

**MINOR SPECIALIZATION COURSE ON  
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Modules (Semester wise):**

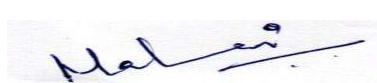
**Sem – 3: U18MAR0206 - PYTHON FOR AI AND ML**

**Sem – 4: U18MAR0207 - MATHEMATICS FOR AI AND ML**

**Sem – 5: U18MAR0208 – MACHINE LEARNING**

**Sem – 6: U18MAR0209 – DEEP LEARNING**

**Sem – 7: U18MAR0210 – INDUSTRIAL PROJECT FOR AI&ML**



**(Signature of Chairperson)**

<b>U18MAR0206</b>	<b>PYTHON FOR AI AND ML</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>2</b>	<b>0</b>	<b>2</b>	<b>0</b>	<b>3</b>

### **Course Outcomes**

**After successful completion of this course, the students should be able to**

CO1: Solving problems using the data structures in python programming

CO2: Handling data for data pre-processing

### **Objectives**

- Learn the Fundamentals of Python Programming
- Provide insights on Data Pre-processing

**Pre-requisite courses:** Nil

### **DIRECT**

1. Continuous Assessment Test I, II
2. Assignment
3. Assignment based on R Software
4. End Semester Examination

### **INDIRECT**

1. Course-end survey

### **Contents**

#### **Introduction**

Introduction-Python Interpreter-Interactive and script mode-Values and types, operators, expressions, statements, precedence of operators, Multiple assignments, comments.

#### **Control Statements and Functions**

Conditional (if), alternative (if-else), chained conditional (if-elif-else)-Iteration-while, for, break, continue, pass – Functions-Introduction, inbuilt functions, user defined functions, passing parameters, return values, recursion, Lambda functions

#### **Strings, Lists and Sets**

Strings-String slices, immutability, string methods and operations -Lists-creating lists, list operations, list methods, mutability, aliasing, cloning lists, list and strings, list and functions-list processing-list comprehension, Sets-creating sets, set operations.

#### **Tuples, dictionaries and Numpy**

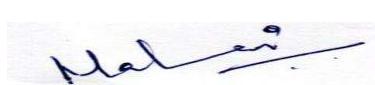
Tuples-Tuple assignment, Operations on Tuples, lists and tuples, Tuple as return value- Dictionaries-operations and methods, Nested Dictionaries. The NumPy ndarray: A Multidimensional Array Object – Universal Functions: Fast Element-wise Array Functions – Data Processing Using Arrays.

#### **Files and Pandas**

File Input and Output with Arrays – Reading and Writing Data from Files – Linear Algebra – Random Number Generation – Random Walks. Introduction to pandas Data Structures – Essential Functionality – Summarizing and Computing Descriptive Statistics – Handling Missing Data.

#### **Reference book**

Wes McKinney, "Python for Data Analysis", O'Reilly Media.2012.

  
 (Signature of Chairperson)

<b>U18MAR0207</b>	<b>MATHEMATICS FOR AI AND ML</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>2</b>	<b>0</b>	<b>2</b>	<b>0</b>	<b>3</b>

### **Course Outcomes**

#### **After successful completion of this course, the students should be able to**

CO1: Understand about data collection process and visualize the data graphically using bar chart and pie chart.

CO2: Analyze the data using inferential statistics which will form the basis for data analysis.

CO3: Analyze the data using predictive statistics: Linear and Non-linear Regression.

CO4: Predict the trend using time series analysis.

CO5: Analyze any data using the above concept using R.

### **Objectives**

- The primary goal of a data analyst is to manipulate and analyze the data.
- To draw inference from the data so that the derived knowledge can be used to make informed decisions.

