

Assignment-1: Medical Diagnosis using Conditional Probability and Bayes' Theorem

Problem:

To determine the probability of a patient having a particular disease given a positive or negative test result, using conditional probability and Bayes' theorem.

Dataset:

Simulated medical test data. Assume the following structure:

Patient ID (S.no)	Test Result	Actual Disease Status
1	Positive(1)	Yes(1)
2	Negative(0)	Yes(1)
3	Positive(1)	No(0)

Create a binary data set using a random function.

Write a Python code to calculate the following:

(i) Theory: Write the procedure how to calculate following (10M)

(ii) Coding: Write the code to calculate following (10M)

- Calculate the probability of having the disease. (1M)
- Calculate the probability of getting a positive test result. (1M)
- Calculate the probability of getting a positive test result given that a patient has the disease. (1M)
- Calculate the probability of having the disease given a positive test result using Bayes' theorem. (2M)
- Calculate confusion matrix, accuracy, sensitivity and specificity. (2M)
- Generate classification reports including precision, recall and F1-score. (1M)
- Analyze the variation in result if dataset size increases or decreases. (2M)

Note : a. Use only standard arithmetic operations for calculations.

b. For dataset creation can use predefined functions.

Objective:

This assignment is designed to give hands-on experience in building, evaluating, and comparing classification using conditional probability and Bayes' theorem on a medical diagnosis task with a synthetically generated random data.

Submission:

Submit the following as part of your assignment:

- Jupyter Notebook or Python script containing your code and comments.
- A report (PDF or Markdown) containing the model evaluation results and your discussion on model performance.