Data Science Workshop

By RUSkilled (A unit of Kyrion Technologies Pvt Ltd)

In this intensive two-session program, we dive into the core principles of Data Science and its practical applications in the real world. Designed for beginners and enthusiasts alike, our workshop aims to provide a solid foundation in Data Science fundamentals, including essential statistics, programming, and domain knowledge, while exploring beginner-friendly machine learning concepts.

Duration: 1 Days (7 hours of learning)

Data Science and Real-World Applications

Session 1: Foundations of Data Science (3 Hours)

- Introduction to Data Science
    - What is Data Science? Definition and importance.
    - Key components: Statistics, Programming, Domain Knowledge.
    - Applications across industries (e.g., Finance, Healthcare, Retail).
- Data Science Workflow
    - Problem definition, data collection, cleaning, analysis, modeling, evaluation, deployment.
    - Real-world workflow examples.
- Exploring Data
    - Types of data: Structured vs. Unstructured.
    - Basic statistics for data analysis: Mean, Median, Mode, Variance, Standard Deviation.
    - Visualizing data: Using tools like Matplotlib or Power BI.
- Hands-On Exercise
    - Analyze a sample dataset (e.g., Titanic dataset or sales data).
    - Perform basic exploration, calculate summary statistics, and create visualizations.

Session 2: Beginner-Friendly Machine Learning Concepts (3 Hours)

- Supervised Learning
    - Regression: Predicting continuous outcomes
    - Classification: Predicting categorical outcomes
    - Hands-On Exercise: Build a simple linear regression model using scikit-learn.
- Basic ML Workflow
    - Splitting data into training and test sets.
    - Training models and evaluating performance metrics.
        - Regression
        - Classification

- Unsupervised Learning Overview
    - Clustering: Grouping unlabeled data (e.g., customer segmentation).
    - Hands-On Exercise: Use K-Means clustering on a small dataset (e.g., Iris dataset).
- Mini Project
    - Build a classification model with the Industry dataset.
    - Preprocess data, train a decision tree or logistic regression model, evaluate performance.

By the end of this workshop, participants will be able to:

1. Understand the definition and importance of Data Science across various industries such as Finance, Healthcare, and Retail.
2. Identify and execute the key steps in the Data Science workflow, including problem definition, data collection, cleaning, analysis, modeling, evaluation, and deployment.
3. Gain hands-on experience with data exploration, employing tools like Matplotlib and Power BI to visualize data and interpret basic statistical measures.
4. Develop foundational skills in supervised and unsupervised learning, building and evaluating models like linear regression, decision trees, and logistic regression.
5. Complete a Mini Project where they will preprocess data, train a model, and evaluate its performance on an industry-specific dataset.

What's Next - How to Continue Learning in Data Science

- Continuing your journey in Data Science after this workshop involves several strategic steps to deepen your understanding and enhance your skills:
- Practice Regularly: Apply the concepts learned during the workshop by working on different datasets and trying out new problem statements.
- Further Education: Consider enrolling in advanced courses or obtaining certifications in Data Science and machine learning.
- Join Online Communities: Engage with other data science learners and professionals in online forums and communities. This can provide ongoing support and resources.
- Stay Updated: Data Science is a rapidly evolving field. Stay updated with the latest tools, techniques, and industry trends by following relevant blogs, podcasts, and publications.
- Build a Portfolio: Work on personal projects or volunteer for internships to build a portfolio of real-world data science projects. This will be crucial for your professional development and potential job opportunities.