World Journal of Engineering Research and Technology



WJERT

www.wjert.org

SJIF Impact Factor: 5.924



ANALYSIS OF MALARIA DIAGNOSIS ON PATIENTS USING DATA MINING CLUSTERING TECHNIQUES

Mani Shanker Chaubey¹*, Adelaja Adebayo Oluwaseun² and Nyaaku Oluwayemisi Esther³

^{1*}MSc Computer Science and Information Technology, South Ural State University, Chelyabinsk, Russia.

²MSc Computer Science and Information Technology, South Ural State University, Chelyabinsk, Russia.

³MSc Computer Science, Lagos State University, Lagos, Nigeria.

Article Received on 15/02/2020

Article Revised on 05/03/2020

Article Accepted on 26/03/2020

*Corresponding Author Mani Shanker Chaubey MSc Computer Science and Information Technology, South Ural State University, Chelyabinsk, Russia.

ABSTRACT

The research was carried out on the malaria patients with some symptoms on high rate that shows positive +ve result while those with some symptoms on low rate that shows negative -ve result. KNIME data mining tool was used to build a comprehensive work flow model consisting of nodes with their respective functions. Fuzzy c-mean, kmean and hierarchical clustering nodes were utilized to produce

grouped subsets termed clusters from the malaria_result.csv file (training-set). A decision tree level classifier was designed from the patient's diagnosis of the malaria symptoms. Data Analysis Knowledge Discovery Process for the clustering was also built. The result obtained in this research shows statistical clustering means such as scatter plots, interactive histogram, clustered data table and interactive tables which will be helpful for future observations and predictions of malaria in health care.

KEYWORDS: KNIME, Fuzzy c-means node, k-mean node, hierarchical node, Knowledge Discovery.

INTRODUCTION

In most medical sectors across the world today especially in Nigeria, large amount of medical

data related to patients suffering from malaria has been retrieved for medical history. The process of mining, analyzing these medical data records of malaria victims can effectively support medical practitioners and the health sector to predict likely occurrence of malaria thereby taking preventive measures to reduce the rate at which people contracted this ailment now and the future. The World Health Organization (WHO) statistical reports proved that malaria is the second leading cause of mortality from vulnerable disease in most African countries after HIV/AIDS.^[1]

The *anopheles mosquitoes* bite on human transmits *genus of unicellular eukaryotes* (obligate parasites) called *plasmodium* is one of the main causes of malaria. There are different types of plasmodium parasites but only five types cause malaria in humans. *Plasmodium falciparum* which is the most common malaria parasite, mainly found in Africa also responsible for most malaria deaths worldwide while the *plasmodium malariae* is quite rare and found in Africa. The *plasmodium ovale* is a fairly uncommon parasite but usually found in countries located in the western part of Africa; this particular parasite can reside in the human's liver for several years without showing the victim any symptoms.^[2] The *plasmodium knowlesi* is found in Southeast Asia. In some countries in Asia and the Southern part of America, the *plasmodium vivax* can be found, it causes milder symptoms than the *plasmodium falciparum* and it can stay in the liver for an approximately 3- 4 years which could result in relapses.^[3]

The World Health Organization, World Bank, UNDP and UNICEF established a new initiative health sector-wide partnership known as the Roll Back Malaria RBM to support researches and discover products such as drugs to combat malaria at global, regional, country and local levels. The Roll Back Malaria movement saved an estimate of 3 million lives across the World; the global mortality rate drastically reduced by 43% while in Sub-Saharan Africa regions (such as Botswana, Gabon, Rwanda, Uganda, Senegal, Tanzania, Zimbabwe, Congo, Ghana and especially Nigeria) the rate decreased by 49%.^[4] This paper focuses on using the data mining clustering technique to effectively analyze the large amount of data and medical examination history obtained from patients who were diagnosed of malaria to transform this data into useful knowledge. Data mining is a process used in extracting rules; predicting certain performances in numerous professions such as medicine and surgery, sciences, military, aviation, education, human resource and information technology.^[5] Data mining can be described as the collection or gathering of pure driven data algorithms to achieve

meaningful knowledge pattern from the raw data. Data mining also known as "data dredging" or "data fishing" or "knowledge mining in data" is a process that involves searching of large information of data or records to discover knowledgeable patterns and utilizing these patterns for predicting future occurrences especially in the medical area which is the main scope of this research.^[6] As a result of technological advancement, it is possible to gather, store and easily retrieve large quantities of medical records containing vital information which could help in diagnosing and treating malaria. Data mining techniques such as clustering can extract intelligent knowledge through data analytics from medical data for convenient treatment and diagnosis. Clustering is a data mining task of assigning a set of objects into groups called *Clusters.* Clustering is a solution for data analytics which involves partitioning a set of data objects into subsets. In clustering techniques, objects in a subset are identical to one another yet dissimilar to objects in other subsets.^[6] Clustering is an unsupervised learning performed on data sets by observations, this is before the semantic of the classes is not known beforehand. The main types of clustering methods include partitioning method, hierarchical clustering, fuzzy C-means clustering, density-based clustering, model-based clustering or grid-based clustering.^[6] The main goal of clustering in data mining application particularly in the medical aspect is group both similar and different set of data objects derived for the medical results into the same clusters and different clusters respectively.^[7] Clustering can also be defined as a process of grouping a given set of unlabeled patterns into a number of clusters such that identical patterns are attributed to one cluster.^[8]

Decision Tree Level Classifer

Decision tree is a tree-like structure consisting of internal nodes, branches, and leaf node, in which each branch denotes an attribute value, each internal node denoted a test on an attribute which is used for and a leaf node represents the predicted classes or class distributions.^[9] The decision tree have *nodes* that form *rooted tree*, this implies that decision tree is a *directed tree* with a node known as *root*.^[10] Decision tree supports a predictive approach in data mining and machine learning.^[11,12] The decision tree level classifier used in this research is an organized hierarchical structure which shows set of condition such as the list of malaria symptoms in the patient (high temperature/ fever, nausea, common cold, headache, body pain, diarrhea and vomiting), the decision and the result of the decision which states YES which depicts the test is positive +*ve*; or NO which depicts the patient is being checked for other malaria symptom, else they are tested negative –*ve*.

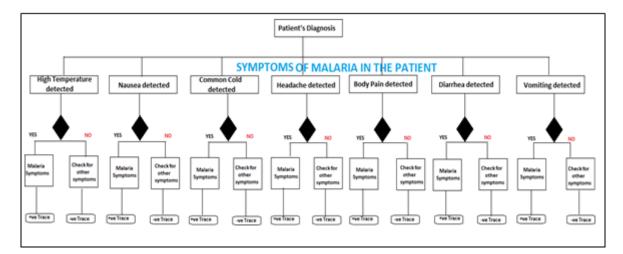


Figure 1: Hierarchical Structure Showing the Malaria Symptom and Results.

Fuzzy C-Means Clustering

The fuzzy c-means clustering algorithm is a well-known unsupervised learning technique that is used to reveal the underlying data structure. Fuzzy clustering allows each data point to belong to several clusters, with a degree of membership to each one.^[13] Fuzzy C-Means clustering is widely applied in medical diagnosis, shape analysis, image processing and analysis, target recognition, geology and engineering. It is a data clustering technique used for analyzing distance between various input data points.^[14] Clusters are formed based on the distance between data points, likewise the centers of the cluster are formed for each cluster.^[15]

Hierarchical Clustering

Hierarchical Clustering approach creates a decomposition of data sets (or objects) in multiple levels of hierarchies using some criterion. It is method of analyzing clusters with the aim of building the cluster in a hierarchical form.^[7] There are two approaches to hierarchical clustering known as the *agglomerative/ bottom-top* and *divisive/ top-bottom*. The agglomerative hierarchical clustering starts with a single data point and merge two or more cluster in a recursive pattern.^[7] The divisive hierarchical clustering starts with big cluster and splits this cluster into smaller clusters in a recursive pattern.^[7]

K-Means Clustering

K-Means is one of typical partitioning clustering approach in which each cluster is represented by the centre of the cluster. It is a method of clustering observations into a specific number of disjoint clusters.^[16] The aim of the algorithm is to minimize the measurement between the centroid of the clusters and a given observation by iteratively

appending the observation to and clusters when the lowest distance is achieved. K-means performance is determined by initialization and appropriate distance measure.^[17]

Data Analysis Framework

The result obtained from the malaria diagnosis conducted on the patient is termed as the RAW data; to derived meaningful information from this *RAW* data, we decided to perform data transformation and cleansing process before analyzing these data with three clustering methods (k-Means, Fuzzy C-means and Hierarchical clustering) respectively. The main target is to easily produce different set of clusters from the respective clustering method which we utilized. The final stage of this data analysis framework is the report patterns generated from the *interactive histograms, interactive tables and scatter plots* in the three clustering approaches employed in this work. From the outputs of the report patterns, we could discover knowledge due the *RAW* data which is being interpreted and effectively useful for most medical centre across the counties where malaria parasite is predominant for convenient diagnosis now and the future.

