Kavayitri Bahinabai Chaudhari North Maharashtra University Jalgaon, Ordinance 181

College Name : - Institute of Management Research and Development, Shirpur

Title of the Course : - Certificate Course in AI Using Python

Aims / Objective : - To prepare learners with practical AI skills using Python, covering data preprocessing, machine learning, deep learning, and NLP for AI- driven applications

Duration of the Course : - 1 Year

Fees Structure : - Rs. 3,000/-

Course Structure :- **1)** Paper I – Fundamental of Artificial Intelligence using Python

2) Paper II – Advance AI with Python

3) Paper III- LAB on AI using Python

Eligibility of the Course: - Student passed in the Higher Secondary Examination (HSC) (10+2) or equivalent examination.

SN	Paper	Name of the subject	Theory / Practical	Teaching Hours	Maximum Marks Allotted			Passing			Credit
					Ext	Int	Total	Ext	Int	Total	
01	Paper I	Fundamental of Artificial Intelligence using Python	THEORY	90	60	40	100	24	16	40	6
02	Paper II	Advance AI with Python	THEORY	90	60	40	100	24	16	40	6
03	Paper III	LAB on AI using Python	PRACTICAL	120	60	40	100	24	16	40	8

Minimum Staff : - 03 Three

Mode of Examination : - English (Theory & Practical)

Detailed Syllabus : - Attached

Paper I
Fundamental of Artificial Intelligence using Python (90 Hrs.)

Unit-I Introduction to Use Cases for AI [20 Hr]

- What is AI?
- Why AI?
- Branches of AI
- Building Intelligent Agents Representative
- AI Use Cases:
    - Digital Personal Assistant and Chatbots
    - Shipping and Warehouse Management
    - Human Health,
    - Knowledge Search
    - Recommender Systems
    - The Smart Home
    - Gaming, Movie Making
- Data Cleansing and Transformation.

Unit-II Machine Learning Pipelines [12 Hr]

- What is a Machine Learning Pipeline
- Problem Definition
- Data Ingestion
- Data Preparation
- Data Segregation
- Model Training.

Unit-III Feature Selection and Feature Engineering [14 Hr]

- Feature Selection
- Feature Engineering
- Outlier Management
- One-hot Encoding
- Log Transform,
- Scaling
- Date Manipulation

Unit-IV Classification and Regression Using Supervised Learning [18 Hr]

- Supervised versus Unsupervised Learning
- What is Classification
- Preprocessing Data
- Label Encoding

- Logistic Regression Classifiers
- The Naive Bayes Classifier
- SupportVector Machines
- Confusion Matrices
- What is Regression?
  Building Single-Variable Regressor
  Building Multivariable Regressor.

Unit-V Predictive Analytics [12 Hr]

- Ensemble Learning
- Decision Trees
- Random Forests
- Dealing with Class Imbalance,

Unit-VI Detecting Patterns with Unsupervised Learning [14 Hr]

- Unsupervised Learning
- Clustering Data with K-Means Algorithm
- Estimating Number of Clusters with the Mean Shift Algorithm
- Estimating Quality of Clustering with Silhouette Scores,
- Gaussian Matrix Models
- Propagation Model

References:

1] Python Machine Learning, Sebastian Raschka, Vahid Mirjalili, Packt Publishing, ISBN: 978-1-78995-575-0
2] AI Crash Course, Hadelin de Ponteves, Packt Publishing, ISBN: 978-1-83864-535-9
3] Artificial Intelligence with Python, Prateek Joshi, Packt Publishing, Packt Publishing, ISBN: 9781786464392

Paper II
Advance AI with Python (90 Hrs.)

Unit-I Building Recommender Systems [14 Hr]

- Extracting the nearest Neighbors
- Building K-Nearest Neighbors Classifier
- Commuting Similarity Scores
- Finding Similar Users Using Collaborative Filtering
- Case Study:
    - Building Movie
    - Recommender System.

