SAVITRIBAI PHULE PUNE UNIVERSITY A PARTIAL PROJECT REPORT ON

"Personality Prediction System Using Resume"

SUBMITTED TOWARDS THE SAVITRIBAI PHULE PUNE UNVERSITY,
PUNE IN PARTIAL FULFILLMENT FOR THE AWARD OF THE DEGREE
OF

BACHELOR OF ENIGINNERING

IN

COMPUTER ENGINEERING

Name: Mr. Awachar Piyush Madhavrao

Name: Mr. Dange Prajwal Nanasaheb

Name: Mr. Kumbhar Sangram Shahaji

Name: Mr. Pawar Ashish Sudam

Exam No:

Exam No:

Exam No:

UNDER THE GUIDENCE OF

Prof. S. S. Shaikh



DEPARTMENT OF COMPUTER ENGINEERING

Sanjivani Rural Education Society's, Sanjivani College of Engineering

[23/2020-21]

CERTIFICATE



This is to certify that the Project Entitled

"Personality Prediction System Using Resume"

Submitted By

Mr. Awachar Piyush Madhavrao
Mr. Dange Prajwal Nanasaheb
Mr. Kumbhar Sangram Shahaji
Mr. Pawar Ashish Sudam

Exam No:
Exam No:
Exam No:

is a bonafide work carried out by students under the supervision of Prof. S. S. Shaikh and it is submitted towards the fulfillment of the requirement of Bachelor of Engineering (Computer Engineering).

Prof. S. S. Shaikh [Internal Guide]

Dr. D. B. Kshirsagar [H.O.D. Computer Dept.]

Dr. A. G. Thakur
[Principal]
Sanjivani College of Engineering, Kopargaon-423603

Signature of Internal Examiner Signature of External Examiner

PARTIAL PROJECT SHEET

A Partial Project Report

On

"Personality Prediction System Using Resume"

Is successfully completed by

Mr. Awachar Piyush Madhavrao
Mr. Dange Prajwal Nanasaheb
Mr. Kumbhar Sangram Shahaji
Mr. Pawar Ashish Sudam

Exam No:
Exam No:
Exam No:

(B.E. Computer)

at

DEPARTMENT OF COMPUTER ENGINEERING

Sanjivani College of Engineering

Kopargaon - 423 603.

SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE ACADEMIC YEAR 2020-2021

Prof. S. S. Shaikh [Internal Guide]

Dr. A. B. Pawar

[Project Coordinator]

Dr. D. B. Kshirsagar

[H.O.D. Computer Dept.]

ABSTRACT

Human Resource Management is apparently supported by and provided with more opportunities by the development of Job Characteristics Model (JCM) which in turn is based on the concept of modern job design. Fortunately, the development in modern information system, digital technologies, the universal access of electronic technology and internet led to the inclination of the global Human Resource Management development and make the system more applicable. Following the trend, the proposed system tries to design a plan to integrate Job Characteristics Model into E-HR system to search for a new model of efficient operation on Human Resource Management in the Internet Age. In this project, we present a set of techniques that makes the whole recruitment process more effective and efficient. We have implemented a system that ranks the candidates based on weight-age policy as well as an aptitude test. Today there is a growing interest in the personality traits of a candidate by the organization to better examine and understand the candidate's response to similar circumstances. Therefore, the system conducts a personality prediction test to determine the personality traits of the candidate. Finally, it presents the results of the candidates to the recruiter who evaluates the top candidates and shortlists the candidate.

ACKNOWLEDGEMENT

It is with the greatest pleasure and pride that we present this report. At this moment of triumph, it would be neglect all those who helped us in the successful partially completion of this project. We are very much thankful to our respected project guide **Prof.S.S Shaikh** and project coordinator **Prof. A. B. Pawar** for his ideas and help proved to be valuable and helpful during creation of partial project report and set us in the right path.

We would also like to thank all the faculties who have cleared all the major concepts that were involved in the understanding techniques behind our project. Lastly, we are thankful to our friends who shared their knowledge in this field with us.

Mr. Piyush Madhavrao Awchar Mr. Prajwal Nanasaheb Dange Mr. Sangram Shahaji Kumbhar Mr. Ashish Sudam Pawar (B.E. COMPUTER)

Contents

1	1.1 1.2 1.3 1.4 1.5	Releva Literat Scope	n 1 m Definition 1 ant Theory 1 ure Review 1 ives 2 ives 2		
2	REC	DUIREN	MENT ANALYSIS 3		
	2.1	_	rement Specification		
		2.1.1	Normal Requirements		
		2.1.2	Expected Requirements		
		2.1.3	Excited Requirements		
	2.2	Validat	tion of Requirement		
		2.2.1	Validation of Normal Requirements		
		2.2.2	Validation of Expected Requirements		
		2.2.3	Validation of Excited Requirements		
	2.3	-	n Requirements		
		2.3.1	Software Requirements		
		2.3.2	Hardware Requirements		
3	SYS	TEM M	MODEL 6		
	3.1		s Model		
		3.1.1	Selected Model:Incremental Model 6		
		3.1.2	Why to use Incremental Models		
		3.1.3	Disadvantages of Incremental Models		
		3.1.4	Breakdown Structure(Module)		
		3.1.5	System Architecture		
	3.2	Project	t Estimation		
		3.2.1	Estimation in KLOC		
		3.2.2	Efforts(E):		
		3.2.3	Development Time (In months):		
		3.2.4	Total Time Required for Project Development:		
		3.2.5	Number of Person Required (N):		
4	SYSTEM DESIGN 13				
•	4.1		t Scheduling And Tracking		
		4.1.1	Task Identification		
		4.1.2	Project Schedule		
		4.1.3	Project Table And Time-Line Chart		

		4.1.4	Time-Line Chart	17
	4.2			
	4.2		sis Modeling	18
		4.2.1	Behavioral Modeling	18
		4.2.2	Functional Model	24
		4.2.3	Architectural Modeling	28
		4.2.4	Mathematical Model	30
		4.2.5	Database Table with Discription	35
			r	
5	Risk	Manag	gement	36
	5.1	-	dentification	36
	5.2		gies used to manage Risk	37
	5.3	_	rojection	37
	5.4			37
	3.4		ility Study	
		5.4.1	Cost Feasibility	38
6	Tech	nnical S	pecification	39
Te	chnic	al Speci	ification	39
	6.1	Softwa	are Requirement Specification and its justification	39
		6.1.1	Elipse	39
		6.1.2	XAMPP	39
		6.1.3	Bootstrap	40
		0.1.5	Bootstrap	+0
7	Con	clusion	and Future Scope	41
References 42				
		7.0.1	Annexure A: Project Planner and Progress Report	43
		7.0.2	Annexure B: Plagiarism Report	50
		7.0.2	rameaute b . 1 agains 11 toport	50

List of Figures

3.1	Incremental Model	7
3.2	Breakdown Structure	9
3.3	System Architecture	11
4.1	Project Work Breakdown Structure(Analysis)	13
4.2	Project Work Breakdown Structure(Implementation)	14
4.3	Expected September Timeline	17
4.4	Actual September Timeline	17
4.5	Expected October Timeline	17
4.6	Actual October Timeline	18
4.7	Expected November Timeline	18
4.8	Actual November Timeline	18
4.9	Use-Case Diagram	19
4.10	State Chart Diagram	20
4.11		21
4.12	Sequence Diagram	22
4.13	Activity Diagram	23
	Data Flow Diagram Level 0	24
4.15	Data Flow Diagram Level 1	25
4.16	Data Flow Diagram Level 2	26
4.17	Control Flow Diagram	27
4.18	Component Diagram	28
4.19	Deployment Diagram	29
4.20	Venn Diagram	33
7.1	Plagiarism Scan Report 1	50
7.2	Plagiarism Scan Report 2	50
7.3	Plagiarism Scan Report 3	50
7.4	Plagiarism Scan Report 4	51
7.5	Plagiarism Scan Report 5	51
7.6	Plagiarism Scan Report 6	51
7.7	Plagiarism Scan Report 7	51

List of Tables

3.1	Estimation of KLOC	12
	Project Task Table	
5.1	Risk Table along with RMMM Plan	38

CHAPTER 1

Introduction

1.1 Problem Definition

There is a huge workload on the human resource department to select the right candidate for a particular job profile which in turn would provide experts workforce for the organization from a large pool of candidates.

1.2 Relevant Theory

There is a huge workload on the human resource department to select the right candidate for a particular job profile which in turn would provide experts workforce for the organization from a large pool of candidates.

