Karmaveer Bhaurao Patil University, Satara

Syllabus for

M. Sc. I Data Science

Under

Faculty of Science and Technology

(As per NEP 2020)

With effect from Academic Year 2024-2025

SYLLABUS FOR MASTER OF SCIENCE DATA SCIENCE

1. **Title: Subject: Data Science**

2. **Year of implementation: June 2023 onward**

3. **Preamble:**

Welcome to the Master of Science in Data Science program! In today's digital age, the world is generating vast amounts of data at an unprecedented rate. Extracting meaningful insights and making informed decisions from this data has become crucial for businesses, organizations, and societies alike. The M.Sc. Data Science program is designed to equip students with the knowledge, skills, and tools necessary to tackle complex data challenges and drive innovation in various domains.

Our program offers a comprehensive curriculum that combines theoretical foundations with practical applications. Through a blend of rigorous coursework, hands-on projects, and industry collaborations, we aim to cultivate a strong foundation in data science principles while emphasizing real-world problem-solving and critical thinking. Students will develop expertise in statistical analysis, machine learning, data visualization, data mining, and other essential areas, gaining proficiency in both the technical and analytical aspects of data science.

One of the unique aspects of our program is its interdisciplinary nature. Data science transcends traditional disciplinary boundaries, and we encourage students from diverse academic backgrounds to join us on this exciting journey. Whether you have a background in computer science, mathematics, statistics, engineering, or any other related field, this program will provide the necessary bridge to advance your skills and thrive in the data-driven landscape.

Our distinguished faculty comprises leading experts in the field of data science, bringing a wealth of industry experience and research expertise. They are committed to fostering a collaborative and engaging learning environment, where students can interact with faculty, fellow students, and industry professionals to gain valuable insights and expand their networks.

Beyond the classroom, we offer numerous opportunities for practical experience and professional development. Students will have access to cutting-edge technologies, state-of-the-art data labs, and industry partnerships, enabling them to work on real-world data problems and gain hands-on experience with industry-standard tools and platforms. Additionally, we organize workshops, seminars and guest lectures to expose students to the latest trends, emerging technologies, and industry best practices.

Programme Outcomes

PO. NO.	Programme Outcomes After completing M.Sc. programme, the students will be able to...
PO-1	Understand the fundamental and advancements of subject
PO-2	Study, Plan and Conduct Experiments in the lab to validate the ideas principals and theories acquired in the classrooms.
PO-3	Define their area of focus academia, research and development.
PO-4	Enhance scientific knowledge of the subject.
PO-5	Purse careers in various fields such as science, engineering, education, banking, business, public services, etc. or become an entrepreneur with precision, analytical thinking, innovative ideas, clarity thought, expression, and systematic approach.

Program Specific Objectives:

The students will be able to …

1. analyze real-world data sets using statistical, mathematical, and machine learning techniques to derive insights and make data-driven decisions.
2. develop the ability to handle, preprocess, and manage large volumes of structured and unstructured data using advanced data engineering techniques.
3. proficient in implementing machine learning algorithms, including supervised and unsupervised learning, deep learning, and reinforcement learning to solve complex problems.
4. apply data science techniques to business processes, enhancing decision-making in various domains such as finance, healthcare, marketing, and operations.

Programme Specific Outcomes

PSO. NO.	Programme Outcomes After completing M.Sc. (Data Science) programme the students will be able to..
PSO-1	Data Analysis and Interpretation: Graduates will be able to apply statistical methods and data mining techniques to analyze and interpret complex data sets.
PSO-2	Data Management: Graduates will be proficient in using various data management tools and techniques for data cleaning, preprocessing, storage, and retrieval.
PSO-3	Machine Learning and Predictive Modeling: Graduates will be skilled in designing, implementing, and evaluating machine learning models and algorithms for predictive analytics.
PSO-4	Big Data Technologies: Graduates will have hands-on experience with big data technologies such as Hadoop, Spark, and NoSQL databases, enabling them to handle and analyze large-scale data
PSO-5	Acquire the ability to conduct experiments and analyses data and present findings, demonstrating a solid understanding of scientific methodologies, implementing improvements.

