**ASSIGNMENT**

**1. What does one mean by the term &quot;machine learning&quot;?**

**2.Can you think of 4 distinct types of issues where it shines?**

**3.What is a labeled training set, and how does it work?**

**4.What are the two most important tasks that are supervised?**

**5.Can you think of four examples of unsupervised tasks?**

**6.State the machine learning model that would be best to make a robot walk through various**

**unfamiliar terrains?**

**7.Which algorithm will you use to divide your customers into different groups?**

**8.Will you consider the problem of spam detection to be a supervised or unsupervised learning**

**problem?**

**9.What is the concept of an online learning system?**

**10.What is out-of-core learning, and how does it differ from core learning?**

**11.What kind of learning algorithm makes predictions using a similarity measure?**

**12.What&#39;s the difference between a model parameter and a hyperparameter in a learning**

**algorithm?**

**13.What are the criteria that model-based learning algorithms look for? What is the most popular**

**method they use to achieve success? What method do they use to make predictions?**

**14.Can you name four of the most important Machine Learning challenges?**

**15.What happens if the model performs well on the training data but fails to generalize the results**

**to new situations? Can you think of three different options?**

**16.What exactly is a test set, and why would you need one?**

**17.What is a validation set&#39;s purpose?**

**18.What precisely is the train-dev kit, when will you need it, how do you put it to use?**

**19.What could go wrong if you use the test set to tune hyperparameters?**

**SOLUTIONS**

1. ***Machine learning is a subset of artificial intelligence that enables computers to learn from data without being explicitly programmed. It is a process of training a computer program to recognize patterns in data and make predictions or decisions based on that data.***
2. ***Machine learning shines in several distinct types of issues, including image recognition, natural language processing, anomaly detection, and recommendation systems.***
3. ***A labeled training set is a dataset that contains input data along with the correct output values, or labels, for that data. It is used to train a machine learning model by feeding it examples of input-output pairs so that it can learn to predict the output for new inputs.***
4. ***The two most important tasks that are supervised are classification and regression.***
5. ***Four examples of unsupervised tasks are clustering, dimensionality reduction, anomaly detection, and association rule learning.***
6. ***The machine learning model that would be best to make a robot walk through various unfamiliar terrains is a reinforcement learning algorithm.***
7. ***The algorithm that will be used to divide customers into different groups will depend on the specific problem and data, but some popular algorithms include k-means clustering, hierarchical clustering, and Gaussian mixture models.***
8. ***Spam detection is typically considered a supervised learning problem since the algorithm needs to be trained on labeled data to distinguish between spam and non-spam messages.***
9. ***An online learning system is a machine learning model that can learn and update its predictions in real-time as new data becomes available. It is particularly useful in situations where the data is constantly changing or where a quick response is needed.***
10. ***Out-of-core learning is a technique used to train machine learning models on datasets that are too large to fit in memory. It differs from core learning, which assumes that the entire dataset can be loaded into memory.***
11. ***A learning algorithm that makes predictions using a similarity measure is known as a nearest neighbor algorithm.***
12. ***A model parameter is a variable that is learned from the training data, while a hyperparameter is a setting that is set before training begins and affects the behavior of the learning algorithm.***
13. ***Model-based learning algorithms look for the best model that fits the data while minimizing a specific cost function. The most popular method they use to achieve success is gradient descent, and they use the learned model to make predictions on new data.***
14. ***Four of the most important Machine Learning challenges are data quality, overfitting, selecting appropriate algorithms, and interpreting results.***
15. ***If the model performs well on the training data but fails to generalize the results to new situations, three different options are to collect more data, use a simpler model, or use regularization techniques.***
16. ***A test set is a dataset that is used to evaluate the performance of a machine learning model after it has been trained on a training set. It is used to estimate the model's ability to generalize to new data.***
17. ***The purpose of a validation set is to evaluate the performance of a model during the training process and to tune hyperparameters to improve its performance.***
18. ***A train-dev kit is a subset of the training data that is used to evaluate the performance of a model during development. It is useful when the training data is small, and it is used to detect overfitting and to fine-tune the model.***
19. ***If you use the test set to tune hyperparameters, you risk overfitting the model to the test set, which can lead to poor generalization performance on new data. It is better to use a validation set to tune hyperparameters and then evaluate the model's performance on the test set.***