A SURVEY PAPER ON APPLICATIONS OF DATA SCIENCE IN VARIOUS AREAS

Satwik P M[1] , Manjusha SreeKumar[2]
[1.] Assistant Professor, [2.]Assistant Professor
Department of Computer Applications, T John College, Bangalore, India

ABSTRACT:
Data Science (DS) stores up facts for reference or analysis. on diverse types of work in bits of instruction and decode facts for checking major association. BD is the examination of utilization of facts which is so gigantic and complex that standard application programs is to a marvelous degree bewildered to guide them. It collects instructions and data from multi solicitations to store up them and promotes the data. DS has helped the budgetary business into the forefront especially taught time. Associations utilize gigantic data to pass on a motivation to its patron. Financial sectors are using DS to reduce the risk of deceit. With the consistency of broad amount of information and the reduced cost of securing it, associations have seen the potential of saddling facts to succeed. In any business, promotional marketing plays a significant role and DM processes have extensive applications in tasks like online promoting and cross-selling. They apply the concept of DM throughout their organizational associations, from supply chains to marketing activities to extracting
.

Key Words: Data Science (DS), Big Data (BD), Data Mining (DM).

INTRODUCTION:

The term Data Science has emerged only recently to specifically designate a new profession that is expected to make sense of the vast stores of big data. But making sense of data has a long history and has been discussed by scientists, statisticians, librarians, computer scientists and others for years. Big data is defined as the study of applications of data which are so intense and complex that traditional data-processing application software are very complicated to deal with them.
The term data science has been popular around for the past last 30 years and was originally used as a substitute for “computer science” in 1960. After 15 years later the term data science was used to define the survey of data processing methods that are used in different applications. In 2001, data science was introduced as an independent discipline. September 1994, Business week publishes a cover story on “Database marketing” companies are collecting mountains of information about you, crunching it to predict how likely you are to buy a product, and using that knowledge to craft a marketing message precisely calibrated to get you to do so.
An earlier flush of enthusiasm prompted by the spread of checkout scanners in the 1980s ended in widespread disappointment: Many companies were too overwhelmed by the sheer quantity of data to do anything useful with the information. Still, many companies, believe they have no choice but to brave the database-marketing frontier. In 1977, the international association for statistical computing (IASC) is established as a section of the ISI, “it is the mission of the IASC to link traditional statistical methodology, modern computer technology and the knowledge of domain experts in order to convert data into information and knowledge”.

APPLICATIONS OF DATA SCIENCE:

- **EDUCATION:**

Education economists are increasingly using the availability of large datasets in the learning process of students. As every step a student makes in the online world can be traced, analyzed and used, there are plenty of opportunities to improve the learning process of students. First, data analytics techniques can be used to enhance the learning process of the student by providing real-time feedback or by enriching the learning experience.

The latter might take place in adaptive learning paths that provide a tailored learning environment for students. Thanks to the use of data analytics, the learning environment of students can better correspond to the characteristics of students in terms of cognitive abilities. While similar adaptive and differentiated learning is difficult to realize in a physical classroom, it is relatively easy to realize in the online-classroom. Second, we see possibilities in using data analytics to measure the performance of instructors.
Data analytics can be used to support the instructor or teacher. By combining comprehensive data of the student with learning outcomes in terms of student success, drop out, or cognitive skills- of earlier cohorts of students, the learning outcomes of the evaluated student can be predicted. Warned by similar indicators, instructors can pay additional attention to those students who are at risk of lagging behind.
Moreover, the instructor can obtain descriptive analytics from the progress students are making in online courses, or in tools in the electronic learning environment. In addition, the instructor can use data analytics to detect fraud from students in a cost-effective way. Finally, for policy makers it is often unclear how schools use the available resources to 'produce' the outcomes. By combining structured and unstructured data from various sources, data analytics might be an answer for governments that aim to monitor the performance of schools.