1.3 Literature Review

1. **Title:** Integrated E-Recruitment System for Automated Personality Mining and Applicant Ranking

Author and Year: Faliagka et al, 2014

Methodology: An automated candidate ranking was implemented by this system. It was based on objective criteria that the candidate's details would be extracted from the candidate's LinkedIn profile. The candidates' personality traits were automatically extracted from their social presence using linguistic analysis. The candidate's rank was derived from individual selection criteria using Analytical Hierarchy Process (AHP), while their weight was controlled by the recruiter (admin). The limitations of the system were that senior positions that required expertise and certain qualifications were screened inconsistently

2. **Title:** An integrated e-recruitment system for automatedpersonality mining and applicant ranking

Author and Year: Evanthia Faliagka, Athanasios Tsakalidis, Giannis Tzimas, 2012 **Methodology:** The purpose of this paper is to present a novel approach for recruiting and ranking jobapplicants in online recruitment systems, with the objective to automate applicant pre-screening. An integrated, company-oriented, e-recruitment system was implemented based on the proposed scheme and its functionality was showcased and evaluated in a real-world recruitment scenario.

3. **Title:** The General Factor of Personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study

Author and Year: Dimitri van der Lindena,c,*, Jan te Nijenhuisb, Arnold B. Bakker, 2010 **Methodology:** Recently, it has been proposed that a General Factor of Personality (GFP) occupies the top of the hierar-chical personality structure. We present a meta-analysis (K= 212, total N= 144,117) on the intercorrela-tions among the Big Five personality factors (Openness, Conscientiousness, Extraversion, Agreeableness,and Neuroticism) to test for the existence of a GFP. In addition, we report a multi-method validity studytesting the relationship between the GFP and supervisor-rated job performance. The meta-analysis pro-vided supporting evidence for the two meta-factors Stability and Plasticity (oraand b, respectively) and aGFP at the highest hierarchal level. The validity study indicated that the GFP has a substantive componentas it is related to supervisor-rated job performance

1.4 Scope

- To understand the personality of the candidate.
- To predict the best candidate for the job profile.
- To help HR department for best possible candidate acquisition.
- To avoid any selection of candidates for job to save time.

1.5 Objectives

- To develop a system to provide a more effective way of short-listing the candidates.
- To determine the key skill characteristic by defining each expert's preferences and ranking decisions.
- To automate the process of requirement specifications and applicant's ranking.
- To conduct online aptitude and personality test
- To produce ranking decisions that would have relatively higher consistency than those of human experts.

CHAPTER 2

REQUIREMENT ANALYSIS

Requirements Analysis or requirement engineering is a process of determining user expectations for a new software or providing update for previous product. This core points must be measurable, relevant and detailed. In software engineering eld this term is also called as functional specifications. Requirements analysis mainly deals with communication with users or customers to determine system feature expectations, requirements and reduce concept as demanded by various software users. Energy should be directed towards ensuring that the system or product conforms to client needs rather than attempting to turn user expectations to t the requirements

A software requirements specification (SRS) is a comprehensive description of the in-tended purpose and environment for software under development. SRS minimizes the time and sort required by developers to achieve desired goals and also minimizes the development cost.

2.1 Requirement Specification

Requirement specification describes the function and performance of the computer based system and constraints which govern its development. It can be a written document, a set of graphical models, a collection of scenarios, or any combination of above. These are of 3 types:

NR: Normal Requirments

ER: Expected Requirements

XR: Excited Requirements

2.1.1 Normal Requirements

These are the requirement which are clearly stated by the customer so all these requirements will be present in project for customer satisfaction.

N1:- To Detect Personality traits.

N2:- To view Job and requirement details.

N3:- To view aptitude results

N4:- To provide the output in job suitable candidate

2.1.2 Expected Requirements

These requirements are implicit type of requirements. These requirements are not clearly stated by the customer but implicitly comes during system design.

E1:- To use the some exiting Personality traits data set.

E2:-To post-process the data.

E3:- To categorize the types of Candidates.

2.1.3 Excited Requirements

These requirements are neither stated by the customer nor expected. But to achieve total customer satisfaction the developer may include certain requirements which enhance the functionality of the product.

X1:-To provide User friendly interface.

X2:-To provide Web Application.

2.2 Validation of Requirement

A validation criterion is probably the most important point, the most often neglected section of software requirement specification. How do we recognize a successful implementation? What classes of tests must be conducted to validate the functions, performance and constraints? We neglect this section because completing its demands through understanding of software requirement, something that we often do not at this stage. Yet, the specification of validation criteria acts as an implicit review of all requirements. It is essential that time and attention to be given to this section.

The requirement checklist as follows,

- 1. Are all requirements consistent?
- 2. Is the requirements really necessary?
- 3. Is each requirement testable?
- 4. Does the requirement model properly reflect the information function and behaviour of the system to be built?

2.2.1 Validation of Normal Requirements

VNR1:- This requirement gets satisfied by the whole System.

VNR2:-This requirement gets satisfied by viewing job details.

VNR3:- This requirement gets satisfied by viewing results.

VNR4:- This requirement gets satisfied by even low skilled user being able to understand the output provided by the system.

2.2.2 Validation of Expected Requirements

VER1:-This requirement gets satisfied only if the exiting dataset available has appropriate data and the data is big enough to analyze, test and train every aspect.

VER2:-This requirement gets satisfied if the data analysis is done correctly and personality traits dataset doesn't contain any impurities.

VER3:-This requirement gets satisfied if the system is able to predict the personality.

2.2.3 Validation of Excited Requirements

VXR1:-This requirement gets satisfied if the user can understand and use the website properly.

VXR2:- This requirement gets satisfied if the system can run and provide correct output.

VXR3:-User can easily handle the system.

2.3 System Requirements

2.3.1 Software Requirements

1. Operating System: Windows 10

2. Language: Python3.7/java

3. GUI Created: HTML,CSS

4. Web Server: Xampp

5. Database: MySQL

2.3.2 Hardware Requirements

1. Ram: 512mb

2. Hard disk: 20GB

CHAPTER 3

SYSTEM MODEL

3.1 Process Model

Process model is an abstract representation of a process. The goal of process model is to provide guidance for systematically coordinating and controlling the tasks that must be performed in order to achieve the end product and the project objective. Incremental model is used as the process model in our system. The incremental build model is a method of software development where the model is designed, implemented and tested incrementally (a little more is added each time) until the product is finished. It involves both development and maintenance. The product is defined as finished when it satisfies all of its requirements. This model combines the elements of the waterfall model with the iterative philosophy of prototyping.

3.1.1 Selected Model:Incremental Model

Incremental Model is a process of software development where requirement are broken down into multiple standalone modules of software development cycle. Each iteration passes through the requirement, design, coding and testing phase and each subsequent release of the system adds function to the previous release until all design functionality has been implemented. When incremental model is used, the rst increment is generally a core product. That is, basic requirements are implemented but many supplementary features remain undelivered. Based on the use and evaluation by the customer of the rst increment the plan for next increment is estimated which includes the additional features and functionality as well as modification of the core product delivered in rst increment. Until the complete product is produced the process is repeated following the delivery of each increment.

Incremental Model contain 5 phases

1. Communication:-

It includes complete estimation and project scheduling and tracking, also includes estimation of project cost and time.

2. Planning:-

Task requires building of one or more representation of the application. It includes analysis and design. It is a multiple process that includes four attributes of program data structure, software architecture, interface representation and procedural details.

3. Modeling:-

Software development process starts with the communication between customer and developer.

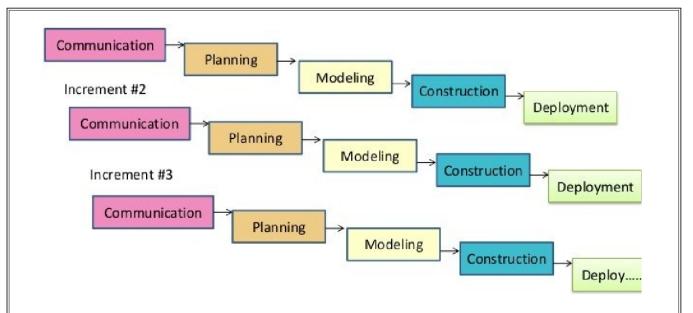


Figure 3.1: Incremental Model

According to need of project, we gather the requirement related to project.

