Practice Quiz - Results

Attempt 1 of Unlimited

Written Oct 19, 2024 11:54 PM - Oct 20, 2024 12:11 AM

Attempt Score **9 / 25 - 36 %**

Overall Grade (Highest Attempt) **9 / 25 - 36 %**

Question 4 — 0 / 1 point

Which of the below observations best describes the bias-variance trade off?

- ✗ ☑ A model with high bias always performs well on test data but poorly on training data due to underfitting.
- ✓ ☐ Both bias and variance can be minimized simultaneously, resulting in perfect model performance.
- ➡ ✓ ☑ A model with high variance performs well on training data but poorly on test data due to overfitting.
- ✓ ☐ A model with high bias consistently performs poorly on both training and test data due to oversimplification.

Question 5 — 0 / 1 point

Which of the below sentences are true in context to a stratified split? Select all that apply.

- ✗ ☐ Stratified splitting retains the original distribution of classes, which helps mitigate the risk of biased model evaluation on imbalanced datasets.
- ✓ ☐ Stratified splitting is only applicable to binary classification problems and cannot be used for multi-class classification.
- ✓ ☐ Stratified splitting will lead to overfitting if the training set is too small relative to the number of classes.
- ✓ ☐ Stratified splitting ensures that the sample sizes of minority classes are increased in the training set, even if the original dataset is imbalanced.

Question 6 — 0 / 1 point

What type of dependent variable does logistic regression predict?

- ✓ ☐ Continuous
- ✗ ☐ Multinomial
- ✓ ☑ Binary
- ✓ ☐ Ordinal

Question 7 — 0 / 1 point

Which of the following is true for non-parametric models? Select all that apply.

- ✗ ☑ They have lower risk of overfitting.
- ✗ ☐ They are capable of fitting many functional forms.

- ✓ ☐ It is easy to explain these models.
- ➡ ✗ ☐ They are slower to train.

Question 9 — 0 / 1 point

For an intended output **z**, and a predicted classifier score y, hinge loss is defined as :

$$\ell(y) = \max(0, 1 - z.y)$$

Which of the following is correct for hinge loss? Select all that apply.

- ➡ ✓ ☑ Hinge loss is not differentiable everywhere.
- ✓ ☐ Hinge loss cannot be minimized using gradient descent
- ➡ ✓ ☑ Hinge loss is zero for observations on the correct side of the margin (when margin corresponds to value 1)
- ➡ ✗ ☐ Hinge loss is convex function.

Question 10 — 0 / 1 point

Which of the following metrics is NOT typically used to evaluate the performance of a logistic regression model?

- ◯ Recall
- ➡ ◯ R-squared
- ◯ Precision

✗ ◉ Confusion matrix

Question 11 — 0 / 1 point

Which of the following statements correctly differentiates between **epistemic** and **aleatoric uncertainty**?

- ➡ ✗ ☐ **Epistemic** uncertainty can be reduced by augmenting data and improving model quality, whereas **aleatoric** uncertainty is irreducible and related to inherent randomness.
- ➡ ✗ ☐ **Aleatoric** uncertainty arises from inherent variability in the data, while **epistemic** uncertainty stems from a lack of knowledge or information about the model.
- ✓ ☐ Both types of uncertainty can be addressed through Ensemble methods.
- ➡ ✗ ☐ **Epistemic** uncertainty is captured by individual predictions while **Aleatoric** uncertainty is captured by model behavior over multiple predictions.

Question 12 — 0 / 1 point

In active learning, what is the primary goal of using entropy-based uncertainty sampling technique?

- ○ To uniformly sample data points across all classes to ensure balanced representation.
- ✗ ◉ To select data points that are closest to the decision boundary for labeling.
- ○ To minimize the overall size of the dataset by removing redundant samples, and improving data quality.
- ➡ ○ To choose data points with the highest uncertainty in their predicted class probabilities, maximizing information gain.

Question 15 — 0 / 1 point

What does the concept of Equalized Odds aim to achieve in machine learning fairness?

- It ensures that all demographic groups have the same overall accuracy in the model's predictions.
- ✗ It focuses on ensuring that all individuals have the same predicted probabilities of positive outcomes, irrespective of actual outcomes.
- It mandates that the model's predictions are independent of sensitive attributes, regardless of their effect on outcomes.
- ➡ It requires that the false positive rates and true positive rates are equal across different demographic groups.

Question 17 — 0 / 1 point

A binary classification model predicts the probability of an individual being invited for an interview. If the probability is above 0.5, the individual is invited for the interview (positive class labelled as 1) otherwise the individual is not invited for the interview (negative class labelled as 0). For two groups of individuals A and B, the below confusion matrix is yielded by the classifier.

Group A	True Label =1	True Label =0
Predicted Label = 1	10	30
Predicted Label = 0	20	40

Group B	True Label =1	True Label =0
Predicted Label = 1	20	20
Predicted Label = 0	10	30

Which of the following is true for **Equal Opportunity** measure of classifier fairness? Assume beta = 0.2 wherever necessary. Select all that apply.

- ✗ ☐ The equal opportunity measure of classifier is concerning, since the difference is greater than beta
- ✓ ☑ The classifier is unfair and biased against group A.
- ✓ ☐ The classifier is fair to both groups.
- ✓ ☐ The equal opportunity measure of classifier is not concerning, since the difference is less than beta

Question 18 **0 / 1 point**

A specific measure of algorithmic fairness evaluates wheather similar individuals receive similar outcomes from the ML model. Which fairness measure is this referring to?

- ✗ ◉ Predictive Parity
- ○ Counterfactual Fairness
- ○ Treatment Equality
- ○ Individual Fairness

Question 19 **0 / 1 point**

Which of the following describes the key difference between L1 and L2 Loss functions?

- ✓ ☐ L1 loss is always preferred in regression models, while L2 loss is not applicable in this context.
- ✗ ☐ L2 loss penalizes larger errors more heavily than L1 loss.
- ✗ ☑ L1 loss is more sensitive to outliers compared to L2 loss, which treats all errors equally.
- ✗ ☐ L1 Loss minimizes the Absolute Error, while L2 Loss minimizes Squared Error.

Question 20 — 0 / 1 point

In the context of a binary classification problem, what does cross-entropy loss measure?

- ○ It measures the average number of bits needed to communicate a set of events.
- ○ It measures the average distance between predicted and actual values, penalizing larger errors more heavily.
- ○ It calculates the variance of the predictions and distributes them evenly across classes.
- ➡ ○ It measures the difference between the predicted probability distribution and the true distribution, and indicates the confidence of predictions.

Question 22 — 0 / 1 point

Gradient Descent is useful if the learning rate is optimally selected. What happens to the algorithm outcome if the learning rate is set too high in gradient descent?

- ➡ ✓ ☑ The algorithm will take larger steps.
- ✓ ☐ The algorithm will perfectly fit the training data.
- ➡ ✓ ☑ The algorithm may overshoot the global minimum and diverge.
- ➡ ✗ ☐ The algorithm may converge faster.

Question 24 — 0 / 1 point

What is the main advantage of using mini-batch stochastic gradient descent (SGD), over batch gradient descent?

- Mini-batches increase the computational cost per iteration compared to batch GD.
- Mini-batches introduce noise that can help escape local minima.
- Mini-batches guarantee convergence to the global minimum.
- Mini-batches provide a more accurate estimate of the gradient.

Question 25 — 0 / 1 point

What are the causes of algorithmic biases? Select all that apply.

- [ ] Using machines generated synthetic data.
- [x] Poorly designed and deployed data annotation process.
- [x] Bad selection of training objective function
- [x] Inadequately selected feature sets.

Done