Predictive Diagnosis of Kidney Stones

Chitesh Mittal (071280)

Dinkey Goyal (071417)

Shruti Mittal (071525)

Karan Thaman (071532)

Under the supervision of

Dr. Satish Chandra

May-2011

Submitted in partial fulfillment of the Degree of

Bachelor of Technology

DEPARTMENT OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT.

CERTIFICATE

This is to certify that the work entitled, **"Predictive Diagnosis of Kidney Stones"** submitted by **Chitesh Mittal(071280), Dinkey Goyal(071417), Shruti Mittal (071525) and Karan Thaman (071532)** in partial fulfillment for the award of degree of Bachelor of Technology in Computer Science Engineering & Information Technology of Jaypee University of Information Technology has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Signature of Supervisor	
Name of Supervisor	Dr.Satish Chandra
Designation	Assistant Professor
Date:	23.May.2011

ACKNOWLEDGEMENT

We would like to extend our gratitude and take this opportunity to thank our esteemed project guide **Dr.Satish Chandra** who helped us time and again and also guided us in times of uncertainty. We would like to thank the Computer Science Department of JUIT for giving us permission to commence this project at the first instance & to do the necessary research work.

We owe our heartiest thanks to **Brig. (Retd.) S.P. Ghrera** (H.O.D CSE/I.T Department) who've always inspired confidence in us to take initiative. He has always been motivating and encouraging.

Date: 23·May·2011

Chitesh Mittal

Dinkey Goyal
Dinkey Goyal

Shruti
Shruti Mittal

Karan
Karan Thaman

Summary

The project "Predictive Diagnosis of Kidney Stones" is about how to predict the probability of kidney stone in human-beings. This prediction is based on certain factors viz. environmental, life-style, dietary etc. These are the factors which influence the kidney stone formation in a human-being. These factors were taken from various research papers and were consulted by the nephrologists. After preparing a list of all the factors, data was collected from various hospitals and from the university students. This data included both the kidney stone patients and the people who were not having the stone in their kidney. Once the data was collected, it was converted into comma separated values (.csv) format. Data collected was of about 500 people.

For applying the data mining techniques on the data, the WEKA 3.6.4 tool was used. So after loading the input file (data in .csv format), pre-processing of data was done. Firstly the clustering technique "Simply k-means" was applied on the data set. Since our data had two classes i.e 0 (for persons not having the kidney stone) and 1 (for the kidney stone patients), so it got trained with the data-set (75% as training set). After applying the algorithm on the remaining test set (25%), it divided the data into two clusters viz. 0 and 1, indicating the number of people in both the clusters. So as a result, it was found that through this technique the prediction can be made only of the whole population and not of any particular individual.

Next technique that was applied was a classification technique – Linear Regression method. It gave a linear regression equation as the output. Putting the values in that equation will tell is there any probability of a person getting kidney stone in future or not. So this was a better technique as compared to the clustering method.

Chitesh MIttal	Dr.Satish Cha ndra
Dinkey Goyal	
Shruti Mittal	Date: 23. May. 2011
Karan Thaman	
Date: 23. May. 2011	

List of Figures

Table of Contents

Chapter 1
Introduction

1.1 Problem Statement:

To design a data mining tools that will predict the probability of kidney stone formation in a human-being in his/her future life-time. This will provide the technicians an easy way of predicting the kidney stones in human beings. It will also help the people, as they will be able to know whether they have got any probability of getting a kidney stone formation in their future, based on some simple factors and not much of clinical testings. Doctors will just have to ask certain questions from the patients regarding their environment, in which they live, the kind of food that they take, about their family history, any kind of medication that they are going through etc. Based on answers given by the persons, probability of kidney stone formation will easily be predicted.

1.2 Methodology:

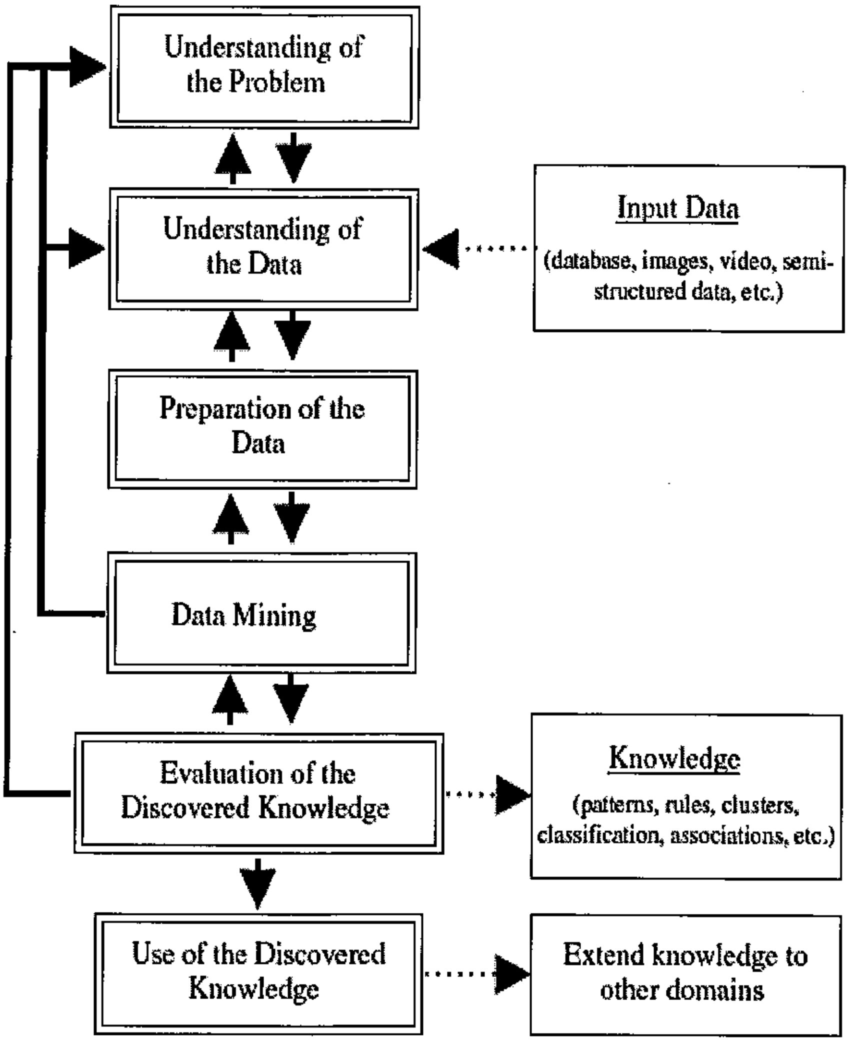

Understanding of the Problem
Understanding of the Data
Input Data
(database, images, video, semi-structured data, etc.)
Preparation of the Data
Data Mining
Evaluation of the Discovered Knowledge
Knowledge
(patterns, rules, clusters, classification, associations, etc.)
Use of the Discovered Knowledge
Extend knowledge to other domains

Fig 1.1 Methodology

1.3 Hardware and Software requirements:

Operating System - Microsoft Windows XP

Software- Weka version 3.6.4

1.4 Literature Survey

Studied various research papers to know about the kidney stones and the various factors that help in enhancing the kidney stone formation in human beings. After this literary survey, we came across a number of factors, so we consulted various nephrologists to know what are the prominent factors among all the factors, which help in stone formation. All the factors were categorized in some groups like the factors based on environmental conditions in which a person lives, the type of diet that he/she takes, the medications he/she is going through etc. The various important factors were categorized as follows:-

History

Gender

Age

Family History

Clinical Data

Blood Pressure

Diabetes

Blood Disease

Previous kidney stone attacks

Urinary tract blockage

Urine Output

Hyperparathyroidism

Gout

Kidney status

Symptoms

Drugs

Environmental & Life-style Factors

Stress

Anorexia

Climate

Bed-ridden

Vit B6 deficiency

Water intake

Oxalate-rich foods

Beverages

Salt-intake

Calcium Supplements

Chapter 2

Kidney Stones in Human Beings

2.1 Introduction

The human body has two kidneys. The kidneys are vital organs that filter waste products from the blood to make urine, which then flows from the kidneys into small tube called ureters into the bladder. The bladder stores urine until it is eliminated from the body through the urethra.

Kidney stones are small bits of hard crystallized material that form in the kidney. These deposits can impair the passage of urine that has the potential to result in infection and kidney damage or failure in severe cases. Kidney stones, also called renal lithiasis, are a common condition and are often made up of calcium, but can also contain uric acid or amino acids (proteins). One or more kidney stones can form in one or both kidneys. Kidney stones begin as small specks and can gradually increase in size. A person with a small kidney stone may be unaware of the condition, and it may pass in the urine out of the body without causing pain or other problems. There are generally no symptoms of large kidney stones that remain in the kidney. However, when a large kidney stone moves out of the kidney into the ureter it causes severe pain, called renal colic.

Other symptoms of a large kidney stone that has moved out of the kidney include hematuria, (blood in the urine), dysuria (difficulty urinating), nausea and vomiting. Most kidney stones generally pass out of the body in the urine. On occasion, a kidney stone can get stuck in a ureter and result in potentially serious, even life-threatening complications in some people.

Diagnostic testing includes blood tests and performing a urinalysis test, which checks for the presence of blood in the urine and other elements and help to differentiate between a kidney stone and a urinary tract infection. Imaging tests, such as ultrasound or CT, are performed to determine the cause of symptoms and locate any possible kidney stones. A urine test that includes collecting urine for 24 hours may be ordered to evaluate the urine for substances that typically form kidney stones.

2.2 Symptoms of Kidney stones

The list of signs and symptoms mentioned in various sources for Kidney stones includes the 21 symptoms listed below:

- No early symptoms - early stages have no symptoms
- Severe abdominal pain
- Kidney pain - sharp pain usually on the back and at the side
  - Severe back pain
  - Side pain
- Nausea
- Vomiting
- Groin pain
- Hematuria
- Kidney infection symptoms - very serious symptoms:
  - Vomiting
  - Blood in urine
  - Smelly urine
  - Cloudy urine
  - Burning when urinating
  - Urinary urgency
  - Fever - if kidney infection develops
  - Chills - if kidney infection develops
  - Kidney damage - severe cases
  - Kidney failure - severe cases

- Reduced urination - if flow of urine is blocked

2.3 List of causes of Kidney stones

Following is a list of causes or underlying that could possibly cause Kidney stones includes:

- Urinary tract infections
- Cystic kidney diseases
- Metabolic disorders
- Hyperparathyroidism
- Renal tubular acidosis
- Cystinuria
- Hyperoxaluria
- Absorptive hypercalciuria
- Hyperuricosuria
- Gout
- Urinary tract blockage

2.4 List of Risk Factors for Kidney stones

The list of risk factors mentioned for Kidney stones in various sources includes:

- Family history of kidney stones
- Sedentary occupations
- Lack of physical activity
- Hypercalcuria
- Acidic urine
- Alkaline urine
- Reduced urine
- Too little water
- Excessive sweating

- Gout
- Kidney infections
- Vitamin A deficiency
- Overactive parathyroid gland

2.5 Treatment List for Kidney stones

The list of treatments mentioned in various sources for Kidney stones includes the following list. Always seek professional medical advice about any treatment or change in treatment plans.

- No treatment - in mild cases; some kidney stones pass naturally in urine
- Extracorporeal shockwave lithotripsy (ESWL)
- Surgery
    - Percutaneous nephrolithotomy
    - Ureteroscopic Stone Removal
- Water - drinking lots of water may help the stone pass naturally
- Pain medication
- Acute management:
    - Analgesia - NSAID's, paracetamol, narcotic anlgesia
    - Management of nausea and vomiting which often accompanies renal colic
    - Fluid management - often IV to treat and prevent dehydration
    - Imaging to determine if obstruction is present
    - Appropriate waiting to determine if the stone will pass spontaneously. If no infection or obstruction is present, it may be appropriate to wait and follow up regularly over a period of weeks
    - Agents that may aid the passage of stones - calcium channel blockers, alpha blockers
    - Alkalization of urine to dissolve uric acid stones

  - Surgical intervention for obstruction - Cystoscopy and stone retrieval. Ureteric stenting
  - Antibiotics if infection is present
- Extracorporeal shockwave lithotripsy
- Percutaneous nephrostolithotomy
- Long term management and prevention of recurrence
  - Ensure adequate fluid intake. > 2L per day, and more for those living in hot climates and those involved in heavy manual labour
  - Moderation of calcium or oxalate intake if calcium or oxalate stones are diagnosed
  - Allopurinol for prevention of uric acid stones

2.6 Misdiagnosis of Kidney stones

These include urinary tract infection, pyelonephritis, appendicitis, sexually transmitted diseases, epididymitis, prostatitis, and pelvic inflammatory disease.

2.7 Kidney stones: Diagnostic Tests

The list of diagnostic tests mentioned in various sources as used in the diagnosis of Kidney stones includes:

- X-rays
- Sonogram
- IVP (intravenous pyelogram)
- Analysis of a passed stone
- 24 hour urine test

2.8 Kidney stones: Types list

The list of types of Kidney stones mentioned in various sources includes:

- Calcium stone - most common type
- Struvite - infection stone
- Uric acid stone
- Cystine stone
- Ureteral stone
- Silent kidney stones - causing no symptoms

2.9 Deaths:

Kidney stones: Hospitalization Statistics

The following are statistics from various sources about hospitalizations and Kidney stones:

- 0.45% (56,987) of hospital episodes were for urolithiasis in England 2002-03 (Hospital Episode Statistics, Department of Health, England, 2002-03)
- 90% of hospital consultations for urolithiasis required hospital admission in England 2002-03 (Hospital Episode Statistics, Department of Health, England, 2002-03)
- 71% of hospital episodes for urolithiasis were for men in England 2002-03 (Hospital Episode Statistics, Department of Health, England, 2002-03)
- 29% of hospital episodes for urolithiasis were for women in England 2002-03 (Hospital Episode Statistics, Department of Health, England, 2002-03)
- 41% of hospital admissions for urolithiasis required emergency hospital admission in England 2002-03 (Hospital Episode Statistics, Department of Health, England, 2002-03)

Chapter 3

Data Mining in Prediction of Diseases

3.1 Introduction

Data Mining: - The nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data.

The terms of above definition are explained as follows:

Pattern

-models or structure in data (traditional sense)

- expression in some language describing a subset of the data or a model applicable to that subset (data comprises a set of facts)

Nontrivial (process)

- it must involve search for structure, models, patterns, or parameters

Valid

- discovered patterns should be valid for new data with some degree of certainty

Novel

- at least to the system and preferably to the user

Potentially useful

- for the user or task

Understandable

- Discovered patterns should be understandable - if not immediately, then after some post processing.

Data mining is a step in the KDD (knowledge discovery in databases) process concerned with the algorithmic means by which patterns or structures are enumerated from the data.

KDD includes the following steps:-

1. SELECTION: - selecting or segmenting the data that are relevant to some criteria.

2. PREPROCESSING (data cleaning):- unnecessary information is removed. This stage reconfigures the data (taken from various sources) to ensure a consistent format. This step consists of removing outliers, dealing with noise and missing values in the data, and accounting for time sequence information and known changes.

3. TRANSFORMATION: - data is made usable and navigable.

4. DATA MINING: - extraction of patterns from the data.

5. INTERPRETATION AND EVALUATION: - patterns are converted into knowledge, which is the used to support decision-making.

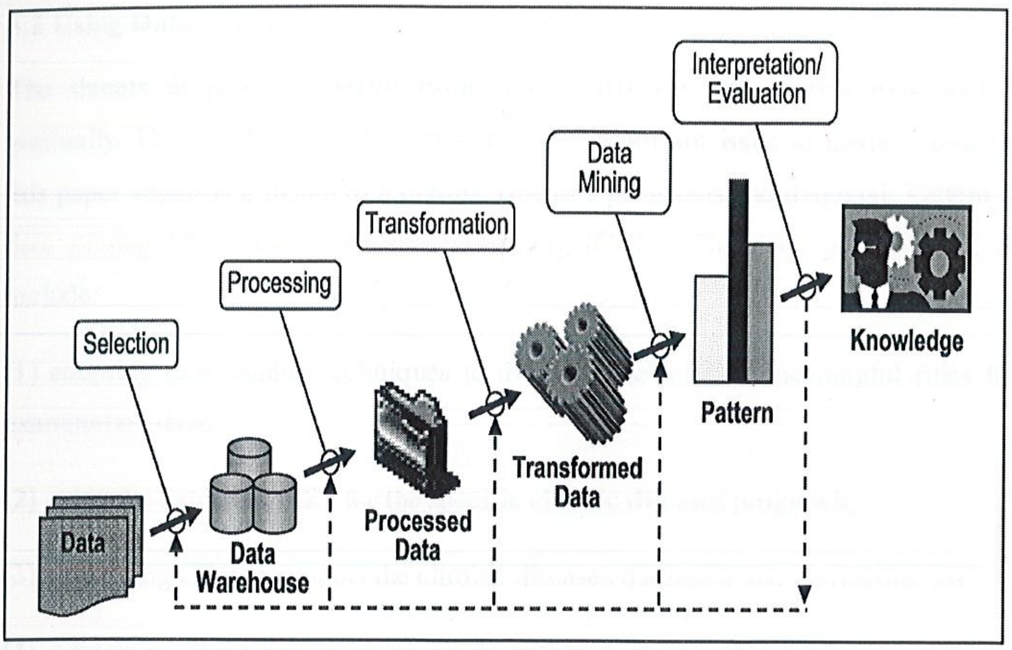


Fig 2.1 The process of KDD

Goals of Data Mining

1. Prediction: - makes use of existing variables in the database in order to predict unknown or future values of interest.

