Data Mining for Students' Employability Prediction

S.M.Madhavie Malika

189467G

Master of Science in Information Technology

Faculty of Information Technology

University of Moratuwa

Sri Lanka

July 2021

Data Mining for Students' Employability Prediction

S.M.Madhavie Malika

189467G

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfilment of the requirement of the Degree of Master of Science in Information Technology.

July 2021

Declaration

We declare that this thesis is our own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is provided.

Student Details:

Name of Student	Signature of Student
S.M.M.Malika	……
	Date: ……...

Supervisor Details:

Name of Supervisor	Signature of Supervisor
Mr.S.C.Premarathne	……
	Date: ……...

Acknowledgements

I would like to express my sincere gratitude towards my supervisor, Mr. S. C. Premaratne, Senior Lecturer at the Faculty of Information Technology, University of Moratuwa for his patient guidance, encouragement, supervision and advice as well as his precious time spent towards the success of this research project.

Further I'm grateful to my husband and my parents, for their enormous support and encouragement given to me during this time and who experienced all the ups and downs of my research.

Finally, I thank all the staff members of Faculty of Information Technology, University of Moratuwa.

Abstract

Assessing student employability enables a method of integrating student abilities and organizations requirements, which is an important aspect for educational institutions. Improving student-evaluation techniques for employability will assist students to have a better knowledge of business organizations and find the right career for them. As a result, improved student employability prediction can assist students in matching their desirability to company requirements and fitting the employment profile of the firm for which they are searching.

The data is gathered through a survey in which students are asked to fill out a questionnaire in which they may indicate their abilities and academic achievement. This information may be used to determine their competency in a variety of skill categories, including soft skills, problem-solving skills and technical abilities and so on.

Data mining has been used in a variety of fields to efficiently assess large volumes of data. The aim of this study to predict student employability by considering different factors such as skills that the students have gained during their diploma level and time duration with respect to the knowledge they have captured when they expect the placement at the end of graduation by using the data mining techniques. Further during this research most specific skills with relevant to each job category also was identified.

In this research for the prediction of the student employability Rapid Miner software has used and different data mining models such as such as KNN, Naive Bayer's, and Decision Tree were evaluated based on classification techniques. The best model was identified among these models for this institute's student's employability prediction. Further associated technique has been used to identify the most associated skills with respect to each job category. So in this research classification and association techniques were used and evaluated. This study will be expanded to get more data by using a qualitative research, and further the employer's aspects of employability will also consider.

Table of Contents

Page

Acknowledgements ii

Abstract iii

Table of Contents iv

List of Figures vii

List of Tables x

Chapter 1 1

Introduction 1

1.1 Prolegomena 1

1.2 Background and Motivation 2

1.3 Problem Statement 3

1.4 Aims and Objectives 3

1.4.1 Aim 3

1.4.2 Objectives 3

1.5 Proposed Solution 4

1.6 Structure of the Dissertation 4

1.7 Summary 4

Chapter 2 5

Review of Literature 5

2.1 Introduction 5

2.2 Data-Mining Techniques 5

2.3 Data-Mining Process 6

2.4 Review of Previous Work 8

2.4.1 Literature Related to Basics of Employability 8

2.4.2 Literature Related to Factors Affecting for Employability Prediction 9

2.4.3 Literature Related to Data Mining Techniques for Employability Prediction 12
2.5 Summary 14

Chapter 3 15
Technologies and Tools Adopted 15
3.1 Introduction 15
3.2 Data Mining Techniques Used 15
3.3 Rapid Miner Studio 16
3.4. Decision Tree 16
3.5 K-Nearest Neighbors 17
3.6 Naïve Bayes 17
3.7 Association techniques 18
3.8 Summary 18

Chapter 4 19
A Novel Approach to Predict Employability 19
4.1 Introduction 19
4.2 Hypothesis 19
4.3 Input 19
4.4 Output 24
4.5 Process and Features 27
4.6 Summary 27

Chapter 5 28
Analysis and Design 28
5.1 Introduction 28
5.2 Data Collection 28
5.3 Design the classification and Association Tools 28
5.4 Summary 29

Chapter 6 30
Implementation 30
6.1 Introduction 30
6.2 Collection of Data (and Basic Functionalities) 30
6.3 Data Pre-Processing 31
6.4 Classification Models 33
6.5 Association Analysis 35
6.6 Summary 36

Chapter 7 37
Evaluation 37
7.1 Introduction 37
7.2 Evaluation of Classification Techniques 37
7.3 Evaluation of Association Techniques 42
7.4 Summary 60

Chapter 8 61
Conclusion and Further Work 61
8.1 Introduction 61
8.2 Limitations 61
8.3 Future Developments 62
8.4 Summary 62

References 63

List of Figures

Page

Figure 2 1. Data-mining process 6
Figure 3.1 Data-mining techniques 16
Figure 4.1 Questionnaire created on Google Forms 20
Figure 4.2 Professional courses followed by the students 21
Figure 4.3 Time duration to find a job after graduation 22
Figure 4.4 Job categories of business finance students 22
Figure 4.5 Types of skills and their level of importance 23
Figure 4.6 Types of skills and their development during within the ATI (institute) 24
Figure 4.7 Feedbacks for the survey questionnaire 25
Figure 4.8 Sample classification output 26
Figure 4.9 Sample classification test set output 26
Figure 5.1 Designing of pattern recognition using classification methods 29
Figure 6.1 Replace missing values 31
Figure 6 2 Discretization process 32
Figure 6.3 End of data preprocessing final processes 32
Figure 6.4 Decision tree model using the Rapid Miner 33
Figure 6.5 Decision tree model with split validation using the Rapid Miner 34
Figure 6.6 Decision tree model with split validation process 34
Figure 6.7 Decision tree model with split validation- training the model 35
Figure 6.8 Identifying associated skills for each job category 36
Figure 7.1 Accuracy of predicting the job category- using KNN algorithm 38
Figure 7.2 Accuracy of predicting the job category- using Naïve Bayes algorithm 39
Figure 7.3 Accuracy of predicting the job category- using Decision Tree algorithm 39
Figure 7.4 Accuracy of determining the time period to get a job – using KNN algorithm 40
Figure 7.5 Accuracy of determining the time period to get a job – using Naïve Bayes algorithm 40

Figure 7.6 Accuracy of determining the time period to get a job – using Decision Tree algorithm 41
Figure 7.7 Job categories mentioned in the survey questionnaire 42
Figure 7.8 Identify the five most important skills..... 43
Figure 7.9 Process diagram to identify most associated-skills 44
Figure 7.10 Process diagram to identify associated-skills cont.. 44
Figure 7.11 Sample frequency set for accountancy job category in the Rapid Miner..... 45
Figure 7.12 Accountancy job category with one associated skill type 45
Figure 7 13 Accountancy job category with two associated skill types 46
Figure 7.14 Accountancy job category with three associated skill types 46
Figure 7.15 Accountancy job category with four associated skill types..... 47
Figure 7.16 Auditing job category with one associated skill type 47
Figure 7.17 Auditing job category with two associated skill types 47
Figure 7.18 Auditing job category with three associated skill types 48
Figure 7.19 The most associated four skill types for auditing job category 48
Figure 7.20 Banking and finance job category with one associated skill type 48
Figure 7.21 Banking and finance job category with two associated skill types 49
Figure 7.22 Banking and finance job category with three associated skill types 49
Figure 7.23 Banking and finance job category with four associated skill types..... 50
Figure 7.24 Customer service job category with one associated skill type 50
Figure 7.25 Customer service job category with two associated skill types 51
Figure 7.26 Customer service job category with three associated skill types 51
Figure 7.27 Customer service job category with four associated skill types..... 52
Figure 7.28 Management and supervision job category with one associated skill type ... 52
Figure 7.29 Management and supervision job category with two associated skill types . 52
Figure 7.30 Management and supervision job category with three associated skill types 53
Figure 7.31 Management and supervision job category with four associated skill types. 53
Figure 7.32 Marketing business services and public relations job category with one associated skill type 54
Figure 7.33 Marketing business services and public relations job category with two associated skill types..... 54

Figure 7.34 Marketing business services and public relations job category with three associated skill types 55
Figure 7.35 Marketing business services and public relations job category with four associated skill types 55
Figure 7.36 Marketing business services and public relations job category with five associated skill types 56
Figure 7.37 Administration job category with one associated skill type 56
Figure 7.38 Administration job category with two associated skill types 57
Figure 7.39 Administration job category with three associated skill types 57
Figure 7.40 Administration job category with four associated skill types 58
Figure 7.41 Administration job category with five associated skill types 58
Figure 7.42 Whole sale and retail trade job category with one associated skill type 58
Figure 7.43 Whole sale and retail trade job category with two associated skill types 59
Figure 7.44 Whole sale and retail trade category with three associated skill types 59
Figure 7.45 Whole sale and retail trade job category with four associated skill types 60

List of Tables

Page

Table 7.1 Accuracy of prediction the job category according to the skills..... 39
Table 7.2 Accuracy of determining the time period to get a job 41

Chapter 1

Introduction

1.1 Prolegomena

Mostly in education sector uses prediction-based analysis to predict student placement as an employability which has become one of the crucial area in the present world. Vast number of students take admission on government institutes with the hope of acquiring their dream job. So, it would be great if students and the institute lecturers can get an idea about placement crucial factors beforehand. Many businesses nowadays place a greater premium on practical and provable abilities rather than academic subjects studied or grades obtained in institutions.

This predictive analysis will be done using the Higher National Diploma in Business Finance (HNDBF) course students who are studying at a higher education institute of government. This business-finance course cater students a solid business and finance foundation. Specific knowledge in a range of subjects, as well as finance related technical skills, management, and business planning, will be provided to students. The program allows learners to build logical, critical, analytical, problem-solving, and soft skills in the process. The main aim of business finance programme is to produce the middle level professionals in business finance. Those middle level finance professional play and important role in the business organization.

These students will be able to get the higher national diploma at the end of two institutional academic periods and half year industry training. So, the diplomates finally have to compete with the graduates who passed out from state universities as well as the private universities and professional organizations.

Therefore, it is very important to identify important skills, determine the time duration to get an expected job with their knowledge and skills after completion of the diploma.

1.2 Background and Motivation

Education is the most important sector for the students in the today's business world. Students' performance and the employability is very important and amongst employability is crucial for the education institutes in the present world when government institutes take enrollment of the fresher graduates.

Universities around the world are becoming more aware of the importance of employability and have responded by implementing in place relevant measures within the implementation of degree programs to help learners gain the required interpersonal and application knowledge that will assist them to find a placement in their chosen subject of practice. Candidate selection in accounting and finance can have a significant impact on a company's operational performance and strategic direction.

Traditional educational setup refers to classroom teaching, which is still the most common method of teaching used by institutions around the world and should be researched further. Students want to enroll in an institution that has a strong academic record and a high rate of employability among its graduates. Predicting academic achievement as well as employability can assist management in identifying students who are at risk of poor educational performance and low employability. The prediction process entails the use of various data mining methods to forecast dependent variables based on independent variables.

Depending on diplomates students' professional skills gained from courses followed other than this business finance course or from other degree programs and soft skills developed the waiting duration to get a job also will vary. Further it is very import to predict what kind of skills they need to develop to get the desired job and as well as the time duration.

Therefore, better prediction model is required by using large set to predict the employability.

1.3 Problem Statement

These business finance students will be able to get the higher national diploma at the end of institutional academic periods and industry training. The diplomates finally have to compete with the graduates who passed out from state universities as well as the private universities and professional organizations.

The lack of identifying skills in diplomates in business finance in higher education institutions is a major problem. So, it is critical to recognize employable abilities of business finance students and determining pattern and identify future trend. No research was found regarding the employability of higher national diplomates in Sri Lanka or the foreign countries.

Mostly in education sector uses prediction-based analysis to predict student placement as employability which has become one of the crucial fields in the present world. It is important for institute lecturers and students to get an idea about placement in future. So, we decided to introduce a method to predict the employability by using model formation in data mining. Ultimately it can be used to predict the student employability.

1.4 Aims and Objectives

1.4.1 Aim

To explore the factors affecting for business finance students' employability using data mining techniques.

1.4.2 Objectives

- To conduct a survey for the passed out Higher National Diploma in Business Finance students who are currently in the business organizations.
- To analyze currently available data mining tools and software tools using literature.
- To develop in-depth knowledge on the employability prediction techniques and analyzing factors affecting for that.
- To identify the most appropriate data mining technique/s that can be used for employability prediction in higher education institute.

- To propose a design for the employability prediction and skill identification based on the research results.

1.5 Proposed Solution

We propose a survey questionnaire and it distribute online using Google Forms and collect employability data of already passed out students who are already in the business organization. It is easy to predict the employability of currently pass out business finance diploma students using this information. We model the employability using the best subset of employability parameters, and then use multiple classification algorithms and association algorithms to select the best factors affect for the employability.

1.6 Structure of the Dissertation

The overall thesis has structured. The first chapter is about introduction to the project with the aims and objectives, background and problem. Chapter two has critically reviewed the literature in the data mining technology, in the employability prediction with reference to the classification and association techniques. The third chapter has allocated for the technologies and tools adopted and data mining technology with relevant to the employability prediction. The fourth chapter has described the approach, inputs, outputs, the process and features. The fifth chapter has allocated to the analysis and design of the solution of this study. The sixth chapter is for the implementation of the solution of this study. The seventh chapter has based on evaluation of the solution. Finally eighth chapter has elaborated the consequences of this study. Further in this chapter has discussed the limitations and future works also.

1.7 Summary

This chapter was discussed background and motivation, problem definition, aim and objectives of this study. In the next chapter past literature has critically discussed and analyzed.

Chapter 2

Review of Literature

2.1 Introduction

The techniques in datamining are used to predict the employability of the business finance students and it is critically reviewed in this chapter. In this perspective we first discuss the data mining techniques general usage. Then identify the issues and concerns inherent in the techniques for data mining in prediction the student employability. Further in this research basics of employability, different factors affects for employability prediction and data mining methods which were used in the previous literature also analyzed. Finally, in this study research problem was addressed. In this chapter data mining techniques are also identifying that could be used to extend the problems solution.

2.2 Data-Mining Techniques

Data mining comprises gathering and analyzing historical data in order to uncover patterns, correlations, and regularity in big data sets. So data mining techniques are effective for a wide range of activities. Clustering, classification, association rule mining and regression are a few of them. Each of these techniques is critical in any situation.

There are a variety of methods for organizing and analyzing historical data. Which strategy to use is largely determined by our goal (e.g., prediction, inference of correlations) and the type of data we have (structured vs. unstructured). We may end up with a certain data arrangement that is suitable for one purpose but not for another. As a result, in order to make data useable, one must be aware of the theoretical models and methodologies utilized in data mining, as well as the potential trade-offs and problems associated with each.

2.3 Data-Mining Process

A process of data-mining has broken down into various steps, as shown in Figure 2.1. These steps are interactive and comprise cleaning the data, Integration of data, selection, transformation, data-mining, pattern examining, and demonstrating the knowledge. Users has participated either directly or indirectly.

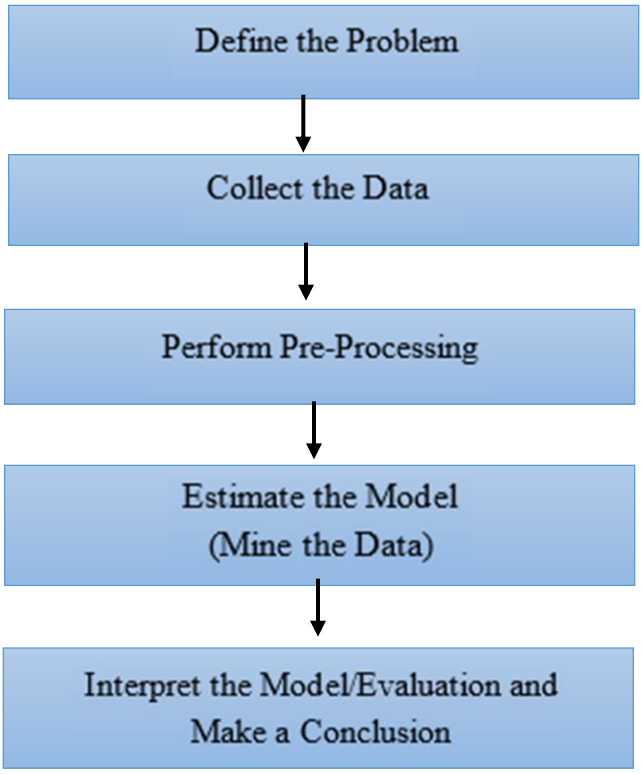


Figure 2 1. Data-mining process

- Define the Problem

  Establish the scope of the business problem and the data exploration project's goals.

- Collect the Data

  It's critical to know that data collection affects theoretical distribution because this information is frequently used in forecasting as well as, thereafter, in ultimate findings analysis. It's also essential for make sure that the information included to

estimate such approach, and information then used to verify and validate the findings, comes from a comparable, unpredictable sample data.

- Perform Pre-Processing

  In data-mining the data cleaning and preparing is very important. In order to be used in various analytical techniques, raw data must be cleaned and pre-processed.

  In most instances, in the real world data is insufficient, unclear, and inaccurate. It's possible that the information found in databases is data of relevance, absent input parameters and so forth. There are occasions when data contains inaccuracies or outliers. The identification of a worker, for example, could be maintained in many database table or files. So, the information is ambiguous. If the data set isn't clean, the data analysis findings will not be reliable or specific.

