Department of Data Science
Programme Outcomes and Programme Specific Outcomes
MSc Data Science

Programme Outcomes

Upon completion of MSc in Data Science degree, graduates will possess the following data science skills and abilities.

PO1: Possess a theoretical understanding, explain and critically assess the key concepts and techniques from the disciplines defining modern data science and analytics.

PO2: Critically evaluate emerging data analysis technologies and how they can be applied to heterogeneous data at volume, scale and types, in order to get insight for business, scientific or social innovation.

PO3: Analyse in depth how data analysis techniques can be applied to a range of interdisciplinary research areas.

PO4: Effectively use modern data science programming languages and technologies to scrape, clean, organize, explore, visualize, and model large volumes and varieties of data.

PO5: Evaluate, select, combine and apply advanced skills, data science tools and techniques in the related areas of artificial intelligence to the design of solutions to data science and analytics tasks.

PO6: Prepare for careers as data scientists by proposing, planning, developing, evaluating and creating a commercially and/or research-wise relevant project and/or product for business, science and society.

PO7: Develop professional communication skills (e.g., writing, presentations, interviews, email etiquette, etc.), effective time and resource management skills as well as leadership and team working skills towards meeting organizational goals.

PO8: Understand, value and safeguard social, legal and ethical use of data that increasingly challenge and confront data scientists while developing data science systems.

PO9: Learn effectively and independently to acquire new knowledge and skills for the purpose of continuing professional development in related areas of data science.

Programme Specific Outcomes

PSO1: Show mastery over different applications of data analytics namely web analytics, customer analytics, supply chain analytics and social network analytics.

PSO2: Build software applications using new languages and tools such as Neo4J, Tableau, Julia, SpaCy and Rasa

PSO3: Develop a disruptive entrepreneurship spirit and integrate with the data science community.

PSO4: Identify and assess the needs of an organization for a data science task by conducting a needs assessment and communicating data science options and limitations that could meet organizational needs.

BISHOP HEBER COLLEGE (AUTONOMOUS), TIRUCHIRAPPALLI-620 017
M. Sc., Data Science
(Applicable to Candidates admitted from the Academic Year 2020-2021 onwards

Sem	Course	Course Title	Course Code	Hours / Week	Credits	Marks		
						CIA	ESE	Total
I	Core I	Mathematical Foundation for Data Science	P20DS101	5	4	25	75	100
	Core II	Problem Solving using Python and R	P19DS102	5	4	25	75	100
	Core III	NoSQL Database Management	P20DS103	5	4	25	75	100
	Elective I	Essential Statistics for Data Science Design and Analysis of Algorithms Software Engineering	P20DS1:1 P19DS1:A P19DS1:B	5	4	25	75	100
	Core Practical I	Problem Solving using Python and R Lab	P19DS1P1	5	3	40	60	100
	Core Practical II	NoSQL Database Management Lab	P19DS1P2	5	3	40	60	100
II	Core IV	Time Series Analysis and Forecasting	P20DS204	4	4	25	75	100
	Core V	Data and Visual Analytics	P19DS205	4	4	25	75	100
	Core VI	Practical Machine Learning	P19DS206	4	4	25	75	100
	Elective II	Natural Language Processing Multivariate Analysis	P20DS2:2 P19DS2:A	4	4	25	75	100
	Elective III	Health Care Data Analytics Basics of Bioinformatics	P19DS2:3 P19DS2:B	3	3	25	75	100
	Core Practical III	Data and Visual Analytics Lab	P19DS2P3	3	3	40	60	100
	Core Practical IV	Practical Machine Learning Lab	P19DS2P4	3	3	40	60	100
	Core Practical V	Natural Language Processing Lab	P19DS2P5	3	3	40	60	100
	VLO	RI/MI	P17VL2:1 P17VL2:2	2	2	25	75	100
III	Core VII	Principles of Deep Learning	P20DS307	4	4	25	75	100
	Core VIII	Big Data Management and Analytics	P19DS308	4	4	25	75	100
	Core IX	Social Media Analytics	P19DS309	4	4	25	75	100
	Elective IV	Image and Video Analytics Computational Genomics	P20DS3:4 P19DS3:A	4	4	25	75	100
	Core Practical VI	Big Data Management and Analytics Lab	P19DS3P6	5	3	40	60	100
	Core Practical VII	Social Media Analytics Lab	P19DS3P7	4	3	40	60	100
	Core Practical VIII	Principles of Deep Learning Lab	P20DS3P8	5	3	40	60	100
IV	Core X	Customer Relationship Management	P20DS410	5	4	25	75	100
	Elective V	Supply Chain Management Web Development using Python	P19DS4:5 P20DS4:A	5	4	25	75	100
	Core Project	Core Project	P20DS4PJ	40	60	---	---	100
			Total Credits		90			

PROGRAMME ARTICULATION MATRIX

Course Code	Programme Outcomes												
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
P20DS101	H	H	H	-	M	M	M	H	-	H	H	-	-
P19DS102	H	H	H	H	M	M	M	H	-	H	H	H	M
P20DS103	H	H	H	H	M	M	M	H	-	H	H	-	-
P20DS1:1	H	H	H	-	M	M	M	H	-	H	H	-	-
P19DS1P1	H	H	H	H	M	M	M	H	-	H	H	H	M
P19DS1P2	H	H	H	H	M	M	M	H	-	H	H	H	M
P20DS204	H	H	H	H	M	M	M	H	-	H	H	H	M
P19DS205	H	H	H	H	M	M	M	H	-	H	H	H	M
P19DS206	H	H	H	H	M	M	M	H	-	H	H	H	M
P20DS2:2	H	H	H	H	M	M	M	H	M	H	H	H	M
P19DS2:3	H	H	H	H	M	M	M	H	-	H	H	H	M
P19DS2P3	H	H	H	H	M	M	M	H	-	H	H	H	M
P19DS2P4	H	H	H	-	M	M	M	H	-	H	H	H	L
P19DS2P5	H	H	H	H	M	M	M	H	-	H	H	H	M
P20DS307	H	H	H	H	M	M	M	H	-	H	H	H	-
P19DS308	H	H	H	H	M	M	M	H	-	H	H	H	M
P19DS309	H	H	H	H	M	M	M	H	-	H	H	H	-
P19DS3:4	H	H	H	H	M	M	M	H	L	H	H	H	M
P19DS3P6	H	H	H	H	M	M	M	H	-	H	H	H	M
P19DS3P7	H	H	H	H	M	M	M	H	-	H	H	H	M
P20DS3P8	H	H	H	H	M	M	M	H	-	H	H	H	-
P20DS410	L	L	L	M	H	H	H	H	H	H	H	H	M
P19DS4:5	L	L	L	M	H	H	H	H	H	H	H	H	M
P20DS4PJ	H	H	H	H	M	M	M	H	-	H	H	H	M

CORE I: MATHEMATICAL FOUNDATION FOR DATA SCIENCE			
Semester	I	Hours/Week	5
Course Code	P20DS101	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Solve systems of linear equations by use of the matrix	K5	I
CO2	Determining basis and understanding linear mappings of vector spaces	K5	I
CO3	Determine the Orthogonality and Projections	K5	II
CO4	Evaluate eigenvectors and eigenvalues	K5	III
CO5	Explain the properties gradients and PDE	K6	IV
CO6	Summarize different probability distributions	K6	V

2. A. SYLLABUS

UNIT - I: Linear Algebra

Systems of Linear Equations – Matrices - Solving Systems of Linear Equations - Vector Spaces - Linear Independence - Basis and Rank - Linear Mappings - Affine Spaces.

UNIT - II: Analytic Geometry

Norms - Inner Products - Lengths and Distances - Angles and Orthogonality - Orthonormal Basis - Orthogonal Complement - Inner Product of Functions - Orthogonal Projections - Rotations

UNIT - III: Matrix Decompositions

Determinant and Trace - Eigenvalues and Eigenvectors - Cholesky Decomposition – Eigen decomposition and Diagonalization - Singular Value Decomposition Matrix Approximation - Matrix Phylogeny -

UNIT – IV: Vector Calculus

Differentiation of Univariate Functions - Partial Differentiation and Gradients - Gradients of Vector-Valued Functions - Gradients of Matrices - Useful Identities for Computing Gradients - Backpropagation and Automatic Differentiation - Higher-Order Derivatives - Linearization and Multivariate Taylor Series

UNIT – V: Probability and Distributions

Probability and Distributions: Construction of a Probability Space - Discrete and Continuous Probabilities - Sum Rule, Product Rule, and Bayes' Theorem - Summary Statistics and Independence - Gaussian Distribution.

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Mathematics for Data Science	https://www.coursera.org/specializations/mathematics-for-data-science
2	Mathematics for Machine Learning Specialization	coursera.org/specializations/mathematics-machine-learning
3	Topics in Mathematics of Data	https://ocw.mit.edu/courses/mathematics/18-s096-

	Science	topics-in-mathematics-of-data-science-fall-2015/

C. TEXT BOOK(S)

1. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, **"** *Mathematics for Machine Learning***"**, Cambridge Press, 2019 (Chapters 2, 3, 4,5,6)

D. REFERENCE BOOKS

1. Gilbert Strang, **"Introduction to Linear Algebra***",* 3ed, Cambridge Press, 2003.
2. M. D. Weir, J. Hass, and G. B. Thomas**, "Thomas' calculus***"*, Pearson, 2016.

E. WEB LINKS

1. https://elitedatascience.com/learn-math-for-data-science
2. https://machinelearningmastery.com/gentle-introduction-linear-algebra/'

3. SPECIFIC LEARNING OUTCOMES (SLO)

Unit/ Section	Topic	Learning outcomes	Level
I	**Linear Algebra**		
1.1	Vectors	Understand the vectors and its linkage to understanding data features	K1
1.2	System of Linear Equations	Represent the linear equations as matrices	K2
		Evaluate the different matrix operations and types	K5
1.3	Solving system of Linear Equations	Prescribe the solution for a system of linear equations to Row Echelon form	K6
Vector Spaces			
1.4	Vector Spaces	Summarize different properties of vector spaces and subspaces	K6
1.5	Linear Independence	Assess vectors of a vector space as a linear combination	K5
		Explain the properties of linear independent sets	K3
1.6	Basis and Rank	Determine the basis vectors and learning about its properties	K5
		Determine the rank of a matrix	K5
Linear Mappings			
1.7	Linear Mappings	Analyze the properties of linear mappings	K4
		Formulate linear mappings as matrices	K4
		Assess the image and kernel of a linear mapping	K5
1.8	Affine subspaces	Differentiate affine subspaces from vector spaces	K5
		Explain the properties of affine mappings	K5
II	**Analytical Geometry**		
2.1	Norms and distances	Explain the different types of norms and distances	K4
Inner Products			
2.2	Dot Products	Determine the properties of dot products of vectors	K5
2.3	General Inner Products	Determine inner products from bilinear mappings	K3

		Determine symmetric and positive definite matrices from inner products	K4
		Determine lengths and distances from inner products	K3
Angles and Orthogonality			
2.4	Angles	Determine angles of vectors	K3
2.5	Orthogonality	Summarize the orthogonal and orthonormal vectors and its properties	K6
2.6	Orthonormal Basis	Explain orthogonal and orthonormal basis and its properties	K5
2.7	Orthogonal Complement	Describe the orthogonal complement of a vector space	K4
Orthogonal Projections			
2.8	Projections	Assess the Projections onto n-dimensional subspaces	K5
		Formulate the Gram-Schimdt Orthogonalization	K6
Rotations			
2.9	Rotations	Determine the n-dimensional rotations	K5
		Design the rotation matrix	K6
III	**Matrix Decompositions**		
3.1	Determinant and Trace	Determine determinant of a matrix	K3
		Determine the trace of a matrix	K4
Eigen Vector and Spaces			
3.2	Eigen vectors and Spaces	Computing eigen values, eigen vectors and eigen spaces,	K4
Cholesky Decomposition			
3.3	Cholesky Decomposition	Explain the Cholesky Decomposition of a matrix	K5
	Eigen Decomposition and Diagonalization		
3.4	Diagonisable Matrices	Construct the diagonisable matrices	K6
		Compute Eigen Decomposition of a matrix	K4
Singular Value Decomposition			
3.5	SVD Theorem	Understand the underlying principles of the SVD Theorem	K4
3.6	Constructing an SVD	Explain the steps to construct a SVD	K5
		Differentiate the between Eigen decomposition and SVD	K6
Matrix Approximations			
3.7	Spectral Norm	Explain the Spectral norm of a matrix and related properties	K5
IV	**Vector Calculus**		
4.1	Differentiation of Univariate functions	Evaluate the difference quotient of a univariate function	K3
		Assess the properties of the derivative of a function	K2
		Differentiate the Taylor's function	K4
		Assess the differentiation rules	K2
Partial Differential Equations and Gradients			
4.2	Partial Derivatives	Evaluate the partial derivatives of a function	K4
		Assess the properties of Gradients	K5

		Understand the basic rules of Partial differential equation	K2
		Evaluate the chain rule of differentiation	K5
4.3	Gradients of vector valued functions	Assess the gradients of a vector valued function	K5
		Explain the properties of Jacobian Matrix	K5
4.4	Gradient Matrices	Explain the steps for gradient of vectors with respect to matrices	K4
		Summarize the gradient of matrices with respect to Matrices	K6
Backpropagation			
4.5	Gradients in a Deep Network	Formulate the gradients of a deep neural network	K6
4.6	Automatic differentiation	Design the automatic differentiation	K6
Higher Order Derivatives			
4.7	Higher Order derivatives	Compute higher order partial derivatives	K3
		Evaluate the differentiation of multivariate Taylor Series	K5
V	**Probability Distributions**		
5.1	Probability Space	Understand sample space, event and computing probabilities	K3
		Describe the properties of a random variables	K4
Discrete and Continuous probabilities			
5.2	Discrete Probabilities	Assess the joint probabilities	K5
		Compile the features of probability mass function	K6
5.3	Continuous probabilities	Discriminate the probability density function and cumulative distribution function	K5
Bayes Theorem			
5.4	Bayes Theorem	Understand the sum and product rule	K3
		Evaluate the Bayes Theorem	K5
		Analyze the likelihood and posterior probabilities	K4
Summary Statistics and Independence			
5.8	Expected Value	Explain the expected value of a probabilistic function	K5
5.9	Covariance	Assess the Covariance	K5
		Evaluate the variance and correlation of random variables	K5
5.10	Statistical Independence	Describe the statistical independence of two random variables	K4
5.11	Inner products	Explain the inner products of two random variables	K4
Gaussian Distribution			
5.12	Gaussian distribution	Assess the mean and covariance of Gaussian Distribution	K5
		Formulate the marginal and conditional probabilities of Gaussian Distribution	K6
Other named distributions			
5.13	Bernoulli Distribution	Assess the mean and covariance of Bernoulli Distribution	K5

5.14	Binomial Distribution	Compute the mean and covariance of Gaussian Distribution	K6
5.15	Beta Distribution	Analyze the mean and covariance of Beta Distribution	K4

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**			**H**			**L**			**L**			**H**
CO2	**H**	**L**	**H**				**L**			**L**	**M**		
CO3	**M**	**M**		**H**	**M**		**L**		**H**			**M**	**H**
CO4	**M**	**L**		**H**	**M**			**L**	**H**		**M**		
CO5	**H**		**H**	**H**	**M**			**M**					**H**
CO6	**H**		**H**		**H**		**H**	**H**	**H**	**H**		**M**	**H**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. P. S. Eliahim Jeevaraj

CORE II: PROBLEM SOLVING USING PYTHON AND R			
Semester	I	Hours/Week	5
Course Code	P20DS102	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Understand python basic syntax, usage of built in functions	K3	I
CO2	Understand conditional and looping statements and build user defined functions	K3	I
CO3	Explain the concepts of files using Python	K5	II
CO4	Develop object oriented programs in Python	K6	III
CO5	Access and Design the internet and database data	K6	IV
CO6	Understand R basic data structures and develop programs	K5	V

2. A. SYLLABUS

UNIT - I: Python Basics, Functions, Loops and Strings
Variables – Getting Inputs – Conditions – Catching exceptions – Function calls – Built-in functions – Type conversion functions and math functions – Parameters and arguments –While statement – Infinite loops -Continue statement – For loops – Strings -Slice - The in operator – String comparison – String methods- parsing strings – Format operator.

UNIT - II: Files and Lists
Opening files – Text files – Reading files – Searching through files – Writing files – Traversing list – List operations – List slice – List methods – Deleting elements – Built-in list functions – Objects, value and aliasing – List arguments.

UNIT - III: Dictionaries, Tuples and OOP
Dictionaries – Files and dictionaries – Looping and dictionaries – Tuples – Comparing tuples – Tuple assignments – Dictionaries and tuples – Tuples as keys in dictionaries – Creating objects – Encapsulation – Classes as types – Object lifecycle – Instances – Inheritance.

UNIT – IV: Internet Programming
Regular expressions – Character matching – Extracting data – Escape character – Designing simple web browser using sockets – Retrieving images using HTTP – Retrieving web pages using urllib – Reading binary files using urllib – Accessing data from databases

UNIT – V: Programming with R
Variables - Vector, matrix, arrays – List – Data Frames – Functions – Strings – Factors – Loops – Packages –Date and Time – Files - Making packages

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Introduction to Python Programming	https://www.udacity.com/course/introduction-to-python--ud1110
2	Introduction to Python	https://www.coursera.org/projects/introduction-to-python
3	Introduction to Python	https://realpython.com/learning-paths/python3-

		introduction/
4	R Programming	https://www.coursera.org/learn/r-programming

C. TEXT BOOK(S)

1. Allen B. Downey, ―Think Python: How to Think like a Computer Scientist, 2nd edition, Updated for Python O_Reilly Publishers, 2016
2. Charles R. Severance, Python for Everybody: "**Exploring data using Python 3"**, Schroff Publishers, 1ed, 2017, ISBN 978-9352136278.
3. Richard Cotton, "Learning R", O'Reilly, 2013

D. REFERENCE BOOKS

1. Zed Shaw's, Learn Python the Hard Way: A Very Simple Introduction to the Terrifyingly Beautiful World of Computers and Code, Addison-Wesley Professional; 3 edition, 2013
2. Robert Sedgewick, Kevin Wayne, Robert Dondero, Introduction to Programming in Python: An Inter - disciplinary Approach, Pearson India Education Services Pvt. Ltd., 2016.
3. Wesley J Chun, Core Python Programming , 2nd edition, Prentice Hall ,2009
4. Colin Gillespie, Robin Lovelace, and Efficient R Programming: A Practical Guide to Smarter Programming," O'Reilly Media, Inc.", 2016
5. Paul Teetor, R Cookbook-Proven Recipes for Data Analysis, Statistics, and Graphics, O'Reilly Media, 2011

E. WEB LINKS

- https://www.kaggle.com/learn/python
- https://www.dataquest.io/course/introduction-to-data-analysis-in-r/

3. SPECIFIC LEARNING OUTCOMES (SLO)

Unit/ Section	Topic	Learning outcomes	Level
I	**Python Basics, Functions, Loops and Strings**		
1.1	Python basics	• Understand python variables and assignment	K1
	Built in Functions		
1.2	Built in Functions and other important functions	• Understand python built in functions	K1
		• Understand conversion and math functions	K1
1.3	Conditional and looping statements	• Build if, else statements within programs and understand outputs	K3
		• Build while and for loops for understanding looping concept	K3
	User Defined Functions		
1.4	User Defined Functions	• Understanding functions structures	K2
		• Understanding parameters and arguments	K2
	Manipulating strings		
1.5	Handling strings	• Understanding strings data type	K1
		• Applying string slicing	K2
		• Applying string comparison, parsing and string formatting	K3
II	**Files and Lists**		
2.1	Handling files	• Understand syntax to read and write files	K2

		• Using io library functions to check file and folder existence	K3
	Lists		
2.2	List data structure	• Understanding list data structure and operations	K1
		• Applying list slicing and items deletion	K2
	Lists and loops		
2.3	Using lists in loops	• Using list comprehension in programs	K3
III	**Dictionaries, Tuples and OOP**		
3.1	Dictionary data structure	• Understanding dictionary data structure	K1
	Loops and Dictionaries		
3.2	Using dictionaries in loops	• Applying dictionary comprehension in programs	K2
	Tuples		
3.3	Tuples data structure	• Understanding tuples data structure	K1
3.4	Tuple Operations	• Applying tuples operations	K1
	Dictionaries and Tuples		
3.5	Dictionaries and Tuples	• Applying tuples as keys in dictionaries	K2
	Objects and Classes		
3.6	Objects and Classes structure	• Understanding objects and classes	K4
		• Understanding inheritance	K5
		• Creating objects and Classes	K4
IV	**Internet Programming**		
4.1	Regular Expressions	• Understanding pattern matching in strings	K4
		• Applying re module functions for pattern matching in various examples	K5
	Accessing data using urllib		
4.2	Retrieving images in web	• Using HTTP to retrieve images	K3
4.3	Retrieving web pages	• Using urllib module to retrieve web pages	K3
	Accessing data from databases		
4.4	Extracting data from databases	• Using modules to extract data from SQL databases	K4
V	**Programming with R**		
5.1	R data structures	• Understand R data types	K2
		• Understand R data structures	K2
	Functions		
5.2	User defined Functions	• Creating user defined functions	K3
	R libraries		
5.3	R libraries	• Using R libraries	K2
		• Creating user defined R libraries	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4

CO1	**H**	**M**	**H**	**M**	**L**	**M**	**M**	**M**	**M**	**M**	**-**	**H**	**H**
CO2	**H**	**H**	**M**	**L**	**-**	**L**	**-**	**L**	**M**	**M**	**-**	**M**	**-**
CO3	**H**	**H**	**H**	**H**	**M**	**M**	**L**	**-**	**-**	**H**	**H**	**H**	**M-**
CO4	**H**	**H**	**H**	**H**	**H**	**M**	**L**	**M**	**M**	**H**	**H**	**M**	**H**
CO5	**H**	**M**	**-**	**M**	**L**	**M**	**-**	**H**	**M**	**H**	**H**	**M**	**-**
CO6	**M**	**M**	**-**	**H**	**M**	**L**	**-**	**-**	**L**	**H**	**H**	**H**	**M**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project, Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Prof. K. Jemimah

Core III: NoSQL DATABASE MANAGEMENT			
Semester	I	Hours/Week	5
Course Code	P20DS103	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Construct the Queries and sub queries in SQL	K6	I
CO2	Construct queries to retrieve the data from more than one tables using different techniques.	K6	II
CO3	Design and Analyse different techniques and operations in Mango DB	K6	III
CO4	Assess various operators and clause to generate unstructured data	K5	IV
CO5	Choose different indices to retrieve data	K5	IV
CO6	Construct databases using SQL, MongoDB and Neo4J	K6	V

2. A. SYLLABUS

UNIT 1- Structured Query Language-I **12 Hours**

ER Model: Entity types, Attribute types, Relationship types – Weak entity types, Ternary relationship types – Examples of ER model. Enhanced ER model: Specialization/Generalization – Categorization - Aggregation – Examples of EER. Relational DB Process and outcome approach - Simple Queries on one table – First look at joins – Sub queries.

UNIT 2- Structured Query Language-II **12 Hours**

Self Joins: Self relationships, Questions involving Both – Multiple relations between tables – Set operations – Aggregate Operations – Window functions – Efficiency considerations: Indexing and Join Techniques.

UNIT 3- MongoDB-I **12 Hours**

Introduction: MongoDB document, collection and database – Basic Operations – Datatypes – Creating, deleting, updating documents: insert, batch insert, remove, find, findone, update – arrays – insert – Updating multiple documents

UNIT 4- MongoDB-II **12 Hours**

Comparison operators – OR and NOT queries – Querying arrays – Querying on embedded documents – WHERE queries – Limits, skips and sort – Compound Index – Unique index – Sparse Index – Pipeline aggregation: MATCH, PROJECT, GROUP and UNWIND clauses.

UNIT 5- Neo4J and Cypher **12 Hours**

Labelled Property Graph Model – Querying graphs using Cypher: CREATE AND ASSERT, MATCH, WHERE and RETURN clauses– ORDER BY – WITH clause – Case Study: Telent.net Social recommendations application.

