CAREER PREDICTION

BY

Md. Masum Billah
ID: 131-15-2295

AND

Ujjal Raz Bongshi
ID:131-15-2257

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering.

Supervised By

Dr. Syed Akhter Hossain
Professor and Head
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

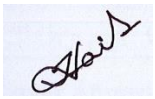
MAY 2017

APPROVAL

This Project titled **"Career Prediction**", submitted by Md. Masum Billah and Ujjal Raz Bongshi to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering (BSc) and approved as to its style and contents.

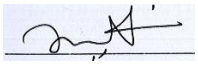
BOARD OF EXAMINERS

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University



Dr. Sheak Rashed Haider Noori — **Internal Examiner**
Associate Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Md. Zahid Hasan — **Internal Examiner**
Associate Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University



Dr. Mohammad Shorifuddin — **External Examiner**
Professor and Chairman
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Dr. Syed Akhter Hossain, Professor and Head, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Dr. Syed Akhter Hossain
Professor and Head
Department of CSE
Daffodil International University

Submitted by:

(Md. Masum Billah)
ID: 131-15-2295
Department of CSE
Daffodil International University

(Ujjal Raz Bongshi)
ID: 131-15-2257
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

Foremost, we are grateful to Allah for the good health and welling that were necessary to complete this project.

We would like to express our sincere gratitude to our supervisor **Prof. Dr. Syed Akhter Hossain,** Professor and Head, Department of Computer Science and Engineering, Daffodil International University, Dhaka, for the continuous support of our project study, for his patience, motivation, enthusiasm and immense knowledge. His guidance helped us in all the time of working and writing of this report. We could not imagine having a better supervisor and mentor for our project.

Beside our supervisor, we would like to thank the rest of the project committee for their encouragement, insightful comments and hard questions.

We would like to thank **Mr. Feroz Mahmud** Director of the Career Development Center of Daffodil International University for his kind and valuable information.

We would like to thank other faculty members and the staffs of Computer Science and Engineering Department of Daffodil International University, for their kind help to finish our project.

And specially we would like to thank those engineers who gave their valuable information as our project data.

We would like to thank our entire course mate and all juniors in Daffodil International University, who took part in this discuss while completing the course work.

Last but not least, we would like to thank our family specially our parents, for giving birth to us at the first place and supporting us spiritually throughout our life.

ABSTRACT

The project “Career Prediction” is a thesis based data mining project. We are looking for a pattern for CSE engineers. This pattern will help a student to know which department will be best for them or in which working sector he or she can work successfully. We collect data from CSE engineers whose are already established in CSE of these categories, 1. Academic Background, 2. Academic Interest, 3. Interested Working Sector, 4. Skill, 5. Personality, 6. Other activities.

After collecting data by using google form we do preprocessing out data by using Data Preprocessing techniques. And then we apply data mining techniques to find a pattern for CSE engineers. We apply Zero classification to classify data and K means clustering to clustered data.

After classification and clustering we can find a class and a cluster for maximum CSE engineers. This cluster will be the pattern for a CSE engineer.

After comparing student’s data with the cluster, a student can easily understand if he or she is perfect for Computer Science and Engineering.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Career is the progress and actions taken by a person throughout a lifetime. After completing higher secondary student hesitate to decide which way to go. It is very crucial moment for his life, because it will decide the career. It would be great if the student could know if he is suitable for the subject he wants to study. Our report analyzes the data of student and tells if the Computer science and engineering is suitable for him. So, that after completing his graduation he might become successful in his career.

If you look, you can see there are a huge percent of students, who are not happy with their department. That's why after a while they change their department or they drop out from study. At the same time, there are also a percent of service holders who are not happy with their service sector or their working career. The main cause is that they could not select their department correctly. A student chooses wrong department without being knowing if she or he is perfect for the department are the main failures.

That's why we want to do this project which will help a student to take decision which department will be the best department for him or her. It is very important decision for a student. Also, important for career. If our decision doesn't match with our capability, then we can't deserve a better outcome from that decision. So, to take a decision the first thing need to know ourselves. In which department, I am good at and which working sector I like and also which work I will be able to do by myself, then I can take the decision.

Our project will help a student to select the Computer Science and Engineering department as him or her studying area and working area.

1.2 Motivation

Now a day's everyone is worried about their career after the graduation. But the most important step before entering career is choosing subject for Bachelor's after Higher Secondary Certificate exam. A student chooses subject with only knowing a little. Most of the students don't know his capability. Our motivation is to introduce a student with his capabilities. When he knows his capabilities, it will be easier for him to take first step toward his career.

We motivated from some of our friend whose already failure as a CSE student.

Some of our senior, who are not happy with their job sector.

We motivated from some of our classmates who are not perfect for his or her department.

Our past motivated us very much. When we were at that situation we were also confused to choose, which subject will be best for us and which working sector will bring happiness for us. It was hard time.

1.3 Rational of the Study

A systematic inquiry that investigates hypotheses, suggests new interpretations of data or texts, and poses new questions for future research to explore is known as research. For finding solution to a problem, research is done. Solution of a problem can be found through research in a structured way. In every research, there should be clear objective. Without the objective, the candidate might get lost from his track of research. While starting the research, the candidate must have the clear idea of his objective of the research. On measuring the variable, such as identify or describe them is generally focused by research objective.

The research objectives are given below:

- ✓ To find the main characteristics of CSE engineers.
- ✓ To find the pattern for students who are eligible and will do better in CSE.
- ✓ To find suitable working field in CSE for the student.
- ✓ To help a student to know his or her inner strength.
- ✓ To decrease the percent of dropout students after HSC.

1.4 Research Questions

When we started our project, there were several questions which arose in our mind. Which were our project related. Among those questions there were some major sectors, with which we completed our research. These sectors are “subjects you like”, “subjects you are good at”, “Interested activities”, “Interested working area”, “Working style”, “Importance of demand”, “Thinking style”, “academic result” and “other activities”.

1.5 Expected Outcome

In this project “Career Prediction” we wanted to find pattern of students who are suitable for choosing CSE related jobs for their career. So, we have collected patterns of already established persons from different working areas of CSE. If these characteristics match at a minimum level, he is suitable to choose CSE for his career.

1.4 Report Layout

This research paper consists of five chapter. Introduction, Literature Review, Data Mining Requirements, Data mining Implementation, Discussion and Conclusion.

Chapter 1: Introduction; Introduction, Motivation, Rational of the Study, Research Questions, Expected Outcome, Report Layout.

Chapter 2: Background; Introduction, Related Works, Research Summary, Scope of the problem, Challenges.

Chapter 3: Research Methodology; Introduction, Research Subject and Instrumentation, Data Collection Procedure, Statistical Analysis, Implementation Requirements.

Chapter 4: Experimental Results and Discussion; Introduction, Experimental Results, Descriptive Analysis, Summary.

Chapter 5: Summary, Conclusion, Recommendation and Implication for Future Research; Summary of the Study, Conclusions, Recommendations, Implication for Further Study.

CHAPTER 2

BACKGROUND

2.1 Introduction

Career prediction is done with survey analysis. Where data is gathered by completing survey by users. Survey is done through google form. Answers of the survey questions can be given in many ways. But to specify and for ease of access most of the answers are taken in a range with rating. If answers were taken from the users, as they want to give, some of them would have given junk data's which would be useless. In this project from the data's we have collected, we are going to find the pattern of the characteristics of CSE students. Which will be helping a fresher either to take CSE background as his career or not.

By the definition of career, we know it Career is the progress and actions taken by a person throughout a lifetime.

There has a deference between a normal person career and a student career.

A normal person's career is depending on his or her working sector and time. But for a student there is a sequence of period. From class one to five a student just learns how to learn.

From class 5 to 8 a student learns about deferent department of study like science, business and arts.

The first step of career starts after passing class 8. At that time, a student chooses group of either Science, Commerce or Arts. Then the area of his career specifies slowly and gradually.

After HSC, the main time has come to take decision, so that which department's knowledge he or she can gather. If a student can choose a right subject to learn deeply then this student will be the best for his or her subject but if can't then that student will be the failure and that student will not get a better job in his sector.

Like a student studying in Computer science and engineering. But this student obtained those characteristic

1. He is not good at math
2. Not good at engineering subject
3. Not good at programming
4. Good at biology
5. Good at chemistry
6. He is not interested to work computer
7. He is interested about health science

So, for those characteristic that student was not perfect for CSE. That student was perfect for Biology, chemistry related any subject

A student career depends on that student's characteristic. To know which department match with his or her characteristic is very important. If a student can find it then that student can select a subject easily and can make a better life more than previous example's student.

Like a student is obtained those characteristics given below

1. Good at math
2. Good at engineering subject
3. Interested to work with computer
4. Expert in puzzle game
5. Interested to know about computer deeply
6. Good at problem solving
7. Concern about new technology
8. Good at pure science
9. Expert in using internet

Then if these students choose CSE as his or her graduation subject then that student will do better. And these students can find easily his or her working area in CSE and will be successful as a CSE engineer.

So, to predict a student career it's very important to know that student's characteristic including

1. Academic Background
2. Academic Interest

3. Skill
4. Interested working area
5. Personality
6. Others

Then it will be easier to predict the career of a student.

2.2 Related Works

In recent past, there were not enough related works are done. But there were a few related works are done and some of those works helped us a lot.

Related work:

1. "Educational Data Mining: A Mining Model for Developing Students' Programming Skills - by Asraful Alam Pathan, Mehedi Hasan, Md. Ferdous Ahmed, and Dewan Md. Farid, Department of Computer Science and Engineering, United International University, Bangladesh
2. "Knowledge Mining for Effective Teaching and Enhancing Engineering Education" - by Dewan Md. Farid, and Hasan Sarwar, Department of Computer Science and Engineering United International University

There are some websites which are working with career prediction.

1. 16Personalities
   https://www.16personalities.com/
2. Mettl
   https://tests.mettl.com/authenticateKey/f1d45793

2.3 Research Summary

Career is the most important subject in every man's life. But most of our people don't have any idea or knowledge about career. As most of them cannot choose proper career for them in their life.

Finding the potential: A single person cannot be good at everything. Every person is good at a specific thing. He must find his potential. If he cannot find his potential, he will remain in his darkroom of ignorance searching for his potential. A man who knows his potential, can work greatly at his working field. He will enjoy and find everything easy at his working field. But most of the people, keep searching for this potential until doing jobs. Career prediction will help him to find his potential. He will know about himself, at where he is good and where he is weak. This will help him to find his desirable and most suitable career for him. At our country, there are more than 70% people who are not happy with their career. They are continuing for the livelihood. If these people would know about their potential, they would enjoy their work and would be at proper place where he should have been.

Reducing the dropouts: After completing the Higher Secondary certificate exam students get admitted to different universities and colleges to study different subjects. But a study shows that about 45% students dropout from study after admitted newly. They dropout for many reasons. Some of them doesn't like the subject or cannot do good in it. At here, career prediction helps one to choose the subject he is good at and will do great at future.

Mentally satisfaction: Everyone in this world likes to do what he loves. Most of them cannot do it. Everyone chooses different types of careers. Most of the people hanker after money. Where there is more money there is more people. They do not think about mentally satisfaction. As a result, they earn good and lives good but not satisfied mentally with career. At here, career prediction helps one to achieve mental satisfaction. His career will be predicted from his given data's. A career that matches with him will be suggested to him. So that he loves what he does. Money doesn't give happiness and satisfaction always.

Physically Satisfaction: We all work according to our physical capabilities. A person who has energy of working 8 hours, cannot work for 12 hours. If he forces his body, he will be sick of wouldn't be happy. Career prediction helps one to get physical satisfaction. Through career prediction one will get suitable career where he will be working for that amount of time which he is capable of. Not more or less than that.

Financial satisfaction: People sometimes earn less, when he can earn more if he chooses right career for his own. They cannot choose right career for their own. As a result, the potential of the person is wasted, he lives with the amount he earns. At here, career prediction helps people so

that they earn money according to their potential. If he has good potential he will earn good and if less he will earn less. By this way, he will not earn lower than his potential. He will financially satisfied, along with mentally satisfied.

2.4 Scope of the Problem

In every research, we all face some problems. It's very much difficult to cope with those problems. The working procedure of us might not match with the future or with the past. This is one of the biggest problems. In our project, we faced several problems.

- First, we were in very much trouble, while setting the attributes for the survey question. It took lot of time to set the correct attributes.
- After setting the survey question it needed to be verified. We needed to know, if the questions are correct to predict career of a student.
- Finally, we faced problem with cluster selection. We selected the cluster considering the present. In future clusters might change or not.

2.5 Challenges

Every important works and hard works always there have some challenges. Our projects also were some challenges. Those challenges make our works interesting and enjoyable. From those challenges, we can be learnt many new things.

The very big challenges were to make the question set. When we were making question, we faced many problem and many case studies. Firstly, we made a question set and applied to our desire engineers but it was totally incorrect. Then we made new question set and checked by our honorable supervisor and some teachers. That was also not fully correct. Then finally we made a question set with the help of CDC of our University, some books, some online tools.

The next challenge was to make the attribute table from question set.

Data collection was also a very big challenge. A good data set provide a good knowledge that's why it was challenging.

Finding pure CSE engineering was also challenging.

Finally, to gather knowledge from data it was most challenging part of our project. Our whole project was depending on outcome that's why it was very important to us. We did this challenging part very carefully.

At last we want to say that we enjoyed those challenges very much.

CHAPTER 3

RESEARH METHODOLOGY

3.1 Introduction

The process used to collect information and data for the purpose of making business decisions. The methodology may include publication research, interviews, surveys and other research techniques, and could include both present and historical information.

As already mentioned my research was based on the career prediction of a student. We are working for a pattern for a perfect Computer Science and Engineering. Due to the latter, this issue was talking about the process how we do this research. After making the question set which was most difficult and challenging part of this project we collect data by using google form. For a pure data set we pre-process data by using data pre-processing techniques of Weka. And then we apply data mining techniques to find the expected outcome of this research.

From the knowledge of mining we make a cluster or make a class for CSE engineers. This cluster or class contain the characteristic of CSE engineers.

3.2 Research Subject and Instrumentation

In our research the main subject were to predict the career of a student. At here there are some sub subjects and they are:

Making question set: In making question set we used several research papers, different websites and different person's suggestion.

Collecting pure data: We used to google form as instrument to collect data from the students.

Data Pre-processing: We used Weka's built in pre-processing technique as pre-processing instrument for data pre-processing.

Data Mining: We used Weka's default data mining algorithms (Zero classification, Simple k-means) to mine from the data.

