EAS 504: Applications of Data Science – Industrial Overview – Spring 2023

-Lecture by Sriganesh Madhvanath, eBay (formerly Conduent)

Name: Tananki Ranga Sai Saran Rohit

UB ID: 50441793

UB Email: rangasai@buffalo.edu

Ques 1 : Describe the market sector or sub-space covered in this lecture :

The market sector or sub-space covered in this lecture are how Business Process Services (BPS) operate and the need for analytics in this sector. BPS focuses to automate key operational activities by combining advanced algorithms and knowledgeable workers. BPS is a process where a firm works with someone who has significant experience in a certain area of our industry, who are referred as strategic partners. BPS helps in adding significant value , or direct business impact ,beyond the bottom line. Over the years, the industry has evolved significantly, and moulded itself from a service provider to a strategic partner for future driving growth. This has been majorly achieved due to rise in number of outsourcing models that the industry now offers. BPS industry also helps their clients to achieve long term strategic goals and also offers customizations of business models. In BPaaS (BPO as a Stack),there is more flexibility offered for the client to adapt best to their business environment with latest cloud based industry solutions , intelligent automation, analytics and also deployment of plug and play services .The end goal is to drive savings, increase business value by using the latest technology like platform based solutions, automation by using Artificial Intelligence and robotics which would result not only in cost arbitrage but also intellectual arbitrage. This would help both the organizations to develop long term strategic and symbiotic partnership which results to grow and expand their businesses .The global third party BPS industry in today's market has more than 200 service providers and is valued at more than 150 billion USD .

Ques 2 : What data science related skills and technologies are commonly used in this sector?

Companies, government agencies, universities and third parties who are interested to solve their business needs ,they develop state of art solutions that typically integrate descriptive, predictive and prescriptive elements. Data science also helps when dealing with large and complex data in various forms (audio,text,video etc.) and multidisciplinary approach. Also, the application of data to business problems raised by the customers may be informed by that data when data analytics is used for services. In Business Process Services BPS, data science is used to answer business problems and aid in the making of right business decisions. Few data science skills and technologies which are widely used are Machine Learning,Dyanamic Control and Systems, Statistical Modeling, Data Mining, Bigdata technologies (Hadoop, Apache Spark), Spatio-temporal modeling, Operations Research and Optimization (includes cost reduction and operational efficiency), Graph Algorithms, Text Analytics, Multimedia Analysis, Behavioral economics ,User-Centered Design, Ethnography: Understanding user behavior and expectations, Domain expertise. All these data science skills and technologies mentioned above necessitates a multi-disciplinary team helps in solving business challenges in this sector .

Ques 3 : How are data and computing related methods used in typical workflows in this sector? Illustrate with an example.

The applications of information data and computing related techniques are best explained using an example such as credit risk assessment. We would begin by deciding which features we want to use to predict the risk of issuance of credit card .The issuance of new credit card will be based on various features like salaried/business, spending history , cibil score, Age of Customer, Education level of customer, Tenure with current employer (in years),Number of years in same address, Customer Income, Debt to income ratio, Credit to Debt ratio, prompt payment history,etc…Once after collecting all the necessary data, we perform data cleaning, pre-process the data, apply feature reduction and various scaling techniques ,also deal with missing and null values to make data look consistent and then apply a suitable model to get accurate business decisions (i.e. whether to approve or reject issuance of credit card to specific individual).

Ques 4 : What are the data science related challenges one might encounter in this domain?

Some of the data science challenges that can be encountered are data variability - different types of data , quantity - data size, speed - processing speed where data can often be complex and difficult to process and data can contain missing values, which are some of the obstacles that can be encountered in BPS. The first step of any data science project is finding and collecting necessary data assets. However, the availability of suitable data is still one of the most common challenges that organizations and data scientists face as many times data needs to be pulled from different servers and different systems which usually have different formats , and consolidating data from lots of disparate and semi-structured sources is a complex process and

this directly impacts their ability to build robust ML models.Data prearation and cleaning often helps ,but then data scientists spend a lot of time processing and preparing data so that it's consistent and structured enough to be analyzed. Also most of the time, data consistency and data authenticity can be a challenge as data may not always be available in the correct format or may be stored in outdated systems.

Ques 5 : What do you find interesting about the nature of data science opportunities in this domain?

Increasing demand in the use of data is exponentially growing from past few years. For instance, the amount of data in 2010, which was being created every two days, and in 2021 it was being created just for every 40 minutes. The amount of data ,which is being created by applications, sensors ,electronic devices, virtual assistants , smart phones, self driving cars , AI devices etc… had become massive and the need for use of this data to enhance customer experience, take business decisions lead to huge demand in data science and analytics field. As we all know that Data science is the study of data to extract meaningful insights for business. It is a multidisciplinary approach that combines principles and practices from the fields of mathematics, statistics, artificial intelligence, and computer engineering to analyze large amounts of data.i found it very interesting because it is widely used across different verticals and various industries like Finance, Healthcare, Travel industry, Energy ,Manufacturing, Gaming, Education etc…

(i) Please comment on the BPO vs BPS vs BPaaS paradigms and the increasing role of Data Science in the BPS domain.

