



# **Smart Resume CV selector for IT Industry**

## **Software Requirements Specification**

**Project ID: 18-005**  
**Name – Y.I. Kodithuwakku**  
**Student ID – IT14115776**

**B.Sc. Special (Honors) Degree in Information Technology**  
**specialized in Information Technology**  
**Submitted on 2018-05-08**

## **DECLARATION**

I hereby declare that the submitted project Software Requirements Specification document for Smart Resume is an original work done by Y.I. Kodithuwakku. This document is proprietary and an exclusive property of the SLIIT project group 18-005. List of references I referred for the preparation of this document are given as references at the end of the document.

Member : IT14115776 – Y.I. Kodithuwakku

### Approval Block

	Name	Partner	Date	Signature
From				
Verified by				
Approved by				

### Revision History

Version	Date	Summary of Changes	Author
V1.0	2018-5-6	Initial version	Smart Resume
V1.1	2018-5-9	Modified as Supervisor feedback	Smart Resume

## **Table of Contents**

1 INTRODUCTION .....	7
1.1 Purpose .....	7
1.2 Scope.....	7
1.2.1 Objectives.....	9
1.2.2 Other Objectives.....	9
1.2.3 Benefits.....	8
1.3 Definitions, Acronyms, and Abbreviations .....	11
1.4 Overview.....	11
1.4.1 Goals.....	11
1.4.2 Users.....	13
1.4.3 System Overview .....	13
2 OVERALL DESCRIPTION.....	14
2.1 Product perspective.....	15
2.1.1 System interfaces.....	16
2.1.2 User interfaces .....	17
2.1.3 Hardware interfaces.....	18
2.1.4 Software interfaces .....	18
2.1.5 Communication interfaces.....	18
2.1.6 Memory constraints.....	18
2.1.7 Operations .....	19
2.1.8 Site adaptation requirements .....	19
2.2 Product functions.....	19
2.3 User characteristics .....	21
2.4 Constraints .....	21
2.5 Assumptions and dependencies .....	21

2.6 Apportioning of requirements.....	22
3 SPECIFIC REQUIREMENTS .....	22
3.1 External interface requirement.....	22
3.1.1 User interfaces .....	22
3.2.1 Hardware interfaces .....	22
3.2.2 Software interfaces .....	22
3.2.3 Communication interfaces.....	23
3.2 Classes/Objects .....	24
3.3 Performance Requirements.....	24
3.4 Design Constraints .....	25
3.5 Software System Attributes .....	25
3.5.1 Reliability .....	25
3.5.2 Availability .....	26
3.5.3 Security.....	26
3.5.4 Maintainability .....	26
3.6 Other Requirements .....	27
4 SUPPORTING INFORMATION.....	28
4.1 Appendices .....	28
5 REFERENCES .....	30

## **List of figures**

Figure 1: System Overview.....	13
Figure 2: Statistics of analyzed data.....	18
Figure 3: View selected candidates.....	18
Figure 4: View evaluated statistical details.....	20
Figure 5: View candidates analysis details.....	21
Figure 6: Use case diagram – view dashboard.....	21
Figure 7: Class diagram.....	25
Figure 8: Work breakdown structure.....	29
Figure 9: Gantt chart.....	30

## **List of Tables**

Table 1 - Definitions, Acronyms, and Abbreviations.....	11
Table 2 – Classification Table.....	14

## **1 Introduction**

### **1.1 Purpose**

The purpose of this document is to fully describe the functional and non-functional requirements, design constraints, project approach and other factors necessary to provide an in-depth view of the project progress. This SRS document contains a detail overview of project software product, its parameters and goals. This document is targeting the client, designers, developers and other stakeholders as its audience. A description of the problem in focus will be provided relating what the system would do in order to overcome the problem. This document will serve the purpose of providing the potential users to determine if the software specified meets their needs or how the software must be modified to meet their needs.

### **1.2 Scope**

**Smart Resume CV selector – Business intelligence tool** syndicates operational data with analytical tools to present complex and competitive information to decision makers, in order to dynamically select the best CVs for a given attributes of a job vacancy. This research project involves to select the employees who have relevant technical skills according to the vacancy in order to achieve the ultimate project goals of the company. There are four primary components that need to be handled in order to complete **Smart Resume**.

1. BOT Creation
2. ETL Tool
3. Model Building and Prediction
4. Model Evaluation and Dashboard Simulation

This document describes **Model Evaluation and Dashboard Simulation** for IT industry. The scope of the predictive model evaluation and dashboard simulation is limited to implementing the CV Management System having the following functional areas. There are two primary components that need to be handled in order to complete **Model evaluation and Dashboard simulation** component.

1. Predictive Model Evaluation
2. Dashboard Simulation

If the requirement changes in future it is possible to change the specification accordingly. It also covers the details of hardware and software requirements need to implement the predictive model and gives a detailed description of the externally visible behaviour of the predictive model and covers the areas that contain limitations while completing the predictive model for IT Industry.

The proposed **Smart Resume** aims at describing processes of building cost effective, flexible, ease of use and productive enterprise focused Business Intelligence tool.

Driven by fact –Data to decision, **Smart Resume** Business Intelligence (BI) tool make it possible to reach unambiguous decisions in the right time, minimize costs and also provide the opportunity to verify and correlate the decisions with the entire enterprise's strategy.

The user can select only required attributes to analyse the data hence **Smart Resume** will only consume the selected attributes and proceed for predictive analysis. Hence, it will result in high- accuracy and best quality output.

