Assignment-5

Note: use your own words with examples

Naive Bayes Classifiers

1. True/False + Explain your answer
    a. The fundamental Naïve Bayes assumption is that each feature is Independent.
    b. Bayes theorem provides a way of computing posterior probability P(A|B) from P(B), P(A) and P(B|A).

    c. You are given P(B|A). you have all the information you need to calculate P(A|B).
    d. If your model has high accuracy on the training dataset , you can claim with 100% confidence that your model is very good.

2. How does Naive Bayes work?

3. If we have the following dataset, apply Naive Bayes to predict the new features X={color=orange, Calories=2, Tall=long, Tasty=Bad}

Color	**Calories**	**Tall**	**Tasty**	**Fruit**
Yellow	3	Short	Good	Y
Orange	2	Long	Bad	Y
Orange	3	Short	Good	Y
Yellow	3	Short	Good	Y
Orange	2	Short	Bad	N
Yellow	2	Long	Bad	N
Yellow	2	Long	Bad	N
Yellow	2	Short	Good	N

SVM

4. What is the fundamental idea behind Support Vector Machines? What is a support vector?
5. Explain the parameters that you need to change/adapt in the SVM?

Logistic Regression:

6. Assume that our LR model trains on two features data sets (x1, x2) and comes up with the following equation:f(x)=z= ax1+bx2+c, where a=3, , b=15 and c=-15, the model predicts y=1 if z>=0 and otherwise predicts 0.
    a. Plot some data and draw the area where the model predicts 1 and 9.
    b. What is the prediction of the model of x1=-2, x2=2.

7. Assume that we have a dataset and we plot/ visualize the dataset as it is shown below. Explain which **algorithms (LR, LogisticRegression, SVM, and/or Naive Bayes)**that you suggest to use the dataset and explain your answer? Alo explain which algorithm will work better on the datset?