Progressive Education Society's
Modern College of Arts, Science & Commerce, (Autonomous)
Ganeshkhind, Pune 411016

Program Code: MScDS11
Master of Science (Data Science)
(Under Faculty of Science & Technology)

A.Y: 2025 – 2026

Opposite to University of Pune, Pashan Road 411016
e-Mail:moderncollege16@gmail.com
Phone: +917768050797 or +917768020797 Fax:(020) 25650931

Name of Program: Master of Science (Data Science)

Introduction:

M.Sc. (Data Science) is a Two Year Full Time Post Graduate Program. A Master of Science in Data Science is an interdisciplinary degree program designed to provide studies in scientific methods, processes, and systems to extract knowledge or insights from data in various forms, either structured or unstructured, similar to data mining. Our curriculum provides students with a rigorous course of study in big data technologies, applications and practices, a pathway for student internships and full-time employment. Students are prepared to meet the challenges at the intersection between big data, business analytics, and other emerging fields like Artificial intelligence, Deep learning, Neural networks etc.

Program Structure:

- The Program is a Two Year (Four semesters) Full Time Degree Program.
- The Program shall be based on a credit system comprising 88 credits.

Objectives:

1) To expose and provide a strong foundation to the students in the upcoming era of Data Science and Artificial Intelligence
2) The programme aims at providing a rigorous training in fundamental concepts of Statistics, Mathematics, & Computers Science which creates a strong knowledge base in Data Science domain.
3) To provide a complete understanding of the subject by introducing projects from the Second semester on the relevant subject.
4) Focus on blending theory with practical and industry application to enhance understanding and learning.
5) Focus on the overall development of the students to help gain knowledge and skillsets required for further studies after completion of the course.

Eligibility Criteria:

- Graduate degree in Statistics / Mathematics / Computer Science / Computer Application/ Engineering / Technology or any other discipline with a minimum of two years of learning Mathematics or statistics from a recognized university / institution with an equivalent qualification.

- English Language Proficiency.
- Intake: 30 Seats

Medium of Instruction: English

Instructions for Teachers for Internal Evaluation for 20 Marks and 40 Marks:

The purpose of internal evaluation is to assess the depth of knowledge, understanding and awareness. For this purpose, a teacher is expected to use different evaluation methods in order to have rational and objective assessment of the learners and available resources.

External Examination:

There will be a written Examination of 30 marks in 2 hrs. duration and 60 marks of 3 hrs. for every course at the end of each Semester only for major and Elective subjects.

Award of Class:

Letter Grade	Grade Point
O (outstanding)	10
A+ (Excellent)	9
A (Very good)	8
B+ (Good)	7
B (Above average)	6
C (Average)	5
P (Pass)	4
F (Fail)	0
Ab (Absent)	0

Question Paper Pattern:CIE

Max. Marks: 20 (Credit:02, Duration: 50 Min.)				
Question No.	**Question**	**No. of sub questions**	**Marks to each sub question**	**Total Question Marks**
1	Multiple Choice Questions 5	5	1	5
2	Define any 5	6	1	5
3	Attempt any two of the following	3	2	4
4	Attempt any two of the following	3	3	6
Total Marks:				**20**

Max. Marks: 20 (Credit 4, Duration: 50Min.)				
Question No.	**Question**	**No. of sub questions**	**Marks to each sub question**	**Total Question Marks**
1	Multiple Choice Questions 5	5	1	5
2	Define any 5	6	1	5
3	Attempt any two of the following	3	2	4
4	Attempt any two of the following	3	3	6
Total Marks:				**20**

For 2 credits CIE 20 Marks for internal Examination and 20 Marks for CBCS activity (Open book test, Seminars, Online Test, Surprise Test, Preparation of Models, Group Discussions etc.) average of 40 marks will be considered.

Question Paper Pattern: ESE

Max. Marks: 60 (Credit:04, Duration: 3 Hrs.)				
Question No.	**Question**	**No. of sub questions**	**Marks to each sub question**	**Total Question Marks**
1	Attempt All	10	1	10
2	Attempt any 5	7	2	10
3	Attempt any 5	7	3	15
4	Attempt any 3	5	5	15
5	Attempt any 2	4	5	10
Total Marks:				**60**

Max. Marks: 30 (Credit 2, Duration: 2 Hrs.)				
Question No.	**Question**	**No. of sub questions**	**Marks to each sub question**	**Total Question Marks**
1	Attempt All	5	1	5
2	Attempt any 5	7	2	10
3	Attempt any 2	4	5	10
4	Attempt any 1	2	5	5
Total Marks:				**30**

M.Sc. Data Science Curriculum & Syllabus

Programme Specific Outcomes (PSOs)	
After the successful completion of M.Sc. Data Science program the students are expected to	
PSO1	Understand fundamental concepts in statistics, mathematics and Computer Science.
PSO2	Apply and use various analysis tools and software used in data science
PSO3	Perform laboratory-oriented problem solving and be capable in data visualization and interpretation.
PSO4	provide solutions particular domain specific problems by having in depth domain knowledge.
PSO5	Utilize the data science theories for societal and environmental concerns.

Programme Outcomes (POs)	
On successful completion of the M.Sc. Data Science students will be able to	
PO1	Engage in continuous reflective learning in the context of technology and scientific advancement and in various subjects like mathematics, statistics, and data science
PO2	Develop mathematical & Statistical models and their computational solutions to reason and interpret Big data in wide range of domains
PO3	Acquire the skills in handling data science programming tools towards problem solving and solution analysis for domain specific problems.
PO4	Develop scalable techniques for data analysis and decision making in many areas, including machine learning, algorithms, statistics, operations research, databases, complexity analysis, visualization, and privacy and security
PO5	Innovation and Entrepreneurship:Produce innovative IT solutions and services based on global needs and trends.

MSc. (Data Science) Semester -I							
Course Type	**Course Code**	**Course Title**	**Credits**		**Evaluation**		
			TH	**PR**	**CIE**	**ESE**	**Total**
Major Mandatory (4)+(4)+(4)+(2)	DSC51101	Machine Learning	4		40	60	100
	DSC51102	Statistical Techniques for Data Scientist	4		40	60	100
	DSC51103	Computational Mathematics	4		40	60	100
	DSC51104	Practical based on ST & CM	-	2	20	30	50
DSE Electives	DSC51205	Artificial Intelligence	2	-	20	30	50
	DSC51206	Practical based on ML & AI	-	2	20	30	50
RM	DSC51307	Research Methodology	2	-	20	30	50
	DSC51308	Practical on Research Methodology	-	2	20	30	50
		Total	16	6	220	330	550
NON SGPA	HR101	Human Rights-I	1	-	--	25	25

MSc. (Data Science) Semester -II							
Course Type	**Course Code**	**Course Title**	**Credits**		**Evaluation**		
			TH	**PR**	**CIE**	**ESE**	**Total**
Major Mandatory (4)+(4)+(4)+(2)	DSC52101	Database Technologies	4		40	60	100
	DSC52102	Inferential Statistics	4		40	60	100
	DSC52103	Optimization Techniques	4		40	60	100
	DSC52104	Practical on DBT and IS	-	2	20	30	50
DSE Electives	DSC52205	Deep Learning & Neural Networks	2	-	20	30	50
	DSC52206	Project on Data Science	-	2	20	30	50
OJT	DSC52407	On Job Training	-	4	40	60	100
		Total	14	8	220	330	550
NON SGPA	HR102	Human Rights-II	1	--	--	25	25
	CS102	Introduction to Cyber Security	4	--	40	60	100

MSc. (Data Science) Semester -III							
Course Type	**Course Code**	**Course Title**	**Credits**		**Evaluation**		
			TH	**PR**	**CIE**	**ESE**	**Total**
Major Mandatory (4)+(4)+(4)+(2)	DSC63101	Predictive Analysis	4		40	60	100
	DSC63102	Natural Language Processing	4		40	60	100
	DSC63103	Cloud Computing	2		20	30	50
	DSC63104	Data Security & Ethics	2		20	30	50
	DSC63205	Practical on Predictive Analysis	-	2	20	30	50
DSE Electives	DSC63206	Business Information Studies	2	-	20	30	50
	DSC63507	Data Science Tool(Power BI)	-	2	20	30	50
RP	DSC63508	Research Project	-	4	40	60	100
		Total	14	8	220	330	550
NONSGPA	SK201	Skill Development-I	--	2	20	30	50

MSc. (Data Science) Semester -IV							
Course Type	**Course Code**	**Course Title**	**Credits**		**Evaluation**		
			TH	**PR**	**CIE**	**ESE**	**Total**
Major Mandatory (12)	DSC64101	Full time Industrial Training	-	12	120	180	300
DSE Electives	DSC64202	MOOC/Online Courses	4	-	40	60	100
RP	DSC64503	Industrial Project Design	-	6	60	90	150
		Total	4	18	220	330	550
NON SGPA	SK202	Skill Development-II	--	2	20	30	50


First
Semester

Semester- I
Course Type: Major Paper1 (Theory)
Course Code: DSC51101Course Title: - Machine Learning

Teaching Scheme: 4 Hours / Week	**No. of Credits: 4**	**No of Lectures: 60**	**Examination Scheme: CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Familiarity with Probability Theory, Multivariable Calculus, Linear Algebra
- Programming in Python (NumPy, SciPy, Pandas, Matplotlib, Seaborn, SciKit-Learn, Statistics Model, Tensor Flow)

The main objectives of this course are to:

- To introduce students to the basic concepts and techniques of Data Mining and Data Warehousing and Machine Learning.
- To study machine algorithms for solving practical problems.
- To write python programs using machine learning algorithms for solving practical problems

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Identify the key processes of data mining, data warehousing and knowledge discovery and machine learning.	B1,
CO2	Design data warehouse with dimensional modeling and apply OLAP operations	B2, B6
CO3	To introduce knowledge of machine learning and Identify appropriate machine learning algorithms to solve real world problems	B4,B5,B3

