ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 9(17), DOI: 10.17485/ijst/2016/v9i17/93024, May 2016* ISSN (Online) : 0974-5645

Analyzing and Performing Privacy Preserving Data Mining on Medical Databases

D. Aruna Kumari*, Y. Vineela, T. Mohan Krishna and B. Sai Kumar

Department of Electronics and Computers, K L University, Guntur – 522502, Andhra Pradesh, India; aruna_d@kluniversity.in, vineelayanduru@gmail.com, mohankrishnatadavarthi1@gmail.com, saikumars2904@gmail.com

Abstract

For both the production and consumption of data the internet is becoming a standard whereas the security for private data is gradually decreasing. Therefore, to have a safe transaction in the data, security and privacy would be the key issues to be considered. In recent days, privacy has become a key issue in many data mining and knowledge discovery fields which lead to the development of many Privacy Preserving Data Mining (PPDM) techniques. In our work we use few of these techniques to privately preserve the data holder such as hospital data. In this we use techniques named "Anonymization", "Suppression", "Generalisation" and "Data Hiding" on different fields for the data to be more secure and project the data which is useful to the public. This is a new way of our approach to create awareness among the public to be more attentive and health conscious. The modified data is clustered based on diseases. Based on the end user requirement the private data of the individual is hidden and the required data is projected.

Keywords: Anonymization, Cluster, Data Hiding, Generalisation, Privacy Preserving Data Mining (PPDM), Suppression

1. Introduction

Security of that database is a paramount issue. The mining applied to this database is referred as "mining" or "extracting" knowledge from the huge amounts of data. Discovering interesting knowledge from huge databases or data warehouses or any other repositories is known as "Data Mining". By applying data mining high level information or interesting knowledge is discovered which can be viewed or used for future implementation. One of the most important frontiers for database Systems is considered as data mining and it is also the interdisciplinary action in the information industry. This extraction of useful information from huge collection of databases is applicable in many real time scenarios such as supermarket database analysis, customer relationship analysis, hospital databases etc. This valuable data discovered by mining the huge collection of database is misused. So the data to be private and to be preserved this would be an issue in the privacy concern. This privacy implies that the individual information to be secured privately without viewing or making it available to others. Once the privacy has lost we cannot prevent it from the misuse. Let us consider an example, if the fields like user id or contact fields are known then the data may be misused.

To solve the problem of privacy there are few methods namely PPDM[1] techniques which deals with the issue of protecting the privacy of the sensitive or the individual data. The aim of this PPDM concept is to mine the information from huge collection of databases at the same time protecting the meaningful information. The problem arises at the point where the result sets of the mined data are to be preserved and are to be kept private. For example, when hospital database is considered, the private data of the patient along with the public fields are maintained by the hospital database which is not to be projected publicly. In this concern we use PPDM techniques to maintain privacy for the individual's private data.

* *Author for correspondence*

Normally the data is distributed[1], and this distributed database scenario is classified as[1]:

- Horizontally partitioned data
- Vertically partitioned data

These can be discussed as both centralised and also the distributed environment where the data will be distributed in different sites as shown in the below Figure 1.

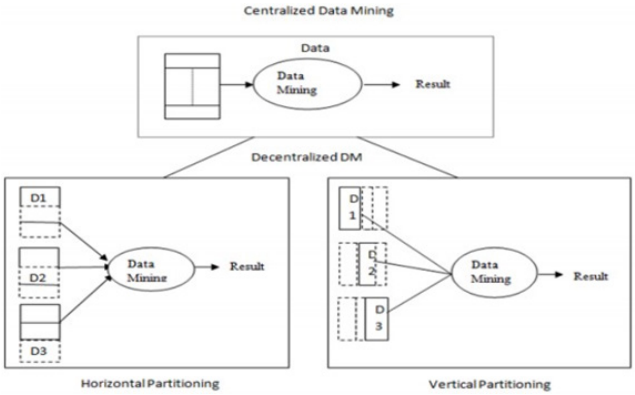


Figure 1. Data distribution.

1.1 Horizontally Partitioned Data

In this approach the database is divided into horizontal partitions. The data from different areas have the records about the same entity or the particular product or person. This info is used for mining the data.

1.2 Vertically Partitioned Data

In this approach the database is divided into vertical partitions. This will have the different attributes with same no. of transactions in the data. This is now extended to the various data mining applications like Naïve Baye's Classifier, k-means Clustering and etc.