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Database Architecture	https://www.youtube.com/watch?v=W6P58yb-edE
2	Normalization	https://www.guru99.com/database-

		normalization.html
3	DynamoDB	https://www.tutorialspoint.com/dynamodb/index.htm
4	Apache HIVE	https://data-flair.training/blogs/apache-hive-tutorial/

C. TEXT BOOKS

1. Clare Churcher. *Beginning SQL Queries: From Novice to Professional*, APress, 2ed, 2016. ISBN 978-1-4842-1954-6
2. Wilfried Lemahieu, Seppe vanden Broucke and Bart Baesens. *Principles of Database Management: The Practical Guide to Storing, Managing and Analyzing Big and Small Data*, Cambridge University Press, 2018. ISBN 978-1-107-18612-5 (Chapter 3 ER diagram only)
3. Kristina Chodorow, *MongoDB: The Definitive Guide*, 2ed, Oreilly Publishers
4. Ian Robinson, Jim Webber and Emil Eifrem. Graph Databases: New Opportunities for connected data. 2ed, Oreilly Publishers. ISBN 978-1491930892.

D. REFERENCES BOOKS

1. Eric Redmond; Jim R. Wilson. *Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement*. Pragmatic Bookshelf. 2012. ISBN: 1934356921Pramod J. Sadalage; Martin Fowler. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Addison-Wesley. 2012 ISBN: 0321826620
2. Adam Fowler. *NoSQL for Dummies*. John Wiley. 2015. ISBN 978-1-118-90574-6
3. Guy Harrison. *Next Generation Databases*. APress. 2016. 978-1-484213-30-8
4. Thomas M. Connolly and Carolyn E. Begg. Database Systems: "*A Practical Approach to Design, Implementation, and Management",* 6th Edition, Pearson, 2015.

E. WEB LINKS

1. https://www.simplilearn.com/introduction-to-nosql-databases-tutorial-video
2. https://www.w3schools.com/sql/

3. SPECIFIC LEARNING OUTCOMES (SLO)

Unit/ Section	Course Content	Learning outcomes	Level
I	**Structured Query Language-I**		
1.1	ER Model: Entity types, Attribute types, Relationship types	Construct Entity Relationship Model to structure the relational Data base	K2
1.2	Weak entity types, Ternary relationship types – Examples of ER model	Classify the Strong and Weak entity sets	K4
1.3	Specialization/Generalization Categorization	Categorize the data into specialized and Generalized	K4
1.4	Aggregation – Examples of EER	Summarize the data in the table	K2
	Relational DB Approach		
1.5	Simple Queries on one table	Build simple SQL query	K3
1.6	Sub queries	Compile the sub query to retrieve data from the table	K6
1.7	First look at joins	Explain the join concept	K5
II	**Structured Query Language-II**		
2.1	Joins-Types of Joins	Assess the various joins to get data from more than one table	K5
2.2	Multiple relations between tables	Determine the relationship among the tables in the database	K5

2.3	Set operations	Experiment with set operators	K3
	Group Functions		
2.4	Aggregate Operations	Analyse the aggregate function in group by clause	K4
2.5	Window functions	Explain the window functions	K2
2.6	Efficiency considerations: Indexing	Apply indexing technique for effective performance	K3
2.7	Join Techniques	Adapt joins technique to get data	K6
III	**MongoDB-I**		
3.1	Introduction: MongoDB	Relate SQL with NoSQL Data base	K1
3.2	Document, collection and database	Illustrate the document, collection in NoSQL	K2
3.3	Basic Operations	Construct NoSQL query with basic operations	K3
3.4	Datatypes	Categorize the datatypes in NoSQL	K4
	Creating, deleting, updating documents		
3.5	Insert, batch insert, remove, find, find one, update	Construct document and perform CURD operations	K6
3.6	Arrays	Adapt array to store the data in table	K6
3.7	Updating multiple documents	Examine the updating for multiple document	K4
IV	**MongoDB-II**		
4.1	Comparison operators – OR and NOT queries	Assess OR and NOT operators	K5
4.2	Querying arrays	Develop query using Array in NoSQL	K3
	Embedded Document		
4.3	Querying on embedded documents	Compile the NoSQL query for embedded document	K6
4.4	WHERE queries – Limits, skips and sort	Explain to filter the collections in the document	K5
4.5	Compound Index	Apply the compound index	K3
4.6	Unique index – Sparse Index	Explain the types of Index	K2
4.7	Pipeline aggregation: MATCH, PROJECT, GROUP and UNWIND clauses.	Evaluate the pipeline functions	K5
V	**Neo4J and Cypher**		
5.1	Introduction to Graph Database	Explain how to access nodes and relationships in a native graph database	K2
5.2	Data Modelling with Graphs	Apply different models with graph	K3
Querying Graphs			
5.3	An Introduction to Cypher	Illustrate the cypher	K2
5.4	Create and Assert, Match, Where and Return Clause	Apply various operations in graph database	K3
5.5	ORDER BY	Examine the order by in graph database	K4
5.6	WITH clause	Construct the graph data model to design the questions in the form of cypher queries using with clause	K6
5.7	Case Study: Telent.net Social recommendations application.	Analyse the graph model in social recommend applications	K4

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	H	M	H	M	L	M	M	M	M	M	-	H	H
CO2	H	H	M	L	-	L	-	L	M	M	-	M	-
CO3	H	H	H	H	M	M	L	-	-	H	H	H	M-
CO4	H	H	H	H	H	M	L	M	M	H	H	M	H
CO5	H	M	-	M	L	M	-	H	M	H	H	M	-
CO6	M	M	-	H	M	L	-	-	L	H	H	H	M

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. M. Lovelin Pon Felichiah

ELECTIVE-1: ESSENTIAL STATISTICS FOR DATA SCIENCE			
Semester	I	Hours/Week	5
Course Code	P20DS1:1	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Experiment the methods of descriptive statistics and variability.	K5	I
CO2	Assess and examine the different tests of the statistical inferences	K5	II
CO3	Build the nonparametric statistics methods	K6	III
CO4	Classify and Construct the different types of regression methods for data analytics	K6	IV
CO5	Analyze the different properties of the regression methods.	K6	V
CO6	Evaluate all possible regression on given data sets.	K6	V

2. A. SYLLABUS

Unit I: Descriptive Statistics
Introduction to Statistics - Organizing Data Using Tables and Graphs- Measures of Central Tendency: Mode – Median – Mean. Measures of Variability: Variability – Range - Interquartile Range - Standard Deviation.

Unit II: Inferential Statistics – I
Sampling Distribution of Means: Sampling Distribution - Central Limit Theorem. Hypothesis Testing: Hypothesis Testing Steps - Effect Size for a Z-Test - Assumptions – Errors – Power. One-Sample t Test: t-Statistics – t- Distributions - One-Sample t Test – Effect Size – Assumptions. Two-Sample t Test: Independent Samples Design: Calculations – Hypothesis Testing – Effect Size – Assumptions. Two-Sample t Test: Related Samples Design: Calculations – Hypothesis Testing – Effect Size – Assumptions.

Unit III: Inferential Statistics - II
Confidence Interval versus Point Estimation: Introduction- Point Estimates - Confidence Intervals – One Sample t- Test - Two-Sample t Test: Independent Samples Design – Repeated Measure Design - Degree of Confidence Vs. Degree of Specificity One-Way Analysis of Variance: Introduction – Variance – F- statistics – Hypothesis Testing with F- Statistic - F- Distribution Table - Notations for ANOVA - Calculations – Hypothesis Testing – Effect Size – Assumptions. Chi-Square: Chi-Square - Chi-Square Statistic – Assumptions- Goodness of Fit - Goodness of Fit for Known Proportions- Goodness of Fit for No Preference – Test of Independence - Nonparametric Statistics for Ordinal Data: Mann-Whitney U Test - Kruskal-Wallis H Test. Correlation: Introduction – Scatter Plot - Pearson Product Moment Correlation - Hypothesis Testing - Coefficients of Determination and Non determination – Interpretation and Uses of The Pearson Correlation.

Unit IV: Regression Analysis - I
Regression Model - Goals of Regression Analysis - Statistical Computing in Regression Analysis - Simple Linear Regression – Multiple Linear Regression – Logistic Regression – Poisson Regression

Unit V: Regression Analysis - II

Detection of Outliers and Influential Observations: Detection of Outliers in Multiple Linear Regression - Detection of Influential Observations in Multiple Linear Regression - Test for Mean-shift Outliers - Graphical Display of Regression Diagnosis. Model Selection: Effect of Underfitting and Overfitting - All Possible Regressions – Stepwise Selection. Model Diagnostics: Test Heteroscedasticity - Detection of Regression Functional Form

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Statistical Thinking for Data Science and Analytics	https://www.edx.org/course/statistical-thinking-for-data-science-and-analytic
2	Linear Regression for Business Statistics	https://www.coursera.org/learn/linear-regression-business-statistics
3	Learning Statistics with R	https://learningstatisticswithr.com/
4	15 Types of Regression	https://www.listendata.com/2018/03/regression-analysis.html

C. TEXT BOOKS

1. Cheryl Ann Willard, *"Statistical Methods: An Introduction to Basic Statistical Concepts and Analysis"* , Routledge, 2020. (Unit – I – III)
2. Xin Yan & Xiaogang Su, *"Linear Regression Analysis : Theory and Computing*"**,** World Scientific Pulishing Ltd, 2009. (Unit – IV: Chapter 1,2,3, 8.5,8.6; Unit – V: Chapters 4.2,4.3, 4.4, 4.5, 5.1-5.3, 6.1, 6.2)

D. REFERENCE BOOKS

1. John.E.Freund, Irwin Miller, Marylees Miller *"Mathematical Statistics with Applications* ", 8th, Prentice Hall of India, 2014
2. Ross, Sheldon. M, *"Introduction to Probability and Statistics for Engineers and Scientists*", Academic Press, 2009
3. D.C Montgomery, E.A Peck and G.G Vining, *"Introduction to Linear Regression Analysis***"**, John Wiley and Sons, 2003.
4. S. Chatterjee and AHadi, **"***Regression Analysis by Example***"**, 4th Ed., John Wiley and Sons, Inc, 2006

E. WEB LINKS

- https://www.listendata.com/2018/03/regression-analysis.html
- https://www.coursera.org/learn/linear-regression-business-statistics

3. SPECIFIC LEARNING OUTCOMES(SLO)

Unit/ Section	Course Content	Learning outcomes	Blooms Taxonomy Level of Transaction
I	**Descriptive Statistics**		
1.1	Introduction to Statistics	Identify the types of Measurement	K3
1.2	Organizing Data Using Tables and Graphs	Construct a grouped frequency table.	K3
1.3	Measures of Central Tendency: Mode – Median – Mean	Determine mean, median, mode.	K5
1.4	Measures of Variability: Variability	Analyze the common measures of variability.	K5
1.5	Range	Determine the range on scores.	K5
1.6	Interquartile Range	Determine the interquartile range.	K5
1.7	Standard Deviation	Compare population standard	K4

		deviation and sample standard deviation.	
II	**Inferential Statistics – I**		
	Sampling Distribution of Means		
2.1	Sampling Distribution	Recall Sampling distribution of mean.	K2
2.2	Central Limit Theorem	Discuss the central Limit Theorem.	K6
	Hypothesis Testing		
2.3	Hypothesis Testing Steps	Apply hypothesis testing on given data.	K3
2.4	Effect Size for a Z-Test	Outline the effect size for Z-Test	K3
2.5	Assumptions	Recall the assumptions for the *z*-test	K1
2.6	Errors	List two types of error that are possible on decisions.	K1
2.7	Power	Analyze the power of Statistical test	K4
	One-Sample t Test		
2.8	t-Statistics	Recall t-Statistics	K1
2.9	t- Distributions	Illustrate t-Distributions	K2
2.10	One-Sample t Test	Apply One Sample T Test to find sample is representative of a population	K3
2.11	Effect Size	Determine the effect size for One sample T Test	K3
2.12	Assumptions	Recall the assumptions for One Sample T Test	K1
	Two-Sample t Test: Independent Samples Design		
2.13	Calculations	Recall Standard Error of Difference for Two Sample T test of independent Sample.	K1
2.14	Hypothesis Testing	Make use of Two-sample T Test in hypotheses testing for independent sample	K3
2.15	Effect Size	Determine the effect size for Two Sample T test of independent Sample.	K5
2.16	Assumptions	List the Assumptions of the independent samples *t* test	K1
	Two-Sample t Test: Related Samples Design		
2.17	Calculations	Recall Standard Error of Difference for Two Sample T test of related Sample.	K1
2.18	Hypothesis Testing	Make use of Two-sample T Test in hypotheses testing for related samples.	K3
2.19	Effect Size	Determine the effect size for Two Sample T test of related Sample.	K5
2.20	Assumptions	List assumptions of the repeated measures for Two Sample T test of related Sample.	K2
III	**Inferential Statistics – II**		
	Confidence Interval versus Point Estimation		
3.1	Point Estimates	Recall point estimates for sample T tests.	K1

3.2	Confidence Intervals	Outline confidence interval for sample t tests.	K2
3.3	One Sample t- Test	Apply confidence interval for one sample t test.	K3
3.4	Two-Sample t Test: Independent Samples Design	Apply confidence interval for two sample test of independent samples.	K3
3.5	Repeated Measure Design	Identify confidence interval for two sample test of Repeated Measure Design.	K3
3.6	Degree of Confidence Vs. Degree of Specificity	Relate degree of confidence to degree of specificity	K2
	One-Way Analysis of Variance		
3.7	Introduction	Define one-way analysis of variance.	K1
3.8	Variance	List the types of total variance.	K2
3.9	F- statistics	Illustrate F-Statistics.	K2
3.10	Hypothesis Testing with F-Statistic	Apply hypothesis testing with F-Statistic on given scenario.	K3
3.11	F- Distribution Table	List the features of F-Distribution table	K2
3.12	Notations for ANOVA	Make use of the notions for analysis of variance.	K3
3.13	Calculations	Explain the calculations needed for conducting an analysis of variance	K2
3.14	Hypothesis Testing	Apply hypothesis protocol for ANOVA	K3
3.15	Effect Size	Recall the popular measure of effect size for ANOVA	K1
3.16	Assumptions	List the assumption for ANOVA	K1
	Chi-Square		
3.17	Chi-Square	Recall Chi-Square.	K1
3.18	Chi-Square Statistic	Recall Chi-Square Statistic	K1
3.19	Assumptions	List the assumptions of Chi-Square	K1
3.20	Goodness of Fit - Goodness of Fit for Known Proportions	Apply goodness of fit for known proportions.	K3
3.21	Goodness of Fit for No Preference	Apply goodness of fit for no preference.	K3
3.22	Test of Independence	Make use of Chi-Square in testing for independence of variables	K3
	Nonparametric Statistics for Ordinal Data		
3.23	Mann-Whitney U Test	Apply Mann-Whitney U Test on given problem.	K3
3.24	Kruskal-Wallis H Test	Apply Kruskal-Wallis H Test on given problem	K5
	Correlation		
3.25	Introduction	Contrast positive correlation with negative correlation	K2
3.26	Scatter Plot	Illustrate Scatter Plots.	K3
3.27	Pearson Product Moment Correlation	Examine linear relationships between variables using Pearson Product Moment Correlation.	K4
3.28	Hypothesis Testing	Apply hypotheses testing with the relationship between two variables	K5

		in a population	
3.29	Coefficients of Determination and Non determination	Illustrate Coefficients of Determination and Non determination	K3
3.30	Interpretation and Uses of The Pearson Correlation.	Explain the uses of the Pearson Correlation.	K5
IV	**Regression Analysis – I**		
4.1	Regression Model	Classify the types of regression model.	K4
4.2	Goals of Regression Analysis	List the goals of regression analysis	K1
4.3	Statistical Computing in Regression Analysis	outline statistical softwares that have been developed to make the regression analysis	K2
4.4	Simple Linear Regression	Show that the least squares estimator b*1* is an unbiased estimate of β	K3
4.5	Multiple Linear Regression	Discuss Least Squares Estimates of the Multiple Regression Pa-Rameters.	K2
4.6	Logistic Regression	Evaluate logistic regression model.	K6
4.7	Poisson Regression	Evaluate Poisson regression model.	K6
V	**Regression Analysis – II**		
	Detection of Outliers and Influential Observations		
5.1	Detection of Outliers in Multiple Linear Regression	Define standardized residual.	K1
5.2	Detection of Influential Observations in Multiple Linear Regression	Explain Detecting influential observations in multiple linear regression.	K2
5.3	Test for Mean-shift Outliers	Apply mean-shift outliers to calculate outliers when *i*th observation is suspected to be an outlier	K3
5.4	Graphical Display of Regression Diagnosis	Demonstrate the methods of graphical display of regression diagnosis.	K3
	Model Selection		
5.5	Effect of Underfitting and Overfitting	Justify that over fitting or under fitting affects the generalization ability of a model to new observations.	K5
5.6	All Possible Regressions	Apply all possible regression on data sets .	K3
5.7	Stepwise Selection	Apply stepwise selection algorithms on data set.	K3
	Model Diagnostics		
5.8	Test Heteroscedasticity	Evaluate White's test for Heteroscedasticity.	K6
5.9	Detection of Regression Functional Form	Calculate the Nonlinear Effect via Additive Models.	K5

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**M**			**M**	**L**	**M**			**M**				
CO2	**M**	**H**				**M**	**L**			**H**			**H**
CO3	**M**					**M**			**M**				
CO4	**M**		**M**						**H**	**H**			**M**
CO5	**M**	**H**	**M**	**M**		**H**			**H**	**H**		**M**	**M**
CO6	**M**	**H**		**M**		**H**							

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. A. Santhanasamy

CORE PRACTICAL I: PROBLEM SOLVING USING PYTHON AND R LAB			
Semester	I	Hours/Week	5
Course Code	P20DS1P1	Credits	3

1. COURSE OUTCOMES

S.No.	Course Outcomes	Level	Activity
CO1	Write simple Python programs using Python data structures	K6	1-5
CO2	Manipulate files using Python	K6	6
CO3	Develop object oriented programs in Python	K6	7-8
CO4	Access internet and database data	K6	9-12
CO5	Write R programs for data visualization	K6	13
CO6	Creating dashboards using Tableau	K6	14-15

2. SYLLABUS

Develop applications that will demonstrate the following Python and R programming features

- Functions
- String processing
- List processing
- Dictionaries
- Tuples
- File processing
- Regular Expressions
- OOP
- Retrieving webpages from web
- Data visualization in Matplotlib, Seaborn and R
- Database programming
- Concurrent programming

3. SPECIFIC LEARNING OUTCOMES (SLO)

Exercise	Topic	Learning outcomes	Level
1	Python Basics and Conditions	Applying conditional statements to programs	K6
2	Python Loops	Applying loops to Python Programs	K6
3	Python Functions and Modules	Creating user defined functions and using modules	K6
4	Python String Processing	Manipulating strings in Python	K6
5	List Processing in Python	Using lists in programs	K6
6	Python File Processing	Reading and writing files	K6
7	Python Regular Expressions	Applying pattern matching to strings	K6
8	Object Oriented Bank in Python	Creating Python classes and objects	K6
9	Functional Programming	Using map, filter and reduce	K6

		functions in programs	
10	Retrieving Data from Web and Parsing	Retrieving data from webpages using urllib	K6
11	Database Programming Using Sqlite3	Extracting data from SQL databases using python libraries	K6
12	2D and 3D Data Visualization Using Seaborn	Creating data visualizations using seaborn library	K6
13	Animated Data Visualization Using R	Creating interactive visualizations using R libraries	K6
14	Dashboard Visualization Using Tableau	Creating dashboards and repots using Tableau	K6
15	Concurrent Programming in Python	Creating concurrent programs for multiprocessing	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1		H			H	H	H	H				H	
CO2		H			H	H	H	H		H	H	H	H
CO3		H			H	H	H	H		H		H	H
CO4		H			H	H	H	H		H	H	H	
CO5		H			H	H	H	H				H	H
CO6		H			H	H	H	H					

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Prof. K. Jemimah

CORE PRACTICAL II: NoSQL DATABASE MANAGEMENT LAB			
Semester	I	Hours/Week	5
Course Code	P20DS1P2	Credits	3

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Exercise Covered
CO1	Build a Table using SQL Queries and perform the basic operation	K6	1,2
CO2	Construct a SQL queries to evaluate various operators	K6	3,4
CO3	Evaluate the result using Subquery and Join techniques	K6	5,6,7,8
CO4	Assess the basic Queries in NoSQL using Mango DB	K5	9,10,11
CO5	Review the SQL, NoSQL and Neo 4J Graph data base	K5	12
CO6	Design a Graph database for Movie and Flight Data	K6	13,14

2. SYLLABUS

S.No	List of Exercises
1	Designing and Querying My Restaurant Database
2	India Weather Analytics Using Historical Data Part-I
3	India Weather Analytics Using Historical Data Part-II
4	Retail Sales Analytics Part-I
5	Retail Sales Analytics Part-II
6	Retail Sales Analytics Part-III
7	University Course Enrolment Data Analytics
8	Retail Sales Analytics Part-IV
9	Student Information System Design using MongoDB Part-I
10	Student Information System Design using MongoDB Part-II
11	Ecommerce Product CatLog Design Using MongoDB
12	Neo4J Play Ground Exercise
13	Designing Movie Graph Database using Neo4J
14	Designing Flight Graph Database Using Neo4J

3. SPECIFIC LEARNING OUTCOMES (SLO)

Exercise	Course Content	Learning outcomes	Level
1	Designing and Querying My Restaurant Database	Create a new table, insert tuples satisfying the constraints and perform queryprocessing.	K6
2	India Weather Analytics Using Historical Data Part-I	Evaluate the selection, filtering and aggregate functions to analyse the historical data of India Weather Information	K5
3	India Weather Analytics Using Historical Data Part-II	Determine further India Weather Dataset with additional query operators	K5

		such as GROUPBY, HAVING and ORDERBY	
4	Retail Sales Analytics Part-I	Create analytics on retail sales of a particular enterprise	K6
5	Retail Sales Analytics Part-II	Assess multiple tables in retail sales dataset	K5
6	Retail Sales Analytics Part-III	Compile Group function and Aggregate function in retail dataset	K6
7	University Course Enrolment Data Analytics	Design document in NoSQL for university course enrollment	K6
8	Retail Sales Analytics Part-IV	Develop queries for subquery	K6
9	Student Information System Design using MongoDB Part-I	Create student information system design using mongo DB	K6
10	Student Information System Design using MongoDB Part-II	Analyze various operators in mango DB	K4
11	Ecommerce Product CatLog Design Using MongoDB	Create a Collection in mongo DB for product catalog for heterogenous data	K6
12	Neo4J Play Ground Exercise	Outline the Neo4J for movie graph dataset	K2
13	Designing Movie Graph Database using Neo4J	Create a graph database for Actors and Movies in which actors played roles. You will write queries in Cypher and find answers to all queries.	K6
14	Designing Flight Graph Database Using Neo4J	Develop graph data base for Flight Transportation system	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	H	M	L	M	M	M	L	-	-	H	M	M	-
CO2	H	M	H	M	M	H	M	M	L	H	H	-	M
CO3	H	M	M	L	H	M	-	-	L	M	M	M	-
CO4	H	L	H	H	H	H	M	M	L	H	M	M	H
CO5	H	M	L	H	H	L	L	M	M	H	H	H	M
CO6	H	M	M	L	H	L	L	L	M	H	H	-	-

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Practical Components): Closed Book
2. Cooperative Learning Report, Assignment, Group Discussion, project Report, Field Visit Report, Seminar.
3. Pre/Post Test, Viva, Report for each Exercise.
4. Lab Model Examination & End Semester Practical Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. M. Lovelin Pon Felciah

CORE IV: TIME SERIES ANALYSIS AND FORECASTING			
Semester	II	Hours/Week	4
Course Code	P20DS204	Credits	4

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

S.No.	Course Outcomes	Level	Unit
CO1	Solve the stationarity, trending and detrending of time series data	K6	1
CO2	Assess the features of the ARMA Models and estimation techniques	K5	2
CO3	Explain the ARIMA models and SARMA Models	K6	3
CO4	Summarize the characteristics of Spectral behaviour and periodic behaviour of the time series	K6	4
CO5	Compile the behaviour of smoothing in DLMS	K6	5
CO6	Design the Timeseries models using R for different time series data	K6	All

2. A. SYLLABUS

UNIT I - BASIS TIME SERIES MODELS 12 HOURS

Examples of Nature of Time series data – Time series statistical models – Measures of dependence - Stationary. Time series regression – Detrending and differencing – Smoothing a time series

UNIT II - AR MODELS, FORECASTING AND ESTIMATION 12 HOURS

Auto Regressive models – Moving Average models - ARMA models – Auto Correlation Function - Partial Auto Correlation Function – Forecasting algorithms – **Estimation:** Yule-Walker, Method of moments, MLE and LSE

Unit III - ARMA AND GARMA MODELS 12 HOURS

Basics of ARIMA models: random models with drift, Steps to fitting ARMA model – **Multiplicative Seasonal ARIMA models:** Mixed, SARMA – Generalized Auto Regressive Conditionally Heteroscedastic (GARCH) models

UNIT IV - SPECTRAL ANALYSIS 12 HOURS

Cyclical Behaviour and Periodicity: concepts, Periodic Series, Star Magnitude - **The Spectral Density:** Periodic stationary process–Periodogram: Spectral analysis as ANOVA, Principal Component Analysis

UNIT V - STATE SPACE MODELS 12 HOURS

Dynamic Linear Models – Examples of DLMs – Filtering DLM – **Smoothing DLM:** Kalman, Lag One covariance – Forecasting DLM – Maximum Likelihood Estimator for DLMs

B. TOPICS FOR SELF -STUDY

S.No.	Topics	Web Links
1	Forecasting hierarchical or grouped time series	**https://otexts.com/fpp2/hierarchical.html**
2	Autoregression Models for Time Series Forecasting With Python	**https://machinelearningmastery.com/\autoregression-models-time-series-forecasting-python/**
3	Time Series ARIMA Model using R	**https://sites.google.com/site/econometricsacademy/econometrics-models/time-series-arima-models**
4	Simple Exponential Smoothing for Time Series	**https://towardsdatascience.com/simple-exponential-smoothing-749fc5631bed**

	Forecasting	

C. TEXT BOOKS

1. Shumway and Stoffer. Time Series Analysis and its applications, with examples in R. 4ed, Springer. 2016.

D. REFERENCES BOOKS

1. Brockwell& Davis. Introduction to Time Series and Forecasting, 3rd edition, Springer. 2016
2. Cryer& Chan. Time Series Analysis with Applications in R, Springer. 2008
3. Prado & West. Time Series: Modeling, Computation, and Inference Chapman & Hall. 2010
4. Petris, Petrone, Campagnoli. Dynamic Linear Models with R, Springer. 2009
5. Ruppert& Matteson. Statistics and Data Analysis for Financial Engineering with R examples, 2ed, Springer. 2016

E. WEB LINKS

1. https://machinelearningmastery.com/autoregression-models-time-series-forecasting-python/
2. **https://sites.google.com/site/econometricsacademy/econometrics-models/time-series-arima-models**

3. SPECIFIC LEARNING OUTCOMES

Unit/ Section	Course Content	Learning outcomes	Level
I	**BASIS TIME SERIES MODELS**		
1.1	Nature of Time series data	Analyze the different types of Times series and its characteristics.	K4
1.2	Time series statistical models	Compare the different statistical model of times series data.	K6
1.3	Measures of dependence	Assess the measure of dependence for different statistical model of time series data.	K5
1.4	Stationary	Evaluate the stationarity property for time series models.	K5
1.5	Time series regression	Investigate the regression for time series data	K4
1.6	Detrending and differencing	1.Formulate the detrending model for Time series data. 2. Devise the differencing method for time series data.	K6
1.7	Smoothing a time series	Construct the smoothing filters for time series models	K6
II	**AR MODELS, FORECASTING AND ESTIMATION**		
2.1	Auto Regressive models	Compare the features of AR Models	K4
2.2.	Moving Average models	Analyze the characteristics of MA Models	K5
2.3	ARMA models	Summarize the working methods of ARMA Models	K6
2.4	Auto Correlation Function - Partial Auto Correlation Function	Explain the role of Autocorrelation and partial auto correlation function for time series	K6
2.5	Forecasting algorithms	Compile the forecasting algorithm for time series data.	K6
2.6	**Estimation:** Yule-Walker, Method of moments, MLE and LSE	Specify the features of different estimation algorithms of time series data.	K6
III	**ARMA AND GARMA MODELS**		
3.1	Basics of ARIMA models:	Explain the basics of ARIMA models of Time	K3

	random models with drift, Steps to fitting ARMA model	series data	
3.2	**Multiplicative Seasonal ARIMA models:** Mixed, SARMA	Compile the features of the SARIMA and Multiplicative SARIMA model for time series data.	K6
3.3	Generalized Auto Regressive Conditionally Heteroscedastic (GARCH) models	Evaluate the characteristics of GARCH Model.	K5
IV	**SPECTRAL ANALYSIS**		
4.1	Cyclical Behaviour and Periodicity: concepts, Periodic Series, Star Magnitude	1. Examine the concepts of periodicity 2. Evaluate the cyclical behaviour of the time series. 3. Assess the properties of Star Magnitude	K3 K6 K5
4.2	Periodic stationary process	Discriminate the periodic stationary process over the stationary process	K5
4.3	Periodogram	Outline the periodogram of the time series data	K6
4.4	Spectral analysis as ANOVA	Construct the ANOVA for the spectral analysis of Time series data	K6
4.5	Spectral analysis as Principal Component Analysis	Develop the PCA for the spectral analysis of time series data.	K6
V	**STATE SPACE MODELS**		
5.1	Dynamic Linear Models-Examples of DLMs	Illustrate the Dynamic Linear Models (DLM)	K3
5.2	Filtering DLM	Create the filtering of DLM for the time series data	K6
5.3	**Smoothing DLM:** Kalman Filter	Evaluate the role of Kalman Filter in smoothing.	K5
5.4	Lag One covariance	Construct the smoothing filter using Lag One covariance	K6
5.5	Forecasting DLM	Design the forecasting algorithm using Kalman filter for DLMs	K6
5.6	Maximum Likelihood Estimator for DLMs	Evaluate the features of the MLE for DLMs	K6

4. MAPPING

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**M**	**-**
CO2	**M**	**H**	**M**	**H**	**M**	**M**	**-**	**M**	**M**	**M**	**M**	**-**	**-**
CO3	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**H**	**H**	**M**	**-**	**-**
CO4	**H**	**H**	**L**	**H**	**M**	**H**	**-**	**M**	**H**	**H**	**H**	**-**	
CO5	**H**	**M**	**H**	**M**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**-**	**-**
CO6	**H**	**H**	**H**	**M**	**M**	**H**	**H**	**-**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Peer Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course evaluation survey
2. Faculty feedback about the course.

Name of the Course Coordinator: Dr. P. S. Eliahim Jeevaraj

CORE V: DATA AND VISUAL ANALYTICS			
Semester	II	Hours/Week	4
Course Code	P20DS205	Credits	4

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

S.No.	Course Outcomes	Level	Unit
CO1	Test the NumPy functions for array processing	K6	I
CO2	Create time series plots using the Date and Time classes	K6	II
CO3	Interpret the Plotting of the dataset and time series	K6	III
CO4:	Perform data aggregation and group operations	K6	IV
CO5	Create and use Series and Data Frames for data wrangling	K6	V
CO6	Create various plots using Matplotlib and Seaborn	K6	

2. A. SYLLABUS

Unit-1. NumPy and Pandas Basics

Why Python for Data analysis – Essential Python libraries – ndarray – Universal functions – Data processing using arrays – File I/O with arrays – Random number generation – Series, Data Frames – Indexing, re-indexing, sorting, ranking – Summarizing descriptive statistics – Handling missing data – Hierarchical indexing

Unit-2. Data Loading and Wrangling

Data Loading: reading and storing data in text format, binary format – Data Wrangling: Combining and merging data sets – Reshaping – Pivoting – Data transformation – String manipulation

Unit-3. Plotting and Visualization using Matplotlib

Figures – Subplots – Colors – Ticks – Label – Legends – Annotation – Saving plots to file – Plots: Line, Bar, Histogram, Density Plots – Scatter Plots

Unit-4. Data Aggregation and Group Operations

Iterating over groups – Selecting columns – Grouping with Series and functions – Data aggregation: Column wise aggregation, returning aggregated data – General-Split-Apply-Combine – Quantile and bucket analysis – Pivot table and cross tabulation

Unit-5. Time Series

Date and Time – Time Series – Date Range, Frequencies and Shifting – Periods and period arithmetic – Resampling and frequency conversion – Time Series Plotting

B. TOPICS FOR SELF-STUDY

S.No	Topic Title	Web Link
1	Facebook Data Analysis	https://www.kaggle.com/sd2beatles/deep-analysis-sql-and-statistical-test-included
2	Clothing Fit Dataset for Size Recommendation	https://www.kaggle.com/agrawaladitya/step-by-step-data-preprocessing-eda
3	UCI Adult dataset	https://www.kaggle.com/kashnitsky/a1-demo-pandas-and-uci-adult-dataset
4	Wikipedia Time series analysis	https://www.kaggle.com/kashnitsky/a9-demo-time-series-analysis

C. TEXT BOOKS

1. Wes. Mc Kinney, "*Python for Data Analysis"*, 2nd Edition, Schroff Publishers, 2013. ISBN 9789352136414

D. REFERENCES

1. Cyrille Rossant. "Learning I*Python for interactive Computing and data visualization"*, First edition [Packt]
2. Jake VanderPlas ,Python Data Science Handbook - Essential Tools for Working with Data, O'Reily, 2016
3. Zhang.Y ,An Introduction to Python and Computer Programming, Springer Publications,2016

E. WEB LINKS

- https://www.kaggle.com/agrawaladitya/step-by-step-data-preprocessing-eda
- https://www.kaggle.com/kashnitsky/a1-demo-pandas-and-uci-adult-dataset

3. SPECIFIC LEARNING OUTCOMES

Unit	Topic	Topic Learning Outcome	Level
I	**NumPy and Pandas Basics**		
1.1	Python for Data analysis	Why Dataset analysis? And Why and What Python	K1
1.2	Essential Python libraries	Illustrate Python libraries which is used for Data Science	K2
1.3	Universal functions	Utilize **universal function** (or ufunc for short) is a **function** that operates on ndarrays in an element-by-element fashion, supporting array broadcasting, type casting, and several other standard features	K3
1.4	Data processing using arrays	Discover data processing tasks without writing complex loops	K4
1.5	File I/O with arrays	Test for file read and write using array	K4
1.6	Random number generation	Analyze pseudo-random number generator for various distributions.	K4
1.7	Series, Data Frames	Appraise the single list with index. Examine a dataframe using collection of series that can be used to analyse the data	K5,K4
1.8	Indexing, re-indexing, sorting, ranking	Find NA/NaN in locations having no value in the previous index. Plan to index and reindex using Indexing, reindexing, sorting, ranking.	K1, K3
1.9	Summarizing descriptive statistics	Examine summarizing and organizing the data so they can be easily understood.	K4
1.10	Handling missing data	Determine missing values for a number of reasons such as observations that were not recorded and data corruption.	K5
1.11	Hierarchical indexing	Discover to incorporate multiple index levels within a single index.	K4

II	Data Loading and Wrangling		
2.1	Data Loading: reading and storing data in text format, binary format	Motive the ability to read, manipulate, and write data to and from CSV files using Python is a key skill to master for any data scientist or business analysis.	K4
2.2	Data Wrangling: Combining and merging data sets	Measure the processing of data in various formats like - merging, grouping, concatenating etc. for the purpose of analysing or getting them ready to be used with another set of data.	K5
2.3	Reshaping	Examine 'reshape()' function, that takes a single argument that specifies the new shape of the array.	K4
2.4	Pivoting	Analyze Pivoting for reshape a DataFrame by column/index values.	K4
2.5	Data transformation	how we can combine data from different sources into a unified dataframe	K1
2.6	String manipulation	List the manipulation of string like concatenation, isupper(), join(), lower(), etc.	K4
III	**Plotting and Visualization using Matplotlib**		
3.1	Figures	Create graph using figure() in Python.	K6
3.2	Subplots	Create subplots by the use of subplot() function in pyplot module.	K6
3.3	Colors	Discuss colouring plot by python colour code.	K6
3.4	Ticks	Create Ticks value to show specific points on the coordinate axis.	K6
3.5	Label	Assess plot axis label	K5
3.6	Legends	Analyze legend for describing area and elements of the graph	K4
3.7	Annotation	Utilize annotate() function to draw an arrow connecting two points on the plot.	K3
3.8	Saving plots to file	Recommend savefig() function to save plot in to file	K5
3.9	Plots: Line, Bar, Histogram, Density Plots	List basic graphics primitives to draw plot	K4
3.10	Scatter Plots	Evaluate the data as a collection of points.	K5
IV	**Data Aggregation and Group Operations**		
4.1	Iterating over groups	Discover data cluster using Iterating over groups	K4
4.2	Selecting columns	Select multiple columns using loc, iloc, etc...	K3
4.3	Grouping with Series and functions	Create series group using groupby() function	K6
4.4	Data aggregation: Column wise	Examine summarization using computing aggregations like sum() , mean() , median() , min() , and max() , in	K4

	aggregation	which a single number gives insight into the nature of a potentially large dataset.	
4.5	Data aggregation: returning aggregated data	Plan statistical method for data aggregation.	K3
4.6	General-Split-Apply-Combine	Create group by three step Split-Apply-Combine.	K6
4.7	Quantile and bucket analysis	Examine quantile() function to get values at the given quantile over requested axis.	K4
4.8	Pivot table and cross tabulation	Create cross-tabulation table for show the frequency with which certain groups of data appear	K6
V	**Time Series**		
5.1	Date and Time	Examine date and time parameter for time series.	K4
5.2	Time Series	Discover statistical descriptive by statistical tests and several linear model classes: autoregressive, AR, autoregressive moving- average, ARMA, and vector autoregressive models VAR.	K4
5.3	Date Range	Model a large range of dates for various offsets are pre-computed	K3
5.4	Frequencies and Shifting	Evaluate percentage change from sample to sample.	K5
5.5	Periods and period arithmetic	Dissect the time elapsed between two values of the same magnitude.	K4
5.6	Resampling and frequency conversion	Survey the Convenience method for frequency conversion and resampling of time series	K4
5.7	Time Series Plotting	Visualize trends in counts or numerical values over time.	K4

4. MAPPING

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**H**	**-**
CO2	**M**	**H**	**M**	**H**	**M**	**M**	**-**	**M**	**M**	**M**	**M**	**-**	**H**
CO3	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**H**	**H**	**M**	**H**	**H**
CO4	**H**	**H**	**L**	**H**	**M**	**H**	**-**	**M**	**H**	**H**	**H**	**-**	**H**
CO5	**H**	**M**	**H**	**M**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**H**	**H**
CO6	**H**	**H**	**H**	**M**	**M**	**H**	**H**	**-**	**M**	**H**	**H**	**H**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Peer Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).

4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course evaluation survey
2. Faculty feedback about the course.

Name of the Course Coordinator: Dr. B. Karthikeyan

CORE VI: PRACTICAL MACHINE LEARNING			
Semester	II	Hours/Week	4
Course Code	P19DS206	Credits	4

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

CO#	Course Outcome	Level	Unit
CO1	Perceive the Types of ML and develop Perceptron model	K6	I
CO2	Develop a supervised ML model for the given business problem	K5	II
CO3	Assess the pre-processing methods and reduce dimensions of data	K6	III
CO4	Evaluate the training and the testing of the designed ML model	K6	IV
CO5	Develop an unsupervised ML model for the given business problem	K5	V
CO6	Deploy machine learning models into production environment	K6	All

2. A. SYLLABUS

Unit-1. ML Basics and Perceptron

Three types Machine Learning – Three steps of ML process – Perceptron neural network – Adaline neural network – Stochastic gradient descent neural network

Unit-2. Supervised Learning classifiers

Logistic regression – Support vector machines – Kernel SVM – Decision Trees – K-Nearest Neighbour classifier – Random Forest – Linear Regression–Sentiment Analysis of Movie Reviews using Logistic Regression - Developing a web application with Flask

Unit-3. Pre-processing and Dimensionality Reduction

Pre-processing: Missing data, categorical data, feature scaling, feature selection. Dimensionality reduction: Principal Component Analysis, Linear Discriminant Analysis, Kernel PCA

Unit-4. Model evaluation

Pipelines - K-fold cross validation - Grid search - Confusion matrix, Precision, Recall, ROC curves, Scoring metrics –Majority vote classifier – Bagging, Bootstrapping, Adaptive Boosting

Unit-5. Unsupervised Learning classifiers and Multilayer NN

K-Means, K-Means++, Finding optimal no. of classifiers - Agglomerative Hierarchical clustering, Density based clustering -Multilayer Neural Network: Feed forward, Back Propagation Training, Multilayer Perceptron

B. TOPICS FOR SELF STUDY

SNo	Topic Title	Web Link
1	Kaggle Machine Learning	https://www.kaggle.com/learn/overview
2	IBM ML with Python: Practical introduction	https://www.edx.org/course/machine-learning-with-python-a-practical-introduct
3	Predictive Analytics using Machine Learning	https://www.edx.org/course/predictive-analytics-using-machine-learning
4	Google AI	https://ai.google/education/

C. TEXT BOOK(S)

1. Sebastian Raschka, “Python Machine Learning”, First Edition, [PACKT], 2015.

D. REFERENCES BOOK(S)

1. Andreas C Muller and Sarah Guido, Introduction to Machine Learning with Python, Shroff Publishers, ISBN 978935213451

2. Joel Grus, “Data Science from Scratch”, First Edition, O’Reilly,2015
3. Gavin Hackeling, “Mastering machine learning with scikit-learn”, First Edition, [PACKT] , 2014

E. WEB LINKS

- https://www.edx.org/course/predictive-analytics-using-machine-learning
- https://www.edx.org/course/machine-learning-with-python-a-practical-introduct

3. SPECIFIC LEARNING OUTCOMES

Unit	Topic	Topic Learning Outcomes	Level
I	**Machine Learning Basics and Perceptron**		
1.1	Three types Machine Learning	Select a machine learning model, given business, scientific and societal use cases.	K2
1.2	Three steps of ML process	Explain machine learning steps based on the given use cases.	K2
1.3	Perceptron neural network	Draw Perceptron Neural Network for AND, OR and NOT logic gate operations. Create Perceptron in *sklearn* for a simple dataset that contains 4 samples for 2 numerical input features and corresponding y values, perform training and testing for an unknown sample.	K3 K5
1.4	Adaline neural network	Differentiate Adaline from Perceptron neural network.	K2
1.5	Stochastic gradient descent neural network	Differentiate Perceptron, Adaline and Stochastic Gradient Descent Neural Networks. Create a dataset, perform training, testing and print error rates for SGD Neural Network using *sklearn*, for the given use case.	K2 K6
II	**Supervised Learning Classifiers**		
2.1	Types of Supervised ML and ML pipeline	Identify the type of supervised ML, given use cases. Identify the steps of ML pipeline for classification and regression problems.	K2 K2
2.2	Classes and methods of ML models available in *sklearn* package	Import and instantiate ML models using sklearn. Call methods and properties of ML models in sklearn.	K3 K3
2.3	Linear Regression in Scikit Learn	Identify input features and target from dataset, preprocess data, split dataset for training & testing, create LinearRegression model using *sklearn*, perform training and testing and print MSE, SSE and R2 errors, for the given regression problem. Create Ridge Regression model in sklearn, for the given use case, by following ML pipeline steps. Create LASSO Regression model in sklearn, for the given use case, by following ML pipeline steps. Create Polynomial Regression model in sklearn to represent non-linearity assumption, for the given use case, by following ML pipeline steps.	K6 K6 K6 K6
2.4	Perceptron using Scikit Learn	Identify input features and target from dataset, preprocess data, split dataset for training & testing, create Perceptron using *sklearn*, perform training and testing and print classification accuracy for the given classification	K6

		problem.	
2.5	Logistic regression	Compare LR against Perceptron based on the given use cases. Compute Sigmoid activation function value given weights and values of an input sample. Explain how LR can be used to predict probability values, with use cases. Choose if the given ML model suffers from Overfitting or Underfitting; Also select if the model has high/low variance or high/low bias Explain how regularization solves overfitting issue of a ML model. Create LR model in *sklearn* by following ML system design pipeline and compare against Perceptron and select the best model, for a given use case.	K2 K3 K2 K4 K3 K5
2.6	Support vector machines and Kernel SVM	Compare SVM against LR and Perceptron. Interpret the parameters and their values, given *sklearn* syntax of SVC classifier, Interpret parameters and their values, Ggiven *sklearn* syntax of SVC for Kernel SVM Explain the functions of Linear and RBF kernels. Create SVM model in *sklearn* by following ML system design pipeline and select the best model among Perceptron, LR and SVM models, for the given use case.	K2 K2 K2 K2 K6
2.7	Decision Trees	Find the best split of DT node using Entropy value. Find the best split of DT node using Gini Index value. Create a DT manually using ID3 algorithm for the specified depth, given a dataset. Create a DT manually using C4.5 algorithm for the specified depth, given a dataset. Create syntax for *sklearn* DecisionTreeClassifier class, given parameter values. Create syntax for *sklearn* DecisionTreeRegressor class, given parameter values, Create Decision Tree model in *sklearn* by following ML system design pipeline, compare its performance against other ML models and select the best model, for a given use case.	K4 K4 K4 K4 K4 K5

2.8	Random Forest	Select Random Forest or Decision Tree approach based on the business objective. Create manually CART decision tree for the given use case. Create manually Random Forest using CART trees for the given use case. Create syntax for *sklearn* RandomForestClassifier class, given parameter values. Create syntax for *sklearn* RandomForestRegressor class, given parameter values, Create Random Forest classification model in *sklearn* by following ML system design pipeline, compare its performance against Decision Tree classification model and select the best model, for a given use case. Create Random Forest regression model in *sklearn* by following ML system design pipeline, compare its performance against Decision Tree regression model and select the best model, For a given use case.	K5 K5 K5 K4 K4 K5 K5
2.9	K-Nearest Neighbour classifier	Classify ML methods into parametric and non-parametric categories. Classify ML methods into Easy Learners and Lazy Learners. Select the best value for **k** for KNN classifier. Create a KNN model syntax using *sklearn,* given values for number of neighbors and distance metric. For the given use case, build dataset, create KNN model and evaluate its performance.	K2 K2 K3 K5 K5
2.10	Sentiment Analysis of Movie Reviews using Logistic Regression	Create feature vectors manually by computing term frequency, inverse document frequency and TF-IDF values, for the given use case. Create feature vectors automatically with TfidfVectorizer class for the given use case. Create a sentiment analysis system using Logistic Regression model for the movie reviews dataset.	K4 K4 K6
2.11	Developing a web application with Flask	Create a simple website and deploy a machine learning model using Flask.	K6
III	**Pre-processing and Dimensionality Reduction Methods**		
3.1	Handling missing data and categorical data	Find missing values and replace with mean / median / mode values for numerical data. Apply LabelEncoder to ordinal attributes in order to represent string values to integers. Apply One Hot Encoder to nominal attributes to represent categorical data.	K3 K4 K4