4. Construction:-

It includes coding and testing steps. A, Coding: Design details are implemented using appropriate programming language. B. Testing: It carried out to check whether ow of coding is correct, to check out errors of program.

5. Deployment:-

Include delivery of the partially completed or implemented project and taken feedbacks. The feed backs is considered while reconstruction of the project.

Advantages:-

- 1. Generate working software quickly and easily during the software life cycle.
- 2. The model is more flexible-less costly to change scope and requirements.
- 3. It is easier to test and debug during a smaller iteration.
- 4. In this model customer can respond to each built.
- 5. Initially lower deliver cost

When to Use Incremental Model:-

- 1. This model can be used when the requirement of the complete system are clearly known and understood.
- 2. Major requirements are must be dened, however some details can get evolve with time.
- 3. There is need to get the product to the market early.
- 4. A new technology is being used and resources with needed skill set are not available.
- 5. There are some high risk features and goals.
- 6. Unavailability of staff to complete implementation of project within established business deadline.

3.1.2 Why to use Incremental Models

The main aim of using the model is the reason that we have to add more features in the existing modules to increase project reliability and usability. Using this model we can adapt to the changing requirements of the customer which helps in developing the project in relatively small amount of time. The next increment implements customer suggestions plus some additional requirements in the previous increment. The process is repeated until the project is completed.

3.1.3 Disadvantages of Incremental Models

- Needs good planning and design.
- Needs a clear and complete definition of the whole system before it can be broken down and built incrementally.
- Total cost is higher than waterfall.
- Each iteration phase is rigid and does not overlap each other.

3.1.4 Breakdown Structure(Module)

Breakdown structure of the Personality prediction using Resume is shown below. In this the main system is divided into five parts:-

1. Registration:-

(a) User:User has to enter information to register, after registration user can login.

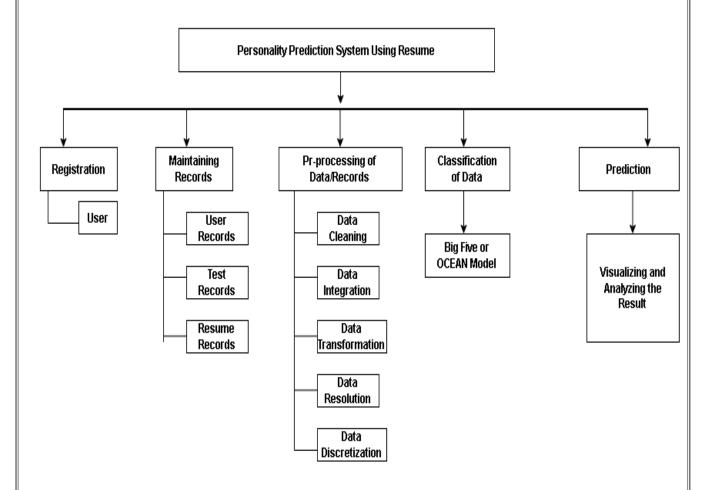


Figure 3.2: Breakdown Structure

2. Maintaining Records:-

- (a) User Records: It contains all information of the user.
- (b) Resume Records: It contains all the data of the Resume uploaded by the user.
- (c) Test Records: It contains all the test related data set by the Admin.

3. Pre-processing of Data:-

(a) Data Cleaning: Data is Cleansed through processes such as filling in missing values, smoothing the noisy data or resolving the inconsistencies in the data.

(b)	Data Integration: Data with different representations are put together and conflict within
	the data are resolved

- (c) Data Transformation: Data is normalized, aggregated and generalized.
- (d) Data Reduction: This step aims to reduce representation of data.
- (e) Data Discretization: Involves the reduction of a number values of a continuous attribute by dividing the range of attribute interval.
- 4. Classification of Data:- Data will be classified using Big 5 or OCEAN Model.
- 5. **Prediction:-** Here the Final prediction of personality trait takes place, Keeping all the aspects in mind like test results, Resume, Personality Trait Dataset.

3.1.5 System Architecture

System Architecture and module-wise details:-

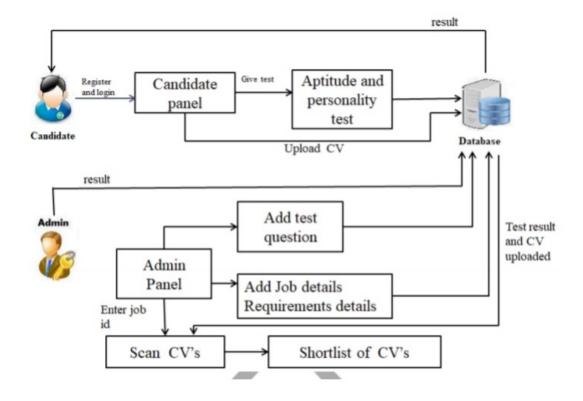


Figure 3.3: System Architecture

3.2 Project Estimation

3.2.1 Estimation in KLOC

In this section various calculation and estimations related to project has been calculated. The figure shows the system modules. The number of lines required for implementation of various modules can be estimated as follows.

3.2.2 Efforts(E):-

The efforts required in person of month for implementation can be estimated as follows $E = 3.2(KLOC)^{2}1.05$

 $E = 3.2(3)^1.05$

Table 3.1: Estimation of KLOC

Sr. no.	Modules	Estimated KLOC		
1	Database Operations	0.3		
2	Classification	0.4		
3	Data Pre-processing	0.9		
4	Prediction	0.9		
5	Visualition	0.5		
Total		3		

E = 10.42 person/Months

3.2.3 Development Time (In months):-

D = E/N

D = 10.42 / 5

D = 2.02 Months

Development time for project requirement analysis and design requires 2.02 months implementation and testing requires 3.97 months.

D = 2.02 + 3.97 Months

D = 6.00 Months

3.2.4 Total Time Required for Project Development:-

Total time required for the successful development of project is 6.00 months.

3.2.5 Number of Person Required (N):

Number of Person= Efforts / Development Time

Number of Person= 10.42 / 2.02

Number of Person= 4 Person.

Four Persons are required to complete the project within given time span successfully.

D1: Mr. Awachar Piyush Madhavrao

D2: Mr. Dange Prajwal Nanasaheb

D3: Mr.Kumbhar Sangram Shahaji

D4: Mr. Pawar Ashish Sudam

CHAPTER 4

SYSTEM DESIGN

4.1 Project Scheduling And Tracking

We have selected the appropriate process model, we have identified the software Engineering tasks that, we have to performed, we estimated the amount of work and the number of people and we know the deadline.

Project Work Breakdown Structure(Analysis)

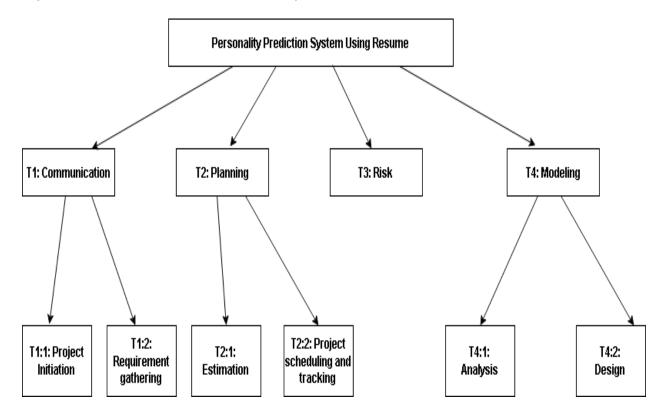


Figure 4.1: Project Work Breakdown Structure(Analysis)

Project Work Breakdown Structure(Implementation)

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus ,it can be considered to be the most critical stage in the achieving a successful new system and in giving the user ,confidence that the new system will work and be effective.

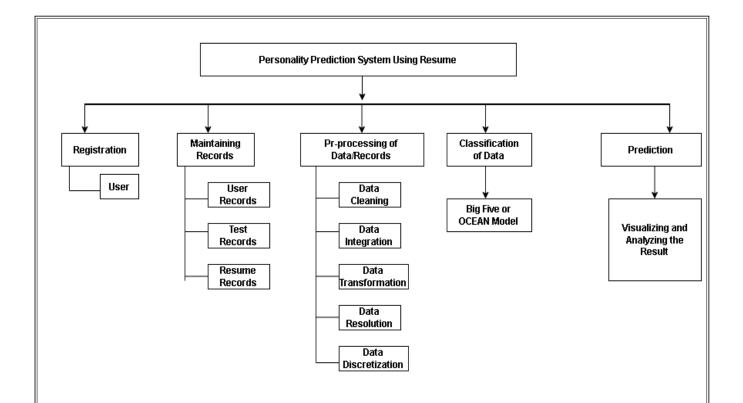


Figure 4.2: Project Work Breakdown Structure(Implementation)

4.1.1 Task Identification

Following analysis and design tasks are to be carried out in process of analysis and design of project. All project modules are divided into following tasks.

