FOReSiGHT

SYLLABUS

Course Title: Data Science

Course Description:

This online course on Data Science provides a comprehensive introduction to the fundamental principles, tools, and methodologies that underpin the rapidly evolving field of data analysis and interpretation. Designed for learners with diverse backgrounds, this course aims to equip participants with essential skills to harness the power of data and transform it into actionable insights. From understanding data types and formats to employing statistical techniques and data visualization tools, students will gain a solid foundation in data manipulation and analysis.

 By the course's end, students will have acquired the necessary expertise to navigate the ever-expanding data landscape, making them adept data practitioners prepared to tackle challenges across various industries and disciplines.

Course Objectives:

By the end of this course, students will be able to:
1. Understand and differentiate between various data types, structures, and formats.
2. Apply statistical methods and techniques to analyze and interpret data.
3. Utilize popular data science tools and programming languages, such as Python and R, for data manipulation and analysis.
4. Design and interpret data visualizations to communicate findings effectively.
5. Understand the principles of machine learning and its applications in data analysis.
6. Evaluate the ethical considerations and challenges in data science.
7. Apply data science methodologies to real-world problems across different industries.

Key Topics:

Module 1: Introduction to Data Science
- What is Data Science?
- Role of a Data Scientist
- Overview of the Data Science Process

Module 2: Data Types and Structures
- Understanding Data Types: Quantitative vs. Qualitative
- Data Structures: Tabular, Hierarchical, Network
- Data Formats: CSV, JSON, XML

Module 3: Data Exploration and Preprocessing
- Data Cleaning and Transformation
- Handling Missing Values
- Feature Engineering and Selection

FOReSiGHT

Module 4: Statistical Analysis
- Descriptive Statistics
- Inferential Statistics
- Hypothesis Testing and p-values

Module 5: Data Visualization
- Principles of Effective Visualization
- Tools for Data Visualization: Matplotlib, Seaborn, ggplot
- Interactive Visualizations

Module 6: Introduction to Machine Learning
- Supervised vs. Unsupervised Learning
- Common Algorithms: Regression, Classification, Clustering
- Model Evaluation and Validation

Module 7: Advanced Topics in Data Science
- Text Analysis and Natural Language Processing
- Time Series Analysis
- Deep Learning and Neural Networks

Module 8: Ethics and Challenges in Data Science
- Data Privacy and Security
- Bias and Fairness in Machine Learning
- Future Trends in Data Science

Course Materials:

Course Slides / Reading List

Proposed Assessment:

1. Quizzes (40% of final grade)
   - Objective: Test students' understanding of each module's content.
   - Format: Multiple-choice, true/false, and short-answer questions.
   - Frequency: At the end of each module.
2. Hands-on Assignments (30% of final grade)
   - Objective: Provide students with practical experience in data manipulation, analysis, and visualization.
   - Format: Students will be given datasets and tasks related to the topics covered in each module.
   - Assessment: Assignments will be graded based on accuracy, creativity, and thoroughness.
3. Final Project (30% of final grade)
   - Objective: Assess students' ability to apply data science methodologies to a real-world problem.
   - Format: Students will choose a dataset and conduct a comprehensive data analysis, culminating in a report and presentation of their findings.
   - Assessment: Projects will be graded on the depth of analysis, clarity of insights, and presentation skills.