Data Analytics for Business Process Services

Introduction

The lecture discusses "Data Analytics for Business Process Services" and how it has evolved the Business Process Services (BPS) sector from cost optimization to a data-driven strategic domain. It explores data-driven decision-making across BPS and the public sector, showcasing the integration of data and computing methods into business processes. Despite challenges like data privacy and quality, these hurdles stimulate innovation within the sector. Data scientists in BPS play a pivotal role at the intersection of technology, strategy, and social impact, offering the potential to revolutionize public services and enhance resource allocation efficiency. Two case studies illustrate diverse data science techniques, from customer churn prediction to improving customer care through analytics dashboards. This report essentially captures the transformative journey of data science in BPS, from addressing challenges to harnessing opportunities for meaningful change.

Base Questions

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

The lecture primarily centres on the services industry, specifically Business Process Services (BPS), and delves into the integration of data analytics within this domain. Within the business landscape, organizations comprise both core and non-core business processes. Business Process Outsourcing (BPO) is the practice of outsourcing specific non-core activities, aiming to reduce costs. This practice is encapsulated under the term "BPS" or Business Process Services. BPS encompasses a wide spectrum of services, encompassing critical functions such as customer support, finance and accounting, human resources, and supply chain management.

Outsourcing goes beyond cost-cutting; it leverages external expertise, improves operational efficiency, and enhances overall business value. The sector now emphasizes data-driven automation across the entire value chain, evolving from its cost optimization origins in non-core and back- office operations. During the lecture, an additional topic discussed was Business Process as a Service (BPaaS). This paradigm allows clients and companies to opt for cloud-based industry solutions. Such solutions empower organizations to adapt swiftly to the dynamic business environment by deploying plug-and-play solutions, incorporating intelligent automation, and employing analytics [1].

In the contemporary landscape, the objective of service providers has shifted towards forging long-term strategic partnerships that foster growth and expansion for both organizations involved [1].

2. What data science-related skills and technologies are commonly used in the public sector?

The public sector uses a variety of data science-related skills and technologies to solve complex problems and make better decisions. Some of the most common skills and technologies include:

- Data mining and machine learning: These techniques are used to extract patterns and insights from large datasets.

- Statistical modeling: This is used to develop mathematical models that can be used to predict future outcomes.
- Spatio-temporal modeling: This is used to analyze data that is collected over time and space.
- Operations research and optimization: This is used to solve problems that involve optimizing resources.
- Graph algorithms: These are used to analyze and process data that is represented as graphs.
- Text analytics: This is used to extract insights from text data.
- Multimedia analysis: This is used to extract insights from multimedia data such as images, videos, and audio.
- Big data technologies: These are used to store, process, and analyze large datasets.
- User-centered design: This is used to design solutions that are user-friendly and effective.
- Domain expertise: This is knowledge of the specific domain that is being studied.

Data scientists in the public sector require versatile skills and technologies, with the specific ones chosen based on the problem at hand. Common skills encompass technical areas such as data mining, machine learning, modeling, and more. Equally crucial are effective communication, collaboration, and adaptability skills, allowing data scientists to convey insights to diverse audiences and work collaboratively across disciplines. While applications vary across agencies and organizations, these skills and technologies remain prevalent in the diverse landscape of the public sector.

