Quiz 2 Answer Key

Question 1: What defines a support vector in an SVM model?

1. A data point that lies on the margin boundary (Correct)

2. A data point that lies within the margin

3. A data point that is misclassified

4. A data point that is farthest from the decision boundary

Question 2: In an SVM, setting a very large value for parameter C typically results in:

1. A narrower margin with fewer classification errors (Correct)

2. A wider margin with potentially more classification errors

3. No change to margin width

4. A model that ignores misclassifications

Question 3: What is the difference between a hard margin and a soft margin in SVM?

1. A soft margin allows some misclassifications controlled by C (Correct)

2. Hard margin uses kernel functions while soft margin does not

3. Hard margin works only for non-linear data

4. A soft margin enforces zero training error

Question 4: When might you prefer Manhattan (L1) distance over Euclidean (L2) distance in KNN?

1. When you want to reduce the influence of large individual coordinate differences (Correct)

2. When you want to penalize larger distances more heavily

3. When you have categorical features

4. When you need to compute cosine similarity

Question 5: What is the purpose of the "elbow method" in KMeans clustering?

1. To select the optimal number of clusters by plotting WCSS vs. K (Correct)

2. To initialize centroids using K-nearest means

Quiz 2 Answer Key

3. To measure cluster purity

4. To determine the maximum number of iterations

Question 6: Given the covariance matrix [[4, 2],[2, 3]], what is the largest eigenvalue (approximately

1. 5.56 (Correct)

2. 1.44

3. 3

4. 6

Question 7: In a 2D KNN classification with points A:(0,0) class 0, B:(1,0) class 1, C:(0,1) class 1, wh

1. 1 (Correct)

2. 0

3. Depends on distance weighting

4. Cannot decide without more data

Question 8: What is the entropy of a dataset with 9 positive and 5 negative samples? (Use log)

1. 0.94 (Correct)

2. 0.85

3. 0.78

4. 1

Question 9: When performing PCA, standardizing all features before computing the covariance mat

1. PCA to effectively use the correlation matrix instead of the covariance matrix (Correct)

2. PCA to ignore variables with small variance

3. PCA to produce identical components as without standardization

4. No change in the eigenvalue ratios

Question 10: In a decision tree split where the class probability at a node is p=0.7 for positive and 0

Quiz 2 Answer Key

1. 0.42 (Correct)

2. 0.21

3. 0.49

4. 0.58