



*Certificate of merit and participation will be given to successful participants.*

**Online Executive Development Program on  
DATA SCIENCE WITH R & PYTHON  
Module II: :Machine Learning with R Programming**  
organized by  
**Department of Statistics, Pondicherry University.**

**16 - 17 May, 2020 8.30 a.m to 7 p.m**  
**Inauguration: 16 May 2020 @ 8.30. a.m. &**  
**Valedictory : 17 May, 2020 @ 6.30. p.m.**

**Training Co-ordinator**



**Prof. Tirupathi Rao Padi,**  
Department of Statistics,  
Pondicherry University

**Chief Guest**



**Prof. S. Balakrishnan,**  
Director, Studies

**Keynote Speaker**



**Prof. V. Mariappan,**  
Dept of Banking Technology,  
SOMS, Pondicherry University

**Convener**



**Prof. P. Dhanavanthan,**  
Dean, RSMS,  
Pondicherry University

**Resource Persons**

**Ms. Rajya Lakshmi,**  
Data Scientist,  
STAR Consultancy Group-India

**Prof. Tirupathi Rao Padi,**  
Department of Statistics,  
Pondicherry University

**Mr. Srikanth CS,**  
Consultant from Farma  
Industry, STAR Consultancy  
Group-India

**Valedictory Address**



**Prof. S. Sudalai Muthu,**  
Dept of Banking Technology,  
SOMS, Pondicherry University

**For details:**

[http://www.pondiuni.edu.in/sites/default/files/Online\\_EDP-DS-STP%20\(Revised\)\\_Information%20Brochure\\_o.pdf](http://www.pondiuni.edu.in/sites/default/files/Online_EDP-DS-STP%20(Revised)_Information%20Brochure_o.pdf)

**For Queries: email id. : traopadipu@gmail.com | Whatsapp No. +91-9629862241**

**SCHEDULE OF MODULE-II:**

**MACHINE LEARNING WITH R PROGRAMMING**

<b>Day /Session/ Time</b>	<b>Course Contents</b>	<b>Details of the contents</b>
<b>Day-1 ( 2<sup>nd</sup>Saturday): 8.30 AM to 6.30 PM</b>		
<b>08.30 AM to 09.00 AM: Registration &amp; Inauguration</b>		
Day-I/ 08.30 AM - 09.00 AM	Registration & Inauguration	Declaration on Orientation of Workshop/ Skilled Training / Capacity Building by the coordinator and introduction of resource persons
Day-I/Session-1 09.00AM to 11.00AM	Basic Statistics	<ol style="list-style-type: none"> <li>1. Measures of Central Tendency</li> <li>2. Measures of Dispersion, Skewness&amp; Kurtosis</li> <li>3. Ordinal/Positional Measures</li> <li>4. Relative/Ratio Measures</li> <li>5. Frequency Tables</li> <li>6. Correlations and Regression</li> <li>7. Association Measures</li> </ol>
<b>11.00 AM to 11.15 AM: Morning Tea Break</b>		
Day-I/Session-2 11.15AM to 1.15PM	Data Science	<ol style="list-style-type: none"> <li>1. Introduction: What is Data Science? - Current landscape of perspectives - Skill sets needed</li> <li>2. Data Science hype – Why now?</li> <li>3. Supervised learning</li> <li>4. Unsupervised Learning</li> <li>5. Topics covered in Supervised learning</li> <li>6. Topics covered in unsupervised learning</li> </ol>
<b>01.15 PM. to 02.00PM: Lunch Break</b>		
Day-I/Session-3 02.00 PM to 04.00PM	Requirements of Good Machine learning System and ML in R – I	<ol style="list-style-type: none"> <li>1. Data preparations with R for ML</li> <li>2. Analysing data</li> <li>3. Patterns identification</li> <li>4. Prediction</li> <li>5. Conclusion</li> </ol>
<b>04.00 PM to 04.15PM: Evening Tea Break</b>		
Day-I/ Session-4 04.15PM to 05.45PM	ML in R – II	<ol style="list-style-type: none"> <li>1. Classification</li> <li>2. Regression</li> </ol>
Day-I/Session-5 5.45 PM to 6.30 PM	Interaction of Participants	<ol style="list-style-type: none"> <li>1. Discussions, Case studies, data sets of participants, Statistical Analysis Planning, etc.</li> </ol>

**Day-2 ( 2<sup>nd</sup>Sunday): 9.00 AM to 7.00 PM**

Day-II/Session-1 09.00AM to 11.00AM	ML in R – III Un-Supervised Learning - I	<ol style="list-style-type: none"><li>1. Concepts in Unsupervised learning</li><li>2. Unstructured data</li><li>3. Why Linear Regression?</li><li>4. Logistic regression</li><li>5. Clustering</li></ol>
<b>11.00 AM to 11.15 AM: Morning Tea Break</b>		
Day-II/Session-2 11.15AM to 01.15PM	ML in R – IV Un-Supervised Learning-II	<ol style="list-style-type: none"><li>1. Text Mining</li><li>2. Sentimental analysis</li><li>3. Targeted Marketing</li></ol>
<b>01.15 PM to 02.00PM: Lunch Break</b>		
Day-II/Session-3 02.00 PM to 04.00PM	Data Analytics with R-1	<ol style="list-style-type: none"><li>1. Test Dataset</li><li>2. Validation dataset</li><li>3. Training Dataset</li></ol>
<b>04.00 PM to 04.15PM: Evening Tea Break</b>		
Day-II/ Session-4 04.15 PM to 05.15PM	Data Analytics with R-2	<ol style="list-style-type: none"><li>1. Statistical model – Logistic Regression model with Business Problem</li><li>2. Validation of the model – Logistic Regression</li><li>3. Interpretation of the summary on Logistic Regression</li></ol>
Day-II/ Session-5 5.15PM to 6.00 PM	Interaction	<ol style="list-style-type: none"><li>1. Discussion on the Data sets of the participants, Statistical Data Analysis Planning</li></ol>
Day-II/ Session-6 6.00 PM to 6.30 PM	Examination & Screening Test	<ol style="list-style-type: none"><li>1. Online examination, evaluations, and declaration of results and grades</li></ol>
<b>06.30 PM to 7.00 PM: Valedictory, Issue of Certificates &amp; Feedback</b>		