**Abstract**

All students dream to obtain a job offer in their hands before they leave their college. A placement chance predictor helps students to have an idea about where they stand and what to be done to obtain a good placement. A placement predictor is a system that could predict the possibility or the type of company a pre-final year student has chances to be placed. Thus a prediction system could help in the academic planning of an institution for future years. With the emergence of data mining and deep learning, many predictor models were introduced by analyzing the previous year student’s dataset. This paper presents a literature survey on different placement prediction models for pre-final year engineering graduate students.

**Existing System:**

According to statistics 1.5 million engineers are graduating each year in India. The demand and need for qualified graduates in field of IT industry is rising day by day. But most of the students are unaware about the needs of the IT industry. The number of the student graduates who satisfies the requirements and quality of a company is very low. Placements are one of the biggest challenges faced by a student in the lifetime. It is the responsibility of the institutions to provide maximum placement chance to its students. Also the placement cell and teachers of an institute should take proper steps in order to produce a set of students suitable for each company’s requirements. A placement prediction system can be used to identify the capability of a particular student for the specified job.

**Proposed System:**

All companies in the IT sector spend a large amount of its total capital in recruiting the students to its company. Thus it is necessary to find an alternative process of filtering to reduce the capital cost that is used for this process. Effective filtering of students could be performed by applying various deep learning tools on the student details. This system defined the meaning of deep learning in the field of education as a method of identifying, discovering and capturing the unknown similarities or patterns from a dataset by using an ensemble combination of various analytical approaches. It is possible for an educational institute to exploit this data mining feature to figure out the recruitment policy of a company from previous year placement statistics and student dataset. So the placement cell of the institution could prepare a placement predicted list for the present students. Thus it is very important to conduct a study on various placement prediction systems. This paper presents a survey on different placement prediction system models and its application for the students.

**SYSTEM CONFIGURATION:**

**Hardware requirements:**

Processer : Any Update Processer

Ram : Min 4 GB

Hard Disk : Min 100 GB

**Software requirements:**

Operating System : Windows family

Technology : Python 3.6

IDE : PyCharm

**Implemented by**

Development team      :          Cloud Technologies

Website                       :           <http://www.cloudstechnologies.in>

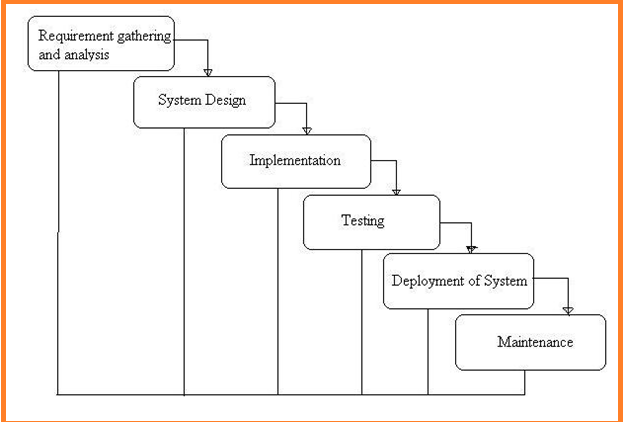
Call on                       :           8121953811

Whatsapp : 8522991105

**INTRODUCTION**

According to statistics 1.5 million engineers are graduating each year in India. The demand and need for qualified graduates in field of IT industry is rising day by day. But most of the students are unaware about the needs of the IT industry. The number of the student graduates who satisfies the requirements and quality of a company is very low. Placements are one of the biggest challenge faced by a student in the lifetime. It is the responsibility of the institutions to provide maximum placement chance to its students. Also the placement cell and teachers of an institute should take proper steps inorder to produce a set of students suitable for each company’s requirements. A placement prediction system can be used to identify the capability of a particular student for the specified job. All companies in the IT sector spends a large amount of its total capital in recruiting the students to its company. Thus it is necessary to find an alternative process of filtering to reduce the capital cost that is used for this process. Effective filtering of students could be performed by applying various data mining and machine learning tools on the student details. Luan [1] defined the meaning of data mining in the field of education as a method of identifying, discovering and capturing the unknown similarities or patterns from a dataset by using an ensemble combination of various analytical approaches. It is possible for an educational institute to exploit this data mining feature to figure out the recruitment policy of a company from previous year placement statistics and student dataset. So the placement cell of the institution could prepare a placement predicted list for the present students. Thus it is very important to conduct a study on various placement prediction systems. This paper presents a survey on different placement prediction system models and its application for the students.

**SYSTEM ANALYSIS**



**What is Waterfall Model?**

Waterfall Model is a sequential model that divides software development into different phases. Each phase is designed for performing specific activity during SDLC phase. It was introduced in 1970 by Winston Royce.

**Requirements:**

The first phase involves understanding what needs to design and what is its function, purpose, etc. Here, the specifications of the input and output or the final product are studied and marked.

**System Design:**

The requirement specifications from the first phase are studied in this phase and system design is prepared. System Design helps in specifying hardware and system requirements and also helps in defining overall system architecture. The software code to be written in the next stage is created now.

**Implementation:**

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

**Integration and Testing:**

All the units developed in the implementation phase are integrated into a system after testing of each unit. The software designed, needs to go through constant software testing to find out if there are any flaws or errors. Testing is done so that the client does not face any problem during the installation of the software.

**Deployment of System:**

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

**Maintenance:** This step occurs after installation, and involves making modifications to the system or an individual component to alter attributes or improve performance. These modifications arise either due to change requests initiated by the customer, or defects uncovered during live use of the system. The client is provided with regular maintenance and support for the developed software.

**FUNCTIONAL REQUIREMENTS**

* Admin Login
* Questions Dataset Upload
* Data Set Upload
* Student Login
* Student Registration
* Take Exam
* View Dataset
* Accuracy Measuring
* Prediction

**NON-FUNCTIONAL REQUIREMENTS**

## What is Non-Functional Requirement?

**NON-FUNCTIONAL REQUIREMENT** (NFR) specifies the quality attribute of a software system. They judge the software system based on Responsiveness, Usability, Security, Portability and other non-functional standards that are critical to the success of the software system. Example of nonfunctional requirement, *“how fast does the website load?”* Failing to meet non-functional requirements can result in systems that fail to satisfy user needs.

Non-funional Requirements allows you to impose constraints or restrictions on the design of the system across the various agile backlogs. Example, the site should load in 3 seconds when the number of simultaneous users are > 10000. Description of non-functional requirements is just as critical as a functional requirement.

* Usability requirement
* Serviceability requirement
* Manageability requirement
* Recoverability requirement
* Security requirement
* Data Integrity requirement
* Capacity requirement
* Availability requirement
* Scalability requirement
* Interoperability requirement
* Reliability requirement
* Maintainability requirement
* Regulatory requirement
* Environmental requirement

## Examples of Non-functional requirements

Here, are some examples of non-functional requirement:

1. Users must change the initially assigned login password immediately after the first successful login. Moreover, the initial should never be reused.
2. Employees never allowed to update their salary information. Such attempt should be reported to the security administrator.
3. Every unsuccessful attempt by a user to access an item of data shall be recorded on an audit trail.
4. A website should be capable enough to handle 20 million users with affecting its performance
5. The software should be portable. So moving from one OS to other OS does not create any problem.
6. Privacy of information, the export of restricted technologies, intellectual property rights, etc. should be audited.

## Advantages of Non-Functional Requirement

Benefits/pros of Non-functional testing are:

* The nonfunctional requirements ensure the software system follow legal and compliance rules.
* They ensure the reliability, availability, and performance of the software system
* They ensure good user experience and ease of operating the software.
* They help in formulating security policy of the software system.

## Disadvantages of Non-functional requirement

Cons/drawbacks of Non-function requirement are:

* None functional requirement may affect the various high-level software subsystem
* They require special consideration during the software architecture/high-level design phase which increases costs.
* Their implementation does not usually map to the specific software sub-system,
* It is tough to modify non-functional once you pass the architecture phase.

## KEY LEARNING

* A non-functional requirement defines the performance attribute of a software system.
* Types of Non-functional requirement are Scalability Capacity, Availability, Reliability, Recoverability, Data Integrity, etc.
* Example of Non Functional Requirement is Employees never allowed to update their salary information. Such attempt should be reported to the security administrator.
* Functional Requirement is a verb while Non-Functional Requirement is an attribute
* The advantage of Non-functional requirement is that it helps you to ensure good user experience and ease of operating the software
* The biggest disadvantage of Non-functional requirement is that it may affect the various high-level software subsystems.

**LITERATURE SURVEY**

**Introduction to the problem domain terminology**

Today’s Artificial Intelligence (AI) has far surpassed the hype of blockchain and quantum

computing. This is due to the fact that huge computing resources are easily available to

the common man. The developers now take advantage of this in creating new Machine

Learning models and to re-train the existing models for better performance and results.

The easy availability of High Performance Computing (HPC) has resulted in a sudden

increased demand for IT professionals having Machine Learning skills.

In this tutorial, you will learn in detail about:

* What is the crux of machine learning?
* What are the different types in machine learning?
* What are the different algorithms available for developing machine learning
* models?
* What tools are available for developing these models?
* What are the programming language choices?
* What platforms support development and deployment of Machine Learning
* applications?
* What IDEs (Integrated Development Environment) are available?
* How to quickly upgrade your skills in this important area?

[1] “Data Mining Approach for Predicting Student and Institution's Placement Percentage”, Professor. Ashok M Assistant Professor Apoorva A ,2016 International Conference on Computational Systems and Information Systems for Sustainable Solutions In this paper author has used the data mining technique for the prediction of the student’s placement. For the prediction of student's placement author has divided the data into the two segments, first segment is the training segment which is historic data of passed out students. Another segment consists of current data of students, based on the historic data author has designed the algorithm for calculating the placement chances. Author has used the various data mining algorithms such as decision tree, Naive Bayes, neural network and the prosed algorithm were applied, and decision are made with the help of confusion matrix.

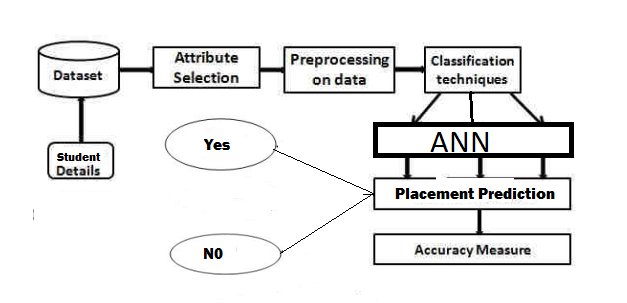
[2] “Student Placement Analyzer: A Recommendation System Using Machine Learning”, Senthil Kumar Thangavel , Divya Bharathi P, Abijith Sankar, International Conference on Advanced Computing and Communication Systems (ICACCS -2017), Jan. 06 - 07, 2017, Coimbatore, INDIA In this paper author is concern about the challenges face by any institute regarding the placement. The placement prediction is very complex when the number of the entities increases in any institute. With the help of machine learning this complex problem of prediction can be easily solved. In this paper all the academic record of student is taken into consideration. Various classification and data making algorithms are used such as Naïve Bayes, Decision Tree, SVM and Regressions. After the prediction of the students can be placed in of the given category that is core company, dream company or support services.

[3] "A Placement Prediction System Using K-Nearest Neighbors Classifier", Animesh Giri, M Vignesh V Bhagavath, Bysani Pruthvi, Naini Dubey, Second International Conference on Cognitive Computing and Information Processing (CCIP), 2016 The placement prediction system predicts the probability of students getting placed in various companies by applying K-Nearest Neighbors classification. The result obtained is also compared with the results obtained from other machine learning models like Logistic Regression and SVM. The academic history of student along with their skill sets like programming skills, communication skills, analytical skills and team work is considered which is tested by companies during recruitment process. Data of past two batches are taken for this system.

[4]"Class Result Prediction using Machine Learning", Pushpa S K, Associate Professor, Manjunath T N, Professor and Head, Mrunal T V, Amartya Singh, C Suhas, International Conference On Smart Technology for Smart Nation, 2017 In this paper, the result of a class is predicted using machine learning. Performance of students in past semester along with scores of internal examinations of the current semester is considered to predict whether the student passes or fails in the current semester before attempting the final examination. The author uses SVM, Naive Bayes, Random Forest Classifier and Gradient Boosting to compute the result. Boosting is an ensemble learning algorithm which combines various learning algorithm to obtain better predictive performance.

[5]“Student Placement Analyzer : A Recommendation System Using Machine Leraning”, Apoorva Rao R, Deeksha K C, Vishal Prajwal R, Vrushak K, Nandini, JARIIE-ISSN(O)-2395-4396Now-a-days institutions are facing many challenges regarding student placements. For educational institutions it is much difficult task to keep record of every single student and predict the placement of student manually. To overcome these challenges, concept of machine learning and various algorithms are explored to predict the result of class students. For this purpose, training data set is historical data of past students and this is used to train the model. This software system predicts placement status in 5 categories viz dream company, core company, mass recruiter, not eligible and not interested in placements. This system is also helpful to weaker students. Institutions can provide extra care towards weaker students so that they can improve their performance. By use Naïve Bayes algorithm all the data will be monitor and appropriate decision will be provided.

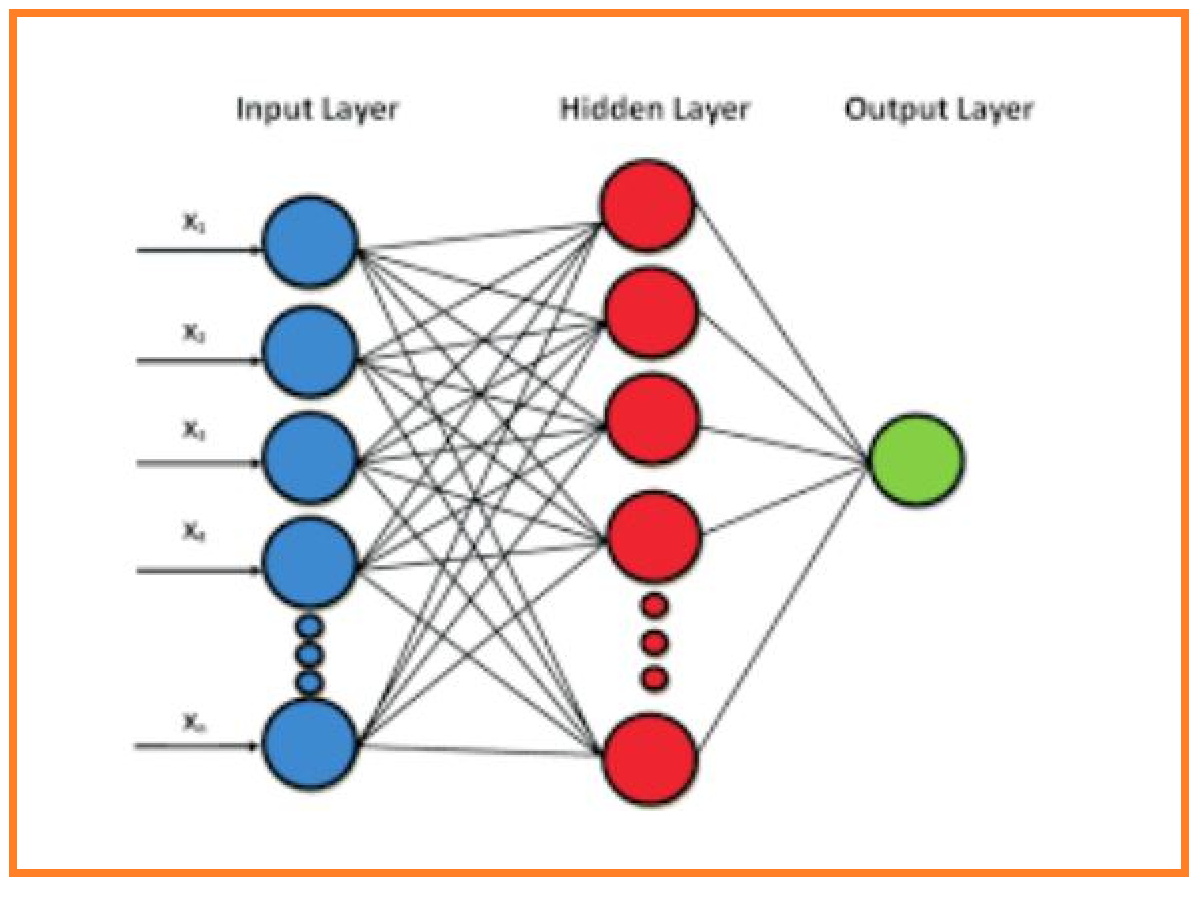
**IMPLEMENTATION**

****

**Module Description**

In Our System to find Placement Chance we are going to keep a Exam in our application on areas like Coding Skills, Aptitude Skills, Technical Skills Based on the Results we are going to Find the Placement using Deep learning ANN algorithm

**Neural networks**



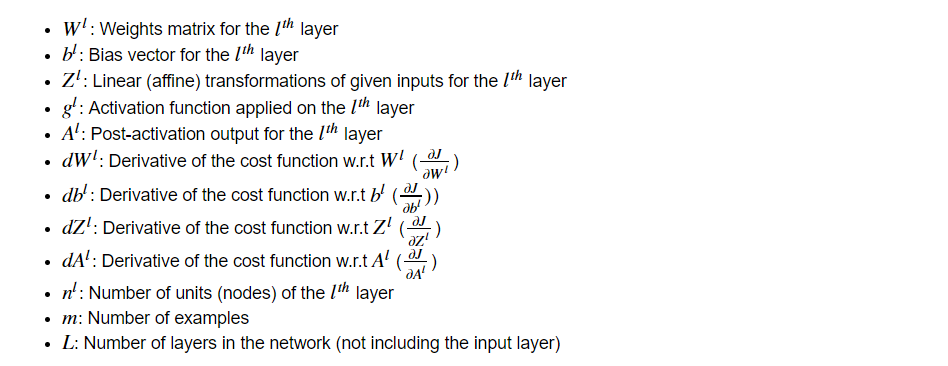
Output layer. Therefore, we can define neural network as information flows from inputs through hidden layers towards the output. For a 3-layers neural network, the learned function would be: f(x) = f\_3(f\_2(f\_1(x))) where:

f\_1(x): Function learned on first hidden layer

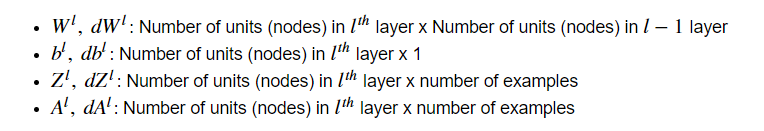
f\_2(x): Function learned on second hidden layer

f\_3(x): Function learned on output layer

Lets first introduce some notations that will be used throughout the post:



Next, we’ll write down the dimensions of a multi-layer neural network in the general form to help us in matrix multiplication because one of the major challenges in implementing a neural network is getting the dimensions right.