In the dashboard, visualizer will give the detailed view of data for the user to identify the candidate information with graphs. This will helps the user to recruit the best employees for the company.



### 1.2.1 Objectives

#### **Objective 1: Introduce Predictive Model to Filter CVs**

Introduce an intelligent system to filter CVs based on the attributes given by user and apply this system to find out who are the most suitable feasible or optimal candidates in IT industry. The component relies on historical data of the procedures performed in the past. Depending on the job characteristics given an optimal or feasible team will be generated. The optimal CV filter is the one with the lowest probability of unfavourable outcomes an optimal solution is a theoretically proven solution (candidates with less practical knowledge). But it might not be the logically suitable solution and we might have to come up with the feasible candidates. Hence the tool has the option of providing the most feasible (candidates with possible and more practical knowledge) solution as well.

#### **Objective 2: Adaptable and easy going**

The **Smart Resume** enables users to find CVs by selecting relevant attributes for job vacancy depending on their company or industry. Most of the BI tools in the current market have a very complex interface where professional knowledge is required to perform the tasks. **Smart Resume** will not require technical knowledge or professional expertise to interact and it has been designed in a simple way with less number of controls in order to increase the adaptability and user-friendliness.

### 1.2.2 Other Objectives

#### **Objective 3: Build a solution to select most optimal CVs for IT industry**

Since optimal candidates mean who have less practical knowledge, **Smart Resume - Model evaluation and dashboard simulation component** will select the most optimal candidates for IT industry considering attributes related to the

job vacancy and also employee characteristics that company need in IT sector. Depending on the employee characteristics, technical skills and other details listed on CV, most optimal CVs will be selected and they will be participate to the interview. The prediction is expected to be highly accurate and the final decision is represented as a simulation.

#### **Objective 4: Build a solution to select most feasible CVs for IT industry**

Since the feasible candidates mean who have more practical knowledge, **Smart Resume - model evaluation and dashboard simulation component** will select the most feasible candidates for IT industry considering attributes related to the job vacancy and also employee characteristics that company need in IT sector. Depending on the employee characteristics, technical skills and other details listed on CV, most feasible team will be selected. The prediction is expected to be highly accurate and the final decision is represented as a simulation.

#### **Objective 5: Affordable price**

Nowadays a lot of IT companies spend a huge cost when recruiting employees to the company. For a growing small scale or medium scale companies, these BI tools are too expensive and not affordable. But these companies can invest a small amount of investment and get affective results towards the future goals in the company. Our plan is to serve this community and help them to uplift their business with the help of **Smart Resume predictive model evaluation and dashboard simulation component** separately as a predictive modelling tool for the immerging businesses.

### **1.2.3 Benefits**

- Since it is driven by fact: data to decision, it helps to make mathematical model based strategic decision which gives the precise and accurate end result.
- No need of expert knowledge in order to deal with this component
- It helps to get faster answer to business operations

### 1.3 Definitions, Acronyms, and Abbreviations

Table 1 provides the definitions, acronyms & abbreviations that will be used throughout this document and any other document referencing this one:

Term	Definition
SRS	Software Requirements Specification
GUI	Graphical User Interface
IT	Information Technology
ETL	Extract, Transform, Load
BI	Business Intelligence
CV	Curriculum Vitae
ROC	Receiver Operating Characteristics
TP	True Positive – correctly predicted that suitable for job vacancy
TN	True Negative – correctly predicted that not suitable for job vacancy
FP	False Positive – Incorrectly predicted that suitable for job vacancy
FN	False Negative – Incorrectly predicted that not suitable for job vacancy
AUC	Area under curve

Table 1 - Definitions, Acronyms, and Abbreviations

### 1.4 Overview

This SRS document intends to cover all the functional and non-functional requirements of the **Smart Resume - model evaluation and dashboard simulation component**. Each of them has been discussed clearly in detail. All are described under three chapters.

The first chapter provides a full description of the project to the users who are interested in evaluation of big data prediction. The purpose of the SRS, particular audience, what will this component do, how the component going to perform their actions, general

objectives, goals, benefits that can gain, definitions, acronyms, abbreviations and overview of the system.

And also contains user characteristics which can be useful to identify the intended customer and potential user. The second chapter concerns details of each of the system functions and actions in full for the software developers' assistance. These two sections are cross-referenced by topic; to increase understanding by both groups involved.

Inside of predictive model evaluator and dashboard simulator component and how it is subdivided into smaller components and the processes are taking place to develop the component are described in chapter 2. A detailed overview of product perspective, user interfaces, system interfaces, hardware interfaces, communication interfaces, memory constraint, operations, site adoption requirements with the use of use case, given in the rest of the part of this chapter.

The third chapter includes all the supporting information such as references, indexes and appendices. Finally, this will focus on the component deployment and usage procedures in the real world environment and how it will meet the requirements of the Stakeholders. This document can be used as a guide by the development team in the development phase.

Today in Sri Lanka, in both apparel industry and IT industry since there are a huge number of graduates, employee recruitment is a critical factor. Selecting applied CVs in an effective way has a big impact on the decision-making process and productivity of the company by the employees who will be recruited after interviewing. Unfortunately looking for optimal or near optimal candidates and feasible or near feasible candidates are a costly task for humans due to the exponential number of CVs. So **Smart Resume - model evaluator and dashboard simulator component** focus on depending on the relationship of the attributes (Internal and External) this will dynamically predict the most optimal or feasible candidates on the Dashboard.

