Project Name:
SMART MONITORING AND PREDICTING STUDENT SKILL USING MACHINE LEARNING APPROACH

Abstract

Educational Data Mining (EDM) and Learning Analytics (LA) research have emerged as interesting areas of research, which are unfolding useful knowledge from educational databases for many purposes such as predicting students success. Accurately predicting students future performance based on their ongoing academic records is crucial for effectively carrying out necessary pedagogical interventions to ensure students on-time and satisfactory graduation, predicting student performance in completing degrees (e.g. college programs) is much less studied and faces new challenges: (1) Students differ tremendously in terms of backgrounds and selected courses; (2) Courses are not equally informative for making accurate predictions; (3) Students evolving progress needs to be incorporated into the prediction. In this project, we develop a novel machine learning method for predicting student performance in degree programs that is able to address these key challenges.

The proposed method has two major features. First, a belayed structure comprising of multiple base predictors and a cascade of ensemble predictors is developed for making predictions based on students evolving performance states. Second, a data-driven approach based on latent factor models and probabilistic matrix factorization is proposed to discover course relevance, which is important for constructing efficient base predictors.

INDEX

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Making higher education affordable has a significant impact on ensuring the nations economic prosperity and represents a central focus of the government when making education policies [1]. Yet student loan debt in the United States has blown past the trillion-dollar mark, exceeding Americans combined credit card and auto loan debts [2]. As the cost in college education (tuitions, fees and living expenses) has skyrocketed over the past few decades, prolonged graduation time has become a crucial contributing factor to the evergrowing student loan debt. In fact, recent studies show that only 50 of the more than 580 public four-year institutions in the United States have on-time graduation rates at or above 50 percent for their full-time students [2]. To make college more 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 s to build a system that can continuously keep track of students academic performance and accurately predict their future performance, such as when they are likely to graduate and their estimated final GPAs, given the current progress. Although predicting student performance has been extensively studied in the literature, it was primarily studied in the contexts of solving problems in Intelligent Tutoring Systems (ITSs) [3][4][5][6], or completing courses in classroom settings or in Massive Open Online Courses (MOOC) platforms [7][8]. However, predicting student performance within a degree program (e.g. college program) is significantly different and faces new challenges.

First, students can differ tremendously in terms of backgrounds as well as their chosen areas (majors, specializations), resulting in different selected courses as well as sequences in which they take the courses. On the other hand, the same course can be taken by students in different areas. Since predicting student performance in a particular course relies on the student past performance in other courses, a key challenge for training an effective predictor is how to handle heterogeneous student data due to the different areas and interests. In contrast, solving problems in ITSs often follow routine steps which are the same for all students [9]. Similarly, predictions of students performance in courses are often based on in-course assessments which are designed to be the same for all students [7].

Second, students may take many courses but not all courses are equally informative for predicting students future performance. Utilizing the students past performance in all courses that

he/she has completed not only increases complexity but also introduces noise in the prediction, thereby degrading the prediction performance. For instance, while it makes sense to consider a students grade in the course Linear Algebra for predicting his/her grade in the course Linear Optimization, the students grade in the course Chemistry Lab may have much weaker predictive power. However, the course correlation is not always as obvious as in this case. Therefore, discovering the underlying correlation among courses is of great importance for making accurate performance predictions.

Third, predicting student performance in a degree program is not a one-time task; rather, it requires continuous tracking and updating as the student finishes new courses over time. An important consideration in this regard is that the prediction needs to be made based on not only the most recent snapshot of the student accomplishments but also the evolution of the student progress, which may contain valuable information for making more accurate predictions. However, the complexity can easily explode since even mathematically representing the 2 evolution of student progress itself can be a daunting task. However, treating the past progress equally as the current performance when predicting the future may not be a wise choice either since intuition tells us that old information tends to be outdated .

1.2 MOTIVATION OF THE PROJECT

- More students are enrolling in college and professional degree programs than ever before. However, current degree programs are often one size fits all; such programs ignore the heterogeneity of students in terms of backgrounds, abilities, learning styles and career goals.

- As the cost in college education (tuitions, fees and living expenses) has skyrocketed over the past few decades, prolonged graduation time has become a crucial contributing factor to the ever growing student loan debt.

- Because of ever-increasing student/teacher ratios, students are often left struggling to find their own path through degree programs. The combination leads to poor learning outcomes, low engagement, dissatisfaction and high dropout rates.

CHAPTER 2

LITERATURE SURVEY

2.1 LITERATURE SURVEY

1. Nguyen Thai-Nghe, Lucas Drumond, Tomas Horvffath, and Lars Schmidt-Thieme, University of Hildesheim "Multi-Relational Factorization Models for Predicting Student Performance.

   In this paper we propose to exploit such multiple relationships by using multi-relational MF methods. Experiments on three large datasets show that the proposed approach can improve the prediction results. Predicting student performance (PSP) is the problem of predicting how well a student will perform on a given task. It has gained more attention from the educational data mining community recently. Previous works show that good results can be achieved by casting the PSP to rating prediction problem in recommender systems, where students, tasks and performance scores are mapped to users, items and ratings respectively.

2. Beverly Park Woolf, Ryan Baker, Worcester Polytechnic, Institute Erwin P. Gianchandani "From Data to Knowledge to Action: Enabling Personalized Education.

   We describe how data analytics approaches have the potential to dramatically advance instruction for every student and to enhance the way we educate our children. The Internet, intelligent environments, and rich interfaces (including sensors) allow us to capture much more data about learners than ever before and the quantities of data are growing at a rapidly accelerating rate.

3. Nazeema Alli, Rahim Rajan, and Greg Ratliff 3. How personalized learning unlocks student Success.

   EDUCAUSE Review is the general-interest, bimonthly magazine published by EDUCAUSE. With a print publication base of 22,000, EDUCAUSE Review is sent to EDUCAUSE member representatives as well as to presidents/chancellors, senior academic and administrative leaders, non-IT staff, faculty in all disciplines, librarians, and corporations. It takes a broad look at current developments and trends in information technology, what these mean for higher education, and how they may affect the college/university as a whole.

CHAPTER 3

OVERVIEW OF SYSTEM

3.1 PROBLEM STATEMENT

- To address the aforementioned challenges, we proposed a novel algorithm for predicting students performance in college programs given his/her current academic records. In Proposed studies shows that academic performances of the students are primarily dependent on their past performances. Our investigation confirms that past performances have indeed got a significant influence over students performance. Further, we confirmed that the performance of neural network increases with increase in dataset size.
  System will comprise the tracking of detailed information of a student regarding his academics and curricular activity and would predict the right learning Courses using a algorithm over the information tracked meeting the ambition or the goal for a student.

3.1.1 Goals and objectives

- To develop a system for automatic identification the basic goals and ambitions of the student.
- To automate decision making for selection of course.
- To apply efficient method for effective measurement of student performance to recommend course to student based on result.
- To track student performance and motivate him/ her towards their goals.
- Applying various kinds of techniques to effectively calculate students performance and recommend course according to the result.
- To track students performance and motivate him/her towards their goals and providing them with a right path.

3.1.2 Statement of scope

- A description of the software with Size of input, bounds on input, input validation, input dependency, i/o state diagram, Major inputs, and outputs are described without regard to implementation detail.

- The scope identifies what the product is and is not, what it will and wont do, what it will and wont contain.

3.2 OVERVIEW

In the last decade, school conducts examination manually. It has so many problems. The existing systems are very time consuming. It is difficult to analyze the exam manually. Results are not precise as calculation and evaluations are done manually. Result processing after summation of exam takes more time as it is done manually. So we introduce a Preschool examination Portal system, which is fully computerized. Existing system is a large man power process and is difficult to implement. It provides an easy to use environment for both Test Conductors and Students appearing for Examination. The main objective of Preschool examination Portal system is to provide all the features that an Examination System must have, with the ”interfaces that don’t Scare it’s Users!”

CHAPTER 4

SYSTEM REQUIREMENT SPECIFICATION

4.1 INTRODUCTION

The introduction of the Software Requirements Specification (SRS) provides an overview of the entire SRS with purpose, scope, definitions, acronyms, abbreviations, references and overview of the SRS. The aim of this document is to gather and analyze and give an in-depth insight of the complete Disease Prediction by Machine Learning from Healthcare Communities system with minimized time consumption by defining the problem statement in detail.

4.1.1 Purpose and Scope of Document

Purpose: We going to develop a the proposed method which has two major features. First, a belayed structure comprising of multiple base predictors and a cascade of ensemble predictors is developed for making predictions based on students evolving performance states. Second, a data-driven approach based on latent factor models and probabilistic matrix factorization is proposed to discover course relevance, which is important for constructing efficient base predictors.

4.1.2 Overview of responsibilities of Developer

developer should be able :

1. To have understanding of the problem statement.
2. To know what are the hardware and software requirements of proposed system.
3. To have understanding of proposed system.
4. To do planning various activates with the help of planner.Designing, programming, testing etc.

4.1.3 User Classes and Characteristics

1. **Candidate / User:**
    - Can create registration and login the system which he select.
    - Can fill the all details regarding education profile (Inputs).
    - Can fill the all details regarding extra activities.

- Can check predicted courses recommended course materials and company details also.
- Can logout or if any query will solve through Chatbot model.

2. **Staff:**

- Can login the system.
- Can check the queries of users and reply back.
- Can schedule the tests as per course.
- Can view the test marks detail.
- Can logout the system.

3. **Admin:**

- Can login the system.
- Can activate the candidates / users account.
- Can check the all users details and his marks.
- Can send the course related materials.
- Can logout the system.

4.1.4 Assumption and Dependencies

- End User application will be developed in Windows OS.
- Database will be in MySQL.
- All application code shall conform to the java standard.
- All scripts shall be written in JavaScript.
- Application design pattern shall be Singleton.
- Developer IDE shall be Eclipse Kepler.

4.2 FUNCTIONAL REQUIREMENTS

- The application is user friendly.
- It provides an easy interface to user.
- The accessibility or response time of the application should be fast.
- Performance of the system is appropriate.

4.3 EXTERNAL INTERFACE REQUIREMENTS

4.3.1 User Interface

- Capture input parameter.
- Processing through server.
- Generate or predict course.
- Descision making on predicted course.
- Processing it.
- Recommend course material.
- Predict exam regarding to course selection.
- Capture input and course score and alanyse performance.
- Outcome
- Time and money saving.

4.3.2 Hardware Interface

- System : Intel (R) Core (TM) i5-7200U CPU @ 2.50GHz 2.70GHz
- Hard Disk : 1TB
- RAM : 8GB

4.3.3 Software Interface

- Operating System : Windows 7 (Above).
- Coding Language : Java/J2EE.
- Data Base : Mysql5.1
- IDE : Eclipse Kepler.

4.3.4 Communication Interface

- Our Project belongs to web based, so connecting user at online with request and response form. For that HTTP protocol we are going to use. That is provided by tomcat server 7/8.
- HTTP protocol: The Hypertext Transfer protocol is an application protocol for distributed, collaborative, and hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web. Hypertext is structured text that uses logical links (hyperlinks) between nodes containing text.

4.4 NON-FUNCTIONAL REQUIREMENTS

4.4.1 Performance Requirement

- System can produce results faster on 4GB of RAM.
- It may take more time for peak loads at main node.
- The system will be available 100% of the time. Once there is a fatal error, the system will provide understandable feed back to the user.

4.4.2 Safety Requirements

- Only administrators have access to the database of each individual user.
- All data will be backed-up every day automatically and also the system administrator can back-up the data as a function for him.
- This makes it easier to install and updates new functionality if required.
- For the safety purpose backup of the database must be required.

4.4.3 Security Requirements

- Our System is being developed in Java. Java is an object oriented programming language and shall be easy to maintain. The system is designed in modules where errors can be detected and fixed easily.
- For the security purposes and to avoid illegal use of the system, while using this application user must do following things:
  - At the time of deploying this software user have to register to system.
  - To use software user have to login and logout each time.

4.4.4 Software Quality Attributes

- The system considers following non-functional requirements to provide better functionalities and usage of system.
- **Availability :** The system shall be available during 24 hours of a day.
- **Usability :** The system is designed keeping in mind the usability issues considering the end-users who are developers/programmers. It provides detailed help which would lead to better and faster learning. Navigation of system is easy.
- **Consistency :** Uniformity in layout, screens, Menus, colours scheme, format.
- Performance : The performance of the system should be fast and as per user requirement. From this system we will get expected outcome in less time and less space since efficiency is higher. Speed is totally depending on the response of the database and connection type.
- **Extendibility :** Prevention in the system should be done in the system by which we make changes in the system later on.
- **Reusability :** Files of any type can be used by the system for any number of times during transformation.
- **Reliability :** Protection of data from malicious attack or unauthorized access.
- **Security :** The system provides security to the randomly generated private key by performing encryption to it for encrypting patient data and thus protects from other nodes in the network. The network is free from malicious node and misbehaving node attacks.

- **Reliability :** Our system can provide user an efficient search each time. So the user can reliable on the system. Because system can guarantees user to provide his/her interested data every time in least amount of time.

4.5 SYSTEM REQUIREMENTS

4.5.1 Database Requirements

- **MySQL :**

  MySQL, the most popular Open Source SQL database management system, is developed, distributed, and supported by Oracle Corporation.

  The MySQL Web site (http://www.mysql.com/) provides the latest information about MySQL software. MySQL is database software which is used to store all database related activities regarding to our project and it is easily stored and retrieve the data.

4.5.2 Software Requirements

With most programming languages, you either compile or interpret a program so that you can run it on your computer. The Java programming language is unusual in that a program is both compiled and interpreted. With the compiler, first you translate a program into an intermediate language called Java byte codes the platform-independent codes interpreted by the interpreter on the Java platform. The interpreter parses and runs each Java byte code instruction on the computer. Compilation happens just once; interpretation occurs each time the program is executed. The following figure illustrates how this works.

4.6 HARDWARE REQUIREMENTS

- System : Intel (R) Core (TM) i5-7200U CPU @ 2.50GHz 2.70GHz
- Hard Disk : 1TB
- RAM : 8GB

CHAPTER 5

PROJECT PLAN

5.1 PROJECT ESTIMATES

We are using waterfall model for our project estimation.