### **Pre-requisite courses: Nil**

### **DIRECT**

1. Continuous Assessment Test I, II
2. Assignment
3. Assignment based on R Software
4. End Semester Examination

### **INDIRECT**

1. Course-end survey

### **Data Collection and Data Visualization**

**5 Hours**

Different types of scales: nominal, ordinal, interval and ratio - Collection of Primary data: concept of a questionnaire and a schedule, Secondary data - Types of data: Qualitative and quantitative data; Time series data and cross section data, discrete and continuous data- Classification - Tabulation & Diagrammatic representation using bar diagrams and pie chart - Univariate frequency distribution of discrete and continuous variables. Cumulative frequency distribution – Graphical representation of frequency distribution by Histogram, frequency polygon, Stem and leaf diagram -Box plot

### **Inferential Statistics**

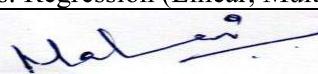
**7 Hours**

Develop an intuition how to understand the data, attributes, distributions - Procedure for statistical testing, etc. - Test of Hypothesis (Concept of Hypothesis testing, Null Hypothesis and Alternative Hypothesis) - Cross Tabulations (Contingency table and their use, Chi-Square test, Fisher's exact test), - One Sample T-test (Concept, Assumptions, Hypothesis, Verification of assumptions, Performing the test and interpretation of results) - Independent Samples T-test - Paired Samples T-test - One way ANOVA (Post hoc tests: Fisher's LSD, Tukey's HSD). - z-test and F-test

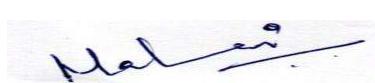
### **Predictive analysis - Linear Regression**

**7 Hours**

- Regression basics: Relationship between attributes using Covariance and Correlation - Relationship between multiple variables: Regression (Linear, Multivariate) in prediction -

  
(Signature of Chairperson)

Residual Analysis - Identifying significant features, feature reduction using AIC, multi-collinearity - Non-normality and Heteroscedasticity - Hypothesis testing of Regression Model - Confidence intervals of Slope - R-square and goodness of fit			
<b>Predictive analysis – Non-Linear Regression</b>	<b>6 Hours</b>		
Multiple Linear Regression - Polynomial Regression - Regularization methods - Lasso, Ridge and Elastic nets - Categorical Variables in Regression			
Non-Linear Regression - Logit function and interpretation - Types of error measures (ROCR) - Logistic Regression in classification			
<b>TIME SERIES</b>	<b>5 Hours</b>		
Forecasting models - Trend analysis - Cyclical and Seasonal analysis - Smoothing; Moving averages; Box-Jenkins, Holt-winters, Auto-correlation; ARIMA - Examples: Applications of Time Series in financial markets			
<b>STATISTICAL LAB USING R-PROGRAMMING</b>			
<b>List of Experiments</b>			
1. Introduction, Basic data representation. 2. Data presentation methods - Bar Chart, Pie Chart. 3. Importing data from MS-Excel. 4. Data manipulation 5. Mean, median, mode. 6. Standard deviation, five number summary, box plot. 7. Scatter diagram, correlation. 8. Regression. 9. ANOVA – One-way classification. 10. ANOVA – Two-way classification.			
<b>Theory: 30 Hours</b>	<b>Practical: 30 Hours</b>	<b>Tutorial: 0 Hours</b>	<b>Total : 60 Hours</b>
<b>TEXTBOOKS</b>			
1. Johnson R. A., Miller & Freund's, "Probability and Statistics for Engineers", 9 <sup>th</sup> Edition, PearsonEducation, Delhi, 2018. 2. Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, "An Introduction to StatisticalLearning with Applications in R", Springer, 2017.			
<b>REFERENCES</b>			
1. Walpole R. E., Myers S.L. & Keying Ye, "Probability and Statistics for Engineers and Scientists", Pearson Education Inc, 9 <sup>th</sup> Edition, 2012. 2. Charles Henry Brase and Corrinne Pellillo Brase, "Understandable Statistics: Concepts and Methods", Cengage Learning, 12 <sup>th</sup> Edition, 2018. 3. Richard A. Johnson and Dean W. Wichern, "Applied Multivariate Statistical Analysis", 6 <sup>th</sup> Edition, Pearson Education, Asia, 2012. 4. Anderson, T. W, "An Introduction to Multivariate Statistical Analysis", John Wiley and Sons, 2003 5. Gupta S.C and Kapoor V.K, "Fundamentals of Mathematical Statistics", 11 <sup>th</sup> extensively revised edition, Sultan Chand & Sons, 2007.			