#### **1.4.1 Goals**

- Evaluate the prediction models to predict the optimal and feasible CVs.
- Increase model accuracy and performance.
- Develop informative dynamic dashboard.

#### **1.4.2 Users**

- IT Companies

#### **1.4.3 System Overview**

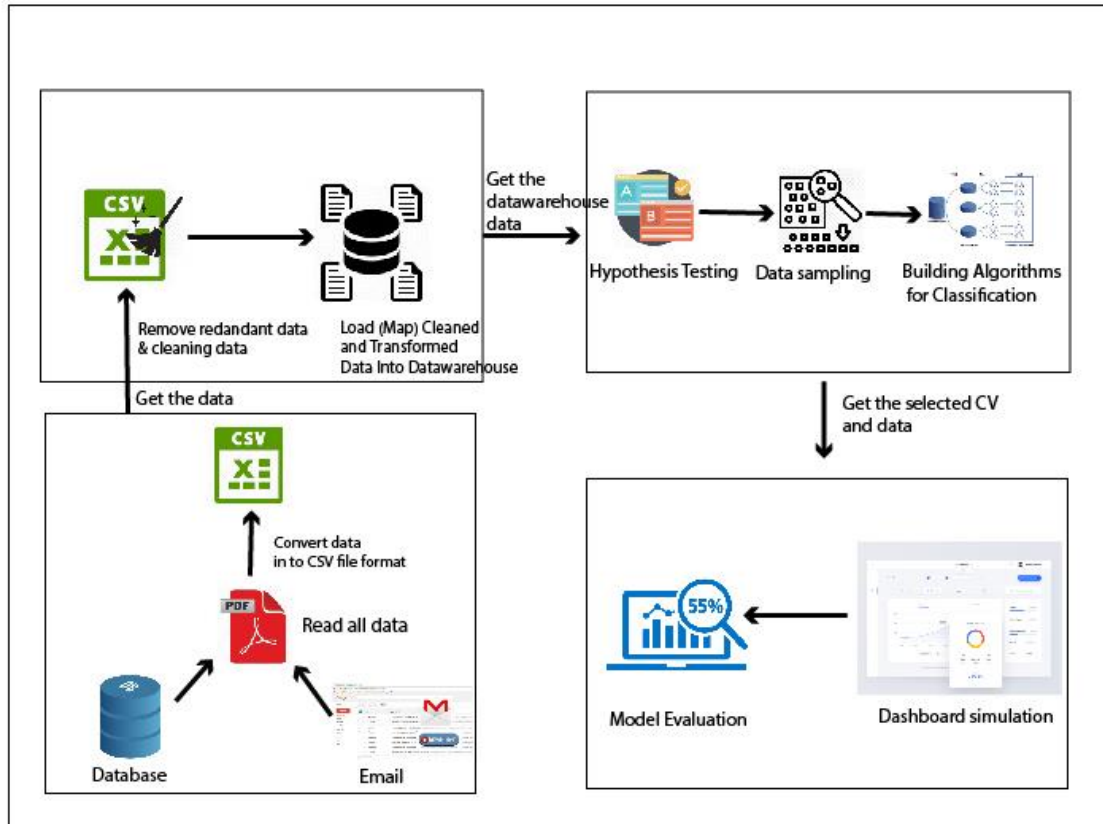


Figure 1: System Overview

## 2 Overall Descriptions

**Smart Resume** - Business intelligence tool syndicates operational data with analytical tools to present complex and competitive information to decision makers, in order to select the best candidates for a given attributes of a job vacancy.

**Smart Resume - model evaluation and dashboard simulation component** involve the predictive model evaluation, analysis, implementation and dashboard simulation of CV insights in regards to the data set available. The predictive model is an “expression of relationships between variables in the form of an equation”.

The evaluation of the predictive model is carried out using confusion matrix, ROC curve, and Model accuracy methods. Classification table describes the performance of a classification model.

	<b>Predicted:</b> <b>0</b>	<b>Predicted:</b> <b>1</b>
<b>Actual:</b> <b>0</b>	TN	FP
<b>Actual:</b> <b>1</b>	FN	TP

Table 2: Classification Table

Every observation in the testing set is represented in exactly one box. It's a 2x2 matrix because there are 2 response classes. **Classification Accuracy** metric need to be calculated using classification matrix.

$$\text{Classification Accuracy} = (TP+TN) / (TP+TN+FP+FN)$$

Receiver Operating Characteristic (ROC) Curves provide a graphical representation of the range of possible cut points with their associated sensitivity vs. 1-specificity. This illustrates the merit of the particular predictor/predictive model, making it possible to identify different cut-points for specific applications – depending on the ‘cost’ of misclassification. Estimates of the area under the curve provide an indication of the utility of the predictor and a means of comparing (testing) two or more predictive models. ROC Curves plot the true positive rate (sensitivity) against the false positive rate (1-specificity) for the different possible cut points of a diagnostic test. Each point on the ROC curve represents a sensitivity/specificity pair.

Finally, after processing all data set based on predefined predictive models, it generates comprehensive, self-descriptive and exploratory dashboard which provide fact – driven visualization of the data. The result will be displayed as a graphs, charts, hierarchy with selected employees for the job interview. Further analyzed details of a particular

employee also can be viewed as well as descriptive reports also available in the dashboard.

## **2.1 Product perspective**

### **Products Available in the Market:**

#### **1) Oracle BI**

- provides the industry's first integrated, end-to-end Enterprise Performance Management System
- Only a predictive analysis and all the users don't have to interact with the system.

