Data Science Masters Program

About CertAdda

CertAdda is a leading elearning platform providing live instructor-led interactive online training. We cater to professionals and students across the globe in categories like Big Data & Hadoop, Business Analytics, NoSQL Databases, Java & Mobile Technologies, System Engineering, Project Management and Programming.

We have an easy and affordable learning solution that is accessible to millions of learners. With our students spread across countries like the US, India, UK, Canada, Singapore, Australia, Middle East, Brazil and many others, we have built a community of over 1 million learners across the globe.

About the Program

Data Science Masters Program makes you proficient in tools and systems used by Data Science Professionals. It includes training on Statistics, Data Science, Python, Apache Spark & Scala, Tensorflow and Tableau. The curriculum has been determined by extensive research on 5000+ job descriptions across the globe.

Data Science Masters Program

Index

Python Statistics for Data Science Course

Module 1: Understanding the Data

Learning Objectives:

At the end of this module, you should be able to:

- Understand various data types
- Learn Various variable types
- List the uses of variable types
- Explain Population and Sample
- Discuss sampling techniques
- Understand Data representation

Topics:

- Introduction to Data Types
- Numerical parameters to represent data
- Mean
- Mode
- Median
- Sensitivity
- Information Gain
- Entropy
- Statistical parameters to represent data

Hands-on/Demo:

- Estimating mean, median and mode using Python
- Calculating Information Gain and Entropy

Module 2: Probability and its uses

Learning Objectives:

At the end of this module, you should be able to:

Understand rules of probability
Learn about dependent and independent events
Implement conditional, marginal, and joint probability using Bayes Theorem
Discuss probability distribution
Explain Central Limit Theorem

Topics:

Uses of probability
Need of probability
Bayesian Inference
Density Concepts
Normal Distribution Curve

Hands-on/Demo:

Calculating probability using python
Conditional, Joint and Marginal Probability using Python
Plotting a Normal distribution curve

Module 3: Statistical Inference

Learning Objectives:

At the end of this Module, you should be able to:

Understand concept of point estimation using confidence margin
Draw meaningful inferences using margin of error
Explore hypothesis testing and its different levels

Topics:

Point Estimation
Confidence Margin
Hypothesis Testing
Levels of Hypothesis Testing

Hands-on/Demo

Calculating and generalizing point estimates using python

Estimation of Confidence Intervals and Margin of Error

Module 4: Testing the Data

Learning Objectives:

At the end of this module, you should be able to:

Understand Parametric and Non-parametric Testing
Learn various types of parametric testing
Discuss experimental designing
Explain a/b testing**Topics**
Parametric Test
Parametric Test Types
Non- Parametric Test
Experimental Designing
A/B testing

Hands-on/Demo

Perform p test and t tests in Python
A/B testing in Python

Module 5: Data Clustering

Learning Objectives:

At the end of this module, you should be able to:

Understand concept of association and dependence
Explain causation and correlation
Learn the concept of covariance
Discuss Simpson's paradox
Illustrate Clustering Techniques

Topics:

Association and Dependence
Causation and Correlation
Covariance
Simpson's Paradox

Clustering Techniques

Hands-on/Demo

Correlation and Covariance in Python
Hierarchical clustering in Python
K means clustering in Python

Module 6: Regression Modelling

Learning Objectives:

At the end of this module, you should be able to:

Understand the concept of Linear Regression
Explain Logistic Regression
Implement WOE
Differentiate between heteroscedasticity and homoscedasticity
Learn concept of residual analysis

Topics

Logistic and Regression Techniques
Problem of Collinearity
WOE and IV
Residual Analysis
Heteroscedasticity
Homoscedasticity

Hands-on/Demo

Perform Linear and Logistic Regression in Python
Analyze the residuals using Python

Course Curriculum

Module 1: Understanding the Data

Learning Objectives

At the end of this Module, you should be able to understand various data types, learn various variable types, list the uses of variable types, explain population and sample, discuss sampling techniques, and understand data representation

Topics

Introduction to Data Types
Numerical parameters to represent data
Mean
Mode
Median
Sensitivity
Information Gain
Entropy
Statistical parameters to represent data

Hands-on/Demo

Estimating mean, median and mode using R
Calculating Information Gain and Entropy

Module 2: Probability and its Uses

Learning Objectives

At the end of this Module, you should be able to understand rules of probability, learn about dependent and independent events, implement conditional, marginal and joint probability using Bayes Theorem, discuss probability distribution and explain Central Limit Theorem.