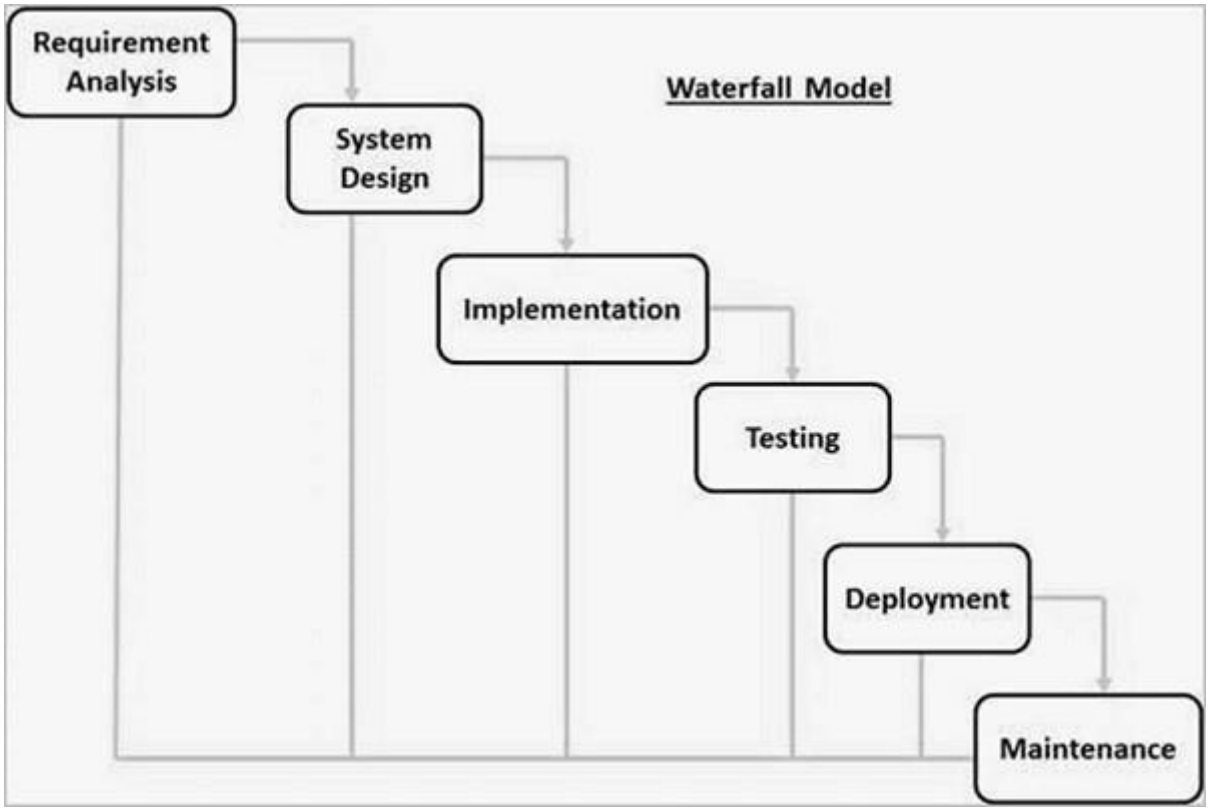


Figure 5.1: Waterfall Model

1. **Requirement gathering and analysis:**

In this step of waterfall we identify what are various requirements are need for our project such are software and hardware required, database, and interfaces.

2. **System Design:**

In this system design phase we design the system which is easily understood for end user i.e. user friendly. We design some UML diagrams and data flow diagram to understand the system flow and system module and sequence of execution.

3. **Implementation:**

In implementation phase of our project we have implemented various module required of successfully getting expected outcome at the different module levels. With inputs from system

design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality which is referred to as Unit Testing.

4. Testing:

The different test cases are performed to test whether the project module are giving expected outcome in assumed time. All the units developed in the implementation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures.

5. Deployment of System:

Once the functional and nonfunctional testing is done, the product is deployed in the customer environment or released into the market.

6. Maintenance:

There are some issues which come up in the client environment. To fix those issues patches are released. Also to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment.

All these phases are cascaded to each other in which progress is seen as flowing steadily down-wards like a waterfall through the phases. The next phase is started only after the defined set of goals are achieved for previous phase and it is signed off, so the name ”Waterfall Model”. In this model phases do not overlap.

5.1.1 Reconciled Estimates

5.1.1.1 Cost Estimate

Not applicable.

5.1.1.2 Time Estimates

Approximately 11 months.

5.1.2 Project Resources

Windows 7 (Above) , eclipse IDE, 2.70GHZ CPU speed, 8 GB RAM, High speed internet connection.

5.2 RISK MANAGEMENT W.R.T. NP COMPLETE ANALYSIS

Np Complete:

Using PTTP algorithm it is possible to find out the optimized solution in polynomial time by reducing the actual problem P

5.2.1 Risk Analysis

The risks for the Project can be analyzed within the constraints of time and quality.

ID	Risk Description	Probability	Impact		
			Schedule	Quality	Overall
1	Correctness	Medium	High	High	Medium
2	Availability	High	High	High	High

Table 5.1: Risk Table

5.2.2 Risk Probability definitions:

Probability	Value	Description
High	Probability of occurrence is	> 75%
Medium	Probability of occurrence is	26 – 75%
Low	Probability of occurrence is	< 25%

Table 5.2: Risk Probability definitions

5.2.3 Risk Impact definitions:

Impact	Value	Description
Very high	> 10%	Schedule impact or Unacceptable quality
High	5 – 10%	Schedule impact or Some parts of the project have low quality
Medium	< 5%	Schedule impact or Barely noticeable degradation in quality Low Impact on schedule or Quality can be incorporated

Table 5.3: Risk Impact definitions

5.2.4 Overview of Risk Mitigation, Monitoring, Management

Following are the details for each risk.

Risk ID	1
Risk Description	Time extend
Category	Environment
Source	Time Predicition
Probability	Medium
Impact	High
Response	Mitigate
Strategy	Break schedule
Risk Status	Occurred

5.3 PROJECT SCHEDULE

The project schedule starts from Aug 2017 and ends in March 2018. The Different Phases identified are:

1. Requirement Analysis.
2. Requirement Specification.
3. System Design.
4. Detailed Design.
5. Coding.
6. Testing.

The Gantt chart as shown below represents the approximate schedule followed for the completion of each phase:

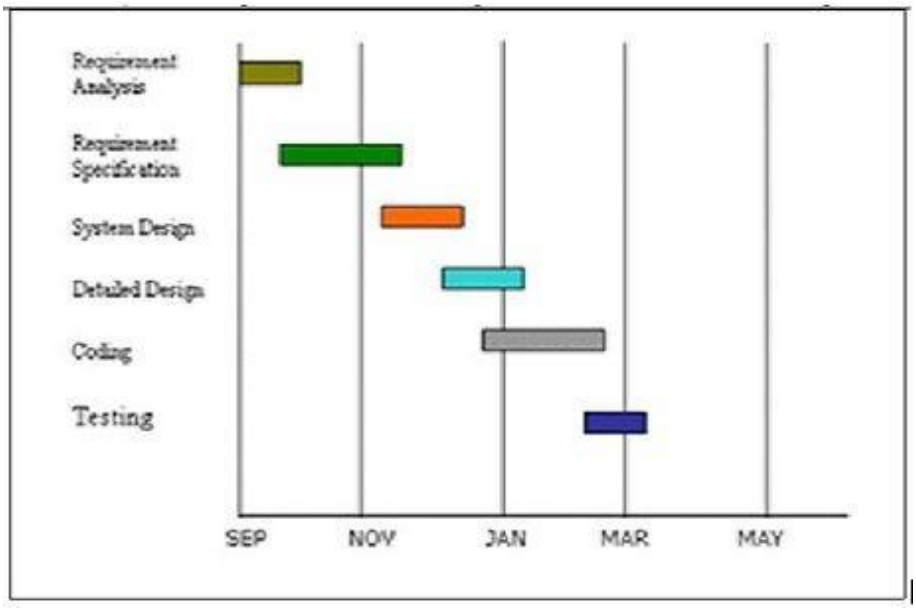


Figure 5.2: Project Schedule

5.3.1 Project task set

Major Tasks in the Project stages are:

- Task 1: Correctness.
- Task 2: Availability.
- Task 3: Integrity.

5.3.2 Task network

Project tasks and their dependencies are noted in this diagrammatic form.

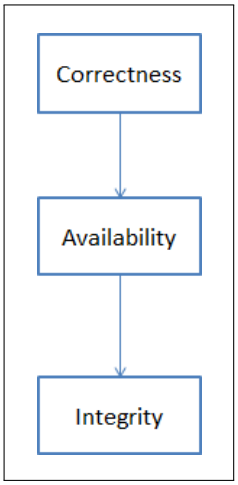


Figure 5.3: Task network

5.3.3 Timeline Chart

A project timeline chart is presented. This may include a time line for the entire project. Above points should be covered in Project Planner as Annex C and you can mention here Please refer Annex C for the planner

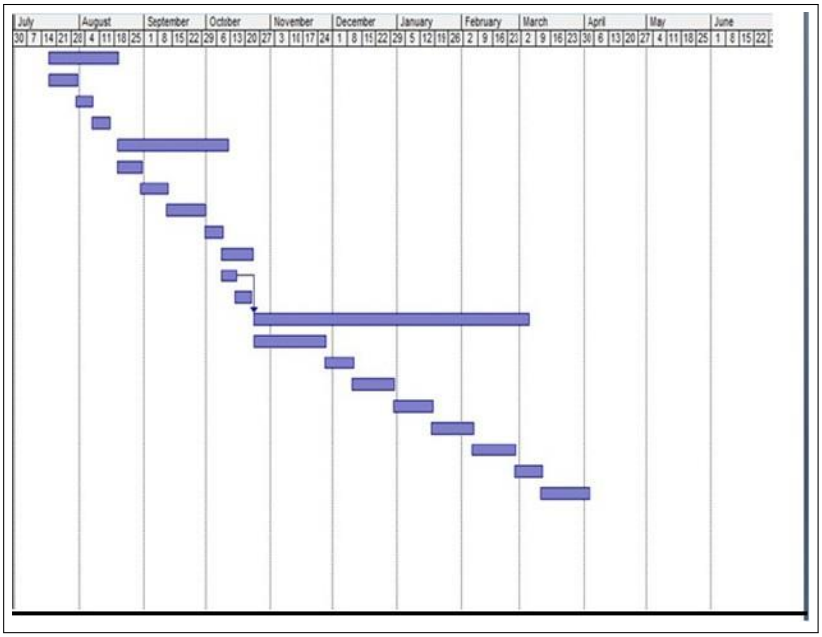


Figure 5.4: Timeline Chart

5.4 TEAM ORGANIZATION

Team consists of only one member and proper planning mechanism is used and roles are defined.

5.4.1 Team structure

The team structure for the project is identified. Roles are defined.

Sr. No.	Member	Responsibilities
1	Kadu Bhakti & Kawade Rutuja	Project analysis, Developer
2	Shinde Supriya	Requirement Gathering, middleware
3	Tribhuvan Priyanka	Testing and Design

5.4.2 Management reporting and communication

Well planning mechanisms are used for progress reporting and inter/intra team communication are identified as per requirements of the project.

CHAPTER 6

PROJECT IMPLEMENTATION

6.1 INTRODUCTION

To overcome this problem of existing system, we proposed a system which is First, students can differ tremendously in terms of backgrounds as well as their chosen areas (majors, specializations), resulting in different selected courses as well as sequences in which they take the courses. Second, students may take many courses but not all courses are equally informative for predicting students future performance. Utilizing the students past performance in all courses that he/she has completed not only increases complexity but also introduces noise in the prediction, thereby degrading the prediction performance. Third, predicting student performance in a degree program, it does tracking and updating as the student finishes new courses over time.

6.2 OVERVIEW OF PROJECT MODULES

1. **Candidate / User:**
    - Can create registration and login the system which he select.
    - Can fill the all details regarding education profile (Inputs).
    - Can fill the all details regarding extra activities.
    - Can check predicted courses recommended course materials and company details also.
    - Can logout or if any query will solve through Chatbot model.
2. **Staff:**
    - Can login the system.
    - Can check the queries of users and reply back.
    - Can schedule the tests as per course.
    - Can view the test marks detail.
    - Can logout the system.
3. **Admin:**
    - Can login the system.
    - Can activate the candidates / users account.

- Can check the all users details and his marks.
- Can send the course related materials.
- Can logout the system.

6.3 TOOLS AND TECHNOLOGIES USED

- **JAVA:** Java is one of the most popular and widely used programming language and platform. A platform is an environment that helps to develop and run programs written in any programming language.

  Java is fast, reliable and secure. From desktop to web applications, scientific supercomputers to gaming consoles, cell phones to the Internet, Java is used in every nook and corner.

  Java is a programming language and computing platform first released by Sun Microsystems in 1995. There are lots of applications and websites that will not work unless you have Java installed, and more are created every day. Java is fast, secure, and reliable. From laptops to datacenters, game consoles to scientific supercomputers, cell phones to the Internet, Java is everywhere!

  Java is a general-purpose, concurrent, object-oriented, class-based, and the runtime environment(JRE) which consists of JVM which is the cornerstone of the Java platform. This blog on What is Java will clear all your doubts about why to learn java, features and how it works.

- **XAMPP:** XAMPP stands for Cross-Platform (X), Apache (A), MySQL (M), PHP (P) and Perl (P). It is a simple, lightweight Apache distribution that makes it extremely easy for developers to create a local web server for testing purposes. Everything you need to set up a web server server application (Apache), database (MySQL), and scripting language (PHP) is included in a simple extractable file. XAMPP is also cross-platform, which means it works equally well on Linux, Mac and Windows. Since most actual web server deployments use the same components as XAMPP, it makes transitioning from a local test server to a live server is extremely easy as well. Web development using XAMPP is especially beginner friendly.

- **JDK:** The Java Development Kit (JDK) is an implementation of either one of the Java Platform, Standard Edition, Java Platform, Enterprise Edition, or Java Platform, Micro Edition platforms released by Oracle Corporation in the form of a binary product aimed at Java developers on Solaris, Linux, macOS or Windows. The JDK includes a private JVM and a few other resources to finish the development of a Java Application. Since the introduction of the Java platform, it has been by far the most widely used Software Development Kit (SDK).[citation needed] On 17 November 2006, Sun announced that they would release it under the GNU General Public License (GPL), thus making it free software. This happened in large part on 8 May 2007, when Sun contributed the source code to the OpenJDK.

- **Apache:** Apache is the actual web server application that processes and delivers web content to a computer. Apache is the most popular web server online, powering nearly 54 percent of all websites.

- **MySQL:** Every web application, howsoever simple or complicated, requires a database for storing collected data. MySQL, which is open source, is the worlds most popular database management system. It powers everything from hobbyist websites to professional platforms like WordPress. You can learn how to master PHP with this free MySQL database for beginners course.

  .

6.4 ALGORITHM DETAILS

6.4.1 Machine Learning

Machine Learning is such a field which gives an ability to learn without being explicitly programmed. They mainly focus on the prediction. Statistics and Machine Learning are closely related fields. They can be divided into three categories: 1) Supervised Learning 2) Unsupervised Learning 3) Reinforcement Learning

6.4.2 SVM

This set of notes presents the Support Vector Machine (SVM) learning algorithm. SVMs are among the best (and many believe are indeed the best) ”off-the-shelf supervised learning algorithms. To tell the SVM story, well need to first talk about margins and the idea of separating data with a large gap”. Next, well talk about the optimal margin classifier, which will lead us into a digression on Lagrange duality. Well also see kernels, which give a way to apply SVMs efficiently in very high dimensional (such as infinitedimensional) feature spaces, and finally, well close off the story with the SMO algorithm, which gives an efficient implementation of SVMs.

6.4.3 k-Means Clustering

k-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed apriori. The main idea is to define k centers, one for each cluster. These centers should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as barycenter of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this loop we may notice that the k centers change their location step by step until no more changes are done or in other words centers do not move any more. Finally, this algorithm aims at minimizing an objective function know as squared error function given by:

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\left\| x_i - v_j \right\| \right)^2$$

where,

'$||x_i - v_j||$' is the Euclidean distance between x_i and v_j.

'c_i' is the number of data points in i^{th} cluster.

'c' is the number of cluster centers.

Algorithmic steps for k-means clustering

Let X = {x1,x2,x3,..,xn} be the set of data points and V = {v1,v2,.,vc} be the set of centers.

