22CSPE404	PREDICTIVE DATA ANALYTICS		SEMESTER VII		
PREREQUISTIES		**CATEGORY**	**PE**	**Credit**	**3**
NIL		**Hours\Week**	**L**	**T**	**P**	**TH**
			3	**0**	**0**	**3**

Course Objectives:

1.	Develop and use various quantitative and classification predictive models based on various regression and decision tree methods
2.	Understanding of how to formulate predictive analytics questions
3.	Learn how to search, identify, gather and pre-process data for the analysis

UNIT I	**OVERVIEW OF PREDICTIVE ANALYTICS**	**9**	**0**	**0**	**9**

Predictive Analytics – Predictive Analytics vs. Business Intelligence – Predictive Analytics vs. Statistics – Predictive Analytics vs. Data Mining – Challenges in Using Predictive Analytics. Setting up the Predictive Modeling project: Predictive Analytics Processing Steps: CRISP-DM – Defining Data for Predictive Modeling – Defining the Target Variable – Defining Measures of Success for Predictive Models.

UNIT II	**DATA UNDERSTANDING**	**9**	**0**	**0**	**9**

Single Variable Summaries – Data Visualization in One Dimension – Histograms – Multiple Variable Summaries – Data Visualization, Two or Higher Dimensions. Data Preparation: Variable Cleaning – Feature Creation.

UNIT III	**DESCRIPTIVE MODELING**	**9**	**0**	**0**	**9**

Data Preparation Issues with Descriptive Modeling – Principal Component Analysis – Clustering Algorithms. Interpreting Descriptive Models: Standard Cluster Model Interpretation.

UNIT IV	**PREDICTIVE MODELING**	**9**	**0**	**0**	**9**

Decision Trees – Logistic Regression – Neural Networks – K-Nearest Neighbor – Naive Bayes – Linear Regression – Other Regression Algorithms. Assessing Predictive Models: Batch Approach to Model Assessment.

UNIT V	**MODEL ENSEMBLES**	**9**	**0**	**0**	**9**

Motivation for Ensembles – Bagging – Boosting – Improvements to Bagging and Boosting – Interpreting Model Ensembles. Model Deployment: General Deployment Considerations – Case Study.

Total (45 L)= 45 Periods

Text Book:	
1.	Dean Abbott, "Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst, JohnWiley & Sons, Inc., 2014
Reference Books:	
1.	John D.Kelleher, Brain Mac Namee, Aoife D"Arcy, "Fundamentals of Machine Learning for Predictive Data Analytics", MIT Press,2015
2.	Gopal M, "Applied Machine Learning", McGraw Hill Education, 2018
E-Reference:	
1	https://archive.nptel.ac.in/courses/111/106/111106164/

COURSE OUTCOMES: Upon completion of this course, the students will be able to:		Bloom's Taxonomy Mapped
CO1	Explore the processing steps of predictive analysis for solving real time problems	L3
CO2	Make use of data for modeling project	L3
CO3	Utilize various descriptive modeling algorithms	L3
CO4	Implement different types of predictive modeling algorithms	L3
CO5	Apply predictive analytics concepts to real world applications	L3

COURSE ARTICULATION MATRIX

COs/ POs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	1	1	1										3	1
CO2	1	1	1										3	1
CO3	1	1	1										3	1
CO4	1	1	1										3	1
CO5	1	1	1										3	1
Avg	1	1	1										3	1
3/2/1-indicates strength of correlation (3- High, 2-Medium, 1- Low)														