National Conference on Recent Advances in Computer Science and IT (NCRACIT)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 1 | ISSN : 2456-3307

Data Analytics for Formulation of Effective Public Policy

Moonis Ali[1], Ali Hussain[2], Faisal Rasheed Lone[3]

[123]Department of Computer Science & Engineering, University of Kashmir, North Campus Delina, Baramulla, J&K, India
moonixali@gmail.com[1], alihussainmeer@gmail.com[2], faisal_ln@yahoo.co.in[3]

ABSTRACT

Data Analytics is a subfield of computer science that provides us with vast resources and techniques to visualise, analyze, and derive meaningful insights from data. This is possible due to the availability of large volume of data from various data sources that may vary from sensors to social media platforms. The present paper introduces the subject of Data Analytics and contextualizes various techniques of Data Analytics with respect to the policy making process. An overview has also been presented about the various applications of Data Analytics in various fields and how inferences derived from these cases can be incorporated in policy making to aid in crisis management, health sectors, and in lowering the crime rates. At the end some some suggestions have been made while discussing the future scope of this field.

Keywords: Data Analytics, Data, Policy Making

I. INTRODUCTION

Data analytics is a widely popular and emerging field that can be defined as the "the science of transforming data into useful insights for better decision making". Data analytics employs scientific methods as well as advanced use of information technology techniques that support processing of highly sophisticated data and its subsequent analyses. Technically, the following steps can define a data analytics process: Data sourcing (obtainment of data), Data cleansing (detection & correction of incorrect/inaccurate data), and Data Analysis, Data modelling & visualisation and interpretation of data.

Thus, modern data analytics is often said to be the intersection of advanced statistics, computer science and particular domain expertise.

With ever-increasing reliance on technologies for effective delivery of services and products, many policy makers across the world are turning their attention to the emerging field of Data Analytics. Whereby public policy makers to gain relevant and valuable insights so as the process of formulating policies can use data becomes more efficient, more targeted and ultimately brings positive change in the lives of citizens. Due to the availability of huge amounts of data, the use of appropriate technology becomes very important to process a particular type of data and discover meaningful insights in it.

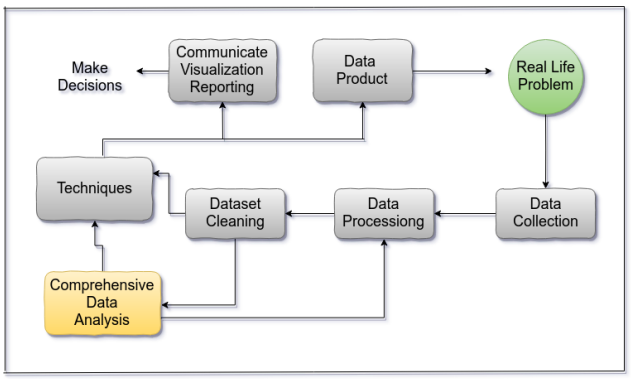

Figure 1. Data analytics based decision-making process.

II. TYPES OF DATA ANALYTICS TECHNIQUES

Analytics techniques assist us in visualising as well as deriving insights from the data. These methods should be able to adapt to increasingly convoluted problems and at the same time be adept to find solutions for them. Generally, these data analytics techniques (approaches) are categorised into the following four types:

- Descriptive analytics
- Diagnostic analytics
- Predictive analytics
- Prescriptive analytics

A. Descriptive Analytics

This technique uses business intelligence and basic statistics to make sense of the previous (historic) data. A method like this describes the past using aggregated or detailed data. Charts, tables and graphs are used as an aid to comprehension.

B. Diagnostic Analytics

Analysis of different types of phenomenon that occur in the data from different perspectives by making use of data mining techniques of correlation to understand why/how things have occurred. Here, the contextualisation of facts and errors or differences is done. Various visualisation methods are used to pinpoint variances, occurence of outliers and changing trends.

C. Predictive Analytics

Statistical techniques and models are used to forecast future trends.This method takes help of mathematical calculations to predict future trends and events by making use of historical data, patterns in the data and gives an estimation of the likelihood of their accuracy.

D. Prescriptive Analytics

Here various optimisation algorithms and simulations are performed to decide the future course of action. Prescriptive analytics is responsible for effectuation of predictive models into actions and decisions.

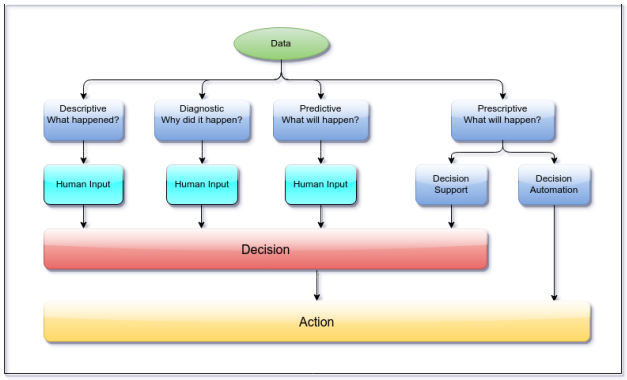

Figure 2. Overview of Different Data Analytics Techniques.

III. VARIOUS DATA SOURCES

Depending on the type and origin of data, various sources of Data can be categorised mainly into three categories viz

- Social Networks (human-sourced information)
- Traditional Business processes
- Machine-sourced Data and Internet of Things (IoT)

A. Social Networks (human-sourced information)

Well documented record of various human experiences in a digitised format in form of text data, photographs, videos and animations. This data is termed as "human sourced data" and is ubiquitous from personal devices to social media platforms. The chief property of this data is that it is ungoverned and structured loosely.

