**Project design - Phase - I**

**Proposed solution**

The proposed solution for the cancer mortality and incidence rates classification project involves several steps. Firstly, a dataset containing patient information and health conditions is obtained. The dataset is then preprocessed to handle missing values, encode categorical variables, and scale numerical features if necessary. Exploratory data analysis techniques are applied to gain insights from the dataset through visualisations and summary statistics. A suitable classification algorithm, such as Logistic Regression, Decision Trees, Random Forest, or SVM, is chosen to develop the machine learning model. The dataset is split into training and testing sets, and the model is trained using the training data. The model's performance is evaluated using appropriate metrics on the testing data. Finally, a Flask-based web application is developed to create a user interface where users can input their information and receive predictions on cancer mortality and incidence rates based on the trained model.

**Solution Architecture**

The Solution Architecture is divided into several components , as outlined below:

* **Dataset and Data Preprocessing:**
  + Obtain a dataset that includes patient information and health-related features.
  + Clean the dataset by handling missing values and transforming categorical variables into numerical representations.
  + Scale numerical features if needed.
* **Exploratory Data Analysis (EDA):**
  + Explore the dataset to gain insights and identify patterns and relationships.
  + Create visualisations and summary statistics to better understand the data.
* **Machine Learning Model Development:**
  + Choose an appropriate classification algorithm like Logistic Regression, Decision Trees, Random Forest, or SVM.
  + Split the dataset into training and testing sets to train and evaluate the model.
  + Train the selected model using the training data and adjust its parameters if necessary.
* **Flask Web Application Development:**
  + Develop a web application using Flask to provide a user-friendly interface for the cancer mortality and incidence rates classification system.
  + Design and implement routes and views to handle user requests and responses.
  + Apply form validation and data preprocessing techniques to handle user inputs effectively.
  + Utilize the trained machine learning model to make predictions based on user inputs.
  + Display the predictions to the user through the web interface.