Pharma Base Analytical Dashboard

ROSHNI THAKKAR [1], TANISHA MEHTA [2], ARUNDHATI PANDE [3], HARISH PADOLE [4],
APOORVA SHARMA [5], DINESH BANABAKODE [6]

[1, 2, 3, 4, 5, 6] *Dept.of Information Technology, S.B. Jain Institute of Technology Management and Research Nagpur, Maharashtra, India*

Abstract -- This paper reviews about the analysis of Cancer disease in Nagpur district. Analysis on the number of patients suffering from different types of cancer, mainly lung cancer, oral cancer, breast cancer and cervical cancer. The main aim behind this dashboard is to display the analysis on cancer disease in Nagpur district in an efficient manner. The technology used for developing this dashboard is sequel server and tableau. ETL processes in data warehousing are applied to implement grids. The four grids display the heat map representation of the region, bar graph representation of medicines availability, demographic representation according to the gender and information about the hospitals. With the help of survey and authenticated dataset is created and then various data mining algorithms are applied on the data.

Indexed Terms: data mining, cancer, analysis.

I. INTRODUCTION

In this project an Analytical Dashboard is implemented that will brief an analysis on population of Nagpur region affected by a cancer disease. There are four types of cancer considered in the project and they are lung cancer, oral cancer, breast cancer and cervical cancer. The population of Nagpur region affected by cancer and also the hospitals and medicines availability of the disease are mentioned.

This will further drive the analysis on availability and requirements of medicines with the help of bar graph representation. After that the third grid displays the gender based analysis through the demographic representation. The percentage of people suffering from cancer, are bifurcated by their gender. In the last grid information about the hospitals treating that specialized cancer will be displayed.

The end-users receive generalized information about the number of population affected by a particular cancer, medicines consumed by patients from a particular disease and its requirements, the demographic representation of genders affected by cancer and lastly grid information about hospitals in Nagpur region. The hospitals, health ministry, patients and pharmacy can be the end users of this dashboard.

II. LITERATURE REVIEW

In "Chemo Explorer: A Dashboard for the Visual Analysis of Chemotherapy Response in Breast Cancer Patients" by N. Karall, M.E. Gröller, R.G. Raidou, it is discussed that how there are different alternatives can be employed in breast cancer chemotherapy. This research worked introduced a web based dashboard which facilitates the comparison and analysis of publicly available breast cancer chemotherapy response and also provide the initial basis for clinical research working on chemotherapy optimization.

In "Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research" by Beat A. Schwendimann, Maria Jesus RodriguezTriana, Andrii Vozniuk, Member, IEEE, Luis P. Prieto, Member, IEEE, Mina Shirvani Boroujeni, Adrian Holzer, Denis Gillet, and Pierre Dillenbourg, presents the state-of-the-art of research on learning dashboards in the fields of learning analytics and educational data mining. It also helps in understanding how visualization is important. And the research on learning dashboards is still at a young age and increasing. This helps us to understand about visualization and learning analytics.

"Study on Business Intelligence Tools for Enterprise Dashboard Development" by K.Gowthami, M.R. Pavan Kumar, presents detailed information on Business Intelligence (BI) tools for developing a

dashboard. And also it focuses on some potential business intelligence tools like Spago BI, Power BI, Tableau, etc. With this paper we get brief idea about tableau, how tableau is efficient for developing an interactive dashboard.

III. OBJECTIVE

The main objective behind this project is to enable the end users to keep track on population affected by a disease and medicine required to cure it. This will result in effective decision making. The end users of the dashboard will be health ministry and hospital inventory who can easily identify problems highlighted through KPI's (key performance indicator).

IV. METHODOLOGY

ETL is used for creating different grids in the dashboard. In computing, **extract, transform, load** (**ETL**) refers to a process in database usage and especially in data warehousing. Data extraction involves extracting data from homogeneous or heterogeneous sources, while data transformation processes data by transforming them into a proper storage format/structure for the purposes of querying and analysis; finally, data loading describes the insertion of data into the final target database such as an operational data store, a data mart, or a data warehouse.
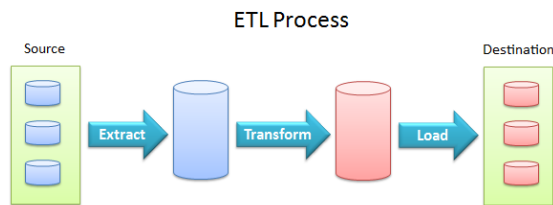

The project is completed by surveying different cancer specialized hospitals in Nagpur region. The survey is done according to number of patients affected by different kinds of cancer mainly lung cancer, oral cancer, breast cancer and cervical cancer.

a) Survey:
- Oral cancer:

Oral cancer is major public health problem in the Indian subcontinent, where it ranks among the top three types of cancer in the country. Nagpur is the oral cancer capital of the world. 50% cancer in men and 20% cancer in female is caused due to tobacco use according to Indian Council of Medical Research (ICMR) tobacco-related cancers are head and neck cancers which cancers of oral cavity, tongue, oesophagus.
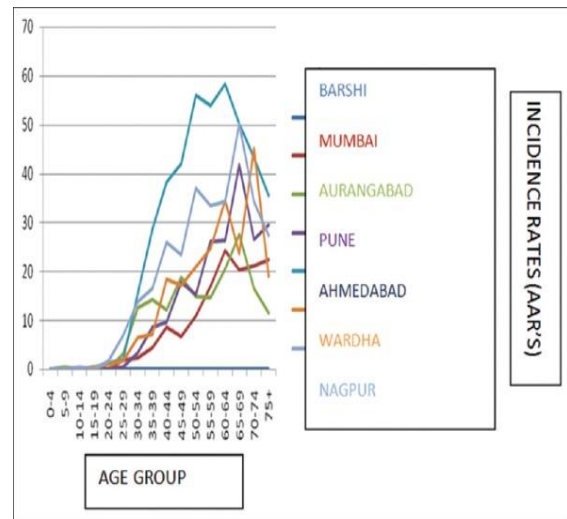
- Lung cancer:

Nagpur is on the second position in terms of lung cancer. The main cause is due to tobacco consumption. Also large number of patients from central India visit Nagpur for cancer treatment. A train from Jabalpur to Nagpur is labelled as "cancer train" as the large number of patients come for treatment.

- Breast Cancer

Breast cancer is the commonest cancer in women in Nagpur. Breast cancer accounts for31.9%of all cancers in women in Nagpur. Nagpur is on third position in terms of breast cancer.

- Cervical Cancer:

Cervical cancer, mainly caused by Human Papillomavirus infection, is the leading cancer in Indian women and the second most common cancer in women worldwide. Though there are several methods of prevention of cervical cancer, prevention

by vaccination is emerging as the most effective option, with the availability of two vaccines. Several studies have been published examining the vaccine's efficacy, immunogenicity and safety. Questions and controversy remain regarding mandatory vaccination, need for booster doses and cost-effectiveness, particularly in the Indian context.

b) Dashboard:

After the survey on cancer in hospitals the dashboard will consist four major grids namely the heat map representation, the bar graph representation of medicines and its availability, the demographic representation according to the gender and last but not the least information about the hospitals in Nagpur region. As the dashboard is strictly concentrated to cancer, there will be a drop down given which will have the four cancer names. Once a particular cancer is selected the dashboard will change dynamically and display all the grid information accordingly.

•   Heat map representation:

Heat map representation is the first grid of the dashboard. This will display the heat map restricted to Nagpur district only. With the help of heat map representation, it will become easier to understand that in which reason the people are more affected by cancer. For example, if we select oral cancer the dashboard will display the details related to oral cancer. Likewise, the het map will highlight the areas in Nagpur region affecting from oral cancer.

•   Medicines Availability:

In the second grid information related to medicines availability will be displayed in bar graph representation. Here, we will be showing the analysis of the medicines that are used for the treatment and again this will be as per the type of cancer the user is selecting. Once the type of cancer is selected then the bar graph representation will show the medicines available and medicines required. With all these representations it will become easier to provide or avail medicines in the areas with its necessities. And the areas with excess medicines surplus can provide medicines in the areas needed.

•   Demographic Analysis

The third grid is the demographic analysis of the cancer. This analysis will be done on the basis of gender and according to the age groups. The number of male and female affected by any particular type of cancer will be displayed. This analysis will be shown according to the age group. For suppose we take female affected by cervical cancer then it will visualise the data as per the age groups affected by it. Similarly, if we take oral cancer then the third grid will display the number of males and females affected by it in different age groups.

•   Hospital information

In the final grid the information related the hospitals treating that cancer will be displayed. This will have various attributes like hospital name, doctor's name, address, contact number, reviews etc. All this information will again help to display the data of the hospitals. With further analysis we can also add more predictive modules but that will be strictly depending upon the real time data. The disease selector plays the main role in the in the dashboard. The complete will change dynamically according to the disease.

The system architecture of the dashboard is shown below:
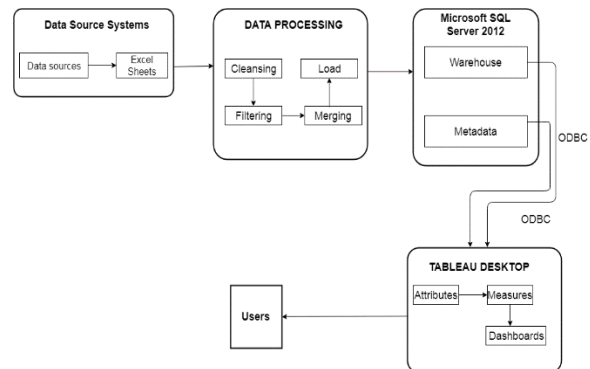
c) Tableau:

Business Analytic or Business Intelligence (BI) is becoming more needed by the top management of any enterprise to visualize, analyse and prepare the strategic planning for the future. One of the important tools is tableau which will be used in the project.

