ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

Predicting the Diabetes using Duo Mining Approach

V.V.Jaya Rama Krishnaiah[1], D.V.Chandra Sekhar[2], Dr. K.Ramchand H Rao[3], Dr. R Satya Prasad[4]

Associate Professor, Dept of Computer Science, ASN COLLEGE, Tenali, India[1]

Associate Professor, Dept of Computer Science, TJPS College, Guntur,India[2]

Professor, Department of CSE, ASN college of Engineering, Tenali, India[3]

Associate Professor, Department of CSE, Acharya Nagarjuna University, Guntur, India[4]

ABSTRACT: The amount of information related to biomedical databases is growing so rapidly that the rate at which researchers can convert it into knowledge cannot keep in pace. In this information age it is easy to store large amounts of data electronically, but the proliferations of documents available are structured and unstructured form. We need a tool to extract information and identification of the key concepts from biomedical documents and Bio-medical databases. Duo Mining is a proposed tool, which mines the data from diabetic type bio-medical documents and databases and converts into the knowledge format, and this paper suggests Data mining architecture in addition with Knowledge discovery process.

***Keywords*: Duo mining, Type1 Diabetics, Data Mining, Text mining, Biomedical Literature**

I. INTRODUCTION

Massive amounts of biomedical literature are readily available online and offline to for research in many forms: text abstracts, Medline with more than 16 million biomedical abstracts, full text research articles, databases of blood sugar protein interactions, dictionaries of gene and protein names, and other electronic databases. Huge amounts of valuable knowledge and useful information are embedded in these resources and available to be properly extracted, discovered, and utilized. There is a great need for computational techniques to utilize and extract the useful knowledge from these resources. A number of systems and software tools have been developed to utilize these extensive resources. Biomedical research has shown that duo-Mining can be effective in this field, making text mining and Data Mining increasingly important and necessary for biology and medicine.

A. *Statement Of The Problem*

An incredible wealth of biological information is stored in documents at many hospitals and scientific journals of the state. Summaries of such articles are available in the various databases of different forms . However, retrieving and processing this information is very difficult due to the lack of formal structure and unstructured data in these documents. Automatically extracting information from biomedical text holds the promise of easily consolidating large amounts of biological knowledge in computer-accessible form. A number of recent projects have focused on the manual development of information extraction (IE) systems for extracting information from biomedical literature and documents. Unfortunately, manual engineering of IE systems for particular applications is a tedious and time-consuming process.. Recently, several machine learning methods have been used to develop Medline IE systems

II. METHODOLOGY

Duo-Mining is the variation of data and text mining. It has demonstrated especially well for the Bio Medical Study, in order to take better decisions. As separate capabilities, of the pattern finding technologies of data mining and text mining have been around for years. However, it is only recently that research have been started to use the two in acycle - and have discovered that it is a combination that is worth more than the sum of its parts [1].

The following Figure 1 shows the Duo Mining process

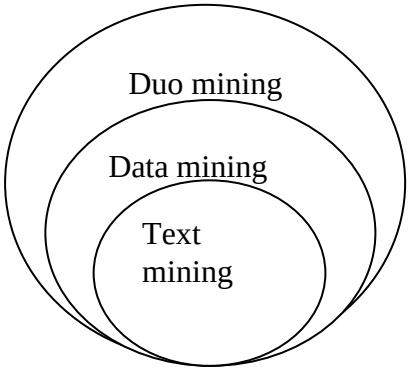


Figure 1: Duo Mining Process

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

Data Mining and Text Mining are similar because they both "mine" large amounts of data, and looking for significant patterns. However, what they evaluate is quite different. Instead of only being able to analyze the structured data they collect from transactions. New developments in text mining technology that go beyond simple searching methods are the key to information discovery which is generally work on the unstructured data.

A. *Data Mining Process*

Data mining is the process of extracting patterns from large data sets with the intent of finding knowledge. Data mining is seen as an increasingly important tool by modern business to transform data into business intelligence giving an informational advantage. Now a days the Data Mining is used everywhere in a collection of data and its availability and accessibility. It is currently used in a wide range of the profiling practices, such as marketing, surveillance, fraud recognition, Bio Medical Analysis and methodical discovery. To prepare data, different methods like Data Integration, Data Reeducation, Data Transformation and Binning were used.