#### **2) Birst BI**

- All the users don't have to interact with the system.
- Only the predictive analysis and visualize the data in abstract way.

#### **3) Jobscan**

- A tool that gives job seekers an instant analysis of how well their resume is tailored for a particular job.
- Only the optimal result can be found.
- Visualize the predictive analysed data in an abstract way.

#### **4) Smart Recruit**

- No predictive analysis only visualization of the data



Even though there are existing proposed products in the market area, they do not address most of the problems that the **model evaluation and dashboard simulation** is going to address. The following table shows a comparison of features between the existing products or applications and the proposed solution “**Smart Resume**”.

Features	Oracle BI	Birst	Jobscan	Smart Recruit	Smart Resume
BOT(Automated)					✓
ETL Tool					✓
Optimal Solution			✓		✓
Feasible Solution					✓
All the user does not have to interact with the system	✓	✓	✓	✓	✓
Visualization of data in an abstract way		✓	✓	✓	✓
Predictive Analytics	✓	✓	✓		✓

### 2.1.1 System interfaces

- Web – Desktop connectivity Interface.
- JVM

### 2.1.2 User interfaces

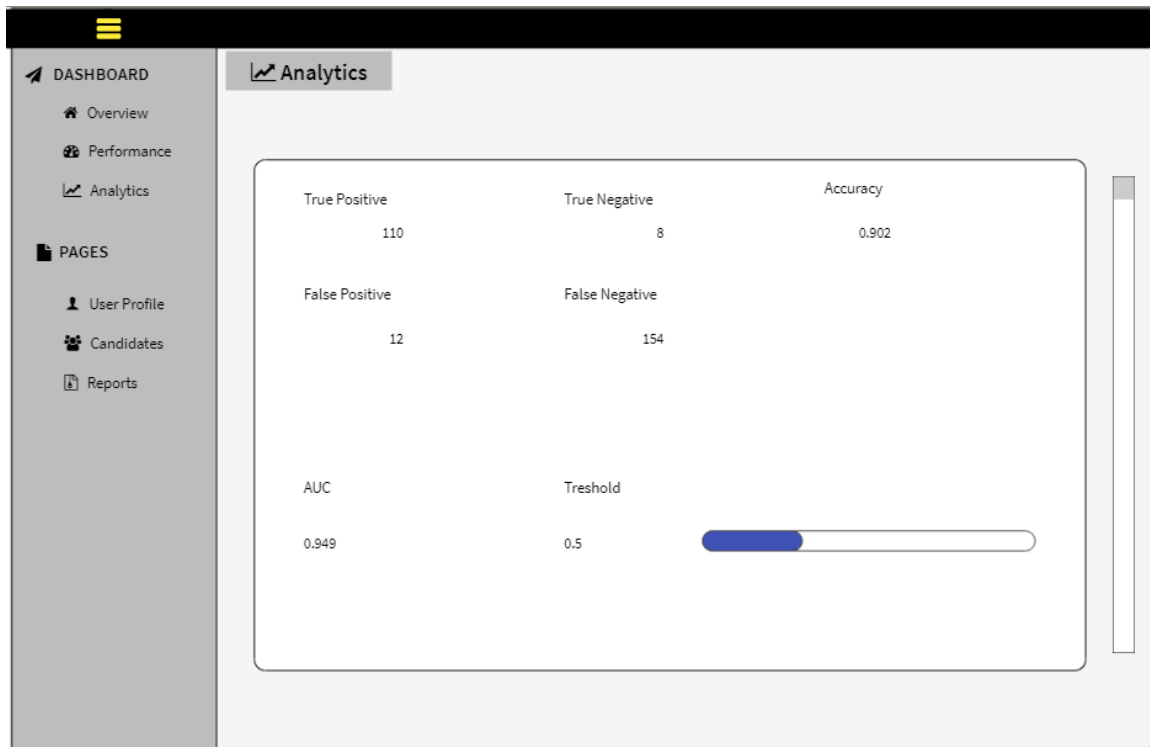


Figure 2: Statistics of the analyzed data

The Candidates dashboard displays a table of selected candidates with the following data:

Name	Age	Marital Status	Contact	Gender	Email	Education
George	27	Unmarried	+64854670	Male	gj@gmail.com	BSc
Kelly	26	Married	+642585740	Female	kthompson@g...	Msc
Samson	31	Unmarried	+646854000	Male	sperera@gmail...	BSc
George	27	Unmarried	+64854670	Male	gj@gmail.com	BSc
Nelly	26	Married	+642585740	Female	kthompson@g...	Msc
Samson	31	Unmarried	+646854000	Male	sperera@gmail...	BSc
George	27	Unmarried	+64854670	Male	gj@gmail.com	BSc
Kelly	26	Married	+642585740	Female	kthompson@g...	Msc
Samson	31	Unmarried	+646854000	Male	sperera@gmail...	BSc
George	27	Unmarried	+64854670	Male	gj@gmail.com	BSc
Kelly	26	Married	+642585740	Female	kthompson@g...	Msc
Samson	31	Unmarried	+646854000	Male	sperera@gmail...	BSc

Figure 3: View Selected candidates

### 2.1.3 Hardware interfaces

**Smart Resume - model evaluation and dashboard simulation component** do not require any special Hardware interfaces apart from a computer

#### **2.1.4 Software interfaces**

Following software interfaces are required in order to use the **Smart Resume - model evaluation and dashboard simulation component**:

- My-SQL
- Apache Server
- Latest updated web browser that supports Python and PHP  
Ex: Google Chrome, Mozilla Firefox, Apple Safari, Internet Explorer, Opera web browser.