M.Sc. -I Semester-I

Level	Semester	Course code	Course Title	No. of hours Per Week	Credits
6	I	MDST 411	Foundation for Data Science - I	4	4
		MDST 412	Programming using R	4	4
		MDST 413	Fundamentals of Data Science I	4	4
		MDST 414 E-I	Distributed Database Concept	2	2
		MDST 414 E-II	IOT for Data Science I	2	2
		MDST 415	Research Methodology	2	2
		MDSP 416	Practical Course I: Lab I Based On (MDST 411,412,413)	4	2
		MDSP 417	Practical Course II: Lab II Based On (MDST 414 E-I)	4	2
		MDSP 418	Practical Course II: Lab III Based On (MDST 415)	4	2
Total					**22**

M.Sc. -I Semester-II

Level	Semester	Course code	Course Title	No.of hours Per Week	Credits
6	II	MDST 421	Foundation for Data Science - II	4	4
		MDST 422	Python Programming	4	4
		MDST 423	Data Preparation Analysis	4	4
		MDST 424 E-I	AI for Data Science	2	2
		MDST 424 E-II	IOT for Data Science II	2	2
		MDST 425	Research Project	4	4
		MDSP 426	Practical Course I: Lab I Based On (MDST 411,412,413)	4	2
		MDSP 427	Practical Course II: Lab II Based On (MDST 424 E-II)	2	2
Total					**22**

Evaluation Structure

Semester-I

Course	Course Category	Course Title	Course Code	Internal Evaluation				ESE	Total Marks	Credits
				CCE- I	Mid-Semester	CCE-II	Activity			
DSC	T	Foundation for Data Science - I	MDST 411	10	10	10	10	60	100	04
	T	Programming using R	MDST 412	10	10	10	10	60	100	04
	T	Fundamentals of Data Science I	MDST 413	10	10	10	10	60	100	04
	P	Lab I Based On (MDST 411,412,413)	MDSP 416	--	--	--	--	50	50	02
DSE (1 Theory Papers Out of Two)	T	Distributed Database Concept	MDST 414	05	05	05	05	30	50	02
	T	IOT for Data Science I	MDST 414	05	05	05	05	30	50	02
	P	Lab II Based On (MDST 414 E-I)	MDSP 417	--	--	--	--	50	50	02
RM	T	Research Methodology	MDST 415	05	05	05	05	30	50	02
	P	Lab III Based On (MDST 415)	MDSP 418	--	--	--	--	50	50	02
Total									**550**	**22**

Semester-II

Course	Course Category	Course Title	Course Code	Internal Evaluation				ESE	Total Marks	Credits
				CCE- I	Mid-Semester	CCE-II	Activity			
DSC	T	Foundation for Data Science - II	MDST421	10	10	10	10	60	100	04
	T	Python Programming	MDST422	10	10	10	10	60	100	04
	T	Data Preparation Analysis	MDST423	10	10	10	10	60	100	04
	P	Lab I Based On (MDST 411,412,413)	MDSP 426	--	--	--	--	50	50	02
DSE (1 Theory Papers Out of Two)	T	AI for Data Science	MDST424	05	05	05	05	30	50	02
	T	IOT for Data Science II	MDST424	05	05	05	05	30	50	02
	P	Lab II Based On (MDST 424 E-II)	MDSP 427	--	--	--	--	50	50	02
RP	P	Research Project	MDSP 425	--	--	--	--	100	100	04
Total									**550**	**22**

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: Major	**Marks:** 60
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 411: Foundation for Data Science - I

Course Objectives:

1) Understand the concept of Descriptive statistics.
2) Study correlation and regression.
3) Understand the applications probability theory.
4) Study the probability distribution.

