ECE/CS/ME 539 – Fall 2024 — Activity 10

1.

Using the table provided, answer the following questions to construct the decision tree.

Outlook	**Temperature**	**Humidity**	**Wind**	**Played football(yes/no)**
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

Table 1: Weather conditions and football play decisions

(a) Compute the entropy of the entire dataset for the target variable "Played football (yes/no)". Show your work.

(b) Calculate the entropy for the subsets of the "Outlook" attribute. Use these subsets to determine the information gain for each possible split: "Sunny", "Overcast", and "Rain". Which attribute value of "Outlook" gives the maximum information gain?

(c) After selecting the "Outlook" attribute and performing the first split, repeat the calculation of entropy, information gain, and resulting splits for the next attribute, focusing on the "Sunny" subset. Should the next split be on "Humidity" or "Wind"? Provide the calculations and rationale behind your choice.

(d) Perform the necessary calculations to determine the splits for the remaining subsets of "Outlook" (i.e., "Overcast" and "Rain"), considering "Wind" and "Humidity" as potential attributes for splitting. Include all your entropy and information gain calculations.

(e) Based on your calculations, construct the decision tree.