

# DATA MINING APPROACH TO PREDICT DENGUE FEVER

Swaty Behl Department of Computer Science & Engineering Punjabi University Patiala India Gaurav Gupta Department of Computer Science & Engineering Punjabi University Patiala India

Abstract: Data mining approached are used everywhere in the industry. It permits users to analyze data from many different perspectives. It is also used to classify and build relationships among data. So it is widely used in today's world in every field. It has given very fruitful results in healthcare field also. Many diseases such as Hih BP, sugar, dengue, COVID, heart attacks etc are taking the lives of community. In this paper dengue data has been analyzed using data mining techniques. Association and classification techniques have been applied to predict dengue fever. The results generated by these techniques will predict the dengue fever in individuals and they can consult doctor and take necessary medicine.

Keywords: Data Mining; Classification; Association; Dengue; Prediction; Medical Science.

## 1. INTRODUCTION

The process of data mining analyzes data from different dimensions and summarizing it into useful information. Basically data mining is used to analyze data. It permits users to analyze data from many different perspectives. Scientifically, data mining is a process for finding out an interconnection or patterns in huge amount of relational databases. It is a process of arranging data to find out patterns and identifying relationships. Data mining techniques can be executed quickly on real software and hardware platforms to improve the value of current information resources. These patterns cannot be determined by using traditional methods because of the bulkiness of data.

Data mining is one the important and motivating range of research with the equitable of discovery intelligible knowledge from large scale data sets [8]. Healthcare is most popular field application in data mining to identify hidden and admired knowledge in health data. Data mining is used in health industry to discover fraud in health insurance, treatment at lower cost, cure the disease [9].

#### 1.1 Association

Association rule is one of important aspect of data mining. Association rule mining is used to predict correlations, patterns, associations or causal structures from data sets found in several types of databases such as relational databases, transactional databases, and other forms of data repositories. Given set of transactions, the goal of association rule mining is to find the rules which permit us to predict the occurrence of specific item depend on occurrences of the another items in the transaction. In many application areas it is required to discover the frequent patterns, frequent patterns are those data items that frequently occur in the data set. Classical problem of association data mining is super market analysis. Super market contains huge amount of data of their customers and it is required to maintain the data in the well-mannered form.

The association rule mining on "Market Basket Data" is Boolean association rule mining in which only Boolean values are considered. The aim of the association rule mining is to extract interesting patterns in the large datasets of items. Association rules are often used to analyze sales transactions. For example, it might be noted that customers who buy cereal at the grocery store often buy milk at the same time. In fact, association analysis might find that 85% of the checkout sessions that include cereal also include milk. This relationship could be formulated as the following rule: Cereal implies milk with 85% confidence



# 1.2 Classification

Classification consists of predicting a certain outcome based on a given input. In order to predict the outcome, the algorithm processes a training set containing a set of attributes and the respective outcome, usually called goal or prediction attribute. The algorithm tries to discover relationships between the attributes that would make it possible to predict the outcome. Next the algorithm is given a data set not seen before, called prediction set, which contains the same set of attributes, except for the prediction attribute – not yet known. The algorithm analyses the input and produces a prediction. The prediction accuracy defines how "good" the algorithm is. For example, in a medical database the training set would have relevant patient information recorded previously, where the prediction attribute is whether or not the patient had a heart problem.

		Training set	
Age	Heart rate	<b>Blood pressure</b>	Heart problem
65	78	150/70	Yes
37	83	112/76	No
71	67	108/65	No

		Prediction set	
Age	Heart rate	<b>Blood pressure</b>	Heart problem
43	98	147/89	?
65	58	106/63	?
84	77	150/65	?

Among several types of knowledge representation present in the literature, classification normally uses prediction rules to express knowledge. Prediction rules are expressed in the form of IF-THEN rules, where the antecedent (IF part) consists of a conjunction of conditions and the rule consequent (THEN part) predicts a certain predictions attribute value for an item that satisfies the antecedent. Using the example above, a rule predicting the first row in the training set may be represented as following:

# IF (Age=65 AND Heart rate>70) OR (Age>60 AND Blood pressure>140/70) THEN Heart problem=yes

Knowledge Discovering in Databases is a process of discover knowledge in data and recap the "high-level" application of specific data mining technique. It is used for relevance interest to researcher's awareness, databases, specific, artificial intelligence and data visualization. The unified aim of the KDD process is to abstract data from data warehouse in the framework of huge databases.

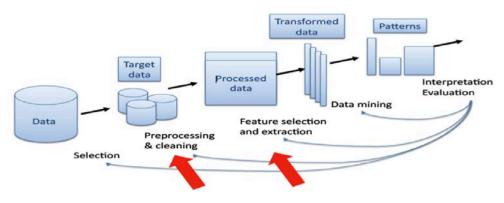


Figure-1: Knowledge Discovering in Databases (KDD)



## 1.3 Dengue Fever

Dengue is a dangerous disease. Dengue fever is one of the important reasons for the increasing deaths globally. It is a mosquito borne life threatening disease caused by breeding of Aedes mosquito. In India numbers of deaths are caused by dengue disease. Dengue is a dangerous disease with concern that is widespread in tropical and sub-tropical countries. It is a disease caused by breeding of Aedes Mosquito and also named as break bone fever. Dengue disease infected by Dengue Hemorrhagic Fever (DHF) whose symptoms including breeding, low level of blood platelets, low blood pressure, metallic taste in mouth, headache, muscle and joint pain and rashes etc. Dengue fever is painful and life threatening mosquito borne disease caused by female mosquitoes. Dengue virus is spread by Aedes mosquito. The mosquito becomes affected when it bites a person with dengue virus in their blood and it can be breed to other normal person. There is no particular medicine or antibiotic convenient to treat it. Dengue fever occurs in form of cycles and this cycle is present inside the body of an affected person for two weeks or less than two weeks. The various symptoms that are associated with dengue disease are:

- Fever Temperature
- Period of Fever
- Muscle and Joint Pain
- Rashes
- Low Heart Rate
- Blood Pressure
- Pain behind Eyes
- Headache

These various symptoms indicate serious illness and are sign that we must consult a doctor.