#### **2.1.5 Communication interfaces**

Following communication interfaces are required in order to use the **Smart Resume predictive model evaluator and dashboard simulation component**:

- Internet Connection - Required Connection bandwidth might differ time to time. Since large data load is travelling through the network, having a high bandwidth internet connection will help a lot for the users.
- Database connection interface

#### **2.1.6 Memory constraints**

- Ram of 2GB or higher

#### **2.1.7 Operations**

## Operations of Top Level Management

- Configure the characteristics that suitable for job role.
- Select suitable attributes and generate optimal or feasible candidates.
- Interpret results with the dashboard.
- Viewing reports

### 2.1.8 Site adaptation requirements

Following site adoption requirements are identified regarding the implementation of **Smart Resume - model evaluation and dashboard simulation component**:

- The server must have MySQL installed on it.
- Server machine must be running Apache server in order to deploy the web application
- The user machine should have Java Virtual Machine installed.

### 2.2 Product functions

Use case name	View evaluated statistical details.
Pre – condition	Database connection is active. Web application should be running correctly
Actor	User
Main success scenario	1. Navigate to application 2. Click on dashboard 3. Click Statistics
Post – condition	View all the metrics calculated
Alternatives	

Figure 4: View evaluated statistical details

Use case name	View candidates analysis details
Pre – condition	Database connection is active. Web application should be running correctly
Actor	User
Main success scenario	4. Navigate to application 5. Click on dashboard 6. Click selected candidate details
Post – condition	View a table of all the attributes with details of each and every candidates get selected
Alternatives	

Figure 5: View candidates analysis details

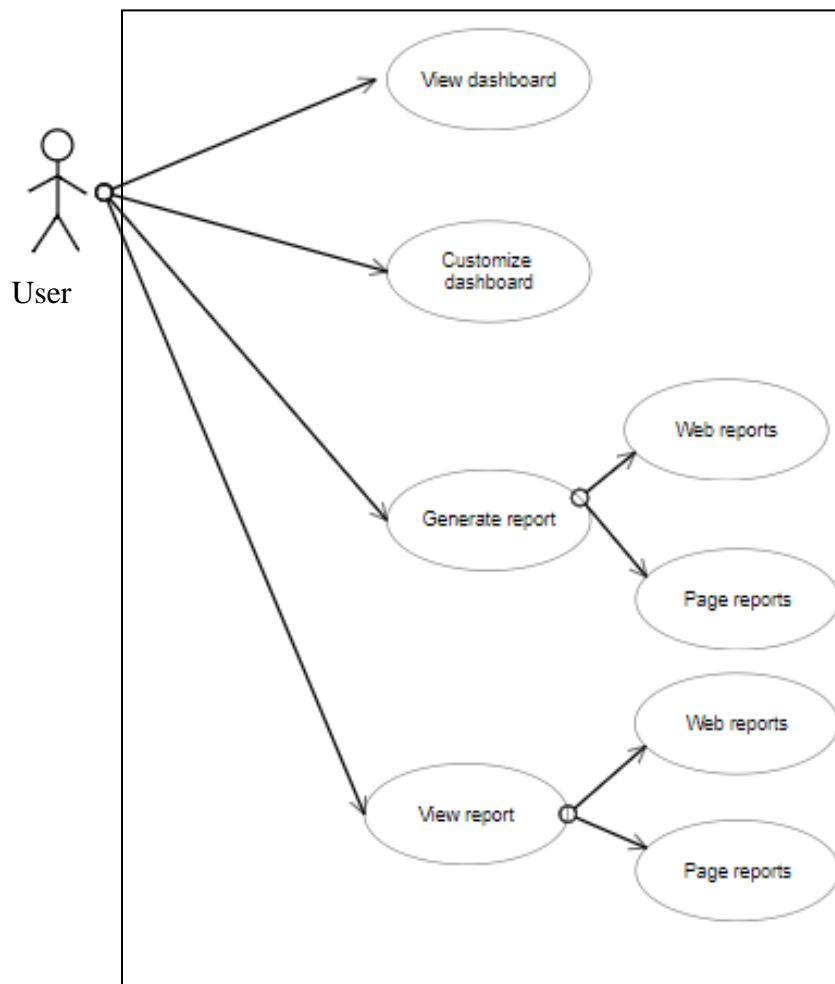


Figure 6: Use case Diagram - View dashboard

### 2.3 User characteristics

User	Privilege	Activities
Top Level Management/Decision Maker	Full access to the system.  Mainly focuses on Visualization layer /Dashboard	1. Extract, Transform and Load data set to data Mart 2. Optimal or Feasible Solution. 3. Interpret predictive results with the dashboard.

Users of the **Smart Resume - model evaluation and dashboard simulation component** are not to be a business intelligence professional to use this system. A user who has an interest and motivation in predictive modelling and its application in business intelligence can use this system. Especially top level management, decision maker, the program manager can use this system who needs to select the best optimal or feasible candidates for a job opportunity based on existing data.

This system is used by several users,

- Top Level Management
- Project Managers

### 2.4 Constraints

A Major constraint that will be facing is the limitation of available time. The project group is expected to complete the **predictive model evaluation and dashboard simulation component** within 4 months.

- All the tools and technologies used for the development should be open source

### 2.5 Assumptions and dependencies

All the users should have basic knowledge using a computer and the internet.

- There is an active internet connection.
- Server is up and running 24x7.
- There's sufficient memory and processing power in all user PC's.

