PG Certificate Course on

Data Science, AI/ML and Data Engineering

8 Months Program | Certificate by E&ICT, IIT Roorkee

+080-4920-2224 reachus@cloudxlab.com www.cloudxlab.com

About IIT Roorkee & EICT

IIT Roorkee (IITR) has been ranked the best among IITs, as per the QS World Best Universities Ranking 2019. E&ICT-IITR provides certification courses with emphasis on hands-on learning in basic and advanced topics and emerging technologies in Electronics and ICT. It is sponsored by the Ministry of Electronics and Information Technology, Govt. of India.

QS WORLD UNIVERSITY RANKINGS

#1st

Among the IITs in the 'Citations per Faculty' parameter
*QS World Rankings

INDIA TODAY

#5

Ranked Engineering College
*India Today 2020

nirf NATIONAL INSTITUTIONAL RANKING FRAMEWORK

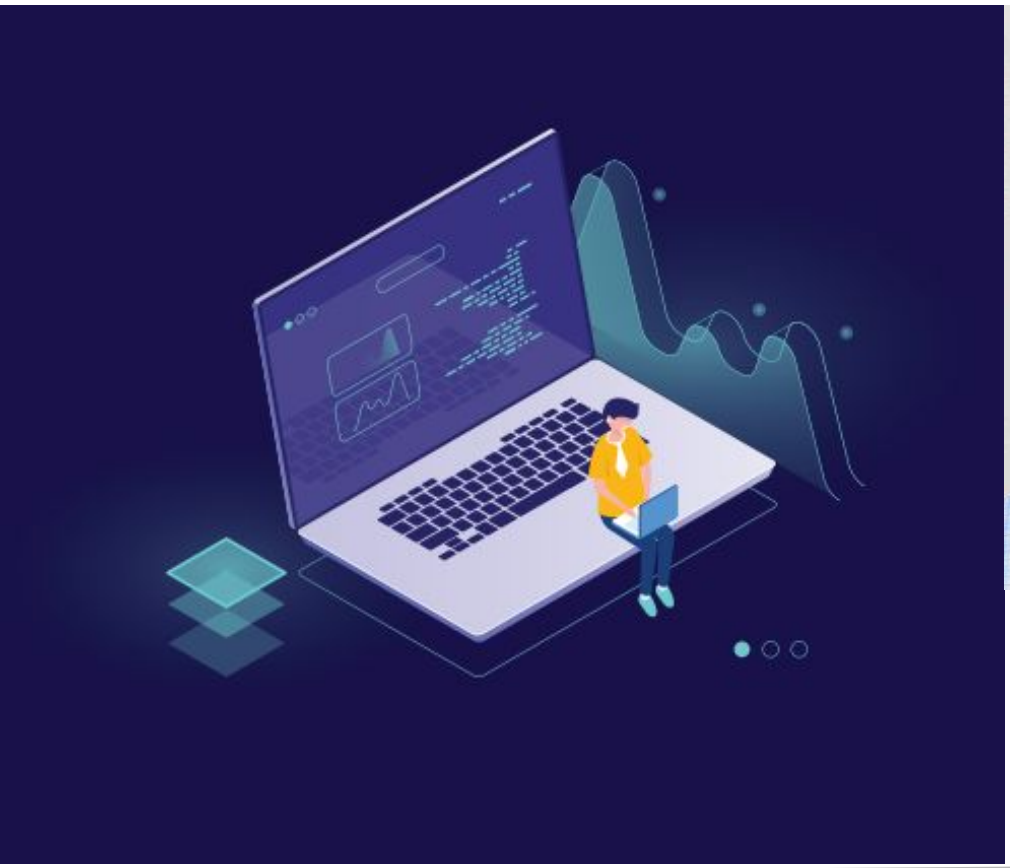
#6

Ranked for IITs
*NIRF 2020

QS WORLD UNIVERSITY RANKINGS

#12

Ranked Best Global Universities in India
*QS World Rankings



Prof. Sanjeev Manhas
Coordinator E-Learning Centre
IIT Roorkee

CloudxLab & Course

Cloudxlab is a team of developers, researchers, and educators who create gamified learning experiences for users. The company upskills engineers in deep tech to make them employable & future-ready. Cloudxlab is proud to collaborate with EICT-IITR to empower learners to solve complex problems with this Data Science course.

Our Students Work At

and more!

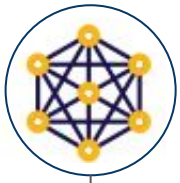
Sandeep Giri

Founder at CloudxLab

+080-4920-2224 reachus@cloudxlab.com www.cloudxlab.com

Why CloudxLab & IIT Roorkee

Earn a certificate from IIT Roorkee.



