COMS 4030A and COMS 7047A
Adaptive Computation and Machine Learning

TEST 19 April 2023

Time: One Hour

Question 1. (6 *marks*)

(a) (2 *marks*) Let $y = \sigma(z) = \frac{1}{1+e^{-z}}$. Show that $\frac{dy}{dz} = y(1-y)$.

(b) (4 *marks*) Consider a neural network that uses the σ activation function at every output node. Suppose the sum-of-squares loss function is used for training the network with backpropagation.
For each node n, we have $y_n = \sigma(z_n)$, where we use y_n as variable for the activation value of n and z_n as variable for the sum of inputs plus bias at n.
Derive the formula for δ_n in the case that n is an output node.
Recall that $\delta_n = \frac{\partial L}{\partial z_n}\big|_{\boldsymbol{xtW}}$, where $\boldsymbol{x}$, $\boldsymbol{t}$ and $\boldsymbol{W}$ are the current values of the inputs, target and weights, and $L = \frac{1}{2}\sum_\ell (y_\ell - t_\ell)^2$ with ℓ ranging over all output nodes.

Question 2. (4 *marks*)

Let n be a node in a hidden layer in a neural network. Let y_n be the variable for the activation value of n and z_n the variable for the sum of inputs plus bias at n.

(a) (2 *marks*) Give the formula for computing δ_n from the δ's in the next layer.

(b) (2 *marks*) Simplify your answer to (a) in the case that n uses a *relu* activation function.

Question 3. (6 *marks*)

Explain what the cross-entropy loss function is and how it is used in training neural networks.

Question 4. (4 *marks*)

(a) (2 *marks*) For a binary classification problem with classes 1 and 0 denoting positive and negative, respectively, complete the following confusion matrix by filling in TP, FP, TN and FN.

		prediction	
		1	0
actual	1		
	0		

(b) (2 *marks*) Give the definitions of (i) accuracy, (ii) precision, (iii) sensitivity and (iv) F1 score.

Question 5. (6 *marks*)

Use the entropy method to create a decision tree with the dataset S given below.
Show all your calculations and draw the final decision tree.

S :

F_1	F_2	target
a	0	Y
c	1	N
c	1	Y
b	0	N
a	0	Y
b	1	N
b	1	Y
c	0	Y
c	1	N
a	1	N
c	0	Y
b	0	N

Question 6. (4 *marks*)

(a) (2 *marks*) Explain the rationale behind the naïve Bayes classifier.

(b) (2 *marks*) Use the naïve Bayes classifier on the following dataset to classify input $(f, A, 0)$.
Show your working.

x_1	x_2	x_3	target
t	B	0	N
t	A	0	Y
t	C	1	Y
f	A	1	N
f	C	0	Y
t	A	1	Y
f	B	0	N
f	C	1	N
f	B	1	N

Total marks: 30