

**IMPACT OF SOFT DRINKS ON HEALTH TO THE LOCAL
PEOPLE OF BANGLADESH BASED ON DATA MINING
TECHNIQUES**

BY

**Mehadi Hasan Sarkar
ID: 161-15-7624**

**Nowrin Tasnin Sinthia
ID: 161-15-7560**

**Mosaddak Rahman Bhuiyan
ID: 161-15-7024**

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

Supervised By

Anup Majumder

Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

Mr. Shah Md. Tanvir Siddiquee

Senior Lecturer

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

December 2019

APPROVAL

This Project title is “**Impact of soft drinks on health to the local people of Bangladesh based on data mining techniques**”, submitted by Mehadi Hasan Sarkar ID: 161-15-7624, Nowrin Tasnin Sinthia ID: 161-15-7560 and Mosaddak Rahman Bhuiyan ID: 161-15-7024 to the Department of Computer Science and Engineering, Daffodil International University, have been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on First Week of December 2019.

BOARD OF EXAMINERS



Dr. Syed Akhter Hossain
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Chairman



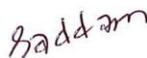
Md. Sadekur Rahman
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



Abdus Sattar
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



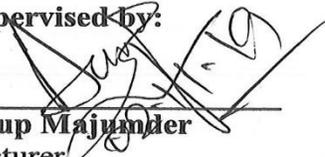
Dr. Md. Saddam Hossain
Assistant Professor
Department of Computer Science and Engineering
United International University

External Examiner

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Anup Majumder**, Lecturer of department of CSE in 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:


Anup Majumder
Lecturer
Department of CSE
daffodil International University

Co-Supervised by:

Shah Md. Tanvir Siddiquee
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:


Mehadi Hasan Sarkar
ID: 161-15-7624
Department of CSE
Daffodil International University


Nowrin Tasnin Sinthia
ID: 161-15-7560
Department of CSE
Daffodil International University


Mosaddak Rahman Bhuiyan
ID: 161-15-7024
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

In the beginning we want to thanks and gratitude to Almighty Allah. It was totally impossible to complete this project without the blessing of Allah.

We are really grateful to Anup Majumder, Lecturer, department of CSE, Daffodil International University, Dhaka. We are inspired to see the enough knowledge and interest of our supervisor in the department of software development. His supervision, advising, patience, friendly attitude, encouragement, constructive criticism really helps us to complete this project successfully. We would like to thank to **Professor Syed Akhter Hossain**, head, department of CSE, daffodil International University, for his encouragement. We also want to thanks to other course teacher and all the staffs of CSE department, Daffodil International University.

ABSTRACT

In this world every life has different nurture. Human are not opposite from that. From the beginning of life man also take changes for stay in this world like other. Those changes have effect on lifestyle. Some effect is good and some are bad. The change of food habits is one of them. The modern food habits are a threat to healthy life. Among other, soft drinks consumption is a huge part of modern food habits. The children are the most targeting part of it. The quality and the ingredient of soft drink has its effect in health heavily. The impact of it on human health is become a subject of depth research. Consumption of soft drink is playing a major part in various diseases like diabetes, obesity, dental, bon disorders and other child diseases. The toxic effect of it has gathered a much attention because of scientific reports and media. So, the objective of this resource is to provide a perfect look of the impact of soft drink on human health and also suggest its alternatives for healthy lifestyle. The enormous chemical load in the environment. Through modernization, industrialization and lifestyle changes. They always prefer to take soft drinks instead the nutritious foods such as fruits, vegetables, and milk products. The quality of the soft drinks is completed with the type and quantity of huge percent of chemicals, including those present water used for such preparations. The impact of soft drinks on human health has become the main subject of intense research. Taking soft drinks plays always role in obesity, diabetes, various types of dental diseases and one disorders and more, among children and adolescents. The toxic effects of soft drinks have gained much attention, due to the frequent scientific reports and media attention. The purpose of this review is to pick up a comprehensive investigation on the effects of soft drinks. To suggest health, as well as healthy living options.

TABLE OF CONTENTS

CONTENTS	PAGE
Approval	I
Declaration	Ii
Acknowledgements	Iii
Abstract	Iv
CHAPTER 1: INTRODUCTION	1-3
1.1 Introduction	1
1.2 Motivation	2
1.3 Objectives	3
1.4 Expected Outcome	3
1.5 Report Layout	3
CHAPTER 2: BACKGROUND STUDY	5-8
2.1 Introduction	5
2.2 Literature Review	6
2.3 Comparative Studies	8
2.5 Challenges	8
CHAPTER 3: METHODOLOGY	10-13
3.1 Introduction	10
3.2 Research Methodology	10
3.3 Data Flow Diagram	13
CHAPTER 4: DESIGN SPECIFICATION	14-20
4.1 Front-end Design	14
4.2 Back-end Design	14
4.3 Experimental Result	19
4.4 Implementation Requirements	20

CHAPTER 5: IMPLEMENTING AND TESTING	21-26
5.1 User Registration with Answer	21
5.2 Implementation of Database	22
5.3 Training Data	24
5.4 Testing Implementation	25
5.5 Acceptance Testing	25
5.6 Test Results and Reports	26
5.7 System Testing	26
CHAPTER 6: CONCLUSION AND FUTURE SCOPE	27
6.1 Discussion and Conclusion	27
6.2 Future plan	27
REFERENCES	28

LIST OF FIGURES

FIGURES	PAGE
Figure 1: Sex Structure of all collection	15
Figure 2: Age collection	16
Figure 3: Preference of soft drinks	16
Figure 4: Amount of drinking structure	17
Figure 5: Data flow diagram of whole system	17
Figure 6: Back-end design	18
Figure 7: Database of stored data	18
Figure 8: Matrix of dataset in YML file	19
Figure 9.1: Implementation of Database	23
Figure 9.2: Implementation of Database	23
Figure 9.3: Implementation of Database	24
Figure 9.4: Implementation of Database	24
Figure 10: Process of training data	25
Figure 11: Testing data	25

List of Table

Table	Page
Table.1: simple database of stored data	14
Table 2: Used methods	19
Table 3: Database of stored data	21

CHAPTER 1

INTRODUCTION

1.1 Introduction

In machine learning approach defines a data mining has successfully been applied to analyzing and taking decision on huge amount of data which are collected from different area. It has several applications such as data streams, generation of trajectory data, education system data mining for future vision data mining. Here we demonstrate impact of soft drinks and how it's becoming serious issues for health and the negative side on human body. Demand of soft drinks is growing a big problem for research area and it has spacious applications in the ground increasing, identification and verification of a person's condition for drinking it, by this technique identification and verification for a affected or detected people and a people who are not affected by this drinks, the variation of those. We have focused on the negative impact of this and with the technique of data mining we find out the growing steps of it and the solution and rules regulation for taking it.