Unit-II AI on Cloud [14 Hr]

- Why are Companies migrating to Cloud?
- Top Cloud Providers-
    - Amazon Web Services (AWS),
    - Microsoft Azure
    - Google Cloud Platform (GCP).

Unit-III Building Speech Recognizer [16 Hr]

- Working with speech signals
- Extracting speech features
- Visualizing audio signals
- Transforming audio signals into frequency domain
- Generating audio signals
- Synthesizing tones to generate music
- Recognizing spoken words

Unit-IV Building NLP Tools [20 Hr]

- Installations of NLP packages
- Tokenizing text data
- Stemming
- Lemmatization
- Dividing text into chunks
- Bag of words model
- Building category predictor
- Constructing gender Identifier
- Building sentiment Analyzer

Unit-V Chatbots **[08 Hr]**

- Concepts Chatbots today
- A well-architected chatbot Platforms
- Creating a chatbot

Unit-VI AI and Big Data **[18 Hr]**

- Crawling
- Indexing
- Ranking
- Worldwide datacenters
- Distributed lookups
- Custom softwares
- The three V's of Big data
- Big data and machine Learning –
    - Apache Hadoop
    - Apache Spark
    - Apache Impala,
- NoSQL databases –
    - Apache Cassandra
    - MangoDB
    - Redis
    - Neo4j

References:

1] Python Machine Learning, Sebastian Raschka, Vahid Mirjalili, Packt Publishing, ISBN: 978-1-78995-575-0
2] AI Crash Course, Hadelin de Ponteves, Packt Publishing, ISBN: 978-1-83864-535-9
3] Artificial Intelligence with Python, Prateek Joshi, Packt Publishing, Packt Publishing, ISBN: 9781786464392

Paper III
LAB on AI using Python (120 Hrs.)

1. Installation of Python & Packages

- Install Python on Windows/Ubuntu.
- Install essential packages using pip or conda.
- Load datasets using Pandas and NumPy.

2. Data Preparation

- Perform data cleansing, filtering, and aggregation.
- **Handling Missing Values**
  - Remove missing values (dropna()) or impute using mean, median, or mode (fillna()).
  - Use ML models to predict missing values.

3. Feature Scaling, and Inconsistent Values

Feature Scaling

- Normalize data using Min-Max Scaling (MinMaxScaler).
- Standardize data using Z-score Scaling (StandardScaler).

Handling Inconsistent Values

- Detects duplicates and incorrect formats (duplicated(), unique()).
- Standardize categorical values and correct data formats (str.lower(), pd.to_datetime()).

4. Feature Engineering & Selection

- Apply feature scaling techniques (Normalization, Standardization).
- Select important features using univariate selection, correlation heatmaps, wrapper-based, and filter-based methods.

5. Feature Engineering Techniques

Outlier Management

- Detect outliers using statistical methods (e.g., IQR, Z-score).
- Handle outliers by removing them or transforming values (e.g., capping, log transformation).

One-Hot Encoding

- Convert categorical variables into numerical form using pd.get_dummies() or OneHotEncoder from sklearn.
- Example: "Red", "Blue", "Green" → [1,0,0], [0,1,0], [0,0,1].

Log Transformation

- Apply log transformation to reduce skewness in data.
- Formula: $X' = \log(X+1)$ (to avoid log(0)).

- Implement using np.log1p().

6. Implement Logistic regression classifier.

7. Implement Naïve Bayes classifier.

8. Implement classifier using Support Vector Machines.

9. Build a decision tree classifier and evaluate performance of a classifier by printing classification report.

10. Build random forest and extremely random forest classifiers and analyze the output.

11. Build K-nearest classifier

12. Use confusion matrixes to describe performance of a classifier.

13. Implement K-Means algorithm for clustering.

14. Visualizing audio signals.

15. Transform audio signals to the frequency domain.

16. Generate audio signals.

17. Installation of NLTK and tokenizing text data.

18. Converting words to their base forms using stemming, lemmatization.

19. Extracting the frequency of terms using Bag of Words model.