CSE475 HW2 Due: Monday, 03/24/2025

1. Please note that you have to typeset your assignment using either LATEX or Microsoft Word for Problem 1 and Problem 2. If you have difficulty typing mathematical equations or symbols, you can take photos of handwritten answers and then insert the photos to the file. Only legible handwritten answers will be accepted and graded. You need to produce an electronic version (in PDF form) for Problem 1 and Problem 2, and name the PDF file "CSE475-HW2-LastName-FirstName.pdf". Please name your code files for Problem 3 and Problem 4 "CSE475-HW2-P3-LastName-FirstName.ipynb" and "CSE475-HW2-P4-LastName-FirstName.ipynb" respectively. Please put the PDF file and the two code files to a zipped file named CSE475-HW2-LastName-FirstName, and submit the zipped file to Canvas.

2. If you have any questions on the homework problems, you should post your question on the Canvas discussion board (under HW2 Q&A), instead of sending emails to the instructor or TA. We will answer your questions there. In this way, we can avoid repeated questions, and help the entire class stay on the same page whenever any clarification/correction is made.

1. [10pts] Assuming that the features $(\mathrm{D}, \mathrm{B}, \mathrm{W})$ are conditionally independent given N (that is, the Naïve Bayes Assumption holds), the problem (deciding if $\mathrm{N} = \mathrm{Yes}$) can be solved by applying Naïve Bayes classifier. There are 3 reasons to buy a new Phone: (i) Dropped Phone (ii) Broken Phone (iii) Want to Upgrade Phone. The following table shows the data collected regarding this problem (Yes/No in problem 3 are replaced with TRUE/FALSE).

D	B	W	N
TRUE	FALSE	TRUE	**FALSE**
FALSE	TRUE	FALSE	**TRUE**
TRUE	FALSE	FALSE	**FALSE**
TRUE	TRUE	TRUE	**TRUE**
FALSE	TRUE	FALSE	**FALSE**
FALSE	FALSE	FALSE	**FALSE**
FALSE	FALSE	TRUE	**TRUE**
TRUE	TRUE	FALSE	**FALSE**

Find the probability of buying a new phone if $\mathrm{D} = \mathrm{True}$, $\mathrm{B} = \mathrm{True}$ and $\mathrm{W} = \mathrm{False}$, i.e. $P(\mathrm{N} = \mathrm{Yes}|\mathrm{D} = \mathrm{True}, \mathrm{B} = \mathrm{True}, \mathrm{W} = \mathrm{False})$.

[Hint: Note that Naïve Bayes classifier can compute the required posterior probability and then it decides if $\mathrm{N} = \mathrm{Yes}$. Instead of deciding the label for N, you are required to compute such posterior probability. By the Naïve Bayes Assumption, we have $P(\mathrm{D}, \mathrm{B}, \mathrm{W}|\mathrm{N}) = P(\mathrm{D}|\mathrm{N})P(\mathrm{B}|\mathrm{N})P(\mathrm{W}|\mathrm{N})$ and $P(\mathrm{D}, \mathrm{B}, \mathrm{W}, \mathrm{N}) = P(\mathrm{D}|\mathrm{N})P(\mathrm{B}|\mathrm{N})P(\mathrm{W}|\mathrm{N})P(\mathrm{N})$].

2. [10pts] Below are data points of classes A and B. Also, there is an unknown test data point for which the class is unknown. Use K Nearest Neighbors classifier with Euclidean distance (i.e., L2 distance) to classify the unknown point as one of the two classes.

Class A = [(1,1), (2,1), (4,1), (2,2), (1,2.5)]
Class B = [(3,1), (3, 1.5), (3, 2.5), (4,4), (2,4)]
Test points = (3.8, 1.2)
K = 3

3. [10pts] Implement the K Nearest Neighbors (KNN) predict function in “HW2_P3.ipynb”. Submit your Jupyter Notebook file (the. ipynb file) containing your code and the outputs produced by your code (note that .ipynb file can contain both the code and the outputs, and please name your code file CSE475-HW2-P3-LastName-FirstName.ipynb).

4. [10pts] Please refer to “HW2_P4.ipynb”, and name your code file CSE475-HW2-P4-LastName-FirstName.ipynb.