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

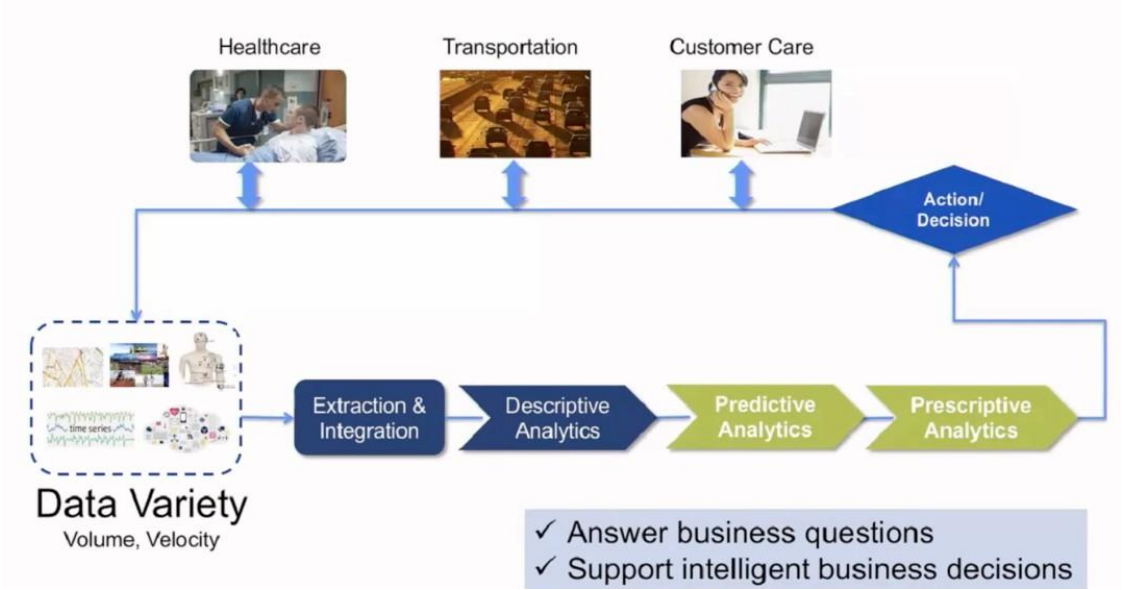

Fig.1 The work-flow of data and computing methods in the BPO domain

In the Business Process Services (BPS) sector, data and computing methods drive the workflow:

1. **Business Domain and Data Collection** : Identify the business domain and collect data from sources like healthcare or customer care, handling large volumes of data generated rapidly.

2. **Extraction and Integration** : Extract data from diverse sources, including sensors and cloud systems, and integrate them into a unified system.

3. **Descriptive Analytics** : Analyze data to understand "What happened?" Clean and explore data for historical trend insights.

4. **Diagnostic Analytics** : Shift focus to "Why it happened?" Identify patterns and root causes to address underlying issues.

5. **Predictive Analytics** : Address "What will happen?" by using statistical and machine learning techniques to forecast trends and potential issues.

6. **Prescriptive Analytics** : Determine "What should I do?" by providing actionable recommendations and strategies based on insights.

7. **Final Decision and Action** : Conclude with informed decision-making and action based on insights and recommendations.

Example : As discussed in the lecture, the computing and data workflow involves the collection of data in various forms, including video recordings and text complaints from sources such as official websites, complaint forums, and platforms like Twitter. Additionally, data can be obtained from sensors and other sources based on specific business requirements. This diverse dataset is then stored and integrated for analysis, which includes exploratory data analysis (EDA).

Exploratory data analysis allows us to gain insights from the dataset through descriptive analysis, visualization, and pattern recognition. It helps us understand frequent issues people are facing and how issue resolution processes are being carried out. Moving beyond descriptive analysis, diagnostic analysis helps uncover the reasons behind patterns observed in the data. For instance, it might reveal that customers are experiencing network issues with their cellular devices, but this problem is specific to a particular region. Predictive analysis takes our understanding a step further by forecasting when similar issues might occur again. In certain domains, such as medical diagnosis and disease control, prescriptive analysis becomes crucial. It helps us formulate preventive measures to avoid the reoccurrence of problems or to enhance patient health. Finally, based on the outcomes of these analyses and the specific business needs or patient health conditions, a final decision is made. This decision could range from improving customer service processes to taking critical actions in medical diagnosis and treatment.

The data workflow involves collecting and integrating data from diverse sources, performing exploratory data analysis to gain insights, and then progressing to diagnostic, predictive, and prescriptive analysis to make informed decisions based on the data's patterns and trends. The BPS sector employs data and computing methods to follow this structured workflow, facilitating informed decision-making and action to address business challenges and opportunities.

4. **What are the data science-related challenges one might encounter in this domain?**

The Domain of BPO faces a lot of challenges while implementing data science related technologies ranging from privacy of the private user data to the quality and high volumes of data which is detailed below :

Data privacy and security: The public sector is responsible for protecting the privacy and security of sensitive data, such as personal information and financial data. This can be a challenge, as data scientists need to access this data in order to extract insights.

Data cleaning: Data cleaning is a time-consuming and challenging task, and it can take up to 70% of a data scientist's time. This is because the data needs to be formatted, cleansed, and standardized before it can be used for analysis.

