INTERNATIONAL MANAGEMENT INSTITUTE BHUBANESWAR
Post Graduate Diploma in Management
Machine Learning Using R and Python (IS613)
PGDM 2020-22 Term III
CREDIT: 2 credits
SESSION DURATION: 60 Minutes

Instructor: Dr. Santanu Das and Dr. Manit Mishra
Telephone: 8658006959/8658988485
E-Mail: santanu@imibh.edu.in/ manit.mishra@imibh.edu.in

Course Introduction: This course aims to provide the students an advanced exposure to R and Python programming languages which have become very popular in the industry. It also aims to prepare the students to undertake advanced data analysis using machine learning techniques.

Pedagogy
The course will be delivered through a mix of lectures and real-life assignments involving thorough analysis and presentations. Exercises will be based on R and Python.

Learning Outcomes
- To provide firsthand exposure to R and Python programming (LO 1).
- To educate students about data handling (LO 2).
- To disseminate knowledge about machine learning techniques (LO 3).
- To strengthen the analytical foundation of students and prepare them for more advanced specialization courses latter (LO 4).

Evaluation Criteria:

Item	**Weightage (%)**	**Learning Outcome(s)**
Project (Group of 2 each)	20%	LO1 to LO 3
Quiz	20%	LO1 to LO 4
Report Writing	20%	LO1 to LO 4
End Term Examination	40%	LO1 to LO 4

Readings:
R for Everyone – Jared Lander, *Pearson Edu*. (2nd Ed)
Treading on Python – Volume 1 Foundations – Matt Harrison

Session No.	Topic	Learning Outcome	Readings
1-2	Basics of R Programming • Installation • Libraries • Basic operations • Vectors, matrix • List, Data frames, Arrays	LO1	
3	Dealing with data • Importing • Creating variables • Extracting variables from a dataset • Exporting data	LO1, LO2	
4-5	Data Visualization • Line, scatterplots, histograms • ggplot2	LO1, LO2	
6-7	Control statements • if else • For and while loops	LO1, LO2	
8-9	Data Manipulation • Apply family • Mutation • Aggregate • Pipe operator	LO1, LO2	
10-11	Data Analysis-I • GLM (Logit, Poisson)	LO3	
12-15	Data Analysis-II • Supervised learning (Random Forest) • Unsupervised learning (PCA, MDS, k-means clustering)	LO3	
16-17	Fundamentals of Python • Anaconda • Jupyter Notebook	LO1-LO2	Textbook
18-20	Data munging and analysis • Pandas	LO1-LO2	Textbook
21-22	PROJECT PRESENTATION		

Guidelines for project: Each student needs to mandatorily undertake a project and make a presentation at the end of the term. Since it is a group project, the contribution of each member needs to be clearly delineated although it is expected that each member should know the project thoroughly. The final report needs to be submitted before the presentation with the following components:

- Title page (format same as that of SIP report)
- Index

- Introduction about the topic (about 500 words)
- Data Source
- Methodology and model used
- Discussion of result
- Conclusion

Plagiarism:
We are committed to upholding the highest standards of academic integrity and honesty. Plagiarism is the use of or presentation of ideas, works that are not one's own and which are not common knowledge, without granting credit to the originator. You may refer the already available content just for your reference and to get the basic ideas. Only 20% of such content is acceptable, above that comes under the definition of Plagiarism which is unacceptable in IMI and will be treated seriously. All such cases will be referred to the appropriate body of the Institute for suitable disciplinary action.