2. Framework of PPDM

In Knowledge Discovery from Databases (KDD) or Data Mining processes the data is gathered from a single or multiple institutions and is stored at a Database. Now the collected data is then transformed to a particular structure which suits the analytical approach which is stored in large data warehouses. After the data is transformed the data mining techniques are applied to that data which results in Knowledge or Information. This framework of PPDM[2] consists of three different levels which are stated in Figure 2.

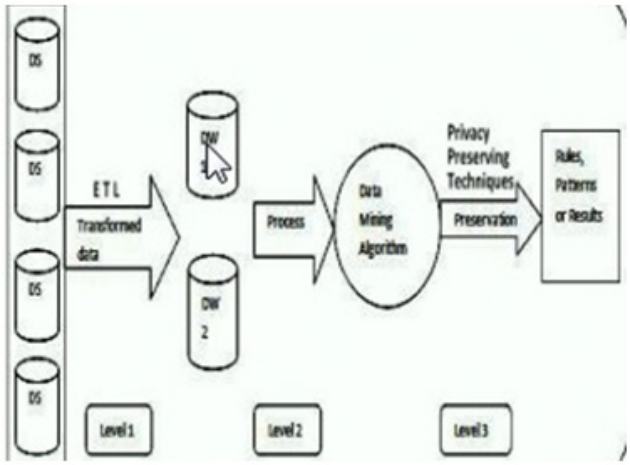


Figure 2. Framework of privacy preserving data mining.

This level 1 consists of databases or raw data where the transactions are present. The second level comprises with the Data Mining techniques which ensure the privacy of data. The third level includes the results of those data mining algorithms and techniques.

2.1 Level 1

This foremost level contains the raw data which is collected from a single or multiple organisations which are to be processed for analytical approach. This will also be the level where we should consider the privacy issue. The main part of this level would be making the raw data to be suited for analytical approach.

2.2 Level 2

Now the data from the warehouses is then applied with different techniques which sterilize the data. This sterilized data cannot be accessed or viewed by any miner. The above stated techniques are: Suppression, Generalisation, Perturbation, Blocking etc. Now the different data mining algorithms are applied to the data which is processed for the knowledge discovery or meaningful information.

2.3 Level 3

The last but not the least level includes the checking process for its sensitivity towards the risks. This information processed in the above level is verified for its sensitivity and then it is used as it can face any disclosure risks. In

this level the mined data undergoes with the privacy preservation techniques which help in protecting the data from the unauthorised users. Now the private data of an individual will be safe and secure and cannot be misused.

3. PPDM Techniques

These PPDM techniques are of two types. These data mining techniques are applied at two different levels. The new approaches obtained will be allowed at the time of mining the data and also at the time of resulting the information or knowledge. These are given below:

- Techniques that preserve the private data in the time of mining process.
- Techniques that preserve the private data mining results which are obtained after applying the data mining techniques.

PPDM techniques can be classified as stated below[2]:

- Data Modification
- Data Distribution
- Data Mining
- Data Hiding
- Privacy Preserving

Our work is under Privacy Preserving Data Mining.

3.1 Data Modification

Data Modification is the technique which is used to modify the original private data of an individual. This technique when applied modifies/changes the original data based on the technique applied. In this Data modification the following are the sub techniques:

- Data Perturbation
- Data Blocking
- Data Sampling
- Data Swapping
- Data Encryption

3.2 Data Distribution

Data Distribution[3,4] is the technique in which the original data is distributed. This distributed data is classified as horizontally portioned data and vertically partitioned data. This is explained earlier. This technique refers to the data which resides in different places.

3.3 Data Mining

Data Mining[5] is the technique in which raw dataset is considered and extract information and arrange it in a meaningful format for the future use. This data mining has the following techniques:

- Association algorithm
- Clustering algorithm
- Decision tree algorithm
- Naive Baye's algorithm
- Time series algorithm
- Linear regression algorithm
- Logistic regression algorithm

3.4 Data Hiding

Data Hiding is also called as Rule Hiding. Data Hiding is the technique in which the original data or the raw data and the grouped data are hidden. This Data Hiding is referred to protect the private data of an individual which includes data fields like name, contact, address, personal id's etc. Whereas rule hiding is referred as the technique in which the confidential information/knowledge is protected. This is called as Data Hiding.

