

COURSE TITLE	PYTHON FOR DATA SCIENCE AND AI APPLICATIONS
COURSE CODE	01CE0413
COURSE CREDITS	3

Objective:

- 1 This course builds a solid foundation in Python programming, emphasizing data analysis, preprocessing, and visualization for data science and machine learning. It integrates advanced techniques in deep learning, natural language processing, and AI applications to enable students to design and implement intelligent, real-world solutions.

Course Outcomes: After completion of this course, student will be able to:

- 1 Apply core Python programming concepts and to develop functional applications
- 2 Apply data cleaning, preprocessing, and visualization techniques to analyze datasets and extract meaningful insights
- 3 Analyze the advanced machine learning techniques, including model selection and hyperparameter tuning, to optimize predictive models
- 4 Evaluate the performance of deep learning models by incorporating image preprocessing methods such as data augmentation and transfer learning.
- 5 Create innovative AI applications using Generative Adversarial Networks (GANs) and reinforcement learning to address real-world challenges

Pre-requisite of course:Fundamentals of computer programming

Teaching and Examination Scheme

Theory Hours	Tutorial Hours	Practical Hours	ESE	IA	CSE	Viva	Term Work
1	0	4	50	30	20	25	25

Contents : Unit	Topics	Contact Hours
1	Module 1: Python Programming and Object-Oriented Development Introduction to Python - Basic Syntax, Data Types, Control Structures, Functions, Modules, Error Handling - Lists, Tuples, Sets,, Dictionaries - String Manipulation, File Operations - Classes, Objects, Inheritance, Polymorphism, Encapsulation - Design Patterns (Singleton, Factory, Observer), NumPy, Pandas, Matplotlib, Seaborn, scikit-learn	5
2	Module 2: Python for Data Science Data Cleaning, Preprocessing (Pandas, Data Imputation, Scaling, Encoding, Normalization, Outlier Detection), Feature Selection, Feature Extraction, Feature Engineering (scikit-learn), Exploratory Data Analysis (EDA), Data Visualization (Matplotlib, Seaborn).	6

Contents : Unit	Topics	Contact Hours
3	Module 3: Python for Machine Learning Machine Learning Basics - Supervised Learning (Linear Regression, Decision Trees, Random Forests), Unsupervised Learning (K-means, PCA), Model Selection, Model Evaluation, Hyperparameter Tuning	6
4	Module 4: Introduction to Deep Learning and NLP Image Preprocessing (Resizing, Normalization, Data Augmentation), Neural Networks, TensorFlow, Keras - Transfer Learning, Pretrained Models (VGG, ResNet), NLP Basics (NLTK, SpaCy), Transformer Models (BERT, GPT) - Computer Vision (OpenCV)	6
5	Module 5: Cutting-Edge Technologies Advanced Visualization in Python (Plotly, Bokeh, Dash), Reinforcement Learning (Q-learning), Basics of Generative Adversarial Networks (GANs) and LLM, Time Series Analysis.	5
Total Hours		28

Suggested List of Experiments:

Contents : Unit	Topics	Contact Hours
1	Practical 1 Write a Python Program to Implement a Simple Text-Based Calculator with Basic Arithmetic Operations.	4
2	Practical 2 Write a Python Program to Implement a Word Frequency Counter Using Dictionaries and Text Processing	4
3	Practical 3 Write a Python Program for Titanic Dataset Preprocessing and Imputation Using Mean/Median and KNN Algorithms	4
4	Practical 4 Write a Python Program for Exploratory Data Analysis (EDA) and Visualization of the Iris Dataset Using Seaborn and Matplotlib	4
5	Practical 5 Write a Python Program to Build and Evaluate a Regression Model for Boston Housing Using Linear Regression and Random Forest Regression	4
6	Practical 6 Write a Python Program for Feature Engineering on the Wine Quality Dataset Using Scaling and One-Hot Encoding	4
7	Practical 7 Write a Python Program to Apply Feature Selection on the Adult Income Dataset Using Recursive Feature Elimination (RFE).	4
8	Practical 8 Write a Python Program for Dimensionality Reduction on the MNIST Dataset Using Principal Component Analysis (PCA)	4

Suggested List of Experiments:

Contents : Unit	Topics	Contact Hours
9	Practical 9 Write a Python Program for Hyperparameter Tuning in Heart Disease Prediction Using GridSearchCV	4
10	Practical 10 Write a Python Program to Preprocess Images and Train a CNN on CIFAR-10 Using TensorFlow and Keras	4
11	Practical 11 Write a Python Program to Implement Transfer Learning for Dogs vs. Cats Classification Using Pre-Trained Models (VGG, ResNet)	4
12	Practical 12 Write a Python Program for Sentiment Analysis on IMDB Reviews Using NLP Techniques with NLTK and SpaCy	4
13	Practical 13 Write a Python Program to Build a Generative Adversarial Network (GAN) for Image Generation.	4
14	Practical 14 Write a Python Program for Time Series Forecasting Using LSTM Networks on Air Quality Data.	4
Total Hours		56

Textbook :

- 1 Python Object-Oriented Programming: Build robust and maintainable object-oriented Python applications and libraries, Steven F. Lott, Dusty Phillips, 4th Edition Kindle Edition, 2021

References:

- 1 Data Science using Python: A Step-by-Step Practical Approach for Beginners, Data Science using Python: A Step-by-Step Practical Approach for Beginners, Dr. Vishal Goyal Dr. Monika Bansal, Dr. Munish Jindal, Dr. Harmandeep Kaur, Paperback, 2021
- 2 Machine Learning using Python, Machine Learning using Python, Manaranjan Pradhan, U Dinesh Kumar, Paperback, 2019
- 3 Getting started with Deep Learning for Natural Language Processing (NLP), Getting started with Deep Learning for Natural Language Processing (NLP), Sunil Patel, Paperback, 2021
- 4 Generative AI for Everyone: Deep learning, NLP, and LLMs for creative and practical applications, Generative AI for Everyone: Deep learning, NLP, and LLMs for creative and practical applications, Karthikeyan Sabesan, Sivagamisundari, Nilip Dutta,, Paperback, 2024

Suggested Theory Distribution:

The suggested theory distribution as per Bloom's taxonomy is as follows. This distribution serves as guidelines for teachers and students to achieve effective teaching-learning process

Distribution of Theory for course delivery
--

Remember / Knowledge	Understand	Apply	Analyze	Evaluate	Higher order Thinking / Creative
0.00	0.00	35.00	35.00	20.00	10.00

Instructional Method:

- 1 The course delivery method will depend upon the requirement of content and need of students. The teacher in addition to conventional teaching method by black board, may also use any of tools such as demonstration, role play, Quiz, brainstorming, MOOCs etc.
- 2 The internal evaluation will be done on the basis of continuous evaluation of students in the laboratory and class-room.
- 3 Practical examination will be conducted at the end of semester for evaluation of performance of students in laboratory.
- 4 Students will use supplementary resources such as online videos, NPTEL videos, e-courses, Virtual Laboratory.

Supplementary Resources:

- 1 YouTube - Python Crash Course: https://youtu.be/K5KVEU3aaeQ?si=Oev_rXWXo2s_6MsA
- 2 Coursera - Python for Data Science <https://www.coursera.org/learn/python-data-analysis>
- 3 Machine Learning: <https://www.youtube.com/watch?v=0D-PCtUJInE>
- 4 Andrew NG Courses: <https://www.andrewng.org/courses/>