PHASE 2

Processes of Big Data Analytics using IBM Cloud Computing

Data Collection and Ingestion:

- Collect and ingest your data into IBM Cloud. This can include structured and unstructured data from various sources.
- Use IBM Cloud services like IBM DataStage or IBM Cloud Pak for Data to help with data integration and data quality.

Data Storage:

- Store your data in a suitable data repository. IBM Cloud offers options like IBM Db2 on Cloud, IBM Cloud Object Storage, and IBM Cloud Databases.

Data Preparation and Cleaning:

- Use tools like IBM Watson Studio, IBM Data Refinery, or Jupyter Notebooks to clean, preprocess, and transform your data.
- Address missing values, outliers, and ensure data quality.

Data Analysis and Modeling:

- Utilize analytics tools such as IBM Watson Studio, which supports various programming languages like Python and R.
- Develop and train machine learning models or run data analytics jobs.

Scalability and Performance:

- Leverage IBM Cloud's auto-scaling and elastic resources to handle the performance requirements of big data analytics.

Data Visualization:

- Use tools like IBM Cognos Analytics or Watson Analytics to create visualizations and dashboards for data insights.

Machine Learning and AI:

- Explore IBM Watson Machine Learning for deploying and managing machine learning models.
- IBM Watson AutoAI can assist in automating the model building process.

Real-time Analytics:

- If needed, set up real-time analytics using services like IBM Streaming Analytics to process and analyze streaming data.

Security and Compliance:

- Ensure data security by leveraging IBM Cloud's built-in security features.
- Comply with relevant data protection regulations.

Deployment and Monitoring:

- Deploy your analytics solution in the IBM Cloud environment.
- Continuously monitor the performance and the accuracy of your models.

Optimization:

- Use IBM Cloud Monitoring and Analytics to identify performance bottlenecks and optimize your big data analytics solution.

Cost Management:

- Keep track of your cloud resource usage to manage costs effectively. IBM Cloud provides tools for cost management and optimization.

Integration:

- Integrate your analytics results with other business systems and applications as needed.

Documentation and Training:

- Ensure that your team is well-versed in the tools and technologies used for big data analytics on IBM Cloud.
- Document the processes and best practices for future reference.

Feedback and Iteration:

- Collect feedback and insights from your analytics results to make improvements and iterate on your analytics solution.

INNOVATION :

Augmented Analytics using IBM CLOUD COMPUTING

- As AI and machine learning models become more complex, the need for explainability has grown. Innovations in making AI decisions more transparent and interpretable are crucial, particularly in regulated industries.

Block chain for data security

- Blockchain technology is being explored for data security and integrity, particularly in scenarios where trust and immutability are essential.

Explainable AI

- As AI and machine learning models become more complex, the need for explainability has grown. Innovations in making AI decisions more transparent and interpretable are crucial, particularly in regulated industries.

Hybrid and Multi-Cloud Solutions

- Innovations in hybrid and multi-cloud solutions provide flexibility in choosing where to store and analyze data, optimizing costs and performance.

File Naming Convention: CAD_BigDataAnalytics_Phase2