Business process outsourcing (BPO):

Business process outsourcing (BPO) stands for Business Process Outsourcing and the industry provides support to other companies. As the name suggests, it is an opportunity to outsource tasks to other companies to make their processes more efficient. (BPO) is the process by which a company outsources front- and back-office non-business activities to other organizations with the aim of saving money and time. The business services organization would manage the business according to the instructions of the outsourcing company.

Business process services (BPS):

Business process services are services that are outsourced to organizations that specialize in outsourcing. Some business processes that are not an organization's core competency can be outsourced to companies that specialize in those functions. This gives these companies greater

flexibility in managing both finances and operations. Business Process Services (BPS) services are provided to organizations needed by companies with expertise in one or more business areas. These business activities may include handling procurement ,health records ,financial services, records management,inventory, insurance, human resources, etc. ADP, Conduent, Accenture, Teleperformance, Xerox, Convergys,Paychex, and other major BPS companies offer various models from which the customer can choose according to their needs.

Business Process as a Service (BPaaS):

Business Process as a Service (BPaaS) as a provision of cloud-sourced business process outsourcing services designed for multi-tenancy. Services are often automated, and where human process operators are required, there is a clear lack of a dedicated customer-facing workforce. Business Process as a Service (BPaaS) is a cloud-based business process solution that can be used by customers. These solutions provide HR, infrastructure and applications in addition to business process functions. One of the main advantages of BPaaS is that companies do not have to worry about infrastructure, equipment or maintenance because these things are managed by the organization. They could concentrate on their business apart from these worries. As a result, overall costs and expenses are reduced; Additional benefits include greater efficiency, business value, and flexibility in business operations. Cognizant, Accenture, Wipro, and others are examples of companies that offer BPaaS.

Increasing role of Data Science in the BPS domain:

BPS companies like Accenture, Cognizant ,eBay biggest asset is the amount of data they have access to. eBay, for example, has data on the buying behavior of millions and millions of buyers, which it can use to target specific customers. So, with so much information at a company's disposal, the role of data science becomes really important. There is huge untapped potential and huge price competition in this industry to compete against competitors as the industry now offers a model based on performance and profit sharing. The ultimate goal is to save costs, and the job of Data Science is to provide the basics of how to do that in this age of big data. A better business process can always provide a competitive advantage in terms of (a) speed and accuracy of decision-making and (b) process analytics. By comparing other business practices with your own business processes, you can identify inefficiencies that can be improved. This clearly shows the increasing demand of Data Science in the BPS domain.

(ii) Pick two of the case studies from the lecture to discuss how different data science techniques are used to solve these problems.

Case Study 1 - CUrb(Customer lssue Analysis in the Urban Context):

CUrb aims to enable public transport agencies to listen to their customers' concerns through formal and informal channels such as social media (facebook,twitter,apps,customer care) and visualize and act on them in an automated and integrated way to enhance customer experience. This basically involves identifying and classifying urban traffic customer problems and analyze them over time opinion mining to determine problem severity interactive visualization of problems in the form of heat maps showing problem intensity and occurrence tracking problems over time to extract relevant patterns intelligent routing of composition query data to concerned agency departments. spotting and categorization of customer issues in urban transportation, and analysis of their time opinion mining to determine severity of issues (for suppose problems with bus breakdown, problems with potholes, rude behavior of driver, delay in bus schedule , delay in bus arrival ,etc..) . lot of people doesn't call the transport department instead post it on social media platform like Facebook or twitter .So, the Transport department is interested to listen to multiple channels and they need the severity levels to be categorized too. Classifiers and opinion mining as well are used in this use case . Interactive visualization ,Routing is also used in this use case .Stanford NL tool kit is used for Natural language processing .Also suggestion for new services of bus patterns are extracted in this example.

Case Study 2 - Public Transportation :

The customers of these services can be public transport agencies, and the main purpose of using data science from the point of view of the transport company is to increase revenue, reduce congestion, increase the system of all lines, increase user experience. But if you look at the perspective of passengers, you had to look at other characteristics, for example, the total travel time - which includes the arrival time of commuters ,walking time to the stop, waiting time, the travel time and buffer time at times. Efficiency is about minimizing total time. The parameters used are overall startup time, reliability, closure level connectivity. The data used by are - Stops and routes, service schedule and schedule compliance at time points. The data is also categorized into less than 10 minutes, less than 20 minutes and less than 30 minutes .also formed other cluster of core network bus stops with actual services and scheduled services .here in this use case, we are applying Network connectivity. This information is used to measure the connection area covered by the transport during a certain time period, such as 30 minutes. The main contribution of the single number to capture user centric connectivity. As a result Connectivity measure from travelers' perspective , they were able to get well connected network, reliable service schedule and on time services .For the transport authority, they were able to extract Spatial model to predict total number of transit riders per TAZ (Transportation Analysis Zone),Significant variables per TAZ: Population, Number of workers, Average network connectivity, Extended notions of connectivity, Area covered by core network for scaling the transportation in future and also to develop reliable transport network.