The following Figure 2 shows the Data Mining Mechanism

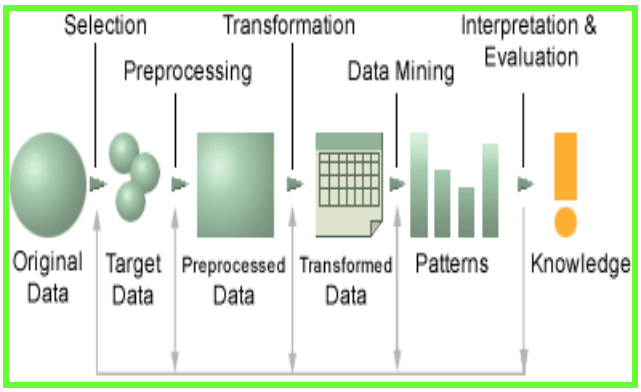


Figure 2: Data Mining Process

Generally Data mining have four classes of responsibilities engage in it are as follow:

Clustering: Is the undertaking of discovering groups and structures in the data that are in some way or another "comparable", without using known structures in the data. For Example, different categories of Diabetic Parents with respect to the BMI Ratio.
Classification: Is the task of generalize known structure to apply to new data. For example, defining the different classes of diabetic patents with respect to Family History, Cholesterol Ratio and BMI
Association rule learning: Searches for associations between variables. Using association rule learning, one can predict feature extractions. For example, how many of the diabetic patents were struggling with Kidney Problems, Retinopathy.
Data visualization: using graphical methods to show patterns in data.

B. *Text Mining*

Text mining is used to extract useful information from text based files, Text mining is also like data mining, as the application of algorithms and methods from the field's machine learning and statistics to texts with the goal of finding useful patterns. It consists of the analysis of multiple text documents by extracting key phrases, concepts etc. and in the preparation of the text processed in that manner for further analysis with numeric data mining techniques(e.g., to determine co-occurrences of concepts, key phrases, names, addresses, product names etc.). To prepare data into the structured fashion, we different methods like Information Extraction, Information Retrieval and Text encoding methods like stemming, text processing, filtering are used.
Basic Steps involved in Text Mining as follows

Document Acquisition Phase: a collection of documents is collected in a repository. Generally, documents, coming from various sources, are of different types. For example Lab Reports.

Document Preprocessing: relevant features of documents, that describe their content, are extracted, and then desired features were created.

Text Mining: documents are treated using a text mining method. In literature several approaches has been proposed

Results Interpretation and Refinement: results of the text mining phase are submitted to performance evaluation using quality indexes [2] and to the interpretation and refinement of a human operator.
The following figure 3 indicates the Text Mining Process.

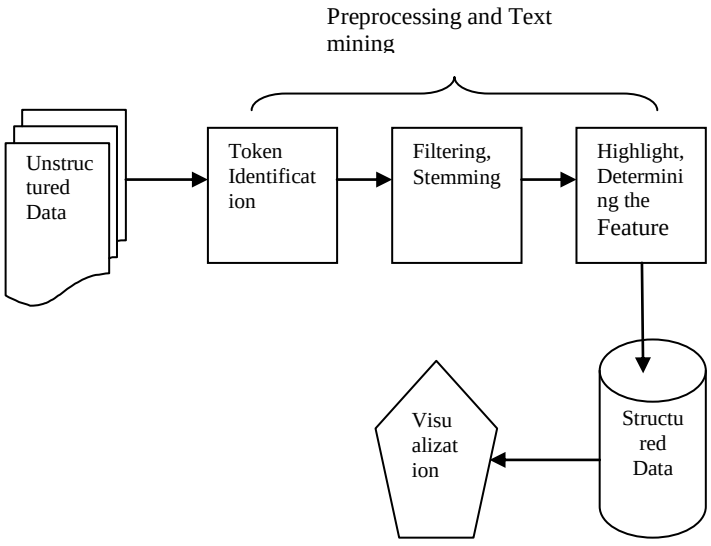


Figure 3: Phases in text mining.