1.2 Motivation

Data mining technology is being used in many sectors like researchers use data mining approaches like multi-dimensional databases, machine learning, soft computing, data visualization and statistics. Mining can be used to predict the volume of patients in every category. We want to use data mining for collect data and try to measure the negative effect of soft drinks on human body for long term take it. The negative impact of different types drinks became the very big issues all over the world specially this sub-continent. Diabetes Heart diseases, bone and teeth diseases are frequently induced due to excessive consumption of soft drinks particularly children and aged person. We know our Health ministry has already taken many necessary steps to aware peoples about this silent problem but this problem becoming the big headache for our upcoming generation in our country.

Diseases mentioned above induce a negative impact of our country chiefly less developing country. Our proposed system is capable to mention that this types of soft drinks is harmful for health and it's like the major and serious upcoming issue in Bangladesh for which we have to pay a lot.

As for the purpose, we are extremely motivated to establish a system that measure the negative effect of this on body and which is becoming cause for our death, and also we can say that it is the very conscious health issue for us.

1.3 Objectives

Our main concern about is to measure the amount of negative impact on body for taking regular soft drinks and mention the cycle of the steps of how it's works.

We are attempting to construct a Data mining system that identify the diseases based on taking the amount of soft drinks very quickly from regular backgrounds. We want to minimize the negative effects of this and aware to local people that the dangerous impact of it.

The excess drinking of carbonated soft drinks harms the teenage population because it contains no vital nutrients and harms their general and oral health. At one time the data mining technique is done; our following intent is to train our system with more criteria of diseases. By using Local Binary Patterns for each image a feature vector is to be computed.

By adopting this proposal, we can raise public awareness about negative effect of soft drinks and also increase consciousness of taking the rate of it.

By the establishment of our system we want to give a useful message to them who are addicted to drink different types of soft drinks.

Moreover, our main concern is to achieve a control and effective system and save the valuable lives of different ages people specially youth.

1.4 Expected Outcomes

This report is access by quality a board range is issues which is relation to soft drinks using up to resource and which is reflecting the information and the literature available. The report is not a complete review but aims to raise awareness of some of the border issues associated with reducing the consumption of soft drinks. It provides an overview of current knowledge surrounding the relationship between soft drink consumption and weight status and other health implications, and reports on the nature and extent of soft drink consumption. The system saves the valuable life specially the children who are the victim for this. To develop our country our proposal would work efficiently. It completely reduces tendency of taking soft drinks which are generated by the system as well as it helps to teach them about awareness.

1.5 Report Layout

Chapter 1: Introduction- We explain here about The introduction, Objective, motivation of the work, expected outcome and also the layout of the report of our project.

Chapter 2: Background Study- Background part of our proposed method is discussed in chapter 2. We also explain literature review, related works, argumentation of the problems and claims about our project in this portion.

Chapter 3: Methodology- In this chapter we include overall procedures that we have used to build this proposed system. Methods are explained here step by step.

Chapter 4: Design Specification- Here the section is for to discuss about the graphical interface and display to the users.

Chapter 5: Implementation and Testing- Our empirical results are shown along with the performance analysis that is achieved by the proposed system. We have also covered results summary in this chapter.

Chapter 6: Conclusions and Future plans- In this part of the report we have discussed about conclusion and future aims of our proposed system for general people.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction

Impact of soft drinks on human body based on Data Mining Technique where we have developed the system by using six very important theorem of Data Mining and where at first, we used Ten cross validation method to start the first step. It enables to accessing how the result of a statistical analysis will generalize to an independent data set. Here it is mainly used in settings where the goal is prediction, and one want to estimate how accurately a predictive will perform in practice. This Data Mining process is categorized into two steps which are Raw data collection and the second one is data preprocessing. Its main concern is to raise awareness of this system of what the negative effect of this Drinks if someone takes it regularly or takes it with a break.

This technique generally holds some rules like collecting the Raw data from the root level. And then arrange the data and preprocessing the data. After preprocess the data Split data set and then Data Training and Data Testing and Finally train the data Algorithm. We use a larger and better training sample to improve this system and to get the accurate result.

This Data Mining Technique can be done by following those Data Mining Steps. Data Mining is very convenient and simple to perform than the other technique to summarize the Big Data.

This paper deals with the impact on health by taking soft drinks by using six algorithms and identify that which algorithm is more effective to give the accurate result. And also give from those algorithms which classifier and describes the accuracy by local patterns.

2.2 Literature Review

Charles Spence, Xiaogang Wan [1] suggested that in our daily lives we usually take drinks from mirrors, cups, cups, cans, bottles, and straws. In this article, we consider the effect that the physical and easy-to-consider nature of drinks may have on people's perception of content. The same color of the drink sometimes sets the taste expectations completely depending on the type of deposit and the cultural background of the participants. He recently led a lab study among participants in the United Kingdom where they showed that social drinkers use beer 60% slower than 12 FL Oz full glasses from the straight side than from a volume flute vessel with curved sides. Interestingly, no such effects were observed in participants using carbonated soda (7-Up) from the same pair of glasses, suggesting a rather complex interaction between glassware and content. Containers or containers have a wide range of drinks, from wine to coffee and from hot chocolate to soft drinks and beer.

Takehiro Yamakoshi¹, Kenta Matsumura¹ and Shota Hanaki² & Peter Rolfe^{3, 4} [2] referred that Increasing the ratio of healthier foods available could cheer up healthier consumption, but witness to date is limited in opportunity and quality. Unhealthy mood forms, including excess fuel intake, are the root of non-infectious rags. Contributor, which is currently the cause of most untimely deaths worldwide. Intervening to increase the ratio of lower energy options obtainable in organizational settings may be a hopeful strategy to minimize energy bought and in turn consumed. Increasing the ratio of lower energy food options obtainable in worksite cafeterias – while aiming to keep the total number of options constant may minimize energy purchased, without affecting revenue. While larger studies are warranted to exactly estimate the impact of altering healthier product availability, this intervention could be thought in a wide range of organizational settings to promote healthier food environments. Given that impact varied by site, future research should establish the most effective ways of implementing this hopeful intervention to ensure that potential benefits to both public health and employee health are maximized.