2. Description:- find patterns describing the data and the subsequent presentation for user interpretation.

3.2 Using Data mining for diagnosis of diseases

The threats to people's health from chronic diseases are always exist and increasing gradually. How to decrease these threats is an important issue in medical treatment. Thus, this paper suggests a model of a chronic diseases prognosis and diagnosis system integrating data mining (DM) and case-based reasoning (CBR). The main processes of the system include:

(1) adopting data mining techniques to discover the implicit meaningful rules from health examination data,

(2) using the extracted rules for the specific chronic diseases prognosis,

(3) employing CBR to support the chronic diseases diagnosis and treatments, and

(4) expanding these processes to work within a system for the convenience of chronic diseases knowledge creating, organizing, refining, and sharing. The experiment data are collected from a professional health examination center, MJ health screening center, and implemented through the system for analysis. The findings are considered as helpful references for doctors and patients in chronic diseases treatments.

Chapter 4
Implementation

After studying all the kidney stone related factors (mentioned in chapter1), created a excel file listing all of them. This file was given in several hospitals to be filled up by the patients in accordance with the certain codes mentioned in the list only. The codes were given to different attributes, after studying the research papers.

Attributes and their codes:

- Status: It included the information whether the person is having kidney stone or not. The persons having the kidney stone were supposed to fill the code '1' and those who were not having it had the code '0'.

- Gender: Kidney stones are more likely to develop in men than in women. Males were given code '0' and females '1'.

- Age: Since kidney stone is more prevelant in age of 24-70, so it was divided into several ranges, ranging from 0 to 100. Range of 0-20 was coded as '0', 21-40 as '1', 41-60 as '2', 61-80 as '3' and 81-100 as '4'.

- Stress: It mentioned whether the life of the person is stressful or not. So the person having stressed were given code '1' and those having no stress were given code '0'.

- Bed-ridden: The people who were bed-ridden were coded as '1' and those who were not were coded as '0'.

- Anorexia: refers to a lack or loss of appetite, resulting in the inability to eat. So the people having the problem of anorexia were coded as '1' and ones having no such problem were coded as '0'.

- Climate: People who become dehydrated and/or live in hot climates are at risk for kidney stones because they lose more body water and produce smaller amounts of urine that contains a higher concentration of substances that form kidney stones, such as calcium and amino acids. Since there is a large variation in climate of different regions. So it was also categorized and was given codes accordingly. The people who live in coastal region were coded as '0', those living in mountain area were codes as '1', those who live in the region having very high temperature and very cold temperature were given codes as '2' and '3' respectively.

- Family History: Kidney stones develop more frequently in individuals with a family history of kidney stones than in those without a family history. So, the people having this problem in their family history were coded as '1', else '0'.

- Vitamin B deficiency: The people having the deficiency were coded '1' and the ones not suffering from this deficiency were given code '0'.

- Water intake: Less intake of water is directly associated with the formation of kidney stones, as less water intake means less urine output. So one who consumes less than 1l/day has been given the code '0'i.e they have alarming chances of getting the kidney stone, between 1 to 4l/day is coded as'1' meaning that they have less chances of kidney stone formation and the one who consumes more than 4l/day has negligible chances of kidney stone, that's why has been coded '2'.

- Oxalate rich items: There are a number of oxalate rich items which one consumes in daily life. So if a person is consuming the oxalate rich food-item, e.g black tea, spinach, coffee, oranges, potatoes, strawberries, nuts, poultry etc. then the code given is 1 else 0.

- Salt intake: Higher amount of salt in the diet enhances the kidney stone formation in the body. So, if one consumes less than 2300mg/day, there are no chances of stone formation, hence coded '0'otherwise '1'.

- Calcium intake: Women who take calcium supplements have a 20% higher risk for stones. Therefore, consumption of calcium supplement greater than 1000-1200mg/day, has been coded '0' as chances of stone formation is less whereas '1' is given to the consumption greater than 1200mg/day.

- Hyperparathyroidism: The people having this disease is given code '1', otherwise '0'.

- Blood Pressure: Low b.p has been coded as '0', normal as '1' and high as '2'.

- Diabetes: The diabetic person will mark the code '1' and the one who is not having diabetes will mark **'0'**.

- Previous kidney stone attacks: If a person has got any kidney stone attacks previously, then the risk of stone attacks in future also increases. So the one having previous attacks have been coded as '1', otherwise '0'.

- Urinary tract blockage: If the urinary tract is blocked, then urine output will get reduced, so if there's any kind of urinary tract blockage in a person, the chance of kidney stone formation will get increased, hence given the code '1'. But if there's no urinary tract blockage, the code is '0'.

- Frequency of Urination: If the urinary tract is blocked, the frequency of urination will automatically get increased. So if the frequency is between 1 to 5, its normal and there's no symptom of getting a stone, hence coded '0'. The frequency between 6 to 10 is coded as '1' and greater than 11 as '2'.

- Urine Output: Proper discharge of urine is very much important to avoid the kidney stone formation. So, volume less than 2 L per day is coded as '1' as this can cause stone formation, but volume greater than this is coded as '0' as chances of stone formation is almost negligible.

- Kidney status: People with kidney deformities or anomalies, such as horseshoe kidney, are also at risk. If the person has only one kidney or if he has abnormally shaped kidney, then the chances of kidney stone formation is increased, hence given the codes '0' and '1' respectively. But if he/she has two normal kidneys, then chances of stone reduce, hence given the code '2'.

- Symptoms: There are several symptoms of kidney stone formation in a human being viz. pain, red/cloudy urine, nausea, vomiting, swelling, foul smelling urine, chills, sweating and fever. So if any of the above mentioned symptoms are apparent in any person, the code is '1' else '0'.

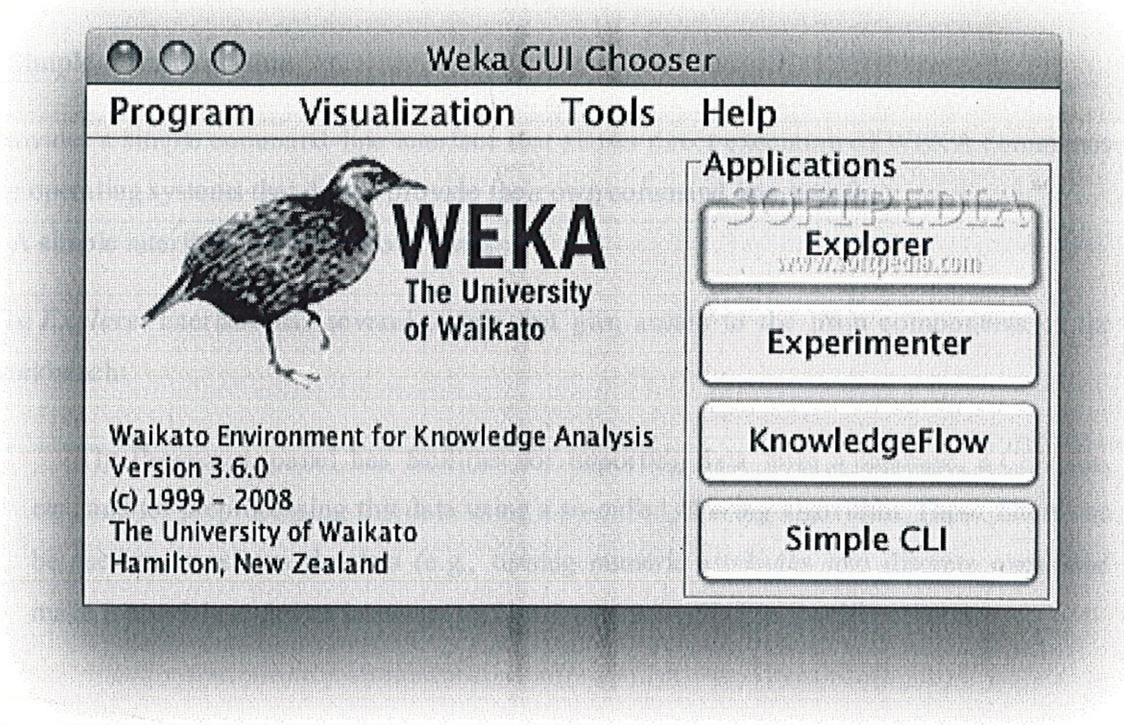
- Drugs: It is very important to know that what kind of medicines a person is taking. So here's some medicines viz. Indinavir, Antacids, Antibiotics, Aspirin, Corticosteroids, Laxatives, Allopurinol, Theophylline coded 0-7 respectively.

Collected the data from various hospitals regarding all the above mentioned factors, so that machine can be trained accordingly, to predict the result.

4.1 Preparation of the data: - After collecting the data from various hospitals, created a excel file to train our machine. We used Weka software, to make our machine learn. **Weka** (Waikato Environment for Knowledge Analysis) is a popular suite of machine learning software written in Java, developed at the University Of Waikato, New Zealand. Weka is free software available under the GNU General Public License.It has the following schemes:

- Schemes for classification include:
  - decision trees, rule learners, naive Bayes, decision tables, locally weighted regression, SVMs, instance-based learners, logistic regression, voted perceptrons, multi-layer perceptron
- Schemes for numeric prediction include:
  - linear regression, model tree generators, locally weighted regression, instance-based learners, decision tables, multi-layer perceptron
- Meta-schemes include:
  - Bagging, boosting, stacking, regression via classification, classification via regression, cost sensitive classification
- Schemes for clustering:
  - EM and Cobweb

WEKA Application Interfaces



Waikato Environment for Knowledge Analysis
Version 3.6.0
(c) 1999 - 2008
The University of Waikato
Hamilton, New Zealand

- **Explorer**

An environment for exploring data with WEKA (the rest of this documentation deals with this application in more detail).

– preprocessing, attribute selection, learning, visualization

- **Experimenter**

An environment for performing experiments and conducting statistical tests between learning schemes.

– testing and evaluating machine learning algorithms

- **Knowledge Flow**

This environment supports essentially the same functions as the Explorer but with a drag-and-drop interface. One advantage is that it supports incremental learning.

– visual design of KDD process

– Explorer

• **Simple Command-line**

Provides a simple command-line interface that allows direct execution of WEKA commands for operating systems that do not provide their own command line interface.
– A simple interface for typing commands

The ***Explorer*** interface has several panels that give access to the main components of the workbench:

- The *Preprocess* panel has facilities for importing data from a database, a CSV file, etc., and for preprocessing this data using a so-called *filtering* algorithm. These filters can be used to transform the data (e.g., turning numeric attributes into discrete ones) and make it possible to delete instances and attributes according to specific criteria.

- The *Classify* panel enables the user to apply classification and regression algorithms (indiscriminately called *classifiers* in Weka) to the resulting dataset, to estimate the accuracy of the resulting predictive model, and to visualize erroneous predictions, ROC curves, etc., or the model itself (if the model is amenable to visualization like, e.g., a decision tree).

- The *Associate* panel provides access to association rule learners that attempt to identify all important interrelationships between attributes in the data.

- The *Cluster* panel gives access to the clustering techniques in Weka, e.g., the simple k-means algorithm. There is also an implementation of the expectation maximization algorithm for learning a mixture of normal distributions.

- The *Select attributes* panel provides algorithms for identifying the most predictive attributes in a dataset.

- The *Visualize* panel shows a scatter plot matrix, where individual scatter plots can be selected and enlarged, and analyzed further using various selection operators.

The key features responsible for Weka's success are:

- It provides many different algorithms for data mining and machine learning
- is open source and freely available
- it is platform-independent
- it is easily useable by people who are not data mining specialists
- it provides flexible facilities for scripting experiments
- it has kept up-to-date, with new algorithms being added
- It is extensible
- Can be integrated into other java packages
- It has GUIs (Graphic User Interfaces), which can run individual experiments or build KDD phases

- Cons of Weka
  - ➢ Lack of proper and adequate documentations
  - ➢ Systems are updated constantly (Kitchen Sink Syndrome)

Since Weka can only read input in .arff or .csv format, so we converted our data,which was initially in .xls format, into .csv format.

4.2 Use of Data Mining technique: - In Explorer, we imported our .csv file and preprocessed it.

A comma-separated values or character-separated values (.CSV) file is a simple text format for a database table.

All the attribute of the table were listed under "Attributes". We applied two different techniques on our data viz. k-means (clustering) and linear regression (classification) on our data.

i) Simple k-means is one of the simplest clustering techniques. It is a data mining/machine learning algorithm used to cluster observations into groups of related observations without any prior knowledge of those relationships. The algorithm clusters observations into k groups, where k is provided as an input parameter. It then assigns each observation to clusters based upon the observation's proximity to the mean of the cluster. It is commonly used in medical imaging, biometrics and related fields.

Here's how the algorithm works:

1. The algorithm arbitrarily selects k points as the initial cluster centers ("means"). 2. Each point in the dataset is assigned to the closed cluster, based upon the Euclidean distance between each point and each cluster center.

3. Each cluster center is recomputed as the average of the points in that cluster.

4. Steps 2 and 3 repeat until the clusters converge. Convergence may be defined differently depending upon the implementation, but it normally means that either no observations change clusters when steps 2 and 3 are repeated or that the changes do not make a material difference in the definition of the clusters.

ii) Regression analysis is a statistical forecasting model that is concerned with describing and evaluating the relationship between a given variable (usually called the dependent variable) and one or more other variables (usually known as independent variables). Regression analysis helps us understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed. Most commonly, regression analysis estimates the conditional expectation of the dependent variable given the independent variables — that is, the average value of the dependent variable when the independent variables are held fixed.

Regression analysis is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning. Regression analysis is also used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships. In restricted circumstances, regression analysis can be used to infer causal relationships between the independent and dependent variables.

Percentage split was taken as 75% which implied that 75% of our data was taken as training set and the remaining 25% as the test set.

It showed the 'type' as 'numeric' as our whole data is in numeric form only.

4.3 To Classify with weka GUI

1. Run weka GUI (in Unix: java –jar weka.jar)

2. Click 'Explorer'

3. 'Open file...'

