

Auburn University Assignment 4

COMP 5630/ COMP 6630/ COMP 6630 - D01 (Spring 2025)
Machine Learning

SUBMIT THE CODE IN AN IPYNB FILE (using Google Collab). NO OTHER FORMAT IS ACCEPTED NOT EVEN .PY

1 Naïve Bayes (25 Points)

See the attached housing data (Asssignment4 Data.xlsx). Each tab in the Excel file contains training and test splits. Your goal is to construct a Naïve Bayes classifier for this data.

1. Compute and show the conditional probability distribution for each feature. Explain how you got these values and show your work. *Note: You are expected to do this part of the question by hand. Explain how you got the probability distribution for at least two features in detail.*
2. Using your conditional probability table, write a Python code that will compute the probabilities for each example in the test data. Your program should output the probabilities of each class as well as the final classification based on the MAP rule. *Note: You should hard-code the conditional probabilities from the previous step into your code.*

2 Decision Tree (25 Points)

Using the same housing data (Asssignment4 Data.xlsx), construct a decision tree classifier. You can use the implementation available on Sci-Kit Learn.

Perform the following experiments and briefly (2-4 sentences) answer the questions.

1. Use the default parameters.
 - (a) What is the accuracy on the training set?
 - (b) What is the accuracy on the test set?
2. What is the effect of restricting the maximum depth of the tree? Try different depths and find the best value.
3. Why does restricting the depth have such a strong effect on the classifier performance?
4. For test data point, perform inference on decision tree
Local Price = 9.0384
Bathrooms = 1
Land Area = 7.8
Living area = 1.5 #
Garages = 1.5 #
Rooms = 7
Bedrooms = 3 Age
of home = 23