



Computer Science paper-IV
Data Analytics
[Discipline Specific Course]

Semester: VI	Credits: 02	Subject Code: BS62204	Lectures: 36
---------------------	--------------------	------------------------------	---------------------

Course Outcomes:
At the end of this course, the learner will be able to:
<ul style="list-style-type: none"> • Apply appropriate analytic techniques and tools to analyze data, create models, and identify insights that can lead to actionable results. • Develop understanding of key technologies in data analytics. • Use various data mining patterns to access the data . • Analyze various sources of text data and explain the text and social media analytics.

Unit 1: Introduction to Data Analytics	06
<ul style="list-style-type: none"> • Concept of data analytics • Data analysis vs Data analytics • Types of analytics • Diagnostic Analytics, Predictive Analytics , Prescriptive Analytics, Exploratory Analysis, Mechanistic Analysis • Mathematical models - Concept Model evaluation: metrics for evaluating classifiers - Class imbalance - AUC, ROC (Receiver-Operator Characteristic) curves, Evaluating value prediction models 	

Unit 2: Machine Learning Overview	10
<ul style="list-style-type: none"> • Introduction to Machine Learning, deep learning, Artificial intelligence • Applications for machine learning in data science • The modeling process <ul style="list-style-type: none"> ○ Engineering features and selecting a model, Training the model, Validating the model, Predicting new observations ○ Types of machine learning Supervised learning, Unsupervised learning, Semi-supervised learning, ensemble techniques • Regression models <ul style="list-style-type: none"> ○ Linear Regression, Polynomial Regression, Logistic Regression • Concept of classification, clustering and reinforcement learning. 	

Unit 3: Mining Frequent Patterns, Associations, and Correlations	10
<ul style="list-style-type: none"> • What kind of patterns can be mined <ul style="list-style-type: none"> ○ Class/Concept Description: Characterization and Discrimination, Mining Frequent Patterns, Associations, and Correlations, Classification and Regression for Predictive Analysis, Cluster Analysis, Outlier Analysis • Mining frequent patterns - Market Basket Analysis. • Frequent Itemsets, Closed Itemsets, and Association Rules • Frequent Itemset Mining Methods <ul style="list-style-type: none"> ○ Apriori Algorithm ○ Generating Association Rules from Frequent Itemsets . ○ Improving efficiency of apriori algorithm 	

Board of Studies	Name	Signature
Chairperson (HoD)	Ms. Ashwini Kulkarni	



o Frequent pattern growth (FP-growth) algorithm	
---	--

Unit 4: Social Media and Text Analytics	10
<ul style="list-style-type: none"> ● Overview of social media analytics ● Social Media Analytics Process, ● Seven layers of social media analytics, accessing social media data ● Key social media analytics methods <ul style="list-style-type: none"> o Social network analysis Link prediction, Community detection, Influence maximization, Expert finding, Prediction of trust and distrust among individuals ● Introduction to Natural Language Processing <ul style="list-style-type: none"> o Text Analytics : Tokenization, Bag of words, Word weighting : TF-IDF, n-Grams, stop words, Stemming and lemmatization, synonyms and parts of speech tagging ● Sentiment Analysis ● Document or text summarization ● Trend analytics Challenges to social media analytics 	

Recommended Reference Books:
<ul style="list-style-type: none"> ● Cathy O'Neil and Rachel Schutt , "<i>Doing Data Science</i>", O'Reilly, 2015. ● Chirag Shah , <i>A Hands-On Introduction to Data Science</i>, Washington Cambridge University Press ● Field Cady, John Wiley & Sons , <i>The Data Science Handbook</i>, Inc, 2017 ● Gypsy Nandi, Rupam Sharma , <i>Data Science Fundamentals and Practical Approaches</i>, BPB Publications, 2020

Journals:
<ul style="list-style-type: none"> ● Springer.com

Websites:
<ul style="list-style-type: none"> ● Github ● Kaggle.com ● https://www.analyticsvidhya.com/

E-resources:
<ul style="list-style-type: none"> ● MOOCS -Data camp, Swayam etc..

Board of Studies	Name	Signature
Chairperson (HoD)	Ms. Ashwini Kulkarni	