4. Select 'Classify' tab

5. 'Choose' a classifier

6. Confirm options

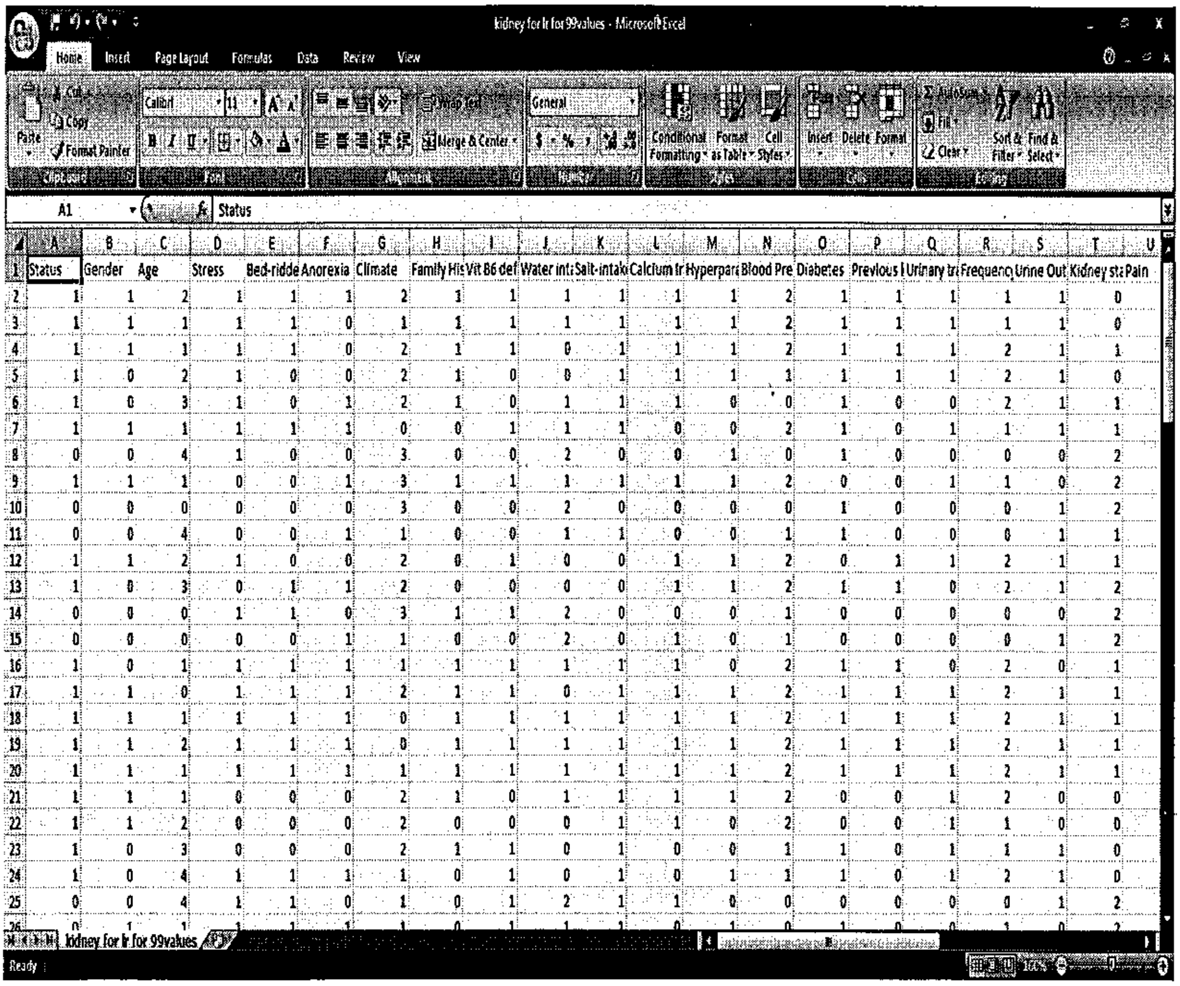
7. Click 'Start'

8. Wait...

9. Right-click on Result list entry

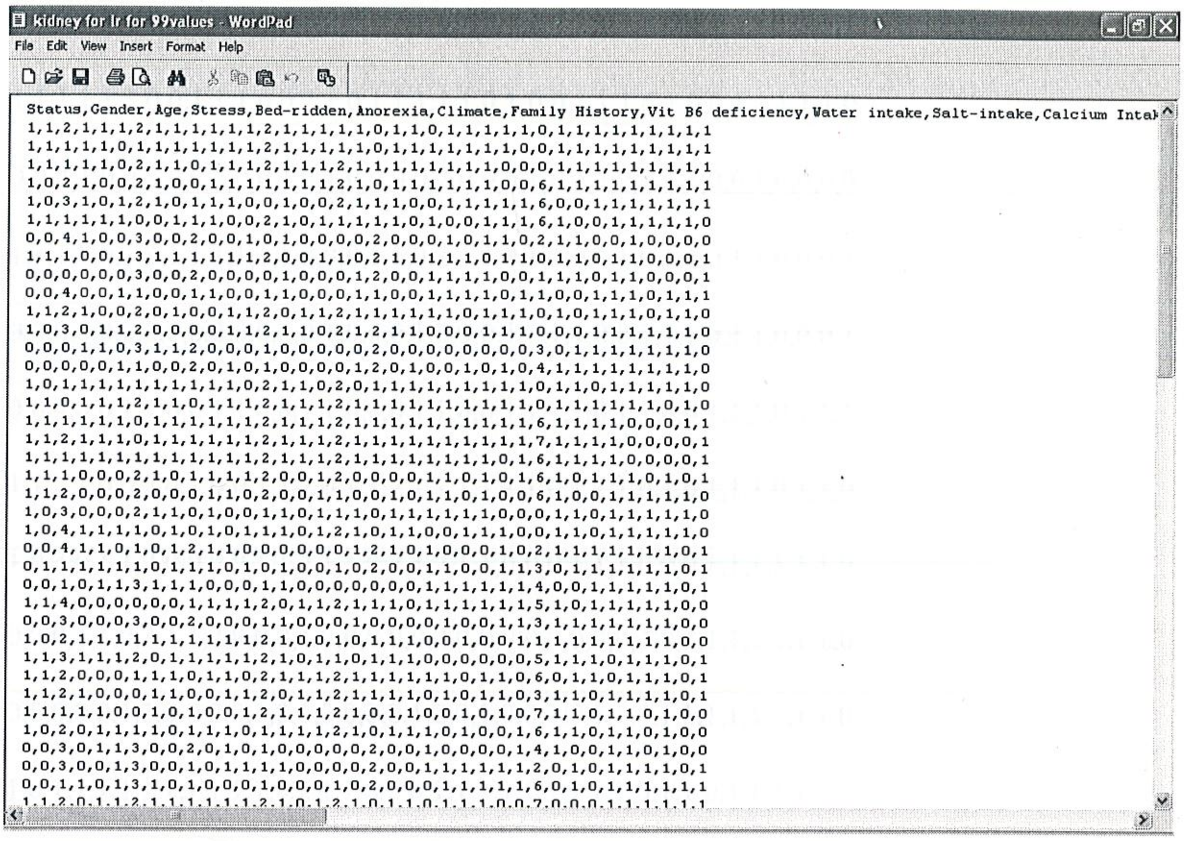
    1. 'Save result buffer'

    2. 'Save model'



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Status	Gender	Age	Stress	Bed-ridde	Anorexia	Climate	Family His	Vit B6 def	Water int	Salt-intak	Calcium Ir	Hyperpar	Blood Pre	Diabetes	Previous I	Urinary tr	Frequenc	Urine Out	Kidney st	Pain
2	1	1	2	1	1	1	2	1	1	1	1	1	1	2	1	1	1	1	1	0	
3	1	1	1	1	1	0	1	1	1	1	1	1	1	2	1	1	1	1	1	0	
4	1	1	1	1	1	0	2	1	1	0	1	1	1	2	1	1	1	2	1	1	
5	1	0	2	1	0	0	2	1	0	0	1	1	1	1	1	1	1	2	1	0	
6	1	0	3	1	0	1	2	1	0	1	1	1	0	0	1	0	0	2	1	1	
7	1	1	1	1	1	1	0	0	1	1	1	0	0	2	1	0	1	1	1	1	
8	0	0	4	1	0	0	3	0	0	2	0	0	1	0	1	0	0	0	0	2	
9	1	1	1	0	0	1	3	1	1	1	1	1	1	2	0	0	1	1	0	2	
10	0	0	0	0	0	0	3	0	0	2	0	0	0	0	1	0	0	0	1	2	
11	0	0	4	0	0	1	1	0	0	1	1	0	0	1	1	0	0	0	1	1	
12	1	1	2	1	0	0	2	0	1	0	0	1	1	2	0	1	1	2	1	1	
13	1	0	3	0	1	1	2	0	0	0	0	1	1	2	1	1	0	2	1	2	
14	0	0	0	1	1	0	3	1	1	2	0	0	0	1	0	0	0	0	0	2	
15	0	0	0	0	0	1	1	0	0	2	0	1	0	1	0	0	0	0	1	2	
16	1	0	1	1	1	1	1	1	1	1	1	1	0	2	1	1	0	2	0	1	
17	1	1	0	1	1	1	2	1	1	0	1	1	1	2	1	1	1	2	1	1	
18	1	1	1	1	1	1	0	1	1	1	1	1	1	2	1	1	1	2	1	1	
19	1	1	2	1	1	1	0	1	1	1	1	1	1	2	1	1	1	2	1	1	
20	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	1	1	
21	1	1	1	0	0	0	2	1	0	1	1	1	1	2	0	0	1	2	0	0	
22	1	1	2	0	0	0	2	0	0	0	1	1	0	2	0	0	1	1	0	0	
23	1	0	3	0	0	0	2	1	1	0	1	0	0	1	1	0	1	1	1	0	
24	1	0	4	1	1	1	1	0	1	0	1	0	1	1	1	0	1	2	1	0	
25	0	0	4	1	1	0	1	0	1	2	1	1	0	0	0	0	0	0	1	2	
26	0	1	1	1	1	1	1	0	1	1	1	0	1	0	1	0	0	1	0	2	

Fig 2.2 Data in spreadsheet format



kidney for lr for 99values - WordPad
File Edit View Insert Format Help

```
Status,Gender,Age,Stress,Bed-ridden,Anorexia,Climate,Family History,Vit B6 deficiency,Water intake,Salt-intake,Calcium Intak
1,1,2,1,1,1,2,1,1,1,1,1,1,2,1,1,1,1,1,0,1,1,0,1,1,1,1,1,0,1,1,1,1,1,1,1,1
1,1,1,1,1,0,1,1,1,1,1,1,1,2,1,1,1,1,1,0,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,1
1,1,1,1,1,0,2,1,1,0,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,0,0,0,1,1,1,1,1,1,1,1
1,0,2,1,0,0,2,1,0,0,1,1,1,1,1,1,1,2,1,0,1,1,1,1,1,1,0,0,6,1,1,1,1,1,1,1,1
1,0,3,1,0,1,2,1,0,1,1,1,0,0,1,0,0,2,1,1,1,0,0,1,1,1,1,1,6,0,0,1,1,1,1,1,1
1,1,1,1,1,1,0,0,1,1,1,0,0,2,1,0,1,1,1,1,1,0,1,0,0,1,1,1,6,1,0,0,1,1,1,1,1,0
0,0,4,1,0,0,3,0,0,2,0,0,1,0,1,0,0,0,0,2,0,0,0,1,0,1,1,0,2,1,1,0,0,1,0,0,0,0
1,1,1,0,0,1,3,1,1,1,1,1,1,2,0,0,1,1,0,2,1,1,1,1,1,0,1,1,0,1,1,0,1,1,0,0,0,1
0,0,0,0,0,0,3,0,0,2,0,0,0,0,1,0,0,0,1,2,0,0,1,1,1,1,0,0,1,1,1,0,1,1,0,0,0,1
0,0,4,0,0,1,1,0,0,1,1,0,0,1,1,0,0,0,1,1,0,0,1,1,1,1,0,1,1,0,0,1,1,1,0,1,1,1
1,1,2,1,0,0,2,0,1,0,0,1,1,2,0,1,1,2,1,1,1,1,1,1,0,1,1,1,0,1,0,1,1,1,0,1,1,0
1,0,3,0,1,1,2,0,0,0,0,1,1,2,1,1,0,2,1,2,1,1,0,0,0,1,1,1,0,0,0,1,1,1,1,1,1,0
0,0,0,1,1,0,3,1,1,2,0,0,0,1,0,0,0,0,0,2,0,0,0,0,0,0,0,0,3,0,1,1,1,1,1,1,1,0
0,0,0,0,0,1,1,0,0,2,0,1,0,1,0,0,0,0,1,2,0,1,0,0,1,0,1,0,4,1,1,1,1,1,1,1,1,0
1,0,1,1,1,1,1,1,1,1,1,1,0,2,1,1,0,2,0,1,1,1,1,1,1,1,1,0,1,1,0,1,1,1,1,1,0
1,1,0,1,1,1,2,1,1,0,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,0,1,0
1,1,1,1,1,1,0,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,6,1,1,1,1,0,0,0,1,1
1,1,2,1,1,1,0,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,7,1,1,1,1,0,0,0,0,1
1,1,1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,0,1,6,1,1,1,1,0,0,0,0,1
1,1,1,0,0,0,2,1,0,1,1,1,1,2,0,0,1,2,0,0,0,0,1,1,0,1,0,1,6,1,0,1,1,1,0,1,0,1
1,1,2,0,0,0,2,0,0,0,1,1,0,2,0,0,1,1,0,0,1,0,1,1,0,1,0,0,6,1,0,0,1,1,1,1,1,0
1,0,3,0,0,0,2,1,1,0,1,0,0,1,1,0,1,1,1,0,1,1,1,1,1,0,0,0,1,1,0,1,1,1,1,1,0
1,0,4,1,1,1,1,0,1,0,1,0,1,1,1,0,1,2,1,0,1,1,0,0,1,1,1,0,0,1,1,0,1,1,1,1,1,0
0,0,4,1,1,0,1,0,1,2,1,1,0,0,0,0,0,1,2,1,0,1,0,0,0,1,0,2,1,1,1,1,1,1,1,0,1
0,1,1,1,1,1,1,0,1,1,1,0,1,0,1,0,0,1,0,2,0,0,1,1,0,0,1,1,3,0,1,1,1,1,1,1,0,1
0,0,1,0,1,1,3,1,1,1,0,0,0,1,1,0,0,0,0,0,0,1,1,1,1,1,4,0,0,1,1,1,1,1,0,1
1,1,4,0,0,0,0,0,0,1,1,1,1,2,0,1,1,2,1,1,1,0,1,1,1,1,1,5,1,0,1,1,1,1,1,0,0
0,0,3,0,0,0,3,0,0,2,0,0,0,1,1,0,0,0,1,0,0,0,0,1,0,0,1,1,3,1,1,1,1,1,1,1,0,0
1,0,2,1,1,1,1,1,1,1,1,1,1,2,1,0,0,1,0,1,1,1,0,0,1,0,0,1,1,1,1,1,1,1,1,1,0,0
1,1,3,1,1,1,2,0,1,1,1,1,1,2,1,0,1,1,0,1,1,1,0,0,0,0,0,5,1,1,1,0,1,1,1,0,1
1,1,2,0,0,0,1,1,1,0,1,1,0,2,1,1,1,2,1,1,1,1,1,0,1,1,0,6,0,1,1,0,1,1,1,0,1
1,1,2,1,0,0,0,1,1,0,0,1,1,2,0,1,1,2,1,1,1,1,0,1,0,1,1,0,3,1,1,0,1,1,1,1,0,1
1,1,1,1,1,0,0,1,0,1,0,0,1,2,1,1,1,2,1,0,1,1,0,0,1,0,1,0,7,1,1,0,1,1,0,1,0,0
1,0,2,0,1,1,1,1,0,1,1,1,0,1,1,1,1,2,1,0,1,1,1,0,1,0,0,1,6,1,1,0,1,1,0,1,0,0
0,0,3,0,1,1,3,0,0,2,0,1,0,1,0,0,0,0,2,0,0,1,0,0,0,0,1,4,1,0,0,1,1,0,1,0,0
0,0,3,0,0,1,3,0,0,1,0,1,1,1,1,0,0,0,0,2,0,0,1,1,1,1,1,1,2,0,1,0,1,1,1,1,0,1
0,0,1,1,0,1,3,1,0,1,0,0,0,1,0,0,0,1,0,2,0,0,0,1,1,1,1,1,6,0,1,0,1,1,1,1,1,1
1,1,2,0,1,1,2,1,1,1,1,1,1,2,1,0,1,2,1,0,1,1,0,1,1,1,0,0,7,0,0,0,1,1,1,1,1,1
```

Fig 2.3 Data in .csv format

Status, Gender, Age, Stress, Bed-ridden, Anorexia, Climate, Family History, Vit B6 deficiency, Water intake, Salt-intake, Calcium Intake, Hyperparathyroidism, Blood Pressure, Diabetes, Previous kidney stone attacks, Urinary tract blockage, Frequency of Urination, Urine Output, Kidney status, Pain, Cloudy/ Red-colored Urine, Nausea, Vomitting, Sweating, Foul Smelling Urine, Chills, Fever, Drugs, Chocolate, Blacktea, Spinach, Coffee, Oranges, Sweet potatoes, Strawberries, Nuts, Poultry

1,1,2,1,1,1,2,1,1,1,1,1,1,2,1,1,1,1,1,0,1,1,0,1,1,1,1,1,0,1,1,1,1,1,1,1,1

1,1,1,1,1,0,1,1,1,1,1,1,1,2,1,1,1,1,1,0,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,1

