Learnbay

Co-developed with

IBM

FULL STACK DATA SCIENCE AND ARTIFICIAL INTELLIGENCE PROGRAM

For Managers and Business Leaders

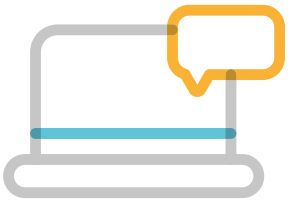
Online and Classroom Mode

300+ hrs of interactive session by industry expert	3 years of flexible subscription	15+ Industrial Projects & Capstone Projects	Domain knowledge & project life-cycle expertise

Program Highlights

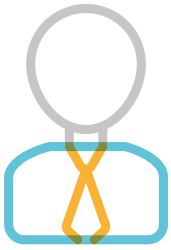
Learnbay offers Data Science and Artificial Intelligence Certification Program which is co-developed and **Certified with IBM**. Course features 15+ real world industry projects and capstone projects under the mentor-ship and guidance of Data Science and AI experts.

Course is especially designed for **business leaders, project managers** having **8+ years of experience** in any domain. Our course is best suited for professionals looking to change their current domain and start a new career in Data science and Artificial Intelligence in senior and leadership role.



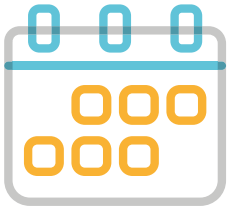
Live Sessions By Expert

- Live Faculty led Online Training.
- Classroom Training available in Bangalore.
- 300+ hrs of Interactive sessions.



Project Based Learning

- 15+ Real World Industry Projects.
- 2 **Capstone Projects.**
- Mentor-ship & Guidance By Expert.

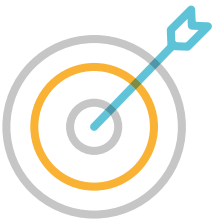


3 Years Subscription

- Flexibility to attend multiple batches from different trainers.
- Life time access to Recordings.

Special Support to Non Programmers

- Learn Python from scratch.
- Special classes for Non programming background aspirants.
- Real time Use Cases from multiple domain.



Certification from IBM in Data Science and AI

- IBM certified Data science and AI program.
- Industry Accredited Global Certification Course.
- Co-developed With IBM.

Analytics & AI project management

- Learn end to end all relevant tools to manage and deliver any data science and AI projects from scratch
- Work On 15+ Real Time Project from BFSI, Retail, Healthcare, Manufacturing

Top Rated Training Institute in India For Data Science And AI Certification

4.8 ★★★★★

300+ user Review

Quora

Top Rated

4.9 ★★★★★

Become IBM Certified Data Science & AI Expert

Program Details

Program Eligibility

Work Experience :

- Professionals working as **project managers and team leads** with **8 to 15 years of experience** in any domain (technical or non technical)

Who Should Apply

- Team managers, Project managers
- Technical lead
- Professionals having 8 to 15 years of work experience, looking to start their career in **data science & analytics manager.**

Click here to apply for profile review

About Instructors

Our instructors are **working professionals** graduated from premier institutes like BITS Pilani, IIT Roorkee and working in companies as **Data Scientist/Machine Learning Engineer and Artificial Intelligence expert.**

Instructors Working in

Microsoft, Mu Sigma, SIGMOID, fractal, AMERICAN EXPRESS, INMOBI, Morgan Stanley, SAMSUNG, HSBC

Course Prerequisite

There is **no Prerequisite** for this course as we cover programming and statistics from basics. We provide special classes & support for professionals from **non-technical background.**

Course Duration

- **Weekday Batches : 11 Months**

  Monday - Friday - 2 hours everyday

- **Weekend Batches : 13 Months**

  Saturday & Sunday - 3.5 hours everyday

Course Fee

INR. 95,000/- (+ 18% GST)

To know more about applicable discount, next batch details...

