

A PROPOSAL BY
LEGIONSOFTECHNOLOGIES

WELCOME ALL

A PROPOSAL BY
LEGIONSOFTECHNOLOGIES

PROJECT
INTERNSHIP PROJECT

LegionSoft
Technologies



PYTHON DATA SCIENCE & INTERNSHIP



07/2023

COURSE SYLLABUS

INTRODUCTION TO PYTHON PROGRAMMING

- Overview of Python as a programming language
- Installing Python and an integrated development environment (IDE)
- Basic syntax and data types in Python
- Variables, data structures, and operators
- Control flow statements: if-else, loops, and functions

WORKING WITH DATA IN PYTHON

- Introduction to the NumPy library for numerical computing
- Creating and manipulating arrays
- Basic mathematical operations on arrays
- Indexing and slicing arrays
- Common array functions and methods

DATA MANIPULATION WITH PANDAS

- Overview of the Pandas library for data manipulation and analysis
- Introduction to Series and DataFrame data structures
- Loading data into Pandas DataFrame
- Data cleaning techniques: handling missing data, duplicates, and outliers
- Filtering and sorting data in a DataFrame

TIMEFRAME

- Daily Classes

COURSE SYLLABUS

EXPLORATORY DATA ANALYSIS (EDA)

- Data visualization with Matplotlib library
- Creating various types of plots: line, bar, scatter, histogram, box plot, etc.

DATA CLEANING TECHNIQUES

- Dealing with missing values: identifying and handling missing data
- Imputation techniques: mean, median, mode, and advanced methods
- Handling duplicates: identifying and removing duplicate records
- Outlier detection and treatment strategies
- Data scaling and normalization techniques

INTRODUCTION TO MACHINE LEARNING

- Overview of machine learning and its applications
- Supervised learning vs. unsupervised learning vs. reinforcement learning
- Understanding the machine learning workflow
- Training set, validation set, and test set
- Evaluating machine learning models

TIMEFRAME

- Daily Classes

COURSE SYLLABUS

LINEAR REGRESSION

- Introduction to linear regression
- Simple linear regression vs. multiple linear regression
- Training and evaluating linear regression models in Python

LOGISTIC REGRESSION

- Introduction to logistic regression
- Understanding binary and multiclass classification problems
- Training and evaluating logistic regression models in Python

DATA MANIPULATION WITH PANDAS

- Overview of the Pandas library for data manipulation and analysis
- Introduction to Series and DataFrame data structures
- Loading data into Pandas DataFrame
- Data cleaning techniques: handling missing data, duplicates, and outliers
- Filtering and sorting data in a DataFrame

TIMEFRAME

- Daily Classes

COURSE SYLLABUS

MODEL EVALUATION AND VALIDATION

- Cross-validation techniques: k-fold cross-validation, stratified k-fold cross-validation
- Performance metrics for classification: accuracy, precision, recall, F1 score
- Performance metrics for regression: mean squared error (MSE), R-squared

UNSUPERVISED LEARNING TECHNIQUES

- Introduction to clustering: k-means clustering, hierarchical clustering
- Preprocessing and feature scaling for unsupervised learning
- Evaluating clustering results using metrics like silhouette coefficient
- Visualizing and interpreting results of dimensionality reduction techniques

TIMEFRAME

- Daily Classes

INTERNSHIP SYLLABUS

INTRODUCTION TO NATURAL LANGUAGE PROCESSING (NLP)

- Overview of NLP and its applications
- Tokenization: breaking text into words or smaller units
- Text preprocessing techniques: removing stop words, stemming, and lemmatization
- Introduction to NLTK (Natural Language Toolkit) library in Python
- Performing basic text analysis and visualization

TEXT CLASSIFICATION WITH MACHINE LEARNING

- Introduction to text classification
- Bag-of-Words (BoW) model and TF-IDF (Term Frequency-Inverse Document Frequency) representation
- Training and evaluating text classification models using algorithms like Naive Bayes, SVM, and logistic regression
- Handling imbalanced datasets and performance evaluation metrics for text classification

TIMEFRAME

- Daily Classes

INTERNSHIP SYLLABUS

INTRODUCTION TO COMPUTER VISION

- Overview of computer vision and its applications
- Image representation: pixels, color spaces, and image formats
- Basic image processing operations: resizing, cropping, and filtering
- Introduction to OpenCV (Open Source Computer Vision Library) in Python
- Performing basic image manipulation and analysis

IMAGE CLASSIFICATION WITH MACHINE LEARNING

- Introduction to image classification
- Overview of deep learning and convolutional neural networks (CNNs)
- Transfer learning and pre-trained models for image classification
- Training and evaluating image classification models using CNNs in Python

TIMEFRAME

- Daily Classes

INTERNSHIP PROJECT

SAMPLE MACHINE LEARNING INTERNSHIP PROJECT

- Overview of a machine learning internship project
- Identifying a real-world problem and formulating it as a machine learning task
- Data collection and preprocessing for the project
- Developing and evaluating machine learning models to solve the problem
- Documentation and presentation of the project's findings and results

INTERNSHIP PROJECT

- Python Projects for Automation
- NLP Projects
- Computer Vision Projects
- Exploratory Data Analysis Project Work

TIMEFRAME

- Daily Classes

INTERNSHIP ROLES

ROLES AND RESPONSIBILITIES DURING INTERNSHIP

- Collect, clean, and preprocess data from various sources to ensure data quality.
- Assist in exploratory data analysis to uncover patterns, trends, and insights.
- Develop and implement data models, algorithms, and statistical techniques.
- Collaborate with the Data Science team to design and execute data experiments.
- Perform predictive and prescriptive analysis to solve business problems.
- Support the development of machine learning and deep learning models.
- Evaluate and validate models using appropriate performance metrics.
- Assist in visualizing and communicating data-driven insights to stakeholders.
- Stay up-to-date with the latest advancements and best practices in data science.
- Adhere to data privacy and security guidelines throughout the project lifecycle.
- Participate in team meetings, discussions, and knowledge-sharing sessions.
- Complete assigned tasks and deliver high-quality work within specified deadlines.
- Seek feedback and actively engage in learning and professional development.