ISSN : 2278 – 1021

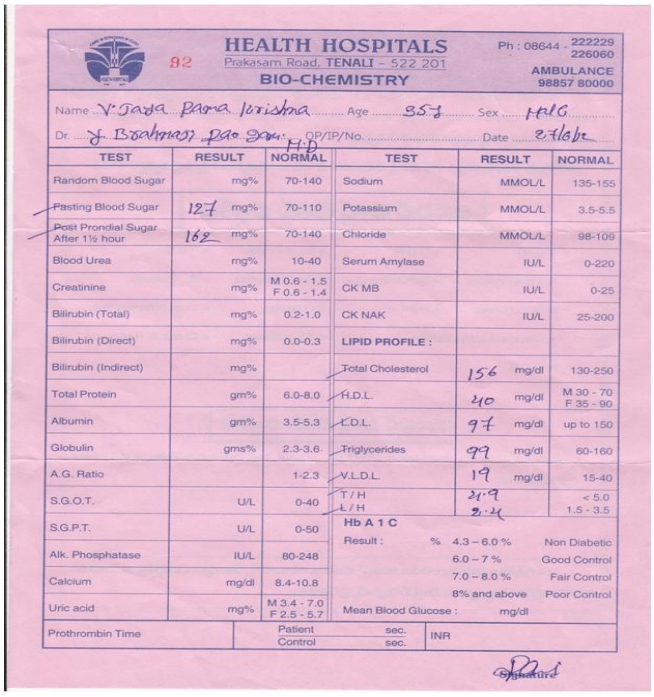
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

III. PROPOSED TOOL: DUO MINING TOOL FOR DIABETIC PATIENTS.

Design and Develop a Duo-Mining Tool for extraction and classification of Diabetic patients information of structured and Unstructured from various sources, developed by using Java Software. The tool developed to extract data from unstructured text then, First, text mining technologies needed for large amounts of text to analyze - several page memos as shown in the figure 1, for example - while call logs are sometimes just snippets in comparison. Second, "stemming," a popular technique in text analysis in which various forms of a word are distilled into one word. Stemming provides better understanding of data and datasets related to Diabetics, Which are then stored in the structured form.

A. *Steps in Proposed Duo mining Tools*

1) Browsing the Bio Medical Literature (Text, Symbols and etc): Extracts the text or flat files from the defined source, like figure 4



HEALTH HOSPITALS
Prakasam Road, TENALI - 522 201
Ph : 08644 - 222229, 226060
AMBULANCE 98857 80000
BIO-CHEMISTRY

Name V. Jaya Rama Krishna Age 35 Sex M/F
Dr. J. Brahmaji Rao Garu OP/IP/No. Date 27/6/12

TEST	RESULT	NORMAL	TEST	RESULT	NORMAL
Random Blood Sugar	mg%	70-140	Sodium	MMOL/L	135-155
Fasting Blood Sugar	127 mg%	70-110	Potassium	MMOL/L	3.5-5.5
Post Prondial Sugar After 1½ hour	162 mg%	70-140	Chloride	MMOL/L	98-109
Blood Urea	mg%	10-40	Serum Amylase	IU/L	0-220
Creatinine	mg%	M 0.6 - 1.5 F 0.6 - 1.4	CK MB	IU/L	0-25
Bilirubin (Total)	mg%	0.2-1.0	CK NAK	IU/L	25-200
Bilirubin (Direct)	mg%	0.0-0.3	LIPID PROFILE :		
Bilirubin (Indirect)	mg%		Total Cholesterol	156 mg/dl	130-250
Total Protein	gm%	6.0-8.0	H.D.L.	40 mg/dl	M 30 - 70 F 35 - 90
Albumin	gm%	3.5-5.3	L.D.L.	97 mg/dl	up to 150
Globulin	gms%	2.3-3.6	Triglycerides	99 mg/dl	60-160
A.G. Ratio		1-2.3	V.L.D.L.	19 mg/dl	15-40
S.G.O.T.	U/L	0-40	T / H	21.9	< 5.0
S.G.P.T.	U/L	0-50	L / H	2.4	1.5 - 3.5
Alk. Phosphatase	IU/L	80-248	Hb A 1 C Result : %	4.3 – 6.0 %	Non Diabetic
Calcium	mg/dl	8.4-10.8		6.0 – 7 %	Good Control
Uric acid	mg%	M 3.4 - 7.0 F 2.5 - 5.7		7.0 – 8.0 %	Fair Control
				8% and above	Poor Control
			Mean Blood Glucose :	mg/dl	
Prothrombin Time	Patient	sec.	INR		
	Control	sec.			

Figure 4: Text Report

2) *Preprocessing and Visualization:* By using Stemming and filtering methods, the feature is extracted and text classifications were determined. For Stemming process in Duo Mining Process, Matching mechanism was used. The following Algorithm is used for determining the starting and ending point of each pattern.

Do for each word in the text
For each match found = True

i. Determine the Start index of the Match
ii. Determine the End index of the Match
iii. Highlight the text with the color
iv. Count the number of words
v. Aggregate them into a Group

Loop.