Live Chat on Whatsapp

PAYMENT MODE

INTEREST FREE INSTANT LOAN WITHOUT CREDIT CARD

Aadhaar Card, Pan Card and 3 Months Salary Slip required

NO COST EMI UPTO 9 MONTHS ON MAJOR CREDIT CARDS

ICICI, HDFC, RBL, Standard Chartered, Axis bank,Kotak credit cards

Tools you will learn

PYTHON

STATISTICS

MACHINE LEARNING

DEEP LEARNING

NATURAL LANGUAGE PROCESSING

MONGO DB

POWER BI

TABLEAU

HADOOP

PROJECT MANAGEMENT IN DATA SCIENCE AND AI

DOMAIN TRAINING IN BFSI, HR, RETAIL, ETC

Banking

E- Commerce

Manufacturing

Healthcare

Aerospace

Finance

Insurance

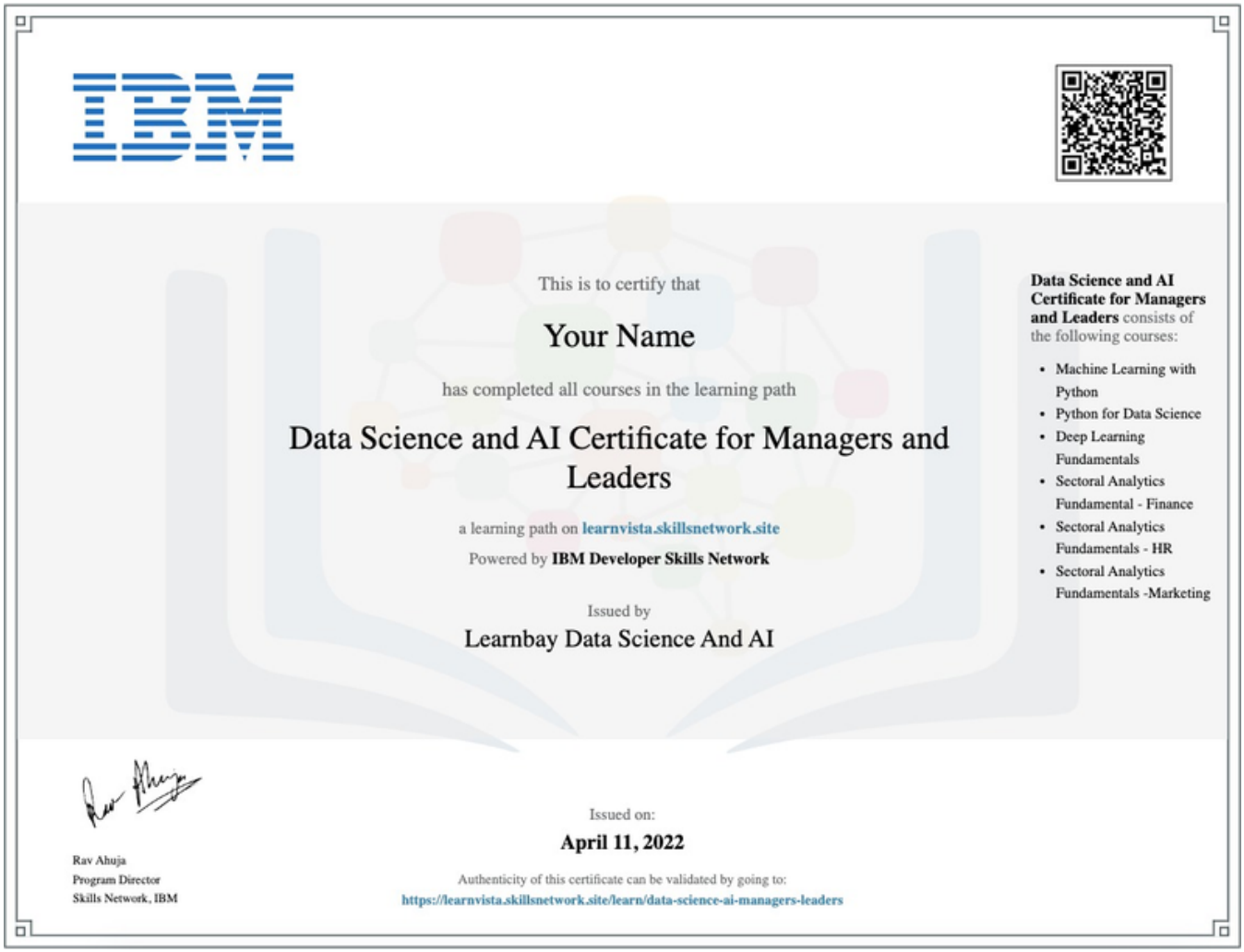
Supplychain

Telecom

Retail

Global Certification from IBM

- Become an industry expert with **Data Scientist & AI Master's Program** in collaboration with IBM. Upon completion of this Program, you will **receive the certificate from IBM** which will help you to increase your professional value.



- Get **Industry-renowned global certification** in Data Science and Artificial Intelligence. Our certification is **recognized globally** and industry wide in companies like JP Morgan, Morgan Stanley, Wells Fargo, Antuit , Genpact, Cognizant, Delloite, E&Y, Tredence Analytics, Mu-sigma, Microsoft, Fossil and other **top MNCs and also Banking & Finance companies**.

Download Certificate

Capstone Project Certification by IBM

Certificate Benefits

- Complete your training with the internationally recognized certificate.
- Validate your Data Science and AI skills with IBM Course Completion Certificate.
- Get acknowledged in IT sector by adding IBM Certificate to your profile.

Why enroll for this program?

Learnbay is specialized in providing personalized courses in Data Science and Artificial Intelligence. We are headquartered in **Bengaluru**, the IT hub of India, and are partnered with IBM since 2019. Our courses have so far helped several talented aspirants from different parts of the world to launch their career in Data Science and AI successfully.
Since the beginning from 2015, we have strongly believed in **quality**, we would never take a chance in compromising with anything lesser than the best quality. Thus, our trainers are without exception, highly experienced field experts.

- Learnbay offers instructor-led interactive program with live doubt solving session as learning from recorded videos can be boring.
- Get hands-on experience with 16+ real time projects and 3 capstone projects, as learning data science would be incomplete without knowing it's practical approach.
- As per the industrial requirement 2 or 3 modules is not sufficient, hence we offer a Full Stack program specially crafted for working professionals.
- Get 1:1 doubt clearing session with expert after your live class. Flexibility to batches, get back up classes and attain session from multiple instructors.
- We over our program with complete flexibility to attain Live/Classroom Session for 3 years and Life - Time Access of LMS. With access to change batches, instructors, etc.

Recording Sample

Python Recording

Statistics

Machine Learning

Deep Learning

NLP

R Programming

Real Time Projects

SUBSCRIBE

SUBSCRIBE US TO WATCH MORE
DATA SCIENCE AND AI VIDEOS

Follow us on **LinkedIn** to stay updated with latest study materials, interview guidance and job updates.

Transition Process

One should be interested in learning programming, statistics & mathematics and Business Knowledge.

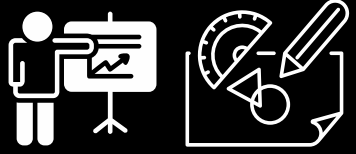
Technical + Managerial Expertise

Step 1

LEARNING PHASE

Learn updated tools & modules from Industrial Experts and understand from basic to advance.

Step 2



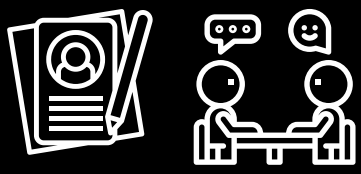
PROJECTS

Show 6-12 months of learning exp. by performing Industrialial projects

Step 3

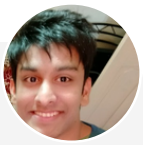
PROFILE GROOMING

Get Interview Ready with experts. Resume Build Up Sessions, Mock Interview & Referral Program



Get placed in top MNC's and Data Science AI Start Up companies as a technical lead, senior Data Scientist, AI Expert

Success Stories



Shezan Baig

Working at Accenture AI

Learnbay is one of the most remarkable data science institutes I've come across. In comparison to other institutes in India, it offers a data science course at a low cost. Excellent value for money. I would strongly advise everyone to attend this institute. All of the trainers are excellent in their own subject, but the Stats & ML trainer in particular is outstanding.

View LinkedIn profile

Viraj Ghodke

Working at Affine Analytics

For me, learning using the Learnbay platform has been a great experience. The teaching and management team are very helpful. They are constantly willing to clarify each individual's doubts and meet the needs of working professionals. I owe my gratitude to the trainers and the whole Learnbay team for assisting me in getting placed.

View LinkedIn profile

Pooja Sharma

Working at Learnvista

I have done Data Science and AI certification and got placed within 8 months. Journey was really tough for me because i was from mechanical domain. But the mentors were really helpful and they have good industrial knowledge. Facility of recording classes is very useful.

View LinkedIn profile

Shubhangi J. Waghmare

Working at Infrasoft Technologies

The offering here is best in the industry I would say both cost and curriculum wise. One advantage joining here is you can access their resources for lifetime unlike others where you have accessibility only for a year or so. Most importantly, there is continuous assistance for recruitment. Well,one enrolls for any course and ends up getting a handsomely paying job.

