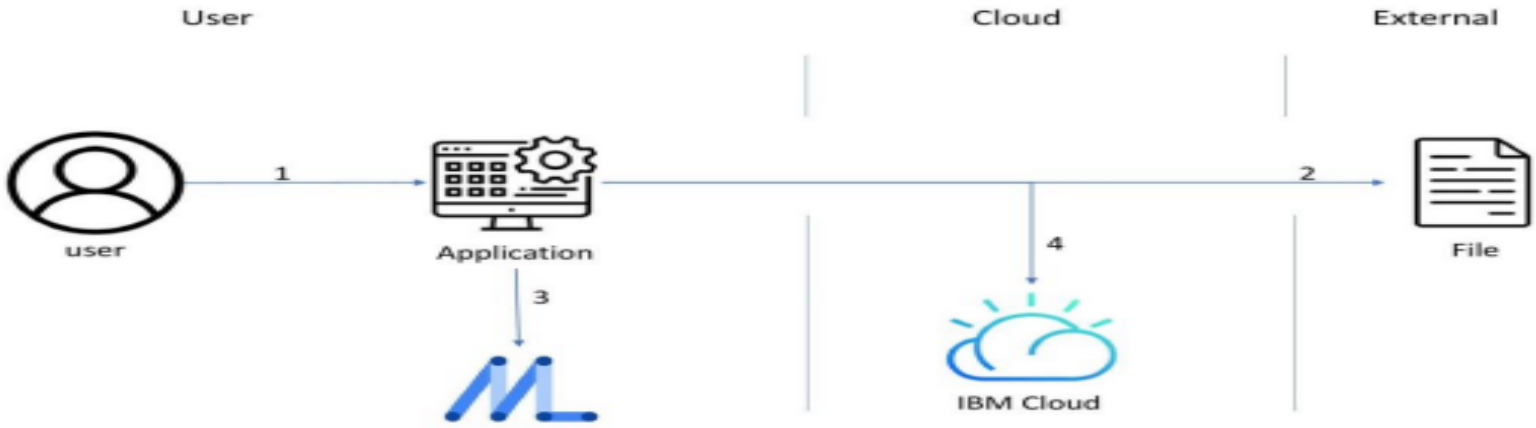
Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	06 May 2023
Team ID	NM2023TMID17607
Project Name	Cancer Mortality & Incidence Rates Classification Using Machine Learning

Technical Architecture:



1. The user login into the webpage with his mail id and password.

2. The User will select the input in the application and will load the same for processing.

3. The input data is given to the machine learning model, and the model will process the data set with the preloaded algorithm and predict the result

4. Dataset needed is stored in the IBM Cloud storage as it is a continuous learning model it updates itself with current dataset.

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 Mode	Fraud detection model using ML algorithms	Python, Scikit-learn.
3.	Machine Learning Module	To detect Cancer Mortality & Incidence Rates Classification Using Machine Learning	Python, Scikit-learn.
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	Data Type, Configurations etc.	Relational or NoSQL database (e.g.MySQL, MongoDB)
6.	Cloud Database	Database Service on Cloud	IBM Cloud Object Storage
7.	External API-1	File storage requirements	APIs from data providers
8.	Infrastructure (Server / Cloud)	Purpose of External API used in the application	IBM Cloud, Kubernetes

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Scikit-Learn Flask	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.
3.	Scalable Architecture	Micro-services architecture cloud-native architecture	Kubernetes,Docker,Istio,Prometheus,Helm.
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	CDNs, Redis, Memcached, JMeter, indexing, database partitioning.