

## Predictive Analysis for Diabetes using Tableau

Dhanamma Jagli  
Dept. of MCA  
VESIT, Chembur  
Mumbai, India  
*dhanamma.jagli@ves.ac.in*

Siddhanth Kotian  
Dept. of MCA  
VESIT, Chembur  
Mumbai, India  
*siddhanth.kotian@ves.ac.in*

**Abstract**—As per the present circumstances, in India, Diabetic Mellitus (DM) has turned into a major wellbeing peril. Diabetic Mellitus (DM) is arranged as a Non-Communicable Diseases (NCD), and numerous individuals are experiencing it. Consistently vast volume of diabetic information is creating and consequently it is important to do examination on this information and settle on effective choices. In the healing centers different records, for example, patients' profile data, x-beam reports, different therapeutic tests' reports and so on are saved and this structures huge information. Enormous information examination is the procedure which inspects such huge informational collections and reveals shrouded data, concealed examples to find learning from the information. By applying investigation on medicinal services information, essential choice and expectation can be made. The framework proposed in this paper is a propelled answer for investigating the information in light of the information acquired from the past multiyear and after that giving outcomes to the up and coming year in an effective way. The proposed framework influences utilization of programming to device named Tableau to deliver the examination for up and coming year in light of the past five long periods of information.

**Keywords**-Tableau, data analysis, diabetes, prediction

\*\*\*\*\*

### I. INTRODUCTION

The current framework utilizes the Hadoop structure. Apache Hadoop is java composed open source system which can be utilized to process gigantic informational index on number of bunch of PCs in conveyed way. Hadoop can give dispersed capacity and preparing of information [1]. As indicated by the need of utilization Hadoop can scale from single server to numerous servers. Hadoop Distributed File System is utilized to store and process the information. Through HDFS enormous volume information can be access with less time and endeavors. Parallel preparing of such information is finished with MapReduce system. MapReduce system can be utilized to compose applications that can procedure huge informational indexes in dependable way. In MapReduce condition extensive information can be prepared in disseminated and parallel route on number of item equipment. MapReduce structure contains two stages, first is Map stage, in which the info information is changed over into middle of the road information as key esteem sets and in Reduce stage this moderate information is changed over in to conclusive yield by k implies grouping calculation [2]. The technique comprises of different stages. In the main stage, information from different sources is gathered together into a framework. As the information originates from different sources it is in various structures and also a portion of the qualities are discovered missing, so the second stage manages rounding out the missing qualities and changing over every one of the information into a settled configuration. When every one of the information is in the required configuration it is then tried for finding the examples that exists among them. These examples are then tried with the examples that exist as of now in the framework. In light of the examples existing in the

framework yield is produced by the framework [8]. Tableau is a Business Intelligence(BI) apparatus for visually investigating the data. Clients can likewise make and disseminate an intelligent and shareable dashboard, which delineate the patterns, varieties, and thickness of the data as diagrams and graphs. Tableau have some exceptional highlights, for example, Speed of Analysis, Self-Reliant, Visual Discovery, Blend Diverse Data Sets, Architecture Agnostic, Real-Time Collaboration, Centralized Data and most importantly Forecasting.

### II. BACKGROUND

#### Data Analysis

Data analysis is an essential segment of data mining and (BI) and is critical to picking up the knowledge that drives business choices. Associations and endeavors dissect data from a large number of sources utilizing Big Data administration arrangements and client encounter administration arrangements that use data investigation to change data into noteworthy insights. Data analysis is a procedure of assessing, purifying, changing, and demonstrating data with the objective of finding valuable data, illuminating conclusions, and supporting basic leadership [3].

#### Predictive Analysis

Predictive analytics is the branch of the progressed analytics which is utilized to make expectations about obscure future occasions. Predictive analytics utilizes numerous methods from information mining, insights, demonstrating, machine learning, and artificial knowledge to break down current information to make forecasts about future. It utilizes various information mining, predictive displaying and diagnostic

procedures to unite the administration, data innovation, and demonstrating business procedure to make expectations about future [4].