(Signature of Chairperson)

<b>U18MAR0208</b>	<b>MACHINE LEARNING</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>2</b>	<b>0</b>	<b>2</b>	<b>0</b>	<b>3</b>

### **Course Outcomes**

**After successful completion of this course, the students should be able to**

- CO1: Explain the working of various machine learning algorithms.
- CO2: Apply the appropriate machine learning strategy for any given problem.
- CO3: Apply decision tree algorithm for classification.
- CO4: Apply Bayesian networks algorithm for classification.
- CO5: Apply various clustering algorithms for different datasets.

### **Objectives**

- To have a thorough understanding of the Supervised and Unsupervised learning techniques
- To study the various probability-based learning techniques
- To understand graphical models of machine learning algorithms

**Pre-requisite courses:** Nil

### **DIRECT**

1. Continuous Assessment Test I, II
2. Assignment
3. Assignment based on R Software
4. End Semester Examination

### **INDIRECT**

1. Course-end survey

### **Contents**

#### **Foundations for ML**

ML Techniques overview - Validation Techniques (Cross-Validations) - Feature Reduction/Dimensionality reduction - Principal components analysis (Eigen values, Eigen vectors, Orthogonality)

#### **Clustering**

Distance measures - Different clustering methods (Distance, Density, Hierarchical) - Iterative distance-based clustering; - Dealing with continuous, categorical values in K-Means - Constructing a hierarchical cluster - K-Medoids, k-Mode and density-based clustering - Measures of quality of clustering

#### **Classification**

Naïve Bayes Classifier - K-Nearest Neighbors - Support Vector Machines - Decision Trees - Ensembles methods: Bagging & boosting and its impact on bias and variance, Random forest, Gradient Boosting Machines and XGBoost

#### **Association Rule mining**

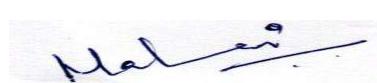
- The applications of Association Rule Mining: Market Basket, Recommendation Engines, etc. - A mathematical model for association analysis; Large item sets; Association Rules - Apriori: Constructs large item sets with mini sup by iterations; Interestingness of discovered association rules; - Application examples; Association analysis vs. classification - FP-trees

### **Reference book**

Jiawei Han, Micheline Kamber, Jian Pei, "Data Mining Concepts and Techniques", Morgan Kaufman Publications, Third Edition, 2011.

  
**(Signature of Chairperson)**

<b>U18MAR0209</b>	<b>DEEP LEARNING</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>
		<b>2</b>	<b>0</b>	<b>2</b>	<b>0</b>
<b>Course Outcomes</b>					
<b>After successful completion of this course, the students should be able to</b>					
CO1: Understand the fundamental principles, theory and approaches for learning with deep neural networks					
CO2: Apply neural networks in applications like - object detection, face recognition, neural style transfer					
CO3: Understand the variations of neural network for sequence data, apply RNN in applications like - Sentiment classification, Language translation, Speech Recognition and Trigger word detection.					
CO4: To learn the fundamentals of natural language processing					
<b>Objectives</b>					
<ol style="list-style-type: none"> <li>1. Learn the Fundamentals of Neural Networks</li> <li>2. Work with pattern recognition systems</li> <li>3. Introduce major deep learning algorithms, the problem settings, and their applications to solve real world problems</li> </ol>					
<b>Pre-requisite courses: U22MAR0208/ U22MAR0208</b>					
<b>DIRECT</b>					
<ol style="list-style-type: none"> <li>1. Continuous Assessment Test I, II</li> <li>2. Assignment</li> <li>3. Assignment based on R Software</li> <li>4. End Semester Examination</li> </ol>					
<b>INDIRECT</b>					
1. Course-end survey					
<b>Contents</b>					
<b>Introduction</b>					
Introduction to Deep Learning – Neural Network Basics – Convolutional Neural Networks – Recurrent Neural Networks - Introduction to Tensor Flow - Image Processing					
<b>Convolutional Neural Networks</b>					
Perceptron – Working of CNN – Computer Vision, Edge Detection, Convolutions, Padding and Strided Convolutions Pooling Layers, Fully Connected Layer, CNN Case Studies: LeNet5, AlexNet					
<b>Recurrent Neural Networks</b>					
Working of RNN – Applications of RNN- Loss Function – Back propagation – LSTM – Bidirectional RNN - Deep Neural Networks – Building Deep Neural Networks for face recognition					
<b>Natural Language Processing</b>					
NLP tasks in syntax, semantics, and pragmatics. Applications such as information extraction, question answering, and machine translation					
<b>Reference book</b>					
<ol style="list-style-type: none"> <li>1. Ian Good fellow, Yoshua Bengio and Aaron Courville, “Deep Learning”, The MIT Press.</li> <li>2. James Allen, “Natural Language Processing with Python”, O’Reilly Media, July 2009.</li> </ol>					