3.3 Data Collection Procedure

We used "Google Form" to collect data. The google form link is https://goo.gl/forms/Wzjbmozy6EWBIYdD3 and the form is

Give rating to the subjects you like. In this section one will give data about which subject's he like most.

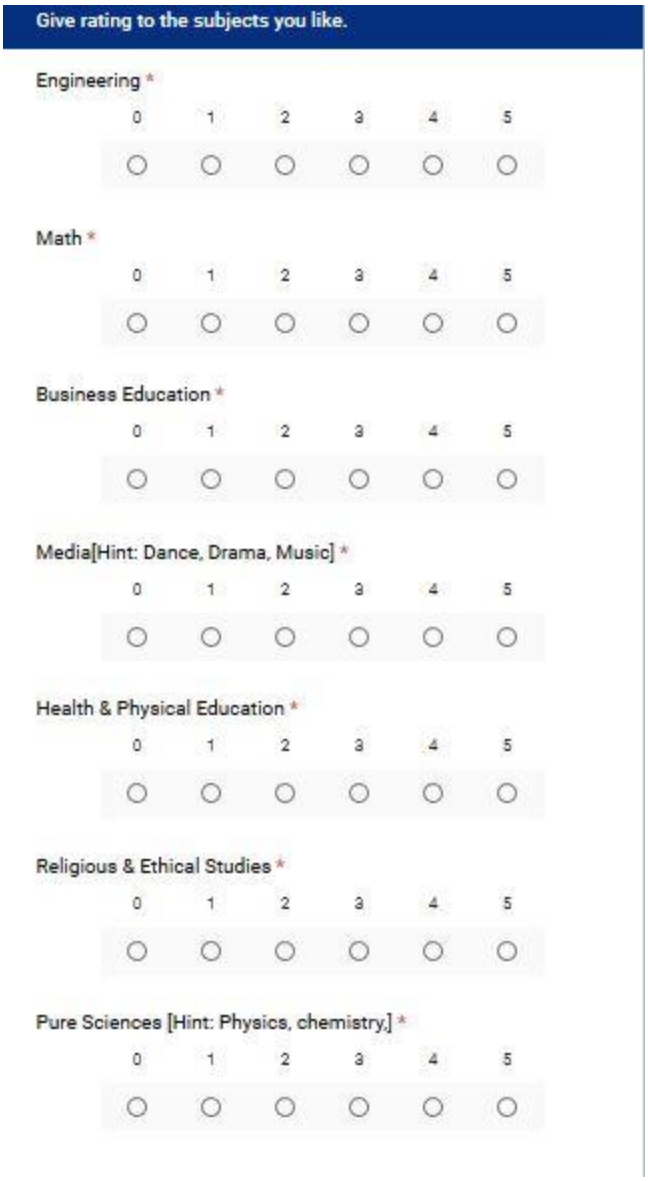


Figure 3.1: Google Form for subject you like

One can like more than 1 subjects and can give rating from scale 0-5, about how much he likes the subject.

Subjects you are good at. In this section one will give data about, which subjects he is good at.

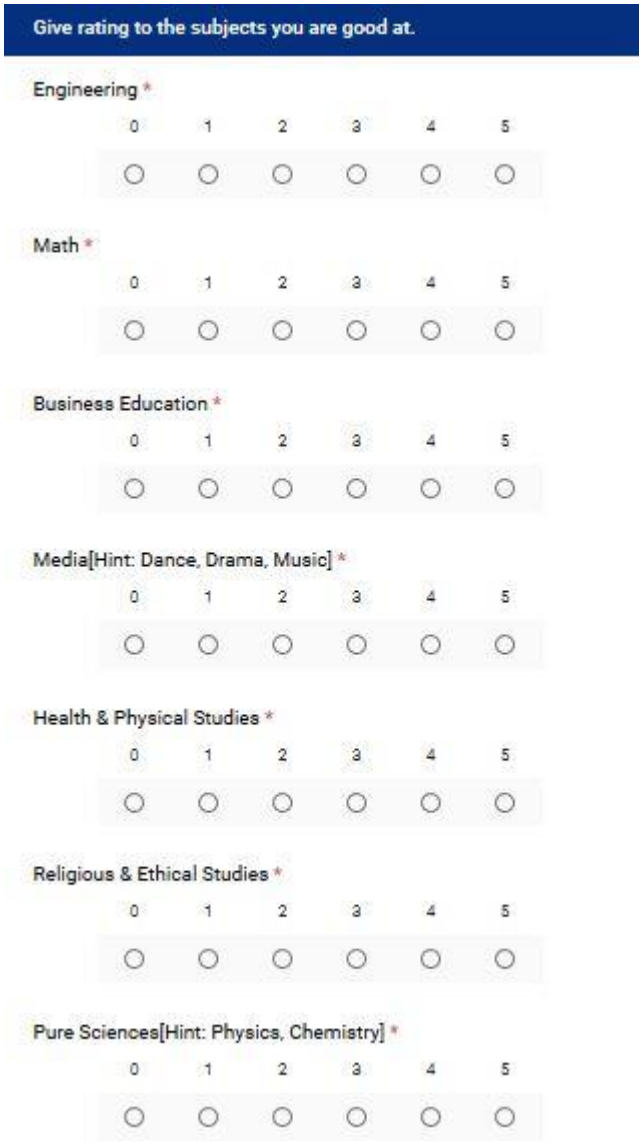


Figure 3.2: Google Form for subject you good

One can be good at more than 1 subjects and can give rating from scale 0-5, about how much he is good at the respected subjects.

Which activities you are interested in. Interest differs from person to person. One of the major

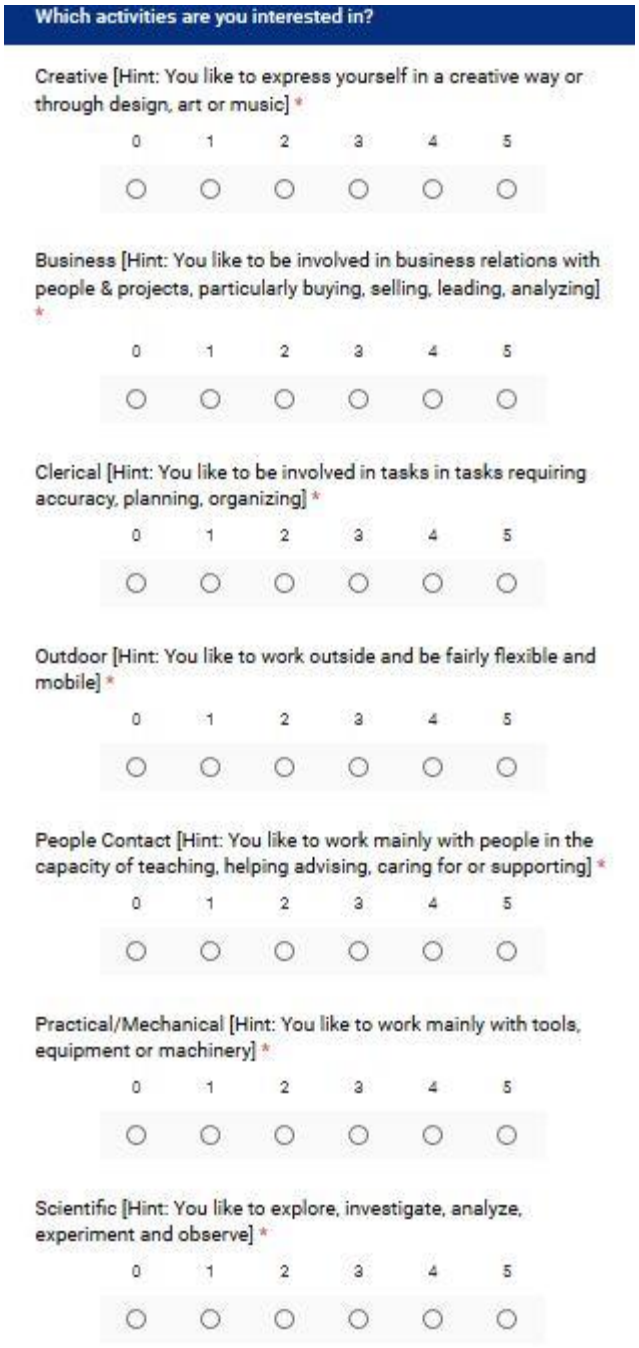


Figure 3.3: Google Form of Interested activities

characteristic of one person is interest. A person does great at anything at which he is interested. In this section, the data of person's interest is taken.

Which field would you like to work in. Not all people would like to work at same field. One

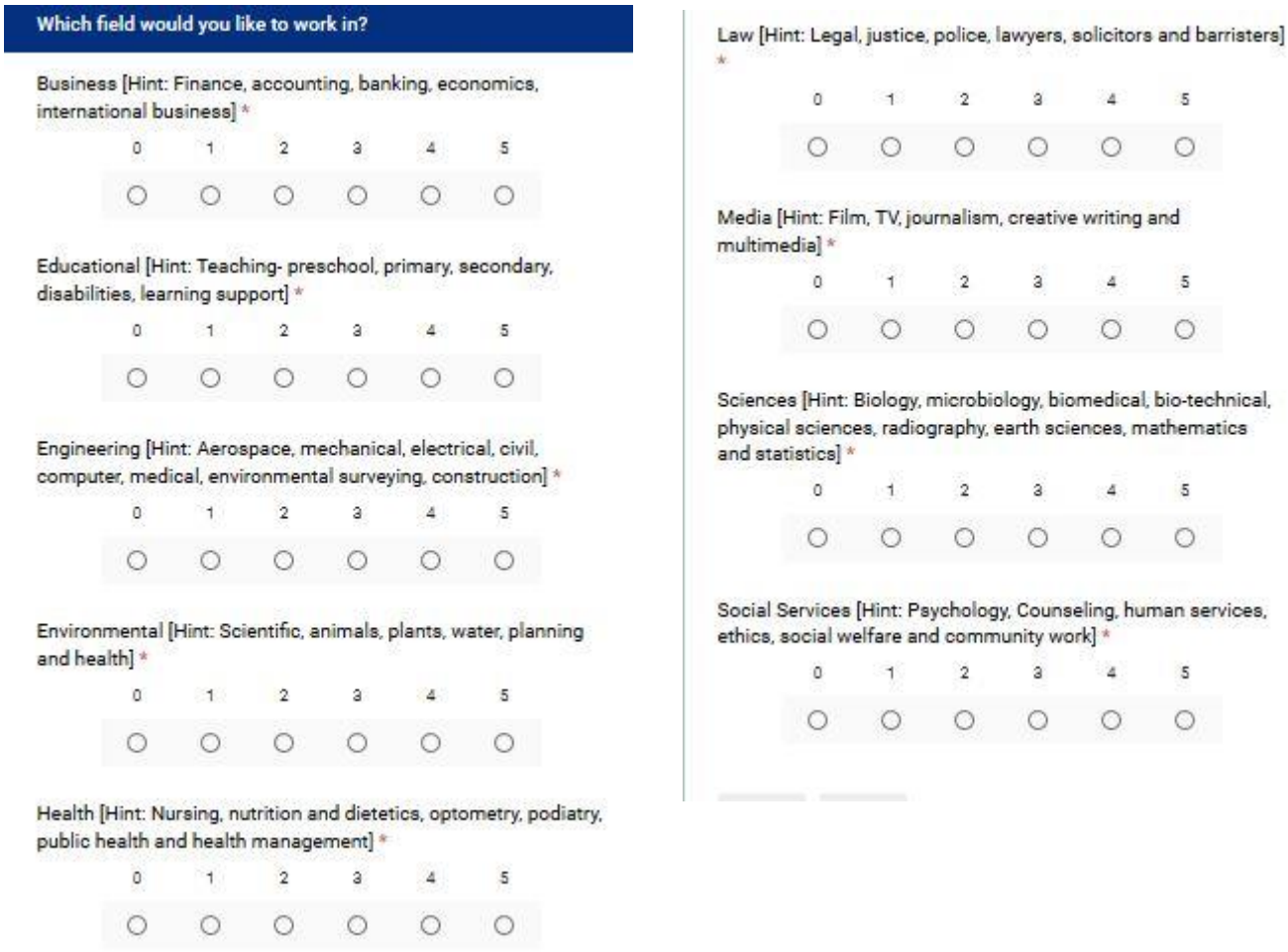


Figure3.4: Google Form of Interested working field

should choose career at the field they would like to or enjoy. This section collects data about the working field they like to work.

How would you describe your working style? Every people have different styles. Not every

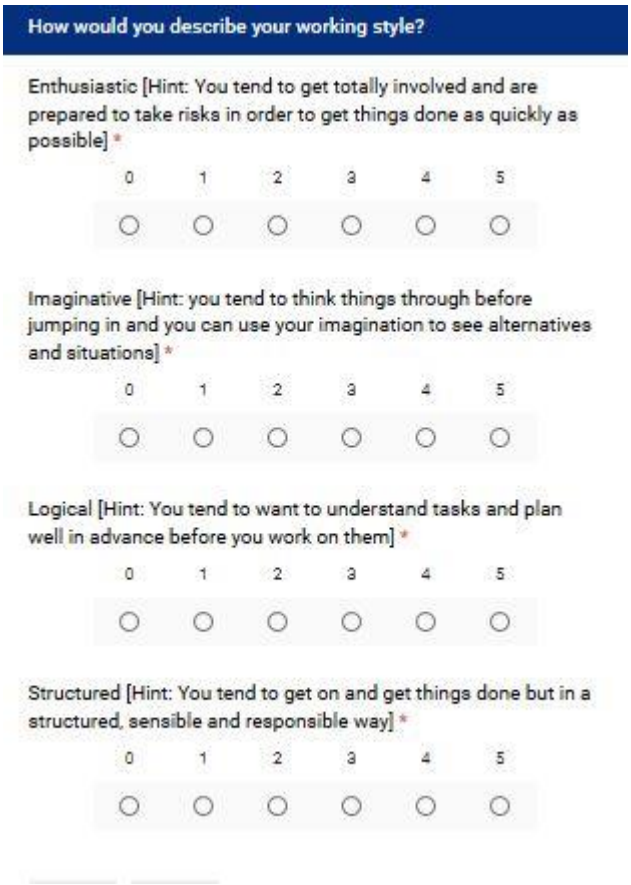


Figure3.5: Google Form of working style

people work in same style. So, this section collects data in four different categories. These are Enthusiastic, Imaginative, Logical and Structured. We will be able to know how much of this working style is in within a person through this section.