Penny Breeze^{1*}, Robert Womack², Robert Pryce¹, Alan Brennan¹, Elizabeth Goyder¹ [3] Proposed that Evaluation of the impact of a local sugar sweetened beverages (SSB) health prosperity and 20p price increase in leisure center meeting places and estimate the impact on consumption. Demand for SSB drinks at leisure center meeting places is highly

respondent to the policy, especially for child-friendly and high caffeine energy drinks, compared with other SSB tax policy evaluations. The policy also increased purchases of carbonated non-SSB. The impact of obesity on associated chronic diseases has led to calls for a comprehensive public health approach to tackling the current trends in physical inactivity and high calorie diets. Leisure centers present an opportunity for broad health education and promotion influencing positive behavior changes both in their service users and in their local communities. Model outcomes included all cold drink sales per attendance, SSB sales per attendance, non-SSB sales per attendance, non-SSB sales per attendance (excluding water and juice), non-SSB fruit juice, non-SSB water and volume of SSB sold per attendance. Price elasticity measures the responsiveness of the quantity demanded of a good to a change in its price. Price elasticity is the ratio between the percentage change in the quantity demanded, q , and the corresponding percent change in price, p .

Lenny R. Vartanyan, Ph.D., Marlene B. Schwartz, Ph.D., and Kelly D. Brownell, Ph.D. [4], suggested that soda consumption was very noticeable and was the focus of public health and public policy. Soft drinks are prohibited in schools in the UK, France, and the USA.

Don Zhuravin [5] suggested that the FDA allows the food and beverage industry to create natural orange juice nutrients during packaging. Your body will receive 10 tablespoons of sugar and 1 ounce of fat. Consumption of natural orange juice is 2-3 times higher than that of orange juice. In clinical studies, the vitamin C content of NOW reconstituted frozen orange juice was compared to freshly squeezed orange juice. Vitamin C in the form of synthetic ascorbic acid exceeds the content of vitamin C in natural orange juice in terms of the content of ascorbic acid in the blood.

Casey M. Morgan E. Gram, Lynn M.S. Deirdre was her certificate. Cruz, Cheryl O.M., Anderson, Lydia A., Joseph Kurs, and Lawrence J. Apple [6] are more widespread and related to glucose metabolism, diabetes, and pancreatic waste. There are several mechanisms through which baking soda can damage urine. In soda (both diet and habitat) phosphorus is accompanied by color and taste. Phosphorus may affect serum phosphorus

content and the growth rate of fibroblast. Soda diet can increase dietary acidity due to its phosphorus content.

This increases the weight and hence the risk of kidney disease. The dose is related to dietary intake and the risk of renal failure at the final stage.

2.3 Comparative Studies

The variable results have a significant impact on the methods used in the studies on the impact of soft drinks consumption and on some of these methodological factors on the results of the studies. The results of the studies suggest different research methods and definitions of key variables such as body weight. Further research would help to clarify a more consistent methodology (ideally experimental schemes) to determine the impact of soft consumption on diet and health outcomes.

2.4 Challenges

Our project “Impact of soft drinks on the health to the local people based on data mining techniques” is quite challenging work for us. In our system we have to survey among local people as well as show all the saved data about them which is very challenging for us. Requirement Creating a dataset by collecting data was trained firstly the people who drink soft drinks to know approximate amount. Showing exact proper information is so challenging that we had needed the proper information for collecting. Server controlling which contains a lot of information was much challenging for us.

Time Scheduling

Time scheduling refers a set of techniques used to develop and offer schedules that says when the whole work will be done. Our main concern was about time scheduling to complete this project in the meantime. If the experiment was not prepared on time it will be a big trouble for us. We have divided our working time and project work among us to complete the whole projects.

Cost reducing

Every decision in project development affects cost so it was another challenging task to reduce cost and increase our profits.

Increasing Communication

During the time of developing our project we had faced many questions. For achieving solutions, we communicated with supervisor.

Skills for the projects

Required skills to complete our whole project work were taken properly.

Chapter 3

RESEARCH METHODOLOGY

3.1 Introduction:

In previous chapter we discussed about the related works of data mining system. We have also explained about some difference types of procedure that were used in pervious. In this chapter we will discuss about our procedure that we have used to develop this system. When we study about related works of our system then we came to know about some difference algorithm which can be used for developing this technique and measure the system. After that study we decided to select a specific algorithm. We will discuss briefly about our used algorithms in the net part of this chapter.

3.2 Methodology:

Here, we are working with six algorithms

- A. Naive Bayes
- B. Multilayer Perceptron
- C. Logistic
- D. K-Star
- E. J48
- F. Random Forest

A. Naïve Bayes: Based on Naïve Bayes technique this classification technique has been used on this work. This technique assume that the presence of any other feature is unrelated with the particular feature in a class. This model is easy to build and also particularly useful for large data and also very large data sets. With the simplicity of Naïve Bayes it is known to outperform and also highly sophisticated with its classification methods. This method is very easy and fast to predict the class of test for data set. It is also performing well where there is multi class prediction. At the time of assumption of independent it holds, a Naïve Bayes classifier performing better with compare to the other models as like logistic regression and you need for it less data.

There are four types of classification

1. Real time prediction
2. Multi class prediction
3. Test classification
4. Recommendation system

B. Multilayer Perceptron: A multilayer perceptron is the feed forward neural network which create a set of outputs from a set of inputs. The MLP is characterized by many layers of input nodes linked as a directed graph between the input and output layers. MLP is a method of specific learning.

C. Logistic: Logistic is another technique use by machine learning from the statistics field. It is the best method for classify binary problems. Logistic is named for the function used at the base of the method, the logistic function.

There is the step for preparing data for logistic.

1. Binary Output Variable.
2. Remove Noise
3. Gaussian Distribution.
4. Remove Correlated Inputs.
5. Fail to Converge.

D. K-star: K-star is a classifier that is the class of a test instance is used upon the class of those training items similar to it as determined by some similarity function. It creates from other instance-based learners in that it uses an entropy-based distance function.

E. J48: Classification is the process of building a model of classes from a set of records that contain class labels. Decision Tree Algorithm is to find out the way the attributes-vector behaves for a number of instances. In the WEKA data mining tool, J48 is an open source Java implementation of the C4.5 algorithms.

F. Random Forest: The random forest is a classification algorithm consisting with many decision trees. It uses bagging and feature randomness when building each individual

tree to try to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree.