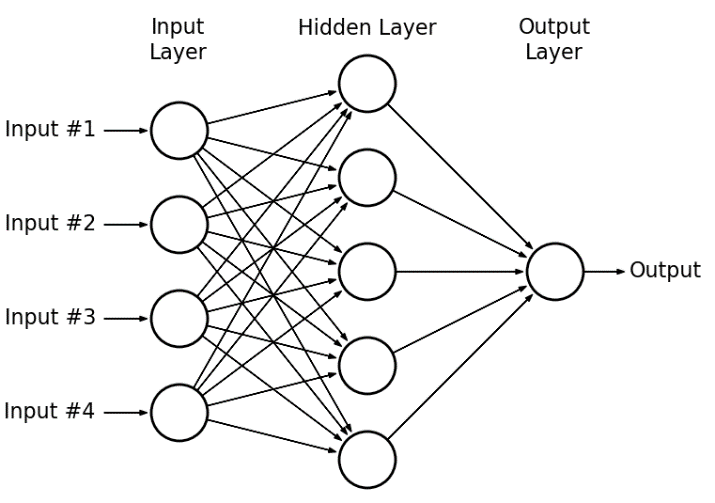
**Multi-Layer Perceptrons**

The field of artificial neural networks is often just called neural networks or multi-layer perceptrons after perhaps the most useful type of neural network. A perceptron is a single neuron model that was a precursor to larger neural networks.

It is a field that investigates how simple models of biological brains can be used to solve difficult computational tasks like the predictive modeling tasks we see in machine learning. The goal is not to create realistic models of the brain, but instead to develop robust algorithms and data structures that we can use to model difficult problems.

The power of neural networks come from their ability to learn the representation in your training data and how to best relate it to the output variable that you want to predict. In this sense neural networks learn a mapping. Mathematically, they are capable of learning any mapping function and have been proven to be a universal approximation algorithm.

The predictive capability of neural networks comes from the hierarchical or multi-layered structure of the networks. The data structure can pick out (learn to represent) features at different scales or resolutions and combine them into higher-order features. For example from lines, to collections of lines to shapes.



**SYSTEM DESIGN**

**UML DIAGRAMS**

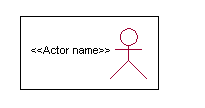
The System Design Document describes the system requirements, operating environment, system and subsystem architecture, files and database design, input formats, output layouts, human-machine interfaces, detailed design, processing logic, and external interfaces.

**Global Use Case Diagrams:**

Identification of actors:

**Actor:** Actor represents the role a user plays with respect to the system. An actor interacts with, but has no control over the use cases.

Graphical representation:



<<Actor name>>

An actor is someone or something that:

Interacts with or uses the system.

* Provides input to and receives information from the system.
* Is external to the system and has no control over the use cases.

Actors are discovered by examining:

* Who directly uses the system?
* Who is responsible for maintaining the system?
* External hardware used by the system.
* Other systems that need to interact with the system.

Questions to identify actors:

* + Who is using the system? Or, who is affected by the system? Or, which groups need help from the system to perform a task?
  + Who affects the system? Or, which user groups are needed by the system to perform its functions? These functions can be both main functions and secondary functions such as administration.
  + Which external hardware or systems (if any) use the system to perform tasks?
  + What problems does this application solve (that is, for whom)?
  + And, finally, how do users use the system (use case)? What are they doing with the system?

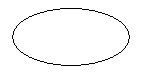
The actors identified in this system are:

1. **System Administrator**
2. **Customer**
3. **Customer Care**

Identification of usecases:

**Usecase:** A use case can be described as a specific way of using the system from a user’s (actor’s) perspective.

**Graphical representation:**



A more detailed description might characterize a use case as:

* Pattern of behavior the system exhibits
* A sequence of related transactions performed by an actor and the system
* Delivering something of value to the actor

Use cases provide a means to:

* capture system requirements
* communicate with the end users and domain experts
* test the system

Use cases are best discovered by examining the actors and defining what the actor will be able to do with the system.

Guide lines for identifying use cases:

* For each actor, find the tasks and functions that the actor should be able to perform or that the system needs the actor to perform. The use case should represent a course of events that leads to clear goal
* Name the use cases.
* Describe the use cases briefly by applying terms with which the user is familiar.

This makes the description less ambiguous

Questions to identify use cases:

* What are the tasks of each actor?
* Will any actor create, store, change, remove or read information in the system?
* What use case will store, change, remove or read this information?
* Will any actor need to inform the system about sudden external changes?
* Does any actor need to inform about certain occurrences in the system?
* What usecases will support and maintains the system?

**Flow of Events**

A flow of events is a sequence of transactions (or events) performed by the system. They typically contain very detailed information, written in terms of what the system should do, not how the system accomplishes the task. Flow of events are created as separate files or documents in your favorite text editor and then attached or linked to a use case using the Files tab of a model element.

A flow of events should include:

* When and how the use case starts and ends
* Use case/actor interactions
* Data needed by the use case
* Normal sequence of events for the use case
* Alternate or exceptional flows

Construction of Usecase diagrams:

Use-case diagrams graphically depict system behavior (use cases). These diagrams present a high level view of how the system is used as viewed from an outsider’s (actor’s) perspective. A use-case diagram may depict all or some of the use cases of a system.

A use-case diagram can contain:

* actors ("things" outside the system)
* use cases (system boundaries identifying what the system should do)
* Interactions or relationships between actors and use cases in the system including the associations, dependencies, and generalizations.

Relationships in use cases:

**1. Communication:**

The communication relationship of an actor in a usecase is shown by connecting the actor symbol to the usecase symbol with a solid path. The actor is said to communicate with the usecase.

**2. Uses:**

A Uses relationship between the usecases is shown by generalization arrow from the usecase.

**3. Extends:**

The extend relationship is used when we have one usecase that is similar to another usecase but does a bit more. In essence it is like subclass.

**SEQUENCE DIAGRAMS**

A sequence diagram is a graphical view of a scenario that shows object interaction in a time-based sequence what happens first, what happens next. Sequence diagrams establish the roles of objects and help provide essential information to determine class responsibilities and interfaces.

There are two main differences between sequence and collaboration diagrams: sequence diagrams show time-based object interaction while collaboration diagrams show how objects associate with each other. A sequence diagram has two dimensions: typically, vertical placement represents time and horizontal placement represents different objects.

**Object:**

An object has state, behavior, and identity. The structure and behavior of similar objects are defined in their common class. Each object in a diagram indicates some instance of a class. An object that is not named is referred to as a class instance.

The object icon is similar to a class icon except that the name is underlined:

An object's concurrency is defined by the concurrency of its class.

**Message:**

A message is the communication carried between two objects that trigger an event. A message carries information from the source focus of control to the destination focus of control. The synchronization of a message can be modified through the message specification. Synchronization means a message where the sending object pauses to wait for results.

**Link:**

A link should exist between two objects, including class utilities, only if there is a relationship between their corresponding classes. The existence of a relationship between two classes symbolizes a path of communication between instances of the classes: one object may send messages to another. The link is depicted as a straight line between objects or objects and class instances in a collaboration diagram. If an object links to itself, use the loop version of the icon.

**CLASS DIAGRAM:**

Identification of analysis classes:

A class is a set of objects that share a common structure and common behavior (the same attributes, operations, relationships and semantics). A class is an abstraction of real-world items.

There are 4 approaches for identifying classes:

1. Noun phrase approach:
2. Common class pattern approach.
3. Use case Driven Sequence or Collaboration approach.
4. Classes , Responsibilities and collaborators Approach
5. **Noun Phrase Approach:**

The guidelines for identifying the classes:

* + Look for nouns and noun phrases in the usecases.
  + Some classes are implicit or taken from general knowledge.
  + All classes must make sense in the application domain; Avoid computer

implementation classes – defer them to the design stage.

* + Carefully choose and define the class names After identifying the classes we have to eliminate the following types of classes:
* Adjective classes.

1. **Common class pattern approach:**

The following are the patterns for finding the candidate classes:

* + Concept class.
  + Events class.
  + Organization class
  + Peoples class
  + Places class
  + Tangible things and devices class.

1. **Use case driven approach:**

We have to draw the sequence diagram or collaboration diagram. If there is need for some classes to represent some functionality then add new classes which perform those functionalities.

1. **CRC approach:**

The process consists of the following steps:

* + Identify classes’ responsibilities ( and identify the classes )
  + Assign the responsibilities
  + Identify the collaborators.

Identification of responsibilities of each class:

The questions that should be answered to identify the attributes and methods of a class respectively are:

1. What information about an object should we keep track of?
2. What services must a class provide?

Identification of relationships among the classes:

Three types of relationships among the objects are:

Association: How objects are associated?

Super-sub structure: How are objects organized into super classes and sub classes?

Aggregation: What is the composition of the complex classes?

Association:

The **questions** that will help us to identify the associations are:

1. Is the class capable of fulfilling the required task by itself?
2. If not, what does it need?
3. From what other classes can it acquire what it needs?

Guidelines for identifying the tentative associations:

* A dependency between two or more classes may be an association. Association often corresponds to a verb or prepositional phrase.
* A reference from one class to another is an association. Some associations are implicit or taken from general knowledge.

Some common association patterns are:

Location association like part of, next to, contained in…..

Communication association like talk to, order to ……

We have to eliminate the unnecessary association like implementation associations, ternary or n-ary associations and derived associations.

Super-sub class relationships:

Super-sub class hierarchy is a relationship between classes where one class is the parent class of another class (derived class).This is based on inheritance.

Guidelines for identifying the super-sub relationship, a generalization are

1***.* Top-down*:***

Look for noun phrases composed of various adjectives in a class name. Avoid excessive refinement. Specialize only when the sub classes have significant behavior.

2.**Bottom-up*:***

Look for classes with similar attributes or methods. Group them by moving the common attributes and methods to an abstract class. You may have to alter the definitions a bit.

3.**Reusability*:***

Move the attributes and methods as high as possible in the hierarchy.

4. **Multiple inheritances*:***

Avoid excessive use of multiple inheritances. One way of getting benefits of multiple inheritances is to inherit from the most appropriate class and add an object of another class as an attribute.

**Aggregation or a-part-of relationship:**

It represents the situation where a class consists of several component classes. A class that is composed of other classes doesn’t behave like its parts. It behaves very difficultly. The major properties of this relationship are transitivity and anti symmetry.

The **questions** whose answers will determine the distinction between the part and whole relationships are:

* Does the part class belong to the problem domain?
* Is the part class within the system’s responsibilities?
* Does the part class capture more than a single value?( If not then simply include it as an attribute of the whole class)
* Does it provide a useful abstraction in dealing with the problem domain?

There are three types of aggregation relationships. They are:

**Assembly:**

It is constructed from its parts and an assembly-partsituation physically exists.

**Container:**

A physical whole encompasses but is not constructed from physical parts.

**Collection member:**

A conceptual whole encompasses parts that may be physical or conceptual. The container and collection are represented by hollow diamonds but composition is represented by solid diamond.

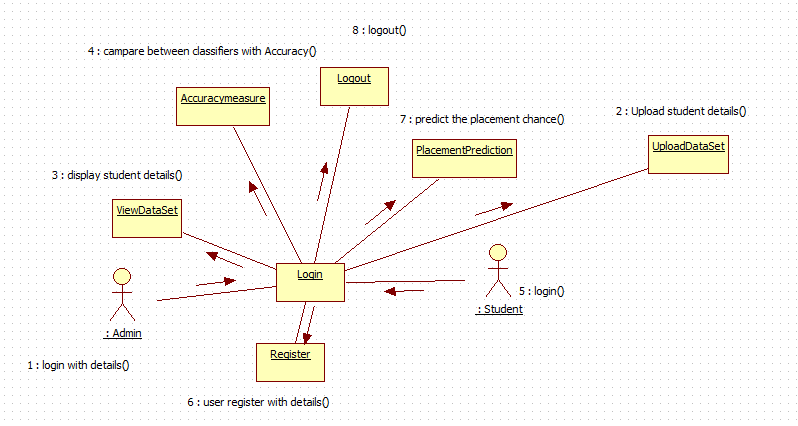
**USECASE DIAGRAM**

****

Seq



Collo



State



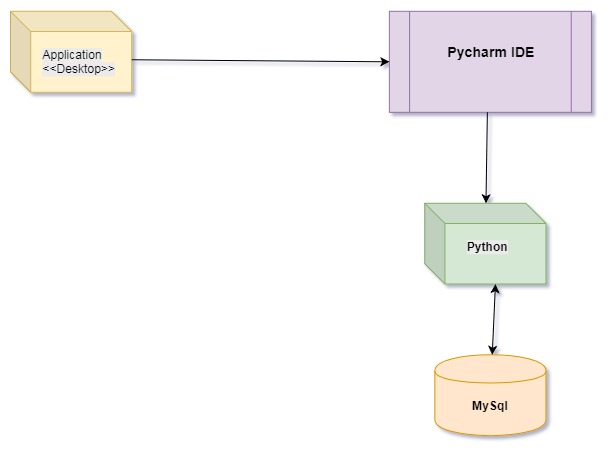
Actibvity



Compo



Python Mysql Deployemnt



**Data Flow Diagrams**

A data flow diagram is graphical tool used to describe and analyze movement of data through a system. These are the central tool and the basis from which the other components are developed. The transformation of data from input to output, through processed, may be described logically and independently of physical components associated with the system. These are known as the logical data flow diagrams. The physical data flow diagrams show the actual implements and movement of data between people, departments and workstations. A full description of a system actually consists of a set of data flow diagrams. Using two familiar notations Yourdon, Game and Sarson notation develops the data flow diagrams. Each component in a DFD is labeled with a descriptive name. Process is further identified with a number that will be used for identification purpose. The development of DFD’S is done in several levels. Each process in lower level diagrams can be broken down into a more detailed DFD in the next level. The lop-level diagram is often called context diagram. It consists a single process bit, which plays vital role in studying the current system. The process in the context level diagram is exploded into other process at the first level DFD.

The idea behind the explosion of a process into more process is that understanding at one level of detail is exploded into greater detail at the next level. This is done until further explosion is necessary and an adequate amount of detail is described for analyst to understand the process.

Larry Constantine first developed the DFD as a way of expressing system requirements in a graphical from, this lead to the modular design.

A DFD is also known as a “bubble Chart” has the purpose of clarifying system requirements and identifying major transformations that will become programs in system design. So it is the starting point of the design to the lowest level of detail. A DFD consists of a series of bubbles joined by data flows in the system.

**DFD Symbols:**

In the DFD, there are four symbols

1. A square defines a source(originator) or destination of system data
2. An arrow identifies data flow. It is the pipeline through which the information flows
3. A circle or a bubble represents a process that transforms incoming data flow into outgoing data flows.
4. An open rectangle is a data store, data at rest or a temporary repository of data

Process that transforms data flow.

Source or Destination of data

Data flow

Data Store

**Constructing a DFD:**

Several rules of thumb are used in drawing DFD’S:

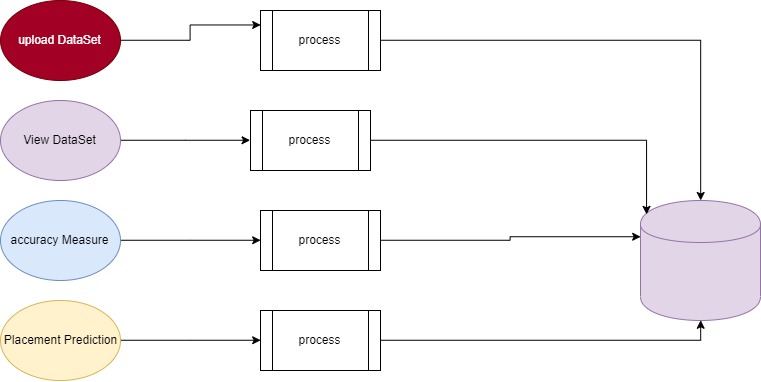
Process should be named and numbered for an easy reference. Each name should be representative of the process. The direction of flow is from top to bottom and from left to right. Data traditionally flow from source to the destination although they may flow back to the source. One way to indicate this is to draw long flow line back to a source. An alternative way is to repeat the source symbol as a destination. Since it is used more than once in the DFD it is marked with a short diagonal. When a process is exploded into lower level details, they are numbered. The names of data stores and destinations are written in capital letters. Process and dataflow names have the first letter of each work capitalized. A DFD typically shows the minimum contents of data store. Each data store should contain all the data elements that flow in and out. Questionnaires should contain all the data elements that flow in and out. Missing interfaces redundancies and like is then accounted for often through interviews.

**Silent Feature of DFD’s**

1. The DFD shows flow of data, not of control loops and decision are controlled considerations do not appear on a DFD.
2. The DFD does not indicate the time factor involved in any process whether the dataflow take place daily, weekly, monthly or yearly.
3. The sequence of events is not brought out on the DFD.