What's important to you? Life is not a bed of roses. Every single person in this world has

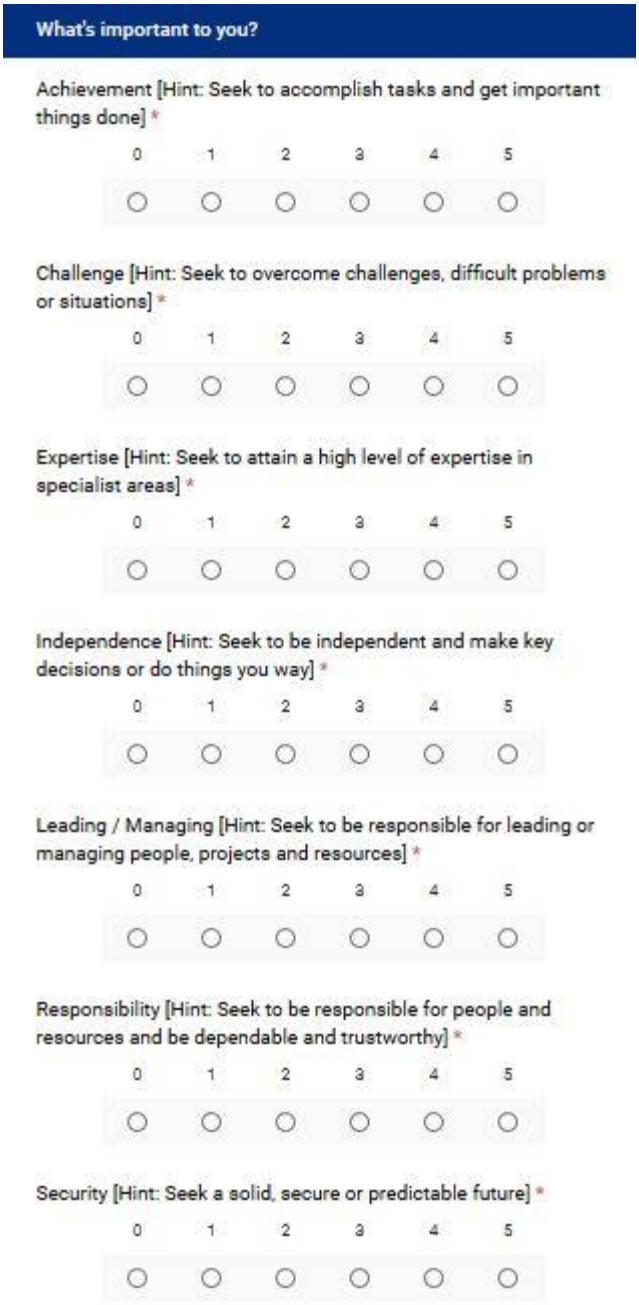


Figure 3.6: Google Form of What's important to you

expectation. Different persons give priorities to different things. At this section, how much importance a person at different thing's is collected. This section collects data about achievement, challenge, expertise, independence, leading, responsibility and security.

How do you usually think about things? Thinking of different persons cannot be same. Every

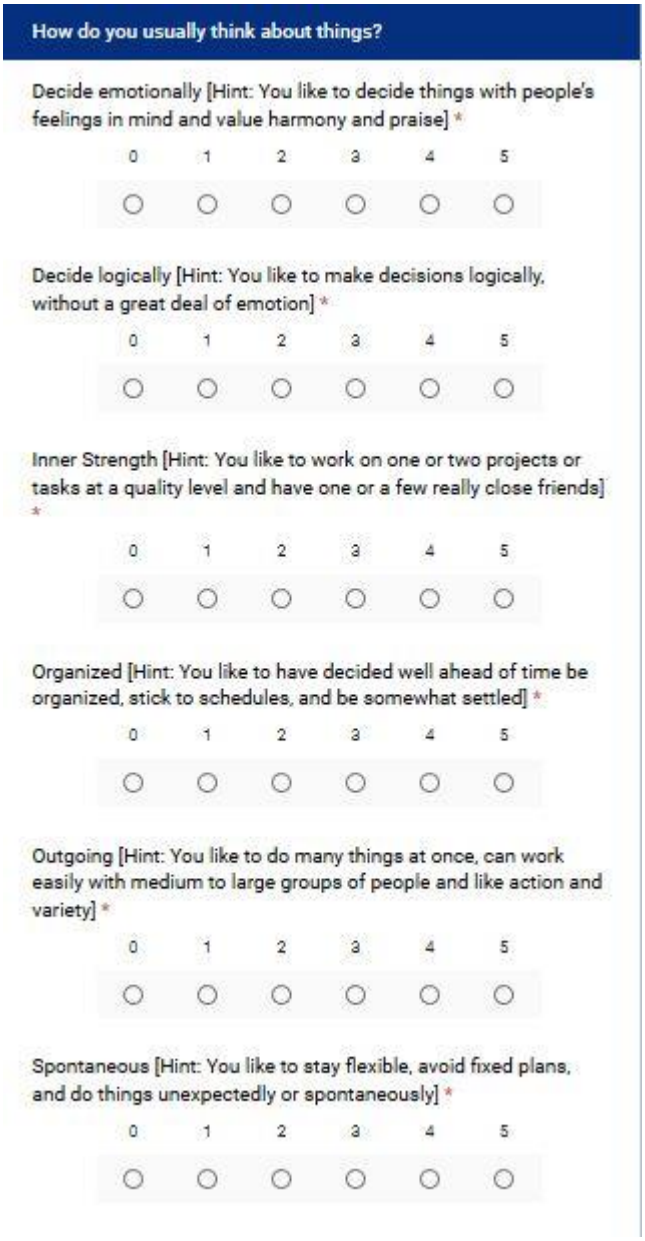


Figure 3.7: Google Form of Thinking style

person thinks in different way. This section collects data about how he thinks about different matters. This section collects data about how much he decides emotionally, logically, inner strength, organized, outgoing and spontaneous thinking of a person.

Academic Results. It very important that how much a student was good at his previous days.

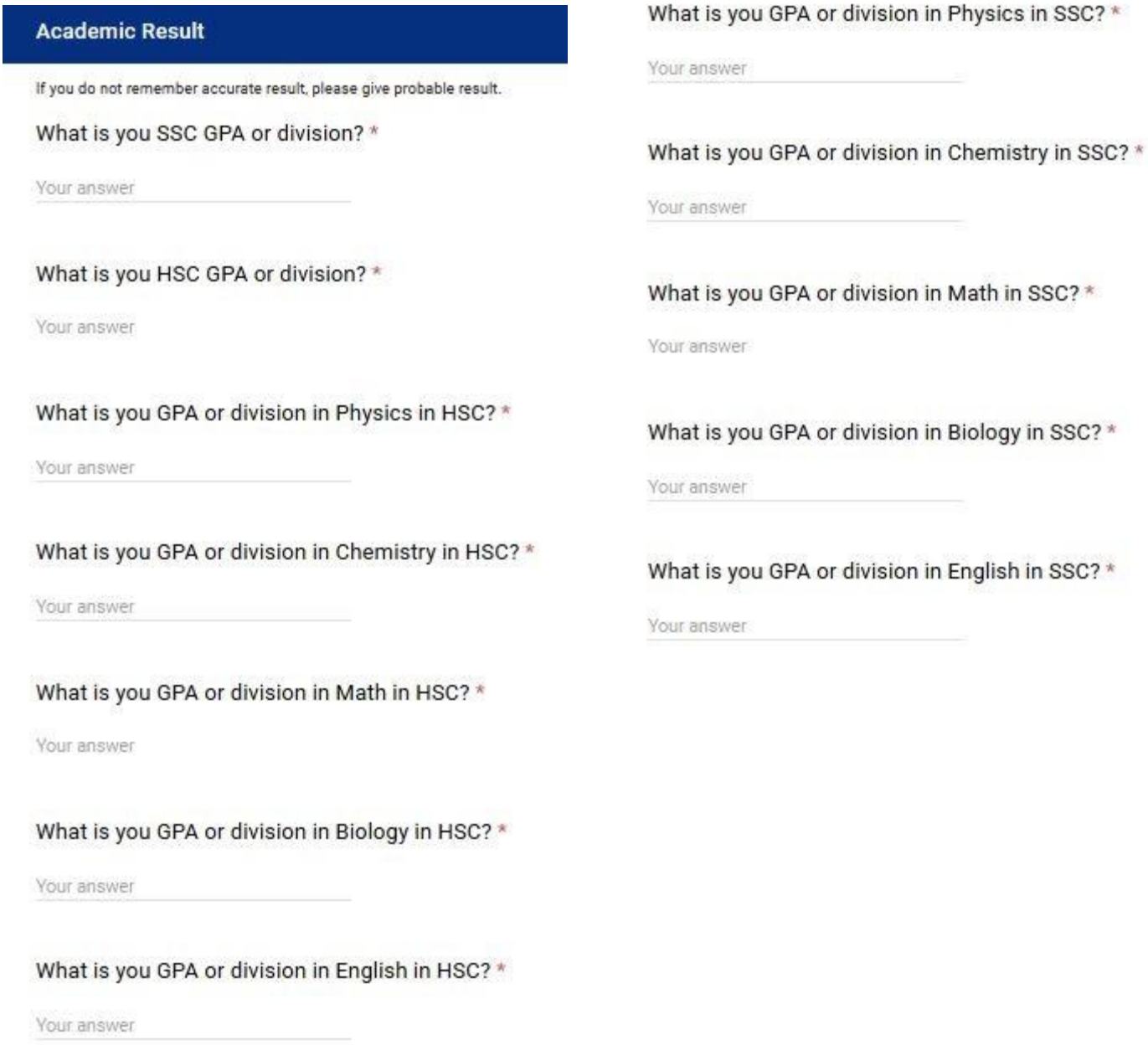


Figure 3.8: Google Form of Academic results

These records will be collected from academic results. In this section persons, academic result of HSC and SSC is collected.

Other activities. All the other necessary information's are collected through this section.

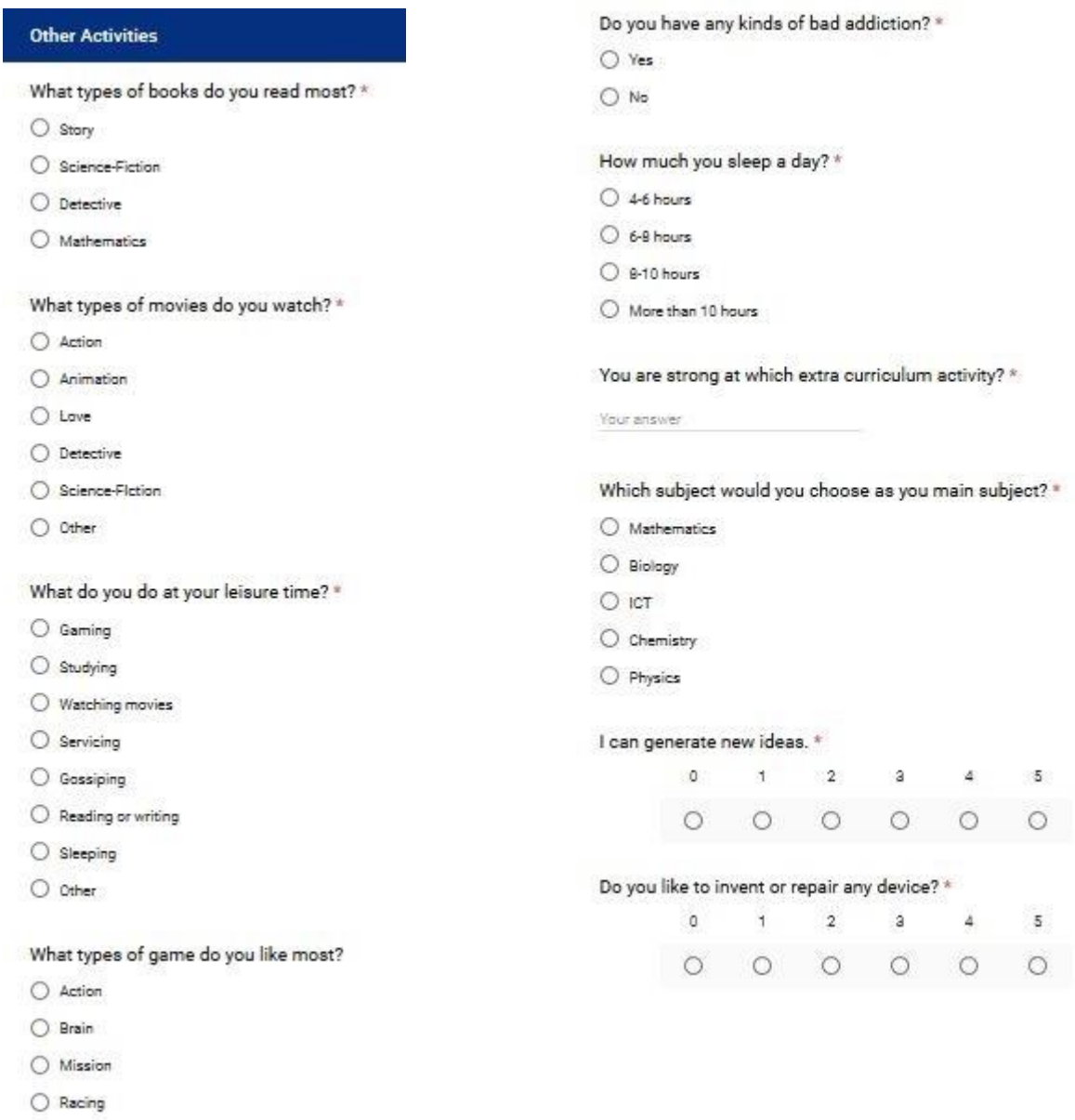


Figure 3.9: Google Form of Other activities

Information's like what book, movies, games a person's likes. What he does at his leisure time, how much he sleeps a day and how would he like to describe himself. After collecting data, we pre-processed data to cut off all unexpected and unwanted data generated by google form. A pure data mining depends on a pure data set.

