CSE 4392/5369 Special Topic: Natural Language Processing

Homework 3 - Spring 2026

Due Date: Feb 9, 2026, 11:59 PM

As we delved into the Naive Bayes classifier, a foundational algorithm for text classification tasks such as spam detection, sentiment analysis, and more.

This homework is designed to reinforce your understanding of the Naive Bayes algorithm's principles and applications. You will engage in a practical exercise to manually calculate classification probabilities and make predictions using a predefined Naive Bayes model.

Problem 1 - 30%

Given the below training documents d_1 to d_5, and their class labels, compute $P(c \mid d_6)$, after applying add-1 smoothing. Please write down all the intermediate steps.

	Doc	Words	Class
Training	1	excellent definitly good	P
	2	not bad	P
	3	not good enough	N
	4	so good	P
	5	so so bad	N
Test	6	so so good	?

Problem 2 - 30%

You are given the below training set. Compute the conditional probabilities matrix P with add-1 smoothing.

	Email ID	Content	Class
Training	1	Limited offer, apply now	Spam
	2	Your subscription offer expires soon	Spam
	3	Meeting scheduled for tomorrow	Ham
	4	Congratulations! You won a prize	Spam
	5	Can we reschedule the meeting?	Ham
Test	6	Your prize offer expires tomorrow	?

Problem 3 - 40%

How do we measure the performance of a Naive Bayes classifier? In this lecture, we learned how to evaluate a classifier. Please use the knowledge you gained from this class to train and test a Naive Bayes classifier using the dataset provided below. What method are you using? How does it perform on this dataset? Why?

	Doc	Words	Class
Training	1	leaf floats gently water	1
	2	sky reflects blue leaf	0
	3	water ripples under sky	1
	4	star glimmers night gently	0
	5	night embraces cold star	1
	6	glimmers reflects cold floats	0
	7	ripples night blue water	1
	8	leaf under star sky	0
	9	cold water night floats	1
Test	10	gently embraces blue ripples	0

Submission Format: Submit the `.pdf` version of your homework (typed submissions are preferred; Scanned images must be readable) via Canvas. File must be named `lastname_studentID_hw3.pdf`.