

	Course Title: DATA SCIENCE		
	Course Code: 18CSE024	No. of Credits: 3: 0: 0 (L-T-P)	No. of lecture hours/week: 3
	Exam Duration: 3 hours	CIE + Assignment + SEE = 45 + 5 + 50 = 100	Total No. of Contact Hours: 42
Course Objectives:	Description		
	<ol style="list-style-type: none">1. Determine the appropriate natural language processing, machine learning and deep learning models to solve the business-related challenges.2. Indicate proficiency with statistical analysis of data to derive insight from results and interpret the data findings visually.3. Demonstrate skills in data management by obtaining, cleaning and transforming the data.4. Discuss how social networks appraise the ways in which the social clustering shape individuals and groups in contemporary society.		
Unit No	Syllabus Content		No of Hours
1.	Visualizing Data , matplotlib, Bar Charts, Line Charts, Scatterplots, Linear Algebra , Vectors, Matrices, Statistics , Describing a Single Set of Data, Correlation, Simpson’s Paradox, Some Other Correlational Caveats, Correlation and Causation, Probability , Dependence and Independence, Conditional Probability, Bayes’s Theorem, Random Variables, Continuous Distributions, The Normal Distribution, The Central Limit Theorem.		08
2.	Hypothesis and Inference , Statistical Hypothesis Testing, Example: Flipping a Coin, p-Values, Confidence Intervals, p-Hacking, Example: Running an A/B Test, Bayesian Inference, Gradient Descent , The Idea Behind Gradient Descent Estimating the Gradient, Using the Gradient, Choosing the Right Step Size, Using Gradient Descent to Fit Models, Minibatch and Stochastic Gradient Descent, Getting Data , stdin and stdout, Reading Files, Scraping the Web, Using APIs, Example: Using the Twitter APIs, Working with Data , Exploring Your Data, Using NamedTuples, Dataclasses, Cleaning and Munging, Manipulating Data, Rescaling, An Aside: tqdm, Dimensionality Reduction.		08
3.	Machine Learning , Modeling, What Is Machine Learning?, Overfitting and Underfitting, Correctness, The Bias-Variance Tradeoff, Feature Extraction and Selection, k-Nearest Neighbors , The Model, Example: The Iris Dataset, The Curse of Dimensionality, Naive Bayes , A Really Dumb Spam Filter, A More Sophisticated Spam Filter, Implementation, Testing Our Model, Using Our Model, Simple Linear Regression , The Model, Using Gradient Descent, Maximum Likelihood Estimation, Multiple Regression , The Model, Further Assumptions of the Least Squares Model, Fitting the Model, Interpreting the Model, Goodness of Fit, Digression: The Bootstrap, Standard Errors of Regression Coefficients, Regularization, Logistic Regression , The Problem, The		09

TEXT BOOKS:	
1. Joel Grus, “Data Science from Scratch” , 2 nd Edition, O’Reilly Publications/Shroff Publishers and Distributors Pvt. Ltd., 2019. ISBN-13: 978-9352138326.	
REFERENCE BOOKS:	
1. Emily Robinson and Jacqueline Nolis, “Build a Career in Data Science” , 1 st Edition, Manning Publications, 2020. ISBN: 978-1617296246.	
2. Aurélien Géron, “Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems” , 2 nd Edition, O’Reilly Publications/Shroff Publishers and Distributors Pvt. Ltd., 2019. ISBN-13: 978-1492032649.	
3. François Chollet, “Deep Learning with Python” , 1 st Edition, Manning Publications, 2017. ISBN-13: 978-1617294433	
4. Jeremy Howard and Sylvain Gugger, “Deep Learning for Coders with fastai and PyTorch” , 1 st Edition, O’Reilly Publications/Shroff Publishers and Distributors Pvt. Ltd., 2020. ISBN-13: 978-1492045526.	
5. Sebastian Raschka and Vahid Mirjalili, “Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2” , 3 rd Edition, Packt Publishing Limited, 2019. ISBN-13: 978-1789955750	
SELF-STUDY REFERENCES/WEBLINKS:	
1. Natural Language Processing https://www.youtube.com/watch?v=xvqsFTUsOmc	
2. Network Analysis https://www.youtube.com/watch?v=K5xiFDClgjo	
3. Recommender Systems https://www.youtube.com/watch?v=39vJRxIPSxw	
COURSE COORDINATOR:	Dr.Gowrishankar S.