GEEKSforGEEKS

MACHINE LEARNING ROADMAP

Join Us Now

DAY 1

Introduction to Machine Learning

Topics

1. Understand the basics of machine learning.
2. Learn about supervised, unsupervised, and reinforcement learning.
3. Familiarize yourself with common machine learning terminology.

Practice Questions:

1. Define machine learning and distinguish between supervised, unsupervised, and reinforcement learning.
2. Provide examples of real-world applications for each type of machine learning.
3. Explain the difference between classification and regression tasks in machine learning.

DAY 2

Python Basics

Topics

1. Learn Python fundamentals such as variables, data types, and basic operations.
2. Explore control structures like loops and conditional statements.
3. Understand functions and modules in Python.

Practice Questions:

1. What are the advantages of using Python for machine learning over other programming languages?
2. Write a Python function to calculate the factorial of a given number.
3. Explain the difference between Python lists and tuples.

DAY 3

NumPy and Pandas

Topics

1. Install NumPy and Pandas libraries.
2. Learn about NumPy arrays and basic operations.
3. Understand Pandas data structures like Series and DataFrame.

Practice Questions:

1. Create a NumPy array containing integers from 1 to 10 and calculate its mean and standard deviation.
2. Read a CSV file into a Pandas DataFrame and display the first 5 rows.
3. Explain the purpose of broadcasting in NumPy.

DAY 4

Data Visualization with Matplotlib and Seaborn

Topics

1. Install Matplotlib and Seaborn libraries.
2. Create basic plots using Matplotlib.
3. Explore advanced visualization techniques with Seaborn.

Practice Questions:

1. Create a line plot using Matplotlib to visualize the trend of a stock price over time.
2. Plot a histogram of a dataset using Seaborn and customize the color and bin size.
3. Compare the distribution of two different features in a dataset using a box plot.

DAY 5

Linear Regression

Topics

1. Understand the concept of linear regression.
2. Implement linear regression using Python libraries.
3. Evaluate and interpret the results of linear regression.

Practice Questions:

1. Implement simple linear regression using Python and NumPy on a sample dataset.
2. Interpret the meaning of the coefficients in a linear regression model.
3. Evaluate the performance of a linear regression model using metrics such as mean squared error or R-squared.

DAY 6

Logistic Regression

Topics

1. Learn about logistic regression and its applications.
2. Implement logistic regression for classification problems.
3. Evaluate model performance using accuracy, precision, and recall.

Practice Questions:

1. Explain the difference between logistic regression and linear regression.
2. Implement logistic regression using scikit-learn on a binary classification problem.
3. Interpret the odds ratio in the context of logistic regression coefficients.

DAY 7

Decision Trees and Random Forests

Topics

1. Understand decision trees and ensemble methods.
2. Implement decision tree and random forest classifiers.
3. Tune hyperparameters for better model performance.

Practice Questions:

1. Build a decision tree classifier using scikit-learn on a sample dataset and visualize the resulting tree.
2. Explain how random forests combine multiple decision trees to improve predictive performance.
3. Discuss the concept of feature importance in random forests and how it can be used for feature selection.

DAY 8

Model Evaluation Techniques

Topics

1. Learn about cross-validation and its importance.
2. Implement k-fold cross-validation.
3. Understand bias-variance tradeoff and overfitting/underfitting.

Practice Questions:

1. Explain the purpose of cross-validation in machine learning model evaluation.
2. Implement k-fold cross-validation on a dataset using scikit-learn.
3. Discuss the impact of bias and variance on model performance and how to address them.

DAY 9

Support Vector Machines (SVM)

Topics

1. Understand the theory behind support vector machines.
2. Implement SVM for classification problems.
3. Explore kernel tricks and SVM applications.

Practice Questions:

1. Describe the concept of a support vector in SVMs and its role in defining the decision boundary.
2. Implement SVM classification using scikit-learn on a sample dataset.
3. Discuss the importance of kernel functions in SVMs and provide examples of commonly used kernels.

DAY 10

K-Nearest Neighbors (KNN)

Topics

1. Learn about the K-nearest neighbors algorithm.
2. Implement KNN for classification and regression.
3. Understand the impact of choosing different values of K.

Practice Questions:

1. Explain how the K-nearest neighbors algorithm works for both classification and regression.
2. Implement KNN classification using scikit-learn on a sample dataset.
3. Discuss the impact of choosing different values of K on the performance of the KNN algorithm.

DAY 11

Dimensionality Reduction with PCA

Topics

1. Understand the concept of dimensionality reduction.
2. Implement Principal Component Analysis (PCA).
3. Explore applications of PCA in feature extraction and visualization.

Practice Questions:

1. Describe the goal of dimensionality reduction and how PCA achieves it.
2. Implement PCA using scikit-learn on a high-dimensional dataset and visualize the reduced dimensions.
3. Discuss the trade-off between explained variance and the number of principal components retained.

DAY 12

Clustering with K-Means

Topics

1. Learn about unsupervised learning and clustering.
2. Implement the K-means clustering algorithm.
3. Evaluate clustering performance using metrics like silhouette score.