Learn Data Science, AI/ML & Data Engineering from IIT Roorkee professor and industry experts and become expert in Data Science domain

Online cloud lab for hands-on for real-world experience

Best-in-class support Throughout your learning journey

Lifetime course access

Work on real-world projects.

Interact with the international community of peers via the discussion forum.

Mentors / Faculties

Tharun Kumar Reddy Bollu
Faculty ECE Dept,
IIT Roorkee

Course Instructor
Know More

Raksha Sharma
Faculty CSE Dept,
IIT Roorkee

Course Instructor
Know More

Sandeep Giri
Founder at CloudxLab
Past: Amazon, InMobi, D.E.Shaw

Course Developer
Know More

Abhinav Singh
Co-Founder at CloudxLab
Past: Byjus

Course Developer
Know More

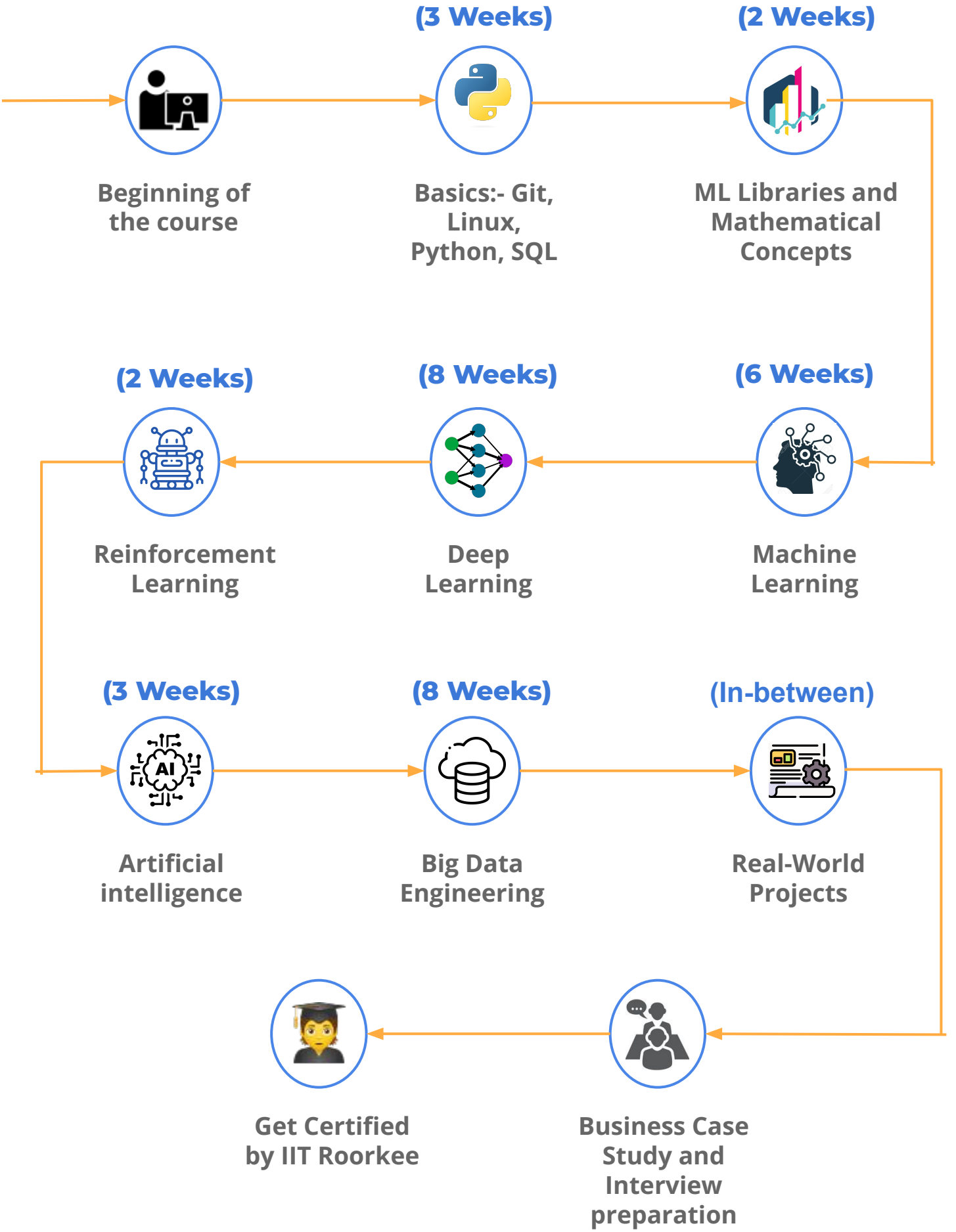
Praveen Pavithran
Co-Founder at Yatis
Past: YourCabs, Cypress Semiconductor