B1 - Remember; **B2** - Understand; **B3** - Apply; **B4** - Analyze; **B5** - Evaluate;**B6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Introduction to Machine Learning .** ● Concepts of Data Science ● Machine Learning ● Why learn and what is learning? ● What is Machine Learning? ● Traditional Programming Vs. Machine Learning ● Machine Learning Process ● Types of Data ● Key Elements of Machine Learning (Representation, Evaluation and Optimization) ● Feature Selection techniques (filter and wrapper method). ● Descriptive and Inferential Statistics: Probability, Normal Distribution, ● Distance Measures (Euclidean and Manhattan), Correlation Techniques, Introduction to Hypothesis Testing. ● Creating our own dataset, Importing the dataset, Handling Missing Data, Cross validation methods, Feature Scaling (Normalization and standardization).	10	CO1

	• Type of Learning- Supervised, Unsupervised and Semi-Supervised Learning		
2	**Introduction to Data Mining** • Definition Data mining and difference between Data mining and SQL • Data Mining issues • Stages of the Data Mining Process (KDD) • Data Mining Techniques/Tasks • Knowledge Representation Methods • Applications of Data mining • Data Pre-processing o Data Cleaning o Data Transformation o Data Reduction o Data Discretization	5	CO1
3	**Introduction to Data Warehousing** • Introduction to Data Warehouse • Data Warehouse Architecture and its components • Data Modeling with OLAP o Introduction o Difference between OLTP and OLAP o Data Mart o Fact Table, Dimension Table, OLAP cube o Different OLAP Operations • Schema Design o Introduction o Star and Snow-Flake Schema • Introduction to Machine Learning • Introduction to Pattern Matching • Case study based on Schema Design	10	CO1,CO2
4	**Supervised Machine Learning** • Introduction to Supervised Machine Learning • Regression in Machine Learning: o Types of Regression Algorithms-Linear Regression, Polynomial. o Evaluation methods for regression: MAE, RMSE, R-Square error • Classification in ML: Classification Algorithms o Logistic Regression o Decision Tree o k-Nearest Neighbors o Naive Bayes o Support Vector Machines (SVM) o Evaluation methods: Confusion matrix, precision, recall, F1-score, Accuracy, ROC-AUC curve, • Advantages and Disadvantages of supervised ML	14	CO3

5	**Unsupervised Machine learning** ● Introduction to Unsupervised Learning o Clustering Methods o Density-Based Methods o Hierarchical-Based Methods o Partitioning-Based Methods o Grid-Based Methods ● Clustering Algorithms o K-means clustering o Hierarchical Clustering (Agglomerative, Divisive), Dendrogram o Selecting optimal number of clusters: Within Clusters Sum of Squares (WCSS) by Elbow Method, Silhouette Score ● Introduction to Association rules o Association Algorithms- Apriori Algorithm o Dimensionality Reduction – Singular value decomposition, Principle Component Analysis(PCA) ● Anomaly detection (outlier detection) ● Independent Component Analysis ● Advantages and Disadvantages of Unsupervised ML	14	CO3
6	**Ensemble Learning** ● Model Combination Schemes ● Bagging: Random Forest Trees ● Boosting: Adaboost, Gradient boost ● Voting ● Stacking ● Error-Correcting Output Codes ● Gaussian mixture models ● The Expectation-Maximization (EM) Algorithm	7	CO3

Reference Books

1. Data Mining: Concepts and Techniques, Han, Elsevier ISBN:9789380931913/ 9788131205358
2. Margaret H. Dunham, S. Sridhar, Data Mining – Introductory and Advanced Topics, Pearson Education
3. Tom Mitchell, ―Machine Learning‖, McGraw-Hill, 1997
4. Christopher M. Bishop, ―Pattern Recognition and Machine Learning‖, Springer 2006
5. Raghu Ramkrishnan, Johannes Gehrke, Database Management Systems, Second Edition, McGraw Hill International
6. Ian H. Witten, Eibe Frank Data Mining: Practical Machine Learning Tools and Techniques
7. Mitchell, Tom M."Machine learning.WCB."(1997).
8. Alpaydin,E.2010. Introduction to Machine Learning. 2 nd edition, MIT.
9. Ethem Alpaydin: Introduction to Machine Learning, PHI 2 nd Edition-2013.
10. Andreas C. Müller & Sarah Guido: Introduction to Machine Learning with Python A Guide for Data Scientists- O'Reilly publications

11. Rogers, Simon, and Mark Girolami. A first course in machine learning.CRCPress,2015.
12. Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. The elements of statistical learning.Vol.1. Springer, Berlin: Springer series in statistics, 2001.
13. Witten, Ian H., and Eibe Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2005.

Related Online Content (MOOCS, SWAYAM, NPTEL, WEBSITES etc.)

- http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-KaufmannSeries-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-DataMining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf
- https://data-flair.training/blogs/data-mining-tutorial/
- http://infolab.stanford.edu/~ullman/mmds/book.pdf
- https://archive.org/details/IntroductionToDataMining
- https://www.academia.edu/6489220/Data_Mining_ebook

Online Courses:

Data Mining : https://onlinecourses.swayam2.ac.in/cec19_cs01/preview
https://onlinecourses.nptel.ac.in/noc20_cs12/preview
https://www.classcentral.com/course/swayam-data-mining-13982
Data Science For Engineers : https://onlinecourses.nptel.ac.in/noc19cs60
Introduction to Data Science in Python: https://www.coursera.org/learn/python-data-analysis
Python for Data Science : https://onlinecourses.nptel.ac.in/noc21cs33

.

Semester- I
Course Type: Major Paper2 (Theory)
Course Code: DSC51102 Course Title: - Statistical Techniques for Data Scientist

Teaching Scheme: 4 Hours / Week	**No. of Credits: 4**	**No. of Lectures: 60**	**Examination Scheme: CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Mathematics at HSC level

The main objectives of this course are to:

- Learn concepts of mathematical logic for analyzing propositions and proving theorems.
- Use sets for solving applied problems and use the properties of set operations algebraically.
- Work with relations and investigate their properties.
- Investigate functions as relations and their properties.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Describe basic features of the data.	B1,B2
CO2	Summarize data using different quantitative measures.	B3,B2
CO3	do logical thinking and the learning ability of students will be enhanced.	B2,B4
CO4	learn counting skills which are essential for the study of probability distributions and achieve the skills essential to become a good programmer and data scientist	B3,B5

B1 - Remember; **B2** - Understand; **B3** - Apply; **B4** - Analyze; **B5** - Evaluate; **B6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Descriptive Statistics** • Measures of Central Tendency: Mean,combined mean Median(ogive curve), Mode(histogram) • Partition Values: Quantiles (Box Plot) • Measures of Dispersion: Variance, Standard Deviation, Coefficient of variation, quartile deviation • Skewness: Concept of skewness, measures of skewness • Kurtosis: Concept of Kurtosis, Measures of Kurtosis • correlation and regression • Numerical problems	18	CO1,CO2
2	**Introduction to Probability** • Probability - classical definition, probability models, axioms of probability, probability of an event.	6	CO3

	• Concepts and definitions of conditional probability, multiplication theorem P(A∩B) =P(A). P(B|A) • Bayes' theorem (without proof) • Concept of Posterior probability, problems on posterior probability.		
3	Probability and Sensitivity • Definition of sensitivity of a procedure, specificity of a procedure. Application of Bayes' theorem to design a procedure for false positive and false negative. • Concept and definition of independence of two events. • Numerical problems related to real life situations.	6	CO3
4	**Introduction to Random Variables** • Definition of discrete random and continuous random variable. • Concept of Discrete and Continuous probability distributions. (p.m.f. and p.d.f.) • Distribution function • Expectation and variance • Numerical problems related to real life situations	10	CO3
5	**Special Distributions I** • Discrete Uniform Distribution • bernoulli Distribution • Binomial Distribution • Poisson Distribution • Negative Binomial Distribution	10	CO4
6	**Special Distributions II** • Geometric Distribution • Continuous Uniform Distribution • Exponential Distribution • Gamma Distribution • Normal Distribution • Log Normal Distribution • Pareto Distribution (For all the probability distributions its pmf/pdf, p-p plot, q-q plot, generation of probabilities and random samples using R software is expected.)	4	CO4

Reference Books

1.Fundamentals of Applied Statistics (3rd Edition), Gupta and Kapoor, S.Chand and Sons, New Delhi, 1987.
2. An Introductory Statistics, Kennedy and Gentle.
3. Statistical Methods, G.W. Snedecor, W.G. Cochran, John Wiley & sons, 1989.
4. Introduction to Linear Regression Analysis, Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, Wiley
5. Modern Elementary Statistics, Freund J.E., Pearson Publication, 2005.
6. Probability, Statistics, Design of Experiments and Queuing theory with applications Computer Science,
7. Trivedi K.S., Prentice Hall of India, New Delhi,2001.
8. A First course in Probability 6th Edition, Ross, Pearson Publication, 2006.
9. Introduction to Discrete Probability and Probability Distributions, Kulkarni M.B., Ghatpande S.B., SIPF Academy, 2007.
10. A Beginners Guide to R, Alain Zuur, Elena Leno, Erik Meesters, Springer, 2009
11. Statistics Using R, Sudha Purohit, S.D.Gore, Shailaja Deshmukh, Narosa, Publishing Company

Related Online Content (MOOCS, SWAYAM, NPTEL, WEBSITES etc.)	
1	https://www.mooc-list.com/tags/statistics
2	https://onlinecourses.swayam2.ac.in/cec24_mg20/preview

Semester- I **Course Type: Major Mandatory Paper 3(Theory)** **Course Code: DSC51103 Course Title: - Computational Mathematics**			
Teaching Scheme: **4 Hours / Week**	**No. of Credits: 4**	**No. of Lectures:** **60**	**Examination Scheme:** **CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Problem solving Skills