  Data-preprocessing typically comprises of two regular activities:

  - Outlier Recognition (and elimination) :
    In relation to the most of observations, outliers are figures which are out of the ordinary. Outliers are often the consequence of measuring, coding, or recording errors, but they can also be natural, anomalous results. Such non-representative samples can have a significant impact on the model created later.
    Outliers can be dealt with in two ways: as part of the preprocessing phase, detect and eventually eliminate outliers. Also, provide strong modeling methods which will not be affected by outliers.

  - Scaling, encoding, and choosing characteristics:
    Variable scaling and different forms of encoding are examples of data preparation steps. For example, within the applied technique, one feature with a range of [0, 1] and another with a range of [100, 1000] will not have the same weight. They will also have a varied impact on the final data-

mining outcomes. As a result, it's best to measure them and give those categories the similar mass for future study.

Preprocessing methods, in general, incorporating prior knowledge into implementation, give the optimal depiction for a data-mining technique.

- Evaluation of Model (Mining Data)

  Choosing and implementing a suitable data-mining method is a key duty. In fact, most deployments are based on different of models, with selecting the simplest one being a separate procedure.

- Interpret the Model and Make a Conclusion

  Data-mining models should, in most situations, assist in decision-making. If the models are simple then usually easier to understand, but they are also less accurate. When use high-dimensional models, modern data-mining approaches produce very accurate findings. The crucial work of interpreting these models is treated as a separate endeavor, with specific procedures for validating outcomes.

  To display the results of data mining to users, a variety of knowledge representation and visualization strategies are employed.

2.4 Review of Previous Work

2.4.1 Literature Related to Basics of Employability

Education is the most essential area for students. When government institutions take on fresher graduates, students' performance and employability are incredibly important, and employability is critical for education institutes in the current world. Employability has been defined in a variety of ways, including the capability to acquire a career, get a career

in a certain duration following the diploma awarding, the capacity of skillfulness of oneself according to workplace requirements, and student's willingness to continue graduate learning at work [1].

Institutions must provide learners with the essential competences to join the job marketplace and develop their capacities to meet particular workplace demands in order to maximize the possibility of graduates getting respectable employment that match their education and training. Education institutes around the world are becoming more aware of the importance of employability and have responded by implementing academic theoretical knowledge and the skills and those will help them to find a career in their stream of studied [2, 3]. Candidate selection in accounting and finance may have a significant impact on a company's operational success and strategic direction, since both academic theoretical knowledge and the skills are critical [4, 5, 6].

Learning using cases, team's work on solving problems and interactive quizzes and project-based learning are among the approaches of instruction and learning used during the teaching space, according to a prior study based on a Middle East study [7]. Students need to attend industrial trainings, placements, and contests outside of the classroom, which enable new graduates to see the employment market.

2.4.2 Literature Related to Factors Affecting for Employability Prediction

According to the literature required graduate qualities as stated by companies, proficient bookkeeping groups and administration mechanisms were not matched with program records and graduate opinion. Those researchers stated that colleges really need to directly integrate curriculum and evaluation with the abilities that businesses really demand [3].

In instructing and learning setting, it is critical to examine graduation [3, 8]. Most of the trainees in accounting not have sufficient writing as well as spoken abilities and characteristics for effective communication in the accounting industry [2]. The most significant abilities required in entry-level accountants, according to practitioners, are writing and spoken communication duties such as generating relevant documents and

engaging in informal discussions. It is vital to recognize the needed written, spoken, presenting abilities in accounting graduates [2].

According to findings of this study employers demand a variety of common abilities, that graduates were not sufficiently studied in their bookkeeping diploma program,. Against this context of skill convergence, the most significant areas of skill dispersion, according to employers, were management abilities, speaking abilities, team skills as well as graduates' personal abilities [5].

Learners' capability relating for assessment principles to the structural, syntactic and demonstration modules of writing correspondence improved significantly [4]. Students need to get basic professional skills in accounting education, such as investigative ability, critical thoughtful abilities and communication which are essential for career advancement in the area [9].

Poor English language proficiency has affected to lower employability and English language knowledge has a greater impact for the employability [10, 11]. Another study [12] found that knowledge of GPA and English language fluency are necessary for students to continue working, and that female applicants outperformed than male ones in campus placement.

Researchers [13] suggest that problem-based learning be developed and implemented in accounting programs, based on data obtained through student surveys across terms wise in the year. Students said problem-based learning was usually beneficial, especially in terms of fostering questioning, collaboration, and problem-solving abilities.

Prior to commencing a work in the accounting in the industry, accounting graduates are expected to know how to use Excel functions. Their findings show that accounting students undervalue the significance and use of several Excel features. Their findings also show that new recruits' perceptions of practice certain operations in their bookkeeping profession differ from students' perceptions in how to utilize some of the same Excel functions [14].

In the literature there are fifteen graduate abilities and features recognized and which are important for administration learning pursuing a finance as well as accounting program at

institutes [15]. Communication, technical skills, creativity, interpersonal skills, teamwork, problem-solving, leadership, thought of critically, managing time, self-management, innovation skills, planning and organizing, economic skills and investigative abilities were the most often stated of these skills. Past out learners with strong communicating [15] and personal abilities are the most marketable, and listening and oral skills are given special consideration when interviewing and recruiting graduates for accountancy positions [16, 17]. Accounting and finance graduates should be able to spot trends in complicated financial records, as well as test and predict numerous methods and solutions [5, 18, 19].

Leadership abilities are mentioned in much of the research as another major aspect for business and management students, including finance and accounting students. The capacity to encourage and guide people to achievement is considered as leadership [15]. Self-management also include the willingness to take on new responsibilities, the willingness to improve one's performance based on comments, time management skills, flexibility, resilience, reflective learning and assertiveness [18]. The learner's assurance in dealing with individual and work-associated difficulties, is frequently connected to self-management [19]. The capacity of applicants to managing the timing at work as well as operate under stress is rewarded by larger organizations [20].

Institutions are using cutting-edge technologies such as data mining to process the massive amounts of data created in the classroom, which includes academic, behavioral, demographic, and faculty data. The data collected in a learning environment can provide a wealth of information on the educational process [21]. It has been determined that, as the paradigm shifts from product based to service based, the curriculum and delivery mode of lectures must develop in order to improve employability with accordance to what the industry is expecting [22]. In a research study [23] which was conducted using Malaysian graduates, it was identified that the employers expect from employees some personal attributes such as commitment, loyalty, honesty, enthusiasm, positive self-esteem, integrity, reliability, common sense, presentation, and a great sense of humor. Further they expected motivation, adaptability, a balanced attitude toward work and personal life, and the ability to deal with tension.

Employability is influenced by the employee's general personality as well as career self-efficacy and competence. According to a study [24] which express the above idea, the employers prioritize soft skills, teamwork, and some other factors. The authors created an employability prediction model based on real source data that included academic, social and emotional skills [25].

There are five methods to developing employable skills through entrepreneurship [26]. One of those is personal development, in which students must recognize the purpose of studying, such as developing a career plan. Apart from that, use what you have learned to study theoretical and cognitive activities, as well as transferring skills from academia to the job. Another is skill development, which allows you to practice and receive credit for your progress in areas such as people, personal and task. Furthermore, work-based learning is an important component of any degree, as it allows for personal growth, practical learning, and skill development. Finally, there will be motivation to participate in continuous professional development activities when it comes to career management.

When evaluating an applicant's employability, practical abilities, methodological abilities [27] and software awareness are critical [28, 29]. In general, these abilities indicate to the organization authorities that a new pass out student has learned the precise competencies required to be fruitful in a given position. It's worth noting that practical abilities are more framework-dependent than easy-going abilities and solving problems abilities.

2.4.3 Literature Related to Data Mining Techniques for Employability Prediction

The commercial value of an institution is determined by the employability of its graduates. Research is necessary to build comprehensive models for employability and to develop a system that can predict employability. Techniques like data mining can be used to extract valuable and relevant information from massive volumes of data. Classification and prediction, clustering and association rule mining are just a few of the activities that data mining can perform. Furthermore, building classification models from an input data set is very useful approaches in m mining data. Classification analysis are frequently related to create models that anticipate future data patterns.

In the past, academics have evaluated many parameters for prediction and attempted to forecast the outcome of a single course at a specific university using various algorithms. The authors considered learners from two distinct institutions. Institutions, one of which is international and one of which is small, to examine the impact of research on both independently. In the above two situations, the decision tree was identified as more accurate model than the Bayesian Network [30]. Some researchers [31] came at the conclusion that the combination of Decision tree and Nave Bayes are ideal for prediction of student performance in the event of a New Zealand Polytechnic. Decision tree algorithm and Nave Bayes are two examples of data classification algorithms in data mining but there more classifications methods [32].

The writers of a research study [24] looked at graduate students at Maejo University in Thailand to see if they were in employment within a year of graduation. The study also highlights characteristics of talented graduates. Bayesian Network and decision tree are the approaches followed. In this study the finalized the conclusion that the Bayesian algorithm is a superior methods since the performance achieved is 99.77 % correctness related to the J48 method 98.31% correctness.

Random Forest methods obtain maximum correctness of 99% when related to further classification techniques such as KNN, decision tree [33], Logistic Regression, Naive Bayes, and Support Vector Machine (SVM) [34] as evaluated on the learners' competencies dataset. As a result, the Random Forest algorithm has utilized to the dataset from the learners' job skills in order to forecast their employability [35].

Further the Decision Tree is among the greatest commonly used methods because that generates a Decision Tree out of the dataset by means of humble calculations based primarily on the gain-ratio calculation that routinely assigns masses to characteristics and allows the investigator to identify the greatest operative characteristics based on the forecast goal. A decision tree would be created as a consequence of this approach, with classification rules derived from it [36], as well as another classification called the Nave Bayes classifier. A target class is predicted using this categorization. It is based on probability calculations, specifically the Bayesian theorem. The classifier's findings are

more accurate and efficient as a result of this application, and further responsive to fresh facts introduced to the database [36].

It is apparent that Decision Tree and Support Vector Machine surpass Nave-Bayes and KNN in accurately forecasting student employability with 98% accuracy. It is also apparent that characteristics such as aptitude, communication skills, and reasoning abilities, having a mentor, family economic condition, and teaching quality, among others, have a major influence on students' employability [37]. This valuable information at the start of the course will support higher educational institutions in making excellent content during academic planning in order to improve their learning strategies in order to prepare students with career development skills and to conduct proper pedagogical intervention strategies with students from the start of the program in order to make them employable by the end of the program.

The research [38] was carried out with the help of 228 Finance and Economics students from Zhejiang University's School of Information in China. The above sample of dataset has used to investigate the connection between fresh graduates' employment choices and their educational achievement. The association rule mining is accomplished using the Apriori method, which is based on R programming.

2.5 Summary

This chapter discussed other related works to employability prediction analysis with a specific reference to data mining. The employability of students who graduate from an institution determines the college's institution market worth. So, research is necessary to build comprehensive models for employability and to design a system that can forecast employability. We have defined the research problem and also identified the factors addressing the research problem. Further according to the previous literature, we also identified the possible data mining methods such as classification and association which can be used for solving the research problem. In the next chapter technologies and tools used has discussed.

Chapter 3

Technologies and Tools Adopted

3.1 Introduction

In the previous chapter we discussed different findings which were done in the area of employability predictions, factors affects for the employability, issues as well as datamining methods suitable for the research study. We defined our research problem and also identified factors that we are going to evaluate using data mining classification as the technology to address the problem mainly and association techniques to assess the different kind of skills relevant to job categories. This chapter highlights the effectiveness of selected technology.

3.2 Data Mining Techniques Used

Data mining has been used in a variety of fields due to its ability to quickly evaluate large amounts of data. This has been done by various data analysis methods. Data mining techniques are effective for a wide range of activities and clustering, classification, association rule mining and regression are a few of them. In any case, each of these approaches is essential. In career development, data mining techniques are used to look for meaningful interactions, associations, and fluctuations in attributes in data. So among these techniques, classification will be used to build employability prediction model in this research. Classification use to assign new data to the relevant classes based on previous categorized data and help to do predictions for the future. Naive Bayes, KNN & Decision Tree will be applied as classification algorithms with relevant to this research. Further Association technique will be used to identify the specific skills relevant to different job categories. In *figure 3.1* shows the classification techniques used in this research.

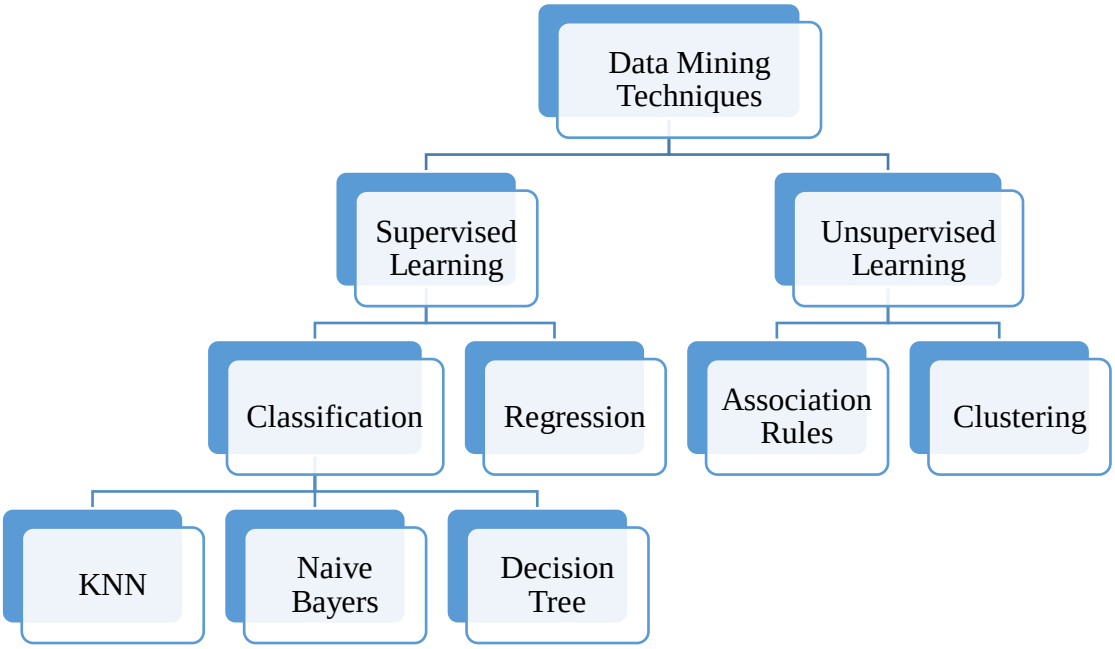


Figure 3.1 Data-mining techniques

3.3 Rapid Miner Studio

Rapid Miner Studio is a free and open source software for facts or text mining and. Further it is a powerful and productive data mining software that can handle large datasets. It supports Macintosh, Linux, and UNIX operating systems in addition to Windows. It is featured as a stand-alone data or text analysis application.

It provides a graphical user interface for machine learning and that functions as data preparation, deep learning, predictive analytics and text mining. In this research this software has been utilized in the data mining process to obtain more precise results.

3.4. Decision Tree

This method is among the utmost effective and extensively applied approaches for both prediction and classification. Decision tree is a more powerful method and it can be most widely used for classification and prediction. It is a flowchart which is very simple to understand the data and very comfortable with human level thinking to make some good

interpretations. It has a structure like a tree with internal nodes, branches and leaves. It is non- parametric and distribution-free method which can be used for high dimensional data to get more accurate results. Using feature selection measures, each feature which related to the given dataset has been ranked with scores. The most widely utilized selection criteria are Information Gain Ratio, Gain, and Gini Index. Best scored feature can be selected to split the dataset in to possible fine way. This best feature is considered as a decision node and divide the data set in to subsets. By repeating the same process, decision tree can be built up.

3.5 K-Nearest Neighbors

The K-Nearest-Neighbors (k-NN) technique has a basic and straightforward characteristics, easy-to-apply supervised machine learning method which may be used for regression and classification. The K-NN algorithm is a non-parametric algorithm, meaning it does not mark any simplifying assumptions on data.

The K-NN approach keeps all available facts and recognizes a fresh data point depending on how similar it is to the existing data. This algorithm is applied in two steps. The algorithm must initially be trained to recognize specific classes as a first step. Then, based on the classification, it can be utilized to generate an accurate prediction.

3.6 Naïve Bayes

The Bayes theorem is used to create the Naive Bayes classification algorithm. According to this, a given feature is assumed to be independent of the other features in a class. This classifier is a widely used and simple statistical technique that may be used to classify a huge number of data while taking probability into account. It brings more efficient and accurate outcome than complex methods. As a result, it's extensively used in email filtering, recommendation systems, and sentiment analysis, among other things.

Naive Bayes is an extremely quick and flexible method that may be used for both binary and multiclass classification. GaussianNB, MultinomialNB, and BernoulliNB are the

three most common types of algorithms. This technique may also be quickly trained on small datasets and used to massive amounts of data.

3.7 Association techniques

Association analysis is a data mining technique that serves as a foundation for a variety of other data mining techniques. It is a form of unsupervised learning strategy that looks for data items that are dependent on each other and maps them accordingly. It finds interesting associations between variables in a database using a set of criteria. The most important aspects in machine learning is association rule learning, which is used in Market Basket analysis, Web usage mining, continuous production, and other applications, etc. In this research association analysis has used to identify the most associated skills with respect to the job category.