View LinkedIn profile

Pawan Yadav

Working at Oracle

I have done Data Science certification and i placed in Oracle. Journey was really tough for me because i was from core electronics domain. Mentors are really helpful and they have good knowledge. Personally i liked teaching style of Trainer Nishant. Facility of recording classes is very useful.

View LinkedIn profile

Suman Karmakar

Working at IBM

It was a good and effective course with dedicated faculties for modules.You get flexibility to attend classes from multiple instructors.Very Supportive environment for learning.

View LinkedIn profile

Success Stories

Srikanth Saurav

Working at Mediamarksaturn

Machine Learning concepts & Statistics are very well explained by Utkarsh. Best thing was completing the syllabus on-time as they have promised. Trainers are clearing the doubts . Got multiple joining offers from different MNCs for Data Science and AI developer

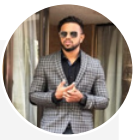
View LinkedIn profile

Afrin Sultana

Working at Fossil

It's a very good place to start with..LB does what it says.They have good faculties for machine learning, statistics, python and some good project sessions as well. Krishna and Abhisekh helped till I got placed. I have got multiple offers after doing the course from here and some extra effort from my end as well.So nothing is bad about it.In one word I would say it's excellent.

Watch Transition Interview



Keerti Bafna

Working at Antuit

I joined the Data Science batch of September. The trainer was Amritansh. And since then i have evolved in Machine Learning drastically . The trainer is very educated and teaches passionately The staff is supporting and you can re-attend and switch classes anytime

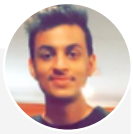
Watch Transition Interview

Rahul Anand

Working at Affine Analytics

Learnbay is one of the best institutes in Bangalore. The faculty members are experienced working professionals and they help you to build the concepts in order to achieve your goals. The whole course and practical sessions are very helpful specially in the field of data science.

Watch Transition Interview



Neelesh Dugar

Working at Act21 Softwares

Very well designed and structured. I really appreciate him and would want to put some light on Utkarsh Kulshrestha. Cheers to you guys! I had an amazing experience at Learnbay, which got me where I am today. Thank you to each one of you and also Abhishek who is handling very well. All the best guys!!

Deevraj

Working at Mindtree

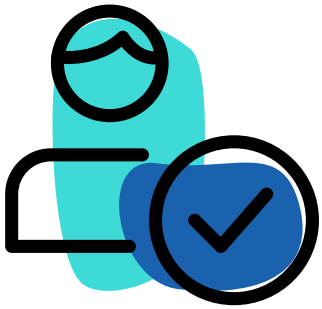
The quality of content is very nice mainly the instructor concentrating on the practical part, live project sessions make you feel confident to attend interviews. Multiple batch options, access for any instructor class videos or materials. Totally positive environment around. One can join here with no second thought.

Read more reviews

Watch Transition Videos

How to apply?

Eligibility Check



Qualification: BE, B. Tech, ME, M.Tech. BCA, MCA (Any Branch), MBA, Etc. All technical or managerial degree. Professionals having 8 to 15 years of experience in any domain. To know more about Eligibility Whatsapp Us

Whatsapp Now

Talk to our admission executive & get your profile reviewed

Attend Personalised Career Counselling and profile review session with expert. This session will help you to understand whether your profile is suitable for Data Science and AI certification course.

Apply For Profile Review

Pay and Enrol For this Program

Contact our Admission Officer for discount coupon. Apply the discount coupon and enrol for IBM certified Program.

Payment Mode: Debit Card, Credit Card, UPI, Bank Transfer, Interest Free Loan, No Cost EMI (Credit Card)

Pay and Enroll for the program

Program Syllabus

MODULE 0

Basic programming fundamentals for Non-programmers

3 days (6 hours)

TERM 1

Core Python + Numpy + Pandas + Matplotlib + Seaborn

(40 hours) : : 1.5 Month Weekday :: 1.5 Months Weekend

TERM 2

Statistics + Machine Learning + Capstone Project

(70 hours) : : 2 Months Weekday :: 2.5 Months Weekend

TERM 3

(SQL + MongoDB) + (Tableau + PowerBI) + Cloud Deployment of ML Model using GCP +(Hadoop basics & Apache Spark) + R Programming

(70 hours) : : 2 Months Weekday :: 2.5 Months Weekend

TERM 4

Deep Learning using Tensor-flow + Natural Language Processing & Text Analytics + Capstone Project

(48 hours) : : 1.75 Months Weekday :: 2.25 Months Weekend

TERM 5

Project Life Cycle + Domain Knowledge + Capstone project - 3 (End to end - Strategy, manage and deliver)

(60 hours) : : 1.75 Months Weekday :: 2 Months Weekend

About Instructors :

- All our instructors are working professionals , working in MNCs, banking domains, startups as ML engineer, data scientist & AI expert.**(min 6 years relevant exp.)**
- **Instructor for Term 5** are leading Data Science and ML Team in MNC and have **15 years of industry experience.**

Special Programming Classes | 6 hours

Module 0

Chapter 1: Introduction to Programming (2 hrs)

What is a programming language ?
Source code Vs bytecode Vs machine code
Compiler Vs Interpreter
C/C++, Java Vs Python

Chapter 2: Jupyter notebook basics (1 hrs)

Different type of code editors in pythonIntroduction to Anaconda and jupyter notebookFlavours of python.

Chapter 3: Python Programming Basics (1 hrs)

Variable Vs identifiers Vs strings
Operators Vs operand Procedure oriented Vs modular programming

Chapter 4: Statistics basics (2 hrs)

Introduction to statisticsMean, median, mode, Standard deviation, AverageIntroduction to probability, permutations and combinationsIntroduction to linear Algebra

Special classes for non-programmers:
Module 0
All the basic concepts regarding programming language and python will be covered in this class. Aspirants from non-programming background must attain this session. Those who are familiar with basic programming fundamentals can skip this module.

Python for Data Science | 40 hours

Module 1
Term 1

1. Programming Basics & Environment Setup

Installing Anaconda ,Anaconda Basics and Introduction
Get familiar with version control, Git and GitHub.
Basic Github Commands.
Intro to Jupyter Notebook environment.
Basics Jupyter notebook Commands.
Programming language basics.

2. Python Programming Overview

Python Overview
Python 2.7 vs Python 3
Writing your First Python Program
Lines and Indentation,Python Identifiers
Various Operators and Operators Precedence
Getting input from User,Comments,Multi line Comments.

