Instructions:

1. **You can use any library for implementation of the question.**

2. **Please submit a .py file containing the required code named according to your Roll number. For example if your roll number is M20CS057 . The .py file must be named as M20CS057.py. Ipynb file won't be accepted.**

3. **You also have to submit a report in pdf format named by your Roll number.For example : M20CS057.pdf. We will not accept the submissions in the word format.**

4. **Put both the files in a folder named as your roll number (e.g., M20CS057), zip it, and upload it in google-classroom.**

5. **The report must contain all the results clearly. It should contain the link to your colab file and you must ensure that the code is accessible through this link otherwise marks will be deducted.**

6. **Question 2 is not required to be submitted.**

7. **The submissions that do not follow any of the above mentioned guidelines would result in the deduction of all the marks.**

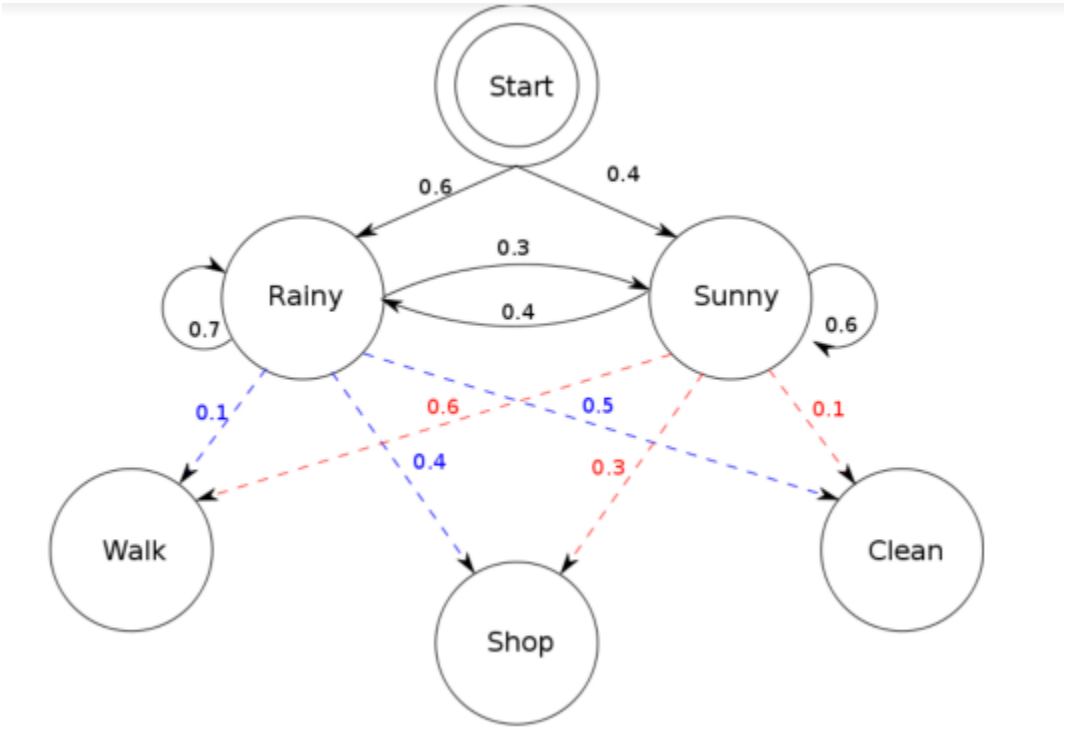
8. **There is zero tolerance to plagiarism.**

Question 1 (Due: 17 Nov 2021, 11:59 PM)

1. Download the dataset from the link below. https://drive.google.com/file/d/1s2IhEwbbSAGEtVuPpLwLq_wQ8P3Svf0l/view?usp=sharing

2. Split the dataset into training and test data in the ratio 70:30.
3. Preprocess the data, and perform classification using the Naive Bayes classifier with inbuilt libraries.
4. Implement the Naive Bayes classifier from scratch using each feature individually with appropriate bin-size, and evaluate the classification accuracies.
5. Implement the Naive Bayes classifier from scratch using all the available features with appropriate bin-sizes, and evaluate the classification accuracy.
6. Analyze the effect of bin-size on the classification accuracy in (5).
7. Compare and analyze the accuracies of the classifier built from scratch in (5) and the one using inbuilt libraries in (3).

Question 2 (Non-graded; submission not required)



Consider the above state-transition model along with the emission probabilities.
(1) What is the probability of the days of the next week to be ‘sunny’, ‘sunny’,‘rainy’, ‘sunny’, ‘sunny’, ‘rainy’, ‘rainy’?
(2) What is the probability of observing the action ‘Walk’ on every day of the next week?
(3) Solve the above problems using some inbuilt library (e.g., ‘hmmlearn’).