# 2. LITERATURE REVIEW

This section reviews the current research papers and articles to present the existing knowledge on this particular topic.

Mani Shankar, Mayank Pahadia, Divyang Srivastava, Ashwin T S, G. Ram Mohan Reddy (2015) A Novel Method for Disease Recognition and Cure Time Prediction Based on Symptoms: The study of this paper describes a novel model for identification of disease and prediction of their cure time by considering only the symptoms, they allow different coefficients for each symptoms of a disease and filter out a dataset with rating assigned to each symptom by the user. The disease get identify based on calculate numerical value using reinforcement learning. Algorithm used into account the similarly between the symptoms of one user and other user how have suffered from the same disease. For implementing this approach recommender system was used to read an input to find out pattern, which is based on the dataset given to train the system. This paper provides the solution that takes the symptoms with rating given by the patient to predict the possible disease and possible cure time of the disease.

Kashish Ara Shakil, Shadma Anis and Mansaf Alam (2017) Dengue Disease Prediction Using Weka Data Mining Tool: Study of this paper describes dengue is Dangerous disease. It is caused by Aedes mosquito. In this paper, they use available dataset for dengue including the symptoms like fever temperature, WBC, platelets, severe headache, vomiting metallic taste, joint pain, appetite, diarrhea, hematocrit, haemoglobin. This paper discusses various algorithm approaches available for data mining that have been used for the predication of dengue disease. Data mining is a wellknown technique used for classification of diseases such as dengue, diabetics and cancer by many health organizations. The proposed approach used WEKA for evaluation. It was concluded that naïve bayes and J48 are the best performance algorithms for classification because they achieve maximum accuracy (100%) with 99 correctly classified instances and least mean absolute error and it takes least time foe building this model through explorer and knowledge flow results.

**P.Manivannan, Dr. P. Isakki Devi (2017) Dengue Fever Prediction using K-Medoid Clustering Algorithm:** The paper describes about dengue is a life threatening disease infection by female mosquitoes. It is searched in hot regions. This paper include the symptoms of dengue like severs joints pains, headache, rashes, thrombocytopenia, leucopenia and muscle ache. It is also called break bone fever because it is due to muscle ache and joint pains. This paper, they include four levels namely preprocessing, attribute selection, k-medoid clustering and prediction of dengue



fever by using of dengue fever by using R 3.3.2 tool for preprocessing the household of dengue dataset. The aim of this paper, they predict those people they affected by dengue categorization of age by using k-medoid clustering algorithm, which has been implemented. The k-medoid algorithm is just like k-means clustering algorithm. The proposed algorithm is used for systematic method of selecting the initial medoids. The performance of algorithm of algorithm is varying on initial medoids. K-medoid clustering is improving the efficiency of output.

V.Krishnaiah, Dr.G.narsimha, Dr.N. Subhash Chandra (2018) Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques: The paper describes the cancer is most considerable cause of death for both man and women. Initial detection of cancer can be beneficial in curing the disease completely. Preliminary of techniques to discover the appearance of cancer module in initial stage is expanding. Commonly lung cancer disease is misdiagnosed. Previous diagnosis of lung cancer saves immense lives, failing may lead to other severe headache causing abruptly fatal end. Early detection depends on the cure rate and diagnosis of the disease. In this paper they include following symptoms dyspnea, hemoptysis, chronic coughing or change in regular coughing pattern, wheezing, chest pain or pain in abdomen, cachexia, dysphonia, clubbing of the finger nails, dysphasia, pain in shoulder, chest, arm, bronchitis or pneumonia, decline in health and unexplained weight loss. Data mining and knowledge discovery of database is diverse applications in business and scientific domain. The main purpose of this paper to proposed a model for early detection and correct diagnosis of a disease and they will consult doctor for saving life of the patient.

Nandini. V and Sriranjitha. R and Yazhini. T. P (2019) Dengue Detection and Prediction System Using Data Mining with Frequency Analysis: In this paper author(s) defines clinical documents and conditions of patients. Dengue is widespread disease in today's world. This main focus of this paper is performing named entity recognition to extract disorder mentions, time expressions and other related characteristic from clinical data. They proposed a model, which predict about presence or absence of dengue disease and analysis of relationship of occurrence with consider symptoms. The proposed model produces accuracy and estimate tool for medical experts. The proposed system is effective for medical repositories. POS tagging is used for tag the discharge detail of the patient.

Shameem Fathima, Nisar Hundewale(2019) Comparison of Classification Techniques-SVM and Naives Bayes to predict the Arboviral Disease-Dengue: In this paper, author(s) define the analysis of several data mining techniques to predict the arboviral disease dengue. It analyzes the real dataset from super specialty hospitals diagnostic laboratories where blood samples were collected. The data set contain 5000 records with 29 parameters. This paper includes data mining techniques SVM and Naïve bayes classifier and evaluates the comparison results between the method by using dengue disease diagnosis. They collect the data record forms in Chennai and tirunelveli from India.

**N.Subitha, Dr.A.Padmapriya (2020) Diagnosis for Dengue Fever Using Spatial Data Mining:** In this paper they describe spatial data and used by accurate method for rule mining. Basically association rule mining is used for finding interestingness patterns from spatial database by using K-means algorithm in which patterns are not exactly stored. Spatial association rule mining is used to calculate several spatial correlations among a huge number of spatial objects. Finding interesting optimization method is called progressive refinement. It adopted by spatial association analysis. In this paper, we have implementation the automatic dengue test kit system to reorganization the dengue fever by using microscope blood image report. Microscope image report as input and signals are filtered and the feature characteristics are extracted by neural networks. Classification and back propagation network gives 98% correct results in short time.