Course Outcomes:

1) Comprehend the concepts of descriptive statistics.
2) Apply recent concepts in correlation and regression.
3) Utilize probability theory.
4) Imbibe concepts of probability distribution.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Descriptive Statistics and Nature of Data** **1.1** Measures of Central Tendency, Measures of Variation, **1.2** Quartiles and Percentiles, **1.3** Moments, Skewness and Kurtosis., Data Classification, Tabulation, **1.4** Frequency and Graphic Representation, Sampling Techniques.	15
Module -2:	**Module -2: Correlation and Regression** **2.1** Correlation, Scatter Diagram, Karl Pearson's Correlation Coefficient, Rank Correlation, **2.2** Multiple Correlation, Partial Correlation – Concept, Definition, Formulae, Properties, **2.3** Correlation Coefficient for Bivariate Frequency Distribution, Regression, Regression Coefficients, Fitting of Regression Lines, Method of least Square, **2.4** Multiple Regression: Concept of Multiple Regression. Yule's Notations, Fitting of Multiple Regression Plan, Partial Regression confidents, interpretations, Logistic Regression.	15
Module -3:	**Module -3: Probability Theory** **3.1** Random Experiment, Sample Space, Events, Axiomatic Definition of Probability , Addition Theorem, **3.2** Multiplication Theorem, Baye's Theorem Applications. Independence of Events, Definition of Probability in terms of odds ratio, **3.3** Distribution Function Continuous and Discrete Random Variables – Distribution Function of a Random Variable, **3.4** Probability Mass Functions and Probability Density Functions–Characteristic Functions –Central Limit Theorems.	15
Module -4:	**Module -4: Probability Distributions** **4.1** Probability Distributions – Recurrence Relationships – Moment Generating Functions, **4.2** Cumulate Generating Functions Continuous Probability Distributions Rectangular Distribution, **4.3** Binomial Distribution, Poisson Distribution Continuous Probability Distributions, **4.4** Uniform Distribution - Normal Distribution – Exponential Distribution.	15

Reference Books:-

1) Gelman, A., Hill, J., & Vehtari, A., Regression and Other Stories (Cambridge, UK: Cambridge University Press), (2020).

2) Shumway, R. H., & Stoffer, D. S., Time Series Analysis and Its Applications: With R Examples (New York: Springer). (2017).

3) Efron, B., & Hastie, T. Computer Age Statistical Inference: Algorithms, Evidence, and Data Science (Cambridge, UK: Cambridge University Press), (2016).

4) Good fellow, I., Bengio, Y., & Courville, A. Deep Learning (Cambridge, MA: MIT Press), (2016).

5) Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. ,Bayesian Data Analysis (Boca Raton, FL: CRC Press), (2013).

Evaluation Pattern:

Total Marks: 100/60

Internal Continuous Evaluation:	**End Semester Examination:**
• CCE-I – 10 Marks • CCE-II – 10 Marks • Mid Semester – 20 Marks • Activity – 10 Marks	• Question -1 (02 Marks = 2 X 6=12 Marks) • Question -2 (06 Marks = 6 X 2=12 Marks) • Question -3 (06 Marks = 6 X 2=12 Marks) • Question -4 (06 Marks = 6 X 2=12 Marks) • Question -5 (04 Marks = 4 X 3=12 Marks) • Question -6 (04 Marks = 4 X 3=12 Marks) • Question -7 (04 Marks = 4 X 3=12 Marks)

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: Major	**Marks:** 60
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 412: Programming using R

Course Objectives:

1) Master the use of the R and R Studio interactive environment.
2) Explore and understand how to use the R documentation.
3) Read Structured Data into R from various sources.
4) Understand the different data types and data structure in R.

