Assignment II: MIS 64038 Analytics in Practice

Case 2: Fraud Detection in Banking

The deployment stage of the analytics project can be linked to the issue the bank is experiencing. It appears that there was a delay in processing ATM withdrawal requests because the algorithm was not designed for real-time processing. The requests timed out as a result, preventing consumers from making withdrawals from their accounts.

Reasons:

1. It's probable that when building the algorithm, the project team did not take the method's performance needs into account.
2. It's also likely that the group did not fully test the algorithm in a real-world setting to guarantee that it would function as intended.
3. Additionally, it's possible that the team disregarded the necessity to optimize the algorithm to match the production environment's performance needs.

Recommendations:

1. Performance requirements to be taken into account during development: When developing the algorithm, the project team should take the algorithm's performance needs into account. The algorithm must be created and optimized to meet the performance requirements of the production environment in order to accomplish this.
2. Test in a real-world setting: Before implementing the algorithm, it is crucial to test it in a real-world setting to make sure it will work as intended. This aids in finding and fixing problems before they have an impact on clients.
3. After the algorithm is put into use, it is crucial to continuously check how it performs in real-world situations in order to spot and address any performance issues as soon as they arise.
4. Optimize the algorithm: To avoid delays in processing ATM withdrawal requests, the algorithm should be optimized for real-time processing. The code, architecture, and hardware must all be optimized in order to boost the algorithm's performance.

Case 3: Amazon Rekognition

Several project elements that may have been disregarded or neglected during development and deployment are to blame for Amazon Rekognition's problem with misidentifying dark-skinned females.

Reasons:

1. Biases in the training data: According to the MIT researchers, lighter-skinned people, especially men, tend to be overrepresented in the datasets used to train Rekognition and other facial recognition systems. The lack of variety in the data used to train these systems can be blamed for this bias. By gathering a more varied collection of photographs that features people from various racial and ethnic backgrounds, genders, ages, and other aspects, this issue may have been solved.
2. Insufficient testing:Prior to the product's launch, it's conceivable that Amazon did not sufficiently examine Rekognition's performance on a wide variety of faces. This could have entailed testing the program on people with various skin tones, levels of facial hair, and other physical traits to make sure it functions the same way in all populations.

3. Lack of openness: The bias found in the MIT study might have been caused by a lack of transparency in the development and training of Rekognition's algorithms. Researchers might have been able to spot and fix these problems earlier had Amazon given more details about how the system functions.

Recommendations:

1. Diversify training data: To ensure that facial recognition systems are accurate and unbiased, it's important to collect a diverse set of images that includes people from different racial and ethnic backgrounds, genders, ages, and other factors. This can help to reduce the risk of biases in the training data and improve the performance of the system across different groups.
2. Conduct rigorous testing: Before releasing a facial recognition system, it's important to test it thoroughly on a diverse range of faces to ensure that it performs equally well across different groups. This can involve testing the system on individuals with varying skin tones, facial hair, and other physical characteristics to identify and address any biases or inaccuracies.
3. Be transparent about algorithms: It's important for companies to be transparent about how their facial recognition algorithms are designed and trained. This can help researchers to identify and address any biases or inaccuracies, and also help to build trust with users and customers.
4. Consider ethical implications: Facial recognition systems can have significant ethical implications, particularly when they are used by law enforcement or other government agencies. Companies should carefully consider the potential consequences of their products and services, and engage in open and transparent discussions about the ethical implications of facial recognition technology.
5. Implement governance policies: Companies should develop governance policies to ensure that their facial recognition systems are developed and deployed responsibly. This can involve establishing guidelines for data collection, testing, and deployment, as well as policies for ensuring transparency and accountability.

Case 4: IBM Watson in Healthcare

It appears that several aspects of the project were ignored or not adequately addressed, leading to the failure of IBM Watson's implementation in the healthcare industry.

Reasons:

1. Planning Phase: The rush to market and aggressive marketing of the product without ensuring that it was competent and capable of delivering the promised results suggests that the planning phase may have been rushed or not given sufficient attention.
2. Data Preparation and Modeling Phase: The use of biased data from a single hospital to train Watson limited the system's ability to provide personalized and effective treatment recommendations for patients outside of that hospital's network. This suggests that the data preparation and modeling phase may have been insufficiently thorough, with too little attention paid to ensuring that the data used to train the system was representative and unbiased.
3. Implementation Phase: The lack of collaboration with hospitals to ensure that Watson was functioning properly and providing accurate recommendations suggests that the implementation phase may have been inadequate, with insufficient attention paid to integrating the system into the healthcare environment and ensuring that it was meeting the needs of clinicians and patients.