In Duo Mining Tool, we determined the different matches like BMI, Cholesterol, and Glucose in the Medical Diagnosis report. Different color shades were used to highlight the text in the Annotation Box

3) Knowledge Preparation: It is Important Step in the Duo Mining, with this phase the extracted information with respect to the features like BMI, Cholesterol along with Patent Name, weight and other features were converted into the structured format. For this tool, we use the MS-Access Database for storing the data.

4) Visualization: Different data mining techniques were used for visualizing the data like clustering, classification or decision trees.

The following Figure 5 shows the Duo Mining tool for making the Diabetic patent information.

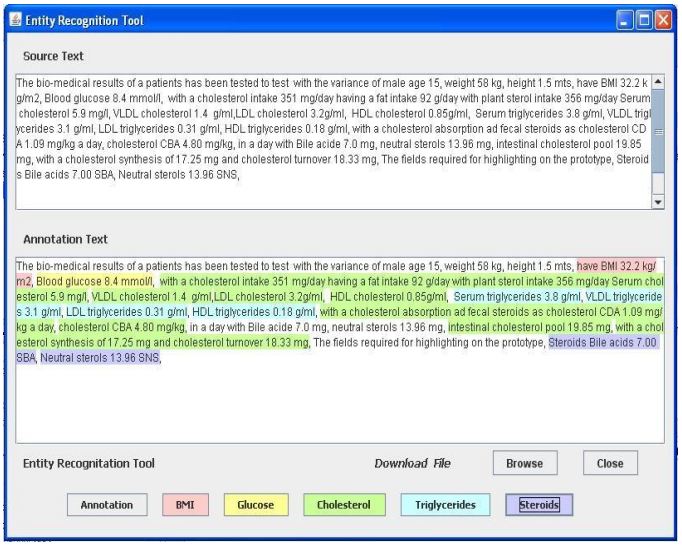


Figure 5: Sample Window, which reads the features of a Patent for Duo Mining

The Predication of Diabetics is done by evaluating the following algorithms

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

i. Calculating BMI value
ii. Calculating Fat Test
iii. HBA1 C Test Calculation

- *Algorithm for BMI value*

```
bmi = (pounds / ((feet/100)*(feet/100)))
TotalInches=eval(feet*12)+eval(inches)
Meters=TotalInches/39.36
Kilos=pounds/2.2
Square=Meters*Meters
Bmi=Kilos/Square
if (bmi < 18.5)
bmitype = "Underweight"

if ((bmi >=18.50) && (bmi < 25))
bmitype = "Normal"

if ((bmi >=25) && (bmi < 30))
 bmitype = "Overweight"

if (bmi >= 30 && bmi<35)
bmitype = "Obese Class I"

if (bmi >= 35 && bmi<40)
bmitype = "Obese Class II"

if (bmi >= 40)
bmi type =”Obese Class III”
```

- *Algorithm for Calculating Fat Test*

LDL Cholesterol = Total Cholesterol - HDL - (TG / 5)

If LDL <=170 “normal”
If LDL > 170 “abnormal” leading to
By checking the LDL values for the given biomedical datasets , The prediction of diabetics is done.

- *Algorithm for Calculating HBA1C*

a = value of fast blood sugar

b = a+46.7

c = 28.7

d = b/c*100;

A1C = Math.floor(d)/100;

if a1c is in btwn 4.3 && 6.0% no diabetic

if a1c is in btwn 6.0% and 7% diabetic and good control

if a1c is in btwn 7.0% and 8.0% high stage of diabetic

>8% very high diabetic

The following Table1 gives the partial view of the structured database

Diabetic Patent Information							
Patient code	**Age**	**Sex**	**Cholesterol**	**HDL-Chol**	**VLDL**	**LDL**	**BMI**
1	31	M	152	37	34	81	32.2
2	31	M	172	45	28	99	33.2
3	31	F	190	42	30	118	30.89
4	31	M	265	36	29	200	34.9
5	31	F	185	42	32	111	30.9
6	32	F	162	46	32	84	33.9
7	32	M	187	42	37	108	33.95
8	32	M	192	45	38	109	34.9
9	32	M	159	41	34	84	30.8
10	32	M	174	45	31	98	29.99
11	32	M	220	40	26	154	30.8
12	33	M	269	36	30	203	30.3
13	33	M	197	38	37	122	29.34
14	33	M	156	39	30	87	22.1
15	34	M	306	39	30	237	20.89
16	34	F	202	35	27	140	33.09