Course Developer
Know More

Venkat Karun
Staff Software Engineer
Google

Course Advisor
Know More

Learning Path



Course Curriculum

1: Foundations

- Linux for Data Science/ Machine Learning
- Getting Started with Git
- Python Foundations
- Machine Learning Prerequisites(Including Numpy, Pandas and Linear Algebra)
- Getting Started with SQL
- Statistics Foundations

Course on Machine Learning

1: Machine Learning Applications & Landscape

- Introduction to Machine Learning
- Machine Learning Application
- Introduction to AI
- Different types of Machine Learning - Supervised, Unsupervised

2: Building end-to-end Machine Learning Project

- Machine Learning Projects Checklist
- Get the data
- Launch, monitor, and maintain the system
- Explore the data to gain insights
- Prepare the data for Machine Learning algorithms
- Explore many different models and short-list the best ones
- Fine-tune model

Course Curriculum

3. Classification

- Training a Binary classification
- Multiclass,Multilabel and Multioutput Classification
- Performance Measures
- Confusion Matrix
- Precision and Recall
- Precision/Recall Tradeoff
- The ROC Curve

4. Training Models

- Linear Regression
- Gradient Descent
- Polynomial Regression
- Learning Curves
- Regularized Linear Models
- Logistic Regression

5. Support Vector Machines

- Linear SVM Classification
- Nonlinear SVM Classification
- SVM Regression

Course Curriculum

6: Decision Trees

- Training and Visualizing a Decision Tree
- Making Predictions
- Estimating Class Probabilities
- The CART Training Algorithm
- Gini Impurity or Entropy
- Regularization Hyperparameters
- Instability

7: Ensemble Learning and Random Forests

- Voting Classifiers
- Bagging and Pasting
- Random Patches and Random Subspaces
- Random Forests
- Boosting and Stacking

8: Dimensionality Reduction

- The Curse of Dimensionality
- Main Approaches for Dimensionality Reduction
- PCA
- Kernel PCA
- LLE
- Other Dimensionality Reduction Techniques

Course Curriculum

Course on Deep Learning and Reinforcement Learning

1: Introduction to Artificial Neural Network

- From Biological to Artificial Neurons
- Implementing MLPs using Keras with TensorFlow Backend
- Fine-Tuning Neural Network Hyperparameters

2: Convolutional Neural Networks and Computer Vision

- The Architecture of the Visual Cortex
- Convolutional Layer
- Pooling Layer
- CNN Architectures
- Classification with Keras
- Transfer Learning with Keras
- Object Detection
- YOLO

3: Recurrent Neural Network

- Recurrent Neurons and Layers
- Basic RNNs in TensorFlow
- Training RNNs
- Deep RNNs
- Forecasting a Time Series
- LSTM Cell
- GRU Cell

Course Curriculum

4: Natural Language Processing

- Introduction to Natural Language Processing
- Creating a Quiz Using TextBlob
- Finding Related Posts with scikit-learn
- Generating Shakespearean Text Using Character RNN
- Sentiment Analysis
- Encoder-Decoder Network for Neural Machine Translation
- Attention Mechanisms
- Recent Innovations in Language Models

5: Training Deep Neural Networks

- The Vanishing / Exploding Gradients Problems
- Reusing Pretrained Layers
- Faster Optimizers
- Avoiding Overfitting Through Regularization
- Practical Guidelines to Train Deep Neural Networks

6: Custom Models and Training with TensorFlow

- A Quick Tour of TensorFlow
- Customizing Models and Training Algorithms
- Tensorflow Functions and Graphs

7: Loading and Preprocessing Data with TensorFlow

- Introduction to the Data API
- TFRecord Format
- Preprocessing the Input Features
- TF Transform
- The TensorFlow Datasets (TDFS) Projects

Course Curriculum

8: Autoencoders and GANs

- Efficient Data Representations
- Performing PCA with an Under Complete Linear Autoencoder
- Stacked Autoencoders
- Unsupervised Pre Training Using Stacked Autoencoders
- Denoising Autoencoders
- Sparse Autoencoders
- Variational Autoencoders
- Generative Adversarial Networks

9: Reinforcement Learning

- Learning to Optimize Rewards
- Policy Search
- Introduction to OpenAI Gym
- Neural Network Policies
- Evaluating Actions: The Credit Assignment Problem
- Policy Gradients
- Markov Decision Processes
- Temporal Difference Learning and Q-Learning
- Deep Q-Learning Variants
- The TF-Agents Library