1) Randomly select c cluster centers.

2) Calculate the distance between each data point and cluster centers.

3) Assign the data point to the cluster center whose distance from the cluster center is minimum of all the cluster centers

4) Recalculate the new cluster center using:

$$v_i = (1/c_i) \sum_{j=1}^{c_i} x_i$$

Where, ci represents the number of data points in ith cluster.

5) Recalculate the distance between each data point and new obtained cluster centers.

6) If no data point was reassigned then stop, otherwise repeat from step 3).

CHAPTER 7

DETAILED DESIGN DOCUMENT

7.1 INTRODUCTION

This document specifies the design that is used to solve the problem of Product.

7.2 ARCHITECTURAL DESIGN

First, students can differ tremendously in terms of backgrounds as well as their chosen areas (majors, specializations), resulting in different selected courses as well as sequences in which they take the courses. Second, students may take many courses but not all courses are equally informative for predicting students future performance. Utilizing the students past performance in all courses that he/she has completed not only increases complexity but also introduces noise in the prediction, thereby degrading the prediction performance. Third, predicting student performance in a degree program, it does tracking and updating as the student finishes new courses over time.

The main purpose behind the proposed system is to provide a comprehensive computerized system, which can capture, collate and analyze the data from these wards and evaluate the impact of the program.

The proposed method has two major features. First, a bilayered structure comprising of multiple base predictors and a cascade of ensemble predictors is developed for making predictions based on students evolving performance states. Second, a data-driven approach based on latent factor models and probabilistic matrix factorization is proposed to discover course relevance, which is important for constructing efficient base predictors. Through extensive simulations on an undergraduate student dataset collected over three years at UCLA.

7.2.1 Internal software data structure

Real time queue

7.2.2 Global data structure

Data Streams

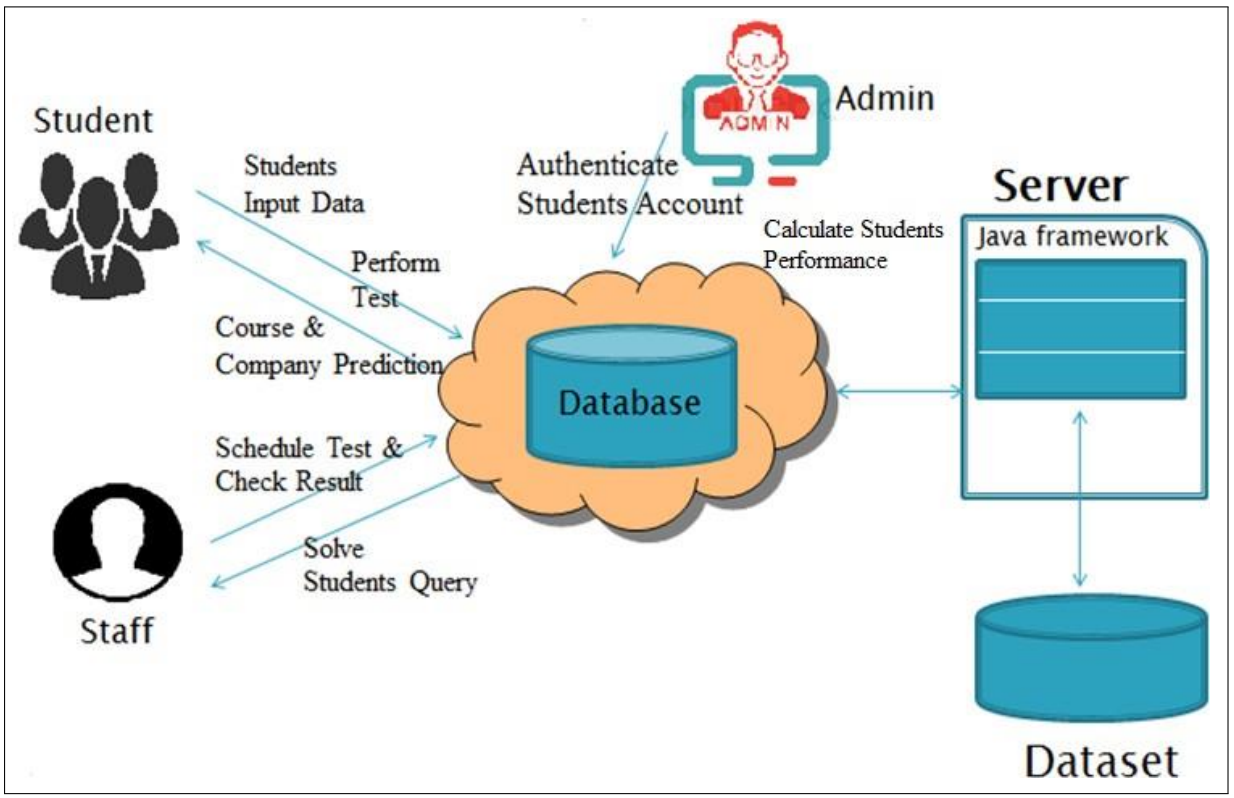


Figure 7.1: Architecture diagram

7.2.3 Temporary data structure

Processing items

7.2.4 Database description

Processed and analyzed data.

7.3 MATHEMATICAL MODEL

Input:

1. Let S be the set of student

S= {s1, s2, s3,… , sn}

2. Let Q be the set of question generated randomly in exam.

Q = {Q1, Q2, Q3,..,Qn}.

3. Let Re bet the academic record monitor by system

Re={Re1, Re2, Re3,…,Ren}.

4. A is the Admin.

Process:

1. S will Register and login at the time of Exam. 2. S will select Ss to solve Q.

   Ss = sections of paper.

   Ss = {Ss1, Ss2, Ss3,... ,Ssn}.

3. S will select O for each Q.

   O is set of selected answer option.

   O = {O1, O2, O3,. ...,On}.

4. R of S will be generated after solving Q.

   R is result.

   R= {R1, R2, R3,. .., Rn}.

5. Re will be generate after R and S. 6. G will be generate and Display as Per Re of S to analyze the IQ of particular S.

   G is Graph.

   G= {G1, G2, G3,. ..., Gn}.

Output:

   X={R,G, Re}

   Students Performance Prediction.

