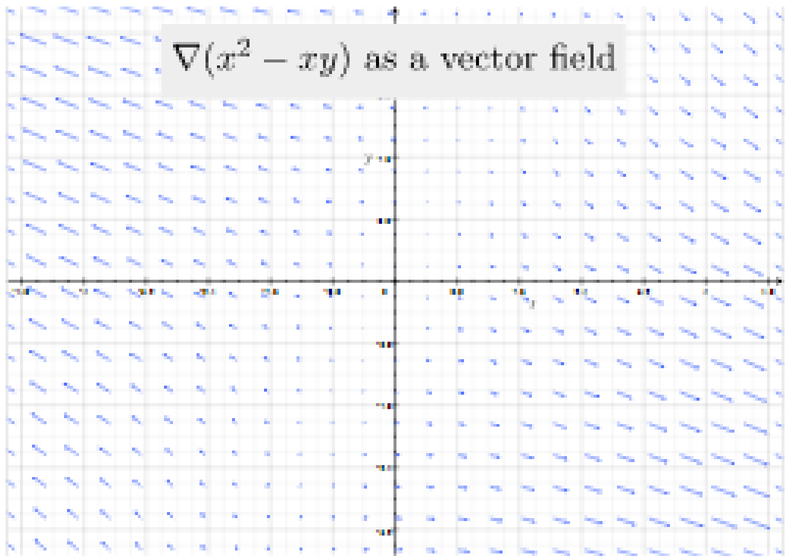

Figure 5.1.4 Gradient field

5.1.4 IMPACT FACTOR

To calculate Impact factor, the very first thing needed is to give each option of the questions weight based on the nature of the question. There are total 10 questions and each question carries 4 options having weights in the range of 1-4. Depending on the type of questions the weight is assigned to the options where 4 is the maximum weight and 1 minimum.

If user doesn't give answer of any question, the weight will be considered 0 for that particular question and accordingly Impact factor will be calculated. After receiving all the answers of the questions, addition is done and it is divided by 10 to normalize. This is done because the impact factor range is 0-9,so to fit our addition of weights, those are divided by 10.

As we are having 2 parts in our project: 1.marks input, 2.Questionaries. Here impact factor is used to give the range of predicted marks based on questionnaires. So normalized value is used to calculate gradient factor which by then gives the range of marks which are to be predicted.

The impact factor that is used to calculate the average of weights obtained from questions and gradient calculated from marks are collaborated and added together to get Deviation Factor.

Deviation factor gives the final output which is integration of above two methods. The result is shown in range.