## 2.6 Apportioning of requirements

The requirements described in sections 1 and 2 of this document are referred to as primary specifications; those in section 3 are referred to as requirements (or functional) specifications. The two levels of requirements are intended to be consistent. Inconsistencies are to be logged as defects. In the event that a requirement is stated within both primary and functional specifications, the component will be built from functional specification since it is more detailed.

## 3 Specific requirements<sup>(1)</sup> (for “Object Oriented” products)

### 3.1 External interface requirements

#### 3.1.1 User interfaces

- **Visualization Interface**

This is the interface where all the visualization is done. This will display graphical graphs to indicate a number of candidates, who have selected and who have not selected. After selecting attributes and click on “Generate” button the best candidates for the job will be displayed in the dashboard as a table, with relevant details. The user can click on particular candidates name and view analyzed statistical details of that candidate.

#### 3.1.2 Hardware interfaces

No special hardware interfaces are used for the system

#### 3.1.3 Software interfaces

- **My-SQL:**

My-SQL is used as the database management system. As the **model evaluation and dashboard simulation component** need to access the database, which is created by the model builder, My-SQL would be used regularly for major operations of the system.

- **Apache Server:**

As the visualization dashboard component of the system is planned to be developed using PHP, Apache server will be used as the web server. Dashboard and reports will be available through the web application which is developed on top of apache server.

- **Jython library:**

Jython library is used as the interface between python and Java. All the user interfaces and validations are done using java. Most of the data processing and complex mathematical calculations are done using python. Hence Jython acts as the connector between these two layers.

- Web browsers like Google Chrome, Mozilla Firefox, Opera, Apple Safari, and Internet Explorer to test the visualizations and dynamic reports.

#### 3.1.4 Communication interfaces

- **Internet Connection:**

Although the data processing is done offline, dashboard and reports will be available through the PHP web application which can be accessed by authorized members. So in order to access the web application user needs an internet connection.

- **Database connection interface:**

Database connection interface is used to exchange data between the application and the database. It acts as the adapter which converts the database queries into application data and wise-versa.



### 3.2 Classes/Objects

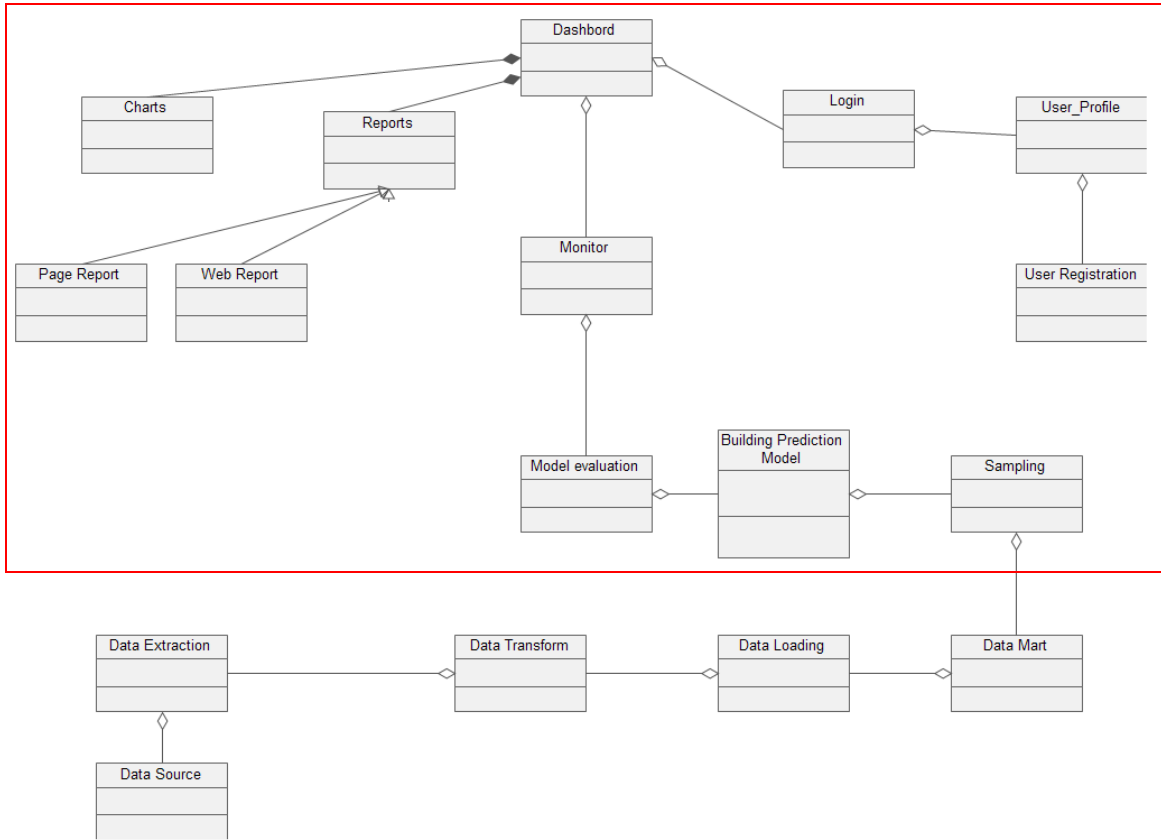


Figure 7: Class Diagram

### 3.3 Performance requirements

It is expected that the proposed system will perform all the requirements stated under the functional requirements section. Performance can be identified as the most important factor of this **model evaluation and dashboard simulation component** as this component process a large data set in order to give the best candidates and demands high responsiveness at the user-end .The following were the proposed performance goals of the component and should be achieved through these modules:

- The system should be able to accommodate a minimum of 50,000 records in the database.
- The predictive list should be generated within 2 hours.