3.5 Privacy Preserving

Privacy preserving[6] techniques are those which are used to preserve the private data of an individual. These Privacy Preserving techniques are again classified based on few types:

- Cryptography based
- Reconstruction based and
- Heuristic based

4. Problem Definition

Biomedical involves the applications of the natural sciences, especially the biological and physiological to clinical medicine. It is a discipline in biological, medicine to improve human health by integrating the medical usages to help the clinical practices. This can be done through the advancement of medical sciences and the database management system which creates huge number of databases in the medical world. Here we are introducing data mining technique through extracting the useful information from the raw bio medical data which helps to discover and manage the large heterogeneous data. This data can be used to make the work of doctors and practitioners slightly less. The advancements in this field is not only limited to this sector but can be implemented into various sectors.

Now a day's security has become the key issue regarding any type of data we here are with a solution to have more secured data. Our work Analysing and Performing Privacy Preserving Data Mining[7] on Medical Databases deals the key issue i.e., having security for the preserved data. The main aim of our work is to secure the private data (hospital) of an individual while projecting it to an end user. This is implemented using different techniques namely Anonymization, Generalization, Suppression and Data Hiding. These are the four different techniques which we implement on the individual fields in the data base so that the original data of the patient is hidden and the required fields such as disease name, medicine used, period of time required to cure, precautions to be taken etc. are projected to the end user. After modifying the original data using PPDM techniques, the modified data is then made into clusters based on disease and the age group using Weka tool. Now the clustered data is used to easily identify the fields based on the end user requirement. The requirements of the end user are considered using a web page.

5. Proposed Technique

Considering single hospital dataset which contains the different fields like patient id, Name, DOB, Age, Gender, Address, Contact, Disease, Prescription, period of treatment, precautions etc. As this data includes the personal fields like address, contact, DOB etc. these private fields are to be secured and are not to be known. As our main aim is to project the data to the public and to create awareness among them about the diseases occurring and to provide precautions they should look about in order to avoid diseases. This data includes the private fields of the individuals, so that they should be privately preserved and only the public data to be projected either to the public or practitioner or the doctors. For this to be done we came up with an idea which is shown in Figure 3 of applying a "Hybrid Technique" to the private data fields and to hide them from the public.

This Hybrid technique is the combination of two or more PPDM techniques. In our work we apply three different techniques namely "Anonymization", "Generalization", "Suppression", and "Data Hiding" on the different fields of the private data so that they will not be identified by anyone.

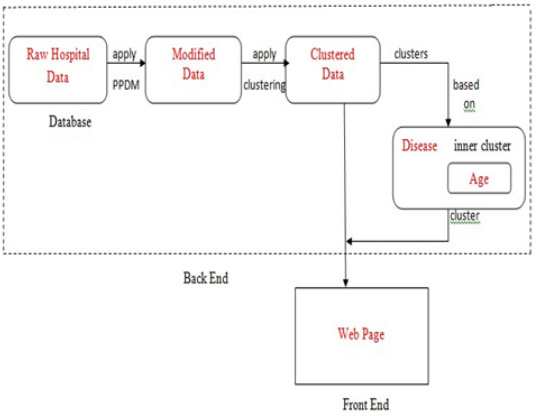


Figure 3. Explanation of proposed work.

After applying the PPDM techniques the modified data is then clustered into groups based on diseases and age groups. The difference between the original data and the modified data is shown. The modified data clusters[8] are done using "Weka" tool. This clustered data is then connected to the front end called Web page.

This Web page consists the information about the Hospital and also the search box in which the end use based on his requirement of the disease can search the details. He can even specify the age group when required. Based on his own requirement the data required would be projected in which the personal details of the patient would be hidden.

6. Proposed Solution

The Hospital data which is considered contains the private fields of the patient in which they are to be modified by applying PPDM techniques so that those private fields cannot be accessed by either public or practitioners. That data to be modified we need to apply few PPDM[8] techniques. In our work we implement four different techniques.

Classification of PPDM techniques based on few criteria are:

- Perturbation
- Randomization
- Anonymization
- Condensation
- Cryptography
- Hybrid

These are the techniques classified under Privacy Preserving Data Mining. These are explained below.

In which we use the techniques namely Anonymization, Generalization, Suppression and Data Hiding.