Course Outcomes:

1) Install, Code and Use R Programming Language in R Studio IDE to perform basic tasks on Vectors, Matrices and Data frames. Describe key terminologies, concepts and techniques employed in Statistical Analysis.
2) Define, Calculate, Implement Probability and Probability Distributions to solve a wide variety of problems
3) Conduct and interpret a variety of Hypothesis Tests to aid Decision Making.
4) Explain, Analyze, Interpret Correlation and Regression to analyses the underlying relationships between different variables.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Fundamental of R Programming** **1.1** Introduction to R – Help Functions in R – Vectors – Vectorized Operations, **1.2** Functions in R – Packages in R , Data Types, Subletting, writing data, Reading from csv files, **1.3** Creating a vector and vector operation, Initializing data frame, Flow control: For loop, **1.4** If condition, debugging toolset- directing R Output.	15
Module -2:	**Module -2: Matrices, Arrays and Lists** **2.1** Matrix Operations, Adding and Deleting Rows and Columns, **2.2** Higher Dimensional Arrays – Lists – General List Operations, **2.3** Accessing List Components and Values, **2.4** Applying functions to Lists.	15
Module -3:	**Module -3: Data Frames** **3.1** Creating Data Frames, Matrix-like Operations on a Data Frame, **3.2** Merging Data Frames, Applying functions to Data Frames, **3.3** Factors and Tables – Common Functions used with Factors, **3.4** Working with Tables.	15
Module -4:	**Module -4: Data manipulation and Visualization** **4.1** List Management, Data Transformation, Merging Data Frames, **4.2** Outlier Detection, Combining multiple vectors, **4.3** Creating bar chart and dot plot, **4.4** Creating histogram and box plot, plotting with base graphics, Plotting and coloring in R.	15

Reference Books:-

1) Matloff, N. Parallel Computing for Data Science: With Examples in R, C++, and CUDA (Boca Raton, FL: CRC Press), (2019).

2) Lander, J. P. R for Everyone: Advanced Analytics and Graphics (Addison-Wesley Data & Analytics Series), (2017).

3) Silge, J., & Robinson, D., Text Mining with R: A Tidy Approach (Sebastopol, CA: O'Reilly Media),

<table>
<tr><td colspan="2">(2017).
4) Wickham, H., & Grolemund, G., R for Data Science: Import, Tidy, Transform, Visualize, and Model Data (Sebastopol, CA: O'Reilly Media), (2016).
5) Peng, R. D., R Programming for Data Science (Boca Raton, FL: CRC Press), (2016).</td></tr>
<tr><td colspan="2">Evaluation Pattern:</td></tr>
<tr><td colspan="2">Total Marks: 100/60</td></tr>
<tr><td>Internal Continuous Evaluation:
• CCE-I – 10 Marks
• CCE-II – 10 Marks
• Mid Semester – 20 Marks
• Activity – 10 Marks</td><td>End Semester Examination:
• Question -1 (02 Marks = 2 X 6=12 Marks)
• Question -2 (06 Marks = 6 X 2=12 Marks)
• Question -3 (06 Marks = 6 X 2=12 Marks)
• Question -4 (06 Marks = 6 X 2=12 Marks)
• Question -5 (04 Marks = 4 X 3=12 Marks)
• Question -6 (04 Marks = 4 X 3=12 Marks)
• Question -7 (04 Marks = 4 X 3=12 Marks)</td></tr>
</table>

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: Major	**Marks:** 60
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 413 Fundamentals of Data Science - I

Course Objectives:

1) Understand the recommendation system and two basic architectures for a recommendation system.
2) Develop the fundamental knowledge and understand concepts to become a data science professional.
3) Learn statistical methods and machine learning algorithms required for Data Science.
4) Visualize data and use for communicating stories from data.

