Machine Learning BITS F464	03-09-2022
Lecture 1-3: Self Information, Entropy, Decision Tree	
Lecturer: Harikrishnan N B	*Student:*

FDTA: Devale Tanmay Abhijity, Biyani Param Hemant Kumar

READ THE FOLLOWING CAREFULLY:

Honour Code for Students: I shall be honest in my efforts and will make my parents proud.
Deadline for Assignment Submission: 02:00 PM, 09 September 2022. (No assignments will be accepted after the deadline. Students shall submit an hard copy of the assignment to the FDTA (Tanmay). The submission of the assignment is in the class. The submission should be on A4 size blank paper. The FDTA will submit the assignment to instructor-in-charge on 09 Sept 2022 (Friday) before 03:00 PM).
Question 1- 4 carries 5 marks, Question 5 carries 10 mark - show the step by step calculations.

Question 1

Plot the functions (a) $f(p) = -p\log_2(p)$, (b) $g(p) = -(1-p)log_2(1-p)$, (c) $H(p) = f(p) + g(p)$. Take p varying from 0.001 to 0.999 for the plots. What are the maximum values of these functions and at what value of p does the functions achieve maximum? Derive the location of the maxima analytically.

Question 2

An experiment consists of tossing 2 coins simultaneously. One of the coins is unbiased, but the other coin has HEADS on both sides. List the sample space and find the self-information of all possible outcomes of this random experiment. What is the Shannon Entropy of this experiment? (remember that Shannon Entropy has units:)

Question 3

An unbiased die with 3 sides numbered 1, 2 and 3 is rolled once. If 1 appears, then another unbiased die with six sides is freshly rolled, else (if a 2 or a 3 appeared in the first roll of the first die) do nothing. (a) Write the sample space S and the associated probabilities of all the outcomes of S. (b) What is the probability that a 4 appears in any of the rolls? (c) Let V be the event that a prime number appears in any of the rolls. What is the probability of V ? (d) What is the Shannon Entropy of this experiment?

Question 4

Experiment X_1 has an entropy of $H(X_1) = 5$ shannons, experiment X_2 has an entropy of $H(X_2) = 5$ nats and experiment X3 has an entropy of $H(X_3) = 3.5$ trits. Arrange these in descending order of entropies (justify your answer).

Question 5

Alice and Bob loves playing tennis. As a BITS F464 Machine Learning student help Alice and Bob to develop a decision tree algorithm to decide whether they should play tennis or not given the attributes: Outlook, Temperature, Humidity, and Windy. The data recordings are provided in Table -2.1.

Table -2.1: Weather Data Set (all nominal).

Item No.	Outlook	Temperature	Humidity	Windy	Play
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rainy	Mild	High	Weak	Yes
5	Rainy	Cool	Normal	Weak	Yes
6	Rainy	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rainy	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rainy	Mild	High	Strong	No