6.1 Anonymization

This is one of the PPDM techniques. In our present work this is the technique which we are implementing. This includes other sub parts like Generalization and Suppression. These are the three techniques used on different attributes of hospital database. All the private details of the patient are hidden using these techniques and the public data is projected out using the help of web page.

This Anonymization is the technique where it hides the data. For example, let us consider the zip code attribute. This should be hidden and should not be accessed by other users. This is shown in the below Table 1.

Table 1. Anonymization

Zip code	Anonymized data
53756489	5375****
52220123	5222****
45673456	4567****

In this way the private data can be hidden and can be protected.

6.2 Generalization

This is the other technique we use to hide the private data of the patient. This is the technique in which it generalizes the data. For example, when considered a supermarket data base the products are to be generalized as shown in the below Table 2.

Table 2. Generalization

Product	Generalized data
Apple	Fruit
L'Oreal	Shampoo
Santoor	Soap

6.3 Suppression

This is the technique in which we suppress the original data so that the end user cannot access the original data. For example, considering the private attribute like address, it should be hidden. As being a private field it should not be predicted by the end user. So the original data is suppressed to maintain privacy as shown in the below Table 3.

Table 3. Suppression

Address	Suppressed data
5-6-3,Gandhi Chowk,Tenali	Tenali
Ft.no:401, Sai Nadh Colony,Gorantla, Guntur.	Guntur

6.4 Data Hiding

In this we also use a technique called "Data Hiding[9]" which is used to hide the private fields of patient like Name and etc. The entire attribute is hidden from the database. This is one of the PPDM technique used to preserve private data and maintain the privacy.

6.5 Hybrid

This is a technique in which we combine two or more techniques to provide privacy to the data. As these PPDM techniques are many in numbers, this is a new approach in which we combine more number of techniques to provide more security to the private data. This technique results in more accuracy to the data fields.

In our present work we combine 4 different techniques and so it is called as "Hybrid Technique[10]", the different techniques namely Anonymization, Generalization and Suppression. These Generalization and suppression are the part of Anonymization. The other technique namely Data Hiding is used to hide the private data fields, here which is used for the field named "Patient name."

Considering a hospital database which has the private fields of patient details like "Pid, Name, Dob, Age, Gender, Address, Contact, Disease, Prescription, Treatment Duration and Precautions" in which the fields like Patient name, Contact, Address and Dob are private fields in which privacy is needed and are to be hidden. The other fields are public data and can be projected to the public. So, to the fields where we need to provide security PPDM techniques are applied. The above stated three techniques are here applied to the private details of the patient.

- Anonymization : Contact
- Generalisation : Dob
- Suppression : Address
- Data Hiding: Patient Name

The original data would be modified using the above specified techniques and the difference between the original and modified data s shown below. The modified data is clustered based on "diseases" using "Weka" tool. This clustered data is connected to the front end. This web page includes the search box for the end user in which he can specify the disease he requires to get the information. This result gives you the clear idea about the disease, symptoms, precautions, medicine and etc. This data also includes the private fields of the patient which are secure as PPDM techniques have been applied to the data. This data when projected publicly, private fields will be preserved securely and cannot be predicted by others.

7. Results

In this work the original database is modified using PPDM techniques which are stated above. Figure 4 shows you the original database which is stored and the PPDM techniques are applied for the datasets.

After applying the techniques the result is shown in Figure 5.