**Data Flow:**

1. A Data Flow has only one direction of flow between symbols. It may flow in both directions between a process and a data store to show a read before an update. The later is usually indicated however by two separate arrows since these happen at different type.
2. A join in DFD means that exactly the same data comes from any of two or more different processes data store or sink to a common location.
3. A data flow cannot go directly back to the same process it leads. There must be at least one other process that handles the data flow produce some other data flow returns the original data into the beginning process.
4. A Data flow to a data store means update (delete or change).
5. A data Flow from a data store means retrieve or use. A data flow has a noun phrase label more than one data flow noun phrase can appear on a single arrow as long as all of the flows on the same arrow move together as one package.

****

**About MySQL:**

**MySQL** is a relational database management system (RDBMS)[]](http://en.wikipedia.org/wiki/MySQL#cite_note-1) that runs as a server providing multi-user access to a number of databases.  The SQL phrase stands for Structured Query Language.Free-software-open source projects that require a full-featured database management system often use MySQL. For commercial use, several paid editions are available, and offer additional functionality. Applications which use MySQL databases include: TYPO3, Joomla, WordPress, phpBB, Drupal and other software built on the LAMP software stack. MySQL is also used in many high-profile, large-scale World Wide Web products, including Wikipedia, Google  , Facebook, and Twitter.

MySQL is the world's most popular open source database software, with over 100 million copies of its software downloaded or distributed throughout it's history. With its superior speed, reliability, and ease of use, MySQL has become the preferred choice for Web, Web 2.0, SaaS, ISV, Telecom companies and forward-thinking corporate IT Managers because it eliminates the major problems associated with downtime, maintenance and administration for modern, online applications.

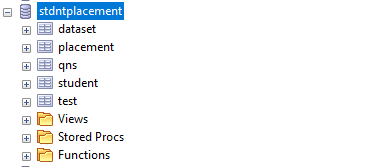
Many of the world's largest and fastest-growing organizations use MySQL to save time and money powering their high-volume Web sites, critical business systems, and packaged software including industry leaders such as Yahoo!, Alcatel-Lucent, Google, Nokia, YouTube, Wikipedia, and Booking.com.

The flagship MySQL offering is MySQL Enterprise, a comprehensive set of production-tested software, proactive monitoring tools, and premium support services available in an affordable annual subscription.

MySQL is a key part of LAMP (Linux, Apache, MySQL, PHP / Perl / Python), the fast-growing open source enterprise software stack. More and more companies are using LAMP as an alternative to expensive proprietary software stacks because of its lower cost and freedom from platform lock-in.

MySQL was originally founded and developed in Sweden by two Swedes and a Finn: David Axmark, Allan Larsson and Michael "Monty" Widenius, who had worked together since the 1980's. More historical information on MySQL is

**DATABASE TABLES**



CREATE DATABASE /\*!32312 IF NOT EXISTS\*/`stdntplacement`

CREATE TABLE `dataset` (

`CodingSkills` varchar(100) DEFAULT NULL,

`AptitudeSkills` varchar(100) DEFAULT NULL,

`TechnicalSkills` varchar(100) DEFAULT NULL,

`Projects` varchar(100) DEFAULT NULL,

`Internships` varchar(100) DEFAULT NULL,

`AcademicPerformance` varchar(100) DEFAULT NULL,

`Placed` varchar(100) DEFAULT NULL

)

CREATE TABLE `placement` (

`rno` varchar(500) DEFAULT NULL,

`cs` varchar(500) DEFAULT NULL,

`as` varchar(500) DEFAULT NULL,

`ts` varchar(500) DEFAULT NULL,

`prjcts` varchar(500) DEFAULT NULL,

`is` varchar(500) DEFAULT NULL,

`ap` varchar(500) DEFAULT NULL,

`res` varchar(500) DEFAULT NULL

)

CREATE TABLE `qns` (

`sno` int(11) NOT NULL,

`qn` varchar(1000) NOT NULL,

`A` varchar(500) NOT NULL,

`B` varchar(500) NOT NULL,

`C` varchar(500) NOT NULL,

`D` varchar(500) NOT NULL,

`E` varchar(500) DEFAULT NULL,

`Ans` varchar(10) DEFAULT NULL,

`Level` int(11) DEFAULT NULL

)

CREATE TABLE `student` (

`name` varchar(100) DEFAULT NULL,

`stdid` varchar(100) DEFAULT NULL,

`pwd` varchar(100) DEFAULT NULL,

`email` varchar(100) DEFAULT NULL,

`mno` varchar(100) DEFAULT NULL

)

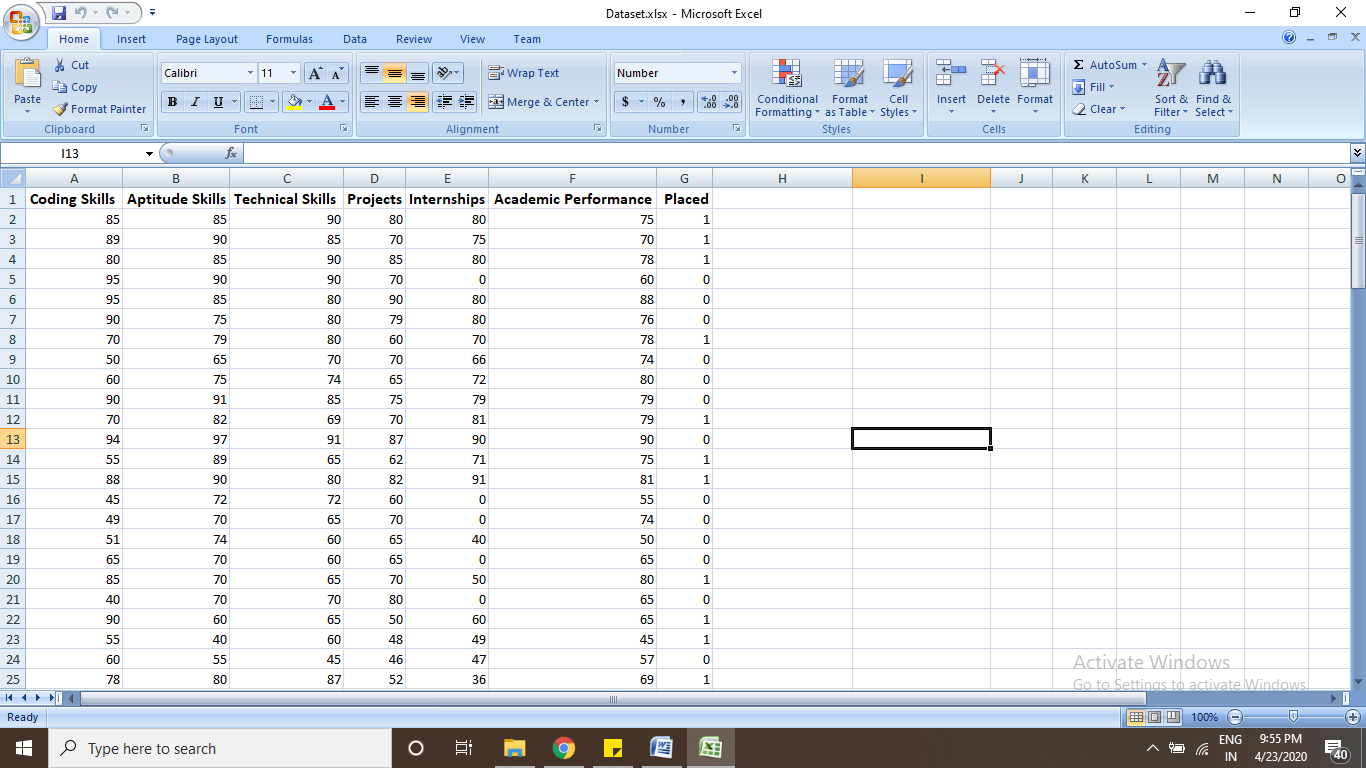
CREATE TABLE `test` (

`testcode` varchar(10) DEFAULT NULL,

`score` varchar(100) DEFAULT NULL

)

**DATASET**

****

Dataset attributes used in our project

* Coding Skills
* Aptitude Skills
* Technical Skills
* Projects
* Internships
* Academic Performance
* Placed

**TECHNOLOGIES USED**

**History of Python**

Python was developed by Guido van Rossum in the late eighties and early nineties at the National Research Institute for Mathematics and Computer Science in the Netherlands.Python is derived from many other languages, including ABC, Modula-3, C, C++, Algol-68, SmallTalk, and Unix shell and other scripting languages.

Python is copyrighted. Like Perl, Python source code is now available under the GNU General Public License (GPL).

Python is now maintained by a core development team at the institute, although Guido van Rossum still holds a vital role in directing its progress.

**Input as CSV File**

Reading data from CSV(comma separated values) is a fundamental necessity in Data Science. Often, we get data from various sources which can get exported to CSV format so that they can be used by other systems. The Panadas library provides features using which we can read the CSV file in full as well as in parts for only a selected group of columns and rows.

The CSV file is a text file in which the values in the columns are separated by a comma. Let's consider the following data present in the file named input.csv.You can create this file using windows notepad by copying and pasting this data. Save the file as input.csv using the save As All files(\*.\*) option in notepad.

import pandas as pd

data= pd.read\_csv('path/input.csv')

print(data)

**Operations using NumPy**

NumPy is a Python package which stands for 'Numerical Python'. It is a library consisting of multidimensional array objects and a collection of routines for processing of array.

Using NumPy, a developer can perform the following operations −

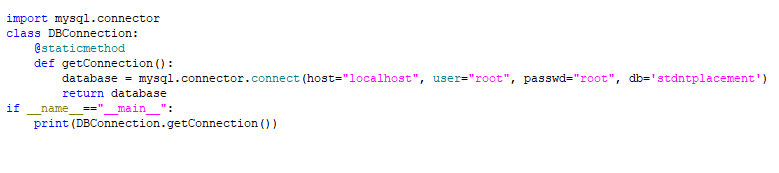
* Mathematical and logical operations on arrays.
* Fourier transforms and routines for shape manipulation.
* Operations -related to linear algebra. NumPy has in-built functions for linear algebra and random number generation.

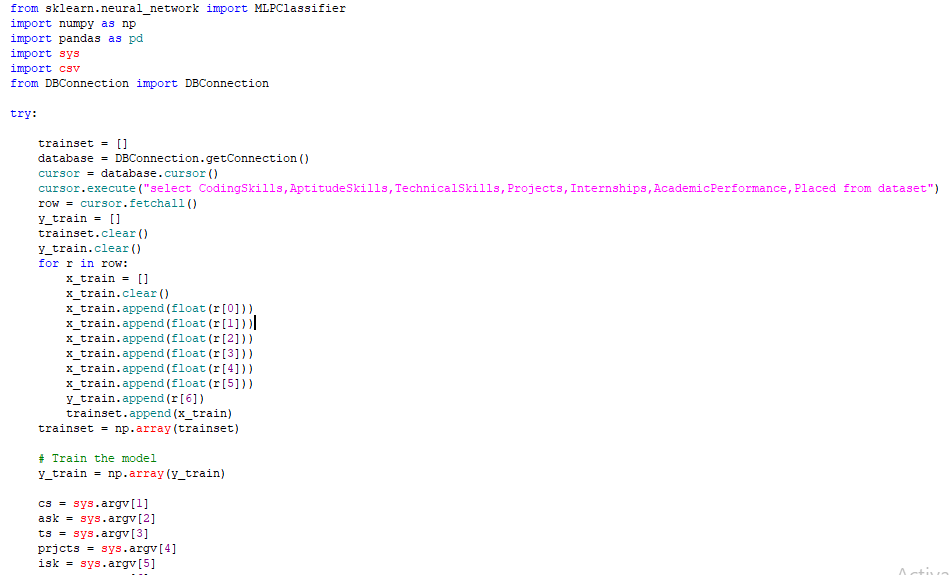
## Key Features of Pandas

* Fast and efficient DataFrame object with default and customized indexing.
* Tools for loading data into in-memory data objects from different file formats.
* Data alignment and integrated handling of missing data.
* Reshaping and pivoting of date sets.
* Label-based slicing, indexing and subsetting of large data sets.
* Columns from a data structure can be deleted or inserted.
* Group by data for aggregation and transformations.
* High performance merging and joining of data.
* Time Series functionality.

**SAMPLE CODE**

**DBConnection.py**

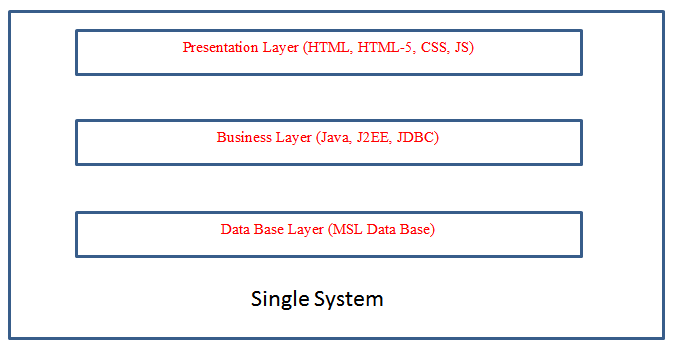
**** **Predict.py**

****

**About Project Software’s**

JAVA, Apache Server, MSQL, EDIT ++

In our web Application Development we are using one tier architecture as total applicant will be developed in single system with all the three layers of application development like presentation layer where we use our web technologies to make of GUI of the application like HTML, HTML-5, CSS, JS Etc. and in second layer we have to make our business logic or called as implementation of application where we are using java, J2EE and also we use JDBC to connect from our Business layer to data base layer and final our data base layer where we develop the Data structure of the application



Single tire Architecture Project Development

**How we used java in our Project Development**

**Installation and Setup in our system**

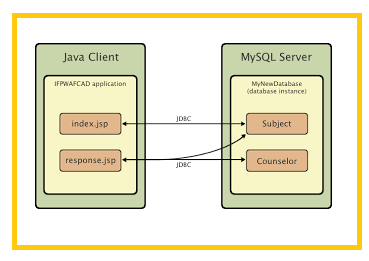
The Software we download form the oracle website as it’s an open source as per the software we have installed it in our system and for we have set the system path of java in our OS location We have used the main logic of our algorithm by core java concepts only for web application we have used all JSP concepts and to connect data base we have used JDBC with all this concept we have done the application in Single tire Architecture Project

**Data Storage in MYSQL**

We have taken open source software MYSQL from the provide website and run in our system we used for creating our project data base related tables as per project requirement’s even for user friendly access of my sql we used Software called SQL Yog where we can do all the operation of mysql by click & use

**About the role of apace tomcat webserver**

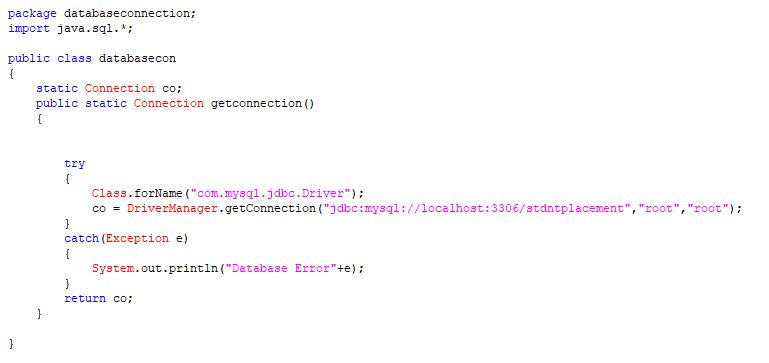
As our project is a web applicant we need webserver so for that we used again open sour software where our total project source code will be in webapps of the server form that location the application run into web browser where users can see the implementation of the total project



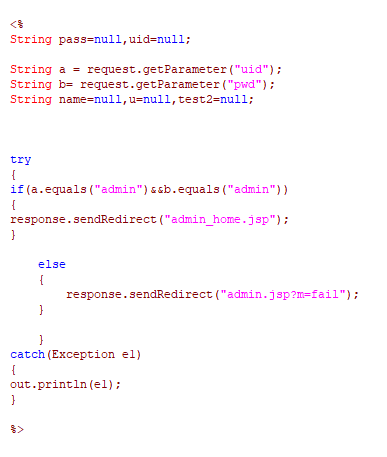
**Application Development Structure**

**SAMPLE CODE**

**databasecon.java**

****

**alogin.jsp**

****

**exam.jsp**

****

**SOFTWARE TESTING**

Software testing is one of the main stages of project development life cycle to provide our cessation utilizer with information about the quality of the application and ours, in our Project we have under gone some stages of testing like unit testing where it’s done in development stage of the project when we are in implementation of the application after the Project is yare we have done manual testing with different Case of all the different modules in the application we have even done browser compatibility testing in different web browsers in market, even we have done Client side validation testing on our application

**Unit testing**

The unit testing is done in the stage of implementation of the project only the error are solved in development stage some of the error we come across in development are given below

**TESTING**

Testing is the debugging program is one of the most critical aspects of the computer programming triggers, without programming that works, the system would never produce an output of which it was designed. Testing is best performed when user development is asked to assist in identifying all errors and bugs. The sample data are used for testing. It is not quantity but quality of the data used the matters of testing. Testing is aimed at ensuring that the system was accurately an efficiently before live operation commands.

**Testing objectives:**

The main objective of testing is to uncover a host of errors, systematically and with minimum effort and time. Stating formally, we can say, testing is a process of executing a program with intent of finding an error.

A successful test is one that uncovers an as yet undiscovered error.

A good test case is one that has probability of finding an error, if it exists.

The test is inadequate to detect possibly present errors.

The software more or less confirms to the quality and reliable standards.

**Levels of Testing:**

In order to uncover present in different phases we have the concept of levels of testing.

**The basic levels of Testing:**

Client needs acceptance testing

Requirements system testing

Design integration testing

Code unit testing

Figure: Levels of Testing

**Code testing:**

This examines the logic of the program. For example, the logic for updating various sample data and with the sample files and directories were tested and verified.

**Specification Testing:**

Executing this specification starting what the program should do and how it should performed under various conditions. Test cases for various situation and combination of conditions in all the modules are tested.