Recommendations:

To be successful, analytics projects must carefully consider all phases of the project lifecycle and ensure that each phase is given sufficient attention and resources to achieve the desired outcomes.

Case 5: AI for University Admission

The problem can be associated with the following phases of the analytics project:

1. Discovery Phase: The problem may be associated with the discovery phase, where the researchers may not have adequately considered all of the factors that contribute to success on the entrance exam.
2. Data Preparation Phase: The problem may be associated with the data preparation phase, where the researchers may not have adequately prepared and curated the data used to train the AI system.
3. Modeling and Evaluation Phase: The problem may also be associated with the modeling and evaluation phase, where the researchers may not have adequately tested the AI system on a wide range of questions and scenarios to ensure that it is capable of passing the entrance exam.

Reasons:

1. Domain Knowledge: The researchers may have underestimated the importance of domain knowledge in developing an AI system that can pass the entrance exam for the University of Tokyo. The entrance exam likely tests for a wide range of knowledge and skills, and the researchers may not have adequately considered all of the factors that contribute to success on the exam.
2. Data Quality: The researchers may have also overlooked the importance of data quality in training the AI system. If the system is not trained on high-quality data that accurately reflects the types of questions and content on the exam, it may not be able to perform well on the exam.
3. Complexity: The researchers may have underestimated the complexity of the task of passing the entrance exam for the University of Tokyo. While AI systems are capable of performing complex tasks, such as playing chess or Go, passing a highly selective entrance exam may require a level of reasoning and understanding that is beyond the capabilities of current AI technology.

Recommendations:

1. Domain Expertise: It is important to have experts in the relevant domain on the team to provide input on the requirements and challenges of the task at hand. In the case of the Todai project, it may have been helpful to have educators or examiners from the University of Tokyo involved in the project.
2. Quality Data: To train an AI system effectively, it is important to have high-quality data that accurately reflects the task at hand. In the case of the Todai project, it may have been helpful to collect a wide range of exam questions and answers, as well as related materials such as textbooks, lectures, and notes.
3. Diversify Data: In order to improve the AI's ability to grasp the meaning in a broad spectrum, it may be helpful to diversify the data used to train the system. This could include incorporating materials from a wider range of sources,or using natural language processing to extract relevant information from unstructured data such as news articles or social media posts.
4. Experimentation: It may be helpful to conduct a range of experiments to test the AI system's performance on different types of questions and scenarios. This can help identify areas where the system needs improvement,and inform the development of new strategies and techniques.

5. Collaboration: Collaboration between different experts and organizations can help to bring diverse perspectives and expertise to the project. This can help to identify blind spots and develop more effective solutions.

Case 6: Mars Orbiter

The loss of the Mars orbiter by NASA was primarily due to a mix-up in the units of measurement used by Lockheed Martin's engineering team and NASA's internal team. This issue can be attributed to a failure in the data preparation phase and communication of the analytics project.

Reasons:

1. Communication: The communication gap between Lockheed Martin's engineering team and NASA's internal team was a critical factor that contributed to the loss of the Mars orbiter. The use of different metrics by the two teams resulted in a significant error that had costly consequences. This highlights the importance of effective communication and collaboration among team members, particularly when working on complex projects with high stakes. It is essential to establish clear communication protocols and ensure that all team members are on the same page to prevent costly mistakes like this from occurring in the future.
2. The data preparation phase: involves collecting, cleaning, and formatting data before analysis can begin. In this case, it appears that there was a lack of standardization in the units of measurement used by the two teams, which should have been addressed in the data preparation phase.

Furthermore, the internal review panel at NASA's Jet Propulsion Laboratory noted that the loss was due to an "end-to-end process problem." This suggests that there were likely issues throughout the entire analytics project lifecycle, including data collection, preparation, analysis, and implementation. It's possible that there were errors in the data analysis phase that could have caught the issue with the units of measurement, but it's also possible that the issue was not caught due to a lack of proper checks and balances in the implementation phase.

The review panel also noted that Congressional budget constraints played a role in the error. This suggests that there were issues in the planning and resource allocation phases of the project. The project may have been underfunded, leading to a lack of resources for proper checks and balances throughout the project lifecycle.

Overall, the loss of the Mars orbiter by NASA can be attributed to failures in multiple phases of the analytics project, including data preparation, analysis, implementation, planning, and resource allocation. However, the primary issue appears to be a failure in the data preparation phase due to the lack of standardization in the units of measurement used by the two teams.

Recommendation:

Organizations should standardize measurement units, implement strict checks and balances, invest enough resources, and perform regular evaluations to avoid similar mistakes in future analytics projects.