Assignment -1

Failed Analytics Project Real-World Cases

IBM Watson for Oncology 2013

Introduction: -

IBM Watson for Oncology was one of the most prominent and ambitious projects aimed at using AI to cancer treatment. Aiming to give physicians and patients individualized, evidence-based cancer treatment recommendations, IBM Watson for Oncology was a system that combined natural language processing, machine learning, and big data analytics. One of the top centers in the world for cancer research and treatment, Memorial Sloan Kettering Cancer Center (MSKCC), and IBM worked together to build the system. After its 2013 introduction, the system has been implemented in several hospitals and healthcare facilities worldwide.

The IBM Watson for cancer project was one of the most extensive and well-known initiatives to employ AI for cancer. Aiming to give physicians and patients individualized, evidence-based cancer treatment recommendations, IBM Watson for Oncology was a system that combined natural language processing, machine learning, and big data analytics.

Unfortunately, due to the project's reliance on obsolete and inadequate data, inconsistent and erroneous suggestion generation, and opposition from the medical community, it was unable to provide any meaningful results. Following the cessation of use by a number of hospitals and health groups, the project was abandoned in 2018.

Case Study

Cancer affected 1.5 million Indians in 2016; by 2020, that number is expected to rise to 1.7 million. The nation also has a severe problem with knowledge and affordability; only 12.5% of patients seek treatment in the early stages. Consequently, the number of Indian deaths from cancer was estimated to reach 880,000 in 2020, up from 736,000 in 2016. The stark shortage—there are only 1,000 oncologists in the nation, or a ratio of [1:1500]—complicates these issues. The US, on the other hand, has a ratio of [1:100].

Big hospital networks in India, including Apollo, Manipal, and Fortis, who together cater to around 60% of the country's organized private market, have the duty and opportunity to implement cutting-edge patient care strategies. In 2016, Manipal, which holds around 12.5% of the organized private market share, declared the nationwide rollout of IBM Watson in six hospitals. The plan is to employ a cognitive computing platform to give doctors in India individualized, evidence-based cancer care alternatives by giving them information and insights. For patients with breast, colorectal, and lung cancer who are registered with a Manipal hospital nationwide, Manipal oncologists have access to IBM Watson for Oncology.12 oncologists have been trained to use this solution to help patients in better diagnosis and administer treatment accordingly. In the long term, Manipal hopes to gain 3 benefits for its oncologists.

Discovery phase:

IBM Watson for cancer's discovery phase is the first step in utilizing cognitive computing and artificial intelligence (AI) to support cancer treatment for medical practitioners. A system called IBM Watson for Oncology was created by IBM and Memorial Sloan Kettering Cancer Center (MSK) to offer cancer patients suggestions for evidence-based treatment. Large volumes of structured and unstructured data are gathered and analyzed by IBM Watson for Oncology throughout the discovery phase. These sources include patient records, clinical trials, medical literature, and the knowledge of expert oncologists.

Data Collection and Integration: The system gathers a lot of information about treating cancer, such as medical histories of patients, genetic information, pathology reports, treatment protocols, medication information, and data from clinical trials. Medical databases, academic articles, electronic health records (EHRs), and other sources are the sources of this data.

Natural Language Processing (NLP): To extract pertinent information from unstructured text, including clinical notes, research papers, and medical literature, IBM Watson for Oncology makes use of NLP capabilities. By interpreting and comprehending human language, NLP enables the system to draw conclusions and information from a wide range of textual data sources.

information Acquisition: IBM Watson for Oncology gains domain-specific information about cancer treatment protocols, medication efficacy, treatment recommendations, and best practices by working with oncology specialists from MSK and other top cancer facilities. To facilitate evidence-based decision-making, this information is incorporated within the cognitive computing architecture of the system.

Machine Learning and Training: To examine and draw conclusions from patterns in the data, such as patient outcomes, treatment responses, and clinical recommendations, the system makes use of machine learning algorithms. Over time, IBM Watson for Oncology enhances the precision of its therapy suggestions and constantly sharpens its understanding by training on expert advice and historical data.

Validation and Testing: To evaluate IBM Watson for Oncology's efficacy, precision, and dependability, extensive validation and testing procedures are carried out during the discovery stage. This entails assessing the system's capacity to produce treatment recommendations that are clinically meaningful and in line with consensus guidelines and assessments made by knowledgeable oncologists.

All things considered, IBM Watson for Oncology's discovery phase is a thorough and iterative process that includes knowledge acquisition, machine learning, integration, validation, and data gathering and processing. IBM Watson for cancer uses AI and cognitive computing technologies to enhance the decision-making abilities of oncologists, enhance patient outcomes, and progress the cancer care domain.

Preparation Phase:

The process of preparing IBM Watson for Oncology for usage in a particular hospital or institution is known as the preparation phase. The steps involved are as follows:

1) Installing and setting up Watson for Oncology's hardware and software, including servers, databases, and interfaces.

2)Tailoring the system to the institution's standards, local preferences, and practices; this includes changing treatment regimens, medication names, and doses.