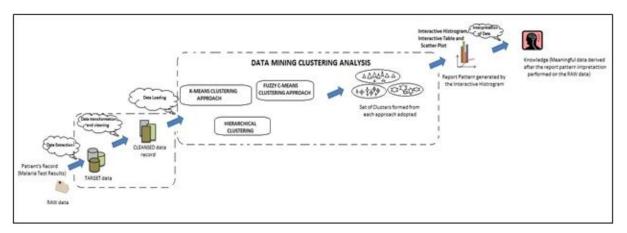


Figure 2: Data Analysis Knowledge Discovery Process on the Malaria Results.

RESULTS AND DISCUSSION

The goal of this research was to build a workflow model which reads the malaria patients data from **malaria_result.csv** file, perform three major clustering (fuzzy c-means, k-means and hierarchical) for the chosen attributes (Patient Medical File No, High Temperature, Nausea, Common Cold, Headache, Body Pain, Diarrhea, vomiting and the test result) hence displaying the results. The workflow model was build and the nodes were connected using the KNIME data mining software. Figure 3 shows the workflow model:

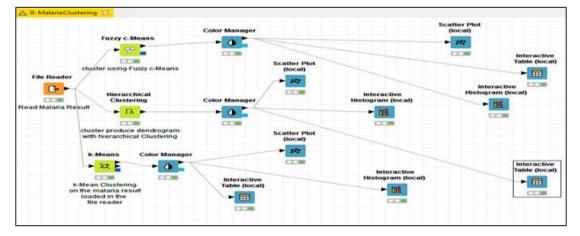


FIGURE 3: Workflow Model that read the malaria result and perform clustering.

3.0 Fuzzy C-Means Node Clustering on the Malaria Results

Fuzzy c-means clustering on the malaria data produced three clusters (cluster 0, cluster 1 and cluster 2) with their clusters centers on each attributes. Other results such as the within and between Cluster Variation, the value of partition coefficient, the value of the partition entropy, the Xie-Beni (XB) index and Fuzzy Hyper volumes for the three clusters were computed as shown in Figure 4:



FIGURE 4: Statistics View of the Fuzzy C-Means Clustering on the Malaria Results.

3.1 Hierarchical Clustering on the Malaria Results

Hierarchical clustering on the malaria data produced a dendrogram of the clusters on the row attributes, the distance between these clusters and the clustered data table produced by the dendrogram.

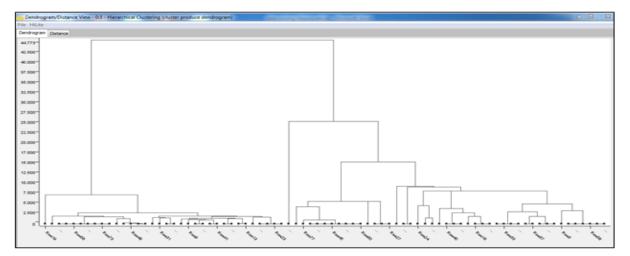


Figure 5: Hierarchical Clustering Dendrogram produced on the malaria result.

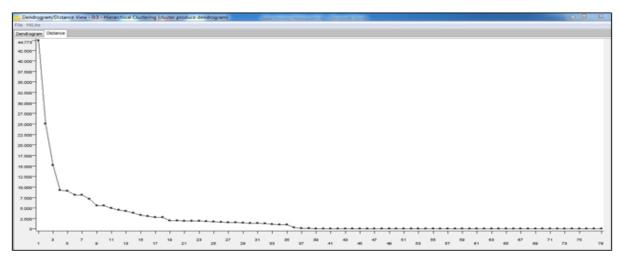


Figure 6: Hierarchical Clustering distance between the clusters produced from the data.

