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

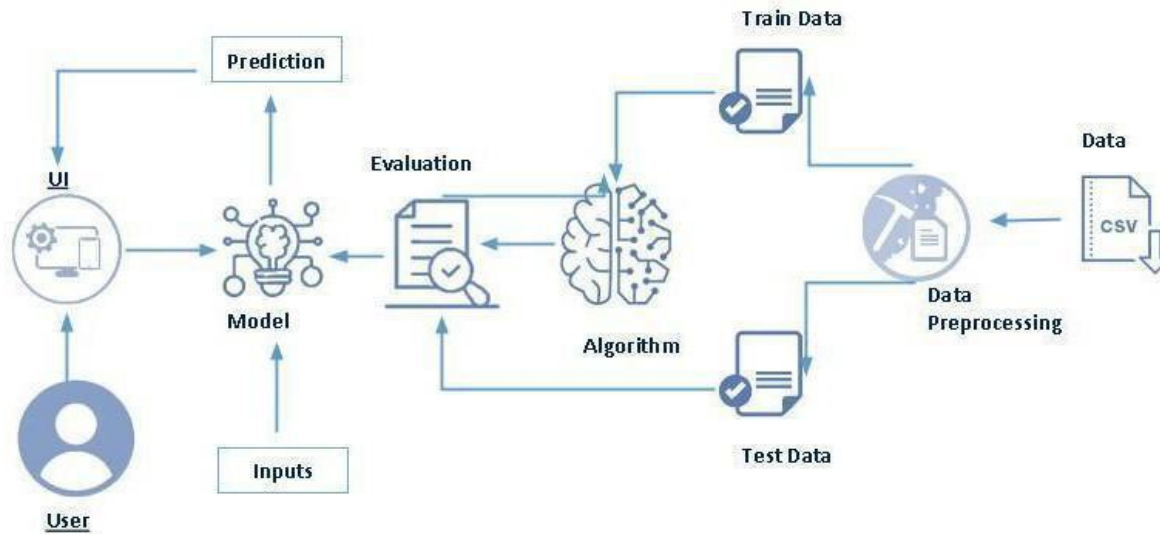| Date	13 May 2023
Team ID	NM2023TMID01937
Project Name	Project – Audit AI: A Machine Learning for Detecting Fraud in Audit Data
Maximum Marks	4 Marks

Technical Architecture:

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	Graphical User Interface (GUI) for user interaction	HTML, CSS, JavaScript, Flask
2.	Machine Learning Model	Fraud detection model using ML algorithms	Python, Scikit-learn, TensorFlow
3.	Data Processing Logic-1	Pre-processing and data cleaning	Python, Pandas
4.	Data Processing Logic-2	Feature engineering and selection	Python, Scikit-learn
5.	Data Storage	Storage for audit data and processed data	Relational or NoSQL database (e.g. MySQL, MongoDB)
6.	Cloud Storage	Storage Service on Cloud	IBM Cloud Object Storage
7.	External API-1	Integration with financial data providers	APIs from financial data providers (e.g. Bloomberg, Reuters)
8.	Infrastructure (Server / Cloud)	Deployment on Cloud	IBM Cloud, Kubernetes

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Scikit-Learn Flask	Python Python
2.	Security Implementations	Data encryption Secure user authentication and access controls Secure APIs	AES-256 algorithm OAuth2 or JWT for secure user authentication secure APIs to prevent unauthorized access to data
3.	Scalable Architecture	microservices architecture cloud-native architecture	Docker, Kubernetes, Memcached, Linkerd, Azure Autoscale.

S.No	Characteristics	Description	Technology
			Docker, NGINX, AWS, Amazon RDS, Memcached, AWS Lambda
4.	Availability	Multiple availability zones, using load balancers to distribute traffic across servers, and implementing failover mechanisms to ensure that if one server goes down, the application can continue to run on another server.	Technology used for ensuring availability can include clustering, failover mechanisms, and monitoring tools.
5.	Performance	Handle multiple requests in parallel, Caching mechanisms, Content Delivery Networks, Load testing tools, Query optimization techniques, Asynchronous processing	CDNs, Redis, Memcached, JMeter, indexing, database partitioning