Data quality: The data used in the public sector is often of poor quality, which can make it difficult to extract insights and make accurate predictions. This is because the data may be incomplete, inaccurate, or inconsistent.

Lack of domain expertise: Data scientists often lack the domain expertise to understand the nuances of the public sector. This can make it difficult to interpret the data and develop solutions that are relevant to the specific needs of the organization.

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

The BPS domain offers data scientists a unique opportunity to work on challenging and impactful projects. Here are three specific aspects of data science opportunities that I found captivating in this domain:

Privacy and security challenges : In the BPS domain, data scientists face significant privacy and security challenges. They must balance data insights with safeguarding sensitive information. To address this, they employ techniques like encryption, access controls, and data anonymization, which removes personally identifiable information (PII) while retaining data utility. Staying current with evolving privacy laws and regulations is essential for compliance and ensuring people's privacy protection.

Interdisciplinary collaboration : Collaboration is pivotal in BPS data science. Data scientists team up with experts, policymakers, and stakeholders from diverse domains. This collaboration facilitates a deeper understanding of problems and leads to more effective solutions. For instance, when optimizing healthcare, data scientists collaborate with medical professionals to comprehend healthcare system challenges. This joint effort enhances the likelihood of impactful, patient-centric solutions.

Impactful outcomes in public services : The most rewarding aspect of BPS data science is its potential for real-world impact. Data scientists in this field contribute to enhancing public services,

optimizing resource allocation, and streamlining government operations. For instance, they may use data to improve the delivery of social services or to devise crime prevention strategies.

The prospect of positively affecting people's lives serves as a powerful motivator for data scientists in the BPS domain. They are eager to apply their skills to effect meaningful change. This domain offers a unique opportunity for challenging yet fulfilling projects, where creativity, resourcefulness, and a strong commitment to public service are essential traits for success.

Additional Questions

(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) entails hiring a third-party for non-core business functions like customer service or IT support. It can be domestic or international and aims to cut costs or enhance efficiency. Business process services (BPS), a specialized BPO form, focuses on efficient process management. BPS providers use tech and data science to improve processes and often help companies redesign them for efficiency. Business process as a service (BPaaS) is a cloud-based BPS delivery model. BPaaS providers offer pre-configured solutions adaptable to businesses, saving time and costs while increasing agility. Data science is increasingly vital in BPS. Data scientists use skills like data mining, machine learning, and statistical modeling to enhance processes, identify opportunities, and boost operations.

Some of the ways that data science is being used in the BPS domain include:

- Developing recommender systems to personalize product recommendations for customers
- Detecting fraud and abuse in financial transactions
- Improving customer service by automating tasks and providing real-time insights into customer behavior

The use of data science in the BPS domain is still in its early stages, but it has the potential to revolutionize the way that businesses operate. By using data science to improve the efficiency and effectiveness of their business processes, businesses can save money, improve customer service, and gain a competitive edge.

Here are some specific examples of how **data science is being used in the BPS domain:**

- eBay uses data science to recommend products to customers based on their past purchases and browsing history. This helps eBay to increase sales and improve customer satisfaction.
- Amazon uses data science to predict which products are likely to be purchased together. This helps Amazon to optimize its inventory and shipping costs.

- Bank of America uses data science to detect fraud in credit card transactions. This helps Bank of America to protect its customers from financial losses.

These are just a few examples of how data science is being used in the BPS domain. As data science continues to evolve, we can expect to see even more innovative and effective ways to use data to improve the performance of business processes.

(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: Churn Prediction
Objective: Identifying the customer who can leave a service or subscription and retain them.
Techniques: Random forest algorithm, Trees (Decision Trees), transduction sums.
Key Insights: Case uses hyper-graphs and NPS score for deploying retention strategies.

This case study addresses the high cost of customer churn for a particular service provider. The main objectives are to identify customers at risk of churning and develop effective retention strategies. This challenge is divided into three phases: prediction, characterization, and prevention.

The data used in this case study includes three categories: customer profile data, customer network data, and customer service usage history. Customer profile data includes information such as age, gender, nationality, and address. Customer network data captures relationships and interactions between customers. Customer service usage history includes information such as joining dates, churning dates, churn types, returns, and usage patterns. The goal of the prediction phase is to identify common attributes or features among customers who are likely to churn. To achieve this, a hyper-graph representation is used, which connects multiple nodes based on shared attributes such as age, service preferences, and regional affiliations.

