JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY ANANTAPUR

B.Tech-CSE - II Sem   **L   T   P   C**

                                                                                                                      3   0   0   3

(20A05101T) PYTHON PROGRAMMING & DATA SCIENCE

(CSE, AI & DS,CSE (AI), CSE(IoT), CSE (Data Science), CSE(AI & ML), IT**)**

Course Objectives

- To learn the fundamentals of Python.
- To discuss the concepts of Functions and Exceptions.
- To familiarize with Python libraries for Data Analysis and Data Visualization.
- To introduce preliminary concepts in Pattern Recognition and Machine learning.
- To provide an overview of Deep Learning and Data Science models.

Unit-I

Introduction to Python: Features of Python, Data types, Operators, Input and output, Control Statements.

Strings: Creating strings and basic operations on strings, string testing methods. Lists, Dictionaries, Tuples.

Learning outcomes:

At the end of this unit, the students will be able to
- List the basic constructs of Python. (L1)
- Apply the conditional execution of the program (L3)
- Design programs for manipulating strings (L6)
- Use the data structure lists, Dictionaries and Tuples (L3)

Unit-II

Functions: Defining a function, Calling a function, returning multiple values from a function, functions are first class objects, formal and actual arguments, positional arguments, recursive functions.

Exceptions: Errors in a Python program, exceptions, exception handling, types of exceptions, the except block, the assert statement, user-defined exceptions.

Learning outcomes:

At the end of this unit, the students will be able to
- Solve the problems by applying the modularity principle. (L3)
- Classify exceptions and explain the ways of handling them. (L4)

Unit-III

Introduction to NumPy, Pandas, Matplotlib.

Exploratory Data Analysis (EDA), Data Science life cycle, Descriptive Statistics, Basic tools (plots, graphs and summary statistics) of EDA, Philosophy of EDA. Data Visualization: Scatter plot, bar chart, histogram, boxplot, heat maps, etc.

Learning outcomes:

At the end of this unit, the students will be able to
- Demonstrate various mathematical operations on arrays using NumPy (L2)
- Analyze and manipulate Data using Pandas (L4)
- Creating static, animated, and interactive visualizations using Matplotlib. (L6)

Unit-IV

Introduction to Pattern Recognition and Machine Learning: Patterns, features, pattern representation, the curse of dimensionality, dimensionality reduction. Classification—linear and non-linear. Bayesian, Perceptron, Nearest neighbor classifier, Logistic regression, Naïve-Bayes, decision trees and random forests; boosting and bagging.Clustering---partitional and hierarchical; k-means clustering. Regression.

Cost functions, training and testing a classifier. Cross-validation, Class-imbalance – ways of handling, Confusion matrix, evaluation metrics.

Learning outcomes:

At the end of this unit, the students will be able to
- Define Patterns and their representation (L1)
- Describe the Classification and Clustering (L2)
- illustrate cost functions and class imbalance (L3)

Unit-V

Introduction to Deep Learning: Multilayer perceptron. Backpropagation. Loss functions. Hyperparameter tuning, Overview of RNN, CNN and LSTM.

Overview of Data Science Models: Applications to text, images, videos, recommender systems, image classification, Social network graphs.

At the end of this unit, the students will be able to
- Describe RNN, CNN and (L2)
- Explain the applications of Data Science (L2)

Textbooks:

1. Allen B. Downey, "Think Python", 2nd edition, SPD/O'Reilly, 2016.
2. Cathy O'Neil, Rachel Schutt, Doing Data Science, Straight Talk from the Frontline. O'Reilly, 2013.
3. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.

References:

1. Michael Nielsen, Neural Networks and Deep Learning, Determination Press, 2015.
2. Francois Chollet, Deep Learning with Python, 1/e, Manning Publications Company, 2017
3. EMC2: Data Science and Big Data Analytics, EMC Education Services, EMC 2 , Wiley Publication, 2015.
4. V. Susheela Devi and M. Narasimha Murty. Pattern Recognition – An Introduction. Universities Press (Indian Edition; there is an expensive Springer version of the same)
5. Goodfellow and YoshuaBengio and Aaron Courville. Deep Learning. MIT Press. Book available online at https://www.deeplearningbook.org/.
6. J. Leskovec, A. Rajaraman, J.D. Ullman. Mining of Massive Datasets. Cambridge University Press. (Indian Edition; Online pdf is available for download)

Course Outcomes:

1. Apply the features of Python language in various real applications. (L3)
2. Identify the appropriate data structure of Python for solving a problem (L2)
3. Demonstrate data analysis, manipulation and visualization of data using Python libraries (L5)
4. Enumerate machine learning algorithms. (L1)
5. Analyze the various applications of Data Science. (L4)
6. Design solutions for real-world problems using Python. (L6)