- **HEALTHCARE:**

Machine learning, and other data science techniques are used in many ways in healthcare. From image processing that detects abnormalities in the x-rays or MRI's to algorithms that pull from electronic medical records to detect diseases, the risk of disease, or the progression of the disease, the application of the Machine Learning techniques can easily improve both the health care and getting value of data science.
Firstly, in a wholistic view, data science requires considering where in the clinician's workflow a machine learning algorithm should be used and the ultimate value it will provide to the clinician's goals. As you learn the workflow of the clinician's and they learn what inside your algorithm and cannot provide the original problem post may be refined or changed completely. Make sure that the clinician is educated on the limitations of the algorithms, and make sure you are educated on the resources available to the clinician. The healthcare sector receives great benefit from the data science application in medical imaging. There is a lot of research in this area, and one of the major studies in Big data analytics in health care, published in Biomed Research International.
According to the study, popular imaging techniques include magnetic resonant imaging ,x-ray computed tomography,mammography,and so on. Numerous methods are used to tackle the difference the modality resolution and dimension of these images. Many more efficiently and provide the most accurate interpretations. The deep learning-based algorithms increase the diagnostic accuracy by learning from the previous examples and then suggest better treatment solutions. The most popular image processing techniques focus on enhancement, segmentation and diagnosing that allows deep analysis of organ anatomy, and detection of diverse disease conditions.
The most promising applications aimed to detect tumors, artery stenosis, organ delineationetc. Different methods and frameworks contribute to medical imaging in various aspects. Hadoop, a popular analytical framework, employs map reduce to find the optimal parameters for tasks like lung texture classification. It applies Machine learning, methods, support vector machines (SVM), content based medical image indexing and wavelet analysis for solid texture classification.

- **FINANCE:**

In recent years the ability of data science and machine learning to cope with number of principal financial tasks has become an especially important point at issue. Companies want to know more what improvements the technologies bring and how they can reshape their business strategies. Firstly, automating risk management. Risk management is an enormously important areas for financial institutions, responsible for company'ssecurity, trust worthiness and strategic decisions. The approaches to handling risk management have charged significantly over he past years, transforming the nature of finance sector'snever machine learning modes today define the vectors of business development.
There are many origins from which risks can come such as competitors, investors, regulators, or company's customers. Also,risks can differ in importance and potential losses. Therefore, the main steps are identifying prioritizing and monitoring risks which are the perfect tasks for machine learning. With training on the huge

amount of customer data, financial lending and insurance results allegoricalnot only increase the risk scoring models but also enhance cost efficiency and sustainability. Secondly amongst the most important applications of data science and artificial, intelligence (AI) in risk management is identifying the credit worthiness of potential customers. To establish the appropriate credit amount for a customer, companies use machine learning algorithms that can analyze past spending behavior and patterns. This approach is also useful which working with new customers or the ones with a brief credit history.

Althoughdigitalization and automation of risk management process in finance are in the early stages the potential is extremely huge. Financial institutions still need to prepare for this change by automating core financial process, improving analytical skills of the finance team and making strategic technology investments. But as soon as the company starts to move on in this direction, the profit will not make itself wait. Finally, fraud detectionsit's an obligation for financial firms to guarantee the highest level of security to its users. The main challenge for companies is to find a good fraud detecting system with criminals always backing new ways and setting up new traps. Only qualified data scientists can create perfect algorithms for detection and prevention of any anomalies in user behavior or ongoing working process in this diversity of frauds, for instance alerts for unusual financial purchases for a particular user or large cash withdrawals will lead to blocking those actions with the customer confirms them. In the stock market, machine learning tools can identify paillons in trading data that might indicate manipulations and alert staff to investigate.

However, the greatest thing of such algorithms is the ability of self-teaching, becoming more and more effective and intelligent over time.

- **MANUFACTURING:**

In a world where corporate executives have become accustomed to business analytics on demand, manufacturing remains a blind spot. The unique nature of manufacturing data makes it a big challenge to illuminate this blind spot. Senior corporate leaders have a data-rich picture of most areas of their business-finance,sales,supply chain,HR, etc... In all these areas, rich software tools have developed over decades that let executives see the big picture, zoom in on the smallest details and see every level in between. Sophisticatedanalytics tolls let executives flag problems and spot opportunities for improvement.