3)Teaching staff members how to create patient cases, go over treatment alternatives, and give feedback using the system in an effective and efficient manner.

4)Testing the system includes testing the suggestions, confirming the connectivity, and ensuring the quality, correctness, and dependability of the data.

Several weeks or months may pass during IBM Watson for Oncology's preparation phase, contingent upon the institution's size and complexity, resource availability, and degree of customization needed. This phase's mission is to guarantee that Watson for Oncology can offer the best possible assistance for cancer care while being completely integrated and in line with the workflow and goals of the institution.

Model Planning Phase:

Creating and refining machine learning models for cancer treatment requires careful consideration, which starts with IBM Watson for cancer's model planning stage. Data selection, feature engineering, model selection, interpretability, validation, performance measurements, bias reduction, ongoing learning, and adherence to legal and ethical requirements are all included in this phase. Patient records, data from clinical studies, medical journals, and professional advice are examples of data sources. Preprocessing removes errors like missing values and outliers from raw data and formats it so that it may be used for model training.

Choosing and altering pertinent variables from the data is known as feature engineering, and it is used to improve the models' capacity to recognize patterns and provide precise predictions. Selecting the right machine learning algorithms—such as decision trees, support vector machines, or neural networks—that are in line with the objectives of the system is known as model selection. For medical professionals to comprehend and have faith in the system's suggestions, it is imperative that the recommendations be comprehensible.

Metrics for validation and performance evaluate how well machine learning models work and how well they generalize to new data. Methods like fairness-aware machine learning are used to lessen inequities and biases. To keep the model current and relevant, it incorporates ongoing learning and upgrades.

The model planning stage of Watson for Oncology is a painstaking, repetitive procedure that prioritizes interpretability. The model planning step is not complete without ethical and regulatory compliance, which complies with established norms, patient privacy rules, and industry standards. To put it briefly, IBM uses continuous learning, bias reduction, validation, and AI to deliver safe and reliable AI-driven insights for cancer therapy.

Recovery from the data breach for the IBM Waston for Oncology.

Discovery Phase:

The processes involved in IBM Watson for Oncology's data breach recovery procedure are several. The initial phase is incident identification, which entails figuring out which breach resulted in the compromising of private information. The following stage is scope evaluation, which entails determining the types of data exposed, the amount of the data exposure, and the possible effects on patients, healthcare providers, and the integrity of the system. In order to put appropriate remediation procedures in place, a detailed root cause analysis is carried out to determine the reason of the breach.

It is imperative to do a thorough inventory of any compromised data, encompassing sensitive information such as treatment plans and patient data. During the discovery phase, prompt and clear communication is essential for managing fallout and preserving confidence. Legal and regulatory compliance is particularly crucial since IBM Watson for Oncology needs to make sure that privacy laws, laws protecting healthcare data, and other pertinent requirements are followed.

To get proof of the breach, forensic analysis is carried out. It involves looking through network traffic, system logs, and other digital artifacts to track down the attackers' movements, locate entry points, and estimate the amount of unapproved access. Mitigation and containment strategies are implemented to reduce the extent of the breach's harm.

An extensive analysis of the incident response technique, including an assessment of the efficacy of security controls, incident detection mechanisms, and response protocols, rounds off the discovery phase. These actions are essential for ensuring a thorough recovery process and averting more security breaches.

Data sourcing for the IBM Waston for Oncology

To support the system's capacity to offer evidence-based therapy recommendations, IBM Watson for Oncology data source entails gathering and integrating a wide variety of healthcare data. IBM Watson for Oncology uses data from several sources so that it may use a wide range of facts for analysis. The following are important details about IBM Watson for Oncology data sourcing:

1)Organized clinical data, including demographics, diagnosis, staging, biomarkers, therapies, and results for each patient.

2) Unstructured clinical data, including results from genetic sequencing, pathology and imaging reports, and clinical notes. Natural language processing methods are utilized to consider this data.
3) Evidence-based cancer therapy alternatives that consider the client's experience, pertinent US research, the patient profile, and a comparison to consensus guidelines in the US. The Memorial Sloan Kettering Cancer Center's (MSKCC) doctors have selected and organized this data, applying their expertise to teach Watson for Oncology.
4) Additional data includes clinical studies, published medical literature, and the prescribing guidelines provided by makers of cancer medications. For Drug Discovery, another IBM Watson Health application, Watson retrieves and ranks this data.

What data was taken away from IBM Waston for Oncology?

Through a project called IBM Watson for Oncology, patients with cancer might receive individualized therapy suggestions via artificial intelligence (AI). But despite several obstacles and critiques, the project was ultimately auctioned off for components in 2022.

A few of the information extracted from IBM Watson for Oncology is as follows:

1) The genetic information and medical records of thousands of cancer patients from different hospitals and research facilities that collaborated with IBM .

2) Clinical guidelines and expert views from New York's Memorial Sloan Kettering Cancer Center (MSK), whose clinicians taught Watson to identify important information related to various cancer types.

3) IBM created a range of data and analytics technologies for the treatment of cancer, including population health, clinical trials, and medication development.

