GANPAT UNIVERSITY									
FACULTY OF ENGINEERING AND TECHNOLOGY									
Programme	Bachelor of Technology					Branch/Spec.	Computer Science & Engineering (BDA)		
Semester	VI					Version	1.0.0.0		
Effective from Academic Year			2021-22			Effective for the batch Admitted in			June 2019
Subject code	2CSE608		Subject Name			PREDICTIVE MODELING			
Teaching scheme						Examination scheme (Marks)			
(Per week)	Lecture(DT)		Practical(Lab.)		Total		CE	SEE	Total
	L	TU	P	TW					
Credit	3	0	2	0	5	Theory	40	60	100
Hours	3	0	4	0	7	Practical	60	40	100

Pre-requisites:

Data Structures, Mathematics , Probability and Statistics

Objectives of the Course:

Upon Completion of the course, the students will be able to

- Apply predictive modelling using various techniques using SPSS Modeler
- Design and analyze appropriate predictive models
- Build statistical models for analysis
- Demonstrate data mining process life cycle

Practicals are defined based on the following topics:

Unit	Content	Hrs.
1	**Introduction to Predictive Analytics** Introduction to Predictive Analytics and its use cases. CRISP – DM methodology and the skills required for successfully implementing Predictive Analytics / Machine Learning Use Cases	4
2	**Introduction to IBM SPSS Modeler** SPSS Modeler interface, and the terminologies such as streams,nodes, palettes	4
3	**Collecting, Understanding and Analysis of Data** Data Understanding stage: Collecting Initial Data and Describing Data, exploring the data and assessing the quality of data.	6
4	**Integrate Data** Integrating datasets which are typically stored in different tables / databases.	3
5	**Identifying Relationships and Modeling** Modeling techniques and algorithms in Predictive Analytics	4
6	**Using Functions in IBM SPSS Modeler** Inbuilt functions in IBM SPSS Modeler.	4
7	**Field Transformations: Derive, Binning, Reclassify** Three nodes to cleanse and enrich data: Derive, Binning, Reclassify	4
8	**Additional Field Transformations: Filler, Transform** additional nodes for Data Preparation: Filler and Transform	8

9	**Sequence Data, Sampling, Balancing and Partitioning Data** concept of Sequence Data, and how it can be handled, Sampling, Balancing and Partitioning Data	**8**

Practical content
Practicals will be based on various algorithms implementation using SPSS

Text Books:	
1.	Kattamuri S. Sarma, "Predictive Modeling with SAS Enterprise Miner: Practical Solutions for Business Applications", SAS Publishing.

Reference Books:	
1	Alex Guazzelli, Wen-Ching Lin, Tridivesh Jena, James Taylor, "PMML in Action Unleashing the Power of Open Standards for Data Mining and Predictive Analytics, Create Space Independent Publishing Platform.
2	Ian H. Witten, EibeFrank , "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann.
3	Eric Siegel , "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die", Wiley.
4	Conrad Carlberg, "Predictive Analytics: Microsoft Excel", Que Publishing.

Course Outcomes:	
COs	Description
CO1	Apply predictive modelling using various techniques using SPSS Modeler
CO2	Design and analyze appropriate predictive models
CO3	Build statistical models for analysis
CO4	Demonstrate data mining process life cycle

Mapping of CO and PO:

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	2	2	2	0	0	0	1	1	1	1	0
CO2	3	3	3	3	3	2	2	2	1	2	2	1
CO3	3	1	3	3	2	1	1	3	1	2	2	2
CO4	3	3	3	2	2	2	1	1	1	0	0	0