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

Diabetic Patent Information							
Patient code	**Age**	**Sex**	**Cholesterol**	**HDL-Chol**	**VLDL**	**LDL**	**BMI**
17	34	F	178	32	27	119	32.22
18	34	M	286	40	26	220	30.78
19	34	F	184	34	30	120	30.65
20	34	M	192	36	31	125	30.33
21	34	M	210	40	28	142	30.65
22	34	M	242	41	29	172	30.65
23	35	M	198	46	34	118	31.44
24	35	F	156	48	34	74	32.89
25	35	M	205	37	36	138	30.78
26	32	M	165	39	26	100	22.1
27	34	M	158	38	28	92	21.9
28	36	F	162	42	26	94	24.9
29	37	M	178	41	22	115	23.09
30	39	M	160	40	26	94	29

Table 1: List of Patient Information for the Age Group 31-40

The following Figure 6 describes about the BMI Charts for the age group 31-40 years

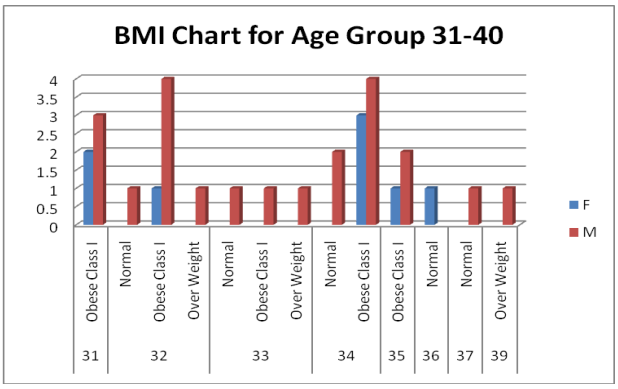


Figure 6: BMI Chart for Age Group 31-40

The following Table 2: Illustrates about the BMI Index and HBA1C Index values. HBA1C indicates about the warnings on feature diabetics and BMI index specifies about the category of Obesity of a patent.

DIABETIC INFORMATION 31-40 YEARS							
Patient code	**Age**	**Sex**	**Blood Glucose**	**BMI Index**	**BMI**	**HBA1C Value**	**HBA1C INDEX**
1	31	M	187	Obese Class I	32.2	8.14285714285714	Very High Diabetic
2	31	M	218	Obese Class I	33.2	9.22299651567944	Very High Diabetic
3	31	F	192	Obese Class I	30.89	8.31707317073171	Very High Diabetic
4	31	M	215	Obese Class I	34.9	9.1184668989547	Very High Diabetic
5	31	F	235	Obese Class I	30.9	9.81533101045296	Very High Diabetic
6	32	F	217	Obese Class I	33.9	9.18815331010453	Very High Diabetic
7	32	M	244	Obese Class I	33.95	10.1289198606272	Very High Diabetic

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

DIABETIC INFORMATION 31-40 YEARS							
Patient code	Age	Sex	Blood Glucose	BMI Index	BMI	HBA1C Value	HBA1C INDEX
8	32	M	184	Obese Class I	34.9	8.0383275261324	Very High Diabetic
9	32	M	209	Obese Class I	30.8	8.90940766550523	Very High Diabetic
10	32	M	184	Over Weight	29.99	8.0383275261324	Very High Diabetic
11	32	M	179	Obese Class I	30.8	7.86411149825784	Very High Diabetic
12	33	M	197	Obese Class I	30.3	8.49128919860627	Very High Diabetic
13	33	M	201	Over Weight	29.34	8.63066202090592	Very High Diabetic
14	33	M	245	Normal	22.1	10.1637630662021	Very High Diabetic
15	34	M	245	Normal	20.89	10.1637630662021	Very High Diabetic
16	34	F	217	Obese Class I	33.09	9.18815331010453	Very High Diabetic
17	34	F	185	Obese Class I	32.22	8.07317073170732	Very High Diabetic
18	34	M	198	Obese Class I	30.78	8.52613240418119	Very High Diabetic
19	34	F	181	Obese Class I	30.65	7.93379790940766	Very High Diabetic
20	34	M	152	Obese Class I	30.33	6.92334494773519	High Stage of Diabetic
21	34	M	180	Obese Class I	30.65	7.89895470383275	Very High Diabetic
22	34	M	192	Obese Class I	30.65	8.31707317073171	Very High Diabetic
23	35	M	205	Obese Class I	31.44	8.77003484320558	Very High Diabetic
24	35	F	218	Obese Class I	32.89	9.22299651567944	Very High Diabetic
25	35	M	211	Obese Class I	30.78	8.97909407665505	Very High Diabetic
26	32	M	203	Normal	22.1	8.70034843205575	Very High Diabetic
27	34	M	85	Normal	21.9	4.58885017421603	No Diabetic
28	36	F	99	Normal	24.9	5.07665505226481	No Diabetic
29	37	M	89	Normal	23.09	4.72822299651568	No Diabetic
30	39	M	90	Over Weight	29	4.76306620209059	No Diabetic

