School of Computing and Information Systems
The University of Melbourne
COMP90049 Introduction to Machine Learning (Semester 2, 2022)

Week 4

1. What is optimization? What is a "loss function"?
2. Given the following dataset, build a Naïve Bayes model for the given training instances.

ID	*Outl*	*Temp*	*Humi*	*Wind*	PLAY
A	s	h	n	F	N
B	s	h	h	T	N
C	o	h	h	F	Y
D	r	m	h	F	Y
E	r	c	n	F	Y
F	r	c	n	T	N
G	o	m	n	T	?
H	?	h	?	F	?

3. Using the Naïve Bayes model that you developed in question 2, classify the given test instances.
   (i). No smoothing.
   (ii). Using the "epsilon" smoothing method.
   (iii). Using "Laplace" smoothing ($\alpha = 1$)
4. For the following set of classification problems, we want to design a Naive Bayes classification model.
   A. You want to classify a set of images of animals in to 'cats', 'dogs', and 'others'.
   B. You want to classify whether each customer will purchase a product, given all the products (s)he has bought previously.

   Answer the following questions for each problem:
   (i). what are the instances, what are the features (and values)?
   (ii). explain which distributions you would choose to model the observations, and
   (iii). explain the significance of the Naive Bayes assumption.
5. [OPTIONAL] Given the following dataset,
   (i). Build a Naïve Bayes model for the given training instances (1-4, above the line).
   (ii). Estimate the probability of the test instance (5, below the line)

$X_1(Headache)$	$X_1(Sore)$	$X_1(Temp)$	*Y(Diagnosis)*
0.8	0.4	39.5	Flu
0	0.8	37.8	Cold
0.4	0.4	37.8	Flu
0.4	0	37.8	Cold
0.8	0.8	37.8	? (Flu)