3.4 Statistical Analysis

After analyzing the result, it shows some statistics as follow:

The statistical result for subject CSE engineers like are:

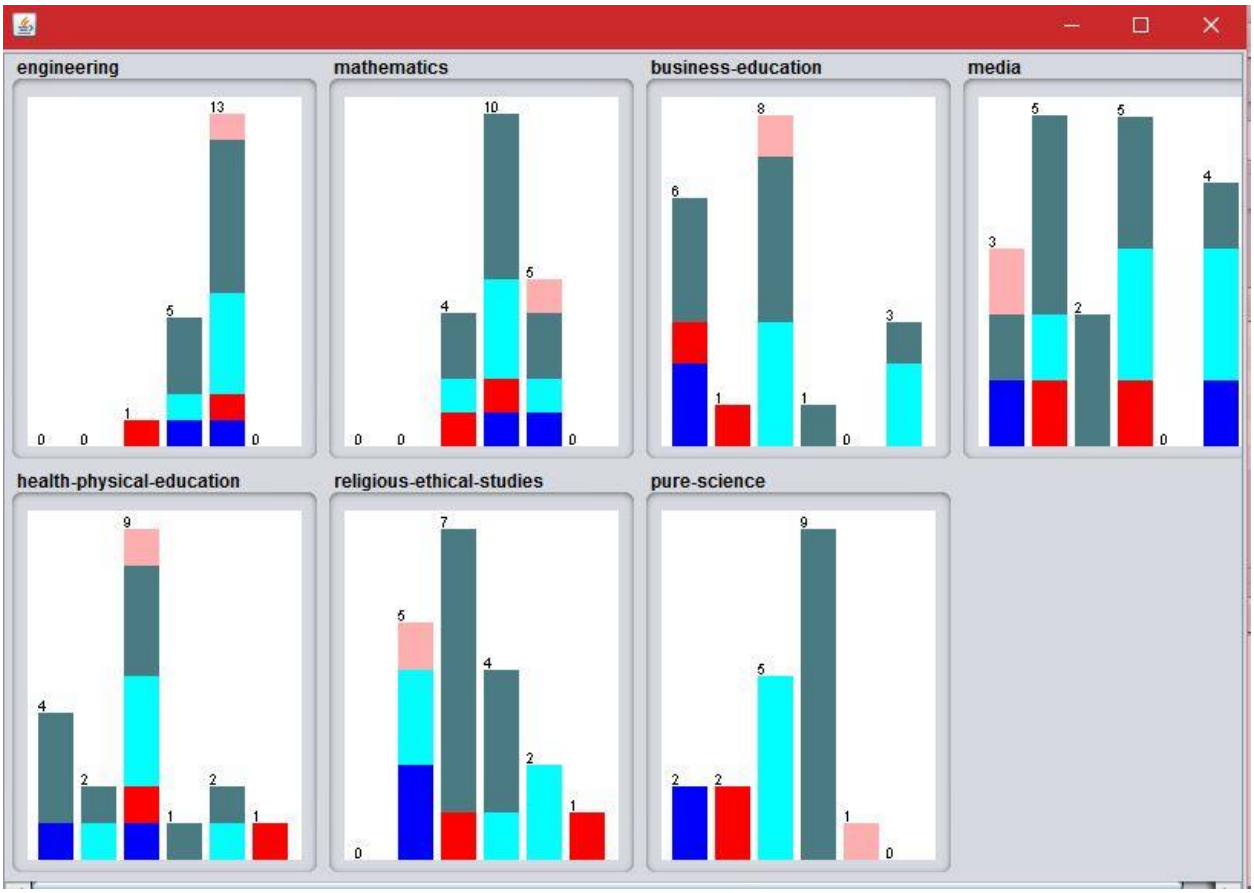


Figure 3.10: The statistical result for subject you like

The statistical result for subject CSE engineers good at:

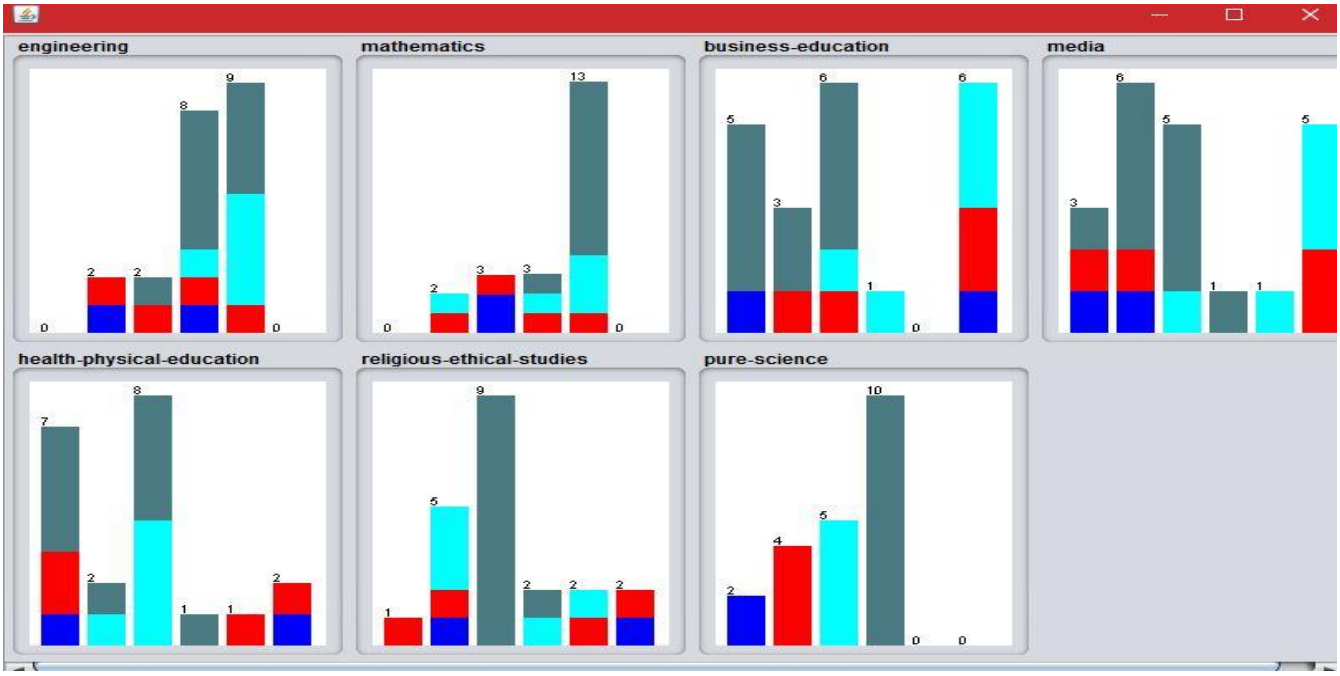


Figure 3.11: The statistical result for subject you good

The statistical result for interested activities of CSE engineers:

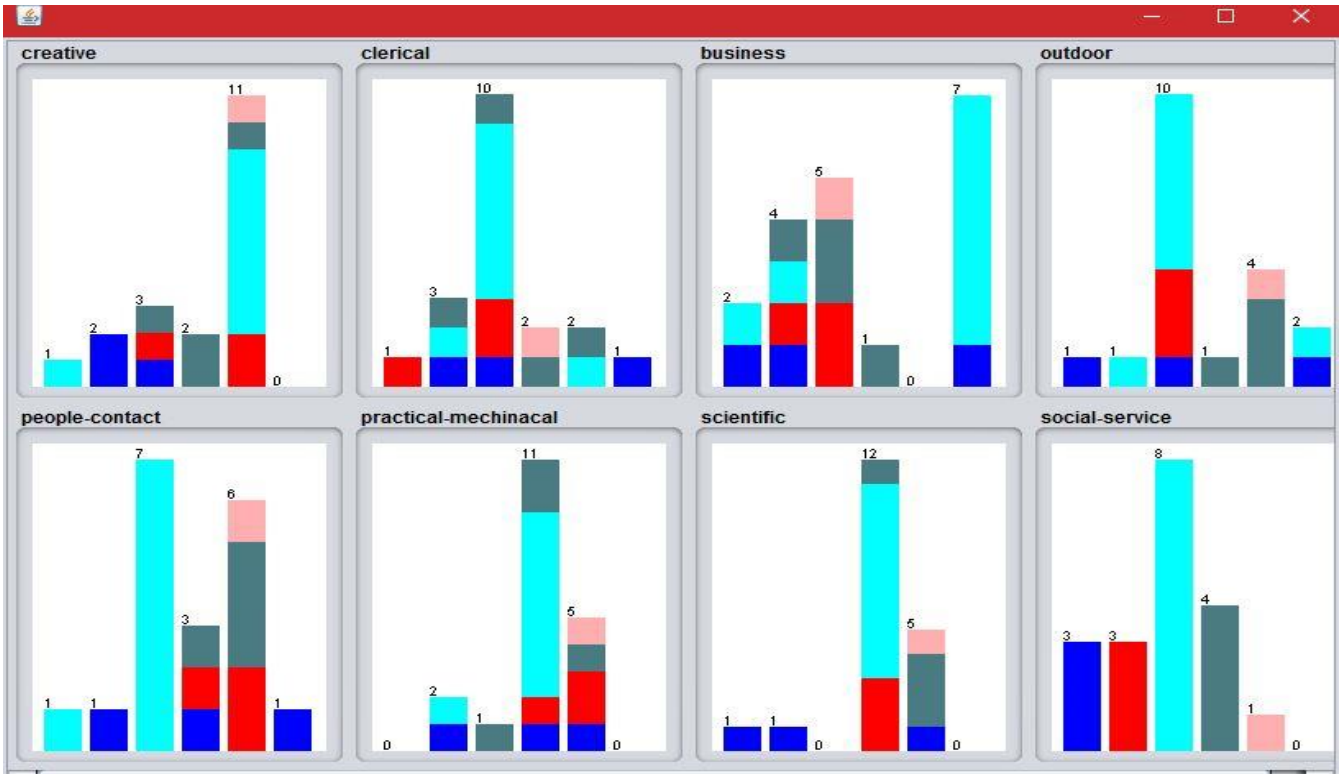


Figure 3.12: The statistical result for interested activities

The statistical result for interested working field for CSE engineers:

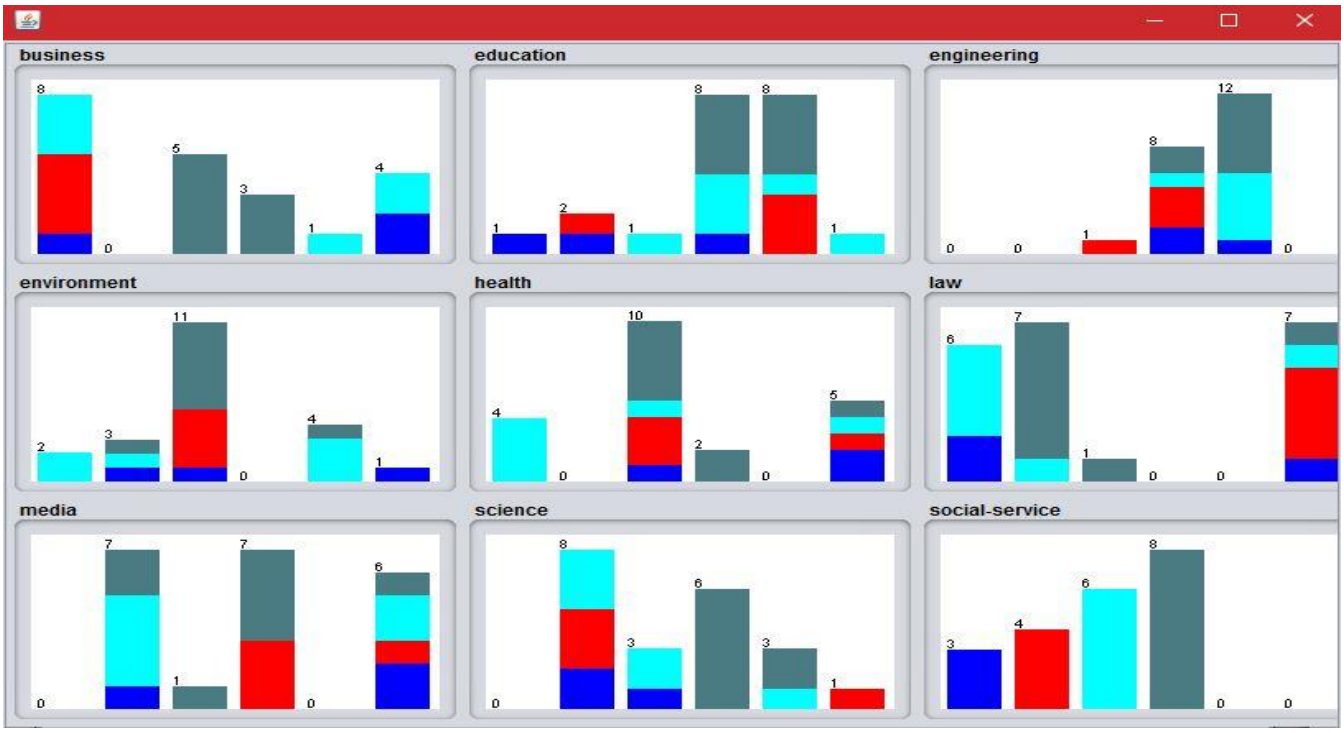


Figure 3.13: The statistical result for interested working field

The statistical result for interested working style of CSE engineers:

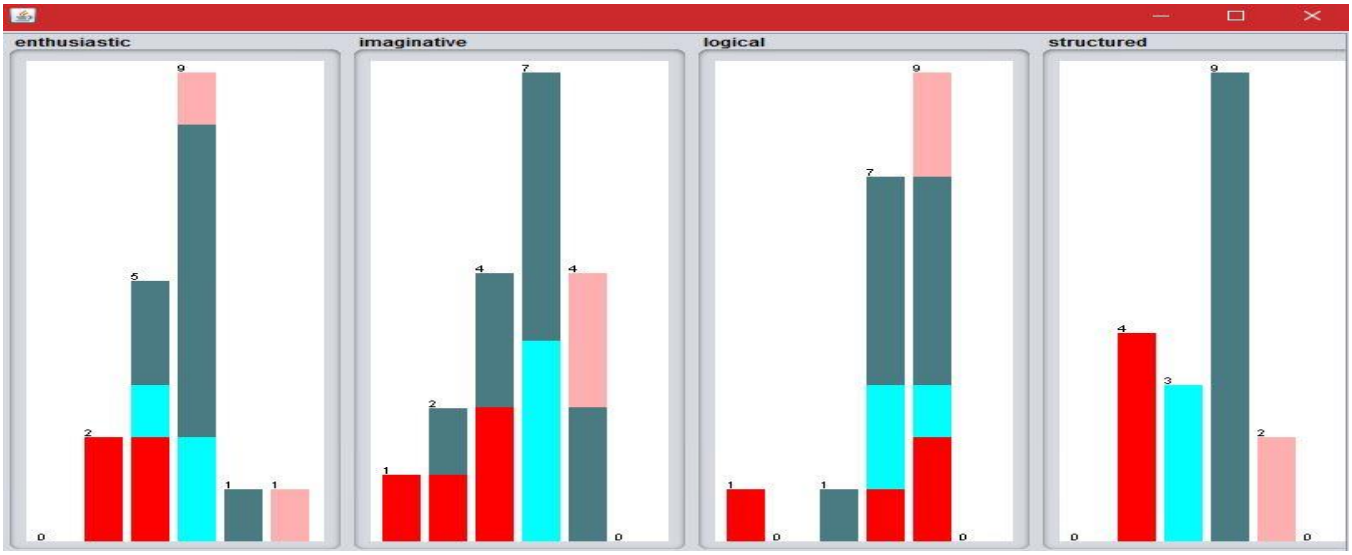


Figure 3.14: The statistical result for working style

The statistical result for what's important for CSE engineers:

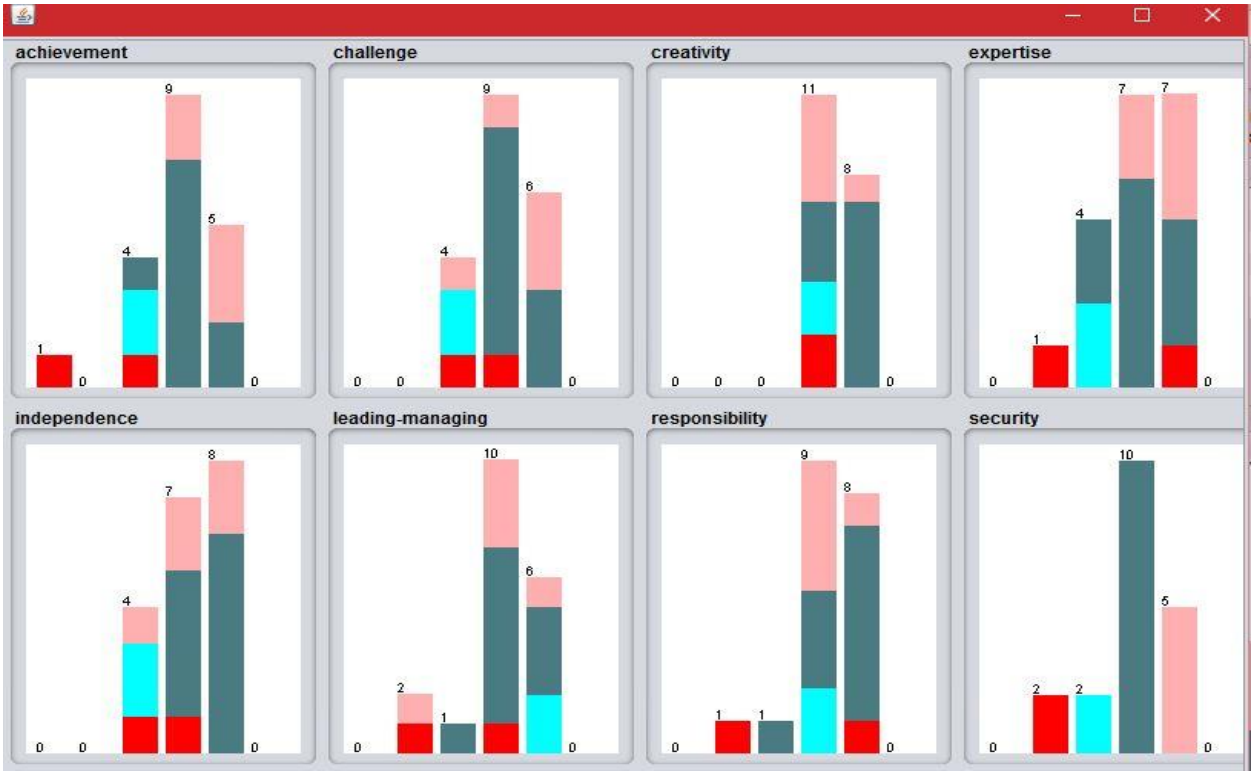


Figure 3.15: The statistical result for what's important

The statistical result for thinking style of CSE engineers:

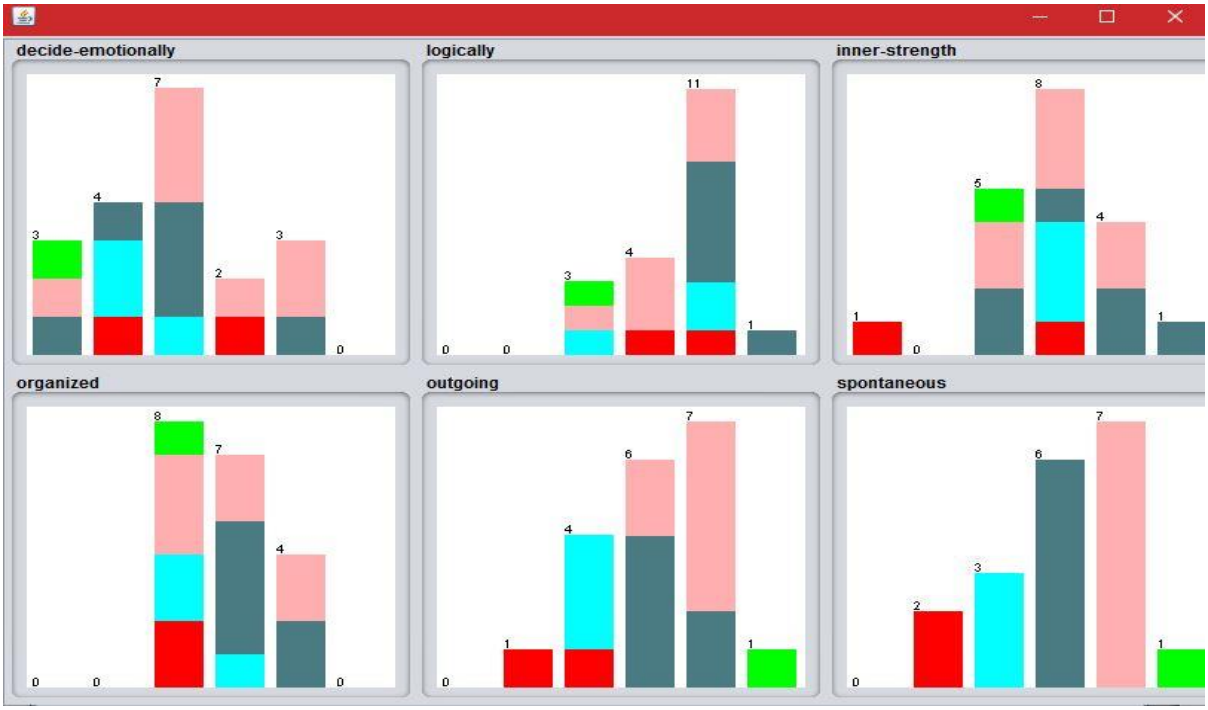


Figure 3.16: The statistical result for thinking style

The statistical result for all other activities of CSE engineers:

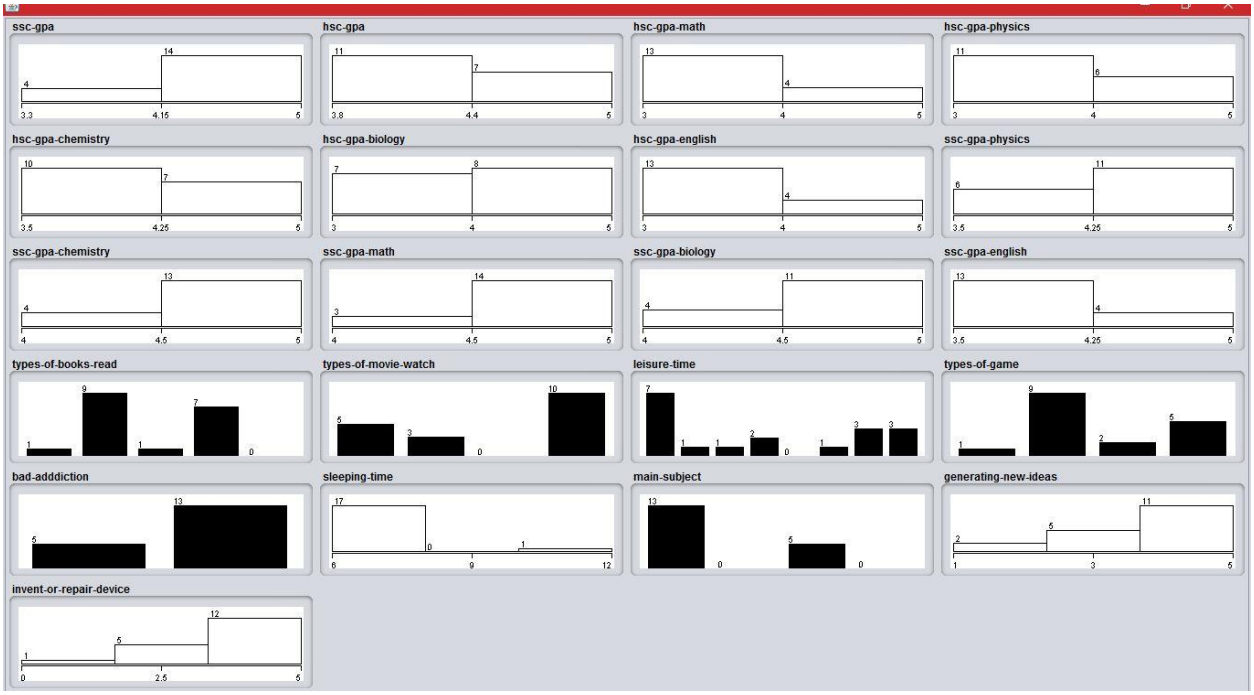


Figure 3.17: The statistical result for other activities

3.5 Implementation Requirements

To implement this project to students we need three contents

1. Develop a system: We need to develop a system where students will be able to know about their career. So that every student gets proper instruction to choose next step for his career. I this system there are there are three part. One for taking input, another one is for comparing input with already developed cluster and final one to apply the result.
2. Publicity: When a student knows about our system then that student can apply it. That's why we need to publish this project in different media. By those publication, it will be clear to our student.
3. Applying: We need to apply and collect data for further development.

CHAPTER 4
EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

Every research is done with patience and in the end the result of the research is found. Experimental results decide whether the research is successful or not. In life guidance is a must. After coming to this world, we are guided by our parents, in school and college we are guided by our teachers. In every step of our life guidance is needed. If one gets the guidance he needed, he can do great things. Sometimes without proper guidance a child might be spoilt. After passing HSC, a student takes first step towards his career. If he gets proper guidance he can shine in his career. But in most cases students don't gets this guidance as a result they fail to achieve a successful career. Career prediction helps a student to give proper guidance a student needs. Career prediction helps a student to take decision about his first step of his career.

4.2 Experimental Results

After doing the research we got our expected results.

We use Simple K-Means for clustering data.

Simple K-Means

In the clustering problem, we are given a training set $x(1),...,x(m)$, and want to group the data into a few cohesive "clusters." Here, we are given feature vectors for each data point $x(i) \in \mathrm{R}n$ as usual; but no labels $y(i)$ (making this an unsupervised learning problem). Our goal is to predict k centroids **and** a label $c(i)$

for each data point. The k-means clustering algorithm is as follows:

1. Initialize **cluster centroids** $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{R}^n$ randomly.
2. Repeat until convergence: {

   For every i, set

   $$c^{(i)} := \arg\min_j ||x^{(i)} - \mu_j||^2.$$

   For each j, set

   $$\mu_j := \frac{\sum_{i=1}^{m} 1\{c^{(i)} = j\} x^{(i)}}{\sum_{i=1}^{m} 1\{c^{(i)} = j\}}.$$

   }

Euclidean Distance:

The notation $\|x-y\|$

means Euclidean distance between vectors *x* and *y*

The clustering default setting values

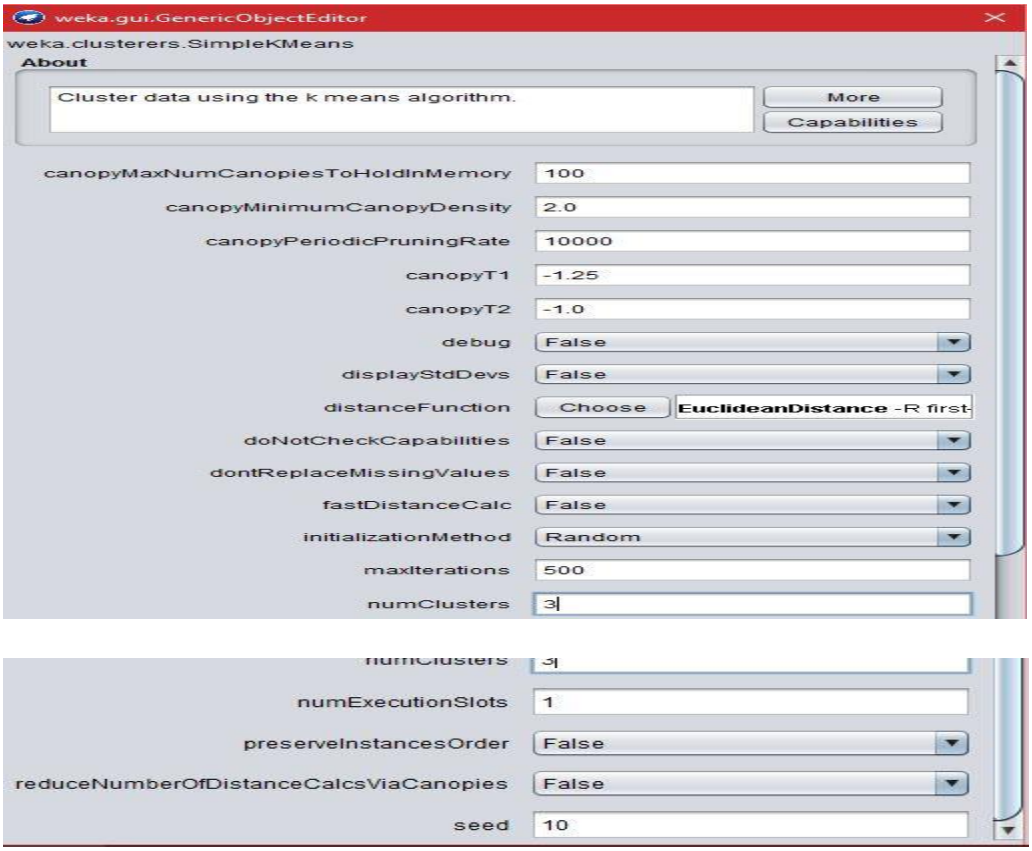


Figure 4.1: Setting of Weka for clustering

This is the cluster value for the subject CSE engineers like most.

```
Final cluster centroids:
                                    Cluster#
Attribute                 Full Data         0          1          2
                            (10.0)       (5.0)      (3.0)      (2.0)
=============================================================================
engineering                     e5          e5         e4         e5
mathematics                     m4          m4         m4         m3
business-education              b3          b1         b3         b0
media                          md4         md2        md4        md0
health-physical-education       h3          h1         h2         h3
religious-ethical-studies       r3          r3         r4         r2
pure-science                    p4          p4         p3         p3



Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       5 ( 50%)
1       3 ( 30%)
2       2 ( 20%)
```

Figure 4.2: Cluster result for subject you like

This is the cluster value for subject CSE engineers good at:

```
Final cluster centroids:
                                        Cluster#
Attribute                  Full Data          0          1          2
                              (14.0)     (12.0)      (1.0)      (1.0)
=============================================================================
engineering                      e5         e5         e3         e5
mathematics                      m5         m5         m4         m2
business-education               b3         b1         b2         b3
media                           md3        md2        md3        md3
health-physical-education        h3         h3         h2         h2
religious-ethical-studies        r3         r3         r3         r4
pure-science                     p4         p4         p4         p3



Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      12 ( 86%)
1       1 (  7%)
2       1 (  7%)
```

Figure 4.3: Cluster result for subject you good

This is the cluster value for interested activities of CSE engineers:

```
Final cluster centroids:
                                        Cluster#
Attribute                  Full Data          0          1          2
                              (10.0)      (3.0)      (5.0)      (2.0)
=============================================================================
creative                         cr5        cr5        cr5        cr5
clerical                         cl3        cl3        cl3        cl4
business                          b0         b3         b0         b3
outdoor                           o3         o3         o3         o5
people-contact                   pc3        pc5        pc3        pc5
practical-mechinacal             pm4        pm5        pm4        pm5
scientific                        s4         s4         s4         s5
social-service                   ss3        ss2        ss3        ss4



Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       3 ( 30%)
1       5 ( 50%)
2       2 ( 20%)
```

Figure 4.4: Cluster result for interested activities

This is the cluster value for interested working field of CSE engineers:

```
                          Cluster#
Attribute         Full Data          0          1          2
                     (12.0)      (7.0)      (2.0)      (3.0)
============================================================
business                  b1         b1         b3         b1
education                ed4        ed4        ed5        ed4
engineering              eg5        eg4        eg5        eg5
environment              en3        en3        en3        en5
health                    h3         h3         h3         h1
law                       l2         l2         l2         l1
media                     m4         m4         m2         m2
science                   s2         s4         s4         s2
social-service           ss4        ss4        ss4        ss3



Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       7 ( 58%)
1       2 ( 17%)
2       3 ( 25%)
```

Figure 4.5: Cluster result for interested working field

This is the cluster value for working style of CSE engineers:

```
Missing values globally replaced with mean/mode

Final cluster centroids:
                           Cluster#
Attribute       Full Data          0          1          2          3
                   (10.0)      (2.0)      (4.0)      (2.0)      (2.0)
=============================================================================
enthusiastic           e4         e4         e3         e4         e3
imaginative           im4        im5        im4        im4        im3
logical                l5         l5         l4         l5         l5
structured            st4        st4        st3        st4        st2



Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       2 ( 20%)
1       4 ( 40%)
2       2 ( 20%)
3       2 ( 20%)
```

Figure 4.6: Cluster result for working style

This is the cluster value for what important for CSE engineers:

```
Missing values globally replaced with mean/mode

Final cluster centroids:
                                Cluster#
Attribute           Full Data         0          1          2
                       (10.0)      (2.0)      (2.0)      (6.0)
============================================================
achievement                a4         a5         a3         a4
challenge                 ch4        ch5        ch3        ch4
creativity                cr5        cr5        cr4        cr5
expertise                 ex3        ex5        ex3        ex4
independence              in5        in5        in3        in5
leading-managing          lm4        lm4        lm5        lm4
responsibility             r5         r4         r4         r5
security                   s4         s4         s3         s4



Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       2 ( 20%)
1       2 ( 20%)
2       6 ( 60%)
```

Figure 4.7: Cluster result for what's important to you

This is the cluster value for thinking style of CSE engineers:

```
Cluster 2: e3,l5,is3,o4,ot4,sp4

Missing values globally replaced with mean/mode

Final cluster centroids:
                                Cluster#
Attribute             Full Data         0          1          2
                         (10.0)      (3.0)      (3.0)      (4.0)
=================================================================
decide-emotionally           e3         e2         e1         e3
logically                    l5         l5         l5         l5
inner-strength              is3        is4        is4        is3
organized                    o3         o3         o4         o3
outgoing                    ot4        ot3        ot5        ot4
spontaneous                 sp4        sp3        sp5        sp4



Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       3 ( 30%)
1       3 ( 30%)
2       4 ( 40%)
```

Figure 4.8: Cluster result for thinking style

This is the cluster value for All other activities of CSE engineers:

		Cluster#		
Attribute	Full Data	0	1	2
	(9.0)	(4.0)	(1.0)	(4.0)
ssc-gpa	4.4489	4.06	3.8	5
hsc-gpa	4.3667	4.05	4.3	4.7
hsc-gpa-math	3.875	3.5938	3.5	4.25
hsc-gpa-physics	4.0625	3.7656	3.5	4.5
hsc-gpa-chemistry	4.3125	3.9531	4	4.75
hsc-gpa-biology	4.5625	4.6406	3.5	4.75
hsc-gpa-english	4.25	4.5625	5	3.75
ssc-gpa-physics	4.5	3.875	5	5
ssc-gpa-chemistry	4.875	4.7188	5	5
ssc-gpa-math	4.875	4.7188	5	5
ssc-gpa-biology	4.875	4.7188	5	5
ssc-gpa-english	4.125	4.0313	5	4
types-of-books-read	science-fiction	science-fiction	story	science-fiction
types-of-movie-watch	science-fiction	action	science-fiction	science-fiction
leisure-time	gaming	gaming	gossiping	r or w
types-of-game	brain	brain	brain	brain
bad-adddiction	no	no	no	no
sleeping-time	7.3333	8	8	6.5
main-subject	mathematics	mathematics	mathematics	mathematics
generating-new-ideas	3.8889	3.5	4	4.25
invent-or-repair-device	4.2222	4.75	5	3.5

Figure 4.9: Cluster result for other activities

4.3 Descriptive Analysis

For our expected result, we apply Weka's pre-processing techniques, zero classification for Classify and Simple K-Means for Clustering.

We chose Simple K-Means because

It's easy to us. Number of cluster is given by manual. May need to scale data. Good initial method.

After using all of techniques of data mining we get 8 different cluster for 8 different categories.

For first category:

Cluster Title: Subject CSE engineers like most are

Characteristics: Which subject CSE engineers like most and which one is least.

1. 50% of CSE engineers mostly like mathematics and Engineering subject. They give rating 5 out of 5 for mathematics and engineering subject.
2. 30% of CSE engineers like pure science (Physics, Chemistry, Biology) with rating 3 to 4.
3. Maximum CSE engineers dislike business study. They give rating for it 1 to 2.

For Second category:

Cluster Title: Subject CSE engineers are good at.

Characteristics: Which subject CSE engineers are good and which one is bad.

1. 86% of CSE engineers are very good at mathematics and Engineering subject. They give rating 5 out of 5 for mathematics and engineering subject.
2. 86% of CSE engineers good at pure science (Physics, Chemistry, Biology) with rating 4.
3. Maximum CSE engineers are not good at business study. They give rating for it 1 to 2.

For third category:

Cluster Title: Interested working activities of CSE Engineers

Characteristics: In which activity CSE engineers like most

1. 80% of CSE engineers mostly like creative activity. They give rating 5 out of 5 for creative activity.
2. 70% of CSE engineers mostly like practical and mechanical activities. They give rating 4 to 5 out of 5 for creative activity.
3. Maximum CSE engineers dislike business activities and social services. They give rating for it 1 to 2.

For fourth category:

Cluster Title: Which working field of CSE Engineers like most.

Characteristics: In which activity CSE engineers like most

1. 58% of CSE engineers mostly like engineering working field. They give rating 5 out of 5 for engineering field.
2. 75% of CSE engineers like educational field. They give rating 4 to 5 out of 5 for educational field.
3. Maximum CSE engineers dislike business field and law. They give rating for it 1 to 2.

For fifth category:

Cluster Title: Working style of CSE Engineers

Characteristics: In which style CSE engineers done their work

1. 60% of CSE engineers working style are logical. They give rating 5 out of 5 for logical working style. 40% engineers give rating 4 for logical working style
2. 20% of CSE engineers working style are imaginative. They give rating 5 out of 5 for imaginative working style.

For sixth category:

Cluster Title: For which important CSE engineers done their job?

Characteristics: What is most important for a CSE engineer.

1. In 60% of CSE engineer's creativity, independence and responsibility are most important. They give rating 5 for creativity, independence and responsibility.
2. In 20% of CSE engineer's challenge, expertise and security are also important. They give rating 4 for challenge, expertise and security.

For seventh category:

Cluster Title: Thinking style of CSE engineers

Characteristics: How a CSE engineers think about a new problem?

1. Around 100% CSE engineers thinking style are logical. They give rating 5 for logically thinking.
2. Maximum CSE engineers give rating for emotionally thinking 1 to 2 .

For eighth category:

Cluster Title: Academic result and other activities?

Characteristics: Academic result and others activity.

1. Academic result of CSE engineers is mixed up with good and bad result.
2. 80% CSE engineer's academic result of mathematics is 4.7 to 5 GPA.
3. 60% CSE engineer's English result is good.
4. Maximum CSE engineer's result of pure science is good.

Others activity:

1. 80% CSE engineer's like science fiction book and movie.
2. Maximum CSE engineer's pass their leisure time by playing game.
3. Maximum CSE engineer's play brain games.
4. Maximum CSE engineer's average sleeping time is 7 and half hours.
5. Maximum CSE engineer's choice mathematics as their main subject.
6. Maximum CSE engineers have no bad addictions.
7. They are expert to generating new idea and invent or repair device.

So, when a student has those characteristics he or she can choose CSE as his or her bachelor subject or CSE related any subject.

4.4 Summary

Behind every story of success, there are many stories of failures. When one becomes success all the pain of being failed vanishes on the glory of success.

The first and main success of this project is to complete within time and with expectations.

To complete this project, we needed to learn many things. Those were also a great success for us. Such as

1. Study about characteristic of several types of person
2. Study about characteristic of CSE engineers
3. Several data mining technique
4. Several algorithms
5. Learn R language
6. Study many research papers

We could find the main common characteristic of CSE engineers successfully.

Our project will help a student to know his or her inner strength.

Applying our project cluster a student can take a decision to choose CSE as his or her graduation subject. It will help to decrease the percentages of drop out student of CSE department.

Beside those success, we learn how to work with a big project and it will us to our next work.

Failures:

The failure of our project was a little. Failures are the pillars of success. Every career prediction might not be fully correct. Someone might be good for one type of career, but it doesn't mean he might not do great at another career. It might be possible that, that person would do great in other career, because of his recent changes. Which he developed after his Higher Secondary Certificate exam. There are so many more cases like this. A person who would be a great business man. Who has all characteristics to become a businessman, but cannot become one because he has no money to start the business. These are the most probable failures of career prediction.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

5.1 Summary of the Study

Career is an occupation which he undertakes for a significant period of time. One person's career starts immediately after joining in work. But before starting career, one person has to take some decisions towards building his career. This decision make big impact on his career. After the HSC exam one student have to take big decision of his career life. He has to choose the subject. Our project mainly works at here. Career prediction helps one student to choose the best subject for his career. There is a huge number of people who are not happy with their career. But they are onto it, because they need money or other reasons. After HSC if a student takes the help of career prediction and provides data, he will surely choose the right subject.

This project was done basically after watching the rate of unhappy persons in career. A person who has great potential in designing, should do designing, not coding. Career prediction helps one to choose the subject which will lead him to his dreamy career. So that after completing graduation and post-graduation a person can become happy with his career.

In this project, we set the attribute first. It took a lot of time to set the attribute. These attributes define the characteristics of an CSE engineer. These attributes are academic background, academic interest, skill, interested working area, personality and others.

After setting the attributes we make the question set regarding the attributes. The answer of this questions defines the characteristics of the person. Then we collect responses of these questions via google form.

After collecting data, we pre-processed it with Weka's default pre-processing algorithms. Then we applied simple K-means and zero classification data mining algorithms and also find the cluster of each attribute.

It was very important and challenging research for us. The biggest challenge was to set the question set. Data collection was also a big challenge. A good data provides good knowledge.

5.2 Conclusions

One person's career is dependent in different kinds of characteristics. One becomes successful if he chooses his career according to his characteristics. We have found eight different characteristics for CSE engineers. If one persons have these characteristics, he will do good in CSE related careers. These characteristics are 1. The subject you like most, 2. The subject you are good at, 3. Interesting activities, 4. Interesting working field, 5. Working style, 6. What is important to you, 7. Thinking style, 8. Academic result and other activities.

The mining knowledge about CSE engineer's characteristics are given below:

1. Subject CSE engineers like: Maximum CSE engineers like math and engineering subjects and dislike business studies.
2. Subject CSE engineers good at: Maximum CSE engineers are good at math and engineering subjects and bad at business studies.
3. Interested activities of CSE engineers: Most of the CSE engineers like creative activity then practical and mechanical activities.
4. Working field CSE engineers like most: Maximum CSE engineers like engineering working fields then they like educational and other fields.
5. Working style of CSE engineers: Most of CSE engineers working style are logical and they are also imaginative.
6. What is important for CSE engineers: For most of the CSE engineer's creativity, independence and responsibility are most important.
7. Thinking style of CSE engineers: Almost 100% of the CSE engineers think logically rather than thinking emotionally.
8. Academic result and other activities of CSE engineers: The academic result is mixed up with bad and good results, but we found that for more than 80% students result in 4.7 out of 5 in mathematics.

Those characteristics are common and important characteristic of a CSE engineer. Because we took data from some well-established CSE engineers. Maximum engineers have those characteristics.

5.3 Recommendations

In this research, we have find out the characteristics of CSE engineers by setting the attributes first. Then we make the question set and collect the responses from qualified and who are successful in their CSE related career. A student gives his data, career prediction compares the characteristics and then gives the result if he is suitable for CSE related careers.