1,1,1,1,1,0,2,1,1,0,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,0,0,0,1,1,1,1,1,1,1,1

1,0,2,1,0,0,2,1,0,0,1,1,1,1,1,1,1,2,1,0,1,1,1,1,1,1,0,0,6,1,1,1,1,1,1,1,1

1,0,3,1,0,1,2,1,0,1,1,1,0,0,1,0,0,2,1,1,1,0,0,1,1,1,1,1,6,0,0,1,1,1,1,1,1

1,1,1,1,1,1,0,0,1,1,1,0,0,2,1,0,1,1,1,1,1,0,1,0,0,1,1,1,6,1,0,0,1,1,1,1,1,0

0,0,4,1,0,0,3,0,0,2,0,0,1,0,1,0,0,0,0,2,0,0,0,1,0,1,1,0,2,1,1,0,0,1,0,0,0,0

1,1,1,0,0,1,3,1,1,1,1,1,1,2,0,0,1,1,0,2,1,1,1,1,1,0,1,1,0,1,1,0,1,1,0,0,0,1

0,0,0,0,0,0,3,0,0,2,0,0,0,0,1,0,0,0,1,2,0,0,1,1,1,1,0,0,1,1,1,0,1,1,0,0,0,1

0,0,4,0,0,1,1,0,0,1,1,0,0,1,1,0,0,0,1,1,0,0,1,1,1,1,0,1,1,0,0,1,1,1,0,1,1,1

1,1,2,1,0,0,2,0,1,0,0,1,1,2,0,1,1,2,1,1,1,1,1,1,0,1,1,1,0,1,0,1,1,1,0,1,1,0

1,0,3,0,1,1,2,0,0,0,0,1,1,2,1,1,0,2,1,2,1,1,0,0,0,1,1,1,0,0,0,1,1,1,1,1,1,0

0,0,0,1,1,0,3,1,1,2,0,0,0,1,0,0,0,0,0,2,0,0,0,0,0,0,0,0,3,0,1,1,1,1,1,1,1,0

0,0,0,0,0,1,1,0,0,2,0,1,0,1,0,0,0,0,1,2,0,1,0,0,1,0,1,0,4,1,1,1,1,1,1,1,1,0

1,0,1,1,1,1,1,1,1,1,1,1,0,2,1,1,0,2,0,1,1,1,1,1,1,1,1,1,0,1,1,0,1,1,1,1,1,0

1,1,0,1,1,1,2,1,1,0,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,0,1,0

1,1,1,1,1,1,0,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,1,6,1,1,1,1,0,0,0,1,1

1,1,2,1,1,1,0,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,1,7,1,1,1,1,0,0,0,0,1

1,1,1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,0,1,6,1,1,1,1,0,0,0,0,1

1,1,1,0,0,0,2,1,0,1,1,1,1,2,0,0,1,2,0,0,0,0,1,1,0,1,0,1,6,1,0,1,1,1,0,1,0,1

1,1,2,0,0,0,2,0,0,0,1,1,0,2,0,0,1,1,0,0,1,0,1,1,0,1,0,0,6,1,0,0,1,1,1,1,1,0

1,0,3,0,0,0,2,1,1,0,1,0,0,1,1,0,1,1,1,0,1,1,1,1,1,1,0,0,0,1,1,0,1,1,1,1,1,0

1,0,4,1,1,1,1,0,1,0,1,0,1,1,1,0,1,2,1,0,1,1,0,0,1,1,1,0,0,1,1,0,1,1,1,1,1,0

0,0,4,1,1,0,1,0,1,2,1,1,0,0,0,0,0,0,1,2,1,0,1,0,0,0,1,0,2,1,1,1,1,1,1,1,0,1

0,1,1,1,1,1,1,0,1,1,1,0,1,0,1,0,0,1,0,2,0,0,1,1,0,0,1,1,3,0,1,1,1,1,1,1,0,1

0,0,1,0,1,1,3,1,1,1,0,0,0,1,1,0,0,0,0,0,0,1,1,1,1,1,4,0,0,1,1,1,1,1,0,1

1,1,4,0,0,0,0,0,0,1,1,1,1,2,0,1,1,2,1,1,1,0,1,1,1,1,1,1,5,1,0,1,1,1,1,1,0,0

0,0,3,0,0,0,3,0,0,2,0,0,0,1,1,0,0,0,1,0,0,0,0,1,0,0,1,1,3,1,1,1,1,1,1,1,0,0

1,0,2,1,1,1,1,1,1,1,1,1,1,2,1,0,0,1,0,1,1,1,0,0,1,0,0,1,1,1,1,1,1,1,1,0,0

1,1,3,1,1,1,2,0,1,1,1,1,1,2,1,0,1,1,0,1,1,1,0,0,0,0,0,5,1,1,1,0,1,1,1,0,1

1,1,2,0,0,0,1,1,1,0,1,1,0,2,1,1,1,2,1,1,1,1,1,0,1,1,0,6,0,1,1,0,1,1,1,0,1

1,1,2,1,0,0,0,1,1,0,0,1,1,2,0,1,1,2,1,1,1,1,0,1,0,1,1,0,3,1,1,0,1,1,1,1,0,1

1,1,1,1,1,0,0,1,0,1,0,0,1,2,1,1,1,2,1,0,1,1,0,0,1,0,1,0,7,1,1,0,1,1,0,1,0,0

1,0,2,0,1,1,1,1,0,1,1,1,0,1,1,1,1,2,1,0,1,1,1,0,1,0,0,1,6,1,1,0,1,1,0,1,0,0

0,0,3,0,1,1,3,0,0,2,0,1,0,1,0,0,0,0,2,0,0,1,0,0,0,0,1,4,1,0,0,1,1,0,1,0,0

0,0,3,0,0,1,3,0,0,1,0,1,1,1,1,0,0,0,0,2,0,0,1,1,1,1,1,1,2,0,1,0,1,1,1,1,0,1

0,0,1,1,0,1,3,1,0,1,0,0,0,1,0,0,0,1,0,2,0,0,0,1,1,1,1,1,6,0,1,0,1,1,1,1,1,1

1,1,2,0,1,1,2,1,1,1,1,1,1,2,1,0,1,2,1,0,1,1,0,1,1,1,0,0,7,0,0,0,1,1,1,1,1,1

0,1,2,1,1,0,2,0,1,2,0,0,0,1,0,0,0,0,1,2,0,0,1,1,0,1,0,1,8,1,0,0,1,1,1,1,1,0

1,1,1,1,0,1,1,1,1,1,1,1,1,2,1,1,1,1,1,1,1,1,1,0,1,1,1,0,1,0,1,1,1,1,1,1,0

1,1,1,1,0,0,2,1,1,0,1,0,1,2,0,1,1,1,1,2,1,1,1,1,0,0,1,1,0,1,1,1,1,1,1,1,1,0

1,1,2,0,0,0,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,0,1,0,1,1,1,1,1,1,1,1,0

1,1,3,0,1,1,2,1,1,1,1,1,1,2,1,1,1,2,1,2,1,0,0,0,1,1,0,0,0,1,1,1,1,1,1,1,1,1

0,1,2,1,1,0,3,0,1,2,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,1,3,1,0,1,1,1,1,1,1,1

0,0,2,1,1,0,2,1,0,2,0,0,1,0,0,0,0,0,0,2,0,0,0,0,0,1,1,0,3,1,0,1,1,1,1,1,0,1

0,0,2,0,0,1,1,0,0,2,1,0,0,0,0,0,0,0,0,2,0,0,1,1,1,1,1,0,2,1,1,1,1,1,1,1,0,1

```
1,0,2,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,1,2,1,1,1,1,1,0,1,1,2,1,1,1,1,1,1,1,1,

1,0,1,1,1,1,2,1,1,0,0,1,1,2,1,1,1,2,1,2,1,1,1,1,0,1,0,0,2,1,1,1,1,1,1,1,1,1

1,0,3,0,1,0,0,0,1,0,1,1,1,2,0,1,1,2,1,1,1,1,1,1,0,1,0,1,3,0,1,1,1,1,1,1,1,1

0,0,3,0,0,0,2,0,0,2,0,0,0,1,1,0,0,1,0,2,0,0,0,0,1,1,0,0,4,1,1,1,1,1,1,1,1,1

1,1,1,1,1,1,2,1,1,0,1,1,1,1,1,0,1,2,1,1,1,1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1

1,1,1,1,1,1,2,1,1,0,1,1,1,1,1,0,1,2,1,1,1,1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1

1,1,1,1,1,1,2,1,1,0,1,1,1,1,1,0,1,2,1,1,1,1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1

1,1,1,1,1,1,2,1,1,0,1,1,1,1,1,0,1,2,1,1,1,1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1

1,1,2,1,1,1,2,1,0,0,1,1,1,1,1,0,1,1,1,0,1,1,1,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1

1,1,2,1,1,1,2,1,0,1,,1,1,1,1,0,1,1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1,0,1,1,1

1,1,2,0,1,1,2,1,0,1,1,1,1,1,1,0,1,1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1,0,1,1,0

1,0,2,0,1,1,1,1,0,1,1,1,1,1,1,1,1,2,1,0,1,1,1,1,0,1,1,1,2,1,0,1,1,1,0,0,1,0

1,0,1,0,0,0,0,0,0,1,1,0,1,0,1,1,1,2,1,0,1,1,1,1,0,1,1,1,6,1,0,0,1,0,0,1,1,0

1,0,3,0,0,0,0,1,0,1,1,1,1,0,1,1,1,1,1,0,1,1,1,1,1,1,1,1,6,1,0,1,1,1,0,1,1,0

1,1,3,1,0,0,0,1,1,0,1,1,1,0,1,1,1,2,1,1,1,1,1,1,1,1,0,1,6,1,0,1,1,1,0,0,0,0

1,1,0,1,0,0,0,1,1,0,0,0,1,0,1,1,0,2,1,1,1,1,1,1,1,1,0,1,1,1,0,1,1,0,1,0,0,0

1,1,0,1,1,1,1,1,1,0,0,0,1,1,1,1,0,2,1,1,1,0,1,1,1,1,0,1,0,1,1,0,1,0,1,0,0,1

1,1,0,1,1,1,1,1,1,0,0,0,1,1,1,1,0,2,1,1,1,0,0,1,1,1,1,1,0,1,1,0,1,0,1,1,0,1

1,1,3,1,1,1,2,0,1,0,1,1,0,1,1,1,1,2,1,1,1,1,0,1,1,1,1,1,0,1,0,0,1,0,1,1,0,1

1,0,3,1,1,1,2,0,1,0,1,1,0,1,0,0,1,2,1,2,1,1,1,0,0,0,1,1,1,1,1,0,1,1,1,1,1,1

0,0,3,1,1,1,3,0,1,2,0,1,0,0,0,0,0,0,1,2,0,1,1,1,0,0,1,1,6,1,1,1,1,1,1,1,1,1
```

0,0,3,1,1,1,3,0,1,2,0,0,0,1,0,0,0,1,1,2,0,1,1,0,0,1,0,1,5,1,1,1,1,1,1,1,1

0,0,1,1,0,1,1,0,0,1,1,0,1,0,1,0,0,1,0,2,0,1,0,0,1,1,0,1,2,1,0,1,1,1,0,1,1,1

1,0,4,1,0,1,2,1,0,1,1,1,1,1,0,1,1,2,1,2,1,0,0,1,1,1,0,0,2,1,0,0,1,1,0,1,1,0

1,1,4,1,0,1,2,1,0,1,1,1,1,1,0,1,1,2,1,1,1,1,0,1,1,1,1,0,2,1,0,0,0,1,0,1,1,0

1,1,1,0,0,1,2,1,0,0,1,1,1,0,0,1,1,2,1,2,1,1,0,1,1,1,1,0,2,1,0,0,0,1,0,1,1,0

1,0,1,0,0,0,1,1,0,0,1,1,0,0,1,0,1,2,1,2,1,1,1,1,1,1,0,6,1,0,1,0,1,1,1,1,0

1,0,2,0,1,0,1,1,0,0,1,1,0,1,1,0,1,2,1,2,1,1,1,0,0,0,1,1,0,1,0,1,1,1,1,0,1,1

1,0,3,1,1,0,0,1,0,0,1,1,0,1,1,1,1,2,1,2,1,1,1,0,0,0,1,1,0,1,0,1,1,1,1,0,1,1

1,0,2,1,1,0,0,1,1,0,1,0,0,1,1,1,0,0,1,2,1,1,1,1,1,0,0,1,0,1,0,1,1,1,1,0,1,1

0,0,3,1,1,1,2,0,1,1,1,0,1,0,1,0,0,1,0,1,0,1,0,1,1,0,0,1,6,1,0,1,0,1,1,0,1,1

0,0,0,1,1,0,2,0,1,2,0,0,0,0,1,0,0,1,0,2,0,1,0,1,1,0,1,0,6,1,0,1,0,1,1,1,1,1

1,0,4,0,1,0,1,1,1,1,1,0,0,0,0,1,1,2,1,2,1,1,1,0,0,1,1,0,4,1,1,1,1,1,1,1,1,1

1,0,3,0,1,1,2,1,1,0,1,1,1,1,1,1,0,1,1,2,1,1,0,1,0,1,1,0,6,1,0,1,1,1,1,1,1,1

1,1,2,0,0,0,1,1,1,0,1,1,0,2,1,1,1,2,1,1,1,1,1,1,0,1,1,0,6,1,0,1,1,1,1,1,1,0

0,0,3,0,0,1,3,0,0,1,0,1,1,1,1,0,0,0,0,2,0,0,1,1,1,1,1,1,2,1,0,0,1,1,1,1,1,0

1,0,1,1,1,1,1,1,1,1,1,1,0,2,1,1,0,2,0,1,1,1,1,1,1,1,1,0,1,0,0,0,1,0,1,1,0

1,1,0,1,1,1,2,1,1,0,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,0,1,0,0,0,1,0,1,0,0

1,1,1,1,1,1,0,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,1,6,1,0,0,0,1,0,1,0,0

1,1,1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,0,1,6,1,0,0,0,0,0,1,0,1

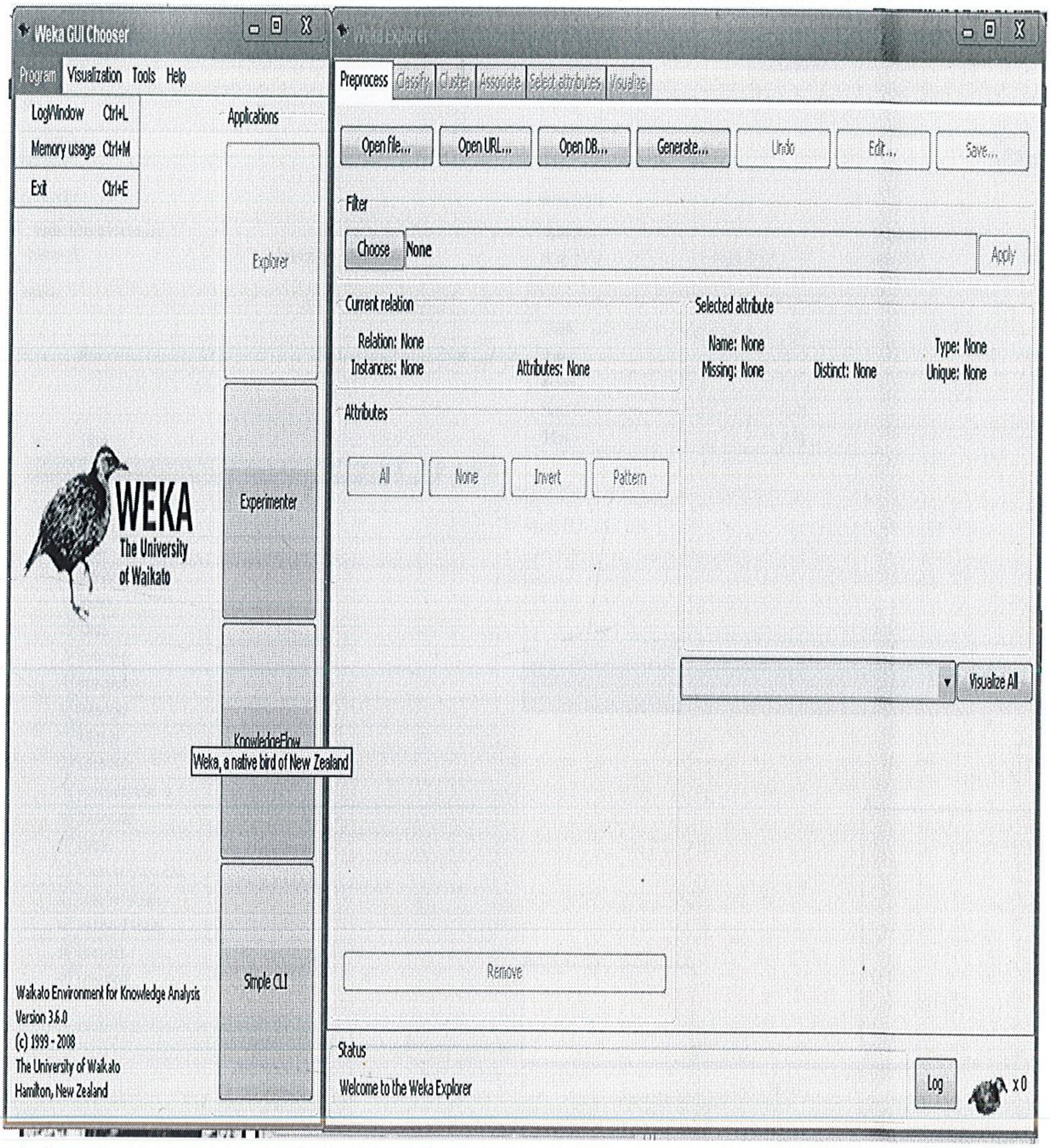
0,0,4,1,1,0,1,0,1,2,1,1,0,0,0,0,0,0,1,2,1,0,1,0,0,0,1,0,2,1,0,0,1,0,1,1,0,1