Practice Questions:

1. Explain the concept of clustering and how K-means algorithm partitions data into clusters.
2. Implement K-means clustering using scikit-learn on a sample dataset and visualize the resulting clusters.
3. Discuss the challenges of choosing the optimal number of clusters in K-means and potential solutions.

DAY 13

Natural Language Processing (NLP) Basics

Topics

1. Understand the basics of NLP.
2. Learn about tokenization, stemming, and lemmatization.
3. Explore text preprocessing techniques.

Practice Questions:

1. Describe the preprocessing steps involved in preparing text data for NLP tasks.
2. Implement tokenization, stemming, and lemmatization using NLTK or spaCy on a sample text.
3. Discuss the importance of text normalization in NLP and provide examples of normalization techniques.

GEEKSforGEEKS

DAY 14

Text Classification with Naive Bayes

Topics

1. Learn about the Naive Bayes classifier.
2. Implement text classification using Naive Bayes.
3. Evaluate classifier performance using metrics like accuracy and F1-score.

Practice Questions:

1. Explain the principle behind the Naive Bayes classifier and its assumption of conditional independence.
2. Implement text classification using the Multinomial Naive Bayes classifier in scikit-learn on a text dataset.
3. Discuss the strengths and weaknesses of the Naive Bayes classifier for text classification tasks.

DAY 15

Sentiment Analysis

Topics

1. Understand sentiment analysis and its applications.
2. Implement sentiment analysis using NLP techniques.
3. Explore different approaches to sentiment analysis.

Practice Questions:

1. Describe the goal of sentiment analysis and its applications in analyzing textual data.
2. Implement sentiment analysis using lexicon-based approaches or machine learning classifiers on a sample text dataset.
3. Discuss the challenges of sentiment analysis, such as handling sarcasm and context, and potential solutions.

DAY 16

Introduction to Neural Networks

Topics

1. Understand the basics of neural networks.
2. Learn about activation functions and feedforward neural networks.
3. Implement a simple neural network using Python libraries.

Practice Questions:

1. Explain the basic architecture of a feedforward neural network and the role of input, hidden, and output layers.
2. Implement a simple neural network using a library like TensorFlow or Keras to solve a classification problem.
3. Discuss the concept of activation functions and their importance in neural networks.

GEEKSfORGEEKS

DAY 17

Deep Learning with TensorFlow / Keras

Topics

1. Install TensorFlow and Keras libraries.
2. Learn about deep learning concepts like layers, loss functions, and optimizers.
3. Implement a deep learning model for image classification.

Practice Questions:

1. Describe the difference between TensorFlow and Keras and their roles in deep learning development.
2. Implement a deep learning model using TensorFlow/Keras for image classification on a sample dataset like MNIST.
3. Discuss common deep learning optimization techniques like stochastic gradient descent and Adam optimization.

DAY 18

Convolutional Neural Networks (CNNs)

Topics

1. Understand the architecture of convolutional neural networks.
2. Implement a CNN for image classification tasks.
3. Fine-tune CNN hyperparameters for better performance.

Practice Questions:

1. Explain the architecture of a convolutional neural network (CNN) and the purpose of convolutional and pooling layers.
2. Implement a CNN using TensorFlow/Keras for image classification on a dataset like CIFAR-10 or Fashion MNIST.
3. Discuss the concept of transfer learning and how pre-trained CNN models can be utilized for new tasks.

DAY 19

Recurrent Neural Networks (RNNs)

Topics

1. Learn about recurrent neural networks and their applications.
2. Implement a simple RNN for sequential data analysis.
3. Explore long short-term memory (LSTM) networks.

Practice Questions:

1. Describe the architecture of a recurrent neural network (RNN) and its ability to handle sequential data.
2. Implement a basic RNN using TensorFlow/Keras for sequence prediction on a dataset like stock prices or text.
3. Discuss common challenges with traditional RNNs like the vanishing gradient problem and solutions like Long Short-Term Memory (LSTM) networks.

DAY 20

Transfer Learning

Topics

1. Understand transfer learning and its advantages.
2. Implement transfer learning using pre-trained models.
3. Fine-tune pre-trained models for specific tasks.

Practice Questions:

1. Explain the concept of transfer learning and its benefits in deep learning applications.
2. Implement transfer learning using pre-trained models like VGG or ResNet on a custom dataset for image classification.
3. Discuss strategies for fine-tuning pre-trained models and selecting appropriate layers for transfer learning.

DAY 21

Reinforcement Learning Basics

Topics

1. Learn about reinforcement learning and its components.
2. Understand Markov Decision Processes (MDPs).
3. Implement a basic reinforcement learning algorithm.

Practice Questions:

1. Describe the basic components of a reinforcement learning problem, including agents, environments, and rewards.
2. Implement a simple reinforcement learning algorithm like Q-learning for solving a grid-world problem.
3. Discuss the trade-off between exploration and exploitation in reinforcement learning and methods to balance them.