SCREENSHOTS

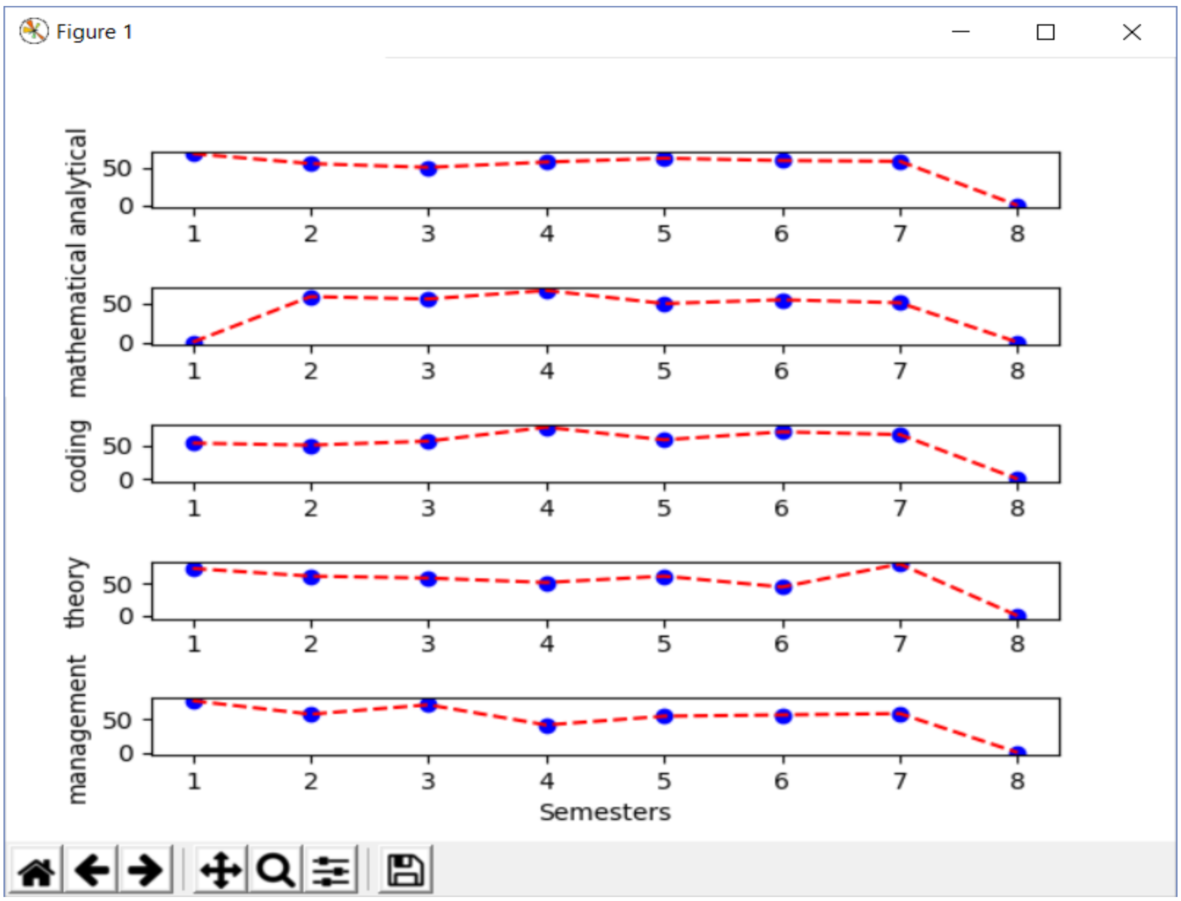

Figure 5.1.5 Student graph