3.2	Feature scaling and feature selection	Compute normalized values using min max scaling. Compute standardized values using standard scaling. Apply MinMaxScaler and StandardScaler to preprocess data. Compute important features using L2 and L1 regularization methods. Compute important features using Random Forest algorithm.	K3 K3 K5 K5 K4
3.3	Principal Component Analysis	Compute Covariance matrix, Eigen vectors and Eigen values of a given matrix. For *load_digits* dataset from *sklearn*, reduce original dimension (1797x64) into low dimension (1797x7) using PCA algorithm and print its shape. Create PCA model in sklearn for a small N+1 dimensional matrix, print N principle component vectors and the variance each principle components holds (called explained_variance_ratio value in *sklearn*). Create PCA model in sklearn and visualize data in low dimensions using matplotlib for the given use case (For ex. Breast Cancer dataset from *sklearn*).	K4 K5 K5 K5
3.4	Linear Discriminant Analysis	Explain Singular Valued Decomposition with an example. Create LinearDiscriminantAnalysis model for a small N+1 dimensional matrix, print N components and explained_variance_ratio value. Create LinearDiscriminantAnalysis model in sklearn and visualize data in low dimensions using matplotlib for the given use case (For ex. Iris dataset in *sklearn*).	K2 K4 K5
3.5	Kernel PCA	Create a syntax for KernelPCA for the values of the input parameters. Reduce original dimension (1797x64) into low dimension (1797x7) using KernelPCA algorithm and print its shape, for load_digits dataset from sklearn. Create and visualize *make_moons* dataset using PCA and KernelPCA models.	K4 K4 K5
IV	**Model Evaluation**		
4.1	PipelineE	Create Pipeline in sklearn, given the requirements for transformers and estimators. Create a Pipeline for a ML model, perform training and testing and show its performance values, for the given business use case.	K4 K6
4.2	Holdout validation	Divide the dataset for training, validation and testing based on the performance requirement.	K3

4.3	K-fold cross validation	Explain the working of Kfold cross validation, Stratified Kfold CV, Leave one out CV and Shuffle split CV. Create a ML model and compute the CV score (which may be classification accuracy or error) using *sklearn.model_selection*, for the given use case. Create a Pipeline for a ML model and compute the CV classification accuracy or error using *sklearn.model_selection*, for the given use case.	K4 K6 K6
4.4	Grid search	Create a syntax for a GridSearchCV model for the given parameter values. Create and perform GridSearchCV search, for the given use case, with various parameters values for the chosen ML model and select best parameter values; then create that ML model with the best parameter values and show performance results.	K4 K5
4.5	Confusion matrix, Precision and Recall	Create a confusion matric based on the performance values of a ML system. Compute manually precision and recall values of a ML system. For example, given the sequence of predictions of an email spam classifier. Evaluate the performance of a ML model using precision, recall and fscore values in sklearn.	K5 K3 K4
4.6	ROC curve	Plot ROC curve with AUC values for the ML models, compare performances of many ML models using sklearn and give recommendations to business clients.	K5
4.7	Majority vote classifier	Explain the concepts of majority voting classifier. Given a simple dataset of 5 samples with 2 input numerical features for X matrix and output vector y values, create VotingClassifier in sklearn with atleast 2 ML classifiers, such as Logistic Regression and SVC, perform training and testing and verify output y manually. Create a VotingClassifier for the given classification problem and report the performance results.	K2 K6 K5
4.8	Bootstrapping	Explain bootstrapping with random replacement policy with examples.	K2
4.9	Bagging	Create BaggingClassifier model with 10 decision trees, perform training and testing for a simple dataset. Create BaggingClassifier model with Bootstrapping feature with 10 decision trees, perform training and testing for a simple dataset. Create BaggingClassifier model with 10 SVC classifiers, perform training and testing for a simple dataset. Create BaggingClassifier model, perform training and testing for the business use case. Compare its performance against other ML classifiers.	K5 K5 K5 K6

4.10	Adaptive Boosting	Create AdaBoostClassifier with 5 estimators for a simple dataset. Create GradientBoostingClassifier with 10 estimators for a simple dataset. Create XGBoost classifier with 10 estimators for a simple dataset. Create AdaBoostClassifier model, perform training and testing for the business use case. Compare its performance against Gradient boosting and bagging ensemble classifiers.	K5 K5 K5 K6
V	**Unsupervised Learning Classifiers and Multilayer NN**		
5.1	K-Means	Identify and recommend the appropriate ML strategy (which may be classification, regression or clustering approach) for the given business or scientific or societal application. Given a small dataset with 2 numerical features with 4 samples and values for 2 centroids, apply manually KMeans algorithm and predict 2 clusters (You can use Euclidean distance). Create the syntax of KMeans model in sklearn (assume input samples X are already available) and no. of clusters to predict. Also, perform training and testing on X. Print clusters and Cluster Sum of Squared Error, called Inertia. Create KMeans model and predict the optimal number of clusters using Elbow method, given input samples X. Create KMeans model and predict the optimal number of clusters and evaluate the quality of clusters using Silhouette Coefficients, given input samples X. Create and develop KMeans clustering system in *sklearn* for the business use case and provide recommendations to users.	K4 K5 K5 K5 K5 K6
5.2	K-Means++	Create and develop KMeans++ clustering system in *sklearn* for the business use case and provide recommendations to users.	K6
5.3	Agglomerative Hierarchical clustering	Perform manually hierarchical clustering using single linkage and show clusters of students, for a one dimensional data representing marks of students (say, 5 students), Perform manually hierarchical clustering using complete linkage and show clusters of students, for a one dimensional data representing marks of students (say, 5 students). Perform manually hierarchical clustering using single linkage, draw Dendrogram and choose the number of clusters, for a one dimensional data representing marks of students (say, 5 students). Plot Dendrogram for the given input samples and choose the clusters using *scipy* package. Apply agglomerative clustering using *sklearn* package for the given business use case and predict N clusters.	K4 K4 K4 K4 K6
5.4	Density based clustering	Create clusters by applying DBSCAN algorithm in sklearn and visualize clusters, for a given dataset.	K5

5.5	Feed Forward Multilayer Neural Network	Draw a Single layer neural network with input layer and output layer. The dataset represents details of 5 persons. The 3 input features are 'smoking, 'obesity' and 'exercise' with values 0 or 1. The target or output feature is 'diabetic' which can have a value 0 or 1. Draw a Multilayer neural network for XOR operations and differentiate from Perceptron. Draw the architectural diagram of Multilayer Perceptron neural network with bias input nodes, given input samples X and target output y values and the size of the hidden layer. Create syntax for MLPClassifier in sklearn given parameter values. Create MLPClassifier, perform preprocessing, training and testing. Print the performance values of classification metrics. Also print the learnt weight matrix and bias vector values, for the given business use case.	K4 K4 K3 K6
5.6	Back Propagation Training	Compute predicted output and propagate error for one iteration, given a simple Multilayer Perceptron having just one hidden layer with values for input sample, weights and desired output. Design a ML system using MLPClassifier and compare its performance against other classifiers, for the given use case.	K4 K6

4. Mapping

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	H	L		L									
CO2	H	M	L	H								H	H
CO3	H	M		H	H	H						H	H
CO4	H	M		H	H	H						H	
CO5	H					H		H		M			H
CO6	H						H		H		H		

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Peer Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course evaluation survey
2. Faculty feedback about the course.

Name of the Course Coordinator: Dr. K. Rajkumar

Elective II: NATURAL LANGUAGE PROCESSING			
Semester	II	Hours/Week	4
Course Code	P20DS2:2	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Experiment with text pre-processing and classification	K6	I
CO2	Create language models and POS tagging	K6	II
CO3	Evaluate the context free grammars and parse sentences	K6	III
CO4	Validate the meaning of sentences	K6	IV
CO5	Explain the Dependency parsing	K6	IV
CO6	Design NLP applications	K6	V

2. A. SYLLABUS

UNIT I-Text Pre-processing, Classification and Evaluation 12 Hours

Text tokenization, normalization, Lemmatization and Stemming – Minimum Edit Distance between two strings – Bigram, Trigram and NGrams – Perplexity – Smoothing Methods – Naïve Bayes Classifier – Binary Naïve Bayes for Sentiment Analysis and Spam Detection – Precision, Recall, F1-score, Cross Validation - Bootstrapping.

UNIT II- Language Modeling 12 Hours

Term Document Matrix and Word Word Matrix – Cosine Similarity – TFIDF weighting – Positive PMI – Skip Gram with Negative Sampling – Neural Network Unit – Feed Forward Neural Network – Neural Network for Language Model - English word classes for POS Tagging – Penn Treebank POS Tagset – POS Tagging Process – HMM POS Tagger – Viterbi Decoding algorithm for HMM.

UNIT III- Context Free Grammars 12 Hours

Context Free Grammars: Definition – Grammar rules for English – Tree Banks as Grammars – Grammar equivalence and normal forms – Lexical Grammars – Syntactic Parsing: Types of Ambiguities - CKY Parsing Algorithm – Chunking – Probabilistic CFG for Disambiguation and Language Modeling – Probabilistic CKY Parsing Algorithm for PCFGs – Probabilistic Lexicalized CFGs: Collins Parser – Parser Evaluation Methods.

UNIT IV- Dependency Parsing 12 Hours

Dependency Parsing: Dependency Relations from Universal Dependency Set – Dependency Tree – Shift Reduce Parsing – Arc Eager Transition Based Parsing – Edge Factored Parsing model using MST – First Order Logic for meaning representation: elements, variables, quantifiers, lambda notation, inference – Event Representation – Event Time Representation – Representation and Inference using Description Logic

UNIT V- NLP Applications 12 Hours

NLP Applications: Named Entity Recognition – Word Sense Disambiguation - Factiod Question Answering – Rule based and corpus based chatbots – Dialogue State Architecture for chatbots.

B. TOPICS FOR SELF-STUDY:

S.No.	Topics	Web Links
1	Natural language processing with Deep Learning	https://www.youtube.com/watch?v=OQQ-W_63UgQ&list=PL3FW7Lu3i5Jsnh1rnUwq_TcylNr7EkRe6
2	Latent structure models for NLP	https://deep-spin.github.io/tutorial/acl.pdf
3	Chatbots	https://www.analyticssteps.com/blogs/learn-everything-about-machine-learning-chatbots
4	Language Interpretability Tool (LIT)	https://github.com/PAIR-code/lit

C. TEXT BOOK

1. Jurafsky and Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, 3ed. 2020.

D. REFERENCES BOOKS

1. Steven Bird, Ewan Klein, and Edward Loper , Natural Language Processing with Python– Analyzing Text with the Natural Language Toolkit.
2. Indurkhya, Nitin and Fred Damerau, Handbook of Natural Language Processing, 2ed, 2010, Chapman & Hall/CRC.
3. Christopher Manning and HinrichSchutze, Foundations of Statistical Natural Language Processing, MIT Press.

E. WEB LINKS

1. https://london.ac.uk/sites/default/files/study-guides/introduction-to-natural-language-processing.pdf
2. http://www.datascienceassn.org/sites/default/files/Natural%20Language%20Processing%20with%20Python.pdf

3. SPECIFIC LEARNING OUTCOMES (SLO)

Unit/ Section	Course Content	Learning outcomes	Level
I	**Text Pre-processing**		
1.1	Text tokenization, normalization, Lemmatization and Stemming	Apply various methods to pre-process the text	K3
1.2	Minimum Edit Distance between two strings	Compare the distance between the two string	K4
1.3	Bigram, Trigram and NGrams	Evaluate the precision for words using grams	K5
1.4	Perplexity , Smoothing Methods	Choose to evaluate the probability of each test sentence	K3
	Classification and Evaluation		
1.5	Naïve Bayes Classifier	Analyse feature vector in a sentence	K4
1.6	Binary Naïve Bayes for Sentiment Analysis and Spam Detection	Categorize the sentence for detecting spam and to analyse sentiment of words	K4
1.7	Precision, Recall, F1-score, Cross Validation - Bootstrapping.	Evaluate the measure for words	K5
II	**Language Modelling**		
2.1	Term Document Matrix and Word	Illustrate the space vector to	K2

	Word Matrix	represent features	
2.2	Cosine Similarity	Explain the cosine similarity	K2
2.3	Skip Gram with Negative Sampling	Analyse the gram sampling	K4
	Neural Network for Language Model		
2.4	Neural Network Unit – Feed Forward Neural Network	Outline the neural network unit	K2
2.5	Neural Network for Language Model	Apply the model in neural network for language	K3
2.6	English word classes for POS Tagging	Categorize the words for POS tagging	K4
2.7	POS Tagging Process – HMM POS Tagger – Viterbi Decoding algorithm for HMM.	Choose proper tag for context related words	K5
III	**Context Free Grammars**		
3.1	Definition – Grammar rules for English	Define grammar rules	K1
3.2	Tree Banks as Grammars	Illustrate tree grammars	K2
3.3	Grammar equivalence and normal forms	Construct grammar equivalence	K3
3.4	Lexical Grammars – Syntactic Parsing	Examine the parsing grammar	K4
	Types of Ambiguities		
3.5	CKY Parsing Algorithm	Explain the parsing algorithms	K2
3.6	Chunking – Probabilistic CFG for Disambiguation and Language Modelling	Analyze CFG for language modelling	K4
3.7	Probabilistic CKY Parsing Algorithm for PCFGs	Adapt CKY algorithm for parsing	K6
3.8	Probabilistic Lexicalized CFGs: Collins Parser – Parser Evaluation Methods	Evaluate the probability for various parsing methods	K5
IV	**Dependency Parsing**		
4.1	Dependency Relations from Universal Dependency Set	Explain dependency parsing	K2
4.2	Dependency Tree	Construct dependency tree	K3
4.3	Shift Reduce Parsing	Examine the shift reducing parsing	K4
4.4	Arc Eager Transition Based Parsing	Outline the Arc eager based parsing	K2
4.5	Edge Factored Parsing model using MST	Make use of Edge model for MST	K3
4.6	First Order Logic for meaning representation	Explain first order logic	K5
4.7	Representation and Inference using Description Logic	Adapt description logic	K6
V	**NLP Applications**		
5.1	Named Entity Recognition	Classify entities referred according to their type	K4
5.2	Word Sense Disambiguation	Explain the WSD	K5
5.3	Rule based and corpus based chatbot	Analyze the rule-based corpus	K4
5.4	Dialogue State Architecture for chatbots.	Build architecture for chatbots	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**M**	**-**
CO2	**M**	**H**	**M**	**H**	**M**	**M**	**-**	**M**	**M**	**M**	**M**	**-**	**-**
CO3	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**H**	**H**	**M**	**-**	**-**
CO4	**H**	**H**	**L**	**H**	**M**	**H**	**-**	**M**	**H**	**H**	**H**	**-**	
CO5	**H**	**M**	**H**	**M**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**-**	**-**
CO6	**H**	**H**	**H**	**M**	**M**	**H**	**H**	**-**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator : Dr. Janani Selvaraj

ELECTIVE-3: HEALTH CARE DATA ANALYTICS			
Semester	I	Hours/Week	3
Course Code	P19DS2:3	Credits	3

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Explain the Coding system of Electronic Health Records	K5	I
CO2	Develop the algorithms for Biomedical Analysis	K5	I
CO3	Design the Algorithm for Healthcare Data using NLP, SMA	K6	III
CO4	Construct the predictive models for Healthcare Data	K6	IV
CO5	Analyze the role of Analysis in Pervasive Health	K5	V
CO6	Design the Computer-Assisted Medical Image Analysis Systems	K6	V

2. A. SYLLABUS

Unit-1. Introduction
Introduction to Healthcare Data Analytics- Electronic Health Records–Components of EHR-Coding Systems- Benefits of EHR- Barrier to Adopting EHR Challenges-Phenotyping Algorithms.

Unit-2. Analysis
Biomedical Image Analysis- Mining of Sensor Data in Healthcare- Biomedical Signal Analysis-Genomic Data Analysis for Personalized Medicine.

Unit-3. Analytics
Natural Language Processing and Data Mining for Clinical Text- Mining the_Biomedical- Social Media Analytics for Healthcare.

Unit-4. Advanced Data Analytics
Advanced Data Analytics for Healthcare– Review of Clinical Prediction Models- Temporal Data Mining for Healthcare Data- Visual Analytics for Healthcare- Predictive Models for Integrating Clinical and Genomic Data- Information Retrieval for Healthcare- Privacy-Preserving Data Publishing Methods in Healthcare.

Unit-5. Applications
Applications and Practical Systems for Healthcare– Data Analytics for Pervasive Health- Fraud Detection in Healthcare- Data Analytics for Pharmaceutical Discoveries- Clinical Decision Support Systems- Computer-Assisted Medical Image Analysis Systems- Mobile Imaging and Analytics for Biomedical Data

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Electronic Data Warehouse (EDW)	https://www.osplabs.com/healthcare-analytics/

2	Big Data in Health care	https://www.wipro.com/healthcare/advanced-healthcare-data-analytics/
3	Health Care Economics	https://healthcare.business.uconn.edu/certificate-health-care-analytics/

C. TEXT BOOK(S)

1. Chandan K. Reddy and Charu C Aggarwal, "Healthcare data analytics", Taylor &Francis, 2015
2. Hui Yang and Eva K. Lee, "Healthcare Analytics: From Data to Knowledge to Healthcare Improvement, Wiley, 2016.

D. REFERENCE BOOKS

1. Vikas Kumar, Healthcare Analytics Made Simple: Techniques in healthcare computing using machine learning and Python, Packt Publishing, 2018
2. Ross M. Mullner, Edward M. Rafalski, Healthcare Analytics Foundations and Frontiers, Routledge, 2019.

E. WEB LINKS

- https://onlinedegrees.sandiego.edu/classes/advanced-health-care-analytics/
- https://www.hci.net.in/courses/advanced-post-graduate-diploma-healthcare-decision-analytics/

3. SPECIFIC LEARNING OUTCOMES (SLO)

Unit/ Section	Topic	Learning outcomes	Level
I	Introduction		
1.1	Introduction to Healthcare Data Analytics	• Understand the concepts of Healthcare Data Analytics	K3
1.2	Electronic Health Records	• Explain the concepts of EHR	K4
1.3	Components of EHR	• List the components of EHR • Discuss the features of EHR	K2 K4
1.4	Coding Systems	• Development of Coding Systems algorithm	K5
1.5	Benefits of EHR	• List the benefits of EHR	K2
1.6	Barrier to Adopting EHR Challenges	• Discuss the Challenges in EHR	K4
1.7	Phenotyping Algorithms.	• Design the Phenotyping Algorithms	K6
II	**Analysis - I**		
2.1	Biomedical Image Analysis	• Explain the Biomedical Image Analysis	K5
2.2	Mining of Sensor Data in Healthcare	• Explore the mechanism to get the data from sensor • Design the algorithm for Mining the sensor data	K4 K6
	Biomedical Signal Analysis	• Explain the Biomedical Signal Analysis	K5
2.3	Genomic Data Analysis for Personalized Medicine	• Develop the personalized medicine system using Genomic Data	K6
III	**Analysis - II**		
3.1	Natural Language Processing and Data	• Develop the algorithms for mining text in HER using NLP.	K5

	Mining for Clinical Text	• Design the data mining tool for Clinical text data	K6
3.2	Mining the Biomedical	• Design the mining algorithm for Biomedical data	K6
3.3	Social Media Analytics for Healthcare.	• Develop the algorithms for Health care data using Social Media Analysis	K5
IV	Advanced Data Analytics		
4.1	Advanced Data Analytics for Healthcare– Review of Clinical Prediction Models	• List the techniques for Advanced Data Analytics for Healthcare • Assess the Clinical Prediction Models	K2 K6
4.2	Temporal Data Mining for Healthcare Data	• Design the temporal data mining algorithms for Healthcare Data	K6
4.3	Visual Analytics for Healthcare	• Develop the Visual Analytics for Healthcare	K5
4.4	Predictive Models for Integrating Clinical and Genomic Data	• Design the predictive model using Integrating Clinical and Genomic Data	K6
4.5	Information Retrieval for Healthcare	• Apply the Information Retrieval for Healthcare	K4
4.6	Privacy-Preserving Data Publishing Methods in Healthcare.	• Explain the Data publishing methods in Healthcare	K5
V	**Programming with R**		
5.1	Applications and Practical Systems for Healthcare	• List the applications of the health care data analysis	K2
5.2	Data Analytics for Pervasive Health	• Develop the algorithm for Pervasive Health using DA	K5
5.3	Fraud Detection in Healthcare	• Explain the Fraud Detection in Healthcare	K4
5.4	Data Analytics for Pharmaceutical Discoveries	• Develop the algorithm for Pharmaceutical Discoveries using DA	K5
5.5	Clinical Decision Support Systems	• Design the Clinical Decision Support Systems	K6
5.6	Computer-Assisted Medical Image Analysis Systems	• Construct the Computer-Assisted Medical Image Analysis Systems	K6
6.7	Mobile Imaging and Analytics for Biomedical Data	• Design the System for Mobile Imaging and Analytics to the Biomedical Data	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**H**	**M**	**L**	**M**	**M**	**M**	**M**	**M**	**-**	**H**	**H**
CO2	**H**	**H**	**M**	**L**	**-**	**L**	**-**	**L**	**M**	**M**	**-**	**M**	**-**
CO3	**H**	**H**	**H**	**H**	**M**	**M**	**L**	**-**	**-**	**H**	**H**	**H**	**M-**
CO4	**H**	**H**	**H**	**H**	**H**	**M**	**L**	**M**	**M**	**H**	**H**	**M**	**H**

CO5	**H**	**M**	**-**	**M**	**L**	**M**	**-**	**H**	**M**	**H**	**H**	**M**	**-**
CO6	**M**	**M**	**-**	**H**	**M**	**L**	**-**	**-**	**L**	**H**	**H**	**H**	**M**

5. COURSE ASSESSMENT METHODS

DIRECT:

5. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
6. Open Book Test.
7. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project, Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
8. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. M. Lovelin Pon Felciah

CORE III: DATA AND VISUAL ANALYTICS LAB			
Semester	II	Hours/Week	3
Course Code	P19DS2P3	Credits	3

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

CO#	Course Outcome	Level	Activity
CO1	Create data analytics systems using Numpy	K6	1, 2
CO2	Create data wrangling systems using Pandas	K6	3-6, 10
CO3	Create data visualization systems using Seaborn	K6	7
CO4	Crate time series analysis systems using Pandas time series	K6	8
CO5	Create interactive dashboards using Tableau	K6	11
CO6	Build and deploy end to end data analytics product at client site	K6	9, 12

2. SYLLABUS

Activity	Lab Activity Description
1	Data Analytics using NumPy
2	Data Analytics using Advanced NumPy
3	Pandas Indexing and Selection
4	Pandas Grouping and Aggregation
5	Pandas Concatenation, Merging and Join
6	Data Cleaning in Pandas
7	Data Visualization using Seaborn
8	Pandas Time Series Analysis
9	Exploratory Data Analysis on Cardiovascular Data
10	Advanced Data Wrangling in Pandas
11	Interactive Dashboard Creation in Tableau
12	Telecom data analytics system

3. SPECIFIC LEARNING OUTCOMES

Activity#	Lab Activity	Learning Outcome	Level
1	Data Analytics using NumPy	Create data analytics platform using NumPy	K6
2	Data Analytics using Advanced NumPy	Create advanced data analytics platform using NumPy	K6
3	Pandas Indexing and Selection	Perform Pandas Indexing and Selection operations	K6
4	Pandas Grouping and Aggregation	Perform Pandas Grouping and Aggregation operations	K6
5	Pandas Concatenation, Merging and Join	Perform Pandas Concatenation, Merging and Join operations	K6
6	Data Cleaning in Pandas	Perform data cleaning on large data	K6
7	Data Visualization using Seaborn	Develop Data Visualization systems using Seaborn	K6
8	Pandas Time Series Analysis	Perform Pandas Time Series Analysis	K6
9	Exploratory Data Analysis on Cardiovascular Data	Develop Exploratory Data Analysis platform on Cardiovascular Data	K6
10	Advanced Data Wrangling in Pandas	Perform Advanced Data Wrangling in Pandas	K6

11	Interactive Dashboard Creation in Tableau	Create Interactive Dashboard using Tableau	K6
12	Telecom data analytics system	Build and deploy Telecom data analytics system	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**L**	**M**	**M**	**M**	**L**	**-**	**-**	**H**	**M**	**M**	**-**
CO2	**H**	**M**	**H**	**M**	**M**	**H**	**M**	**M**	**L**	**H**	**H**	**-**	**M**
CO3	**H**	**M**	**M**	**L**	**H**	**M**	**-**	**-**	**L**	**M**	**M**	**M**	**-**
CO4	**H**	**L**	**H**	**H**	**H**	**H**	**M**	**M**	**L**	**H**	**M**	**M**	**H**
CO5	**H**	**M**	**L**	**H**	**H**	**L**	**L**	**M**	**M**	**H**	**H**	**H**	**M**
CO6	**H**	**M**	**M**	**L**	**H**	**L**	**L**	**L**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Practical Components): Closed Book
2. Cooperative Learning Report, Assignment, Group Discussion, project Report, Field Visit Report, Seminar.
3. Pre/Post Test, Viva, Report for each Exercise.
4. Lab Model Examination & End Semester Practical Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. B. Karthikeyan

CORE VI: PRACTICAL MACHINE LEARNING LAB			
Semester	II	Hours/Week	3
Course Code	P19DS2P4	Credits	3

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

CO#	Course Outcome	Level	Activity
1	Practice data and file formats; visualize data and familiarize Colab and Azure	K6	1
2	Build and deploy systems for business problems based on regression models	K6	3,4
3	Build and deploy systems for business problems based on classification models	K6	2,5,7,8
4	Build and deploy systems for business problems based on predictive analytics	K6	6
5	Build and deploy systems for business problems based on tree models	K6	9,10
6	Build and deploy systems for business problems based on clustering models	K6	11

2. SYLLABUS

Activity	Lab Activity Description
13	WarmUp: Familiarity with Data and Visualization
14	Pizza Liking Prediction using kNN
15	Fuel Amount Prediction using Linear Regression
16	House Price Prediction using LR with Regularization
17	Diabetes Classification using Logistic Regression
18	Predictive Analytics for Hospitals
19	Loan Approval Classification using SVM
20	Animal Classification using Decision Trees
21	Employee Hopping Prediction using Random Forests
22	Patients Physical Activities Prediction using Boosting
23	Shopping Mall Customer Segmentation using Clustering

TOPICS FOR SELF STUDY

S.No	Topic Title	Web Link
1	Stock price prediction	https://www.kaggle.com/darkknight91/ge-stock
2	Wake up word detection for Alexa	https://github.com/Picovoice/wake-word-benchmark
3	Jane Street Market prediction	https://www.kaggle.com/c/jane-street-market-prediction
4	HuBMap – Hacking the kidney	https://www.kaggle.com/c/hubmap-kidney-segmentation

3. SPECIFIC LEARNING OUTCOMES

Activity#	Lab Activity	Learning Outcome	Level
1	WarmUp: Familiarity with Data and Visualization	Open, process and visualize various data and files using CoLab and Azure platforms	K6
2	Pizza Liking Prediction using kNN	Build kNN model, perform training and prediction and compute accuracy values	K6
3	Fuel Amount Prediction	Perform preprocessing; build LR model,	K6

	using Linear Regression	perform training and prediction; compute MSE and R2 error; compare performance against KNN regressor and SGDregressor models and interpret results	
4	House Price Prediction using LR with Regularization	Perform One Hot Encoding, build LR model, compute RMSE error and compare performance against SGD Regressor, RidgeCV and LassoCV and interpret results	K6
5	Diabetes Classification using Logistic Regression	Create heatmap, build Logistic Regression model, print ROC curve and compare performance against LogisticRegressionCV with L1 and L2 and interpret results	K6
6	Predictive Analytics for Hospitals	Perform prediction, Apply Forward Selection, plot AUC scores and Plot Gain curves and Life curves and interpret results	K6
7	Loan Approval Classification using SVM	Perform EDA, Create LinearSVC model, Print accuracy, confusion matrix and classification report and compare LinearSVC model with SVC and SGDClassifier models	K6
8	Animal Classification using Decision Trees	Create ID3 Decision Tree using Entropy metric, Create CART Decision Tree using Gini metric and Visualize graph using graphviz	K6
9	Employee Hopping Prediction using Random Forests	Create RandomForestClassifier, perform training and testing; Print feature importance values; and Select the best number of trees based on out-of-bag error values	K6
10	Patients Physical Activities Prediction using Boosting	Build GradientBoostingClassifier, fit and predict on test data; Find the best no. of decision trees and learning rate using GridSearch and Cross Validation; Build AdaBoost, LogisticRegressionCV and VotingClassifier; Interpret results and parameter values	K6
11	Shopping Mall Customer Segmentation using Clustering	Perform Skew analysis; Build KMeans model; Apply Elbow method; Perform Cluster Analysis; Perform PCA; Build MeanShift clustering and Agglomerative clustering; Visualize clusters using Dendrogram	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	H	M	L	M	M	M	L	-	-	H	M	M	-
CO2	H	M	H	M	M	H	M	M	L	H	H	-	M
CO3	H	M	M	L	H	M	-	-	L	M	M	M	-
CO4	H	L	H	H	H	H	M	M	L	H	M	M	H
CO5	H	M	L	H	H	L	L	M	M	H	H	H	M
CO6	H	M	M	L	H	L	L	L	M	H	H	-	-

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Practical Components): Closed Book
2. Cooperative Learning Report, Assignment, Group Discussion, project Report, Field Visit Report, Seminar.
3. Pre/Post Test, Viva, Report for each Exercise.
4. Lab Model Examination & End Semester Practical Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. K. Rajkuamr

CORE VI: NATURAL LANGUAGE PROCESSING LAB			
Semester	II	Hours/Week	3
Course Code	P19DS2P5	Credits	3

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

CO#	Course Outcome	Level	Activity
CO1	Design systems to perform NLP preprocessing and document similarity	K6	1 - 5
CO2	Design NLP systems for spam filtering	K6	6
CO3	Design NLP systems for sentiment analysis	K6	7
CO4	Design NLP systems using tagging and named entity recognition	K6	8 - 10
CO5	Design NLP systems using Context free grammars	K6	11 - 14
CO6	Design NLP systems using SpaCy	K6	15

2. SYLLABUS

Activity	Lab Activity Description
1	Understanding Large Text Files
2	Computing Bigram Frequencies
3	Computing Document Similarity using VSM
4	Computing Document Similarity using Word2Vec
5	Stemming and Lemmatization on Movie Dataset
6	Spam Filtering using Multinomial Naïve Bayes
7	Sentiment Analysis on Movie Reviews
8	Exploring Part of Speech Tagging on Large Text Files
9	Building Bigram Tagger
10	Named Entity Recognition on Food Recipes Dataset
11	Building Parse Trees
12	Building and Parsing Context Free Grammars
13	Improving Grammar to Parse Ambiguous Sentences
14	Word Sense Disambiguation with Improved Lesk
15	Text Processing using SpaCy

Topics for Self Study

S.No	Topic Title	Web Link
1	Text preprocessing in languages other than English	https://github.com/morkapronczay/meetup-talk-text-preproc
2	Cross-classification of translationese	http://cl.haifa.ac.il/projects/translationese/index.shtml
3	Distinguishing between human and machine translation	http://cl.haifa.ac.il/projects/pmt/index.shtml
4	Native Language Identification	https://github.com/ellarabi/reddit-l2

3. SPECIFIC LEARNING OUTCOMES

Activity#	Lab Activity	Learning Outcome	Level
1	Large Text Files Processing	Understand Large Text Files	K6
2	Bigram Frequencies	Compute Bigram Frequencies	K6
3	Document Similarity using VSM	Compute Document Similarity using VSM	K6

4	Document Similarity using Word2Vec	Compute Document Similarity using Word2Vec	K6
5	Stemming and Lemmatization	Perform Stemming and Lemmatization on Movie Dataset	K6
6	Spam Filtering	Perform Spam Filtering using Multinomial Naïve Bayes	K6
7	Sentiment Analysis	Develop system for Sentiment Analysis on Movie Reviews	K6
8	Part of Speech Tagging on Large Text Files	Explore Part of Speech Tagging on Large Text Files	K6
9	Bigram Tagger	Build Bigram Tagger	K6
10	Named Entity Recognition	Perform Named Entity Recognition on Food Recipes Dataset	K6
11	Parse Trees	Build Parse Trees	K6
12	Context Free Grammars	Build and Parse Context Free Grammars	K6
13	Parsing Ambiguous Sentences	Improve Grammar to Parse Ambiguous Sentences	K6
14	Word Sense Disambiguation	Perform Word Sense Disambiguation with Improved Lesk	K6
15	Text Processing using SpaCy	Perform text processing using SpaCy	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**L**	**M**	**M**	**M**	**L**	**-**	**-**	**H**	**M**	**M**	**-**
CO2	**H**	**M**	**H**	**M**	**M**	**H**	**M**	**M**	**L**	**H**	**H**	**-**	**M**
CO3	**H**	**M**	**M**	**L**	**H**	**M**	**-**	**-**	**L**	**M**	**M**	**M**	**-**
CO4	**H**	**L**	**H**	**H**	**H**	**H**	**M**	**M**	**L**	**H**	**M**	**M**	**H**
CO5	**H**	**M**	**L**	**H**	**H**	**L**	**L**	**M**	**M**	**H**	**H**	**H**	**M**
CO6	**H**	**M**	**M**	**L**	**H**	**L**	**L**	**L**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Practical Components): Closed Book
2. Cooperative Learning Report, Assignment, Group Discussion, project Report, Field Visit Report, Seminar.
3. Pre/Post Test, Viva, Report for each Exercise.
4. Lab Model Examination & End Semester Practical Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. Janani Selvaraj

CORE VII: PRINCIPLES OF DEEP LEARNING			
Semester	III	Hours/Week	4

Course Code	P20DS307	Credits	4

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

S.No.	Course Outcomes	Level	Unit
CO1	Examine the basics of Tensorflow and its models	K6	I
CO2	Explain the characteristics of Convolutional Neural Networks	K6	II
CO3	Depict the architecture and use of the Autocoders	K6	III
CO4	Evaluate the Sequence analysis using Tenserflow	K5	III
CO5	Summarize the features of the Recurrent Neural Network	K6	IV
CO6	Construct the CNN using Deep reinforcement learning	K6	V

2. SYLLABUS

UNIT I - TENSORFLOW BASICS 12 HOURS

TensorFlow: variables, operations, placeholder Tensors, sessions – Navigating variable scopes and shared variables – Managing models over CPU and GPU – Logistic Regression in TensorFlow– Training Logistic Regression model – Visualizing using Tensor Board – Building multilayer model in TensorFlow

UNIT II - CONVOLUTIONAL NEURAL NETWORKS 12 HOURS

Shortcomings of Feature Selection – Width, height and depth of layers – Filters and feature maps – Describing convolutional layer – Max pooling - Architectural Description of Convolution Networks – Recognizing handwritten digits using CNN for MNIST dataset –Image preprocessing pipelines - Training with Batch normalization

UNIT III - AUTOENCODERS AND SEQUENCE ANALYSIS 12 HOURS

Embedding – Principal Component Analysis - Architecture of Autoencoders – Implementing autoencoders in TensorFlow–Denoising - Word2Vec framework for language modelling. Sequence Analysis: seq2seq problem – Dependency parsing – Beam search

UNIT IV - RECURRENT NEURAL NETWORKS 12 HOURS

Single neuron and fully connected recurrent layer – Challenges of vanishing gradients - LSTM architecture – TensorFlow primitives for RNN models – Implementing Sentiment analysis Model – Solving seq2seq tasks with RNN – Augmenting RNN with Attention – Designing Neural Translation Network

UNIT V - DEEP REINFORCEMENT LEARNING 12 HOURS

Reinforcement Learning: Markov Decision Processes, Policy, Future return, Discounted future return, Balancing Explore-Exploit dilemma, Annealed e-Greedy – Policy learning and Value learning - Solving Pole Cart problem with Policy Gradients - QLearning -Deep QNetworks – Deep Q Recurrent Networks – UNREAL Learning

TOPICS FOR SELF -STUDY

S.No.	Topics	Web Links
1	**Keras Tutorial**	**https://keras.io/getting_started/**
2	**Keras Tutorial: Deep Learning in Python**	**https://www.datacamp.com/community/tutorials/deep-learning-python**
3	**Machine Learning with Tensorflow**	**https://www.python-course.eu/tensor_flow_introduction.php**
4	**From Solving Equations to Deep**	**https://www.toptal.com/machine-**

	Learning: A TensorFlow	**learning/tensorflow-python-tutorial**

Text Books

1. Nikhil Buduma, Nicholas Locascio. Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms. O'Reilly Media. 2017.
2. Ian Goodfellow, YoshuaBengio, Aaron Courville. Deep Learning (Adaptive Computation and Machine Learning series). MIT Press, 2017.

References

1. Francois Chollet. Deep Learning with Python. 1ed, Manning Publications, 2017. ISBN 978-1617294433.

Web Links

1. www.tensorflow.org / tutorials
2. https://www.tensorflow.org/tutorials/generative/autoencoder
3. https://towardsdatascience.com/reinforcement-learning-with-python-part-1-creating-the-environment-dad6e0237d2d

3. SPECIFIC LEARNING OUTCOMES

Unit/ Section	Course Content	Learning outcomes	Level
I	**TENSORFLOW BASICS**		
1.1	TensorFlow: variables, operations, placeholder Tensors, sessions	Explain the features of the TensorFlow	K4
1.2	Navigating variable scopes and shared variables	Characterize the navigating variable and shared variable of TenserFlow.	K4
1.3	Managing models over CPU and GPU	Develop the models using CPU and GPU	K6
1.4	Logistic Regression in TensorFlow - Training Logistic Regression model	Evaluate the Logistic Regression using TensorFlow	K5
1.5	Visualizing using Tensor Board	Depict the model using Tensor Board	K6
1.6	Building multilayer model in TensorFlow	Design the multilayer model in TensorFlow	K6
II	**CONVOLUTIONAL NEURAL NETWORKS**		
2.1	Shortcomings of Feature Selection	Evaluate the disadvantages of the conventional feature selection	K5
2.2.	Width, height and depth of layers	Analyze the Width, height and depth of layers in CNN	K5
2.3	Filters and feature maps	Assess the filters and feature maps in CNN	K5
2.4	Describing convolutional layer	Explain the role of convolutional layers in CNN	K6
2.5	Max pooling	Diagnose the max pooling methods of CNN	K4
2.6	Architectural Description of Convolution Networks	Design the architecture of CNN	K6
2.7	Recognizing handwritten digits using CNN for MNIST dataset	Formulate the CNN for Recognizing handwritten digits from MNIST dataset	K6
2.8	Image preprocessing pipelines	Categorize Image preprocessing pipelines	K6
2.9	Training with Batch	Analyze the training of CNN using batch	K4

	normalization	normalization	
III	**AUTOENCODERS AND SEQUENCE ANALYSIS**		
3.1	Embedding	Explain the features of embedding	K5
3.2	Principal Component Analysis	Assess the characteristics of PCA	K5
3.3	Architecture of Autoencoders	Design the architecture of Autoencoders	K6
3.4	Implementing autoencoders in TensorFlow	Construct the autoencoders with TensorFlow	K6
3.5	Denoising	Describe the denoising methods for autoencoders	K4
3.6	Word2Vec framework for language modelling.	Develop the Autoencoders for Word2Vec framework for language modelling	K6
3.7	Sequence Analysis: seq2seq problem	Formulate the Sequence Analysis using TensorFlow	K6
3.8	Dependency parsing	Explain the steps of Dependency parsing	K4
3.9	Beam search	Interpret the Beam Search method for sequence analysis.	K5
IV	**RECURRENT NEURAL NETWORKS**		
4.1	Single neuron and fully connected recurrent layer	Explain the characteristics of single Neuron Construct the Fully connected recurrent layer	K4 K6
4.2	Challenges of vanishing gradients	Evaluate the challenges of vanishing gradients	K4
4.3	LSTM architecture	Explain the components of LSTM architecture.	K5
4.4	TensorFlow primitives for RNN models	Describe the TensorFlow primitives for RNN models	K2
4.5	Implementing Sentiment analysis Model	Design the Sentiment analysis Model using TensorFlow	K6
4.6	Solving seq2seq tasks with RNN	Formulate the solution for seq2seq tasks	K6
4.7	Augmenting RNN with Attention	Assess the augmenting RNN	K4
4.8	Designing Neural Translation Network	Design the Neural Translation Network using RNN	K6
V	**DEEP REINFORCEMENT LEARNING**		
5.1	Reinforcement Learning: Markov Decision Processes, Policy, Future return, Discounted future return, Balancing Explore	Describe the Reinforcement Learning Explain the Markov Decision Process. Compare the different types of return	K3 K5 K5
5.2	Exploit dilemma, Annealed e-Greedy	Assess the exploit dilemma of DRL Characterize the Annealed e-Greedy	K5 K4
5.3	Policy learning and Value learning	Distinguish the Policy learning and Value learning	K5
5.4	Solving Pole Cart problem with Policy Gradients	Prescribe the solution for pole cart problem using policy gradients	K6
5.5	QLearning	Explain the properties of QLearning	K4
5.6	Deep QNetworks	Assess the features of Deep QNetworks from conventional Neural network	K5
5.7	Deep Q Recurrent Networks	Compare the Deep QRecurrent Networks over Deep QNetworks	K6
5.8	UNREAL Learning	Explain the characteristics of unreal learning	K5

4. MAPPING

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**M**	**-**
CO2	**M**	**H**	**M**	**H**	**M**	**M**	**-**	**M**	**M**	**M**	**M**	**-**	**-**
CO3	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**H**	**H**	**M**	**-**	**-**
CO4	**H**	**H**	**L**	**H**	**M**	**H**	**-**	**M**	**H**	**H**	**H**	**-**	
CO5	**H**	**M**	**H**	**M**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**-**	**-**
CO6	**H**	**H**	**H**	**M**	**M**	**H**	**H**	**-**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Peer Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course evaluation survey
2. Faculty feedback about the course.

Name of the Course Coordinator: Dr. K. RAJKUAMR

CORE VIII: BIG DATA MANAGEMENT AND ANALYTICS			
Semester	III	Hours/Week	4
Course Code	P20DS308	Credits	4

1. Course Outcomes

On successful completion of this course, students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Perceive Big Data concepts and technologies	K6	I
CO2	Evaluate the Storing and manipulation of data using HDFS	K6	II
CO3	Construct the very large datasets using Pig	K6	III
CO4	Create MapReduce using Spark	K6	IV
CO5	Formulate Data Warehousing operations using Hive	K6	V
CO6	Create applications using Hadoop	K6	All

2. A. SYLLABUS

Unit-1. Introduction to Big Data

What is Big data – Industrial examples of Big Data: Digital Marketing, fraud, risk, trading, healthcare, medicine, advertising – Big Data Technology: Hadoop, cloud, BI, crowdsourcing analytics – Business Analytics:

Unit-2. MapReduce-I and HDFS

MapReduce model: Weather dataset, Analyzing data with Hadoop, Combiner functions, Hadoop streaming with Python. **Hadoop Distributed File System**: Block, Namenode, Datanode, Caching – File system operations in command line – Java Interface to Basic Hadoop - Reading data and writing data – Anatomy of File Write

Unit-3. MapReduce-II

Steps of developing MapReduce application - Working of MapReduce: Running Jobs, failure, Shuffle and sort, Task execution - MapReduce Types: Input formats - Output formats - MapReduce features: Counters, Sorting, Joins

Unit-IV. Exploring large datasets using Pig

Structure, Statements, Expressions, Types, Schemas, Functions, Macros - User-Defined Functions: Filter UDF, Eval UDF, Load UDF - Data Processing Operators: Loading and Storing Data, Filtering Data, Grouping and Joining Data, Sorting Data, Combining and Splitting Data

Unit-5. Data Warehousing using Hive

Comparison with Traditional Databases - HiveQL: Data Types, Operators and Functions - Tables:Managed Tables and External Tables, Partitions and Buckets, Storage Formats, Importing Data, Altering Tables, Dropping Tables - Querying Data: Sorting and Aggregating, MapReduce Scripts, Joins, Subqueries, Views - User-Defined Functions: Writing a UDF, Writing a UDAF – 6 Elements of Big Data Security

B. TOPICS FOR SELF-STUDY

C. TEXT BOOKS

1. Michael Minelli, Michele Chambers and Ambiga Dhiraj. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, 1ed, Wiley CIO Series, 2013. ISBN 9781118147603

2. Tom White Hadoop: The Definitive Guide, Fourth Edition, O'reilly Media, 2015.
3. Six Elements of Securing Big Data. MapR Ebook https://mapr.com/big-data-security-6-elements/

REFERENCES

1. Nathan Marz and James Warren, Big Data Principles and Practice of Scalable Real Time Data Systems, Manning Publications. 2015
2. Michael Berthold, David J. Hand, Intelligent Data Analysis, Springer, 2007
3. Bill Franks, Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics, John Wiley& sons, 2012.
4. Glenn J. Myatt, Making Sense of Data, Volume I and II. John Wiley & Sons, 2007.
5. Mark Grover, Ted Malaska, Jonathan Seidman, Gwen Shapira. Hadoop Application Architecture, Shroff Publishers.2015
6. Chris Eaton, Dirk DeRoos, Tom Deutsch, George Lapis, Paul Zikopoulos, Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, McGrawHill Publishing, 2012.

3. SPECIFIC LEARNING OUTCOMES

Unit	Topic	Topic Learning Outcome	Level
I	**Introduction to Big Data**		
1.1	What is Big data	Why need to handle big dataset	K1
1.2	Industrial examples of Big Data: Digital Marketing	Outline big data in marketing	K2
1.3	fraud, risk management	List the usage of the big data in fraud, risk management.	K1
1.4	trading	Explain time to time data analysis	K1
1.5	healthcare,	Why big data in healthcare	K1
1.6	medicine	How big data is in medicine	K1
1.7	advertising	How to advertise with big data	K1
1.8	Big Data Technology: Hadoop	Discover storing data and running applications on clusters of commodity hardware.	K4
1.9	Cloud	Build on demand services using internet.	K3
1.10	BI	Survey information retrieval from available huge amount of data	K4
1.11	crowdsourcing analytics	Develop outsourcing for quality and to handle large amount of data	K3
1.12	Business Analytics	Build statistical report by business analytics.	K3
II	**MapReduce-I and HDFS**		
2.1	**MapReduce model**: Weather dataset	Create combined report for weather from shuffler.	K6
2.2	Analyzing data with Hadoop	Analyze a huge collection of data that comprises both structured data found in traditional databases and unstructured data like text documents, video and audio.	K4
2.3	Combiner functions	Assess an optional class that operates by accepting the inputs from the Map class and thereafter passing the output key-value pairs to the Reducer class.	K5
2.4	Hadoop streaming with	Create stream using programming language that can read from standard input and write to standard output.	K6

	Python.		
2.5	**Hadoop Distributed File System**: Block	Create sequence of blocks from file which is to store.	K6
2.6	Namenode	Develop the centerpiece of an HDFS file system. It keeps the directory tree of all files in the file system, and tracks where across the cluster the file data is kept. It does not store the data of these files itself.	K3
2.7	Datanode	Build a DataNode stores data in the [Hadoop File System]. A functional filesystem has more than one DataNode, with data replicated	K3
2.8	Caching	Plan the Centralized cache management which is an explicit caching mechanism that allows users to specify paths to be cached by HDFS.	K3
2.9	File system operations in command line	Make use of HDFS command for import file and mapreduce	K3
2.10	Java Interface to Basic Hadoop	Adapt java interface for implement FileSystem represents like client interface to a filesystem in Hadoop, and there are several concrete implementations	K6
2.11	Reading data and writing data	Criticize WORA(Write once Read many) models	K5
2.12	Anatomy of File Write	Define file format	K1
III	**MapReduce-II**		
3.1	Steps of developing MapReduce application	Divide MapReduce as three stages, namely map stage, shuffle stage, and reduce stage.	K4
3.2	Working of MapReduce: Running Jobs	Create mapper's jobs to process the input data.	K6
3.3	failure	Determine TaskTracker to marks the task when failed.	K5
3.4	Shuffle and sort	Create shuffler for transfer mapper intermediate output to the reducer	K6
3.5	Task execution	Create task from MapReduce for parallel processing	K6
3.6	MapReduce Types: Input formats	Select different input format for block.	K5
3.7	Output formats	Select different output format for Shuffler and reducer.	K5
3.8	MapReduce features: Counters	Measure occurrences of any events.	K5
3.9	Sorting	Build sorting algorithm to automatically sort the output key-value pairs from the mapper by their keys.	K6
3.10	Joins	Select join for map the partitioned and sorted according to the keys.	K5
IV	**Exploring large datasets using Pig**		
4.1	Structure	Use pig structure for data processing	K1
4.2	Statements	List pig statements for data processing	K1
4.3	Expressions	Use Pig expression to manipulate data.	K1
4.4	Types	Examine four types of data model	K4

4.5	Schemas,	Define dataset schema.	K1
4.6	Functions	Use EVAL functions, Math functions, String functions and Pig built-in functions for data processing.	K1
4.7	Macros	Create the code modular and makes Pig Latin code shareable	K6
4.8	User-Defined Functions: Filter UDF	Create UDF for conditions in filter statements in data processing and return Boolean value.	K6
4.9	Eval UDF	Create UDF for FOREACH-GENERATE in data processing	K6
4.10	Load UDF	Create UDF Load function top on Hadoop for InputFormat to read data.	K6
4.11	Data Processing Operators: Loading and Storing Data	Elaborate Load Operator and Store Operator for Reading and Storing Data.	K6
4.12	Filtering Data	Select the required tuples from a relation based on 'condition'.	K5
4.13	Grouping and Joining Data	Make up cluster of data using group. Create Combine record using Join.	K6
4.14	Sorting Data	Create data in systematic order like ascending or descending order.	K6
4.15	Combining and Splitting Data	Select combine for join two or more relations. Select Split to split two or more relations.	K6
V	**Data Warehousing using Hive**		
5.1	Comparison with Traditional Databases	List difference between RDBMS and HIVE	K1
5.2	HiveQL: Data Types	Use Hive data types	K1
5.3	Operators and Functions	Recall Hive operations operators and functions for data storage	K1
5.4	Tables:Managed Tables and External Tables	Create Hive vertical table for manipulate data. Describes the metadata / schema on external files using hive	K6
5.5	Partitions and Buckets	Create partitions and these partitions can be further subdivided into more manageable parts known as Buckets or Clusters.	K6
5.6	Storage Formats	Compose storage format for input block from HDFS	K6
5.7	Importing Data	Create a directory in HDFS to hold the file and import CSV files into Hive tables.	K6
5.8	Altering Tables	Change the existing table like table name, column name, comment, and table properties.	K6
5.9	Dropping Tables	Delete the table/column data and their metadata	K6
5.10	Querying Data: Sorting and Aggregating	Create Querying data for sorting using Order By Create aggregate using AVG, SUM, or MAX functions.	K6
5.11	MapReduce Scripts	Create Hive script using gedit for MapReduce	K6
5.12	Joins	Select query for join two or more tables	K5
5.13	Subqueries	Create a subquery for evaluated and returns a result set.	K6
5.14	Views	Evaluate user requirements by generating views.	K6

5.15	User-Defined Functions: Writing a UDF	create custom functions to process records or groups of records	K6
5.16	Writing a UDAF	create custom Aggregate functions to process records or groups of records	K6
5.17	6 Elements of Big Data Security	Justify big data security by the use of six steps.	K5

4. Mapping

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**M**	**-**
CO2	**M**	**H**	**M**	**H**	**M**	**M**	**-**	**M**	**M**	**M**	**M**	**-**	**-**
CO3	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**H**	**H**	**M**	**-**	**-**
CO4	**H**	**H**	**L**	**H**	**M**	**H**	**-**	**M**	**H**	**H**	**H**	**-**	
CO5	**H**	**M**	**H**	**M**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**-**	**-**
CO6	**H**	**H**	**H**	**M**	**M**	**H**	**H**	**-**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. B. Karthikeyan

CORE IX: SOCIAL MEDIA ANALYTICS

Semester	III	Hours/Week	4
Course Code	P19DS309	Credits	4

1. Course Outcomes

On successful completion of this course, students will be able to:

CO#	Course Outcome	Level	Unit
CO1	Explain the essentials of graphs for social networks	K6	I
CO2	Measure social network nodes and simulate social network models	K6	II
CO3	Evaluate the community analysis of social networks	K6	III
CO4	Measure and model information diffusion and homophily in social networks	K6	IV
CO5	Develop recommender systems and predict user behaviours	K6	V
CO6	Build and deploy end to end products into production environment	K6	All

2. A. SYLLABUS

Unit-1. Introduction to SMM and Graph Mining

What is social media mining – New challenges for mining. Graph Essentials: Graph basics – Graph representation – Types of graphs – Connectivity in graphs – Special graphs – Graph algorithms

Unit-2. Social Network Models

Network Measures: Centrality – Transitivity, reciprocity – Balance and status – Similarity. Network Models: Properties – Random graphs – Small world models – Preferential attachment model

Unit-3. Data Mining Basics and Community Discovery

Data Mining Essentials: Data Preprocessing – Supervised Learning Algorithms – Unsupervised Learning Algorithms. Community Analysis: Community detection – Community evolution – Community evaluation

Unit-4. Information Diffusion and Influence in Social Media

Information Diffusion: Herd behaviour – Information cascades – Diffusion of innovations – Epidemics. Influence and Homophily: Measuring Assortativity – Measuring and modelling influence – Measuring and modelling homophily – Distinguishing influence and homophily

Unit-5. Recommendation and Behaviour Analysis in Social Media

Recommendation in Social Media: Challenges – Classical recommendation algorithms – Recommendation using social context – Evaluating recommendations. Behaviour Analysis: Individual behaviour – Collective behaviour. Events Analytics in Social Media.

B. TOPICS FOR SELF STUDY

SNo	Topic Title	Web Link
1	Creating graphs using NetworkX for Airline data set	https://www.analyticsvidhya.com/blog/2018/09/introduction-graph-theory-applications-python/
2	Implementation of Movie recommender system	https://www.geeksforgeeks.org/python-implementation-of-movie-recommender-system/
3	Diving into GraphQL and Neo4j with Python	https://medium.com/elements/diving-into-graphql-and-neo4j-with-python-244ec39ddd94
4	DataCamp Network analysis using Python	https://www.datacamp.com/courses/introduction-to-network-analysis-in-python

C. TEXT BOOK(S)

1. Reza Zafarani, Mohammad Ali Abbasi, and Huan Liu. Social Media Mining: An Introduction, Cambridge University Press, 2014

D. REFERENCES BOOKS

1. Matthew A. Russell. Mining the Social Web. 3rd Edition. O'Reilly Media. 2019
2. Jennifer Golbeck. Analyzing the Social Web. Morgn Kaufmann. 2013. ISBN 978-0124055315
3. Ricardo Baeza-Yates and BerthierRibeiro-Neto. Modern Information Retrieval: The Concepts and Technology behind Search. 2ed. ACM Press Books, 2011. ISBN 978-0321416919
4. Charu C. Aggarwal. Social Network Data Analytics. Springer. 2011

E. WEB LINKS

- https://medium.com/elements/diving-into-graphql-and-neo4j-with-python-244ec39ddd94
- https://www.geeksforgeeks.org/python-implementation-of-movie-recommender-system/

3. SPECIFIC LEARNING OUTCOMES

Unit	Topic	Topic Learning Outcomes	Level
I	**Introduction to SMM and Graph Mining**		
1.1	Challenges and methodologies for mining	Identify challenges and methodologies for social media mining	K1
1.2	Types of SM and marketing opportunities that exist in SM	List social media types and identify marketing opportunities	K2
1.3	Graph basics	Compute degree and degree distribution of directed and undirected graphs	K3
1.4	Graph representation	Find adjacency list and edge list Given these lists, create a graph	K4 K6
1.5	Types of graphs	Given business problem, create weighted and signed graphs for the social network	K6
1.6	Connectivity in graphs	Find all connectivity in graphs and compute diameter	K4
1.6	Special graphs	Create various special graphs such as MST, Steiner tree, planner graph, bipartite graph and regular graph for the given business problem	K6
1.8	Graph traversals	Apply BFS and DFS traversal methods for the given social network	K3
1.9	Shortest path algorithms	Compute shortest paths using Dijkstra's and Prim's algorithms based on the business use case	K3
1.10	Network flow algorithms	Analyze maximum messages a social network can handle. Analyze maximum matching between products and users.	K4 K4
II	**Social Network Measures and Models**		
2.1	Centrality measures	Apply centrality measures and predict the most central important nodes from social networks	K6
2.2	Transitivity	Apply transitivity measures and analyse linking behaviour of nodes	K4
2.3	Reciprocity	Analyze reciprocity of the given social network	K4
2.4	Balance and status	Determine consistency of relationship in signed graphs	K6
2.5	Similarity measures	Apply similarity measures and predict similar nodes	K6

		in a social network	
2.6	Properties of real world networks	Discuss the properties of real world networks	K6
2.7	Random graph model	Discuss the types, evolution and properties of random graph model	K6
2.8	Small world model	Discuss the properties of small world model Compare the properties of random graph and small world models	K6
2.9	Preferential attachment model	Discuss the properties of small world model Compare the functionalities of random graph, small world and preferential attachment models Figure out the differences between random graphs, regular lattices, and small-world models	K6 K4 K4
III	**Data Mining Basics and Community Discovery**		
3.1	Data pre-processing steps	Given a business problem, identify various features Explain data pre-processing steps, given a use case Select a ML methodology based on the given problem scenario	K1 K2 K4
3.2	Decision tree learning	Create a decision tree given a dataset representing the use case	K6
3.3	Naïve bayes classifier	Create a Naïve bayes classifier given a dataset representing the use case	K6
3.4	Nearest neighbour classifier	Create a KNN classifier given a dataset representing the use case	K6
3.5	Supervised learning evaluation methods	Compare the evaluation measures for supervised ML classifiers	K4
3.6	KMeans clustering	Predict clusters using KMeans given an use case	K6
3.7	Unsupervised learning evaluation methods	Evaluate quality of clusters from unsupervised ML classifiers	K6
3.8	Member based community detection	Detect communities by applying node similarity, node degree and node reachability methods	K4
3.9	Group based community detection	Discuss the methods to detect group communities from social networks	K2
3.10	Community evolution	Explain how communities evolve over time in social networks	K2
3.11	Community evaluation	Given members of communities, analyze precision, recall, Fscore, purity and NMI measures	K4
IV	**Information Diffusion and Influence**		
4.1	Information cascades	Given a network with activation probabilities, analyse final set of activated nodes using ICM method Describe the independent cascade model Explain the objectives of cascade maximization	K4 K2 K4
4.2	Diffusion of innovations	Compare innovation diffusion models	K4
4.3	Epidemics	Discuss the mathematical relationship between the SIR and the SIS models Defend why in SIR model, the probability that an individual remains infected follows a standard exponential distribution Compute in SIRS model, the length of time that an infected individual is likely to remain infected before he or she recovers Given a business or societal problem, select the	K6 K6 K3 K4 K6

		appropriate information diffusion model Sumarize intervention approaches for information diffusion models	
4.4	Measuring Assortativity	Compute assortativity for ordinal and nominal attributes of social network nodes	K3
4.5	Measuring influence	Illustrate the types of influence measures in blogosphere and twitter	K4
4.6	Modelling influence	Select all activated nodes with Linear Threshold Model	K4
4.7	Measuring homophily	Estimate homophily for nominal and ordinal attributes in a social network	K6
4.8	Modelling homophily	Explain the variation of independent cascade model to model homophily	K4
4.9	Distinguishing influence and homophily	Determine the source of assortativity in social networks	K6
V	**Recommendation and Behaviour Analysis**		
5.1	Challenges of recommender systems	Describe the challenges of recommendation systems	K1
5.2	Content based recommendation system	Differentiate content-based recommendation from collaborative filtering	K4
5.3	User based collaborative filtering	Predict missing ratings using user based CF	K6
5.4	Item based collaborative filtering	Predict the most similar items using item based CF	K6
5.5	Model based collaborative filtering	Predict ratings and items using SVD based CF	K6
5.6	Group based recommendation	Find and recommend items to group of users	K3
5.7	Recommendation using social context	Predict ratings and items leveraging social context Provide examples where social context can help improve classical recommendation algorithms in social media	K6 K3
5.8	Evaluation of recommender systems	Evaluate the accuracy of predictions	K5
5.9	Evaluating relevancy of recommendation	Evaluate the relevancy of recommendations	K5
5.10	Evaluating ranking of recommendation	Evaluate the ranking of recommendations	K5
5.11	Individual behaviour analysis, modelling and prediction	List the features for User Community-Joining Behavior Explain the methods for predicting individual behaviours	K1 K4
5.12	Collective behaviour analysis, modelling and prediction	Outline a method for predicting Box office Revenue for Movies	K5

4. MAPPING

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4

CO1	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**M**	**-**
CO2	**M**	**H**	**M**	**H**	**M**	**M**	**-**	**M**	**M**	**M**	**M**	**-**	**-**
CO3	**H**	**M**	**H**	**H**	**H**	**H**	**-**	**-**	**H**	**H**	**M**	**-**	**-**
CO4	**H**	**H**	**L**	**H**	**M**	**H**	**-**	**M**	**H**	**H**	**H**	**-**	
CO5	**H**	**M**	**H**	**M**	**H**	**H**	**-**	**-**	**M**	**H**	**H**	**-**	**-**
CO6	**H**	**H**	**H**	**M**	**M**	**H**	**H**	**-**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator: Dr. M. Lovelin Pon Felciah

ELECTIVE IV: IMAGE AND VIDEO ANALYTICS			
Semester	III	Hours/Week	4
Course Code	P19DS3:4	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Elaborate the fundamental principles of image and video analysis	K5	I
CO2	Choose the mathematical basic equation to transform images into different domain for performing smoothing and sharpening operations	K6	II
CO3	Evaluate a statistical model to solve Image Enhancement, Segmentation and Compression problems	K	III
CO4	Select most relevant information from the original image to construct a feature vector such as texture, color and shape	K6	IV
CO5	Design suitable Classifier for Object Detection, Tracking and Recognition	K5	IV
CO6	Decide suitable image and video analysis approaches for developing solutions to solve real time applications	K6	V

2. A. SYLLABUS

Unit-1 Image Representation and Processing

Digital image representation- Visual Perception- Sampling and Quantization- Basic Relations between Pixels- Mathematical Tools Used in Digital Image Processing: Fundamental Operations – Vector and Matric Operations- Image Transforms (DFT, DCT,DWT, Hadamard).

Unit-2 Image Filtering

Fundamentals of spatial filtering: spatial correlation and convolution-smoothing, blurring-sharpening- edge detection - Basics of filtering in the frequency domain: smoothing-blurring-sharpening--Histograms and basic statistical models of image.

Unit-3 Colors and Compression

Color models and Transformations – Image and Video segmentation-Image and video demonising-Image and Video enhancement- Image and Video compression.

Unit-4 Object Detection and Tracking

Object detection and recognition in image and video-Texture models Image and Video classification models- Object tracking in Video.

Unit-5 Applications

Applications and Case studies- Industrial- Retail- Transportation & Travel- Remote sensing-Video Analytics in WSN: IoT Video Analytics Architectures.

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Pattern Recognition and Application	https://nptel.ac.in/courses/117/105/117105101/
2	Practical Machine Leaning with Tensor Flow (Video)	https://nptel.ac.in/courses/106/106/106106213/
3	Object Representation and Description	https://www.youtube.com/watch?v=yxID4fgz1C0

4	Deep Learning for Computer Vision	https://www.kaggle.com/learn/computer-vision

C. TEXT BOOKS

1. R.C. Gonzalez and R.E. Woods. Digital Image Processing. 3rd Edition. Addison Wesley, 2007.

.

D. REFERENCES BOOKS

1. Pratt, W.K. Digital image processing: PIKS scientific inside. 4ed. New York: John Wiley, 2007.
2. W. Härdle, M. Müller, S. Sperlich, A. Werwatz. Nonparametric and Semi parametric Models. Springer, 2004.
3. Rick Szelisk. Computer Vision: Algorithms and Applications. Springer 2011.
4. Jean-Yves Dufour. Intelligent Video Surveillance Systems. Wiley, 2013.
5. Caifeng Shan, FatihPorikli, Tao Xiang, Shaogang Gong. Video Analytics for Business Intelligence. Springer, 2012.
6. AsierPerallos, Unai Hernandez-Jayo, Enrique Onieva, Ignacio Julio GarcíaZuazola. Intelligent Transport Systems: Technologies and Applications. Wiley, 2015.
7. BasudebBhatta. Analysis of Urban Growth and Sprawl from Remote Sensing Data. Springer, 2010

E. WEB LINKS

1. https://www.coursera.org/learn/digital
2. https://nptel.ac.in/courses/106/105/106105032

3. SPECIFIC LEARNING OUTCOMES (SLO)

Unit/ Section	Course Content	Learning outcomes	Level
I	**Image Representation and Processing**		
1.1	Digital image representation	Discuss the fundamental steps involved in Image processing system	K2
		Describe the image representation method	K2
1.2	Visual Perception	Explain the human visual perception system with necessary diagrams.	K4
		Analyze the image formation takes place in eye and state the principle operation of brightness adaption and discrimination	K4
1.3	Sampling and Quantization	Design the image digitization process by sampling and quantization	K5
1.4	Basic relations between Pixels	Analyze the basic relationships between pixels	K4
		Distinguish the following terms: i) Adjacency ii) Connectivity iii) Region iv)Boundary	K4
1.5	Mathematical Tools Used in Digital Image Processing	Examine the following mathematical operations on digital image i) Array versus Matrix operation ii) Linear versus Nonlinear Operations	K4
1.6	Image Transforms (DFT, DCT,DWT, Hadamard)	Create a MATLAB script to construct the forward and inverse 2D DFT for the given image f(m,n)	K5
		Compare the following two properties of 2D-DFT	K2

		i) Convolution ii) Correlation	
		Design the basis function of Haar Transform for N=8	K5
		Use the Hadamard kernel matrix 4 x 4 for the image segment and perform transform with matrix multiplication method	K3
		Construct 2D DCT for the image of size 2 X 2 and verify the output after inverse DCT	K5
		Determine the approximation and detailed coefficient of the Harr Transform which takes an argument as 2 – dimensional digital signal 'S'.	K6
		Create discrete cosine transform(DCT) matrix for N=4	K5
		Design second-level decomposition of the input image using a Haar wavelet	K5
		Construct the Haar transform $T = HFH^T$ of the 2 x 2 image F(m,n) also find the inverse Haar transform $F = H^TTH$ of the obtained result.	K5
		Construct the subband modeling using DWT	K5
II	**Image Filtering**		
2.1	Fundamentals of spatial filtering	Analyse the impact of convolving the image f(x,y) with the mask h(x,y)that performs averaging operation which results in blurring the image	K4
		Compare linear and Non-Linear spatial Filtering Techniques	K2
		Explain Image Negative and Log transformation Techniques	K2
		Determine the output image if f(m, n) and h(m, n) are linearly convolved with zero padding of the original image	K6
		Examine the behavior of Spatial-domain low-pass filtering of the input image using different window sizes like 3 × 3, 5 × 5 and 7 × 7	K4
		Verify the effect of a 5 × 5 uniform averaging filter to a digital image N times.	K6
2.2	Spatial correlation and convolution	Formulate the 2D linear convolution between the signal x(m, n) and h(m,n) and comment on the observed result.	K5
		Compare the following properties of two-dimensional convolution (i) Commutative property (ii) Associative property (iii) Distributive property	K2
		Determine the correlation between the two image matrices x1[m,n] and x2[m,n]	K6
		Verify that convolution with a 2D separable filter can be accomplished by	K6

		performing two one dimensional Convolutions	
2.3	Smoothing, blurring	Justify the statement Mean filter is an effective tool to minimize salt and pepper noise through simple example	K6
		Invent the new value of the pixel(2,2) if smoothing is done using a 3x3 neighborhood using the following filters a) Mean filter b) Weighted average filter c) Median filter d) Min and Max filter	K5
		Discuss the limiting effect of repeatedly applying a 3x3 low-pass spatial filter to a digital image. Ignore border effects. Is this effect different from applying 5x5 filter	K6
		Analyze 3 x 3 mean filter in the frequency domain and prove that it behaves like a low pass filter	K4
		Determine the convolution process using 3x3 mask in the portion of pixels(2x2) of original image size 5x5 and write the filtered image.	K6
		Verify the smoothing behavior of Gaussian filter with varying levels of smooth factor σ.	K6
		Show the output impact in applying full-scale contrast stretch to the image 4 × 4, 4bits/pixel image.	K3
2.4	Sharpening- edge detection	Discuss the effect of first order derivative gradient operators for image sharpening	K6
		Invent the discontinuity in the image using canny edge detector and give justification why it outperforms than gradient edge detectors and implement using MATLAB code	K5
		Discuss the behaviour of the second order Derivative for a step and ramp edges	K6
		Construct the LOG filter to detect isolated points and line in an image	K3
2.5	Basics of filtering in the frequency domain: smoothing-blurring-sharpening	Verify that convolution in spatial domain is equal to multiplication in the frequency domain using MATLAB code	K6
		Use MATLAB code to perform a two-dimensional Butterworth low-pass filter of the given image for two different cut-off frequencies	K3
		Construct a filter for image smoothing in frequency domain	K3
		Explain the various high pass filters used in frequency domain	K4
2.6	Histograms and basic statistical models of image	Justify your answer can two different images have the same histogram	K6

		Justify Histogram processing is called as an efficient tool for graphical representation of the total distribution in a given digital image.	K6
		Design a statistical model of Histogram Equalization to the given image by rounding the resulting image pixels to integers	K5
		Determine the histogram equalization as an idempotent operation for the 5 x 5 image segment. Plot the graph before and after equation.	K6
		Determine the visual appearance of the resulting image by applying global histogram equalization for the grey level image f(x,y) of size 256x256 with $1<x,y<256$, which has the following intensities $f(x,y) = r+1$ if $1\leq x,y\leq 12$ and $f(x,y)=r$ if $13\leq x,y\leq 16$, otherwise $f(x,y)=r+3$.	K6
		Determine the mean and standard deviation of the image. If an image has gray levels ranging from 0 to 19.	K6
III	**Colors and Compression**		
3.1	Color models and Transformations	Describe the representation of three-color components red, green and blue for the given color image.	K2
		Classify the foreground and background from the given RGB image and segment it using the Global thresholding method	K3
		Verify the gamma correction for the given color image for different values of gamma and comment on the output result.	K6
		Formulate the CMY coordinates from the given color image represented in terms of RGB components	K5
		Construct a statistical model of histogram equalization of the given RGB image	K5
		Discuss the additive and subtractive color model also implement Python OpenCV code to extract color components	K6
		Justify the result for the color transform model to read the color image. Convert the RGB format to YIQ format (NTSC). Filter only the Y component (high-pass filtering). Do not disturb the I and Q components. Then convert the filtered Y component, I component and q component back to the RGB format and check the result.	K6
		Construct the median filter for the color image corrupted by salt-and-pepper noise and try to restore the corrupted image	K5
3.2	Image and Video	Apply the Region based split-and-merge	K3

	segmentation	technique to segment the given image	
		Discuss the Morphological operations opening and closing for the given binary image	K2
		Predict the number of black pixels in the resultant image for the given input binary image if hit-or-miss transformation is performed with the structure element [0 1 0,1 1 1,0 1 0].	K5
		Explain watershed segmentation tends to over-segmentation problem in images. Mention the solution to overcome the problem	K4
		Formulate the gradient magnitude and the direction of the gradient for the pixel f(x,y)=2x2.	K5
		Determine the hit and miss transformation of Morphological operator on a binary array that represents a portion of a black-and-white image and perform the operations on this piece of image. Assume that all the pixels that surround this segment contain a black background	K6
		Verify that the Prewitt edge detector along a horizontal direction can be obtained by convolving two one-dimensional signals [1, 1, 1] and [−1, 0, 1] T and then scaling the result by a factor of 1/3	K6
		Construct the linear filter masks for the following operations: (a) Detecting horizontal lines (b) Detecting vertical edges	K5
		Distinguish between image segmentation based on thresholding with image segmentation based on region-growing techniques.	K4
		Design an Automatic thresholding of grey level image using otsu's thresholding	K5
		Discuss Multiple object segmentation in video using Graph Cut	K6
		Devise a technique to detect outlier from motion segmentation in video	K5
		Apply the Region based split-and-merge technique to segment the given image	K3
3.3	Image and video demonising	Discuss the tools available for image and video demonizing	K2
3.4	Image and Video enhancement	Explain the various video enhancement techniques	K2
		Distinguish the following enhancement operations: i) Contrast stretching ii) Bit-plane slicing	K4
		Analyze the behavior of piecewise linear transformation and grey level transform for image enhancement	K4

