GRAStech

Data Science Program using Machine Learning

Duration – 4 Months(Weekdays) & 5 Months(Weekends)

Contents in detail:

1. Introduction to Data Science

- **Overview of Data Science**
  - Definition and Scope
  - Applications and Industry Use Cases
- **Data Science Pipeline**
  - Data Collection
  - Cleaning, Exploration
  - Modeling
  - Deployment

2. Excel Basic to Advanced

1. Introduction to Excel
2. Excel Interface
3. Basic Functions and Formulas
4. Basic Data Entry and Formatting
5. Advanced Formulas
6. Data Management
7. Conditional Formatting
8. Data Analysis and Visualization
9. Charts and Graphs
10. Pivot Tables
11. Data Validation
12. Complex Formulas

3. Mathematics for Data Science

- **Linear Algebra**
  - Vectors, Matrices, and Operations
  - Eigenvalues and Eigenvectors
- **Probability and Statistics**

- Probability Distributions
- Statistical Inference
- Hypothesis Testing and Confidence Intervals

4. Programming for Data Science

- **Python for Data Science**
  - Python Basics to core and Data Structures
  - Libraries: NumPy, pandas, Matplotlib, Seaborn
- **SQL and Database Management**
  - SQL Queries and Operations
  - DDL
  - DML
  - Relational Databases

5. Data Exploration and Visualization

- **Data Visualization**
  - Creating Visualizations with Matplotlib, Seaborn, ggplot2
  - Advanced Visualization Techniques (Interactive Dashboards)

6. Machine Learning

- **Supervised Learning**
  - Regression (Linear Regression, Polynomial Regression)
  - Classification (Logistic Regression, Decision Trees, Random Forests)
  - Model Evaluation and Hyperparameter Tuning
- **Unsupervised Learning**
  - Clustering (K-means, Hierarchical Clustering)
  - Dimensionality Reduction (PCA, t-SNE)
- **Advanced Machine Learning**
  - Ensemble Methods (Boosting, Bagging)
  - Support Vector Machines (SVM)
  - Neural Networks Basics

7. NLP

1. Introduction to NLP

- **What is NLP?**

    - Definition and Scope
    - Applications and Use Cases
- **History and Evolution**
- **NLP Pipeline**
    - Text Processing and Understanding

2. Text Preprocessing

3. Text Generation and Summarization

- **Text Generation**
- **Text Summarization**

4. Sentiment Analysis and Opinion Mining

1. Use of cloud computing

1. Data Science and Business Intelligence

1. Capstone Project

- **Capstone Project**
    - End-to-End Data Science Project
    - Real-World Problem Solving