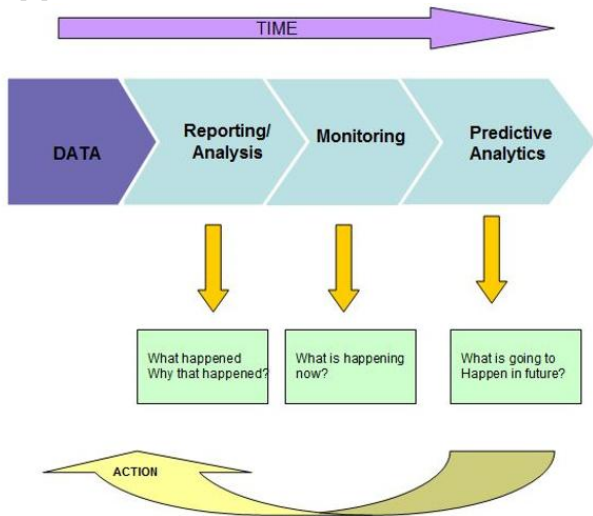


Figure 1. Predictive Analysis

### III. LITERATURE SURVEY

The survey was done mainly on the software to be used for predicting the results based on the previous year's data. Study found various method and software suitable for the prediction. The expected system must be not only fast but also accurate. After this the next decision was on the number of years of data to be provided to the software to get the most accurate results. So, after a detailed study Tableau was chosen as the software. Further study led to the conclusion of selecting previous data of five years to provide accurate results [5]. The data for the previous five years from all the sources available is collected and stored in an excel format. The data sheet is then exported into a comma separated value format. (.csv). The csv file is fed to the Tableau software. The software uses the data to predict the amount of diabetes patients in the next year accurately. A report is generated after all the analysis has been done by the software. The report is then stored for future use.

### IV. IMPLEMENTATION DETAILS

The implementation of the system uses Tableau software for producing accurate data for the next year using five years of previous data.

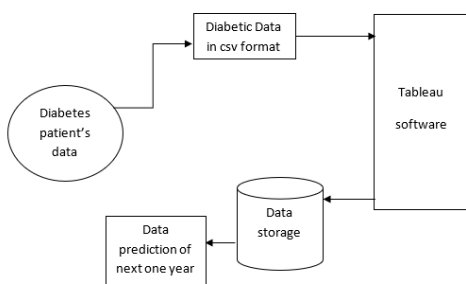


Figure 2. Implementation flow for predictive analysis using Tableau

### V. SOFTWARE SPECIFICATIONS

Tableau is a BI instrument for outwardly investigating the information. Clients can make and circulate an intelligent and

shareable dashboard, which portray the patterns, varieties, and thickness of the information as diagrams and outlines [6]. Tableau can interface with documents, social and Big Data sources to procure and process information. The product permits information mixing and ongoing coordinated effort, which makes it extremely interesting. It is utilized by organizations, scholarly researchers, and numerous administration associations for visual information investigation. It is additionally situated as a pioneer BI and Analytics Platform in Gartner Magic Quadrant. As a main information perception device, Tableau has numerous alluring and exceptional highlights [7]. Its ground-breaking information disclosure and investigation application enables you to answer vital inquiries in a flash. You can utilize Tableau's simplified interface to picture any information, investigate distinctive perspectives, and even consolidate different databases effectively. It doesn't require any complex scripting [9]. Any individual who comprehends the business issues can address it with a perception of the applicable information. After examination, offering to others is as simple as distributing to Tableau Server. The information is given to the scene programming in csv organize. In light of past five long stretches of the information, Tableau gives a point by point expectation of the information for the coming multi year.

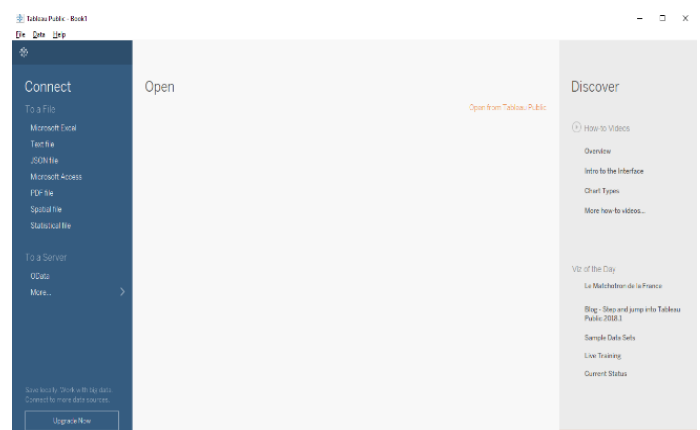


Figure 3. Screenshot of Tableau software.