1,1,2,1,1,1,2,1,1,1,1,1,1,2,1,1,1,1,1,0,1,1,0,1,1,1,1,1,0,1,0,1,1,0,1,1,0,1

```
1,1,1,1,1,0,1,1,1,1,1,1,1,2,1,1,1,1,1,0,1,1,1,1,1,1,1,0,0,1,0,1,1,1,1,1,1,0

1,1,1,1,1,0,2,1,1,0,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,0,0,0,1,1,1,1,1,1,1,1,0

1,0,3,1,0,1,2,1,0,1,1,1,0,0,1,0,0,2,1,1,1,0,0,1,1,1,1,1,6,1,1,1,1,1,1,1,1,0

0,0,4,0,0,1,1,0,0,1,1,0,0,1,1,0,0,0,1,1,0,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,0

1,1,1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,1,1,1,1,1,1,1,1,0,1,6,1,1,1,1,1,1,1,1,1

1,0,4,1,1,1,1,0,1,0,1,0,1,1,1,0,1,2,1,0,1,1,0,0,1,1,1,0,0,1,1,1,1,1,1,1,1,1

1,1,3,1,1,1,2,0,1,1,1,1,1,2,1,0,1,1,0,1,1,1,0,0,0,0,0,0,5,1,1,1,1,1,0,1,1,0

0,0,3,0,0,1,3,0,0,1,0,1,1,1,1,0,0,0,0,2,0,0,1,1,1,1,1,1;2,1,1,0,1,1,0,1,1,0

0,0,1,1,0,1,3,1,0,1,0,0,0,1,0,0,0,1,0,2,0,0,0,1,1,1,1,1,6,1,1,0,1,1,1,0,1,0

1,1,1,1,0,0,2,1,1,0,1,0,1,2,0,1,1,1,1,2,1,1,1,1,0,0,1,1,0,1,1,0,1,1,1,0,1,0

1,1,2,0,0,0,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,0,1,0,1,1,0,1,1,1,0,1,1
```

Input file in .csv format.

Run Weka and select the Explorer



Weka, a native bird of New Zealand

Waikato Environment for Knowledge Analysis
Version 3.6.0
(c) 1999 - 2008
The University of Waikato
Hamilton, New Zealand

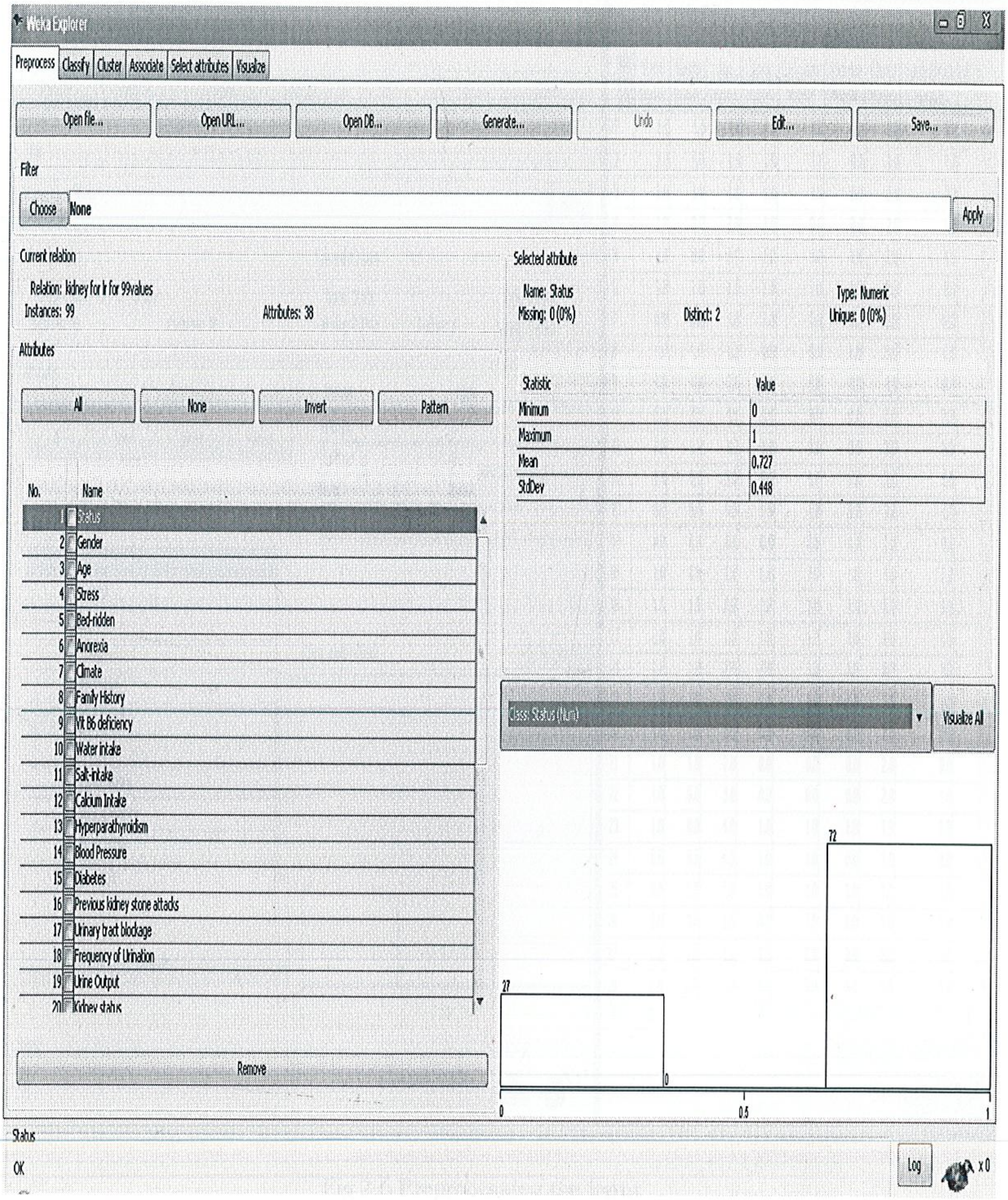
Filter: None

Relation: None
Instances: None
Attributes: None

Name: None
Missing: None
Distinct: None
Type: None
Unique: None

Welcome to the Weka Explorer

Fig 2.4 Weka Interface



Filter: None

Current relation

Relation: kidney for lr for 99values

Instances: 99 Attributes: 38

Selected attribute

Name: Status Type: Numeric

Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

Statistic	Value
Minimum	0
Maximum	1
Mean	0.727
StdDev	0.448

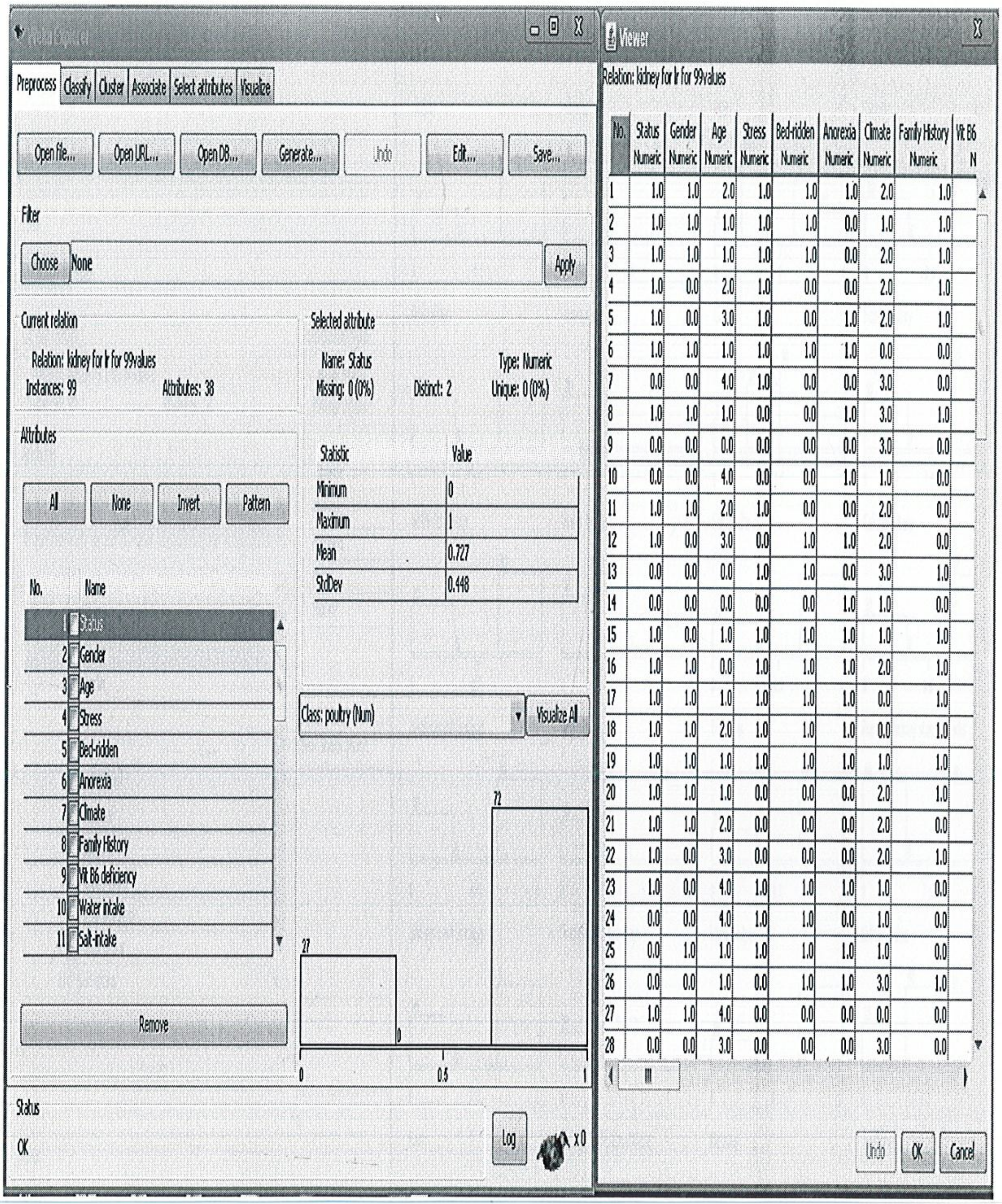
No.		Name
1	☐	Status
2	☐	Gender
3	☐	Age
4	☐	Stress
5	☐	Bed-ridden
6	☐	Anorexia
7	☐	Climate
8	☐	Family History
9	☐	Vit B6 deficiency
10	☐	Water intake
11	☐	Salt-intake
12	☐	Calcium Intake
13	☐	Hyperparathyroidism
14	☐	Blood Pressure
15	☐	Diabetes
16	☐	Previous kidney stone attacks
17	☐	Urinary tract blockage
18	☐	Frequency of Urination
19	☐	Urine Output
20	☐	Kidney status

Class: Status (Num)

Status

OK

Fig 2.5 Loading data into Weka –CSVformat (click on Open file…)



Filter: None

Current relation

Relation: kidney for lr for 99values

Instances: 99 Attributes: 38

Selected attribute

Name: Status Type: Numeric

Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

Statistic	Value
Minimum	0
Maximum	1
Mean	0.727
StdDev	0.448

Attributes

No.	Name
1	Status
2	Gender
3	Age
4	Stress
5	Bed-ridden
6	Anorexia
7	Climate
8	Family History
9	Vit B6 deficiency
10	Water intake
11	Salt-intake

Class: poultry (Num)