But manufacturing data is rarely part of the picture. According to a recent survey by LNS Research sponsored by Sight Machine only 14 % of the respondents have a corporate analytics program that users manufacturing data .

A select group of manufacturing leader have begun taking on these challenges. From our partners, we see an increasing urgency to glean corporate – level insights from manufacturing data. Some companies suing BI SUSTEMS Teradata, IBM Analytics of Tableau are trying to build data pipelines from individual machines all the way up to the corporate level.

At the corperator line of business level, they may be looking at the metrics likes cross plant AKPISs to identify nest practices that could be shared between pants or comparing the performance of contract manufactures. The plant level wants an accurate and immediate assessment of information like the day's yield and scrap rates and may look at metrics like overall equipment effectives (OEE).The manufacturing engineers want to know the causes of machine downtime, so they can improve machine performance.

Getting accurate answers from the factory floor is generally a daunting process. The characteristics of manufacturing data – volume, variety and velocity – make it a real challenge to consolidate , analyze, and take advantage of even on a machine or line level not to mention on a company – wide level.

The companies that figure that out a way to systematically collect, condition model and analyze their data in a scalable, repeatable manner will secure a strong advantage in their industries. The rest will struggle to compete wearing blinders.

CONCLUSION:

Data Science as the name suggests it is a broad field that refers to collective processes, scientific methods, algorithms for extraction of valuable knowledge and insights from raw data. Data science plays a very important role in improving and development in business and other fields. A data scientist who has the precise knowledge of this field can play a major vital role in a company as they are trained to avoid and manage risks, and for delivering relevant products, by the implementation of data science we can empower management to make better decisions, directing actions based on trends, identifying opportunities, decision making with quantifiable, data driven evidence and testing these decisions, and also help with the identification of key groups in business via precision and analysis of disparate sources of data helping the product margins to flourish, hence this plays a very big role as it has many applications in education,gaming, healthcare, finance and manufacturing and many other sectors as Data Science can add value to any business who can use their data well. From statistics and insights across workflows and hiring new candidates, to helping senior staff make better informed decisions. By this data science is valuable to any company in any industry.

REFRENCES:

REFERENCES

[1] McFarlane C. Patientory: A health care peer-to-peer EMR storage network v1.0. Available at: https://patientory.com/patientory_whitepaper.pdf.

[2] Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin CT. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: patient and physician experiences. J Am Med Inform Assoc. 2004; 11:410–17.

[3] Interoperability with Blockchain Deterministic Methods for Connecting Patient Data to Uniform Patient Identifiers.

[4] State of Blockchain and Artificial Intelligence in Fintech http://news.crowdvalley.com/news/state-of-blockchain-and-artificial-intelligence-ai-in-fintech. Accessed 8/4/2018

[5] M. Swan, "Blockchains as an Equality Technology," *Broader Perspective* blog, 2015.

[6] Vitalik Buterin. A next-generation smart contract and decentralized application platform. White Paper, 2014

[7] Ashish K Jha, David Doolan, Daniel Grandt, Tim Scott, and David W Bates. The use of health information technology in seven nations. International Journal of Medical Informatics, 77(12):848{854, 2008.

[8] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., et al. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. Jama, 316(22), 2402.

[9] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancerwith deep neural networks. Nature, 542(7639), 115– 118.

[10] Translating Artificial Intelligence into Clinical Care, Andrew L. Beam, Isaac S. Kohane, JAMA 316, 2368, 2016

[11] Opportunities and Obstacles for Deep Learning in Biology and Medicine, CS Greene et al., bioRxiv preprint first posted online May. 28, 2017;

[12] Peterson, K., Deeduvanu, R., Kanjamala, P., & Boles, K. (2016). A Blockchain-Based Approach to Health Information Exchange Networks.

[13] Ekblaw, A., Azaria, A., Halamka, J. D., & Lippman, A. (2016, August). A Case Study for Blockchain in Healthcare: "MedRec" prototype for electronic health records and medical research data. In Proceedings of IEEE Open & Big Data Conference.

[14]]Travers Ching, Daniel S. Himmelstein, Brett K. Beaulieu-Jones,et al, Opportunities and obstacles for deep learning in biology and medicine,