B. Traditional Business processes

Processing of records, monitoring of transactional events, like registration of customers, manufacturing of goods, taking orders, etc. This form of data sourcing generally includes conventional business data (in both operational as well as Business Intelligence systems). This data is structured, referenced and includes details about transactions, tables and correlations. The data is stored in relational databases and may include data generated by commercial transactions, e-commerce, stock trading/banking and public organisations and agencies.

C. Machine-sourced Data and Internet of Things (IoT)

Data gathered by machines has become ever increasing in recent years owing to the milestones in the evolution of mobile technologies and due to the exponential growth in fabrication of sensors. These sensors record events and measure parameters in the physical environment. The data generated by machines is streamlined, properly structured, and often in real time. With the growth of sensors, this data has become a very important part of the information eco-system and is used by various organisations, as it is very suitable for data processing. Machine generated data includes data from sensors, both permanent (e.g. smart home automation systems, smart weather sensors, pollution sensors, traffic sensors, surveillance systems) and mobile data, and also data obtained from computer systems. [1]

IV. DATA ANALYTICS AND POLICY MAKING

The power to convert conventional information sources such as text, images, and transactional records into meaningful insights has opened the door for policy makers to derive data-driven inferences for various policy issues and allow quantitative analysis to penetrate the policy process more deeply than ever before. This advancement in the realm of technology has generated the opportunity or/and necessity for a more complex, more sophisticated and technology-driven approach to transforming data into policy action that goes beyond traditional empirical research [2].

V. IMPLEMENTATION OF DATA ANALYTICS IN PUBLIC POLICY

A. Healthcare:

Among vast applications of Data Analytics in Medical Sciences its main application is in making predictions related to life expectancy, birth-rate, infant mortality, along with forecasts of congenital diseases and occurrence of different types of cancers. The data sources may include the previous medical record of the patient and data of a large number of populace which has already been mined. Some other factors including the lifestyle, income, data from phones, apps and other personal sensors of the patient can also play an important role. With the copious amount of data available , Data Analytics can also help in discovering new drugs by determining the chemical compounds that can work together. Therefore we can simulate the experimentation of millions of compounds efficiently in less time. IBM has collaborated with Teva Pharmaceuticals to discover new treatment options for central nervous system and respiratory diseases using Machine Learning algorithms such as predictive and visual analytics that run on IBM Watson Health Cloud [3]. Apart from suggesting the treatment choices, Data Analytics also helps in detecting the diseases at the early stages, and can therefore aid the overall healthcare system if implemented properly in a healthcare policy.

B. Crime Prevention

Data related to crimes and criminals constitutes a large amount of structured data that is available to governments across the world. This tremendous volume of data can be analyzed to study the trends and patterns of frequently occurring crimes and incidents. Data Analytics can empower governments, law enforcement agencies to monitor illegal activities, crime hotspots and lower the crime rate to a greater extent. Use of insightful data can also shed light on the existing policies of policing and if those policies

are effective in reducing criminal footprints from neighbourhood and dealing with gang related violence. With the application of Data driven systems in crime analysis and prediction, Data Analytics offers a great opening in solving and detecting crime patterns across various domains.

C. Disaster/ Crisis Management

Governments in the present era are incorporating intelligent data driven policies that can help their citizen in the time of crisis or disaster.With the advancement of computational ability and data analytics techniques we are able to analyse big data and predict disasters. Disasters such as epidemic outbreak can be monitored and predicted from the data which is collected from various sources such as satellites, historical data, updates from social media. Machine learning algorithms like SVM (Support Vector Machines) and ANN(Artificial Neural networks) have been used .For example, in prediction of Malaria outbreaks, taking into account data such as temperature, average monthly rainfall, total number of positive cases and other data points [4]. Use of platforms like ProMED-mail which a real-time reporting program for keeping an eye on emerging diseases and providing the real-time outbreak analysis report [5]. Such web based data analysis programs can help the government in deploying the adequate amount of aid to the places which are at higher risk. Similarly there is another platform known as AIDR (Artificial Intelligence for Disaster Response), which is designed to perform automatic classification of crisis-related microblog communications. The objective of AIDR is to classify messages that people post during disasters into a set of user-defined categories of information (e.g., “needs”, “damage”, etc.) For this purpose, the system continuously ingests data from Twitter, processes it (i.e., using machine learning classification techniques) and leverages human-participation (through crowdsourcing) in real-time” [6]. These tools can be hugely beneficial during the time of disaster and can be collectively incorporated to form smart data driven systems for disaster management programs and policies.

VI.CONCLUSION

Big data and Data analytics fit perfectly into several steps of the policy making process and add extensive value to it. Especially when the conventional methods of policymaking shift toward a more data driven approach. The use of data analytics can spawn new growth in the area of active governance and help in formulation of effective public policies, which can ultimately improve the lives of citizens. The emergence of new Data Analytics techniques along with the evolution of new data sources can provide a huge boost to government’s world over to form targeted policies tailored for specific needs of areas and population segments thus improving the overall systems of governance.

VII. REFERENCES

[1] Big data analytics for policy making Report: A study prepared for the European Commission DG INFORMATICS (DG DIGIT)- 2016-2017

[2] https://mag.uchicago.edu/university-news/data-science-meets-public-policy

[3] https://www.ibm.com/blogs/research/2017/04/machine-learning-models-drug-discovery/

[4] Malaria Outbreak Prediction Model Using Machine Learning Vijeta Sharma1 ,Ajai Kumar2 ,Lakshmi Panat3 ,Dr. Ganesh Karajkhede4 ,Anuradha lele5

[5] http://www.promedmail.org/aboutus/

[6] M. Imran, C. Castillo, J. Lucas, P. Meier, and S. Vieweg, “AIDR: Artificial intelligence for disaster response,”