Course Code	Course Title			
116h54C501	**Data Analytics**			
	TH	**P**	**TUT**	**Total**
Teaching Scheme(Hrs.)	**03**	**--**	**--**	**03**
Credits Assigned	**03**	**--**	**--**	**02**

Examination Scheme	Marks							
	CA		**ESE**	**TW**	**O**	**P**	**P&O**	**Total**
	ISE	**IA**						
	30	**20**	**50**	**--**	**--**	**--**	**--**	**100**

Course prerequisites (if any):
- Concepts of DBMS
- Data Mining
- Knowledge of basic Machine Learning algorithms and programming language (C/C++/Java/ Python).

Course Objectives

Introduction to the fundamental concepts of Data Analytics through real world case studies

Comprehension of data analytics and visualization for decision-making in Geospatial, healthcare and text mining business applications

Course Outcomes

At the end of successful completion of the course the student will be able to

CO1	Understand basic concepts of data analytics to solve real-world problems
CO2	Apply the data analytics in the field of geospatial system
CO3	Preform the social data analytics
CO4	Preform Time series Analytics and forecasting
CO5	Apply the data analytics in the field of Health care.

Module No.	Unit No.	Details	Hrs.	CO
1		**Introduction to Data Analytics**	**05**	**CO1**
	1.1	Introduction to Data Analytics, Different types of data analytics: Descriptive analytics, Diagnostics Analytics, Predictive analytics, Prescriptive analytics		
	1.2	# Self-Learning: LinkedIn analytics, Netflix Analytics, Cricket and FIFA Analytics.		
2		**Data Analytics in GIS**	**10**	
	2.1	Introduction, Definition of GIS, Evolution of GIS , components of GIS		
	2.2	Vector Data Model : Topology, Non topological Vector models, Attribute Data in GIS, Attribute Data Entry, Vector Data Query, Manipulation of Fields and Attribute Data Raster Data Model : Elements of Raster Data Model, Types of Raster Data, Raster Data Structure, Raster Data Query, Data Compression, Data Conversion, Integration of Raster and Vector data		**CO2 CO1**
	2.3	Terrain Analysis: Data for Terrain Mapping and analysis, Terrain Mapping, slope and aspect, Surface curvature, Raster vs TIN, View shed and water shed analysis .		
	2.4	GIS application Case study: A real world problem and its step by step procedure using open source software tools.		
		# Self-learning: QGIS, Hadoop, GeoSpark R PostgreSQL, PostGIS, Python		
3		**Graph Analytics**	**10**	
	3.1	Introduction to the Social Network, Clustering of Social-Network Graphs, Direct Discovery of Communities		**CO3**
	3.2	Partitioning of Graphs, Finding Overlapping Communities, Simrank, Counting Triangles, Neighborhood Properties of Graphs		
	3.3	# Self-learning: GraphX tools of Apache.		
4		**Time series Analytics and forecasting**	10	
	4.1	Introduction, Finding and Wrangling Time Series Data, Exploratory Data Analysis for Time Series, Simulating Time Series Data, Storing Temporal Data,		**CO4**
	4.2	Statistical Models for Time Series, State Space Models for Time Series, forecasting methods, Testing for randomness, Regression based trend model, random walk model, moving average forecast, exponential smoothing forecast, seasonal models		
5		**Data analytics in Health Care Systems**	**10**	**CO5**
	5.1	Introduction, Components of HER, Benefits of EHR-		

		Barrier to Adopting HER challenges		
	5.2	Mining Sensor Data in Medical Informatics Challenges in Healthcare Data Analysis Sensor Data Mining Applications		
	5.3	Natural Language Processing and data mining for clinical text data: Mining Information from Clinical Text, Challenges of Processing Clinical Reports, Clinical Applications		
		Self-learning: Introduction to Social media analytics for healthcare,		
		Total	**45**	

 # Students should prepare all Self Learning topics on their own. Self-learning topics will enable students to gain extended knowledge of the topic. Assessment of these topics may be included in IA and Laboratory Experiments.

Recommended Books:

Sr. No.	Name/s of Author/s	Title of Book	Name of Publisher with country	Edition and Year of Publication
1.	Michael J. de Smith, Michael F. Goodchild and Paul A. Longley	Geospatial Analysis: A Comprehensive Guide to Principles, Techniques, and Software Tools,	Wiley, Second Edition	2019
2.	Anil Maheshwari	*Data Analytics*	Mc Graw Hill	2017
3.	James, G., Witten, D., Hastie, T., Tibshirani, R.	*An introductionto statistical learning with applications in R*	Springer	2013
4.	Chandan K. Reddy and Charu C Aggarwal	*Healthcare data analytics*	Taylor & Francis	2015
5.	Hui Yang and Eva K. Lee	*Healthcare Analytics: From Data to Knowledge to Healthcare Improvement*	Wiley	2016
6.	Mohammed J. Zaki and Wagner Miera Jr.	*Data Mining and Analysis: Fundamental Concepts and Algorithms*	Cambridge University Press	2014
7.	U. Dinesh Kumar	*Business Analytics*	Wiley	2017
8.	Kang-tsung Chang	"Introduction to Geographical Information Systems",	Tata McGraw Hill	Third Edition, 2003
9	Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman	Mining Massive dataset	Dreamtech Press	2nd edition 2016
10	Albright and Winston	Bussiness Analytics	Cengage Publication	5th edition, 2015
11	Aileen Nielsen	Practical Time Series Analysis	O'Reilly Media, Inc.	1st edition October 2019

Course Code	Course Title							
116h54L501	**Data Analytics**							
	TH			**P**		**TUT**		**Total**
Teaching Scheme(Hrs.)	-			**02**		**--**		**02**
Credits Assigned	-			**01**		**--**		**01**
	Marks							
Examination Scheme	**CA**		**ESE**	**TW**	**O**	**P**	**P&O**	**Total**
	ISE	**IA**						
	-	-	-	**25**	**25**	**--**	**--**	**50**

Term-Work:

Term work will consist of experiments/ tutorials covering entire syllabus of the course 'Data Analytics'. Students will be graded based on continuous assessment of their term work.