In future if anyone wants to find pattern for other subject related careers, then they can follow our research paper and take idea. At here our recommendation will be that, at first the work should be started with the attribute collection. Then the implementation of the attributes through the question set. Then after collecting the data, data mining algorithms can be applied. There are many algorithms. But for starting Simple K-means and zero classification algorithm is good.

5.4 Implication for Further Study

In this project, we find a pattern for CSE engineers. A student can compare him characteristic with that pattern and easily take decision for choosing CSE as his or her bachelor's subject. There is some implication of further study.

1. We can find the pattern for rest of department. For finding pattern for others department we can use this project as a template.
2. We can develop a system software for this project. Which system take input from user and after mining system shows the characteristic of user and show the perfect department for him or his.
3. Anyone can use our project as a template for their further study.

Reference:

[1] Learn about Career definition, available at <<http://www.businessdictionary.com/definition/career>>, last accessed on 02-05-2017 at 2:10pm.

[2] Learn about Research methodology, available at<<http://www.businessdictionary.com/definition /research-methodology.html>>, last accessed on 02-05-2017 at 2:11pm.

[3] Learn about K-means clustering algorithm, available at << http://stanford.edu/~cpiech/cs221 /handouts/kmeans.html >>, last accessed on 02-05-2017 at 2:11pm.

[4] Learn about Conclusion, available at << https://www.papermasters.com/conclusions _recommendations.html>>, last accessed on 02-05-2017 at 2:12pm.

APPENDIX

1. Give rating to the subjects you like, from 0 to 5.
    - Engineering ( )
    - Math ( )
    - Business Education ( )
    - Media ( )[Hint: Dance, Drama, Music]
    - Health & physical education ( ) [Hint: Health education, Physical education]
    - Religious and ethical studies ( )
    - Pure Sciences ( ) [Hint: Physics, Chemistry]

Figure A1: Question's attributes for subject you like most

2. Give rating to the subjects you are good at, from 0 to 5
    - Engineering ( )
3. Which activities are you interested in? Give rating from 0-5
    - Creative ( )[Hint: You like to express yourself in a creative way or through design, art or music]
    - Business ( ) [Hint: You like to be involved in business relations with people and projects, particularly buying, selling, leading, analyzing]
    - Clerical ( ) [Hint: You like to be involved in tasks requiring accuracy, planning, organizing]
    - Outdoor ( ) [Hint: You like to work outside and be fairly flexible and mobile]
    - People contact ( ) [Hint: You like to work mainly with people in the capacity of teaching, helping advising, caring for or supporting]
    - Practical/Mechanical ( ) [Hint: You like to work mainly with tools, equipment or machinery]
    - Scientific ( ) [Hint: You like to explore, investigate, analyze, experiment and observe]

Figure B2: Question's attributes for subject you are good at

4. Which field would you like to work in? Give rating from 0-5
    - Business ( ) [Hint: Finance, accounting, banking, economics, international business]
    - Educational ( ) [Hint: Teaching- preschool, primary, secondary, disabilities, learning support]
    - Engineering ( ) [Hint: Aerospace, mechanical, electrical, civil, computer, medical, environmental surveying, construction]
    - Environmental ( ) [Hint: Scientific, animals, plants, water, planning and health]
    - Health ( ) [Hint: Nursing, nutrition and dietetics, optometry, podiatry, public health and health management]
    - Law ( ) [Hint: Legal, justice, police, lawyers, solicitors and barristers]
    - Media ( ) [Hint: Film, TV, journalism, creative writing and multimedia]
    - Sciences ( ) [Hint: Biology, microbiology, biomedical, biotechnical, physical sciences, radiography, earth sciences, mathematics and statistics]
    - Social services ( ) [Hint: Psychology, Counselling, human services, ethics, social welfare and community work]

Figure C3: Question's attribute for field of working

5. How would you describe your working style? Give rating from 0-5
    - Enthusiastic ( ) [Hint: You tend to get totally involved and are prepared to take risks in order to get things done as quickly as possible]
    - Imaginative ( ) [Hint: you tend to think things through before jumping in and you can use your imagination to see alternatives and situations]
    - Logical ( ) [Hint: You tend to want to understand tasks and plan well in advance before you work on them]
    - Structured ( ) [Hint: You tend to get on and get things done but in a structured, sensible and responsible way]

Figure D4: Question's attribute for working style

6. What's important to you? Give rating from 0-5
    - Achievement ( ) [Hint: Seek to accomplish tasks and get important things done]
    - Challenge ( ) [Hint: Seek to overcome challenges, difficult problems or situations]
    - Creativity ( ) [Hint: Seek to innovate or create new ideas and be identified with original input]
    - Expertise ( ) [Hint: Seek to attain a high level of expertise in specialist areas]
    - Independence ( ) [Hint: Seek to be independent and make key decisions or do things you way]
    - Leading/managing ( ) [Hint: Seek to be responsible for leading or managing people, projects and resources]
    - Responsibility ( ) [Hint: Seek to be responsible for people and resources and be dependable and trustworthy]
    - Security ( ) [Hint: Seek a solid, secure or predictable future]

Figure E5: Question's attribute for importance to person

7. How do you usually think about things? Give rating from 0-5
    - Decide emotionally ( ) [Hint: You like to decide things with people's feelings in mind and value harmony and praise]
    - Decide logically ( ) [Hint: You like to make decisions logically, without a great deal of emotion]
    - Inner strength ( ) [Hint: You like to work on one or two projects or tasks at a quality level and have one or a few really close friends]
    - Organized ( ) [Hint: You like to have decided well ahead of time be organized, stick to schedules, and be somewhat settled]
    - Outgoing ( ) [Hint: You like to do many things at once, can work easily with medium to large groups of people and like action and variety]
    - Spontaneous ( ) [Hint: You like to stay flexible, avoid fixed plans, and do things unexpectedly or spontaneously]

Figure F6: Question's attributes about usual thinking

8. What is your SSC GPA?
   Ans: (user input)
9. What is your HSC GPA?
   Ans: (user input)
10. What is your GPA in physics in HSC?
    Ans: (user input)
11. What is your GPA in chemistry in HSC?
    Ans: (user input)
12. What is your GPA in math in HSC?
    Ans: (user input)
13. What is your GPA in Biology in HSC?
    Ans: (user input)
14. What is your GPA in English in HSC?
    Ans: (user input)
15. What is your GPA in Physics in SSC?
    Ans: (user input)
16. What is your GPA in Chemistry in SSC?
    Ans: (user input)
17. What is your GPA in Math in SSC?
    Ans: (user input)
18. What is your GPA in Biology in SSC?
    Ans: (user input)
19. What is your GPA in English in SSC?
    Ans: (user input)

Figure G7: Question's attributes for academic result

Other activities

16. What types of books do you read?
    Ans: 1. Story, 2. Science Fiction, 3. Detective, 4. Mathematics (ans:)
17. What types of movies do you watch?
    Ans: 1. Action, 2. Animation, 3. Love, 4. Detective, 5. Science Fiction(Sci-Fi) (ans:)
18. What do you do at your leisure time?
    Ans: (user input)
19. What types of game do you like?
    Ans: 1. Action 2. Brain 3. Mission 4. Racing 5. (ans:)
20. Do you have any kind of bad addiction?
    Ans: (user input)
21. How much you sleep a day?
    Ans: 1. 4-6 hours, 2. 6-8 hours, 3. 8-10 hours, 4. More than 10 hours.
22. In which extra curriculum activity, you are so strong?
    Ans: (User Input)
23. From given option which subject you will select as your main subject?
    Ans: 1. Mathematics 2. Biology 3. ICT 4. Chemistry
24. I can generate new ideas. Give rating from 0 to 5 ( )
25. Do you like to invent or repair any device? Give rating from 0 to 5( )

Figure H8: Question's attributes for other activities

1.-subject-you-like.arff - Notepad

File Edit Format View Help

```
@relation the-subject-you-like
@attribute engineering {'e1','e2','e3','e4','e5','e0'}
@attribute mathematics {'m1','m2','m3','m4','m5','m0'}
@attribute business-education {'b1','b2','b3','b4','b5','b0'}
@attribute media {'md1','md2','md3','md4','md5','md0'}
@attribute health-physical-education {'h1','h2','h3','h4','h5','h0'}
@attribute religious-ethical-studies {'r1','r2','r3','r4','r5','r0'}
@attribute pure-science {'p1','p2','p3','p4','p5','p0'}

@data
'e5', 'm4', 'b1', 'md2','h1','r3','p4'
'e5', 'm4', 'b0', 'md0','h3','r2','p3'
'e4', 'm3', 'b3', 'md3','h3','r3','p4'
'e5', 'm4', 'b1', 'md2','h1','r3','p4'
'e5', 'm4', 'b1', 'md2','h1','r3','p4'
'e5', 'm3', 'b0', 'md0','h3','r2','p3'
'e4','m4','b3','md4','h2','r5','p3'
'e4','m4','b3','md4','h4','r4','p4'
'e5','m4','b4','md4','h5','r4','p4'
'e5','m4','b3','md4','h3','r4','p3'
'e4','m5','b0','md0','h3','r3','p4'
'e5','m5','b3','md1','h3','r2','p5'
'e5','m5','b1','md0','h3','r2','p1'
'e5','m5','b3','md2','h5','r5','p3'
'e5','m3','b3','md3','h3','r4','p4'
'e4','m4','b1','md1','h1','r2','p1'
'e5','m5','b3','md1','h2','r3','p4'
'e5','m4','b1','md4','h3','r3','p2'
'e3','m3','b2','md2','h0','r0','p2'
```

Figure I9: Dataset for subject you like

2.-subject-you-good.arff - Notepad

File Edit Format View Help

```
%
%
@relation the-subject-you-good
@attribute engineering {'e1','e2','e3','e4','e5','e0'}
@attribute mathematics {'m1','m2','m3','m4','m5','m0'}
@attribute business-education {'b1','b2','b3','b4','b5','b0'}
@attribute media {'md1','md2','md3','md4','md5','md0'}
@attribute health-physical-education {'h1','h2','h3','h4','h5','h0'}
@attribute religious-ethical-studies {'r1','r2','r3','r4','r5','r0'}
@attribute pure-science {'p1','p2','p3','p4','p5','p0'}

@data
'e5', 'm5', 'b1', 'md2','h1','r3','p4'
'e5', 'm5', 'b0', 'md0','h3','r2','p3'
'e4', 'm5', 'b3', 'md3','h3','r3','p4'
'e5', 'm5', 'b1', 'md2','h1','r3','p4'
'e5', 'm5', 'b1', 'md2','h1','r3','p4'
'e5', 'm5', 'b0', 'md0','h3','r2','p3'
'e4', 'm5', 'b3', 'md3','h3','r3','p4'
'e5', 'm5', 'b1', 'md2','h1','r3','p4'
'e5', 'm5', 'b0', 'md0','h3','r2','p3'
'e4', 'm5', 'b3', 'md3','h3','r3','p4'

'e4','m4','b4','md5','h3','r5','p3'
'e4','m5','b3','md4','h4','r4','p4'
'e5','m2','b3','md3','h2','r4','p3'
'e3','m4','b2','md3','h2','r3','p4'
'e4','m5','b2','md1','h3','r3','p4'
'e2','m3','b0','md1','h0','r0','p1'
'e5','m5','b3','md2','h5','r5','p2'
'e3','m3','b0','md1','h1','r2','p2'
'e4','m4','b2','md0','h1','r1','p2'
'e4','m3','b1','md2','h1','r2','p1'
'e2','m2','b0','md0','h0','r0','p2'

%
%
```

Figure J10: Dataset for subject you are good at

3.-which-activities-are-you-intersted-in.arff - Notepad

File Edit Format View Help

```
%
%
@relation interest-activities-working-are

@attribute creative {'cr1','cr2','cr3','cr4','cr5','cr0'}
@attribute clerical {'cl1','cl2','cl3','cl4','cl5','cl0'}
@attribute business {'b1','b2','b3','b4','b5','b0'}
@attribute outdoor {'o1','o2','o3','o4','o5','o0'}
@attribute people-contact {'pc1','pc2','pc3','pc4','pc5','pc0'}
@attribute practical-mechinacal {'pm1','pm2','pm3','pm4','pm5','pm0'}
@attribute scientific {'s1','s2','s3','s4','s5','s0'}
@attribute social-service {'ss1','ss2','ss3','ss4','ss5','ss0'}

@data
'cr5','cl3','b0','o3','pc3','pm4','s4','ss3'

'cr5','cl3','b3','o3','pc5','pm5','s4','ss2'
'cr5','cl4','b3','o5','pc5','pm5','s5','ss4'
'cr5','cl4','b3','o5','pc5','pm5','s5','ss5'
'cr4','cl3','b2','o4','pc5','pm4','s4','ss4'
'cr5','cl3','b0','o3','pc3','pm4','s4','ss3'
'cr5','cl3','b0','o3','pc3','pm4','s4','ss3'
'cr5','cl3','b0','o3','pc3','pm4','s4','ss3'
'cr5','cl3','b0','o3','pc3','pm4','s4','ss3'
'cr5','cl3','b3','o3','pc5','pm5','s4','ss2'
```

Figure K11: Dataset for activities interested in

8.-rest-all-attributes.arff - Notepad

File Edit Format View Help

```
@relation rest-of-all-attribute
@attribute  ssc-gpa numeric
@attribute  hsc-gpa numeric
@attribute  hsc-gpa-math numeric
@attribute  hsc-gpa-physics numeric
@attribute  hsc-gpa-chemistry numeric
@attribute  hsc-gpa-biology numeric
@attribute  hsc-gpa-english numeric
@attribute  ssc-gpa-physics numeric
@attribute  ssc-gpa-chemistry numeric
@attribute  ssc-gpa-math numeric
@attribute  ssc-gpa-biology numeric
@attribute  ssc-gpa-english numeric
@attribute  types-of-books-read {'story','science-fiction','detective','mathematics','religious'}
@attribute  types-of-movie-watch {'action','animation','love','science-fiction'}
@attribute  leisure-time {'gaming','studing', 'nothing','movie','servicing','gossiping','r or w','sleeping'}
@attribute  types-of-game  {'action','brain','mission','racing'}
@attribute  bad-adddiction {'yes','no'}
@attribute  sleeping-time numeric
@attribute  main-subject{'mathematics','biology','ict','chemistry'}
@attribute  generating-new-ideas numeric
@attribute  invent-or-repair-device numeric

@data
'4.44','3.8','3.5','3.5','4','5','5','3.5','5','5','5','4','science-fiction','science-fiction','gaming','brain','no','8','mathematics','3','5'

'5','4.9','4','5','5','5','3.5','5','5','5','5','3.5','mathematics','science-fiction','r or w','brain','no','6','mathematics','5','3'


'5','4','4','3','4','4','4','5','5','5','5','4','science-fiction','science-fiction','sleeping','racing','no','6','mathematics','3','4'

'3.8','4.3','3.5','3.5','4','3.5','5','5','5','5','5','5','story','science-fiction','gossiping','brain','no','8','mathematics','4','5'

'5','5','5','5','5','5','4','5','5','5','5','5','science-fiction','action','movie','racing','yes','8','mathematics','4','4'


'3.3','4.7',?,?,?,?,?,?,?,?,?,?,'science-fiction','action','gaming','brain','no','8','ict','4','4'
```

