

Data Science and its future

Tirth Raj Kumar¹, Vikash Kumar Sony², Amit Kumar Sharma³

^{1,2,3}Dept. of IT, C.V. Raman College of Engineering, Bhubaneswar
stirthraj.12345@gmail.com¹, vikashkrsoni007@gmail.com², amitsharmanhes@gmail.com³

Abstract: Two type of data, structured and unstructured that are come from more than single branch. Similar to data mining, data science is part that uses scientific approaches, algorithms, system and process to retrieve knowledge form data. Data science is combination of multiple tool, algorithm, statistics, and machine learning principle through aim to find hidden patterns from collected raw data.

Data science is intersection of computer science, domain expertise and mathematics. It is inter related to statistical research, data processing and machine learning.

In data science, a lot of attention courtesy and promising to useful prediction from a vast amount of data. Statistical, computational and human are viewpoint of data science. This paper focuses on Data Science and its future.

Index terms: Data Science; Static Science

I. INTRODUCTION

Data is a raw material, which makes no sense without elaborating it, simply likes "what is it?". Data science combines more than one field of study. Whose goal is to extract value from data in all its forms. Data science is a process. This does not mean that it is mechanical and devoid of creativity. However, when the data processing steps are deepened, from the reduction of data sources and data cleaning to machine learning and finally to visualization, as shown in the following figure, the raw data goes through a single process .

Whenever we gather a large amount of information from different sources and therefore we need analysis of preparation, visualization, administration and conservation, this term refers to the science of data. Data science comes from the help of methods such as data search, data consultation, real-world application, scientific method, visualization, statistical modelling, statistical computation and data technology. Data science is useful for business, marketing and customer satisfaction.

II. HISTORY

The term "data science" has appeared in various contexts over the last thirty years, but it has not become a fixed term until recently. In a first use, Peter Nauru used it as a substitute for information technology in 1960. Naur then introduced the term "datalogy". In 1974, Naur published Concise Survey of Computer Methods, which freely

used the term data science in his study of contemporary methods of data processing that are used in a wide range of applications.

1962 John W. Tukey writes in "The Future of Data Analysis":

1974 Peter Naur publishes a Concise Survey of Computer Methods in Sweden and the United States.

1977 The International Statistical Computing Association (IASC) is established as a Section of the ISI. "The IASC's mission is to link traditional statistical methodology, modern information technology and expert domain knowledge to convert data into information and knowledge."

1989 Gregory Piatetsky-Shapiro organizes and presides over the first Knowledge Discovery workshop in the Database (KDD). In 1995 the annual conference ACM SIGKDD on Data Mining (KDD) and Knowledge Discovery.

III. CONCEPT OF DATA SCIENCE

It simple concept to bring together through statistics, logic of related data analysis and machine learning to understand facts of data. Therefore, we use multiple techniques and theories from many field like mathematics, information science, statistics and computer science.

IV. BENEFITS OF DATA SCIENCE

Data science is the benefit for the organization and industry because it help to empowerment and decision-making for products and services about customer needs and necessity. Its benefit to according to industry and company that comes TV series, product and movies that customer interested to watch and buy .So Company gets benefit and success our goals in our field. Like Netflix comes video that based on the viewer history. Similarly, company analyze this type of trends in social activity.

V. PROCESS

Around the world, we collect raw data from the area of decision-making that data processes and processes to clean up the data set that it did not need any more after the data went to the exploration data analysis. At the end of the analysis, go to the following process models and algorithms. Data arrive directly through the clean Dataset. In the end, we received the report that makes the decision.

VI. ALGORITHMS

It is a process or set of rules or set up to complete a task. It is one of the main concepts or basic components of computing: the basis of elegant and efficient code design, data processing and preparation, and software engineering.

In the science of data, three algorithms are mainly used:

- Data preparation, munging and process algorithms
- Optimization algorithms for parameter estimation including the gradient of the stochastic gradient, the least squares, the Newton method
- Machine learning algorithms

VII. STATIC SCIENCE OF DATA

Data science, Sigma Six, analysis, business intelligence, are all different sides of the same multi-side polygon. Everyone has different tools, vocabularies, projects and certifications.

However, they are all at the service of the company to reduce costs and increase revenues. The practical tools help companies be more effective in what they do!

Furthermore, each comes with its own set of new management practices and leadership needs to understand the real value that can be achieved by using these tools correctly.

VIII. HOW DATA SCIENCE WORKS

Data scientist work in field that have highly educated, and depth information of analytical tools like-R Programming.

In data science need technical knowledge python, Hadoop Platform, SQL database, apache spark machine learning, AI and data visualization. In non-technical knowledge intellectual curiosity, business acumen, communication skills and teamwork.

Data scientist collect all data, that related to your output. And Extract data that are help for output decision.

After that data mining process start, remove unwanted data and meaningless data from datasets. Make a pattern

and graph for understanding relation between data elements. Then come to every member, it all have a different pattern from different condition and situations. Summarize the data and extract best result from this patterns. So that result uses in business tactics.

IX. THE ROLE OF DATA SCIENCE TODAY

Data science has helped business sector to enter this scenario intelligent technology. Companies use big investment in right product and right place using data science. Banking Departments are use big data for improve success in detecting fraud and scam.

In this time you go for loan then bank demand your account activity after the verifying pass your loan. Your activity predicts about monthly income and earnings.

Companies like Netflix extract large volumes of data of user's watched history. That help to make interested TV programs and companies produce and host. Companies use algorithms of related product and videos that recommended by your view history. In this time amazon and flipkart even recommend products of related to search and watched. In YouTube recommend videos based on your watch history.

REFERENCES:

- [1] www.wikipedia.com
- [2] <http://bigdata-madesimple.com>
- [3] <https://searchenterpriseai.techtarget.com>
- [4] <http://businessoverbroadway.com>
- [5] B. K. Mishra, D. Hazra, K. Tarannum and M. Kumar, "Business Intelligence using Data Mining techniques and Business Analytics," 2016 International Conference System Modeling & Advancement in Research Trends (SMART), Moradabad, 2016, pp. 84-89, ISBN: 978-1-5090-3543-4, doi: 10.1109/SYSMART.2016.7894496, Publisher: IEEE

