ECE 477: Assignment #1: HDSS

1. (2 pts) What are advantages and disadvantages of CDSSs?
2. (1 pt) What is the difference between Random Tree and Random Forest?
3. (2 pts) What are the two major parts of an HDSS and their functions? Give a brief description of its tiers.
4. (2 pts) Describe the seven transitions present in an HDSS.
5. (2 pts) Describe the six stages of Tier-1 PHDS.
6. (2 pts) Describe the decision flow of HDSS Tier-2.
7. (2 pts) Describe the decision flow of HDSS Tier-3.
8. (2 pts) Describe the decision flow of HDSS Tier-4.
9. (2 pts) Describe a disease diagnosis module (DDM).
10. (3 pts) Describe the disease diagnosis module (DDM) design procedure.
11. (20 pts) **Coding project**

    For this project, you will train and compare various classifiers (decision tree, k-nearest neighbor, Naive Bayes, and logistic regression) to determine whether a patient has breast cancer. We will use the Diagnostic Wisconsin Breast Cancer dataset from the UCI machine learning repository (see details at https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic).

    The purpose of this first assignment is to recall the basic ML classifiers, Python data science libraries (Numpy, Pandas, Sklearn), and ML concepts (training-validation-test splits, training accuracy, validation accuracy), that we will draw upon in further assignments.

    See the Jupyter Notebook for more details.

    GitHub repository for ECE 477 coding projects: https://github.com/jha-lab/ECE477-2025