Name:

NetID:

Discussion Section:

Course Heading:

Keep your answers concise and to the point.
Long responses won’t earn extra credit.

1. Given the following confusion matrix, select ALL the statements that accurately define sensitivity and specificity using TP, TN, FP, and FN. (5pt)

	Predicted Positive	Predicted Negative
True	TP	FN
False	FP	TN

A. Sensitivity = TP / (TP + FN)
B. Specificity = TN / (TN + FP)
C. Sensitivity = TP / (TP + FP)
D. Specificity = TN / (TN + FN)
E. None of the above

A.B

2. In SVD, the data matrix A is decomposed as A = USV'. Suppose A is a 6x22 matrix. What are the dimensions of U, V respectively? (5pt)

U 6x6 V 22x22

3. Select **ALL** correct statements about Singular Value Decomposition (SVD), where $A = USV^T$: (5pt):
   A. The matrix **U** is an orthogonal matrix.
   B. The diagonal matrix **S** contains ¥.
   C. The matrix **V** contains the eigenvectors of $A^T A$

A,C
B: The diagonal matrix **S** contains the **singular values** of A, the **square roots of the eigenvalues** of A^T A

4. Construct an optimal decision tree (with depth of 2) based on the following input data to predict patient survival. Partial credit for any reasonable decision tree. (10 pt)

Tumor degree	Tumor Size	Age	Patient survival
I	10	40	Yes
II	10	60	Yes
I	13	40	No
II	50	50	No

5. Given the input and kernel, set stride = 2 and no paddings for convolution, and max pooling with 2x2 filters and stride = 1, then manually complete the convolution step and the max pooling step of a CNN.

The matrix:

12	20	30	0	120	30
8	12	2	0	3	0
34	70	37	4	15	35
112	100	25	12	5	0
200	20	7	0	2	1
147	80	20	14	75	60

The kernel:

1	-1
0	1

(a) Filling the matrix for the output of the convolution step (10 pt):

Answer:

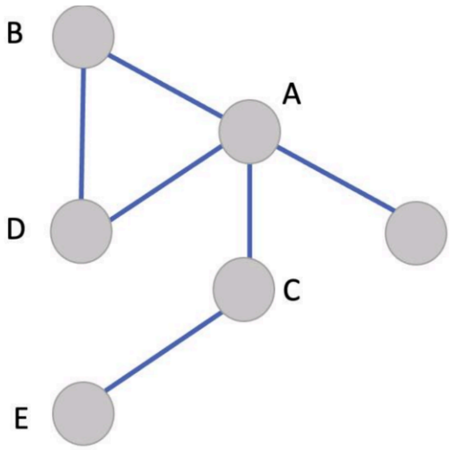
4	30	-108
64	45	-20
60	21	61

(b) Filling the matrix for the output of the max pooling step (5pt):

Answer:

64	45
64	61

5. Given the following network, answer the questions:（10 pt in total, 2pt each）



(a) The degree of node A?

4

(b) The clustering coefficient of node A?

1/6

(c) The clustering coefficient of node B?

1

(d) The clustering coefficient of node C?

0

(e) The shortest path length from B to E?

3

6. Write down two commonly used evaluation metrics for image segmentation (5pt)

Dice index
Jaccard index
Or anything reasonable

7. The following equation describes a perceptron. x is the input, and z is the linear transformation of x. y is the output of the activation. Derive the gradient of y with respect to the weights w. (15 pt)