(Signature of Chairperson)

<b>U18MAR0210</b>	<b>INDUSTRIAL PROJECT FOR AI&amp;ML</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>2</b>	<b>0</b>	<b>0</b>	<b>2</b>	<b>3</b>
<b>Course Outcomes</b>						
To provide sufficient hands-on learning experience related to the design, development and analysis of suitable product / process so as to enhance the technical skill sets in the chosen field.						
<b>Objectives</b>						
<ul style="list-style-type: none"> <li>• Creating a chat bot</li> <li>• To design an innovative application using NLP components</li> </ul>						
<b>Pre-requisite courses:</b> U18MAR0003						
<b>DIRECT</b>						
<ol style="list-style-type: none"> <li>1. Continuous Assessment Test I, II</li> <li>2. Assignment</li> <li>3. Assignment based on R and python Software</li> <li>4. End Semester Examination</li> </ol>						
<b>INDIRECT</b>						
1. Course-end survey						
<b>Contents</b>						
<p>This course intends to develop programming and presentation skills. Each student can select an area for their project in consultation with the faculty. It involves programming, implementation, testing and performance analysis in different application specific contexts. Students will be required to make an in-class presentation, project demonstration and a project report. The course will be evaluated by a panel of (at least) two faculty members.</p>						



(Signature of Chairperson)

<b>U18MAC0301</b>	<b>MATHEMATICAL MODELLING: ANALYSIS AND APPLICATIONS</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>

### Course Outcomes

**After successful completion of this course, the students should be able to**

CO1: Provides introduction of mathematical modeling and analysis in biological sciences.

CO2: Covers the fundamentals of deterministic models in both discrete and continuous time domain.

CO3: Understand the principles mathematical modelling of infectious diseases.

### Pre-requisite courses: Calculus

#### CO/PO Mapping:

(S/M/W indicates strength of correlation) S-Strong, M-Medium, W-Weak

COs	PROGRAMME OUTCOMES													
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	S	S												
CO2	S	M										M		
CO3	S											M		

### DIRECT

1. Mid-Semester Examination
2. End Semester Examination

### INDIRECT

1. Course-end survey

### MATHEMATICAL MODELLING

**4+1 Hours**

Overview of mathematical modeling and types of mathematical models, Introduction to population dynamics, solution methods of linear difference equations and discrete time model-Introduction to Mathematical Modeling-Discrete time linear models in Population dynamics-I-Discrete time linear models in Population dynamics-II-Discrete time linear age structured models

### SIR MODEL-I

**4+1 Hours**

Modelling the Basics- mathematical modelling of infectious diseases- some of the basic concepts in building compartmental models, including to interpret and represent rates, durations and proportions.

Anatomy of an Epidemic- a simple epidemic of a perfectly immunising infection in a stable population- basic concepts of infectious disease epidemiology, including the basic reproduction number ( $R_0$ ), and its implications for infectious disease dynamics.

*Malini*

<b>SIR MODEL -II</b>	<b>4+1 Hours</b>	
Combining Modelling and Insights -different scenarios for reproduction numbers.		
Dynamics of Susceptibles- Susceptibility to infection is the fuel for an infectious disease; understanding the dynamics of susceptibility can offer important insights into epidemic dynamics, as well as priorities for control- three important mechanisms by which susceptibility can change over the course of an epidemic: (i) population turnover, (ii) vaccination, (iii) immunity waning over time-modelling vaccination.		
<b>Theory: 15 Hours</b>	<b>Tutorial: 0 Hours</b>	<b>Total : 15 Hours</b>
<b>REFERENCES</b>		
<ol style="list-style-type: none"> <li>1. James D. Murray, " Mathematical Biology ", Springer New York, 2013.</li> <li>2. Nicholas F. Britton, "Essential Mathematical Biology", Springer London, 2003.</li> <li>3. Michael D Alder., "An Introduction to Mathematical Modelling", HeavenForBooks, 2001.</li> <li>4. <a href="https://www.coursera.org/learn/ordinary-differential-equations#syllabus">https://www.coursera.org/learn/ordinary-differential-equations#syllabus</a></li> <li>5. <a href="https://www.coursera.org/learn/developing-the-sir-model/home/info">https://www.coursera.org/learn/developing-the-sir-model/home/info</a></li> <li>6. <a href="https://nptel.ac.in/courses/111/107/111107113/">https://nptel.ac.in/courses/111/107/111107113/</a></li> </ol>		