Pid	Pname	Gender	DoB	Age	Address	Contact	Disease	Symptoms	Medicine	Treatment
85001	A.Arjun Kumar	Male	1991-04-05	24	7-1-309/a/29,BK guda Hyderabad	9885578316	Migrane	headache,vomitings	Imitex	have good sleep and rest
85002	M.Bobby	Male	1985-01-01	31	6-2-23/2,sali pet,Tenali	8985432514	Asthma	Breathing problem	Predrisone	Aware of dusty area
85003	B.Anusha Devi	Female	1999-02-03	17	Ft:401,sai nadh colony,guntur	9493922345	thyroid	weakness, obesity	thyronorm	balanced diet,exercises
85004	H.Dhanush	Male	1971-07-06	45	3-23-4/1 1 town ,vijayawada	761686513	sunstroke	giddiness	plavix	good glucose levels
85005	T.Eswar	Male	1991-03-04	35	8-25-6/2,Gayatri,Vijayawada	9134863548	tyroid	high fever,vomitings	floxin	avoid junk foods
85006	T.Ravi Kumar	Male	1987-04-05	28	29-6-12/3,autonagar,vijayawada	9705621476	hepatitis	joint pains,Fatigue	interferon alfa 2B	avoid HBV
85007	S.Ramesh	Male	1957-03-04	59	FT/G8 ablock,vijayawada	9032713637	heart attack	chest pain	anti coagulants	cholestrol control
85008	M.Suma	Female	1972-11-13	44	Ft.201,devi chowk,tenali	7382057397	Diabetes	high glucose levels	Glycomet Gp1,Gp2	avoid sweets,cholestrolcontrol
85009	T.Mohan Krishna	Male	2001-07-13	15	26-4/3a,swathi road,bhavanipuram,vijayawada	9876234534	small pox	itching,rashes,high fever	nill	nill
85010	L.Sushmitha	Female	2006-01-01	10	1-26-8,kanakavari totha,guntur	9246151456	Fever	Body pains,Headache	Dolo	healthy diet
85011	T. Vijaya Lakshmi	Female	2004-05-07	12	16-25-21, Old City , Hyderabad	9876532454	Cold and Flue	Headache	Dcold	Nil
85012	M Anusha	Female	2007-07-21	9	12-4-23/12, Siva Colony, Mangalagiri	9886754342	Fever	Body pains, headache	Dolo	Healthy diet
85013	T. Arun	Male	1989-03-02	26	12-3-21, Gnadhinagar, Vijayawada	9550930940	Migrane	Headache	Imitex	Nil
85014	A. Sai Kumar	Male	1993-03-06	22	12-24-89, Second Floor, 1 Town, Vijayawada	9491395770	Asthma	Breathing Problem	Predrisone	Away from dusty area
85015	M. Chandra	Male	1961-03-21	55	2-3-14/B, Thadithota, Rajahmundry	9567834234	Diabetes	High Glucose Level	Glycomet GP1, GP2	Cholestrol Control
85016	S. Ankita	Female	1993-04-24	22	12-3-47, Rao Colony, M G peta, Cuddapah	8686238697	Migrane	Headache	Imitex	Nil
85017	V. Hemalatha	Female	1995-04-09	21	2-115, Jonnalagadda, Guntur	9030104785	Migrane	Headache	Imitex	Nil
85018	P. Hruday	Male	1984-05-19	32	6-63-A/2, Koodareddy Colony, Nellore	9701609180	Diabetes	High Glucose Levels	Glycomet GP1	Maintain proper diet
85019	A. Sravya	Female	1985-07-07	31	2-3-3, Hemanth Nagar, Mylavaram	8500325325	Thyroid	High fever, Vomitings	Floxin	Avoid junk food

Figure 4. Original database.

Pid	Gender	DOB	Age	Address	Contact	Disease	Symptoms	Medicine	Treatment
85001	male	1991	24	hyderabad	988557****	migrane	headache,vomitings	imitex	nill
85002	male	1985	31	tenali	898543****	Asthma	Breathing problem	Predrisone	Aware of dusty area
85003	female	1999	17	guntur	942392****	thyroid	weakness,obesity	thyronorm	Exercise,Balanced Diet
85004	male	1971	45	vijayawada	761686****	SunStroke	giddiness	plavex	maintain good glucose levels
85005	male	1991	35	vijayawada	913486****	typhoid	high fever,vomitings	floxin	avoid junk food
85006	male	1987	28	vijayawada	970562****	hepatitis	fatigue,joint pains	interferonalpha2b	avoid HBV
85007	male	1957	59	vijayawada	970898****	heart attack	chest pain	anti coagulants	cholestrol control,daily diet
85008	female	1972	44	Tenail	738205****	Diabates	High glucoselevels	Glycomet Gp1	avoid sweet,chloestra
85009	male	2001	15	vijayawada	987623****	small pox	itching rashes,	Nill	Nill
85010	Female	2006	10	Guntur	924615****	fever	body pains,head ache	Dolo	healthy diet
85011	Female	2004	12	hyderabad	987653****	Cold and flu	head ache	imitex	nill
85012	Female	2007	9	mangalagiri	988675****	fever	body pains,head ache	dolo	healthy diet
85013	male	1989	26	khammam	955093****	migrane	head ache	imitex	nill
85014	Male	1961	55	Tenail	956173****	Asthma	Breathing problem	Predrisone	Aware of dusty area
85015	male	1961	55	tenali	956783****	diabates	high glucose levels	Glycomet Gp 1, Gp2	balanced diet, exercise
85016	Female	1993	22	kadapa	868623****	Migrane	headache,vomitings	Imitex	Nill
85017	Female	1995	21	Guntur	903010****	Migrane	headache,vomitings	imitex	Nill
85018	Male	1984	32	Nellore	970160****	Diabetes	High Glucose Levels	Glycomet Gp-1,Gp-2	Maintain Proper diet
85019	Female	1985	31	Mylavaram	850032****	Typhoid	High Fever,Vomitings	Floxin	Avoid Junk food

