**COURSE OUTLINE**

Predictive Analytics (Using R & Python)

Term – 5.01

PGDM (BA, AI &ML) 2019-21

|  |  |
| --- | --- |
| Faculty Name | Dr. Hemachandran K |
| Email ID |  |

**Brief Description and Relevance of the Course**

By utilizing predictive analytics we can solve real world problems and business problems. Who would like to build best predictive models to make predictions will find this course more useful. This course is aimed to all students and experts in Data Science, who would like to focus on the practical implementation of predictive analytics using R and Python.

Topic covers:- Different Stages in predictive analytics process and overview of the libraries in  
R and python programming, Problem understanding and Data preparation, Dataset Understanding, Predicting numerical values with machine learning algorithms, Predicting categories with machine learning algorithms, Develop machine learning models, Ensembles, Introduction to Neural Nets, Model evaluation, Model tuning and improving model performance & Case studies.

**Intended Learning Outcomes**

At the end of the course, students will be able to:

1. Understand the fundamentals of Predictive Analytics in R and Python Programming1

2. Analyze and visualize different kinds of data using R and Python

3. Work with different types of data and perform Data Manipulation and Preparation

4. Perform machine learning models on datasets using machine learning libraries

5. Build a model in R and Python

6. Tune and improve the model performance in R and Python

**Recommended and Reference Textbooks**

1. Eric Mayor (2015). Learning Predictive Analytics with R. PACKT Publishing

2. Ashish Kumar (2016). Learning Predictive Analytics with Python. PACKT Publishing

3. Thomas W. Miller (2014). Modelling Techniques in Predictive Analytics with Python and R: A Guide to Data Science. FT Press Analytics

4. Marketing Data Science (2015). Modelling Techniques in Predictive Analytics with R and Python

5. Alvaro Fuentes (2018).Hands-On Predictive Analytics with Python: Master the complete predictive analytics process, from problem definition to model deployment. PACKT Publishing

6. Joseph Babcock (2016).Mastering Predictive Analytics using python. PACT publishing

**Software Requirements for the Course**

* Microsoft Office
* MacOS 10.15 / Windows 10/ Linux VM
* Python – Jupyter Notebook – Anaconda Navigator (or)
* PyCharm – The Python IDE
* R Studio – Anaconda Navigator (or)
* R Studio IDE

**Session-Wise Topics and Reading/References**

**\*Note\* –**

**Readings**: Each session is preceded by a reading assignment. It is important to keep on top of the reading, which will be assumed during the lecture and discussion in class. You should set aside 2 hours to compete each reading. We do not expect you to fully understand everything before coming to class, but the goal is to prepare for class, familiarize yourself with new terminology and definitions, and to determine which part of the subject you want to hear more about. We encourage you to bring questions to class about material that is confusing. Other students might share your confusion.

In addition, you will be provided with hand-outs by the designated lecturer on a subject related to the session. You are expected to read the paper before class and collate your queries before arriving the lecture hall. The query should be thought provoking about the assigned paper.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sn.** | **Topic** | **Intended Learning Outcomes** | **Reading/Reference** |
| 1 | Introduction to predictive modelling, Applications, Python packages and R packages | 1. Understand the fundamentals of Predictive Analytics in R and Python Programming1 | 1. Chapter 1 - Eric Mayor (2015). Learning Predictive Analytics with R. PACKT Publishing.  2. Chapter 1 - Ashish Kumar (2016) Learning Predictive Analytics with Python. PACKT Publishing |
| 2-3 | Data visualization using lattice and Manipulate data using R | 1. Analyze and visualize different kinds of data using R and Python  2. Work with different types of data and perform Data Manipulation and Preparation | 1. Chapter 2 & 3 -  Eric Mayor (2015). Learning Predictive Analytics with R. PACKT Publishing |
| 4-5 | Visualization and Exploratory data analysis, Data Cleaning and Data Wrangling | 1. Analyze and visualize different kinds of data using R and Python  2. Work with different types of data and perform Data Manipulation and Preparation | 2. Chapter 2 & 3 - Ashish Kumar (2016) Learning Predictive Analytics with Python. PACKT Publishing |
| 6-7 | Statistical concepts for predictive modelling | 1. Work with different types of data and perform Data Manipulation and Preparation  performance in R and Python | 2. Chapter 4 - Ashish Kumar (2016) Learning Predictive Analytics with Python. PACKT Publishing |
| 8-9 | Predicting categories with machine learning algorithms: model evaluation, tuning and performance measures | 1. Perform machine learning models on datasets using machine learning libraries  2. Build a model in R and Python  3. Tune and improve the model performance in R and Python | 2. Chapter 8 - Ashish Kumar (2016) Learning Predictive Analytics with Python. PACKT Publishing  5. Chapter 5 - Alvaro Fuentes (2018).Hands-On Predictive Analytics with Python: Master the complete predictive analytics process, from problem definition to model deployment. PACKT Publishing |
| 10-11 | Develop machine learning models with the KNN, Naive Bayes and CART algorithms using Python’s scikit-learn | 1. Understand the fundamentals of Predictive Analytics in R and Python Programming1  2. Perform machine learning models on datasets using machine learning libraries  3. Build a model in R and Python | 3. Chapter 12 A - Thomas W. Miller (2014). Modelling Techniques in Predictive Analytics with Python and R: A Guide to Data Science. FT Press Analytics  4. Chapter 14A - Marketing Data Science (2015): Modelling Techniques in Predictive Analytics with R and Python |
| 12-13 | **Ensembles : - combine multiple algorithms and improve results** | 1. Perform machine learning models on datasets using machine learning libraries  2. Build a model in R and Python | 2. Chapter 5 & 6 - Ashish Kumar (2016) Learning Predictive Analytics with Python. PACKT Publishing  3. Chapter 12 A - Thomas W. Miller (2014). Modelling Techniques in Predictive Analytics with Python and R: A Guide to Data Science. FT Press Analytics  4. Chapter 14A - Marketing Data Science (2015): Modelling Techniques in Predictive Analytics with R and Python |
| 14-15 | Cluster Analysis and Dimensionality Reduction with Principal Component Analysis using R programming | 1. Perform simple regressions and classifications on datasets using machine learning libraries  2. Build a model in R and Python | 1.Chapter 4 & 5 -  Eric Mayor (2015). Learning Predictive Analytics with R. PACKT Publishing |
| 16 | Introduction to Neural Nets : Introducing neural network models Introducing Tensorflow and Keras | 1. Perform simple regressions and classifications on datasets using machine learning libraries  2. Build a model in R and Python | 6. Chapter-7 : Joseph Babcock (2016).Mastering Predictive Analytics using python. PACT publishing |
| 17-18 | Regression and Classification with neural networks and dark art of training neural network models | 1. Perform simple regressions and classifications on datasets using machine learning libraries  2. Build a model in R and Python | 6. Chapter-7 : Joseph Babcock (2016).Mastering Predictive Analytics using python. PACT publishing |
| 19-20 | Case Studies:  Predictive analysis- Strategy and Success of Swiggy | 1. Understand the fundamentals of Predictive Analytics in R and Python Programming1  2. Analyze and visualize different kinds of data using R and Python  3. Work with different types of data and perform Data Manipulation and Preparation  4. Perform simple regressions and classifications on datasets using machine learning libraries  5. Build a model in R and Python  6. Tune and improve the model performance in R and Python |  |