3. Strings, Decisions And Loop Control

Working With Numbers, Booleans and Strings,String types and formatting, String operations
Simple if Statement, if-else Statement if-elif Statement. Introduction to while Loops. Introduction to for Loops,Using continue and break.

Class hands-on :

6 programs/coding exercise on string, loop and conditions in classroom

4. Python Data Types

List,Tuples,Dictionaries
Python Lists,Tuples,Dictionaries Accessing Values,Basic Operations
Indexing, Slicing, and Matrixes
Built-in Functions & Methods
Exercises on List,Tuples And Dictionary

- *Program to convert tuple to dictionary*
- *Remove Duplicate from Lists*
- *Python program to reverse a tuple*
- *Program to add all elements in list.*

5. Functions And Modules

Introduction To Functions - Why Defining Functions. Calling Functions
Functions With Multiple Arguments.
Anonymous Functions - Lambda
Using Built-In Modules,User-Defined Modules,Module Namespaces,
Iterators And Generators

Class hands-on :

8+ Programs to be covered in class from functions, Lambda, modules, Generators and Packages.

6. File I/O And Exceptional Handling and Regular Expression

Opening and Closing Files
open Function,file Object Attributes
close() Method ,Read,write,seek.
Exception Handling, try-finally Clause
Raising an Exceptions,User-Defined Exceptions. Regular Expression- Search and Replace. Regular Expression Modifiers. Regular Expression Patterns,re module. *10+ Programs to be covered in class from File IO, Reg-ex and exception handling.*

Python for Data Science | 40 hours

Module 1
Term 1

7. Data Analysis Using Numpy And Pandas

Introduction to **Numpy**. Array Creation, Printing Arrays, Indexing, Slicing and Iterating, Shape Manipulation - Changing shape,stacking and spliting of array. Vector stacking, Broadcasting.

Pandas : Introduction to Pandas Importing data into Python Pandas Data Frames, Indexing Data Frames, Basic Operations With Data frame, Renaming Columns, Subletting and filtering a data frame.

8. Data Visualisation using Python: Matplotlib and Seaborn

Matplotlib: Introduction,plot(),Controlling Line Properties,Subplot with Functional Method, MUltiple Plot, Working with Multiple Figures,Histograms

Seaborn :

Intro to Seaborn And Visualizing statistical relationships , Import and Prepare data .Plotting with categorical data and Visualizing linear relationships Seaborn Exercise

Real time Use cases in Python to be Covered in Class

- ***3 Case Study on Numpy, Pandas , Matplotlib***
- ***1 Case Study on Pandas And Seaborn***

Python Assignments

Assignment 1 (Week 1):

10 Coding exercises on Python Basics - Variables, Operators, Strings, Loops

Assignment 2 (Week 2):

10 Python programs practice & set on List, Tuples, Dictionaries & matrices operations

Assignment 3 (Week 3):

10 Coding exercises on Functions, File And Regular Expression

Assignment 4 (Week 4):

15 Programs and Practice set Questions on Numpy and Pandas

Assignment 5 (Week 5):

2 Case Studies using Numpy Pandas and Matplotlib.

Statistics for Data Science | 24 hours

Module 2
Term 2

1. Fundamentals of Math and Probability

Basic understanding of linear algebra, Matrics, vectors. Addition and Multimplication of matrics.
Fundamentals of Probability
Probability distributed function and cumulative distributed function.
Problem solving using R for vector manipulation
Problem solving for probability assignments

2. Descriptive Statistics

Describe or sumarise a set of data
Measure of central tendency and measure of dispersion.
The mean,median,mode, curtosis and skewness. Computing Standard deviation and Variance.
Types of distribution.

Class Handson:

5 Point summary BoxPlot
Histogram and Bar Chart
Exploratory analytics R Methods

3. Inferential Statistics

What is inferential statistics
Different types of Sampling techniques
Central Limit Theorem
Point estimate and Interval estimate
Creating confidence interval for population parameter
Characteristics of Z-distribution and T-Distribution. Basics of Hypothesis Testing. Type of test and rejection region. Type of errors in Hypothesis yesting,

Type-I error and Type-II errors
P-Value and Z-Score Method
T-Test, Analysis of variance(ANOVA) and Analysis of Co variance(ANCOVA)
Regression analysis in ANOVA

Class Hands-on:

Problem solving for C.L.T
Problem solving Hypothesis Testing
Problem solving for T-test, Z-score test
Case study and model run for ANOVA, ANCOVA

4. Hypothesis Testing

Hypothesis Testing
Basics of Hypothesis Testing
Type of test and Rejection Region
Type o errors-Type 1 Errors,Type 2 Errors. P value method,Z score Method.
The *Chi-Square* Test of Independence.
Regression. Factorial Analysis of Variance. Pearson Correlation Coefficients in Depth. Statistical Significance, Effect Size

5. Data Processing & Exploratory Data Analysis

What is Data Wrangling, Data Pre-processing and cleaning?
How to Restructure the data?
What is Data Integration and Transformation
EDA : Finding and Dealing with Missing Values.What are Outliers? Using Z-scores to Find *Outliers*. Bivariate Analysis,Scatter Plots and Heatmaps.
Introduction to Multivariate Analysis

Machine Learning Algorithm | 46 hours

Module 3
Term 2

Introduction To Machine Learning

What is Machine Learning?
What is the Challenge?
Introduction to Supervised Learning,
Introduction to Unsupervised Learning
What is Reinforcement Learning?
Machine Learning applications
Difference between Machine Learning and Deep Learning

1. Supervised Learning

Support Vector Machines
Linear regression
Logistic regression
Naive Bayes
Linear discriminant analysis
Decision tree
k-nearest neighbor algorithm
Neural Networks (Multilayer perceptron)
Similarity learning

2. Linear Regression

Introduction to Linear Regression
Linear Regression with Multiple Variables
Disadvantage of Linear Models
Interpretation of Model Outputs
Understanding Covariance and Colinearity
Understanding Heteroscedasticity

Case Study - Application of Linear Regression for Housing Price Prediction

3. Logistic Regression

Introduction to Logistic Regression.-
Why Logistic Regression .
Introduce the notion of classification
Cost function for logistic regression
Application of logistic regression to multi-class classification.
Confusion Matrix, Odd's Ratio And ROC Curve
Advantages And Disadvantages of Logistic Regression.