For all those algorithms we have to find out

1. Accuracy
2. Sensitivity
3. Specificity
4. Precision
5. False Positive Rate
6. False Negative Rate

1. Accuracy: It's mean to the closeness of mathematical value and the standard value.

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN}$$

2. Sensitivity: The actual positive part of the calculation.

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

3. Specificity: The actual Negative part of the calculation.

$$\text{Specificity} = \frac{TN}{TN+FP}$$

4. Precision: It's mean to the closeness of mathematical value to each other.

$$\text{Precision} = \frac{TP}{TP+FP}$$

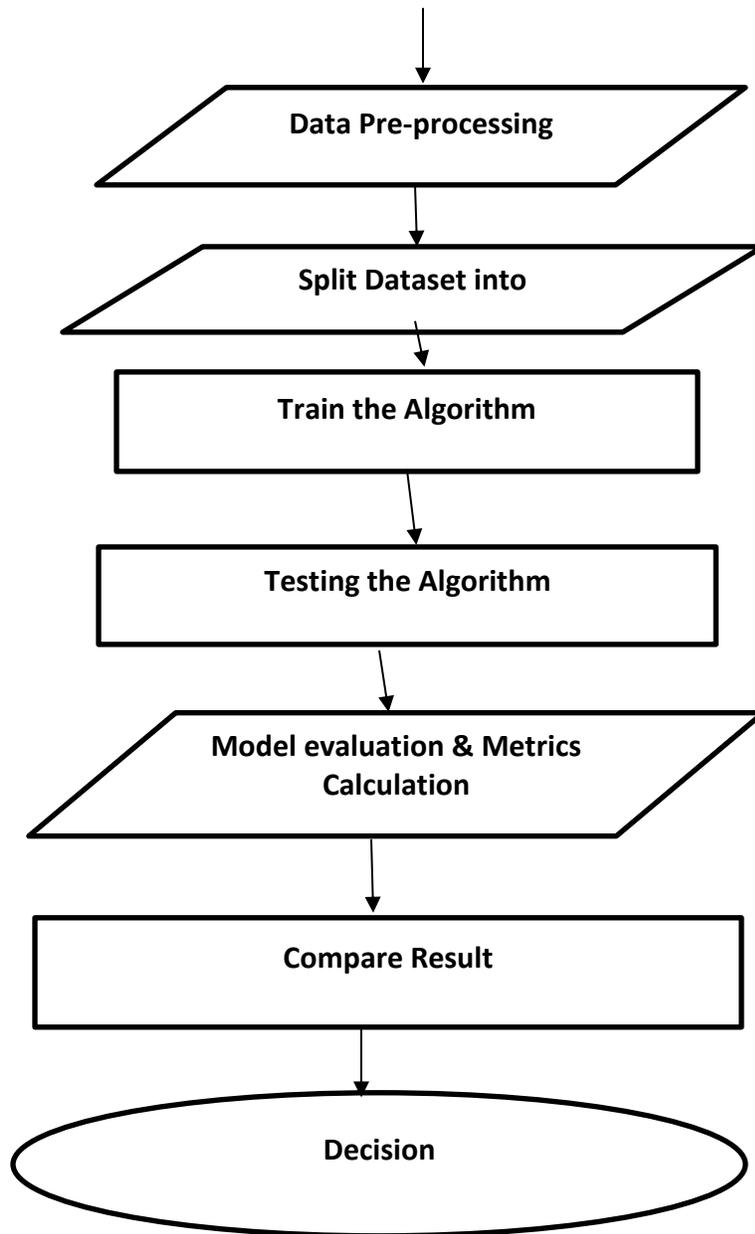
5. FPR: It's an error of data reporting which indicates percentage of a condition.

$$\text{FPR} = \frac{FP}{FP+TN}$$

6. FNR: It's also an error of data reporting when in reality it is not present.

$$\text{FNR} = \frac{FN}{FN+TP}$$

3.3 Research Methodology of Process:



This diagram is describing the whole procedure of this system

CHAPTER 4

DESIGN SPECIFICATION AND RESULT

4.1 Front-end design:

The Admin will put specific user mail to get the all necessary information from the different users. This user id needs to accurate the data what they have. No need to put same user id for more than one user. Admin will take all the answers from the given questions and arrange them with a pattern.

4.2 Back-end design:

Table 1: simple database of stored data

E-mail	Sex	Age	Like	Amount (month)	Starting Age	Color	Season	With Food	Amount (day)	Problem
mehadi7624@gmail.com	male	24	No	(100-500ml)	9	white	Extremely summer	Fast food	200ml	Obesity
Asma1122@gmail.com	female	17	Yes	(1000-3000ml)	6	color	Any time	Fast food	250ml	Asthma
feroza@gmail.com	female	55	No	(100-500ml)	40	white	Extremely summer	Fast Food	100ml	Diabetic
Prima9772@Gmail.com	female	22	Yes	(1000-3000ml)	10	color	Summer	Fast Food	250ml	Obesity
adnan@gmail.com	male	12	Yes	(1000-3000ml)	5	color	Summer	Any Time	250ml	Asthma

Here, firstly we have collected the raw data from local people after then those data have been labeling. After labeling raw data, we have to process those data that's why we processed those data by doing pre-processing. Now we have a dataset which is processed, and we have to split this dataset into training data and testing data. While we splitting the dataset, it was doing by ten-cross validation method. After splitting the dataset, we have to train the algorithm and also testing the algorithm. Then model evaluation and metrics calculation have to be implemented. After doing all of these we can take decision from this.

We have found from our pie chart, (Here, we have divided sex into (Male=M) and (Female=F)).

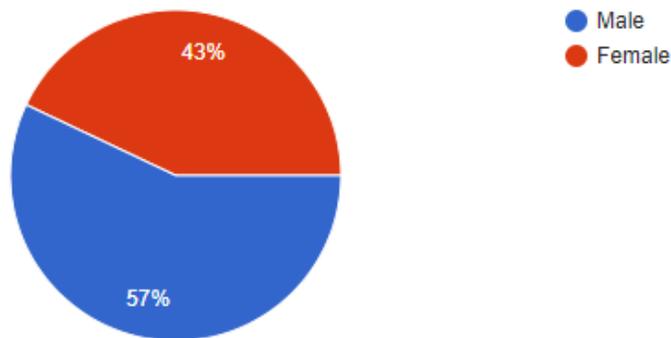


Figure 1: Structure of all collection based on male and female

This pie chart is for identifying the gender of the person by male and female. Because it's important to know with the help of this system the ratio of people who taking soft drinks and how the negative effect create on their health with their gender.

