

Programme	BSc Statistics				
Course Code	STA3MN203 (P)				
Course Title	Random variables and CART				
Type of Course	Minor				
Semester	III				
Academic Level	200 - 299				
Course Details	Credit	Lecture per week	Tutorial per week	Practical per week	Total Hours
	4	3	-	2	75
Pre-requisites	Basic knowledge of set theory and probability theory				
Course Summary	This course offers a comprehensive understanding of random variables, distributions, and statistical learning methods like classification and regression trees, bagging, random forest, with hands-on experience in R				

### Course Outcomes (CO):

CO	CO Statement	Cognitive Level*	Knowledge Category#	Evaluation Tools used
CO1	Grasp random variables, distributions.	U	C	Instructor-created exams / Quiz
CO2	Summarize discrete, continuous distributions and analyze data to help entrepreneurial decisions using critical thinking skills.	Ap	P	Practical Assignment / Observation of Practical Skills
CO3	Calculate probabilities and statistical parameters for various standard distributions.	Ap	P	Seminar Presentation / Group Tutorial Work
CO4	Introduce to statistical learning, variables, models and critically evaluate ethical implications of statistical methods aligning with human values.	U	C	Instructor-created exams / Home Assignments
CO5	Understand bagging, random forest.	Ap	P	One Minute Reflection Writing assignments/ Instructor-created exams
CO6	Implement classification, regression trees in R.	Ap	P	Viva Voce/ Instructor-created exams

\* - Remember (R), Understand (U), Apply (Ap), Analyse (An), Evaluate (E), Create (C)  
# - Factual Knowledge(F) Conceptual Knowledge (C) Procedural Knowledge (P) Metacognitive Knowledge (M)

### Detailed Syllabus:

Module	Unit	Content	Hrs (45+ 30)	Marks (70)
<b>I</b>	<b>Random variables</b>		<b>14</b>	<b>20</b>
	1	Random variable	2	
	2	Distribution function	2	
	3	Discrete random variable	2	
	4	Probability mass function	2	
	5	Discrete distribution function	2	
	6	Continuous random variable	2	
	7	Probability density function	2	
	Unit 1: 5.1 Ref[2] Unit 2: 5.2 Ref[2] Unit 3: 5.3 Ref[2] Unit 4: 5.3.1 Ref[2] Unit 5: 5.3.2 Ref[2] Unit 6: 5.4 Ref[2] Unit 7: 5.4.1 Ref[2]			
<b>II</b>	<b>Standard distributions</b>		<b>15</b>	<b>20</b>
	8	Bernoulli distribution	2	
	9	Binomial distribution	4	
	10	Poisson distribution	4	
	11	Normal distribution	4	
	12	Importance of Normal distribution	1	
	Unit 8: 7.1 Ref[2] Unit 9: 7.2 Ref[2] Unit 10: 7.3 Ref[2] Unit 11: 8.2 Ref[2] Unit 12: 8.2.13 Ref[2]			
<b>III</b>	<b>Statistical learning</b>		<b>10</b>	<b>20</b>
	13	An introduction to Statistical learning	1	
	14	Input and output variables	1	
	15	Response and predictor variables	1	
	16	Supervised and unsupervised learning	1	
	17	Classification versus regression	1	
	18	Classification and regression trees (CART)	2	
	19	Trees versus linear models	2	
	20	Advantages and disadvantages of trees	1	
	Unit 13: 2.1 Ref[1] Unit 14: 2.1 Ref[1] Unit 15: 2.1 Ref[1] Unit 16: 2.1.4 Ref[1] Unit 17: 2.1.5 Ref[1] Unit 18: 8.1.1, 8.1.2 Ref[1] Unit 19: 8.1.3 Ref[1] Unit 20: 8.1.4 Ref[1]			
<b>IV</b>	<b>Bagging</b>		<b>6</b>	<b>10</b>
	21	An introduction to Bagging	3	

	22	Random forest	3	
	Unit 21: 8.2.1 Ref[1] Unit 22: 8.2.2 Ref[1]			
<b>V</b>	<b>PRACTICUM</b>		<b>30</b>	
	Do practice problems in R software from any 5 units of the given list and one additional problem decided by the teacher-in-charge, related to the content of the course. Other units listed here may be used as demonstrations of the concepts taught in the course.			
	1	Fitting classification trees		
	2	Pruning trees		
	3	Use the function–lm.fit		
	4	Use the function–names()		
	5	Use the function–predict()		
	6	Plotting of least square regression line-abline()		
	7	Plotting of least square regression line- plot()		
	8	residuals() function		
	Sections from References: Unit 1: 8.3.1 Ref[1] Unit 2: 8.3.1 Ref[1] Unit 3: 3.6.2 Ref[1] Unit 4: 3.6.2 Ref[1] Unit 5: 3.6.2 Ref[1] Unit 6: 3.6.2 Ref[1] Unit 7: 3.6.2 Ref[1] Unit 8: 3.6.2 Ref[1]			
Books and References:				
1. G. James, D. Witten, T. Hastie, and R. Tibshirani. (2013), An Introduction to Statistical Learning: with Applications in R, Springer.				
2. Gupta, S.C. and Kapoor, V.K. (1997) Fundamentals of Mathematical Statistics. Sultan Chand and Sons, New Delhi				

**Mapping of COs with PSOs and POs :**

	PSO 1	PSO 2	PSO 3	PSO4	PSO 5	PSO6	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	2	2	-	-	-	2	2	1	1	-	-	-
CO 2	2	2	-	3	2	2	2	1	2	-	3	-
CO 3	3	-	-	-	-	3	3	2	-	3	-	-
CO 4	2	-	-	-	-	2	2	3	1	-	-	3
CO 5	2	-	2	-	2	-	2	-	2	-	-	-
CO 6	-	-	-	2	-	2	1	2	-	3	-	-

**Correlation Levels:**

Level	Correlation
-	Nil
1	Slightly / Low
2	Moderate / Medium
3	Substantial / High

**Assessment Rubrics:**

- Quiz / Assignment/ Quiz/ Discussion / Seminar
- Midterm Exam
- Programming Assignments (20%)
- Final Exam (70%)

**Mapping of COs to Assessment Rubrics :**

	Internal Exam	Assignment	Project Evaluation	End Semester Examinations
CO 1	✓	✓		✓
CO 2	✓	✓		✓
CO 3	✓			✓
CO 4	✓	✓		✓
CO 5		✓		✓
CO 6	✓			