4. Decision Trees

Decision Tree - data set
How to build decision tree?
Understanding Kart Model
Classification Rules- Overfitting Problem.
Stopping Criteria And Pruning. How to Find final size of Trees?
Model A decision Tree.
Naive Bayes. Random Forests and Support Vector Machines. Interpretation of Model Outputs

Case Study:

1 Business Case Study for Kart Model
2 Business Case Study for Random Forest
3 Business Case Study for SVM

Case Study:To classify an email as spam or not spam using logistic Regression.

Machine Learning Algorithm | 46 hours

Module 3
Term 2

5. Unsupervised Learning

Hierarchical Clustering
k-Means algorithm for clustering - groupings of unlabeled data points.
Principal Component Analysis(PCA)-Data. Independent components analysis(ICA)
Anomaly Detection
Recommender System-collaborative filtering algorithm
Case Study- Recommendation Engine for e-commerce/retail chain

6. Natural language Processing

Introduction to natural Language Processing(NLP).
Word Frequency Algorithms for NLP
Sentiment Analysis
Case Study :
Twitter data analysis using NLP

7. Introduction to Time Series Forecasting

Basics of Time Series Analysis and Forecasting ,Method Selection in Forecasting
Moving Average (MA) Forecast Example,Different Components of Time Series Data ,Log Based Differencing, Linear Regression For Detrending

8. ARIMA and Multivariate Time Series Analysis

Introduction to ARIMA Models,ARIMA Model Calculations,Manual ARIMA Parameter Selection,ARIMA with Explanatory Variables
Understanding Multivariate Time Series and Their Structure,Checking for Stationarity and Differencing the MTS
Case Study : Performing Time Series Analysis on Stock Prices

Important Note :
All Machine Learning Algorithms are covered in depth with real time case studies for each algorithm.
Once 60% of ML is completed, **Capstone Project will be released for the batch.**

SQL & MongoDB | 24 hours

Module 4
Term 3

1. RDBMS And SQL Operations :

Introduction To RDBMS. Single Table Queries - SELECT,WHERE,ORDER BY,Distinct,And ,OR. Multiple Table Queries: INNER, SELF, CROSS, and OUTER, Join, Left Join, Right Join, Full Join, Union
Advance SQL Operations:
Data Aggregations and summarizing the data
Ranking Functions: Top-N Analysis
Advanced SQL Queries for Analytics

2. NoSQL Databases :

Topics - What is HBase?
HBase Architecture, HBase Components,
Storage Model of HBase,
HBase vs RDBMS
Introduction to Mongo DB, CRUD
Advantages of MongoDB over RDBMS
Use cases

3. Programming with SQL :

Mathematical Functions
Variables
Conditional Logic
Loops
Custom Functions
Grouping and Ordering
Partitioning
Filtering Data
Subqueries

4. MongoDB Overview :

Where MongoDB is used?
MongoDB Structures
MongoDB Shell vs MongoDB Server
Data Formats in MongoDB
MongoDB Aggregation Framework
Aggregating Documents
What are MongoDB Drivers?

5. Basics and CRUD Operation :

Databases, Collection & Documents
Shell & MongoDB drivers
What is JSON Data
Create, Read, Update, Delete
Finding, Deleting, Updating, Inserting Elements
Working with Arrays
Understanding Schemas and Relations

6. Introduction to MongoDB :

What is MongoDB?
Charateristics and Features
MongoDB Ecosystem
Installation process
Connecting to MongoDB database
Introduction to NoSQL
Introduction of MongoDB module
What are ObjectIds in MongoDb

Tableau & PowerBI | 22 hours

Module 5
Term 3

1. Introduction to Tableau :
Connecting to data source
Creating dashboard pages
How to create calculated columns
Different charts
Hands-on :
Hands on on connecting data source and data cleansing
Hands on various charts

2. Visual Analytics :
Getting Started With Visual Analytics
Sorting and grouping
Working with sets, set action
Filters: Ways to filter, Interactive Filters
Forecasting and Clustering
Hands-on :
Hands on deployment of Predictive model in visualization

3. Dashboard and Stories :
Working in Views with Dashboards and Stories
Working with Sheets
Fitting Sheets
Legends and Quick Filters
Tiled and Floating Layout
Floating Objects

4. Mapping :
Coordinate points
Plotting Latitude and Longitude
Custom Geocoding
Polygon Maps
WMS and Background Image

5. Getting Started With Power BI :
Installing *Power BI Desktop* and Connecting to Data
Overview of the Workflow in Power BI Desktop
Introducing the Different Views of the Data Mode
Query Editor Interface
Working on Data Model

6. Programming with Power BI :
Working with Timeseries
Understanding aggregation and granularity
Filters and *Slicers in Power BI*
Maps, Scatterplots and BI Reports
Connecting Dataset with Power BI
Creating a Customer Segmentation Dashboard
Analyzing the Customer Segmentation Dashboard

Big Data & Spark Analysis | 14 hours

Module 6
Term 3

1. Introduction To Hadoop :
Distributed Architecture – A Brief Overview
Understanding Big Data
Introduction To Hadoop ,Hadoop Architecture
HDFS ,Overview of MapReduce Framework
Hadoop Master – Slave Architecture
MapReduce Architecture
Use cases of MapReduce

2. Apache Spark Analytics :
What is Spark
Introduction to Spark RDD
Introduction to Spark SQL and Dataframes
Using R-Spark for machine learning
Hands-on:
installation and configuration of Spark

Using R-Spark for machine learning programming

3. Apache Spark Analytics :
Getting to know PySpark
Pyspark Introduction
Pyspark Environment Setup
pySpark – Spark context
RDD , Broadcast and Accumulator
Sparkconf and Sparkfiles
Spark MLlib Overview ,Algorithms and utilities in Spark Mlib

Hands-on:
Map reduce Use Case 1 : Youtube data analysis
Map reduce Use Case 2: Uber Data Analytics

Hands-on:
Spark RDD programming

Hands-on:
Spark SQL and Dataframe programming

R programming | 10 hours

Module 7
Term 3

1. Introduction To R :

Installation Setup
Quick guide to RStudio User Interface
RStudio's GUI3
Changing the appearance in RStudio
Installing packages in R and using the library
Development Environment Overview
Introduction to R basics
Building blocks of R
Core programming principles
Fundamentals of R

2. Programming with R :

Creating an object
Data types in R
Coercion rules in R
Functions and arguments
Matrices
Data Frame
Data Inputs and Outputs with R
Vectors and Vector operation
Advanced Visualization
Using the script vs. using the console

3. Manipulating Data :

Data transformation with R - the Dplyr package - Part
Data transformation with R - the Dplyr package - Part
Sampling data with the Dplyr package
Using the pipe operator in R
Tidying data in R - gather() and separate()
Tidying data in R - unite() and spread()

4. Visualizing Data :

Intro to data visualization
Introduction to ggplot2
Building a histogram with ggplot2
Building a bar chart with ggplot2
Building a box and whiskers plot with ggplot2
Building a scatterplot with ggplot2

Why learn R programming?