3.8 Summary

In this chapter explains the technologies and tools which have been applied to conclude this study. In here, Data mining is applied as the key approach for implementing the classification data model and association rule mining also has used. The next chapter shows a novel approach to predict employability by using the presented technologies and tools in this chapter. In the next chapter approach to address the problem of predicting employability has discussed.

Chapter 4

A Novel Approach to Predict Employability

4.1 Introduction

Previous chapter presented the technology applied to answer the investigation problem in this study. This chapter described the approach to solve the problem of predicting employability accurately by using data mining techniques. So in this chapter approach of the current research is presented by highlighting hypothesis, input, output, process, users and features of the solution.

4.2 Hypothesis

We hypothesize that the issue of an unavailable proper mechanism to predict diploma holders' employability and this can be achieved by using classifier analysis. We are going to use various classification technologies such as Naïve Bayes, Decision Tree and K-Nearest Neighbor and then finally pick up the most accurate classifying techniques based on the data model.

4.3 Input

For conducting this research, business finance students of a government higher education institute were considered and those students were the already passed out students from academic years 2014 to 2017. To the research a survey questionnaire was created in Google Form and it was shared among these passed out students. So this survey questionnaire was distributed online and collected the feedbacks to an online Google Sheet for data analysis purpose. This questionnaire was mainly divided into three sub sections such as background factors, work situation and rating of work-related skills and competences as in figure 4.1.

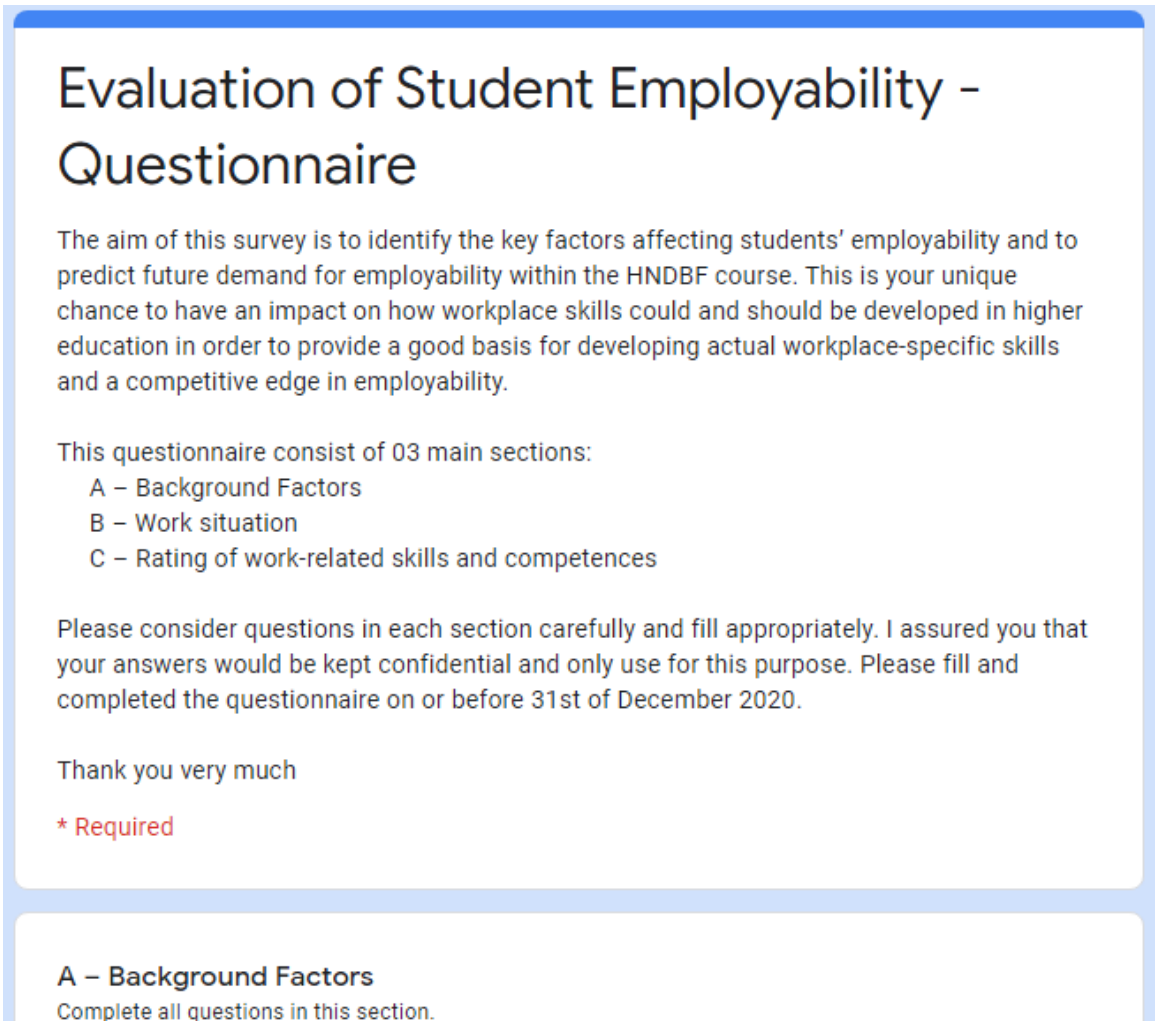


Figure 4.1 Questionnaire created on Google Forms

By using online questionnaires all the employability related attributed was collected. Some of the important attributes used in this research were shown in following figures.

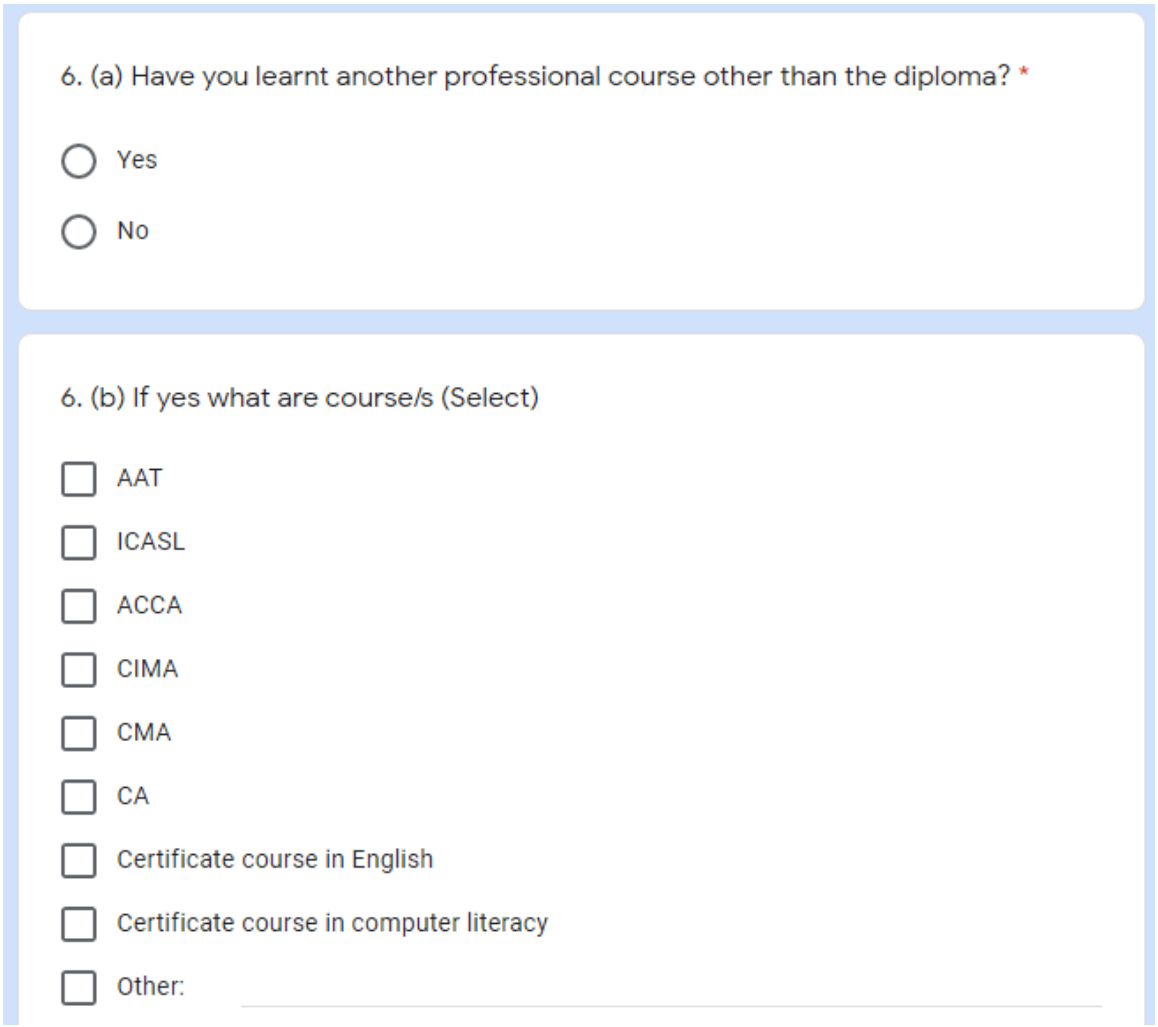


Figure 4.2 Professional courses followed by the students

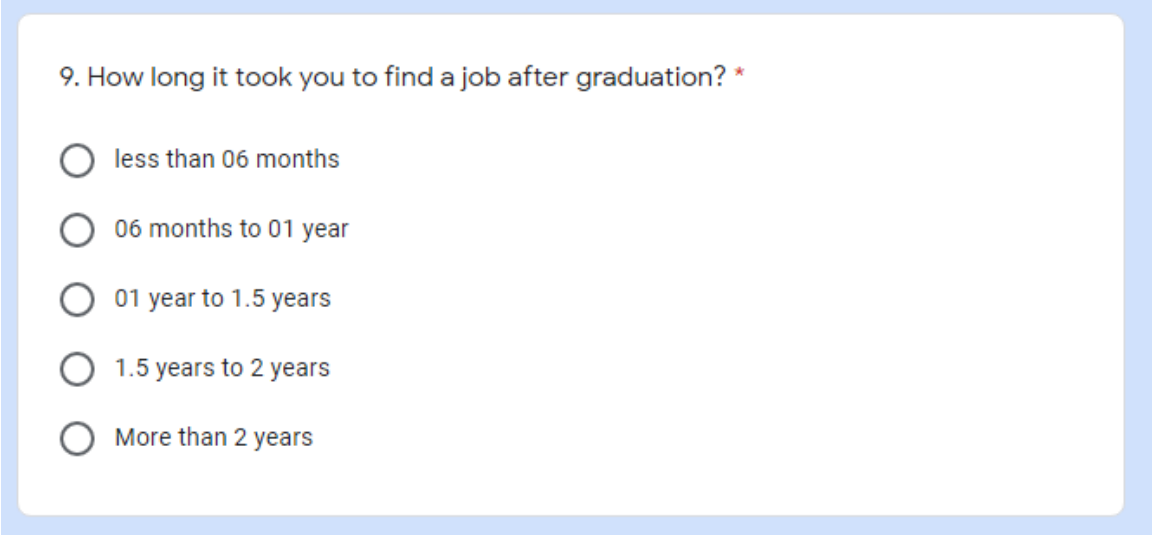


Figure 4.3 Time duration to find a job after graduation

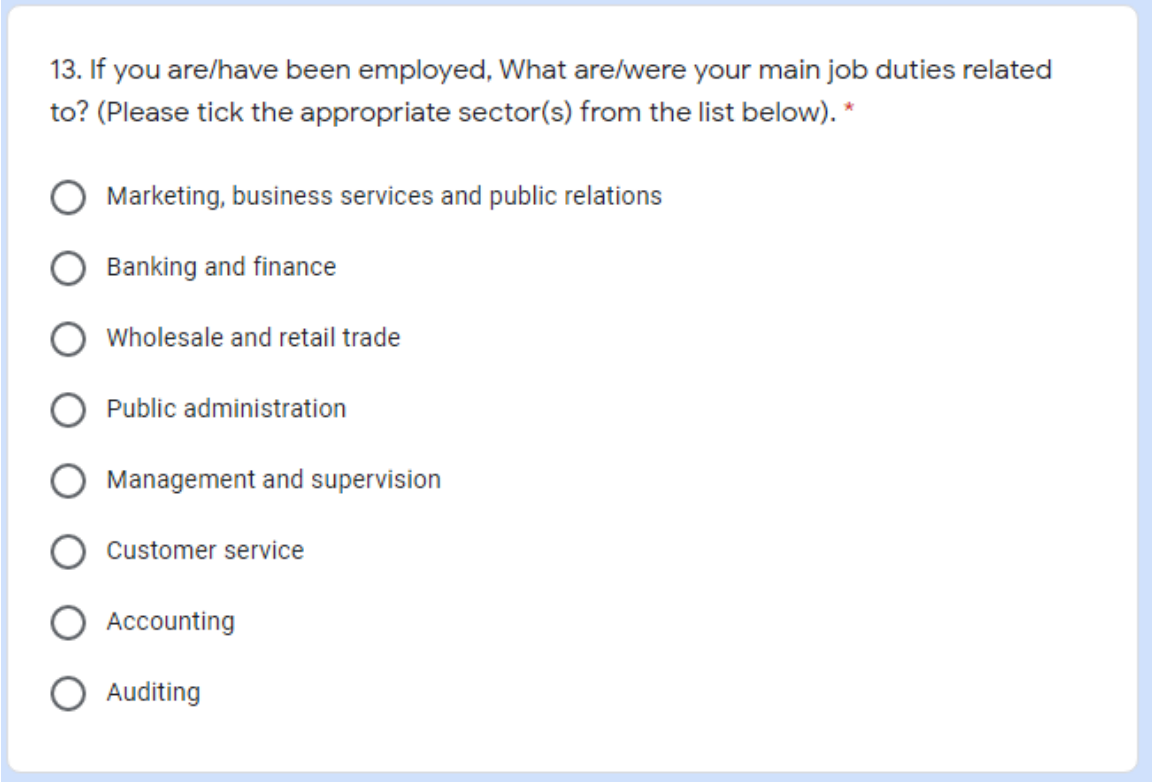
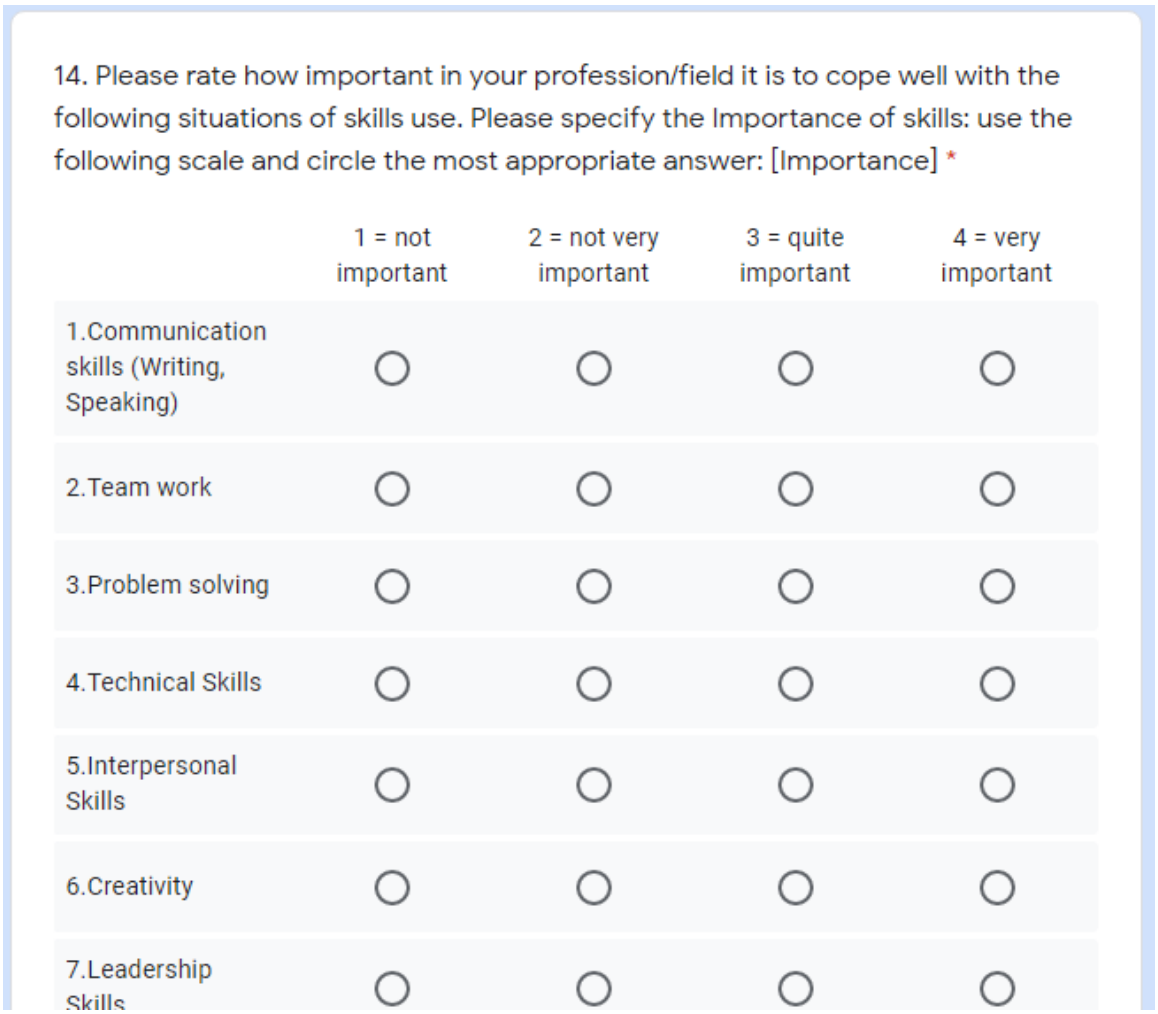