The fate of these data following IBM's sale of Watson for Oncology is unknown. A portion of them could have been given to private equity company Francisco Partners, the new owners. It's possible that some were kept by IBM or given back to the original suppliers. These data may include sensitive information about patients and their treatments, raising further questions about their security and privacy.

Modeling Phase:

To identify solutions and tactics to restore normal operations and thwart future assaults, a variety of methodologies and procedures are applied during the modeling phase of the IBM Watson for Oncology data breach recovery process.

Some of the measures the **IBM Waston for Oncology took to Recovery from Data Breach.**

1) Specify the goals and parameters of the recovery, including the target system state, the amount of time allotted, the resources needed, and the parties participating. Sort the recovery tasks (data restoration, system repairs, software updates, security control implementation, and functionality testing) into categories and rank them.

2) The right recovery methods, such as encryption, malware removal, backup restoration, data recovery, and system reconfiguration, should be chosen and used.
3) Measure the system's performance, availability, dependability, and security, and compare the outcomes to the goals and baseline in order to assess the efficacy and efficiency of the recovery process.
4) Recoverable outcomes and lessons learned should be recorded and disseminated, including event information, recovery activities, costs, impacts, gaps, and suggestions.

A multidisciplinary team of experts, including data scientists, security analysts, system administrators, and incident responders, is needed for the modeling phase of the IBM Watson for Oncology data breach recovery process. These experts must be able to work together and coordinate their efforts to achieve the best possible outcome. To identify, respond to, and recover from the cyberattack, the modeling phase also necessitates the use of a variety of techniques and technologies, including data analysis, machine learning, incident response, and data resilience. These are a few instances of such devices and technologies:

The IBM Storage Flash System family's artificial intelligence capabilities are made possible by IBM Flash Core Module technology, which also combines with Storage Defender to offer end-to-end data resilience across primary and secondary workloads using AI-powered sensors meant to detect cyber threats sooner.

Using inline data corruption detection software and cloud-based artificial intelligence, IBM Storage Defender software scans all incoming data down to block level granularity without affecting performance as it is being written. This helps organizations become more adept at identifying and responding to ransomware and other cyberattacks that threaten their data.

IBM Watson Studio offers a range of modeling approaches and processes, including data preparation, feature engineering, model selection, model optimization, and model assessment. This enables data scientists and developers to create, implement, and trust models.

By studying incident response methodologies and security models and completing a case study on a recent cyberattack and breach, the Coursera Cybersecurity Capstone: Breach Response Case Studies course helps students gain insight into cybersecurity breaches and work in the cybersecurity field as cybersecurity analysts or specialists.

An intricate and crucial procedure requiring meticulous planning, carrying out, and assessing is the modeling phase of IBM Watson for Oncology's recovery from the data breach. In addition to returning the system to its pre-attack condition, the modeling phase tries to strengthen its security

and resilience against potential attacks. Insights and lessons gained from the modeling phase may also assist the company improve its cybersecurity preparedness and posture.

Reason for it’s Failure:

Personalized and evidence-based therapy recommendations for cancer patients were to be provided by IBM Watson for Oncology, an ambitious initiative that used artificial intelligence (AI). Nevertheless, several obstacles and complaints that the project encountered contributed to its demise, including:

Absence of proof and data: Watson for Oncology was dependent on a small, skewed dataset consisting mostly of fictitious instances and professional judgments from Memorial Sloan Kettering Cancer Center (MSK), a single organization. Real-world patient data, clinical trial results, or outcomes were not available to it. To further confirm the precision, security, and efficacy of Watson's suggestions, IBM did not provide any peer-reviewed research.

Cancer's complexity and variability: Watson for Oncology found it difficult to manage the intricacies and variations of cancer situations, including tumor heterogeneity, the presence of other illnesses, the way cancer changes over time, and the values and preferences of both patients and physicians. Watson frequently made irrelevant, out-of-date, or incorrect recommendations because he was unable to comprehend the context and subtleties of patient information.

AI mismatch with clinical practice: Watson for Oncology was intended to assist oncologists in making decisions, not to make decisions for them. But a lot of physicians thought Watson's output and UI were unclear, unreliable, and laborious1. Medical professionals who disapproved of AI or who valued their own expertise and judgment also opposed Watson1. Additionally, Watson failed to take into consideration the price, accessibility, and availability of medicines in various nations and areas.

These are a few of the primary causes behind IBM Watson for Oncology's downfall. The experiment shows that there are several technological, clinical, and societal issues that need to be resolved before implementing AI in the healthcare industry may be as easy or uncomplicated as it may seem.

CONCLUSION:

IBM Watson for Oncology is a game-changing tool in the field of cancer therapy that uses artificial intelligence to evaluate patient data, suggest evidence-based treatments, and keep track of ongoing research. Regulating standards and ethical data processing are two issues that still need to be resolved. Its success depends on developer, healthcare provider, and regulatory body collaboration as well as user trust in the technology. The entire promise of AI in cancer treatment will be realized by addressing obstacles, improving techniques, and increasing ethical issues. With continued development and patient-centered methods, IBM Watson for Oncology has the potential to revolutionize cancer treatment.