<b>U18MAC0302</b>	<b>Resource Management Techniques using TORA</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>0</b>	<b>0</b>	<b>2</b>	<b>0</b>	<b>1</b>

### **Course Outcomes**

**After successful completion of this course, the students should be able to**

CO1: Formulate linear programming problems mathematically and solve them using TORA software

CO2: Solve transportation and assignment model problems using TORA software

CO3: Analyse queueing models using TORA software

CO4: Use TORA to solve problems in game theory.

### **Pre-requisite courses: Nil**

<b>CO/PO Mapping</b>												
COs	Programme Outcomes(POs)											
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1		M			S							
CO2	W				M							
CO3		W			M							
CO4	W				M							

### **DIRECT**

1. Mid-Semester Examination
2. End Semester Examination

### **INDIRECT**

1. Course-end survey

### **Linear Programming** **4+1 Hours**

Linear Programming Problem – Formulation - Application – Integer programming problem - Solution of LPP and IPP using TORA software-Sensitivity analysis for LPP.

### **Transportation Model and Assignment Model** **4+1 Hours**

Transportation Model- Assignment model – Balanced, unbalanced and restricted problems– Determining optimal solution using TORA software

### **Queueing Theory and Game Theory** **4+1 Hours**

**Queueing Theory** - Characteristics of a queueing system - Markovian Queues – Single and Multi-server Models – Little's formula – Machine Interference Model – Self Service Queue – Determining characteristics of queues using TORA software.

**Game Theory** – Two person zero sum game – Payoff matrix - Saddle point – Pure and Mixed strategies - Solving problems using TORA software.

*Malini*

<b>Theory: 15 Hours</b>	<b>Tutorial: 0 Hours</b>	<b>Total : 15 Hours</b>
<b>REFERENCES</b>		
<ol style="list-style-type: none"> <li>1. Jaisankar S., "Operations Research – Decision Models Approach", Excel Publications, New Delhi, 2010</li> <li>2. Taha H.A., "Operations Research: An introduction", 10<sup>th</sup> edition, Pearson Education , 2017</li> <li>3. Kanti Swarup , P. K. Gupta , Man Mohan , "Operations Research", Sultan Chand and Sons, 2010</li> </ol>		



<b>U18MAC0303</b>	<b>MATHEMATICS FOR CYBER SECURITY</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>

### Course Outcomes

**After successful completion of this course, the students should be able to**

CO1: Understand the concepts and results of Number Theory used in coding theory and cryptography.

CO2: Understand basic concepts of various algebraic structures and theorems like Euler's theorem for designing security algorithm.

CO3: Understand coding theory which will be useful for data compression, information hiding.

**Pre-requisite courses:** Nil

CO/PO Mapping												
(S/M/W indicates strength of correlation) S-Strong, M-Medium, W-Weak												
COs	Programme Outcomes(POs)											
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S										
CO2	S	M										
CO3	S											

### DIRECT

1. Mid-Semester Examination
2. End Semester Examination

### INDIRECT

1. Course-end survey

<b>NUMBER THEORY</b>	<b>4+1 Hours</b>
----------------------	------------------

Introduction - Divisibility - Greatest common divisor - Prime numbers - Fundamental theorem of arithmetic - Euclidean algorithm - Number systems in different bases (in particular binary and hexadecimal).

