IRJSE International Research Journal of Science & Engineering
Int. Res. J. of Science & Engineering, 2020; Special Issue A7: 843-859 SJIF Impact Factor 6.68
ISSN: 2322-0015
RESEARCH ARTICLE **OPEN ACCESS**

Application of Preprocess and Association on Chronic Kidney Disease Data Analysis for Clustering and Classification algorithm

Walse Rajesh S[1], Kurundkar GD[2] and Bhalchandra Parag[1]

[1]School of Computational Sciences, Swami Ramanand Teerth Marathwada University, Nanded, India
Department of Computer Science, Shri Gurubuddhi Swami Mahavidyalaya, Purna Dist. Parbhani, S.R.T.M. University, Nanded, India
Email: [1]rajeshwalse@gmail.com | [2]gajanan.kurundkar@gmail.com

Manuscript Details
Available online on http://www.irjse.in
ISSN: 2322-0015

Cite this article as:
Walse Rajesh S, Kurundkar GD and Bhalchandra Parag. Application of Preprocess and Association on Chronic Kidney Disease Data Analysis for Clustering and Classification algorithm, *Int. Res. Journal of Science & Engineering,* February, 2020, Special Issue A7 : 843-859.

ABSTRACT

In WEKA preprocess, in this research, we are explaining how to explore and analyze the Chronic Kidney related data in the Apriori associator. In this algorithm, the minimum matrix or confidence value in the association rule mining supporter, confidence number of cycles performing the role of preparing Rules. This research is carried out by formatting and found ten best rules. The rules create x belongs to y attributes; the constant output of Apriori is to set the best rules by using its value and over caste, and its output shows the rules in the form of the model, we can say model rule. Based on rules, describe every rule — the class implementing an Apriori type algorithm. During the execution of the Apriori algorithm, minimum support 0.75 (301 instances), if minimum metric <confidence> level is less than 0.9, then Rule not suitable to apply further; therefore, this does not rule. The association is always on applied data. Moreover, Apriori is to applied to Nominal data or Binary Data.

Keywords: Data Mining, Clustering, Classification, preprocess, Association, Apriori, WEKA

INTRODUCTION

Many computer science techniques of data mining and machine learning used to study the influence of various parameters and make predictions of the based on different data sets. The data mining technique is the process of identifying the hidden patterns from large and complex data. It may provide a crucial role in decision

making for complex not only agriculture but also health-related problems.

This research paper experiment was carried out on chronic kidney disease patient, Todays chronic kidney disease patient in India is increased day today because of their lifestyle, homemade food is the best food, earlier chronic or acute kidney disease patient is very less, and kidney fails growth rate is very less i.e., on the basis of different health problem maximum kidney problem was increased on the basis of Hypertension, and Blood sugar are the kidney patients. Therefore, nowadays lots of Nephrologist doctors and Medicine Industries is to be used to predict kidney problems of patients on the basis of different Machine Learning algorithms, many researchers are working on kidney patients database, once doctor has predicted the leading cause of kidney damage based on Machine Learning tools by applying Clustering and Classification techniques, once doctor predict the cause. Definitely, this will help kidney patients and control further damages of the kidney. Data mining plays an active role in predicting future kidney-related health problems. The main intention of the Experiment is using the WEKA tool first preprocess the data based on Current relation, attributes, selected attributes, class, and visualize each and every attribute using their Missing value, distinct, type, and unique. Similarly, the researcher uses the association Apriori algorithm to discover new rules; therefore, this type of formed rules and preprocess will help to propose the best model with higher accuracy for Clustering and Classification of data.

Data Fusion sampling multi-resolution analysis – De-noising Feature-Extraction Normalization-- Dimension reduction - Classification Clustering – Visualization Validation

Data mining is the procedure of using huge data sets to infer important hidden knowledge. (fig.1) shows that knowledge discovery data mining process is divided into seven methods:

- Data cleaning
- Data Integration
- Data Selection
- Data transformation
- Data Mining
- Pattern estimation
- Knowledge display

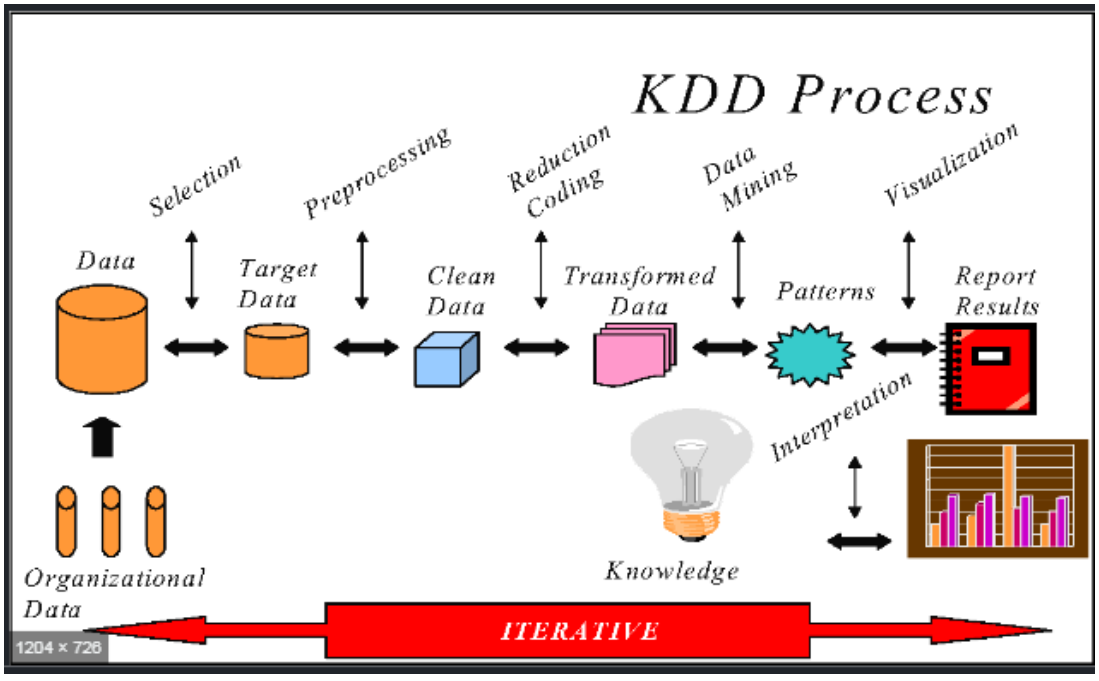


Figure 1: KDD Process in Data Mining

Work Flow Diagram:

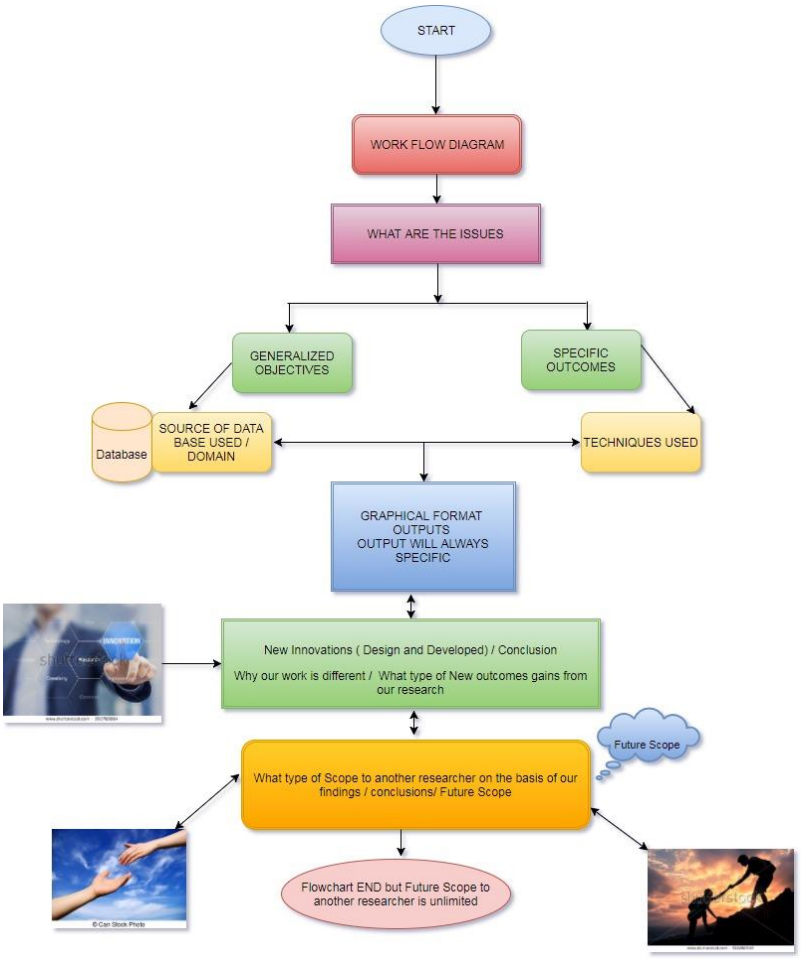


Figure 2: shows the research work flow

II Methodology:

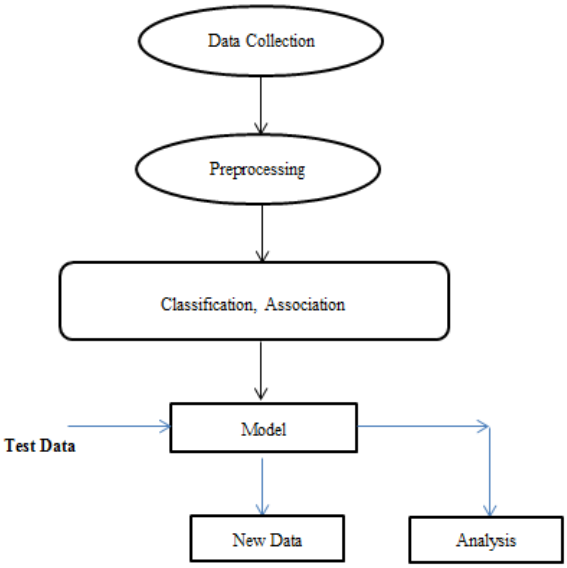


Figure 3: Flowchart displays the KDD

Experiment setup:

The Chronic kidney failure disease dataset has collected from standard data set on the dataset training database using WEKA to explore the application and process and classify the database irrespective of their attributes and their class. On that basis, we are getting the exact graphical output based on their classification of all 26 attributes of 401 instances.

Again, we are applying association for discovering association rule. These generated rules will give the direction for further research as well as the best rule for further clustering and classification of data for prediction means how to apply relation to chronic kidney disease patients based on setting rules and as well as based on probability. If the maximum probability, choose that rule, if the Minimum metric <confidence> level is <= 0.9, then the rule is not to fit for applying further.

1. Data set :

The clinical data of 401 records considered for analysis has taken from the UCI Machine Learning Repository. The data obtained after cleaning and removing missing values. The data has implemented using Rapid Miner Tool. There are 25 attributes in the dataset. The numerical attributes include age, blood pressure, blood glucose random, blood urea, serum creatinine, sodium, potassium, hemoglobin, packaged cell volume, WBC count, RBC count. The nominal attributes include specific gravity, albumin, sugar, RBC, Pus cell, Pus cell clumps, bacteria, hypertension, diabetes mellitus, coronary artery disease, appetite, pedal edema, anemia and class (CKD and Non-CKD)

Number of instances: 400, number of attributes: 25 class (CKD, Non-CKD)

Missing attribute value: yes

Class distribution (63% for CKD and 37% for not CKD).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Serial No.	Age	Blood Pre	Specific gr	Albumin	Sugar	Red Blood	Pus Cell	Pus Cell C	Bacteria	Blood Glu	Blood Ure	Serum Cre	Sodium	Potassium	Hemoglot	Packed Ce	White Blo	Red Blood	Hyperten	Diabet
2	Sr. No.	'age'	'bp'	'sg'	'al'	'su'	'rbc'	'pc'	'pcc'	'ba'	'bgr'	'bu'	'sc'	'sod'	'pot'	'hemo'	'pcv'	'wbcc'	'rbcc'	'htn'	'dm'
3	1	48	80	1.02	1	0	?	normal	notpreser	notpreser	121	36	1.2	?	?	15.4	44	7800	5.2	yes	yes
4	2	7	50	1.02	4	0	?	normal	notpreser	notpreser	?	18	0.8	?	?	11.3	38	6000	?	no	no
5	3	62	80	1.01	2	3	normal	normal	notpreser	notpreser	423	53	1.8	?	?	9.6	31	7500	?	no	yes
6	4	48	70	1.005	4	0	normal	abnormal	present	notpreser	117	56	3.8	111	2.5	11.2	32	6700	3.9	yes	no
7	5	51	80	1.01	2	0	normal	normal	notpreser	notpreser	106	26	1.4	?	?	11.6	35	7300	4.6	no	no
8	6	60	90	1.015	3	0	?	?	notpreser	notpreser	74	25	1.1	142	3.2	12.2	39	7800	4.4	yes	yes
9	7	68	70	1.01	0	0	?	normal	notpreser	notpreser	100	54	24	104	4	12.4	36	?	?	no	no
10	8	24	?	1.015	2	4	normal	abnormal	notpreser	notpreser	410	31	1.1	?	?	12.4	44	6900	5	no	yes
11	9	52	100	1.015	3	0	normal	abnormal	present	notpreser	138	60	1.9	?	?	10.8	33	9600	4	yes	yes
12	10	53	90	1.02	2	0	abnormal	abnormal	present	notpreser	70	107	7.2	114	3.7	9.5	29	12100	3.7	yes	yes
13	11	50	60	1.01	2	4	?	abnormal	present	notpreser	490	55	4	?	?	9.4	28	?	?	yes	yes
14	12	63	70	1.01	3	0	abnormal	abnormal	present	notpreser	380	60	2.7	131	4.2	10.8	32	4500	3.8	yes	yes
15	13	68	70	1.015	3	1	?	normal	present	notpreser	208	72	2.1	138	5.8	9.7	28	12200	3.4	yes	yes
16	14	68	70	?	?	?	?	?	notpreser	notpreser	98	86	4.6	135	3.4	9.8	?	?	?	yes	yes
17	15	68	80	1.01	3	2	normal	abnormal	present	present	157	90	4.1	130	6.4	5.6	16	11000	2.6	yes	yes
18	16	40	80	1.015	3	0	?	normal	notpreser	notpreser	76	162	9.6	141	4.9	7.6	24	3800	2.8	yes	no
19	17	47	70	1.015	2	0	?	normal	notpreser	notpreser	99	46	2.2	138	4.1	12.6	?	?	?	no	no
20	18	47	80	?	?	?	?	?	notpreser	notpreser	114	87	5.2	139	3.7	12.1	?	?	?	yes	no
21	19	60	100	1.025	0	3	?	normal	notpreser	notpreser	263	27	1.3	135	4.3	12.7	37	11400	4.3	yes	yes
22	20	62	60	1.015	1	0	?	abnormal	present	notpreser	100	31	1.6	?	?	10.3	30	5300	3.7	yes	no
23	21	61	80	1.015	2	0	abnormal	abnormal	notpreser	notpreser	173	148	3.9	135	5.2	7.7	24	9200	3.2	yes	yes
24	22	60	90	?	?	?	?	?	notpreser	notpreser	?	180	76	4.5	?	10.9	32	6200	3.6	yes	yes
25	23	48	80	1.025	4	0	normal	abnormal	notpreser	notpreser	95	163	7.7	136	3.8	9.8	32	6900	3.4	yes	no

