A Report on Six-days Skill Development Training Program on
"Data Dive: Python and Machine Learning Essentials"
Organized by Skill Development Cell
In Association with Department of Computer Science & Technology
From 25.11.2024 to 30.11.2024

Report Submitted by: Mr. P. Firoze Khan, Assistant Professor, Department of CST SDC-APSSDC Coordinator.
Resource person Details: Mr. P. Lakshmiramana, Assistant Professor, Department of CST and Mr. V. Naveen, Assistant Professor, Department of CST.
Total Participants: 70-Students and Faculties.
Venue: Computer Lab (CB: 10)
Mode of Conduct: Offline
Report Received on 03.12.2024.

The APSSDC- Skill Development Cell, Madanapalle Institute of Technology and Science, Andhra Pradesh**,** Madanapalle in association with the Department of Computer Science & Technology, MITS has organized a **Six-days Skill Development program** from 25.11.2024 to 30.11.2024 on **"Data Dive: Python and Machine Learning Essentials"** In this program, about 64 participants participated and made the event grant success**.**

A summary of the skill development program is as follows:

Mr. P. Firoze Khan, Assistant Professor in CST SDC-APSSDC Coordinator, Welcomed the Dr. K. Dinesh, Professor & Head of the Dept. Computer Science & Technology and welcomed the resource person. Mr. P. Lakshmiramana, gave a brief introduction to the six-day skill development program. Dr. K. Dinesh, Professor & Head of the Dept. Computer Science & Technology inaugurated the program with his motivational speech, introduced the resource person and he handed over the session to the resource person. The 70 students from the second year and faculties from the Department of CST participated in this six-day hands-on training program which was run by two resource persons. (Each per 3 days).

Day-1 (25.11.2024)

Morning session: Foundations of Python for Data Analysis

- **Introduction to Python:** Installation and Setup (Anaconda, Jupyter Notebook, or VS Code)

- Python Fundamentals: Data Types, Variables, and Operators
- Control Structures: Loops and Conditional Statements
- Hands-on Practice: Writing simple Python programs

Afternoon session: Python for Data Handling

- Functions: Writing and Using Functions
- working with Python Libraries: numpy, pandas
- Data Handling: Importing, Cleaning, and Exploring Data
- Hands-on Practice: Working with a dataset (e.g., CSV file analysis)

Day-2 (26.11.2024)

Morning session: Data Visualization

- **Overview of Data Visualization**

- Libraries: matplotlib, seaborn
- Visualizing Data: Histograms, Scatter Plots, Boxplots
- Hands-on Practice: Visualizing patterns in a dataset

Afternoon session: Python for Data Handling

- **Basics of Descriptive Statistics:** Mean, Median, Mode, Variance, Standard Deviation

- Correlation Analysis and Covariance
- Introduction to Inferential Statistics
- Hands-on Practice: Statistical insights from datasets

Day-3 (27.11.2024)

Morning session: Introduction to Machine Learning

- What is Machine Learning? Applications and Types (Supervised, Unsupervised, Reinforcement)
- Steps in a Machine Learning Workflow
- Introduction to scikit-learn library

- **Dataset Preparation:** Splitting Data (Train/Test), Feature Scaling

Afternoon session: Supervised Learning Basics

- Linear Regression: Concept and Implementation
- Evaluation Metrics: MAE, MSE, R^2 Score
- Hands-on Practice: Predicting continuous variables

During Training Sessions:
On the Fourth Day of the program, Mr. P. Firoze Khan, Assistant Professor in CST SDC-APSSDC Coordinator, Welcomed the resource person. Mr. V. Naveen, AP/CST to address the session for remaining days.

Day-4 (28.11.2024)

Morning session: Classification Techniques:

- **Introduction to Logistic Regression:**

- Binary vs. Multiclass Classification
- Evaluation Metrics: Confusion Matrix, Precision, Recall, F1-Score
- Hands-on Practice: Predicting binary outcomes

Afternoon session: Decision Trees and Random Forests

- Understanding Decision Trees
- Introduction to Random Forests and Feature Importance
- Hands-on Practice: Classification with Decision Trees/Random Forests

Day-5 (29.11.2024)

Morning session: Unsupervised Learning and Feature Engineering

- **Introduction to Clustering: Applications and Use Cases**

- K-Means Clustering: Concept and Implementation
- Evaluation of Clustering Models (Silhouette Score)
- Hands-on Practice: Grouping data points based on patterns

Afternoon session: Feature Engineering and Dimensionality Reduction

- **Importance of Feature Engineering**

- Handling Missing Values, Encoding Categorical Data
- Introduction to PCA (Principal Component Analysis)
- Hands-on Practice: Feature transformation and dimensionality reduction

Day-6 (30.11.2024)

Morning session: Model Optimization and Deployment

- **Hyper parameter Tuning: Grid Search and Random Search**

- Cross-Validation: K-Fold Cross-Validation
- Model Saving and Loading: joblib, pickle
- Hands-on Practice: Optimizing a machine learning model

Afternoon session: Deploying Machine Learning Models

- **Introduction to Model Deployment Concepts**

- Basics of Flask/Streamlit for ML Deployment
- Deploying a simple model as a web app
- Wrap-up: Review, Feedback, and Q&A.

Outcomes: Students can be able to

- Python Programming Proficiency
- Data Analysis and Visualization Skills
- Machine Learning Foundations
- Supervised Learning Models
- Unsupervised Learning Techniques
- Model Optimization and Deployment
- Practical Problem-Solving Skills