7.4 UML DIAGRAMS:

7.4.1 Data Flow Diagram

7.4.1.1 Level 0 Data Flow Diagram

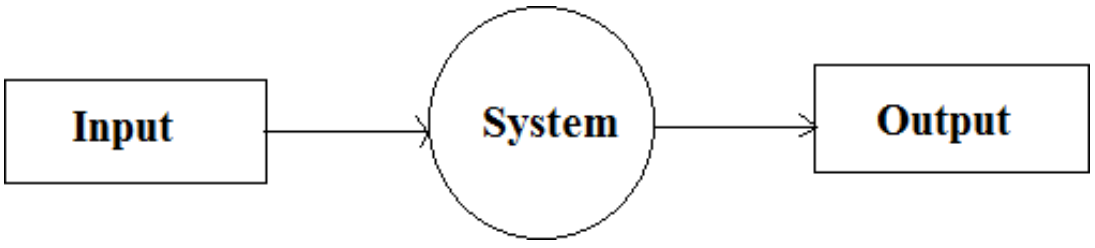


Figure 7.2: DFD0

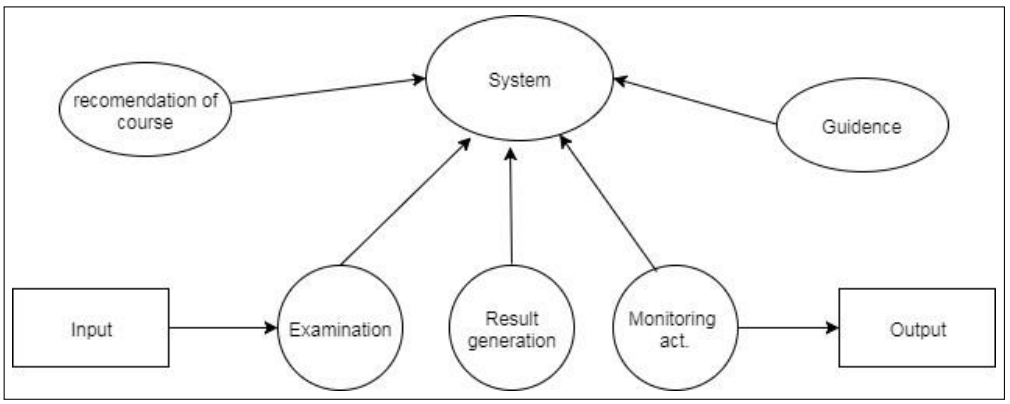


Figure 7.3: DFD1

7.4.1.2 Level 1 Data Flow Diagram

7.4.1.3 Level 2 Data Flow Diagram

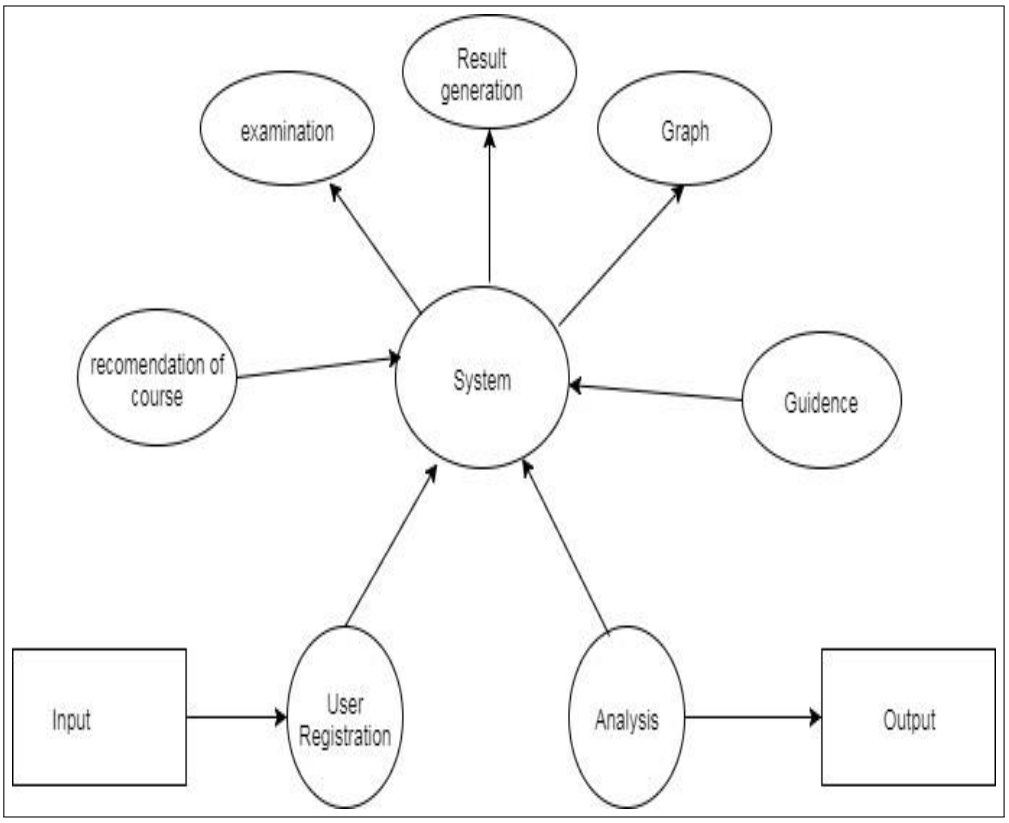


Figure 7.4: DFD2

7.5 ENTITY RELATIONSHIP DIAGRAMS

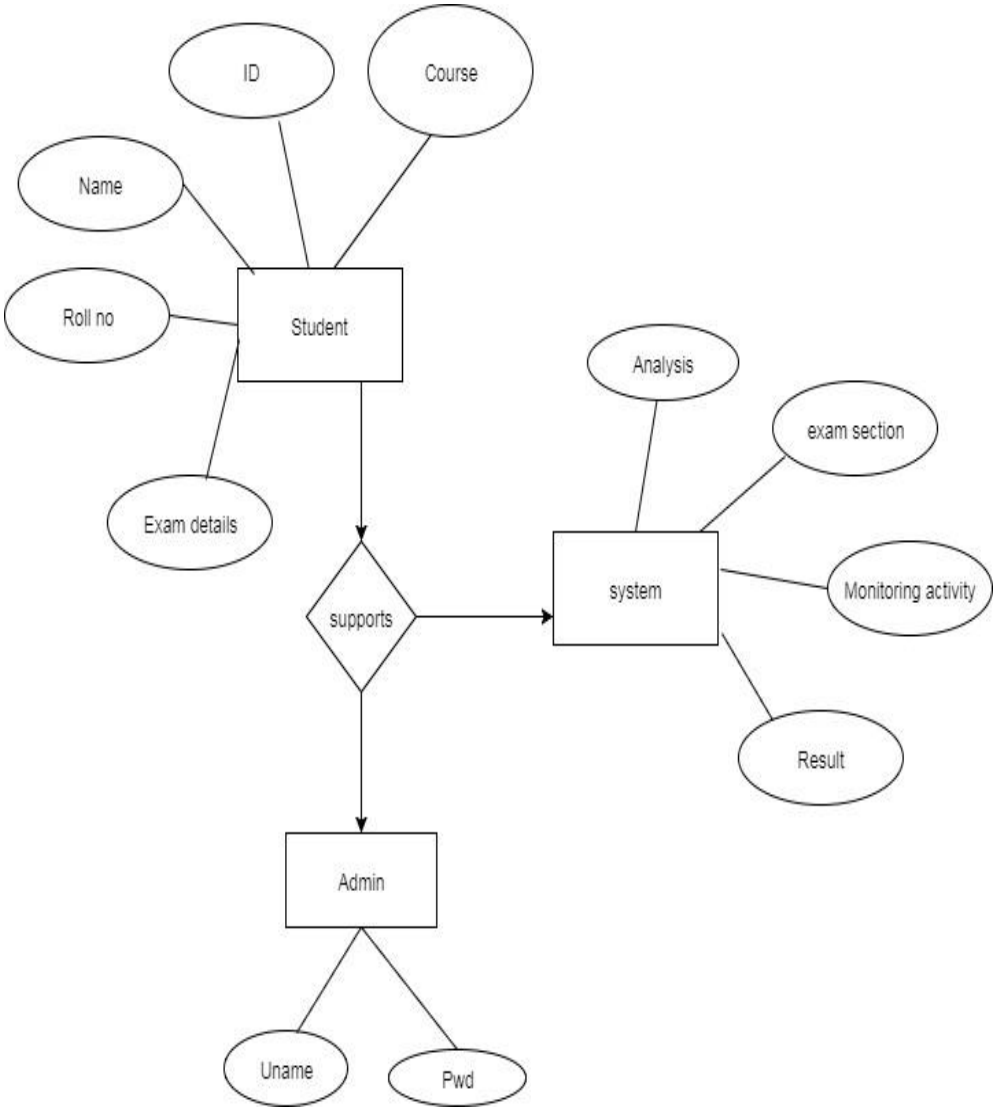


Figure 7.5: Entity Relationship Diagram

7.6 UML DIAGRAM

7.6.1 Use case Diagram (Student)

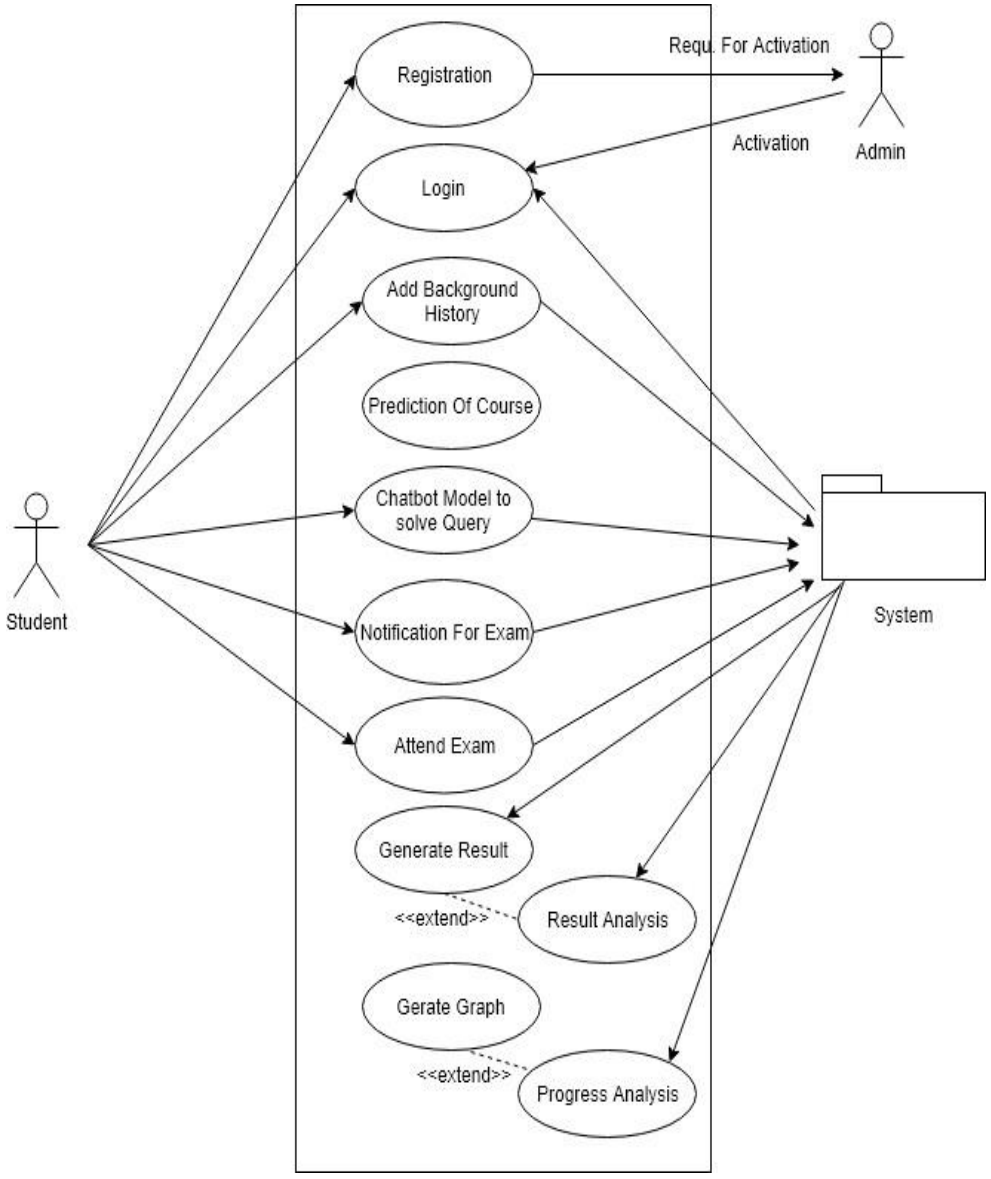


Figure 7.6: Use case Diagram (Student)

7.6.2 Use case Diagram (Faculty)

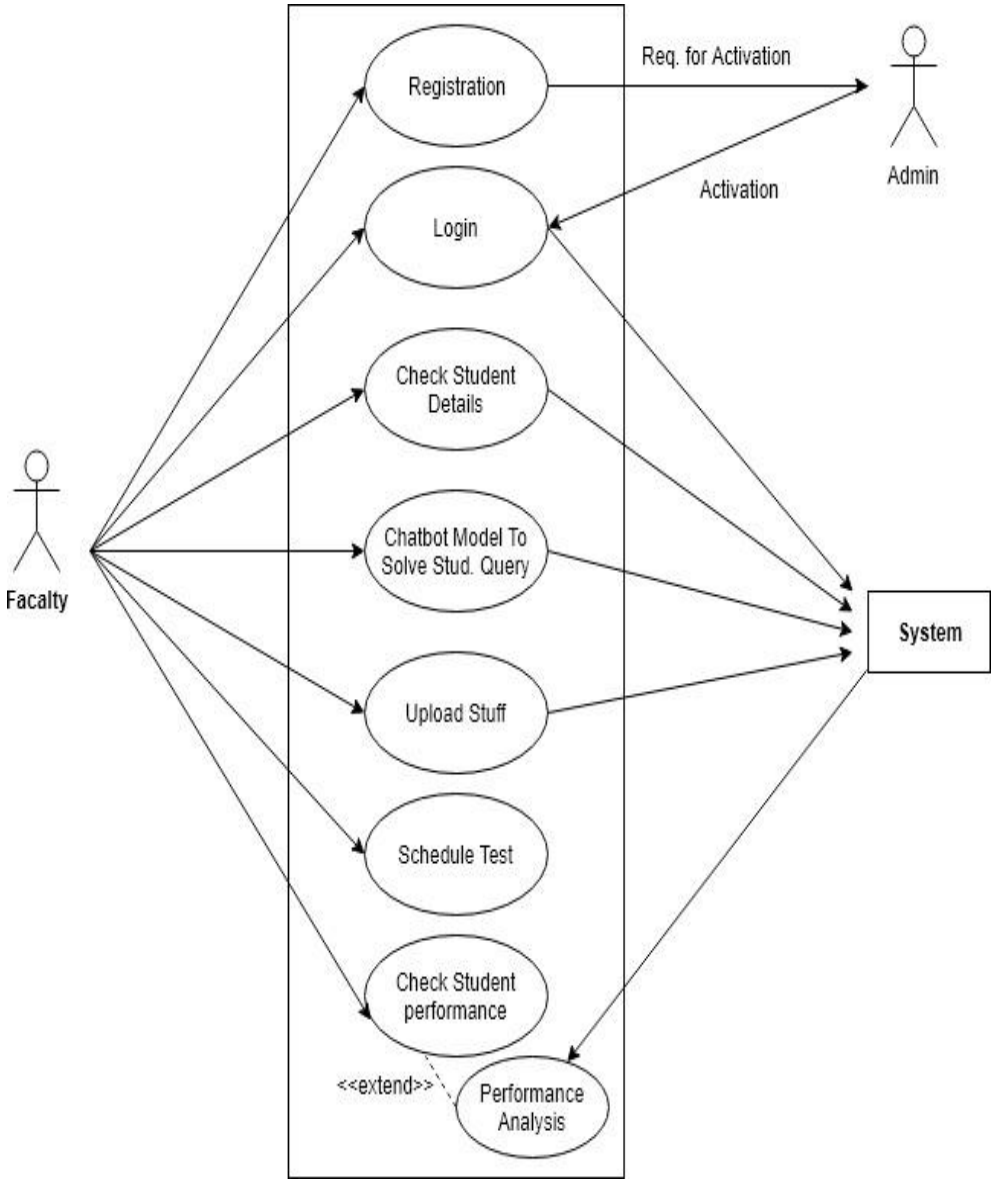


Figure 7.7: Use case Diagram (Faculty)