Figure 4: shows the .csv chronic kidney data set

2. Tool used: WEKA is a collection of machine learning algorithms for performing the data mining task.

Figure 5: shows the main page of WEKA tool

Material and Method:

1. Relevant Information: (total No of 25 Attributes)

age	-	age
bp	-	blood pressure
sg	-	specific gravity
al	-	albumin
su	-	sugar
rbc	-	red blood cells
pc	-	pus cell
pcc	-	pus cell clumps
ba	-	bacteria
bgr	-	blood glucose random
bu	-	blood urea
sc	-	serum creatinine
sod	-	sodium
pot	-	potassium
hemo	-	hemoglobin
pcv	-	packed cell volume
wc	-	white blood cell count
rc	-	red blood cell count
htn	-	hypertension
dm	-	diabetes mellitus
cad	-	coronary artery disease
appet	-	appetite
pe	-	pedal edema
ane	-	anemia
class	-	class

1. .Number of Instances: 400 (250 CKD, 150 notckd)
2. .Number of Attributes: 24 + class = 25 (11 numeric , 14 nominal)
3. Attribute Information :

    1. Age(numerical) age in years

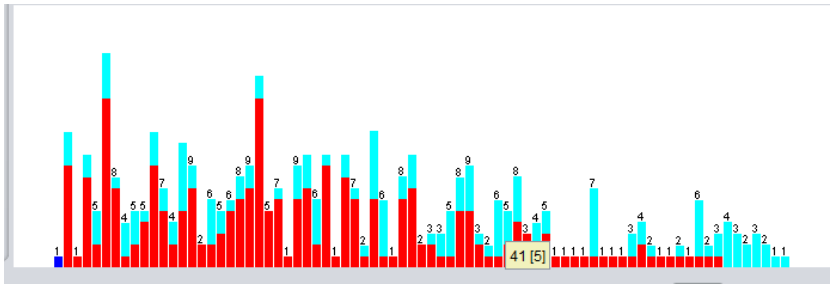
2. Blood Pressure(numerical)
   bp in mm/Hg
3. Specific Gravity(nominal)
   sg - (1.005,1.010,1.015,1.020,1.025)
4. Albumin(nominal)
   al - (0,1,2,3,4,5)
5. Sugar(nominal)
   su - (0,1,2,3,4,5)
6. Red Blood Cells(nominal)
   rbc - (normal,abnormal)
7. Pus Cell (nominal)
   pc - (normal,abnormal)
8. Pus Cell clumps(nominal)
   pcc - (present,notpresent)
9. Bacteria(nominal)
   ba - (present,notpresent)
10. Blood Glucose Random(numerical)
    bgr in mgs/dl
11. Blood Urea(numerical)
    bu in mgs/dl
12. Serum Creatinine(numerical)
    sc in mgs/dl
13. Sodium(numerical)
    sod in mEq/L
14. Potassium(numerical)
    pot in mEq/L
15. Hemoglobin(numerical)
    hemo in gms
16. Packed Cell Volume(numerical)
17. White Blood Cell Count(numerical)
    wc in cells/cumm
18. Red Blood Cell Count(numerical)
    rc in millions/cmm
19. Hypertension(nominal)
    htn - (yes,no)
20. Diabetes Mellitus(nominal)
    dm - (yes,no)
21. Coronary Artery Disease(nominal)
    cad - (yes,no)
22. Appetite(nominal)
    appet - (good,poor)
23. Pedal Edema(nominal)
    pe - (yes,no)
24. Anemia(nominal)
    ane - (yes,no)
25. Class (nominal)
    class - (ckd,notckd)

5. Missing Attribute Values: Yes(Denoted by "?")
6. Class Distribution: (2 classes)

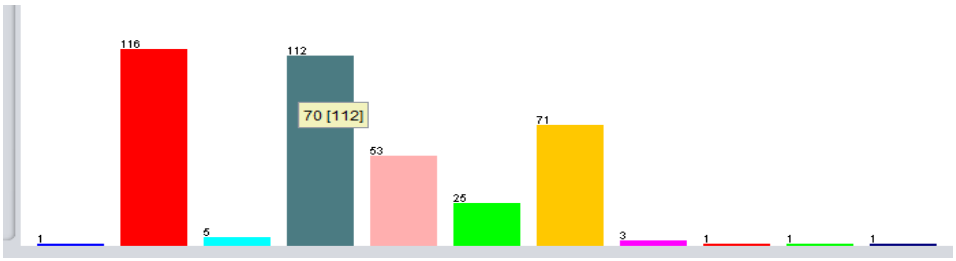
Class	Number of instances
ckd	250
notckd	150

RESULTS AND DISCUSSION:

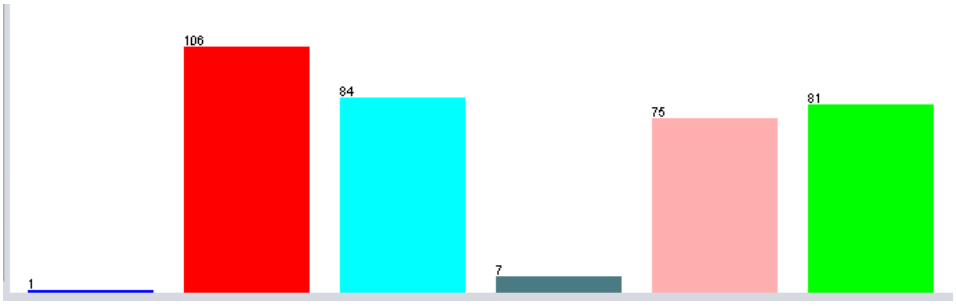
Preprocess for Classification :
Relation :- Chronic_kideny_disease Attributes : 26
Instances : 401 and sum of weights : 401
[1] Attribute : Age Missing : 9(2%) , Distinct : 77, Type : Nominal, , Unique: 17(4%).



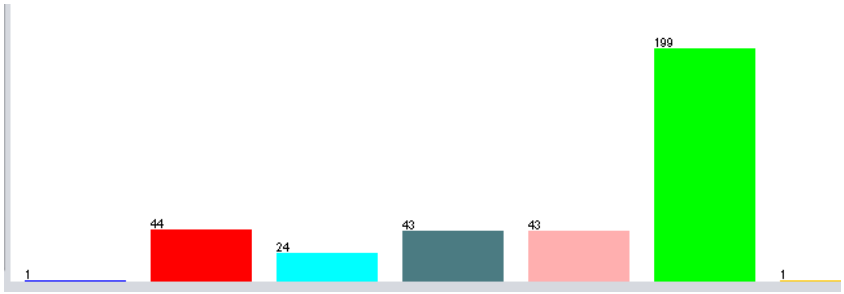
[2] Attribute : Blood Pressure : 12(3%) , Distinct : 11, Type : Nominal, , Unique: 4(1%)



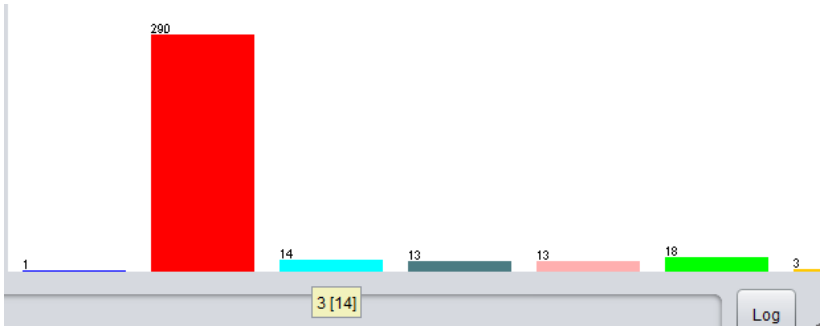
[3] Attribute: Specific Gravity: 47(12%) , Distinct : 6, Type : Nominal, , Unique: 1(0%)