Figure L12: Dataset for rest of all attributes

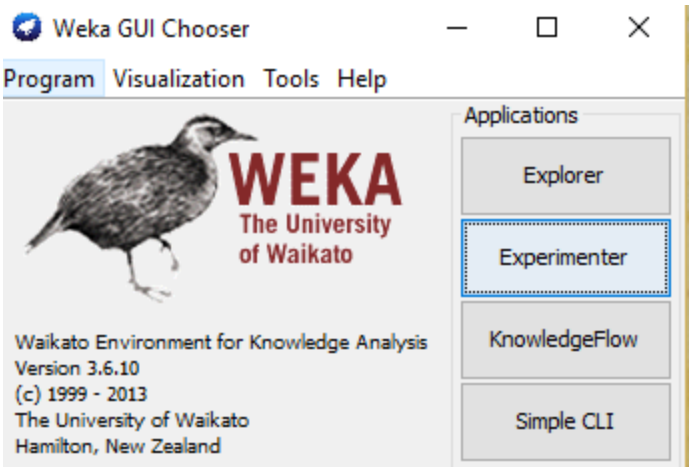


Figure M13: Weka GUI

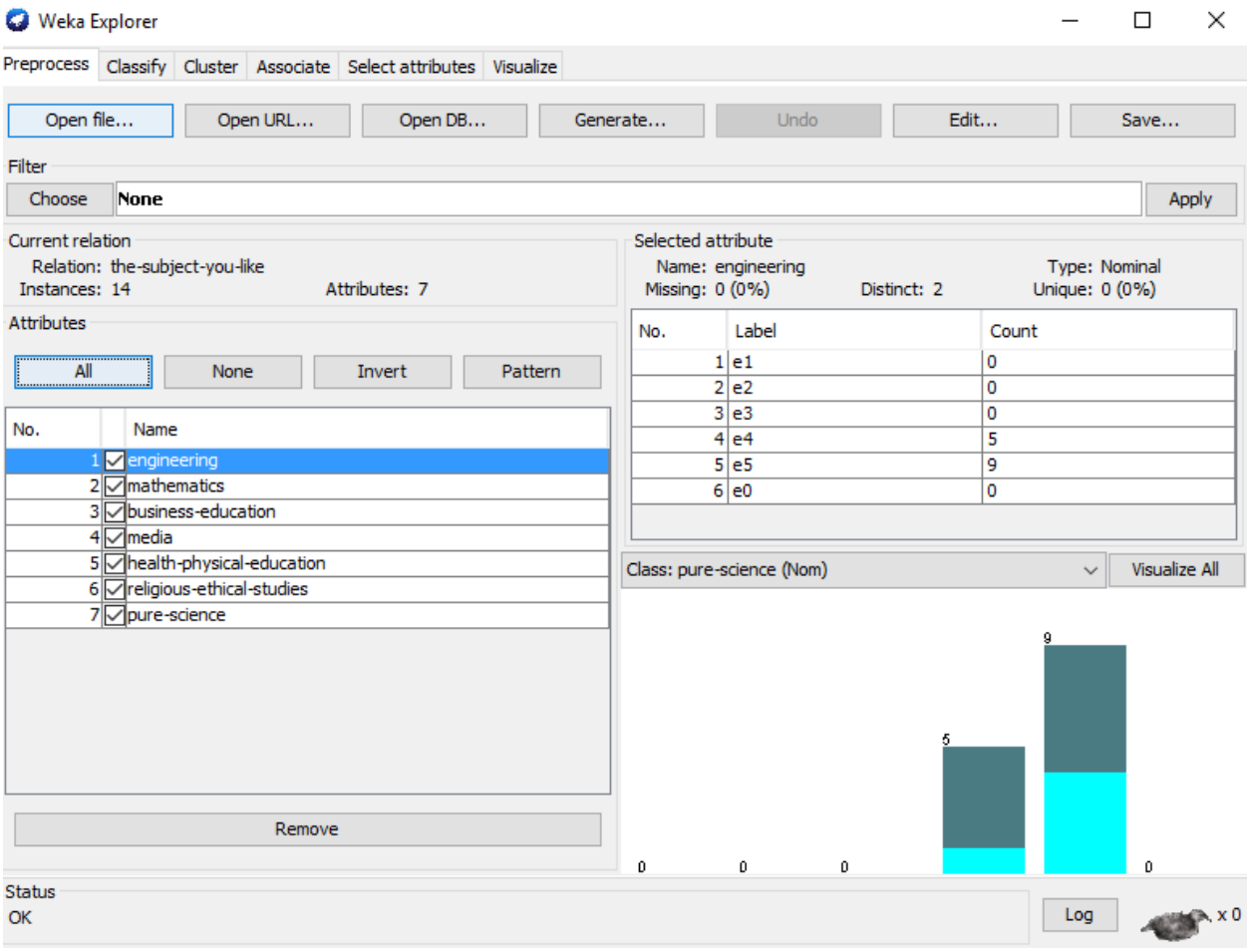


Figure N14: Preprocessing

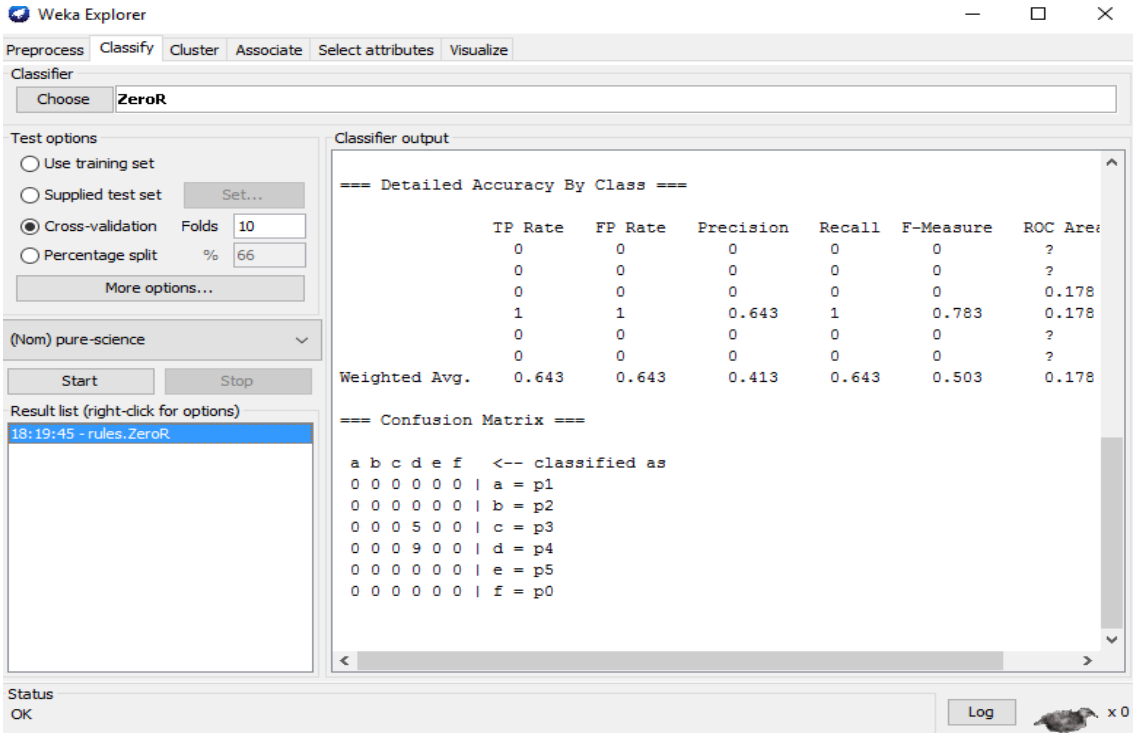


Figure O15: Classifying

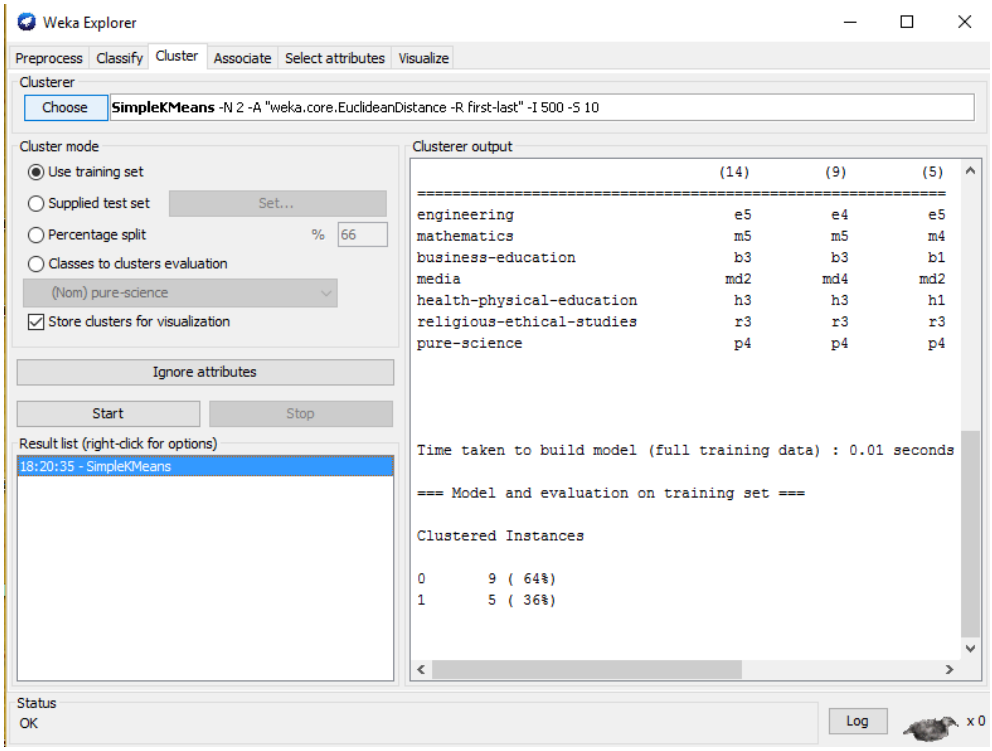


Figure P16: Clustering

For first category:

Cluster Title: Subject CSE engineers like most are

Characteristics: Which subject CSE engineers like most and which one is least.

1. **50% of CSE engineers mostly like mathematics and Engineering subject. They give rating 5 out of 5 for mathematics and engineering subject.**
2. **30% of CSE engineers like pure science (Physics, Chemistry, Biology) with rating 3 to 4.**
3. **Maximum CSE engineers dislike business study. They give rating for it 1 to 2.**

Figure Q17: Result for attribute subject you like most

For Second category:

Cluster Title: Subject CSE engineers are good at.

Characteristics: Which subject CSE engineers are good and which one is bad.

1. **86% of CSE engineers are very good at mathematics and Engineering subject. They give rating 5 out of 5 for mathematics and engineering subject.**
2. **86% of CSE engineers good at pure science (Physics, Chemistry, Biology) with rating 4.**
3. **Maximum CSE engineers are not good at business study. They give rating for it 1 to 2.**

Figure R18: Result for attribute Subject you good at

For third category:

Cluster Title: Interested working activities of CSE Engineers

Characteristics: In which activity CSE engineers like most

1. **80% of CSE engineers mostly like creative activity. They give rating 5 out of 5 for creative activity.**
2. **70% of CSE engineers mostly like practical and mechanical activities. They give rating 4 to 5 out of 5 for creative activity.**
3. **Maximum CSE engineers dislike business activities and social services. They give rating for it 1 to 2.**

Figure S19: Result for attribute Interested working activities

For fourth category:

Cluster Title: Which working field of CSE Engineers like most.

Characteristics: In which activity CSE engineers like most

1. **58% of CSE engineers mostly like engineering working field. They give rating 5 out of 5 for engineering field.**
2. **75% of CSE engineers like educational field. They give rating 4 to 5 out of 5 for educational field.**
3. **Maximum CSE engineers dislike business field and law. They give rating for it 1 to 2.**

Figure T20: Result for attribute Working field

For fifth category:

Cluster Title: Working style of CSE Engineers

Characteristics: In which style CSE engineers done their work

1. **60% of CSE engineers working style are logical. They give rating 5 out of 5 for logical working style. 40% engineers give rating 4 for logical working style**
2. **20% of CSE engineers working style are imaginative. They give rating 5 out of 5 for imaginative working style.**

Figure U21: Result for attribute Working style

For sixth category:

Cluster Title: For which important CSE engineers done their job?

Characteristics: What is most important for a CSE engineer.

1. **In 60% of CSE engineer's creativity, independence and responsibility are most important. They give rating 5 for creativity, independence and responsibility.**
2. **In 20% of CSE engineer's challenge, expertise and security are also important. They give rating 4 for challenge, expertise and security.**

Figure V22: Result for attribute Most important for engineers

For seventh category:

Cluster Title: Thinking style of CSE engineers

Characteristics: How a CSE engineers think about a new problem?

1. **Around 100% CSE engineers thinking style are logical. They give rating 5 for logically thinking.**
2. **Maximum CSE engineers give rating for emotionally thinking 1 to 2.**

Figure W23: Result for attribute thinking style

For eighth category:

Cluster Title: Academic result and other activities?

Characteristics: Academic result and others activity.

1. **Academic result of CSE engineers is mixed up with good and bad result.**
2. **80% CSE engineer's academic result of mathematics is 4.7 to 5 GPA.**
3. **60% CSE engineer's English result is good.**
4. **Maximum CSE engineer's result of pure science is good.**

Figure X24: Result for attribute academic result

Others activity:

1. **80% CSE engineer's like science fiction book and movie.**
2. **Maximum CSE engineer's pass their leisure time by playing game.**
3. **Maximum CSE engineer's average sleeping time is 7 and half hours.**
4. **Maximum CSE engineer's choice mathematics as their main subject.**
5. **They are expert to generating new idea and invent or repair device.**

Figure Y25: Result for attribute other activities