Age:

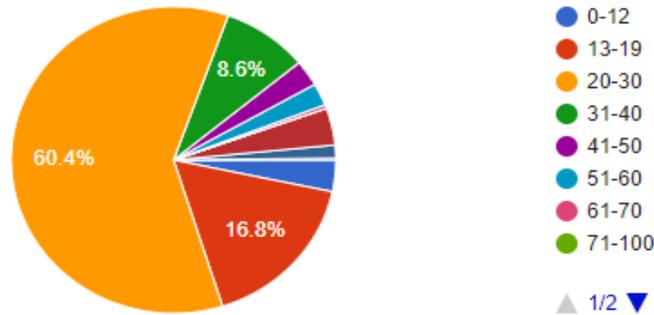


Figure 2: Age collection

Here, we have arranged the numeric value which is denoted by (below 10=0),(11-20=1),(21-30=2),(31-40=3),(41-50=4),(51-60=5),(61-70=6) and (more than 70=7).

Soft drinks preference:

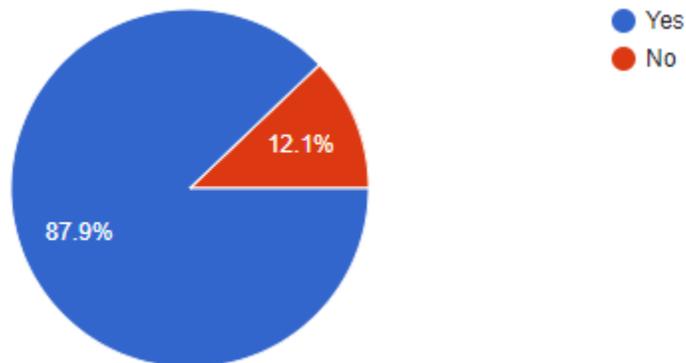


Figure 3: Preference of soft drinks.

Here, we have decided to split preference into two categories; one is YES and another one is NO.

Monthly drinking amount:

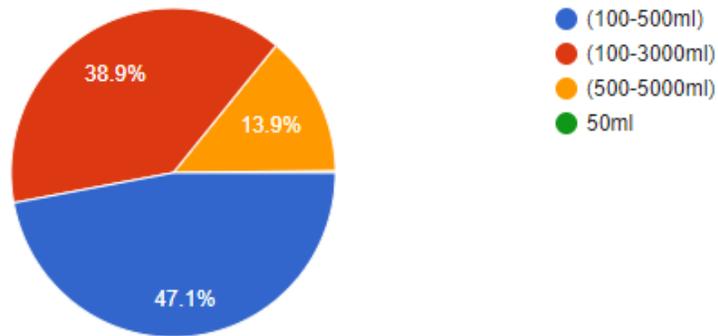


Figure 4: Amount of drinking structure

Here, we assumed that the amount of drinking is divided into three categories; $\{(100-500)=0\}, \{(500-3000)=1\}, \{(3000-5000)=2\}$

Drinking starting age:

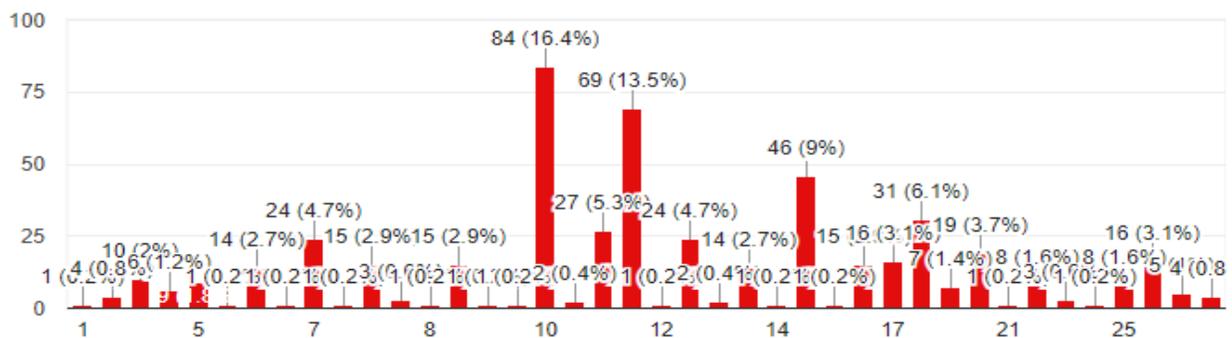


Figure 5: Starting age structure

Here, this data is nominal that is also known as nominal scale which is a type of data that is used to label variables without providing any quantitative value.

Soft drinks color preference:

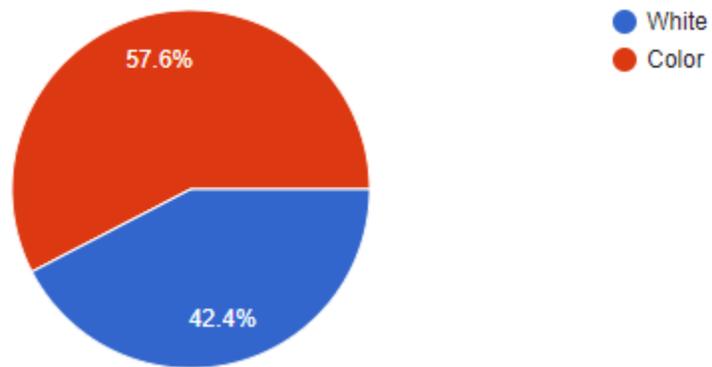


Figure 6: Color of soft drinks structure

Here, we consider two colors which one is White=0 and another one is Color=1.

Season of more drinking:

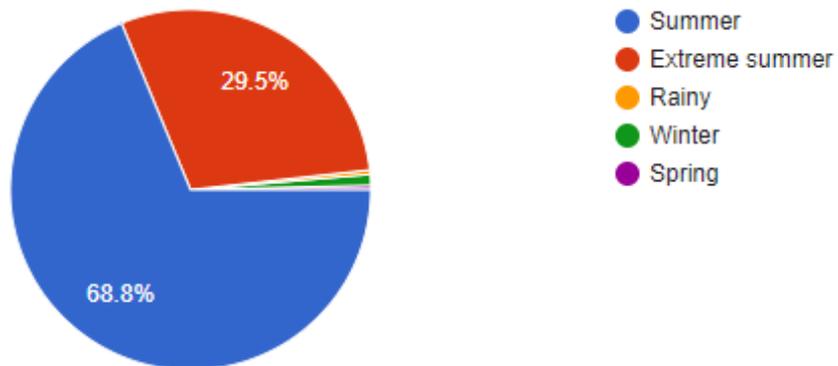


Figure 7: Season for drinking Soft drinks

Here, the season will be denoted like this; summer=0, extreme summer=1, spring=2, rainy=3, winter=4 and anytime=5