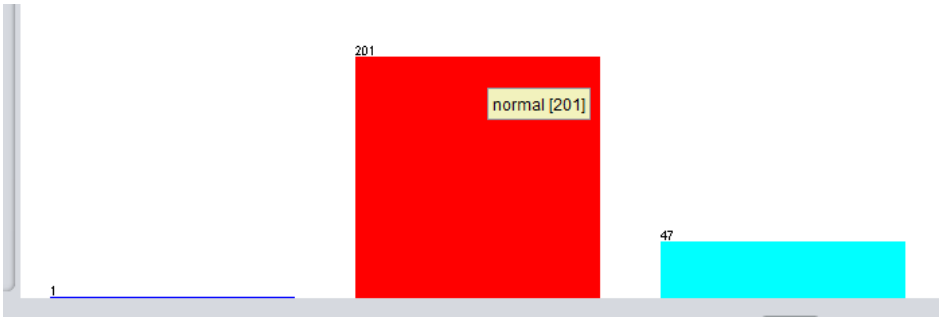
[4] Attribute: Albumin: 46(11%) , Distinct : 7, Type : Nominal, , Unique: 2(0%)



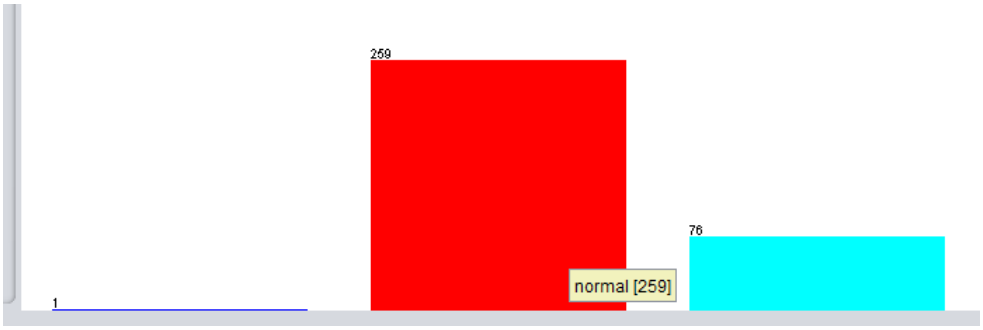
[5] Attribute: Sugar: 49(12%) , Distinct : 7, Type : Nominal, , Unique:1(0%)



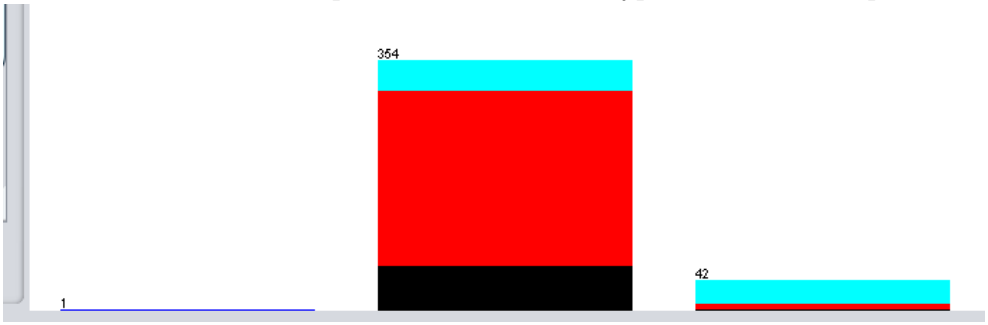
[6] Attribute: Red blood cell: 152(38%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



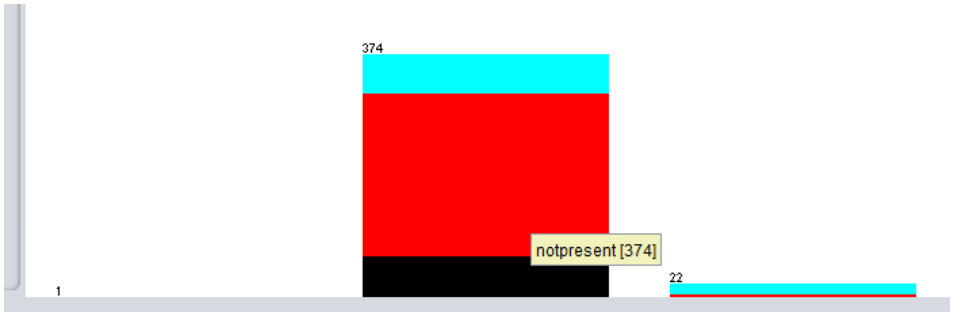
[7] Attribute: Pus Cell: 65(16%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



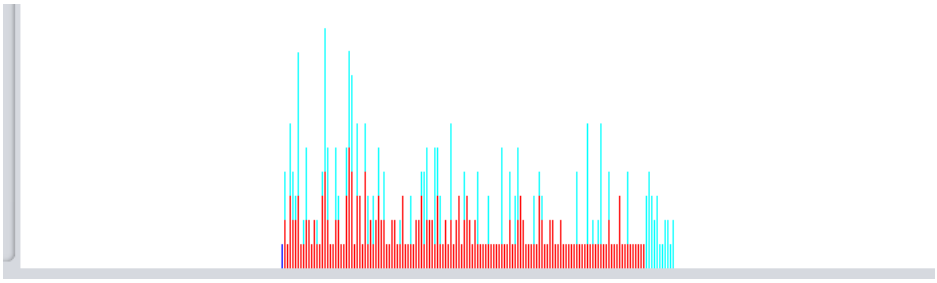
[8] Attribute: Pus Cell clumps: 4(1%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



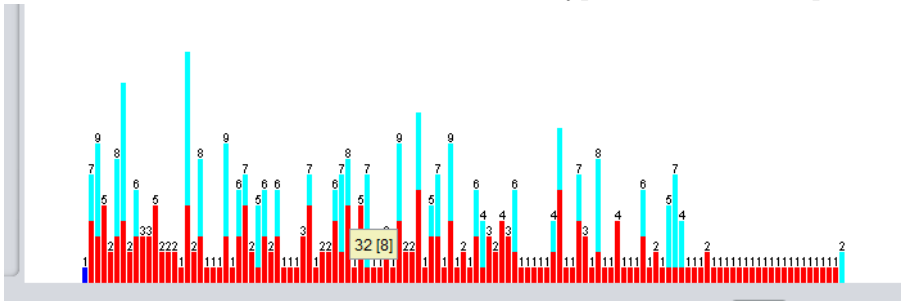
[9] Attribute: Bacteria: 4(1%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



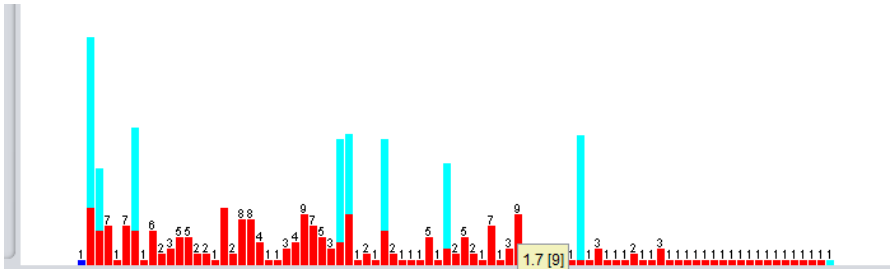
[10] Attribute: Blood Glucose Random: 44(11%) , Distinct : 147, Type : Nominal, , Unique:66(16%)



