

<b>Programme Name</b>	Executive Certification in Advanced Data Science & Applications		
<b>Certificate Issued by</b>	CODE & Pravartak-IITM		
<b>Certificate Type</b>	Participants who successfully meet the evaluation criteria and satisfy the requisite attendance criteria will be awarded a 'Certification of Completion' - <b>Executive Certification in Advanced Data Science &amp; Applications</b>		
<b>Name of the Faculties who shall teach in the Programme.</b>	<ol style="list-style-type: none"> <li>1) Ganapathy Krishnamurthi,</li> <li>2) Balaji Srinivasan</li> </ol>		
<b>Programme Introduction and Objectives</b>	<p><b>Introduction:</b> Data science techniques and associated methods in Artificial intelligence and Machine learning have now at the forefront of revolution in various traditional fields. Consequently, increasing number of professionals in the field of scientific computing, software engineering and development, Business are looking to increase their understanding of the fundamental techniques and ideas driving this field. The current program aims to empower professionals to move to the forefront of this revolution.</p> <p><b>Objectives:</b></p> <ol style="list-style-type: none"> <li>1. Provide a thorough introduction to the various methods in the field of data analytics, Machine Learning &amp; Artificial Intelligence and its mathematical foundations</li> <li>2. Provide contextual understanding using case studies from various business verticals</li> </ol> <p>Self-study by applications of the various techniques to real life data from various business verticals</p> <p><b>Learning Outcome:</b>  Basic proficiency in Python, R and mySQL  Basic proficiency in Machine Learning frameworks such as Pytorch and TensorFlow  Ability to formulate and program models for computer vision tasks  Ability to formulate and program models for natural language processing (NLP) tasks  Ability to implement business applications such as recommender systems, customer segmentation  Ability to formulate and program predictive models using advanced time series techniques  Developing the ability to understand emerging paradigms in advanced data analytics</p> <p><b>Projected Career Growth:</b>  The candidate will be able to incorporate the techniques learnt in this case for his/her business case and to interface with other teams using such techniques. It will also help candidates to cross over into a career in data analytics and Machine learning. Middle level executives can benefit by understanding the strength and limitations of these methods in their business case.</p>		
<b>Eligibility</b>	<ul style="list-style-type: none"> <li>• Qualification: Graduate / 4-year Engineering Degree / B.Sc+M.Sc from a recognized university (UGC/AICTE/DEC/AIU/State Government/recognized international universities)</li> <li>• Minimum Experience : 3 years preferably in software engineering and /or other disciplines involved in computational work</li> <li>• Industry Targeting (Preference): IT,Software, Engineering Research business analytics, Finance etc</li> </ul>		
<b>Duration</b>	10 months 120 hrs		
<b>Course Schedule</b>	Frequency: Once a week Day: Sunday Time: 10 am to 1 pm	<b>No. of sessions:</b>	40 sessions including in-campus session

<p><b>Pedagogy</b></p>	<p><b>Insight on Capstone project</b></p> <p>No capstone project</p> <p>Minor project will start halfway through the course. The participants are expected to present their projects in the campus visit</p> <p><b>Number of case studies along with their key details</b></p> <p>For the program there will be 5 case studies. Some example case studies are (subject to change)</p> <ol style="list-style-type: none"> <li>1. Recommender systems</li> <li>2. Time series analysis</li> <li>3. Anomaly detection</li> <li>4. Online Advertisement</li> <li>5. Finding similar items</li> <li>6. Mining data streams</li> </ol> <p>There will also be homework assignments that will also serve as mini case studies. These will include studies like above with data for homework, in addition to other topics in Machine Learning</p> <p><b>Number of simulation</b></p> <p>No Simulations</p> <p><b>Number of Projects</b></p> <p>None</p>						
<p><b>Programme Content</b></p>	<ul style="list-style-type: none"> <li>• <b>Content:</b></li> </ul> <p>Module1 – Mathematical Preliminaries  Module 2 – Tools  Module 3 – Algorithms – supervised &amp; unsupervised learning methods &amp; Statistical methods  Module 4 – Business applications &amp; implementations</p> <p>(Syllabus breakdown with modules and key pointers on sub-modules)</p> <ul style="list-style-type: none"> <li>• <b>Tools to be used: Python, R, Matlab</b></li> <li>• Learning Level of the tools: <b>Moderate proficiency will be helpful</b></li> </ul>						
<p><b>Assessment</b></p>	<p>Homework, Final exam, project</p>						
<p><b>Fees</b></p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%;"></th> <th style="width: 50%;"><b>Total Fees (Rs.)</b></th> </tr> </thead> <tbody> <tr> <td>Application Fees</td> <td>Rs.1500/- + GST</td> </tr> <tr> <td>Programme Fee</td> <td>Rs.200,000/- + GST</td> </tr> </tbody> </table>		<b>Total Fees (Rs.)</b>	Application Fees	Rs.1500/- + GST	Programme Fee	Rs.200,000/- + GST
	<b>Total Fees (Rs.)</b>						
Application Fees	Rs.1500/- + GST						
Programme Fee	Rs.200,000/- + GST						