- Based on the derived predictive lists, reports must be generated within 10 seconds.

### 3.4 Design constraints

A Major constraint that will be facing is the limitation of available time. The project group is expected to complete the predictive model component within 4 months.

- The **predictive model evaluation and dashboard simulation component** will be only focusing on IT industry for this 4- month implementation phase.
- The predictive model evaluation, testing and all other operations will be carried out using some old datasets.
- Performance, accuracy, security, reliability, scalability and other goals should be achieved.

### 3.5 Software system attributes

#### 3.5.1 Reliability

Reliability is the ability of the predictive model run with a minimum number of failures. The predictive model has to go through a testing i.e. application must test by fixing each and every possible bug. Each and every sub-component of the **model evaluation and dashboard simulation component** will be tested and finally integrated system also be tested and to make sure the desired output is obtained. Predictive model output also has to be tested to make sure the output is meaningful and optimal. Since the **model evaluation and dashboard simulation component** is a component to give the optimal and feasible prediction result .Hence, the reliability of the component is much expected.

### 3.5.2 Availability

Availability is a key performance measure in the predictive analysis. **Model evaluation and dashboard simulation components** are in working condition at a given time and that should be able to deliver the required services for intended users that range from top-level management to other decision makers. All and necessary information for user requested can be viewed and available at any given time.

### 3.5.3 Security

Security is a crucial non-functional requirement of any system. This must keep all critical details securely. There is a different type of users are using the system. Each and every user has their own access levels within the system and they can only perform tasks according to those access levels. So good policies need to be developed and followed. Data is classified according to different security and handling requirements. All critical data is secured during storage and transmission through proper access controls. Along with this, expected to implement VPN for Visualization tier or this can be high security can be achieved through authenticating and authorizing users.

### 3.5.4 Maintainability

High maintainability is one of the key virtues of stable and highly productive products. Even in predictive model implementation and visualization component implementation, we are more focused on creating a highly maintainable component. The standards of the coding practices will be followed throughout component implementation and that will minimize the bugs as much as possible. The code is commented wherever it is necessary, especially in critical and complex code segments which will help the developers or maintain team for modifications in future. So considering the future potential use of the component, this component is designed in a way that it assists for updates of the component in future. That will lead to maintaining the stability of the **model evaluation and dashboard simulation component**.

### **3.6 Other requirements**

- Data acquisition:

The data for the predictive model building, evaluation and testing should be acquired from valid sources, otherwise it can affect the accuracy of the model.

- Use of source open source technologies:

The product should be developed only using open source technologies.