Drinking amount with food:

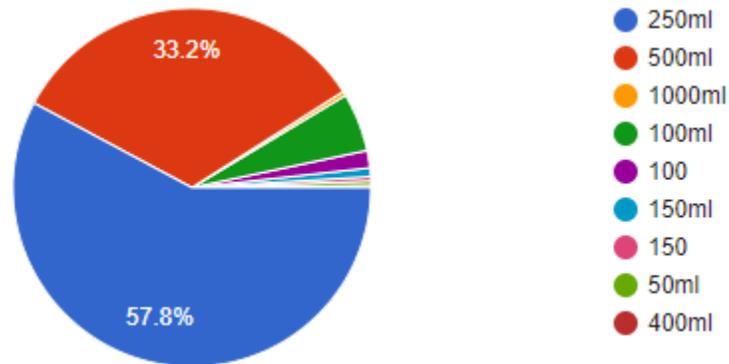


Figure 8: Drinking amount with any kind of food

When our system will get the question's answer from the different users who give their data, it will automatically rearrange on the table like the table no 1.

4.3: Experimental result:

The experiment was executed by using Weka on windows 10 operating system.

Table 2: Used methods

Classifier Name	True Positive (TP)	False Negative(FN)	False Positive (FP)	True Negative (TN)	Accuracy
Naive Bayes	421	12	71	7	83.7573%
Multilayer Perceptron	411	22	57	21	84.5401%
Logistic	430	3	69	9	85.91%
K- Star	422	11	53	25	87.4755%
J48	422	11	60	18	86.1057%

Random Forest	413	20	53	25	85.71%
---------------	-----	----	----	----	--------

4.4 Implementation Requirement:

When we started to implement this project, it was very new to us. That was hard to cover up this challenge. For implementing the system, we had to maintain a big survey and need a computer to implement the code. Besides it requires some soft skills.

- a. Naïve Bayes
- b. Multilayer perceptron
- c. Logistic
- d. K-Star
- e. J48
- f. Random forest

The database handling should be careful so that database is not to be heavy for data redundancy.

CHAPTET 5

IMPLEMENTATION AND TESTING

5.1 User Registration

For data mining, at first we have to register user with collecting his/her data using. By putting unique user data is separated from another id and data.

Table 3: Database of stored data.

Gmail	Sex	How old	Like	Amount (month)	Age start	Color	Season	With Food	Amount (day)	Problem
mehadi7624@gmail.com	male	24	No	(100-500ml)	9	white	Extremely summer	Fast food	200ml	Obesity
Asma1122@gmail.com	female	17	Yes	(1000-3000ml)	6	color	Any time	Fast food	250ml	Asthma
feroza@gmail.com	female	55	No	(100-500ml)	40	white	Extremely summer	Fast Food	100ml	Diabetic
Prima9772@Gmail.com	female	22	Yes	(1000-3000ml)	10	color	Summer	Fast Food	250ml	Obesity
adnan.psc@gmail.com	male	12	Yes	(1000-3000ml)	5	color	Summer	Any Time	250ml	Asthma

5.2 Implementation of database

Database contains a set of information and person name of each email along. Database will be dynamic in future. Here the database for implement the whole system we categories around parts. Where the first column is sex, by which we can easily identify that the person is male or female. Because it's important to know with the help of this system the ratio of people who taking soft drinks and how the negative effect create on their health with their sex. And after that the second part of our database is Age. In this area we defined the person whose data we collect by dividing their age ten after ten and mention that as like below $10=1$, $11-20=2$ as same up to 100. It will help us which ages people are more addicted to getting soft drinks. After that we arranged our database by collecting the preference of the people. If the interested to taking soft drinks it defines by Yes on the other hand if they are not interest to taking this, then it's define by No. It helps us to find the ration of people who are interested or not. And then we measure that people how much take soft drinks in a month. $100-500\text{ml}$ we defined it by 1 and $500-3000$ is 2 and the last $3000-5000$ is 3. And then age, from which age people starting to getting this. we collect the age of all those people for betterment our implementation. Then we divided different color of soft drinks by mentioning two colors, white is white which is $=0$ and all the other is color which $=1$. Then we divide a whole year by 5 parts where there been included summer, extremely summer etc. These all are mentioned by 0,1,2 up to 5. As well as we included that, with which food people would like to take soft drinks and also how much they have taken in one time when they getting food. And the very important part of this implementation is to find the problem status of those people from whom we collect Raw data. By analyzing all those data their we have been taken a decision is that the people feels a problem or not. Though we collect data specifically which problem the person has been faced, but we only mention that if he is sufferer or not.

Database contains a set of information and person name of each email along. Database will be dynamic in future