- T1 : Searching for project definitions.
- T2: Required literature collection.
- T3: Give presentation of topic.
- T4: Allocation of responsibilities.
- T5: Literature review.
- T6: Synopsis submission.
- T7 : Requirements gathering and validation.
- T8: Process model and project Estimation.
- T9: Risk analysis and management.
- T10 : UML Modeling.
- T11 : Feasibility management of project using mathematical modeling.
- T12: Completion of partial project report.

• T13 : Pre-processing module implementation.

• T14: Classification module implementation.

• T15 : Prediction module implementation.

• T16: Database configuration.

• T17: Implementation of Web application.

• T18 : Integrating all modules.

• T19: Integration and Troubleshooting.

• T20 : Report Generation.

• T21 : Deployment.

Each task is assigned to one or more team members, where

D1: Mr. Awachar Piyush Madhavrao D2: Mr. Dange Prajwal Nanasaheb D3: Mr.Kumbhar Sangram Shahaji D4: Mr. Pawar Ashish Sudam

4.1.2 Project Schedule

In Project management, the scheduling consists of a list of projects terminal elements with intended start and end dates. Terminal elements with intended start and finish dates. Terminal elements are the lowest Elements in a schedule, which is not further subdivided. Those items are often estimated in terms of the resource requirements, budget, cost and duration, linked by dependencies and schedule events. Table describes the schedule for the project development. It also highlights all the tasks to be carried out along with their duration, and developers assigned to accomplish those tasks.

Table 4.1: Project Task Table

Task	Days	Dependencies	Developer Assigned
T1	5	-	D1,D2,D3,D4
T2	7	T1	D1,D2,D3,D4
T3	5	T1,T2	D1,D2,D3,D4
T4	4	T3,T1	D1,D2,D3,D4
T5	4	T3,T4	D3,D1
Т6	5	T3,T5	D1
T7	7	T6	D1,D4
Т8	16	T6	D1,D4
Т9	4	T7	D3
T10	4	T7	D1,D2,D3,D4
T11	3	T7	D1,D1,D4
T12	7	T7,T8,T9,T10	D1,D2,D3,D4

4.1.3 Project Table And Time-Line Chart

Table 4.2: Project Schedule Time Chart

Task Id	Exp.Start Time	Act.Start Time	Exp.End Time	Act.End Time	Developers
T1	07/09/2020	07/09/2020	25/09/2020	25/09/2020	D1,D2,D3,D4
T2	25/09/2020	25/09/2020	26/09/2020	26/09/2020	D1,D2,D3,D4
T3	26/09/2020	26/09/2020	1/10/2020	1/10/2020	D1,D2,D3,D4
T4	1/10/2020	1/10/2020	06/10/2020	06/10/2020	D1,D2,D3,D4
T5	06/10/2020	06/10/2020	11/10/2020	11/10/2020	D1
T6	11/10/2020	11/10/2020	14/10/2020	14/10/2020	D1
T7	14/10/2020	14/10/2020	20/10/2020	20/10/2020	D1,D4
T8	20/10/2020	20/10/200	30/10/2020	30/10/2020	D1,D4
T9	30/10/2020	30/10/2020	05/11/2020	05/11/2020	D3
T10	05/11/2020	05/11/2020	10/11/2020	10/10/2020	D1,D2,D3,D4
T11	10/11/2020	10/11/2020	20/11/2020	20/11/2020	D1,D4
T12	20/11/2020	20/11/2020	01/01/2021	01/01/2021	D1,D2,D3,D4
T13	2/2/2021	-	12/02/2021	-	D1,D2,D3,D4
T14	12/2/2021	-	22/02/2021	-	D1,D2,D3,D4
T15	22/2/2021	-	02/03/2021	-	D1,D2,D3,D4
T16	02/03/2021	-	06/03/2021	-	D1,D2,D3,D4
T17	06/03/2021	-	18/03/2021	-	D1,D2,D3,D4
T18	18/03/2020	-	06/04/2021	-	D1,D2,D3,D4
T19	06/04/2020	-	19/04/2021	-	D1,D2,D3,D4
T20	19/04/2020	-	24/04/2021	-	D1,D2,D3,D4
T21	24/04/2020	-	01/05/2021	-	D1,D2,D3,D4

4.1.4 Time-Line Chart

Searching for project defination

Timeline chart shows the progress of project development in various phases. Timeline chart or the project is divided into 6 months accordingly given in figures below

Project Schedule



Figure 4.3: Expected September Timeline

Project Schedule



Figure 4.4: Actual September Timeline

Project Schedule



Figure 4.5: Expected October Timeline

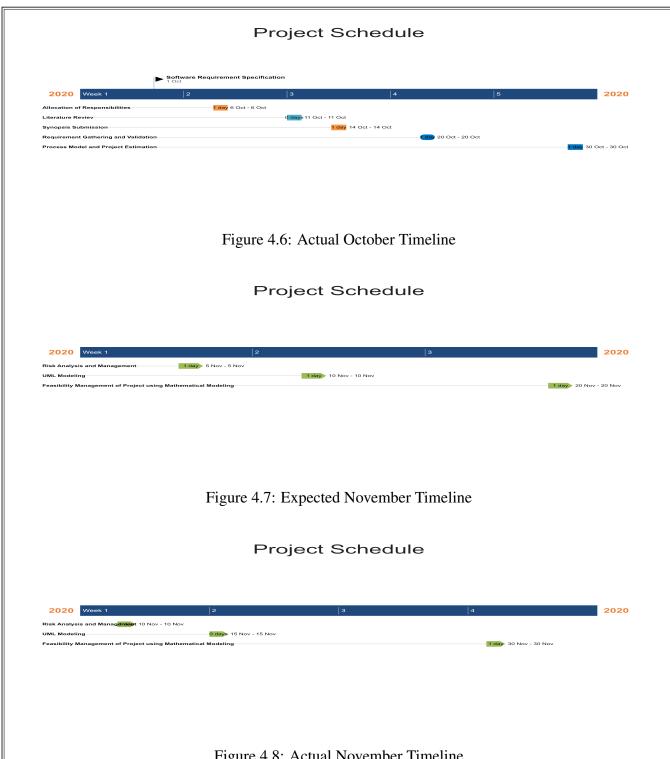


Figure 4.8: Actual November Timeline

Analysis Modeling 4.2

Behavioral Modeling 4.2.1

Use Case: A use case involves a sequence of interactions between the initiator and the system, possibly involving other actors.



State Chart Diagram: A use case involves a sequence of interactions between the initiator and the system, possibly involving other actors.