**Unit testing:**

In the unit testing we test each module individually and integrate with the overall system. Unit testing focuses verification efforts on the smallest unit of software design in the module. This is also known as module testing. The module of the system is tested separately. This testing is carried out during programming stage itself. In the testing step each module is found to work satisfactorily as regard to expected output from the module. There are some validation checks for fields also. For example the validation check is done for varying the user input given by the user which validity of the data entered. It is very easy to find error debut the system.

Each Module can be tested using the following two Strategies:

1. Black Box Testing
2. White Box Testing

**BLACK BOX TESTING**

**What is Black Box Testing?**

Black box testing is a software testing techniques in which **functionality of the software under test (SUT) is tested without looking at the internal code structure**, implementation details and knowledge of internal paths of the software. This type of testing is based entirely on the software requirements and specifications.

**In Black Box Testing we just focus on inputs and output of the software system** without bothering about internal knowledge of the software program.



The above Black Box can be any software system you want to test. For example : an operating system like Windows, a website like Google ,a database like Oracle or even your own custom application. Under Black Box Testing, you can test these applications by just focusing on the inputs and outputs without knowing their internal code implementation.

**Black box testing - Steps**

Here are the generic steps followed to carry out any type of Black Box Testing.

* Initially requirements and specifications of the system are examined.
* Tester chooses valid inputs (positive test scenario) to check whether SUT processes them correctly. Also some invalid inputs (negative test scenario) are chosen to verify that the SUT is able to detect them.
* Tester determines expected outputs for all those inputs.
* Software tester constructs test cases with the selected inputs.
* The test cases are executed.
* Software tester compares the actual outputs with the expected outputs.
* Defects if any are fixed and re-tested.

**Types of Black Box Testing**

There are many types of Black Box Testing but following are the prominent ones -

* **Functional testing** – This black box testing type is related to functional requirements of a system; it is done by software testers.
* **Non-functional testing** – This type of black box testing is not related to testing of a specific functionality, but non-functional requirements  such as performance, scalability, usability.
* **Regression testing** – Regression testing is done  after code fixes , upgrades or any other system maintenance to check the new code has not affected the existing code.

**WHITE BOX TESTING**

White Box Testing is the testing of a software solution's internal coding and infrastructure.It focuses primarily on strengthening security, the flow of inputs and outputs through the application, and improving design and usability.White box testing is also known as **clear, open, structural, and glass box testing**.

It is one of two parts of the **"box testing" approach** of software testing. Its counter-part, blackbox testing, involves testing from an external or end-user type perspective. On the other hand, Whitebox testing is based on the inner workings of an application and revolves around internal testing. The term "whitebox" was used because of the see-through box concept. The clear box or whitebox name symbolizes the ability to see through the software's outer shell (or "box") into its inner workings. Likewise, the "black box" in "black box testing" symbolizes not being able to see the inner workings of the software so that only the end-user experience can be tested

## What do you verify in White Box Testing ?

White box testing involves the testing of the software code for the following:

* Internal security holes
* Broken or poorly structured paths in the coding processes
* The flow of specific inputs through the code
* Expected output
* The functionality of conditional loops
* Testing of each statement, object and function on an individual basis

The testing can be done at system, integration and unit levels of software development. One of the basic goals of whitebox testing is to verify a working flow for an application. It involves testing a series of predefined inputs against expected or desired outputs so that when a specific input does not result in the expected output, you have encountered a bug.

**How do you perform White Box Testing?**

To give you a simplified explanation of white box testing, we have divided it into **two basic steps**. This is what testers do when testing an application using the white box testing technique:

**STEP 1) UNDERSTAND THE SOURCE CODE**

The first thing a tester will often do is learn and understand the source code of the application. Since white box testing involves the testing of the inner workings of an application, the tester must be very knowledgeable in the programming languages used in the applications they are testing. Also, the testing person must be highly aware of secure coding practices. Security is often one of the primary objectives of testing software. The tester should be able to find security issues and prevent attacks from hackers and naive users who might inject malicious code into the application either knowingly or unknowingly.

**Step 2) CREATE TEST CASES AND EXECUTE**

The second basic step to white box testing involves testing the application’s source code for proper flow and structure. One way is by writing more code to test the application’s source code. The tester will develop little tests for each process or series of processes in the application. This  method requires that the tester must have intimate knowledge of the code and is often done by the developer. Other methods include manual testing, trial and error testing and the use of testing tools as we will explain further on in this article.

**System testing:**

Once the individual module testing is completed, modules are assembled and integrated to perform as a system. The top down testing, which began from upper level to lower level module, was carried out to check whether the entire system is performing satisfactorily.

There are three main kinds of System testing:

1. Alpha Testing
2. Beta Testing
3. Acceptance Testing

**Alpha Testing:**

This refers to the system testing that is carried out by the test team with the Organization.

**Beta Testing**:

This refers to the system testing that is performed by a selected group of friendly customers

**Acceptance Testing:**

This refers to the system testing that is performed by the customer to determine whether or not to accept the delivery of the system.

**Integration Testing:**

Data can be lost across an interface, one module can have an adverse effort on the other sub functions, when combined, may not produce the desired major functions. Integrated testing is the systematic testing for constructing the uncover errors within the interface. The testing was done with sample data. The developed system has run successfully for this sample data. The need for integrated test is to find the overall system performance.

**Output testing:**After performance of the validation testing, the next step is output testing. The output displayed or generated by the system under consideration is tested by asking the user about the format required by system.

**UNIT TESTING**

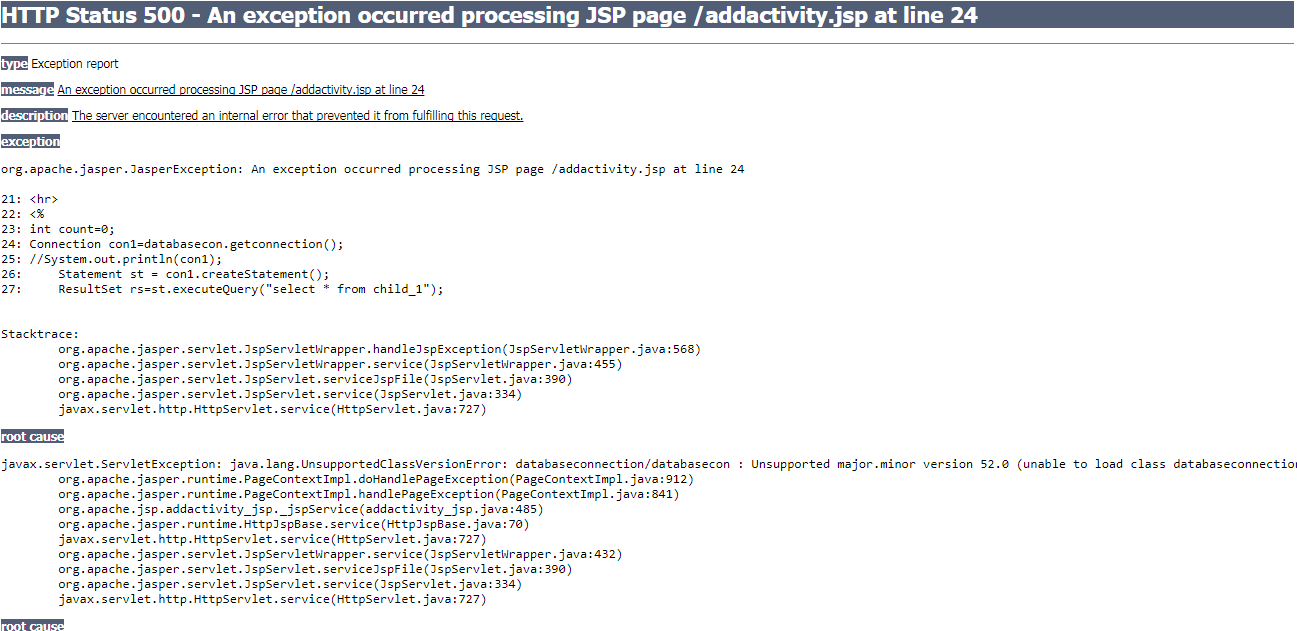
**Testing done when application is in development stage**

When the pyqt5 module is not found the error will come as shown in fig

****

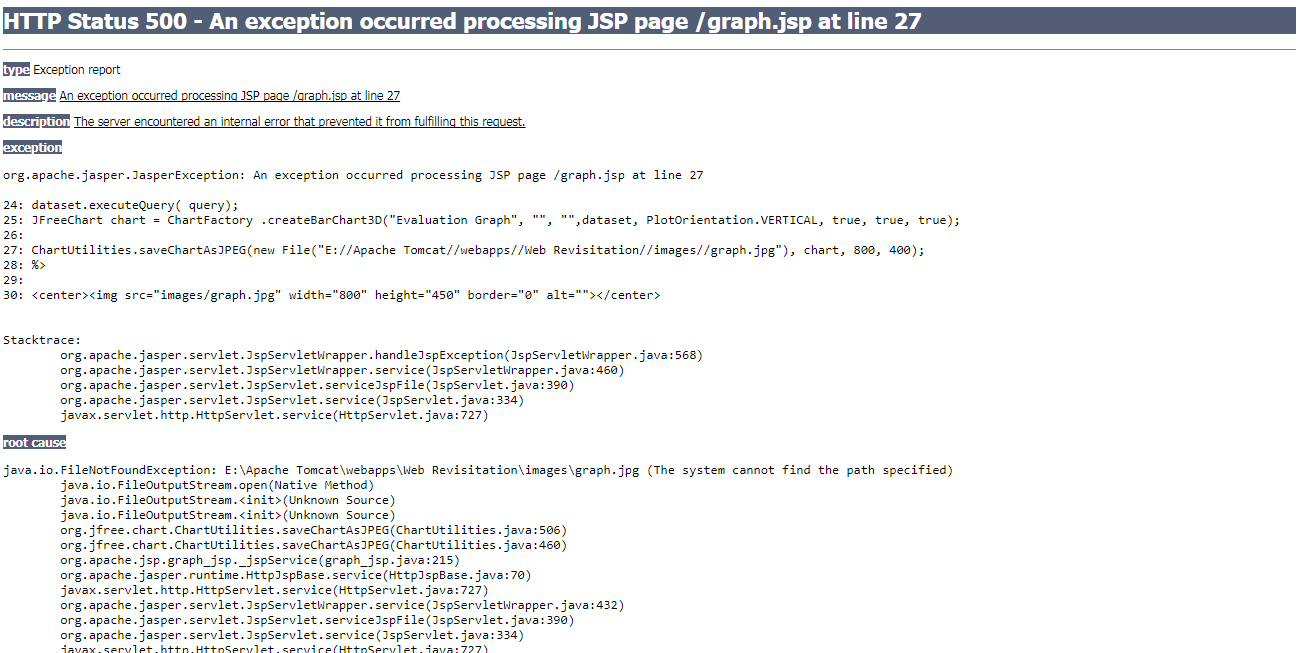
When the mysql module is not found the error will come as shown in fig

**Class version Error in our application**

****

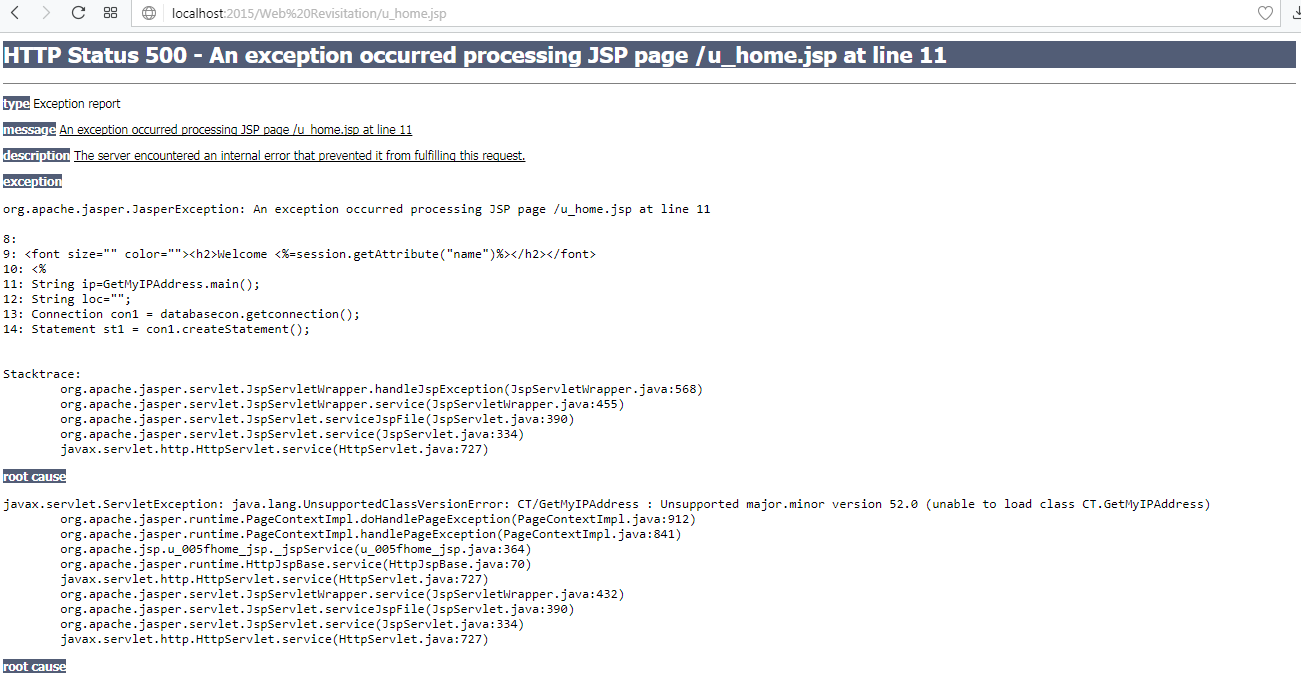
This Error Come when we move our application from one system to other and mainly when we version issues in the software’s we us

**Path related error in our application**

****

This Error Came when I have Performance Metrics to show in graph when I missed my server directly path in the system so we got this error in the applicant in development stage

**Server Connection Error**

****

**BROWSER COMPATIBILITY TESTING**

****Cross browser compatibility testing has been gaining a lot of traction in recent years and there is a reason for it. While technology is evolving rapidly, people aren’t. A significant amount of people are resistant to changes, or more specifically, “have an aversion to upgrading their tech”. In this scenario, it’s browser compatibility testing that enables companies to ensure that no customer is left behind or has an experience that is not desired. So even though browsers like Google Chrome and Firefox dominate the market, there are people using their older versions, or other browsers. And their numbers are too high to be ignored.

#### ****What is cross browser compatibility testing?****

Cross browser compatibility testing is a non-functional form of testing, which emphasizes on availing your website’s basic features and functionality to users on different browser-OS combinations, devices, and assistive tools.

#### ****How does it impact your application?****

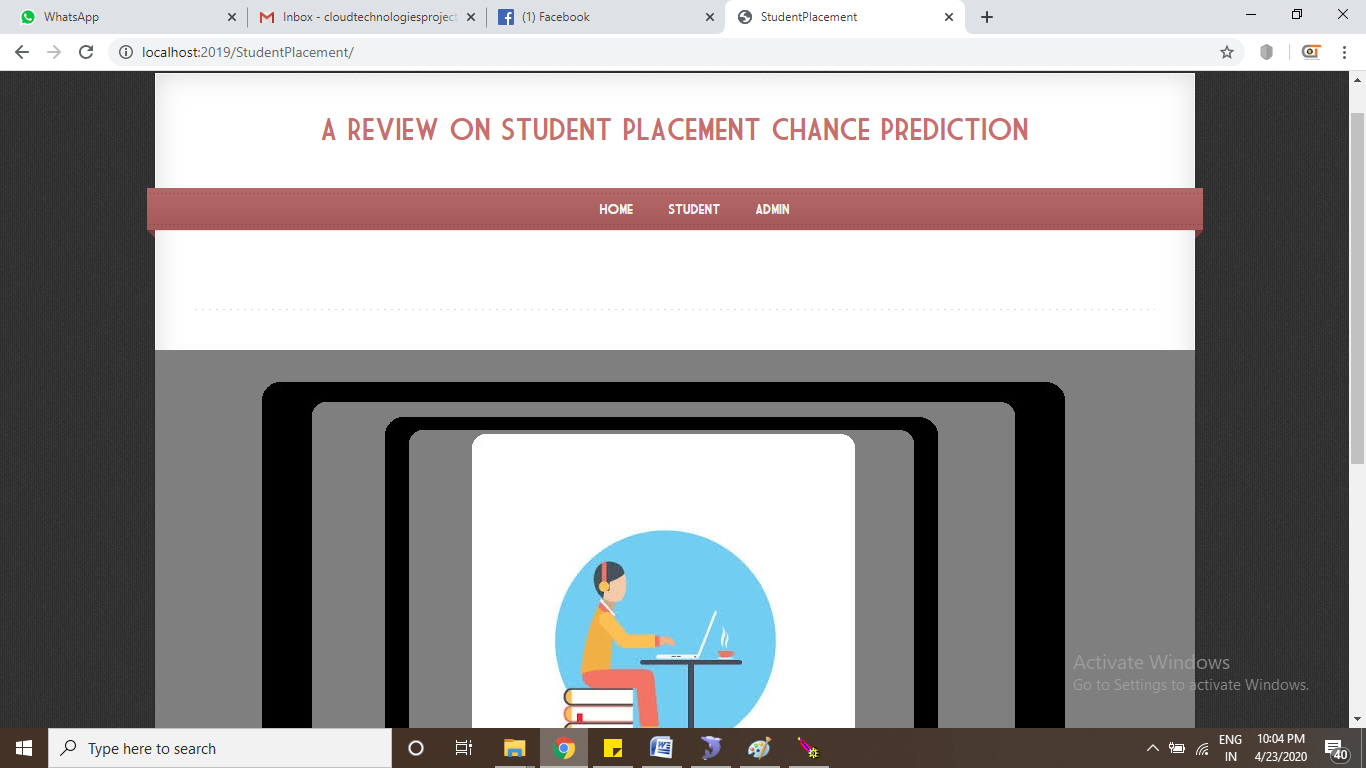
Not all browsers and devices work on the same configuration; they face browser compatibility issues on different levels. This inconsistency is the reason why you might observe the lack of application uniformity across browsers and devices. You would not want a section of your prospective users to not be able to access the application features.