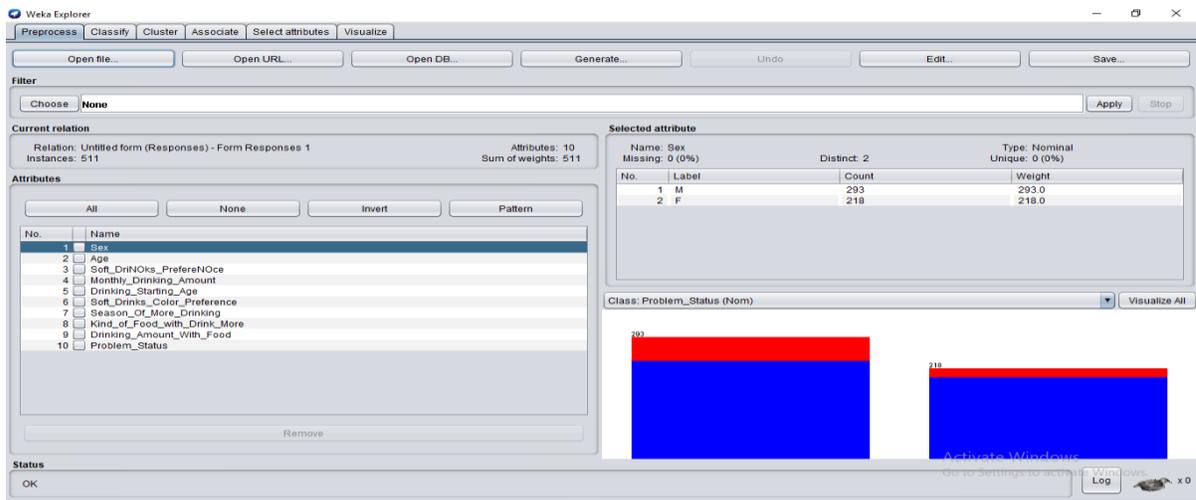


Figure 9.1: Implementation of Database

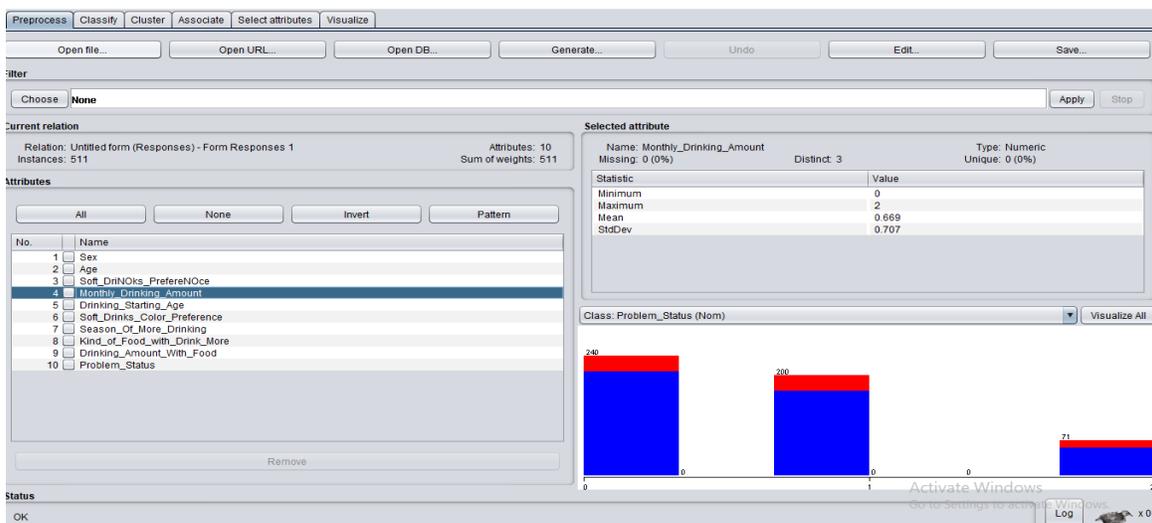


Figure 9.2: Implementation of Database

The confusion matrix:

Here this table is often used to elaborate the performance of the model table where from the data set true value is known.

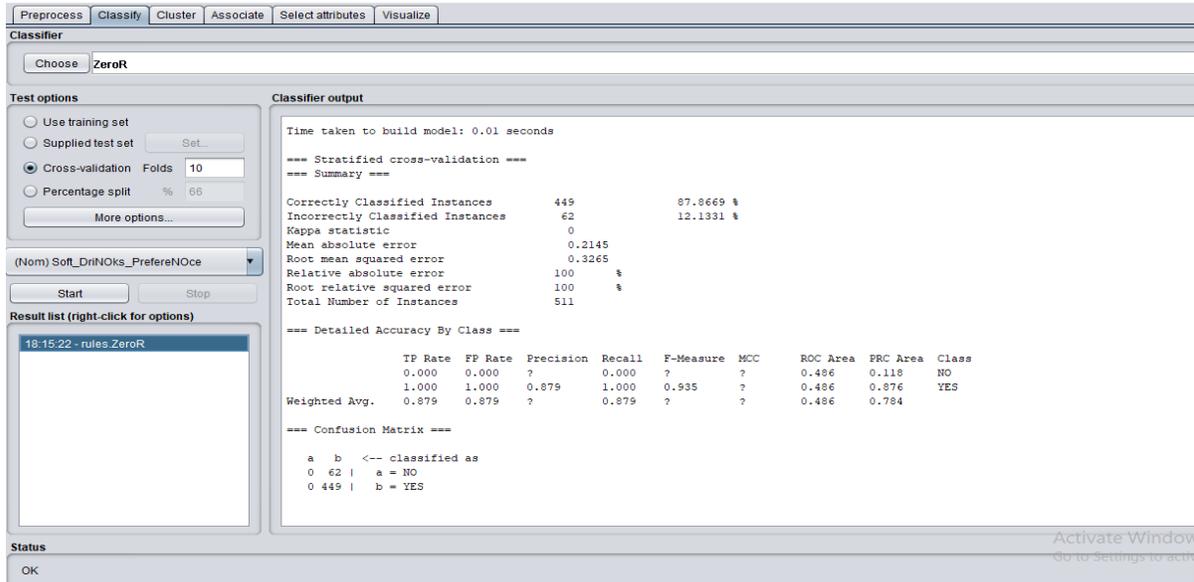


Figure 9.3: Implementation of Database

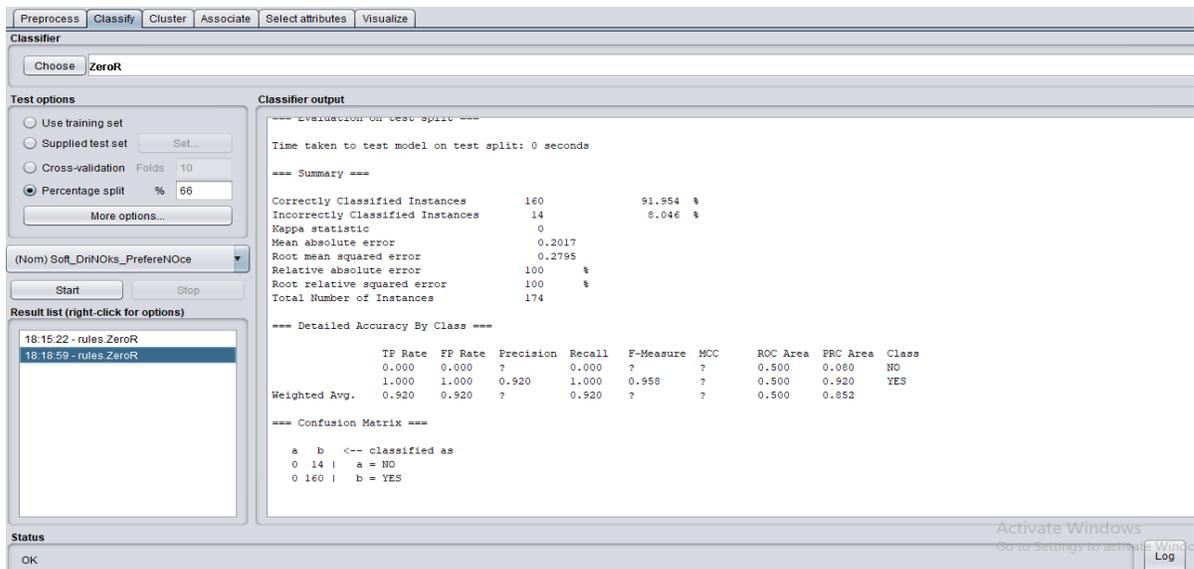


Figure 9.4: Implementation of Database

5.3 Training data

Training data is the model of dataset which is reason to fit the data and modify the data set model where test data is used only to access data. Training data is basically fit the parameter. You use training dataset, that contains outcomes to train the machine. You then use testing dataset that has no outcomes to predict outcomes. If we divide 100 percentage of data into 66 percentage of data for training, then 33 percentage of data will be for testing.