## 4 Supporting information

### 4.1 Appendices

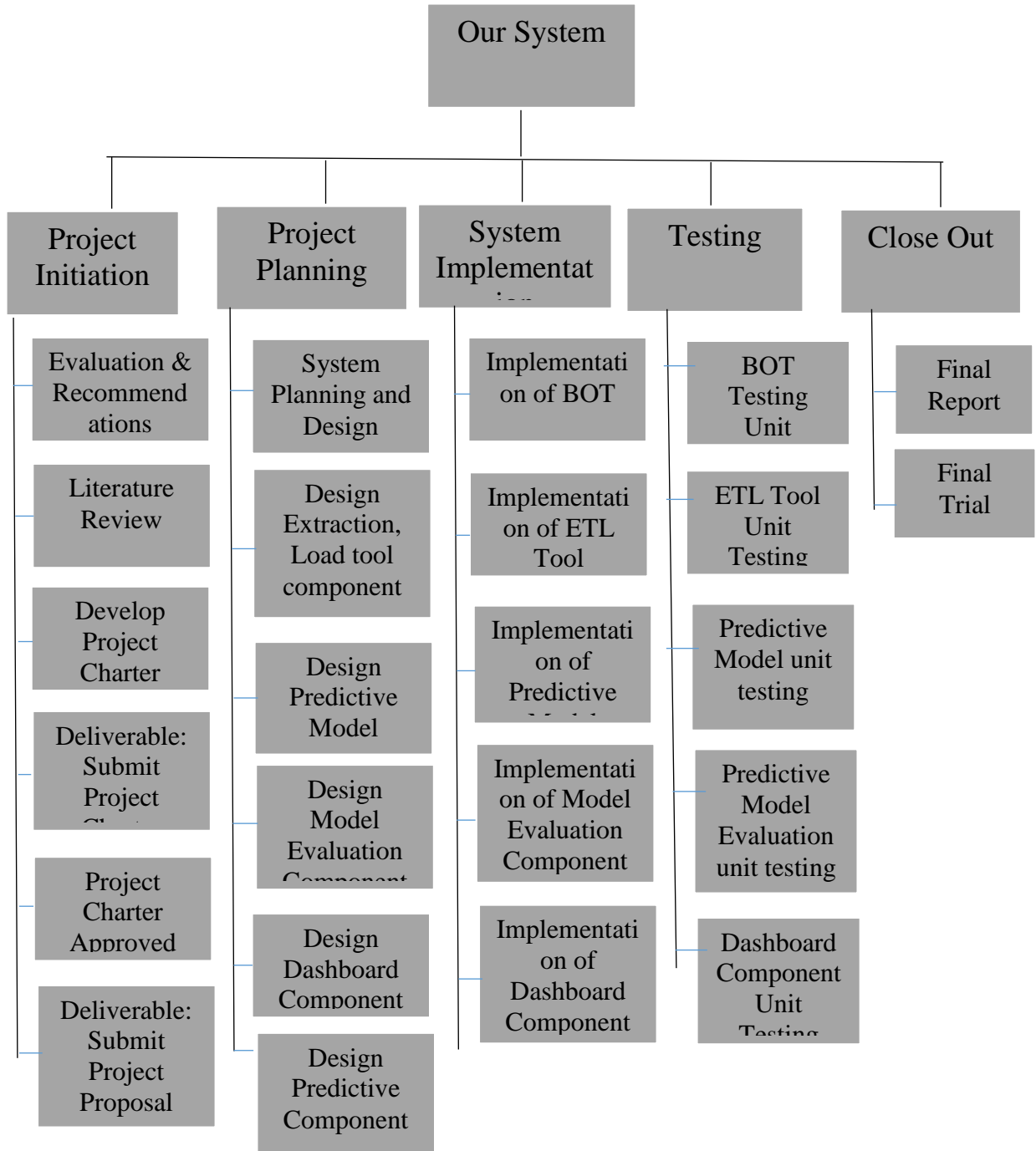


Figure 8: Work breakdown structure

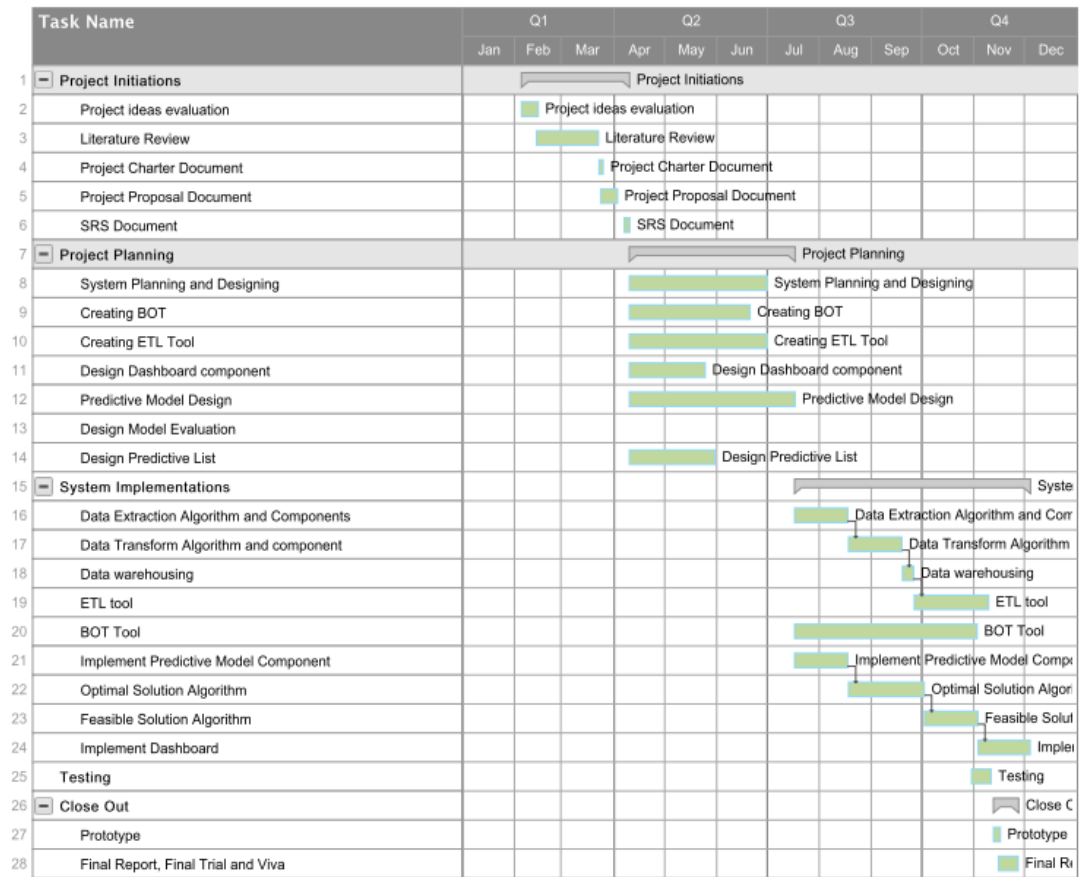


Figure 9: Gantt chart

## 5 References

- [1] Q. A. Al-Radaideh and E. Al Nagi, "Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance", *Thesai.org*, 2012. [Online]. Available: <https://thesai.org/Downloads/Volume3No2/Paper%2025%20-%20Using%20Data%20Mining%20Techniques%20to%20Build%20a%20Classification%20Model%20for%20Predicting%20Employees%20Performance.pdf>. [Accessed: 01-Apr- 2018].
- [2] H. M. and S. M.N., "A REVIEW ON EVALUATION METRICS FOR DATA CLASSIFICATION EVALUATIONS", *Pdfs.semanticscholar.org*, 2015. [Online]. Available:<https://pdfs.semanticscholar.org/6174/3124c2a4b4e550731ac39508c7d18e520979.pdf>. [Accessed: 01- Apr- 2018].
- [3]N. Gayatri, N. S. and A. Reddy, "Performance Analysis of Datamining Algorithms for Software Quality Prediction - IEEE Conference Publication", *Ieeexplore.ieee.org*. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/5329377/>. [Accessed: 06- May- 2018].