Topics

      Uses of probability
      Need of probability
      Bayesian Inference
      Density Concepts
      Normal Distribution Curve

Hands-on/Demo

      Calculating probability using R
      Conditional, Joint and Marginal Probability using R
      Plotting a Normal distribution curve

Module 3: Statistical Inference

Learning Objectives

At the end of this module, you should be able to understand the concept of point estimation using confidence margin, demonstrate the use of level of confidence and confidence margin, draw meaningful inferences using margin of error and explore hypothesis testing and its different levels

Topics

      Point Estimation

      Confidence Margin

      Hypothesis Testing

      Levels of Hypothesis Testing

Hands-on/Demo

      Calculating and generalizing point estimates using R

Estimation of Confidence Intervals and Margin of Error

Module 4: Testing the Data

Learning Objectives

At the end of this module, you should be able to understand parametric and non-parametric testing, learn various types of parametric testing and explain A/B testing

Topics

Parametric Test

Parametric Test Types

Non- Parametric Test

A/B testing

Hands-on/Demo

Perform P test and T tests in R

Module 5: Data Clustering

Learning Objectives

At the end of this module, you should be able to understand the concept of association and dependence, explain causation and correlation, learn the concept of covariance, discuss Simpson's paradox, and illustrate clustering techniques.

Topics

      Association and Dependence

      Causation and Correlation

      Covariance

      Simpson's Paradox

      Clustering Techniques

Hands-on/Demo

      Correlation and Covariance in R

      Hierarchical clustering in R

      K means clustering in R

Module 6: Regression Modelling

Learning Objectives

At the end of this module, you should be able to: Understand the concept of Linear Regression, Explain Logistic Regression, Implement WOE, Differentiate between heteroscedasticity and homoscedasticity and Learn concept of residual analysis

Topics

- Logistic and Regression Techniques
- Problem of Collinearity
- WOE and IV
- Residual Analysis
- Heteroscedasticity
- Homoscedasticity

Hands-on/Demo

- Perform Linear and Logistic Regression in R
- Analyze the residuals using R
- Calculation of WOE values using R

Module 1: Introduction to Data Science

Learning Objectives

Get an introduction to Data Science in this module and see how Data Science helps to analyze large and unstructured data with different tools.

Topics

What is Data Science?

What does Data Science involve?

Era of Data Science

Business Intelligence vs Data Science

Life cycle of Data Science

Tools of Data Science

Introduction to Big Data and Hadoop

Introduction to R

Introduction to Spark

Introduction to Machine Learning

Module 2: Statistical Inference

Learning Objectives

In this module, you will learn about different statistical techniques and terminologies used in data analysis.

Topics

>   What is Statistical Inference?
>
>   Terminologies of Statistics
>
>   Measures of Centers
>
>   Measures of Spread
>
>   Probability
>
>   Normal Distribution
>
>   Binary Distribution

Module 3: Data Extraction, Wrangling and Exploration

Learning Objectives

Discuss the different sources available to extract data, arrange the data in structured form, analyze the data, and represent the data in a graphical format.

Topics

>   Data Analysis Pipeline
>   What is Data Extraction
>   Types of Data
>   Raw and Processed Data
>   Data Wrangling
>   Exploratory Data Analysis
>   Visualization of Data

Hands-on/Demo

>   Loading different types of dataset in R
>   Arranging the data

Plotting the graphs

Module 4: Introduction to Machine Learning

Learning Objectives

Get an introduction to Machine Learning as part of this module. You will discuss the various categories of Machine Learning and implement Supervised Learning Algorithms.

Topics

What is Machine Learning?
Machine Learning Use-Cases
Machine Learning Process Flow
Machine Learning Categories
Supervised Learning algorithm: Linear Regression and Logistic
Regression

Hands-on/Demo

Implementing Linear Regression model in R
Implementing Logistic Regression model in R

Module 5: Classification Techniques

Learning Objectives

In this module, you should learn the Supervised Learning Techniques and the implementation of various techniques, such as Decision Trees, Random Forest Classifier, etc.

Topics

What are classification and its use cases?

What is Decision Tree?

Algorithm for Decision Tree Induction

Creating a Perfect Decision Tree

Confusion Matrix

What is Random Forest?

What is Navies Bayes?

Support Vector Machine: Classification

Hands-on/Demo

Implementing Decision Tree model in R

Implementing Linear Random Forest in R

Implementing Navies Bayes model in R

Implementing Support Vector Machine in R

Module 6: Unsupervised Learning

Learning Objectives

Learn about Unsupervised Learning and the various types of clustering that can be used to analyze the data.

Topics

What is Clustering & its use cases

What is K-means Clustering?

What is C-means Clustering?

What is Canopy Clustering

What is Hierarchical Clustering?

Hands-on/Demo

Implementing K-means Clustering in R

Implementing C-means Clustering in R

Implementing Hierarchical Clustering in R

Module 7: Recommender Engines

Learning Objectives

In this module, you should learn about association rules and different types of Recommender Engines.

Topics

What is Association Rules & its Use Cases?

What is Recommendation Engine & its Workings?

Types of Recommendations

User-Based Recommendation

Item-Based Recommendation

Difference: User-Based and Item-Based Recommendation

Recommendation Use Cases

Hands-on/Demo

Implementing Association Rules in R

Building a Recommendation Engine in R

Module 8: Text Mining

Learning Objectives

Discuss Unsupervised Machine Learning Techniques and the implementation of different

algorithms, for example, TF-IDF and Cosine Similarity in this module.