Nor Shafikah Roslan, Zulkiflee Abd Latif, and Nazri Che Dom (2020) Dengue Cases Distribution Based in Land Surface Temperature and Elevation: It describes Malaysia is one of the countries that dengue fever crisis. The dengue virus in Malaysia remains high. Cure of dengue is unavailable. In this paper, they use association between land surface temperature, elevation and study of dengue fever. Spatial application is geographical information system and remote sensing image. These applications are used under control of dengue outbreaks as analysis of potentially high risk areas in urban areas. Dengue dataset can disturbed in patterns and factors are visualized. In this paper, it includes the range of temperature from 14-18 degree C at lower level and 35-40 degree C at the upper level. Dengue fever is spread between 10 degree C and not more than 30 degree temperature. It includes four classes of temperature 12 degree C to above 30 degree C. dengue incidence was distributed results at low elevation. It emphasized lower elevation less than 500 meter. It was found dengue is occurring in low land area.

#### 3. **RESULTS AND ANALYSIS**

Various responses from 100 Patients have been collected according the questionnaire framed. The data has been collected from patients suffering and not suffering from dengue disease. The collected response was analyzed on

© 2013, IJMDEBM, All Rights Reserved

http://www.ijmdebm.org/



various parameters as presented below. In order to carry out an implementation R Studio is used as data mining tool. The main aim of the research is to predict the disease based on symptom.

The following are the screenshots to illustrate the results produced from association rule mining in R studio.

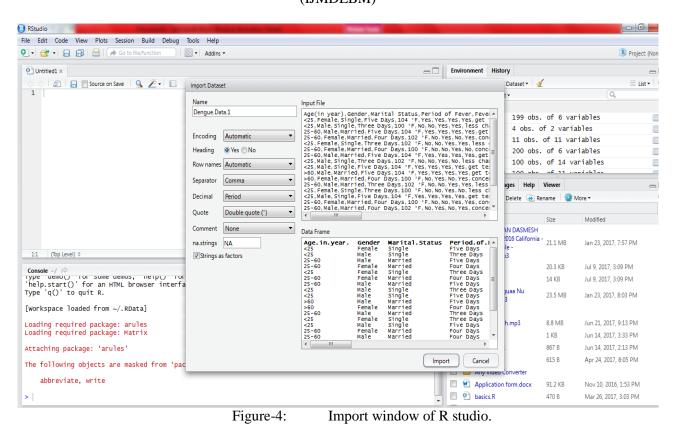
Riste				1018	9. X
File Edit Code View Plots Session Build Debug Tools Help 9. • 🐨 • 🕞 😝 😂 📣 Constanting View 1 🔯 • Addes •				8 hije	nz (Nore)
9) United x	-0	Environment History			-0
D D D Downenier 9 Z-1 D	Har H Distor + 2	🞯 日 🔄 mant Dataset • 🕑		= p	M+103
1		🚯 Global Environment •		Q.	
		O Dengue. Data.1 100 obs. of 11 v			- C3 -
		O dengue6 10 obs. of 7 var			
		O denguedata10 100 obs. of 11 v O results 10 obs. of 7 var			
		Oresults 10 des. of 7 var Orean 200 obs. of 6 va			
		Oshapshoti 9 obs. of 11 var			- 5
		Astorks 100 obs of 6 us			
		Files Plots Packages Help Viewer			-
		Q New Folder O Celete - Rename S More	4		6
					1
		+ Name	Str	Modified	12
		California - Dhadrianwale - YouTube.mp3	21.1 ME	Jan 20, 2017, 7:57 PM	
11 (boleweit :	K forigt 1	E 🐑 AData	18.4 KB	Jul 9, 2017, 247 PM	- 11
	-	E D Ahistory	13.3 KB	Jul 9, 2017, 247-PM	- 9
Console -/ IP Type versus, the point versus, metp(2) by onintrine netp, or		Aij De Ganguaa Nu Lahritaa.mp3	23.5 MB	Jan 23, 2017, 8-03 PM	
'help.start()' for an HTML browser interface to help. Type 'q()' to quit R.		E G Adobe			
		Akaal Purekh.mp3	8.8 MB	Jun 21, 2017, 9:13 PM	
[workspace loaded from -/.RData]		AlgoLA	1.01	Jun 14, 2017, 3:33 PM	
Loading required package: arules Loading required package: Matrix		Algorithm R	867.8	Jun 14, 2017, 2:13 PM	
Sector		D 9 Aralysis R	615.0	Apr 24, 2017, 8:05 PM	
Attaching package: 'arules'		🖾 🤐 Any Video Converter			
The following objects are masked from 'package:base':		Application form docx	95.2.65	Nov 30, 2006, 3.53 PM	
abbreviate, write		E 2 besch	470 B	Mar 26, 2017, 2:05 PM	
		Bhai Guriobal Singh, Bhai Amandeep Singh. - Sukhmani Sahib (Vyakhya Sahit) mp3	188.6 ME	May 25, 2017, 6:57 PM	61
>]		EL (2) hookt #	164.8	Teb 25, 2017 2355 RM	÷.,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,

Figure-2:

R studio tool window.