The main objectives of this course are to:

- introduce the foundations of mathematics need for data science

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	effectively use matrix algebra tools to analyse and solve systems of linear equations	B1,B2
CO2	use some numerical methods to solve linear systems of equations	B3,B4
CO3	work on vector maps.	B5, B6

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	Contents	No of Lectures	CO targeted
1.	**Systems of Linear Equations and Matrices** • Introduction to system of Linear equations • Gaussian Elimination Method. • Matrix and Matrix Operations • Inverses and Algebraic Properties of Matrices • Diagonal ,Triangular and symmetric Matrices • Linear Equations in Linear Algebra • Systems of linear equations • Row reduction and Echelon forms • Vector Equations • Matrix equations Ax=b-Solution set of linear systems • Applications of linear systems •	10	CO1
2	**General Vector Spaces** • Real Vector Spaces • Subspaces • Linear Independence • Coordinates and Basis • Dimension	20	CO1

	• Change of Basis • Row Space, Column Space and Null Space • Rank , Nullity and Fundamental Matrix Spaces • Linear Independence • Subspaces of Rn • Dimension and Rank		
3	**Eigenvalues and eigenvectors and Inner product spaces** • Eigenvalues and Eigenvectors • Eigen vectors and Upper Triangular matrices • Eigenspaces and Diagonal Matrices • Inner Products and Norms • Linear functionals on Inner Product spaces • length and orthogonality • Orthogonal sets • Orthogonal projections • The Gram-Schmidt Process • Least square problem • Singular value decomposition • Householder Transformation and QR decomposition • Non Negative Matrix Factorization • bidiagonalization	18	CO2
4	**Linear Transformations** • General Linear Transformations • Isomorphism • Compositions and Inverse Transformations • Matrices for General Linear Transformations • The inverse of a matrix • Characterizations of Invertible Matrices • Partitioned Matrices • Matrix factorizations • Similarity	12	CO2,CO3

Text Books
1. Elementary Linear Algebra with Supplemental Applications – Howard Anton and Chris Rorres Eleventh Edition.

Reference Books
1. Introduction to Linear Algebra , 6th Edition , Gilbert Strang. 2. Linear Algebra and its Applications , 5th Edition, David C. Lay.

Related Online Content (MOOCS, SWAYAM, NPTEL, WEBSITES etc.)
1. https://mandal.ku.edu/math290/m290NotesChFour.pdf 2. https://math.emory.edu/~lchen41/teaching/2020_Fall/Chapter_6.pdf 3. https://sam.nitk.ac.in/courses/MA904/vector%20spaces.pdf **Online Courses:** 1. https://www.classcentral.com/subject/vector-spaces

2. Best Vector Calculus Courses Online with Certificates [2024] | Coursera
3. https://www.coursera.org/courses?query=vector%20calculus

Semester- I Course Type: Major Mandatory Paper 4 (Practical) Course Code: DSC51104 Course Title: - Practical Based on ST & CM			
Teaching Scheme: 2 Hours / Week	**No. of Credits: 2**	**No. of Practicals: 10**	**Examination Scheme: CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Problem Solving Skills

The main objectives of this course are to:

- To understand the working of R software.
- to learn mathematics concepts of vectors spaces, linear maps and matrix methods

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Write basic commands in R	B6,B3,B1
CO2	Write programs using R	B2,B3,B4,B5
CO3	Describe and visualize basic features of the data.	B3
CO4	Study linear algebra and its mathematical problems	B5

B1 - Remember; B2 - Understand; B3 - Apply; B4 - Analyze; B5 - Evaluate; B6– Create

UNIT	Contents	No of Sessions	CO targeted
1	Basic commands in R	1	CO1,CO2
2	Diagrammatic and graphical data representation	1	CO1,CO3
3	Summary statistics	1	CO1,CO2
4	Correlation and regression	1	CO1,CO2
5	Probability graphs	1	CO1,CO2
6	Computation of probabilities for different probability distributions	1	CO1,CO2
7	Programming in R Statements: if and if… else, for loop, cat and print Commands.	1	CO1,CO2
8	Field Project	3	CO1,CO3
9	vector spaces-I	1	CO4
10	vector spaces-II	1	CO4
11	Eigenvalues and eigenvectors	1	CO4
12	linear Maps	1	CO4

13	Basic Matrix Methods for Applications-I	1	CO4
14	Basic Matrix Methods for Applications-II	1	CO4
15	Miscellaneous	1	CO4

Semester- I **Course Type: DSE Elective(Theory)** **Course Code: DSC51205 Course Title: - Artificial Intelligence**			
Teaching Scheme: **2 Hours / Week**	**No. of Credits: 2**	**No. of Lectures:** **30**	**Examination Scheme:** **CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Concepts of Data structures and Design and Analysis of algorithms.
- Strong data analytics skills.

The main objectives of this course are to:

- To learn various types of algorithms useful in Artificial Intelligence (AI).
- To convey the ideas in AI research and programming language related to emerging technology.
- To understand the numerous applications and huge possibilities in the field of AI that goes beyond the normal human imagination.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand AI and its applications	B2
CO2	Learn the applications of A* and AO* algorithms.	B2,B1,B3
CO3	Learn different knowledge representation techniques.	B6,B5,B6

B1 - Remember; **B2** - Understand; **B3** - Apply; **B4** - Analyze; **B5** - Evaluate; **B6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Introduction to Artificial Intelligence** ● Introduction and Intelligent systems ● What Is AI, The Foundations of ● Artificial Intelligence, ● The History of Artificial Intelligence ● Applications of AI ● Early work in AI and related fields ● AI problems and Techniques	05	CO1
2	**Searching** ● Defining AI problems as a State Space Search: example, Search and Control Strategies, Problem	10	CO2

	● Characteristics, Issues in Design of Search Programs, Production System. ● Blind Search Techniques :-BFS, DFS,DLS, Iterative Deepening, Search, Bidirectional Search, Uniform cost Search. ● Heuristic search techniques: Generate and test ,Hill Climbing, Best First search, Constraint Satisfaction, Mean-End Analysis, A*,AO*.		
3	**Knowledge Representation** ● Representations and Mappings ● Approaches to Knowledge representation, Knowledge ● Representation method ● Propositional Logic, Predicate logic ● Representing Simple facts in Logic ● Resolution, Forward and backward chaining ● Game Playing- Minimax Search Procedures ● Adding alpha-beta cutoffs.	10	CO3
4	**Slot and Filler Structures :** ● Weak Structures ● Semantic Networks ● Frame ● Strong Structures ● Conceptual Dependencies ● Script	5	CO3

Reference Books

1. Computational Intelligence, Eberhart, Elsevier Publication
2. Artificial Intelligence: A New Synthesis,Nilsson, Elsevier Publication
3. Artificial Intelligence, Meenal Jabde & Charushila Pati,l Nirali Publication

Semester- I Course Type: DSE Elective (Practical) Course Code: DSC51206 Course Title: - Practical Based on ML & AI			
Teaching Scheme: **2 Hours / Week**	**No. of Credits: 2**	**No. of Practicals:** **10**	**Examination Scheme:** **CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Problem Solving Skills

The main objectives of this course are to:

- understand various AI techniques

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	apply various supervised and unsupervised algorithms in machine learing	B6,B3,B1
CO2	apply and Design and implement search strategies and develop logical reasoning systems .	B2,B3,B4,B5

B1 - Remember; **B2** - Understand; **B3** - Apply; **B4** - Analyze; **B5** - Evaluate; **B6**– Create

UNIT	Contents	No of Sessions	CO targeted
1	Practical on Supervised Learning algorithms - Naive Bayes algorithm and SVM and Decision Tree	2	CO1
2	Practical on Un-Supervised Learning Algorithms- K Means and KNN	1	CO1
3	Practical on Linear Regression and Polynomial Regression	1	CO1
4	Practical on Apriori Algorithm	1	CO1
5	Practical on Ensemble Learning	1	CO1
6	Write a program to implement Breadth First Search Traversal.	1	CO2

Semester- I Course Type: DSE Elective (Practical) Course Code: DSC51206 Course Title: - Practical Based on ML & AI			
Teaching Scheme: **2 Hours / Week**	**No. of Credits: 2**	**No. of Practicals:** **10**	**Examination Scheme:** **CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Problem Solving Skills

The main objectives of this course are to:

- understand various AI techniques

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	apply various supervised and unsupervised algorithms in machine learing	B6,B3,B1
CO2	apply and Design and implement search strategies and develop logical reasoning systems .	B2,B3,B4,B5

B1 - Remember; **B2** - Understand; **B3** - Apply; **B4** - Analyze; **B5** - Evaluate; **B6**– Create

7	Write a program to implement Depth First Search Traversal.	1	CO2
8	Write a program to implement A*.	1	CO2
9	Write a program to implement basic search strategies for 4 queen Problem.	1	CO2
10	Write a program to implement uniform cost search.	1	CO2
11	Write a program to implement Water Jug Problem	1	CO2
12	Implement two player game using minimax search algorithm	1	CO2
13	Write a program to calculate n from initial state and goal state of 8-puzzle problem. 'n' is the number of tiles that are not at expected place according to the goal state.	2	CO2

Semester- I Course Type: RM (Theory) Course Code: DSCS51307 Course Title: - Research Methodology			
Teaching Scheme: **2 Hours / Week**	**No. of Credits: 2**	**No. of Lectures:** **30**	**Examination Scheme:** **CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Basic knowledge of computer science subject
- Reading and analytical skills

The main objectives of this course are to:

- Enable students to undertake independent research of a business problem, and to analyze and present their findings.
- Familiarize students with the basic techniques of collection, analysis, interpretation and presentation of data.
- Formulate a research proposal for a business project.
- Obtain Knowledge of research Methodology