DAY 22

Q-Learning

Topics

1. Learn about Q-learning and its applications.
2. Implement Q-learning for simple reinforcement learning problems.
3. Understand exploration-exploitation tradeoff.

Practice Questions:

1. Explain the Q-learning algorithm and its approach to learning optimal policies in reinforcement learning.
2. Implement Q-learning using Python for solving a simple environment like the OpenAI Gym Taxi problem.
3. Discuss the limitations of Q-learning in handling large state spaces and potential solutions like function approximation.

GEEKSFORGEEKS

DAY 23

Deep Q-Learning

Topics

1. Learn about deep Q-learning and its advantages.
2. Implement deep Q-learning algorithms like Deep Q-Networks (DQN).
3. Explore extensions such as Double DQN and Dueling DQN.

Practice Questions:

1. Describe the concept of deep Q-learning and its extension of Q-learning using neural networks.
2. Implement Deep Q-Networks (DQN) using TensorFlow/Keras for solving Atari games or similar environments.
3. Discuss techniques to improve stability and performance in deep Q-learning, such as experience replay and target networks.

DAY 24

Policy Gradient Methods

Topics

1. Understand policy gradient methods for reinforcement learning.
2. Implement basic policy gradient algorithms like REINFORCE.
3. Explore advanced techniques like Actor-Critic methods.

Practice Questions:

1. Explain the principles behind policy gradient methods and their advantages over value-based methods.
2. Implement the REINFORCE algorithm using TensorFlow/Keras for training a policy network on a custom environment.
3. Discuss common challenges in policy gradient methods like high variance and methods to address them like baselines and variance reduction techniques.

GEEKSforGEEKS

DAY 25

Advanced Topics: Generative Adversarial Networks (GANs)

Topics

1. Learn about GANs and their applications in generating synthetic data.
2. Implement a basic GAN architecture.
3. Explore applications of GANs in image generation and data augmentation.

Practice Questions:

1. Describe the architecture of a Generative Adversarial Network (GAN) and the roles of the generator and discriminator networks.
2. Implement a basic GAN using TensorFlow/Keras for generating synthetic images on a dataset like MNIST or CIFAR-10.
3. Discuss challenges in training GANs like mode collapse and strategies to overcome them like Wasserstein GANs.

DAY 26

Advanced Topics: Variational Autoencoders (VAEs)

Topics

1. Understand variational autoencoders and their applications.
2. Implement a VAE for unsupervised learning tasks.
3. Explore applications of VAEs in generating structured data.

Practice Questions:

1. Explain the concept of variational autoencoders (VAEs) and their use in unsupervised learning and generative modeling.
2. Implement a VAE using TensorFlow/Keras for generating synthetic data on a custom dataset like faces or handwritten digits.
3. Discuss the trade-offs between VAEs and GANs in terms of training stability, sample quality, and interpretability.

DAY 27

Deployment and Model Serving

Topics

1. Learn about deploying machine learning models to production.
2. Explore frameworks like Flask and FastAPI for building APIs.
3. Deploy a machine learning model using cloud platforms like AWS or Azure.

Practice Questions:

1. Describe the process of deploying a machine learning model to production, including considerations for scalability, latency, and reliability.
2. Implement a simple Flask or FastAPI application for serving a trained machine learning model as a RESTful API.
3. Discuss best practices for model versioning, monitoring, and updating in production environments.

DAY 28

Model Monitoring and Maintenance

Topics

1. Understand the importance of model monitoring and maintenance.
2. Learn about tools and techniques for monitoring model performance.
3. Implement a basic monitoring system for deployed models.

Practice Questions:

1. Explain the importance of model monitoring and maintenance in production machine learning systems.
2. Implement a basic monitoring system for tracking model performance metrics like accuracy and latency over time.
3. Discuss common issues that can arise in deployed machine learning models and strategies for debugging and troubleshooting.

DAY 29

Ethics and Bias in Machine Learning

Topics

1. Learn about ethical considerations in machine learning.
2. Understand sources of bias in machine learning models.
3. Explore techniques for mitigating bias in machine learning systems.

Practice Questions:

1. Discuss the ethical considerations involved in designing and deploying machine learning systems, including issues related to fairness, privacy, and transparency.
2. Describe common sources of bias in machine learning models and data, such as selection bias and algorithmic bias.
3. Discuss approaches for mitigating bias in machine learning systems, including data preprocessing techniques, algorithmic fairness measures, and diverse model training.

DAY 30

Review and Project

Topics

1. Review key concepts covered in the past 29 days.
2. Work on a machine learning project or participate in a Kaggle competition.
3. Reflect on your learning journey and identify areas for further improvement.

Practice Questions:

1. Reflect on your learning journey over the past 29 days and identify key concepts and skills you've acquired.
2. Work on a machine learning project or participate in a Kaggle competition to apply your knowledge and skills to a real-world problem.
3. Present your project or competition results to peers or mentors, discussing your approach, challenges faced, and lessons learned.