Figure 4.4 Job categories of business finance students



14. Please rate how important in your profession/field it is to cope well with the following situations of skills use. Please specify the Importance of skills: use the following scale and circle the most appropriate answer: [Importance] *

	1 = not important	2 = not very important	3 = quite important	4 = very important
1.Communication skills (Writing, Speaking)	○	○	○	○
2.Team work	○	○	○	○
3.Problem solving	○	○	○	○
4.Technical Skills	○	○	○	○
5.Interpersonal Skills	○	○	○	○
6.Creativity	○	○	○	○
7.Leadership Skills	○	○	○	○

Figure 4.5 Types of skills and their level of importance


15. Please also rate to what level the skills needed in these situations were developed during your ATI study. Please specify the level of skills: use the following scale and circle the most appropriate answer: [Level to which developed during diploma study] *

	1 = not developed at ATI	2 = not developed very well at ATI	3 = developed quite well at ATI	4 = developed very well at ATI
1.Communication skills (Writing, Speaking)	○	○	○	○
2.Team work	○	○	○	○
3.Problem solving	○	○	○	○
4.Technical Skills	○	○	○	○
5.Interpersonal Skills	○	○	○	○
6.Creativity	○	○	○	○

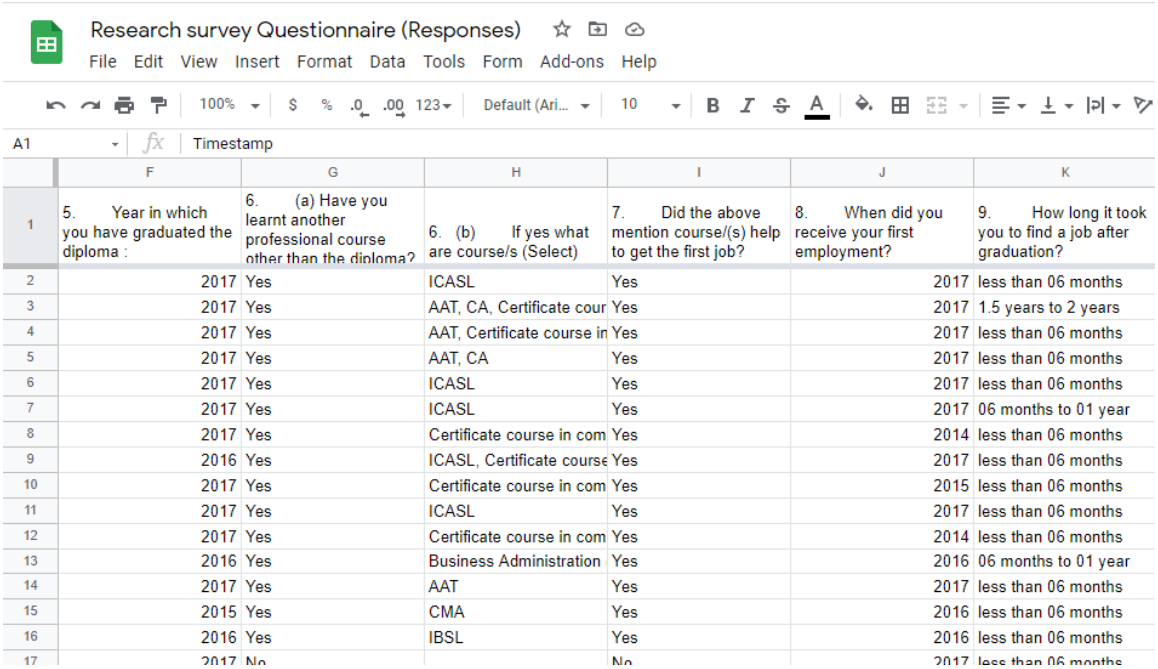
Figure 4.6 Types of skills and their development during within the ATI (institute)

The above are the some of the attributes used for preparation of the questionnaire and then those were highly uses for the data mining analysis. These values went through multiple stages of pre-processing prior to mining and used as inputs to Rapid Miner.

4.4 Output

As the main output of this process, we obtained the students feedback sheet which was collected using an online Google sheets as shown in the figure 4.7. And further we could be able to receive the job categories and time durations to get the particular job with regards to the skills that the students have acquired. Moreover, types of skills need to develop to

get a particular job also can be identified. These kinds of outputs were received after the classification and association in the data mining.



Research survey Questionnaire (Responses)

File Edit View Insert Format Data Tools Form Add-ons Help

A1 Timestamp

	F	G	H	I	J	K
1	5. Year in which you have graduated the diploma :	6. (a) Have you learnt another professional course other than the diploma?	6. (b) If yes what are course/s (Select)	7. Did the above mention course/(s) help to get the first job?	8. When did you receive your first employment?	9. How long it took you to find a job after graduation?
2	2017	Yes	ICASL	Yes	2017	less than 06 months
3	2017	Yes	AAT, CA, Certificate cour	Yes	2017	1.5 years to 2 years
4	2017	Yes	AAT, Certificate course in	Yes	2017	less than 06 months
5	2017	Yes	AAT, CA	Yes	2017	less than 06 months
6	2017	Yes	ICASL	Yes	2017	less than 06 months
7	2017	Yes	ICASL	Yes	2017	06 months to 01 year
8	2017	Yes	Certificate course in com	Yes	2014	less than 06 months
9	2016	Yes	ICASL, Certificate course	Yes	2017	less than 06 months
10	2017	Yes	Certificate course in com	Yes	2015	less than 06 months
11	2017	Yes	ICASL	Yes	2017	less than 06 months
12	2017	Yes	Certificate course in com	Yes	2014	less than 06 months
13	2016	Yes	Business Administration	Yes	2016	06 months to 01 year
14	2017	Yes	AAT	Yes	2017	less than 06 months
15	2015	Yes	CMA	Yes	2016	less than 06 months
16	2016	Yes	IBSL	Yes	2016	less than 06 months
17	2017	No		No	2017	less than 06 months

Figure 4.7 Feedbacks for the survey questionnaire

We then received prediction accuracy in terms of the used train and test data as given following two figures.

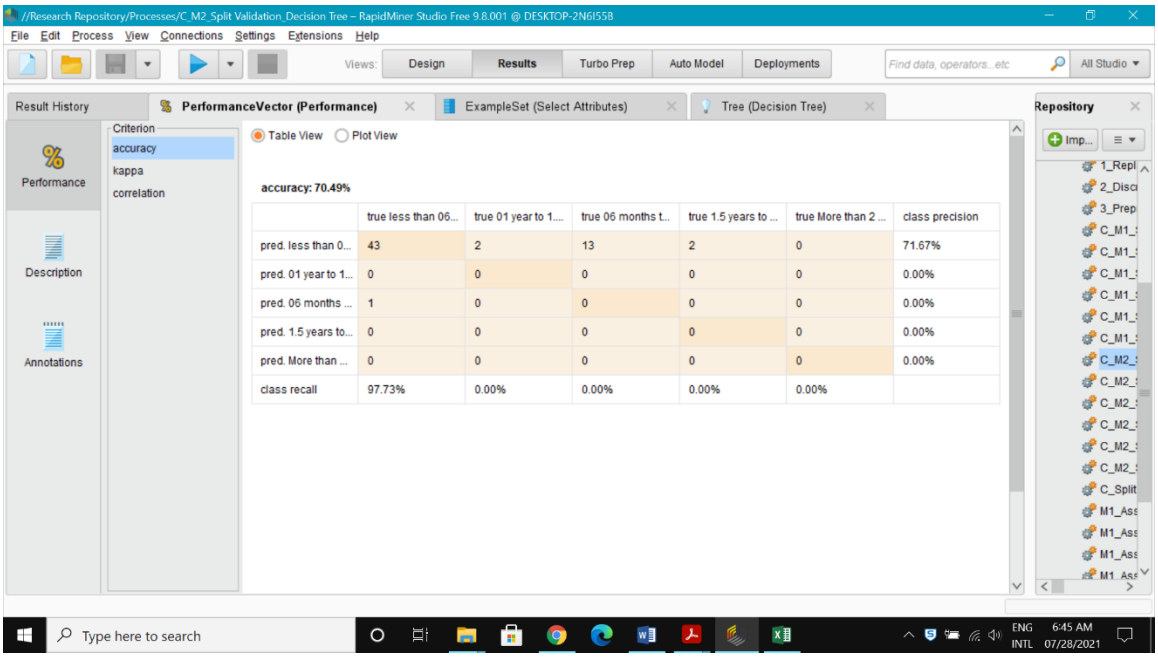


Figure 4.8 Sample classification output

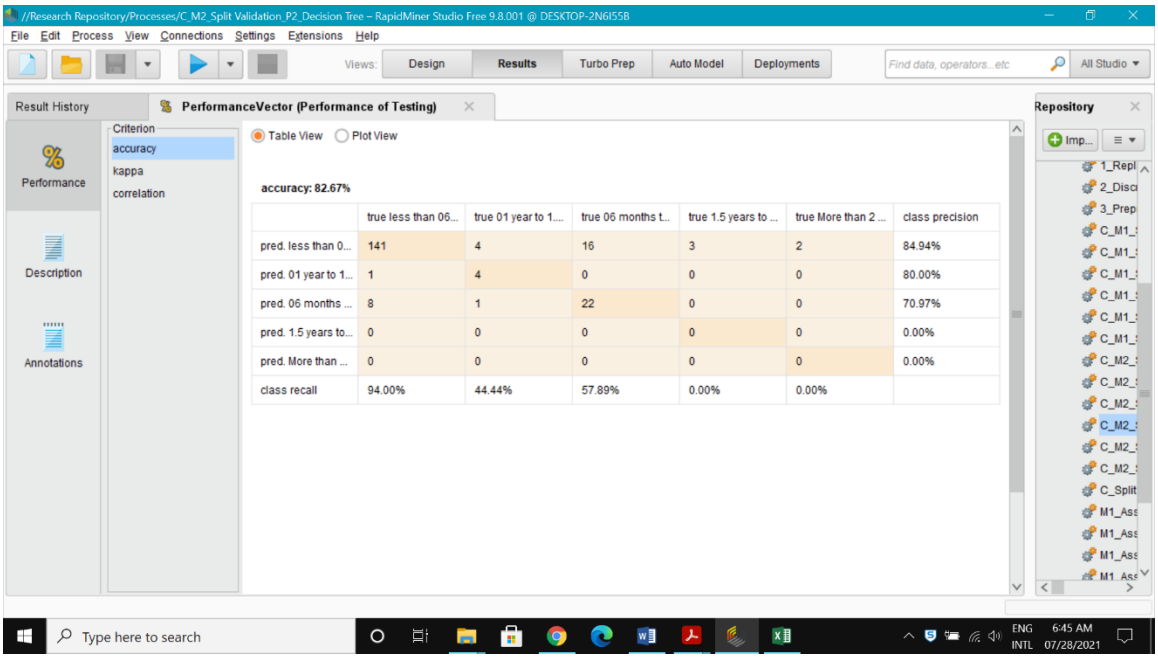


Figure 4.9 Sample classification test set output

4.5 Process and Features

Rapid Miner Studio is a free and open source application and also a sophisticated and efficient data mining software that can handle enormous datasets. Data Collection and Pre-processing of data have been done by this tool. Collected data is used to data mining process to build a classification model. These models were created by using rapid miner software. Then compare the accuracy by applying several classification techniques and evaluate the model. After that, based on evaluated classification model, we can predict the students' employability. Further job specific employability skills were identified by using association mining techniques used in Rapid Miner software.

4.6 Summary

This chapter presented the machine learning approach for conducting this research. It pointed out learning techniques including hypothesis, input, output and the research process. Further for the research problem, the possible approach were also identified. The design of the presented solution has discussed in the next chapter.

Chapter 5

Analysis and Design

5.1 Introduction

In previous chapter the approach followed to address the problem were discussed and the hypothesis, input, output and the process were highlighted. This chapter, analysis and design of the model has discussed.

5.2 Data Collection

Data collection has completed by using a survey questionnaire in the Google Forms. To design the survey questionnaire past literature was referred. All the materials referred to prepare the questionnaire was discussed in the literature review chapter. Then the prepared questionnaire was distributed among student batches from 2014 to 2017 and collected the students' feedbacks to the Google sheet. There are 202 student feedbacks were collected for the employability prediction using data mining techniques.

5.3 Design the classification and Association Tools

According to the literature discussed in the chapter two, classification and association algorithms were selected to analyze the dataset in data mining. Basically, the data set for the employability prediction was collected and that has used for the classification and association. First data set was pre-processed and cleaned for noisy data. Then pre-processed data set was divided in to two categories as training and testing while 70% for the training and 30% for the testing the data.

After analyzing previous literatures three algorithms were identified and those were KNN, Naïve Bayes and decision tree algorithms. So these were used for classification in data mining and then to predict the employability of business finance students. Further these algorithms were applied on training data set separately one by one and created the models

individually. After that these models were tested using testing set. Finally depending on the performance of the testing models the best algorithm was identified and it was used to predict the business finance students' employability.

Association analysis in the data mining also used in this research study to identify the most associated skills for each and every job category. Identifying job specific skills is very important for the students when they search jobs in competitive job market.

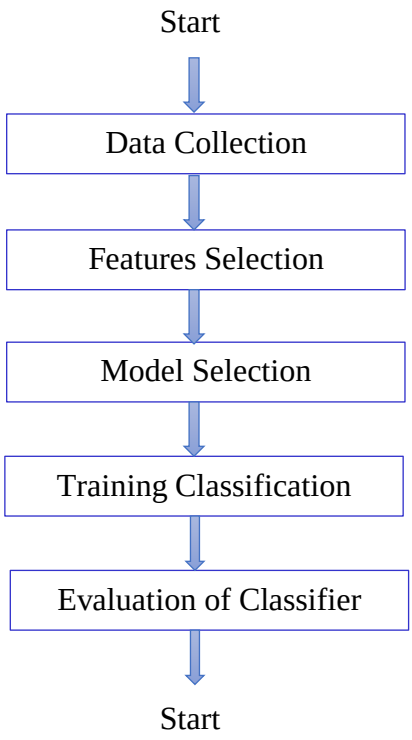


Figure 5.1 Designing of pattern recognition using classification methods

5.4 Summary

This chapter discussed the design of the survey questionnaire then design of classification and association methods also. In next chapter implementation of this solution has been described.

Chapter 6

Implementation

6.1 Introduction

In chapter five, the overall design of the proposed solution was described. This chapter describes the implementation details of each of the sections mentioned in the previous chapter.

6.2 Collection of Data (and Basic Functionalities)

Previous literature was referred and according to that survey questionnaire was prepared with regards to the business finance students' employability prediction. The most important features for the employability were included in the survey questionnaire. In the questionnaire questions were examined by categorizing the questions in to three sections.

Some of the most important features are job category, time duration to get a job after graduation, professional courses followed, level of importance for the selected skills and to what extend these skills were developed within the institute. These are some of the important attributed assessed within the questionnaire. Here fifteen skills were selected to get the feedback for the employability prediction.

This questionnaire consisted of three sub sections such as background factors, work situation and rating of work-related skills and competences. Finally the prepared questionnaire was distributed among student batches from 2014 to 2017 and collected the students' feedbacks to the Google sheet. There are 202 student feedbacks were recorded in the Google sheet for the employability prediction using data mining techniques.

6.3 Data Pre-Processing

In the data mining process data cleaning and preparing is an important step. In order to be used in various analytical methods, basic data must be cleaned and pre-processed. There are occasions when data contains errors or outliers.

In this research when distributing an online survey questionnaire most of the important questions marked as required. And when there are optional questions that did not mentioned required those questions were having missing values. So students have completed the questionnaire with less missing values. In this stage all the missing values were cleaned and prepared the data set for the classification.