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand research terminology & be aware of the ethical principles of research, ethical challenges and approval processes.	B1,B2
CO2	Identify the components of a literature review process	B3
CO3	Describe quantitative, qualitative and mixed methods approaches to research	B3,B4
CO4	Critically analyze published research	B5, B6

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	Contents	No of Lectures	CO targeted
1	**Introduction** • Introduction to Research: Definition, meaning, • characteristics of Research • Types of Research, • Introduction Nature of qualitative and quantitative research • Research in functional areas of management • Process of Research • Formulation of Research Design	2	CO1
2	**Literature Review** • Meaning and sources of Research problem, characteristics of good Research problem, Research process, outcomes, • Application of Research • Literature Review Process – Role, importance, sources, • methods, software tools for literature review. Formulation of Research Problem • Meaning, Need, Types of research • Research design – Exploratory, Descriptive, components of research design and features of good research design.	3	CO2
3	**Sampling & Data Collection** • Sampling: Sampling design: Meaning, logic and application of sampling. • Data Collection Methods: Types and Sources of Data • Data Analysis: Organization and presentation of data, Data Analysis Methods	2	CO2,CO3
4	**Presentation and Publication of Research** • Research Proposal, Research Paper, Research Thesis, Research Report, Report Writing, Format for writing reports • bibliography, and references	3	CO4

Reference Books
1. Research methodology techniques and methods by C R Kothari New age International publishers. 2. Probability and Statistics for Engineers and Scientists , Sheldon Ross Elsevier Academic Press 3. Research Methodology, R.Panneerselvam, PHI,New Delhi 2005 4. Researching Information Systems and Computing, Oates B J, Sage Publications Online E-resources 1. https://onlinecourses.nptel.ac.in/noc22_ge08/preview 2. https://ccsuniversity.ac.in/bridge-library/pdf/Research-Methodology-CR-Kothari.pdf

Semester- I Course Type: RM (Practical) Course Code: DSC51308 Course Title: - Practical based on Research Methodology			
Teaching Scheme: 2 Hours / Week	**No. of Credits: 2**	**No. of Practicals: 10**	**Examination Scheme: CIE: 20 Marks ESE: 30 Marks**
Prerequisites: Student should have basic knowledge of:			

- Problem solving Skills & Basic knowledge of Computer Science Subjects

The main objectives of this course are to:

- To enable students to publish research paper

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand research terminology	B1,B2
CO2	publish research paper/ research article	B3,B4

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	Contents	No of Sessions	CO targeted
1	• ICT Tools for Research • Use of Internet in Research: Browsing the internet through standard features, Accessing and Downloading information, • E-resources for research, Impact Factor: E - Information, H- Index, citation index, , e-journals etc	3	CO1,CO2
2	• Identification of Research Problem and Literature Survey • Various Referencing Styles	3	CO1,CO2
3	• Publishing a research paper	4	Co1, CO2

Second Semester

Semester- II
Course Type: Major Mandatory Paper 1(Theory)
Course Code: DSCS52101 Course Title: - Database Technologies

Teaching Scheme: 4 Hours / Week	**No. of Credits: 4**	**No. of Lectures: 60**	**Examination Scheme: CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Knowledge of file system concepts
- Strong foundation of Related database Concepts (Basic & Advanced)
- A firm foundation of any RDBMS package

The main objectives of this course are to:

- Provide an overview of the concept of NoSQL technology.
- Make the student capable of making choice of what database
- technologies to use on their application needs
- Provide an insight to the different types of NoSQL databases

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Differentiate between RDBMS and NOSQL technologies	B2
CO2	Understand various NOSQL technologies, their need, and applications	B1
CO3	Learn new concepts of data modelling, clustering, polyglot persistence, Version stamps, map reduce, schema migrations	B3,B4

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Overview of DBMS & NOSQL** • RDBMS Overview • Transaction, ACID Properties, • Database recovery techniques, DB Failure • Why NoSQL? • The value of relational databases • Impedance mismatch • Attack of clusters • The emergence of NOSQL	8	CO1
2	**Aggregate Data Models & Distribution Models** • Aggregates • Example of relations and aggregates • Single server • Sharding • Master slave replication • Peer to peer replication • Combining sharding and replication	8	CO2,CO3
3	**Consistency** • Update Consistency • Read Consistency • Relaxing Consistency • CAP theorem • Relaxing durability • Quorums	8	CO3
4	**Version stamps & Map Reduce** • Business & System transactions • Version stamps on multiple nodes	8	CO3

	• Basic Map Reduce • Example of Map Reduce		
5	**Schema Migrations & Polyglot Persistence** ● Schema Changes ● Schema changes in RDBMS ● Schema changes in NOSQL databases ○ Incremental migration ○ migrations in graph databases ○ changing aggregate structure ● Polyglot Persistence (Multi model types) ● Disparate data storage needs ● polyglot data store usage ● choosing right technology ● Enterprise concerns with polyglot persistence	8	CO3
6	**Implementation with NOSQL databases** ● Document Databases (Mongodb) ● Graph databases (Neo4j)	20	CO3

Reference Books

1. NoSQL Distilled,Pramod Sadalge, Martin Fowler,Pearson publication
2. NoSQL for Dummies, A Willy Brand,Pearson publication
3. Multidisciplinary Subjects for Research-XV(implementation of NOSQL document databases – MONGODB) Dipali Meher, Meenal Jabde, Redshine Publication

Online E-resources

- https://www.udemy.com/topic/nosql/
- https://www.udemy.com/topic/neo4j/

Semester- II
Course Type: Major Mandatory Paper 2 (Theory)
Course Code: DSCS52102 Course Title: - Inferential Statistics

Teaching Scheme: 4 Hours / Week	**No. of Credits: 2**	**No. of Lectures: 60**	**Examination Scheme: CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Distribution Theory
- Time series
- Sampling methods

The main objectives of this course are to:

- To predict the uncertainty of the sample or sample to sample variations

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Identify sampling methods from the pattern of the observed data	B6
CO2	Predict the future behavior of the time series data	B5,B4
CO3	Analyze sample data and identify the parameters and their probability distributions.	B4
CO4	Validate the hypothesis to ensure that the entire research process remains scientific and reliable	B1,B3
CO5	Hypothesize and test an assumption regarding population parameters using sample data	B3,B6
CO6	Predict different models of forecasting of time series data	B5

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Sampling** • Introduction to Sampling • Simple random Sampling • Stratified Random Sampling • Cluster Sampling • Concept of Sampling Error	05	CO1
2	**Sampling Distributions** • Introduction to Sampling distributions • Student's t distribution • Chi square distribution • Snedecor's F distribution • **I**nterrelations among t, chi-square and F distributions • Central Limit Theorem (Various Versions) and its applications.	15	CO5
3	**Testing of hypothesis** • Definitions: population, statistic, parameter, standard error of estimator.	15	CO4,CO5

	• Concept of null hypothesis and alternative hypothesis, critical region, level of significance, type I and type II error, one sided and two sided tests, p-value. • Large Sample Tests • Tests based on t, Chi-square and F-distribution		
4	**Analysis of Variance** • One Way ANOVA • Two Way ANOVA • Application of ANOVA to test the overall significance of Regression.	10	
5	**Time Series** • Meaning and Utility. • Components of Time Series. • Additive and Multiplicative models. • Methods of estimating trend: moving average method, least squares method and exponential smoothing method. (single, double and triple) • Elimination of trend using additive and multiplicative models. • Simple time series models: AR (1), AR (2). 5.7 Introduction to ARIMA Modelling	15	CO2,CO6

Reference Books
1. Fundamentals of Applied Statistics (3rd Edition), Gupta and Kapoor, S.Chand and Sons, New Delhi, 1987. 2. An Introductory Statistics, Kennedy and Gentle. 3. Statistical Methods, G.W. Snedecor, W.G. Cochran, John Wiley & sons, 1989. 4. Introduction to Linear Regression Analysis, Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, Wiley 5. Modern Elementary Statistics, Freund J.E., Pearson Publication, 2005. 6. Probability, Statistics, Design of Experiments and Queuing theory with applications Computer Science, Trivedi K.S., Prentice Hall of India, New Delhi,2001. 7. A First course in Probability 6th Edition, Ross, Pearson Publication, 2006. 8. Introduction to Discrete Probability and Probability Distributions, Kulkarni M.B., Ghatpande S.B., SIPF Academy, 2007. 9. A Beginners Guide to R, Alain Zuur, Elena Leno, Erik Meesters, Springer, 2009 10. Statistics Using R, Sudha Purohit, S.D.Gore, Shailaja Deshmukh, Narosa, Publishing Company

Related Online Content (MOOCS, SWAYAM, NPTEL, WEBSITES etc.)
https://www.mooc-list.com/tags/statistical-inference https://onlinecourses.nptel.ac.in/noc20_ma19/preview

.

Semester- II
Course Type: Major Mandatory Paper 3(Theory)
Course Code:DSC52103 Course Title: - Optimization Techniques

Teaching Scheme: 4 Hours / Week	**No. of Credits: 2**	**No. of Lectures: 60**	**Examination Scheme: CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Strong mathematical background (calculus, linear algebra), Analytical thinking, A basic programming skills (preferably), Familiarity with operations research concepts.