Figure 4. Original database.

The difference btween the original datasets and the modified datasets are shown in Figures 6, 7.

Pid	DoB	Age	Address	Contact
85001	1991-04-05	24	7-1-309/a/29,BK guda Hyderabad	9885578316
85002	1985-01-01	31	6-2-23/2,sali pet,Tenali	8985432514
85003	1999-02-03	17	Ft:401,sai nadh colony,guntur	9493922345
85004	1971-07-06	45	3-23-4/1 1 town ,vijayawada	761686513
85005	1991-03-04	35	8-25-6/2,Gayatri,Vijayawada	9134863548
85006	1987-04-05	28	29-6-12/3,autonagar,vijayawada	9705621476
85007	1957-03-04	59	FT:G8 ablock,vijayawada	9032713637
85008	1972-11-13	44	Ft.201,devi chowk,tenali	7382057397

Figure 6. Before Applying PPDM techniques.

Pid	DOB	Age	Address	Contact
85001	1991	24	hyderabad	988557****
85002	1985	31	tenali	898543****
85003	1999	17	guntur	942392****
85004	1971	45	vijayawada	761686****
85005	1991	35	vijayawada	913486****
85006	1987	28	vijayawada	970562****
85007	1957	59	vijayawada	970898****
85008	1972	44	Tenail	738205****

Figure 7. After PPDM techniques.

As the data is retrieved from from the front end based on end user requirement the Figure 8 shows the front end.

Figure 9 shows the page in which end user checks

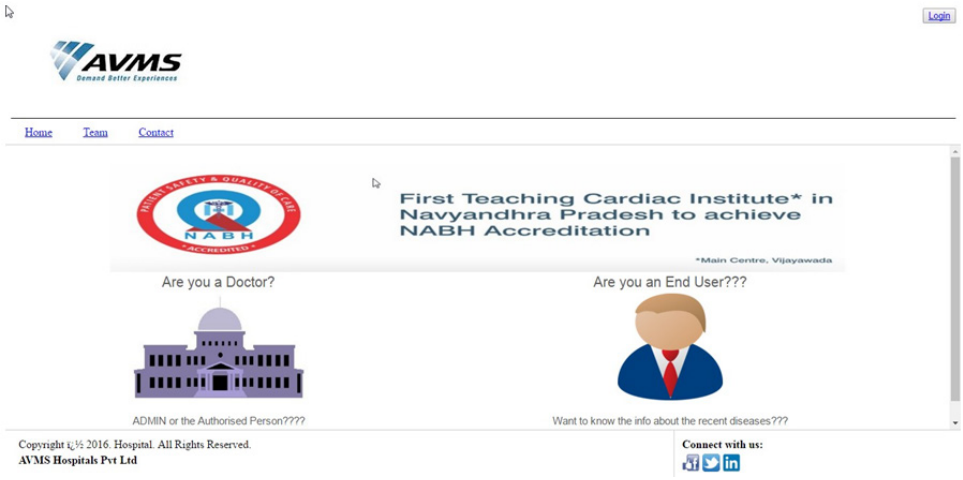


Figure 8. Front end web page.