] 📲 😴 🔹 🔚 🔚 📄 🌈 Go to file/function 🔤 🗏 🔛 Addins 🔹						툏 Project
) Untitled1 ×	-0	Envir	onm	nent History		
🗇 🖒 🔝 🔚 🔄 Source on Save 🛛 🔍 🎽 🕶	🕀 Run 🛛 🔁 🕞 Source 🔹 🗏	3		📑 Import Dataset 🔹 🔏		≡ Lis
1		ore ory	engu engu engu esu l ran	u u vedata10 100 obs. of 7 var 10 obs. of 7 var 10 obs. of 7 var 200 obs. of 6 var 200 obs. of 6 var	iables /ariables iables ariables	Q
		0 sh		shot1 9 obs. of 11 var ks 200 obs. of 6 va		
					ariahlac	
		Files		Plots Packages Help Viewer		
				Folder 🧕 Delete 🕞 Rename 🛛 🎡 Moi	e *	
		- 1		▲ Name	Size	Modified
				(POEM)DHAN DASMESH PITA - 17.9.2016 California - Dhadrianwale - YouTube.mp3	21.1 MB	Jan 23, 2017, 7:57 PM
1:1 (Top Level) ¢	R Script ‡		3	RData	18.4 KB	Jul 9, 2017, 2:47 PM
=			0	.Rhistory	13.3 KB	Jul 9, 2017, 2:47 PM
ponsole ~/ 🔗			D	Aaj De Ganguaa Nu Lahntaa.mp3	23.5 MB	Jan 23, 2017, 8:03 PM
elp.start()' for an HTML browser interface to help. pe 'q()' to quit R.	<u>^</u>			Adobe		
			D	Akaal Purakh.mp3	8.8 MB	Jun 21, 2017, 9:13 PM
iorkspace loaded from ~/.RData]			9	Algo1.R	1 KB	Jun 14, 2017, 3:33 PM
bading required package: arules			9	Algorithm.R	867 B	Jun 14, 2017, 2:13 PM
bading required package: Matrix			0	Analysis.R	615 B	Apr 24, 2017, 8:05 PM
taching package: 'arules'			1	Any Video Converter		
e following objects are masked from 'package:base':				Application form.docx	91.2 KB	Nov 10, 2016, 1:53 PM
abbreviate, write	=		_	basics.R	470 B	Mar 26, 2017, 3:03 PM
· · · · · · · · · · · · · · · · · · ·			D	Bhai Guriqbal Singh, Bhai Amandeep Singh - Sukhmani Sahib (Vvakhva Sahit).mp3	Ji 188.6 MB	May 25, 2017, 6:57 PM
	*			- sukrimani sanib (vyaknya sanit).mps	166 B	Eab 25, 2017, 2-55 DM





Ŧ			📌 Go to file/funct	tion	Addins 🔻					🗷 Project: (1
Դս	ntitled1 ×	Dengue.Da	tal×						-0	Environment History
5								Q		🕂 📊 📑 Import Dataset 🗸 📃 List 🗸
	Age.in.year. 🗘	Gender ‡	Marital.Status	Period.of.Fever	Fever.Temperature	Rashes.Red.Spots	Pain.Behind.Eyes	Muscle.and.Joint.Pain	Low.Heart.Rate	🕆 🌗 Global Environment 🕶 🔍
1	<25	Female	Single	Five Days	104 T	Yes	Yes	Yes	Yes ^	Dengue.Da. 100 obs. of 11 variables
2	<25	Male	Single	Three Days	100 T	No	No	Yes	Yes	<pre> @ dengue6 10 obs. of 7 variables </pre>
3	25-60	Male	Married	Five Days	104 'T	Yes	Yes	Yes	Yes =	<pre>Odenguedat 100 obs. of 11 variables</pre>
4	25-60	Female	Married	Four Days	102 T	No	Yes	No	Yes	<pre>oresults 10 obs. of 7 variables</pre>
5	<25	Female	Single	Three Days	102 T	No	No	Yes	Yes	Oryan 200 obs. of 6 variables
6	25-60	Female	Married	Four Days	100 T	No	No	Yes	No	<pre>③ shapshot1 9 obs. of 11 variables ④ stocks 200 obs of 6 variables</pre>
7	25-60	Male	Married	Five Days	104 T	Yes	Yes	Yes	Yes	=
8	<25	Male	Single	Three Days	102 T	No	No	Yes	No	
9	<25	Male	Single	Five Days	104 T	Yes	Yes	Yes	Yes	Year         Year <t< td=""></t<>
10	>60	Male	Married	Five Days	104 °F	Yes	Yes	Yes	Yes	▲ Name Size Modified
11	>60	Female	Married	Four Days	100 T	No	Yes	No	Yes	(POEM)DHAN
12	25-60	Male	Married	Three Days	102 T	No	No	Yes	Yes	DASMESH PITA - 17.9.2016 California 21.1 MB Jan 23, 2017, 7:5
13	<25	Female	Single	Three Days	100 T	No	No	Yes	No	- Dhadrianwale - PM
14	<25	Male	Single	Five Days	104 T	Yes	Yes	Yes	Yes	YouTube.mp3
15	25-60	Female	Married	Four Days	100 T	No	No	Yes	Yes	RData 18.4 KB Jul 9, 2017, 2:47
16	25-60	Male	Married	Four Days	102 T	No	Yes	No	Yes	Rhistory 13.3 KB Jul 9, 2017, 2:47
17	25-60	Female	Married	Five Days	104 T	Yes	Yes	Yes	Yes	Aaj De Ganguaa Nu Lahntaa.mo3 23.5 MB PM
18	25-60	Female	Single	Three Days	100 T	No	No	Yes	No	Lahntaa.mp3 PM
19	25-60	Male	Married	Four Days	102 T	No	No	Yes	Yes	In 21 2017 0.1
20	25-60	Female	Married	Five Days	104 "F	Yes	Yes	Yes	Yes	Akaal Purakh.mp3 8.8 MB PM
21	>60	Male	Married	Five Days	104 °F	Yes	Yes	Yes	Yes	Algo1.R 1 KB Jun 14, 2017, 3:3
22	. en	Fomalo	Married	Four Dave	102 T	No	No	Var	Ver t	<ul> <li>▼ PM</li> <li>PM</li> <li>■ Algorithm.R 867 B</li> <li>PM</li> </ul>

Figure-5:

Database with number of rows, columns and attributes.