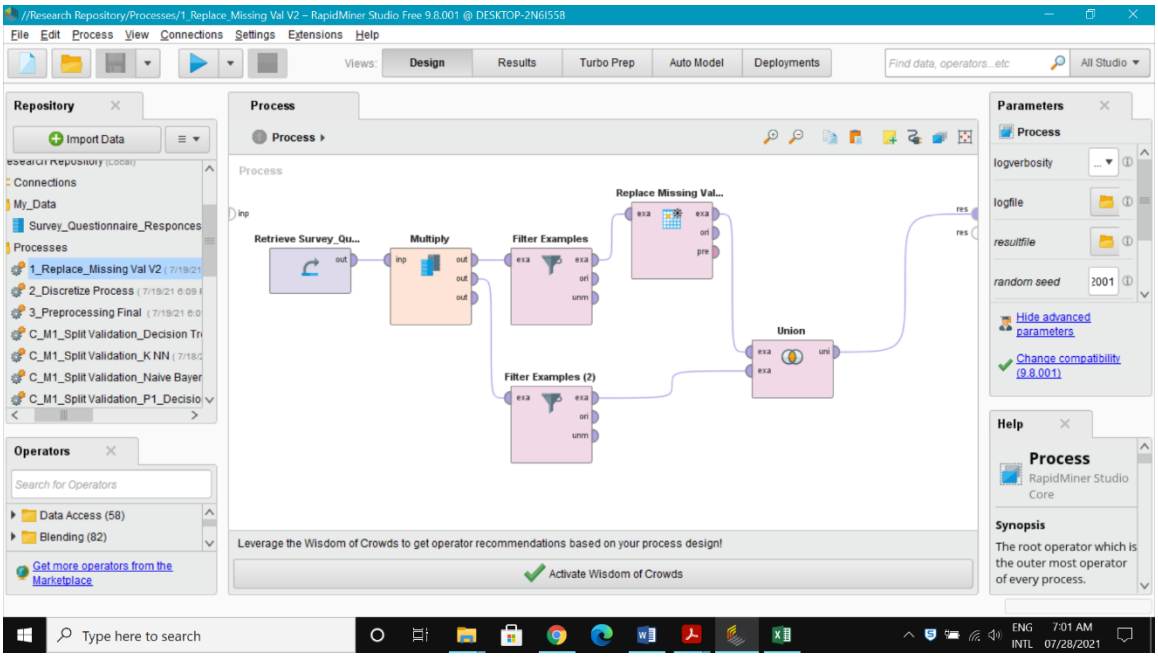


Figure 6.1 Replace missing values

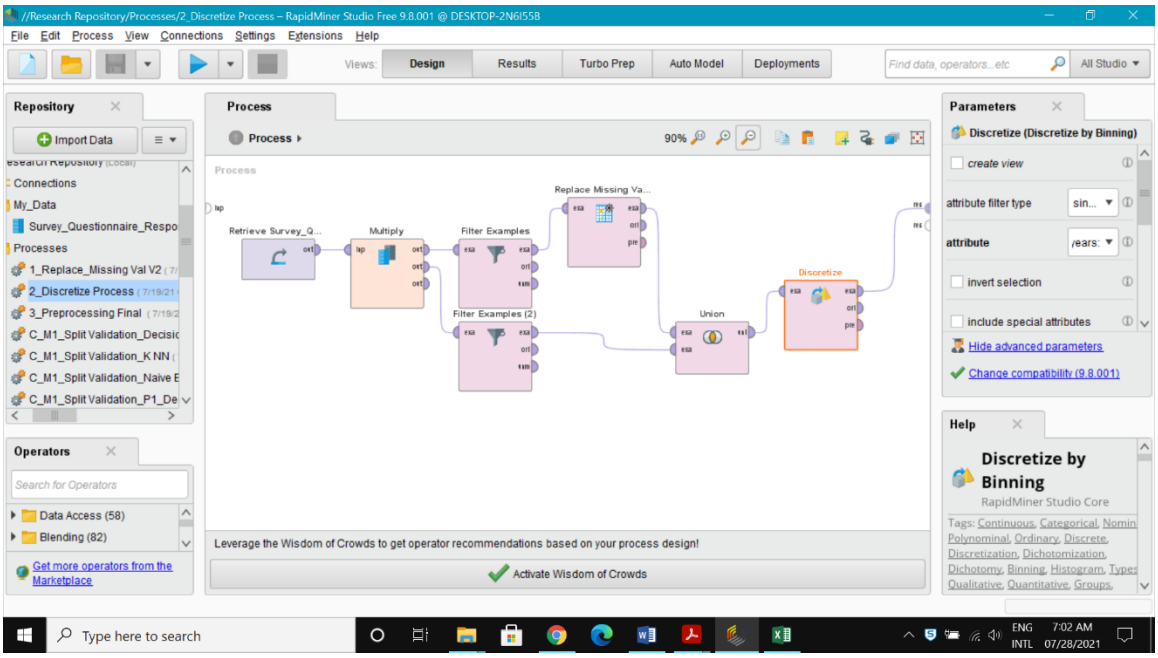


Figure 6 2 Discretization process

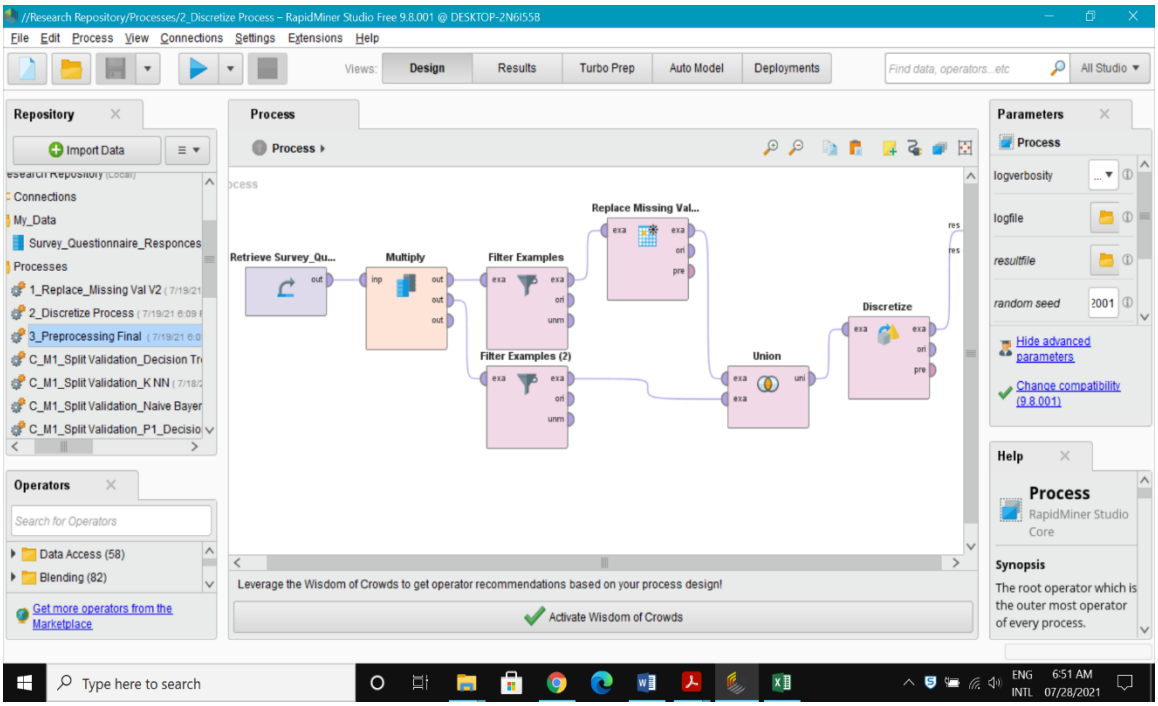


Figure 6.3 End of data preprocessing final processes

6.4 Classification Models

Rapid Miner Studio is the software used for Data mining. Naive Bayes, K-NN and Decision Tree were the classification algorithms used in this research study. Then model accuracy was identified individually. The data set collected using the Google Form was split in to two categories such as training 70% and testing set 30%. Using training dataset classifications models were created and then using the testing dataset, already created models were tested and identified the best model. By using the model, relevant and optimized features have been selected and used for the prediction of the student employability.

In this research study there were two aspects/ modules were considered for the classification techniques. These were the main prediction modules analyzed. Those predictions implemented under classifications methods in this study were given below.

- Predict the job category according to the skills use and their level of importance.
- Determine time period to get a job after graduation with relevant to the skills available

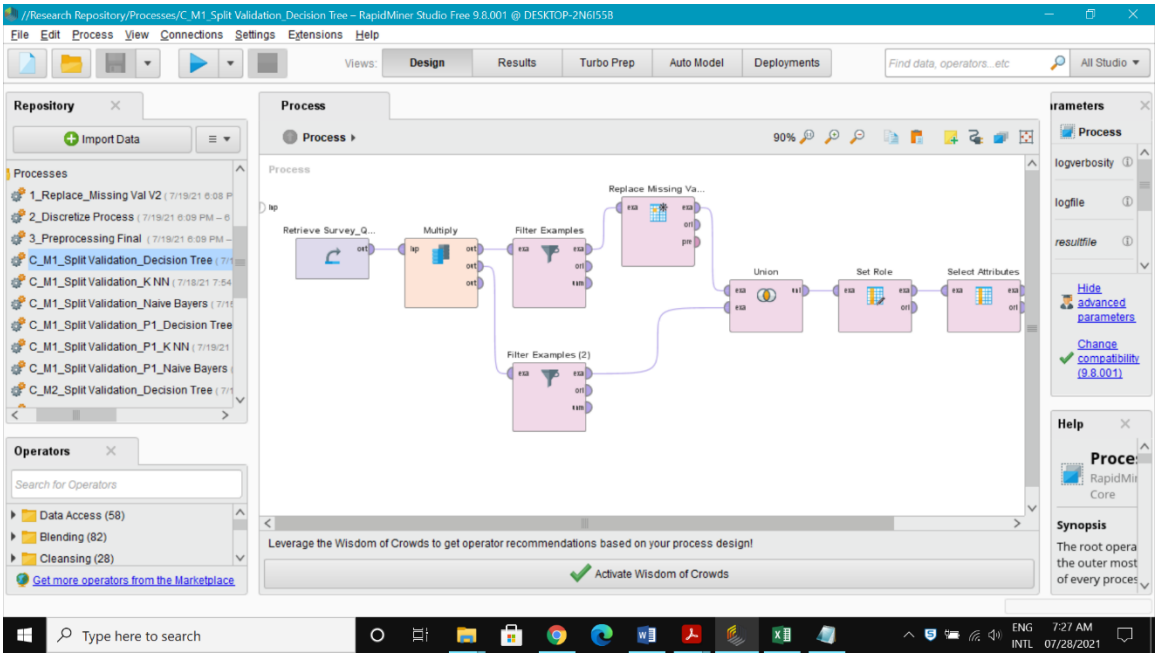


Figure 6.4 Decision tree model using the Rapid Miner

In Rapid Miner when creating the classification model split validation has been used. Using split validation relevant classification models were created by allocating the 70% from the dataset as in the figure 6.5 and 6.6.

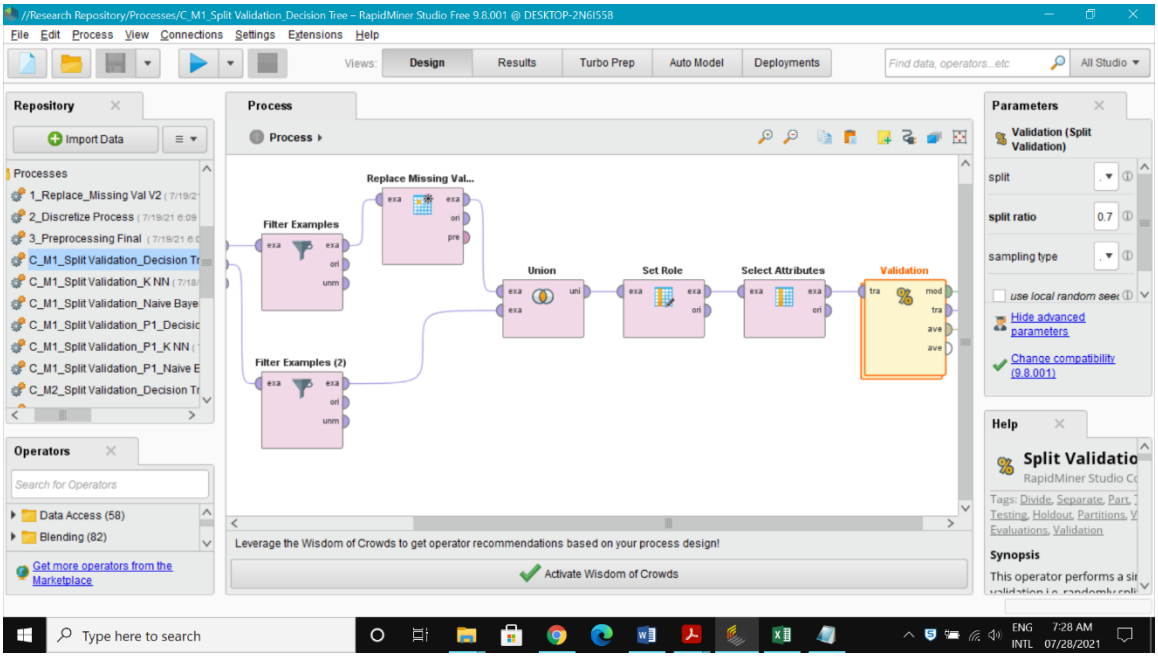


Figure 6.5 Decision tree model with split validation using the Rapid Miner

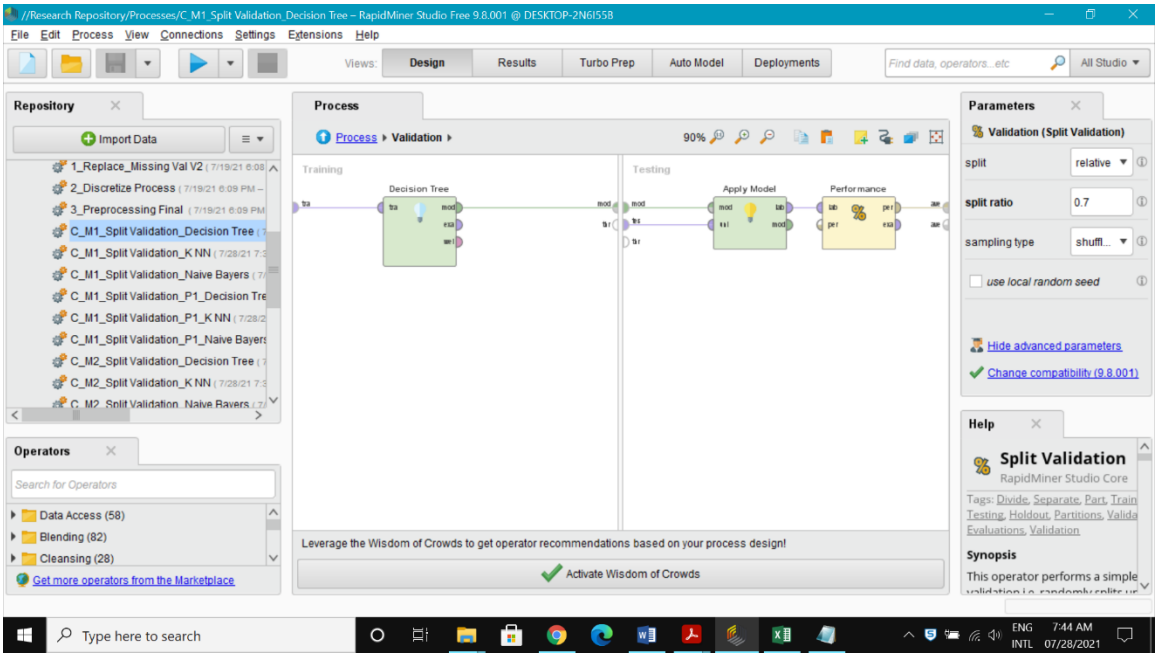


Figure 6.6 Decision tree model with split validation process

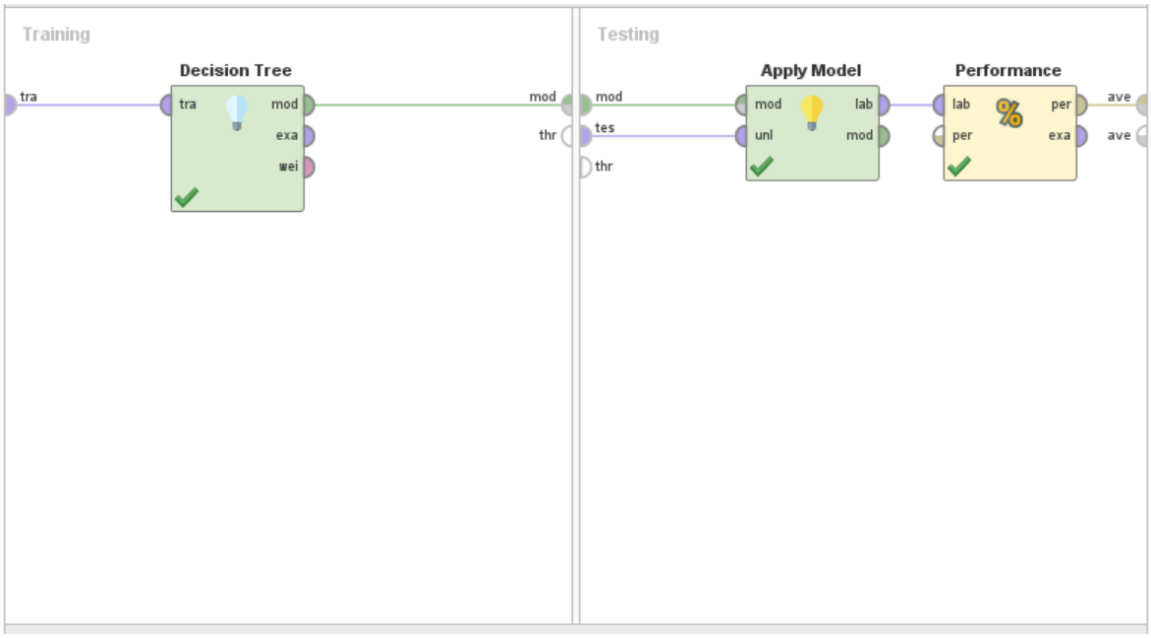


Figure 6.7 Decision tree model with split validation- training the model

6.5 Association Analysis

The association techniques were also used to find the relation between job categories and the types of skills need to have when students search a placement after gratuation. For and example if a student need to be a accountant after the gratuation then he or she need to develop some particular skills before get the required career.

Therefore in the association analysis mainly the following criteria was implemented using the Rapid Miner.

- Identify the most associated skills with respect to the job title/ position.

So this kind of association were implemented in this research study as in figure 6.8.

