Quiz 1

Total points 7/10

CSL7620: Machine Learning | AY 2024-25 | Sem I

The respondent's email (**m24cse029@iitj.ac.in**) was recorded on submission of this form.

✓ **A key characteristic of unsupervised learning is: ***

1/1

- ○ Learning from rewards
- ◉ Discovering hidden patterns in data ✓
- ○ Predicting future outcomes
- ○ None of the others

For a given hypothesis space, if the VC dimension is d, which of the following is/are correct? *0/1

- [x] The hypothesis space cannot shatter any set of d+1 points ✓
- [x] The hypothesis space can perform correct binary classification for any set of d points ✗
- [x] The hypothesis space must not consist of only 2D straight lines if d>3 in 2D Euclidean plane ✓
- [ ] None of the others

Correct answer

- [x] The hypothesis space cannot shatter any set of d+1 points
- [x] The hypothesis space must not consist of only 2D straight lines if d>3 in 2D Euclidean plane

✓ **Which of the following statements correctly describes how reinforcement learning works?** *1/1

- [ ] Reinforcement learning does not require interaction with the environment to learn optimal actions.
- [ ] The agent learns to make decisions by directly performing the optimal actions without feedback.
- [x] The learning process in reinforcement learning is driven by experiences through rewards and penalties from the environment. ✓
- [x] Reinforcement learning involves controlling an agent through a sequence of good decisions. ✓

✓ **Using VC-Dimension, we can get an idea about which of the following?** * 1/1

- ○ The number of features in the dataset
- ◉ A model's capacity of classification ✓
- ○ The size of the dataset required for training
- ○ Probability of some limit of test error of a model

✓ **Which of the following techniques are used for data normalization? *** 1/1

- [x] Z-score Normalization ✓
- [ ] One-Hot Encoding
- [x] Min-Max Scaling ✓
- [ ] Log Transformation

✓ **What is the primary objective of a regression model? *** 1/1

- ( ) To reduce the number of features in the dataset
- (•) To find a curve that best fits the data points ✓
- ( ) To predict categorical labels
- ( ) To cluster similar data points together

✗ **Which of the following scenarios usually is indicative of high variance in a model?** *0/1

- [ ] None of the others
- [x] Large differences between training and testing accuracy ✓
- [ ] High VC dimension
- [ ] More inductive bias

Correct answer

- [x] Large differences between training and testing accuracy
- [x] High VC dimension

✓ **Which of the following statements is/ are usually true about bias and variance?** *1/1

- ○ Increasing the complexity of a model always reduces both bias and variance.
- ○ Bias and variance are independent of each other.
- ◉ High variance usually leads to overfitting ✓
- ○ High bias leads to high variance in a model.

Which of the following statements is/ are usually true about inductive learning? *0/1

- [x] Inductive learning assumes that the patterns in the training data will hold true for unseen data. ✓
- [x] Overfitting is a common problem in inductive learning when a model learns the training data too well, including noise. ✓
- [ ] Inductive learning models can be evaluated based on their ability to predict outcomes on new, unseen data.
- [ ] Inductive learning can only be applied to supervised learning scenarios.

Correct answer

- [x] Inductive learning assumes that the patterns in the training data will hold true for unseen data.
- [x] Overfitting is a common problem in inductive learning when a model learns the training data too well, including noise.
- [x] Inductive learning models can be evaluated based on their ability to predict outcomes on new, unseen data.

✓ **Consider the following scenario for Stochastic Gradient Descent (SGD):** *1/1

You are training a linear regression model using SGD on a single data point. The equation for the output is $y=\theta(0)+\theta(1)x$. The learning rate is set to 0.1. Suppose at a certain iteration, the model parameters $\theta(0)$ and $\theta(1)$ are $\theta(0)=2.0$ and $\theta(1)=1.5$, respectively. The current data point you use to update the model is $x=4$, and the true label y is 10.

Which of the following are correct statements after one SGD update for the parameters $\theta(0)$ and $\theta(1)$?

- ○ New value of $\theta(0)$ and $\theta(1)$ will be 1.8 and 1.3, respectively.
- ◉ New value of $\theta(0)$ and $\theta(1)$ will be 2.2 and 2.3, respectively. ✓
- ○ New value of $\theta(0)$ and $\theta(1)$ will be 1.8 and 0.7, respectively.
- ○ New value of $\theta(0)$ and $\theta(1)$ will be 2.2 and 1.7, respectively.