#### (IJMDEBM)

• Use apriori algorithm with support = 0.2, confidence = 0.5, minimum length = 4, it shows values of parameters and results in 108 rules in following Figure-5 and Figure-7..

```
Apriori
Parameter specification:
 confidence minval smax arem aval originalSupport maxtime support minlen maxlen target
                                                                                                                                                   ext
                                  1 none FALSE
             0.5
                       0.1
                                                                           TRUE
                                                                                            5
                                                                                                      0.2
                                                                                                                   4
                                                                                                                              10 rules FALSE
Algorithmic control:
  filter tree heap memopt load sort verbose
       0.1 TRUE TRUE FALSE TRUE
                                                      2
                                                             TRUE
Absolute minimum support count: 20
set item appearances ...[2 item(s)] done [0.00s].
set transactions ...[19 item(s), 100 transaction(s)] done [0.00s].
sorting and recoding items ... [18 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 8 done [0.00s].
writing ... [108 rule(s)] done [0.00s].
creating 54 object
creating S4 object ... done [0.00s].
```

Figure-6

	lhs		rhs	support	confidence	lift
[1]	{Rashes.Red.Spots=Yes,					
	Pain.Behind.Eyes=Yes, Low.Heart.Rate=Yes}	=>	{DengueStatus=positive}	0.52	1.0000000	1.923077
[2]	{Rashes.Red.Spots=Yes,					
	Muscle.and.Joint.Pain=Yes, Low.Heart.Rate=Yes}	=>	{DengueStatus=positive}	0.52	0.9629630	1 851852
[3]	{Rashes.Red.Spots=Yes,		(benguebeacas-posterie)	0.52	0.9029090	1.051052
	Pain.Behind.Eyes=Yes, Muscle.and.Joint.Pain=Yes}		{DengueStatus=positive}	0.52	0.9811321	1 996703
[4]	{Pain.Behind.Eyes=Yes,	=>	{Deriguescacus=positive}	0.52	0.9811321	1.886792
	Muscle.and.Joint.Pain=Yes,					
[5]	Low.Heart.Rate=Yes} {Rashes.Red.Spots=Yes,	=>	{DengueStatus=positive}	0.52	0.9629630	1.851852
191	Pain.Behind.Eyes=Yes,					
	Muscle.and.Joint.Pain=Yes, Low.Heart.Rate=Yes}		{DengueStatus=positive}	0.52	1.0000000	1 022077
[6]	{Period.of.Fever=Five Days,	=>	{Deriguescacus=positive}	0.52	1.0000000	1.9230//
	Fever.Temperature=104 °F,					
[7]	FeverStatus=get tested} {Fever.Temperature=104 °F,	=>	{DengueStatus=positive}	0.51	1.0000000	1.923077
L/ J	Rashes.Red.Spots=Yes,					
601	FeverStatus=get tested}	=>	{DengueStatus=positive}	0.51	1.0000000	1.923077
[8]	{Fever.Temperature=104 °F, Low.Heart.Rate=Yes,					
	FeverStatus=get tested}	=>	{DengueStatus=positive}	0.51	1.0000000	1.923077
[9]	{Fever.Temperature=104 °F, Pain.Behind.Eyes=Yes,					
	FeverStatus=get tested}	=>	{DengueStatus=positive}	0.51	1.0000000	1.923077
[10]	{Fever.Temperature=104 °F, Muscle.and.Joint.Pain=Yes,					
	FeverStatus=get tested}	=>	{DengueStatus=positive}	0.51	1.0000000	1.923077
			Figure-7			
			I Iguic-7			

- ✤ In Rule [1], if Rashes.Red.Spots=Yes, Pain.Behind.Eyes=Yes, Low.Heart.Rate=Yes then it predicts DengueStatus=positive with support=0.52 and confidence=1.00.
- If we increase the value of parameters with support = 0.3, confidence = 0.9, minimum length = 4, it shows results in 99 rules, that means order of rules get decreased. It is depicted in Figure-8.

And the second s

Volume-10, Issue-3, July-September, 2022 International Journal for Multi Disciplinary Engineering and Business Management (IJMDEBM)

```
Apriori
Parameter specification:
confidence minval smax arem aval originalSupport maxtime support minlen maxlen target
                                                                         ext
           0.1 1 none FALSE
                                     TRUE 5
                                                       4
      0.9
                                                   0.3
                                                               10 rules FALSE
Algorithmic control:
 filter tree heap memopt load sort verbose
   0.1 TRUE TRUE FALSE TRUE
                          2
                              TRUE
Absolute minimum support count: 30
writing ... [99 rule(s)] done [0.00s].
creating 54 object ... done [0.00s].
                                        Figure-8
```

• If we increase the value of parameters with support = 0.4, confidence = 1.0, minimum length = 4, it shows 96 rules as in Figure -9, that means order of rules get decreased.

Apriori Parameter specification: confidence minval smax arem aval originalSupport maxtime support minlen maxlen target ext 1 0.1 1 none FALSE TRUE 5 0.4 4 10 rules FALSE Algorithmic control: filter tree heap memopt load sort verbose 0.1 TRUE TRUE FALSE TRUE 2 TRUE Absolute minimum support count: 40 set item appearances ...[2 item(s)] done [0.00s]. set transactions ...[26 item(s), 100 transaction(s)] done [0.02s]. sorting and recoding items ... [16 item(s)] done [0.00s]. creating transaction tree ... done [0.00s]. writing ... [96 rule(s)] done [0.00s]. Figure-9

The following figures predict the classification result in Tanagra tool.

			and the second second	Concession, name	<b>W</b>			
le Diagram Window H	Help							
i 🍻 📰   12								
O da ne	ning diagram (angly)							
				Components				
Data visualization PLS	Statistics Clustering	Nonparametric statistics Spv learning	Instance selection Meta-spv learning	Feature construction Spv learning assessment	Feature selection Scoring	Regression Association	Factorial analysis	
PLA	View multiple scat							



TANAGRA 1.4.36		A Designation of the local division in the	and the second second	Conception in the local distance of the loca		AND DOUBLE		- 0 🐰
File Diagram Window H	felp							
D 🥔 🖩   11								
Data mir	ning diagram (empty)							
			your dataset and start flowr Nagram tille ; Defoul tille Jata mining diagram file nar Windowskipstern 32/defau Indaset (* bet * arift * side) ;	ne : It tdm	Sancel Help			
6				Components		20.000		
Data visualization PLS	Statistics Clustering	Nonparametric statistics Spv learning	Instance selection Meta-spy learning	Feature construction Spy learning assessment	Feature selection Scoring	Regression Association	Factorial analysis	
Correlation scatterplot Correlation scatterplot Scatterplot Scatterplot with tabel View dataset	E View multiple scatte	mplot	<u>**</u>		. 6000		* • • • • • • • •	245 PM 6/30/2017

Figure-11

Import database file into tanagra.

