CSP 571 Data Preparation and Analysis

Quiz - 5

Question 1

When performing PCA, features with significantly different variance values due to being in different units may result in biased loadings - this can be corrected typically via

- (●) a. scaling of (centered) features to unit standard deviation/variance.
- ( ) b. scaling of (un-centered) features to unit standard deviation/variance.
- ( ) c. removal of features with high variance.
- ( ) d. removal of features with low variance.

Question 2

When working with sequential data a **Recurrent Neural Network** (**RNN**), each element of the sequence *shares* the component of the model with other elements in the sequence

- (●) a. weights
- ( ) b. pools
- ( ) c. layers
- ( ) d. outputs

Question 3

When training a **Neural Network** with a qualitiative response, the following *loss function* is minimized

- ( ) a. mean-squared error
- ( ) b. logit
- ( ) c. softmax
- (●) d. cross-entropy

Question 4

Training of **Neural Networks** requires a large amount of *training* data in order to estimate the large number of

- ( ) a. layers
- (●) b. parameters
- ( ) c. encodings
- ( ) d. poolings

Question 5 10 points

In a simple feed-forward neural network which processes images, a) the Input Layer consists of single-color (grayscale) 30x30 pixels with values ranging from 0-255, b) the first Hidden Layer consists of 256 units, c) the second Hidden Layer consists of 128 units, and d) the Output Layer consisting of 10 one-hot encoded classes. The number of parameters to be estimated are

- ( ) a. $(30 \times 30) \times 256 + 256 \times 128 + 128 \times 10$
- (●) b. $((30 \times 30) + 1) \times 256 + (256 + 1) \times 128 + (128 + 1) \times 10$
- ( ) c. $(30 + 1) \times 256 + (256 + 1) \times 128 + (128 + 1) \times 10$
- ( ) d. $30 \times 256 + 256 \times 128 + 128 \times 10$

Question 6

Hierarchical clustering will use complete linkage to merge clusters with the minimum dissimilarity value across all pairs of clusters, with the dissimilarity value for each cluster pair calculated via the

- ( ) a. average of all pairwise dissimilarity values between observations within the pair of clusters.
- ( ) b. minimum of all pairwise dissimilarity values between observations within the pair of clusters.
- (●) c. maximum of all pairwise dissimilarity values between observations within the pair of clusters.
- ( ) d. centroid dissimilarity within the pair of clusters.

Question 7

Along with standard convolution filters, a **Convolutional Neural Network** (**CNN**) may contain which of the below types of *hidden layers*

- a. null
- b. pooling
- c. recurrent
- d. sequential

Question 8

When working with unlabeled data, i.e., no response variable, as part of exploratory data analysis we may perform the following type of learning:

- a. inferential.
- b. unsupervised.
- c. supervised.
- d. reinforcement.

Question 9

Selection of the number of principal components (PC) to be used for dimensionality reduction can be done via a scree plot, which shows for each PC the percentage of total or incremental explained

- a. mean.
- b. skewness.
- c. variance.
- d. kurtosis.

Question 10

Within the K-Means algorithm, after updating the centroid of each cluster, the following step is performed:

- a. each observation is assigned to a random cluster.
- b. each observation is assigned to the cluster with the smallest within-cluster squared error.
- c. each observation is assigned to the cluster with the farthest centroid.
- d. each observation is assigned to the cluster with the closest centroid.