--	--

# **Syllabus for Advanced Data Analytics Course**

Ganapathy Krishnamurthi (Department of Engineering Design)

Balaji Srinivasan (Department of Mechanical Engineering)

14th September 2022

## **1) Module 1 – Introduction to the Course (3 hours)**

- a) Overview of Course
- b) Introduction to Data Science
  - i) Definition
  - ii) Application areas
  - iii) Past, Present and Future
- c) The role of computation in Data science
- d) The role of mathematics in Data science

## **2) Module 2 – Computational Skills for Data Science (20 hours) (Campus Visit-10 hours)**

- a) Python
  - i) Install and run Python, use IDE
  - ii) Basic control structures
  - iii) Importing and Manipulating data, tools for plotting
- b) R
  - i) Getting started with R
  - ii) Basics, control structures
  - iii) Functions
  - iv) Importing and manipulating data
- c) SQL
  - i) Getting started with MySQL
  - ii) Creating, inserting, retrieving records
  - iii) Searching
  - iv) Interfacing with Python and R

## **3) Module 3- Basic Tools for Data Science (30 hours)**

This module covers the fundamental mathematical and statistical tools required for data science. All of the techniques will be discussed in the context of data science with relevant examples

- a) Probability and Statistics

- i) Probability and counting, Bayes Theorem
  - ii) Independence, Conditional Probability, Marginal Probability,
  - iii) Random Variables, Probability Distributions
  - iv) Expectation, Variance, Covariance
  - v) Descriptive Statistics
  - vi) Statistical Estimation
  - vii) Hypothesis Testing
  - viii) Predictive and Prescriptive Analytics
  - ix) Practical Considerations in Data Science
- b) Linear Algebra
- i) Vectors and their operations
  - ii) Matrices and their operations
  - iii) Inner Products and norms
  - iv) Matrix Decomposition – Eigenvalues, SVD
  - v) Applications
- c) Calculus and Optimization
- i) Partial Derivatives, Multivariable Calculus
  - ii) Gradient, Jacobian
  - iii) Automatic Differentiation
  - iv) Constrained and Unconstrained Optimization
  - v) Gradient Descent and its variants

#### **4) Module 4- Machine Learning (60 Hours)**

- a) Introduction to Machine Learning (**3 hours**) – From Data Science to Machine Learning, Learning paradigm, Components of a machine learning algorithm, Bias-Variance tradeoff, Model selection, Hyperparameters, Regularization
- b) Supervised Learning (**15 Hours**) - Definition of supervised learning, data requirements, types of supervised learning problems, algorithms for regression and classification using supervised learning
- i) Linear Regression
  - ii) Logistic Regression
  - iii) Multiclass Classification
  - iv) KNN
  - v) Decision tree
  - vi) Random Forest
  - vii) Support Vector Machines
  - viii) Naïve Bayes

- c) Unsupervised Learning (**7 Hours**)- Definition of unsupervised learning, scenarios for unsupervised learning, types of data, examples, algorithms for unsupervised learning
  - (1) Agglomerative clustering
  - (2) K-Means
  - (3) Gaussian Mixture Models
  - (4) Introduction to Generative Models
  
- d) Deep Learning (**35 Hours**)- Artificial Neural networks and its evolution, Backpropagation algorithm, modern applications including text & speech analysis, computer vision and natural language processing using deep learning. State of the art Deep Learning techniques in –
  - i) Deep Neural Networks (DNN)
  - ii) Convolutional Neural Networks (CNN)
  - iii) Recurrent Neural Networks (RNN)
  - iv) Transformers
  - v) Generative Adversarial Networks (GANs)

**5) Module 5-Applications & Miscellaneous (Campus visit- 10 hours)**- The techniques discussed in the previous 4 modules will be applied to problems in various domains here. The students can visit the campus to learn in person from domain experts. **Practical applications will be selected from** (but not restricted to)

- a) Business Intelligence
- b) Business Analytics
- c) Sectorial Analytics (Marketing/ Finance/ Operations/ Supply Chain/ HRM)
- d) Computer Vision
- e) Language Modeling
- f) Applications in Engineering
- g) Healthcare
- h) Decision Making
- i) Project ( In candidates own time - 30 hours)**