Data Science and Machine Learning with Python: From Fundamentals to Advanced Techniques

Learning Agenda:

- *Python Fundamentals and Data Analysis with Libraries*
- *Mathematics for Machine Learning*
- *Introduction to Databases and SQL &NoSQL, Web Scraping Techniques and Tools*
- *Fundamentals of Machine Learning*
- *Deep Learning and Neural Networks*
- *Natural Language Processing Fundamentals*
- *Computer Vision and Image Processing*
- *Big Data Analytics and MLOps*

Module no 01: Python Fundamentals and Data Analysis with Libraries

Week 1 & 2:

- Introduction to Python programming language and Python Environment:
  - Installation of Python and Jupyter Notebook
  - Basic syntax and data types
  - Control flow statements: if-else, loops
  - Functions and modules
  - File handling and input/output operations

Week 3:

- Basics of NumPy for numerical computing:
  - Introduction to arrays and multidimensional arrays
  - Array creation, indexing, and slicing
  - Mathematical operations and broadcasting
  - Array manipulation and reshaping
  - Array aggregation and statistical functions

Week 4 & 5:

- Pandas for data manipulation and analysis:
  - Introduction to pandas Series and DataFrame
  - Data cleaning and preprocessing
  - Data selection and filtering
  - Data aggregation and group operations
  - Handling missing data and duplicates

Week 6:

- Data visualization using Matplotlib, Seaborn, and Plotly:
  - Introduction to data visualization and its importance
  - Basic plotting with Matplotlib: line plots, scatter plots, bar plots
  - Customizing plots: labels, titles, legends, colors
  - Statistical data visualization with Seaborn: histograms, box plots, heatmaps
  - Interactive visualizations with Plotly: line plots, scatter plots, bar plots

Week 7:

- Symbolic mathematics with SymPy:
  - Introduction to symbolic mathematics and its applications
  - Symbolic expressions and variables
  - Manipulating and simplifying symbolic expressions
  - Solving equations and systems of equations

- Calculus operations: differentiation, integration, limits

Week 8:

- Scientific computing with SciPy:
  - Introduction to SciPy library and its modules
  - Linear algebra operations: matrix operations, eigenvalues, eigenvectors
  - Numerical integration and solving differential equations
  - Statistical functions and hypothesis testing
  - Optimization algorithms and curve fitting

Module no 02: Mathematics for Machine Learning

Week 1 & 2:

- Foundation of linear algebra:
  - Vectors and vector spaces
  - Matrices and matrix operations
  - Systems of linear equations and their solutions
  - Vector spaces, subspaces, and bases
  - Linear transformations and their properties

Week 3 & 4:

- Fundamental of calculus concepts and techniques:
  - Limits, continuity, and differentiability
  - Derivatives and their applications
  - Rules of differentiation: chain rule, product rule, quotient rule
  - Optimization problems and applications of derivatives
  - Integration and techniques of integration
  - Applications of integration: area, volume, and definite integrals

Week 5 & 6:

- Introduction to probability theory and statistics:
  - Probability basics: sample space, events, and probability measures
  - Combinatorics and counting principles
  - Probability distributions: discrete and continuous distributions
  - Central limit theorem and the normal distribution
  - Statistical inference: estimation and hypothesis testing
  - Introduction to regression analysis and correlation

Week 7 & 8:

- Optimization algorithms for machine learning:
  - Introduction to optimization problems in machine learning
  - Unconstrained optimization: gradient descent, Newton's method
  - Constrained optimization: Lagrange multipliers, KKT conditions
  - Convex optimization and convexity in machine learning
  - Stochastic optimization: stochastic gradient descent, mini-batch gradient descent
  - Hyperparameter tuning and model selection

Module no 03: Introduction to Databases and SQL & NoSQL, Web Scraping Techniques and Tools

Week 1:

- Introduction to databases and their importance in data storage:
  - Overview of different types of databases: relational, NoSQL, etc.
  - Understanding the role of databases in data management
  - Exploring popular database management systems (DBMS)

Week 2:

- SQL for relational databases:
  - Introduction to structured query language (SQL)
  - Creating and manipulating database tables
  - Querying data using SELECT, INSERT, UPDATE, DELETE statements
  - Joining tables and performing advanced SQL operations

Week 3:

- NoSQL databases and their advantages:
  - Introduction to NoSQL databases (e.g., MongoDB)
  - Understanding the key-value, document, columnar, and graph database models
  - Working with NoSQL databases using their respective query languages

Week 4:

- Data Scraping:
  - Understanding the basics of data scraping
  - Introduction to web scraping and its applications
  - Working with Beautiful Soup for web scraping:
  - Parsing HTML and XML documents
  - Navigating and extracting data from web pages

Week 5:

- Scrapy for advanced web scraping tasks:
  - Overview of the Scrapy framework
  - Creating a Scrapy project and defining spiders
  - Extracting structured data from websites
  - Handling pagination, forms, and dynamic content

Week 6:

- Selenium for dynamic web scraping:
  - Introduction to Selenium WebDriver
  - Automating browser interactions for web scraping
  - Handling JavaScript-driven websites and AJAX requests
  - Extracting data from dynamically generated web pages

Week 7 & 8:

- Understanding of tools like Webhose.io and Google Custom Search API:
  - Introduction to Webhose.io for accessing web data and news articles
  - Utilizing the Google Custom Search API for targeted web searches
  - Integrating additional scraping tools into data collection pipelines

Module no 04: Fundamentals of Machine Learning

Week 1:

- Introduction to machine learning concepts and algorithms:
  - Understanding the fundamentals of machine learning
  - Differentiating between supervised, unsupervised, and reinforcement learning
  - Overview of the machine learning workflow

Week 2:

- Supervised learning algorithms:
  - Linear regression and logistic regression
  - Decision trees and random forests
  - Support vector machines (SVM)
  - Evaluation metrics for supervised learning models

Week 3:

- Unsupervised learning algorithms:
  - Clustering algorithms (k-means, hierarchical clustering)
  - Dimensionality reduction techniques (PCA, t-SNE)
  - Anomaly detection algorithms
  - Evaluation metrics for unsupervised learning models

Week 4:

- Reinforcement learning:
  - Introduction to reinforcement learning concepts
  - Markov decision processes (MDPs)
  - Q-learning and policy gradient methods
  - Application areas of reinforcement learning

Week 5:

- Building machine learning models:
  - Data preprocessing and feature engineering techniques
  - Splitting datasets into training and testing sets
  - Implementing machine learning algorithms using Python and scikit-learn
  - Hyperparameter tuning and model selection

Week 6:

- Evaluating machine learning models:

- Performance evaluation metrics (accuracy, precision, recall, F1-score)
- Cross-validation techniques
- Overfitting, underfitting, and bias-variance tradeoff
- Model interpretation and explainability

Week 7:

- Advanced topics in machine learning:
  - Ensemble learning methods (bagging, boosting, stacking)
  - Deep learning overview and neural networks
  - Transfer learning and pre-trained models
  - Handling imbalanced datasets and class imbalance

Week 8:

- Project and practical applications:
  - Applying machine learning algorithms to real-world datasets
  - Designing and implementing end-to-end machine learning pipelines
  - Presenting and interpreting machine learning results
  - Kaggle Competitions

Module no 05: Deep Learning and Neural Networks

Week 1 & 2:

- Fundamentals of deep learning:
  - Introduction to artificial neural networks (ANN)
  - Activation functions and backpropagation
  - Optimization algorithms for training deep models
  - Regularization techniques (dropout, weight decay)
- Deep Learning Frameworks and Tools:
  - Introduction to popular deep learning frameworks (TensorFlow, Keras)
  - Hands-on experience with building and training deep learning models
  - Using pre-trained models and transfer learning in deep learning projects
  - Deployment considerations for deep learning models

Week 3:

- Convolutional Neural Networks (CNN):
  - Understanding convolutional layers and filters
  - Pooling and stride operations
  - CNN architectures (LeNet, AlexNet, VGG, ResNet)
  - Transfer learning with pre-trained CNN models

Week 4:

- Recurrent Neural Networks (RNN):
  - Introduction to sequential data and time series analysis
  - Understanding the architecture of RNNs and their variations (LSTM, GRU)
  - Training and predicting with RNN models
  - Applications of RNNs in natural language processing and speech recognition

Week 5:

- Autoencoders:
  - Introduction to unsupervised learning and dimensionality reduction
  - Understanding the architecture and training of autoencoders
  - Applications of autoencoders in data compression and anomaly detection
  - Variational Autoencoders (VAE) for generating new data samples

Week 6:

- Generative Adversarial Networks (GAN):
  - Understanding the concept of generative modeling
  - Exploring the architecture of GANs (generator and discriminator)

- Training GAN models and generating synthetic data
- Applications of GANs in image synthesis and style transfer

Week 7:

- Advanced topics in deep learning:
  - Advanced activation functions (ReLU, PReLU, ELU)
  - Batch normalization and layer normalization
  - Hyperparameter tuning in deep learning models
  - Handling overfitting and improving generalization

Week 8:

- Project and practical applications:
  - Applying deep learning techniques to real-world problems
  - Implementing deep learning architectures for specific tasks (image classification, text generation)
  - Fine-tuning and optimizing deep learning models
  - Kaggle Competitions

Module no 06: Natural Language Processing Fundamentals

Module no 07: Computer Vision and Image Processing

Module no 08: MLOps and Big Data Analytics

Week 1:

- MLOps Introduction:
  - Introduction to MLOps and its significance
  - Deploying machine learning models at scale
  - Model versioning and tracking
  - Monitoring and performance evaluation of deployed models

Week 2:

- Continuous Integration and Continuous Deployment (CI/CD) for Model Deployment:
  - Understanding CI/CD pipelines in machine learning projects
  - Automated testing and model evaluation in CI/CD pipelines
  - Implementing CI/CD pipelines using tools like Jenkins or Github
  - Best practices for efficient model deployment and management

Week 3:

- Big Data and its Challenges:
  - Understanding the characteristics and challenges of big data
  - Introduction to distributed computing frameworks (Hadoop, Spark)
  - Data processing with Hadoop and MapReduce

Week 4:

- Apache Spark for Big Data Analytics:
  - Exploring Apache Spark and its ecosystem
  - Distributed data processing with Spark RDDs and DataFrames
  - Performing big data analytics tasks with Spark
  - Handling large-scale data using Spark's advanced features

For further inquiries, please do not hesitate to contact:

Ehtisham Sadiq

03054661042

https://www.linkedin.com/in/ehtisham-sadiq/