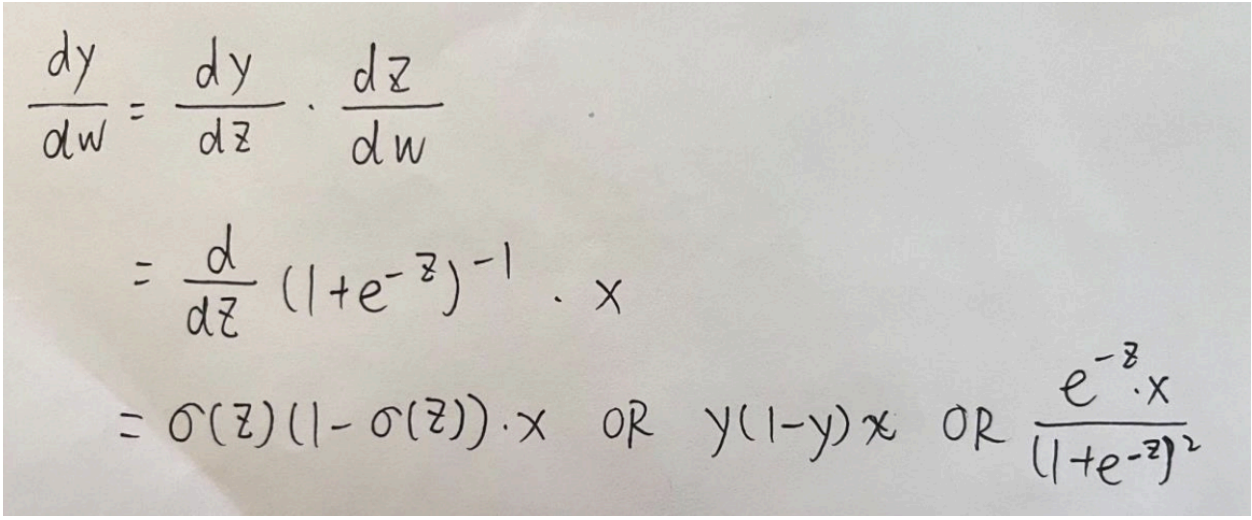
$$z=wx+b$$

$$y=\frac{1}{1+e^{-z}}$$

(a) What is the name of this activation function? (5pt)

Sigmoid

(b) Derive the gradient of y with respect to the weights w. (5 pt)



$$\frac{dy}{dw} = \frac{dy}{dz} \cdot \frac{dz}{dw}$$

$$= \frac{d}{dz}(1+e^{-z})^{-1} \cdot x$$

$$= \sigma(z)(1-\sigma(z)) \cdot x \quad \text{OR} \quad y(1-y)x \quad \text{OR} \quad \frac{e^{-z} \cdot x}{(1+e^{-z})^2}$$

8. Suppose a linear polymer is made up of 12 identical spherical atoms in three dimensions. How many bond angles and dihedral angles does this polymer possess? (5 pt)
Bond angles: 10
Dihedral angles: 9

9.What does it mean if a residue falls outside the allowed regions in a Ramachandran plot? Select ALL correct answers. (5 pt)
A) There may be a problem with the protein structure, such as a misaligned backbone or incorrect torsion angles
B) The residue is likely to be involved in protein-protein interactions
C) The residue is likely to be buried in the interior of the protein
D) The residue might adopts a high-energy or sterically unfavorable backbone conformation, which is unusual in a properly folded protein
A D

10. What is the packing fraction of an all-atom hard-sphere (i.e. explicit hydrogen) representations of residues in protein cores? Name one reason for why residues in protein cores have a smaller packing fraction than the densest packing fraction for identical hard spheres. (5 pt)

0.74 +2pt
flexibility of protein backbone and sidechain/ have voids or channels/ not identical hard spheres +3pt
*****but also, can be 0.55-0.56 in reality (if only gave this number i took off -1pt) ideally they have both numbers and explain why NOT on a lattice

11.True or false: (2.5 pt x4)

FTFFTT

(1) The Reparameterization trick introduced randomness into VAE, which is an issue in back propagation. False

(2) A typical GAN consists of a generator network and a discriminator network True

(3) Wasserstein Variational Autoencoder's loss function consists of two parts: Reconstruction loss and KL divergence. False

(4) Variational autoencoder is a discriminative model. False

(5)
In a GAN, the generator is trained to produce outputs that the discriminator classifies as real.
True

(6) A Wasserstein Variational Autoencoder tries to match the aggregated posterior (the distribution of all encoded latent vectors) to the prior, which can lead to more disentangled representations of the latent space. True

12. The figure shows a picture of a valine amino acid. Circle the four atoms that define the $\chi 1$ side chain dihedral angle for valine. (10 pt)