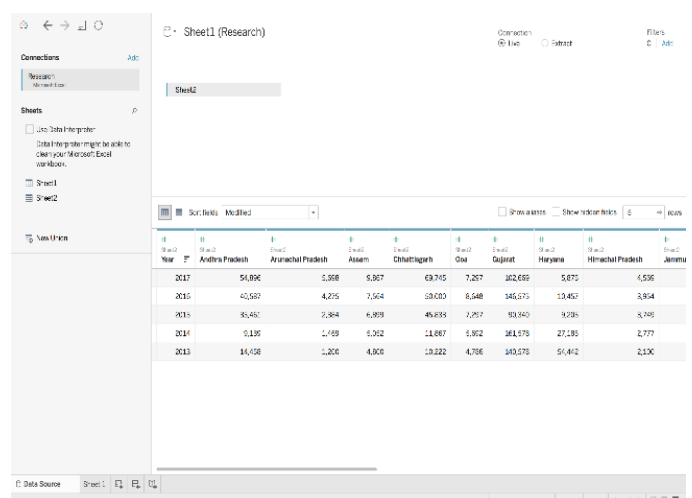


Figure 4. Screenshot of Tableau software.

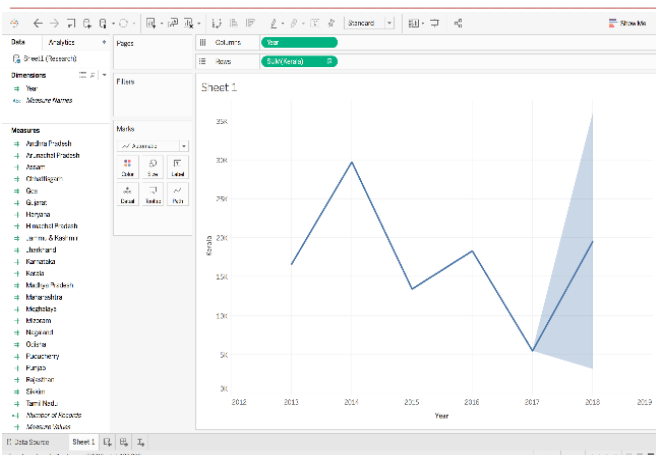


Figure 5. Screenshot of the output generated by Tableau.

## VI. CONCLUSION

Business intelligence has seen a lot of growth in recent years. Various tools have been developed and are in development to make it easier to develop BI applications. Tableau which is an example of one such tool has many different applications in the field of BI. Tableau is a simple, easy and quick to learn software that can be easily used by any personnel/professional from the related domain. Tableau's simplified interface and sophisticated dashboard offer numerous functionality useful for a variety of BI based solutions. Such solutions include predicting diabetic data of patients for a particular year and the same has been illustrated in this paper. This was achieved using previous years data which was in csv format and provided as input to Tableau. Tableau's forecasting functionality helped with the prediction process. This prediction can help keep a tab on diabetic patients and diabetes as a disease. With just a few simple clicks and quick easy steps, it was possible to do a prediction in Tableau.

Hence, Tableau is a reliable, practical and straightforward tool for many use cases in the domain of BI as discussed in the paper.

## REFERENCES

- [1] "User Guide Documentation using Tableau [Tableau Community]", Community.tableau.com, 2018. . Available: <https://community.tableau.com/thread/232730>.
- [2] "Data+Science", Dataplusscience.com,2018.Available: <http://www.dataplusscience.com/TableauReferenceGuide/index.html>.
- [3] SearchDataManagement. (2016). What is data analytics (DA)? - Definition from WhatIs.com. [online] Available at: <https://searchdatamanagement.techtarget.com/definition/data-analytics>.
- [4] En.wikipedia.org. (n.d.). Predictive analytics. [online] Available at: [https://en.wikipedia.org/wiki/Predictive\\_analytics](https://en.wikipedia.org/wiki/Predictive_analytics).
- [5] "Tableau Tip: I'll take you to the candy shop. I'll show you how to make a lollipop.", Vizwiz.com, 2018. Available: <http://www.vizwiz.com/2012/07/tableau-tip-ill-take-you-to-candy-shop.html>.
- [6] "Predictive modeling for presumptive diagnosis of type 2 diabetes mellitus based on symptomatic analysis - IEEE Conference Publication", Ieeexplore.ieee.org, 2018. Available: <https://ieeexplore.ieee.org/document/8079667/>.
- [7] "Predictive analysis of diabetic patient data using machine learning and Hadoop - IEEE Conference Publication", Ieeexplore.ieee.org, 2018. Available: <https://ieeexplore.ieee.org/document/8058253/>.
- [8] "Key Indicators of National Family Health survey (NFHS)", Open Government Data (OGD) Platform India, 2018. Available: <https://data.gov.in/catalog/key-indicators-national-family-health-survey-nfhs>.
- [9] "Predicting Diabetes Using a Machine Learning Approach - DZone AI", dzone.com, 2018. [Online]. Available: <https://dzone.com/articles/predicting-diabetes-using-machine-learning-approac>