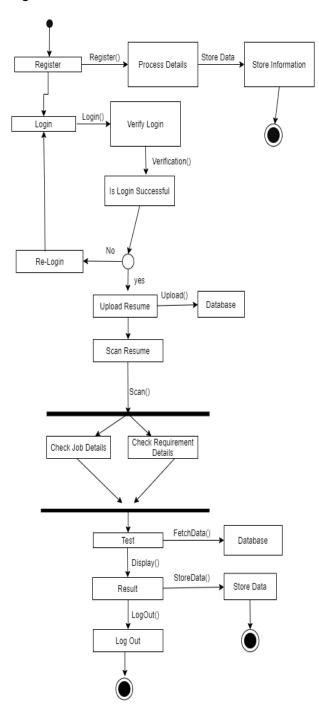
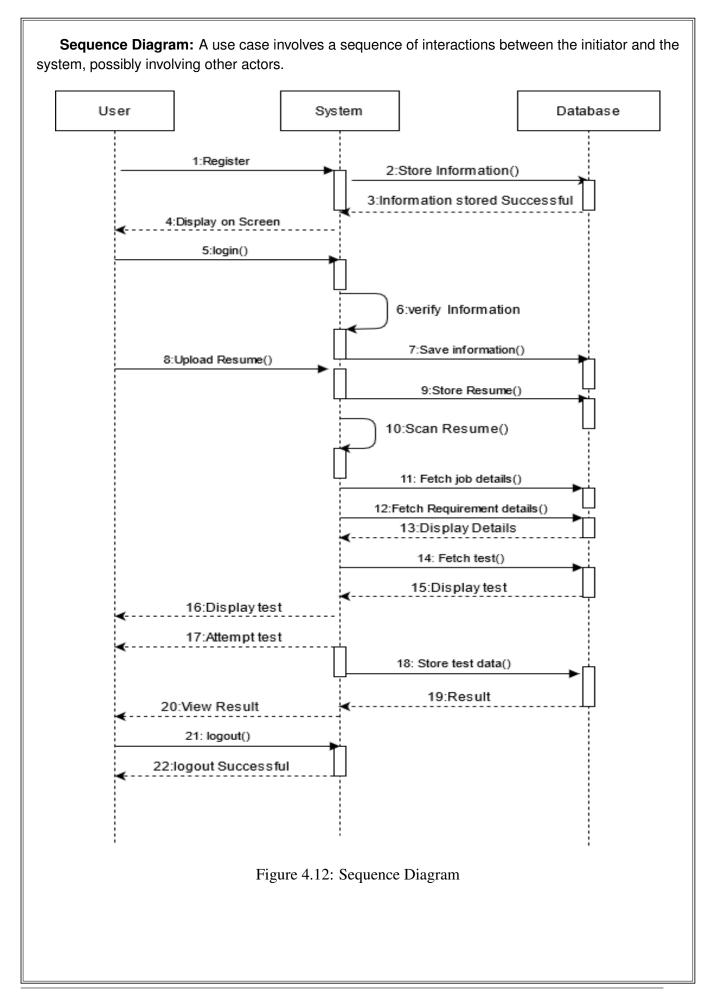


Figure 4.10: State Chart Diagram

Class Diagram: A use case involves a sequence of interactions between the initiator and the system, possibly involving other actors. User Student Database Database - ld: Integer - Id: Integer -Student Record Password : varchar Upload name : varchar - HR Records email: varchar + Login() + Upload() + Fetch() -Dataset Record phone : varchar +Store Data Branch: varchar +Fetch Data Password : varchar +Process Data + Login() Access Data + Register() + Upload() + Fetch() System Dataset DB HR Database Verify - ld: Integer - ld: Integer Processing data : varchar Display Name: varchar + Fetch() +StoreData() Password: varchar + Scan() + Result() +FetchData() + Login() + SetTest() + ViewReult()

Figure 4.11: Class Diagram



Activity Diagram: A use case involves a sequence of interactions between the initiator and the system, possibly involving other actors.

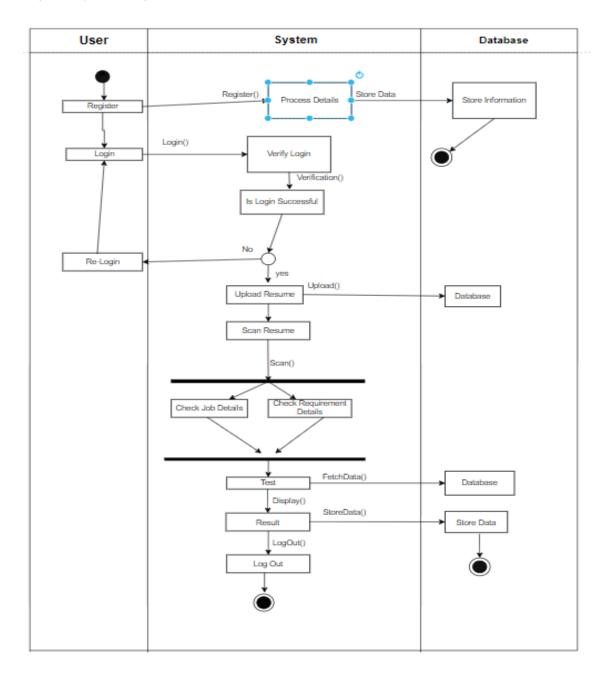


Figure 4.13: Activity Diagram

4.2.2 Functional Model

Data Flow Model: Dataflow diagram is called bubble chart is a graphical technique, which is used in information flow, and transforms through applied when data moves from input to output.

DFD represents system requirement clearly, and identify transforms those become program in design DFD may further partition into different level to show detailed information flow. e.g Level 0, Level 1, Level 2.

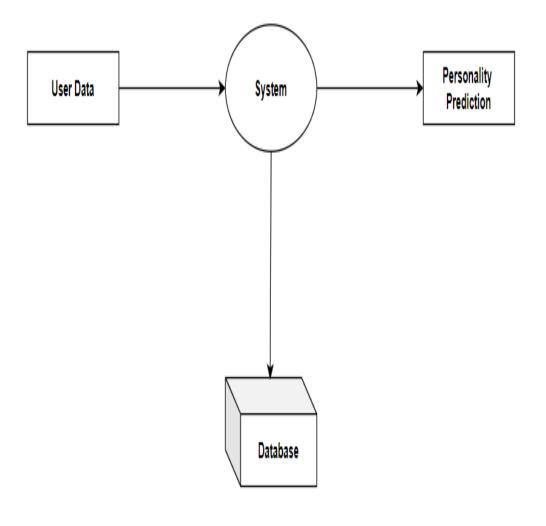
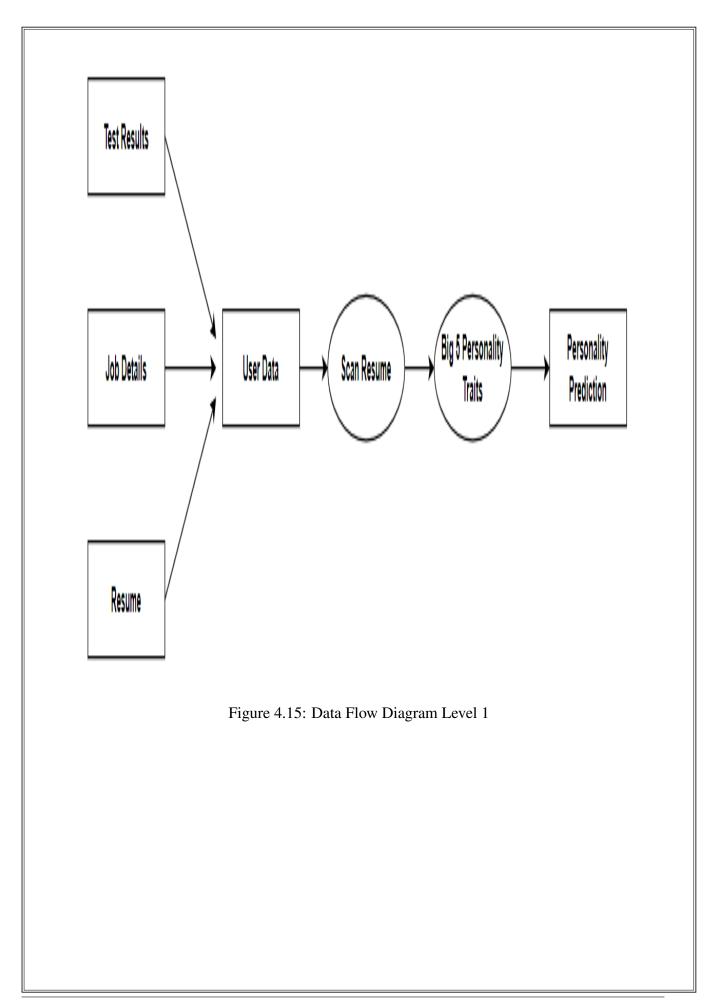
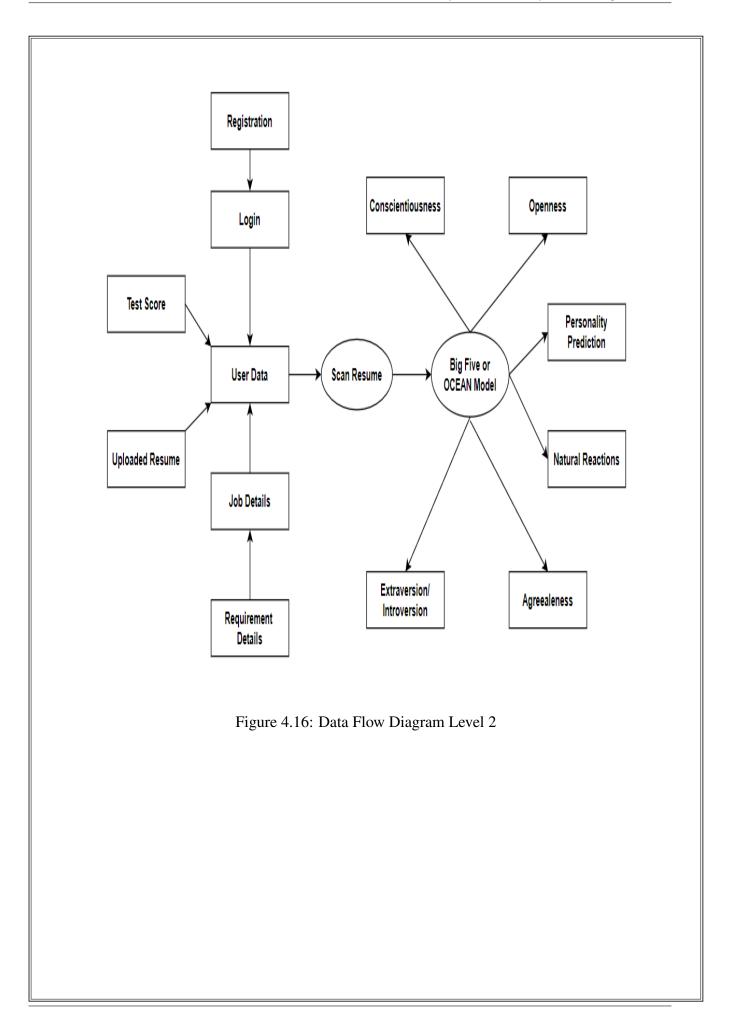