Topics

> The concepts of text-mining
> Use cases
> Text Mining Algorithms
> Quantifying text
> TF-IDF
> Beyond TF-IDF

Hands-on/Demo

> Implementing Bag of Words approach in R
> Implementing Sentiment Analysis on Twitter Data using R

Module 9: Time Series

Learning Objectives

In this module, you should learn about Time Series data, different component of Time Series data, Time Series modeling - Exponential Smoothing models and ARIMA model for Time Series Forecasting.

Topics

> What is Time Series data?
> Time Series variables
> Different components of Time Series data
> Visualize the data to identify Time Series Components
> Implement ARIMA model for forecasting
> Exponential smoothing models

Identifying different time series scenario based on which different Exponential
Smoothing model can be applied

Implement respective ETS model for forecasting

Hands-on/Demo

Visualizing and formatting Time Series data

Plotting decomposed Time Series data plot

Applying ARIMA and ETS model for Time Series Forecasting

Forecasting for given Time period

Module 10: Deep Learning

Learning Objectives

Get introduced to the concepts of Reinforcement learning and Deep learning in this module.
These concepts are explained with the help of Use cases. You will get to discuss Artificial Neural
Network, the building blocks for Artificial Neural Networks, and few Artificial Neural Network
terminologies.

Topics

Reinforced Learning

Reinforcement learning Process Flow

Reinforced Learning Use cases

Deep Learning

Biological Neural Networks

Understand Artificial Neural Networks

Building an Artificial Neural Network

How ANN works

Important Terminologies of ANN's

Python Certification Training for Data Science

Module 1: Introduction to Python

Learning Objectives

You will get a brief idea of what Python is and touch on the basics

Topics

Overview of Python
The Companies using Python
Different Applications where Python is used
Discuss Python Scripts on UNIX/Windows
Values, Types, Variables
Operands and Expressions
Conditional Statements
Loops
Command Line Arguments
Writing to the screen

Hands-on/Demo

Creating "Hello World" code
Variables
Demonstrating Conditional Statements
Demonstrating Loops

Module 2: Sequences and File Operations

Learning Objectives

Learn different types of sequence structures, related operations, and their usage. Also learn diverse ways of opening, reading, and writing to files.

Topics

Python files I/O Functions
Numbers
Strings and related operations
Tuples and related operations
Lists and related operations
Dictionaries and related operations
Sets and related operations

Hands-on/Demo

Tuple - properties, related operations, compared with a list
List - properties, related operations
Dictionary - properties, related operations
Set - properties, related operations

Module 3: Deep Dive – Functions, OOPs, Modules, Errors and Exceptions

Learning Objectives

In this Module, you will learn how to create generic python scripts, how to address errors/exceptions in code and finally how to extract/filter content using regex.

Topics

Functions
Function Parameters
Global Variables
Variable Scope and Returning Values
Lambda Functions
Object-Oriented Concepts
Standard Libraries
The Import Statements
Module Search Path
Package Installation Ways

Errors and Exception Handling
Handling Multiple Exceptions

Hands-on/Demo

Functions - Syntax, Arguments, Keyword Arguments, Return Values
Lambda - Features, Syntax, Options, Compared with the Functions
Sorting - Sequences, Dictionaries, Limitations of Sorting
Errors and Exceptions - Types of Issues, Remediation
Packages and Module - Modules, Import Options, sys Path

Module 4: Introduction to NumPy, Pandas and Matplotlib

Learning Objectives

This Module helps you get familiar with basics of statistics, different types of measures and probability distributions, and the supporting libraries in Python that assist in these operations. Also, you will learn in detail about data visualization.

Topics

NumPy - arrays
Operations on arrays
Indexing slicing and iterating
Reading and writing arrays on files
Pandas - data structures & index operations
Reading and Writing data from Excel/CSV formats into Pandas
matplotlib library
Grids, axes, plots
Markers, colors, fonts and styling
Types of plots - bar graphs, pie charts, histograms
Contour plots

Hands-on/Demo

NumPy library- Creating NumPy array, operations performed on NumPy array
Pandas library- Creating series and dataframes, Importing and exporting data
Matplotlib - Using Scatterplot, histogram, bar graph, pie chart to show
information, Styling of Plot

Module 5: Data Manipulation

Learning Objectives

Through this Module, you will understand in detail about Data Manipulation

Topics

Basic Functionalities of a data object
Merging of Data objects
Concatenation of data objects
Types of Joins on data objects
Exploring a Dataset
Analyzing a dataset

Hands-on/Demo

Pandas Function- Ndim(), axes(), values(), head(), tail(), sum(), std(), iteritems(),
iterrows(), itertuples()
GroupBy operations
Aggregation
Concatenation
Merging
Joining

Module 6: Introduction to Machine Learning with Python

Learning Objectives

In this module, you will learn the concept of Machine Learning and its types.

Topics

Python Revision (NumPy, Pandas, scikit learn, matplotlib)

What is Machine Learning?
Machine Learning Use-Cases

Hands-on/Demo

Machine Learning Process Flow
Machine Learning Categories
Linear regression
Gradient descent
Linear Regression – Boston Dataset

Module 7: Supervised Learning - I

Learning Objectives

In this module, you will learn Supervised Learning Techniques and their implementation, for example, Decision Trees, Random Forest Classifier etc.

Topics

What are Classification and its use cases?
What is Decision Tree?
Algorithm for Decision Tree Induction
Creating a Perfect Decision Tree
Confusion Matrix
What is Random Forest?

Hands-on/Demo

Implementation of Logistic regression
Decision tree
Random forest

Module 8: Dimensionality Reduction

Learning Objectives

In this module, you will learn about the impact of dimensions within data. You will be taught to perform factor analysis using PCA and compress dimensions. Also, you will be developing LDA model.

Topics

Introduction to Dimensionality
Why Dimensionality Reduction
PCA
Factor Analysis
Scaling dimensional model
LDA

Hands-on/Demo

PCA
Scaling

Module 9: Supervised Learning - II

Learning Objectives

In this module, you will learn Supervised Learning Techniques and their implementation, for example, Decision Trees, Random Forest Classifier etc.

Topics

What is Naïve Bayes?
How Naïve Bayes works?
Implementing Naïve Bayes Classifier
What is Support Vector Machine?
Illustrate how Support Vector Machine works?
Hyperparameter Optimization
Grid Search vs Random Search

Implementation of Support Vector Machine for Classification

Hands-on/Demo

Implementation of Naïve Bayes, SVM

Module 10: Unsupervised Learning

Learning Objectives

In this module, you will learn about Unsupervised Learning and the various types of clustering that can be used to analyze the data.

Topics

What is Clustering & its Use Cases?
What is K-means Clustering?
How does K-means algorithm work?
How to do optimal clustering
What is C-means Clustering?
What is Hierarchical Clustering?
How Hierarchical Clustering works?

Hands-on/Demo

Implementing K-means Clustering
Implementing Hierarchical Clustering

Module 11: Association Rules Mining and Recommendation Systems

Learning Objectives

In this module, you will learn Association rules and their extension towards recommendation engines with Apriori algorithm.

Topics

> What are Association Rules?
> Association Rule Parameters
> Calculating Association Rule Parameters
> Recommendation Engines
> How does Recommendation Engines work?
> Collaborative Filtering
> Content-Based Filtering

Hands-on/Demo

> Apriori Algorithm
> Market Basket Analysis

Module 12: Reinforcement Learning

Learning Objectives

In this module, you will learn about developing a smart learning algorithm such that the learning becomes more and more accurate as time passes by. You will be able to define an optimal solution for an agent based on agent-environment interaction.

Topics

> What is Reinforcement Learning
> Why Reinforcement Learning
> Elements of Reinforcement Learning
> Exploration vs Exploitation dilemma
> Epsilon Greedy Algorithm
> Markov Decision Process (MDP)
> Q values and V values
> Q – Learning
> α values

Hands-on/Demo

> Calculating Reward

Discounted Reward
Calculating Optimal quantities
Implementing Q Learning
Setting up an Optimal Action

Module 13: Time Series Analysis

Learning Objectives

In this module, you will learn about Time Series Analysis to forecast dependent variables based on time. You will be taught different models for time series modeling such that you analyze a real time-dependent data for forecasting.

Topics

What is Time Series Analysis?
Importance of TSA
Components of TSA
White Noise
AR model
MA model
ARMA model
ARIMA model
Stationarity
ACF & PACF

Hands-on/Demo

Checking Stationarity
Converting a non-stationary data to stationary
Implementing Dickey-Fuller Test
Plot ACF and PACF
Generating the ARIMA plot
TSA Forecasting

Module 14: Model Selection and Boosting

Learning Objectives

In this module, you will learn about selecting one model over another. Also, you will learn about Boosting and its importance in Machine Learning. You will learn on how to convert weaker algorithms into stronger ones.

Topics

What is Model Selection?
The need for Model Selection
Cross-Validation
What is Boosting?
How Boosting Algorithms work?
Types of Boosting Algorithms
Adaptive Boosting

Hands-on/Demo

Cross-Validation
AdaBoost

Module 1: Introduction to Big Data Hadoop and Spark

Learning Objectives

Understand Big Data and its components such as HDFS. You will learn about the Hadoop Cluster Architecture and you will also get an introduction to Spark and you will get to know about the difference between batch processing and real-time processing.

Topics

    What is Big Data?

    Big Data Customer Scenarios

    Limitations and Solutions of Existing Data Analytics Architecture with Uber Use Case

    How Hadoop Solves the Big Data Problem?

    What is Hadoop?

    Hadoop's Key Characteristics

    Hadoop Ecosystem and HDFS

    Hadoop Core Components

    Rack Awareness and Block Replication YARN and its Advantage

    Hadoop Cluster and its Architecture

    Hadoop: Different Cluster Modes

    Big Data Analytics with Batch & Real-time Processing

    Why Spark is needed?

    What is Spark?

    How Spark differs from other frameworks?

    Spark at Yahoo!

Module 2: Introduction to Scala and Apache Spark

Learning Objectives

Learn the basics of Scala that are required for programming Spark applications. You will also learn about the basic constructs of Scala such as variable types, control structures, collections such as Array, ArrayBuffer, Map, Lists, and many more.

Topics

> What is Scala?
> Scala in other Frameworks
> Basic Scala Operations
> Control Structures in Scala
> Collections in Scala- Array
> Why Scala for Spark?
> Introduction to Scala REPL
> Variable Types in Scala
> Foreach loop, Functions and Procedures
> ArrayBuffer, Map, Tuples, Lists, and more

Hands-On

> Scala REPL Detailed Demo

Module 3: Functional Programming and OOPs Concepts in Scala

Learning Objectives

In this module, you will learn about object-oriented programming and functional programming techniques in Scala.

Topics

Functional Programming

Anonymous Functions

Getters and Setters

Properties with only Getters

Singletons

Overriding Methods

Higher Order Functions

Class in Scala

Custom Getters and Setters

Auxiliary Constructor and Primary Constructor

Extending a Class

Traits as Interfaces

and Layered Traits

Hands On

OOPs Concepts

Functional Programming

Module 4: Deep Dive into Apache Spark Framework

Learning Objectives

Understand Apache Spark and learn how to develop Spark applications. At the end, you will learn how to perform data ingestion using Sqoop.

Topics

Spark's Place in Hadoop Ecosystem

Spark Components & its Architecture

Spark Deployment Modes

Introduction to Spark Shell

Writing your first Spark Job Using SBT

Submitting Spark Job

Spark Web UI

Data Ingestion using Sqoop

Hands On

> Building and Running Spark Application
> Spark Application Web UI
> Configuring Spark Properties
> Data ingestion using Sqoop

Module 5: Playing with Spark RDDs

Learning Objectives

Get an insight of Spark - RDDs and other RDD related manipulations for implementing business logics (Transformations, Actions and Functions performed on RDD).

Topics

> Challenges in Existing Computing Methods
> Probable Solution & How RDD Solves the Problem
> What is RDD, Its Functions, Transformations & Actions?
> Data Loading and Saving Through RDDs
> Key-Value Pair RDDs
> Other Pair RDDs o RDD Lineage
> RDD Lineage
> RDD Persistence
> WordCount Program Using RDD Concepts
> RDD Partitioning & How It Helps Achieve Parallelization
> Passing Functions to Spark

Hands On/Demo

> Loading data in RDDs
> RDD Transformations
> RDD Partitions
> Saving data through RDDs
> RDD Actions and Functions

WordCount through RDDs

Module 6: DataFrames and Spark SQL

Learning Objectives

In this module, you will learn about SparkSQL which is used to process structured data with SQL queries, data-frames and datasets in Spark SQL along with different kind of SQL operations performed on the data-frames. You will also learn about the Spark and Hive integration.

Topics

> Need for Spark SQL
> What is Spark SQL?
> Spark SQL Architecture
> SQL Context in Spark SQL
> User Defined Functions
> Data Frames & Datasets
> Interoperating with RDDs
> JSON and Parquet File Formats
> Loading Data through Different Sources
> Spark – Hive Integration

Hands On/Demo

> Spark SQL – Creating Data Frames
> Loading and Transforming Data through Different Sources
> Stock Market Analysis
> Spark-Hive Integration

Module 7: Machine Learning using Spark MLlib

Learning Objectives

Learn why machine learning is needed, different Machine Learning techniques/algorithms, and SparK MLlib.

Topics

> Why Machine Learning?
> What is Machine Learning?
> Where Machine Learning is Used?
> Face Detection: USE CASE
> Different Types of Machine Learning Techniques
> Introduction to MLlib
> Features of MLlib and MLlib Tools
> Various ML algorithms supported by MLlib

Module 8: Deep Dive into Spark MLlib

Learning Objectives

Implement various algorithms supported by MLlib such as Linear Regression, Decision Tree, Random Forest and many more.

Topics

> Supervised Learning - Linear Regression, Logistic Regression, DecisionmTree, Random Forest
> Unsupervised Learning - K-Means Clustering & How It Workswith MLlib
> Analysis on US Election Data using MLlib (K-Means)

Hands-On

> Machine Learning MLlib
> Linear Regression
> Decision Tree
> K- Means Clustering
> Logistic Regression

Random Forest

Module 9: Understanding Apache Kafka & Apache Flume

Learning Objectives

Understand Kafka and its Architecture. Also, learn about Kafka Cluster, how to configure different types of Kafka Cluster. Get introduced to Apache Flume, its architecture and how it is integrated with Apache Kafka for event processing. At the end, learn how to ingest streaming data using flume.

Topics

- Need for Kafka
- Core Concepts of Kafka
- Where is Kafka Used?
- What is Kafka?
- Kafka Architecture
- Understanding the Components of Kafka Cluster
- Configuring Kafka Cluster
- Need of Apache Flume
- What is Apache Flume?
- Flume Sources
- Flume Channels
- Integrating Apache Flume and Apache Kafka
- Basic Flume Architecture
- Flume Sinks
- Flume Configuration

Hands-On

- Configuring Single Node Single Broker Cluster
- Producing and consuming messages
- Setting up Flume Agent
- Configuring Single Node Multi Broker Cluster
- Flume Commands
- Streaming Twitter Data into HDFS

Module 10: Apache Spark Streaming- Processing Multiple Batches

Learning Objectives

Work on Spark streaming which is used to build scalable fault-tolerant streaming applications. Also, learn about DStreams and various Transformations performed on the streaming data. You will get to know about commonly used streaming operators such as Sliding, Window Operators, and Stateful Operators.

Topics

Drawbacks in Existing Computing Methods
Why Streaming is Necessary?
What is Spark Streaming?
Spark Streaming Features
Spark Streaming Workflow
How Uber Uses Streaming Data
Streaming Context & DStreams
Transformations on DStreams
Describe Windowed Operators and Why it is Useful
Important Windowed Operators
Slice, Window and ReduceByWindow Operators
Stateful Operators

Module 11: Apache Spark Streaming- Data Sources

Learning Objectives

In this module, you will learn about the different streaming data sources such as Kafka and flume. At the end of the module, you will be able to create a spark streaming application.

Topics

Apache Spark Streaming: Data Sources

Streaming Data Source Overview

Apache Flume and Apache Kafka Data Sources

Example: Using a Kafka Direct Data Source

Perform Twitter Sentimental Analysis Using Spark Streaming

Hands-On

Different Streaming Data Sources

Module 12: In Class Project

Learning Objectives

Work on an end-to-end Financial domain project covering all the major concepts of Spark taught during the course.

Module 13: Spark GraphX(Self-Paced)

Learning Objectives

In this module, you will be learning the key concepts of Spark GraphX programming and operations along with different GraphX algorithms and their implementations.

Module 1: Introduction to Deep Learning

Learning Objectives

At the end of this module, you will be able to understand the concepts of Deep Learning and learn how it differs from machine learning. This module will also brief you out on implementing the concept of single-layer perceptron.

Topics

What is Deep Learning?
Curse of Dimensionality
Machine Learning vs. Deep Learning
Use cases of Deep Learning
Human Brain vs. Neural Network
What is Perceptron?
Learning Rate
Epoch
Batch Size
Activation Function
Single Layer Perceptron

Module 2: Getting Started with TensorFlow 2.0

Learning Objectives

At the end of this module, you should be able to get yourself introduced with TensorFlow 2.x. You will install and validate TensorFlow 2.x by building a Simple Neural Network to predict handwritten digits and using Multi-Layer Perceptron to improvise the accuracy of the model.

Topics

Introduction to TensorFlow 2.x
Installing TensorFlow 2.x
Defining Sequence model layers
Activation Function
Layer Types
Model Compilation
Model Optimizer
Model Loss Function
Model Training
Digit Classification using Simple Neural Network in TensorFlow 2.x
Improving the model
Adding Hidden Layer
Adding Dropout
Using Adam Optimizer

Module 3: Convolution Neural Network

Learning Objectives

At the end of this module, you will be able to understand how and why CNN came into existence after MLP and learn about Convolutional Neural Network (CNN) by exploring the theory behind how CNN is used to predict 'X' or 'O'. You will also use CNN VGG-16 using TensorFlow 2 and predict whether the given image is of a 'cat' or a 'dog' and save and load a model's weight.

Topics

Image Classification Example
What is Convolution
Convolutional Layer Network
Convolutional Layer
Filtering
ReLU Layer
Pooling

Data Flattening

Fully Connected Layer

Predicting a cat or a dog

Saving and Loading a Model

Face Detection using OpenCV

Module 4: Regional CNN

Learning Objectives

At the end of this module, you will be able to understand the concept and working of RCNN and figure out the reason why it was developed in the first place. The module will cover various important topics like Transfer Learning, RCNN, Fast RCNN, RoI Pooling, Faster RCNN, and Mask RCNN.

Topics

Regional-CNN

Selective Search Algorithm

Bounding Box Regression

SVM in RCNN

Pre-trained Model

Model Accuracy

Model Inference Time

Model Size Comparison

Transfer Learning

Object Detection – Evaluation

mAP

IoU

RCNN – Speed Bottleneck

Fast R-CNN

RoI Pooling

Fast R-CNN – Speed Bottleneck

Faster R-CNN

Feature Pyramid Network (FPN)

Regional Proposal Network (RPN)

Mask R-CNN

Module 5: Boltzmann Machine & Autoencoder

Learning Objectives

At the end of this module, you should be able to understand what a Boltzmann Machine is and how it is implemented. You will also learn about what an Autoencoder is, what are its various types, and understand how it works.