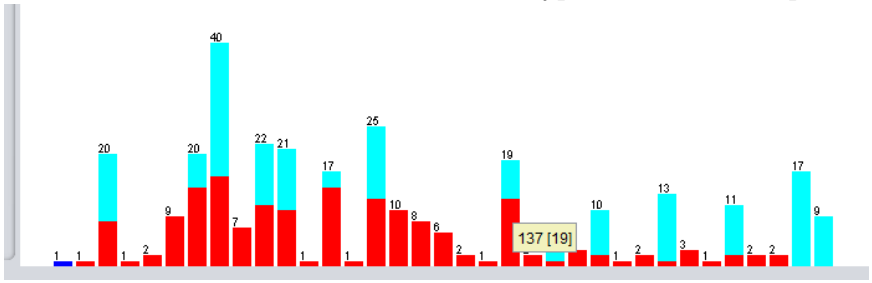
[11] Attribute: Blood Urea: 19(5%) , Distinct : 119, Type : Nominal, , Unique:56(14%)



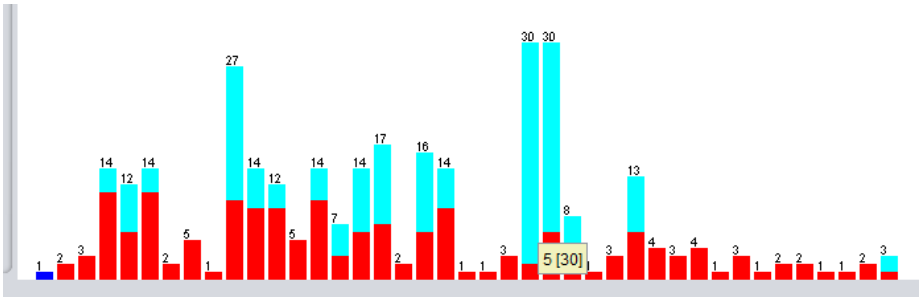
[12] Attribute: Serum Creatinine: 17(4%) , Distinct : 85, Type : Nominal, , Unique:42(10%)



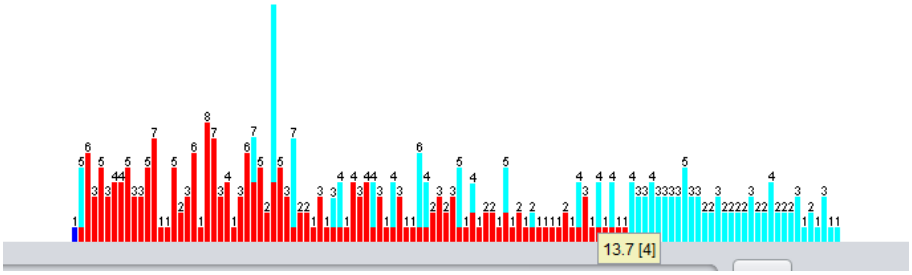
[13] Attribute: Sodium: 87(22%) , Distinct : 35, Type : Nominal, , Unique:8(2%)



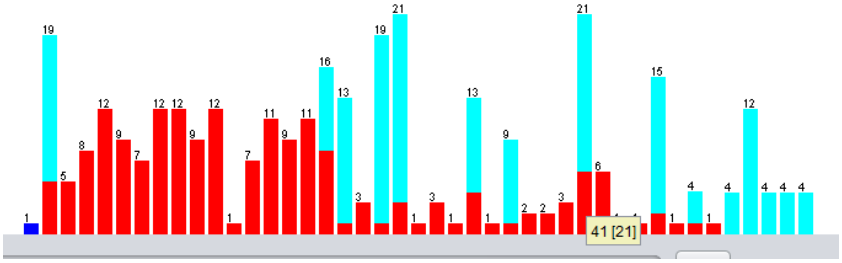
[14] Attribute: Potassium: 88(22%) , Distinct : 41, Type : Nominal, , Unique:9(2%)



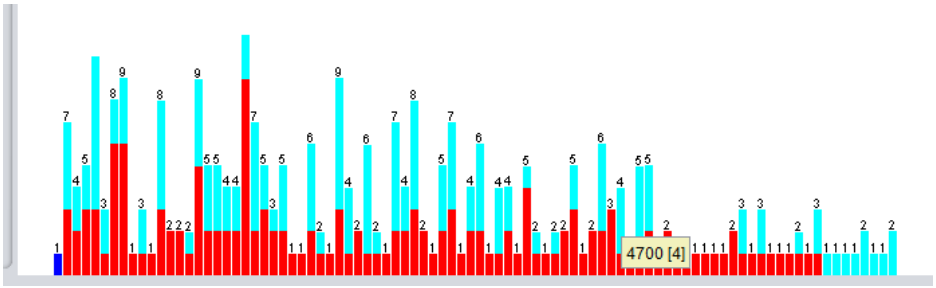
[15] Attribute: Hemoglobin: 52(13%) , Distinct : 116, Type : Nominal, , Unique:29(7%)



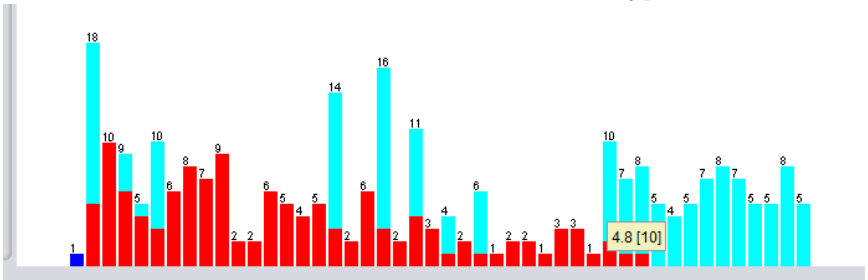
[15] Attribute: Packed Cell Volume: 71(18%) , Distinct : 43, Type : Nominal, , Unique:9(2%)



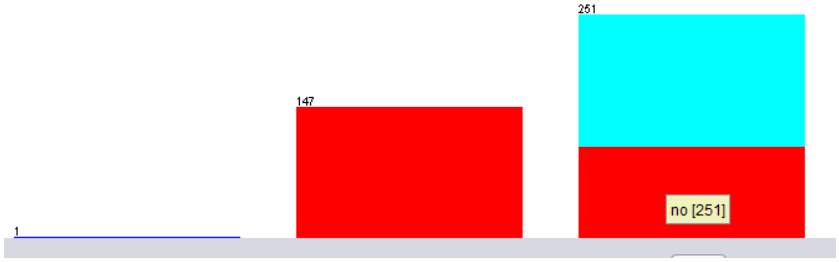
[17] Attribute: White Blood Cell Count: 106(26%) , Distinct : 90, Type : Nominal, , Unique:32(8%)



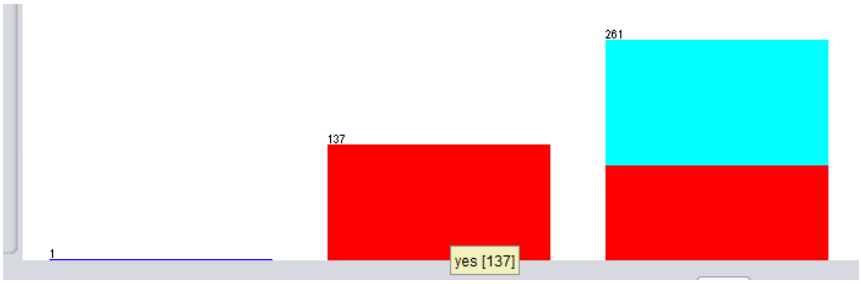
[18] Attribute: Red Blood Cell Count: 131(33%) , Distinct : 46, Type : Nominal, , Unique:4(1%)



