Online Courses

Our courses are going online! As we are faced with the COVID-19 challenge, we are moving all our training courses to live streaming using Microsoft Teams/Zoom platforms until further notice.

Started on	Tuesday, July 2, 2024, 2:30 PM
State	Finished
Completed on	Tuesday, July 2, 2024, 3:48 PM
Time taken	1 hour 17 mins
Grade	**33.50** out of 34.00 (**98.53**%)

Question **1**

Complete

2.00 points out of 2.00

Describe how earlier, symbolic AI algorithms are different from current, connectionist ML algorithms?

Symbolic AI Algorithms: Symbolic AI, also known as classical AI, relies on explicit rules and logical operations. These systems are built using predefined rules and symbolic representations of knowledge. In symbolic AI, knowledge is represented using symbols and structures such as semantic networks, frames, and ontologies. This approach requires human experts to encode domain knowledge into a formal system.

Connectionist Machine Learning (ML) Algorithms: Connectionist ML, exemplified by neural networks and deep learning, relies on patterns and features inferred directly from data. These systems learn from large datasets through training processes that adjust model parameters to minimize error and improve performance. Instead of using predefined rules, connectionist ML models learn hierarchical representations of data.

Comment:

Question **2**

Complete

2.00 points out of 2.00

Provide a high-level overview of how the kNN algorithm works.

The k-Nearest Neighbors (kNN) algorithm works by storing all training data points. When a new data point (x_test) needs to be classified, kNN finds the k nearest neighbors from the training data (x_train) and predicts the class label (y_predict) for the new data point based on the majority vote of its neighbors' labels (y_train).

It is implemented using Sklearn library of python. The results of kNN algorithm depend of choice of various parameters like value of K, value of weight, value of metric, value of algorithm.

Comment:

Question **3**

Complete

4.00 points out of 4.00

Both precision and recall are scoring measures for ML algorithm performance. How are they defined? How do they differ? Name and describe one scoring method, other than precision and recall.

Precision and Recall are two methods used to evaluate the performance of ML algorithms in classification Algorithms tasks. They provide information on different aspects of a model's effectiveness.

Precision: It measures the proportion of positive predictions that are actually correct. It reflects how accurate your model is when it says something is positive.

Precision = True Positives / (True Positives + False Positives)

Recall: It measures the proportion of actual positive cases that the model correctly identifies. It reflects how well your model finds all the relevant positive cases.

Recall = True Positives / (True Positives + False Negatives)

Differences between Precision and Recall: Precision focuses on the quality of positive predictions or avoiding false positives, while Recall focuses on the completeness of positive predictions or avoiding missing true positives. Often, there's a trade-off between precision and recall. Increasing one might decrease the other.

Another scoring method is F1-Score. It provides a balanced view of both precision and recall.

F1 = 2 * (Precision * Recall) / (Precision + Recall)

Comment:

Question **4**

Complete

2.00 points out of 2.00

What is the difference between classification and regression tasks? What is the difference between classification and clustering tasks?

Difference between classification and regression tasks:
Classification classifies data points into predefined categories. Imagine sorting emails into spam or inbox. The output is a discrete label like "spam" or "not spam." Classification algorithms learn to identify patterns that separate these categories. Regression predicts continuous numerical values. For example, forecasting house prices. The output is a number on a spectrum, like a specific estimated value for a house. Regression algorithms uncover relationships between features and the continuous outcome variable.

Difference between classification and clustering tasks:

Clustering is different from both classification and regression. It groups data points together based on their similarities, without relying on predefined categories. Clustering helps uncover hidden structures within data, useful for tasks like customer segmentation or image categorization. Classification sorts data into predefined groups, like labeling emails as spam or not spam. Clustering, however, finds groups by itself based on similarities, like automatically grouping customers with similar purchase history.

Comment:

Question **5**

Complete

2.00 points out of 2.00

According to assessment theory, what type of assessment is used when fitting a ML model? What type of assessment is used when testing a ML model? How do they differ?

In Machine Learning, assessment theory has two main stages with distinct assessment types:

Model Fitting (Training): This stage utilizes formative assessment. Its goal is to iteratively improve the model's ability to learn patterns from the training data. Metrics like training loss or accuracy provide feedback on how well the model is grasping the underlying relationships within the data. This feedback allows adjustments to hyperparameters or model architecture to enhance learning.

Model Testing: Here, summative assessment works. We evaluate the model's generalizability and performance on fresh data. Metrics like test accuracy, precision, recall, or F1-score assess how well the trained model performs on data it hasn't encountered during training. This helps in determining if the model can effectively translate learned patterns to new situations.

Comment:

Question **6**

Complete

3.00 points out of 3.00

Provide a high-level overview of how stratified k-fold cross validation works. Why do we use it?

In ML, stratified k-fold cross-validation fixes an issue with imbalanced datasets. Regular k-fold cross-validation might split the data into folds where some classes are missing or underrepresented. Stratified k-fold fixes this by splitting the data into folds. Ensuring each fold has a similar proportion of classes as the entire dataset.