That is what makes [cross browser testing](https://www.browserstack.com/live) important. If your website is not tested and debugged on different platforms and browsers, it won’t work the same on all of them, causing inconvenience to the users, subsequently impacting your business.

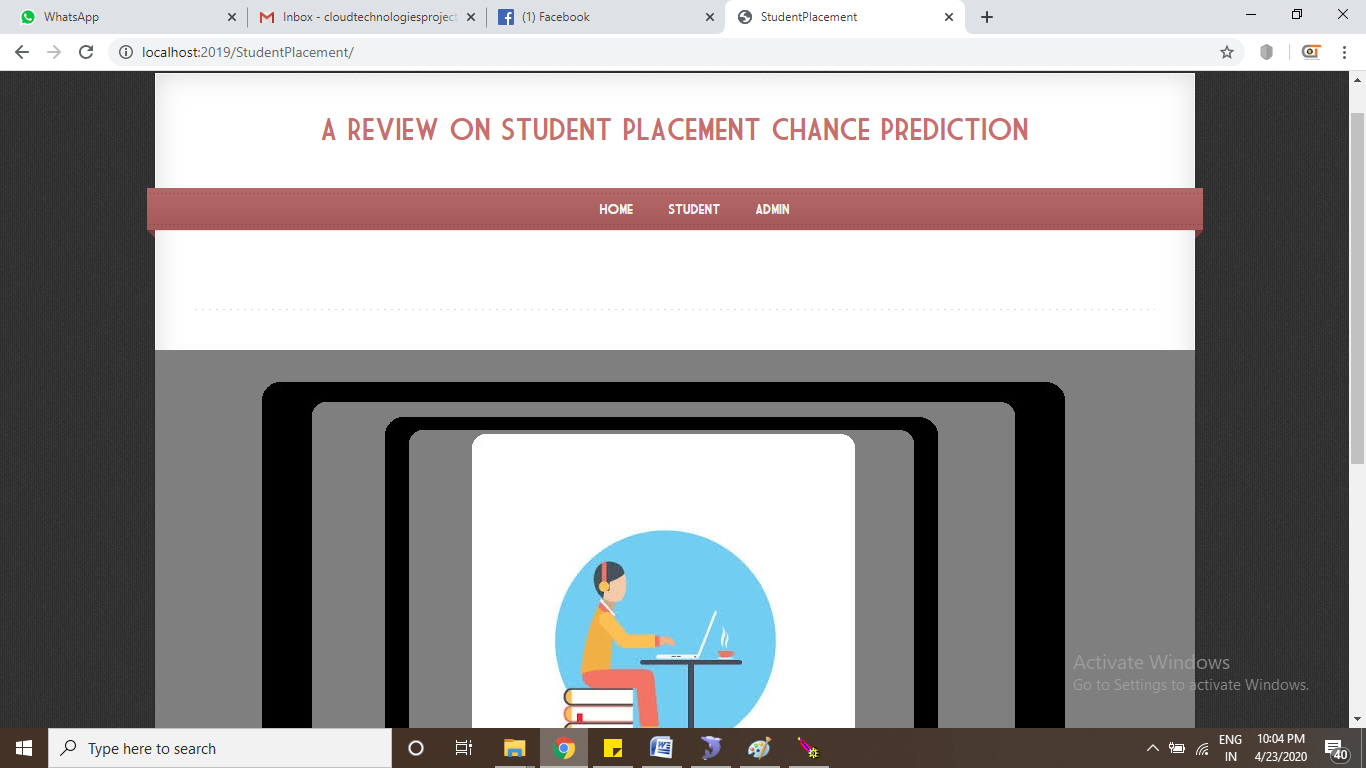
#### ****Which browsers to choose for cross browser testing?****

Since it’s impossible to test on every possible browser-device combination, you need to shortlist the most important ones to test your web application on. As of December 2018, Google Chrome has the largest number of users. It accounts for about 70.95% of the market. Firefox comes second with a market share of 10.05%, while others such as IE, Safari, and Edge has a market share of approx 4-5% each.

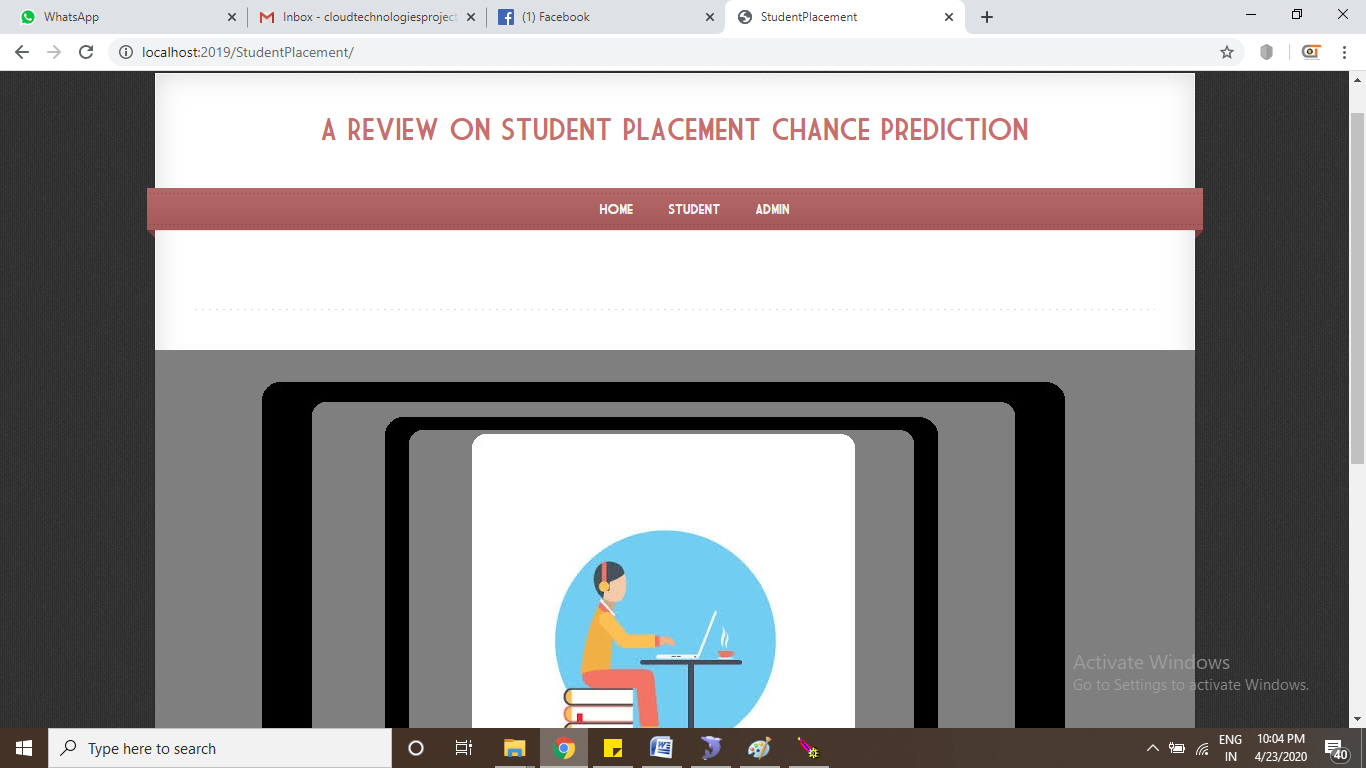
**Result of My Application on UC browser**

****

**Result of my Project in chrome**

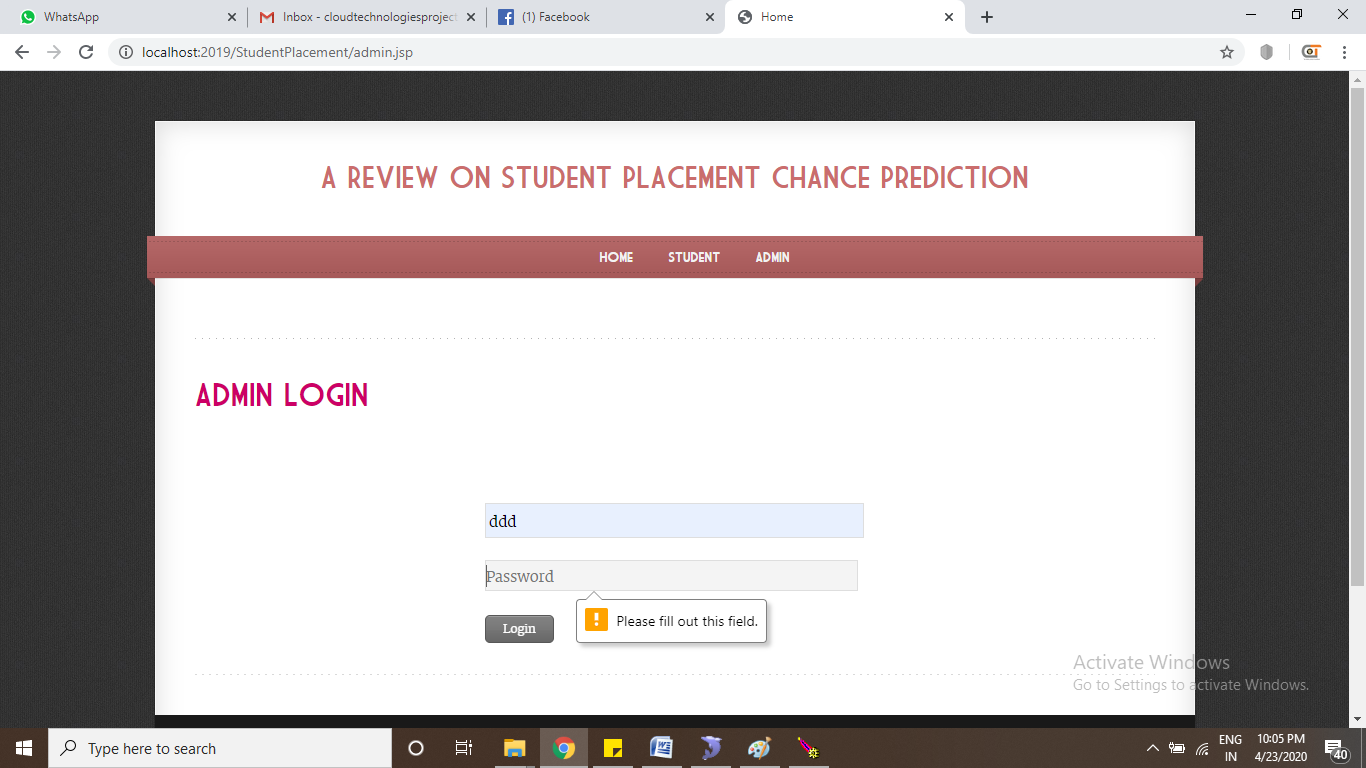
****

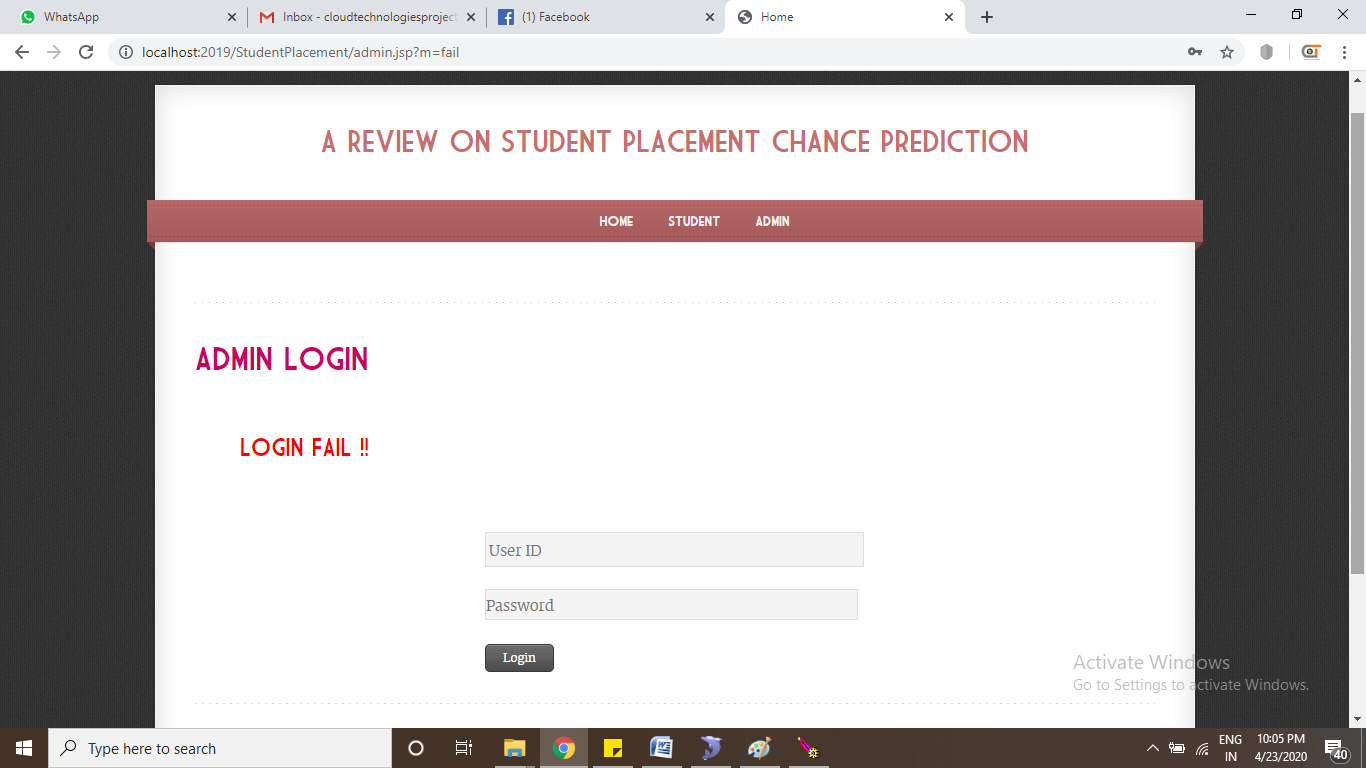
**Result of my Project in Opera**

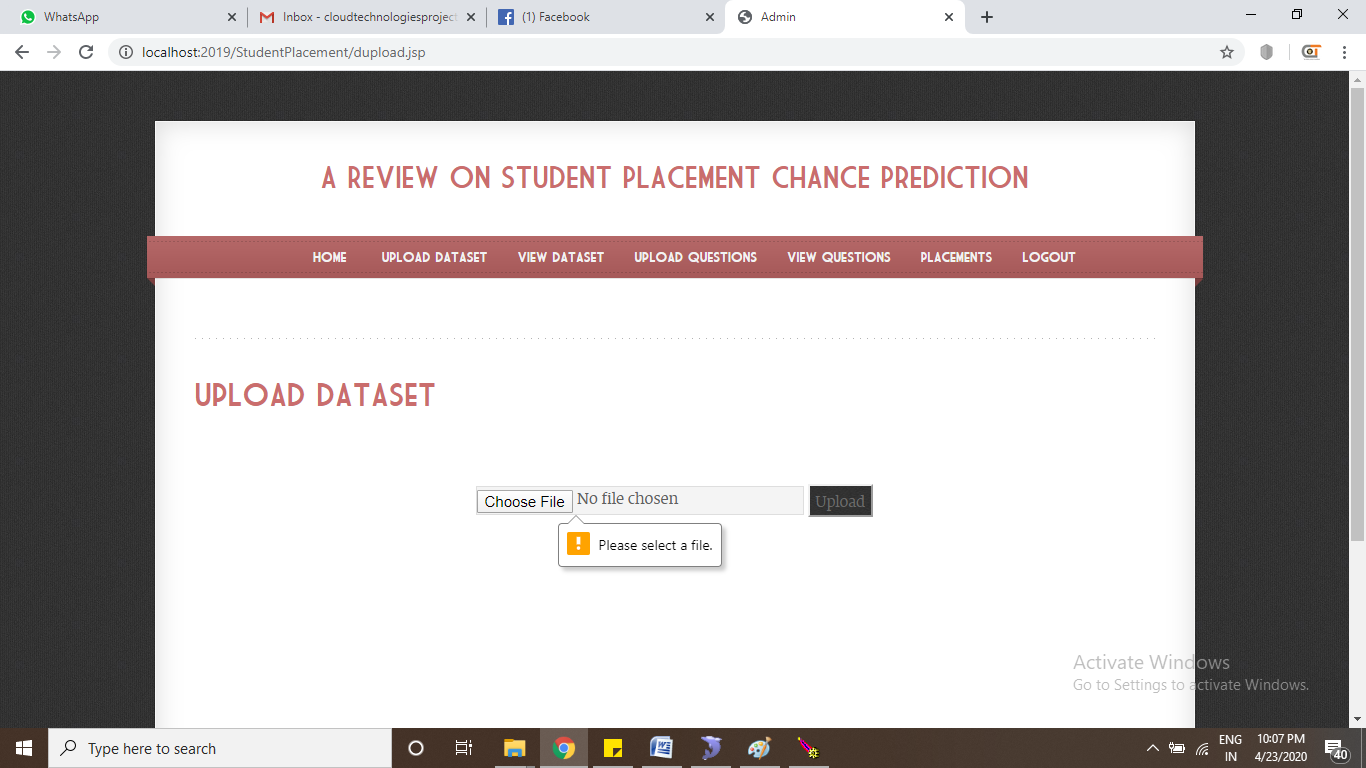
****

**VALIDATION TESTING**

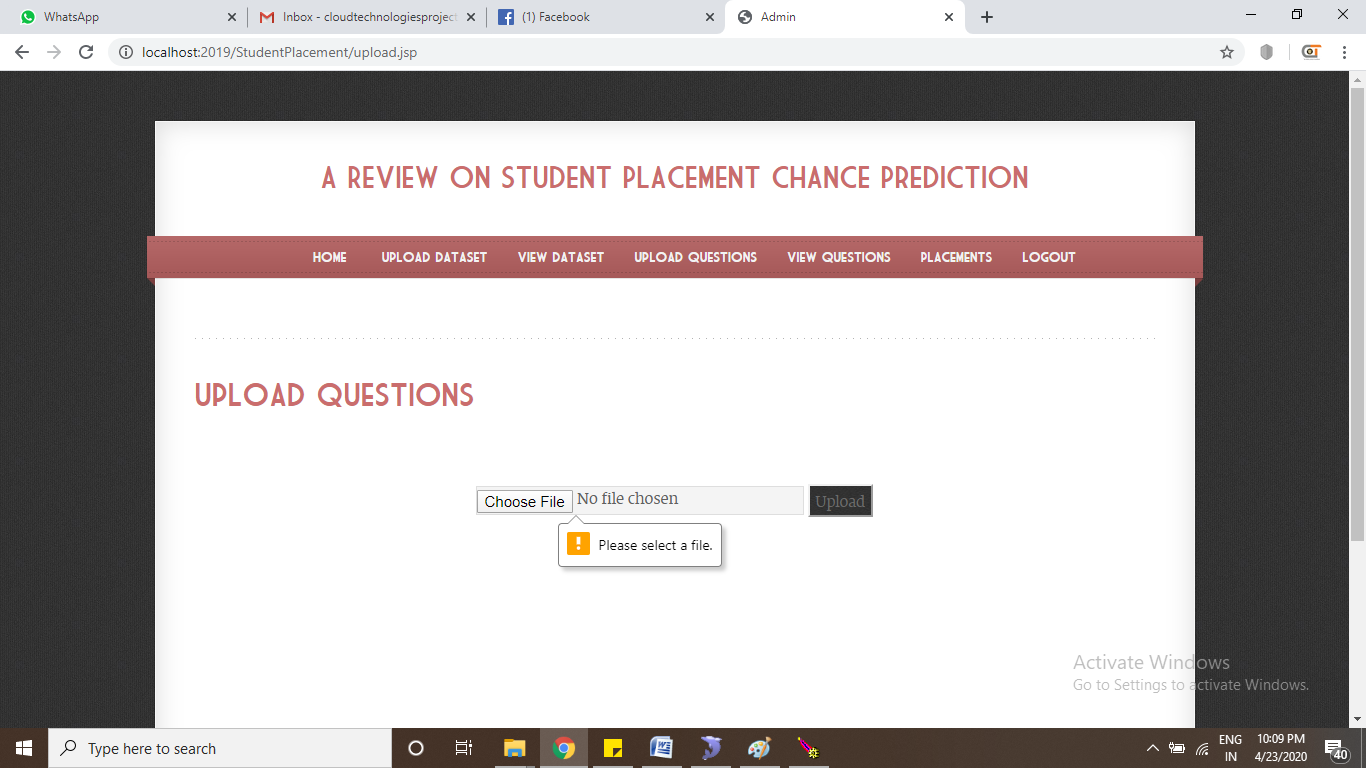
**User ID “NULL” Validation testing**

**Password “NULL” Validation Testing**

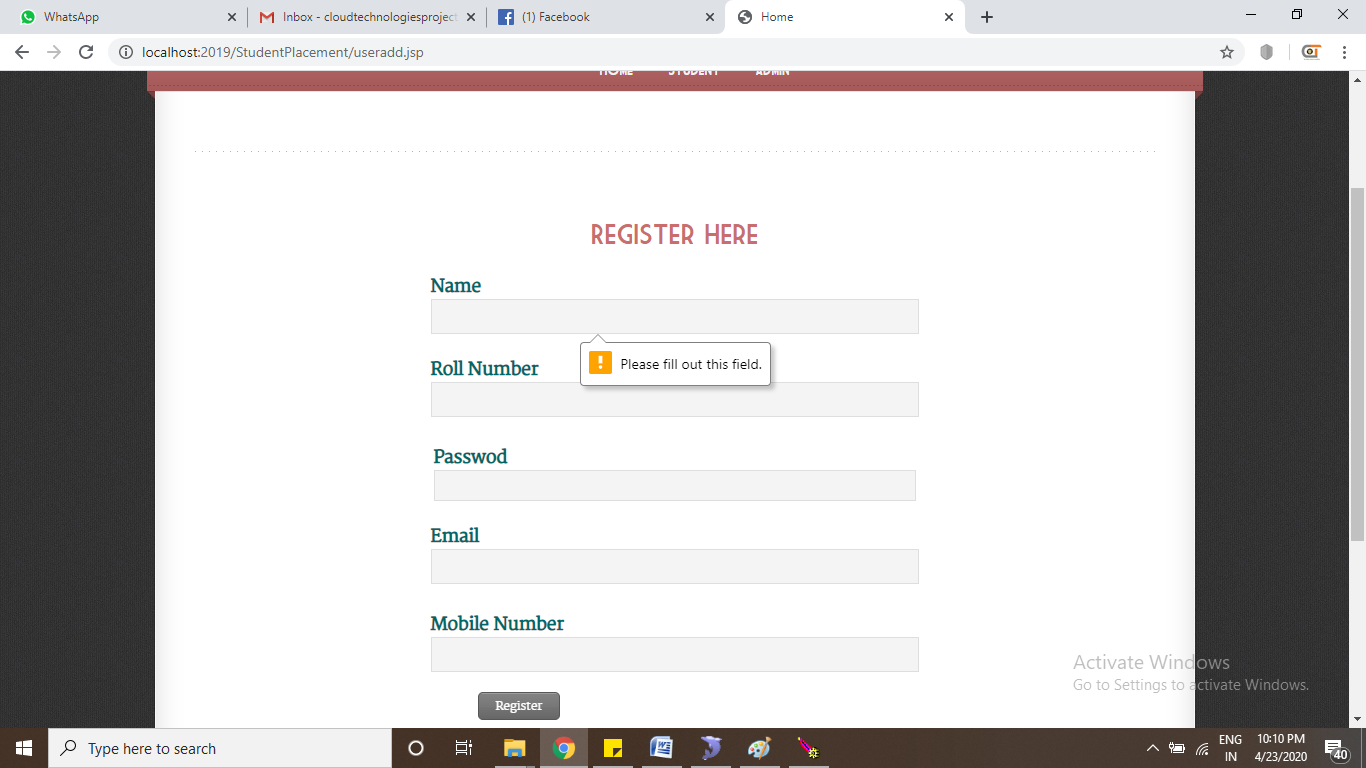
**Invalid User id or Password wrong “Login FAIL”**

****

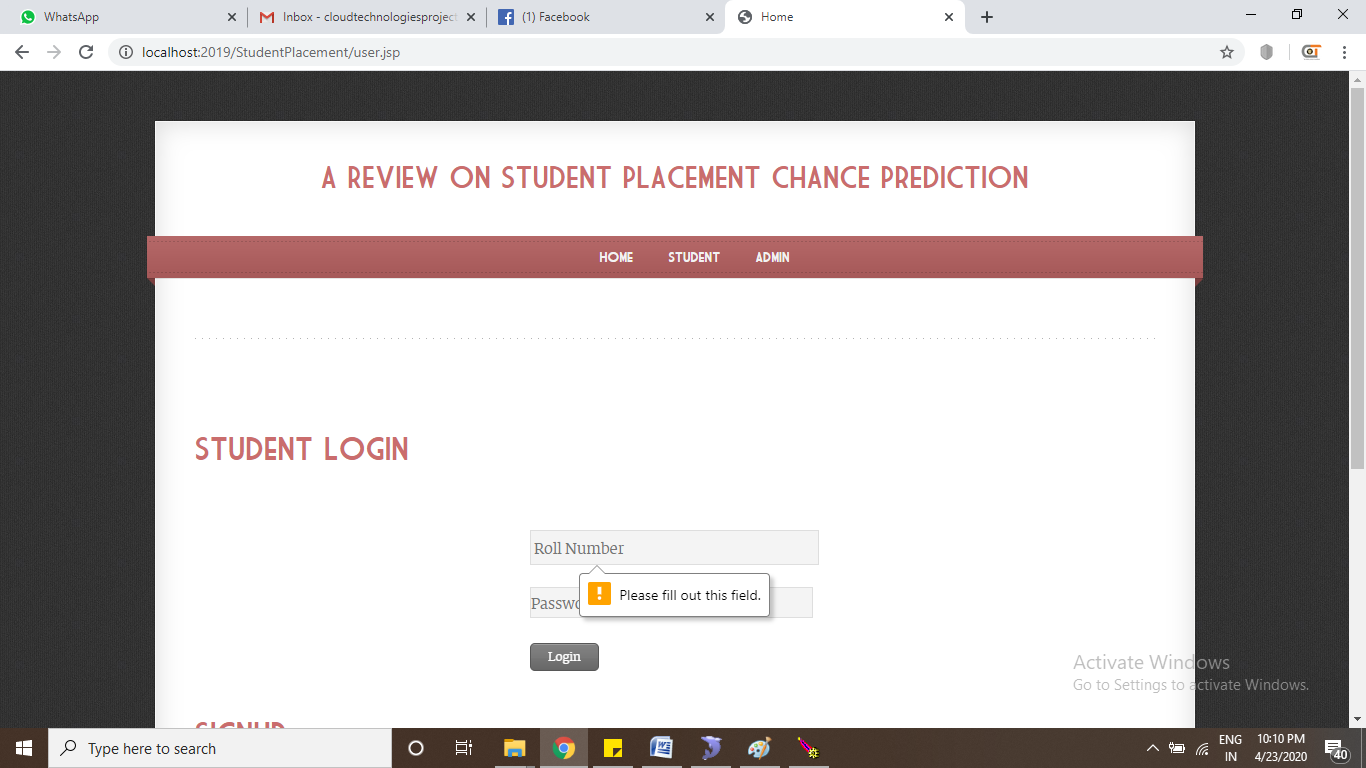
**Dataset Upload “NULL” Validation Testing**

****

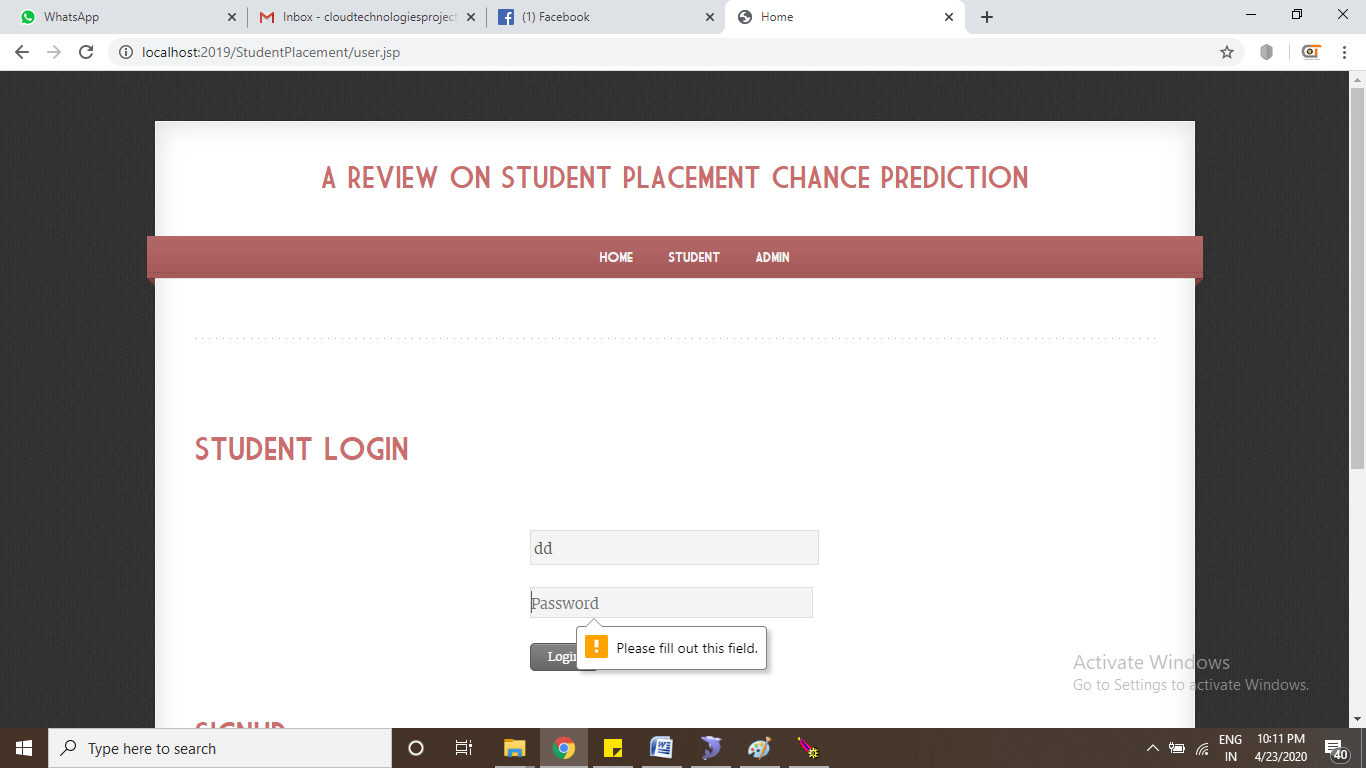
**Upload Questions “NULL” Validation Testing**