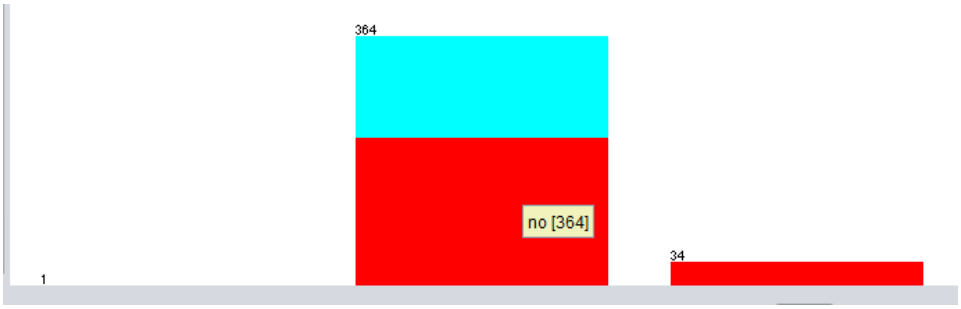
[19] Attribute: Hypertension: 2(0%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



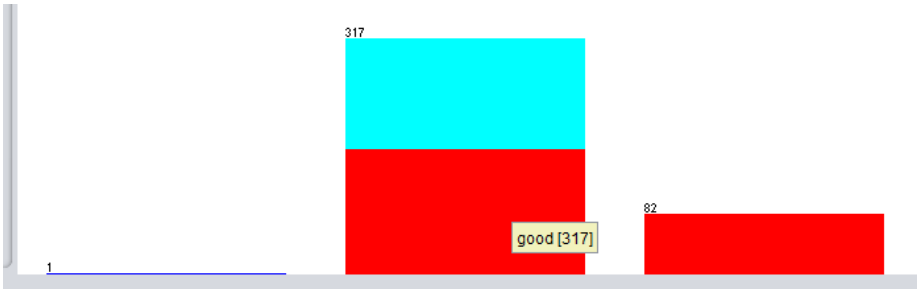
[20] Attribute: Diabetes Mellitus: 2(0%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



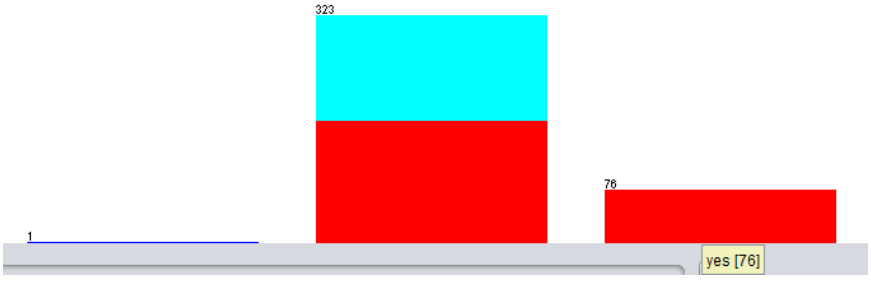
[21] Attribute: Coronary Artery Disease: 2(0%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



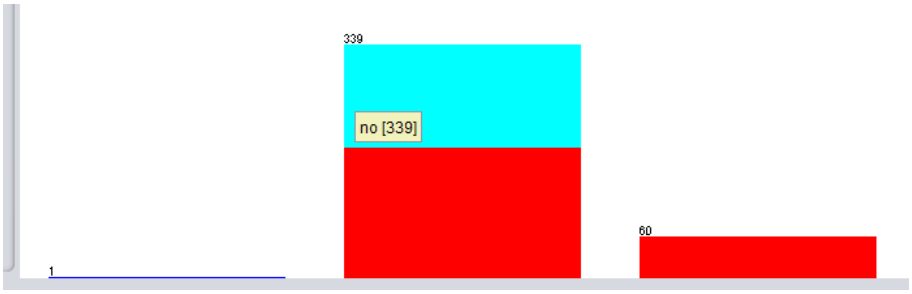
[22] Attribute: Appetite: 1(0%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



[23] Attribute: Pedal Edema: 1(0%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



[24] Attribute: Anemia: 1(0%) , Distinct : 3, Type : Nominal, , Unique:1(0%)



[25] Attribute: Class: 0(0%) , Distinct : 3, Type : Nominal, , Unique:1(0%)

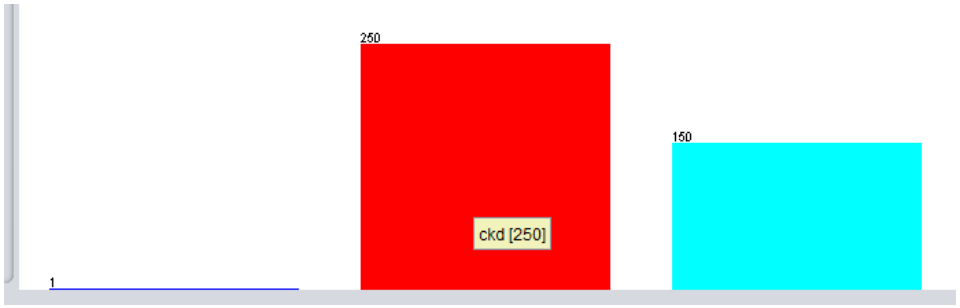


Figure 6: Shows Preprocess datasets of every attributes in graphical format

Figure Shows the performance criteria values. In WEKA, preprocess is one of the option; this is useful for getting detail information or statistics based on the data set. In this research, kidney dataset analysis, which contains 400 rows i.e., instances and 25 attributes, means columns. Here we are getting detail to preprocess steps of all 25 attributes for 400 instances i.e., for records. The researcher has select every attributes to displays type of attributes, the type means Nominal, how many missing values present in the data set for each attribute viz instances, how many distinct values are present in the dataset, distinct means different values, if we select attribute is shown Nom- in front of attribute- nom means nominal type. If data is numeric and select as an attribute, it will show the statistics report in the form of min, max, the mean and standard deviation of Statistics and Value. If we select the attribute as Class, it will show label count and weight in the form of true or false or yes or no.

3.1 Analysis:

Summary of Associate for discovering association rule:

In WEKA preprocess, in this research, we are explaining how to explore and analyze the Chronic Kidney related data in the Apriori associator. In this algorithm, the minimum matrix or confidence value in Association rule mining supporter confidence number of cycle performing the role of preparing Rules. This research is carried out by formatting and found ten best rules. The rules create x belongs to y attributes; the constant output of Apriori is to set the best rules by using its value and over caste, and its output shows the rules in the form of the model, we can say model rule. Based on rules, describe every rule.

During the execution of the Apriori algorithm, minimum support 0.75 (301 instances), if minimum metric <confidence> level is less than 0.9, then Rule not suitable to apply further; therefore, this does not rule. The association is always on applied data. Moreover, Apriori always applied to Nominal data or Binary Data.

```
=== Run information ===

Scheme:       weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Relation:     Chronic_Kidney_Disease_P
Instances:    401
Attributes:   26
              Serial No
              Age
              Blood Pressure
              Specific gravity
              Albumin
              Sugar
              Red Blood cells
              Pus Cell
              Pus Cell Clumps
              Bacteria
              Blood Glucose Random
              Blood Urea
              Serum Creatinine
              Sodium
              Potassium
              Hemoglobin
              Packed Cell Volume
              White Blood Cell Count
              Red Blood Cell Count
              Hypertension
              Diabetes Mellitus
              Coronary Artery Disease
```