We use it because stratified k-fold validation gives a more reliable picture of a model's performance, especially for less frequent classes. It prevents folds with no data from a particular class, leading to a fairer and more generalizable evaluation of the model's overall effectiveness.

Comment:

Question **7**

Complete

3.50 points out of 4.00

Name two different categories of ensemble methods based on Decision Trees. How do they differ? Name one example method from each category.

Two different categories of Ensemble Methods Based on Decision Trees are

Bagging: It involves training multiple decision trees on different subsets of the training data, created by sampling with replacement (bootstrap samples). Each tree is trained independently, and their predictions are averaged (for regression) or voted upon (for classification) to produce the final prediction. Ex: Random Forest.

Boosting: It trains decision trees sequentially, with each tree focusing on correcting the errors made by the previous ones. The trees are not independent; each new tree is trained to improve the overall model by giving more weight to the instances that were previously misclassified. Ex: AdaBoost.

How do they differ:

In Training Process: Bagging trains trees independently and simultaneously on different samples of the data. Boosting trains trees sequentially, with each tree correcting the errors of the previous ones.

In their Focus: Bagging aims to reduce variance by averaging predictions from multiple independent trees. Boosting aims to reduce bias by focusing on difficult instances and iteratively improving the model.

Comment:

Question **8**

Complete

2.00 points out of 2.00

Name two data scaling transformations. How are they defined? How do they differ?

Two data scaling transformations are:

Normalization: Scales features to a fixed range, typically between 0 and 1 (or -1 and 1).

Standardization: Transforms features by subtracting the mean and then dividing by the standard deviation.

Differences:

Normalization sets a specific range for features, making them unitless. Standardization centers and scales data based on the training data's statistical properties (mean and standard deviation).

Comment:

Question **9**

Complete

2.00 points out of 2.00

Provide a high-level overview of how the Random Forest Classifier algorithm works.

The Random Forest Classifier is a powerful ensemble learning technique. It creates a collection of decision trees, each trained on a random subset of the data. Additionally, at each node of a tree, a random subset of features is considered for splitting. This injects randomness and helps prevent overfitting. When a new data point arrives, it's passed through each tree in the forest. Each tree makes a prediction based on its learned rules. Finally, the forest combines the predictions from all the trees. For classification tasks, the most frequent class prediction across all trees becomes the final output. This ensemble approach leads to a more robust and accurate prediction compared to a single decision tree.

Comment:

Question **10**

Complete

4.00 points out of 4.00

What is the difference between underfitting and overfitting an ML model? What is the relation to bias and variance?

Underfitting occurs when a model is too simple and fails to capture the underlying patterns in the training data. This leads to high bias and poor performance on both training and unseen data.

Overfitting happens when a model memorizes the training data too well, including noise and irrelevant details. This results in high variance and good performance on training data, but poor performance on new data.

Relation to bias and variance: Finding the right balance between bias and variance is important. Underfitting represents high bias and low variance, while overfitting represents low bias and high variance. We aim to achieve a sweet spot with moderate bias and variance for optimal generalizability.

Comment:

Question **11**

Complete

3.00 points out of 3.00

How is a ReLU neuron similar to a logistic regression classifier? What is the difference between the two?

A ReLU neuron is similar to a logistic regression classifier as both can be used for binary classification tasks, where the output indicates a probability of belonging to one of two classes of either 0 or 1.

Difference between the two: Logistic regression is a simpler linear model that uses a sigmoid function to transform the linear combination of inputs into a probability between 0 and 1. ReLU neurons introduce non-linearity through the rectified linear unit activation function. This allows them to capture more complex relationships in the data compared to the linear approach of the logistic regression. Logistic regression is a simpler approach, while ReLU neurons offer greater flexibility for modeling complex non-linear relationships when used within neural networks.

Comment:

Question **12**

Complete

2.00 points out of 2.00

What is the purpose of specifying a linkage method for clustering? Name and describe one linkage method.

In hierarchical clustering, specifying a linkage method determines how the distance between clusters is calculated during the merging process. This guides how clusters are formed and ultimately impacts the final cluster structure.

One linkage method is Single Linkage. It defines the distance between two clusters as the minimum distance between any two individual points, one from each cluster. For Ex: Imagine a chain of data points. Single linkage would connect the closest points, potentially forming elongated clusters.

Comment:

Question **13**

Complete

2.00 points out of 2.00

How does hierarchical clustering differ from K-means clustering?

Differences between hierarchical clustering and K-means clustering:

Number of Clusters: K-means requires predefining the desired number of clusters (k). Hierarchical doesn't require specifying k upfront. It builds a hierarchy of clusters, allowing you to decide on an appropriate number later based on the structure or a threshold distance.

Type of Clustering: K-means performs partitional clustering. It assigns each data point to a single, specific cluster. Hierarchical performs hierarchical clustering. It creates a hierarchy or tree structure that shows how data points are grouped at different distance thresholds. You can cut the hierarchy at a desired level to get a specific number of clusters.

Output: K-means provides a set of data points assigned to their respective clusters. Hierarchical provides a dendrogram, which is a tree-like structure visualizing the relationships between data points and how they are progressively merged into clusters based on similarity.

Comment: