

Semester 1: Base SAS

48 Hours

Base SAS
24 Hours

Data Structures | Data Types | Type Casting | Data Import/Export | Loops & Functions | Procedures

Understanding Data
8 Hours

Definition of Data | Meaning of Variable | Understanding Data Types | Measures of Central Tendency in Data | Understanding Skewness in Data | Measures of Dispersion | Understanding Data Distribution

SAS - Store Data Analytics
8 Hours

Extracting Retail Data into SAS | Cleaning the Data | Transforming the Data | Loading the Data for Reporting and Analysis | Reporting and Presenting the Data

Project 1 Submission
8 Hours

Project 1 - Store Data Analytics
Key Learning Outcomes - ETL, Analysis and Reporting using SAS

Semester 2: R Programming

78 Hours

R Basics
32 Hours

Data Structures | Data Types | Data Import/Export | Conditional Statements | Loops & Functions

Business Case: Managing Credit Risk
1 Hours

Meaning of Credit Risk | Impact of Credit Default | Sources of Data for Managing Risk | Understanding Loss given Default | Understanding Default

Linear Regression
4 Hours

Covariance and Correlation in Data | Multivariate Analysis | Assumptions of Linearity | Hypothesis Testing | Limitations of Regression

Case Study on Loss Given Default using Linear Regression using R
4 Hours

Extract the Data in R | Univariate Analysis of Data | Apply Data Transformations | Bivariate Analysis | Multicollinearity in Data | Identify Heteroscedasticity | Modelling of Data | Model Significance Test | Build Model on Training Data Set

Logistic Regression
4 Hours

Reason for using Logistic Regression | The Logistic Transform | Logistic Regression Modelling | Model Optimization | Understanding the ROC Curve

Case Study on Default Modelling using Logistic Regression using R
8 Hours

Extract the Data in R | Univariate Analysis of Data | Apply Data Transformations | Bivariate Analysis | Identify Multicollinearity in Data | Identify Heteroscedasticity | Variable Significance Identification | Model Significance Test | Validate the Model Performance | Drawing the ROC Curve | Estimating the Classification Model Hit Ratio | Isolating the Classifier for Optimum Results

Support Vector Machines
4 Hours

Introduction to SVM | Classification as a Hyper Plane Location Problem | Motivation for Linear Support Vectors | SVM as a Quadratic Optimization Problem | Non Linear SVMs | Introduction to Kernel Functions

Case Study on Default Modelling using Support Vector Machines using R
6 Hours

Build the Model | Training the Model | Testing and Validation | Tuning the Model

Business Case: Intrusion in IT Network
1 Hours

Meaning of Intrusion in IT Network | Cost of Intrusion | Meaning of Intrusion Detection System

Decision Tree & Ensemble Learning
8 Hours

Theory of Entropy & Information Gain | Stopping Rules | Overfitting Problem | Cross Validations for Overfitting Problem | Pruning as a Solution for Overfitting | Ensemble Learning Notion | Concept of Bootstrap Aggregation | Concept of Random Forest

Case Study on Network Intrusion Detection using Decision Trees & Ensemble Learning using R
6 Hours

Validate the Model Performance Perform Cross Validations | Select the Best Split | Prune the Tree | Predict and Validate Performance of Model

Project 2 and 3 Submission
2 Hours

Project 2 Credit Risk Analytics

Key Learning Outcomes - Linear and Logistics Regression, Support Vector Machines using R

Project 3 - Network Intrusion Detection

Key Learning Outcomes - Decision Trees & Ensemble using R

Semester - 3: Python

28 Hours

Data Structures in Python Used for Data Analysis
8 Hours

Data Structures | Data Types | Data Import/Export | Numpy/Pandas | Loops & Functions

Case Study on Default Modelling using Logistic Regression using Python
8 Hours

Univariate Analysis of Data | Apply Data Transformations | Bivariate Analysis | Build Model on Training Data Set | Predict using Testing Data Set | Validate the Model Performance

Case Study on Default Modelling using Support Vector Machines using Python
4 Hours

Build the Model | Training the Model | Testing and Validation | Tuning the Model

Case Study on Network Intrusion Detection using Decision Trees & Ensemble Learning using Python
6 Hours

Extract the Data in Python | Convert Data into Array | Standardize the Independent | Variables | Predict using Testing Data Set | Validate the Model Performance Perform Cross Validations | Select the Best Split | Prune the Tree | Predict and Validate | Performance of Model | Improve Model Performance using Bagging and Random Forest

Project 2 and 3 Submission
2 Hours

Project 2 - Credit Risk Analytics

Key Learning Outcomes - Logistics Regression, Support Vector & Machines using Python

Project 3 - Network Intrusion Detection

Key Learning Outcomes - Decision Trees & Ensemble using Python

Semester 4 : Hive, Spark , Tableau

46 Hours

HIVE Basics
5 Hours

Set Up | Data Loading in HDFS | Creating HIVE Tables | Creating ORC File | Querying HIVE data

Business Case: Pricing Analytics
1 Hours

Understanding Pricing as a Business Function | Importance of Pricing Analytics | Sources of Pricing Data

Case Study on Pricing Analytics using HIVE
8 Hours

Load and Extract Data into HDFS | Transform Data in Hive | Load Data in Hive Tables | Write Queries Prepare Analysis Output | Show the Reports in Zeppelin

SPARK Basics
5 Hours

Set Up | Creating a RDD | Querying a RDD

Business Case: Understanding
Customer Life Cycle
1 Hour

Understanding Customer Life Cycle - Acquisition/Consumption/Saturation/Churn | Impact of Customer Churn | Churn Identification

Case Study on Telecom Customer
Churn using Spark
8 Hours

Case Study on Telecom Customer Churn using Spark | Extract Data in Spark | Perform a Logistic | Stochastic Gradient Descent Model | Predict using Testing Data Set | Validate the Model Performance

Tableau Basics
8 Hours

Introduction to Visualization | Working with Tableau | Data Organization | Advanced Visualization | Mapping | Enterprise Dashboards | Data Presentation

Interview and Resume Preparation
4 Hours

Analytics Experts Guide You on How to Prepare for Technical Interview Round with Tips, Tricks and a Quick Refresher on Concepts Learnt

Mock Interviews - HR and Domain
4 Hours

1:1 or Panel Mock Interviews with Industry Veterans to Clear the HR and Technical Round of Interviews to Give You Confidence to Face Real World Scenarios

Project 4, 5 and 6 Submission
2 Hours

Project 4 - Pricing Analytics

Key Learning Outcomes - Big Data Analytics and Reporting using HIVE

Project 5 - Telecom Churn Analytics

Key Learning Outcomes - In Memory Big Data and Logistics Regression using Spark

Project 6 - Reporting

Key Learning Outcomes - Visualization for Structured and Unstructured Data using Tableau

HANDS-ON PROJECTS

Students get hands-on with industry projects and build a portfolio of demonstrable work



Store Data Analytics
& Reporting



Credit Risk Analytics



Network Intrusion
Detection Analytics



Pricing Analytics in
Bullion/Commodity
Market



Telecom Churn Analytics



Visualizing Vanilla,
Analytical, Un/Structured
Data