KSI Data Research

Python for Machine Learning

Course Outline

Section 01: Download Python Jupyter lab IDE from Anaconda tool and Basic python code for Machine learning.

1. Overview of Python: Definition and history.
2. Unique features of Python and its advantages
3. Installing Python and setting up the development environment (IDE)
4. Writing and running your first Python program
5. Understanding Python identifiers, keywords, and naming conventions
6. Introduction to comments, documentation, and indentation in Python
7. Variables and Python data types, including dynamic typing
8. Type conversion, casting, and variable swapping
9. Basic input/output functions: Receiving input and printing output
10. Control structures: Arithmetic, logical, and comparison operators; if statements; loops (for and while, including break and continue)

Section 02: Strings and lists, Dictionaries and Tuples, Functions, Object-Oriented Programming (OOP)

11. Introduction Overview of strings and lists
12. Slicing and Indexing Techniques
13. Operators and Functions Understanding
14. Overview of dictionaries and tuples
15. Definition and Types of Functions
16. Parameters and arguments work in functions
17. Homework and projects

Section 03: Key concepts and functions of NumPy important for machine learning and AI

18. Learn different ways to create NumPy arrays
19. Determine and manipulate the shape of arrays
20. Indexing and Slicing: Master accessing elements
21. How to specify and convert data types (dtype, astype)
22. Array Operations and Broadcasting
23. Aggregate Functions (sum(), mean(), std(), min(), max(), median())
24. Random Sampling, Sorting, array manipulation

Section 04: Concepts and features of the Pandas library that are crucial for machine learning and AI

25. Understand the core data structures (DataFrame)
26. Data Importing and Exporting ((CSV, Excel, SQL, JSON)
27. Data Inspection, Indexing, Selecting, Data Cleaning
28. Data Manipulation, Grouping Data
29. Merging, Joining and Filtering data
30. Visualization and DataFrame Operations (apply())

Section 05: Understanding Matplotlib library that are essential for data visualization in machine learning and AI

31. How to create basic line plots using plt.plot()
32. Customizing line styles, colors, and markers
33. Understand the concepts of Figures and Axes
34. Create multiple subplots using plt.subplots()
35. Customize plots with titles, labels, and legends
36. Scatter Plots, Bar and Histogram Plots
37. Box Plots, Pie Charts, Annotating Plots
38. Saving Figures, formats (PNG, PDF, SVG)
39. Visualization of data in heatmaps

Section 06: Understanding Seaborn library that are essential for data visualization in machine learning and AI

40. Statistical plots offered by Seaborn distplot, jointplot(), pairplot()
41. Pandas DataFrames in Seaborn
42. Relationships between variables relplot(), scatterplot(), lineplot()
43. Categorical plots catplot(), boxplot(), violinplot(), stripplot().
44. Matrix plot, Pair Plot and Distribution plot
45. Regression Plots and Multi-Plot Grids
46. How to combine Seaborn plots with Matplotlib
47. Modify plots with functions like hue, size, and style

Section 07: Introduction to Machine learning with few resources

48. Basic Concepts of Machine Learning with diagram
49. Types of Machine Learning (supervised, unsupervised)
50. Fundamental or key algorithms (linear regression, logistic regression so on)
51. Evaluation metrics for model performance (accuracy, recall so on)
52. Overfitting and Underfitting, techniques to balance
53. Feature engineering, including data preprocessing
54. Intro to Neural Networks

Section 08: Topics to understand Linear Regression

55. Understand the basic idea that linear regression models
56. Equation of a Line
57. Simple vs. Multiple Linear Regression
58. Least Squares Method
59. Assumptions of Linear Regression
60. Coefficient Interpretation
61. R-squared and Adjusted R-squared
62. Residual Analysis
63. Feature Selection and Engineering
64. Model Evaluation and Validation

Section 09: Applying linear regression to a real-world project

65. Problem definition, Data Collection, Data processing
66. Exploratory Data Analysis (EAD)
67. Feature Selection, Splitting the dataset
68. Model Implementation, model evaluation
69. Model assumption checking and model reporting

Section 10: Binary classification logistic Regression for Machine learning

70. Concept of Logistic Regression
71. Logistic Function, Decision Boundary
72. Binary vs. Multinomial Logistic Regression
73. Model Evaluation Metrics, Regularization Techniques
74. Problem definition, Data Collection, Data processing
75. Exploratory Data Analysis (EAD)
76. Feature Selection, Splitting the dataset
77. Model Implementation, model evaluation

Section 11: Decision Trees and Random Forests, Gradient Boosting

78. Decision tree non-linear classifier
79. Random Forest Ensemble method extension of Decision trees
80. Decision tree non-linear classifier project
81. Random Forest Ensemble method projects
82. Decision tree and Random Forest project solutions

Section 12: Deep Learning Models (Neural Networks)

83. Neural Network Architecture? Feedforward Neural Networks
84. Convolutional (CNNs), Recurrent (RNNs)
85. Long Short-Term Memory (LSTM) Networks
86. Generative Adversarial Networks (GANs)
87. Neural Network Training and Optimization
88. Real-World Applications of Deep Learning in AI