		Determine the output image g(m,n) using logarithmic transformation $g(m,n)=[c\log_{10}(1+f(m,n)]$ by choosing c as (i) c=1 and $c=L/\log_{10}(1+L)$	K6
		Examine the behavior of image arithmetic operation such as addition, subtraction, multiplication and division over an image	K3
		Judge the impact of zeroing least significant and most significant bit planes by reading an eight-bit image, set any of the bit planes 0 to 7 to zero in a user defined manner and reconstruct the image.	K6
3.5	Image and Video Compression	Construct the Huffman tree and find the number of bits needed for encoding a given message. Calculate number of bits using frequency of characters and number of bits required to represent those characters**.**	K5
		Analyze the compression and reconstruction of the 8x8 input images for the 256 × 256 pixel digital image has eight distinct intensity levels also find the minimum number of bits required to code this image in a lossless manner.	K5
		Examine the efficiency of Huffman code for an image clip is formed using six colors—white (W), red (R), yellow (Y), green (G), blue (B) and orange (O). These occur in the clip with the following relative frequencies: {0.5,0.1,0.05,0.05,0.2,0.1}. For the above data, construct a Huffman code that minimizes the average code word length**.**	K4
		Explain MPEG Video compression standard for monochrome and color compression	K4
		Formulate the Peak Signal-to-Noise Ratio (PSNR) for the original and the reconstructed images. Calculate the PSNR expressed in decibels.	K5
		Evaluate encoding of the word a1, a2, a3, a4 using arithmetic code and generate the tag for the given symbol with probabilities: a1= 0.2, a2=0.2, a3=0.4, a4=0.2	K6
		Solve the entropy of the given 2D image given by f(m n).	K3
IV	**Object Detection and Tracking**		
4.1	Object detection and recognition in image and video	Predict the key points in objects using Harries corner detection and SURF. Justify why SURF gives high robustness corresponding to point matching	K5
		Design a system for detecting Criminals using Region based CNN	K5

		Examine Histogram of Gradients in 8×8 cells for object detection	K3
		Discuss Object detection using bounding box technology in Real Time Traffic monitoring system	K6
		Determine the object detection in real time video surveillance system	K6
4.2	Texture models	Discriminate the texture from the given 4 x4 image segment with grey levels (N) =0,1,2, 3 and d={1,0} by assuming direction operator as i) next pixel on right side, ii) next pixel on diagonal, iii) next pixel on perpendicular also calculate its homogeneity and uniformity to construct a test feature vector.	K6
		Discuss the vehicle detection system in real time video based on texture analysis	K6
		Explain Image segmentation using texture extraction	K4
		Invent Statistical texture feature for drugs classification	K5
		Discuss the common statistical features derived from co-occurrence probabilities	K2
		Create a 2D texture mask of size 5 x 5 with the following 1D filter i) E5E5 ii) E5R5/R5E5 iii) S5S5 iv) L5R5/R5L5	K5
4.3	Image and Video classification models	Devise a Model based video classification using SVM	K5
		Discuss the performance evaluation of deep feature learning for RGB image and video classification	K6
		Explain Gaussian mixture models of color and texture features for image classification and Gradient Descent Algorithm	K4
		Design an effective architecture for image classification using CNN	K5
		Create a unified framework for multi-label image classification	K5
		Create a model to classify images into their appropriate class with deep learning using CIFAR-10 dataset	K5
		Predict the classes using SVM classifier for the Breast cancer as Benign, Malignant, or Normal image by applying Otsu thresholding for segmentation, Preprocessing done by applying two-dimensional median filter and histogram equalization for getting more enhanced image. Then extract desired	K5

		features from the images for classification.	
4.4	Object tracking in Video	Compare the various object tracking techniques used in video processing	K4
		Create a model to detecting human motion in video surveillance system	K5
		Distinguish between automatic detection and motion-based object detection in a video	K4
		Construct a model for Motion-Based Multiple Object Tracking with suitable example.	K3
		Choose a suitable object tracking technique to perform Human gesture tracking and recognition	K6
V	**Applications**		
5.1	Applications and Case studies- Retail	Design a system for category analysis in industrial retail using clustering techniques. Suggest a suitable cluster algorithm	K5
		Construct a model for value and store brand identification in food products using Python OpenCV	K3
		Develop a system for product identification method for a mixed-reality web shopping system	K5
		Analyze the RFID Performance Evaluation in a Retail Store	K4
		Construct Image analytics method to monitor retail store	K3
		Construct an Automated Shopping Trolley for Super Market Billing System	K3
		Devise a deep learning pipeline for product recognition on retail store shelves	K5
		Develop an IoT based retail shopping system	K5
		Create a model for RFID Based Smart Shopping and Billing	K5
5.2	Industrial	Identify fault detection in industrial process using suitable image processing Technique.	K4
		Explain Image pattern recognition in industrial inspection	K4
5.3	Transportation & Travel	Propose a technique to detect traffic sign in real time traffic monitoring system	K5
		Predict the presence of Pedestrian in heavy traffic using object tracking method	K5
		Design a system for detecting driver drowsiness using image processing techniques. Suggest a suitable algorithm for each step.	K5
		Determine the discontinuity in the video frame to perform motion segmentation in	K6

		Transportation system	
		Design a system for recognition of number plates in vehicle using image processing techniques. Suggest a suitable algorithm for each step	K5
		Illustrate the applications of vision based intelligent transportation system	K3
5.4	Remote sensing	Design Remote sensing image classification using Deep Learning	K5
		Describe the wiener filter is helpful to reduce the mean square error when Satellite image is corrupted by motion blur and additive noise	K2
		Select the suitable preprocessing techniques to remove the distortion from the images taken from WSN Video surveillance system and reconstruct the same	K6
		Plan spatial resolution requirements for crop identification using optical image sensing	K5
		Propose Remote sensing in precision agriculture	K5
		Design an edge based and texture-based model for segmenting the remote sensing image and give implementation using MATLAB	K5
5.5	Video Analytics in WSN	Design a Distributed visual target-surveillance system in wireless sensor networks	K5
		Predict Rank preserving Discriminate analysis for human behaviour recognition through wireless sensor networks	K5
		Discuss Border patrol through advanced wireless sensor networks	K6
		Propose an intelligent car park management system based on wireless sensor networks	K5
		Design an Intelligent parking IoT application using wireless sensor networks	K5
5.6	IoT Video Analytics Architectures	Devise an efficient algorithm for media-based surveillance system in IoT	K5
		Explain IoT based smart video surveillance system	K4
		Design a Video Analytics – based Intelligent Indoor Positioning System Using IoT	K5

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

Mapping	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	H												
CO2	H	H	L		M	M							M

CO3	H		M	H	M							H	
CO4	H	H	M		H	M						H	
CO5	H	H			H	H			M	M		H	H
CO6	H	H	H	M	H	H			H		H		

5.COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator : Prof. D. Indra Devi

Core Practical VI: Big Data Management and Analytics Lab			
Semester	III	Hours/Week	5
Course Code	P20DS3P6	Credits	3

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

CO	Course Outcome	Level	Exercise
CO1	Develop applications using Hadoop	K6	1,2
CO2	Store and manipulate data using HDFS	K6	3
CO3	Data manipulation using MapReduce	K6	4,5 & 6
CO4	Explore very large datasets using Pig	K6	8,9
CO5	Perform Data Warehousing operations using Hive	K6	10
CO6	Perform data analytics using Spark	K6	7

2. LIST OF EXERCISES

Develop applications for the following tasks

1. Installation and setup of Hadoop
2. File management tasks in Hadoop
3. Benchmarking and stress testing on Hadoop cluster
4. Map Reduce applications for Word Counting
5. Stop word elimination using Map Reduce
6. Weather data analytics using Map Reduce
7. Perform data analytics using Spark
8. Perform sort, group, join, project, and filter operations on Pig
9. Design vector space model for text collection using Pig
10. Create, alter, and drop databases, tables, views, functions, and indexes on Hive

Topics for Self Study

S.No	Topic Title	Web Link
1	HDFS	https://docs.cloudera.com/documentation/enterprise/latest/topics/admin_hdfs_config.html
2	MapReduce	https://archive.cloudera.com/cdh5/cdh/5/hadoop/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html
3	Spark	https://docs.cloudera.com/documentation/enterprise/latest/topics/spark.html
4	Pig	https://docs.cloudera.com/documentation/enterprise/5-9-x/topics/cdh_ig_pig_installation.html
5	Hive	https://docs.cloudera.com/documentation/enterprise/5-8-x/topics/hive.html

3. Specific Learning Outcomes

Exercises	Lab Exercises	Learning Outcome	Level
1	Installation and setup of Hadoop	DFS,FS	K6
2	File management tasks in Hadoop	Place files in DFS	K6
3	Benchmarking and stress testing on Hadoop cluster	Write file in clustered Data Node	K6
4	Map Reduce applications for	Import jar file for MapReduce	K6

	Word Counting		
5	Stop word elimination using Map Reduce	Modify Word Count file as Word Elimination using Eclips	K6
6	Weather data analytics using Map Reduce	Process .csv file using MapReduce	K6
7	Perform data analytics using Spark	Spark using Scala	K6
8	Perform sort, group, join, project, and filter operations on Pig	MapReduce using Apache Tez	K6
9	Design vector space model for text collection using Pig	PigLatin Script	K6
10	Create, alter, and drop databases, tables, views, functions, and indexes on Hive	Data Warehousing	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**L**	**M**	**M**	**M**	**L**	**-**	**-**	**H**	**M**	**M**	**-**
CO2	**H**	**M**	**H**	**M**	**M**	**H**	**M**	**M**	**L**	**H**	**H**	**-**	**M**
CO3	**H**	**M**	**M**	**L**	**H**	**M**	**-**	**-**	**L**	**M**	**M**	**M**	**-**
CO4	**H**	**L**	**H**	**H**	**H**	**H**	**M**	**M**	**L**	**H**	**M**	**M**	**H**
CO5	**H**	**M**	**L**	**H**	**H**	**L**	**L**	**M**	**M**	**H**	**H**	**H**	**M**
CO6	**H**	**M**	**M**	**L**	**H**	**L**	**L**	**L**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Practical Components): Closed Book
2. Cooperative Learning Report, Assignment, Group Discussion, project Report, Field Visit Report, Seminar.
3. Pre/Post Test, Viva, Report for each Exercise.
4. Lab Model Examination & End Semester Practical Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator : Dr. B. Karthikeyan

CORE PRACTICAL VII: SOCIAL MEDIA ANALYTICS LAB			
Semester	III	Hours/Week	4
Course Code	P19DS3P7	Credits	3

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

CO#	Course Outcome	Level	Activity
CO1	Create data analytics systems using the data crawled from Twitter	K6	1 - 4
CO2	Create data analytics systems using the data crawled from Facebook	K6	5, 6
CO3	Create data analytics systems using the data crawled from Linkedin	K6	7
CO4	Create data analytics systems using the data crawled from GitHub	K6	8, 9
CO5	Create data analytics systems using the data crawled from Instagram	K6	10, 11
CO6	Create data analytics systems on bigdata collections	K6	11 - 15

2. SYLLABUS

Activity	Lab Activity Description
1	Real time crawling of tweets from Twitter and predict trending words
2	Extracting text, screen names, and hashtags from tweets. Generating histograms of words, screen names, and hashtags from tweets
3	Sentiment analysis using nltk.sentiment
4	Creating a basic frequency distribution from the words in tweets. Also, finding the most popular tweets in a collection of tweets
5	Counting the total number of page fans from Facebook. Retrieving the Last N items from the feeds of a Facebook Page
6	Finding the number of likes, shares, and comments on a given Facebook post
7	Retrieving your LinkedIn profile and print your last name. Performing Clustering your LinkedIn network based on locations of your connections
8	Finding a list of people who have bookmarked a GitHub repo
9	Computing the degree, betweenness, and closeness centrality measures of a graph
10	Displaying your profile picture from Instagram. Displaying the data of the most recent of your Instagram post
11	Objects detection from images from Instagram posts
12	Using USA Airline flight dataset, perform the following tasks • Install NetworkX package • Display the head (top-5 rows) using DataFrame • Display the nodes and edges • Plot the graph
13	Using USA Airline flight dataset, find the shortest path based on the airtime between the airports AMA and PBI
14	Developing a Movie Recommender System that suggests movie IDs that are most similar to a particular movie ID • Display the head (top-5 rows) of DataFrame • Display mean rating of all movies • Display count rating of all movies • Plot the graph of ratings column
15	Developing a Movie Recommender System that suggests movie IDs that are most similar to a particular movie ID

	• Analyze the correlation of two movies • Suggest similar movies for a given movie

Topics for Self Study

S.No	Topic Title	Web Link
1	Network analysis code and data	www.cl.cam.ac.uk/~dm754/stna/stna-examples.zip
2	NodeXL	http://nodexl.codeplex.com/
3	Pajek	http://pajek.imfm.si/doku.php
4	Folium	https://folium.readthedocs.io/en/latest/
	Graph-Tool	https://graph-tool.skewed.de/

3. Specific Learning Outcomes

Activity#	Lab Activity	Learning Outcome	Level
1.	Twitter data analytics	Crawl tweets at real time from Twitter. Predict trending words from crawled tweets	K6
2.	Twitter data analytics	Extract text, screen names, and hashtags from tweets. Generate histograms of words, screen names, and hashtags from tweets	K6
3.	Twitter data analytics	Perform Sentiment analysis using nltk.sentiment	K6
4.	Twitter data analytics	Create a basic frequency distribution from the words in tweets. Also, find the most popular tweets in a collection of tweets	K6
5.	Facebook data analytics	Count the total number of page fans from Facebook. Retrieve the Last N items from the feeds of a Facebook Page	K6
6.	Facebook data analytics	Find the number of likes, shares, and comments on a given Facebook post	K6
7.	Linkedin data analytics	Retrieve your LinkedIn profile and print your last name. Perform Clustering your LinkedIn network based on locations of your connections	K6
8.	GitHub data analytics	Find a list of people who have bookmarked a GitHub repo	K6
9.	GitHub data analytics	Compute the degree, betweenness, and closeness centrality measures of a graph	K6
10.	Instagram data analytics	Display your profile picture from Instagram. Display the data of the most recent of your Instagram post	K6
11.	Instagram data analytics	Detect objects from images from Instagram posts	K6
12.	Bigdata analytics of Airline data	Find out nodes and edges	K6
13.	Bigdata analytics of Airline data	Find the shortest path	K6
14.	Design of recommender system for Movie data Part-1	Plot graph of ratings	K6
15.	Design of recommender	Analyze correlation of two movies and	K6

	system for Movie data Part-2	suggest similar movies for a given movie	

4.MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	**H**	**M**	**L**	**M**	**M**	**M**	**L**	**-**	**-**	**H**	**M**	**M**	**-**
CO2	**H**	**M**	**H**	**M**	**M**	**H**	**M**	**M**	**L**	**H**	**H**	**-**	**M**
CO3	**H**	**M**	**M**	**L**	**H**	**M**	**-**	**-**	**L**	**M**	**M**	**M**	**-**
CO4	**H**	**L**	**H**	**H**	**H**	**H**	**M**	**M**	**L**	**H**	**M**	**M**	**H**
CO5	**H**	**M**	**L**	**H**	**H**	**L**	**L**	**M**	**M**	**H**	**H**	**H**	**M**
CO6	**H**	**M**	**M**	**L**	**H**	**L**	**L**	**L**	**M**	**H**	**H**	**-**	**-**

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Practical Components): Closed Book
2. Cooperative Learning Report, Assignment, Group Discussion, project Report, Field Visit Report, Seminar.
3. Pre/Post Test, Viva, Report for each Exercise.
4. Lab Model Examination & End Semester Practical Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator : Dr. Janani Selvaraj

CORE PRACTICAL VIII: PRINCIPLES OF DEEP LEARNING LAB			
Semester	III	Hours/Week	5
Course Code	P20DS3P8	Credits	3

1. COURSE OUTCOMES

On successful completion of this course, students will be able to:

CO#	Course Outcome	Level	Activity
CO1	Develop Deep Neural Networks for image classification	K6	1
CO2	Develop Deep Neural Networks for object detection and classification	K6	2,3
CO3	Develop Deep Neural Networks for text analysis	K6	4 - 7
CO4	Develop Deep Neural Networks for word and sentence embedding	K6	8
CO5	Develop machine translation Deep Neural Networks	K6	10
CO6	Develop question answering and dialogue systems for business applications	K6	9, 11

2. LIST OF EXERCISES

Activity	Lab Activity Description
1	Image classification
2	Object detection
3	Object classification
4	Text classification
5	Sentiment analysis
6	Text summarization
7	Text entailment
8	Word and Sentence Embedding
9	Question Answering
10	Machine Translation
11	Dialogue Systems

Topics for Self Study

S.No	Topic Title	Web Link
1	Deep Learning by Andrew Ng	https://www.coursera.org/specializations/deep-learning
2	Introduction to Tensorflow for Deep Learning	https://www.udacity.com/course/intro-to-tensorflow-for-deep-learning--ud187
3	Keras repository for Deep Learning	http://keras.github.com
4	Voice recognition and translation	http://research.baidu.com/warp-ctc/

3. SPECIFIC LEARNING OUTCOMES

Activity#	Lab Activity	Learning Outcome	Level
1	Image classification	Build and deploy DL system for image classification using CNN	K6
2	Object detection	Build and deploy DL system for object detection	K6

3	Object classification	Build and deploy DL system for object classification	K6
4	Text classification	Build and deploy DL system for text classification	K6
5	Sentiment analysis	Build and deploy DL system for sentiment analysis using RNN	K6
6	Text summarization	Build and deploy DL system for text summarization	K6
7	Text entailment	Build and deploy DL system for text entailment	K6
8	Word and Sentence Embedding	Build and deploy DL system for word and sentence embedding	K6
9	Question Answering	Build and deploy DL system for question answering	K6
10	Machine Translation	Build and deploy DL system for machine translation	K6
11	Dialogue Systems	Build and deploy DL system for chatbots	K6

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	H	M	L	M	M	M	L	-	-	H	M	M	-
CO2	H	M	H	M	M	H	M	M	L	H	H	-	M
CO3	H	M	M	L	H	M	-	-	L	M	M	M	-
CO4	H	L	H	H	H	H	M	M	L	H	M	M	H
CO5	H	M	L	H	H	L	L	M	M	H	H	H	M
CO6	H	M	M	L	H	L	L	L	M	H	H	-	-

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Practical Components): Closed Book
2. Cooperative Learning Report, Assignment, Group Discussion, project Report, Field Visit Report, Seminar.
3. Pre/Post Test, Viva, Report for each Exercise.
4. Lab Model Examination & End Semester Practical Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator : Dr. K. Rajkuamr

Core X: Customer Relationship Management			
Semester	IV	Hours/Week	5
Course Code	P20DS410	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Outline the customer relationship development strategies.	K6	I
CO2	Measure customer satisfaction level for the given scenario.	K5	II
CO3	Evaluate the CRM strategy for given Scenario	K6	III
CO4	Explain the functionalities of Sales Force Automation	K5	IV
CO5	Determine the Software application for Customer Service.	K6	V
CO6	Estimate and Analyze CRM Trends	K6	V

2. A. SYLLABUS

Unit-1. Introduction

Introduction to Customer Relationship Management: Concept- Evolution of Customer Relationships: Customers as strangers- acquaintances- friends and partners. Objectives- Benefits of CRM to Customers and Organizations- Customer Profitability Segments- Components of CRM: Information- Process- Technology and People-Barriers to CRM. Relationship Marketing and CRM: Relationship Development Strategies: Organizational Pervasive Approach- Managing Customer Emotions- Brand Building through Relationship Marketing- Service Level Agreements- Relationship Challenges.

Unit-2. Marketing and Data Management

CRM Marketing Initiatives- Customer Service and Data Management: CRM Marketing Initiatives: Cross-Selling and Up-Selling- Customer Retention- Behaviour Prediction-Customer Profitability and Value Modeling- Channel Optimization- Personalization and Event-Based Marketing. CRM and Customer Service: Call Center and Customer Care: Call Routing- Contact Center Sales-Support- Web Based Self Service- Customer Satisfaction Measurement- Call-Scripting- Cyber Agents and Workforce Management. CRM and Data Management: Types of Data: Reference Data-Transactional Data-Warehouse Data and Business View Data- Identifying Data Quality Issues-Planning and Getting Information Quality- Using Tools to Manage Data- Types of Data Analysis: Online Analytical Processing (OLAP) - Clickstream Analysis- Personalization and Collaborative Filtering- Data Reporting.

Unit-3. Strategy

CRM Strategy- Planning: Understanding Customers: Customer Value- Customer Care-Company Profit Chain: Satisfaction- Loyalty- Retention and Profits. Objectives of CRMStrategy- The CRM Strategy Cycle: Acquisition- Retention and Win Back- Complexities of CRM Strategy.

Unit-4. Implementation and Evaluation

CRM Implementation and Evaluation: Planning and Implementation of CRM: Business to Business CRM- Sales and CRM- Sales Force Automation- Sales Process/ Activity Management- Sales Territory Management- Contact Management- Lead Management-Configuration Support-

Knowledge Management CRM Implementation: Steps- Business Planning- Architecture and Design- Technology Selection- Development- Delivery and Measurement.

Unit-5. Evaluation

CRM Evaluation: Basic Measures: Service Quality- Customer Satisfaction and Loyalty-Company 3E Measures: Efficiency- Effectiveness and Employee Change. CRM New Horizons: e-CRM: Concept- Different Levels of E- CRM- Privacy in E-CRM -Software App for Customer Service:# Activity Management- Agent Management- Case Assignment- Contract Management- Customer Self Service- Email Response Management- Escalation- Inbound Communication Management- Invoicing- Outbound Communication Management- Queuing and Routing- Scheduling - Social Networking and CRM - Mobile-CRM - CRM Trends- Challenges and Opportunities - Ethical Issues in CRM.

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Digital Media and Marketing Strategies	https://www.coursera.org/learn/marketing-plan
2	Salesforce in easy steps	https://www.udemy.com/course/learn-salesforce-in-easy-steps-and-get-certified/
3	Customer Relationship Management for Marketers	https://www.edx.org/professional-certificate/iimbx-customer-relationship-management-for-marketers
4	Analyzing Your CRM Data with Reports and Dashboards	https://www.youtube.com/watch?v=lqtiGJlI6t0

C. TEXT BOOK

1. Baran Roger J. & Robert J. Galka. Customer Relationship Management: The Foundation of Contemporary Marketing Strategy, Routledge Taylor & Francis Group. 2014

D. REFERENCE BOOK

1. Anderrson Kristin and Carol Kerr. Customer Relationship Management. Tata McGraw-Hill, 2002.
2. Ed Peelen. Customer Relationship Management. Prentice Hall, 2005.
3. BhasinJaspreet Kaur. Customer Relationship Management. Dreamtech Press, 2012
4. Valarie A Zeithmal, Mary Jo Bitner, Dwayne D Gremler and Ajay Pandit. Services Marketing Integrating Customer Focus Across the Firm. Tata McGraw Hill, 2010.
5. UrvashiMakkar and Harinder Kumar Makkar. CRM Customer Relationship Management. McGraw Hill Education, 2013.

