**INTRODUCTION**

The main goal of data mining is to generate meaningful information or knowledge from the database, usually the term knowledge discovery is used together with data mining. Knowledge discovery is also known as knowledge discovery in database (KDD) and it is the process that seeks new knowledge in some application domain.

There are wide scope of applications of data mining in a diversity of segments namely telecommunication, scientific applications, healthcare and biomedical data analysis, financial and banking data analysis, retail industry, intrusion detection, banking, education, supermarkets, crime agencies, engineering disciplines and so forth.

**Review of Literature**

A researcher H. Shaker, et.al. (2013) point out in his paper entitled “Data Mining and, Techniques, Challenges and Knowledge Discovery: ApplicationsProcess Models in Healthcare”, that Integration of KDD tools with electronic health record might diminish the medical errors, improve patient safety and reduce an unwanted practice variation and improve patient outcome. In large datasets, there are many variables interacting with one another in very subtle ways and KDD helps to find the hidden relationships and patterns within data. This paper also talks about the applications of Knowledge discovery, text mining, and data mining in healthcare by comparing causes, symptoms, courses of treatments, to deliver an analysis and to prove the effectiveness like a prediction of optimum medication dosage. (24)

A research paper entitled “Dengue disease prediction using Weka data mining tool” by SHAKIL K., ANIS S. and ALAMM. (2013) detailed on how the various Data Mining techniques and WEKA tool is used for prediction of Dengue disease. The algorithms Naïve Bayes, J48, SMO, REP Tree and Random tree are compared and performance of these five algorithms is used for prediction of Dengue diseases and researcher had used Dengue dataset for performance evaluation and prediction. (25)

A research paper entitled “Advanced Knowledge-Based Systems: Model, Applications &Research” by Sajja P. and Akerkar R. (2010) represented the various research areas in pure knowledge-based system as KBS development, Knowledge management, Information and query-based system, user interface, expert system, distributed KBS, knowledge grid, multi-agent system, tutoring system, soft computing based KBS, etc. Knowledge discovery is one of the application and current research area of Knowledge-based data mining. (26)

The author’s Patel M., et.al. (2013) had designed Rule-based Expert system for viral infection diseases by collecting information from medical experts. The researcher had focused on knowledge generation in the form of disease, symptoms, and medicines; and designed the system for 12 diseases and
compared it with the actual doctor’s diagnosis in a research
titled “Rule-Based Expert System for Viral Infection
Diagnosis”. (27)

Viral Infectious Diseases

There are various communicable (infectious) diseases such as
Diarrhoea, Viral Hepatitis, Enteric Fever, Malaria, Dengue, Chikungunya, Acute Encephalitis Syndrome(EJE) Meningitis, Measles, Diphtheria, Pertussis, Chicken Pox, Pneumonia, Leptospirosis, Dog Bite, Snake Bite, Adult Tetanus, Neonatal Tetanus, Cholera, Influenza H1N1, Rabies, HIV etc. These
diseases spread through the virus, air, mosquito, water, animal,
direct and indirect contact, etc. The Pune Municipal Corporation (PMC) and Pimpri Chinchwad Municipal Corporation (PCMC) take various initiatives for making the
awareness of viral diseases through Pamphlets, Stickers and
banners, Messages, Radio jingles, Booklets, Teachers and
students meeting, Stickers about VBD on PMPL buses, Society
chairman’s meetings, Audio, video, NGO women’s health
groups, self-help groups, Rallies and public meetings, IMA and
other private practitioners’ meetings, etc.

It has also observed from the following graph no.1. that, the
average number of cases in four years from 2014-2017 of Dog
Bite are 11660 , 7700 of Acute Respiratory Infection (ARI)
/Influenza Like illness (ILI), Dengue /DHF/DSS Suspected
cases are 1598.75,772.75Confirmed Dengue /DHF/DSS, 1294.75 of Actual Diarrheal Disease(including gastro), 874 of Chikungunya, 406 of Dysentery, Influenza H1N1 are 357.5, Enteric Fever is 158.25.

Data Collection and Analysis

The climate, hygiene, poverty, dense population are some
important factors and problems of under developed countries;
and because of this the under developed countries are facing
the problem of infective viral diseases. Based on these
discussion and interview with doctors and medical experts
following are the most common viral infective diseases found
in India and mostly in Maharashtra as: Dengue, Chikungunya,
Influenza (Flu), Cholera, Malaria, Viral Hepatitis (A), Mumps,
Measles, HIV, Chickenpox, Herpes, Swine flu, etc. In the field
survey researcher had gathered the information regarding the
diseases and related parameters in the form of symptoms,
reasons, impact on health, treatment, precautions, etc. from
doctors and experts. Total 141 respondents are selected for the
data collection and filed survey.