****

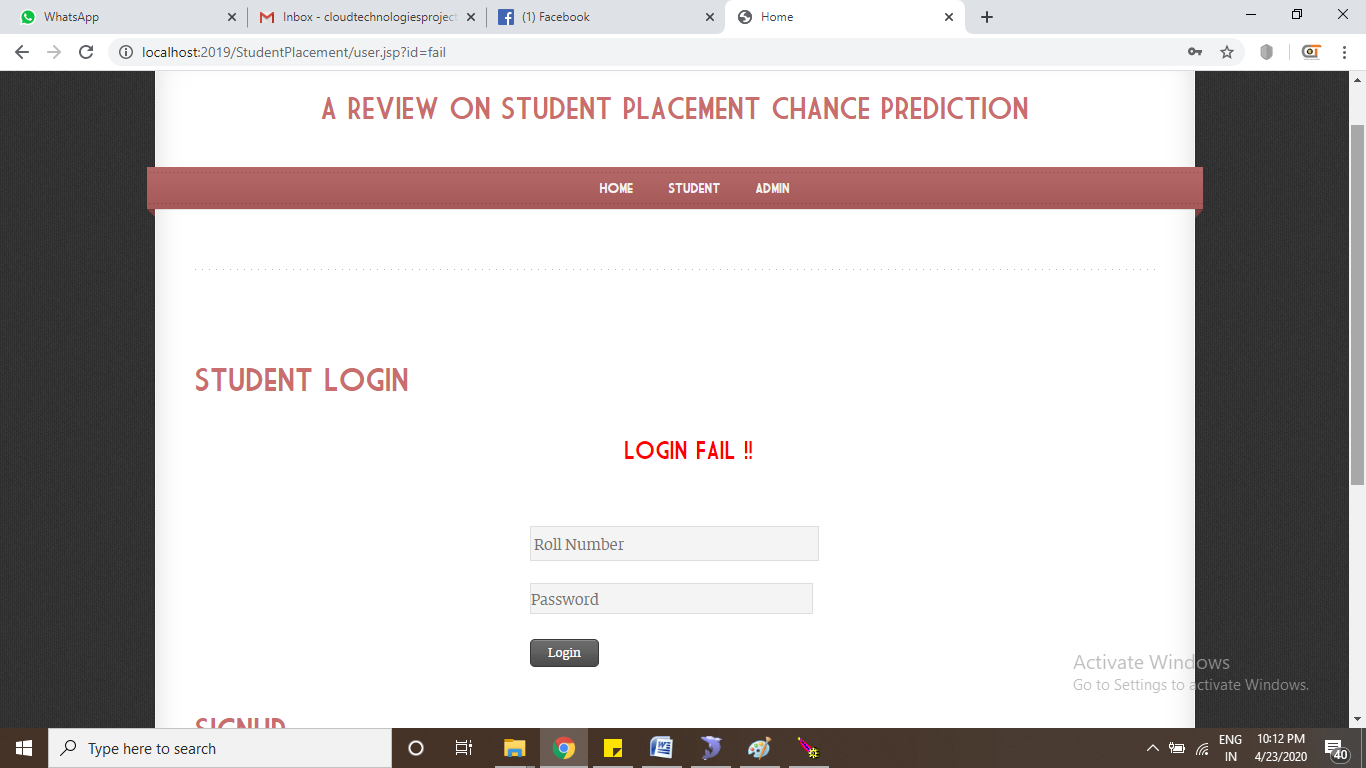
**Register “NULL” Validation testing**

****

**Student Login Roll Number “NULL” Validation Testing**

****

**Student Login Password “NULL” Validation Testing**

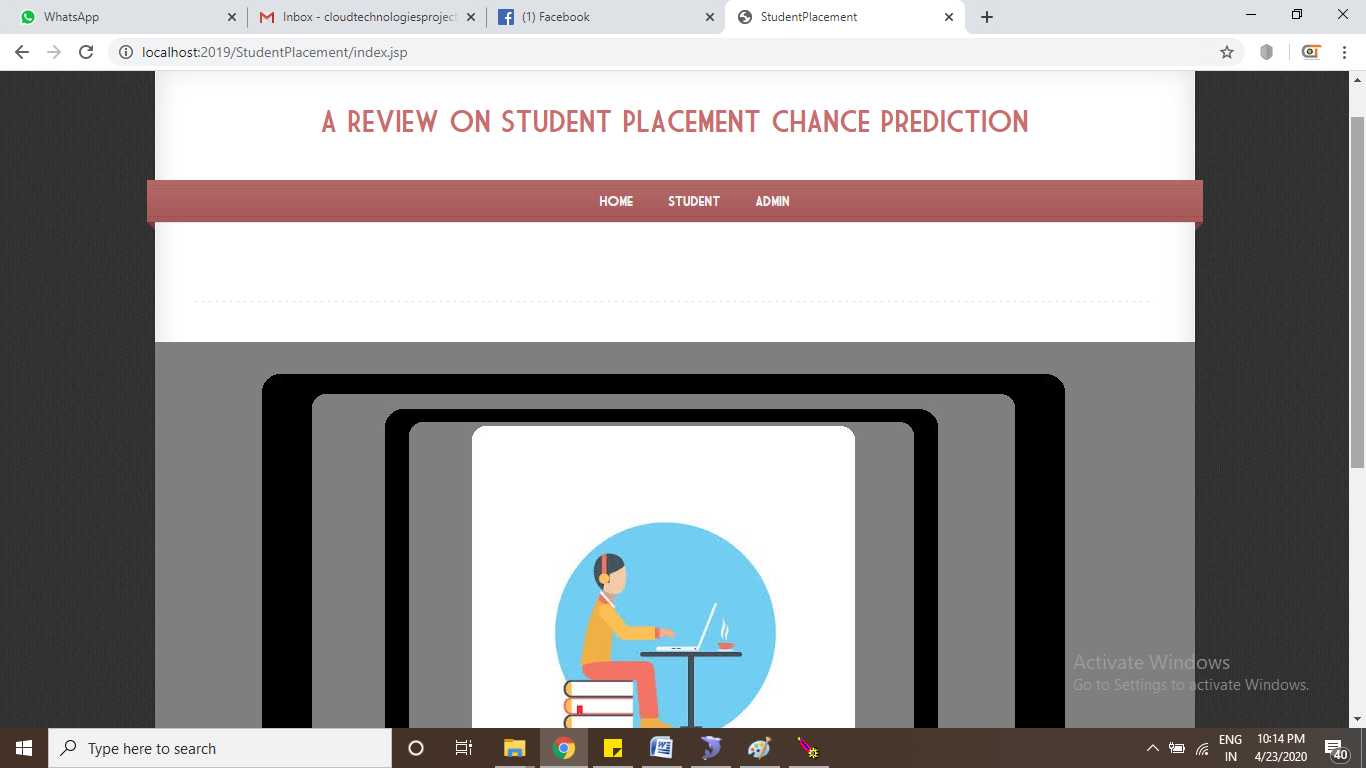
****

**Roll Number or Password wrong “Login Fail”**

**Test Case**

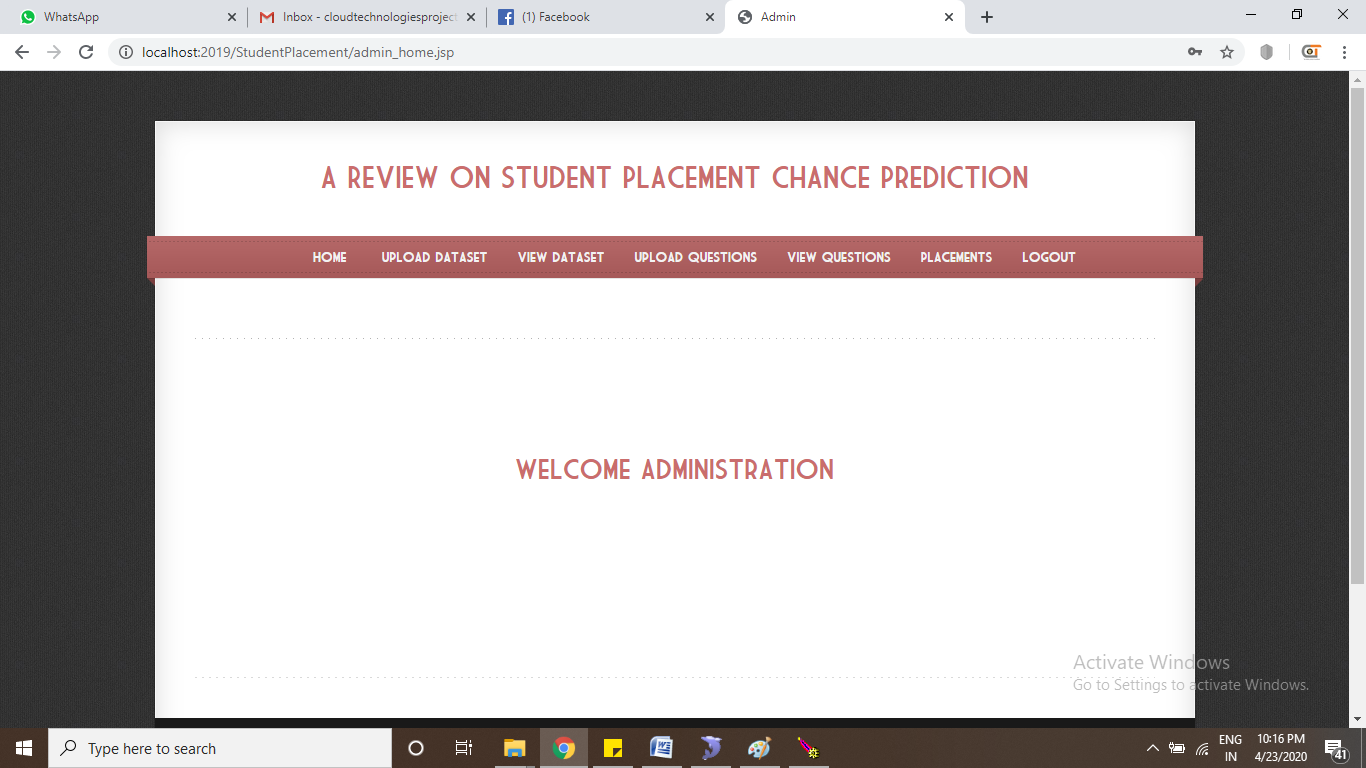
|  |  |  |  |
| --- | --- | --- | --- |
| **SNO** | **Test Case** | **Input** | **Result** |
| 1 | User Name | NULL | YES |
| 2 | Password | NULL | YES |
| 3 | User Name or Password | Invalid | Error Message |
| 4 | Upload Dataset | NULL | YES |
| 5 | Roll Number | NULL | YES |
| 6 | Password | NULL | YES |
| 7 | Name | NULL | YES |
| 8 | Email | NULL | YES |
| 9 | Email | Email Paten | Validation Message |
| 10 | Contact | NULL | YES |
| 11 | Contact | Max Numbers10 | Validation Message |
| 12 | Test Data set upload | NULL | YES |

