Question 1.

Marks : 2.00

Fill your answer within the blanks provided below

If a boosting model has 15 stages and each stage classifier has an error rate of 0.25, what is the final boosted model's error rate? (upto 4 decimal points)

0.0053

Question 2.

Marks : 2.00

Fill your answer within the blanks provided below

For a Gradient Boosting model, if the base learner's error is 0.4 and you use 50 rounds with a learning rate of 0.2, what is the cumulative error reduction?(upto 4 decimal points)

0.0002 (

Question 3.

Marks : 2.00

Fill your answer within the blanks provided below

The gamma parameter in XGBoost is used to prevent overfitting by controlling the complexity of the model.

Question 4.

Marks : 2.00

What is the primary difference between model-based and model-free reinforcement learning?

Answer

- ○ Model-based learning is only applicable to episodic tasks, while model-free learning can be used for continuous tasks. Next
- ○ Model-based learning uses a trial-and-error approach, while model-free learning uses a model of the environment to find the optimal policy.
- ○ Model-free learning is a type of supervised learning, while model-based learning is a type of unsupervised learning.
- ◉ Model-based learning requires a model of the environment (states, rewards, transitions) to find an optimal policy, while model-free learning derives an optimal policy directly from interactions with the environment.

Question 5. **Marks : 2.00**

What role do weak learners play in ensemble methods?

Answer

- ○ They are not used in ensemble methods.
- ○ They contribute equally to the final prediction.
- ◉ They are models that perform poorly on their own.
- ○ They are models with high computational requirements.

Question 6.

Marks : 2.00

Fill your answer within the blanks provided below

To prevent overfitting, ensemble methods often use techniques like regularization and cross-validation .

Question 7. **Marks : 2.00**

In ensemble learning, what is bagging primarily used for?

Answer

- ○ Reducing computational complexity
- ○ Speeding up training time
- ○ Enhancing feature selection
- ◉ Improving generalization by reducing variance

Question 8.

Marks : 2.00

What is Random Forest?

Answer

- ○ Designed for reducing variance only
- ○ Not suitable for high-dimensional data
- ◉ A specific type of ensemble learning method
- ○ Similar to boosting but with fewer computational resources

Question 9. **Marks : 2.00**

Which of the following best describes the exploration-exploitation trade-off in reinforcement learning?

Answer

- ○ The decision of how much to discount future rewards.
- ◉ The trade-off between exploring the environment to find new states and exploiting known states to get rewards.
- ○ The choice between using a model-based or a model-free learning approach.
- ○ The balance between training the agent for a longer or shorter period of time. Back Done

Question 10. **Marks : 2.00**

In Q-learning, what is the purpose of the Q-function (Q(s, a))?

Answer

- (●) To estimate the maximum discounted future reward an agent can expect by taking action 'a' in state 's' and acting optimally thereafter.
- ( ) To calculate the immediate reward received after taking action 'a' in state 's'.
- ( ) To store a history of all actions taken by the agent in the environment.
- ( ) To determine the probability of transitioning to a new state 's' after taking action 'a'.

Question 11. **Marks : 2.00**

Which ensemble learning method is known for its feature importance calculation?

Answer

- ○ Boosting
- ○ Bagging
- ○ Stacking
- ◉ Random Forest

Question 12.

Marks : 2.00

What distinguishes boosting from bagging?

Answer

- ○ Boosting focuses on reducing variance, while bagging focuses on reducing bias
- ◉ Boosting trains models sequentially, while bagging trains them in parallel
- ○ Boosting combines models of different types, while bagging uses identical models
- ○ Boosting requires a larger dataset than bagging

Question 13.

Marks : 2.00

Fill your answer within the blanks provided below

An XGBoost model is trained with the hyperparameters eta=0.1, max_depth=4, subsample=0.9, and 200 rounds. How many trees are created?(round off to closest integer)

200

Question 14. **Marks : 2.00**

How does data leakage affect model performance?

Answer

- [x] It reduces model accuracy
- [x] It decreases model interpretability
- [ ] It improves model accuracy
- [ ] It has no effect on model performance

Question 15. **Marks : 2.00**

Which tools or techniques can be used for feature engineering?
(Select all that apply)

Answer

- [x] Polynomial features
- [x] One-hot encoding
- [x] Label encoding
- [x] Scaling