		an Nama Phoney Manual Plana								
Water all.b	IN TRAFFICE PRODUCTS FILLE AND CO.	D HOUR TENERISATING (PRIVER)	TOT FOR MAN	TEL COMMON COLD	TRA LABORATION COMM	THE BOARD PARTY	D DIAMONA	THE VOIDTHAN	IN PRIMAT	I.B. Chests
C. C	ANT ALD 4	64.3	11.41	13.76	46	48.8	2.3		annamonal +)	sub-radius all
Rosters \$10	A A FT & ACRA	10.03	38. 2.4	19	1	14		4.28	President (-)	Charafone _ 8
Poper IVA	BADY MINIA	58. C	34	1 A.	10	30.	0	1	Compatition 6.3.	coloring and taken and
Property Bills	AND ALDE	(A) (A)	28.27	21.4	18		0.01	10.1	Compation (-)	sharter_1
Russe 5-5	AND ALL M	22.2	39.8	8.4		39.4	0.45	- P.A.	president (-).	shanher _1
R.a.m. 5/5	LOT SALT	22	2. I 2. I	30.4	12	2.4	0.05	10.3	prompto and a later	
Report 2	LOTALES	E.C.	0.0	6.7		22	0.85	12.1	Congelition (-)	(Buthlaw _ A
Rosert. 7	AATT & LOW	0.8	S.1	0.4			0.45	10.4	Company a 1-1	chaster_k
Rame P.S.	8.0.77 8.5.418	6.7	G. 6	50.4	12-	10.4	0.05	6.4	Companyation and	chather 1
Page 198	AAAT ARD 4	10.9	0.3 8.7 8.7	31.4	18	10.A		10. k	Competition (-)	Charles - 4
Wanter Web	A LOT A LOSS	10.33	24. 2/8	21. 3.7	10.91	14.124	0.002	1	Pargation (-)	there have been a
Ruber 47	LOTALES	10.88	38.398	31.47	3.34	30.795	8.24	6.62	THE BATTLE	ububber 1
Allower and		10.25	Mumit	35.47	B-34	0.95	8.24	NO.404	Companying C.S.	
Former D. S.	AUTERDO.	39.83	N. 15 31.15	10.87	18.55	0.98	8.24	0.03	President and the second second	colleged risks
No.ex PR	B-1477 8 8-46	19-20 ·	M. 19.	10.57	28.88	0.08	3.84	10.00	Pargative (-)	ChanlandA
Rapes 17	LAPTRESP LAPTELSE		1.2	1.2	00.5		0.4	0.3	Prespetrice 5-3	chaster_5
Second Second	1.1.17 10 1.2		1.2	1.2			8.4		(respective (r))	chaster _ 1
Rates 140	84/70833		1.2	12.5	10.5 0.1		1.1	18	Pergative 11	discriber _ 1
the second s	\$10T \$30-\$	10 ·····	1	31. 3.9	12.2.4	1.97	1.00	18	President C.S.	chusiles 1
Room W	3.07 1238	3.78	10	31.89	10.38	1.27	0.990	0.23	presidente (1)	uturation 1
Robert 348	BALFY BANK	DOL COL	(8, IF	24.30	10.3	1.79	0.4	10.24	Comparison (1)	
5pm-43	3.3.77 1004	(F)	11	Dr. 8	11.3	1.4	04	0.28	Companying 2-2	distantion
Date: 345	LOT LLES	1	14.9.	(A. 8	34.9	1.5	1.4	10.2	(-) might in (-)	O'LANING1
Rates 45	LAPT SOOT		4.3	10.5	36.3		0.4	10.3	(regettion (.)).	chaster_1
Autor 345 Autor 344	1073333	2. · · · · · · · · · · · · · · · · · · ·	16.8.	1.2	38.3	1.0	0.4	10.8	comparison (1).	chanter_1
Patrice 2	LOTELSE LOTELSE	2-1	12.2	- 54	13	13	0.4	0.8	Company of a	churcher
Salar 2.3	AUT 1254	1.04	11.1	12.2	10.1 10.1	1.3	0.4	10 Z	Compation (.)	city and black
have det	LAPT LEET	Contraction of the second seco	12.2	14.9	39.6		4.88	50.34	THEOREM CO.	Chatcher _
Equat 1.8	AATTOR DR	40000	10.4	18.8			0.0	1.0	COMPANYIN C.S.	Musher . 1
Range II h	AAFT ENGS	4	10.3	DALK.	2.2	12.0	13.1W	10	Printers and	citizen fear
Robust & S	AUT 104	4.7	10.8	8.8	34.9	18.04	0.0	8	president and the second second	shuteter _1
faire 23	JULY \$ 104	A	(B. A.	A. 8	10.5	34.4	6.4		Congentrum (-)	ichaster_a
Rater Life	LOTER A	-10	18	19	88.8	1.8	8	4	passers at a s	chuster_3
Room PP	1.0711.099	44 45	12	19	120	1.2			(analyzed)	(6.48 her _ 2
Palacelli Restore Mill	LAT LOOR		12	15	100	13	5 ·····	17	presentences a pro-	obuster_3
Region W.S.	ALLY ADD-6 ALLY ADD-8	43	12	10	21	12		12	Construct +	Character _ 2
Popper 418	LOTIEN	45	15	125	100	122	G	12	personal and a 2	OLASTAR
Roam W	ALL Y AL DIS	45	5	100	20	125	<u>6</u>	12	present of all	sharter _
A state light	3473188	144	GA	766	1000	DAR.		14	growthy ad all	shutter 2
Rate or S. S.	AUTILES	45	14	1.8	100	102		14	presenter all all	Automat 2
A state of the sta	AUTILES	145	14	1.6	200	4.8	8.0	14	and the set of a	Charles _2
Kana Md	34771113	43.5	14	10.1	10.00	1.4	1.9		And ALL ALL ALL	Churchater
1,000,44	3 APT NOR F	43.9	(H)	TEK B	19.3	1.0	6.0	14	general Anna C. A. 3	ob.ater_2
Kalking B.	SLOT SOME	43.9	141	112.8	19.3		0.9		general Anna (. A.)	infuence2
Anna (C. P.	3.477 1.3008	25.5	A	147.8	44.8	148	5.8	- Basic	printer al.	Acades _1
Lapine 3. W	3.07.1002	11.2		Sec. 1	14		1. X	1.4	p-048/44(+)	Charlotter
Later P3	A STY S S OF	12	1	1000	14	12.4	1.44	2.7	promition (+)	states _2
100 P 2	ALIF LINK	73.4	12	5.64	10	49.3	5.44	12.2	propriet al al a	Autor
and play	AAUT 1 100 h	88.9	8		14	19.3	5.64	9	g-0-041-0-2 + 2	Chamber _
an an at a	A APP A AREA	INST N	28	17.96	499	1296.00	11.05/0	84	Butchild Brush C + 3	city action of
an an 290	LUTILUS.	30-9 () () () () () () () () () (58. · · ·	28.	11.4		8.23		possibles(a)	ob.Attar_1
400-000	AAPP science	89.4	(8.4)	1.82	10		2.3	1.9	and the second second	LAJAMET
alaren (196	340743300	(ep.4)	14.9	18.87	84	14.P.W	2.3	26.9	geographication (+)	placeters
an an 70 P	3.477 9090	89.4	3.8	1.30	16	185.8	5.5	11	positive(+)	prin-actions
Lakardw.	3.07.3.008	819 867.5	G	12	17	17.9	0.0		annews and a g	chatlant_1
total Life	ALT LINE ALL ALL ALL ALL ALL ALL ALL ALL ALL AL	M.F. 1 M.F. 1	12	12	14			1.4	presenter and a state	Autofair_)
anno 1.1	AAT 1222	22.1	14	14 C	14		1.3	3.2	Annantice (+)	chunter
and a state	34/7 1 10.7	10. N	2.0	1.47	110		25	16.1	promitive (+)	chains _
and all a	3071114	29.5	10.0	1.5.9	1.0	17.0	2.3	2.2	promition (+)	Charles _1
ann 24	AAPT mint s	89.8	3.0	18.87	100		2.3	3.9	Annual Control of the	Autor 1
at an in the	ANT AADO	89.5	(g.w/	1.37	10	33.9	2.3	3.9	generative and and	chuster_3
Lance B-B	34774339	58.04	14	24	6.4	1.7.9	8-3	28-38	prover to come (+)	
lation B-P	3471138	38.01	5 <i>4</i> .	54.	1.4	14.8.8	R. R	28.78	proverbire (+)	Aurolatian1
Laters Phil	3.5.77.1.3.98	56.2	58.	LA.	6.4	117.9	2.3	34.4	Another and a lot	ALMANER_1
Labore B	AA/Y on the	56.2	12	58	314		2.3	8.9	\$1000/01/IEE #3	charter_
Aprile 200	3.071109	94.3	10	10	105	32.8	8.3	8.9	promitive 2 + 2	situation _1
an in Sec.	3.07 1013	11.1 	Q.46	Sec. 1		10.5		15	positive(+)	chanter_1
Tables 7.4	1071185			12.74	248	48.5	2.3		and a second sec	charber_1
fater 74	3.0711187	34.2	12.46	2.78	124	140.0			Anterio (Contractor)	in-ener

Figure 7: Clustered Data Table produced from dendrogram hierarchical clustering.

3.2 K-Means Clustering on the Malaria Results

Clusters produced on the malaria result from the k-Means clustering were cluster_0, cluster_1 and cluster_2. K-Means clustering on the training data load generated a cluster view and k-means cluster table.

File Hilite N	avigation View					
Table "default"	- Rows: 3 Spec - Columns: 6 Prop	erties Flow V	ariables			
Row ID	D HIGH TEMPERATURE (FEVER)	D NAUSEA	D COMMON COLD	DHEADACHE	D BODY PAIN	D VOMITING
duster_0	5.331	1.564	1.435	0.332	2.184	0.877
duster_1	1.081	0.54	0.772	1.105	1.256	1.127
duster_2	57.8	3.21	7.05	18.031	16.211	5.647

Figure 8: K-Means Clustered Data Table on the attributes.

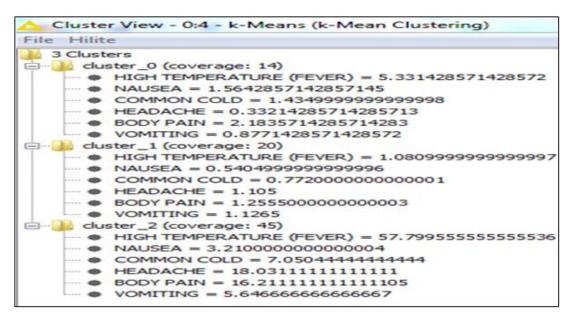


Figure 9: K-Means Cluster View on the attributes.

3.3 Scatter Plot (Fuzzy C-Means, K-Means, And Hierarchical) Clustering.

Fuzzy C-means Clustering Scatter Plot on the attributes.

4					
		100		a second second	
				12	
-				440	
-				2.44	
				-	
				100	
		-		12	
				10	
		1.00			
and the second se				-	
-1-2-24		-		-	
		1		1.5	
4				15	
1		3		A same of	
Allen 12.2.2 genes.		daring trends game		choices into a permit	
1984		1	2 Mar. 1944 (2 4)		
a second se	10 mm	AT the Design of the		1.00000000	- (Ter - 1 - 1
				Terrar and the second	-
And the second sec	-	A DATE OF A DESCRIPTION		Contraction of the local division of the loc	11
	Sec. 12 Sec. 1	a de tarres		1	in the second se
	Sec. of the sec.				
	Art of the second			-	
				-	
			and the second s		
			and the second s		
			and the second s		
	8		,		
			and the second s		
See	8				
			,		

Figure 10: Scatter Plot generated by Fuzzy C-Means on the attributes.

Hierarchical Clustering Scatter Plot on the attributes.

	ter Plot - 0.37 - Scatter Plot Doca0				
Film 3455	Liter Unover-Pilole				
***	1				
8.07					
4.5					
4.0					
3.8-					
3.0-					
2.81					
1.0-	· · ·				
1.8-					
1.01					
	5.				
			80 88 80	64 68 72 76	80 84
Default	Betterge Column Selection Appendicate				
	X Column Y Column	Xatelad	Am (+1)	# P (0)	
	D HOUSE REPORTATION OF D FAMILIER	* v ambu	Ar (+10)	+101	

Figure 11: Scatter Plot generated by Hierarchical clustering on attributes (High temperature and Nausea).

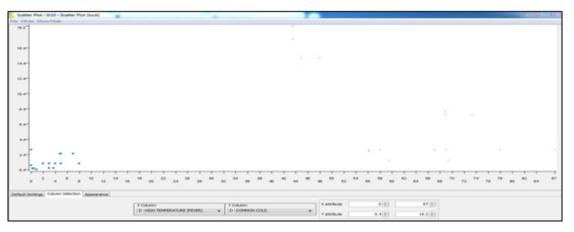


Figure 12: Scatter Plot generated by Hierarchical clustering on attributes (High temperature and Common Cold).

~															•													
		 	 	+		*	1	,	 	 - * *	-	-	-		****										-			-
																							0.940	14				
	•.••	 	 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			 •••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	····	····	••••	••••	•••••••••••••••••••••••••••••••••••••••	····	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	•••••••••••••••••••••••••••••••••••••••	

Figure 13: Scatter Plot generated by Hierarchical clustering on attributes (High temperature and Headache).

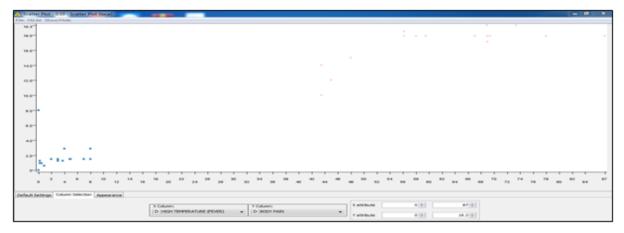


Figure 14: Scatter Plot generated by Hierarchical clustering on attributes (High temperature and Body Pain).

Scatter I				100.00							-	_																													100	
File 141,84	APalente	(below																																								
0.00																																										
0.00-																																										
4.28-																																										
1000																																										
3.00																																										
3.78-																																										
0.00-																																										
0.041																																										
4.000-0																																										
10.000																																										
1.081																																										
1.007																																										
1.00-0-0																																										
1.00**																																										
0.78-	-																																									
120.047																																										
0.00**																																										
0.00-																																										
	100																																									
			1			· · ·	100	1.00		10	11	100				1.0			10		1	1.	10			10				1.1	100	1.1	1	1.7	100							
	1.0				• • •		•	-	-	-	. 88.		24	144	- 10	- 84				40	-	144	-				 		-	**		**	-			1.64	140	.79	-	46.	-	
linfa.it Satt	Arrest 4	Column	frank in	1.00	-	-																																				
								112	1.1								100							1.12	-	1.1		-010				47.0										
								1.1	D-1et	tioni Biles R	and in case	in the second		-			1040	Advanta	i i i i i i	weite		ook m	60. w			-																
								- 12	0.00		-	00.04		1000		10	10,000			- 10 V	0.00			58 (v	-	100		010			- 7	49.0										
1 m									_	_	_			_	_		_	_		_				- 14	-		 _			_	_		1.1	_	_	-	_	_				-

Figure 15: Scatter Plot generated by Hierarchical clustering on attributes (High temperature and Diarrhea).

	Scatter Part 1 10 10 - Scatter Part Docat						
	10.00 (December 10.00)						
	au-						
	68						
					1.1.1.1	A	
	u-						
	na			X (8)			
	· · ·						
	. i						
· · · · · · · · · · · · · · · · · · ·							
		* * * * * * *			** ** ** ** **	· · · · · · · · ·	· ·· ·· ·· ··
		· · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·		· · · · · ·	n ⁿⁿ m ⁿⁿ ₁₂ ⁿ	
		X Columni O Indian Temperanume (Peyero	O VOMITINO	x attribute	0.00	#7(5) *(5)	

Figure 16: Scatter Plot generated by Hierarchical clustering on attributes (High temperature and Vomiting).

K-Means Clustering Scatter Plot on the attributes.

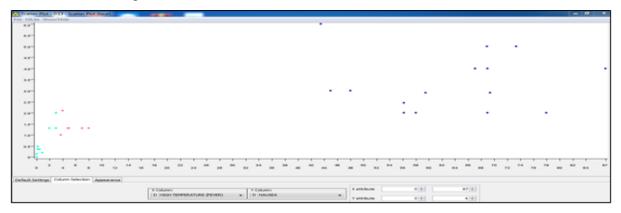


Figure 17: Scatter Plot generated by k-Means clustering on attributes (High temperature and Nausea).

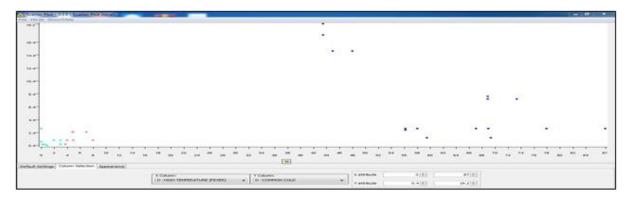


Figure 18: Scatter Plot generated by k-Means clustering on attributes (High temperature and Cold).

A Scatter Plot - 0.19 - Scatter Plat Dacat					-
File Hillie Shore/Hide					
3.89					
3.80-					
8.28-					
1.00				-	
8.00**					
a.74-					
2.00**					
3.24-					
2.00***		1.			
		-			
1.78-					
1.007					
1.98-					
1.00**					
				-	
0.78**					
0.80-					
0.26-					
	10 ¹⁰ 20 ²² 24 ²⁰ 28 ²⁰ 20 ²⁰ 20 ²⁰ 20 ²⁰ 20 ²⁰ 40 ⁴⁰			· ··· ··· ··· ··· ··· ··· ···	** **
Default Settings, Column Selection Appearance					
		for set of the			
	X Column V Column	X albitude	0.01	**)0-1	
	D HERH TEMPERATURE (PEVER) & D DEARSHEA (CONSTANT)	rockers) w	0101	3.69(0)	

Figure 19: Scatter Plot generated by k-Means clustering on attributes (High temperature and Diarrhea).

3.4 Interactive Table And Histogram Fuzzy C-Means Clustering.

Annual Mill	I B Person report.	or want solution comments	man in the second	Call the second	I III stantes.	THE BOARD P.	The designed of	THE OWNER PERSON	I IN TRADING IN	THE HARPENE A	and charmen at	I D -torne of I D -
Bandi	Bury alter		14	- Dia	Data.	10	3			2	Contraction of the local division of the loc	
marks 1	BAUT states	246.8	34	147.W	346.8	80	14.W.	34.	present with a little	1	A	
Barry B	Book all all all all all all all all all al	246.8	24	18	34	10.0	22		Second-secol	8 .	the local division of	20.49
Barry B	Book works		10.2	- 6.0				10. X	AAAAAAA 2.5		1	2. 20.00
Real Property lies	- BART MEND	54.3	- C *	- C. K.		524	22	10.	requestion of a	_		20.000
Barnit	BUT HER					452	0.0	10.0	Annual Annual A			
Bart 7	- 100 Lane		C	Do www.	5.44	11.00	S 24	-	California C.S.	_	_	2 SAAR
Rear and	Burt Laws	2.0	14		14	Sec.	20.0	16 W	Alternative and			
B. (1978)	ALC: NOT LODGE	10.75		. ave	36.94	10.00	14. MAL	20.00	Parameter of the	-		y Laws
B. (p/s, 1)	Burit and a	145	38.	248	_34G	1417	10.1	14	annexity of all			Statement and statement and statement and
Burn Liv	210T 1204	347.4	24.		144	117.00	5.4	34.34	presenter and a la		the second se	 Interime
Acres 5.0	BUT Aller	1.00	19	194	344	288	14	19.00	proton Aliante and all			
BOOK 8.9	BACT LOWE	24.98	24.8	. 24.8	39.3	1.5	34.4	36.30	Swintfield 2.3	the second se		3 30.494
Bab.1.9	BOT SHEET	34.7		- B	144	10.0	11 ·	19.9	prosetty and a d		and the second se	Jacob Schoolse
Barris I.C.	- Borenne		10.5		2.1	11			Pressent 6.8		1	2
Same 1.4	L. Pasts					10.0	22 ·····	10.0	Construction 2.2		1	2
Rain 18	ALC'T A MUSE	and a			5 m	line .	<u> </u>	2.2	and a set of a	_		· (212
East 19	a DY Later	10.000	36.34		1			N 14	Pagaters C.A.			> 1434
Burney (Md	Burr same	242.2			244	Gen	1.44		allocation and all			· Charles
*	BUT LONG	14	10 A	5.4	8.4	10.0	1		Catalantina C.3			2
B. (2/2/2/2)	Boold States	34.5	18. · · · ·	19. au	340	108.0	8.85	10.8	stream in a set of a	1	Statement of the local division of the local	· · · · · · · · · · · · · · · · · · ·
Busie 213	BOAT LUDA		34.3				34.4	1.00	Company of Street State	-		3 Lines
Rest (14	BOUT AND IN	278.4	28	10. Mail	194	14.3	10.00	1.1	and and a state of the			-
* and a *	BOT IT IS IN		34.3	- DA.M.	98.8 ·	31.9	14.4		Frankrik di S.		1	3
Taxa and	BOT SHET	100.0	14 M	DA WY	100.0	34.68	34.4	5.5	and the set of the set		The second value of the se	a jalaka
Report 117	Burl an an	34.1		20.8		13 a	81	2	and the second second			
A second and	107 10 m			1.44	12	100	2.5	22	and the set of		and the second se	See Shares
Raph No.	BART STATE			10.01	so	1.1	£2 ······		Completion ()			5. (hate
Rank No.	Bull Hannah			10.0	SS	12.2	83 ·····	- CA - C	Press of the local division of the local div	-		S (212
diamental distance	And States	142.8		table in	GG-	1.0	1.4	1.00	and the second second second		-	
4 mm 14	B-107 13 17	200.8		10.47	146	Carlos .	11	10.00	anneth-art.to	-	-	Ce Databa
B-000 748	Boold an all		35.8	1.8	34.4	1.0	14 H	20. 8	Parameter de la	The second se		3 States
Read of Lot.	Burger and see	1.0	34.9	24.8	31 ····	14.0	11	201,248	providence (CA)	-		3 20,000
Acres 14	8-14 13.2P		1.8	14.8	24.8	1.4		201.0	Presenting G.S.			3 28.494
Real Property lies	Saur acces	244.4	(A. 9	34.87	46	10.00	38.38.1	8.9	(and the second		_	A Shutte
Barriel 100	5-LPY X00-1		(A. W	14.84P	146	1870.0	2.1	N.P.	anniti all'a		the second se	A Jacobs
Barry 20	- Boot works			1.00		2		12	Press A CA	-	£	A Distant
Rannell .	ALC: NAME	10.4	- 21	- 18.85	12.	354	81	22	2000 m (m) +2		_	1 10.000
Barr. 70	ALC: A DOCT	148		104		110	8.4		Pressence All		(
Bark 53	BLOY antise		50	54.0	10 m	4.8	10.4	Sec. Let	president and a little	_		2 Diana
Barn. 198	Burly How P	244.4		1940.4	10.0	140	1.4		province (r)			and the second second
Barrow 1975	Board \$10.000	1411		100		144	1		Annany miles		6	**************************************
B	BUT 15.85	10.24	36.341	20.07	34.84	1.04	No. and	14.24	Distantion LT	-		P Dealer
B.m. 127	Barriel Barriela	15.93	20.002	10 H F	28.915	10.00	10.000	10. IN F	Amperior 2.4	Married Woman and State		and the second s
B	Dou'V LINKS	34.34	36.87	18 N.T.	24.94	34.44	4.34	la má	Endowed and a second		4	a la suba
A 1011-448	BANY BILLING	140	14	1/8	140	14.4	14		preserve and all		A	
Barry 919	BAT MARK	10	Market Street	199	347	1.8	10	10 m	preserve (a) + ().			
Rank's	8-07-16.80	24.2	(2.44	a. 194	540	10.5	2.2	- C	personal +)		and the second se	and Shade
Read and	5117 21.00 0.07 10.00	100	- 81		240	198	10.44	10.0	present end ed.	_		- Andrew Statement
And and a state of the state of	SAUT LARS	10.0			1.	10.0			present of all	_		- 25.404
Rea 10	LOT LLT.	10 m			10	10.0	1.01		Presentation and a little	_		- ILII
Read Trees	LOV LLNP	2.2		8.2	1	2.4	0.55	- C - C - C - C - C - C - C - C - C - C	Company of Land			1
A second of	And Taxable	204 214			10.0	14.0.0	10.10	D. W.	And and a second second second		_	-
Barn. 58	ALCY LUMB.	94.2	5. MI	12.00	100	140.5	2.2	14	and and the set of the		_	
Barriel 108	Burly and set	248	14		118	11.0.0	10.000 C	16.W	and the second		The local division in which the local division is not the local division of the local di	Mar Inches
Barrella	Brook an en	148	54	144	Saw	24.8	24	a.	present and add	F		and the second s
Allowed at 1	BUT BUT	34.94	34.400	38.817	24.84	30.88	13.34	0.01	Pressource Ld	-	A COLUMN TO A COLUMN	8 (m.ama
* another .	BOUT HILD'S	1. 9	34.3	10. M	TA	30.46	10.45	(0.) ·	And the Party of t	-		1 20.400
Rosenil a	BOT ISSN	15 W	14 A	(A. 4	D	31.4	5.44	38.8	Property and a little	and the second division of the second divisio	A	a jacana
Renality .	BOT MADE	24.2	36.94	10.04	199	199.9	18-5		(construction (construction))		The rest of the local division of the local	
Ranal A.	Bull 11.24	21	5.45	- 22	148	198.8	10.0		Journey and 4-3	-	-	a she
Same a	8-14 11.248	10.0	2.51			10.4	1.00	- C	personal all.		and the owner of the owner, where the ow	
	Brown British				- Contract	22	10.00	- C - L	Falgerrad, E.S.		-	1.404
Report R	B.OT LEBT	24.2		1.87	10.4	10.0	1.00	A 44	Presenting 2.2			30.400
Warm No.	LUT LUB		2.4	10.41	10	103	2.54		aman.a. 1)			0.ells
Range Pa	ALL TIME	22.2	6.84	1. m	100	10.0	10.0	6	present at 1		_	and Diverse
Read Vol.	and the second				144	1.0.0	2.44	10 m	president and a d			-
Rank, P.S.	5.07 11 MI	2.4			11	5.4	2.44	6.1	Privation 3.4	-		1.454
Room Pd.	Sand States	10.24			the second	2.24	14-14	10.00 m	Property at 1.1			Ja.ete
B.44.7%	LUTILAT	24.4	6.6	1.79	144		12.3		present of a		_	-
Rate 75	Bard states	244.2	10		114	110.0	10.0	2.5	Arrest (+)		the second se	and Disate
August 177	BOAT IN ME			100	140	140	1.0		Annual Property and and			
A	ALC: Y LINE L						in the		Departure La			a literatur

Figure 20: Interactive Table generated by Fuzzy C-Means clustering on all attributes.

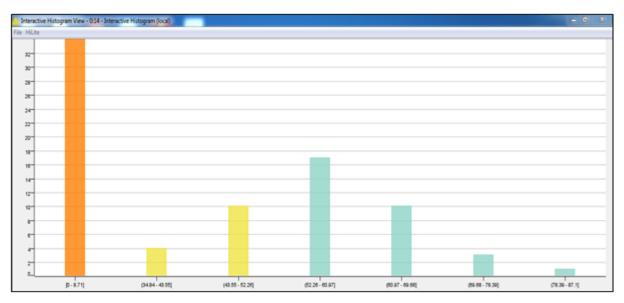


Figure 21: Interactive Histogram View of Fuzzy C-Means clustering.

diates (2)	ATENT MILLOR, FLEND	I IN ADDRESS TATAONER & TANK	Conception in the local division of the loca	1 In column	D HEADADHE	I the second second	D CLANDER	Life I as many themes.	A HOMEN	2 n Chatter
Randi .	ALF-1008		and the second s		100	Life, search and the	M. Contraction	1981	postine(+)	chater_3
diam'r 1	31.071009	H0.E		37.3	49.2	148	5.2		position(x)	thater 2
Barro II	347 A3 83	74.2	14	3	194	1.0.0	20.35	15.46	positive(+)	phater 2
Raw 3	BATTERS.	3	6.3	2.8	36.4		3LX	36.3	progestion (-).	on advant _B
8.am 9	AVALUE:		(A.M.	4.8	36.8	12.4	8.4	1	preparties (.)	invetor_8
Ram 5 Ramb	AMP APER	21	<u>.</u>	2	10	127.8	2.3	1.0	\$00000-00(+)	C. tela.O.
Ran /	14/7 12/2 14/7 12/2	art.s	c	10.000	31.43	1.27	0.04	100	product a)	charlest P
R.andl	AUTLINE		G	G	14.0	10.4	3.1	5.9	post-re(+)	phalm _1
6.and 8	AUT 1398	3.74	19		31.55	18.49	(A. 100)	0.18	megative (.)	chahat 3
A	3.4/T 1.30-4	Get.	16	43	540	1.2	3	14	promition (+)	Shahar 2
Barn 3.5	ALFILDOR	8-7.4	4	54	144	1.7.8	24.18	5.6	providence (+)	sikader 3
4 mm 1.2	3473324			4.8	399	28.9	2	-8.	positive(+)	champs_2
#.me 6.2	31.0-120-0	A. IN	16.3	24.4	36.4	3.5	0.4	0.18	preparation (D)	phater_A
Barr 18	koft social	347	<u>5.</u>		19.8		22	6.8	posting(s)	chahar 3
Barry Like	ALPORTS	C		4.5	23			15	pregature (3	chatter_S
Sec. 1.7	harrow 27	5	1.3	2.2	81	10.0	2.1	10 M	Pergettine CO	Charles A
Rosen Lill	SLP L MON	87.1		5 M	144	1.7	5.1	2.5	provident (n)	in along P
Barr 2.0	SLUTT LOOP	26.93	8.24	S	14.8		5	4.44	magazine (-)	amanter 4
B.ma 20	AAPT LODIN	24.8	18		0.00	118.3	0.89	14	accest-10 +1	Chahar 3
A.m. 24	34491.302	14	54.4	1.2	34.3	2.0	50. W	34.	Inequalities (C)	in.when at
#.am 32	ALT LODG	24.5	24	* 50	30	1.9.3	21.49	1.a		phates 2
Kan 25	AAPT NO.	2	A.Y.	4.2	3(5)	2.9	1.2	18	pergettine ()	stanter 3
B.am 24	3.47 A 16 B	29.4	G	1.00	30	19.3	1.69	8.8	(positive(e)	in any J
April 28	NUTLICE NUTLICE	333.5	2.2	1.3	1.1	2.9	1	2.0	inegative (i)	chaher 2
Barry 27	ALL' LEP	41.5	C	40.8	10.2	148	CC	Q.4.	posting +1	
8.mm 28	31.071.000	34.3	G	14	24	10.4	1.1	5.e	postica(*)	phater_3
8.am 22	1.0F1282	83.5	S *	1.07	Cash.	(1. * *	3. 3	5.5	post-st-1	chahet 3
Rosen 242	3471444	3	1.1	13	313	1.1	0.4	30.3	Inequalities (1.9)	Disabet 8
Barra 3/3	ALT'S SAL		4.8	1.0		11.4		30.8	pegame (.)	on.otor 8
#am 52	AAFTARES.	145.5		49.2	48.5	Ca.4	1.8	le .	promitively all	Challer 3
#.izes.30	3447.1.1.4	34.3	3.4	4.67	145	1.7.8	2.2	3.4	provinces of	in.eter_2
B.mm 34	SAFTAX88	2	1.5	1.2	83	0.5	0.4	20.5	pragadition (c)	physiter_8
Ran 26	31071110	2		1.4	84	22		10.38	pagettine CO	10.400 _ 3
Range Str.	ANP LEAP	01 A	2.4	1.27	144	10.0		5.5	pagante ()	Chahar_1
B.m. 100	AUP 1901	34.5	94	4.87	45	67.8	22	1.4	postine (+)	in.etci_2
8.00.39	Ender Tube	0	<u></u>	12 m		0	8	14	preparation (1)	in-anat_4
8.mm #3	ALF LOOP	84.4	G. 4	1.57	745	1.7.8	5.3	16.4	permitted (a)	phalm 3
B.ma.W5	ALTING .		1.1	4.8	8.5		8.4	8.3	pagestin ()	Chatter 1
8.000 W	317*1868	241	1.2	4.0.	540	CAR .	9	-	permitted (+)	04.what _1
Row-40	3.4/T \$306	>	16.3	1.1	34.5		30.4	0.26	Inequative CD	pitusher_s
8.com/04	AAT SHOT	143.8	*	47.8	49.3	1.4	2.4	34	providence (+).	20.000 J
Barr 10	34573339	140	P	19.	240	19.8	2	17	postive(+)	(charlest _)
B.com (40)	ALAT & STREET	0.21	34.35	1.17	52.25	1.34	0.82	1.34	pegakue (c)	50.a0401_3
B.mar 47	LOTLING	25.54 (0.25	8.47	8.87	23	10.48	1.24	11.4.8	pegabox 4.2	shater_1
B	3471100	10.02	A. A.F	44		142	(2.4%)	person .	pagation (C)	ich.athir_3
8.0+ 30	AAP A KIN	100	G	15	30	1.2	6	C	postical(x)	Date: 3
Barris	3171100	34.3	1.45	3.70	50	18.5	5.1	14	post-rivi	in other of
Acres 5.2	34714539	14s	16	Calle.	44	CAR -	a	4	promition (+)	phater_3
B.July 53	AAF1.004	0.9	8.2	0.4	4	0.4	5.95	2.4	progettine (1)	ithation a
.au 8.	3171158	24.0-4	34.	14	2.4	1.0.0	(A. B.	16.19	(powdrost +)	ckahat_1
Barry 198	A.F. 1.195	6.9	23	11	14	0.8	0.95	[2,4	Intergration (1)	Shaher_4
Rear Sec.	Average .	0.8	84	2.4	dia.	1.5.5	2.7	2.1	pergettine (.)	privation _ 2
Aprel 10 Aprel 10	hard's also	N# 44 14 3	S.+5	1.16	50	12.2	64	2.4	peasing at	10.000 3
B.am.22	34711118	(a)	6 m		144	10.0	0.85	2.4	position(1)	Charlest 3
Ratuba	ALT LINE	(A)	6	45	56	14	5	14	property of all	chanter 3
A state of	31/11/20	0.11	S	A.47	5.81	0.99	6.44	6.44	personal all	DAME
Band.2	ANT A SITS	0.5	88	8.4	14	0.6	0.95	2.1	magazine CO	shaher_1
B.mat.3	ALP'S STR	Q. 8	14.2	1.10	19	20.45	BL 8-5	2.4	Dependent 10	chades 1
Barrell-B	34,073,623		10.40	18.995	240	18.5		4	position (+)	ph.eter_2
8,0465	3171100	<u>51</u>	58.45	3.70	245	14A.S	9.8	24-	providence (+)	a second
*.aute	34.0" 4 128	748.0	26.000	2.00	144	18.8	56.W		(poeth-st(+))	chanter_2
Rand.7	ANT A KINA	Q1.9	8.2	31.4	14	0.4	0.95	84	Inegalive (O	phater_4
Read of Lot of L	AUTIER	See .	2.4	18 du	3.4	12.0	2.24	36.34	paget a Co.	phater_X
Auro 70	3.07 5 528 3.07 1 629	22.2		1.87	14	10.1	5.2+	2.0	permitten (+)	in and
B.m. 73	2071153	34.7	5.45	2.49	100	18.1	2.2	- C	postication (shahes J
Barn 712	SALT'S LOSS	91	9.77	6	44	117.8	0.99	Sec.	prostant at	chatter A
A.ora 72	ALT'S MA	0.4	5.4	8.4	14		in an	5.1	regative (.)	phater_1
Billion Pd	AMTA. 1-10	0.21	8 35	a a .	0.54	0.95	5.24	10.4.2	peganter (-)	chaher_3
8.0m 7%	313713 242	34.3	32.45	1.52	540	148.8	34.2	-	posting ()	ich.adat _2
B.org 76	ALTLOND.	24.1	18	D8	114	12.4.8	2.3	5.6	promitteed with	phater_3
Barn 77	AAPT L DWG	345	18	145	540	1.2	2	16	second etc.	and and and a state of the
Water Pd	BAPT KOO'S	(0.*	9.8	8.4		10.46	(m. mm)	54.8	Insighting (C)	internet_1

3.5 Interactive Table and Histogram K-Means Clustering.

Figure 22: Interactive Table generated by K-Means clustering on all attributes.