Status

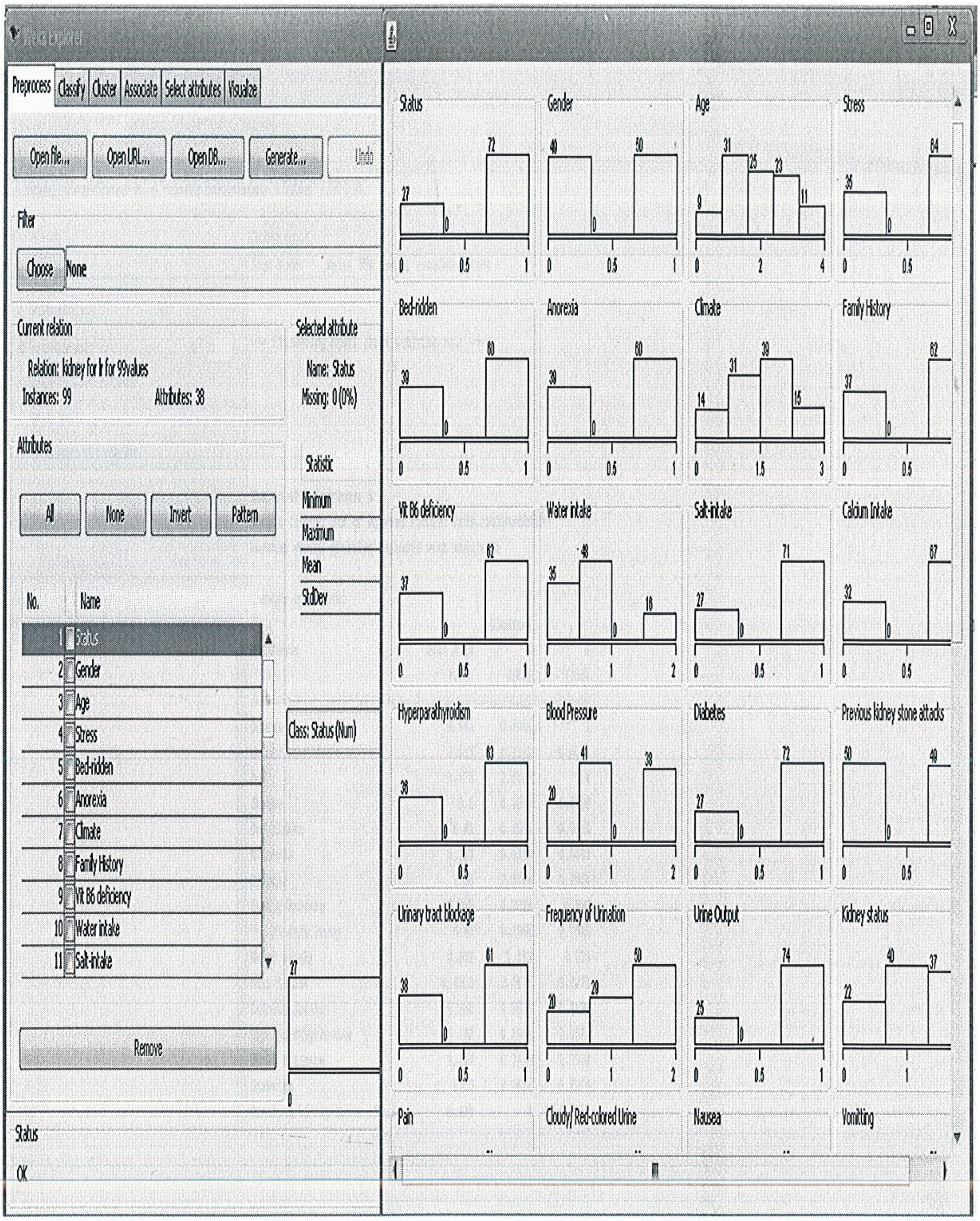
OK

Relation: kidney for lr for 99values

No.	Status Numeric	Gender Numeric	Age Numeric	Stress Numeric	Bed-ridden Numeric	Anorexia Numeric	Climate Numeric	Family History Numeric	Vit B6 N
1	1.0	1.0	2.0	1.0	1.0	1.0	2.0	1.0	
2	1.0	1.0	1.0	1.0	1.0	0.0	1.0	1.0	
3	1.0	1.0	1.0	1.0	1.0	0.0	2.0	1.0	
4	1.0	0.0	2.0	1.0	0.0	0.0	2.0	1.0	
5	1.0	0.0	3.0	1.0	0.0	1.0	2.0	1.0	
6	1.0	1.0	1.0	1.0	1.0	1.0	0.0	0.0	
7	0.0	0.0	4.0	1.0	0.0	0.0	3.0	0.0	
8	1.0	1.0	1.0	0.0	0.0	1.0	3.0	1.0	
9	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	
10	0.0	0.0	4.0	0.0	0.0	1.0	1.0	0.0	
11	1.0	1.0	2.0	1.0	0.0	0.0	2.0	0.0	
12	1.0	0.0	3.0	0.0	1.0	1.0	2.0	0.0	
13	0.0	0.0	0.0	1.0	1.0	0.0	3.0	1.0	
14	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	
15	1.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0	
16	1.0	1.0	0.0	1.0	1.0	1.0	2.0	1.0	
17	1.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0	
18	1.0	1.0	2.0	1.0	1.0	1.0	0.0	1.0	
19	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	
20	1.0	1.0	1.0	0.0	0.0	0.0	2.0	1.0	
21	1.0	1.0	2.0	0.0	0.0	0.0	2.0	0.0	
22	1.0	0.0	3.0	0.0	0.0	0.0	2.0	1.0	
23	1.0	0.0	4.0	1.0	1.0	1.0	1.0	0.0	
24	0.0	0.0	4.0	1.0	1.0	0.0	1.0	0.0	
25	0.0	1.0	1.0	1.0	1.0	1.0	1.0	0.0	
26	0.0	0.0	1.0	0.0	1.0	1.0	3.0	1.0	
27	1.0	1.0	4.0	0.0	0.0	0.0	0.0	0.0	
28	0.0	0.0	3.0	0.0	0.0	0.0	3.0	0.0	

Fig 2.6 Preprocessing the input

Editing data in Weka (click on "Edit"….)



Relation: kidney for lr for 99values

Instances: 99

Attributes: 38

Name: Status

Missing: 0 (0%)

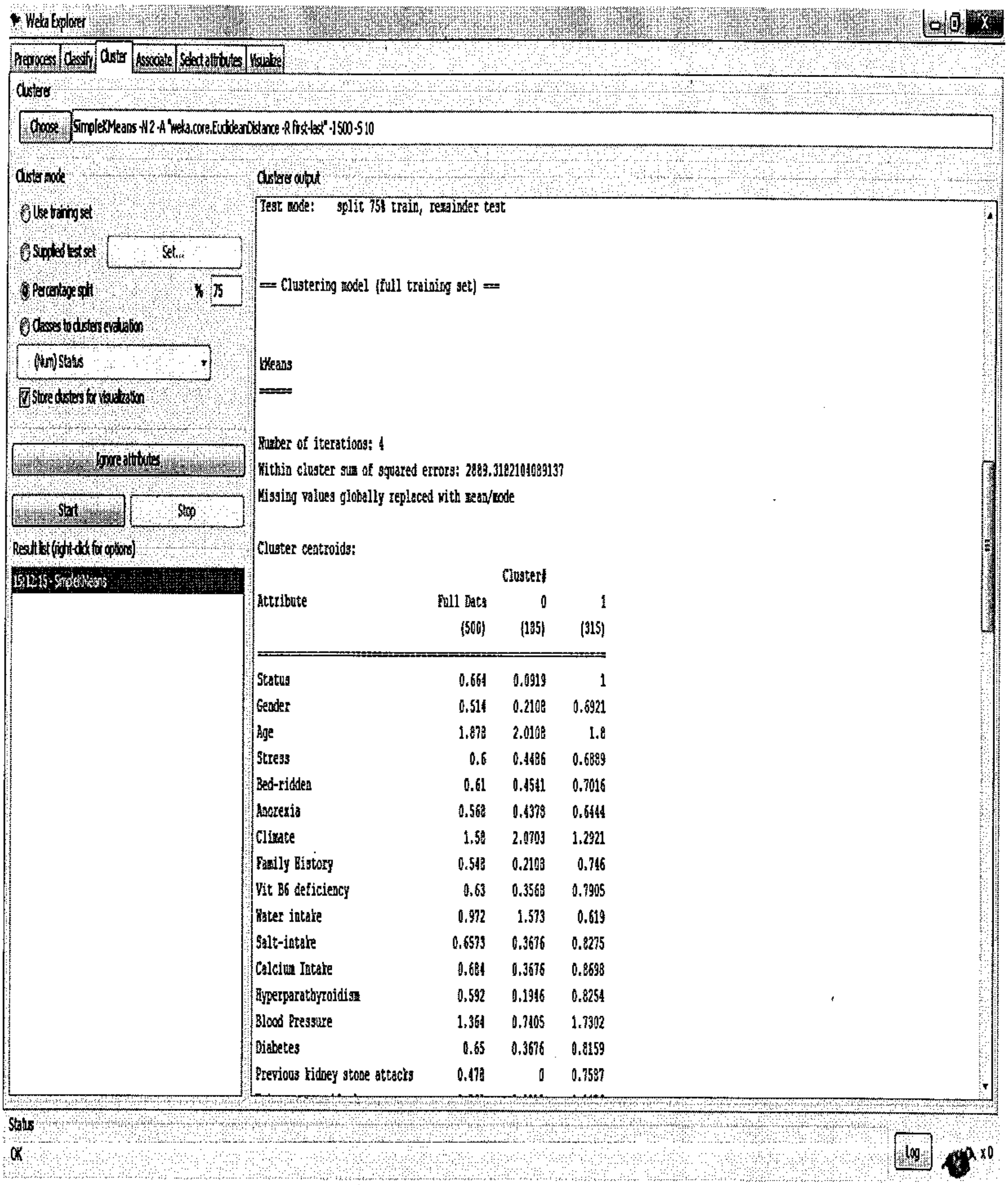
Class: Status (Num)

Status

OK

Fig 2.7 Data pre-processing and visualization. Click on "Visualize All".

Applying k-means clustering tehnique

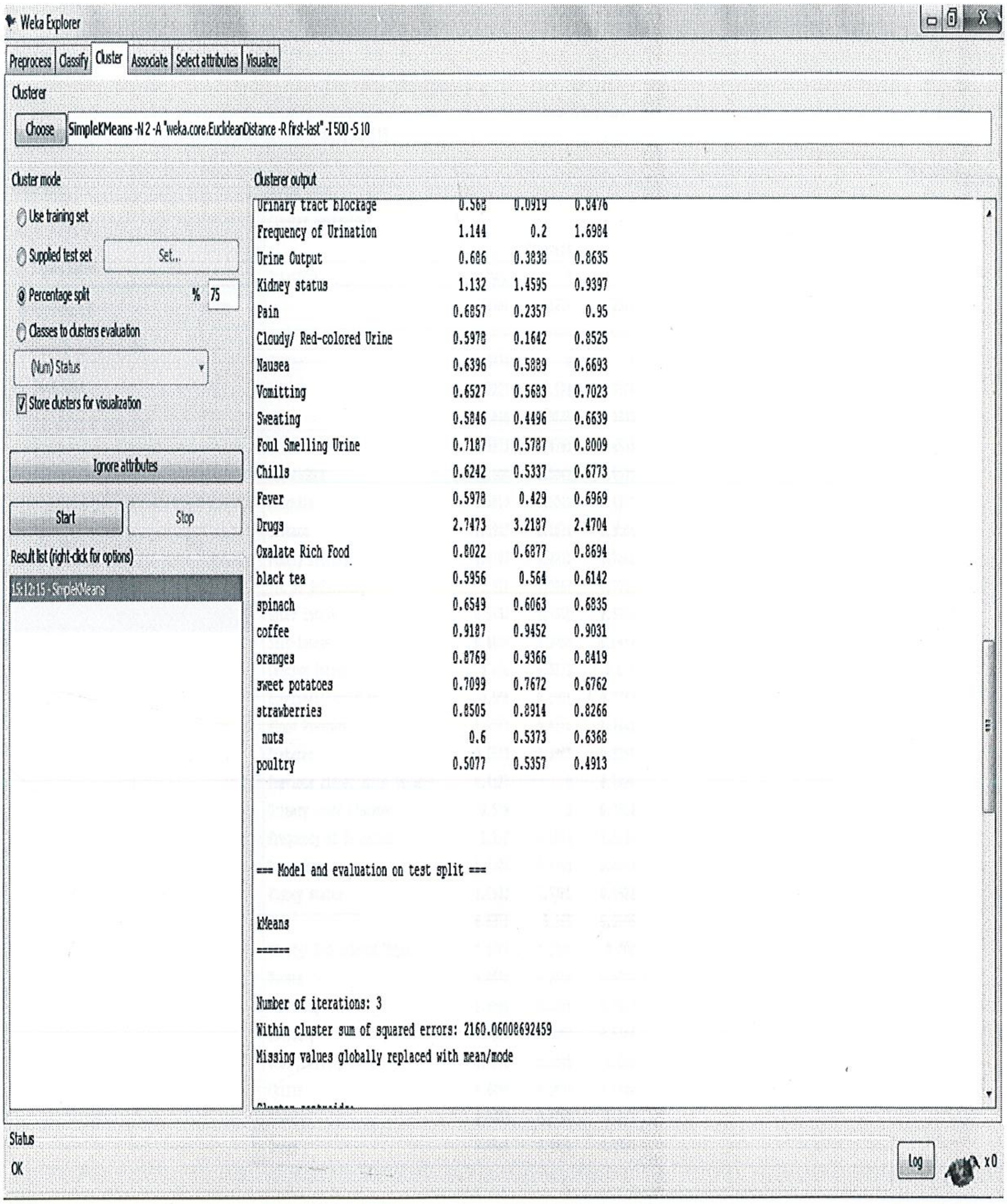


Clusterer: SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10

Cluster mode:
- ○ Use training set
- ○ Supplied test set
- ● Percentage split % 75
- ○ Classes to clusters evaluation
- (Num) Status
- ☑ Store clusters for visualization

```
Test mode:    split 75% train, remainder test



=== Clustering model (full training set) ===



kMeans
======


Number of iterations: 4
Within cluster sum of squared errors: 2889.3182104089137
Missing values globally replaced with mean/mode


Cluster centroids:
                                        Cluster#
Attribute                   Full Data          0          1
                                (500)      (185)      (315)
============================================================

Status                          0.664     0.0919          1
Gender                          0.514     0.2108     0.6921
Age                             1.878     2.0108        1.8
Stress                            0.6     0.4486     0.6889
Bed-ridden                       0.61     0.4541     0.7016
Anorexia                        0.568     0.4378     0.6444
Climate                          1.58     2.0703     1.2921
Family History                  0.548     0.2108      0.746
Vit B6 deficiency                0.63     0.3568     0.7905
Water intake                    0.972      1.573      0.619
Salt-intake                    0.6573     0.3676     0.8275
Calcium Intake                  0.684     0.3676     0.8698
Hyperparathyroidism             0.592     0.1946     0.8254
Blood Pressure                  1.364     0.7405     1.7302
Diabetes                         0.65     0.3676     0.8159
Previous kidney stone attacks   0.478          0     0.7587
```

Fig 2.8(a) Cluster Centroids



SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10

○ Use training set

○ Supplied test set

● Percentage split % 75

○ Classes to clusters evaluation

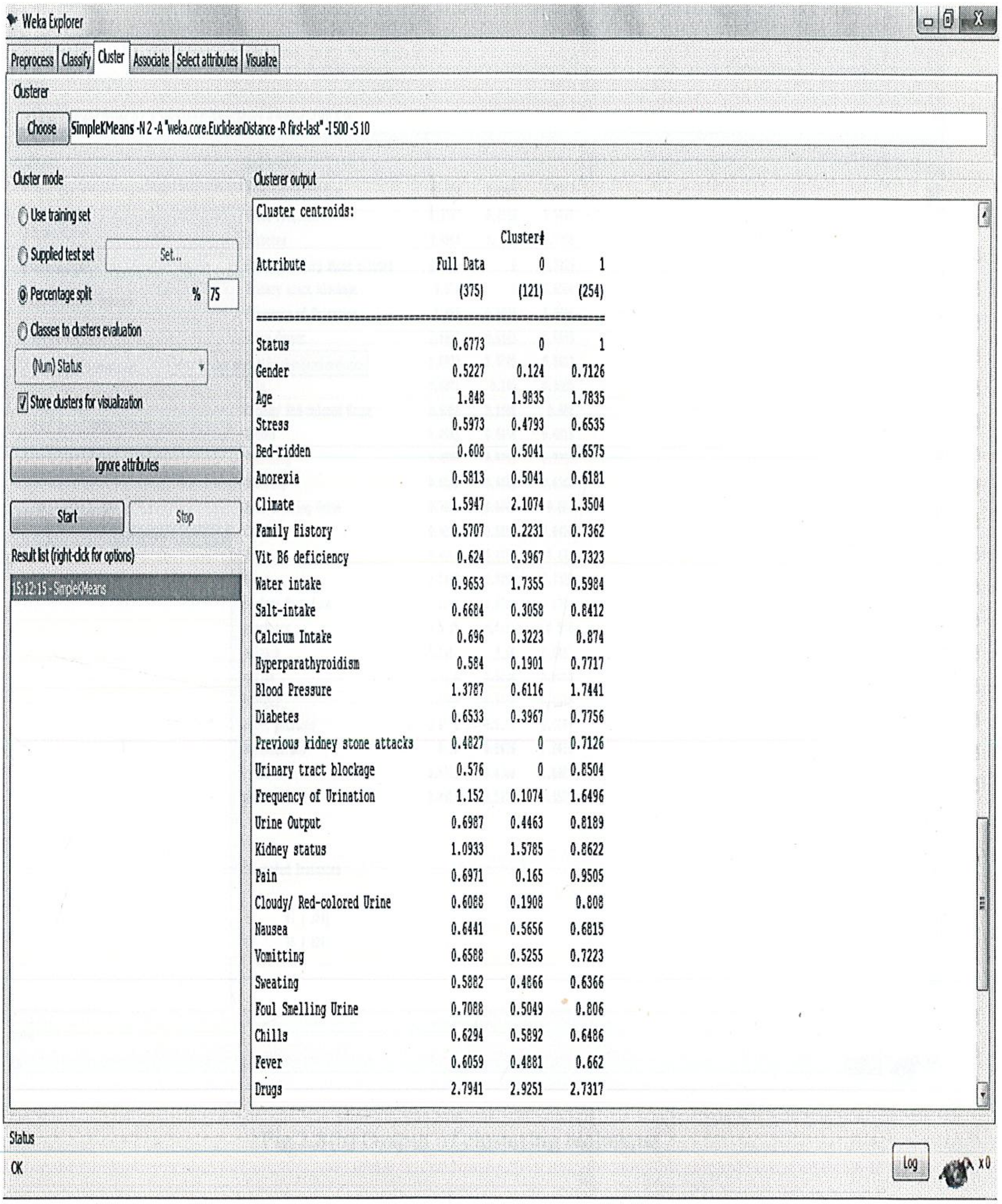
(Num) Status

☑ Store clusters for visualization

15:12:15 - SimpleKMeans

```
Urinary tract blockage             0.568    0.0919    0.8476
Frequency of Urination             1.144       0.2    1.6984
Urine Output                       0.686    0.3838    0.8635
Kidney status                      1.132    1.4595    0.9397
Pain                              0.6857    0.2357      0.95
Cloudy/ Red-colored Urine         0.5978    0.1642    0.8525
Nausea                            0.6396    0.5889    0.6693
Vomitting                         0.6527    0.5683    0.7023
Sweating                          0.5846    0.4496    0.6639
Foul Smelling Urine               0.7187    0.5787    0.8009
Chills                            0.6242    0.5337    0.6773
Fever                             0.5978     0.429    0.6969
Drugs                             2.7473    3.2187    2.4704
Oxalate Rich Food                 0.8022    0.6877    0.8694
black tea                         0.5956     0.564    0.6142
spinach                           0.6549    0.6063    0.6835
coffee                            0.9187    0.9452    0.9031
oranges                           0.8769    0.9366    0.8419
sweet potatoes                    0.7099    0.7672    0.6762
strawberries                      0.8505    0.8914    0.8266
 nuts                                0.6    0.5373    0.6368
poultry                           0.5077    0.5357    0.4913



=== Model and evaluation on test split ===

kMeans
======

Number of iterations: 3
Within cluster sum of squared errors: 2160.06008692459
Missing values globally replaced with mean/mode
```