7.6.3 Sequence Diagram

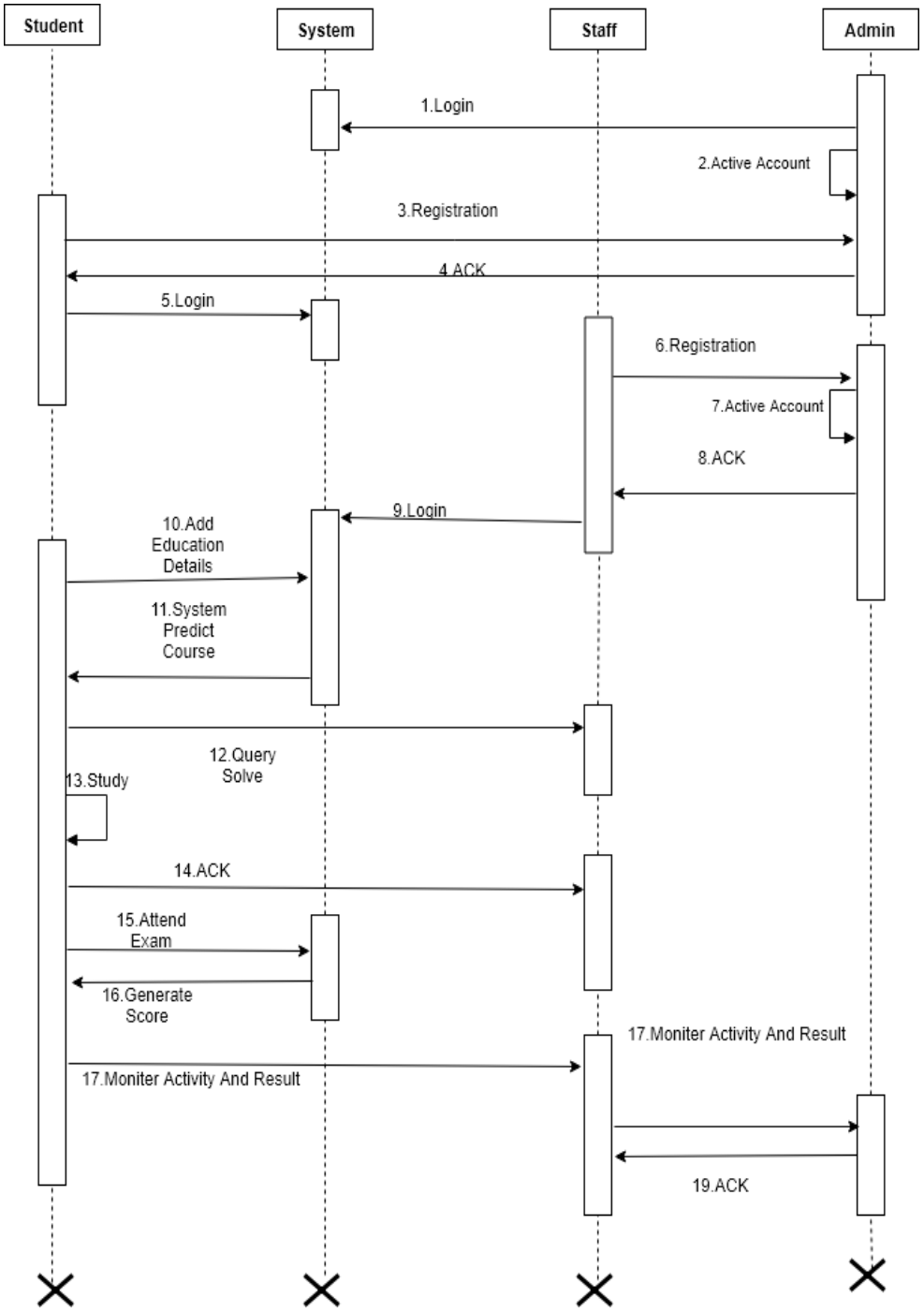


Figure 7.8: Sequence Diagram

7.6.4 Class Diagram

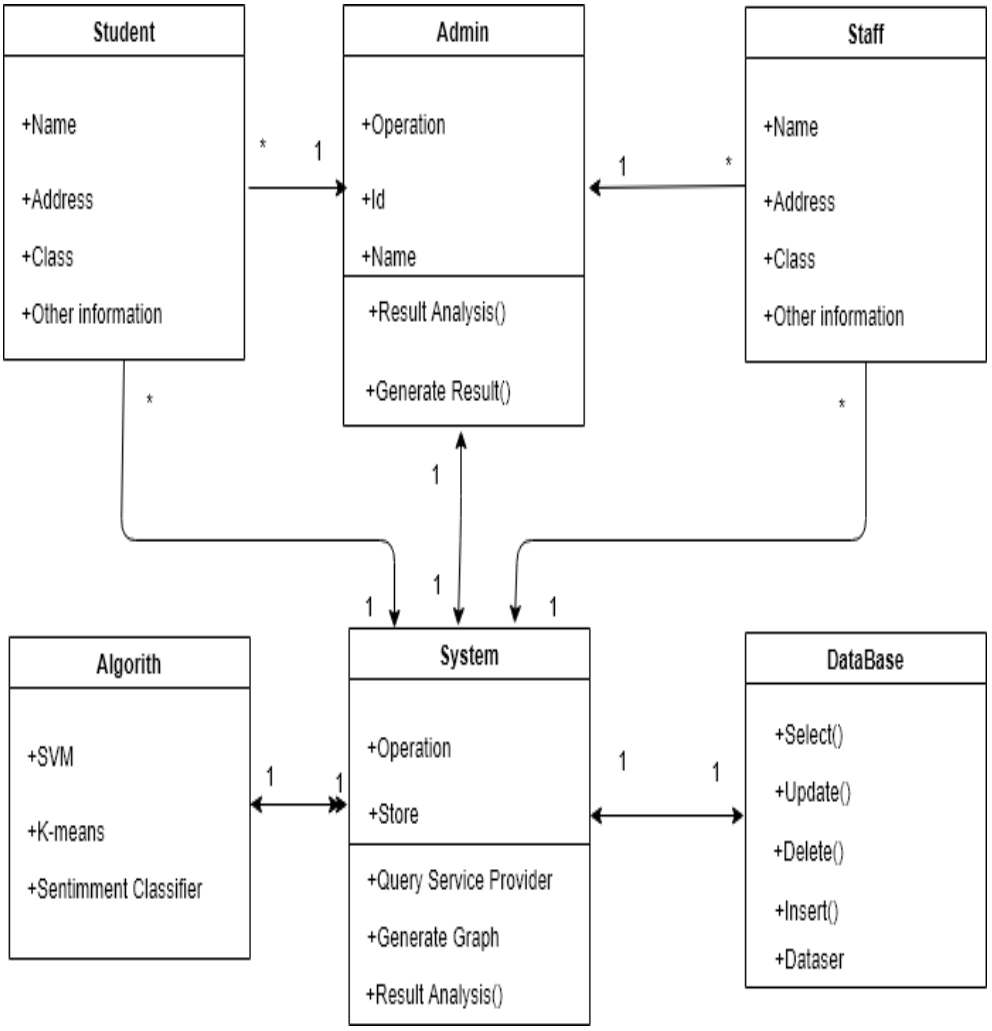


Figure 7.9: Class Diagram

7.6.5 Activity Diagram

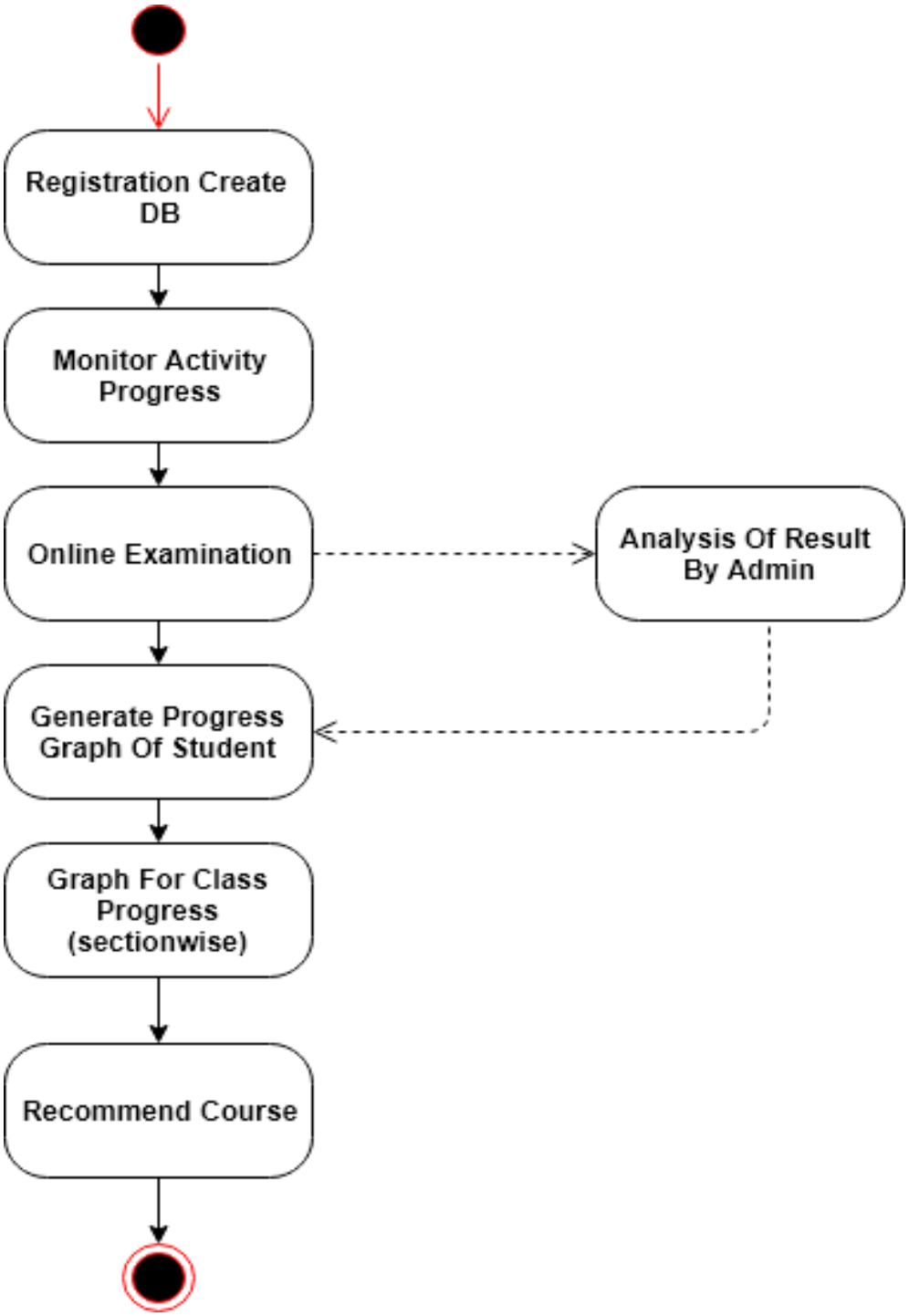


Figure 7.10: Activity Diagram

7.6.6 Component Diagram

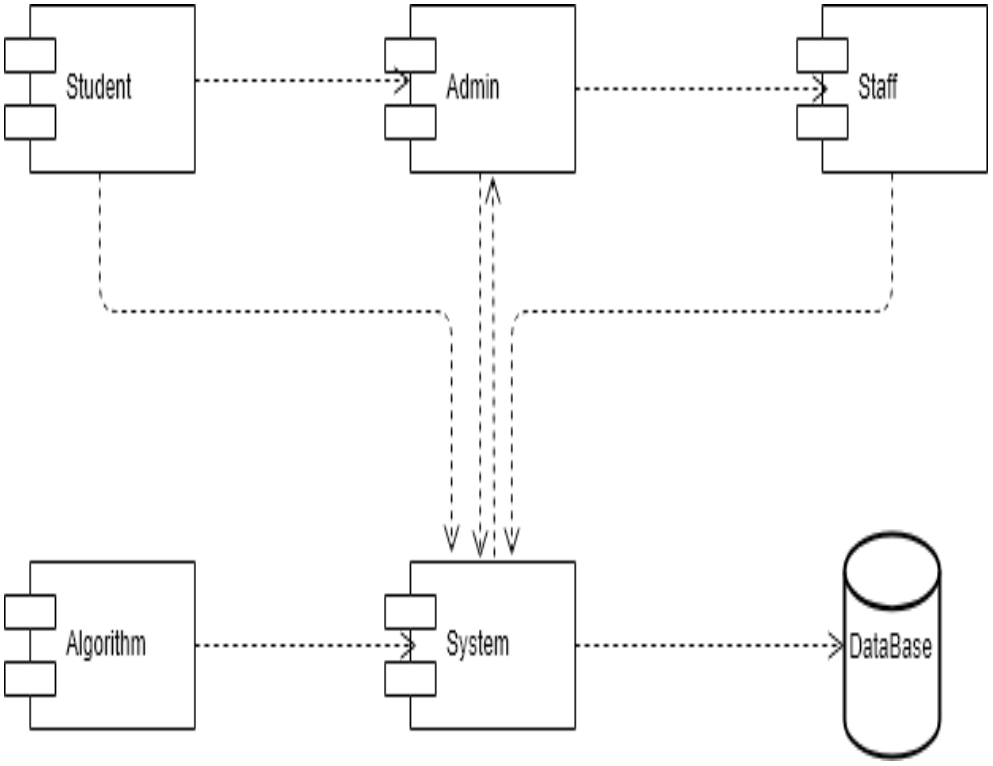


Figure 7.11: Component Diagram

7.6.7 Deployment Diagram:

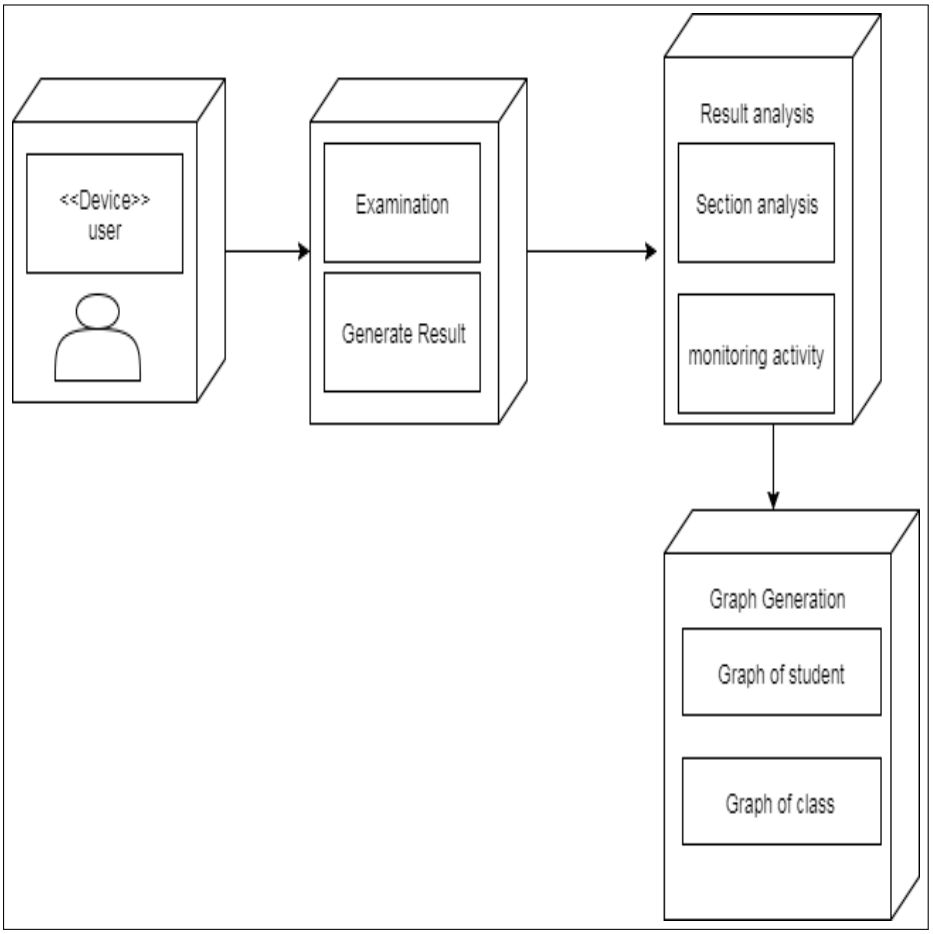


Figure 7.12: Deployment diagram

CHAPTER 8

SOFTWARE TESTING

8.1 TYPE OF TESTING

Types of Testing:

Along with the type of testing also mention the approach to be followed for the testing, that is, Manual Testing or Automated Testing. Use Automated Testing Plan forplanning automation activities in details. The different types of testing that may be carried out in the project are as follows:

- **Unit Testing:**

  Individual components are tested independently to ensure their quality. The focus is to uncover errors in design and implementation, including

  - Data structure in component
  - Program logic and program structure in a component
  - Component interface
  - Functions and operations of a component

- **Integration Testing :**

  A group of dependent components are tested together to ensure their quality of their integration unit. This approach is to do incremental integration to avoid bigbang problem. That is when the entire program is put together from all units and tested as a whole. The big-bang approach usually results in chaos which incremental integration avoids. Incremental integration testing can be done in two different way top down and bottom up. Then there is also the possibility of regression integration.

  The top down integration is when modules are integrated by moving downwards through the control hierarchy, beginning with the main control module. Modules subordinate to the main control module are incorporated into main structure n either depth-first or breadth-first manner. The top down integration verifies major controls or decision points early in the test process. If major control problems do exist, early recognition is essential. Bottom-up integration testing begins construction and testing with the lowest levels in the program structure. Because modules are integrated from the bottom-up,processing required for modules subordinate to a given level is always available and the need for test

stubs is eliminated.

The focus is to uncover errors in:

- Design and construction of software architecture
- Integrated functions or operations at sub-system level
- Interfaces and interaction and/or environment integration

- **System Testing :**

The system software is tested as a whole. It verifies all elements mesh properly to make sure that all system functions and performance are achieved in the target environment. The focus areas are:

- System functions and performance
- System reliability and recoverability (recovery test)
- System behavior in the special conditions (stress and load test)
- System user operations (acceptance test/alpha test)
- Hardware and software integration collaboration
- Integration of external software and the system.

- **Validation Testing:**

Validation can be defined in many ways, but a simple definition is that succeeds when software functions in a manner that can be reasonably expected by the customer. Software validation is achieved through a series of black-box tests that demonstrate conformity with requirements. A test plan outlines the classes of tests to be conducted and a test procedure defines specific test cases that will be used to demonstrate conformity with requirements. Both the plan and procedure are designed to ensure that all functional requirements are satisfied, all behavioral characteristics are achieved, all performance requirements are attained, documentation is correct, and human engineered and other requirements are met.

- **White Box Testing:**

  White-box test design allows one to peek inside the box, and it focuses specifically on using internal knowledge of the software to guide the selection of test data. Synonyms for white-box include: structural, glass-box and clear-box.

  White box testing is much more expensive than black box testing. It requires the source code to be produced before the tests can be planned and is much more laborious in the determination of suitable input data and the determination if the software is or is not correct. This testing is concerned only with testing the software product; it cannot guarantee that the complete specification has been implemented.

- **Black Box Testing:**

  Black-box test design treats the system as a black-box, so it doesnt explicitly use knowledge of the internal structure. Black-box test design is usually described as focusing on testing functional requirements. Synonyms for black box include: behavioral, functional, opaque-box, and closed-box. Black box testing is concerned only with testing the specification; it cannot guarantee that all parts of the implementation have been tested. Thus black box testing is testing against the specification and will discover faults of omission, indicating that part of the specification has not been fulfilled.

- **GUI Testing:**

  Graphical User Interface (GUIs) present interesting challenges for software engineers. Because of reusable components provided as part of GUI development environments, the creation of the user interface has become less time consuming and more precise. But, the same time, the complexity of GUIs has grown, leading to more difficulty in the design and execution of the test cases. Because many modern GUIs have the same look and same feel, a series of test cases can be derived.

8.2 TEST CASES TEST RESULTS

Software to be tested:

After implementation of project software will be tested by tester.

Test Cases

Testing of project problem statement using generated test data (using mathematical models, GUI, Function testing principles, if any) selection and appropriate use of testing tools, testing of UML diagrams reliability.

- **Module-ID:-01**

  Modules to be tested:-Registration

  1. Enter the case insensitive Username click on Submit button.

  Expected: It should display error.

  2. Enter the case sensitive Username click on Submit button.

  Expected: It should accept.

  3. Enter the case insensitive Password click on Submit button.

  Expected: It should display error.

  4. Enter the case sensitive Password click on Submit button.

  Expected: It should accept.

  5. Enter the case insensitive Mobile Number click on Submit button.

  Expected: It should display error.

  6. Enter the case sensitive Mobile Number click on Submit button.

  Expected: It should accept.

  7. Enter the wrong address and click on Submit button.

  Expected: It should display error.

  8. Enter the correct address and click on Submit button.

  Expected: It should accept.

- **Module-ID:-2**

  Modules to be tested:- Login

  1. Enter the correct username and wrong password click on Submit button.

  Expected: It should display error.

  2. Enter the wrong username and correct password and click on Submit button.

  Expected: It should display error.

  3. Enter the correct username and password and click on Login button.

  Expected: It should display welcome page.

Test Case_ID	Description	Test case I/P	Actual Result	Expected result	Test case criteria (P/F)
101	Enter the case insensitive Username click on Submit button.	Username	Error comes	Error Should come	P
102	Enter the case sensitive Username click on Submit button.	Username	Accept	Accept Username	P
201	Enter the case insensitive Password click on Submit button.	Password	Error comes	Error Should come	P
202	Enter the case sensitive Password click on Submit button	Password	Accept	Accept	P
301	Enter the case insensitive Mobile Number click on Submit button	Mobile Number	Error comes	Error Should come	P
302	Enter the case sensitive Mobile Number click on Submit button.	Mobile Number	Accept	Accept	P

4. After login with valid credentials click on back button.

Expected: The page should be expired.

5. After login with valid credentials copy the URL and paste in another browser.

Expected: It should not display the user's welcome page.

6. Check the password with Lower case and upper case.

Expected: Password should be case sensitive.

Test Case_ID	Description	Test case I/P	Actual Result	Expected result	Test case criteria (P/F)
001	Enter the correct username and wrong password click onLogin button.	Username Password	Error comes	Error Should come	P
002	Enter the wrong username and correct password click onLogin button,	Username Password	Error comes	Error Should come	P
003	Enter the correct username and password and click on Login button.	Username Password	Accept	Accept	P

- **Module-ID:-3**

  Modules to be tested:- connect DB

  1. Enter the wrong Username and click on Submit button.

  Expected: It should display error.

  2. Enter the correct Username and click on Submit button.

  Expected: It should accept.

  3. Enter the wrong Host and click on Submit button.

  Expected: It should display error.

  4. Enter the correct Host and click on Submit button.

  Expected: It should accept.

