COURSE HANDOUT (COURSE CURRICULUM)

Course Title : **Data Science**

1. Course Description:

Data science is the study of data to extract meaningful insights for business. It is a multidisciplinary approach that combines principles and practices from the fields of mathematics, statistics, artificial intelligence, and computer engineering to analyze large amounts of data.

2. Skills You Will Gain

- Machine learning
- Communication
- Data wrangling
- Statistics
- Data visualization
- Deep learning
- Cloud computing
- Programming
- Python

3. Who Can Enable Students To:

- To be a data scientist, you'll need to be able to gather and analyze data, then present your findings. This includes technical skills such as programming, manipulating databases, advanced mathematics, and data visualization, along with soft skills like collaboration and public speaking.

4. Course Outcomes: -

- Students will become proficient in the statistical analysis of data and the use of computation tools for data analysis. Students will apply statistical and computational tools to applied problems, and clearly communicate the results in both written reports and oral presentations.

SYLLABUS:

Module 1: Introduction to Data Science
• Selecting rows/observations
• Rounding Number
• Selecting columns/fields
• Merging data
• Data aggregation
• Data munging techniques

Module 2: Introduction to Python
• What is Python?
• Why Python?
• Installing Python
• Python IDEs
• Jupyter Notebook Overview

Module 3: Python Basics
• Python Basic Data types
• Lists
• Slicing
• IF statements
• Loops
• Dictionaries
• Tuples
• Functions
• Array
• Selection by position & Labels

Module 4: Python Packages
• Pandas
• Numpy
• Sci-kit Learn
• Mat-plot library

Module 5: Importing Data
• Reading CSV files
• Saving in Python data
• Loading Python data objects
• Writing data to CSV file

Module 6: Manipulating Data
• Selecting rows/observations
• Rounding Number
• Selecting columns/fields

• Merging data
• Data aggregation
• Data munging techniques

Module 7: Statistics Basics
• Central Tendency
o Mean
o Median
o Mode
o Skewness
o Normal Distribution
• Probability Basics
o What does it mean by probability?
o Types of Probability
o ODDS Ratio?
• Standard Deviation
o Data deviation & distribution
o Variance
• Bias variance Tradeoff
o Underfitting
o Overfitting
• Distance metrics
o Euclidean Distance
o Manhattan Distance
• Outlier analysis
o What is an Outlier?
o Inter Quartile Range
o Box & whisker plot
o Upper Whisker
o Lower Whisker
o Scatter plot
o Cook's Distance

Module 8: Error Metrics
• Classification
o Confusion Matrix
o Precision
o Recall
o Specificity
o F1 Score
• Regression
o MSE
o RMSE
o MAPE

Scope of Course:

- Students will develop relevant programming abilities. Students will demonstrate proficiency with statistical analysis of data. Students will develop the ability to build and assess data-based models. Students will execute statistical analyses with professional statistical software.
- Data science is the domain of study that deals with vast volumes of data using modern tools and techniques to find unseen patterns, derive meaningful information, and make business decisions. Data science uses complex machine learning algorithms to build predictive models.

- **Hands-on Project – 4 Hrs**
- **End Assessment – 1 Hr (50 Questions)**
- **Earn Credit-Based Certificate**
- **Internship Certificate**
- **Live Class through Industry Experts**