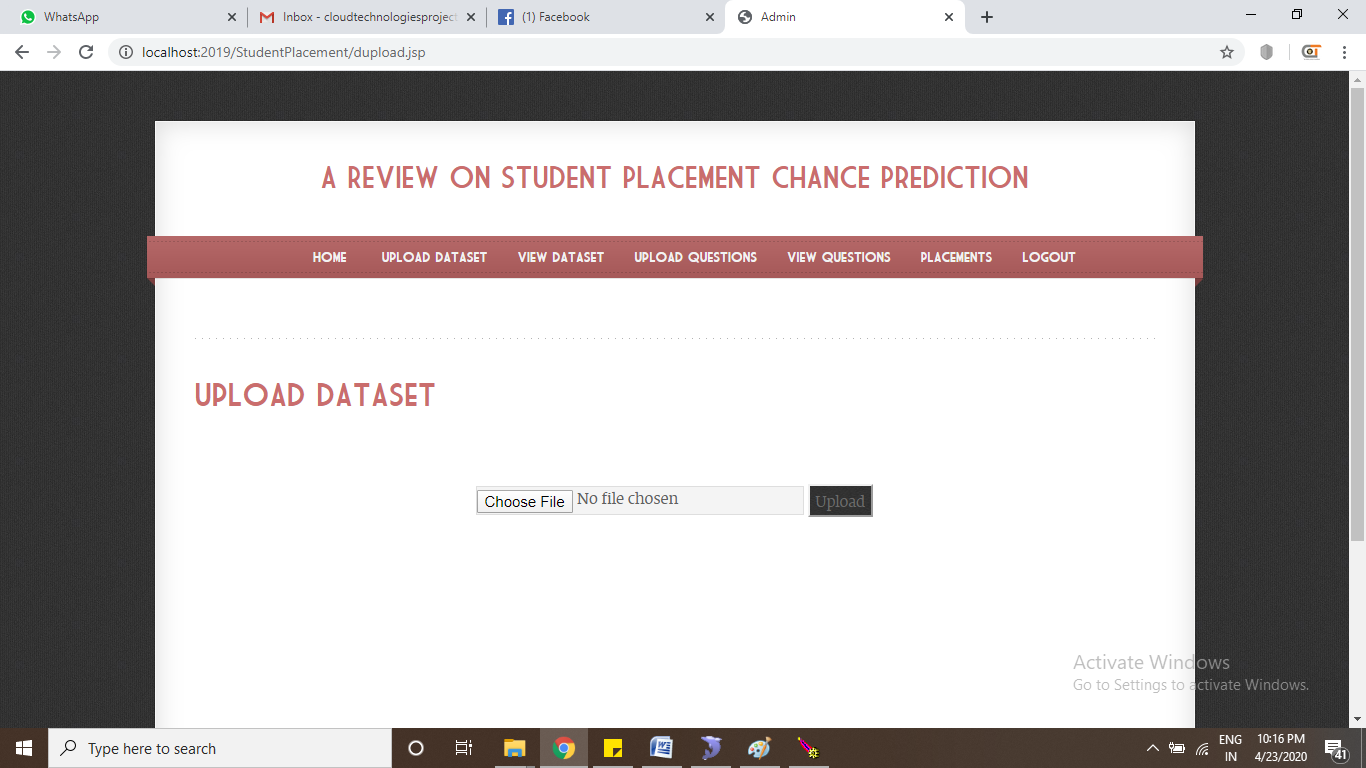
**SCREEN SHORTS**

****

**Home Page**

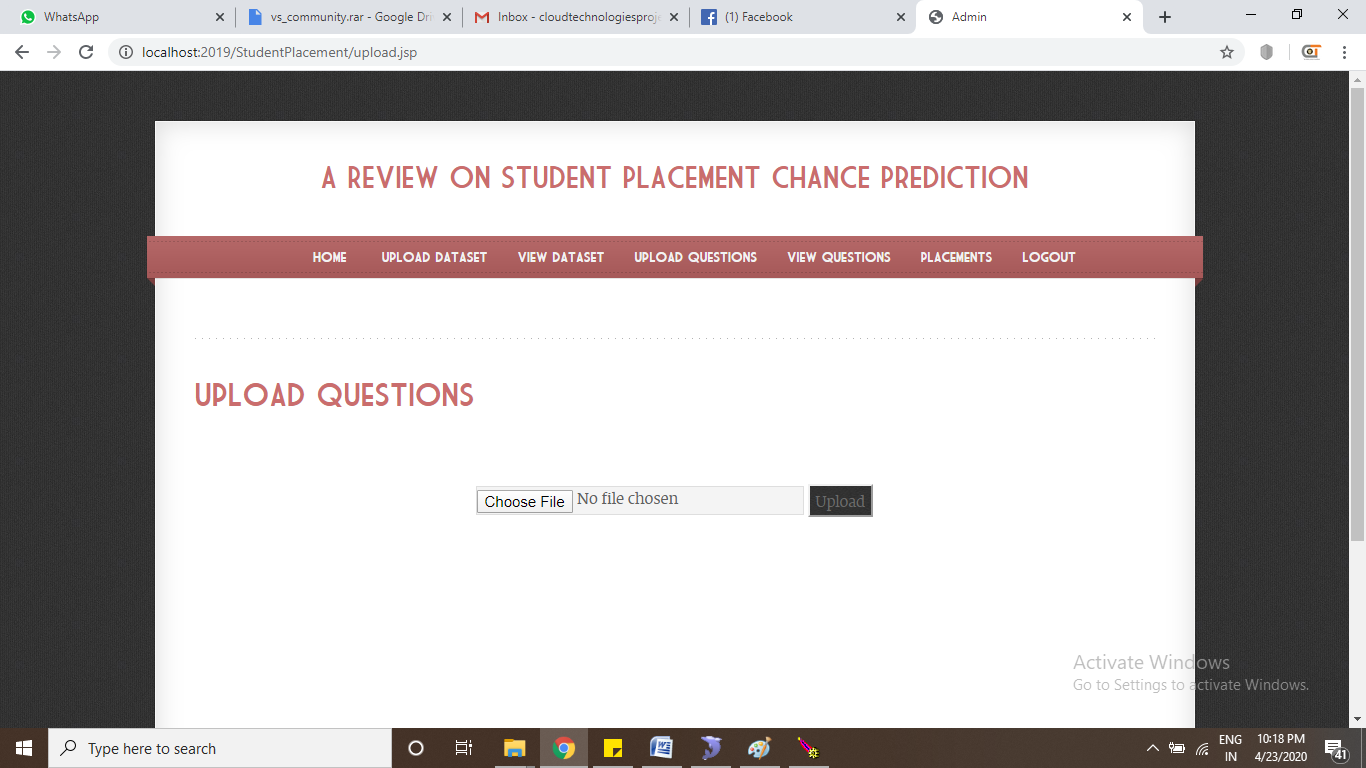
**Admin Login**

**Admin Home**

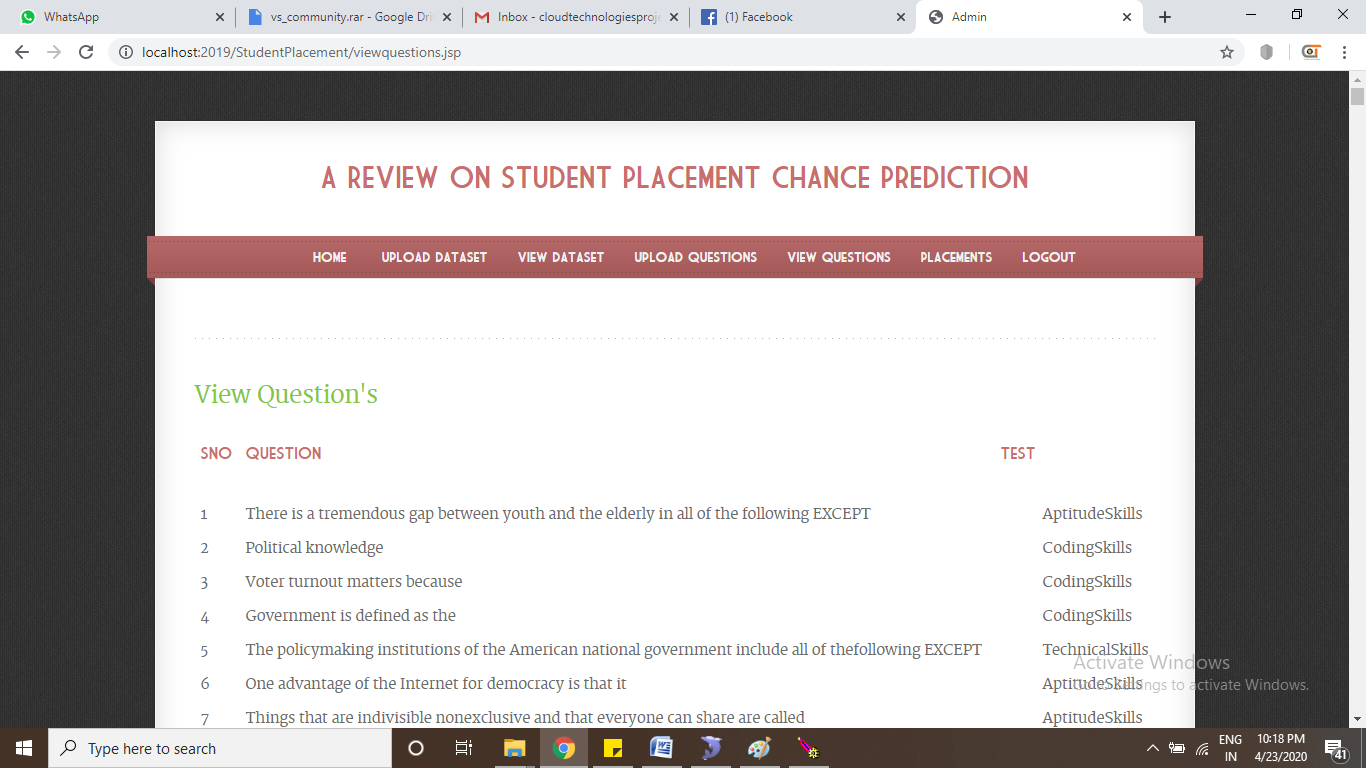
****

**Upload Dataset**

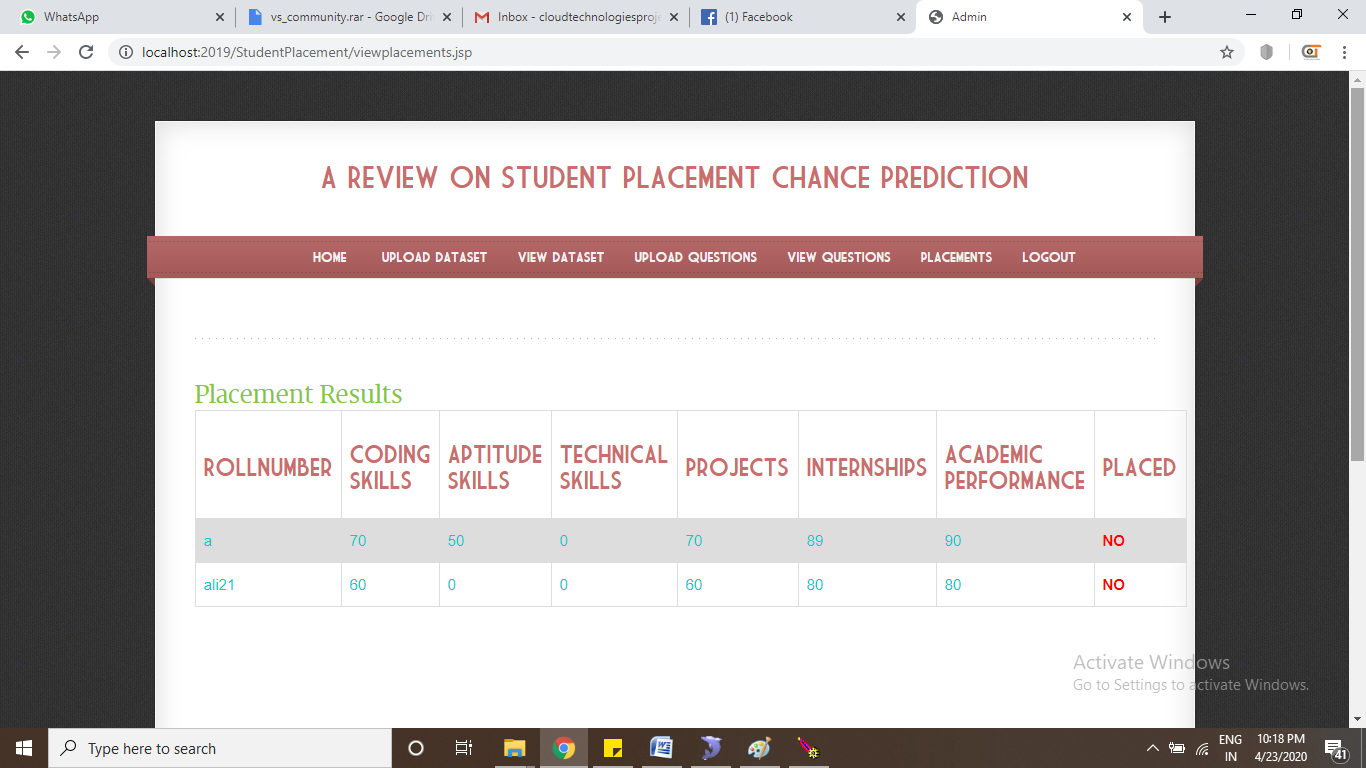
**View Students Dataset**

****

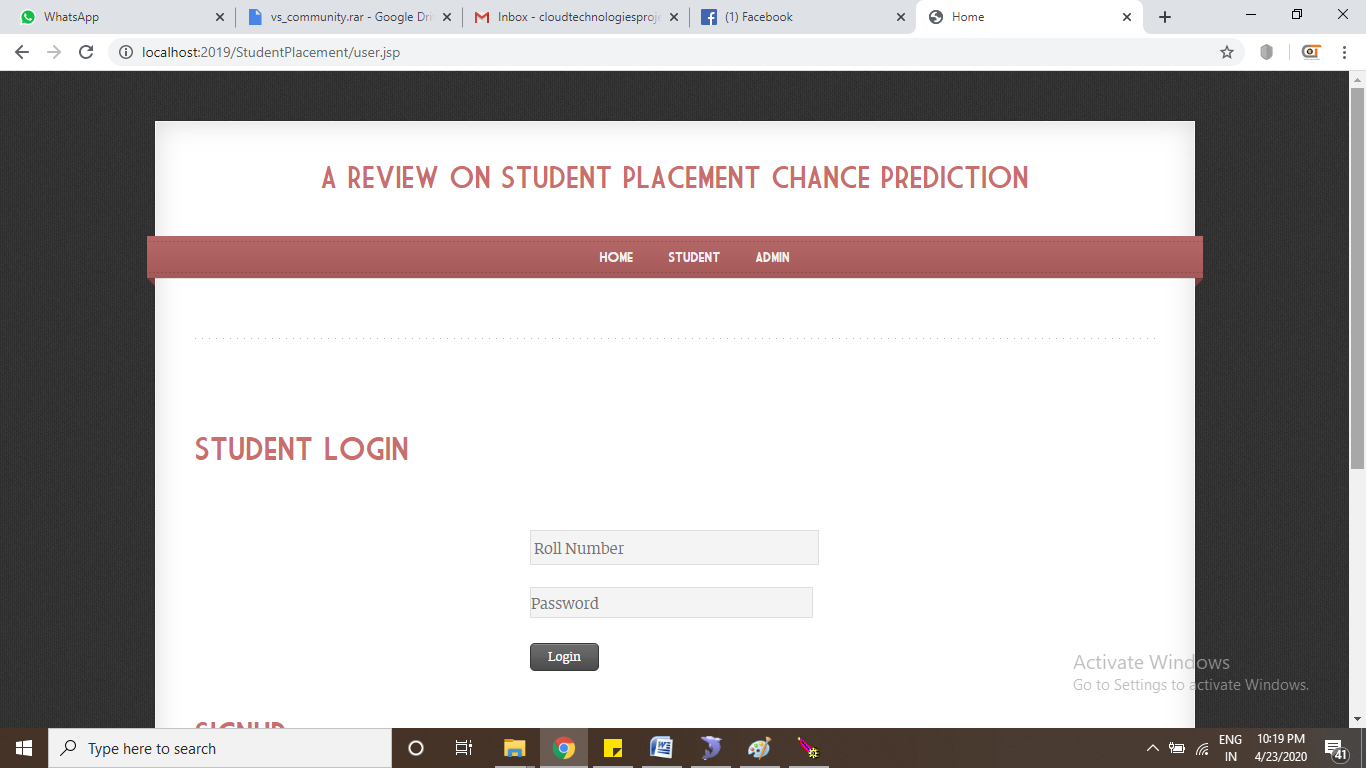
**Upload Questions Data set**

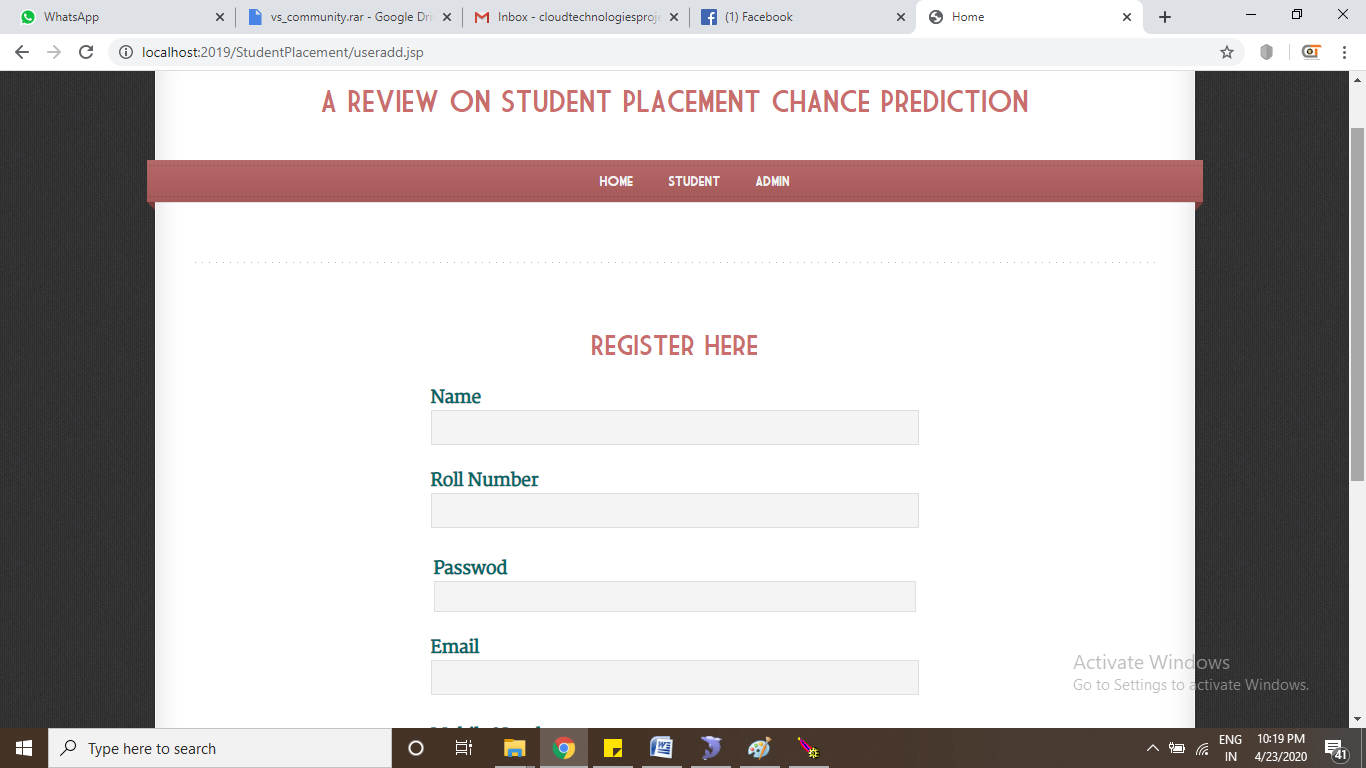
****

**View Question's**

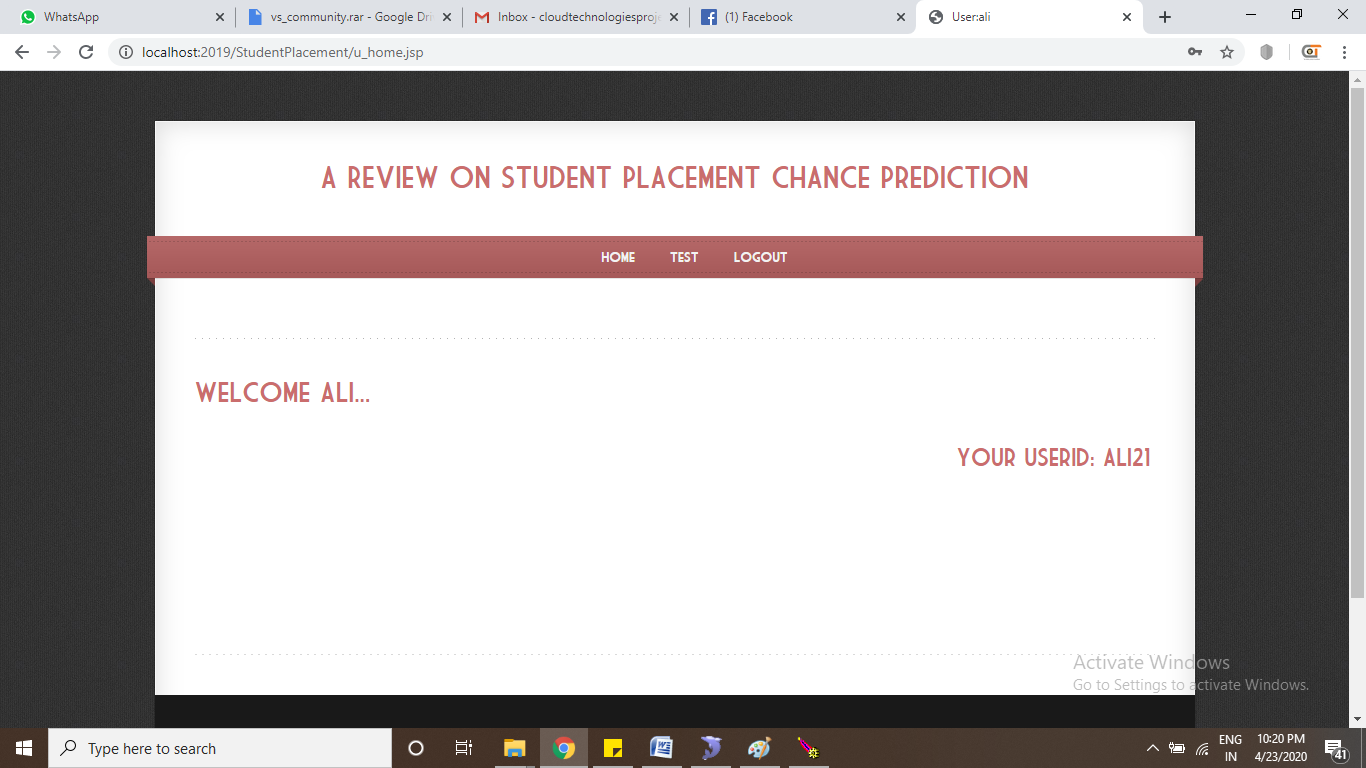
****

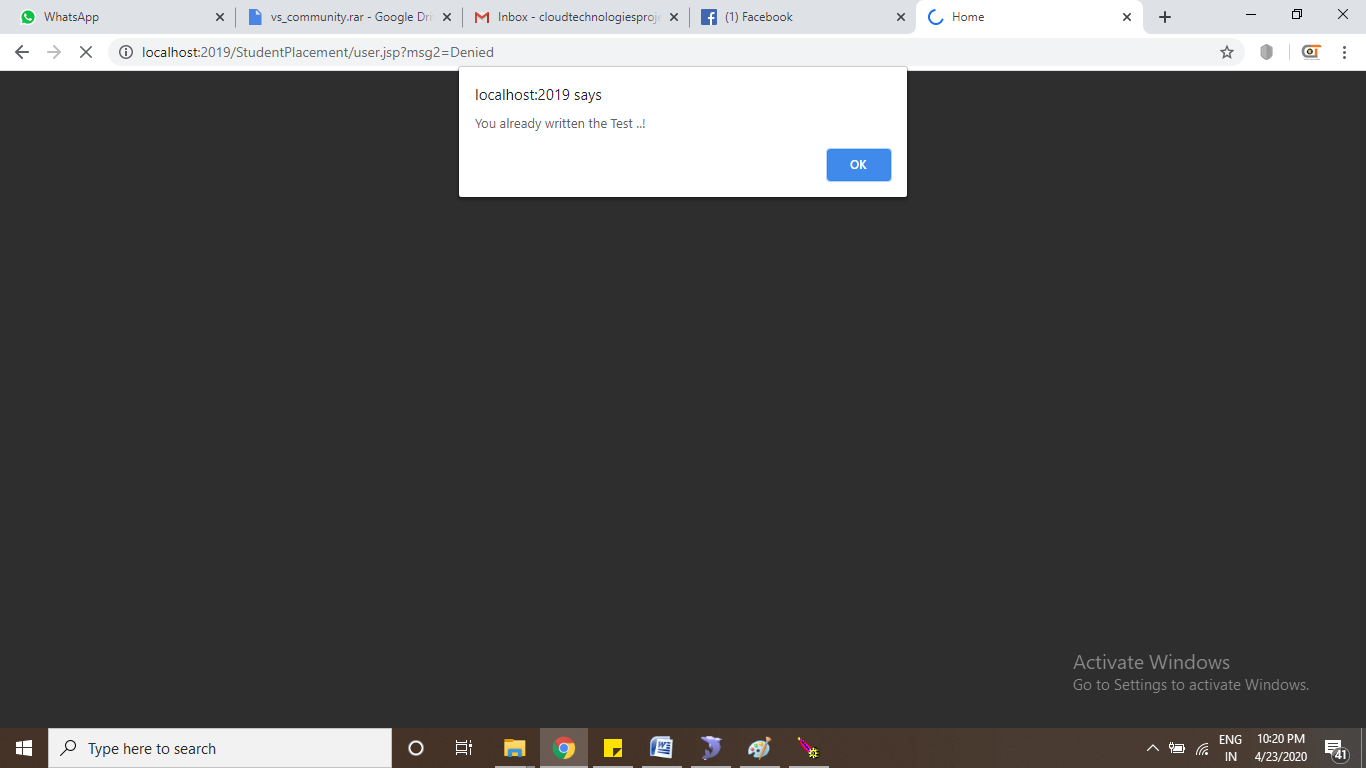
**Placement Results**

**** **Student Login**

****

**Register Here**

**User Home Page**

****

**Exam Page**

**CONCLUSION**

A detailed study was conducted based on different placement prediction models. From the study it is clear that the student dataset containing academic and placement details are a potential source for predicting the future placement chances. This prediction can enlighten students to identify their capabilities and improve accordingly. This system also helps in the academic planning of an institution to prepare proper strategies and improve the placement statistics for the future years.

**Future Work**

It would of great help if we revise and update our curriculum and other extra activities for each semester in accordance with the public, private and government sector requirement. We can also predict which company picks which category of students. Make a list of skill a particular company looking for, then on the basis of that we can train our student. These traits will make prediction process more accurate.

**REFERENCES**

[1] J. Luan, “Data mining and its applications in higher education”, New Dir. Inst. Res, 113:17–36, 2002. [2] A.S. Sharma, S. Prince, S. Kapoor, K. Kumar, “PPS – Placement prediction system using logistic regression”, IEEE international conference on MOOC, innovation and technology in education (MITE), pp 337-341,2014.

[3] Thangavel, S.Bkaratki, P. Sankar, “Student placement analyzer: A recommendation system using machine learning”, Advances in Computing and Communication Systems (ICACCS-2017) International Conference on. IEEE, 2017.

[4] R. Sangha, A. Satras, L. Swamy, G. Deshmukh, “Students Placement Eligibility Prediction using Fuzzy Approach”, International Journal of Engineering and Techniques , Volume 2, Issue 6, Dec 2016.

[5] H. Bhatt, S. Mehta, L. R. D'mello, “Use of ID3 Decision Tree Algorithm for Placement Prediction”, International Journal of Computer Science and Information Technologies (IJCSIT), vol. 6, pp. 4785-4789, 2015.

[6] T. Jeevalatha, N. Ananthi, D. Saravana Kumar, “Performance analysis of undergraduate students placement selection using Decision Tree Algorithms”, International Journal of Computer Applications, vol. 108, pp. 0975-8887, December 2014.

[7] Bharambe, Yogesh, “Assessing employability of students using data mining techniques”, Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on. IEEE, 2017.

[8] P. Khongchai, P. Songmuang, “Random Forest for Salary Prediction System to Improve Students Motivation”, 12th International Conference on SignalImage Technology & Internet-Based Systems (SITIS), pp. 637-642, 2016.

[9] Pruthi, P. Bhatia, “Application of Data Mining in Predicting Placement of Students”, International Conference on Green Computing and Internet of Things (ICGCIoT), 2015.

[10] P. Guleria, M. Sood, “Predicting Student Placements Using Bayesian Classification”, International conference on Image Information Processing, IEEE Computer Society, 2015, pp. 109- 112.