Table 2: BMI and HBA1C Index Values

From the table 2 and Table 1, we can compare the values for BMI and HBA1C index values for all patents with respect to the age groups as shown in the figure 7

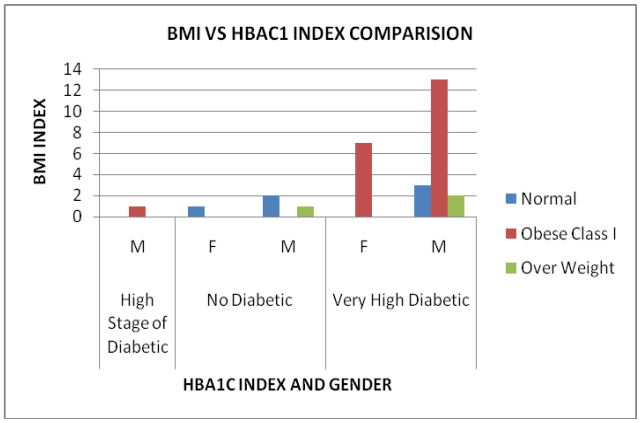


Figure 7: BMI and HBA1C INDEX COMPARISION FOR 31-40 Years Age Group

The following Table 3 indicates about prediction table, discovers the risk of the diabetes with reference to the HBA1C, Bad Cholesterol (LDH) and Body Mass Index (BMI).

BMI	LDH	HBA1C	PREDICTION
0	0	0	NORMAL
0	0	1	NORMAL
0	1	0	SYMPTOMS OF DIABETICS
0	1	1	SYMPTOMS OF DIABETIC
1	0	0	DIABETICS
1	0	1	HIGH
1	1	0	VERY HIGH
1	1	1	VERY HIGH

Table 3: Predication Table for Diabetes

IV.RESULTS

To evaluate the Type-1 Diabetes, we used the Methodology called Entropy based mean clustering (EBM Clustering) [10] proposed by V.V.Jaya Ramakrishniah et.al, for making the clusters. EBM approach is enhanced method for

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

traditional K-mean Clustering, it improves the efficiency of the clustering process by eliminating the empty clusters and highest significance in visibility of data and avoids the missing values, because of this feature, and we can improve the significance of the data.

The following figures 8 to 10 visualize the number of diabetes in the age group 31-40 years. The Diabetic patents are indicated by the red circles and healthy people were indicated with green colored triangles.