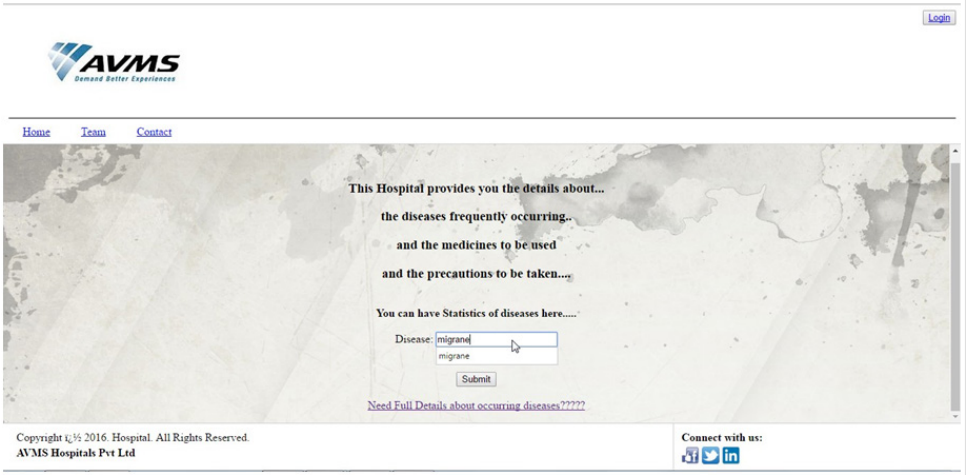


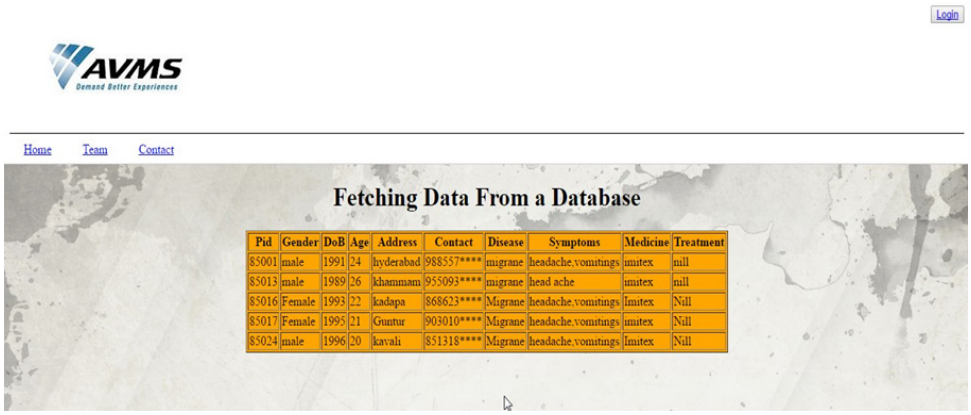
Figure 9. Search Query based on required disease.

for the data based on the disease requirement. Figure 10 shows the result of the quirey of end user.

Figures 11,12 shows the results about the Admin. His login details and the datasets.

8. Conclusion

This work is done to project the live hospital data so that they can take needful measures. But this includes a problem of patient's private data. In our work we overcame this problem by using 3 different PPDM techniques. This will be a useful idea even for the practitioners or for the public to provide awareness. This can be in future implemented by having the database of various hospitals.



Pid	Gender	DoB	Age	Address	Contact	Disease	Symptoms	Medicine	Treatment
85001	male	1991	24	hyderabad	988557****	migrane	headache,vomitings	imitex	nill
85013	male	1989	26	khammam	955093****	migrane	head ache	imitex	nill
85016	Female	1993	22	kadapa	868623****	Migrane	headache,vomitings	Imitex	Nill
85017	Female	1995	21	Guntur	903010****	Migrane	headache,vomitings	imitex	Nill
85024	male	1996	20	kavali	851318****	Migrane	headache,vomitings	Imitex	Nill

Figure 10. Result for the query "Migraine".

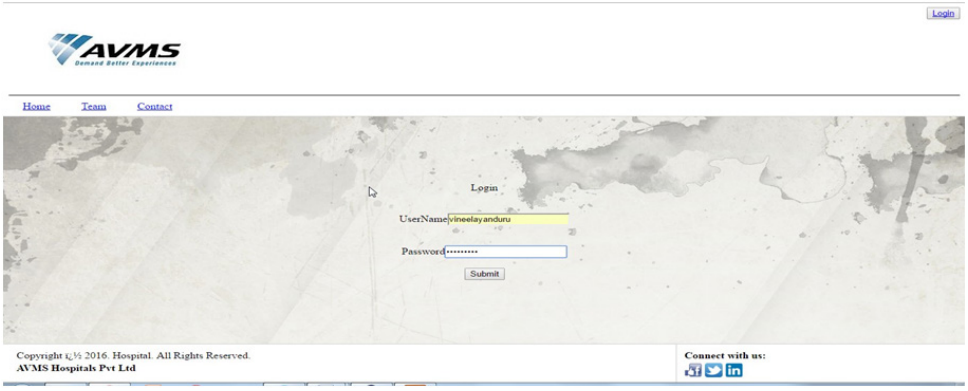


Figure 11. Admin login.