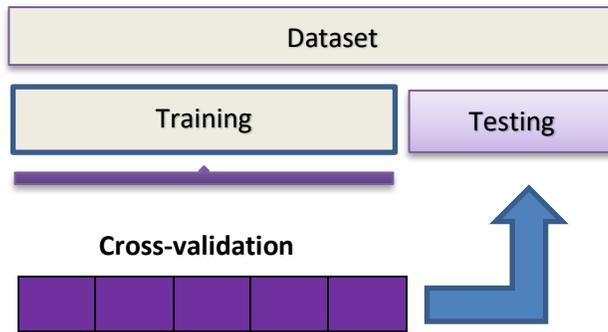


Figure 10: Process of training data

5.4 Testing implementation:

Testing data is unseen where with which testing data can be accessed. Basically test data is data which has been particularly identified for use in tests, typically of a computer program. Some data can be used in a confirmation way.

We use our processed data for further testing actually by using different types of methods e.g : cross –validation and so on .

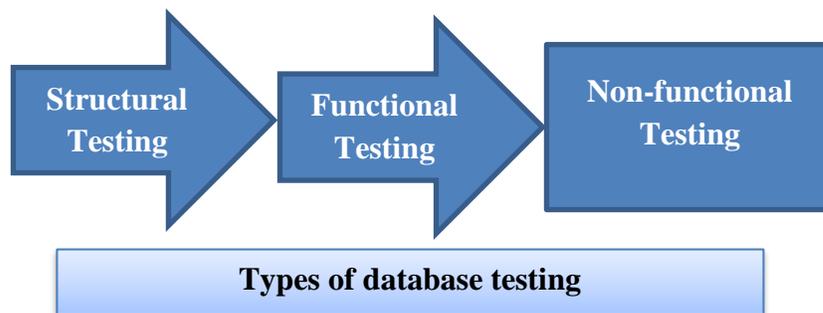


Figure 11: Testing data

5.5 Acceptance Test

This test case is better than other related algorithm which can give accurate result.

5.6 Test result and report

All the tested value was positive. We need to focus on graphical interface our system.

5.7 System testing

The beta testing result was very supportive. End users will understand the system easily and acknowledge will be positive.

CHAPTER 6

CONCLUSION AND FUTURE PLAN

6.1 Discussion and Conclusion

The impact of soft drinks is being a significant issue now-a-days in the whole world. The impact of soft drinks on the health to the local people based on data mining techniques includes health solution for global enforcement agencies to provide human safety. Proper use of this system can be seen in everywhere. This system can be useful to improve the health condition to the local people by showing them the accurate result. Studies have shown that natural products of plants origin e.g. Grape fruits juice contain different organic compounds (e.g. Alkaloids, tannins, pigments, organic and amino acids) and most are known to have inhibitive action. Given the different sources of energy in a diet, it is noticeable that a single source of energy can have such a muscular impact on total energy intake. This research also highlights how we might better motivate learners to perform at their best now and in the future.

6.2 Future scope

Some of the work we have finished but we have a future plan to develop this system is more reliable, understandable and user friendly. To achieve the goal we have to do some work and those are-

- a. Work more with database
- b. Create authentication to use
- c. Specifically mention the disease
- d. Try to measure how much it's unhealthy for people by doing the system more effective to calculate the result

References

- [1] 41, 199-206 2013, Priming and blood sugar level effect on learner cognition in EFL class, Joy JARMAN-WALSH.
- [2] ASEAN Food Journal 14 (2): 69-81 (2007), soft drinks and hard facts: A health perspective 1, *Xavier, R., 1Sreeramanan, S., 2Diwakar, A., 2Sivagnanam, G. and 2Sethuraman, K. R.
- [3] April 2007, Voi 97, No. 4 | American Journal of Public Health, effects of soft drink consumption on nutrition and health: a systematic review and meta-analysis.
- [4] African Health Sciences Vol 13 Issue 3 September 2013, a comparative study of the effect of diet and soda carbonated drinks on the histology of the cerebellum of adult female albino waster rats.
- [5] International journal of engineering applied sciences and technology, 2019, vol.4, issue 3, ISSN no.2455-2143, pages 425-431, corrison behaviors of low-carbon steel in cocoa butter medium
- [6] www.cjasn.org Vol 12 January, 2017, Diet Soda Consumption and Risk of Incident End Stage Renal Disease, Casey M. Morgan E. Grams, *†‡ Lyn M. Steffen, § Deidra C. Crews, *‡ Cheryl A. M. Anderson, †|Lydia A. Bizana ,Josef Coresh, * †** and Lawrence J. Appel*†**
- [7] Gulf medical journal, 2013;2(1);22-26, Consumption pattern of nutritional health drinks and energy drinks among university students in Ajman, UAE Shery Jacob¹, Jamila Tambawel², Fatima Mohammed Trooshi², Yara Alkhoury².
- [8] Food Funct., 2014, 5, 1686, Sensory attributes of soft drinks and their influence on consumers' preferences Noem´ı Redondo, Sonia G´omez-Mart´inez and Ascensi´on Marcos*
- [9] GOOGLE "Cascade Classifier images".

Soft Drinks

ORIGINALITY REPORT

23%

SIMILARITY INDEX

18%

INTERNET SOURCES

11%

PUBLICATIONS

16%

STUDENT PAPERS

PRIMARY SOURCES

1

journals.plos.org

Internet Source

3%

2

Submitted to Daffodil International University

Student Paper

2%

3

dro.dur.ac.uk

Internet Source

2%

4

www.researchgate.net

Internet Source

2%

5

Charles Spence, Xiaoang Wan. "Assessing the Influence of the Drinking Receptacle on the Perception of the Contents", Elsevier BV, 2016

Publication

1%

6

psasir.upm.edu.my

Internet Source

1%

7

towardsdatascience.com

Internet Source

1%

8

Submitted to UNITEC Institute of Technology

Student Paper

1%