Figure 4.14: Data Flow Diagram Level 0





Control Flow Model: The large class of applications having following characteristics required control flow modeling

The application driven by events rather than data

The application that produce control flow information rather than reports or display.

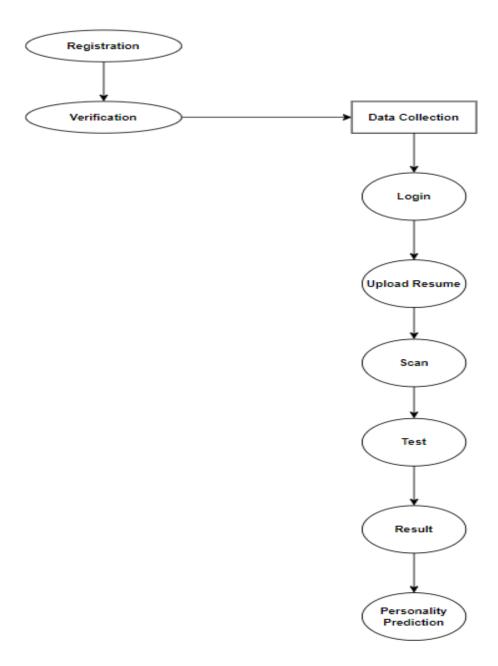


Figure 4.17: Control Flow Diagram

4.2.3 Architectural Modeling

Component Diagram : Component diagrams are used to model physical aspects of a system. Component diagrams are used to visualize the organization and relationship among components in a system. These diagrams are also used to make executable systems.

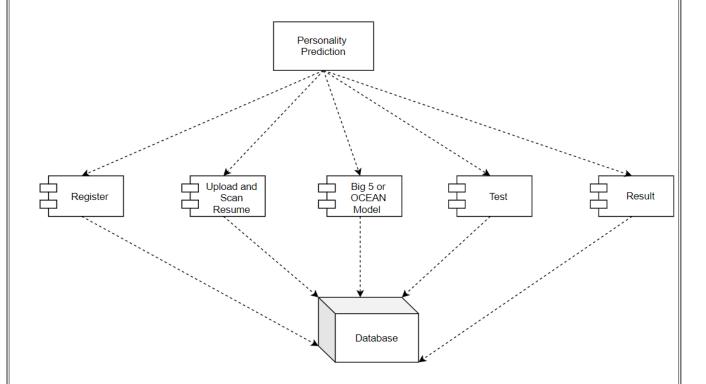


Figure 4.18: Component Diagram

Deployment Diagram: Deployment diagrams are used to visualize the topology of the physical components of a system where the software components are deployed. Deployment diagrams are used for describing the hardware components where software components are deployed.

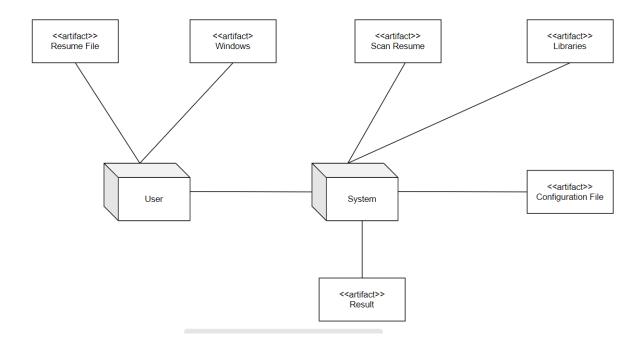


Figure 4.19: Deployment Diagram

4.2.4 Mathematical Model

A mathematical model is a description of a system using mathematical concepts and language. The process of developing a mathematical model is termed mathematical modeling. Mathematical modeling is the art of translating problems from an application area into tractable mathematical formulations whose theoretical and numerical analysis provides insight, answers, and guidance useful for the originating application.

P-Class(Polynomial Class Problem) The problem is in P if there is a known polynomial time algorithm to get that answer. Informally, the class of problem solvable by some algorithm within a number of steps bounded by some fixed polynomial in the length of the input.

NP-Class(Non-Deterministic Polynomial Class Problem) An algorithm is called Non deterministic Polynomial(NP-class) algorithm when for a given input there are more than one paths that the algorithm can follow. Due to this one can not determine which path is to be followed after a particular stage. All NP class problems are basically non deterministic. It is class of computational problems for which solution can be verified by a non-deterministic turing machine in a polynomial time.

NP-Complete(**Non-Deterministic Polynomial Complete Problem**) NP Complete problem is the term used to describe identification problem that are the hardest once in NP in the sense that ,if there were a top-k recommendation algorithm for each problem in NP.NP-complete is a complexity class which represents the set of all problems X in NP for which it is possible to reduce any other NP problem Y to X in a polynomial time.

NP Hard(Non-Deterministic Polynomial Time Hard Problem) NP Hard,in computational complexity theory,is a class of problem that are,informally at least as hard or the hardest problem in NP.A problem it is NP-hard and only if there is NP complete L that is polynomial time duringreducible to H.NP-Hard problems do not have to be in NP.

Set Theory

Set theory is a branch of mathematical logic that studies sets, which informally are collections of objects. Although any type of object can be collected into a set, set theory is applied most often to objects that are relevant to mathematics. The language of set theory can be used to define nearly all mathematical objects. A set is collection of objects which are called the members or elements of that set. If we have a set we say that some objects belong(or do not belong) to this set, are (or are not) in the set. Also, we can say that sets consist of their elements.

According to set theory the relevent mathematical model for our project is designed below.

Let, system as a S set

 $S = \{I, P, R, O\}$

Where,

I is set of all inputs given to system.

P is set of all processes in system

R is set of rules that drives your input set.

o is a set of output expected from system.

Then Input (I) represented as:

 $I = \{I1, I2\}$

Where.

I1:Username.

I2:Password.

Processes(P) is represented as:

 $P = \{P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12\}$

Where,

P1:Register

P2:Login

P3: Verify Id and Password

P4:Upload Resume

P5:Scan Resume

P6:Shortlist Resume

P7:Set question paper

P8:Modify question

P9: Job details and requirements

P10:Test P11:Result

P12:Logout

Rules (R) is represented as:

$$R = \{R1, R2\}$$

Where,

R1: web server should always be available.

R2: system should support web app.

Output(O):

$$\mathbf{O} = \{\mathbf{O1}, \mathbf{O2}\}$$

Where,

O1: Test Results

O2: Personality Traits Prediction

Venn Diagram

Thus,

We can draw the venn diagram from the Inputs,Processes,Rules, and Outputs as shown in figure below.