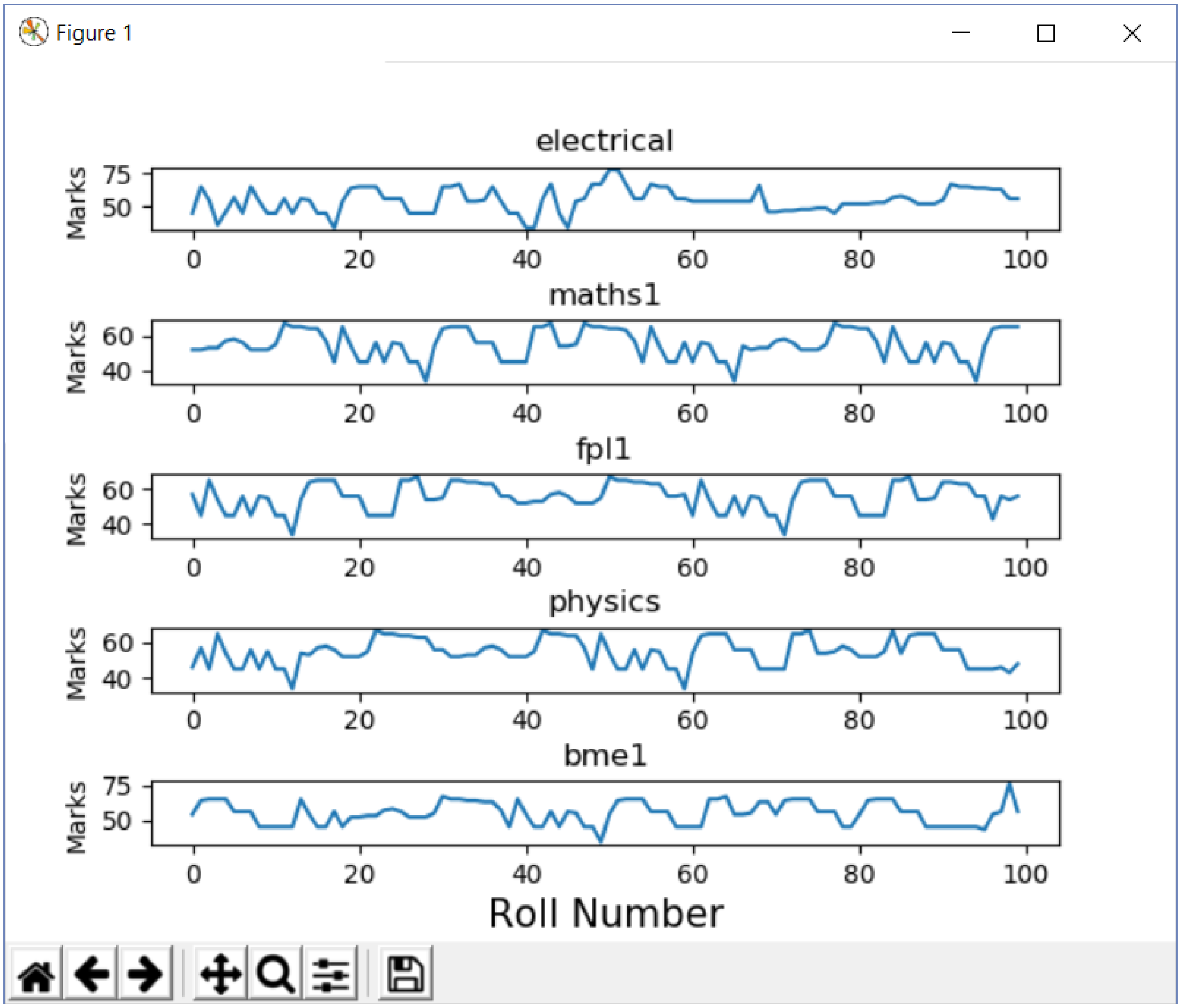

Figure 5.1.6 Subject graph

CHAPTER 6

RESULTS AND EVALUATION