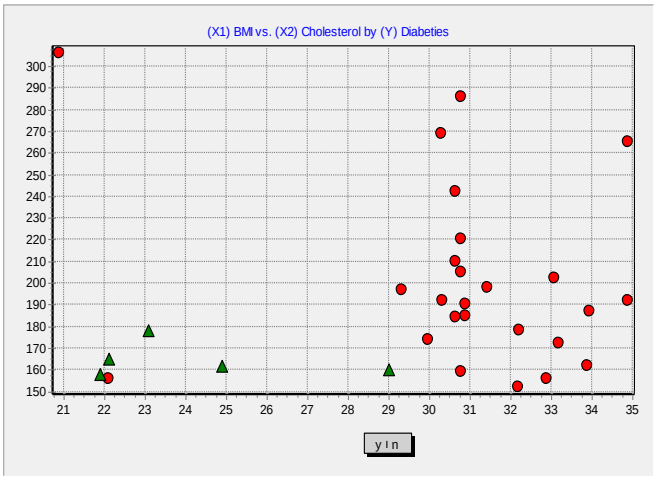


Figure 8: Number of Diabetes with respect to BMI and Cholesterol

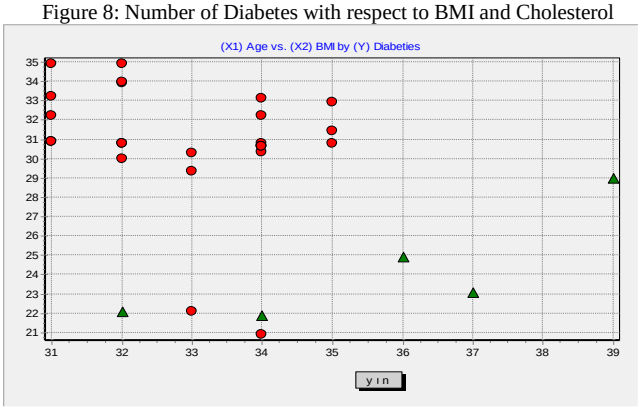


Figure 9: Number of Diabetes with respect to age with respect to BMI

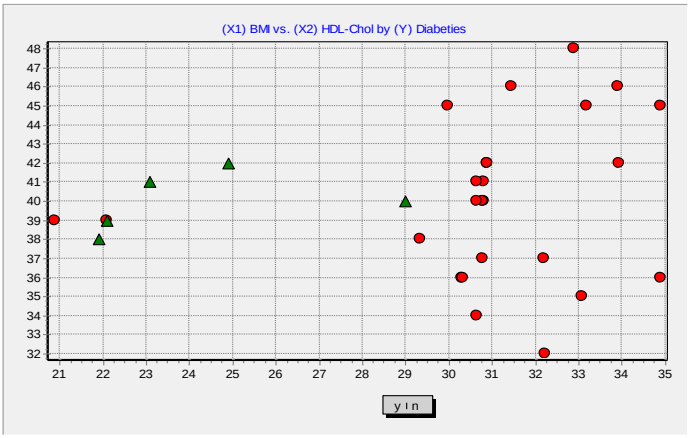


Figure 10: Number of Diabetes with respect to BMI and HDL-Cholesterol

The following Figure 11 indicates the number of male and females with their age groups were affected by the diabetes. Effected people were indicated with Red Bar, and healthy people indicated with Green Bar.

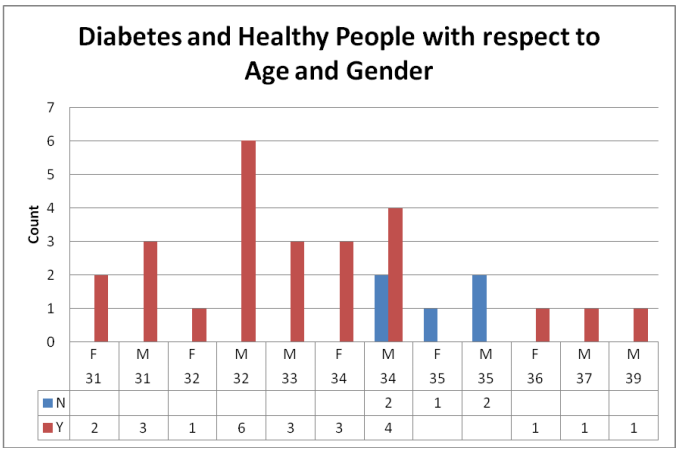


Figure 8: Chart for Number of people effeted with diabetes for the age group 31-40

Total of 130 patients were registered. The mean age was 34.5 yr+/-4.7, mean age at diabetes diagnosis 38.4 yr+/-4.4, and diabetes duration ,Obese group: 34 M/37 F, BMI = 49+/-7 kg/m2 (mean+/-s.d.) age = 36+/-2 y, (26% diabetics, 59% females, mean age +/- SD = 37 +/- 17 years, mean BMI +/- SD Percent body fat, waist circumference, blood pressure, blood glucose, insulin, to make predictions based on data available in the early postoperative period. Values of HbA1c <7.5% were achieved in 12.5%, > or =7.5 and < 8% in 11.3%, > or =8 and <9.5 in 33.5% and > or =9.5 in 40.9% for basic prediction of diabetics and other disease.

V.CONCLUSION

As the amount of unstructured data in our world continues to increase, text mining tools that allow us to sift through this information with ease will become more and more valuable. Text mining tools are beginning to be readily applied in the biomedical field, where the volume of information on a particular topic makes it impossible for a researcher to cover all the material, much less explore related texts. The combination of data and text mining is referred to as "duo-mining" [13]. SAS and SPSS have begun recommending duo-mining to their customers as a way of giving them the edge on consolidated information for better decision making. This process combination has proven to be especially useful to banking and credit card companies. Instead of only being able to analyze the structured data they collect from transactions, they can add call logs from customer services and further analyze customers and spending patterns from the text mining side. These new developments in text mining technology that go beyond simple searching methods are the key to information discovery and have a promising outlook for application in all areas of work.