05 *Gateway to lucrative career*

04 *It has a robust visualization library*

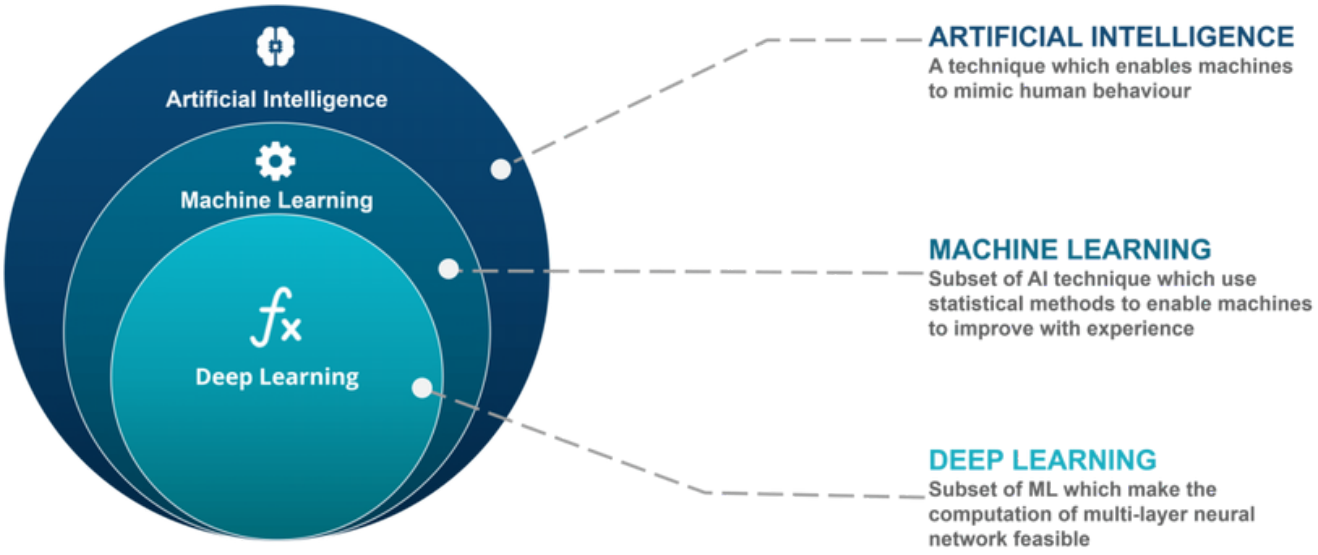
03 *Go to language for Statistics and Data Science*

02 *Latest cutting edge technology*

01 *Independent Platform with free and open-source tools*

Deep Learning | 16 hours

Module 8
Term 4



1. Introduction to Deep Learning And Tensor Flow

Neural Network
Understaing Neural Network Model
Installing TensorFlow
Simple Computation ,Contants And Variables. Types of file formats in TensorFlow. Creatting A Graph - Graph Visualization. Creating a Model - Logistic Regression Model Building using tensor flow
TensorFlow Classification Examples

2. Introduction to Tensor Flow

Installing TensorFlow
Simple Computation ,Contants And Variables
Types of file formats in TensorFlow
Creatting A Graph - Graph Visualization
Creating a Model - Logistic Regression
Model Building
TensorFlow Classification Examples

3.. Understanding Neural Networks With Tensor Flow

Basic Neural Network
Single Hidden Layer Model
Multiple Hidden Layer Model
Backpropagation - Learning Algorithm and visual representation
Understand Backpropagation - Using Neural
Network Example
TensorBoard
Project on backpropagation

4. Convolutional Neural Network(CNN)

Convolutional Layer Motivation
Convolutional Layer Application
Architecture of a CNN
Pooling Layer Application
Deep CNN
Understanding and Visualizing a CNN

Project : Building a CNN for Image Classification

Natural Language Processing | 20 hours

Module 9
Term 4

1. Introduction to NLP & Text Analytics

Introduction to Text Analytics
Introduction to NLP
What is Natural Language Processing?
What Can Developers Use NLP Algorithms For?
NLP Libraries. Need of Textual Analytics
Applications of Natural Language Procession
Word Frequency Algorithms for NLP
Sentiment Analysis

2. Text Pre Processing Techniques

Need of Pre-Processing
Various methods to Process the Text data
Tokenization ,Challenges in Tokenization
Stopping ,Stop Word Removal
Stemming - Errors in Stemming
Types of Stemming Algorithms - Table lookup Approach ,N-Gram Stemmers

3. Distance Algorithms used in Text Analytics

string Similarity
Cosine Similarity Mechanishm - Similarity between Two text documents.
Levenshtein distance - measuring the difference between two sequences.
Applications of Levenshtein distance LCS (Longest Common Sequence) Problems and solutions ,LCS Algorithms

4. Information Retrieval Systems

Information Retrieval - Precision,Recall,F- score
TF-IDF
KNN for document retrieval
K-Means for document retrieval
Clustering for document retrieval

5. Topic Modelling & Dirchlett Distributions

Introduction to Topic Modelling
Latent Dirchlett Allocation
Adavanced Text Analytics & NLP
Introduction to Natural Language Toolkit
POS Tagging
NER

6. Projects And Case Studies

a. Sentiment analysis for twitter, web articles

b. Movie Review Prediction

c. Summarization of Restaurant Reviews

Training & Deployment using GCP | 12 hours

Module 10
Term 4

1. Introduction To GCP Cloud ML Engine

Introduction to Google CloudML Engine
CloudML Engine in Machine Learning WorkFlow
Components of Cloud ML Engine - Google Cloud Platform Console.
gcloud command-line tool and Rest API

2. Training Machine Learning Model

Developing a training application
Packaging a training application
Running and monitoring a training job
Using hyperparameter tuning
Using GPUs for training models in the cloud

3. Deploying Machine Learning Model

Deploying Models ,Understanding training graphs and serving graphs,
Check and adjust model size
Build an optimal prediction graph
Creating input function
creating a model version
Getting Online Prediction

Once all 4 terms are completed we will proceed towards Domain Knowledge Expertise and Project Management Life Cycle in Data Science and Artificial Intelligence.
This module is designed for data science manager aspirants who are targeting senior role in analytics so that you get more insights in multiple domain and also develop skills to manage a data science and AI project end to end in your organisation.