Pid	Pname	Gender	DoB	Age	Address	Contact	Disease	Symptoms	Medicine	Treatment
85002	M.Bobby	Male	1985-01-01	31	6-2-23/2,sali pet,Tenali	8985432514	Asthma	Breathing problem	Predrisone	Aware of dusty area
85003	B.Amsha Devi	Female	1999-02-03	17	Ft:401,sai nadh colony,guntur	9493922345	thyroid	weakness, obesity	thyronorm	balanced diet,exercises
85004	H.Dhanush	Male	1971-07-06	45	3-23-4/1 1 town ,vijayawada	761686513	sunstroke	giddiness	plavix	good glucose levels
85005	T.Eswar	Male	1991-03-04	35	8-25-6/2,Gayatri,Vijayawada	9134863548	tyroid	high fever,vomitings	floxin	avoid junk foods
85006	T.Ravi Kumar	Male	1987-04-05	28	29-6-12/3,autonagar,vijayawada	9705621476	hepatitis	joint pains,Fatigue	interferon alfa 2B	avoid HBV
85007	S.Ramesh	Male	1957-03-04	59	FT:G8 ablock,vijayawada	9032713637	heart attack	chest pain	anti coagulants	cholestrol control
85008	M.Suma	Female	1972-11-13	44	Ft.201,devi chowk,tenali	7382057397	Diabetes	high glucose levels	Glycomet Gp1,Gp2	avoid sweets,cholestrolcontrol
85009	T.Mohan Krishna	Male	2001-07-13	15	26-4/3a,swathi road,bhavanipuram,vijayawada	9876234534	small pox	itching,rashes,high fever	nill	nill
85010	L.Sushmitha	Female	2006-01-	10	1-26-8,kanakavari totha,guntur	9246151456	Fever	Body pains,Headache	Dolo	healthy diet

Figure 12. Admin result.

9. References

1. Vaghashia H, Ganatra A. A survey: privacy preservation techniques in data mining. International Journal of Computer Applications. 2015 Jun; 119(4):20–26.
2. Taneja S, Khanna S, Tilwalia S, Ankita. a review on privacy preserving data mining: techniques and research challenges. International Journal of Computer Science and Information Technologies. 2014; 5(2):2310–15.
3. Keyvanpour MR, Moradi SS. Classification and evaluation the privacy preserving data mining techniques by using a data modification–based framework. International Journal on Computer Science and Engineering. 2011 Feb:1–9.
4. Rajalakshmi V, Mala GSA. Anonymization by data relocation using sub-clustering for privacy preserving data mining. Indian Journal of Science and Technology. 2014 Jul; 7(7):975–80. doi: 10.17485/ijst/2014/v7i7/44454.
5. Hariharan R, Mahesh C, Prasenna P, Kumar RV. Enhancing privacy preservation in data mining using cluster based greedy method in hierarchical approach. Indian Journal of Science and Technology. 2016 Jan; 9(3):1–8. doi: 10.17485/ijst/2016/v9i3/86386.
6. Rathna SS, Karthikeyan T. Survey on recent algorithms for privacy preserving data mining. International Journal of Computer Science and Information Technologies. 2015; 6(2):1835–40.
7. Priyadarsini RP, Valarmathi ML, Sivakumari S. Attribute segregation based on feature ranking framework for privacy preserving data mining. Indian Journal of Science and Technology. 2015 Aug; 8(17):1–9. doi: 10.17485/ijst/2015/v8i17/77584.
8. Sashirekha K, Sabarish BA, Selvaraj A. A study on privacy preserving data mining. International Journal of Innovative Research in Computer and Communication Engineering. 2014 Jul; 2(3):1–5.
9. Trombetta A, Jiang W, Bertino E, Bossi L. Privacy-preserving updates to anonymous and confidential databases. IEEE Transactions on Dependable and Secure Computing. 2011 Jul/Aug; 8(4):578–87.
10. Gionis A, Tassa T. k-Anonymization with minimal loss of information. IEEE Transactions on Knowledge and Data Engineering. 2009 Feb; 21(2):206–09.