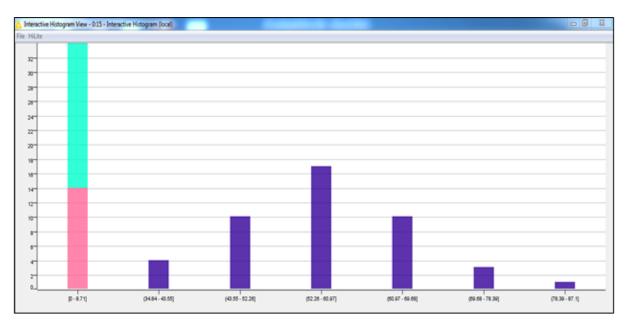


Figure 23: Interactive Histogram View of K-Means clustering.

3.6 Interactive Table and Histogram Hierarchical Clustering.

Aprel 8	B. PATIENT PRESIDENT PRES			D contra	Di rentative		CO COMPANY AL			(b) Chatte
	Sout salar.	266.2	3.48	3.79	148	4.0.0	1.3	14	presidences(+)	internet all
Ran 18	BOW INPO	14.41.2	10.2.4	3	14.3	24	A	4. 10	presentation (-1.	tabative _1
8.0m 78	Bud alles	ja l	10	4.	30	349-			Presignative 5-2	sturier_1
Lan 53	Bur st.m.	16.N	0.3	0.4		30.4	34.83		Pergedice 11	shotter 1
Line 58	BUT 11.36	10.00	10.7	11.4	14	10-A	H #1.	18. W	Prepartive 5.5	infamilier_1
Lon M.	BOAT KRAEP	14.16	0.2	0.4		30.4	in mit.	10. N	Presson 11	pharmer 1
Larve J	LOW HERE .	W.W	10.2	10.4		30.4		10. B	require 1.1	blester_1
Land 3	Buy state			10.4	16	30.4	0.05	20. Y	megative 0.8	infuntion_1
Carroll Y	\$107 \$1.04	2.4	0.3	10.4	14	Mid.	30.85	Contract of A	preparities [1]	blooker_1
1.mm 775	Buf 1145	14.10	10.17	0.4	14	30.4	4 4 1	A 4	Pargadive 5.8	charter_1
Lare 20	South Balling	10.10	0.2	0.4		30-4	14.83	9.0	Parative 21	shorter_k
Local Real	BUT 1140	10.25	30.35	10.2.7	8.94	1.04	8.87	1.00	require 1.1	charter 1
Law 42	Seaf #1.45	10.51	0.58	10.67		30-58	1.34	14 82	Yequitie [1]	storm, 1
100.00	BUT 11.48		8.40	45.8.7	2.51 3.31	51.06	4.84	14. A.S.	presenting 2.1	Johnston _1
Laura 5	\$UT 1100	8-21	10.58	0.57	10 M 1	30.58	3.84	34.817	Pergedres 1.4	charler_6
Lan 74	ALT ILM.		0.26	0.13	2.31	50.05	4.34	8.82	Presidence 11	Scheroline _ 1
Law 1.7	L.T8037		1.8	0.8	10.1	1.5	8.4	0.7	Pagettys 5.8	showing 1
Lan 34	how the set	C	4.9	1.2	0.3	4.8	8.4	23	Paragettine 2.8	sheeting a
Lan A	8.u# #0.82	C			8.x	12.4	83			
turn 14	8×78633	C	1.2	1.8	63 ·····	1.4	8.4	0	PROPERTY 2.5	character _ A
lan 2		C	1. C		10.05	12.0	8.49	2	Yequitie 13	identities _1
	\$10×120×	1 m				4.17		0	reporter 11	t
Lain P	FPA.1300	3.98			10.35	16.89	10 MG	14.275	require 18	sharing 1
Law 2K	And Walk and	1	4.3		0.3	14.15		31.24	Properties 5.2	iduitar_1
1.00.43	but able	(C)	8.18		26.8	4.6	8.4	14.28	presention \$1	johanter_A
1an 38	BUT BLEP	P	5.P.		10.9	11.5	8.4	8.3	Prepetive 23	infantine _1
(mm+1)	\$10 P 200 P	14. · · · · · · · · · · · · · · · · · · ·	4.2	8.2	40.2	4.3	9.4	0.2	THUS TO A LA	information _ 1
an 30	BUT 19.88	2 ·	1.5		20.3	11.75	3.4	14.3	Pergettine 2.8	schuttler _ 1
lais 34	BUTSINS.	18	43		0.3	14.5	36.4	1.2	intergration 2.8	in the second se
Land .	haff sheet	[a.	4.8	(J), %	20.6	4.6	0.4	(A. A.	President to a	johumier_1
1,000 b.3	Bull sides .	4.85	4.3		0.1	4.5	0.4	0.23	inequalities 2.8	infantion _ 1
and the	814F 112P		4.8	6.2	50.8	4.9	1.000	14.34	interpolitive (1.8	Schumiter 1
an 13	Auf#904	34	B. B.	6.P.	0.5	12.9	10.00	10	Pergelitre 59	School are 1
um 2's	BUT (187		11	6.2	30.5	12.0	N. M	2	Pergeduce 2.5	Adverter 1
Law 23	haw same	14	201	1.2	-0.5	12.4		14	Prepartice 2.8	Marrier 1
law 26	8-17 1 1 106		2.1	1.2	0.8	13.0		5	Pergentue 21	Sharelan 1
ALMAN LA	8-1-Y 12/2+			1.4	24	144	G	-	geometrical)	Johnster 2
Lam 27	Bull \$1.40	45	5	1.5	25	4.8	G .	-	generatives, +2	Sharter 1
Land	8.UT \$250.	-8	2	6.8	28	4.2	G	-	anative(s)	Identity of
144 15	8.47 430+	43	3	1.9	24	1.7	G .		anosticat(s)	Mutter 2
Lan W	LUT SHES	at .	2	1.5	28	4.2	G	G	anative at	Charles 1
Lane 40	hulf (1.39	44	-	1.8	24	4.8	G	G	growth (w) all	Marrier 2
1 mm 100	Bull 1120			1.5	24	13	G	G	anative(+)	Internet 3
1.min 50	Barr In Ma	45	C	1.5	100	12	G	C		states _1
Lan b2	Bur 21.54		2	1.9	29		ő	C	(instruct)	shatter 3
auto	BUT ICTS	48	C	1.5	28	12	C	C	3000W-00(+)	MANT _1
Lan 32	BUT MAR	43.8	2	13 2	15.8		6 v	č	30000E140(#)	hater_1
	ENTICES.		<u> </u>			24	13	c	30000-0C+	states_1
1.000 WE	1968 Y	45.5	C	0.5	18.2	85	1.2	C	(energine(+)	Matter 3
Lain b.	LUT AND			12.5	18.2	40	1.8	ē	34444.4(*)	1.1.1.1.m
44.27	BUT ISM	10. B	Č	4.4.8	18.2	144	1.*		3000/01(A)	JAMAN_L
(pro 5.4	8-s-Y \$292	1414 CT	-	2	14	17.9	2.3	3.9	(00000/re(.e)	shatter_2
LANK BOX	BUTCH.	199	d		3.4	17.9	0.30	2.1		hilder _ 2
um 72	BUTLER	THE .	8	2	1.4	17.9	0.95	3.9	(martine(a)	Marter_1
iain.24	BUTTER.	19.4	P	9.96	201	18.3	18.649	8.5	30000-00(+)	photos 1
ate 20	3071071	A. T.	*		100	19.3	4.90	2	providence)	hAnter_1
.a= 312	BOWY AND D	10.0.0	P	1.56	29	49-3	3.99	0.1	(contract)	jahastar "J
an 79	8.17.1129	10.0	4	8	1.0	4-X-x	4.24	9	3440-0L42	34.44 gr _ 1
are 40	bull \$200.5	80.4	2.9	1.54	1.0	127-3		5.9	300409-ref(#)	shorter 1
44.29	B-117 54 30	183.4	2.9	1.57	1.0	47.9	(A. A.	5.4	yanati-4(4)	advantation of
100 KF	Bud about	are, 4	1.4	8.8.0	1.0	(4.3).44	0.9	4.9	promitivat at	hadar 3
ine B	Bud Line	212	4	3	1.4	17.9	3.3	8.9	(0.000 (m) (m)	phatter 1
Later & B	EUVINE.	34.7.4.		2	1.4	1.7.9	34.3	31.97	pendenti +1	shader 1
dand.	\$10 T \$22.5	H 7.1	4	3	1.4	17.0	1111	3.5	post w(w)	phatar "I
Lan & A	BULT LITER.	367.5	4	3	1.4	47.9	34.3	22	annition at	Charles I
are 34	hu Yank?	1000. N	3.4	1.64	4.0	4.5.4	26.18	8.4	print al	phantar_1
am 53	Bull States	294.5	2.4	6.57	1.8	47.9	53	5.4	annetter at	shader 3
Acre. 108	BUW JORI	20.8	3.4	8.8.5	1.4	4.5.9		3.4	provember at a	Autor 2
a-12	kuT stat	00.5	2.9	1.37	1.0	17.9	**	8.4	post and all	shatter 2
and life	LOT 11.PT	No. Cre	6	1	1.4	147.9	2.2	84	panet-st. +1	adapter 7
dan bi	LUT 11.50	766.04			1.4	145.0	2.2	A.+	parameter and and	Shatar 1
an M	Low 11-m	34-2	G		14	47.9	66	S.4	particular and	Shaday 3
and a		Int. 2	G	5	14	10.0	0.0	S.		
	8-14 AD 30	34.7	C		14	17.9	0.0	ST.	posterer al	hater_1
10-128	haven	100. J		5	1.0	43.9	2.2 2.2	07	(ALCOND.) ***	phater_1
Laine B	8107 2013		5					2	30000-00(+)	Autor 2
ton Pi	31074545	56. T	3.46	3.7%	104	44.5	16.9	-	position(*)	phantar_2
an 75	81171147	194.2	2.48	2.79	25	48.8	8.8	ð	34444-14(+)	Advantant _3
Autor Will	\$10Y \$1.52	(Md), 2	2.49	2.79	1010	44.4	(A. 9		post (a)	intuiter_1
and Sill	5071125	(D.)	1.45	1.75	10	45.5	8.9	4	prosition (+)	stater_t