  5. Enter the wrong Database type and click on Submit button.

  Expected: It should display error.

  6. Enter the correct Database type and click on Submit button.

  Expected: It should accept.

Test Case_ID	Test Case	Test case I/P	Actual Result	Expected result	Test case criteria (P/F)
001	Enter the wrong username click on Submit button	Username	Error comes	Error Should come	P
002	Enter the correct username click onLogin button	Username	Accept	Accept	P
001	Enter the wrong Host and click on Submit button.	Host	Error comes	Error Should come	P
002	Enter the correct Host and click on Submit button.	Host	Accept	Accept	P
001	Enter the wrong Database type and click on Submit button.	Database	Error comes	Error Should come	P
002	Enter the correct Database type and click on Submit button	Database	Accept	Accept	P

8.3 TEST RESULTS

Unit Testing:

It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. This is a structural testing, that relies on knowledge of its construction and is invasive.

Unit tests perform basic tests at component level and test a specific business process, applica-

tion, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

CHAPTER 9

RESULTS AND DISCUSSIONS

9.1 OUTCOME

In proposed to predict the performance of students in an academic organization. Present studies shows that academic performances of the students are primarily dependent on their past performances. Our investigation confirms that past performances have indeed got a significant influence over students' performance. Further, we confirmed that the performance of neural networks increases with increase in dataset size. Additionally, this work will also impact curriculum design in degree programs and education policy design in general. Future work includes Extending the performance prediction to elective courses and using the prediction results to recommend courses to students.

Advantages of Proposed System:

- Reducing Time
- Less Manpower
- More Secure
- Easy To Use
- Easy to available course material.
- Easy to learn.
- Easily communicate with staff members and solve query.
- Easy to show track regarding to feature career.

CHAPTER 10

DEPLOYMENT AND MAINTENANCE

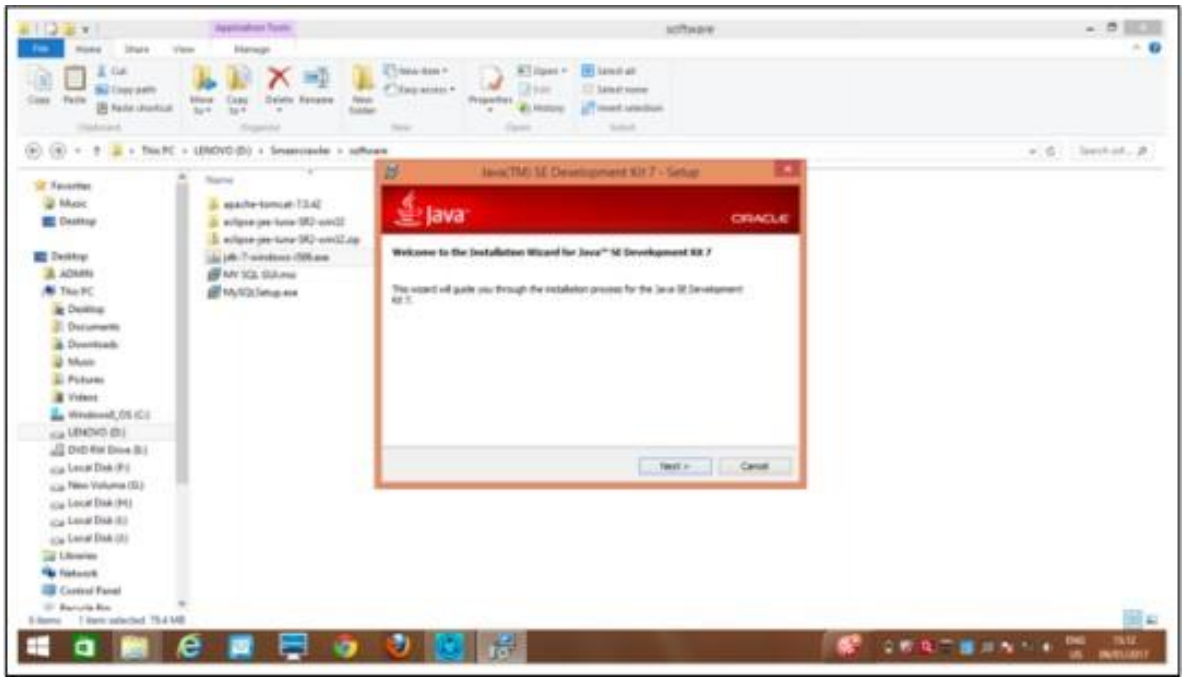
10.1 INSTALLATION AND UN-INSTALLATION

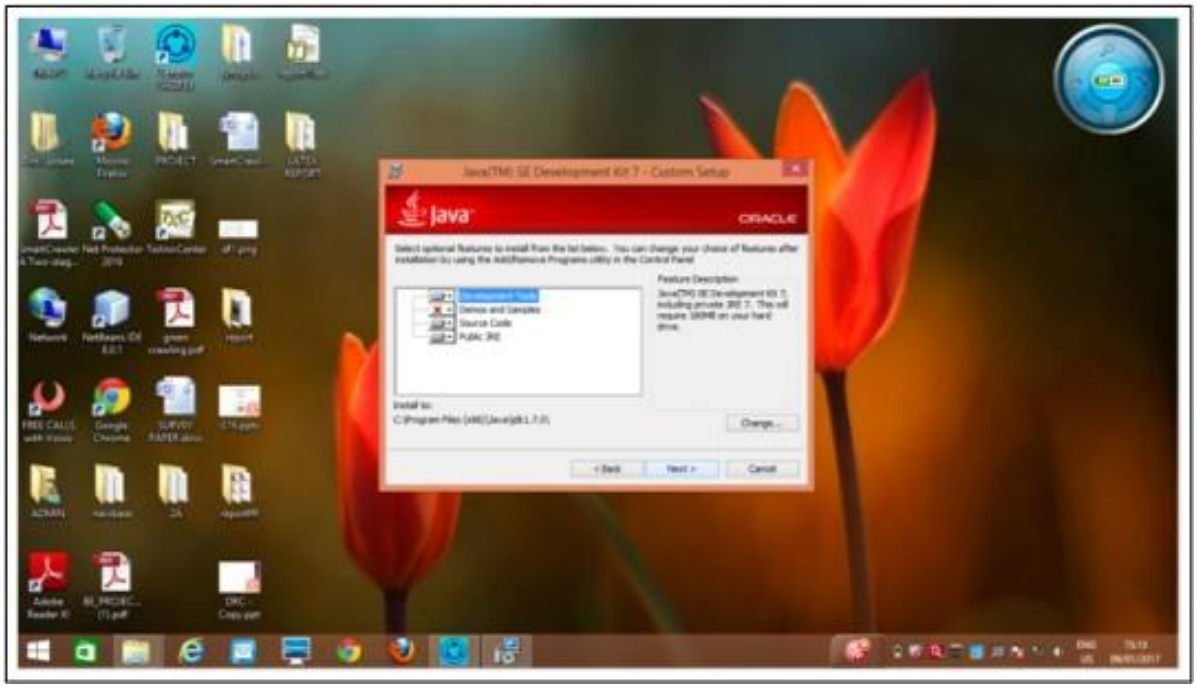
10.1.1 Exlipse

Eclipse Installation:

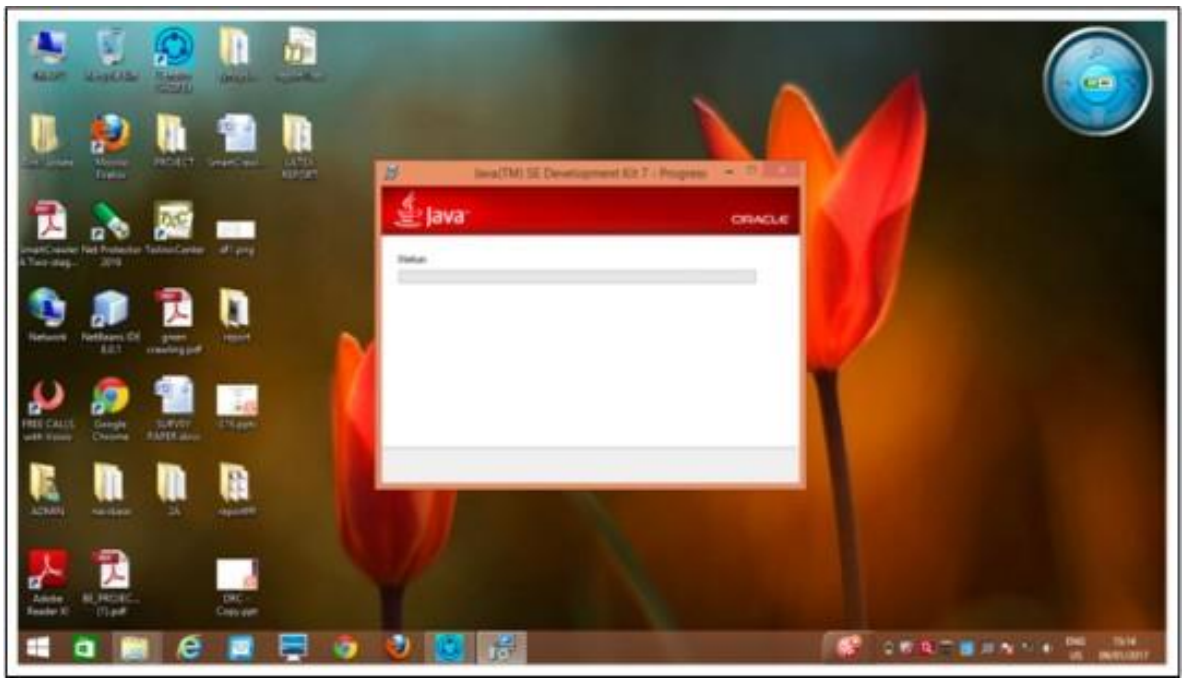
How To Install JDK on Windows Step 0: Un-Install Older Version(s) of JDK/JRE

- Step 1: Download JDK.
- Step 2: Install JDK and JRE.
- Step 3: Include JDKs bin Directory in the PATH.
- Step 4: Verify the JDK Installation.
- Step 5: Write a Hello-World Java Program.
- Step 6: Compile and Run the Hello-World Java Program.




Java
ORACLE


Java
ORACLE



10.1.2 XAMPP Installation Steps

Steps:

Step1: Click on the download link for XAMPP.

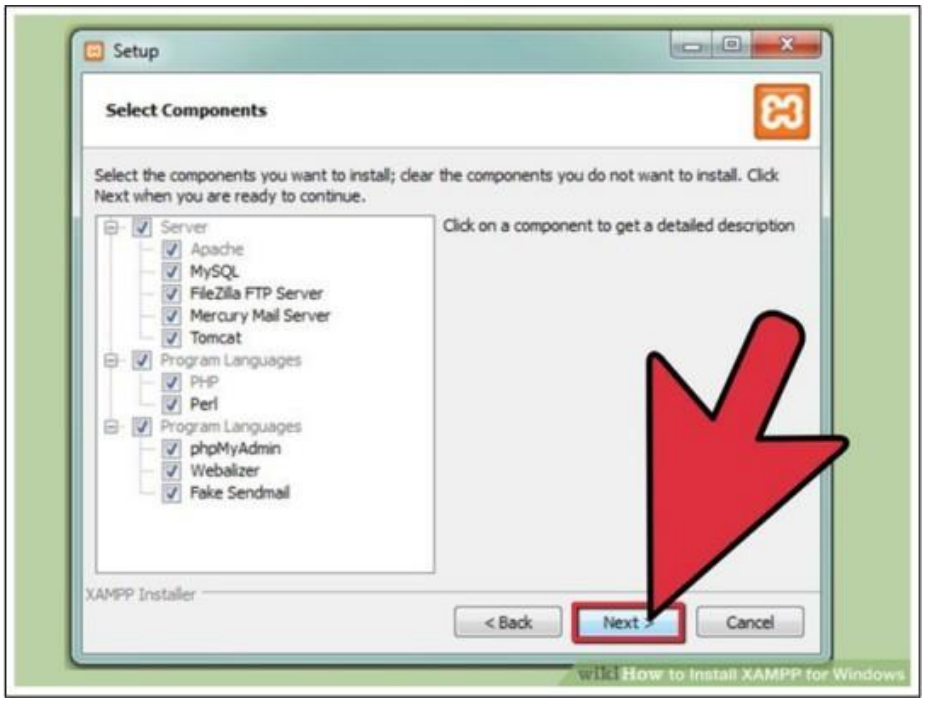
Step2:When prompted for the download, click Save and wait for your download to finish.

Step 3:Open CD or DVD drive from My Computer.

Step 4:Accept the default settings.

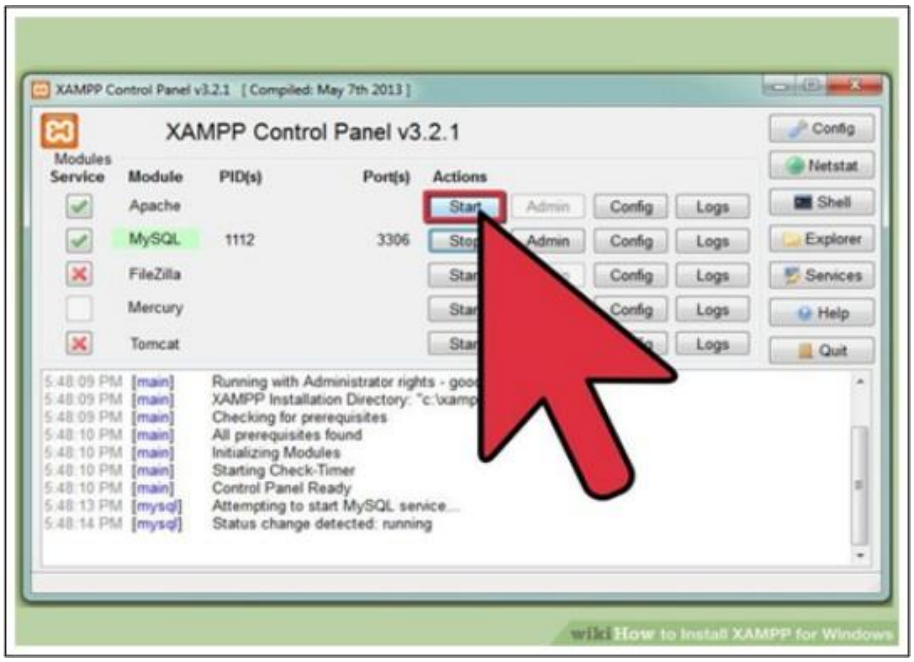
Step 5: When your installation is complete, exit the command window by typing x on the command line.

Step 6:Start the XAMPP Control Panel.