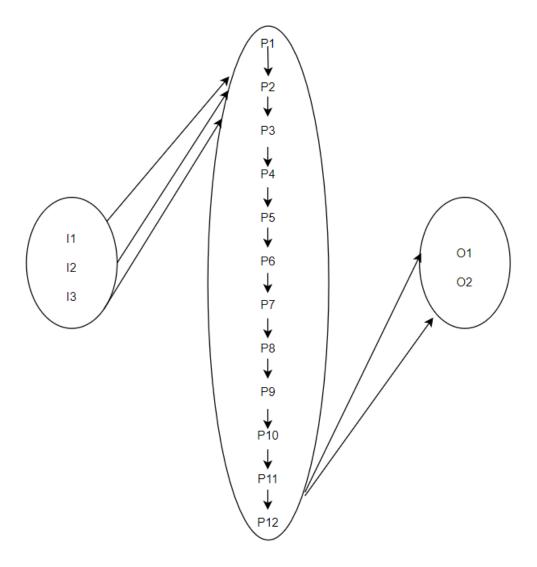


Figure 4.20: Venn Diagram

Space Complexity: Space Complexity of an algorithm is total space taken by the algorithm with respect to the input size. Space complexity includes both Auxiliary space and space used by input.

Space complexity is a parallel concept to time complexity. If we need to create an array of size n, this will require O(n) space. If we create a two dimensional array of size n*n, this will require O(n2) space.

Time Complexity: Time complexity is the number of operations an algorithm performs to complete its task (considering that each operation takes the same amount of time). The algorithm that performs the task in the smallest number of operations is considered the most efficient one in terms of the time complexity. Time Complexity:O(nlog2n)+C.	

4.2.5 Database Table with Discription

Student Database: User database consist of following parameter

- 1. Id: Id is a unique attribute for every student which is auto assigned and auto increment.
- 2. Name: Contains name of the candidate.
- 3. Email: Contains email of the candidate.
- 4. Phone: Contains phone number of the candidate.
- 5. Branch: Contains branch of the candidate.
- 6. Password: Contains password of the candidate.
- 7. Clgname: Contains College name of the candidate.
- 8. Result: Contains result of the candidate.

HR Database: HR database consist of following parameter

Id Name Password

- 1. Id: Id is a unique attribute for every HR which is auto assigned and auto increment.
- 2. Name: Contains name of the HR.
- 3. Password: Contains password of the HR.

Dataset: Dataset consist of following parameter

Id Data

- 1. Id: Id is a unique attribute for every Trait which is auto assigned and auto increment.
- 2. Data: Contains Personality traits.

CHAPTER 5

Risk Management

5.1 Risk Identification

Risk Management involves different kind of risks that might affect the project schedule or quality of the software being develop and monitoring the actions to avoid the risk. Basically risk management is the process of identifying, assessing, responding to, monitoring, and reporting risks. This Risk Management Plan denies how risks associated with the project will be identified, analyzed and managed. Risk management is an ongoing process that continues through the life of a project. effective risk management makes easier to deals with the problem which is occurs while development of process.

1. Product size Risk

R1: As the scope of project is too large, it is difficult to complete the project in given time span with 4 persons.

2. User Related Risk

R2: As the user is not the technical person, so it creates a problem while understanding the exact requirement of the customer.

R3: If user asks for change or gives some unexpected modification in later stages of development then it is difficult to alter the entire system design in accordance with that change.

3. Process Risk

R4: Software Process Model is not followed upto the defined degree.

R5: Quality of a process itself that leads to failures. A low quality processes may not properly work and may break down system.

4. Technical Risks

R6: Lack of training on tools and inexperience.

R7: Images not supported by the system.

R8: Lack of database stability and concurrency.

5. Business impact Risk

R9: Delay in project delivery (violation in time constraints) can hamper the customer economically.

R10: If System is not efficient than the existing system, it will cause economic losses.

6. Development Environment Related Risks

R11: Lack of proper training and less knowledge of programming leads a moderate risk. It will delay product development and deployment.

7. Other Factors

R12: Improper interactions amongst group members can result in misunderstanding, which may cause less work.

5.2 Strategies used to manage Risk

- S1: Formulation and follow up of the project plan on regular basis.
- S2: Keep assigned work under certain deadlines.
- S3: Efficient algorithm must be used for implementation of the system.
- S4: Regular meeting with customers reduce the risk to some extent, design system with extensibility and maintain necessary documentation for the same.
- S5: Redefine software process at a higher degree.
- S6: Proper training on required technical tools for development of project reduce risk.
- S7: Detailed study of required system and proper selection of Software process model
- S8: Ensure that all the members are participating in the design and retain proper interactions amongst group members.
- S9: Study and understanding of project definition, programming language.
- S10: Use of standard database technology which supports concurrency more effectively.
- S11: Each and every module must be tested for its functioning.
- S12: Time constraints must be followed to avoid economical risks.
- S13: Design the system with flexibility to adopt changes in later stages and also maintained all necessary documentation for the same

5.3 Risk Projection

Table 5.1 lists all possible risks which may occur at any stage during development of project. Table also clearly shows the impact of risks and RMMM (Risk Mitigation Monitoring and Management) plan to deal with any such risks. Goal of RMMM plan is to identify as much potential risk as possible. When all risk has been identified, they will then be evaluated to determine their probability of occurrence.

5.4 Feasibility Study

The feasibility of the project is analyzed in this phase and predicting the product with a very general plan for the project and some cost estimates. During system analysis, the feasibility study of the proposed system is to be carried out. This system is having a faster way of getting insight from users data. For feasibility analysis, some understanding of the major requirements for the system is essential.

Table 5.1: Risk Table along with RMMM Plan

Risk	Category	Probability	Impact	RMMM Plan
R1	Product Size	More	High	S1,S2
R2	Customer Related Risk	Less	High	S4
R3	Customer Related Risk	Less	Low	S1,S3
R4	Process Risk	Less	High	S5
R5	Process Risk	Less	Low	S7
R6	Technical Risk	More	High	S5,S6
R7	Technical Risk	Less	Less	S5,S6
R8	Technical Risk	More	High	S10
R9	Business impact Risk	Low	Low	S2
R10	Business impact Risk	Less	High	S4
R11	Development Environment Related Risk	Less	High	S4,S7,S8,S12
R12	Other Factors	Less	Less	S8,S12

5.4.1 Cost Feasibility

This study is carried out to check the economic impact that the system will have on the organization. The expenditures must be justied. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Some hardware products has to be purchased such as Raspberry pi, WiFi module and ECG sensor.

CHAPTER 6

Technical Specification

6.1 Software Requirement Specification and its justification

6.1.1 Elipse

The Eclipse IDE has a very long development history. In November 2001, IBM created the Eclipse Project to implement a Java-based IDE that supports development of embedded Java applications. The initial version of Eclipse derives from Visual Age - a multi-programming language IDE from IBM. In January 2004, the Eclipse Foundation was established as an independent not-for-profit corporation to transparently develop the Eclipse Project as an open and vendor-neutral product. Originally created for developing a Java IDE, the Eclipse Foundation is now developing a wide range of development tools that support many programming languages: C/C++, PHP, Javascript, Python, Rust... However Eclipse is best known as the most widely used IDE for Java development. Because the Eclipse Foundation releases many packages for different programming languages and different domains, in this course the name Eclipse or Eclipse IDE refers to the package Eclipse IDE for Java EE Developers. And as Java programmer, you use this package most of the time. For me, I started using Eclipse in 2004 - Eclipse 3.0 with JDK 1.4. Eclipse is free and open-source, which means you can use it at no cost and access its source code if needed. Today, Eclipse is the most widely used IDE for developing Java applications, with millions of programmers using every day.

6.1.2 XAMPP

XAMPP is an abbreviation for cross-platform, Apache, MySQL, PHP and Perl, and it allows you to build WordPress site offline, on a local web server on your computer. This simple and lightweight solution works on Windows, Linux, and Mac – hence the "cross-platform" part.

Since WordPress isn't a stand-alone application, XAMPP provides two essential components for its installation – Apache, which is used to create the local server, and MySQL which you can use as a database for your website.

You may be wondering why and how developers are using WordPress local server. The answer is simple – it allows them to create a local copy of the site in which they can try out new plugin updates before implementing them on its live version.

This way they can prevent and spot potential errors and issues that might occur.

Now it's time to dive deeper into this XAMPP tutorial and learn how to locally host a website on Windows, Mac, or Linux.

6.1.3 Bootstrap

Bootstrap is a powerful front-end framework for faster and easier web development. It includes HTML and CSS based design templates for creating common user interface components like forms, buttons, navigations, dropdowns, alerts, modals, tabs, accordions, carousels, tooltips, and so on.

Bootstrap gives you ability to create flexible and responsive web layouts with much less efforts.

Bootstrap was originally created by a designer and a developer at Twitter in mid-2010. Before being an open-sourced framework, Bootstrap was known as Twitter Blueprint.