Data Science & AI Project Management Life-cycle | 30 hours

Module 11
Term 5

1. Building AI-ML and Analytics capabilities in organisation /Projects (6 hrs)

Use data science and AI to create and implement business strategies. Incorporate AI on top of existing products and services. Understand the requirements of clients and projects using data driven-methods and perspective. Transform business problems to data analytics Problems. Improve organisation processes using analytics.

2. Building a data science and AI Team (6 hrs)

Defining the data science team, Various job roles in data science - Data analyst, data engineer, data scientist, ML engineer, data science manager. Understand the qualification and expertise of these job roles. Interviewing process for data science job positions. On-boarding the Data Science Team, working with other teams and stakeholders. Challenges and difficulties.

Case Study 1: Strategy for implementing ML in web based retail product (existing project) **Case study 2:** Convert a business problem to data problem from scratch (BFSI domain)

Capstone Project-3: Deliver a machine learning and AI Project from scratch starting from strategy, planning, team building, modeling and deployment. You will be assigned 4 machine learning engineers and you need to deliver this project under mentor-ship of our expert in 3 months time.

3: Understanding the data science workflow and Pipeline (6 hrs)

Organize a data science project workflow from scratch, workflow of a data science project, Identifying various data sources, Cleaning your data, EDA, creating ML model, Model deployment, Monitoring. Building data pipelines, Type of data, evolution of pipelines.

4. Managing data science Projects (6 hrs)

Determine Business Objectives, Determine Data Mining Goal, Understand data, Verify Quality of data,Getting familiar with CRISP-DM methods and process, Business understanding, Data understanding, data preparation, modeling, evaluation, Deployment.

Case study : Solve a retail business problem and follow CRISP-DM process and methods.

5: Architecture of AI & ML Systems (6 hrs)

Strategy for building artificial intelligence systems for business. Understand AI infrastructure requirements, Importance of machine learning system architectures and their various components. Build machine learning system architectures for a chat-bot, recommended systems etc.

Case Study/Practice: Live interview of 2 candidates for data analyst and ML engineer Job Roles.

Banking & Finance analytics Domain
(6 hrs)

If there is one field that is immensely helped by Data Science is the Banking and Finance industry. They deal with abundance of money, which usually attracts the cases of fraud activity, pseudo activity by people around them. It was naturally difficult for banks management to all time protect the money and documents from the fraudulent, but things changed after Data Science happened to them.
According to a survey the fraud cases in banks have been reduced to 40% since the application of Data science. From customer service to fraud detection, banking majorly takes help of Data Science. It is popularly used in Risk Management, customer segmentation, credit risk and analysis.

Let's take a look at the applications of Data Science in Banking:

- Fraud detection
- Managing customer data
- Risk modeling for investment banks
- Real-time and predictive analytics
- Customer segmentation
- Process Automation
- Security
- Underwriting and credit scoring
- Algorithmic trading
- Robo-advisory

Retail Sector
(4 hrs)

In the retail industry, processes like cross-selling and up-selling are practiced by all companies and retailers in order to improve their revenue. Applying Data Science in retailing will help in increase of profits without running A/B tests. Data Science will help this sector majorly in terms of personalizing offers to different customer segments, this strategy has found 80% of active success results.

Let's take a look at the applications of Data Science in Retail:

- Recommendation Engines
- Market basket analysis
- Warranty analytics
- Warranty analytics
- Inventory management
- Customer sentiment analysis
- Fraud detection

Foretelling trends through social media

Healthcare domain
(4 hrs)

The healthcare industry is no more just about a business between a doctor and patient; it is a fundamental field that consist confidential most data of people, which is about health. A survey revealed that healthcare fields store 30 percent of global data. The data which this field consist, could help government in various ways.
Some of the most effective uses of Data Science in healthcare is in medical imaging. Data Science with the backup of Machine Learning makes computers learn to interpret MRI's, X-rays, mammo-graphics and other different medical reports. It identifies the concerning most issues like cancerous cells, tumors, artery issues, organ anomalies and etc, by reading into different layers of the report and through identifying correlations between data.

Let's take a look at the applications of Data Science in Healthcare:

- Medical Image Analysis:
- Genetics and Genomics
- Creation of drugs
- Virtual assistance for patients and customer support
- Predictive Analytics: prognosis and diagnostic accuracy
- Managing customer data
- Industry knowledge
- "How Many Clicks Is Too Many Clicks?" or A/B Testing.
- Research and Clinical Trials
- Optimal staffing
- Monitoring Patient Health
- Improving diagnostic accuracy and efficiency
- Workflow Optimization and Process Improvements
- Reduced healthcare costs

E-Commerce Domain
(4 hrs)

We can really say that it E-commerce fueled hopes and visions of the current technology-driven world, this industry deserves the help of Data Analytics.

Let's take a look at the applications of Data Science in E-commerce:

- Recommendation engines
- Market Basket Analysis
- Leverage on predictive forecasting
- Determine customer behavior and shopping patterns
- Improve customer experience
- Prevent fraud
- Winding up

Supply Chain Management
(4 hrs)

This sector runs differently from other popular ones, other intelligent features of Data Science will get the platform to be presented through such sectors. Its vigilant study of Data will identify the correlations and hidden relations between data sets, its accurate predictions allows to take real-time decisions which will eventually improve the concerned aspect of the sector.

Let's take a look at the applications of Data Science in Supply chain management:

- Efficiency in the sharing and treatment of data:
- Optimized transport and Logistics
- Feedback for Quality Improvement
- Build Long Term Stability:
- Predictive Analytics to Predict Outcomes
- Automating Recruitment Processes
- IoT Features Relevant to Supply Chain
- Blockchain Features Relevant to Supply Chain

Telecom Sector
(3 hrs)

Even though telecom industry is different from Data science, it immensely needs the assistance through Data Science in fields like streamlining the operations, profit building, effective marketing techniques, business strategies, data Visualization, data transfer and etc.