Setup
Select Components
Select the components you want to install; clear the components you do not want to install. Click Next when you are ready to continue.
Server
Apache
MySQL
FileZilla FTP Server
Mercury Mail Server
Tomcat
Program Languages
PHP
Perl
Program Languages
phpMyAdmin
Webalizer
Fake Sendmail
Click on a component to get a detailed description
XAMPP Installer
< Back
Next >
Cancel
wikiHow to Install XAMPP for Windows


Setup
Completing the XAMPP Setup Wizard
Setup has finished installing XAMPP on your computer.
Do you want to start the Control Panel now?
bitnami
< Back
Finish
wikiHow to Install XAMPP for Windows


Open
Scan XAMPP.pr
Open with
Share with
Add to archive...
Add to "XAMPP.rar"
Compress and email...
Compress to "XAMPP.rar" and email
wikiHow to Install XAMPP for Windows


XAMPP Control Panel v3.2.1 [Compiled: May 7th 2013]
XAMPP Control Panel v3.2.1
Modules
Service
Module
PID(s)
Port(s)
Actions
Apache
MySQL
1112
3306
FileZilla
Mercury
Tomcat
Start
Admin
Config
Logs
Netstat
Shell
Explorer
Services
Help
Quit
5:48:09 PM [main] Running with Administrator rights - good
5:48:09 PM [main] XAMPP Installation Directory: "c:\xamp
5:48:09 PM [main] Checking for prerequisites
5:48:10 PM [main] All prerequisites found
5:48:10 PM [main] Initializing Modules
5:48:10 PM [main] Starting Check-Timer
5:48:10 PM [main] Control Panel Ready
5:48:13 PM [mysql] Attempting to start MySQL service...
5:48:14 PM [mysql] Status change detected: running
wikiHow to Install XAMPP for Windows

CHAPTER 11

SUMMARY AND CONCLUSION

11.1 CONCLUSION

Present studies shows that academic performances of the students are primarily dependent on their past performances. Our investigation confirms that past performances have indeed got a significant influence over students' performance. Further, we confirmed that the performance of SVM, K-Means increases with increase in dataset size. Machine learning has come far from its nascent stages, and can prove to be a powerful tool in academia. In the future, applications similar to the one developed, as well as any improvements thereof may become an integrated part of every academic institution. Additionally, this work will also impact curriculum design in degree programs and education policy design in general. Future work includes extending the performance prediction to elective courses and using the prediction results to recommend courses to students.

11.2 FUTURE WORK

- Future work includes extending the performance prediction to elective courses and using the prediction results to recommend courses to students.
- Future work includes extending the performance prediction to elective courses and using the prediction results to recommend courses to students.
- This project can be used in any organization, college as analysis purpose.

11.3 APPLICATIONS

1. Exam systems in schools.
2. In future can use in college with some changes.
3. Any institute or company can register their various types of certificate/non-certificate programs and conduct an online examination for the same.
4. User can select the company, its program, exam schedule in order to give his exam at the selected center.

CHAPTER 12

REFERENCES

1 The White House, Making college affordable, https://www:whitehouse:gov/issues/education/higher-education/making-college-affordable, 2016.

2 Complete College America, Four-year myth: Making collegemore affordable, http://completecollege:or ontent/uploads/2014/11/4-Year-Myth:pdf, 2014.

3 H. Cen, K. Koedinger, and B. Junker, Learning factors analysisageneral method for cognitive model evaluation and improvement, inInternational Conference on Intelligent Tutoring Systems. Springer,2006, pp. 164175.

4 M. Feng, N. Heffernan, and K. Koedinger, Addressing the assessmentchallenge with an online system that tutors as it assesses, User Modelingand User-Adapted Interaction, vol. 19, no. 3, pp. 243266, 2009.

5 H.-F. Yu, H.-Y. Lo, H.-P. Hsieh, J.-K.Lou, T. G. McKenzie, J.-W.Chou, P.-H.Chung, C.-H. Ho, C.-F. Chang, Y.-H.Wei et al., Featureengineering and classifier ensemble for kdd cup 2010, in Proceedingsof the KDD Cup 2010 Workshop, 2010, pp. 116.

6 Z. A. Pardos and N. T. Heffernan, Using hmms and bagged decisiontrees to leverage rich features of user and skill from an intelligenttutoring system dataset, Journal of Machine Learning Research W CP, 2010.

7 Y. Meier, J. Xu, O. Atan, and M. van der Schaar, Personalized gradeprediction: A data mining approach, in Data Mining (ICDM), 2015IEEE International Conference on. IEEE, 2015, pp. 907912.

8 C. G. Brinton and M. Chiang, Mooc performance prediction viaclickstream data and social learning networks, in 2015 IEEE Conferenceon Computer Communications (INFOCOM). IEEE, 2015, pp. 22992307.

9 KDD Cup, Educational data minding challenge, https://pslcdatashop:web: cmu:edu/KDDCup/, 2010.13

10 C. Marquez-Vera, C. Romero, and S. Ventura, Predicting school failureusing data mining, in Educational Data Mining 2011, 2010.

11 Y.-h. Wang and H.-C. Liao, Data mining for adaptive learning in atesl-based e-learning system, Expert Systems with Applications, vol. 38,no. 6, pp. 64806485, 2011.

12 N. Thai-Nghe, L. Drumond, T. Horvath, L. Schmidt-Thiemeet al.,Multi-relational factorization models for predicting student performance,in Proc. of the KDD Workshop on Knowledge Discovery inEducational Data.Citeseer, 2011.

13 A. Toscher and M. Jahrer, Collaborative filtering applied to educationaldata mining, KDD cup, 2010.

14 R. Bekele and W. Menzel, A bayesian approach to predict performanceof a student (bapps): A case with ethiopian students, algorithms, vol. 22,no. 23, p. 24, 2005.

15 N. Thai-Nghe, T. Horvath, and L. Schmidt-Thieme, Factorization modelsfor forecasting student performance, in Educational Data Mining2011, 2010.

16 Y. Meier, J. Xu, O. Atan, and M. van der Schaar, Predicting grades,IEEE Transactions on Signal Processing, vol. 64, no. 4, pp. 959972,Feb 2016.

17 N. Cesa-Bianchi and G. Lugosi, Prediction, learning, and games.Cambridge university press, 2006.

18 Y. Koren, R. Bell, C. Volinskyet al., Matrix factorization techniquesfor recommender systems, Computer, vol. 42, no. 8, pp. 3037, 2009.

19 R. Salakhutdinov and A. Mnih, Probabilistic matrix factorization, inNIPS, vol. 20, 2011, pp. 18.

ANNEXURE A

TEST PLAN

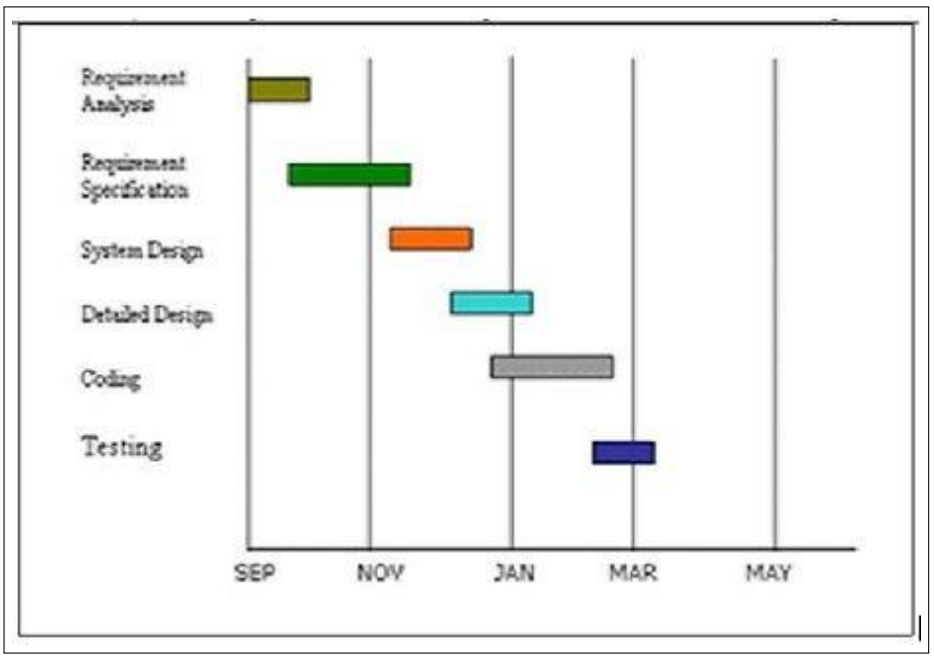


Figure A.1: Test Plan

ANNEXURE B

PAPER PUBLICATION

Smart Monitoring and Predicting Students Skill using Machine Learning Approach

Supriya Shinde[1], Priyanka Tribhuvan[2], Rutuja Kawade[3], Bhakti Kadu[4], Prof. S.D. Jondhale[5]
[1,2,3,4] Students & [5]Asst. Prof. of Department of Computer Engineering,
PREC COE, Loni, Ahmednagar

Abstract: *Data Mining in Educational and learning Analysis studies have shown as exciting areas of analysis that can be advance useful expertise from academic databases for much more uses inclusive of forecasting student's success. The capability to expect predict overall performance of candidates can be useful for activity in contemporary instructional structures. Current techniques have used capabilities that are by and large relevant to academic overall performance, own family earnings and own family property; at the same time as functions inclusion to family consumption and candidates personal data are usually unnoticed. On this survey paper, an attempt is built to analyze introductory function sets by way of gather and retaining college student's information from distinctive universities. Gaining knowledge of analysis, selective and creative class representations are carried out to predict in case a student will be capable to complete his degree education or not as well as the overall performance of students in schooling discipline and after education prediction of tune.*

Index Terms**:** *Data Mining, Machine Learning, Personalized Education, Tracking Students Performance, Course Prediction, and Recommendation System.*

I. INTRODUCTION

To address the aforementioned demanding situations, we proposed a unique algorithm for predicting scholar's or candidates achievements in study field applications given his/her current educational data. In Proposed research indicates that academic performances of the scholars are frequently depending on their past performances. Our investigation confirms that previous performances have indeed got an enormous have an effect on over college students' performance. Further, we showed that the performance of SVM increases with boom in dataset length.

System will include the tracking of unique facts of a student's concerning his academics and curricular activity and might expect the right gaining knowledge of courses the use of an algorithm over the information tracked assembly the ambition or the purpose for a student.

In the previous system, college conducts examination manually. It has such a lot of issues. The prevailing structures are very time ingesting. It's far hard to research the examination manually. Outcomes aren't unique as calculation and reviews are carried out manually. Result processing after summation of study takes extra time as it's far carried out manually. So we introduce a Preschool exam Portal gadget, that's completely automated. Existing machine is a huge man energy procedure and is hard to implement. It gives a smooth to apply environment for each test Conductors and students appearing for study. the primary objective of Preschool examination Portal device is to provide all the features that an examination machine should have, with the" interfaces that don't Scare it's Users!"

II. LITERATURE SURVEY

1. **Multi-Relational Factorization Models for Predicting Student Performance.**

   Author: Nguyen Thai-Nghe, Lucas Drumond, Tomas Horvath, and Lars Schmidt-Thieme, University of Hildesheim

   In this paper we recommend to make the most such a couple of relationships by way of the use of multi-relational MF techniques. Experiments on 3 large datasets show that the proposed technique can enhance the prediction outcomes.

Predicting pupil performance (PSP) is the problem of predicting how well a pupil will carry out on a given project. It has won greater attention from the educational data mining network currently. previous works display that exact outcomes may be done by way of casting the PSP to rating prediction trouble in recommender structures, wherein college students, responsibilities and performance rankings are mapped to customers, gadgets and ratings respectively.

2. **From Data to Knowledge to Action: Enabling Personalized Education**

   Author: Beverly Park Woolf, Ryan Baker, Erwin P. Gianchandani

   We describe how statistics analytics procedures have the capability to dramatically advance training for every scholar and to enhance the manner we train our children. The net, sensible environments, and wealthy interfaces (consisting of sensors) permit us to seize a good deal extra facts approximately novices than ever earlier than and the quantities of information are growing at a hastily accelerating charge.

3. **How personalized learning unlocks student Success.**

   Author: Nazeema Alli, Rahim Rajan, and Greg Ratliff.

   EDUCAUSE assessment is the overall-hobby, bimonthly magazine posted by way of EDUCAUSE. With a print ebook base of 22000, EDUCAUSE review is sent to EDUCAUSE member representatives in addition to to presidents/chancellors, senior educational and administrative leaders, non-IT team of workers, school in all disciplines, librarians, and corporations. It takes a vast examine present day developments and developments in statistics technology, what these mean for higher education, and how they will have an effect on the university/university as a whole.

4. **Personalized grade prediction: A data mining approach, in Data Mining (ICDM)**

   Authors: Yannick Meier, Jie Xu, Onur Atan, Mihaela van der Schaar

   This paper proposes a set of rules that predicts the final grade of each scholar in a category. The algorithm learns on line what is the most desirable prediction and time to issue a prediction based totally on beyond history of students' overall performance in a route. We derive exhibit the performance of our set of rules on a dataset received based on the overall performance of approximately 700 undergraduate students who've taken an introductory virtual signal processing over the past 7 years.

5. **MOOC performance prediction via click stream data and social learning networks.**

   Authors: Christopher G. Brinton; Mung Chiang

   We look at student overall performance prediction in massive Open on-line publications (MOOCs), wherein the objective is to expect whether a person might be correct on First strive (CFA) in answering a question. In doing so, we broaden novel strategies that leverage behavioural statistics amassed by means of MOOC platforms. the usage of video-watching clickstream data from certainly one of our MOOCs, we first extract summary portions (e.g., fraction played, number of pauses) for every user-video pair, and display how positive intervals/units of values for these behaviours quantify that a couple is much more likely to be CFA or no longer for the corresponding question.

6. **Predicting Student Performance Using Personalized Analytics.**

   Authors: Asmaa Elbadrawy; Agoritsa Polyzou; Zhiyun Ren; Mackenzie Sweeney; George Karypis; Huzefa Ra

   To assist resolve the continued trouble of pupil retention, new anticipated overall performance-prediction techniques are needed to facilitate degree making plans and decide who is probably liable to failing or losing a class. personalized multi-regression and matrix factorization techniques primarily based on recommender systems, initially developed for

e-commerce applications, accurately forecast students' grades in future publications as well as on in-magnificence checks.

7. **Combining University Student Self-Regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance**

   Authors: Abelardo Pardo ; Feifei Han ; Robert A. Ellis.

   On this paper authors explore how to integrate statistics about self-regulated mastering abilities with observable measures of on-line hobby in a blended learning course to increase predictive skills of scholar educational performance for the purposes of informing coaching and project layout. A case look at in a direction with 145 students showed that the version of the scholars' very last score for his or her path is better explained whilst factors from both methods are considered. The effects point to the potential of adopting a combined use of self-record and discovered records to benefit a more comprehensive knowledge of successful college scholar learning.

8. **Behaviour-Based Grade Prediction for MOOCs Via Time Series Neural Networks.**

   Authors: Tsung-Yen Yang; Christopher G. Brinton; Carlee Joe-Wong; Mung Chiang

   Authors present a singular technique for predicting the evolution of a scholar's grade in big open online publications (MOOCs). Performance prediction is mainly difficult in MOOC settings because of in step with-scholar assessment response sparsity and the want for personalized fashions. Our approach overcomes these demanding situations by means of incorporating every other, richer shape of statistics collected from every scholar-lecture video-looking clickstreams-into the gadget getting to know feature set, and the usage of that to educate a time collection neural network that learns from both earlier overall performance and clickstream data.

9. **Multi-relational Factorization Models for Student Modelling in Intelligent Tutoring Systems.**

   Authors: Nguyen Thai-Nghe; Lars Schmidt-Thieme.