E. WEBLINKS

- https://www.edx.org/professional-certificate/iimbx-customer-relationship-management-for-marketers
- https://www.coursera.org/learn/marketing-plan

3. SPECIFIC LEARNING OUTCOMES(SLO)

Unit/ Section	Course Content	Learning outcomes	Level
I	**Introduction to Customer Relationship Management**		
1.1	Concepts in CRM	Summarize the concepts of CRM.	K2
1.2	Evolution of Customer Relationships	Recall the evolution of CRM	K1
1.3	Customers as strangers,	Explain the exchange of	K5

	acquaintance, friends and partners	relationships between providers and customers.	
1.6	Objectives	List the objectives of CRM	K1
1.7	Benefits of CRM to Customers and Organizations	Recall the benefits of CRM	K1
1.8	Customer Profitability Segments	Analyze the ways to approach profitability segmentation to help organizations focus their efforts .	K4
	Components of CRM		
1.9	Information, Process, Technology and People	Summarize the four components of a successful business strategy.	K2
1.10	Barriers to CRM	Identify the barriers to CRM for the given scenario.	K3
	Relationship Marketing and CRM		
1.11	Relationship Development Strategies	Outline the relationship development strategies.	K1
1.12	Organizational Pervasive Approach	Explain the pervasive approach to manage the customers.	K5
1.13	Managing Customer Emotions	Identify the types of emotions exhibited by a customer.	K3
1.14	Brand Building through Relationship Marketing	Identify the key elements that help to maximize the brand loyalty for given scenario.	K3
1.15	Service Level Agreements	Illustrate the components of agreement.	K2
1.16	Relationship Challenges	Recognize the key challenges of CRM.	K2
II	**Marketing and Data Management**		
	CRM Marketing Initiatives- Customer Service and Data Management		
2.1	Cross-Selling and Up-Selling	Compare cross-selling and up-selling in sales	K4
2.2	Customer Retention	Calculate the customer retention rate for given scenario.	K3
2.3	Behaviour Prediction	Predict customer behavior pattern	K2
2.4	Customer Profitability and Value Modeling	List the uses of Customer value model	K1
2.5	Channel Optimization	Apply channel optimization strategies for given scenario	K3
2.6	Personalization and Event-Based Marketing	Identify how personalized marketing strategies can be implemented for given scenario.	K3
	CRM and Customer Service:Call Center and Customer Care		
2.7	Call Routing	List the types of call routing.	K1
2.8	Contact Center Sales-Support	Summarize the services of contact centre.	K2
2.9	Web Based Self Service	Outline the benefits of self service in CRM.	K2
2.10	Customer Satisfaction Measurement	Measure customer satisfaction level for the given scenario	K5
2.11	Call-Scripting	Compose call scripting	K6
2.12	Cyber Agents and Workforce Management	List the function of cyber agent and workforce management module.	K1
	CRM and Data Management		

	Types of Data		
2.13	Transactional Data, Warehouse Data and Business View Data	Explain the types of Data.	K2
2.14	Identifying Data Quality Issues	Identify data quality issues in given scenario	K3
2.15	Planning and Getting Information Quality	Summarize the strategies to improve the quality of information.	K2
2.16	Using Tools to Manage Data	List the CRM tools used to manage data.	K1
	Types of Data Analysis		
2.17	Online Analytical Processing (OLAP)	Explain online analytical processing.	K5
2.18	Clickstream Analysis	Analyze clickstream.	K4
2.19	Personalization and Collaborative Filtering	Explain collaborative filtering.	K2
2.20	Data Reporting	Create effective data reports	K6
III	**Strategy**		
	CRM Strategy- Planning: Understanding Customers		
3.1	Customer Value	Estimate the customer value	K5
3.2	Customer Care	Outline the importance of customer care.	K2
	Company Profit Chain		
3.3	Satisfaction	Analyze the effects of satisfaction on profit and growth.	K4
3.4	Loyalty	List the types of loyalty.	K1
3.5	Retention and Profits	Explain the importance of retention	K2
3.6	Objectives of CRM Strategy	List the objectives of CRM strategy	K1
	The CRM Strategy Cycle		
3.7	Acquisition	Apply acquisition strategies for the given scenario	K3
3.8	Retention	Apply the retention strategies in given scenario	K3
3.9	Win Back	Make use of Win Back strategy for given scenario.	K3
3.10	Complexities of CRM Strategy	Analyze the complexities of CRM Strategy.	K4
IV	**Implementation and Evaluation**		
	Planning and Implementation of CRM		
4.1	Business to Business CRM	Summarize the purpose of B2B CRM	K2
4.2	Sales and CRM	Identify the factors of CRM in improving sales.	K3
4.3	Sales Force Automation	Explain the functionalities of SFA.	K5
4.4	Sales Process/ Activity Management	List the features of sales process/activity management module.	K2
4.5	Sales Territory Management	Assess the ways to create sales Territory	K5
4.6	Contact Management	List the features of contact Management module.	K2
4.7	Lead Management Configuration	Outline the importance of lead management configuration.	K2
4.8	Support- Knowledge	Summarize the uses of support-	K2

	Management	knowledge management module.	
	CRM Implementation		
4.9	Steps	List the steps in implementing CRM	K1
4.10	Business Planning	Discuss the business plan for implementing CRM	K6
4.11	Architecture and Design Technology Selection	Outline the importance of technology in implementing CRM	K2
4.12	Development, Delivery and Measurement	Measure the performance of CRM.	K5
V	**Evaluation**		
	Basic Measures		
5.1	Service Quality, Customer Satisfaction and Loyalty	Estimate the basic measures of CRM	K6
	Company 3E Measures		
5.2	Efficiency- Effectiveness and Employee Change	Explain the 3E measure of CRM.	K5
	CRM New Horizons: e-CRM		
5.3	Concept	Summarize the key concepts of e-CRM.	K2
5.4	Different Levels of E- CRM	Explain different levels of E CRM	K2
5.5	Privacy in E-CRM	List the privacy principles of E CRM	K1
	Software App for Customer Service		
5.6	Activity Management	List the features of activity management module.	K2
5.7	Agent Management	Summarize the functions of agent module.	K2
5.8	Case Assignment	Outline the features of Case Assignment module.	K2
5.9	Contract Management	Discuss the functions of contract management module.	K2
5.10	Customer Self Service	List the uses of customer self service module.	K2
5.11	Email Response Management	Outline the features of email response management module.	K2
5.12	Escalation	Utilize the functions of escalation module.	K3
5.13	Inbound Communication Management	Discuss the uses of inbound communication module.	K2
5.14	Invoicing	List the features of invoicing module.	K2
5.15	Outbound Communication Management	Discuss the uses of outbound communication module.	K2
5.16	Queuing and Routing, Scheduling	Explain the functions of queuing and routing module.	K2
5.17	Social Networking and CRM	Identify the ways to integrate social networking with CRM.	K3
5.18	Mobile-CRM	List the benefits of Mobile CRM	K2
5.19	CRM Trends	Analyze CRM trends.	K4
5.20	Challenges and Opportunities	List the challenges and Opportunities of CRM	K2
5.21	Ethical Issues in CRM	Summarize the ethical issues in	K2

		CRM	

4. MAPPING (CO, PO, PSO)

L-Low M-Moderate H- High

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1				**M**		**M**	**M**		**M**	**H**			
CO2	**M**	**M**	**H**		**M**	**H**				**H**		**H**	
CO3			**M**			**M**	**M**			**H**		**H**	**H**
CO4						**M**				**H**		**H**	**H**
CO5				**M**		**H**				**H**			**H**
CO6	**H**	**H**	**H**	**M**	**H**	**H**			**H**	**H**			

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator : Prof. D. Indra Devi

ELECTIVE-5: SUPPLY CHAIN MANAGEMENT			
Semester	IV	Hours/Week	5
Course Code	P20DS4:5	Credits	4

1. COURSE OUTCOMES

After the successful completion of this course the students will be able to

S.No.	Course Outcomes	Level	Unit
CO1	Perceivethe foundations of a supply chain and explore strategies and logistics drivers by which the supply chain of an organization can be managed to enhance its business competitiveness.	K6	I
CO2	Evaluate and Analytically examine the strategic drivers and metrics of supply chain organizations and measure performance improvement	K6	II
CO3	Design and provide a network to support the business decision-making within the context of supply chain management and the real world.	K5	III
CO4	Plan optimized transportation and logistics activities in supply chain operations	K6	IV
CO5	Determine the outsourcing decisions by applying the buy-make framework to manage the benefit and risks of outsourcing	K6	V
CO6	Recommend a proper blend of Logistics and Supply elements thereby confining the wide range of applications in the changing dynamic environment and industry practices	K6	V

2. A. SYLLABUS

Unit-1. Building strategic framework

Supply chain: Definition, 3 streams of knowledge, objectives and importance - Decision phases and process views of a supply chain (SC) - Examples of supply chain - Competitive strategy and SC strategy - 3 steps of achieving strategic fit - Improving SC performance by expanding scope of strategic fit, challenges to achieving strategic fit - Financial measures and drivers of SC performance - Logistical drivers: Roles in SC and decision components - Cross functional drivers: Roles in SC and decision components - Role of infrastructure in SC performance.

Unit-2. Designing SC network

Key factors influencing distribution network design - Design options for a distribution network - Impact of online sales on customer service and cost - Network design decisions: Influencing factors, framework - Capacitated plant location model for network optimization - Gravity location model for network design - Model for demand allocation and locating plants - Global supply chain: Dimensions to evaluate total cost, SC risks, tailored risk mitigation strategies - Discounted cash flow analysis to evaluate network design decision - Decision tree analysis: Basics, Evaluating flexibility at Trip Logistics.

Unit-3. Planning and coordinating demand and supply

Demand forecasting: role, characteristics, components and methods - Static demand forecasting methods - Adaptive demand forecasting methods - Measures of demand forecasting error - Aggregate planning: role, identifying aggregate units, strategies - Aggregate planning using Linear programming - Managing supply and demand to improve synchronization in SC - Lack of SC coordination: Bullwhip effect, effect on performance - Obstacles to coordination in SC - Managerial levers to achieve coordination of demand and supply in SC.

Unit-4. Planning and managing inventories

Cycle inventory terminologies: Lot size, Average flow time, Inventory holding cost, Ordering cost - Computing optimal lot size for single product: Economic order quantity, for Production environment, with Capacity constraint - Lot size based discount schemes: All unit quantity discounts, Marginal unit quantity discount - Trade promotions: Goals, Forward buying, Impact on lot size and cycle inventory - Factors affecting the level of safety inventory - Evaluating required safety inventory: Given a replenishment policy, Desired cycle service level, Desired fill rate - Impact of desired product availability and uncertainty on safety inventory - Impact of supply uncertainty on safety inventory - Factors affecting optimal level of product availability - Managerial levers of inventory to improve SC profitability.

Unit-5. Transportation and cross functional drivers

Modes of transportation in SC - Design options for a transportation network - Transportation and inventory cost trade off - Transportation cost and customer responsiveness trade off - Tailored transportation - Sourcing decisions: In house or Outsource - Sharing risk and reward in SC - Pricing and revenue management for multiple customer segments - Pricing and revenue management for perishable assets - Pricing and revenue management for seasonal demand.

B. TOPICS FOR SELF-STUDY

S.No.	Topics	Web Links
1	Digitization of Supply chain	https://www.coursera.org/lecture/process-improvement/lecture-4-1-digitization-of-the-supply-chain-EFofn
2	Supply chain Analytics	https://nptel.ac.in/courses/110/108/110108056/
3	Artificial Intelligence in Supply Chain Management	https://towardsdatascience.com/artificial-intelligence-in-supply-chain-management-predictive-analytics-for-demand-forecasting-80d2d512f155
4	Logistics and Supply chain Management	http://slmt.in/courses/cilt-international-courses/diploma-in-logistics-and-supply-chain-management-dlsm/

C. TEXT BOOKS

1. Sunil Chopra, Peter Meindl and DV Karla. “Supply Chain Management: Strategy, planning and operation”, 6th edition, Pearson, 2016. ISBN 978-9332548237 (Excluding Excel Examples)

D. REFERENCES BOOKS

1. David Simchi-Levi and Philip Kaminsky. “Designing and managing the supply chain: Concepts, strategies and case studies”, 3rd edition, McGraw Hill, 2007.

E. WEB LINKS

1. http://www.supply-chain.com
2. http://www.transportlink.com
3. http://www.transportlaw.com
4. http://www.apics.org
5. http://www.clm1.org
6. http://www.napm.org

3.SPECIFIC LEARNING OUTCOMES (SLO)

Unit/ Section	Course Content	Learning outcomes	Level
I	**Building strategic framework**		
1.1	Supply chain: Definition, 3 streams of knowledge, objectives and importance	Discuss the goal of supply chain and impact of supply chain decision on success of the firm.	K2

		Describe the various objectives of supply chain	K2
		Illustrate the importance of supply chain management	K3
		Identify the supply chain obstacles.	K4
		Determine Strategic, operational and tactical planning of supply chain	K6
1.2	Decision phases and process views of a supply chain (SC)	Explain decision phases in supply chain	K4
		Describe the cycle and push/pull view of a supply chain	K4
		Determine the underlying theoretical logic for make versus buy decision	K6
1.3	Examples of supply chain	Discuss in what way do supply chain flows affect the success or failure of a firm such as Amazon and list two supply chain decisions that have a significance impact on supply chain Profitability.	K5
1.4	Competitive strategy and SC strategy	Analyze the strategies that are critical to achieving strategic fit for company's overall success.	K4
1.5	3 steps of achieving strategic fit	Explain 'Achieving Strategic Fit' in supply chains with the help of a suitable example.	K4
1.6	Improving SC performance by expanding scope of strategic fit, challenges to achieving strategic fit	Choose strategic fit between its supply chain strategy and its competitive strategy	K5
		Explain the Balanced Score Card approach of supply chain performance measurement.	K2
1.7	Financial measures and drivers of SC performance	Apply the key metrics that track the performance of the supply chain in terms of each driver.	K4
1.8	Logistical drivers: Roles in SC and decision components	Identify the role of major drivers in supply chain	K4
1.9	Cross functional drivers: Roles in SC and decision components	Recommend the ways to boost up the cross functional drivers roles in SC	K6
		Explain the barriers of cross functional drivers	K2
		Analyze cross functional management is effectively managing supply chains	K4
1.10	Role of infrastructure in SC performance	Debate Economic impact of inadequate infrastructure for sc integration	K6
II	**Designing SC network**		
2.1	Key factors influencing distribution network design	Explain the factors influencing distribution network design	K2
		Choose the type of distribution network is typically best suited for commodity items	K6
2.2	Design options for a distribution network	Examine the design options available for a distribution network with option in detail	K3
		Design a suitable distribution network utilized for the specialty chemical company is considering expanding its	K5

		operations into Brazil, when five companies dominate the consumption of specialty chemicals.	
		Construct the role of network design decision in a supply chain	K5
		Plan different design options available for a distribution network with option in detail	K5
2.3	Impact of online sales on customer service and cost	Predict the impact of online sales on consumers and firms. Give evidence from consumer electronics	K5
		Justify is e-business likely to be more beneficial in the early part or the mature part of a product's life cycle.	K6
		Explain the cycle and push/pull view of a supply chain.	K4
2.4	Network design decisions: Influencing factors, framework	Describe planning Networks	K2
		Interpret the objectives & process of Supply Chain Network optimization models	K2
		Asses the outcome and benefits of Supply Chain Network optimization models.	K6
		Analyze the benefits are these using bar codes and scanners for orders entry as opposed to keyboard encoding into a computer database	K4
		Describe the current trends in value addition happened in Indian companies	K2
		Identify factors influencing supply chain network decisions.	K4
		Outline the advantages and disadvantages of distribution network design options	K2
		Propose factors to be considered in deciding whether to make and supply or to buy and supply for blood pressure measuring kits for hospitals in developing rural markets in India.	K5
2.5	Capacitated plant location model for network optimization	Explain optimized network.	K2
		Discuss the various Network optimization models	K2
		Construct the classification of supply chain network design decisions	K5
		Argue the following statement "Some industries are located near the source of raw materials, whereas some near the markets for finished goods"	K6
2.6	Gravity location model for network design	Design network decisions using decision tree and list its importance	K5
		Devise a Framework to make a network design decision	K5
		Identify factors to be considered for locating a centralized kitchen to cook	K4

		food for a restaurant chain. Also suggest an appropriate facility location model. State the assumption if any	
		Discuss the optimization models used for facility location and capacity allocation	K2
2.7	Global supply chain: Dimensions to evaluate total cost, SC risks, tailored risk mitigation strategies	Determine the role of a third party in increasing the supply chain surplus	K6
		Describe global supply chain risk management strategies	K2
		Determine the total cost approach to supply chain risk modelling	K6
		Identify the methods to managing risk to avoid supply chain breakdown	K4
		Discuss the Strategies for supply chain risk management	K4
2.8	Discounted cash flow analysis to evaluate network design decision	Outline uncertainty in network design discounted cash flow analysis	K2
		Determine the uncertainties and risk factors so important in evaluating supply chain design decisions	K6
2.9	Decision tree analysis: Basics, Evaluating flexibility at Trip Logistics	Write the features of decision tree.	
		Asses the benefits of using decision nodes by decision making under uncertainty	K6
		Explain the formation of a decision tree based on the Trips logistics	K2
III	**Planning and coordinating demand and supply**		
3.1	Demand forecasting: role, characteristics, components and methods	Examine the basic approaches to demand forecasting	K3
		Predict the forecast error if demand in 5 tons out to be 125 litres for a grocery store has experienced a weekly demand of oil of 120,127,114,and 122 litres over the last 4 weeks. Forecast demand for period 5 using a four period moving average.	K5
		Evaluate the number of computers the store manager should order in each replenishment lot. Demand for computers in a store is 12,000 units per year. The store incurs a fixed order placement, trasnsportation and receiving cost of Rs.40,000/- each time an order is placed. Each computer costs the store Rs.5000/- and the holding cost is 20%. Also explain the impact of supply chain uncertainty on safety inventory	K6
		Asses the role does forecasting play in the supply chain of a build-to-order manufactures such as dell	K6
		Determine the forecast error if demand in period 5 turns out to be 125 gallons for a super market has experienced weekly demand of milk of 120,127,114 and 122 gallons over the last four weeks. Forecast	K6

		demand for period 5 using a four –period moving average.	
3.2	Static demand forecasting methods	Classify the static and adaptive forecasting methods	K3
		Explain the basic, six step approach to help an organization perform effective forecasting	K4
		Investigate Demand forecasting analysis using time series methods	K4
3.3	Adaptive demand forecasting methods	Determine the house old electricity demand forecasting using adaptive conditional density estimation	K6
		Formulate the adaptive water demand forecasting for near real time management of smart water distribution system	K5
3.4	Measures of demand forecasting error	Invent demand forecast accuracy and forecast error	K5
3.5	Aggregate planning: role, identifying aggregate units, strategies	Outline the operational parameters to identify aggregate plan	K2
		Select the major cost categories needed as input for aggregate planning	K6
		Identify the managerial levers that reduce lot size and cycle inventory in a supply chain without increasing cost.	K4
3.6	Aggregate planning using Linear programming	Explain the role of collaborative planning and forecasting for efficient execution of supply chains.	K6
		Illustrate the role predictive visibility supply chain performance.	K3
		Explain the different types of costs associated with aggregate planning. For each of the cost, enumerate the areas where the cost plays an important role.	K4
		Discuss the major cost categories needed as input for aggregate planning	K6
		Solve aggregate planning using Linear Programming	K3
3.7	Managing supply and demand to improve synchronization in SC	Investigate the Synchronization in supply chains implications for design and management	K4
3.8	Lack of SC coordination: Bullwhip effect, effect on performance	Write a note on the Coordination in a supply chain.	K1
		Analyze the Bullwhip effect in supply chain for the effect on performance	K4
3.9	Obstacles to coordination in SC	List the various obstacles to coordination and how such obstacles can be minimized in supply chain	K1
3.10	Managerial levers to achieve coordination of demand and supply in SC	Design the managerial levers that help to achieve coordination in the supply chain	K5
IV	**Planning and managing inventories**		
4.1	Cycle inventory	Evaluate the number of cartridges that the	K6

	terminologies: Lot size, Average flow time, Inventory holding cost, Ordering cost	store manager should order in each replenishment lot for Demand of cartridges in an electronic store is 1000 units per month. The firm incurs a fixed order placement, transportation and receiving costs of Rs.4000/- each time an order is placed. Each cartridge costs Rs.500/- and the retailer has a holding cost of 20 percent.	
		Explain multi-echelon inventory management in detail with the help of a suitable example.	K4
		Explain how to manage supply chain cycle inventory.	K4
		Construct the role of cycle inventory in a supply chain and how uncertainty in the supply chain managed	K5
4.2	Computing optimal lot size for single product: Economic order quantity, for Production environment, with Capacity constraint	Show how to compute the optimal lot size and cycle length for the given sequence of items in a cycle	K3
		Investigate optimal lot sizes in the economic lot scheduling for production environment	K4
4.3	Lot size based discount schemes: All unit quantity discounts, Marginal unit quantity discount	Distinguish the lot size based and volume based quantity discounts	K2
		Analyze the effect of quantity discounts on lot size and cycle inventory	K4
		Examine the effect of trade promotions on lot size and cycle inventory	K4
4.4	Trade promotions: Goals, Forward buying, Impact on lot size and cycle inventory	Describe how to Managing Multi-Echelon Cycle Inventory	K2
		Explain the impact of trade promotions on lot size and cycle inventory	K4
4.5	Factors affecting the level of safety inventory	State and briefly explain the role of safety inventory in supply chain	K1
4.6	Evaluating required safety inventory: Given a replenishment policy, Desired cycle service level, Desired fill rate	Propose "Relevant deterministic and Stochastic Inventory Models"and explain its relevance in an organization. Also, briefly explain the important features of these models.	K5
4.7	Impact of desired product availability and uncertainty on safety inventory	Evaluate the impact of desired product availability and uncertainty on safety inventory	K6
4.8	Impact of supply uncertainty on safety inventory	Evaluate the Impact of supply uncertainty on safety inventory	K6
4.9	Factors affecting optimal level of product availability	Discuss optimal level of product availability	K2
		Determine the optimal level of product availability	K6
4.10	Managerial levers of inventory to improve SC profitability	Design the managerial levers that help to improve inventory SC profitability	K5
V	**Transportation and cross functional drivers**		
5.1	Modes of transportation in	Discuss the importance of transportation	K6

	SC	in supply chain.	
		Explain the modes of transportation and their performance characteristics	K4
5.2	Design options for a transportation network	Design an option for a transportation network	K5
5.3	Transportation and inventory cost trade off - Transportation cost and customer responsiveness trade off - Tailored transportation	Determine tradeoffs in transportation design network	K6
		Distinguish transportation cost, customer responsiveness tradeoffs and Tailored transportation	K2
5.4	Sourcing decisions: In house or Outsource - Sharing risk and reward in SC	Discuss the importance of in-sourcing and out-sourcing with suitable examples	K6
		Debate Strategic Alliances and Outsourcing	K6
		Describe the ways that a firm such as Wal-Mart form out sourcing decisions	K1
5.5	Pricing and revenue management for multiple customer segments	Explain the importance of pricing in supply chain management and elucidated various type of pricing approaches that generate maximum profit	K6
5.6	Pricing and revenue management for perishable assets	Design Perishable assets for pricing and revenue management	K5
5.7	Pricing and revenue management for seasonal demand.	Explain pricing and revenue management for seasonal demand	K4

4. MAPPING (CO, PO, PSO)

L-Low **M-Moderate** **H- High**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PSO1	PSO2	PSO3	PSO4
CO1	H												H
CO2	H	H	M	H	M	H	M		M	H			H
CO3	H	H	M	H	M	H	M	H	H	H	H	H	
CO4	H	M	M	M		H	M	H	M	H	H	H	M
CO5	H	H	H	M	M	H	M	H	M	H	H		
CO6	H	H	H	M	H	H	H	H	H	H	H		

5. COURSE ASSESSMENT METHODS

DIRECT:

1. Continuous Assessment Test: T1, T2 (Theory & Practical Components): Closed Book
2. Open Book Test.
3. Cooperative Learning Report, Assignment, Group Presentation, Group Discussion, project Report, Field Visit Report, Poster Presentation, Seminar, Quiz (written).
4. Pre-Semester & End Semester Theory Examination

INDIRECT:

1. Course end survey (Feedback)

Name of the Course Coordinator : Prof. D. Indra Devi