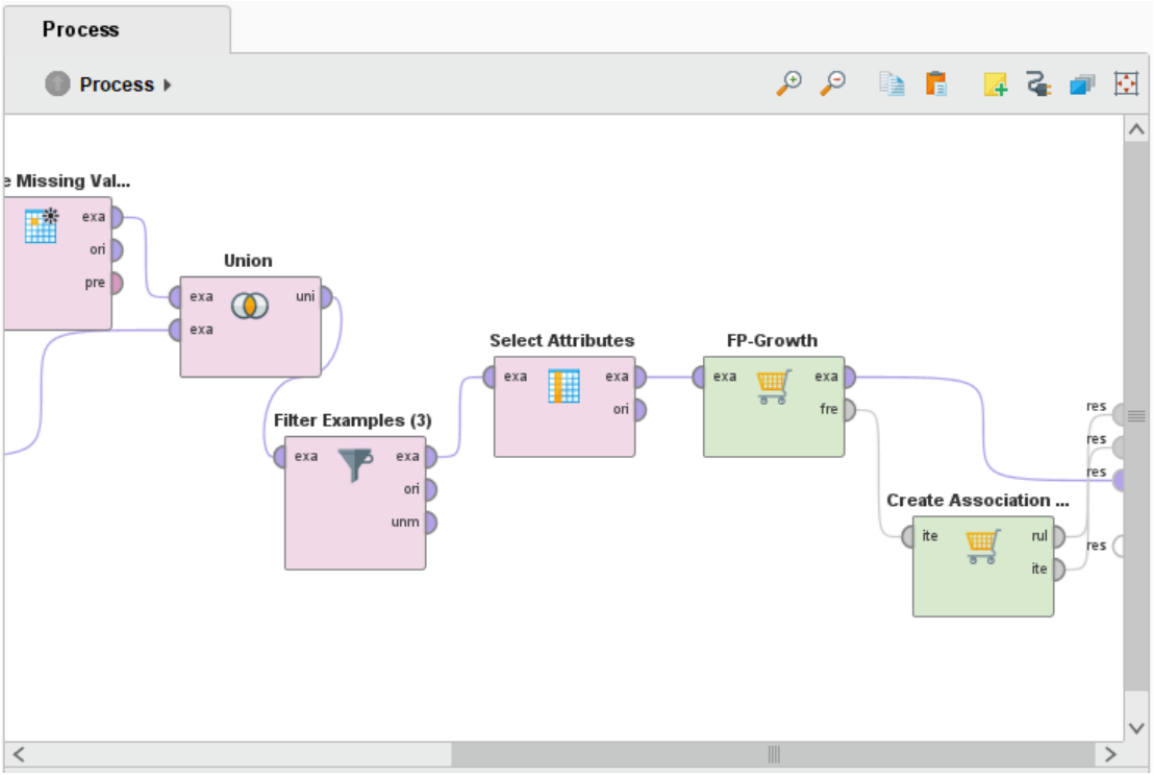


Figure 6.8 Identifying associated skills for each job category

6.6 Summary

This chapter provides an overall implementation information of each method of the proposed solution. Further it has mentioned about Rapid Miner software and data mining techniques such as classification model's development and association rules to find relation with attributes for employability prediction and these were described with align to design. Next chapter, all the modules implemented in the solution has been evaluated.

Chapter 7

Evaluation

7.1 Introduction

In chapter six discussed the details of implementation of all the techniques mentioned in the planned solution. This chapter validates and evaluates the overall solution, data mining methods and data models which were used in classification as well as in the association analysis.

7.2 Evaluation of Classification Techniques

The data set was collected by using the survey questionnaire in Google Form and it was trained using different classification techniques namely Decision Tree, Naïve Bayes and k-NN by using the data mining software names as Rapid Miner tool. The main objective of this classification was to create and recognize the best model with the highest accuracy for the prediction of student employability. Accuracy of the prediction was mainly depended on the model. Accuracy refers to a model's ability to accurately predict the target class of existing and newly unseen data. If the classifier's accuracy is deemed satisfactory, it is applied to categorize upcoming data records in that the class label is unknown. Confusion matrix was used for evaluating a classifier's accuracy, recall and precision.

- Accuracy – this measure the percentage of all decisions that were correct.
- Recall - Portion of the completeness of correct categories that were assigned or classified.
- Precision - Fraction of correctness out of all detected as correct.

In this evaluation stage the aspects/ modules implemented were evaluated with their accuracy and selected the best accuracy model for the future prediction of known data. The main classification modules implemented and evaluated in this research study were given below.

1. Predict the job category according to the skills use and their level of importance.
2. Determine time period to get a job after graduation with relevant to the skills available.

For this evaluation, the below-mentioned classifiers were used. For classification, split-validation techniques were used with 70% for the model training and 30% for the testing dataset was allocated.

Following are the classification models evaluated according to the first module in the classification which is the predicting the job category according to the skills use and their level of importance. KNN, Naïve Bayes and Decision tree algorithms were tested to identify the best model for this prediction.

accuracy: 67.82%

	true Mana...	true Whol...	true Banki...	true Accou...	true Marke...	true Auditing	true Custo...	true Public...	class prec...
pred. Man...	1	0	1	0	0	0	0	0	50.00%
pred. Who...	0	0	0	0	0	0	0	0	0.00%
pred. Ban...	0	0	17	6	3	4	0	0	56.67%
pred. Acco...	5	2	6	96	4	23	1	2	69.06%
pred. Mark...	0	0	0	0	1	0	0	0	100.00%
pred. Audit...	1	0	3	4	0	22	0	0	73.33%
pred. Cust...	0	0	0	0	0	0	0	0	0.00%
pred. Publ...	0	0	0	0	0	0	0	0	0.00%
class recall	14.29%	0.00%	62.96%	90.57%	12.50%	44.90%	0.00%	0.00%	

Figure 7.1 Accuracy of predicting the job category- using KNN algorithm

accuracy: 65.35%

	true Mana...	true Whol...	true Banki...	true Accou...	true Marke...	true Auditing	true Custo...	true Public...	class prec...
pred. Man...	4	0	2	7	0	1	0	0	28.57%
pred. Who...	0	2	0	0	0	1	0	0	66.67%
pred. Ban...	1	0	12	6	1	1	0	0	57.14%
pred. Acco...	2	0	11	78	1	15	0	0	72.90%
pred. Mark...	0	0	2	7	6	3	0	0	33.33%
pred. Audit...	0	0	0	6	0	27	0	0	81.82%
pred. Cust...	0	0	0	0	0	0	1	0	100.00%
pred. Publ...	0	0	0	2	0	1	0	2	40.00%
class recall	57.14%	100.00%	44.44%	73.58%	75.00%	55.10%	100.00%	100.00%	

Figure 7.2 Accuracy of predicting the job category- using Naïve Bayes algorithm

accuracy: 70.79%

	true Mana...	true Whol...	true Banki...	true Accou...	true Marke...	true Auditing	true Custo...	true Public...	class prec...
pred. Man...	5	0	2	2	0	0	0	0	55.56%
pred. Who...	0	0	0	0	0	0	0	0	0.00%
pred. Ban...	0	1	13	4	0	4	0	1	56.52%
pred. Acco...	2	1	10	95	3	19	1	1	71.97%
pred. Mark...	0	0	0	0	5	1	0	0	83.33%
pred. Audit...	0	0	2	5	0	25	0	0	78.12%
pred. Cust...	0	0	0	0	0	0	0	0	0.00%
pred. Publ...	0	0	0	0	0	0	0	0	0.00%
class recall	71.43%	0.00%	48.15%	89.62%	62.50%	51.02%	0.00%	0.00%	

Figure 7.3 Accuracy of predicting the job category- using Decision Tree algorithm

The following summary table is relevant to the classification module 01 which is regarding to the prediction the job category according to the skills use and their level of importance.

Classification Model	Accuracy
KNN	67.82%
Naïve Bayes	65.35%
Decision Tree	70.79%

Table 7.1 Accuracy of prediction the job category according to the skills

According to the algorithms evaluated for the prediction of the job category according to the skills use and their level of importance, the best algorithm identified with the highest accuracy was the Decision Tree.

Following are the classification models evaluated according to the second module in the classification which is the determining of the time period to get a job after graduation with relevant to the skills available. KNN, Naïve Bayes and Decision tree algorithms were tested to identify the best model for this prediction.

accuracy: 78.71%

	true less than 06...	true 01 year to 1....	true 06 months t...	true 1.5 years to ...	true More than 2 ...	class precision
pred. less than 0...	142	6	23	3	2	80.68%
pred. 01 year to 1...	0	3	1	0	0	75.00%
pred. 06 months ...	8	0	14	0	0	63.64%
pred. 1.5 years to...	0	0	0	0	0	0.00%
pred. More than ...	0	0	0	0	0	0.00%
class recall	94.67%	33.33%	36.84%	0.00%	0.00%	

Figure 7.4 Accuracy of determining the time period to get a job – using KNN algorithm

accuracy: 67.33%

	true less than 06...	true 01 year to 1....	true 06 months t...	true 1.5 years to ...	true More than 2 ...	class precision
pred. less than 0...	106	1	14	1	0	86.89%
pred. 01 year to 1...	5	7	0	0	0	58.33%
pred. 06 months ...	17	0	19	0	0	52.78%
pred. 1.5 years to...	7	0	2	2	0	18.18%
pred. More than ...	15	1	3	0	2	9.52%
class recall	70.67%	77.78%	50.00%	66.67%	100.00%	

Figure 7.5 Accuracy of determining the time period to get a job – using Naïve Bayes algorithm

accuracy: 82.67%

	true less than 06...	true 01 year to 1....	true 06 months t..	true 1.5 years to ...	true More than 2 ...	class precision
pred. less than 0...	141	4	16	3	2	84.94%
pred. 01 year to 1...	1	4	0	0	0	80.00%
pred. 06 months ...	8	1	22	0	0	70.97%
pred. 1.5 years to...	0	0	0	0	0	0.00%
pred. More than ...	0	0	0	0	0	0.00%
class recall	94.00%	44.44%	57.89%	0.00%	0.00%	

Figure 7.6 Accuracy of determining the time period to get a job – using Decision Tree algorithm

The following summary table is relevant to the classification module 02 which is regarding to the determining the time period to get a job after graduation with relevant to the skills available.

Classification Model	Accuracy
KNN	78.71%
Naïve Bayes	67.33%
Decision Tree	82.67%

Table 7.2 Accuracy of determining the time period to get a job

According to the algorithms evaluated for the time period to get a job after graduation with relevant to the skills available, the best algorithm identified with the highest accuracy was the Decision Tree.

So, that the Decision Tree has produced the best results in predicting the student employability for this dataset. Other classification techniques as shown in the above tables, did not perform significantly well in this research.

7.3 Evaluation of Association Techniques

The data set collected using the survey questionnaire was implemented using association rules and conducted the evaluation. In this evaluation the most associated skills with respect to the job title/ position were identified. According to the survey questionnaire there were eight (08) job categories identified and data were collected from the past out student from the year 2014 to 2017.

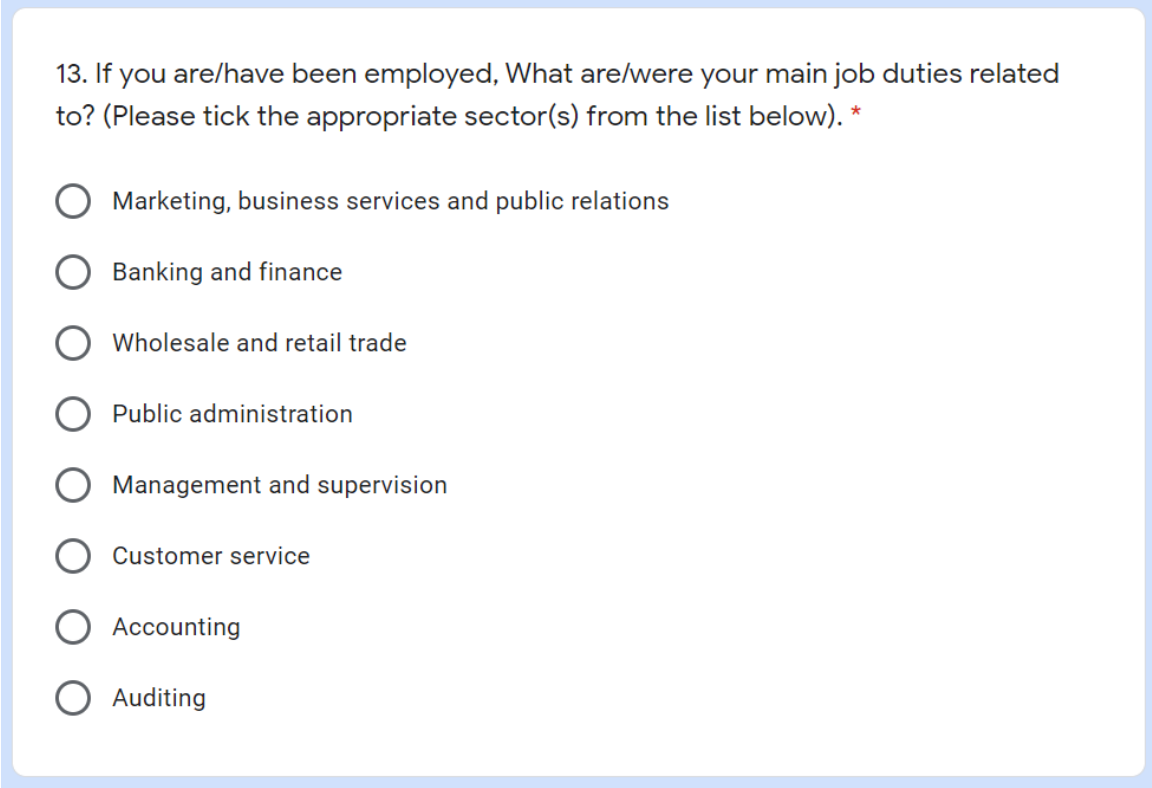


Figure 7.7 Job categories mentioned in the survey questionnaire

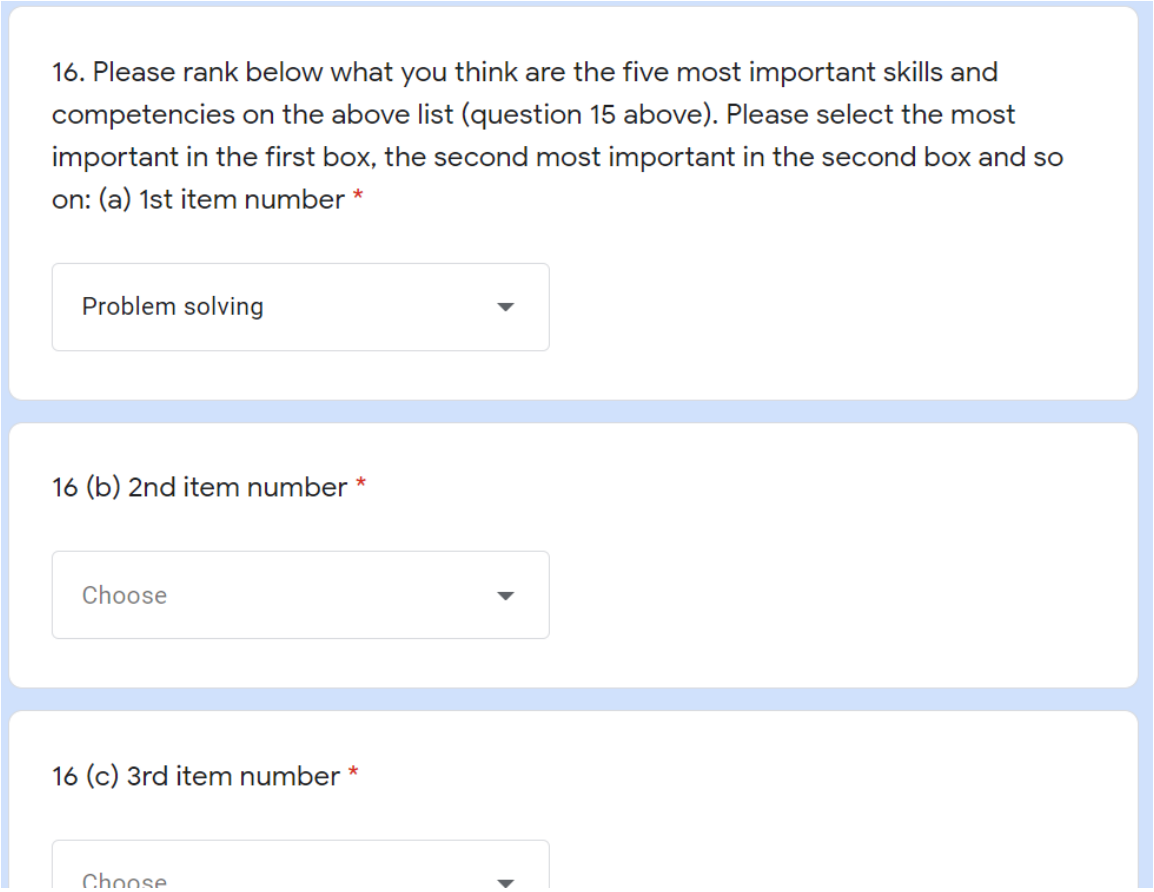


Figure 7.8 Identify the five most important skills

For the association analysis the above-mentioned job categories and the most important job skills were used and evaluated. In the following figures shows that the application of the association mining to the dataset with the use of the Rapid Miner software.