   In this work, authors suggest the use of multi-relational factorization method, which has been successfully implemented in recommender structures region, for pupil modeling inside the wise Tutoring structures. Experiments on big real global records sets display that the proposed technique can enhance the prediction effects and will be used for scholar modeling.

10. **Evaluating student performance using fuzzy inference system in fuzzy ITS.**

    Authors: Pooja Asopa; Sneha Asopa; Nisheeth Joshi; Iti Mathur

    Authors can be used in modeling the unsure conduct of various complex problems and additionally for predicting the uncertainty degree of the students. Systems which have immense skills to offer its newcomers with step-through-step instructions as in keeping with their personal learning fame the usage of system based instructions are known as clever Tutoring machine (ITS). The ITS device which has fuzzy characteristics can be called as fuzzy ITS.

11. **Using neural networks to predict student's performance.**

    Authors: T. Wang ; A. Mitrovic

    In this strategy very frequently effects in issues which might be too clean/tough for the pupil. Here we suggest an intelligent trouble-selection agent, which identifies the precise problem for a pupil in levels. It firstly predicts the quantity of mistakes the pupil will make on a set of troubles, after which inside the second level makes a decision on an appropriate problem for the scholar.

12. **Predicting Student Performance in an ITS Using Task-Driven Features.**

Authors: Ritu Chaturvedi; C. I. Ezeife

This studies proposes a highly accurate version that predicts student fulfillment in assigned responsibilities with a 96% accuracy by the use of functions that are better knowledgeable now not most effective about college students in phrases of the two elements f1 and f2 noted above, however additionally at the assigned undertaking itself (e.g. venture's problem level). with the intention to accomplish this, an instance advice machine (ERS) is designed with a first-class-grained scholar version (to symbolize pupil data) and an excellent-grained area version (to symbolize domain sources including tasks).

13. **Predicting Student Actions in a Procedural Training Environment.**

Authors: Diego Riofrío-Luzcando; Jaime Ramírez; Marta Berrocal-Lobo

This paper authors provides a collective scholar version, that's constructed from beyond pupil logs. These logs are first grouped into clusters. Then, a prolonged automaton is created for every cluster based at the sequences of occasions discovered within the cluster logs. The primary goal of this model is to predict the actions of recent students for improving the tuition comments furnished by way of a clever tutoring system.

14. **Optimizing the Performance of Educational Web Services.**

Authors: Mark Floryan; Beverly Park Woolf.

Authors define web service granularity in terms of the amount of data that can be retrieved from a service in a single request on average. This is important because developers cannot predict if students will be using state of the art hardware. Thus, service-oriented architectures (SOA) with fine service granularity can minimize network communication and allow server machines to perform more work for applications.

15. **Student performance analysis system (SPAS)**

Authors: Chew Li Sa; Dayang Hanani bt. Abang Ibrahim; Emmy Dahliana Hossain; Mohammad bin Hossin

In this system authors said student performance prediction through the rules generated via data mining technique. The data mining technique used in this project is classification, which classifies the students based on students' grade.

III. PROPOSED METHODOLOGY

In proposed to be expecting the performance of college students in a academic organization. Present studies indicates that academic performances of scholars are frequently depending on their beyond performances. Our investigation confirms that beyond performances has certainly been given an extensive having an effect on over students' performance. Similarly, we showed that the performance of neural networks increases with increase in dataset size. Additionally, proposed work may still impact curriculum layout in study record & education coverage layout in widespread. Coming work consists of approaching the achievement prediction to optional publications and the use of the prediction results to advise publications to college students.

IV. SYSTEM ARCHITECTURE OVERVIEW

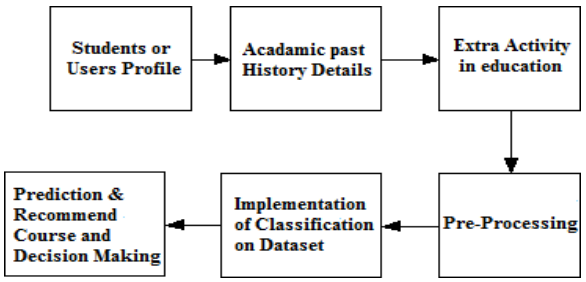


Fig. 1 System Architecture (Block Diagram)

V. CONCLUSION

Present studies indicate that academic performances of the scholars are primarily dependent on their past performances. Our research confirms that past performances have indeed been given a great affect over students' performance. In addition, we confirmed that the performance of neural networks will increase with boom in dataset length. Machine learning gaining knowledge of has come far from its nascent ranges, and can prove to be a powerful device in academia. In the future, programs similar to the one evolved, in addition to any improvements thereof may additionally emerge as an included part of every instructional organization. This challenge may be utilized in any company, university as evaluation reason.

REFERENCES

[1] H. Cen, K. Koedinger, and B. Junker, Learning factors analysis a general method for cognitive model evaluation and improvement, in International Conference on Intelligent Tutoring Systems. Springer, 2006, pp. 164175.

[2] M. Feng, N. Heffernan, and K. Koedinger, Addressing the assessment challenge with an online system that tutors as it assesses, User Modeling and User-Adapted Interaction, vol. 19, no. 3, pp. 243266, 2009.

[3] H.-F. Yu, H.-Y. Lo, H.-P. Hsieh, J.-K.Lou, T. G. McKenzie, J.-W.Chou, P.-H.Chung, C.-H. Ho, C.-F. Chang, Y.-H.Wei et al., Feature engineering and classifier ensemble for kdd cup 2010, in Proceedings of the KDD Cup 2010 Workshop, 2010, pp. 116.

[4] Z. A. Pardos and N. T. Heffernan, Using hmms and bagged decision trees to leverage rich features of user and skill from an intelligent tutoring system dataset, Journal of Machine Learning Research W CP, 2010.

[5] Y. Meier, J. Xu, O. Atan, and M. van der Schaar, Personalized grade prediction: A data mining approach, in Data Mining (ICDM), 2015IEEE International Conference on. IEEE, 2015, pp. 907912.

[6] C. G. Brinton and M. Chiang, Mooc performance prediction via click stream data and social learning networks, in 2015 IEEE Conference on Computer Communications (INFOCOM). IEEE, 2015, pp. 22992307.

[7] C. Marquez-Vera, C. Romero, and S. Ventura, Predicting school failure using data mining, in Educational Data Mining 2011, 2010

Journal of Emerging Technologies and Innovative Research

An International Open Access Journal

www.jetir.org | editor@jetir.org

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Supriya Shinde

In recognition of the publication of the paper entitled

Smart Monitoring and Predicting Students Skill using Machine Learning Approach

Published In JETIR (www.JETIR.org) ISSN UGC Approved (Journal No: 63975) & 5.87 Impact Factor

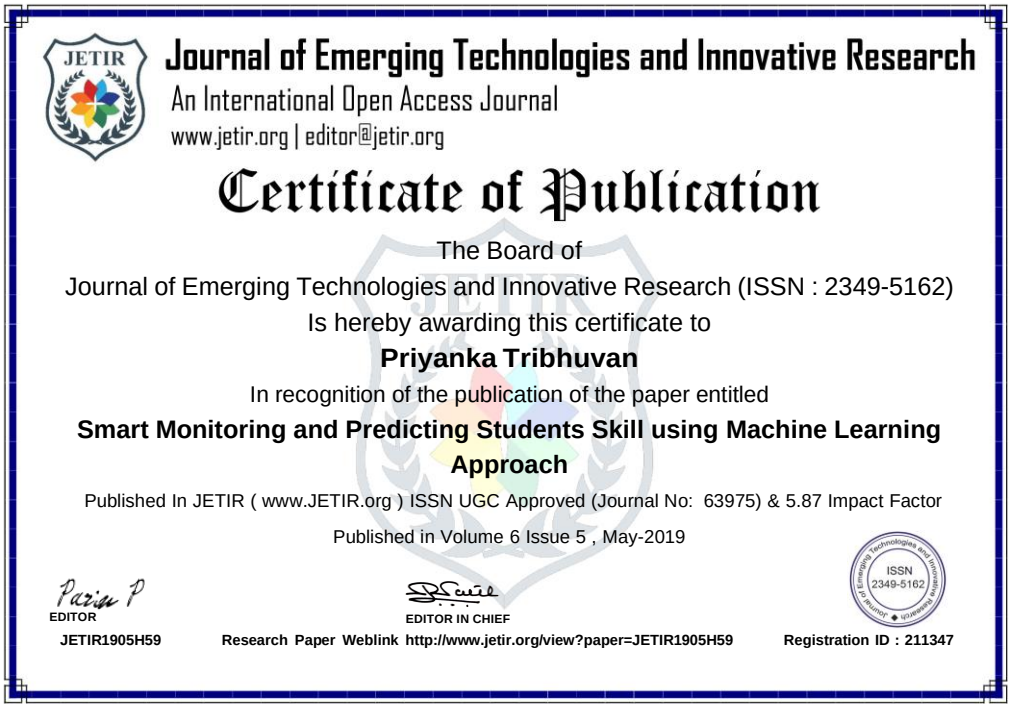
Published in Volume 6 Issue 5 , May-2019

EDITOR

EDITOR IN CHIEF

ISSN 2349-5162

JETIR1905H59 Research Paper Weblink http://www.jetir.org/view?paper=JETIR1905H59 Registration ID : 211347



Journal of Emerging Technologies and Innovative Research

An International Open Access Journal

www.jetir.org | editor@jetir.org

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Priyanka Tribhuvan

In recognition of the publication of the paper entitled

Smart Monitoring and Predicting Students Skill using Machine Learning Approach

Published In JETIR (www.JETIR.org) ISSN UGC Approved (Journal No: 63975) & 5.87 Impact Factor

Published in Volume 6 Issue 5 , May-2019

EDITOR

EDITOR IN CHIEF

ISSN 2349-5162

JETIR1905H59 Research Paper Weblink http://www.jetir.org/view?paper=JETIR1905H59 Registration ID : 211347

Journal of Emerging Technologies and Innovative Research

An International Open Access Journal

www.jetir.org | editor@jetir.org

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Rutuja Kawade

In recognition of the publication of the paper entitled

Smart Monitoring and Predicting Students Skill using Machine Learning Approach

Published In JETIR (www.JETIR.org) ISSN UGC Approved (Journal No: 63975) & 5.87 Impact Factor

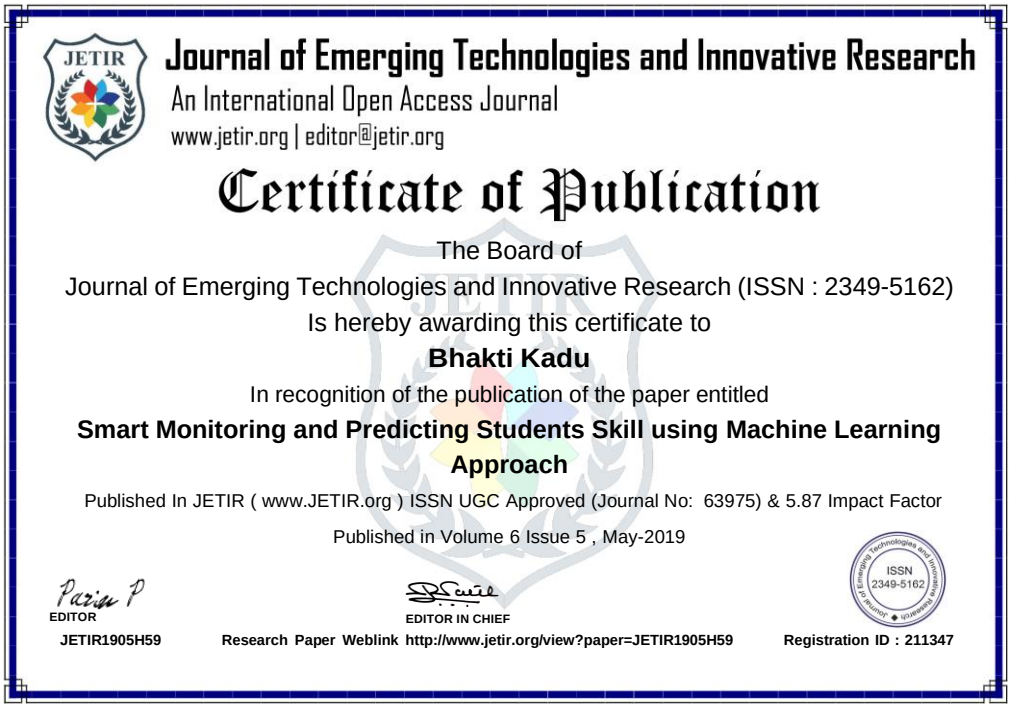
Published in Volume 6 Issue 5 , May-2019

EDITOR

EDITOR IN CHIEF

ISSN 2349-5162

JETIR1905H59 Research Paper Weblink http://www.jetir.org/view?paper=JETIR1905H59 Registration ID : 211347



Journal of Emerging Technologies and Innovative Research

An International Open Access Journal

www.jetir.org | editor@jetir.org

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Bhakti Kadu

In recognition of the publication of the paper entitled

Smart Monitoring and Predicting Students Skill using Machine Learning Approach

Published In JETIR (www.JETIR.org) ISSN UGC Approved (Journal No: 63975) & 5.87 Impact Factor

Published in Volume 6 Issue 5 , May-2019

EDITOR

EDITOR IN CHIEF

JETIR1905H59 Research Paper Weblink http://www.jetir.org/view?paper=JETIR1905H59 Registration ID : 211347

Journal of Emerging Technologies and Innovative Research

An International Open Access Journal

www.jetir.org | editor@jetir.org

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Prof. S.D. Jondhale

In recognition of the publication of the paper entitled

Smart Monitoring and Predicting Students Skill using Machine Learning Approach

Published In JETIR (www.JETIR.org) ISSN UGC Approved (Journal No: 63975) & 5.87 Impact Factor

Published in Volume 6 Issue 5 , May-2019

EDITOR

EDITOR IN CHIEF

ISSN 2349-5162

JETIR1905H59 Research Paper Weblink http://www.jetir.org/view?paper=JETIR1905H59 Registration ID : 211347

ANNEXURE C

INFORMATION OF PROJECT GROUP MEMBERS

1. Name : Kadu Bhakti A.
2. Date of Birth : 05/07/1997
3. Gender : Female
4. Permanent Address : A/P. Pathare BK. Tal. Rahata, Dist-Ahemadnagar.
5. E-Mail : bk.kadu45@gmail.com
6. Mobile/Contact No. : 9527099791

1. Name : Kawade Rutuja S.
2. Date of Birth : 12/07/1997
3. Gender : Female
4. Permanent Address : A/P. Bherdapur, Tal.Shrirampur, Dist-Ahemadnagar.
5. E-Mail : rutujakawade08@gmail.com
6. Mobile/Contact No. : 8669453437

1. Name : Shinde Supriya A.
2. Date of Birth : 30/06/1997
3. Gender : Female
4. Permanent Address : A/P. Belapur, Tal.Shrirampur, Dist-Ahemadnagar.
5. E-Mail : shindesupriya137@gmail.com
6. Mobile/Contact No. : 8149425655

1. Name : Tribhuvan Priyanka A.
2. Date of Birth : 27/10/1996
3. Gender : Female
4. Permanent Address : A/P. Moshi, Tal. Haveli, Dist. Pune
5. E-Mail : priyankatribhuvan60@gmail.com
6. Mobile/Contact No. : 8421623726