You can save a lot of time and effort with Bootstrap. So bookmark this website and continue on.

CHAPTER 7

Conclusion and Future Scope

In this project, we have implemented an organization oriented recruitment system that would assist the human resource department in short listing the right candidate for a specific job profile. The system would be used in many business sectors that will require expert candidate, thus reducing the work load on the human resource department.

References

- [1] F Evanthia, T Athanasios, et al. An integrated e-recruitment system for automated personality mining and applicant ranking. Internet Research 2012; 22: 551568.
- [2] L D. van der, J te Nijenhuis, et al. The General Factor of Personality: A metaanalysis of Big Five intercorrelations and a criterion-related validity study. Journal of Research in Personality 2010; 44: 315-327.
- [3] F. Safia, N Asha, the Impact of Person Organization Fit on Job Satisfaction and Performance of the Employees. Economics and Finance 2014; 11: 122-129.
- [4] A. I Ilke, W Peter, Personality and Job Engagement. Journal of Personnel Psychology 2011; 10: 177-181
- [5] D Tantam, The machine as psychotherapist: impersonal communication with a machine. BJP synch Advances 2017.
- [6] Diseth, A. (2003). The Personality and approaches to learning as predictors of academic achievement. European Journal of Personality, 17, 143-155.

7.0.1 Annexure A: Project Planner and Progress Report

Weekly Assignment Report

Group ID	Project Title	Name of Students	Period (Week)
		1) Awchar Piyush M	
23	Personality Prediction System	2) Dange Prajwal N	1 st Week
	Using Resume	3) Kumbhar Sangram S	
		4) Pawar Ashish S	

Work done:

The first week dated from 07/09/2020 to 25/09/2020 is the starting of our project where we discussed about the problem definition about the project and the domain on which we will be working. We had couple of meetings with our project guide Prof. T. Bhaskar whose outcome was the finalization of the topic and submitting the project title and problem definition along with the domain we will be working in it to project co-ordinator. After the finalization of project topic on 05/10/2020 we extended our work for the completion of synopsis. For the accomplishments of synopsis first we studied the literature review of the topic according to that we define the scope, objective and application of the project. We also designed the tentative model or system architecture of our project and we decided on what basis we meet to proceed for the same. The task of synopsis was divided among the 4 members equally led to completion of synopsis in an efficient manner. After which we submitted it to project guide which he reviewed it and asked us for some changes which we did it and submitted back to guide. After reviewing again, he finally accepted the synopsis which was completed on 22/10/2020.

Guide Remark:

Prof.S.S.Shaikh [Internal Guide]

Group ID	Project Title	Name of Students	Period (Week)
		1) Awchar Piyush M	
23	Personality Prediction System	2) Dange Prajwal N	2 nd Week
	Using Resume	3) Kumbhar Sangram S	
		4) Pawar Ashish S	

Work done:

After the successful completion of title finalization and synopsis we extended our work to complete the software requirement specification (SRS) of the project. We started our work on 26/10/2020 where we started stating the requirements about our project. According to the tentative project model we thought about various software and hardware requirements which are tentative and predictive and developed our SRS. We also prepared the modular design of our project that is the increment model according to the software specification. And finally, on 06/11/2020 we completed our SRS.

Guide Remark:

Prof.S.S.Shaikh [Internal Guide]

Group ID	Project Title	Name of Students	Period (Week)
		1) Awchar Piyush M	
23	Personality Prediction System	2) Dange Prajwal N	3 rd Week
	Using Resume	3) Kumbhar Sangram S	
		4) Pawar Ashish S	

Work done:

After the successful completion of software requirement specification (SRS) we extended our work to next step know as Project planning on 09/11/2020. In this step pf project planning we divided the work of drawing UML diagrams equally among the developer so as to understand the project planning in detail. We also studied the breakdown structure of project model wise which helped us to the better understanding of the project. We draw all the UML diagrams on draw. io and finally completed the assignment on 24/11/2020 and we forwarded to our project guide for verification purpose. After verified from our project guide we finally complete the project planning assignment.

Guide Remark:

Prof.S.S.Shaikh [Internal Guide]

Group ID	Project Title	Name of Students	Period (Week)
		1) Awchar Piyush M	
23	Personality Prediction System	2) Dange Prajwal N	4 th Week
	Using Resume	3) Kumbhar Sangram S	
		4) Pawar Ashish S	

Work done:

After the successful completion of project plan 24/11/2020 we extended our work for the completion of the next assignment which was risk management. In risk management we discussed about various kinds of risks associated with our project and we completed that assignment on 30/11/2020. In the same week we also parallelly started developing the partial project report so that the work gets completed in given deadline and we completed approximately 80 same week and remaining of the work will be completed till the end of next week. And we finally submitted the assignment of risk management on 30/11/2020 to the project guide.

Guide Remark:

Prof.S.S.Shaikh [Internal Guide]

Group ID	Project Title	Name of Students	Period (Week)
		1) Awchar Piyush M	
23	Personality Prediction System	2) Dange Prajwal N	5 th Week
	Using Resume	3) Kumbhar Sangram S	
		4) Pawar Ashish S	

Work done:

In the 6th week of our project we have prepared again next 2 assignments which was somewhat based on the previous assignments. In which 4th assignment which is last it was based on the 2nd assignment and 3rd is individual assignments this 2 assignments were tough and lengthy too.

So we prepared 3rd and 4th assignments in 5th week i.e.

- 1) Requirement Analysis.
- 2) Design.

This was our 3rd 4th assignments it sometime but we did it. We took the help our guide and guided us very well because if him we were able to complete this work smoothly.

Guide Remark:

Prof.S.S.Shaikh [Internal Guide]

Group ID	Project Title	Name of Students	Period (Week)
		1) Awchar Piyush M	
23	Personality Prediction System	2) Dange Prajwal N	6 th Week
	Using Resume	3) Kumbhar Sangram S	
		4) Pawar Ashish S	

Work done:

Our group i.e. Group ID 24 we have participated in the CSI competition so in the 6th week we group members we have done the CSI project competition work which was very important too in that we have uploaded.

- 1) Title and the domain of the project.
- 2) Problem Statement.
- 3) Synopsis ppt.
- 4) Recorded video presentation of max 3 minutes.

All this work we has done in the 6th week under the guidance of our guide sir help us to take part in the CSI project.

Guide Remark:

Prof.S.S.Shaikh [Internal Guide]

Group ID	Project Title	Name of Students	Period (Week)
		1) Awchar Piyush M	
23	Personality Prediction System	2) Dange Prajwal N	7 th Week
	Using Resume	3) Kumbhar Sangram S	
		4) Pawar Ashish S	

Work done:

After completion of everything related to our CSI project.

In the 7th week we started to prepare the Survey Paper of our project as our sir told us to do it sir has send us the sample survey paper according to that we have prepared our survey paper. So 7th week was the Survey Paper's week whole week goes to do that because we faced lot of problem while preparing the Survey Paper. But under the guidance the Prof. T. BHASKAR we have done everything very smoothly.

Guide Remark:

Prof.S.S.Shaikh [Internal Guide]

7.0.2 Annexure B: Plagiarism Report

PLAGIARISM SCAN REPORT

Date	October 06	Words	303	
Exclude URL:				
	100% Unique		0% Plagiarized	

Figure 7.1: Plagiarism Scan Report 1

PLAGIARISM SCAN REPORT



Figure 7.2: Plagiarism Scan Report 2

PLAGIARISM SCAN REPORT

Date	October 06	Words	618	
Exclude URL:				
	93% Unique		7% Plagiarized	

Figure 7.3: Plagiarism Scan Report 3

PLAGIARISM SCAN REPORT Date October 06 Words 779 **Exclude URL:** 90% 10% Plagiarized Unique Figure 7.4: Plagiarism Scan Report 4 PLAGIARISM SCAN REPORT October 06 ~ Date 500 Words **Exclude URL:** 98% 2% Unique Plagiarized Figure 7.5: Plagiarism Scan Report 5 PLAGIARISM SCAN REPORT Date October 06 Words 865 **Exclude URL:** 95% 5% Unique Plagiarized Figure 7.6: Plagiarism Scan Report 6 PLAGIARISM SCAN REPORT Date October 06 Words 733 **Exclude URL:** 100% 0% Unique Plagiarized

Department of Computer Engineering, Sanjivani College of Engineering, Kopargaon

Figure 7.7: Plagiarism Scan Report 7