The main objectives of this course are to:

- To learn history, methods, features and applications of
- To learn the structure of LP model, graphical solution of LPP.
- To study advanced methods like simplex, dual simplex method and transportation problem.
- To learn Game theory and Project Scheduling by PERT-CPM. Course

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand and create from the ground up business intelligence reports of expert caliber.	B2
CO2	Create and employ the same instruments that experts in data science and analysis use.	B6
CO3	Learn about the subtleties and uses of descriptive, predictive, and prescriptive analytics to improve your analytical abilities and make data-driven business choices in real time.	B1
CO4	Analyze a dataset to learn new things and share what you find with impactful data visualizations.	B4
CO5	Apply data visualization techniques to real-world datasets for analysis and decision-making.	B3,B5

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	Introduction to OR	2	CO3, CO5
2	Linear Programming • Introduction • Formulation of LP Models • Examples of LPP Formulations • Graphical Solution of LPP • Formulation of LP Models ○ General Mathematical Model of LPP ○ Guidelines on LPP Formulations ○ Examples of LPP Formulations • Simplex Method ○ Standard form, Canonical form of LP Problem ○ Simplex Algorithm (Maximization case)	16	CO1, CO3

	○ Simplex Algorithm (Minimization case) ○ Revised Simplex Method ○ The Big-M Method (Use of Artificial variables) ○ Types of Linear Programming Solutions ○ Alternative (Multiple) Optimal Solutions ○ Unbounded Solutions (Multiple) Optimal Solutions, Unbounded Solutions		
3	Duality theory and applications • Formulation of Dual LPP • Standard Results on Duality • Advantage of Duality • Dual Simplex Method • Examples of Dual Simplex Method • Sensitivity analysis in LP	10	CO4
4	Transportation Problem • Mathematical Model of TP • Unbalanced Supply and Demand • Method for finding Initial basic feasible Solutions • Degeneracy and it's Resolution • North-west Corner Method • Least Cost Method • Vogel's Approximation Method • Test of Optimality • Modi Method • Assignment Problem ○ Mathematical Models of AP ○ Hungarian Method for Solving AP ○ Alternate Optimal Solutions ○ Maximization case in AP ○ Unbalanced AP	13	CO3
5	Game Theory • Introduction • Two-Person Zero-Sum Games with saddle point • Two-Person Zero-Sum Games without Saddle • Algebraic Method for Solving 2 * 2 • Graphical Method for Solving 2 * 2 • Games Without Saddle Point	8	CO1,CO2
6	Project Scheduling by PERT-CPM • Introduction • Basic Difference Between PERT & CPM • Significant of using PERT/CPM ○ Phases of Project Management ○ PERT/CPM Network Components & Precedence Relationships ○ Rules for AOA Network Construction ○ Errors & Dummies in Network • Critical Path Analysis ○ Forward Pass Method ○ Backward Pass Method ○ Float (Slack) of an Activity & Event ○ Critical Path ○ Case Study	12	CO4

Reference Books

1.Hamdy A. Taha, Operation Research (Eighth Edition, 2009), PrenticeHall of India Pvt. Ltd, New Delhi.
2. J K Sharma, Operations Research (Theory and Applications, second edition, 2006), Macmillan India Ltd.

Semester- II
Course Type: Major Mandatory (Practical)
Course Code: DSC52104 Course Title: - Practical on DBT & IS

Teaching Scheme: 2 Hours / Week	**No. of Credits: 2**	**No. of Practicals: 10**	**Examination Scheme: CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Knowledge of file system concepts
- Knowledge of C and CPP programming concepts
- Strong foundation of Relational database Concepts (Basic & Advanced)
- Knowledge of basic statistical techniques

The main objectives of this course are to:

- To Provide an overview of the concept of NoSQL technology.
- To Make the student capable of making choice of what database technologies to use on their application needs
- To Provide an insight to the MongoDB (Document database) and Neo4j(Graph Database)

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Provide an insight to the different types of NoSQL databases used to real life applications.	B1,B2
CO2	Create and handle databases and queries using various NQSQL technologies like MongoDB and Neo4j.	B3,B4
CO3	able to understand the hypothesis problem and can conclude the results	B5,B6

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Sessions**	**CO targeted**
1	• Database Practicals - Data Query Language(DQL) Statements • Data Query Language(DQL) Statements: (Select statement with • operations like Where clause, Order by, Logical operators, • Scalar functions and Aggregate functions) • Using Virtual Lab IIT Bombay • http://vlabs.iitb.ac.in/vlabs-dev/labs/dblab/labs/index.php Assignment 1: Movie Database 1. Create a database with the name 'Movie'. 2. A 'Film' is a collection of documents with the following fields: a. Film Id b. Title of the film c. Year of release	3	CO1,CO2

	d. Genre / Category (like adventure, action, sci-fi, romantic etc.) A film can belong to more than one genre. e. Actors (First name and Last name) A film can have more than one actor. A film can have more than one director. f. Release details (It consists of places of release, dates of release and rating of the film.) 3. An 'Actor' is a collection of documents with the following fields: a. Actor Id First name c. Last Name d. Address (Street, City, State, Country, Pin-code) e. Contact Details (Email Id and Phone No) f. Age of an actor. Queries: 1. Insert at least 10 documents in the collection Film _ a. Insert at least one document with film belonging to two genres. b. Insert at least one document with film that is released at more than one place and on two different dates. c. Insert at least three documents with the films released in the same year. d. Insert at least two documents with the films directed by one director. e. Insert at least two documents with films those are acted by a pair 'Madhuri Dixit' and 'Shahrukh Khan'. 2. Insert at least 10 documents in the collection Actor. Make sure, you are inserting the names of actors who have acted in films, given in the 'Film' collection. 3. Display all the documents inserted in both the collectAdd a value to the rating of the film whose title starts with 'T'. 4. Add an actor named "_" in the 'Actor' collection. Also add the details of the film in 'Film' collection in which this actor has acted in. 5. Delete the film " ". 6. Delete all actors from an 'Actor' collection who have age greater than >60 7. Update the actor's address where Actor Id is"___" Assignment 2: Model the following Book system as document database Consider Set of books and publishers. Publisher can publish more than one book Book(Book name, Cost, Author, Published Year, Number of Pages) Publisher(name, language, books, city) Queries: a. List all the publishers located in mumbai b. List all the book having pages> 500 c. List all the books having cost 500 d. List all the books published in year 2020 e. List all the books written by "__" and published in 2020		

f. List the books published in english language
g. List the book published in marathi language
Assignment 3:
Model the following hospital database as document database Consider the hospitals in and around pune.Each hospital may have one or more specializations like pediatric, gynac, ortho.A person can recommend or provide review for a hospital. One doctor can be associated with more than one hospital.
Queries:
a. List the names of the hospitals with __ specialization
b. List the names of doctors visiting to birla hospital on monday
c. List the multispeciality hospitals
d. List the names of hospitals having rating >=4
e. List the doctors who are specialized in ortho
f. List the persons who have given ratings to sahyadri hospital

Neo 4J Practical

Assignment 3: Song Database
Consider a Song database, with labels as Artists, Song, Recording_company, Recoding_studio, song author etc.
Relationships can be as follows
Artist → [Performs] → Song →[Written by] → Song_author.
Song → [Recorded in] → Recording Studio →[managed by]
→ recordingCompany Recording Company → [Finances] →
Song
You may add more labels and relationship and their properties,
as per assumptions.
a) List the names of songs written by ":....."
b) List the names of the songs recorded in"..."
c) List the names of record companies who have financed for the song "...."
d) List the names of artist performing the song "....."
e) Name the songs recorded by the studio " ....."
f) List the names of artists who have sung only songs written by " "
g) List the names of artists who have sung the maximum number of songs recorded by "....." studio

Assignment 4: Employee database

Consider an employee database, with a minimal set of labels as follows Employee: denotes a person as an employee of the organization Department: denotes the different departments, in which employees work. Skillset: A list of skills acquired by an employee
Projects: A list of projects in which an employee works. A minimal set of relationships can be as follows:
Works_in :employee works in a department
Has_acquired: employee has acquired a skill
Assigned_to : employee assigned to a project

	Controlled_by: A project is controlled by a department Project_manager : Employee is a project_manager of a Project a) List the names of employees in department "..." b) List the projects along with their properties, controlled by department ".....". c) List the departments along with the count of employees in it. d) List the skillset for an employee " " e) List the projects controlled by a department "..." f) List the names of the projects belonging to departments managed by employee " ..."		
IS Practical			
1	Sampling methods: SRSWOR, SRSWR	1	CO3
2	Tests based on Chi square distribution	1	
3	Tests based on t distribution	1	
4	Tests based on F distribution	1	
5	One way and Two way ANOVA	1	
6	Methods of estimating trend	1	
7	Exponential Smoothing	1	

Semester- II
Course Type: DSE Elective (Theory)
Course Code: DSC52205 Course Title: - Deep Learning and Neural Networks

Teaching Scheme: 2 Hours / Week	**No. of Credits: 2**	**No. of Lectures: 30**	**Examination Scheme: CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Data Mining and Data warehousing
- Computer networks
- Data science

The main objectives of this course are to:

- To understand the basics of neural networks and deep learning
- To learn various types of neural networks
- To understand the concept of reinforcement, recurrent learning methods

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand the concepts of neural networks, and deep learning	B1,B2,B4
CO2	To apply techniques of convolution and recurrent neural networks	B3,B5,B6

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Foundations of Deep Learning** • What is Machine Learning and deep learning? • Supervised and Unsupervised learning, • Bias variance tradeoff • Applications of machine learning • Hyperparameters underfitting /overfitting regularization • Limitations of machine learning • History of deep learning • Advantages and challenges of deep learning • Learning representations from data • Understanding how deep learning works in three figures • Common architectural principles of deep network • Architecture design • Applications of deep learning • Introduction and use of popular industry tools such as Tensor Flow, Keras, PyTorch, Caffe, Shogun	12	CO1
2	**Neural Networks** • Introduction to Neural networks ○ The biological Neuron ○ Artificial Neural Networks	12	CO1