Course Curriculum

Course on Recent Innovations in AI

1: OpenAI's ChatGPT

- Introduction to ChatGPT
- Architecture of GPT
- ChatGPT Architecture and Training

2: Stable Difusion

- Introduction to Stable Diffusion
- Stable Diffusion Components
- Diffusion Model
- Stable Diffusion Architecture and Training

Course on Data Engineering

- Introduction to Hadoop
- Foundation & Environment
- Zookeeper, HDFS, YARN
- MapReduce Basics
- MapReduce Advanced
- Analyzing Data with Pig
- Processing Data with Hive
- NoSQL and HBase
- Importing Data with Sqoop and Flume, Oozie
- Introduction to Spark
- Scala Basics
- Writing and Deploying Spark Applications
- Common Patterns in Spark Data Processing
- Data Formats & Management
- DataFrames and Spark SQL
- Machine Learning with Spark

Projects

1. Churn Email Inbox with Python
   - Churn the mail activity from various individuals in an open source project development team.
2. Predicting the median housing prices in California
   - In this project we will build a machine learning model to predict housing prices. We will learn various data manipulation, visualization and cleaning techniques using various libraries of Python like Pandas, Scikit-Learn and Matplotlib.
3. Noise removal from the images
   - Build a model that takes a noisy image as an input and outputs the clean image.
4. Build a spam classifier
   - Build a model to classify email as spam or ham. First, download examples of spam and ham from Apache Spam Assassin's public datasets and then train a model to classify email.
5. Predict which passengers survived in the Titanic shipwreck
   - The sinking of the RMS Titanic is one of the most infamous shipwrecks in history. In this project, you build a model to predict which passengers survived the tragedy.
6. Predicting Noisy Images using KNN Classifier</h4>
   - We will learn how to predict images from their noisy version. We will use the MNIST dataset for this project. First, we will load the dataset, explore it, and they we will learn how to introduce noise to an image. Next we will train a KNN Classifier to predict the original image from it's noisy version.
7. Credit Card Fraud Detection using Machine Learning
   - Learn how to over-sample the dataset with imbalanced classes using the SMOTE technique and how to use the thus obtained data to build a fraudulent transaction classifier.
8. Building Cat vs Non-Cat Image Classifier using NumPy
   - Use Python and Numpy to build a Logistic Regression Classifier from scratch, and apply it to predict the class of an input image - whether it is a cat or a non-cat.
9. Iris Flowers Classification using Deep Learning and Keras
   - Use Python and Tensorflow 2 Keras to build a dense deep neural network classifier to predict the classes of flowers in the Iris dataset.

Projects

10. Classify Clothes from Fashion MNIST Dataset
    - Build a model to classify clothes into various categories in Fashion MNIST dataset.
11. Sentiment Analysis
    - Sentiment analysis of "Iron Man 3" movie using Hive and visualizing the sentiment data using BI tools such as Tableau.
12. Spark application
    - Write end-to-end Spark application starting from writing code on your local machine to deploying to the cluster.
13. Parse Apache Access Logs using Spark.
    - The logs of a webserver are the gold mines for gaining insights in the user behavior. So learn to parse the text data stored in logs of a web server using the Apache Spark.
14. Analytics Dashboard
    - Build real-time analytics dashboard for an e-commerce company using Apache Spark, Kafka, Spark Streaming, Node.js, Socket.IO and Highcharts
15. MovieLens Project
    - Analyze MovieLens data using Hive
16. Spark MLlib
    - Generate movie recommendations using Spark MLlib
17. Predict bikes rental demand
    - Build a model to predict the bikes demand given the past data.
18. Process the NYSE
    - Process the NYSE (New York Stock Exchange) data using Hive for various insights.

Campus Immersion

- 1:1 with Professors and Industry Experts
- Certificate award ceremony
- Network with your peers
- Showcase your course project to Professors and peers
- Once in a lifetime experience

Campus Immersion Program | Year 2022

Course Details and Fees

Please find more information about the course and fees here:
https://cloudxlab.com/course/165/pg-certificate-course-in-data-science-aiml-data-engineering-by-eict-academy-iit-roorkee

Our Esteemed Customers

For Further Details

Contact us at +080-4920-2224 or +1 412-568-3901 or contact:

For Business

For corporate training and bulk enrollments, write to us at reachus@cloudxlab.com

Headquarters - United States

2035, Sunset Lake Road Suite B-2, 19702
Newark, New Castle
Delaware, United States

R&D Center - India

Issimo Technology Private Limited
1665 27TH Main, 19th Cross Rd, Sector 2,
HSR Layout, Bengaluru, Karnataka 560102