Tableau has been very aggressive in marketing by the most attractive feature of drag and drop products. Anyone can analyse data with Tableau's intuitive

drag & drop products. No programming, just insight. Connecting and visualizing your data in minutes, Tableau is 10 to 100x faster than existing solutions are some of the significant features for the business analysts to choose this tool. Tableau products query relational databases, OLAP cubes, cloud databases, and spreadsheets and then generates a number of graph types. The products can also extract data and store and retrieve from its in-memory data engine. The dashboards are rich in terms of data visualization and very interactive.

Below is an example of dashboard created through tableau.

d) Data Mining:

A more recent innovation in the world of data mining tools and techniques is the Dashboard. A data-mining Dashboard is a piece of software that sits on an end-user's desktop or tablet and reports real-time fluctuations in data as it flows into the database and is manipulated or sorted.

Data mining is the extraction of predictive information from large databases. Recently, there has been a growing emphasis on exploratory analysis of very large datasets to discover useful patterns and/or correlations among attributes. Data mining are helps customers leverage their data more effectively and obtain insightful information that can give them a competitive edge. Simply put, data mining software enables customers to discover previously undetected

facts present in their business-critical data -- data that may consume many gigabytes or terabytes of storage, may reside in files or various DBMS-managed databases, and may be stored on a variety of operating system platforms. Accuracy, efficiency, and an open architecture are important requirements of such data mining software.
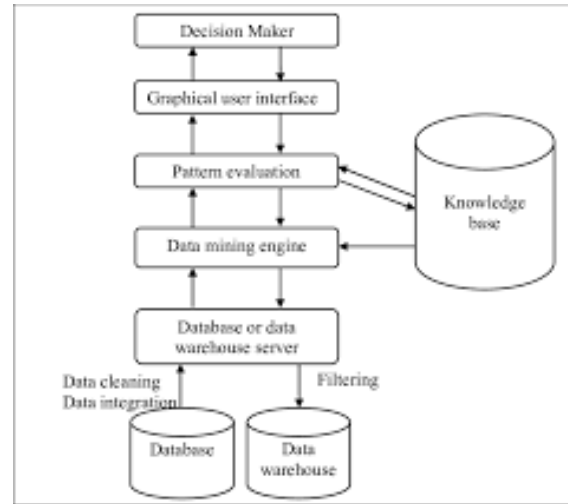
V. FUTURE SCOPE

Dashboard can be used for showing information in an efficient way rather than traditional way. This dashboard can be further used by the health ministry to figure out surplus or deficiency of medicines related to a particular disease. This would help in maintaining statistics related to health care for Nagpur region.

The future scope of the project is that once it is implemented on city level then further it can be expanded on state level.

VI. CONCLUSION

We have introduced an approach by which analysis and prediction can be done for cancer affected people in Nagpur region. The conceptual idea of this work is based on the specific available data set, for which comparisons and analysis are required. Linking to prediction of the treatment outcome of new patients, based on similarities to existing patients, is another relevant direction of the project. The end-users would

receive generalized information about medicines consumed by patients from a particular disease and maintaining a track on availability and requirements of medicines.

REFERENCES

[1]   ChemoExplorer: A Dashboard for the Visual Analysis of Chemotherapy Response in Breast Cancer Patients. N. Karall1 , M.E. Gröller1,2 , R.G. Raidou1 1TU Wien, Austria, 2VRVis Research Center, Austria. Issue: 2018.

[2]   Study on Business Intelligence Tools for Enterprise Dashboard Development. K.Gowthami, M.R. PavanKumar. International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue: 04 | Apr 2017 https://www.irjet.net/archives/V4/i4/IRJET-V4I4721.pdf

[3]   Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research Beat A. Schwendimann, Maria Jesus Rodriguez-Triana, AndriiVozniuk, Member, IEEE, Luis P. Prieto, Member, IEEE, Mina ShirvaniBoroujeni, Adrian Holzer, Denis Gillet, and Pierre Dillenbourg Ieee transactions on learning technologies, vol. 10, no. 1, january-march 2017 https://ieeexplore.ieee.org/document/7542151/

[4]   https://www.sisense.com/dashboard-examples/healthcare/healthcare-analytics/

[5]   https://timesofindia.indiatimes.com/city/ nagpur/ Nagpur-oral-cancer-capital-of-the-world/articleshow/52525048.cms

[6]   http://www.breastcancerindia.net/ statistics/stat_nagpur.html