TANAGRA 1.4.36 - [Dataset (DM DATA.xls)]		
Tile Diagram Component Window Help		_ 8 ×
D 📽 📕 🙀		
Default title	Results	*
[[[]] Dataset (DM DATA.xls)	Download information	
	Workbook information Number of sheets 3	
	Selected sheet Sheet 1	
	Sheet size 251 x 6	
	Dataset size 251 x 6	
	Datasource processing	
	Computation time 16 ms	
	Allocated memory 9 KB	
	Dataset description	=
	6 attribute(s) 250 example(s)	
	Attribute Category Informations	
	ROLE Discrete 3 values	
	GENDER Discrete 2 values	
	LOCALE Discrete 2 values	
	AGE Discrete 3 values	
	EDUCATION Discrete 3 values Al Discrete 3 values	
	N DISCIECE S YOURS	
	Components	
	rametric statistics Instance selection Feature construction Feature selection Regression Factorial a	analysis
PLS Clustering	Spv learning Meta-spv learning Spv learning assessment Scoring Association	
Correlation scatterplot Export dataset	icatterplot 🔝 Scatterplot with label 📓 View dataset 👫 View multiple scatterplot	
🙆 🧷 📋 🖸 🔕	· •	8:00 PM → 111 8:00 PM 7/26/2017
	Figure-12 Dataset description	1,100,2017



TANAGRA 1.4.36 - [Datase File Diagram Compo		the second	-				
🗅 📽 🖬 🔛	nene window neip						
⊡-IIII Dataset (DM DATA.x	Default title  s)				Results		
L → Si Define status 1		Workp Define Numbe Define Selecte Sheet s Datase Compu Alocati	d information attribute statuses attribute statuses attributes :	ОК Са	Illustative Clear selected		E
Data visualization	Statistics	Nonparametric statistics	Instance selection	Components Feature construction	Feature selection	Regression	Factorial analysis
PLS	Clustering	Spv learning	Meta-spv learning	Spv learning assessment	Scoring	Association	
↔ Correlation scatterplot	Export dataset	🔏 Scatterplot	🖾 Scatterplot with	label 🔛 View dataset	í.∶View multiple s	catterplot	
👌 🥖 📑	00	🙆 🖉		Contraction of the second	100.0		▲ 🔒 隆 🌗 and 8:02 PM 7/26/2017

Figure-13

Define status on parameters

TANAGRA 1.4.36 - [Dataset (DM DATA.xls)]	- Charles Backer	a x
Tile Diagram Component Window Help		_ 8 ×
Default title	Results	*
⊡-IIII Dataset (DM DATA.xls)		
	Download information	
	Workg	
	Number Define attribute statuses	
	Selecte	
	Sheet s	
	Datase Attributes : Target Input Illustrative	
	Comput DIGENDER LOCALE	
	AGE EDUCATION	
		E
	6 attrib	
	250 exa	
	Attrib	
	ROLE	
	GENDE	
	LOCAL 🗄 🗄 🧱 Clear all Clear selected	
	AGE	
	EDUCA OK Cancel Help	
		-
Data visualization Statistics Nonpar	Components arametric statistics Instance selection Feature construction Feature selection Regression Factorial analysis	
	Spv learning Meta-spv learning Spv learning assessment Scoring Association	
	Scatterplot 🖉 Scatterplot with label 🕅 View dataset	
		8:04 PM
🚱 🖉 🚞 🖸 🧐		7/26/2017
	Figure-14 Setting up input values	



1 🙀	
Default title	Results
lataset (DM DATA.xls)	
Define status 1	Download information
	Workh-thistestatuses
	Numbel Come activities activities
	Sheet s
	Datace Attributes : Target Input Illustrative
	Datas DROLE ROLE
	Allocat D LOCALE
	6 attrib
	250 exa
	Attrib
	ROLE
	LOCAL <u>B</u> <u>B</u> <u>Clear all</u> <u>Clear selected</u>
	EDUCA
	Al OK Cancel Help
PLS Clustering lation scatterplot	Spv learning         Meta-spv learning         Spv learning assessment         Scoring         Association <u>C</u> _Scatterplot <u>C</u> _Scatterplot with label <u>Spv learning</u> View dataset <u>L</u> _Scatterplot
lation scatterplot	Scatterplot       Scatterplot with label       Wew dataset       L View multiple scatterplot         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatterplot with label       Image: Scatterplot with label       Image: Scatterplot with label         Image: Scatte
lation scatterplot	Scatterplot       Scatterplot with label       Wiew dataset       Scatterplot         Image: Scatterplot       Scatterplot       Scatterplot
lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
lation scatterplot	
lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset Liview multiple scatterplot
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset Liview multiple scatterplot
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset Liview multiple scatterplot
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset E: View multiple scatterplot
Lation scatterplot Export dataset	Scatterplot Constant of the second se
Lation scatterplot Export dataset	Scatterplot Scatterplot with label View dataset E: View multiple scatterplot
Lation scatterplot Export dataset	Scatterplot With label Wiew dataset Wiew multiple scatterplot

♥●■Figure-16Supervised Learning C-RT

MIN SIZE OF NODE TO SPLIT means the minimum number of instances needed for performing a splitting of a node. For our dataset, we do not perform a split if there are less than 10 instances.

PRUNING SET SIZE means the part of the learning set used for the post pruning process. The default value is 33% i.e. 67% of the learning set is used for the growing phase (67% of 300 = 201 instances), 33% for the pruning phase (33% of 300 = 99 instances).