OK



Clusterer

SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10

Cluster mode

○ Use training set

○ Supplied test set

● Percentage split % 75

○ Classes to clusters evaluation

(Num) Status

☑ Store clusters for visualization

Result list (right-click for options)

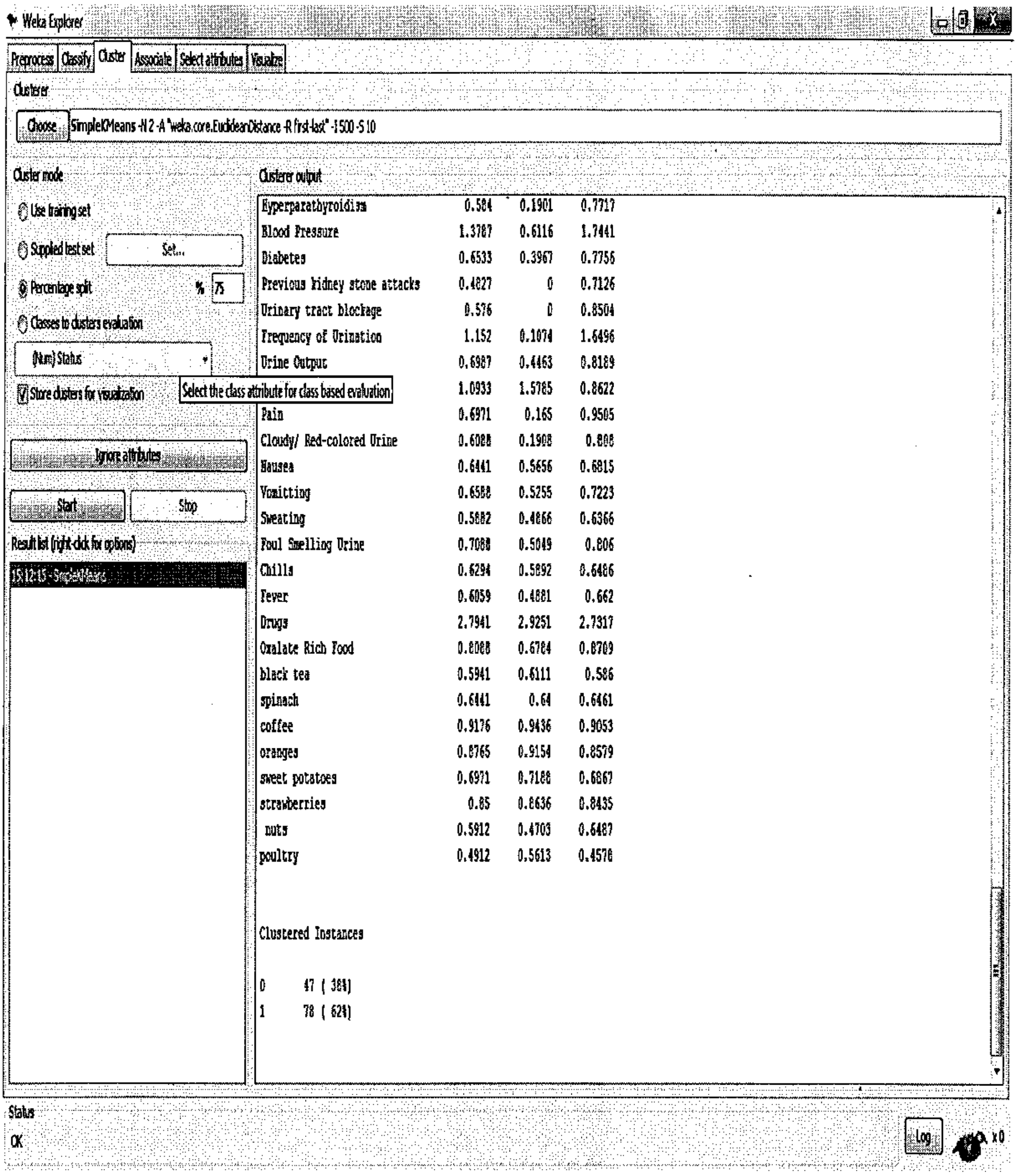
15:12:15 - SimpleKMeans

Clusterer output

```
Cluster centroids:
                                             Cluster#
Attribute                      Full Data          0          1
                                   (375)      (121)      (254)
==============================================================
Status                            0.6773          0          1
Gender                            0.5227      0.124     0.7126
Age                                1.848     1.9835     1.7835
Stress                            0.5973     0.4793     0.6535
Bed-ridden                         0.608     0.5041     0.6575
Anorexia                          0.5813     0.5041     0.6181
Climate                           1.5947     2.1074     1.3504
Family History                    0.5707     0.2231     0.7362
Vit B6 deficiency                  0.624     0.3967     0.7323
Water intake                      0.9653     1.7355     0.5984
Salt-intake                       0.6684     0.3058     0.8412
Calcium Intake                     0.696     0.3223      0.874
Hyperparathyroidism                0.584     0.1901     0.7717
Blood Pressure                    1.3787     0.6116     1.7441
Diabetes                          0.6533     0.3967     0.7756
Previous kidney stone attacks     0.4827          0     0.7126
Urinary tract blockage             0.576          0     0.8504
Frequency of Urination             1.152     0.1074     1.6496
Urine Output                      0.6987     0.4463     0.8189
Kidney status                     1.0933     1.5785     0.8622
Pain                              0.6971      0.165     0.9505
Cloudy/ Red-colored Urine         0.6088     0.1908      0.808
Nausea                            0.6441     0.5656     0.6815
Vomitting                         0.6588     0.5255     0.7223
Sweating                          0.5882     0.4866     0.6366
Foul Smelling Urine               0.7088     0.5049      0.806
Chills                            0.6294     0.5892     0.6486
Fever                             0.6059     0.4881      0.662
Drugs                             2.7941     2.9251     2.7317
```

Status

OK



SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10

○ Use training set
○ Supplied test set
● Percentage split % 75
○ Classes to clusters evaluation
(Num) Status
☑ Store clusters for visualization

Select the class attribute for class based evaluation

15:12:15 - SimpleKMeans

```
Hyperparathyroidism               0.584    0.1901   0.7717
Blood Pressure                    1.3787   0.6116   1.7441
Diabetes                          0.6533   0.3967   0.7756
Previous kidney stone attacks     0.4827        0   0.7126
Urinary tract blockage            0.576         0   0.8504
Frequency of Urination            1.152    0.1074   1.6496
Urine Output                      0.6987   0.4463   0.8189
                                  1.0933   1.5785   0.8622
Pain                              0.6971    0.165   0.9505
Cloudy/ Red-colored Urine         0.6088   0.1908    0.808
Nausea                            0.6441   0.5656   0.6815
Vomitting                         0.6588   0.5255   0.7223
Sweating                          0.5882   0.4866   0.6366
Foul Smelling Urine               0.7088   0.5049    0.806
Chills                            0.6294   0.5892   0.6486
Fever                             0.6059   0.4881    0.662
Drugs                             2.7941   2.9251   2.7317
Oxalate Rich Food                 0.8088   0.6784   0.8709
black tea                         0.5941   0.6111    0.586
spinach                           0.6441     0.64   0.6461
coffee                            0.9176   0.9436   0.9053
oranges                           0.8765   0.9154   0.8579
sweet potatoes                    0.6971   0.7188   0.6867
strawberries                        0.85   0.8636   0.8435
 nuts                             0.5912   0.4703   0.6487
poultry                           0.4912   0.5613   0.4576


Clustered Instances

0       47 ( 38%)
1       78 ( 62%)
```

OK

Fig 2.8(b) Output of clustering technique

Applying 'Linear Regression' technique from 'functions'.

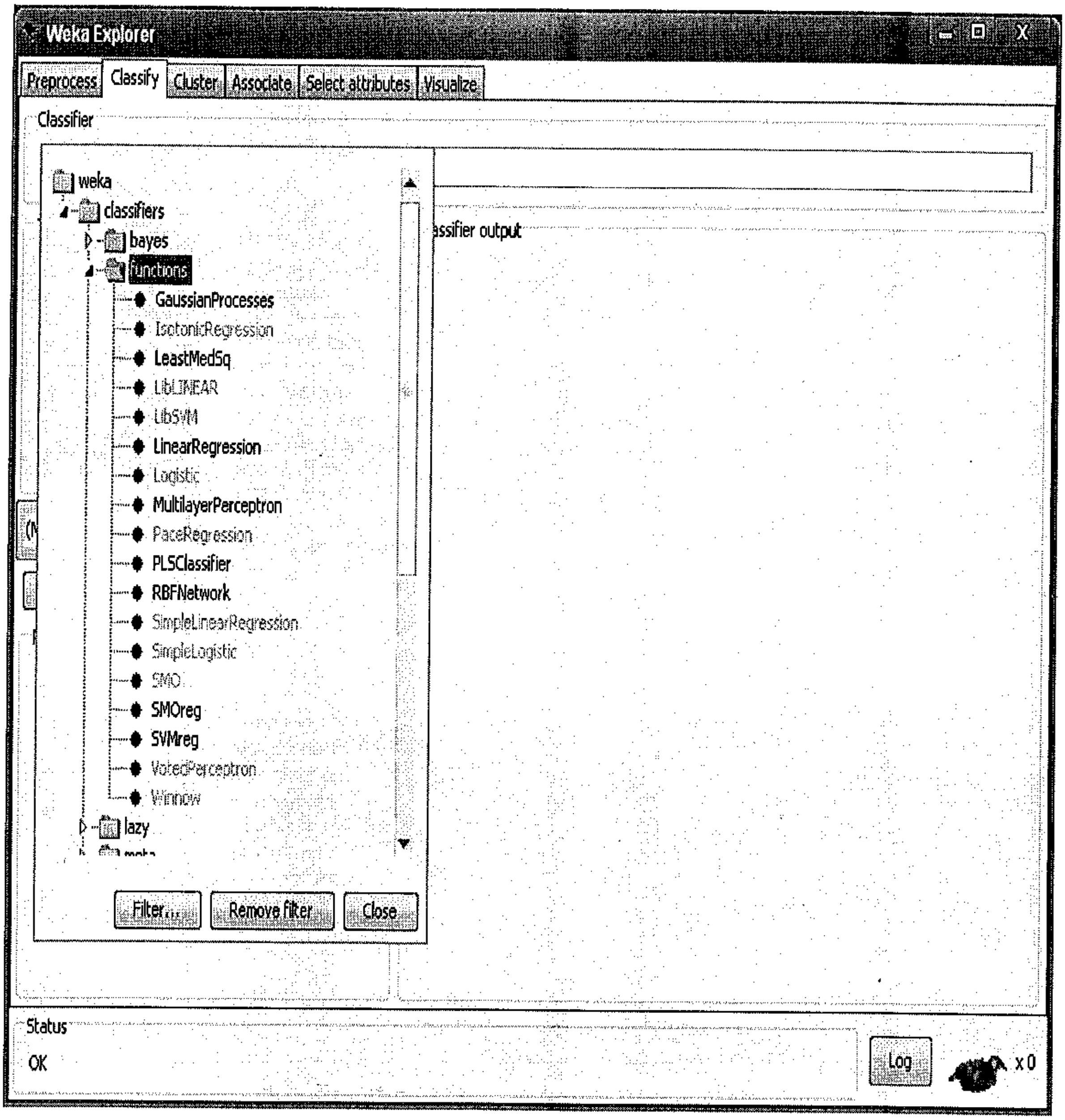
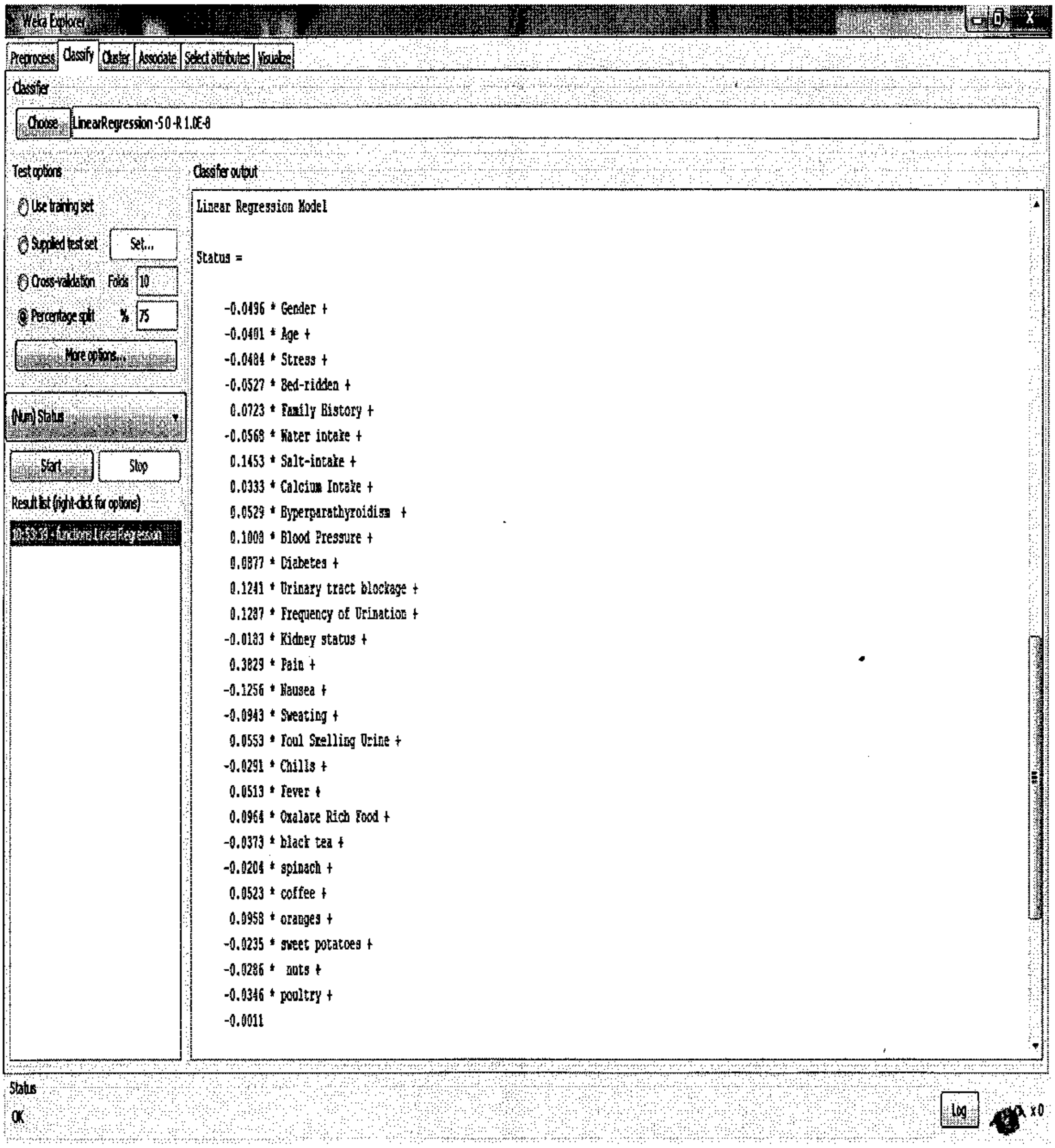


Fig. 2.9 Selection "linear regression" from "functions".

Selecting 75% percentage split on initial 38 attributes and applying linear regression on it.



