COMP 6721 Applied Artificial Intelligence (Winter 2022)

Worksheet #3: Naïve Bayes Classifier

Joint Probabilities. Given the following joint probability distribution:

P(Toothache ∩ Cavity)

evidence

hypothesis

	Toothache	~Toothache
Cavity	0.04	0.06
~Cavity	0.01	0.89

Compute the probability that someone has a cavity, given a toothache:

$P(\text{cavity}|\text{toothache}) =$ ....

Bayes' Theorem. Assume students come to the lecture either by car (event A) or by metro. Event B means the student arrives on-time for the lecture. One student uses the car 70% of the time, i.e., $P(\text{car}) = P(A) = 0.7$. In this case, the student is 80% on-time, i.e., $P(\text{ontime}|\text{car}) = P(B|A) = 0.8$. Also, this student is on-time in general in 60% of all cases, i.e., $P(\text{ontime}) = P(B) = 0.6$. Today the student arrived on time. How likely is it that this student came by car? Apply Bayes' Theorem:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \text{..}$$

AI Fraud Detection. You just built your first AI system for detecting fraudulent credit card transactions (event B). In your company, 0.01% of all transactions are fraudulent, i.e., $P(B) = 0.0001$. Event A is "system detected fraud". You tested your system with existing training data and determined that it finds fraudulent cases with a 96% success rate, i.e., $P(A|B) = 0.96$. Unfortunately, it also sounds an alarm in 1% of non-fraudulent cases, i.e., $P(A|\overline{B}) = 0.01$ ($\overline{B}$ is the complement of B).

So, when your system sounds the fraud alarm, in how many percent of the cases was it a false alarm?

....

Hint: You will need $P(A)$, which you can compute using $P(A) = P(A|B) \cdot P(B) + P(A|\overline{B}) \cdot P(\overline{B})$.

AI Weather Prediction. Now we can build a weather-predicting AI using Bayes' theorem:

Assume we have 3 hypothesis...

- H_1: *weather will be nice* $P(H_1) = 0.2$
- H_2: *weather will be bad* $P(H_2) = 0.5$
- H_3: *weather will be mixed* $P(H_3) = 0.3$

And 1 piece of evidence with 3 possible values

- E_1: today, there's a *beautiful sunset*
- E_2: today, there's a *average sunset*
- E_3: today, there's *no sunset*

$P(E_x \mid H_i)$	E_1	E_2	E_3
H_1	0.7	0.2	0.1
H_2	0.3	0.3	0.4
H_3	0.4	0.4	0.2

Today we observe an *average sunset* (E_2). What kind of weather will we have tomorrow? Compute the probabilities for each hypothesis (H_1, H_2, H_3) using

$$P(H_i|E_2) = \frac{P(H_i) \cdot P(E_2|H_i)}{P(E_2)}, \text{with } P(E_2) = P(H_1) \cdot P(E_2|H_1) + P(H_2) \cdot P(E_2|H_2) + P(H_3) \cdot P(E_2|H_3) = 0.31$$

1. $P(H_1|E_2) =$ ....
2. $P(H_2|E_2) =$ ....
3. $P(H_3|E_2) =$ ....

So, tomorrow's weather will be: ....

Email Spam Detector. Let's train an email spam detector using a *Multinomial Naïve Bayes Classifier*, so it can classify future emails for you into the classes *spam* & *ham*. Here is your training data:

c_1: **SPAM** documents:

- d_1: "cheap meds for sale"
- d_2: "click here for the best meds"
- d_3: "book your trip"

c_2: **HAM** documents:

- d_4: "cheap book sale, not meds"
- d_5: "here is the book for you"

1. Record the *count* of each word per class below. Ignore words from the documents that are not in the table:

	best	book	cheap	sale	trip	meds	#words
c_1: SPAM							
c_2: HAM							

2. Now compute the conditional probabilities $P(w_j|c_i)$ for each word/class, as well as the prior probability $P(c_i)$ for each class, based on your training data:

	best	book	cheap	sale	trip	meds	$P(c_i)$
c_1: SPAM							
c_2: HAM							

3. Now you have a new email coming in:

   - d_6: "the cheap book"

   Is this email *spam* or *ham*? Apply Bayes' Algorithm to find out which class has a higher probability:

   (a) $P(c_1) =$ ....

   (b) $P(c_2) =$ ....

   So, the new email is: ....

Machine Learning System Evaluation. Consider the results from three different ML systems on a binary classification task. Here, X1–X5 are the instances that the systems should have recognized as belonging to a specific class (e.g., spam email, cat photo, fraud transaction). The remaining 495 instances do not belong to this class:

	Target	*system 1*	*system 2*	*system 3*
	X1 √	X1 ×	X1 √	X1 √
	X2 √	X2 ×	X2 ×	X2 √
	X3 √	X3 ×	X3 √	X3 √
	X4 √	X4 ×	X4 √	X4 √
	X5 √	X5 ×	X5 ×	X5 √
	X6 ×	X6 ×	X6 ×	X6 √
	X7 ×	X7 ×	X7 ×	X7 √
	... ×	...	... ×	... ×
	... ×	...	... ×	... ×
	X500 ×	X500 ×	X500 ×	X500 ×

Evaluate the performance of the three systems using the measures *Accuracy, Precision,* and *Recall*:

	system 1	system 2	system 3
Accuracy			
Precision			
Recall			