EXPERIMENTAL SETUP

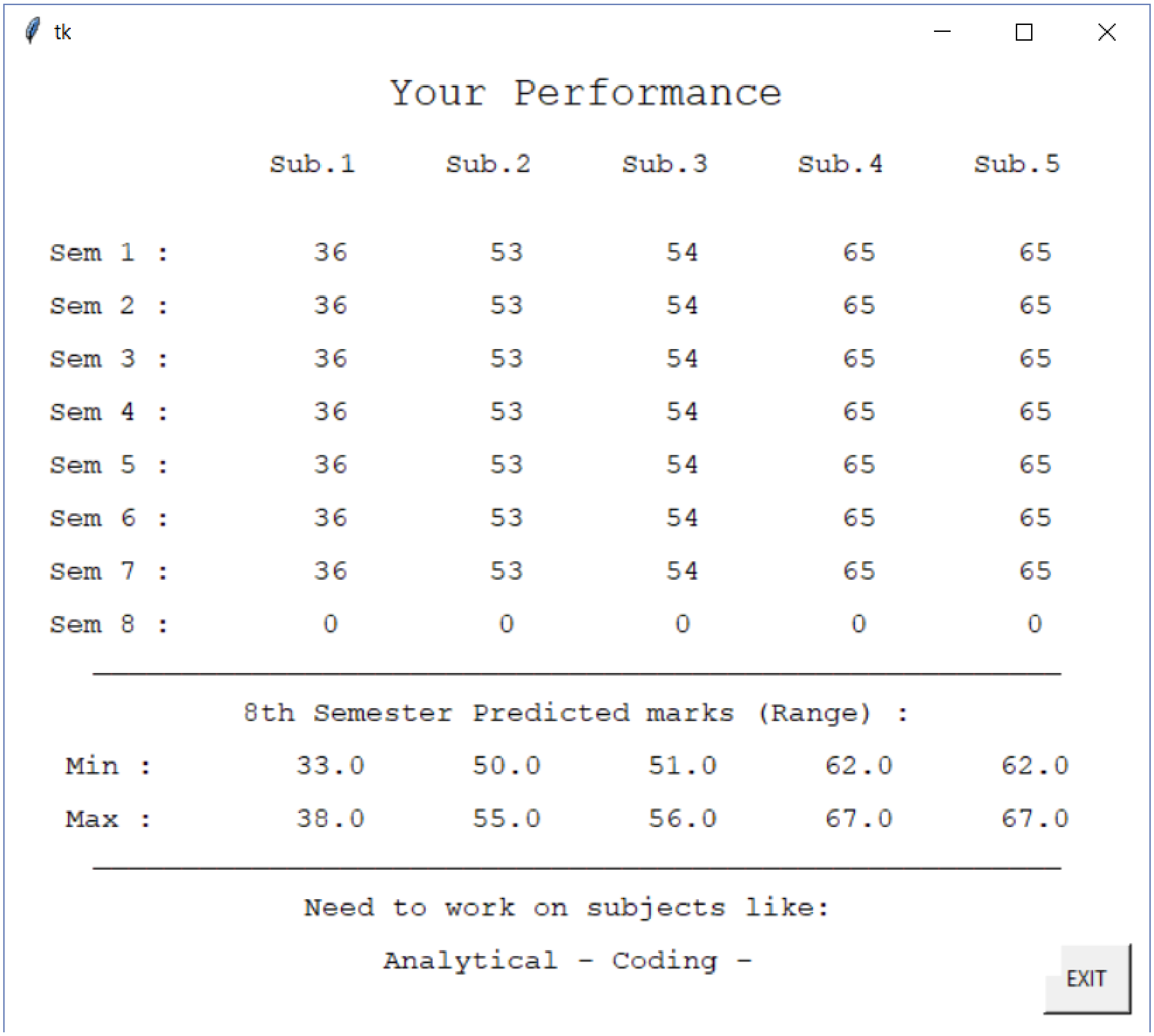

Figure 6.1 Final output

.