	○ The perceptron • Multilayer feedforward networks • Training neural networks • Back propagation and forward propagation • Activation functions ○ linear, Sigmoid, Tannh,Hard Tanh, Softmax, Rectified linear ○ loss functions- Introduction • hyperparameters ○ learning rate ○ regularization ○ Momentum ○ Sparsity in hyperparameters ○ Deep feedforward networks, example of EXOR hidden units, • Cost function • Error back propagation • Gradient based learning • Sentiment analysis • Deep learning with PyTorch, Jypter, Colab		
	Convolution Neural Network • Introduction • CNN,RNN,GAN and its types • Applications of GAN • Reinforcement Learning	6	CO2

Reference Books
1. https://www.deeplearningbook.org/contents/intro.html 2.http://ndl.ethernet.edu.et/bitstream/123456789/88552/1/2018_Book_NeuralNetworksAndDeepLearning.pdf

Semester- II Course Type: DSE Elective (Theory) Course Code: DSC52206 Course Title: - Project on Data Science			
Teaching Scheme: **2 Hours / Week**	**No. of Credits: 2**	**No. of Sessions:** **30**	**Examination Scheme:** **CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- **Data science Subjects**

The main objectives of this course are to:

- able to apply various data science techniques to data and find inference

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	apply and use various data science techniques and software to data and find inference	B3, B6

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	Contents	No of Lectures	CO targeted
	Students will learn various data science software and use them and apply them to data and prepare project. and present that project at the time of examination.		CO1

Third Semester

Semester- III
Course Type: Major Mandatory Paper 1 (Theory)
Course Code: DSC63101 Course Title: - Predictive Analysis

Teaching Scheme: 4 Hours / Week	**No. of Credits: 4**	**No. of Lectures: 60**	**Examination Scheme: CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Statistics and Probability
- Basic Data Science concepts
- Data preprocessing and visualization techniques
- Fundamentals of Machine Learning
- Programming skills in Python or R

The main objectives of this course are to:

- Introduce the fundamental concepts and processes involved in predictive analysis.
- Explain different predictive modeling techniques and their applications.
- Develop the ability to prepare and analyze data for prediction.
- Enable students to select appropriate models and evaluate their performance.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand the fundamentals and applications of predictive analysis.	B1,B2,B4
CO2	Perform data cleaning, preprocessing, and feature engineering for predictive modeling.	B3,B5,B6
CO3	Apply various predictive modeling techniques such as regression, decision trees, and neural networks.	B3, B4, B5
CO4	Analyze and forecast data using time series models and Evaluate, validate, and deploy predictive models effectively.	B3,

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Fundamentals of Predictive Analysis** • Predictive Analysis Introduction • Data analysis • Concept of Prediction • Why Predictive Analysis? • Steps/Lifecycle of Predictive Analysis • Applications of Predictive Analysis • Predictive modeling and its features • Types of Predictive models • Tools for Predictive Analysis • Descriptive Vs Prescriptive Vs Predictive Analysis • AI Vs Predictive Analysis	**10**	CO1

2	**Data Preparation** • Types of data: structured, unstructured, and semi-structured • Cleaning and pre-processing data (handling missing values, outliers, and noise) • Feature engineering: selection, extraction, and transformation • Data scaling and normalization	**6**	CO2
3	**Predictive Modeling Techniques** • Factors for choosing Predictive Modeling Techniques • Decision tree • Concept and working of Decision Trees • CHAID and CART algorithms • Random Forests • Regression Techniques • Simple and Multiple Linear Regression • Assumptions and Residual Analysis • Logistic Regression • Neural Networks (Basics) • Introduction and basic structure of ANN • Difference between Artificial and Biological neurons • Overview of CNN, RNN, and LSTM	**14**	CO3
4	**Introduction to Time Series Analysis and Forecasting** • Introduction to Time Series • Meaning and Importance of Time Series • Trend, Seasonal, Cyclical, and Irregular Components • Time Series Patterns and Randomness • Additive vs. Multiplicative Models • Exponential Smoothing (Simple & Weighted) • Simple Time Series Models: AR(1), AR(2) • Introduction to AR, MA, and ARIMA Models • Choosing the Best Forecast Model • Regression Analysis for Forecasting	**12**	CO4
5	**Predictive Model Evaluation and Deployment** • Introduction • Model Validation • Rule Induction Using CHAID • Automating Models for Categorical and Continuous targets • Comparing and Combining Models • Evaluation Charts for Model Comparison • MetaLevel Modeling • Deploying Model • Assessing Model Performance • Updating a Model	**12**	CO4
6	**Predictive Analytics using Machine Learning Models** • Supervised learning vs. unsupervised learning • Machine Learning algorithms • Regression Models • Classification Models • Model Diagnostics	**4**	CO4

Reference Books
1.https://www.predictiveanalyticsworld.com/book/pdf/Predictive_Analytics_by_Eric_Siegel_Excerpts.pdf

Online Courses
https://onlinecourses.swayam2.ac.in/imb20_mg19/preview https://onlinecourses.nptel.ac.in/noc23_ma46/preview

Semester- III Course Type: Major Mandatory Paper 2(Theory) Course Code: DSC63102 Course Title: - Natural Language Processing			
Teaching Scheme: 4 Hours / Week	**No. of Credits: 4**	**No. of Lectures: 60**	**Examination Scheme: CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Python Programming and Data Handling
- Linear Algebra, Probability, and Statistics
- Machine Learning Algorithms
- Basic Concepts of Artificial Intelligence

The main objectives of this course are to:

- Introduce the fundamental concepts, scope, and challenges in Natural Language Processing (NLP).
- Develop understanding of text processing and linguistic representation techniques.
- Study syntactic, semantic, and pragmatic levels of NLP analysis.
- Apply machine learning and deep learning techniques to NLP tasks.
- Enable students to design and evaluate end-to-end NLP applications.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Recall and understand fundamental NLP concepts, components, and applications.	B1,B2
CO2	Apply text preprocessing and linguistic feature extraction techniques to textual data.	B3
CO3	Analyze syntactic, semantic, and statistical NLP models.	B4
CO4	Evaluate NLP algorithms and select optimal models for various tasks.	B5

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	Contents	No of Lectures	CO targeted
1	**Introduction to Natural Language Processing** • Introduction of NLP. Importance and scope in Data Science • NLP pipeline and architecture • Challenges in NLP: ambiguity, variability, resource limitations • Levels of language analysis: phonology, morphology, syntax, semantics, pragmatics • Linguistic resources: corpora, lexicons, and ontologies (WordNet, FrameNet)	8	CO1

	• Applications: Information retrieval, machine translation, chatbots, question answering		
2	**Text Processing and Representation** • Text preprocessing: tokenization, sentence segmentation, stemming, lemmatization • Stop-word removal, part-of-speech (POS) tagging • Regular expressions for text processing • Statistical text representation: Bag of Words (BoW), N-grams, TF-IDF • Word embeddings: Word2Vec (CBOW, Skip-gram), GloVe, FastText • Contextual embeddings: ELMo, BERT, GPT overview	10	CO2
3	**Syntax and Parsing** • Syntax and grammar: formal grammars, phrase structure rules • Constituency and dependency grammars • Parsing techniques: top-down and bottom-up parsing • Probabilistic parsing and context-free grammars (PCFGs) • POS tagging models: Hidden Markov Models (HMM), Conditional Random Fields (CRF) • Evaluation metrics for tagging and parsing (precision, recall, F1-score)	10	CO3
4	**Semantics and Pragmatics** • Lexical semantics: synonymy, antonymy, hyponymy, polysemy • Word sense disambiguation (Lesk algorithm, supervised WSD) • Semantic similarity measures and vector semantics • Named Entity Recognition (NER) techniques • Sentiment analysis (lexicon-based and machine learning-based) • Pragmatic analysis: discourse structure, anaphora • resolution, co-reference resolution	10	CO3,CO4
5	**Machine Learning for NLP** • Supervised and unsupervised learning in NLP • Text classification: Naïve Bayes, Logistic Regression, SVM • Document clustering using K-Means, hierarchical clustering • Feature engineering and dimensionality reduction for text • Sequence labeling tasks (NER, POS tagging) using CRFs and LSTMs • Model evaluation and error analysis	10	CO4
6	**Deep Learning for NLP** • Neural language models: feedforward, RNN, LSTM, GRU • Attention mechanism and sequence-to-sequence models • Transformer architecture and encoder-decoder framework • Pretrained models: BERT, GPT, RoBERTa, T5 • Transfer learning and fine-tuning for downstream tasks • Case studies: ○ Machine Translation (seq2seq, Transformer)	12	CO5

	○ Text Summarization (Extractive vs. Abstractive) ○ Chatbot development using Transformer models		

Reference Books
1.Eisenstein, J. (2019). *Introduction to Natural Language Processing*. MIT Press. 2.Palash Goyal, Sumit Pandey, Karan Jain (2018). *Deep Learning for Natural Language Processing*. Apress. 3.Manning, C., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press.

Related Online Content (MOOCS, SWAYAM, NPTEL, WEBSITES etc.)
https://onlinecourses.nptel.ac.in/noc23_cs45/preview?utm_source=chatgpt.com https://nptel.ac.in/courses/106101007?utm_source=chatgpt.com

Semester- III
Course Type: Major Mandatory Paper 3(Theory)
Course Code: DSC63103 Course Title: - Cloud Computing

Teaching Scheme: 2 Hours / Week	No. of Credits:2	No. of Lectures: 30	Examination Scheme: CIE: 20 Marks ESE: 30 Marks

Prerequisites: Student should have basic knowledge of:

- Operating System concepts
- Fundamentals of Computer Networks
- Good Understanding of Object-Oriented Programming Concepts

The main objectives of this course are to:

- To understand the principles and paradigm of Cloud Computing
- To appreciate the role of Virtualization Technologies
- Ability to design and deploy Cloud Infrastructure
- Understand cloud security issues and solutions

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Appreciate the need for cloud computing and make decisions on using specific cloud service type and deployment models	B1,B2,B4
CO2	Identify virtualization technologies of a cloud platform.	B3,B5,B6
CO3	Make choices on selection of appropriate cloud service based on application requirements.	