Figure 7: Shows the Associate Apriori discover Association rule output

```
Apriori
=======

Minimum support: 0.75 (301 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 5

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Size of set of large itemsets L(2): 9

Size of set of large itemsets L(3): 1

Best rules found:

 1. Pus Cell Clumps=notpresent Coronary Artery Disease=no 328 ==> Bacteria=notpresent 319    <conf:(0.97)> lift:(1.04) lev:(0.03) [13] conv:(2.21)
 2. Pus Cell Clumps=notpresent 354 ==> Bacteria=notpresent 342    <conf:(0.97)> lift:(1.04) lev:(0.03) [11] conv:(1.83)
 3. Appetite=good 317 ==> Bacteria=notpresent 301    <conf:(0.95)> lift:(1.02) lev:(0.01) [5] conv:(1.26)
 4. Pedal Edema=no 323 ==> Bacteria=notpresent 306    <conf:(0.95)> lift:(1.02) lev:(0.01) [4] conv:(1.21)
 5. Coronary Artery Disease=no 364 ==> Bacteria=notpresent 344    <conf:(0.95)> lift:(1.01) lev:(0.01) [4] conv:(1.17)
 6. Anemia=no 339 ==> Bacteria=notpresent 318    <conf:(0.94)> lift:(1.01) lev:(0) [1] conv:(1.04)
 7. Pus Cell Clumps=notpresent Bacteria=notpresent 342 ==> Coronary Artery Disease=no 319    <conf:(0.93)> lift:(1.03) lev:(0.02) [8] conv:(1.31)
 8. Pedal Edema=no 323 ==> Coronary Artery Disease=no 301    <conf:(0.93)> lift:(1.03) lev:(0.02) [7] conv:(1.3)
 9. Bacteria=notpresent Coronary Artery Disease=no 344 ==> Pus Cell Clumps=notpresent 319    <conf:(0.93)> lift:(1.05) lev:(0.04) [15] conv:(1.55)
10. Pus Cell Clumps=notpresent 354 ==> Coronary Artery Disease=no 328    <conf:(0.93)> lift:(1.02) lev:(0.02) [6] conv:(1.21)
```

Figure 8: Shows the Associate Apriori discover Association rule output

=== Run information ===

Scheme: weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Relation: Chronic_Kidney_Disease_P
Instances: 401
Attributes: 26
Serial No
Age
Blood Pressure
Specific gravity
Albumin
Sugar

Red Blood cells
Pus Cell
Pus Cell Clumps
Bacteria
Blood Glucose Random
Blood Urea
Serum Creatinine
Sodium
Potassium
Hemoglobin
Packed Cell Volume
White Blood Cell Count
Red Blood Cell Count
Hypertension
Diabetes Mellitus
Coronary Artery Disease
Appetite
Pedal Edema
Anemia
Class

=== Associator model (full training set) ===

Apriori
=======

Minimum support: 0.75 (301 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 5

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Size of set of large itemsets L(2): 9

Size of set of large itemsets L(3): 1

```
Best rules found:

 1. Pus Cell Clumps=notpresent Coronary Artery Disease=no 328 ==> Bacteria=notpresent 319    <conf:(0.97)> lift:(1.04) lev:(0.03) [13] conv:(2.21)
 2. Pus Cell Clumps=notpresent 354 ==> Bacteria=notpresent 342    <conf:(0.97)> lift:(1.04) lev:(0.03) [11] conv:(1.83)
 3. Appetite=good 317 ==> Bacteria=notpresent 301    <conf:(0.95)> lift:(1.02) lev:(0.01) [5] conv:(1.26)
 4. Pedal Edema=no 323 ==> Bacteria=notpresent 306    <conf:(0.95)> lift:(1.02) lev:(0.01) [4] conv:(1.21)
 5. Coronary Artery Disease=no 364 ==> Bacteria=notpresent 344    <conf:(0.95)> lift:(1.01) lev:(0.01) [4] conv:(1.17)
 6. Anemia=no 339 ==> Bacteria=notpresent 318    <conf:(0.94)> lift:(1.01) lev:(0) [1] conv:(1.04)
 7. Pus Cell Clumps=notpresent Bacteria=notpresent 342 ==> Coronary Artery Disease=no 319    <conf:(0.93)> lift:(1.03) lev:(0.02) [8] conv:(1.31)
 8. Pedal Edema=no 323 ==> Coronary Artery Disease=no 301    <conf:(0.93)> lift:(1.03) lev:(0.02) [7] conv:(1.3)
 9. Bacteria=notpresent Coronary Artery Disease=no 344 ==> Pus Cell Clumps=notpresent 319    <conf:(0.93)> lift:(1.05) lev:(0.04) [15] conv:(1.55)
10. Pus Cell Clumps=notpresent 354 ==> Coronary Artery Disease=no 328    <conf:(0.93)> lift:(1.02) lev:(0.02) [6] conv:(1.21)
```

Figure 9: Best rules found:
Best rules found::

1. Pus Cell Clumps=notpresent Coronary Artery Disease=no 328 ==> Bacteria=notpresent 319 <conf:(0.97)> lift:(1.04) lev:(0.03) [13] conv:(2.21)
2. Pus Cell Clumps=notpresent 354 ==> Bacteria=notpresent 342 <conf:(0.97)> lift:(1.04) lev:(0.03) [11] conv:(1.83)
3. Appetite=good 317 ==> Bacteria=notpresent 301 <conf:(0.95)> lift:(1.02) lev:(0.01) [5] conv:(1.26)
4. Pedal Edema=no 323 ==> Bacteria=notpresent 306 <conf:(0.95)> lift:(1.02) lev:(0.01) [4] conv:(1.21)
5. Coronary Artery Disease=no 364 ==> Bacteria=notpresent 344 <conf:(0.95)> lift:(1.01) lev:(0.01) [4] conv:(1.17)
6. Anemia=no 339 ==> Bacteria=notpresent 318 <conf:(0.94)> lift:(1.01) lev:(0) [1] conv:(1.04)
7. Pus Cell Clumps=notpresent Bacteria=notpresent 342 ==> Coronary Artery Disease=no 319 <conf:(0.93)> lift:(1.03) lev:(0.02) [8] conv:(1.31)
8. Pedal Edema=no 323 ==> Coronary Artery Disease=no 301 <conf:(0.93)> lift:(1.03) lev:(0.02) [7] conv:(1.3)
9. Bacteria=notpresent Coronary Artery Disease=no 344 ==> Pus Cell Clumps=notpresent 319 <conf:(0.93)> lift:(1.05) lev:(0.04) [15] conv:(1.55)
10. Pus Cell Clumps=notpresent 354 ==> Coronary Artery Disease=no 328 <conf:(0.93)> lift:(1.02) lev:(0.02) [6] conv:(1.21)

In the above best found rule, a total of ten rules are found by the Apriori algorithm. Based on research, the maximum of the cycle in terms of probability is maximum, and less metric confidence is the best rule.
The class implementing an Apriori-type algorithm Iteratively less the less support wherever it finds the sufficient bundle of rules with the given minimum confidence
Class -- Binary class, Missing class values, No class, Nominal class Attributes -- Binary attributes, Empty nominal attributes, Missing values, Nominal attributes, Unary attributes, Additional, Minimum number of instances: 1

CONCLUSION

The Chronic kidney disease data has predicted and diagnosed further damage of kidney based on prediction of Data mining Machine Learning algorithms: In our research work, some of the other factors were considered Red Blood Cell count, Hypertension, Diabetes Mellitus, Coronary Artery Disease, Appetite, Pedal Edema, Anemia, near about 25 factors.

This research is carried out by formatting and found ten best rules. The rules create x belongs to y attributes; the primary output of Apriori is to set the best rules by using its value and over caste, and its output shows the rules in the form of the model, we can say model rule. Based on rules, describe every rule — the class implementing an Apriori type algorithm. During the execution of the Apriori algorithm, minimum support 0.75 (301 instances), if the minimum metric <confidence> level is less than 0.9, then Rule not suitable to apply further. Therefore this does not rule. The association is always on applied data. Moreover, Apriori is to applied to Nominal data or Binary Data.

Conflicts of interest: The authors stated that no conflicts of interest.

