Analysis of Banking Domain.

Bank use Big Data in order to make more effective decisions. Big data technologies can help banking to focus on their resources efficiently, making smarter decisions and improve the performance. 3Vs (volume, variety and velocity) are three distinguishing features of Big Data. The value of Big Data to Banking Industry is immense. The data challenges faced by banking domain are Privacy and Security. Big Data provides major steps for banks with privacy and security but it comes with a large red flag concerning privacy and intrusion. Banks can track behaviour patterns of each customer. It is useful in preventing criminal acts, but it poses a significant security danger to the customer if it falls in wrong hands. The second challenge faced by the banking is the massive amount of growing data which they are struggling with. Most legacy system can't cope up with the growing workload as most of the organisations use the outdated systems to store this massive amount of data. Trying to collect, store and analyse the amount of data on an outdated infrastructure can put the entire system at a huge risk. As a result, organizations need to rebuild the system in order to cope up with the increasing data. The third challenge is that the bigger the data the higher is the risk. Where there is data there is the risk, data collected, stored and analysed by banks should be always stored safe. Only 38% of the organisations worldwide are ready to handle the threat which means that 62% are not ready to handle. Therefore, cybersecurity remains one of the most burning issues in banking. The forth challenged faced by banking is the Data quality. If the data is mapped and cleansed which is in the step of data pre-processing, extreme care must be taken so that the original data is not lost. Most of the data collected is useful but there also comes the irrelevant data which need to be sorted. Special care must be given to unstructured and semi-structured data. Many Banks are struggling with maintaining the data quality.

Various Data science practices can be undertaken in order to overcome the challenges In order to overcome the privacy and security challenge Business Intelligence tools (BI) tools can be used as well as various Machine learning algorithms can be used in effective detection and prevention of fraud involving credit cards, accounting, insurance, and more. By doing analytics with Business Intelligence tools we can identify the risks in sanctioning loans to potential customers. In order to solve the issue of growing data organisations need to update their systems up to date to store massive amount of data. Organizations need to rebuild the entire system to solve this problem. To solve the third issue, big data analytics in banking can be used to enhance the cybersecurity and reduce the risk. This can be done by using intelligent algorithms and also can be used for detecting fraud and prevent potentially malicious actions. Forth

challenge can be solved in the step of data pre-processing. While doing cleaning care must be taken to not lose the original values. Because data migrates criteria is created and recorded to ensure data sets accurately reflects the data at the point of acquisition and that no data is lost. For the best data quality banks need to create data quality metadata that includes all the attributes, measures, business rules, mapping, cleansing routines, data element profiles and controls.

Chatbots are used in banking to perform variety of task that range from checking account balance, offering user suggestions and also on how to save money and also helping customers to solve their query. There are various machine learning techniques which help in fraud detection, managing customer data, tracking customer spending patterns, risk management, helping in product offering, lifetime value prediction and also collecting analysing and responding to customer feedback. These are the techniques and the approaches which are currently being used by banking.

Block chain technology can be used to add value. Block chain fundamentally transform banking and financial services. Financial transactions are broken down into encrypted packets or blocks which are then added to the "chain" of computer code and encrypted for enhanced cybersecurity. Robotic process automation can be used in banks for customer service like checking account balance, payment related queries and save the time of human agents. Augmented reality can be implemented to virtually help customers and solve queries of customers, this can solve the problem of humans. Quantum computing can be used by using quantum mechanics to work out complex data operations.

Analysis of Health Domain.

The first data challenge faced in health domain is Data collection which is a messy problem. All the data comes from many different sources. Capturing data that is clean, complete, accurate, and correctly formatted for use in an ongoing problem faced by medical. This problem can be solved by understanding the complete workflow and a complete understanding of why big data is important can solve this issue. Second challenge faced is cleaning. It is more important to clean the data, dirty data can give you the bad results and also lead to bad predictions. This problem can be solved with the help of Data science by using the necessary tools and by removing the noisy data from the dataset. Third challenge is security, Data security is the number one priority for healthcare organizations. This can be overcome by using up-to-date antivirus software,

setting up firewalls, encrypting sensitive data, and using multifactor authentication. Forth challenge is updating, healthcare data is not static. Most of the elements needed relatively frequent updates in order to remain updated. To solve this problem, Data providers must have a clear idea of which datasets need to be manually updated, which are to be updated via software and also ensuring updates which can be done without damaging the dataset. Ensuring that which dataset needs to be duplicate copy so that the data is not lost.

Techniques which is currently used is medical image analysis. In this popular techniques include magnetic resonance imaging (MRI), X-ray, computed tomography etc. Various methods are used to modality, resolution, and dimension of these images. Many more are being developed to improve the image quality, extract data from images more efficiently, and provide the most accurate interpretation. The deep-learning based algorithms increase the diagnostic accuracy by learning from the previous examples and then suggest better treatment solutions. Hadoop which is an analytical framework is used to find the optimal parameters for task like lung texture classification. With the help of machine learning, predictions helps in effectively scheduling to know which patients needs the most care and which needs the least. The AI apps can provide basic health support known as chatbots.

Techniques that can be implemented and add value are 3D printing drugs in dinosaur shapes for kids, Virtual reality which can be used for medical teaching students at university, Brain-computer interfaces bring hopes for the paralyzed. Data science can be used to alert a person with detection whether he/she is vulnerable to disease or not. Using AI- based symptom and cure checker that uses algorithms to diagnose and treat illness. Artificial intelligence can be used to diagnose potentially deadly blood diseases at a very early stage.

Comparison between Banking and Health.

Commonalities:-

- Data cleaning, filtering and transformation is done in the same way in both domains. Data Analysis and visualizations process is almost same in both the domains. In both the domains Machine learning techniques and big data analytics are being used. In both the domains data needs to be taken from different sources and then need to integrate and used for further analysis. Data security is the number one priority in both the domains. In both the domains datasets are updated because banks and health both has customers and patients where people add on every day.

Banks help customers in keeping track of their bank related work whereas in medical patient related data is stored.

- **Differences:-**Data is in various format in both the domains banks use data related to peoples age, gender, account details whereas health related data is data is in video, images and also in structured format.

References:-

https://www.fintechnews.org/top-9-data-science-use-cases-in-banking/

https://www.wowso.me/blog/technology-in-banking

https://healthitanalytics.com/news/top-10-challenges-of-big-data-analytics-in-healthcare

https://www.corporatecomplianceinsights.com/4-health-care-data-challenges-overcome/

https://medium.com/activewizards-machine-learning-company/top-7-data-science-use-cases-in-healthcare-cddfa82fd9e3

https://builtin.com/artificial-intelligence/artificial-intelligence-healthcare

https://www.datasciencecentral.com/profiles/blogs/5-key-big-data-challenges-in-banking-industry

https://www.datarobot.com/blog/5-data-science-challenges-banks-face-and-how-to-overcome-them/

https://easternpeak.com/blog/big-data-in-the-banking-industry-the-main-challenges-and-use-cases/