

Problem of the Week 4

Name: Ryan Rajaie

Student ID: 016105024

Date: 9/14/2025

Motivation

This week's Problem of the Week is provides a reflection on the new classification models we looked at this week. Here, we will compare and contrast the k -nearest neighbor (KNN) and decision tree (DT) models from Lecture 4-2.

Completing this assignment

To complete this assignment, fill in the markdown cells with your responses. When complete, upload this notebook as a Jupyter notebook AND a PDF to Canvas.

Problem 4.1: Model Description

To begin, let's remind ourselves how each model was designed. Write one or two sentence for each model that describes the principle underlying the model construction.

KNN Model: Assign each point the most common classification for the k closest points.

DT Model: Split data into branches (representing the classification) repeatedly using gini impurity until all regions are appropriately classified.

Problem 4.2: Model Hyperparameters

Both the KNN and DT models have hyperparameters that define their structure. These parameters are chosen by the user when constructing the model. List one of the key hyperparameters used for each of the models.

KNN Model: k

DT Model: criterion

Problem 4.3: Model Similarities and Differences

Both the KNN and DT models work similarly to classify regions of a feature space. What is one way that KNNs and DTs are similar? What is one way in which they are different?

Similarity: They are both able to classify data into regions.

Difference: KNN seeks to classify from bottom up while DT seeks to classify from top down.

Problem 4.4: Model Benefits

What is one benefit of using each of the models seen this week compared to other classification models we have seen? Write one sentence for each model.

KNN Model: More intuitive concept, reminds me of k-means clustering. Are they related?

DT Model: Similar to KNN but regions don't have to be split by lines (can be any shape).

Problem 4.5: Model Drawbacks

What is one drawback of using the models we have seen in this week compared to other classification models we have seen? Write one sentence for each model.

KNN Model: Can have weird regions if data isn't standardized.

DT Model: Can be overfit if too many regions.