Frequency of Viral Infective Diseases

The most of the diseases are highly frequent diseases in Pune Municipal Corporation (PMC) and Pimpri Chinchwad Municipal Corporation (PCMC).

The very high and high frequency of diseases such as Dengue, Chikungunya, Malaria, Influenza Flu, Cholera, Hepatitis, Herpes, Diarrhea, Typhoid, Mumps, Rabies, Measles, HIV, Chickenpox and Swine Flu is signified in following graph 4.5. It has witnessed that, 90.8 % respondents are ranked the Influenza Flu as high frequency diseases, 78.7 % for Dengue, 75.2 % for Chikungunya, 72.3 % for Diarrhea, 71.6 % for HIV, 61.7 % from Swine Flu. While as the diseases like Typhoid, Malaria, Herpes, Rabies, Hepatitis, Cholera, Measles, Mumps and Chickenpox are rated as 17.1 %, 16.3 %, 16.3 %, 11.4 %, 8.5 %, 7.8 %, 7.1 %, 6.3 %, 5.7 % of respondents consecutively.

Sources of Elevated Risk of Viral Diseases Infection

The viral diseases are highly spread through water, air,
mosquitos, direct and indirect contact, etc. There is risk of viral
infection at various sources such as public transport,
entertainment enterprises, shops, malls, schools, hospitals,
home, family, work places etc. These sources of diseases
infection are considered through the interview and discussion with the doctors and medical experts.
Therefore it has been concluded that, 57.45 % respondents are declares that viral infection are spread more in public transport, 47.52 % respondents articulates that entertainment enterprises is also a source of viral infections, 43.97 % utters that in shops and malls there are chances of viral infections, 64.54 % reveals that there are more viral infections in schools, in hospitals there are 34.75 % infections, at home and with family there are 29.08 % chances of infection and 40.43% infections are spread at work place.

Hence it has been perceived that, there are more viral infections in school as compare to public transport, entertainment enterprises, shops and mall, work places, hospitals, home and with family sequentially.

Need of Diseases Outbreak System

| Table 1 Need of System for Prediction of Viral diseases Outbreaks |
|---------------------------------|------------------|
| **Need of System**              | **Percentage**   |
| Highly Agree                    | 29.8             |
| Agree                           | 48.2             |
| Neutral                         | 12.8             |
| Least Agree                     | 5.7              |
| Disagree                        | 3.5              |

Source: Data Compiled by researcher using SPSS tool

The above table represents that statistics for need of Prediction system for viral infective diseases outbreaks. The proposed system will predict the viral diseases outbreaks with various parameters such as diseases name, symptoms, impact on health, geographical location, precautions, etc. This prediction system is useful for various healthcare stakeholders to take remedial measures and better decision making for controlling the impact and increasing patients of viral diseases.

CONCLUSION

The present study is focused towards the understanding of common viral diseases namely such as Dengue, Chikungunya, Malaria, Influenza Flu, Cholera, Hepatitis, Herpes, Diarrhea, Typhoid, Mumps, Rabies, Measles, HIV, Chickenpox, Swine Flu, etc. are considered for the research.

The study is carried out to understand the viral diseases and related terms such as the season in which the diseases are spread, whether the viral infection is based on climate change, how the doctors maintains the diseases count, whether the doctors receives the outbreaks of viral infections, which are the sources used for outbreak prediction, can doctors predict the demographics of highly spread diseases, what are the sources or places where there are most chances of infections of these diseases, is the healthcare originations or departments need the systems for diseases outbreaks prediction, etc. The study is further enhanced towards the data mining of these viral infectious diseases information and data.

References

Unstructured Data, Cambridge University Press
www.cambridge.org/9780521836579.
13. https://wtlab.um.ac.ir/images/elibrary/text_mining/The
%20Text%20Mining%20HandBook.pdf
14. Jiawei Hon, Micheline Kamber, Jain Pei. (2012). Data
Mining concepts and techniques, 3rd Edition
techniques/ (Retrieved on 28 Nov 2016)
16. Viral Diseases. 
http://www.rightdiagnosis.com/v/viral/intro.htm
(Retrieved on 22 Nov 2016)
http://www.news-medical.net/health/Human-Diseases-
Sciences. http://www.news-medical.net/?tag=Dengue-
Fever. (Retrieved on 22 Nov 2016)
20. https://www.healthgrades.com/conditions/viral-diseases
2017)
(Retrieved on 6 April 2017)

How to cite this article:
Ashwini Manish Brahme and Shiavji D. Mundhe., 2019,A Study of Viral Infectious Diseases Using Data Mining Techniques,
Int J Recent Sci Res. 10(06), pp. 33150-33153. DOI: http://dx.doi.org/10.24327/ijrser.2019.1006.3616

*******