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	Contents	No of Lectures	CO targeted
1	**Introduction to Cloud Computing** • Overview, Layers and Types of Cloud • Desired Features of a Cloud • Benefits and Disadvantages of Cloud Computing	8	CO1

	• Cloud Infrastructure Management, ° Infrastructure as a Service Providers ° Platform as a Service Providers Multitenant Technology ° Cloud-Enabling Technology: Broadband Networks and Internet Architecture, Data Center Technology, Virtualization Technology. ° Infrastructure as a Service, Platform as a Service, Software as a Service, Cloud Deployment Models. ° Cloud economics and benefits,		
2	**Abstraction and Virtualization** • Introduction to Virtualization Technologies • Load Balancing and Virtualization • Understanding Hyper visors, • Virtual Machines -Provisioning and Manageability Virtual Machine • Provisioning in the Cloud Context • Virtualization of CPU, Memory, I/O Devices • Virtual Clusters and Resource Management	7	CO2
3	**Programming, Environments and Applications** • Features of Cloud and Grid Platforms • Programming Support of Google App Engine • Programming on Amazon AWS and Microsoft Azure, Emerging Cloud Software Environments • Applications: Moving application to cloud • Microsoft Cloud Services • Google Cloud Applications • Amazon Cloud Services	8	CO3
4	**Security In The Cloud Security Overview** • Cloud Security Challenges and Risks • Software-as-a-Service Security • Security Governance Risk Management – Security Monitoring • Security Architecture Design • Data Security • Application Security • Virtual Machine Security	7	CO4

Reference Books

1. Cloud Computing: Technologies and Strategies of the Ubiquitous Data Center byBrian J.S. Chee and Curtis Franklin
CRC Press, ISBN :9781439806128
2. Mastering Cloud Computing: Foundations and Applications Programming by Rajkumar Buyya, Christian Vecchiola, S. ThamaraiSelvi McGraw Hill, ISBN: 978 1259029950, 1259029956
3. Distributed and Cloud Computing, From Parallel Processing to the Internet of Things by Kai Hwang, Geoffrey C Fox, Jack G Dongarra, Morgan Kaufmann Publishers, 2012.
4. Cloud Computing by Dr. Satish Ambike, Dr. Rajesh K Dhumal Nirali Publication ISBN: 9789354512520

Semester- III
Course Type: Major Mandatory Paper 4(Theory)
Course Code: DSC63104 Course Title: - Data Security & Ethics

Teaching Scheme: 2 Hours / Week	**No. of Credits:2**	**No. of Lectures: 30**	**Examination Scheme: CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Concept of data and basics of data science

The main objectives of this course are to:

- To understand the basics of Information Security.
- To know the legal, ethical and professional issues in Information Security.
- To equip the students' knowledge on digital signature, email security and web security.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand the basics of data and information security	B3,
CO2	Remember the legal, ethical and professional issues in information security	B1,
CO3	Understand the various authentication schemes to simulate different applications.	B2,B6
CO4	Learn various security practices, web security protocols for E-commerce applications and system security standards	B4,B5

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Introduction** • History • What is Information Security? • Critical Characteristics of Information, NSTISSC Security Model Components of an Information System • Securing the Components • Balancing Security and Access, The SDLC, The Security SDLC	5	CO1

2	**Security Investigation** • Need for Security, • Business Needs, Threats, Attacks, • Legal, Ethical and Professional Issues - An Overview of Computer Security • Access Control Matrix, • Policy-Security policies, Confidentiality policies, • Integrity policies and Hybrid policies	10	CO2
3	**Digital Signature and Authentication** • Digital Signature Schemes and their Variants • Digital Signature Standards • Authentication: Overview • Requirements Protocols • Applications - Kerberos -X.509 Directory Services	5	CO3
4	**E-mail, IP & WEB Security** • E-mail and IP Security • Electronic mail security • Email Architecture ○ PGP – Operational Descriptions ○ Key management- Trust Model- S/MIME. ○ IP Security: Overview- Architecture - ESP, AH Protocols IPSec Modes ○ Security association - Key management. • Web Security ○ Requirements- Secure Sockets Layer ○ Objectives-Layers -SSL secure communication- ○ Protocols - Transport Level Security. ○ Secure Electronic Transaction- Entities DS Verification-SET processing.	10	CO4

Reference Books

1. Harold F. Tipton, Micki Krause Nozaki,, "Information Security Management Handbook, Volume 6, 6th Edition, 2016.
2. Stuart McClure, Joel Scrambray, George Kurtz, "Hacking Exposed", McGraw- Hill, Seventh Edition, 2012.
3. Matt Bishop, "Computer Security Art and Science, Addison Wesley Reprint Edition, 2015.
4. Behrouz A Forouzan, Debdeep Mukhopadhyay, Cryptography And network security, 3rd Edition, . McGraw-Hill Education, 2015.
5. Michael E Whitman and Herbert J Mattord, "Principles of Information Security, Course Technology, 6th Edition, 2017.
6. Stallings William. Cryptography and Network Security: Principles and Practice, Seventh Edition, Pearson Education, 2017.

Semester- III
Course Type: Major Mandatory Paper 5(Practical)
Course Code: DSC63105 Course Title: - Practical on Predictive Analysis

Teaching Scheme: 2 Hours / Week	**No. of Credits:2**	**No. of Practicals: 10**	**Examination Scheme: CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Fundamental concepts of statistics and mathematics
- Basics of programming
- Data handling and visualization
- Fundamental concepts of machine learning

The main objectives of this course are to:

- Introduce fundamental concepts, scope, and applications of Predictive Analytics.
- Develop students' ability to prepare, preprocess, and analyze data for predictive modeling.
- Enable students to understand and apply various predictive modeling techniques such as regression, decision trees, ensemble methods, and neural networks.
- Introduce time series forecasting methods and model selection strategies.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Understand basic concepts of predictive analytics and data preparation.	B1,B2
CO2	Apply predictive modeling techniques to real-world datasets.	B2,B3
CO3	Analyze and evaluate neural network models and their components.	B2,B3
CO4	Apply time series analysis and forecasting methods and model evaluation	B3,B4

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Practicals**	**CO targeted**
1	Data Handling and Imputation	1	CO1
2	Data Pre-processing and Feature Engineering	1	CO1
3	Decision Tree for Classification	1	CO2
4	CHAID and CART Algorithm Implementation	1	CO2
5	Random Forest for Classification/Regression	1	CO2
6	Linear and Logistic Regression Modeling	1	CO2
7	Neural Network (ANN) Implementation	1	CO3
8	Time Series Decomposition & Basic Forecasting	1	CO4

9	AR, MA & ARIMA Model Development	1	CO4
10	Model Evaluation, Comparison & Deployment Basics	1	CO4

Semester- III **Course Type: DSE Electives Paper 6(Theory)** **Course Code: DSC63106 Course Title: - Business Information Studies**			
Teaching Scheme: **2 Hours / Week**	**No. of Credits:2**	**No. of Lectures:** **30**	**Examination Scheme:** **CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Basic Knowledge of Data in the Business

The main objectives of this course are to:

- To practice the application of the concepts related to Business information systems .
- To make students familiar with data management and governance

Expected Course Outcomes:

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

CO1	Understand role of IT in organizations	B2,
CO2	do data management and governance	B3,B5,B6
CO3	Learn Emerging Trends in Business Information	B1,

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Information Systems in Business and Business Analytics** • Information Systems • Role of IT in organizations • Enterprise Resource Planning (ERP) • Customer Relationship Management (CRM) • Information System Categories related to Specific Functional Areas of Business • IS Categories applicable Functional Areas; Office Automation Systems; Communication Systems: Teleconferencing, E Mail, Fax,, SMS, Groupware, Internet, Intranets, Extranets, Knowledge Management, and Group Support Systems; Transaction Processing Systems; MIS and Executive Information Systems; • Decision Support Systems	10	CO1
2	**Data Management** • Database design and management • Meaning of data management	10	CO2

	• Need of data management • Data management process • Big data • Data management system components		
3	**Data Governance** • Data governance meaning • Importance , objectives of data governance • Introduction to Data Governance Tools • Concept of data asset, types of data assets, concept of data steward,	6	CO2
4	**Emerging Trends in Business Information** • Generative AI in business • IoT, blockchain & digital ecosystems • Sustainability & green information systems	4	CO3

References
1. Haag & Cummings: Information Systems Essentials, 3ed; Tata McGraw Hill 2. Nickeson : Business Information Systems, 2nd ed, Prentice Hall of India 3. DAMA-DMBOK: Data Management Body of Knowledge DMBOK (2nd Edition) , Technics Publications LLC 4. Master Data Management and Data Governance by Alex Berson, Larry Dubov , McGraw Hill Publications

Online Course Link
Business Intelligence with PowerBI: From Data to Strategic Decisions - Course
POWER-BI-NOTES/POWER BI INTRO.pdf at main · SagarSahani07/POWER-BI-NOTES · GitHub INTRODUCTION TO MICROSOFT POWER BI: BRING YOUR DATA TO LIFE! Learning Microsoft Power BI https://froy otechnologies.in/Learning/PowerBi/Chapter-1.pdf

Semester- III
Course Type: DSE Electives Paper 7 (Practical)
Course Code: DSC63107 Course Title: - Data Science Tool

Teaching Scheme: 2 Hours / Week	**No. of Credits:2**	**No. of Practicals: 30**	**Examination Scheme: CIE: 20 Marks ESE: 30 Marks**

Prerequisites: Student should have basic knowledge of:

- Basic knowledge of Microsoft Excel (formulas, tables, charts)
- Basic understanding of data types, data formats, and data structures.
- Familiarity with CSV, Excel, and text datasets.

