STRATEGIC INFORMATION TECHNOLOGY & CONSULTING SERVICES

"Summer Project Training and Internship on Machine Learning and Data Science with Python"

Strategic Information Technology & Consulting Services (SITCS) announces its exclusive '*Summer Project Training & Internship'* training on **Machine learning and Data Science for students**. The program covers all aspects related to training and internship in the areas of Machine learning and Data Science covering various algorithms with real world hands-on use cases.

Why the Summer Project Training and Internship

The today's 21st century is witnessing a tremendous need for human expertise in the areas of data analysis and data modeling for supporting effective decision making in various aspects of human living environments. The growing demand for Data Analysts and Data Scientists establishes the emergency for the development and upgradation of technical and analytical skills of the organizational personnel to address various business segments where decision making is becoming critical.

The one month training provides a rigor to the participants by imparting the necessary technical and analytical knowledge who incline to jump start their career in Machine learning and Data Science. The course focuses on providing in depth understanding of the need and importance of analyzing data while leveraging the power of data analytics with real world use cases and scenarios and also opportunity to interact with industry experts.

Target Participants

Graduates from Engineering (B.Tech/BE), Sciences (B.Sc(Maths/Statistics/Electronics/Computer Science,B.Pharmacy), Commerce(B.Com(computers)) can all avail the summer project training and internship opportunity.

Training Objectives

1. To understand the vital nature of data for organisations.
2. To learn the conceptual framework of machine learning.
3. To explore and analyze data using supervised and unsupervised learning techniques.
4. To develop and deploy knowledge learning models using Python.
5. To understand the important concepts and applications of Artificial Intelligence.

Learning Outcomes

SITCS
Consulting | R&D | Training

STRATEGIC INFORMATION TECHNOLOGY & CONSULTING SERVICES

Upon completion of the project training and internship, a participant will be able to:

1. Understand the importance of data, information and knowledge.
2. Understand various use cases and business scenarios of the real world.
3. Get hands on experience in exploring, analyzing and modeling real world business scenarios using Python.
4. Understand machine learning framework and develop and deploy machine learning models developed using Python.
5. Develop analytical skills.
6. Industry interaction
7. Get evaluated and certified as Industry ready professionals

A Brief look at the Training Process

Hello to Week-1

Day - 1: Welcome to Data Science World

- Introduction **Talk by Delegates.**
- Why **Machine learning and Data Science?**- *History and Evolution of Data Science*
- Use Cases on Applications of Data Science.
  - ✓ *Real Time Analytics by Chicago*
  - ✓ *Fraud Detection*
  - ✓ *Click Stream Analysis*
  - ✓ *Sentiment Analysis*
  - ✓ *Customer Loyalty Analytics*
- Business Video Case of Motor and Pump Monitoring through Analytics
- Life of a Data Scientist
- **Why Statistics?**- *Introduction to Statistics*
- **Data and its meaning.**
- **Introduction to Statistics concepts and measure.**
  - About Data, Statistical terminology,
  - Central Tendency, Variance and Spread
  - Simple experiments on Basics of Statistics
    Working on student test scores, average heights

Day - 2: **Basic Probability Concepts of Probability Distribution**- *Introduction to Probability*

- Probability basics, Basic Terminology
- Bayesian Theorem
  - *Use case on Drug Testing*
  - *Use case on Factory Production*
- Confusion Matrix, Basic terms and formulae

- Probability Distributions, Terminology, Types of Probability Distributions
  - ✓ *The Normal Distribution with use case on survey of daily travel time*.
- Discrete Probability Distribution- *Introduction and Types*
  - ✓ *Geometric distribution with use case on employee job skills training.*
  - ✓ *Bernoulli distribution with use case on assessing distribution of individuals.*
  - ✓ *Binomial distribution with use case on survival analysis of Canadian males.*
  - ✓ *Poisson distribution with use case on Plutonium-239 radioactive decay.*

- Continuous Probability Distribution-*Introduction and Types*
  - ✓ *Exponential Distribution*
  - ✓ *t-Distribution*

- **Inferential Statistics:**
  - Central limit Theorem, sampling Distribution.
  - Confidence interval – *Introduction*
    - ✓ *Types of confidence intervals*
    - ✓ *How to construct a confidence interval with video use case on 'Apple Orchard'*
  - Hypothesis testing- *Introduction*
    - ✓ *Proving a hypothesis-The Process*
    - ✓ *Null & Alternate Hypothesis*

- **Statistical Hypothesis testing**
  - t-test- *Introduction*, *types, paired sample t-test by hand with use case on student test scores.*
  - z-test- *Introduction, Two Proportion z-test for comparing performance of two drugs*
  - f-test-Introduction, types, use case for testing two samples variance
  - Chi-square-Introduction*, types, use case on analyzing assembly election, use case on distribution analysis of zodiac signs for visual artists.*
  - ANOVA-*Introduction, types, use case for analyzing what production using one-way* and two-way ANOVA

Day 3: Introduction to Python

Python Environment **Setup** and General Introduction of **Jupyter Notebook**

- Introduction on **Python**.
- **Data Analysis** Python
  (1) Numpy
  (2) Pandas
  (3) Matplotlib

DAY -4: Supervised learning - Regression

General Introduction of **supervised learning** and Difference between **Classification and Regression.**

Regression based learning:

- **Linear Regression**
- **Multiple Regression**

HANDS-ON:

- Use case based learning for Regression
- Evaluation metrics used for Regression based learning.

Day 5: Supervised learning- Classification

- **Classification based learning:**
  - **Logistic Regression**
- **Different kind of Evaluation metrics involved in measure of classification based learning.**
  - **Confusion Matrix**

HANDS-ON:

Working on the above-mentioned concepts by taking dataset using python.

- **Interaction session** with student

Hello to Week-2

Day 6: More into Classifications:

- **K-NN**
- **SVM**
- **Decision Tree**
- **Ensemble:**
  - **Random Forest**
  - **AdaBoost**

Hands-on: Working on the above-mentioned concepts by taking dataset using python

Day 7: Unsupervised Learning:

- **K-means Learning**
- **Dimensionality Reduction**
- **Principal Component Analysis**

Hands-on: Working on above mention concepts by taking dataset

Day 8: Natural Processing Language:

Basic Pre-processing Concepts:

- Tokenization
- Stemming
- Lemmatization
- Regular Expression
- N-grams
- Parts of Speech (POS)
- **Bag-of-words**: Count Vectorizer, Tf-idf Vectorizer

Application: Sentiment Analysis

Day 9: Artificial Intelligence:

- **Basic Neural Networks**

  Hands-on: Use case on Neural Networks using python.

Day 10: Data Science Experience:

- **Introduction of Data Science Experience**
- **Prepare data**
- **Machine learning Flow**
- **Visualization**
- **Predictive Analysis**

Project Work - Week 3

Day 11: Regression Based Project

Day 12: Classification Based Project

Day 13: Clustering and Association Based Project

Day 14: Project review session with Students

Day 15: Discussion Panel about Data Science.

Industrial Interaction – Week 4

Day 16 – Day 20: Interaction with Industrial experts at Select industries.

Students registered for Training programme and completing 3 weeks of successful training shall be formulated into groups and deployed to select industries where they have to work on projects assigned by the industry.