0

- 🕞 🌆 🌒 att

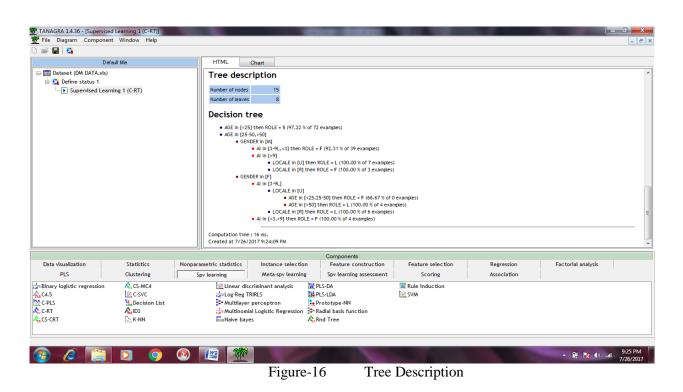


We have the sequences of trees table, with the number of leaves, the error rate calculated on the growing set and pruning set.

The error rate computed on the growing set decreases as the number of leaves increases. We cannot use this information to select the right model. We use the pruning sample to select the "best" model. The tree which minimizes the error rate on the pruning set is highlighted in green.

📽 🖶  👪													
Default title		HTM	HTML Chart										
Im Dataset (DM DATA.xls)     Cefine status 1     Log Supervised Learning 1 (C-RT)		Supervised Learning 1 (C-RT)											
		Parameters											
		Classification tree (C-RT) parameters											
		Size b	efore split		1	0							
		Prunir	ng set size	(%)	3	3							
		x-SE r	ule		1.0	0							
		Random generator			1								
		Show	al tree seq	(even if > 1	5)	0							
					_								
			Results										
		Error rate Values prediction											
				(1997)			0.0880	1					
			/alues prev	diction			Confusion m						
		Value	/alues pre-	diction 1-Precision		5	Confusion m	uatrix	Stam				
		Value S	Alues presented alues presente	diction I-Precision 0.0476	5	5 100	Confusion m	uatrix L D	100				
		Valu S F	Alues pres Recall 1 1.0000 1.0000	diction 1-Precision 0.0476 0.1453	F	0	Confusion m F 0 100	atrix L 0	100				
		Value S	Alues presented alues presente	diction I-Precision 0.0476	F		Confusion m F 0 100 17	uatrix L D	100				
		Valu S F	Alues pres Recall 1 1.0000 1.0000	diction 1-Precision 0.0476 0.1453	F	0	Confusion m F 0 100 17 117	L 0 0 28 28	100 100 50				
Data visualization Statist	cs NonDar	Value S F	Recall         1           1.0000         1.0000           0.5600         1.0000	diction 1-Precision 0.0476 0.1453 0.0000	F	0 5 105	Confusion m F 0 100 17 117 Component	L 0 0 28 28	100 100 50	tion	Regression	Factorial analysis	
PLS Cluster	ng S	Valu S F	Recall 1 1.0000 0.5600	diction 1-Precision 0.0476 0.1453 0.0000	F L Sum	0 5 105 tion ning	Confusion of F 0 100 17 117 Component Feature c Spy learnin	atrix	100 100 50 250 Feature sele Scoring		Regression Association	Factorial analysis	
PLS Cluster Binary logistic regression A, C-R1	ng	Valua S F L Spy learnin	Alues preve e Recall 1 1.0000 0.5600 attistics g C.SVC	diction 1-Precision 0.0476 0.1453 0.0000 Instan Meta-	F L Sum	0 5 105 tion ning	Confusion m F 0 100 17 117 Component Feature c Spv learnin NN	L 0 0 28 28 28 onstruction g assessment	100 100 50 250 Feature sele Scoring	rceptron	Association	St- Radi	al basis functio
PLS Cluster Binary logistic regression & C-RT C4.5 & & C-RT	ng S	Value S F L C Spy learnin	Alues pre-           e         Recall 1           1.0000         1.0000           0.5600         0.5600           ratistics         g           c.5VC         Decision I	diction 1-Precision 0.0476 0.1453 0.0000 Instan Meta-	F L Sum	0 5 105 tion ning K-I K-I	Confusion m F 0 100 17 117 Component Feature c Spv learnin NN mear discrimin	L 0 0 28 28 28 onstruction g assessment	100 100 50 250 Feature sele Scoring		Association	P-Radie	Tree
PLS Cluster	ng S	Valua S F L Spy learnin	Alues pre-           e         Recall 1           1.0000         1.0000           0.5600         0.5600           ratistics         g           c.5VC         Decision I	diction 1-Precision 0.0476 0.1453 0.0000 Instan Meta-	F L Sum	0 5 105 tion ning K-I K-I	Confusion m F 0 100 17 117 Component Feature c Spv learnin NN	L 0 0 28 28 28 onstruction g assessment	100 100 50 250 Feature sele Scoring	rceptron	Association	P-Radie	

Figure-16 Classifier Performance





# (IJMDEBM)

# 4. CONCLUSION AND FUTURE SCOPE

# Conclusion

- > The prediction of dengue disease is carried out by observing various symptoms related to dengue disease.
- From the above results it is concluded that, the main focus of the research is on the association among symptoms of dengue disease like fever temperature, rashes, muscle and joint pain.
- > The importance of knowledge identifies that disease.
- Dengue and from the observed symptoms it is analyzed that dengue fever mostly varies on fever temperature followed by Muscle and Joint pain and Rashes/red spots.
- Private hospital service to treat dengue fever is better than government hospital according to the concluded results.

## Future Scope

The responses gathered here are limited to 100 patients but we increase the simple size by considering more responses to get more accurate results. More powerful tools such as neural network, fuzzy logic or the other data mining techniques can be applied.