Topics

What is Boltzmann Machine (BM)?
Identify the issues with BM
Why did RBM come into picture?
Step by step implementation of RBM
Distribution of Boltzmann Machine
Understanding Autoencoders
Architecture of Autoencoders
Brief on types of Autoencoders
Applications of Autoencoders

Module 6: Generative Adversarial Network(GAN)

Learning Objectives

At the end of this module, you should be able to understand what generative adversarial model is and how it works by implementing step by step Generative Adversarial Network.

Topics

What is Boltzmann Machine (BM)?
Identify the issues with BM
Why did RBM come into picture?
Step by step implementation of RBM
Distribution of Boltzmann Machine
Understanding Autoencoders
Architecture of Autoencoders
Brief on types of Autoencoders
Applications of Autoencoders

Module 7: Emotion and Gender Detection

Learning Objectives

At the end of this module, you will be able to classify each emotion shown in the facial expression into different categories by developing a CNN model for recognizing the facial expression of the images and predict the facial expression of the uploaded image. During the project implementation, you will also be using OpenCV and Haar Cascade File to check the emotion in real-time.

Topics

> What is Boltzmann Machine (BM)?
> Identify the issues with BM
> Why did RBM come into picture?
> Step by step implementation of RBM
> Distribution of Boltzmann Machine
> Understanding Autoencoders
> Architecture of Autoencoders
> Brief on types of Autoencoders
> Applications of Autoencoders

Module 8: Introduction RNN and GRU

Learning Objectives

After completing this module, you should be able to distinguish between Feed Forward Network and Recurrent neural network (RNN) and understand how RNN works. You will also understand and learn about GRU and finally implement Sentiment Analysis using RNN and GRU.

Topics

> What is Boltzmann Machine (BM)?
> Identify the issues with BM
> Why did RBM come into picture?
> Step by step implementation of RBM
> Distribution of Boltzmann Machine
> Understanding Autoencoders
> Architecture of Autoencoders
> Brief on types of Autoencoders
> Applications of Autoencoders

Module 9: LSTM

Learning Objectives

After completing this module, you should be able to understand the architecture of LSTM and the importance of gates in LSTM. You will also be able to differentiate between the types of sequence based models and finally increase the efficiency of the model using BPTT.

Topics

What is Boltzmann Machine (BM)?
Identify the issues with BM
Why did RBM come into picture?
Step by step implementation of RBM
Distribution of Boltzmann Machine
Understanding Autoencoders
Architecture of Autoencoders
Brief on types of Autoencoders
Applications of Autoencoders

Module 10: Auto Image Captioning Using CNN LSTM

Learning Objectives

After completing this module, you should be able to implement Auto Image captioning using pre-trained model Inception V3 and LSTM for text processing.

Topics

Auto Image Captioning
COCO dataset
Pre-trained model
Inception V3 model
Architecture of Inception V3
Modify last layer of pre-trained model
Freeze model
CNN for image processing
LSTM or text processing

Module 1: Data Preparation using Tableau Prep

Learning Objective: Get a brief idea on Data Visualization and Tableau Prep Builder tool.

Topics:

Data Visualization
Business Intelligence tools
Introduction to Tableau
Tableau Architecture
Tableau Server Architecture
VizQL
Introduction to Tableau Prep
Tableau Prep Builder User Interface
Data Preparation techniques using Tableau Prep Builder tool

Hands-On:

Build a simple data flow using Tableau Prep Builder tool
Group and Replace feature using Tableau Prep Builder tool
Pivoting data using Tableau Prep Builder tool
Aggregate data using Tableau Prep Builder tool
Perform Unions and Joins using Tableau Prep Builder tool

Module 2: Data Connection with Tableau Desktop

Learning Objective: Get a brief idea on Tableau UI components and various ways to establish data connection.

Topics:

> Features of Tableau Desktop
> Connect to data from File and Database
> Types of Connections
> Joins and Unions
> Data Blending
> Tableau Desktop User Interface
> Basic project: Create a workbook and publish it on Tableau Online

Hands-On:

> Joins using Tableau Desktop
> Data Blending feature within Tableau
> Create a Workbook and publish it over Tableau Online
> Save a workbook in different formats

Module 3: Basic Visual Analytics

Learning Objective: Understand the importance of Visual Analytics and explore the various charts, features and techniques used for Visualization.

Topics:

> Visual Analytics
> Basic Charts: Bar Chart, Line Chart, and Pie Chart
> Hierarchies
> Data Granularity
> Highlighting
> Sorting
> Filtering
> Grouping
> Sets

Hands-On:

> Basic Charts in Tableau
> Demonstrate Hierarchies, Data Granularity and Highlighting features in Tableau
> Perform Sorting, Filtering and Grouping techniques in Tableau
> Sets in Tableau

Module 4: Calculations in Tableau

Learning Objective: Understand basic calculations such as Numeric, String Manipulation, Date Function, Logical and Aggregate. You will also get introduced to Table Calculations and Level Of Detail (LOD) expressions.