The main objectives of this course are to:

- Understand the concepts of business intelligence and the role of Power BI in analytics.
- Learn how to connect, transform, and clean datasets using Power BI Desktop.
- Create interactive dashboards, reports, and visualizations.

Expected Course Outcomes:

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

CO1	Understand architecture of PowerBI system	B1
CO2	Demonstrate understanding of data modeling principles, relationships, and schema design.	B3
CO3	Import, transform, and load data using Power Query.	B4,B5
CO4	Build dashboards, charts, and advanced visualizations using Power BI Desktop and write and apply DAX calculations	B6

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

UNIT	**Contents**	**No of Lectures**	**CO targeted**
1	**Introduction to Power BI & Data Sources** • Introduction to Power BI ecosystem • Power BI Workflow: Connect– Transform– Model– Visualize– Publish • Installing and configuring Power BI Desktop • Types of data connections- Import mode and DirectQuery • Connecting to- Excel, CSV, Text files, SQL Server, Web APIs, SharePoint	5	CO1
2	**Data Transformation Power Query**	10	CO3,CO4

	• Power Query Editor Interface • Data cleaning- Remove nulls, errors, duplicates, replace values, handle data types • Data shaping- Split column, Merge column, Group by, Transpose and pivot/unpivot • Combining datasets- merge Queries, Append Queries • Custom Columns • Conditional Columns • Introduction to M language		
3	**Data Modeling and DAX(Basic to Advanced)** • Understanding Relationships • Active Vs Inactive relationships • Star Schema Vs Snowflake Schema • Model optimization- Reducing cardinality, Removing unnecessary columns, Creating hierarchies • Calculated Columns Vs Measures • Basic DAX functions- SUM, COUNT, DISTINCTCOUNT, AVERAGE • Logical functions- IF, SWITCH • Aggregation functions- SUMX, AVERAGEX • Calculate() and filter context basics • Row context Vs filter context • Time Intelligence- YTD, QTD, MTD, SAMEPERIODLASTYEAR, DATEADD • Advanced filtering with ALL(), FILTER() • KPI creation	7	CO2
4	**Data Visualization, AI Features & Deployment** • Basic Visuals- Bar/Line/Pie/Table/Matrix • Advanced Visuals- Decomposition Tree, Key Influencers, Smart Narratives, KPI Visual • Drill-down, Drill-through, Bookmarks, Tooltips • Custom visuals form Marketplace • Designing professional dashboards • Mobile layout dashboard building	8	CO4

Semester- III Course Type: DSE Electives Paper 8(Project) Course Code: DSC63108 Course Title: - Research Project			
Teaching Scheme: **4 Hours / Week**	**No. of Credits:4**	**No. of Sessions:** **30**	**Examination Scheme:** **CIE: 40 Marks ESE: 60 Marks**

Prerequisites: Student should have basic knowledge of:

- Data handling and data management
- Excel, Python libraries

The main objectives of this course are to:

- Understand the research and research project.
- Understand how to publish paper in conference.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Carry out a substantial research-based project	B1,B2,B4
CO2	Analyze data and synthesize research findings	
CO3	Report research findings in written and verbal forms	
CO4	Publish research work in reputable journals, present at conferences or in recognized project competitions.	B3,B5,B6

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

Guidelines for Research Work

Guidelines for Research Project Work
1 Each student or group of students must submit a detailed project proposal outlining the research problem, objectives, methodology, and expected outcomes.
2 A mentor will be assigned by college to each group of students to provide guidance and support throughout the research process as well as to do internal assessment.
3 Students are required to conduct a thorough literature review to understand the current state of research in their chosen area.
4 Students should execute the research plan outlined in their proposal, adhering to ethical guidelines and academic standards.

5 Proper documentation of the research process, including experimental setup, data collection methods, and analysis techniques, should be maintained
6 Upon completion of the research work, students must prepare a project report and are encouraged to publish their research work in reputed journals, present at conferences or in recognized project competitions to disseminate their findings.
8 Evaluation will be as per the University guidelines, based on the quality of the research work, adherence to the research plan, presentation skills, and contribution

Semester- IV
Course Type: Major Mandatory Paper 1(Project)
Course Code: DSC64101 Course Title: - Full Time Industrial Training

No. of Credits:12	**Examination Scheme: CIE: 120 Marks**	**ESE: 180 Marks**

Prerequisites: Student should have basic knowledge of:

- Logic Development and Various Programming Languages and Soft Skills

The main objectives of this course are to:

• To provide students with an opportunity to apply theoretical knowledge gained throughout the program in a real-world industrial setting
• To foster professional skills such as teamwork, communication, time management, and problem-solving in an industrial environment.
• To expose students to the practices, technologies, and challenges prevalent in the IT industry or related sectors.
• To enable students to gain hands-on experience by working on projects or tasks relevant to their field of study.
• To facilitate networking opportunities with professionals in the industry, potentially leading to future career prospects.

Expected Course Outcomes:

On the successful completion of the course, student will be able to:

CO1	Apply theoretical concepts learned in the classroom to solve practical problems encountered in an industrial setting.	B3
CO2	Demonstrate proficiency in using industry-standard tools, technologies, and methodologies relevant to their area of specialization.	B2
CO3	Apply analytical and problem-solving skills to address challenges encountered during the industrial training	B3, B5
CO4	Collaborate effectively with team members to achieve project goals and objectives.	B4
CO5	Manage time and resources efficiently to complete assigned tasks and projects within the stipulated timeframe.	B6
CO6	Prepare a comprehensive report documenting their experience, including project details, learnings, and reflections.	B1,B2

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create

Guidelines for Full Time Industrial Training (IT)

1	Students are required to secure an industrial/internship placement in any organization, institution, or IT industry relevant to their field of study.
2	Students must submit the offer letter from the organization within two weeks of starting the industrial training/internship, detailing the terms and duration of the internship.
3	Students must have to work full time in the organization as per their rules and regulations.
4	A mentor will be assigned to each group of students to provide guidance and support throughout the internship period.
5	The industrial training/ internship duration should span a minimum of 360 hours, equivalent to 12 credits.
6	Students may be assigned specific projects or tasks or assignments by the host organization, relevant to their area of specialization.
7	Students should provide regular updates to their mentor through progress reports time to time regarding their progress, challenges faced, and lessons learned during the industrial training.
8	Upon completion of the industrial training/ internship, students must submit a comprehensive report documenting their internship experience, including project/ assignment details, challenges and achievements as per the format specified.
9	Evaluation will be based on the quality content of the internship report, feedback from the host organization, and the overall performance during the internship/ industrial training period.

Evaluation Pattern

• Internal assessment will be carried by college guide/ mentor by continuous evaluation method.

• The final examination or presentation of the work carried during the training/internship period will be in front the panel of examiners as per the schedule given by College.

• There will be a panel of three examiners for the final assessment

1. Industry expert (Appointed by the college)
2. Academic expert (Appointed by the College)
3. College guide/Mentor of the student as an internal examiner

Parameters for Evaluation	Marks
Internal Assessment by Mentor Regular updates(30 Marks) Timely report submission and deliverable (30 Marks) Professional Conduct, Learning and Skill Development (30 Marks) Work Undertaken and Learning Outcomes (30 Marks)	120
External Assessment by Industry expert and Academic expert.	180
Relevance and significance of the project or tasks undertaken (40 Marks) Technical proficiency demonstrated during the internship (40 Marks) Communication skills and presentation of the internship experience (20 Marks) Work Undertaken and Quality of the internship report (40 Marks) Overall performance and contribution to the organization (40 Marks)	
Total Assessment	300

Semester- IV Course Type: Major Mandatory Paper 2(Practical) Course Code: DSC64202 Course Title: - MOOC/Online Courses		
No. of Credits:04	**Examination Scheme: CIE: 40 Marks**	**ESE: 60 Marks**
Prerequisites: Student should have basic knowledge of: • Logic Development and Various Programming Languages and Soft Skills		
The main objectives of this course are to: • acquiring knowledge about specific subject		
Expected Course Outcomes:		
On the successful completion of the course, student will be able to:		
CO1	Acquire knowledge of any specific subject	B4
CO2	Develop critical thinking capacity	B2,B3,B6
CO3	learn to self-regulate their learning process, including setting goals, managing time, and adjusting study strategies as needed.	B1, B5
B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create		
Guidelines for MOOC/ Online Courses		
Students should complete any MOOC course related to Data Science from Swayam/ NPTEL/ MahaSwayam/Coursera website and undergo examination and submit the certificate to department.		

Semester- IV Course Type: Major Mandatory Paper 3(Practical) Course Code: DSC64503 Course Title: - Industrial Project Design		
No. of Credits:06	**Examination Scheme: CIE: 60 Marks**	**ESE: 90 Marks**
Prerequisites: Student should have basic knowledge of: • Logic Development and Various Programming Languages and Soft Skills		
The main objectives of this course are to: • acquiring knowledge about specific subject		
Expected Course Outcomes:		
On the successful completion of the course, student will be able to:		
CO1	Acquire knowledge of any specific subject	B4
CO2	Develop critical thinking capacity	B2,B3,B6
CO3	learn to self-regulate their learning process, including setting goals, managing time, and adjusting study strategies as needed.	B1, B5

B1 - Remember; B**2** - Understand; B**3** - Apply; B**4** - Analyze; B**5** - Evaluate; B**6**– Create
Guidelines for Industrial Project design
Formatting: Use a standard font size (e.g., 12), 1.5 line spacing between headings and body text, and a single tab for new paragraphs. **Margins and Page Size**: Maintain margins and use A4 size paper for your project documentation. **Content Structure**: Include essential sections such as the title page, abstract, table of contents, list of figures, list of tables, acknowledgments, and references. **Project Report Format**: Ensure the report follows a structured format, including chapters for introduction, background, work carried out, experimental results, conclusion, and future scope. **Documentation Importance**: Effective project documentation is crucial for project management, ensuring communication, decision-making, and continuity.