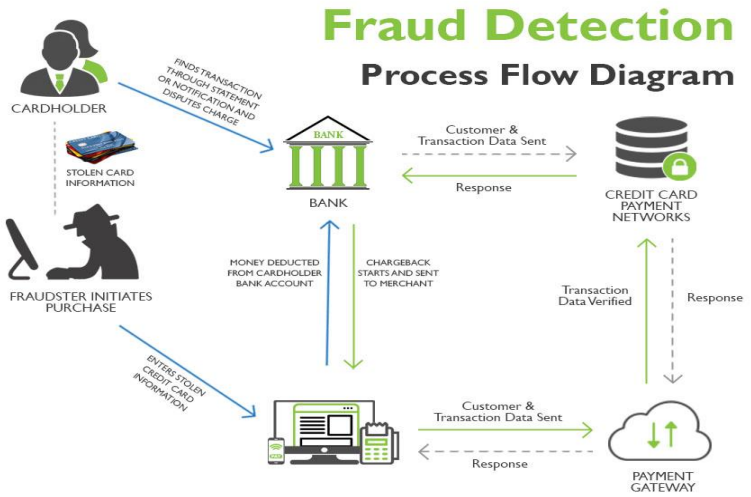
Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	19 October 2022
Team ID	591653
Project Name	Online Fraud Detection
Maximum Marks	4 Marks

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2



Guidelines:

1. Include all the processes (As an application logic / Technology Block)
2. Provide infrastructural demarcation (Local / Cloud)
3. Indicate external interfaces (third party API's etc.)
4. Indicate Data Storage components / services
5. Indicate interface to machine learning models (if applicable)

Table-1 : Components & Technologies:

S.No	Component	Description	Technology

1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Application Logic-1	Login process in the application	Gmail
3.	Verification Process	Login process in the application	Bank Information
4.	Database	Data Type, Configurations etc.	MySQL, NoSQL, etc.
5.	Cloud Database	Database Service on Cloud	Aws.
6.	File Storage	To store user data of app and websites	IBM Block Storage or Other Storage Service or Local Filesystem
7.	External API-1	For the verification user bank details	Bank API, etc.
8.	External API-2	For the verification of person	Aadhar API, etc.
9.	Machine Learning Model	To detect online fraud	Classification Model
10.	Infrastructure (Server / Cloud)	Cloud Server Configuration :	Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Apache Kafka is a distributed streaming platform that can be used to collect and process real-time data. It's often used as the backbone for streaming data in fraud detection systems, enabling you to ingest and analyze data as it flows in. Apache Flink is a stream processing framework that can be used for real-time data analytics. It's suitable for applications where low-latency processing is crucial, such as fraud detection	Apache Kafka, Apache Flink, ELK Stack (Elasticsearch, Logstash, and Kibana), OpenDXL

		The ELK Stack is often used for log analysis and real-time monitoring. Elasticsearch is used for indexing and searching data, Logstash for data collection and transformation, and Kibana for visualization.	
2.	Security Implementations	implement strong user authentication methods such as multi-factor authentication (MFA) to ensure that only authorized users can access accounts and perform transactions. Enforce role-based access control (RBAC) to restrict access to sensitive systems and data to authorized personnel only	e.g. User Authentication and Authorization:
3.	Scalable Architecture	Consider using cloud services like AWS, Azure, or Google Cloud for scalability and elasticity, allowing you to adapt to varying workloads.	Cloud Services
S.No	**Characteristics**	**Description**	**Technology**
4.	Availability	high availability database solutions that offer features like automatic failover, data replication, and low-latency access.	Amazon RDS (Relational Database Service) and Google Cloud SQL.
5.	Performance	Good performance in online fraud detection requires a combination of advanced technologies, effective machine learning models, and robust processes for continuous improvement. Regular testing, optimization, and adaptation to evolving fraud patterns are key to maintaining high-performance levels.	Hadoop Ecosystem: Hadoop, HDFS, and tools like Apache Spark are used for batch processing and analysis of large volumes of historical data. NoSQL Databases: Databases like Apache Cassandra, Amazon DynamoDB, and Google Bigtable offer efficient storage and retrieval for real-time data. Graph Databases: Graph databases like Neo4j can be valuable for analyzing complex relationships in fraud networks.