Topics:

Types of Calculations
Built-in Functions (Number, String, Date, Logical and Aggregate)
Operators and Syntax Conventions
Table Calculations
Level Of Detail (LOD) Calculations
Using R within Tableau for Calculations

Hands-On:

Demonstrate calculations using Built-in Functions in Tableau
Perform Quick Table and Level Of Detail (LOD) calculations in Tableau
Installing R and establishing connection with R within Tableau

Module 5: Advanced Visual Analytics

Learning Objective: Deep dive into Visual Analytics in a more granular manner. It covers various advanced techniques for analysing data that includes Forecasting, Trend Lines, Reference Lines, Clustering, and Parameterized concepts.

Topics:

Parameters
Tool tips
Trend lines
Reference lines
Forecasting
Clustering

Hands-On:

Demonstrate Parameters in Calculations

Perform Data Visualization using Trend lines, Forecasting and Clustering feature in Tableau

In-class Project 1- Domain: Media & Entertainment Industry

Module 6: Level of Detail (LOD) Expressions in Tableau

Learning Objective: Deep dive into advanced analytical scenarios, using Level Of Detail expressions.

Topics:

Use Case I - Count Customer by Order
Use Case II - Profit per Business Day
Use Case III - Comparative Sales
Use Case IV - Profit Vs Target
Use Case V - Finding the second order date
Use Case VI - Cohort Analysis

Hands-On:

All the use cases are Hands-on intensive

Module 7: Geographic Visualizations in Tableau

Learning Objective: Gain an understanding of Geographic Visualizations in Tableau.

Topics:

Introduction to Geographic Visualizations
Manually assigning Geographical Locations
Types of Maps
Spatial Files
Custom Geocoding
Polygon Maps
Web Map Services
Background Images

Hands-On:

Create a Map and assign Geographic locations to the fields
Demonstrate how to create a Map from a Spatial file
Learn how to create a Filled Map, Symbol Map, and a Density Map
Perform Custom Geocoding in Maps
Build a Polygon Map
Establish connection with the WMS Server

Module 8: Advanced Charts in Tableau

Learning Objective: Learn to plot various advanced charts in Tableau Desktop.

Topics:

Box and Whisker's Plot
Bullet Chart
Bar in Bar Chart
Gantt Chart
Waterfall Chart
Pareto Chart
Control Chart
Funnel Chart
Bump Chart
Step and Jump Lines
Word Cloud
Donut Chart

Hands-On:

All the above charts have Hands-on

Module 9: Dashboards and Stories

Learning Objective: Build Dashboards and Stories within Tableau.

Topics:

Introduction to Dashboards
The Dashboard Interface
Dashboard Objects

Building a Dashboard

Dashboard Layouts and Formatting

Interactive Dashboards with actions

Designing Dashboards for devices

Story Points

Hands-On:

Demonstrate how to add objects to a Dashboard

Build a simple Dashboard (using Layouts and Formatting features)

Create Interactive Dashboards using actions

Learn to create Dashboard for devices using Device Designer

Build Stories with Dashboards

In-class Project 2- Domain: Retail Industry

Module 10: Get Industry Ready

Learning Objective: Learn effective ways of designing Dashboards with minimum time investment.

Topics:

Tableau Tips and Tricks

Choosing the right type of Chart

Format Style

Data Visualization best practices

Prepare for Tableau Interview

Hands-On:

Hands-on experience on various tips and tricks with Tableau

In-class Industry Grade Major Project-Domain: Transportation Industry

Module 11: Exploring Tableau Online

Learning Objective: Learn to publish data, interact, modify, and secure the published data on Tableau Online.

Topics:

Publishing Workbooks to Tableau Online

Interacting with Content on Tableau Online

Data Management through Tableau Catalog

AI-Powered features in Tableau Online (Ask Data and Explain Data)

Understand Scheduling

Managing Permissions on Tableau Online

Data Security with Filters in Tableau Online

Hands-On:

Publishing Workbooks to Tableau Online

Interacting with Content on Tableau Online

Managing permissions on Tableau Online

Data security using User-based and Row-level filters

In-class Project

Learning Objective: Learn to create Tableau reports for various industrial scenarios and publish them on Tableau Online. Learn to manage permissions and secure data using filters.

Project Statement:

You have been recruited as a freelancer for a Retail store that supplies Furniture, Office Supplies and Technology products to customers across Europe. You have been asked to create interactive dashboards which can be used to gain insights into the profits for orders over the years.

Auto Insurance Case Study

Learning Objectives:

The capstone project will provide you with a business case. You will need to solve this by applying all the skills you've learned in the courses of the master's program. This Capstone project will require you to apply the following skills

Data Exploration

      Checking Data Size
      Note the important features

Data Wrangling

      Handling Imbalanced Data
      MetaData Creation
      Statistics on the Data
      Identify Missing Variable
      Rectify Missing Variable
      One Hot Encoding
      Scaling: Standard Scaler & Min Max Scaler

Data Exploration

      Data Visualization

Machine Learning

      PCA

      Logistic Regression

      Generating F1 Score Metric

      Linear SVC Classifier

      XG Boost Classifier

      AdaBoost Classifier

Deep Learning Learning

      MLP Classifier

      MLP Classifier with Cross Validation