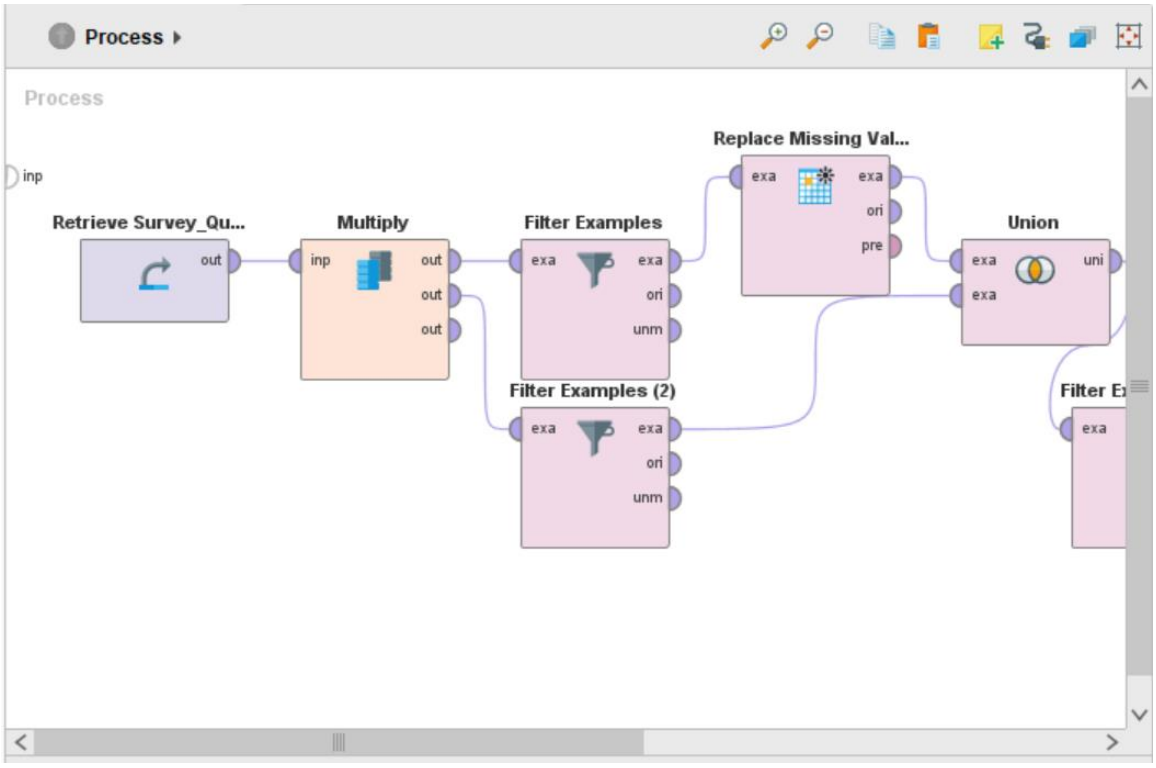


Figure 7.9 Process diagram to identify most associated-skills

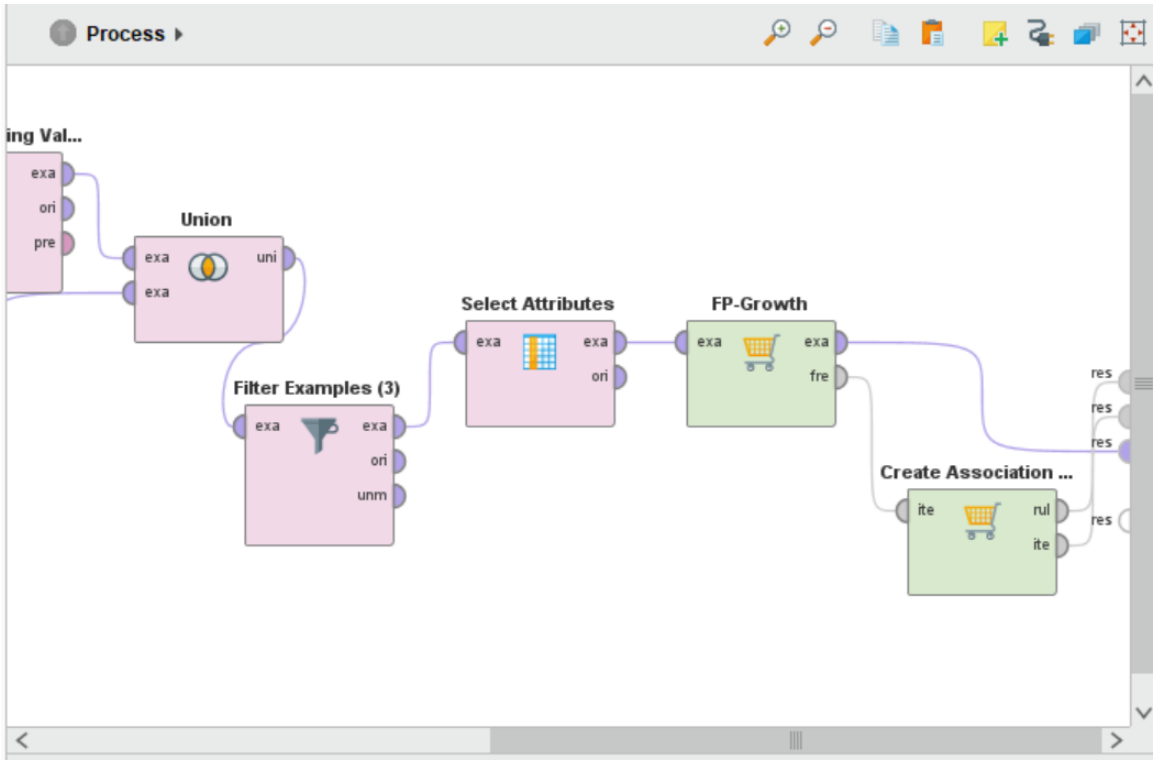


Figure 7.10 Process diagram to identify associated-skills cont..

After applying the association algorithm with respect to each job category the most associated skills could be identified. The higher the support value of the skill that means which has the highest association for the job category. The most associated one skill or two skills or three skills or four skills could be identified according to the support value.

1. Following are the associated skills with respect to the job category, accountancy.

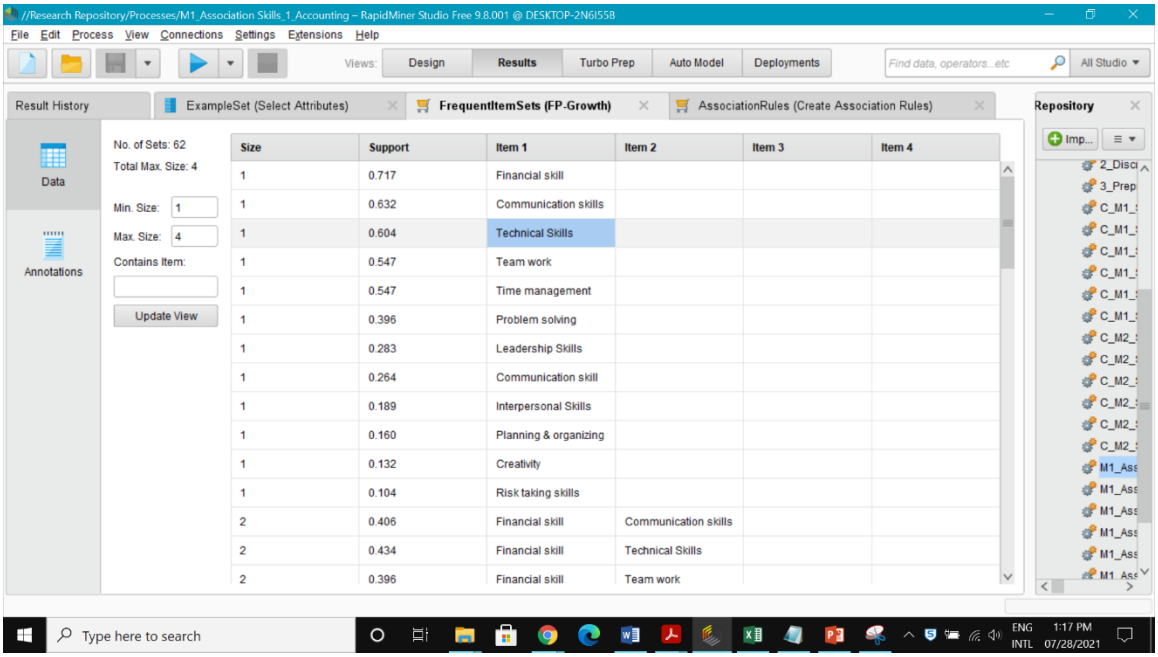


Figure 7.11 Sample frequency set for accountancy job category in the Rapid Miner

Following figure 7.12 is related to the most associated one skill type with respect to the accountancy job category. This the skill level with highest support value.

Size	Support ↓	Item 1
1	0.717	Financial skill
1	0.632	Communication skills
1	0.604	Technical Skills
1	0.547	Team work
1	0.547	Time management

Figure 7.12 Accountancy job category with one associated skill type

Size	Support ↓	Item 1	Item 2
2	0.434	Financial skill	Technical Skills
2	0.434	Financial skill	Time management
2	0.406	Financial skill	Communication ski...
1	0.396	Problem solving	
2	0.396	Financial skill	Team work
2	0.377	Communication ski...	Technical Skills
2	0.349	Communication ski...	Team work
2	0.340	Technical Skills	Time management
2	0.292	Communication ski...	Time management

Figure 7 13 Accountancy job category with two associated skill types

Size	Support ↓	Item 1	Item 2	Item 3
3	0.274	Financial skill	Technical Skills	Time management
1	0.264	Communication skill		
2	0.264	Team work	Time management	
3	0.255	Financial skill	Communication ski...	Technical Skills
2	0.226	Financial skill	Problem solving	
3	0.226	Financial skill	Technical Skills	Team work
2	0.217	Financial skill	Communication skill	
3	0.217	Financial skill	Communication ski...	Team work
3	0.208	Financial skill	Team work	Time management

Figure 7.14 Accountancy job category with three associated skill types

Size	Support ↑	Item 1	Item 2	Item 3	Item 4
4	0.132	Financial skill	Communication ski...	Technical Skills	Team work
4	0.142	Financial skill	Communication ski...	Technical Skills	Time management

Figure 7.15 Accountancy job category with four associated skill types

2. Following are the associated skills with respect to the job category, auditing.

Size	Support ↓	Item 1
1	0.653	Communication ski...
1	0.633	Time management
1	0.571	Financial skill
1	0.551	Team work
1	0.510	Problem solving
1	0.510	Technical Skills

Figure 7.16 Auditing job category with one associated skill type

Size	Support	Item 1	Item 2
2	0.408	Communication ski...	Time management
2	0.408	Communication ski...	Financial skill
2	0.327	Communication ski...	Team work
2	0.347	Communication ski...	Problem solving
2	0.367	Communication ski...	Technical Skills
2	0.265	Communication ski...	Planning & organizi...
2	0.122	Communication ski...	Leadership Skills
2	0.327	Time management	Financial skill

Figure 7.17 Auditing job category with two associated skill types

Size	Support	Item 1	Item 2	Item 3
3	0.204	Communication ski...	Time management	Financial skill
3	0.224	Communication ski...	Time management	Team work
3	0.163	Communication ski...	Time management	Problem solving
3	0.184	Communication ski...	Time management	Technical Skills
3	0.143	Communication ski...	Time management	Planning & organizi...
3	0.184	Communication ski...	Financial skill	Team work
3	0.245	Communication ski...	Financial skill	Problem solving

Figure 7.18 Auditing job category with three associated skill types

4	0.163	Communication skills	Financial skill	Problem solving	Technical Skills

Figure 7.19 The most associated four skill types for auditing job category

3. Following are the associated skills with respect to the job category, banking and finance.

Size	Support ↓	Item 1
1	0.704	Financial skill
1	0.667	Time managem...
1	0.630	Communication...
1	0.556	Team work

Figure 7.20 Banking and finance job category with one associated skill type

Size	Support ↓	Item 1	Item 2
2	0.333	Financial skill	Interpersonal S...
2	0.333	Time managem...	Problem solving
2	0.296	Financial skill	Team work
2	0.296	Time managem...	Communication...
2	0.296	Time managem...	Technical Skills
2	0.296	Time managem...	Interpersonal S...
2	0.296	Communication...	Leadership Skills
2	0.296	Team work	Technical Skills

Figure 7.21 Banking and finance job category with two associated skill types

Size	Support ↓	Item 1	Item 2	Item 3
3	0.222	Financial skill	Time managem...	Technical Skills
3	0.222	Financial skill	Time managem...	Problem solving
3	0.222	Communication...	Technical Skills	Interpersonal S...
3	0.185	Financial skill	Time managem...	Communication...
3	0.185	Financial skill	Communication...	Team work
3	0.185	Time managem...	Communication...	Technical Skills

Figure 7.22 Banking and finance job category with three associated skill types

Size	Support ↓	Item 1	Item 2	Item 3	Item 4
3	0.111	Team work	Technical Skills	Interpersonal S...	
3	0.111	Team work	Interpersonal S...	Leadership Skills	
3	0.111	Technical Skills	Interpersonal S...	Leadership Skills	
4	0.111	Financial skill	Time managem...	Communication...	Technical Skills
4	0.111	Financial skill	Time managem...	Communication...	Interpersonal S...
4	0.111	Financial skill	Time managem...	Technical Skills	Interpersonal S...
4	0.111	Financial skill	Time managem...	Interpersonal S...	Problem solving
4	0.111	Financial skill	Communication...	Technical Skills	Interpersonal S...
4	0.111	Time managem...	Communication...	Team work	Technical Skills
4	0.111	Communication...	Team work	Technical Skills	Interpersonal S...
4	0.111	Communication...	Team work	Interpersonal S...	Leadership Skills

Figure 7.23 Banking and finance job category with four associated skill types

4. Following are the associated skills with respect to the job category, customer service.

Size	Support	Item 1
1	1.000	Adaptability
1	1.000	Communication...
1	1.000	Critical thinking
1	1.000	Problem solving
1	1.000	Technical Skills

Figure 7.24 Customer service job category with one associated skill type

Size	Support ↑	Item 1	Item 2
2	1.000	Adaptability	Communication...
2	1.000	Adaptability	Critical thinking
2	1.000	Adaptability	Problem solving
2	1.000	Adaptability	Technical Skills
2	1.000	Communication...	Critical thinking
2	1.000	Communication...	Problem solving
2	1.000	Communication...	Technical Skills
2	1.000	Critical thinking	Problem solving
2	1.000	Critical thinking	Technical Skills

Figure 7.25 Customer service job category with two associated skill types

Size	Support ↑	Item 1	Item 2	Item 3
3	1.000	Adaptability	Communication...	Critical thinking
3	1.000	Adaptability	Communication...	Problem solving
3	1.000	Adaptability	Communication...	Technical Skills
3	1.000	Adaptability	Critical thinking	Problem solving
3	1.000	Adaptability	Critical thinking	Technical Skills
3	1.000	Adaptability	Problem solving	Technical Skills
3	1.000	Communication...	Critical thinking	Problem solving
3	1.000	Communication...	Critical thinking	Technical Skills
3	1.000	Communication...	Problem solving	Technical Skills

Figure 7.26 Customer service job category with three associated skill types

4	1.000	Adaptability	Communication...	Critical thinking	Problem solving
4	1.000	Adaptability	Communication...	Critical thinking	Technical Skills
4	1.000	Adaptability	Communication...	Problem solving	Technical Skills
4	1.000	Adaptability	Critical thinking	Problem solving	Technical Skills
4	1.000	Communication...	Critical thinking	Problem solving	Technical Skills
5	1.000	Adaptability	Communication...	Critical thinking	Problem solving

Figure 7.27 Customer service job category with four associated skill types

5. Following are the associated skills with respect to the job category, management and supervision.

Size	Support ↓	Item 1
1	0.857	Team work
1	0.714	Communication...
1	0.571	Problem solving

Figure 7.28 Management and supervision job category with one associated skill type

Size	Support ↓	Item 1	Item 2
2	0.286	Team work	Technical Skills
2	0.286	Team work	Planning & orga...
2	0.286	Team work	Time managem...
2	0.286	Communication...	Creativity
2	0.286	Problem solving	Technical Skills
2	0.286	Creativity	Interpersonal S...

Figure 7.29 Management and supervision job category with two associated skill types

Size	Support ↓	Item 1	Item 2	Item 3
3	0.143	Team work	Communication...	Interpersonal S...
3	0.143	Team work	Communication...	Planning & orga...
3	0.143	Team work	Communication...	Time managem...
3	0.143	Team work	Communication...	Adaptability
3	0.143	Team work	Communication...	Risk taking skills
3	0.143	Team work	Problem solving	Creativity
3	0.143	Team work	Problem solving	Interpersonal S...
3	0.143	Team work	Problem solving	Technical Skills
3	0.143	Team work	Problem solving	Planning & orga...

Figure 7.30 Management and supervision job category with three associated skill types

Size	Support ↓	Item 1	Item 2	Item 3	Item 4
4	0.143	Team work	Communication...	Problem solving	Creativity
4	0.143	Team work	Communication...	Problem solving	Interpersonal S...
4	0.143	Team work	Communication...	Problem solving	Technical Skills
4	0.143	Team work	Communication...	Problem solving	Planning & orga...
4	0.143	Team work	Communication...	Problem solving	Time managem...
4	0.143	Team work	Communication...	Problem solving	Risk taking skills
4	0.143	Team work	Communication...	Creativity	Interpersonal S...
4	0.143	Team work	Communication...	Creativity	Technical Skills
4	0.143	Team work	Communication...	Creativity	Adaptability

Figure 7.31 Management and supervision job category with four associated skill types

6. Following are the associated skills with respect to the job category, marketing business services and public relations.