<b>CONGRUENCES AND CLASSICAL THEOREMS</b>	<b>4+1 Hours</b>
---	------------------

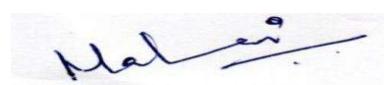
Congruences: Definition - Basic properties of congruences - Residue classes - Integers modulo n, congruence classes - Fermat's little theorem - Euler totient function - Euler's theorem - Chinese remainder theorem

<b>CODING THEORY AND CRYPTOGRAPHY</b>	<b>4+1 Hours</b>
---------------------------------------	------------------

Introduction - Basic concepts: codes, minimum distance - Group codes - Generator matrices and parity check matrices - Error detection and Correction - Symmetric Key Cryptography - Public-Key Cryptography - Basic examples - RSA algorithm

*Malini*

<b>Theory: 15 Hours</b>	<b>Tutorial: 0 Hours</b>	<b>Total: 15 Hours</b>
<b>REFERENCES</b>		
<ol style="list-style-type: none"> <li>1. D. S. Malik, J. Mordeson, M. K. Sen, Fundamentals of abstract algebra, Tata McGraw Hill.</li> <li>2. I. Niven, H.S. Zuckerman and H. L. Montgomery, An introduction to the theory of numbers, John Wiley and Sons, 2004.</li> <li>3. Douglas Stinson, 'Cryptography –Theory and Practice', CRC Press, 2006.</li> <li>4. Joseph A. Gallian, "Contemporary Abstract Algebra", Narosa, 1998</li> <li>5. <a href="https://www.coursera.org/learn/mathematical-foundations-cryptography">https://www.coursera.org/learn/mathematical-foundations-cryptography</a></li> <li>6. <a href="http://homes.soic.indiana.edu/yh33/Teaching/I231-2016/syllabus.html">http://homes.soic.indiana.edu/yh33/Teaching/I231-2016/syllabus.html</a></li> <li>7. <a href="https://nptel.ac.in/courses/106/103/106103015/">https://nptel.ac.in/courses/106/103/106103015/</a></li> <li>8. <a href="https://nptel.ac.in/courses/111/103/111103020/">https://nptel.ac.in/courses/111/103/111103020/</a></li> <li>9. <a href="https://nptel.ac.in/courses/111/101/111101137/">https://nptel.ac.in/courses/111/101/111101137/</a></li> </ol>		



<b>U18MAC0304</b>	<b>R PROGRAMMING FOR ANALYTICS</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>

### **Course Outcomes**

**After successful completion of this course, the students should be able to**

CO1: Understanding basics of R programming and its algorithm.

CO2: Understand and explore the concept of data.

CO3: Perform hypothesis testing and interpret the results.

CO4: Utilize R to apply appropriate techniques to solve business problems.

CO5: Design and carryout a Business analytics project on the provided live data set.

### **Pre-requisite courses: NIL**

<b>CO/PO Mapping</b>												
(S/M/W indicates strength of correlation)      S-Strong, M-Medium, W-Weak												
<b>COs</b>	<b>Programme Outcomes(POs)</b>											
	<b>PO1</b>	<b>PO2</b>	<b>PO3</b>	<b>PO4</b>	<b>PO5</b>	<b>PO6</b>	<b>PO7</b>	<b>PO8</b>	<b>PO9</b>	<b>PO10</b>	<b>PO11</b>	<b>PO12</b>
<b>CO1</b>												
<b>CO2</b>												
<b>CO3</b>												
<b>CO4</b>												
<b>CO5</b>												

### **DIRECT**

1. Mid-Semester Examination
2. End Semester Examination

### **INDIRECT**

1. Course-end survey

### **INTRODUCTION TO R AND DATA STRUCTURES** 2+3 Hours

**Introduction to R** - Understanding the tool user interface, Getting help on tool, Installing Packages

**Understanding Data Structures** - Data Types, Importing Data (CSV, Excel, Fixed Width Formats), Data Manipulation, Combining Data Sets, Sub-setting Data, Data Sorting, Data Aggregations, Re-labelling the Columns

**Exporting Data** - Exploratory Data Analysis, Programming Structures, Charts and Graph.

### **INTRODUCTION TO DATA** 2+3 Hours

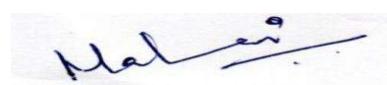
Univariate Descriptive Statistics: Graphs and distribution shapes, Measures of centre and spread, The Normal distribution, Z-scores, Bivariate Distributions: The scatterplot, Correlation, Bivariate Distributions (Categorical Data): Contingency tables, Conditional probability, Examining independence – Overview.

