Week 3 Comprehensive

TOTAL POINTS 10

1. Which of the following indicates whether a doctor or machine is doing well at finding positive examples in a data set? 1 point

- ○ Specificity
- ◉ Sensitivity
- ○ Likelihood Ratio
- ○ Positive Predictive Value

2. Which of the following is used to distinguish the false positive rate from the false negative rate? 1 point

- ○ Negative Predictive Value
- ○ False Negative
- ◉ Specificity
- ○ Sensitivity

3. Which of the following is the best conceptual definition of one dimensional convolution? 1 point

- ○ "Inverting" of a shape, where the inversion matches a feature.
- ◉ "Sliding" of two signals, where a matched feature gives a high value of convolution.
- ○ "Distortion" of one signal, according to the feature shape
- ○ "Intertwining" of two signals, where one wraps around the other to form a feature.

4. Which of the following can a user choose when designing a convolutional layer? (Choose all that are correct.) 1 point

- ☑ Filter stride
- ☐ Filter depth
- ☑ Filter number
- ☐ Filter weights
- ☑ Filter size

5. What is a fully connected readout? 1 point

- ◉ A layer with a single neuron for each output class.
- ○ A layer with connections to all feature maps.
- ○ A layer with ten classifications.
- ○ The vectorization of a pooling layer.

6. Why are nonlinear activation functions preferable? 1 point

- ○ Nonlinear activation functions are preferable because they are used in generalized linear models in statistics.
- ◉ Nonlinear activation functions increase the functional capacity of the neural network by allowing the representation of nonlinear relationships between features in input.
- ○ Nonlinear activation functions are preferable because they have been used historically.
- ○ Nonlinear activation functions are NOT preferable to linear ones, as they lose information in systems with high variance.

7. Which of the following are benefits of pooling? (Choose all that are correct.) 1 point

- ☐ Vectorizes the data.
- ☑ Combats overfitting.
- ☑ Reduces computational complexity.
- ☑ Encourages translational invariance.
- ☐ Decreases bias.

8. How are parameters that minimize the loss function found in practice? 1 point

- ○ Fractal geometry
- ○ Simplex algorithm
- ○ Gradient descent
- ◉ Stochastic gradient descent

9. Which of the following is an advantage of hierarchical representation of image features? 1 point

- ◉ Better leveraging all training data.
- ○ Decreasing variance in the model.
- ○ Eliminating bias.
- ○ Decreasing the computational complexity.

10. Why does transfer learning work? 1 point

- ○ Low-level features are specialized for a particular task, while top-level features are universal to all images.
- ◉ Top-level features are specialized for a particular task, while low-level features are universal to all images.
- ○ All images are composed of pixels with three color channels.
- ○ All layers of filters can be learned by studying the mammalian receptive fields.