Size	Support ↓	Item 1
1	0.750	Creativity
1	0.625	Communication...
1	0.625	Team work

Figure 7.32 Marketing business services and public relations job category with one associated skill type

Size	Support ↓	Item 1	Item 2
2	0.375	Creativity	Interpersonal S...
2	0.375	Communication...	Interpersonal S...
2	0.375	Team work	Interpersonal S...

Figure 7.33 Marketing business services and public relations job category with two associated skill types

Size	Support ↓	Item 1	Item 2	Item 3
3	0.250	Creativity	Communication...	Time managem...
3	0.250	Creativity	Communication...	Financial skill
3	0.250	Creativity	Communication...	create visual ai...
3	0.250	Creativity	Team work	Interpersonal S...
3	0.250	Creativity	Team work	Financial skill
3	0.250	Creativity	Interpersonal S...	create visual ai...
3	0.250	Communication...	Team work	Interpersonal S...
3	0.250	Communication...	Team work	Financial skill
3	0.250	Communication...	Interpersonal S...	create visual ai...

Figure 7.34 Marketing business services and public relations job category with three associated skill types

Size	Support ↓	Item 1	Item 2	Item 3	Item 4
3	0.125	Leadership Skills	Critical thinking	Risk taking skills	
4	0.125	Creativity	Communication...	Team work	Time managem...
4	0.125	Creativity	Communication...	Team work	Planning & orga...
4	0.125	Creativity	Communication...	Team work	create visual aid...
4	0.125	Creativity	Communication...	Team work	Technical Skills
4	0.125	Creativity	Communication...	Interpersonal S...	Time managem...
4	0.125	Creativity	Communication...	Interpersonal S...	Planning & orga...
4	0.125	Creativity	Communication...	Time managem...	Financial skill
4	0.125	Creativity	Communication...	Time managem...	create visual aid...
4	0.125	Creativity	Communication...	Financial skill	Technical Skills

Figure 7.35 Marketing business services and public relations job category with four associated skill types

5	0.125	Creativity	Communication...	Team work	Interpersonal S...	Planning & orga...
5	0.125	Creativity	Communication...	Team work	Interpersonal S...	create visual ai...
5	0.125	Creativity	Communication...	Team work	Time managem...	Financial skill
5	0.125	Creativity	Communication...	Team work	Financial skill	Technical Skills
5	0.125	Creativity	Communication...	Interpersonal S...	Time managem...	create visual ai...
5	0.125	Creativity	Problem solving	Leadership Skills	Critical thinking	Risk taking skills
5	0.125	Team work	Interpersonal S...	Problem solving	Communication...	Planning & orga...
5	0.125	Problem solving	Time managem...	Communication...	Leadership Skills	Analytical Skills

Figure 7.36 Marketing business services and public relations job category with five associated skill types

7. Following are the associated skills with respect to the job category, administration.

1	0.500	Communication...
1	0.500	Communication...
1	0.500	Creativity
1	0.500	Financial skill
1	0.500	Interpersonal S...
1	0.500	Problem solving

Figure 7.37 Administration job category with one associated skill type

Size	Support ↓	Item 1	Item 2
2	0.500	Leadership Skills	Communication...
2	0.500	Leadership Skills	Communication...
2	0.500	Leadership Skills	Creativity
2	0.500	Leadership Skills	Financial skill
2	0.500	Leadership Skills	Interpersonal S...
2	0.500	Leadership Skills	Problem solving
2	0.500	Team work	Communication...
2	0.500	Team work	Communication...
2	0.500	Team work	Creativity

Figure 7.38 Administration job category with two associated skill types

Size	Support ↓	Item 1	Item 2	Item 3
3	0.500	Leadership Skills	Team work	Communication...
3	0.500	Leadership Skills	Team work	Communication...
3	0.500	Leadership Skills	Team work	Creativity
3	0.500	Leadership Skills	Team work	Financial skill
3	0.500	Leadership Skills	Team work	Interpersonal S...
3	0.500	Leadership Skills	Team work	Problem solving
3	0.500	Leadership Skills	Communication...	Creativity
3	0.500	Leadership Skills	Communication...	Problem solving

Figure 7.39 Administration job category with three associated skill types

Size	Support	Item 1	Item 2	Item 3	Item 4
3	0.500	Communication...	Creativity	Problem solving	
3	0.500	Communication...	Financial skill	Interpersonal S...	
4	0.500	Leadership Skills	Team work	Communication...	Creativity
4	0.500	Leadership Skills	Team work	Communication...	Problem solving
4	0.500	Leadership Skills	Team work	Communication...	Financial skill
4	0.500	Leadership Skills	Team work	Communication...	Interpersonal S...
4	0.500	Leadership Skills	Team work	Creativity	Problem solving
4	0.500	Leadership Skills	Team work	Financial skill	Interpersonal S...
4	0.500	Leadership Skills	Communication...	Creativity	Problem solving
4	0.500	Leadership Skills	Communication...	Financial skill	Interpersonal S...

Figure 7.40 Administration job category with four associated skill types

5	0.500	Leadership Skills	Team work	Communication...	Creativity	Problem solving
5	0.500	Leadership Skills	Team work	Communication...	Financial skill	Interpersonal S...

Figure 7.41 Administration job category with five associated skill types

8. Following are the associated skills with respect to the job category, whole sale and retail trade.

1	0.500	Interpersonal S...	
1	0.500	Planning & orga...	
1	0.500	Planning and or...	
1	0.500	Problem solving	
1	0.500	Team work	
1	0.500	Technical Skills	

Figure 7.42 Whole sale and retail trade job category with one associated skill type

Size	Support ↓	Item 1	Item 2
2	0.500	Creativity	Interpersonal S...
2	0.500	Creativity	Planning & orga...
2	0.500	Creativity	Planning and or...
2	0.500	Creativity	Problem solving
2	0.500	Creativity	Team work
2	0.500	Creativity	Technical Skills
2	0.500	Financial skill	Interpersonal S...
2	0.500	Financial skill	Planning & orga...
2	0.500	Financial skill	Planning and or...
2	0.500	Financial skill	Problem solving

Figure 7.43 Whole sale and retail trade job category with two associated skill types

Size	Support ↓	Item 1	Item 2	Item 3
3	0.500	Creativity	Financial skill	Interpersonal S...
3	0.500	Creativity	Financial skill	Planning & orga...
3	0.500	Creativity	Financial skill	Planning and or...
3	0.500	Creativity	Financial skill	Problem solving
3	0.500	Creativity	Financial skill	Team work
3	0.500	Creativity	Financial skill	Technical Skills
3	0.500	Creativity	Interpersonal S...	Planning & orga...

Figure 7.44 Whole sale and retail trade category with three associated skill types

Size	Support ↓	Item 1	Item 2	Item 3	Item 4
3	0.500	Interpersonal S...	Planning & orga...	Technical Skills	
3	0.500	Planning and or...	Problem solving	Team work	
4	0.500	Creativity	Financial skill	Interpersonal S...	Planning & orga...
4	0.500	Creativity	Financial skill	Interpersonal S...	Technical Skills
4	0.500	Creativity	Financial skill	Planning & orga...	Technical Skills
4	0.500	Creativity	Financial skill	Planning and or...	Problem solving
4	0.500	Creativity	Financial skill	Planning and or...	Team work
4	0.500	Creativity	Financial skill	Problem solving	Team work
4	0.500	Creativity	Interpersonal S...	Planning & orga...	Technical Skills
4	0.500	Creativity	Planning and or...	Problem solving	Team work

Figure 7.45 Whole sale and retail trade job category with four associated skill types

7.4 Summary

In this section, detailed evaluation of our build models was done, different types of classifiers were experimented and evaluated. Out of all discussed classifiers, Decision Tree performed well for this dataset. The next chapter has reserved for conclusions of the project and future expansions.

Chapter 8

Conclusion and Further Work

8.1 Introduction

In this research, the best model for the student employability prediction was identified by the evaluation of classification models. Further the most associated skills for each and every job category also evaluated and identified. The accuracy of the tool was very high because it has considered the fifteen most important academic skills and student professional skills also.

According to the results decision tree algorithm was the best technique to forecast the student employability in this research study. Further the most perceived skills for any job was communication skill and followed by team work skill, financial skill, technical skill and time management skill. This was identified by considering the highest support of each job category.

8.2 Limitations

In this research study examined the perception of diploma holders in a government higher education institute of graduate skills in business finance. After a review of past literature a survey questionnaire was established and dispersed among these pass out students using the Google Form in online platform. These are few limitations have identified. The students' perception in a government institute which offer business finance diploma and the generality of the conclusions shall be consider after gathering the facts from other government institutes which offer business finance diplomas.

The other limitation is geographical factors. In some situations, students may get a good career in some districts with relevant to the students' skills and professional qualifications but due to some sociocultural and geographical reasons those good opportunities may lost.

So, further investigations are required to validate the recognized characteristics. This can be conduct using qualitative research or can be applied an interview-based method.

8.3 Future Developments

As a future work the research study will be expanded to get more data and to validate the current data and records by using a qualitative research.

In this research, student's perspective of the graduate skills which are necessary for searching different job categories were considered and collected using the survey questionnaire. But the according to the employer's aspects for different types of job categories what they expected were not considered. So as a future work that employer's aspects will also consider.

8.4 Summary

The final chapter summarized the thesis by explaining the solution given with data mining techniques. Further classification and association techniques were used to evaluate the prediction of employability of the students and addressed the limitation of proposed solution.

References

[1] L. Harvey, "Defining and Measuring Employability, "*Quality in Higher Education*, 7(2), 97–109, 2001.

[2] H. P. Siriwardane, K. Y. Low and Blietz,"Making entry-level accountants better communicators: a Singapore-based study of communication tasks, skills, and attributes," *Journal of Accounting Education,* 33(4): 332–347, 2015.

[3] C. Yap, S. Ryan and J. Yong, "Challenges facing professional accounting education in a commercialized education sector," *Accounting Education*, 23(6): 562–581, 2014.

[4] G. Dale-Jones, P. Hancock and K. Willey, "Accounting students in an Australian university improve their writing: But how did it happen?," *Accounting Education* 22(6): 544–562, 2013.

[5] B. Jackling and P. De Lange, "Do accounting graduates' skills meet the expectations of employers? A matter of convergence or divergence," *Accounting Education: An International Journal* 18(4–5): 369–385, 2009.

[6] M. Osmani, N. Hindi, R. Al-Esmail and V. Weerakkody, "Examining graduate skills in accounting and finance," *Industry and Higher Education*, 31(5), 318–327, 2017.

[7] M. Osmani, V. Weerakkody, N. Hindi, "Graduate attributes in higher education: Examining academics' perception in the Middle East". *Journal of Education for Business* 92(2): 53–64, 2017.

[8] M. Osmani, V. Weerakkody and N. H. Hindi, "Identifying the trends and impact of graduate attributes on employability: a literature review". *Tertiary Education and Management* 21(4): 367–379, 2015.

[9] C. L. Wilkin, “Enhancing the AIS curriculum: integration of a research-led, problem-based learning task”. *Journal of Accounting Education* 32(2): 185–199, 2014.

[10] Z. Othman , F. Musa, N. H. Mokhtar, A. Ya’acob, R. A. Latiff and H. Hussin, "Investigating University Graduates’ English Language Competency towards Employability: A Proposed Research Method," *International Journal of Learning*, 17(7):429–40, Aug 2010

[11] Z. Othman ,F. Musa,N. H. Mokhtar, A. Ya’acob, R. A. Latiff, H. Hussein, A. W. Mohammad and S. Kaur, "Undergraduate Awareness and Readiness Towards Employability: The Significance of English," *British Journal of Arts and Social Sciences,* Vol.2 No.2, 2011.

[12] I. Potgieter and M. Coetzee, "Employability attributes and personality preferences of postgraduate business management students," *SA Journal of Industrial Psychology*, 39(1):1–10, Jan 2013.

[13] T. Stanley and S. Marsden, “Problem-based learning: Does accounting education need it?," *Journal of Accounting Education*, 30(3-4), 267–289, 2012.

[14] L. Ragland and U. Ramachandran, "Towards an understanding of excel functional skills needed for a career in public accounting: Perceptions from public accountants and accounting students," *Journal of Accounting Education*, 32(2), 113–129, 2014.

[15] D. J. Finch, L. K. Hamilton, R. Baldwin and M. Zehner, "An exploratory study of factors affecting undergraduate employability," *Education Training*, 55(7), 681–704, 2013.

[16] F. E. Gray, "Specific Oral Communication Skills Desired in New Accountancy Graduates," *Business Communication Quarterly*, 73(1), 40–67, 2010.

[17] G. Stone, M. Lightbody and R. Whait, "Developing Accounting Students’ Listening Skills: Barriers, Opportunities and an Integrated Stakeholder Approach," *Accounting Education*, 22(2), 168–192, 2013.

[18] A. M. Cox, M. Al Daoud and S. Rudd, "Information Management graduates' accounts of their employability: A case study from the University of Sheffield," *Education for Information*, 30(1-2), 41–61, 2013.

[19] D. Jackson, "Testing a model of undergraduate competence in employability skills and its implications for stakeholders," *Journal of Education and Work*, 27(2), 220–242, 2012.

[20] J. Hernández- March, M. Martín del Peso and S. Leguey, "Graduates' Skills and Higher Education: The employers' perspective," *Tertiary Education and Management*, 15(1), 1–16, 2009.

[21] T. Mishra1, D. Kumar and S. Gupta, "Students' Performance and Employability Prediction through Data Mining: A Survey," *Indian Journal of Science and Technology*, Vol 10(24), June 2017.

[22] M. Mukhtar, Y. Yahya, S. Abdullah, A. R. Hamdan,N. Jailani and Z. Abdullah, "Employability and service science: Facing the challenges via curriculum design and restructuring," in *2009 International Conference on Electrical Engineering and Informatics*, August 2009.

[23] L. A. Shafie and S. Nayan, "Employability awareness among Malaysian undergraduates," *International Journal of Business and Management*, Vol. 5, No. 8, August 2010.

[24] B. Jantawan and C. F. Tsai, "The Application of Data Mining to Build Classification Model for Predicting Graduate Employment," *(IJCSIS) International Journal of Computer Science and Information Security*, Vol. 11, No. 10, October 2013.

[25] T. Mishra1, D. Kumar and S. Gupta, "Students' Employability Prediction Model through Data Mining," *International Journal of Applied Engineering Research*, 11(4):2275–82, 2016.

[26] D. Rae, "Connecting enterprise and graduate employability," *Education and Training*, 49(8/9), 605–619, 2007.

[27] S.Rosenberg, R. Heimler and E. Morote, "Basic employability skills: a triangular design approach," *Education and Training*, Vol. 54 No. 1, pp. 7-20, 2012.

[28] D. R. Laker and J. L. Powell, "The differences between hard and soft skills and their relative impact on training transfer," *Human Resource Development Quarterly*, Vol. 22 No. 1, pp. 111-122, 2011.

[29] M. Pang and H. To Ming, "Preparing business students for the challenge ahead: a case study," *Journal of Teaching in International Business*, Vol. 16 No. 4, pp. 45-64, 2005.

[30] N. T. Nghe, P. Janecek and P. Haddawy, "A comparative analysis of techniques for predicting academic performance," in *37th Annual Frontiers in Education Conference-Global Engineering: Knowledge without Borders, Opportunities without Passports*, 2007 Oct, T2G-7.

[31] I. H. Paris, L. S Affendey and N. Mustapha. "Improving academic performance prediction using voting technique in data mining," *World Academy of Science, Engineering and Technology*, 62:820–3, Feb 2010.

[32] S.F. Shazmeen, M.M.A. Baig, and M.R. Pawar, "Performance Evaluation of Different Data Mining Classification Algorithm and Predictive Analysis," *Journal of Computer Engineering*, vol. 10(6), pp. 01-06, 2013.

[33] S. Elayidom, S. M. Idikkula, J. Alexander and A. Ojha, "Applying Data Mining Techniques for Placement Chance Prediction," in *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, 2009.

[34] M. A. Sapaat, A. Mustapha, J. Ahmad, K. Chamili and R. Muhamad, "A Data Mining Approach to Construct Graduates Employability Model in Malaysia," *International Journal on New Computer Architectures and Their Applications (IJNCAA)*, 1(4): 1086-1098, 2011.

[35] Y. Bharambe, N. Mored, M. Mulchandani, R. Shankarmani and S. G. Shinde, "Assessing employability of students using data mining techniques," in *2017 International Conference on Advances in Computing, Communications and Informatics* (ICACCI), 2017.

[36] T. Fawcett, "ROC Graphs: Notes and Practical Considerations for Data Mining Researchers," *Hewlett-Packard Company*, Palo Alto: CA, 2003.

[37] M. S. Kumar and G. P. Babu, "Comparative Study of Various Supervised Machine Learning Algorithms for an Early Effective Prediction of the Employability of Students," *JES Journal of Engineering Sciences*, Vol 10, Issue 10, Oct 2019.

[38] L. Zhang, X. Tan, S. Zhang and W. Zhang, "Association Rule Mining for Career Choices Among Fresh Graduates," *Applied and Computational Mathematics*. Vol. 8, No. 2, pp. 37-43, 2019.