### **HYPOTHESIS TESTING** 2+3 Hours

Errors in testing, Alpha and critical values - Single sample test, Independent t-test and Dependent t-test, Hypothesis Testing (Categorical Data) :The chi-square test Goodness-of-Fit, Test-of-Independence, Hypothesis Testing (More Than Two Group Means) :The ANOVA, One-way ANOVA Two-way ANOVA, Hypothesis Testing (Quantitative data) :Correlation, Simple (single variable) regression, Multiple regression - - Overview

*Malini*

<b>Theory: 6 Hours</b>	<b>Practical: 9 Hours</b>	<b>Total: 15 Hours</b>
<b>REFERENCES</b>		
<ol style="list-style-type: none"><li>1. R for Beginners - Emmanuel Paradis (<a href="https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf">https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf</a>)</li><li>2. Rtips. Revival 2014! Paul E. Johnson March 24, 2014 (<a href="http://pj.freefaculty.org/R/Rtips.pdf">http://pj.freefaculty.org/R/Rtips.pdf</a>)</li></ol>		

A handwritten signature in blue ink, appearing to read "Malin".

<b>U18MAC0305</b>	<b>DATA ANALYSIS AND FORECASTING USING PYTHON</b>	<b>L</b>	<b>T</b>	<b>P</b>	<b>J</b>	<b>C</b>
		<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>

### Course Outcomes

#### After successful completion of this course, the students should be able to

CO1: Understanding basics of python for performing data analysis.

CO2: Understanding the data, performing preprocessing, processing and data visualization to get insights from data.

CO3: Use different python packages for mathematical, scientific applications and for web data analysis.

CO4: Develop the model for data analysis.

CO5: Evaluate the model performance for data analysis.

### Pre-requisite courses: NIL

CO/PO Mapping												
COs	Programme Outcomes(POs)											
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	S	S										
CO2	S	M										
CO3	S											M
CO4	S	M										M
CO5	S	M										M

### DIRECT

1. Mid-Semester Examination
2. End Semester Examination

### INDIRECT

1. Course-end survey

### INTRODUCTION

2+3 Hours

Data Science – Machine Learning – Supervised and Unsupervised Learning –Introduction to Neural Network – Artificial Neural Networks – Perceptron – Bias - Activation Function – Cost Function – Gradient Descent Optimization – Tensorflow and Keras.

**Application:** MNIST Handwritten Digits Image Classification with Tensorflow and Keras using Python.

### MATHEMATICAL AND SCIENTIFIC APPLICATIONS FOR DATA ANALYSIS

2+3 Hours

Python for Data Analysis – Numpy & Pandas - Understanding and creating N-dimensional arrays, Basic indexing and slicing, Boolean indexing, Fancy indexing, Universal functions - Data processing using arrays, File input and output with arrays.

**Application:** Preprocessing of Data using Python

### DATA FORMATTING AND VISUALIZATION

2+3 Hours

*Malini*

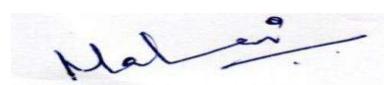
Data Formatting, Exploratory Data Analysis, Filtering and hierarchical indexing using Pandas - Data Visualization: Basic Visualization Tools, Specialized Visualization Tools, Seaborn Creating and Plotting Maps.

**Application:** Data Prediction, Outcomes, Forecasting, Analysis & Visualization on World Map using Python

<b>Theory: 6 Hours</b>	<b>Practical: 9 Hours</b>	<b>Total: 15 Hours</b>
------------------------	---------------------------	------------------------

## REFERENCES

1. David Ascher and Mark Lutz, Learning Python, Publisher O'Reilly Media, 2<sup>nd</sup> Edition, 2003.
2. Reema Thareja, "Python Programming using Problem Solving approach", Oxford University press, 2017.
3. Wes Mckinney "Python for Data Analysis", First edition, Publisher O'Reilly Media, 2012.
4. Allen Downey, Jeffrey Elkner, Chris Meyers,: Learning with Python, Dreamtech Press, 2015.
5. David Taieb , "Data Analysis with Python: A Modern Approach " 1st Edition, Packt Publishing, 5<sup>th</sup> Edition, 2018.
6. <https://www.coursera.org/learn/data-analysis-with-python#syllabus>
7. <https://nptel.ac.in/courses/106/107/106107220/>

A handwritten signature in blue ink, appearing to read "Malini".