REFERENCES

1. Antonio Roberto Formaggio, Matheus Alves Vieira, Camilo DalelesRanno, Object-Based Image Analysis (OBIA) Anand Data Mining (D.M) in

Landsat Time Series for Mapping Soybean in Intensive Agricultural Regions, 978-1-4673-1159-5/12/$31.00 2012 IEEE pg(2257-2260)2012
2. Manish Saggar, Ashish Kumar Agrawal, Abhimanyu LadOptimization of Association Rule Mining using Improved Genetic Algorithms, 0-7803-8566-7/04/$20.00 IEEE pg (3725-3729),2004
3. CarolinFenlon, Luke O Grady, Laurence ShallooRegression techniques for modeling conception in seasonally calving dairy cows, 2016 IEEE 16th International Conference on Data Mining Workshops, 2375-9259/16 2016.174 pg (1191-1196),2016
4. Yaoguang Hu, ZhengjieGuo, Jingqian Wen, Jialin Han Research on knowledge mining for agricultural machinery maintenance based on association rules, 978-4799-8389-6/15/$31.00 2015, IEEE. pg (885-890) 10th Conference on Industrial Electronics and Applications (ICIEA), 2015
5. Niketa Gandhi, Leisa J. Armstrong Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India Using Data Mining Techniques, 2016 IEEE 16th International Conference on Advances in Computer Applications (ICACA) IEEE, 978-I-5090-3770-4/16 pg (357-363), 2016
6. SabriSerkanGulluogluSegmenting Customers With Data Mining Techniques, ISBN: 978-4799-6376-8/15/$31.00 IEEE pg (154-159), 2015
7. Zahid Halim, Dr. Rauf Baig, ShariqBashirSonification: A Novel Approach towards Data Mining, IEEE-ICET—2006 2nd International Conference on Emerging Technologies 1-4244-0502-5/06/$20.0 13-14, November 2006 pg (548-553), 2006
8. Kriti Bhargava, Stepan Ivanov, William Donnelly, ChamilKulatungaUsing Edge Analytics to Improve Data Collection in Precision Dairy Farming, 2016 IEEE 41st Conference on Local Computer Networks Workshops DOI 10,1109/LCNW, 2016.9 pg(137-144), 2016
9. Niketa Gandhi, Leisa J Armstrong review of the application of data mining techniques for decision making in agriculture, 978-1-5090-5256-1/16/$31.00 IEEE, 2nd International Conference on Contemporary Computing and Informatics (ic3i) 2016 pg (1-6), 2016
10. SanyamBharara, A. Sai Sabitha, Abhay Bansal A review on Knowledge extraction for Business operations using Data Mining, 978-1-5090-9/17/$31.00 IEEE, 2017 7th International Conference on Cloud Computing, Data Science & Engineering –Confluence pg (512-518), 2017
11. R. Senthil Kumar, Dr. C. RameshA Study on Prediction of Rainfall Using Data Mining Technique, Department of Computer Science and Engineering Satyabama University Chennai pg(09), Research Scholar
12. Wei Zhang, Shuping Li, Xue Wang, ChunyanXiaApplication of Data Mining in Agricultural Topic Tracking, 978-1-4577-1587-7/11/$26.00 IEEE International Conference on Computer and Network Technology pg (38-41), 2011
13. AyodeleLasisi and Rozaidaghazali, FolaLasisi, TututHerawan, Mustafa Mat DerisKnowledge Extraction of Agricultural Data Using Artificial Immune System, 978-1-4673-7682-2/15/$31.00 IEEE 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery pg (1653-1658), 2015
14. Hisayoshi Kato, Hironori, Hiraishi and Fumio MizoguchiLog summarizing agent for web access data using Data Mining techniques, 0-7803-3/01/$10.0 IEEE 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) pg (2642-2647), 2001
15. R. Sujatha, Dr. P. IsakkiDeviA Study on Crop Yield Forecasting Using Classification Techniques, 978-1-4673-8437-7/16/$31.00 IEEE pg (4), 2016
16. Liquiong Tang, Phillip AbplanalpGPS Guided Farm Mapping and Waypoint Tracking Mobile Robotic System, 978-1-4799-4315 2014 IEEE 9th Conference on Industrial Electronics and Applications (ICIEA), pg (1676-1681), 2014
17. V. Vijay Hari Ram, H. Vishal Regulation of Water in Agriculture Field Using Internet of Things, 978-1-4799-7758-1/15/$31.00 IEEE, International Conference on Technological Innovations in ICT for Agriculture and Rural Development (TIAR 2015), pg (112-115), 2015
18. Srisruthi.S, N. Swarna, G.M. SusmithaRos, Edna Elizabeth Sustainable Agriculture using Eco-

friendly and Energy Efficient Sensor Technology, 978-1-5090-0774-5/16/3$31.00 IEEE International Conference on Recent Trends in Electronics Information Communication, Technology, 2013, India pg (1442-1446), 2016

19. Esther HochsztainA Mining approach to evaluate geoportals usability, 978-1-4673-8111-6/15 $31.00 IEEE International Workshop on Data Mining with Industrial Applications pg (1-7), 2015
20. Lorenzo Di Silvestro, Michael Burch, Margherita Caccamo, Daniel Weiskopf, Fabian Beck, Giovanni GalloVisual Analysis of Time-dependent Multivariate Data from Dairy Farming Industry, 5th International Conference on Information Visualization Theory and applications (IVAPP – Pages 99-106 2014 – International Conference on Information Visualization Theory and Applications – IVAPP- *Italy, Germanypg (99-106),* 2014, ISBN: 978-989-758-055-5
21. M.H. Ariff, I. Ismarani, N. ShamsuddinRFID Based Systematic Livestock Health Management System, 978-1-4700-6100 -1/14/$31.00 IEEE Conference on Systems, Process, and Control (ICSPC-2014) pg (6), 2014
22. Chen Jinyin, Lin Xiang, Zheng Haibing, BaoXintongA novel cluster center fast determination clustering algorithm, ELSEVIER, Applied Soft Computing 57 (2017) 539-555- Journal www.elsevier.com/locate/950c- 2017. Pg (539-555), 2017
23. S. DilliArasu, Dr. R. ThirumalaiselviA Novel imputation method for effective prediction of coronary kidney disease, 978-1-5090-6221-8/17/$ 31.00 C -2017 IEEE 2nd International Conference on Computing and Communications Technologies (ICCT/17) pg (127-136), 2017
24. Zou Chuan, Tang Ying, Bai Li, Zeng Li, Zeng Yuqun, Lu Fuhua, Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangzhou, China Application of clustering analysis to explore syndrome evolution law of peritoneal dialysis patients
25. Veenita Kunwar, Khushboo Chandel, A. Sai Sabitha, Abhay Bansal, ASET, CSE, Amity University, Uttar Pradesh, Noida, India, 978-1-4673-8203-8/16/$31.00C 2016 IEEE ' Chronic Kidney Disease analysis using data mining classification techniques.
    Due Thanh Anh Luong, Dept of the computer. Sci & Engi, uni, at Buffalo Buffalo, NY, USAA K-Means Approach to Clustering disease Progressions IEEE Keywords, Sep 2017,
26. Narander Kumar, Sabita KLhatri, Department of Computer, 3rd IEEE International Conference on Computational Intelligence and Communication Technology (IEEE-CICT 2017) " Implementing WEKA for medical data classification and early disease prediction 978-1-5090-6218-8/17/@31.00 @2017 IEEE
27. AKIhelper: Acute kidney injury diagnostic tool using KDIGO guideone approach, Issariya ubolthamm, Nakornthip Prompon, wirichal Pan-ngum, department of Computer Engineering Chulalongkon University, Bangkok, Thailand, 978-1-2090-0806-3/16/$31.00 Copyright 2016 IEEE 2016 June 26-29, 2016 Okayama Japan https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2845186/,the economics of Dialysis in India, Indian Journal of Nephrology.
28. http://www.indianjnephrol.org Prevalence of chronic kidney disease in India – Where are We heading?