REFERENCES

[1] Schwartz, A. S. And Hearst, M. A. (2003), 'A Simple Algorithm For Identifying Abbreviation Definitions In Biomedical Text', In 'Proceedings Of The 8th Pacific Symposium On Biocomputing', 3rd–7th January, Hawaii, Pp. 451–462. 29. M

[2] Chen, H., Lally, A. M., Zhu, B., And Chau, M. (2003). "Helpfulmed: Intelligent Searching For Medical Information Over The Internet," *Journal Of The American Society For Information Science And Technology*, 54(7), 683-694, 2003. This Article Provides An Overview Of Medical Information Retrieval Techniques On The Internet, Including Web Crawling, Co-Occurrence Analysis, And Document Visualization.

[3] Yang, Y. And Liu, X. (1999). "A Re-Examination Of Text Categorization Methods, In *Proceedings Of The 22nd Annual International ACM Conference On Research And Development In Information Retrieval* (SIGIR'99), 1999, Pp. 42-49.

[4] Mining Biomedical Literature Using Information Extraction ,Ronen Feldman, Yizhar Regev, Michal Finkelstein-Landau, Eyal Hurvitz & Boris Kogan Clearforest Corp, USA & Israel

[5] Alexander Pertsemlidis; TEXT MINING THE BIOMEDICAL LITERATURE,UT Southwestern Medical Center, 5323 Harry Hines, Boulevard, Dallas, Texas 75390-8573

[6] Type - 1 Diabetes Mellitus:Indian And Global Scene – Burden & Challenges.Diabetes Department,Voluntary Health Services,Chennai, Tamil Nadu, India.

[7] Locating Previously Unknown Patterns In Data-Mining Results: A Dual Data- And Knowledge-Mining Method, Mir S Siadaty* And William A Knaus, Address: Department Of Public Health Sciences, University Of Virginia School Of Medicine, Box 800717, Charlottesville, Virginia, 22908, USA Email: Mir S Siadaty* - Mirsiadaty@Virginia.Edu; William A Knaus - Wak4b@Virginia.Edu

[8] V.V.Jaya Rama Krishnaiah, Dr.K.Ramchand H Rao, Dr.R.Satya Prasad, 'Entropy Based Mean Clustering: A Enhanced Clustering Approach, The International Journal Of Computer Science & Applications (TIJCSA), Volume 1, No. 3, May 2012 ISSN – 2278-1080

Biography

V.V.Jaya Rama Krishnaiah received Master's degree in Computer Application from Acharya Nagrajuna University,Guntur, India, Master of Philosophy from Vinayaka University, Salem . He is currently working as Associate Professor, in the Department of Computer Science, A.S.N. Degree College, Tenali, which is affiliated to Acharya Nagarjuna University. He has 14 years teaching experience. He is currently pursuing Ph.D., at Department of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. His research area is Clustering in Databases. He has published several papers in National & International Journals.

D.V. Chandra Shekar, received Master of Engineering with Computer Science & Engineering He is currently working as Associate Professor, in the Department of Computer Science, T.J.P.S COLLEGE (P.G COURSES),Guntur, which is affiliated to Acharya Nagarjuna University. He has 14 years teaching experience and 1 years of Industry experience. He has published 52 papers in National & International Journals.

Dr. R. Satya Prasad Received Ph.D. degree in Computer Science in the faculty of Engineering in 2007 from Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. He have a satisfactory consistent academic track of record and received gold medal from Acharya Nagarjuna University for his out standing performance in a first rank in Masters Degree. He is currently working as Associative Professor in the Department of Computer Science &Engineering,

Acharya Nagarjuna University. His current research is focused on Software Engineering, Image Processing & Database Management System. He has published several papers in National & International Journals.

Dr.K Ramchand H Rao received Doctorate from Acharya Nagarjuna University, Master's degree in Technology with Computer Science from Dr. M.G.R University, Chennai, Tamilnadu, India. He is currently working as Professor and Head of the Department, Department of Computer Science and Engineering, A.S.N. Women's Engineering College, Tenali, which is affiliated to JNTU Kakinada. He has 18 years teaching experience and 2 years of Industry experience at Morgan Stanly, USA as Software Analyst. His research area is Software Engineering. He has published several papers in National & International Journals.