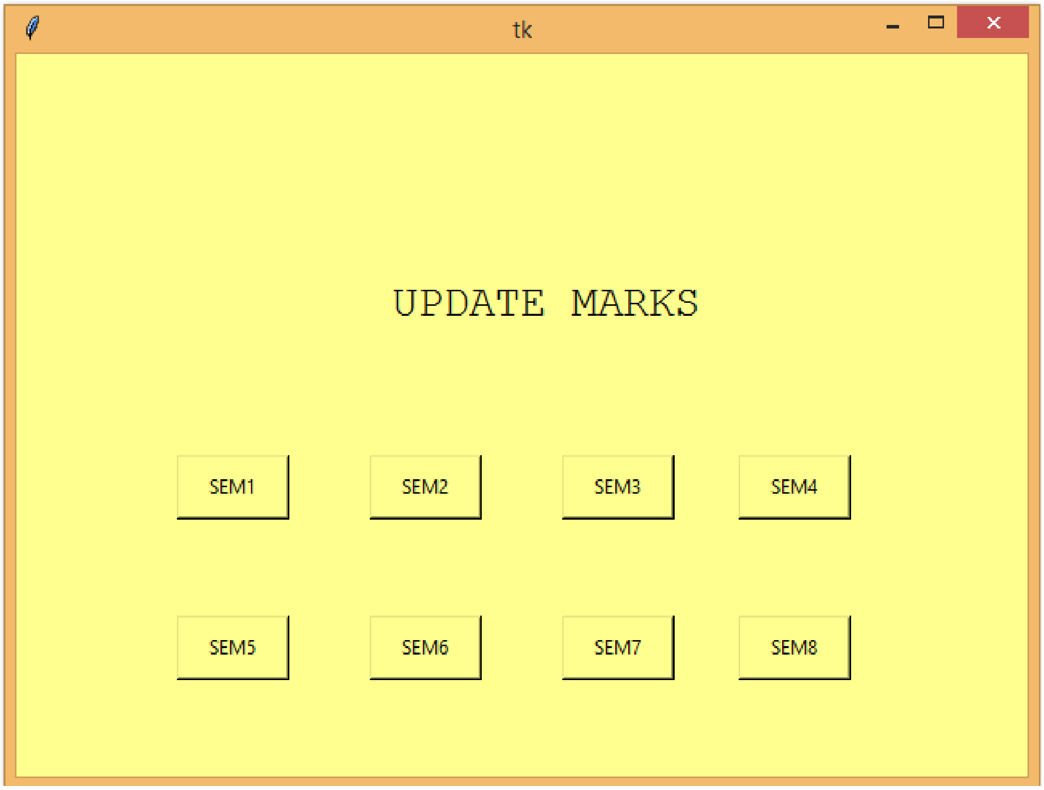

Fig 6.2 teacher should update marks here

TEST CASES

6.1Testing

6.1.1 Strategies used for Testing

1. Unit Testing

Unit testing concentrates verification on the smallest element of the program – the module. Using the detailed design description important control paths are tested to establish errors within the bounds of the module.

In this system each sub module is tested individually as per the unit testing such as campaign, lead, contact etc are tested individually. Their input field validations are tested.

2. Integration testing

Once all the individual units have been tested there is a need to test how they were put together to ensure no data is lost across interface, one module does not have an adverse impact on another and a function is not performed correctly.

After unit testing each and every sub module is tested with integrating each other.

System testing for the current system:

In this level of testing we are testing the system as a whole after integrating all the main modules of the project. We are testing whether system is giving correct output or not. All the modules were integrated and the flow of information among different modules was checked. It was also checked that whether the flow of data is as per the requirements or not. It was also checked that whether any particular module is non-functioning or not i.e. once the integration is over each and every module is functioning in its entirety or not.

In this level of testing we tested the following: -

- Whether all the forms are properly working or not.
- Whether all the forms are properly linked or not.
- Whether all the images are properly displayed or not.
- Whether data retrieval is proper or not.

Specific knowledge of the application's code/internal structure and programming knowledge in general is not required. The tester is aware of *what* the software is supposed to do but is not aware of *how* it does it. For instance, the tester is aware that a particular input returns a certain, invariable output but is not aware of *how* the software produces the output in the first place.

Test Cases

Test cases are built around specifications and requirements, i.e., what the application is supposed to do. Test cases are generally derived from external descriptions of the software, including specifications, requirements and design parameters. Although the tests used are primarily *functional* in nature, *non-functional* tests may also be used. The test designer selects both valid and invalid inputs and determines the correct output without any knowledge of the test object's internal structure.

1] Test case For Teacher Login Page:

Total no of test Cases:-04

Total no of test Cases Passed:-04

Total no of test Cases failed:-00

Total no of test Cases executed:-04

Total no of test Cases pending:-00

Test Case ID	Test Case Procedure	Input Data	Expected Output	Actual Output	Test Status
DST-LG-01	Checking the functionality of Teacher LOGIN Button	1.Enter valid Usernames in textbox 2. Enter valid Password in textbox 3. Click on Teacher LOGIN Button	Teacher Welcome page should be displayed	Teacher Welcome page displayed	Pass
DST - LG-02	Checking the functionality of Teacher LOGIN Button	1.Enter invalid User Name in text box 2. Enter valid Password in password	Teacher Welcome page should not be displayed	Teacher Welcome page is not displayed	Pass

		textbox 3. Click on Teacher LOGIN Button			
DST - LG-03	Checking the functionality of Teacher LOGIN Button	1.Enter valid User name in User name textbox 2. Enter invalid Password in password textbox 3. Click on Teacher LOGIN Button	Teacher Welcome page should not be displayed	Teacher Welcome page is not displayed	Pass
DST - LG-04	Checking the functionality of Teacher LOGIN Button	1.Enter invalid User name in User name textbox 2. Enter invalid Password in password textbox 3. Click on Teacher LOGIN Button	Teacher Welcome page should not be displayed	Teacher Welcome page is not displayed	Pass

2] Test case For Add Mark:

Total no of test Cases:-04

Total no of test Cases Passed:-04

Total no of test Cases failed:-00

Total no of test Cases executed:-04

Total no of test Cases pending:-00

Test Case ID	Test Case Procedure	Input Data	Expected Output	Actual Output	Test Status
DST-AD-01	Checking the functionality of Add Semester wise mark Button	Upload valid Excel file	Upload File Successfully	Upload File Successfully	Pass
DST - AD-02	Checking the functionality of Add Semester wise mark Button	Upload Invalid Excel file	File Not Upload Successfully	File Not Upload Successfully	Pass

DST - AD-03	Checking the functionality of Add Semester wise mark Button	Upload valid Excel file	Upload File Successfully	File Not Upload Successfully	Fail

CHAPTER 7

CONCLUSION AND FUTURE WORK

Present study shows that academic performance of the students is primarily dependent on their past performances. Past Performance indeed got a significant influence over student's present performance. Machine Learning has come far from its nascent stages, and can prove to be a powerful tool in Academic In this we propose new method by using linear regression, exponential moving averages, gradient and impact factor algorithms for predicting Students Performance using their current as well as Past Academic records And taking an input of behavioral based questions which helps to know behavior or the attitude of the Student towards particular thing.

REFERENCES

[1] J. Xu is with the Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL, USA.“ A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs” IEEE, and Mihaela van der Schaar, Fellow, IEEE 2017.

[2] Prediction by Exponentially Weighted Moving Averages and Related D. R. Cox Journal of the Royal Statistical Society. SeriesB(methodological) Vol. 23, No. 2(1961), pp. 414-422 , 2017.

[3] Data Mining Approach for Predicting Student Performance, Research Gate, Edin Osmanbegovic’ &Mirza Suljic , May 2012.

[4] The Utility of Clustering in Prediction Tasks, Shubhendu Trivedi, Zachary A. Pardos and Neil T. Haffernan, e Department of Education IES Math Centre for Mathematics and Cognition grant. Report Date: 05 September 2011.

[5] Student Performance Prediction using Machine Learning, HavanAgarwal , HarshilMavaniDept of Information Technology K.J.Somaiya College of Engineering,Mumbai, IJERT -2015.

[6] Early Prediction of Students Performance using Machine Learning Techniques, Anal Acharya,Department of Computer Science,St Xavier’s College, Kolkata,India. , DevadattaSinha, Department of Computer Science,and Engineering University of Calcutta, Kolkata, India. International Journal of Computer Applications , 2014.

[7] Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue, PoojaThakar Assistant Professor VIPS, GGSIPU Delhi, India, Anil Mehta, Ph.D Professor University of RajasthanJaipur,India, International Journal of Computer Applications, 2015

[8] 8. Using Machine Learning to Predict Student Performance, Murat Pojon, University of Tampere Faculty of Natural Sciences Software DevelopmentM. Sc. Thesis Supervisor: JormaLaurikkala June 2017.

[9] Feng, M., Heffernan, N.T., &Koedinger, K.R. (2010). Addressing the assessment challenge in an Online System that tutors as it assesses. User Modeling and User-Adapted Interaction: The Journal of Personalization Research (UMUAI journal). Volume 19, p.243-26.

[10] Y. Meier, J. Xu, O. Atan, and M. van der Schaar, "Personalized grade prediction: A data mining approach," in Data Mining (ICDM), 2015 IEEE International Conference on. IEEE, 2015, pp. 907–912.

APPENDIX

Dear Aditya Shelar ,

Your manuscript with Registration ID: JETIR209359 has been **Accepted** for publication in the Journal of Emerging Technologies and Innovative Research (www.jetir.org). Your Review Report is as follows:

UGC and ISSN Approved | 5.87 Impact factor Review Results

Registration ID	JETIR209359
Email ID	adityashelar.it@mmcoe.edu.in
Paper Title	Sudent performance prediction using machine learning
Review Status	Accepted
Impact Factor & Licence:	UGC,ISSN,DOI and Creative Common Approved & 5.87 Calculated by Google Scholar
Unique Contents	76 %
Comments	paper accepted