Let's take a look at the applications of Data Science in Telecom sector:

- Customer Retention- Churn Model
- Lifetime value prediction
- Network management and optimization
- Customer segmentation
- Fraud detection
- Predictive analytics
- Product development
- Recommendation engines
- Real-time analytics
- Price optimization
- Future of data in telecom industry
- Important Reviews- Improved Customer Service

Manufacturing Domain
(3 hrs)

Data science is a multidisciplinary area which is responsible for processing and visualizing data of all kinds, big and small. It is the study of statistics and probability which can provide powerful insights for manufacturers when fed enough data into the right data model. Data science is used in manufacturing for a variety of reasons, fueled predominantly by an increase in IoT devices which send efficiency and process data to the cloud.

Let's take a look at the applications of Data Science in Manufacturing sector:

- Real-time Performance Data and Quality
- Fault Prediction and Preventive Maintenance
- Demand Forecasting and Inventory Management
- Supply Chain Optimization
- Price Optimization
- Robotics, Automation and Smart Factory Design
- Product Development and Material Design

Automotive Domain
(3 hrs)

Data science and machine learning are the emerging priorities to be used in the automotive industry in the coming years when it comes to processes and products with automatic learning and optimization techniques.
In every stage of automotive production, data science, machine learning, and inevitably AI can increase productivity, enabling companies to reduce expenses, improve customer experience, and perhaps most importantly evolve new , innovative products.

Let's take a look at the applications of Data Science in Automotive sector:

- Predictive Vehicle Technology
- Self-Driving Technology
- Cars-as-a-Service (CaaS)
- Increased production line performance
- More Fuel-Efficient Rides
- Optimized inventory management
- Transparent supply chain
- Better quality of produced vehicles

Real Time Industrial Projects

1 Project - Loan Default Prediction

Domain - Banking & Finance

The bank wants to improve their services by finding interesting groups of clients. Fortunately, the bank stores data about their clients, the accounts (transactions within several months), the loans already granted, the credit cards issued. This process of loan default prediction can be done with machine learning algorithms.

2 Project - Analyzing Health Data and tracking human activity

Domain - Healthcare

The goal is to breakdown all the data that the Samsung Health app has collected and see what useful insights we can gain by analyzing it.

3 Project - IBM HR Analytics

Domain - Demand/Supply

Applying analytic processes to the human resource department of an organization in the hope of improving employee performance and therefore getting a better return on investment. This is especially concerning if your business is customer facing, as customers often prefer to interact with familiar people.

4 Project- Forecasting Uber Demand

Domain - Demand/Supply

The goal is to create an interactive dashboard using Tableau. This Tableau Dashboard can be used to get historical insights into a neighborhood,
For example,
see its upcoming forecasted demand,
increase the accuracy,
decrease surge pricing events.

5 Project : Clustering Customers

Domain - Retail industry

Big Bazaar has retail outlets across major metropolitan cities in India. With the help of machine learning algorithms we can better understand customer behavior and understand their buying needs better. BigBazaar runs various loyalty programs, festive offers which provide their customer more opportunities to avail discounts.

6 Project - Identify fraudulent credit card transactions.

Domain - Banking & Finance

To recognize fraudulent credit card transactions so that customers are not charged for items that they did not purchase. It involves various processes like Data Cleaning, Data Visualization, Insights generation, Model generation, Feature Engineering and so on.

Real Time Industrial Projects

7 Project - Consumer Reviews of Amazon Products

Domain - E-Commerce

The goal is to analyze Amazon's most successful consumer electronics product launches; discover insights into consumer reviews and assist with machine learning models.
What are the most reviewed Amazon products?
How do the reviews in the first 90 days after a product launch?

amazon

8 Project - Airbnb New User Bookings

Domain - Travel & Hospitality

The goal is to predict which country a new user's first booking destination will be. By accurately predicting where a new user will book their first travel experience, Airbnb can share more personalized content with their community, decrease the average time to first booking, and better forecast demand.

airbnb

9 Project - Netflix Movies and TV Shows

Domain - Media and Entertainment

Explore what all other insights can be obtained from the list of tv shows and movies available on Netflix as of 2019. Understanding what content is available in different countries. Identifying similar content by matching text-based features. Network analysis of Actors / Directors and find interesting insights.

NETFLIX

10 Project - Sales Forecasting

Domain - Retail

This dataset contains the sales for each department from the Walmart dataset containing data of 45 Walmart stores, selected holiday markdown events are also included . These markdowns are known to affect sales, but it is challenging to predict which departments are affected and the extent of the impact.

Walmart

11 Project - BMW Pricing Challenge

Domain - Automation

To find a good statistical model to describe the value of a used car depending on the basic description
How does the estimated value of a car change over time? Can you detect any patterns?
How big is the influence of the factors not represented in the data on the price?

BMW

12 Project - Bosch Production Line Performance

Domain - Manufacturing

To predict internal failures using thousands of measurements and tests made for each component along the assembly line. This would enable Bosch to bring quality products at lower costs to the end user.
The goal is to predict which parts will fail quality control

BOSCH

Real Time Industrial Projects

13 Project - Trending YouTube Video Statistics

Domain - Social Media

The dataset of this project are daily record of the top trending YouTube videos, to generate insights like :
Sentiment analysis in a variety of forms
Categorising YouTube videos based on their comments and statistics
Training ML algorithms like RNNs to generate their own YouTube comments.

14 Project - Identify And Predict Customer churn in telecom industry

Domain - Telecom

The goal is to develop a churn prediction model which assists telecom operators to predict customers who are most likely subject to churn. Also to understand the customer behavior and reasons for churn. Apply multiple classification models to predict the customer churn in telecom industry.

15 Project - Smart Supply Chain for Big Data Analysis

Domain - Supply Chain

A DataSet of Supply Chains used by the company DataCo Global is used for the analysis. Dataset of Supply Chain , which allows the use of Machine Learning Algorithms and R Software.
It also allows the correlation of Structured Data with Unstructured Data for knowledge generation.

16 Project- Generating Chatbot

Domain - Machine Learning

In this project we will build a simple retrieval based chatbot based on NLTK library in python, to perform tasks such as automatic summarization, translation, named entity recognition, relationship extraction, sentiment analysis, speech recognition, and topic segmentation.

Watch the videos to know more about Projects :

HUMAN ACTIVITY

FRAUD DETECTION

CREDIT RISK ANALYSIS

RAPIDO PROJECT

GET IN TOUCH WITH US:

Talk with our counselor
+91 7349 2222 63

Click here to chat with us on
WhatsApp

Write us an email
contacts@learnbay.co

Click here to visit our website

Tap the icon to follow us on social media

Learnbay ,147, 5th Main Rd, Rajiv Gandhi Nagar, HSR Sector 7, Near Salarpuria Serenity, Bengaluru, Karnataka 560102 INDIA

Learnbay