The historical data of customers who have previously churned is compared with that of loyal customers to identify distinctive features. This process involves a feedback loop mechanism, which continuously improves the accuracy and effectiveness of the model. In the context of churn prevention, a survey is introduced to gauge customer sentiment, which is quantified using a Net Promoter Score (NPS) calculation. Improving the NPS score is a key focus for the company, as customers who are willing to recommend the service provider are indicative of success. A key insight that emerges from the analysis is the identification of key individuals within the network, known as "number associates," whose departure has the most significant impact. These highly connected individuals play a vital role in the retention strategy, and decisions on retention are made based on the level of influence these customers exert within the network. The predictive model used in this case study includes decision tree analysis, random forest algorithms, and transduction sums. The company's approach achieves an impressive 93% accuracy in identifying potential churners. The feedback loop continually fine-tunes the model, enabling seamless scaling of the implementation and early detection of churn.

In summary, this case study addresses the critical questions of who is likely to churn, why they might do so, and which customers to target for retention. It leverages the NPS score and network connectivity to inform retention strategies, ultimately leading to a highly effective predictive model with scalable implementation and a focus on early detection.

Case Study 2: Customer Care — Analytics Dashboards
Objective: To improve the quality of service provided to the customers.
Techniques: Sentimental/Emotion Analysis, Voice analysis from Call recording.
Key Features: Dashboard for supervisors to have a bird-eye view on the process, sentiment analysis using Spark and QArt Technologies, issue analysis on social media.

This case study highlights the importance of customer service quality and outlines various methods and tools, including call recording, voice tone analysis, and sentiment analysis, that can be used to improve it.

One key aspect of this strategy is the creation of a dashboard for supervisors that provides a comprehensive overview of ongoing conversations. The dashboard uses a time-based axis and employs color-coded indicators, where red flags indicate problems and blue signifies satisfactory interactions. Each square on the dashboard represents a conversation, and various issues are flagged for further attention. All agents are required to greet customers, inquire about specific details, and flag instances where these steps are omitted. The dashboard also includes a feature to assess whether the conversation is engaging or not.

Furthermore, the dashboard incorporates sentiment analysis, emotional analysis, and tracks trending topics related to customer queries. It is built using Spark and leverages QArt Technologies, which integrates multiple cutting-edge Natural Language Processing and Machine Learning techniques. These techniques include emotion and sentiment classification, topic identification, chat resolution, compliance violation detection, and communication problem detection. Importantly, they are easily configurable and trainable. Another aspect covered in the case study is CUrb (Customer issues Analysis in the Urban Context), which addresses customer issues raised on platforms like Facebook and Twitter. The goal is to capture public sentiment across various channels, involving automatic categorization and opinion mining, as well as the detection of issue severity. Moreover, tracking and routing mechanisms are implemented to direct issues to the appropriate department, such as concerns about potholes being flagged and routed accordingly. The dashboard integrates multiple data sources, analytics modules, and visualizations to provide a holistic view of the situation.

This case study highlights the importance of customer service quality and outlines various methods and tools that can be used to improve it. The use of a comprehensive dashboard, integration with QArt Technologies, and the CUrb approach collectively contribute to enhancing customer service and issue resolution.

Conclusion

The lecture on "Data Analytics for Business Process Services" emphasized the role of data analytics in BPS. BPS encompasses a range of services, adapting through BPaaS and partnerships. In the public sector, data science leverages skills like data mining and machine learning, aided by communication and collaboration. BPS workflows include data collection, descriptive, diagnostic, predictive, and prescriptive analytics, enabling informed decisions. Challenges in BPS involve data privacy, cleaning, quality, and domain understanding, necessitating innovation. Data science in BPS balances privacy and security, fosters collaboration, and drives public service impact. Case studies covered customer churn prediction and improving customer care through analytics dashboards. Data science in BPS is an exciting field that combines innovation and social impact. Challenges offer opportunities, transforming public services through data-driven change.

References

[1] Lecture slides and video.
[2] How Data Science has changed the paradigm of Business Process Services, Shubham Sharma - https://www.linkedin.com/pulse/how-data-science-has-changed-paradigm-business-process-shubham-sharma/
[3] Images — screenshot from lecture videos.