**Performance Evaluation Components for the Course**

|  |  |  |
| --- | --- | --- |
| **Session No.** | **Marks** | **Evaluation Form** |
| **Continuous Evaluation** | | |
| **2** | 10 | Assignment |
| **4 - 18** | 50 | Project: Abstract submission (Project related to Predictive Analysis) |
| Project Review 1 |
| Project Review 2 |
| Project Final Review |
| Publish a paper in Scopus indexed journal |
| **20** | 10 | Class Participation |
| **End Term Examination** | | |
| **After Course Completion** | 30 | Written Test |

**Assignment Schedule**

All assignments are to be presented in class along with the submission of a written report in soft copy format.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name of Assignment** | **Date of Submission** | **Take-home**  **or in-class** | **Individual-work or Team-work** | **Other Instructions** |
| Estimating the quality of wines with regression trees and model trees (Hint: Any two models) | As per Woxsen standards | Take-home | Individual work | Content will be shared during the course. (Based on Real world problems) |

Computational assignments will ask you to develop implementations of algorithms for search, game-solving, constraint satisfaction, knowledge representation and reasoning, and planning, to apply them to different real-world problems, and to analyze the performance. We expect that all code will run, be well-written and be commented appropriately; the course staff is always happy to help explain style and conventions. The written components ask you questions about the concepts and methods that you have learned and to reflect on the performance of your implementations.

Attendance & Punctuality

Learning is an interactive process. Woxsen students are admitted partly based on the experience they bring to the school and what they can add to class discussions. Therefore, attendance is an important aspect of studying here. Students are expected to be present in all the classes. Absence is only appropriate in cases of extreme personal illness, injury, or close family bereavement. Voluntary activities are never valid reasons for missing any class. The faculty, with the assistance of the Faculty Associate, shall keep track of students’ attendance and decide on the nature and extent of penalty for any absence from the class. Penalty may include reduction in grade or repetition of the course.

Late arrival is disruptive to the learning environment; students are expected to be in class before the scheduled commencement time. Students arriving for class after the scheduled commencement time should be turned away unless they have a valid reason to be permitted to attend.

Faculty should not consider attendance of sessions as a component of performance evaluation. The grading system at Woxsen accounts for this.

**External Websites Disclaimer**

Neither the instructor nor Woxsen School of Business is responsible for the content of external websites discussed in the classroom and/or linked to via online course materials, e-mail messages, message boards, or other means. Referred websites are for illustrative purposes only, and are neither warranted nor endorsed by the faculty or Woxsen School of Business. Web pages change frequently, as do ownership of domain names. While every effort is made to ensure proper referencing, it is possible that students may, on occasion, find materials to be objectionable for reasons beyond our control.

**Copyright**

The content provided by the faculty in the class is copy-righted. Students are instructed not to distribute or share content used during courses with external entities without the explicit written consent of the author and/or faculty.

**Student Code of Ethics**

Each student enrolled in this course accepts personal responsibility to uphold and defend academic integrity and to promote an atmosphere in which all individuals may flourish. The Students’ Code of Ethics strives to set a standard of honest behavior that reflects well on students and the school. All students enrolled in these courses are expected to follow the Students’ Code of Ethics, which they have been given at the time of enrolling for the program and pledged to adhere to. Unethical and unfair practices adopted by students may lead to penalties such as having to repeat the course or having the student’s enrollment cancelled.