## REFERENCES

- [1] Alone, V. A., Manusmare, A., & Bhoskar, T. (2017). A Study Based On Women Security System. *International Journal of Science, Engineering and Technology Research*, 6(8), 1241-1243.
- [2] Amoatemaa, A. S., Kyeremeh, D. D. & Arthur, Y. D. (2017). Students Perception of Campus Safety: A Case of Kumasi Campus of University of Education, Winneba, Gha. Asian Reasearch Journal of Arts and Social Sciences, 3 (1), 1-9.
- [3] Ansari, A.H., Partiksha, B., Tejal, M. & Snehal, T. (2017). Women Security System using GSM and GPS. *International Journal of Innovative Research in Science, Engineering and Technology*, 6(3), 3706-3711.
- [4] Bobhate, A., Chutke, D., Udan, M., Pal, K. & Hulke, P. (2018). Hostel Girls' Security System. *International Journal of Advance Research, Ideas and Innovations in Technology*, 4(2), 315-319.
- [5] Bhagwat, S., Funde, M., Sonawane, R., Deore, S. & Ingale, S. (2021). Survey on Women Safety and Alert System. *International Research Journal of Engineering and Technology*, 8(5), 453-456.
- [6] Chakraborty, C., Afreen, A.& Pal, D. (2021). Crime against Women in India: A State Level Analysis. Journal of International Women Studies, 22(5), 1-18
- [7] Daniel, P., Backer, A., Shini, K., Nazar, S., & Joseph, S. (2019). Smart Band for Women Safety. *International Journal of Computer Science Trends and Technology*, 7 (2), 17-21.
- [8] Gadhave, S. N., Kale, S. D., Shinde, S. N., & Bhosale, A. C. (2017). Electronic Jacket For Women Safety. *International Research Journal of Engineering and Technology*, *4* (5), 858-861.
- [9] Kumbhar, S.S., Jadhav, S.K., Nalawade, P.A. & Mutawalli, T.Y. (2018). Women Security System using GSM and GPS. *International Research Journal of Engineering and Technology*, 5(3), 2406-2408.
- [10] Latha, K., Kumar, G.V., Naveen, P., Srikanth, B. & Srinivan, K.V. (2020). Women safety and security system using GSM and GPS. International Journal of Creative Research Thoughts, 8(7), 4618-4622.
- [11] Li, Y. (2021). Application of Big Data Technology in Campus Security Management under the Background of Information Age. *Journal of Physics: Conference Series.* 1881 022097 IOP Publishing, 1-7.
- [12] Likhitha, K.N. & Hemalatha, K.N. (2019). Women Safety Device using GPS and GSM Modem. *International Journal of Innovative Science and Research Technology*, 4(6), 780-782.
- [13] Mahmud, S. R., Maowa, J., & Wibowo, F. W. (2017). Women Empowerment: One Stop Solution for Women. 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE) (pp. 485-489). IEEE.
- [14] Pankan, N., & Radhakrishnan, N. (2016). Gender equality and women empowerment through egovernance: Case study of Akshaya e-centres of Kannur district. *International Journal of Applied Research*, 2 (4), 739-743.
- [15] Parsad, P., Nair, A. & Winster, S.G. (2021). Crime Against Women: Analysis and Prediction. *International Journal of Engineering Research and Technology*, 10(5), 426-429.
- [16] Pragna, B.R., Mahabala, P.P, Punith, N., Pranav, S. & Ram, S. (2018). Women Safety Devices and Applications. *International Journal of Engineering Research & Technology*, 7(7), 175-178.

© 2013, IJMDEBM, All Rights Reserved http://www.ijmdebm.org/



Volume-10, Issue-3, July-September, 2022

# International Journal for Multi Disciplinary Engineering and Business Management (IJMDEBM)

- [17] Ramya, K. & Vimat, T. (2020). Survey on Women Safety Devices. International Research Journal of Engineering and Technology, 7(8), 192-195.
- [18] Rose, S. D. (2013). Challenging Global Gender Violence. *Procedia Social and Behavioral Sciences (Elsevier)*, 82, 61-65.
- [19] Salunkhe, S. P., Sonawane, B. S., & Sonawane, G. N. (2016). A Review on Various Techniques of women security. *International Journal of Current Engineering and Technology*, 6 (5), 1811-1813.
- [20] Saryal, S. (2014). Women's Rights in India: Problems and Prospects. *International Research Journal of Social Sciences*, 3 (7), 49-53.
- [21] Sathyasri, B., Vidhya, U.J., Jothisree, G.V.K., Pratheeba, T. & Ragapriya, K. (2019). Design and Implementation of Women Safety System Based On Iot Technology. International Journal of Recent Technology and Engineering, 7(6S3), 177-181.
- [22] Sell, M., & Minot, N. (2018). What factors explain women's empowerment? Decision-making among small-scale farmers in Uganda. *Women's Studies International Forum (Elsevier)*, 71, 46-55.
- [23] Shimpi, T.R. (2017). Tracking And Security System for Women's using GPS & GSM. International Research Journal of Engineering and Technology, 4(7), 121-123.
- [24] Sui, X. (2018). Study on the Impact Factors of Campus Safety of College Students. Advances in Intelligent Systems Research, 163, 886-889.
- [25] Umadevi, B., Eswaran, P., & Manoharan, N. (2018). Womens Security Solution Using: Iot. *International Journal* of Pure and Applied Mathematics, 119 (10), 1871-1874.
- [26] Uzun, G. O., & Uzunboylu, H. (2015). A Survey Regarding of Domestic Violence Againts Women. *Procedia Social and Behavioral Sciences (Elsevier)*, 190, 24-31.
- [27] Vaghela, S. J. D., Shih, P. C. (2018). WalkSafe: College Campus Safety App. WiPe Paper Geospatial Technologies and Geographic Information Science for Crisis Management Proceedings of the 15th ISCRAM Conference – Rochester, NY, USA May 2018.
- [28] Verma, M., Manoj, M., Rodeja, N., & Verma, A. (2017). Service Gap Analysis of Public Buses in Bangalore With Respect to Women Safety. *Transportation Research Procedia (Elsevier)*, 25, 4322–4329.