Classifier

LinearRegression -S 0 -R 1.0E-8

Test options

○ Use training set
○ Supplied test set
○ Cross-validation Folds 10
● Percentage split % 75

(Num) Status

Result list (right-click for options)

10:53:39 - functions.LinearRegression

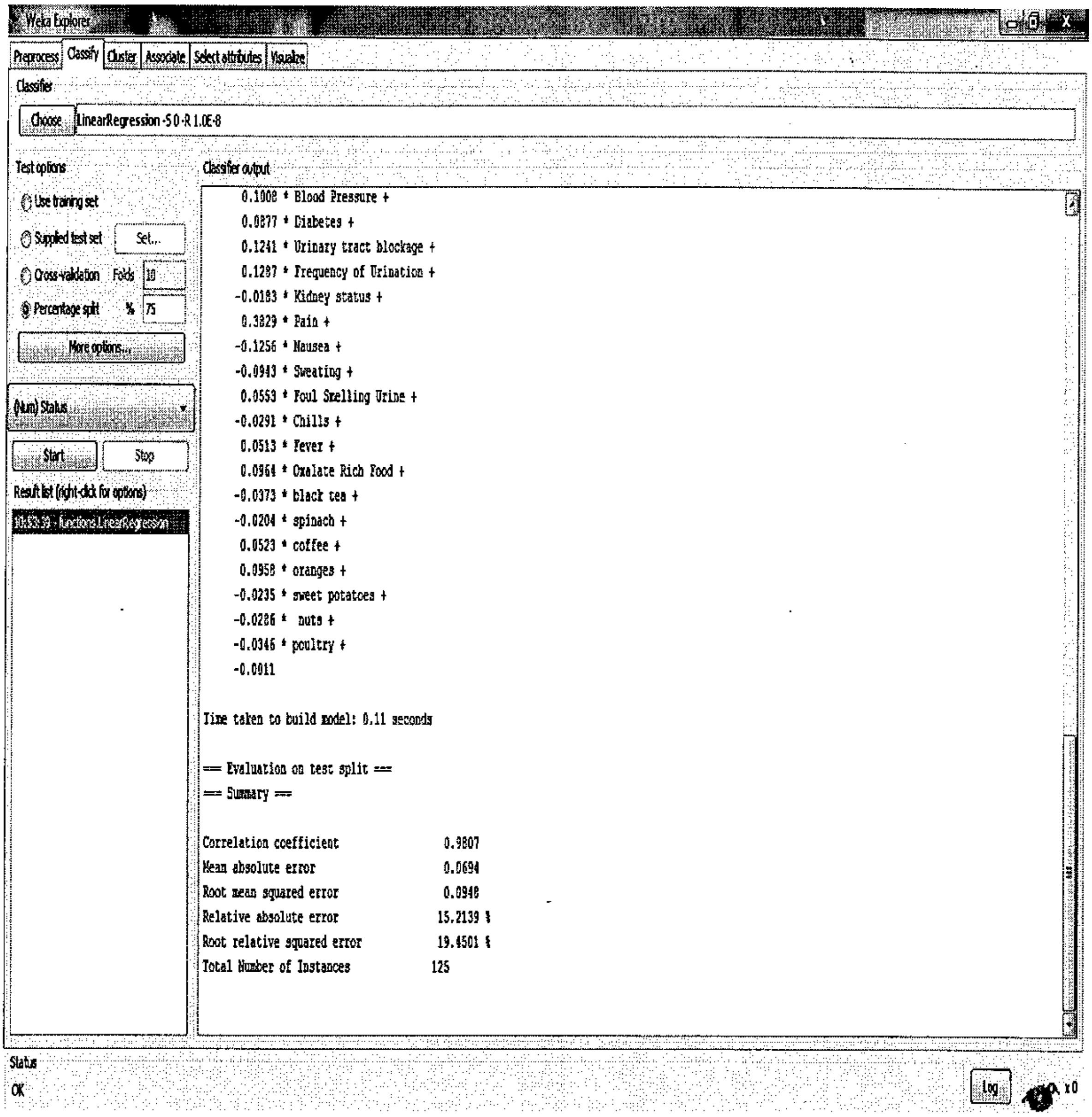
Classifier output

```
Linear Regression Model


Status =


     -0.0496 * Gender +
     -0.0401 * Age +
     -0.0484 * Stress +
     -0.0527 * Bed-ridden +
      0.0723 * Family History +
     -0.0569 * Water intake +
      0.1453 * Salt-intake +
      0.0333 * Calcium Intake +
      0.0529 * Hyperparathyroidism  +
      0.1008 * Blood Pressure +
      0.0877 * Diabetes +
      0.1241 * Urinary tract blockage +
      0.1237 * Frequency of Urination +
     -0.0133 * Kidney status +
      0.3829 * Pain +
     -0.1256 * Nausea +
     -0.0943 * Sweating +
      0.0553 * Foul Smelling Urine +
     -0.0291 * Chills +
      0.0513 * Fever +
      0.0964 * Oxalate Rich Food +
     -0.0373 * black tea +
     -0.0204 * spinach +
      0.0523 * coffee +
      0.0958 * oranges +
     -0.0235 * sweet potatoes +
     -0.0286 *  nuts +
     -0.0346 * poultry +
     -0.0011
```

Status

OK

Fig 2.10 Output of Linear-regression technique



LinearRegression -S 0 -R 1.0E-8

○ Use training set
○ Supplied test set Set...
○ Cross-validation Folds 10
● Percentage split % 75

(Num) Status

10:53:39 - functions.LinearRegression

```
      0.1008 * Blood Pressure +
      0.0877 * Diabetes +
      0.1241 * Urinary tract blockage +
      0.1287 * Frequency of Urination +
     -0.0183 * Kidney status +
      0.3829 * Pain +
     -0.1256 * Nausea +
     -0.0943 * Sweating +
      0.0553 * Foul Smelling Urine +
     -0.0291 * Chills +
      0.0513 * Fever +
      0.0964 * Oxalate Rich Food +
     -0.0373 * black tea +
     -0.0204 * spinach +
      0.0523 * coffee +
      0.0958 * oranges +
     -0.0235 * sweet potatoes +
     -0.0286 *  nuts +
     -0.0346 * poultry +
     -0.0011

Time taken to build model: 0.11 seconds

=== Evaluation on test split ===
=== Summary ===

Correlation coefficient                  0.9807
Mean absolute error                      0.0694
Root mean squared error                  0.0948
Relative absolute error                 15.2139 %
Root relative squared error             19.4501 %
Total Number of Instances              125
```

Status
OK

x 0

Chapter 5

Results and Discussion

5.1 Result of Simple k-mean clustering technique

```
=== Run information ===

Scheme:   weka.clusterers.SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-
last" -I 500 -S 10
Relation:    data for 500
Instances:   500
Attributes:  38
              Status
              Gender
              Age
              Stress
              Bed-ridden
              Anorexia
              Climate
              Family History
              Vit B6 deficiency
              Water intake
              Salt-intake
              Calcium Intake
              Hyperparathyroidism
              Blood Pressure
              Diabetes
              Previous kidney stone attacks
              Urinary tract blockage
              Frequency of Urination
              Urine Output
              Kidney status
              Pain
              Cloudy/ Red-colored Urine
              Nausea
              Vomitting
              Sweating
              Foul Smelling Urine
              Chills
              Fever
              Drugs
              Oxalate Rich Food
              black tea
              spinach
              coffee
              oranges
              sweet potatoes
              strawberries
               nuts
              poultry
Test mode:    split 75% train, remainder test
```

```
=== Clustering model (full training set) ===


kMeans
======

Number of iterations: 4
Within cluster sum of squared errors: 2889.3182104089137
Missing values globally replaced with mean/mode

Cluster centroids:

Attribute               Full Data    0          1
                        (500)        (185)      (315)
=================================================================
Status                  0.664        0.0919     1
Gender                  0.514        0.2108     0.6921
Age                     1.878        2.0108     1.8
Stress                  0.6          0.4486     0.6889
Bed-ridden              0.61         0.4541     0.7016
Anorexia                0.568        0.4378     0.6444
Climate                 1.58         2.0703     1.2921
Family History          0.548        0.2108     0.746
Vit B6 deficiency       0.63         0.3568     0.7905
Water intake            0.972        1.573      0.619
Salt-intake             0.6573       0.3676     0.8275
Calcium Intake          0.684        0.3676     0.8698
Hyperparathyroidism     0.592        0.1946     0.8254
Blood Pressure          1.364        0.7405     1.7302
Diabetes                0.65         0.3676      0.8159
```

Previous attacks	0.478	0	0.7587
Urinary tract blockage	0.568	0.0919	0.8476
Frequency of Urination	1.144	0.2	1.6984
Urine Output	0.686	0.3838	0.8635
Kidney status	1.132	1.4595	0.9397
Pain	0.6857	0.2357	0.95
Cloudy/ Red-colored Urine	0.5978	0.1642	0.8525
Nausea	0.6396	0.5889	0.6693
Vomitting	0.6527	0.5683	0.7023
Sweating	0.5846	0.4496	0.6639
Foul Smelling Urine	0.7187	0.5787	0.8009
Chills	0.6242	0.5337	0.6773
Fever	0.5978	0.429	0.6969
Drugs	2.7473	3.2187	2.4704
Oxalate Rich Food	0.8022	0.6877	0.8694
black tea	0.5956	0.564	0.6142
spinach	0.6549	0.6063	0.6835
coffee	0.9187	0.9452	0.9031
oranges	0.8769	0.9366	0.8419
sweet potatoes	0.7099	0.7672	0.6762
strawberries	0.8505	0.8914	0.8266
nuts	0.6	0.5373	0.6368
poultry	0.5077	0.5357	0.4913

```
=== Model and evaluation on test split ===

k-Means
======

Number of iterations: 3
Within cluster sum of squared errors: 2160.06008692459
Missing values globally replaced with mean/mode

Cluster centroids:
                                  Cluster#
Attribute                          Full Data      0          1
                                     (375)      (121)      (254)
=====================================================================
Status                              0.6773      0          1
Gender                              0.5227      0.124      0.7126
Age                                 1.848       1.9835     1.7835
Stress                              0.5973      0.4793     0.6535
Bed-ridden                          0.608       0.5041     0.6575
Anorexia                            0.5813      0.5041     0.6181
Climate                             1.5947      2.1074     1.3504
Family History                       0.5707     0.2231     0.7362
Vit B6 deficiency                   0.624       0.3967     0.7323
Water intake                        0.9653      1.7355     0.5984
Salt-intake                         0.6684      0.3058     0.8412
Calcium Intake                      0.696       0.3223     0.874
Hyperparathyroidism                 0.584       0.1901     0.7717
Blood Pressure                      1.3787      0.6116     1.7441
Diabetes                             0.6533     0.3967     0.7756
Previous kidney stone attacks       0.4827      0          0.7126
```

Urinary tract blockage	0.576	0	0.8504
Frequency of Urination	1.152	0.1074	1.6496
Urine Output	0.6987	0.4463	0.8189
Kidney status	1.0933	1.5785	0.8622
Pain	0.6971	0.165	0.9505
Cloudy/ Red-colored Urine	0.6088	0.1908	0.808
Nausea	0.6441	0.5656	0.6815
Vomitting	0.6588	0.5255	0.7223
Sweating	0.5882	0.4866	0.6366
Foul Smelling Urine	0.7088	0.5049	0.806
Chills	0.6294	0.5892	0.6486
Fever	0.6059	0.4881	0.662
Drugs	2.7941	2.9251	2.7317
Oxalate Rich Food	0.8088	0.6784	0.8709
black tea	0.5941	0.6111	0.586
spinach	0.6441	0.64	0.6461
coffee	0.9176	0.9436	0.9053
oranges	0.8765	0.9154	0.8579
sweet potatoes	0.6971	0.7188	0.6867
strawberries	0.85	0.8636	0.8435
nuts	0.5912	0.4703	0.6487
poultry	0.4912	0.5613	0.4578

Clustered Instances

0	47 (38%)
1	78 (62%)

5.2 Result of Linear Regression technique

```
=== Run information ===

Scheme:     weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8
Relation:   kidney for lr for 500values
Instances:  500
Attributes: 38
            Status
            Gender
            Age
            Stress
            Bed-ridden
            Anorexia
            Climate
            Family History
            Vit B6 deficiency
            Water intake
            Salt-intake
            Calcium Intake
            Hyperparathyroidism
            Blood Pressure
            Diabetes
            Previous kidney stone attacks
            Urinary tract blockage
            Frequency of Urination
            Urine Output
            Kidney status
            Pain
            Cloudy/ Red-colored Urine
            Nausea
            Vomitting
            Sweating
            Foul Smelling Urine
            Chills
            Fever
            Drugs
            Chocolate
            black tea
            spinach
            coffee
            oranges
            sweet potatoes
            strawberries
            nuts
            poultry
```

```
Test mode:    split 75.0% train, remainder test

=== Classifier model (full training set) ===


Linear Regression Model

Status =
    -0.0496 * Gender +
    -0.0401 * Age +
    -0.0484 * Stress +
    -0.0527 * Bed-ridden +
     0.0723 * Family History +
    -0.0568 * Water intake +
     0.1453 * Salt- intake +
     0.0333 * Calcium Intake +
     0.0529 * Hyperparathyroidism  +
     0.1008 * Blood Pressure +
     0.0877 * Diabetes +
     0.1241 * Urinary tract blockage +
     0.1287 * Frequency of Urination +
    -0.0183 * Kidney status +
     0.3829 * Pain +
    -0.1256 * Nausea +
    -0.0943 * Sweating +
     0.0553 * Foul Smelling Urine +
    -0.0291 * Chills +
     0.0513 * Fever +
     0.0964 * Oxalate Rich Food +
    -0.0373 * black tea +
    -0.0204 * spinach +
     0.0523 * coffee +
     0.0958 * oranges +
    -0.0235 * sweet potatoes +
    -0.0286 *  nuts +
    -0.0346 * poultry +
    -0.0011
```

```
Time taken to build model: 0.11 seconds

=== Evaluation on test split ===
=== Summary ===

Correlation coefficient             0.9807
Mean absolute error                 0.0694
Root mean squared error             0.0948
Relative absolute error            15.2139 %
Root relative squared error        19.4501 %
Total Number of Instances          125
```

Conclusion

Firstly the simple k-means clustering method was applied on our data. The limitation of that method was that the prediction can be made of the whole population only rather than getting the results of an individual. To overcome this limitation, the classification technique linear regression was applied. As the outcome, an equation was found. Putting the values in the equation will give the probability of kidney stone of the particular individual.

Bibliography

1. ERIC N. TAYLOR, MEIR J. STAMPFER, and GARY C. CURHAN, "Dietary Factors and the Risk of Incident Kidney Stones in Men: New Insights after 14 Years of Follow-up", J Am Soc Nephrol 15: 3225–3232, 2004

2. Gary C. Curhan, MD, ScD; Walter C. Willett, MD, DrPH; Eric L. Knight, MD, MPH; Meir J. Stampfer, MD, DrPH, "Dietary Factors and the Risk of Incident Kidney Stones in Younger Women", Arch Intern Med. 2004;164:885-891

3. MaryFran R. Sowers, Mary Jannausch, Craig Wood, Sandra K. Pope, Laurie L. Lachance, and Brenda Peterson, "Prevalence of Renal Stones in a Population-based Study with Dietary Calcium, Oxalate, and Medication Exposures", American Journal of Epidemiology, Vol. 147, No. 10

4. GARY C. CURHAN, WALTER C. WILLETT, ERIC B. RIMM, FRANK E. SPEIZER, and MEIR J.STAMPFER, "Body Size and Risk of Kidney Stones", Journal of the American Society of Nephrology, J Am Soc Nephrol 9: 1645-1652, 1998

5. Eric N. Taylor, MD; Meir J. Stampfer, MD, DrPH; Gary C. Curhan, MD, ScD , "Obesity, Weight Gain, and the Risk of Kidney Stones ", JAMA: The Journal Of The American Medical Association, Vol. 293 No. 4, January 26, 2005

6. Gary C. Curhan, Walter C. Willett, Eric B. Rimm, and Meir J. Stampfer , "A Prospective Study of Dietary Calcium and Other Nutrients and the Risk of Symptomatic Kidney Stones", NEJM: The New England Journal of Medicine, Volume 328:833-838 March 25, 1993 Number 12

7. Michelle J. Seminsa, Andrew D. Shorea, Martin A. Makarya, Thomas Magnusona, Roger Johnsa and Brian R. Matlaga, "The Association of Increasing Body Mass Index and Kidney Stone Disease ", The Journal of Urology, Volume 183, Issue 2, February 2010, Pages 571-575.