Figure 24: Interactive Table generated by Hierarchical clustering on all attributes.

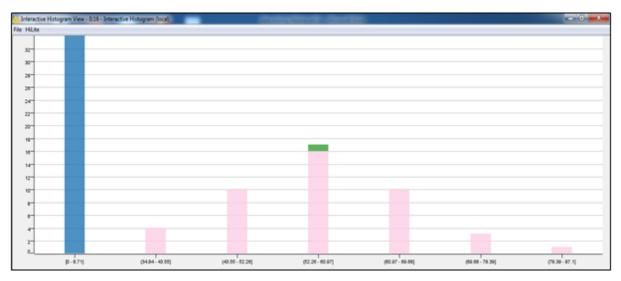


Figure 25: Interactive Histogram generated by Hierarchical clustering.

CONCLUSION

Fuzzy c-means, hierarchical and k-Means clustering approach were implemented to analyze the malaria data (training set). The clustering analysis produced different results respectively. We deducted that the Xie-Beni (XB) index value produced by the fuzzy c-means clustering is 0.27220555023892895. This denotes that the separation between the cluster's distances to the

cluster's centre is a minimal range and compactness between the different clusters is low. The hierarchical clustering on malaria results produced a dendrogram (divisive/top-down tree-like structure), and computed the distances between each clusters. The k-means clustering computed the coverage for cluster_0, cluster_1 and cluster_2 as values of 14, 20 and 45 as shown in figure We could conclude that the cluster_2 with the highest value (45) shows that patients within the attributes ranges tends to have a positive trace of malaria symptoms while cluster_0 and cluster_1 with values of 14 and 20 tends to read negative trace of malaria symptoms. The results can help the medical sectors to predict future occurrences of malaria in most countries.

REFERENCES

- Dr Margaret Chan, "WHO Global Malaria Programme, World Malaria Report", WHO Press, Geneva, Switzerland, 2013; 1-286.
- 2. David Tordrup, Jakod Virenfeldt et al, "Variant Plasmodium ovale isolated from patient infected in Ghana" Malaria Journal BioMed Central, January, 2011; 1-5.
- 3. National Health Service (NHS), "*Causes of Malaria*", Department of Health, Website: https://www.nhs.uk/conditions/malaria/causes/.
- The World Health Report, "Rolling Back Malaria (The Challenge Of Malaria)", 1999; 49-63.
- Al-Radaideh Q.A & Al Nagi E, "Using Data Mining Techniques to build a classification model for predicting employee's performance." International Journal of Advanced Computer Science and Applications, 2012; 3(2): 144-151.
- Alkadhwi Ali Hussein Oleiwi and Adelaja Oluwaseun Adebayo, "Data Mining Application Using Clustering Techniques (K-Means Algorithm) In The Analysis Of Student's Result" ISSN: 2458-925X, May-2019; 5(5): 2587-2593.
- Alkadhwi Ali Hussein Oleiwi, Adelaja Oluwaseun Adebayo and Ali Alkattan Hussein, *"Data Mining using Hierarchical Clustering Techniques on the position of employees in an information technology firm."*. Global Scientific Journals, June-2019; 7(6): 429-437.
- 8. Ming Chaun Hung and Don Lin Yang, "An Efficient Fuzzy C-means Clustering Algorithm", IEEE, 2001; 225-232.
- Kiruthika Devi. S, Krishnapriya.S and Dristipona Kalita, "Prediction of Heart Disease using Data Mining Techniques", Indian Journal of Science and Technology, Vol 9(39), DOI: 10.17485/ijst/2016/v9i39/102078, October 2016; 1-5.
- 10. DavinderKaur, Rajeev Bedi and S.K Gupta, "Review of decision tree data mining

algorithms: ID3 and C4.5", July 2015; 5-8.

- 11. Lior Rokach and Oded Maimon, "Decision Trees", Data mining and Knowledge discovery handbook, 166-192.
- Adelaja Oluwaseun Adebayo and Mani Shanker Chaubey, "Data Mining Classification Techniques on the Analysis of Student's Performance", Global Scientific Journals, Vol 7, Issue 4, ISSN 2320-9186, April-2019; 79-95.
- 13. Open for Innovation KNIME, "*Fuzzy c-Means Clustering*", KNIME AG, Technoparkstrasse 18005 Zurich, Switzerland.
- 14. Yong. Y, Chongxun.Z, and Pan.L, "A novel Fuzzy C-Means Clustering Algorithm for image thresholding", Measurement Science Review, 2004; 4(1).
- 15. Rui. A and Sousa. J. M. C, "Comparison of fuzzy clustering algorithms for Classification", International Symposium on Evolving Fuzzy Systems, 2006; 112-117.
- 16. K-Means clustering algorithm, Website: http://kom.aau.dk/group/04gr742/pdf/ kmeans_worksh eet.pdf.
- 17. Chen Ke, "Machine Learning K-means clustering", University of Manchester, 1-22.