Course Outcomes:

1) Apply data science processes to an e-commerce data and demonstrate the use of estimation methods for analyzing this data.
2) Compare and apply appropriate machine learning algorithms for classification.
3) Compare and choose one data visualization method for effective visualization of data.
4) Apply standard clustering methods to analyze social network graph.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Data Science Concepts** **1.1** What is Data Science , importance of data science, **1.2** Big data and data Science, The current Scenario, **1.3** Industry Perspective Types of Data: Structured vs. Unstructured Data, **1.4** Quantitative vs. Categorical Data, Big Data vs. Little Data, Data science process, Role Data Scientist.	15
Module -2:	**Module -2: Machine Learning Algorithms** **2.1** Machine Learning Algorithms: Linear Regression, **2.2** K-nearest Neighbors (k- NN), K-mean, Spam Filters, **2.3** Naive Bayes, and Wrangling: Naive Bayes, **2.4** Comparing Naive Bayes to k-NN, Scraping the Web: APIs and Other Tools.	15
Module -3:	**Module -3: Data Visualization** **3.1** Data visualization: Introduction, **3.2** Types of data visualization, Data for visualization: Data types, **3.3** Data encodings, Retinal variables, **3.4** Map ping variables to encodings, Visual encodings.	15
Module -4:	**Module -4: Social Network Analysis** **4.1** Social Networks as Graphs, Varieties of Social Networks, **4.2** Graphs With Several Node Types, **4.3** Clustering of Social-Network Graphs: Distance Measures for Social-Network Graphs, **4.4** Applying Standard Clustering Methods, Betweenness, The Girvan-Newman Algorithm, Using Betweenness to Find Communities.	15

Reference Books:-

1) Laura Igual and Santi Segui, Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications,Springer; 1st ed. 2017 edition.
2) Cao, L., Data Science: An Introduction (Cambridge, MA: MIT Press), (2018).
3) Donoho, D. L. ,50 Years of Data Science (New York: Springer), (2017).
4) Wickham, H., ggplot2: Elegant Graphics for Data Analysis (New York: Springer), (2017).

<table>
<tr><td colspan="2">5) VanderPlas, J., Python Data Science Handbook: Essential Tools for Working with Data (Sebastopol, CA: O'Reilly Media), (2016).</td></tr>
<tr><td colspan="2">Evaluation Pattern:</td></tr>
<tr><td colspan="2">Total Marks:60</td></tr>
<tr><td>Internal Continuous Evaluation:
• CCE-I – 10 Marks
• CCE-II – 10 Marks
• Mid Semester – 20 Marks
• Activity – 10 Marks</td><td>End Semester Examination:
• Question -1 (02 Marks = 2 X 6=12 Marks)
• Question -2 (06 Marks = 6 X 2=12 Marks)
• Question -3 (06 Marks = 6 X 2=12 Marks)
• Question -4 (06 Marks = 6 X 2=12 Marks)
• Question -5 (04 Marks = 4 X 3=12 Marks)
• Question -6 (04 Marks = 4 X 3=12 Marks)
• Question -7 (04 Marks = 4 X 3=12 Marks)</td></tr>
</table>

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: E I	**Marks:** 30
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDST 414 E I : Distributed Database Concepts.

Course Objectives:

1) Understand the various aspects in Distributed Data.
2) Understand query processing and optimization in Distributed Database.
3) Management of distributed data with different levels of transparency.
4) Understand how to use database management tools in resolving deadlock situations.

Course Outcomes:

1) Design distributed database for any Real-World application.
2) Write query for data manipulation on Distributed Database.
3) Manage Transaction using fragmentation.
4) Handle deadlock situation in Distributed Database.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Fundamental of Distributed Database Design** **1.1** What is Distributed Database System (DDBS), Features of DDBS, Design issue in DDBS, **1.2** Distributed DBMS architecture:- Client/server System, Peer- to-Peer, Multi-Database system, Levels of distribution transparency, **1.3** Types of Data Fragmentation, Integrity Constraints in Distributed Databases, Framework of Distributed Databases Design, Design of Database Fragmentation, Allocation of fragments, Transparencies in Distributed Database Design. Structured vs. Unstructured Data, **1.4** Quantitative vs. Categorical Data, Big Data vs. Little Data.	15
Module -2:	**Module -2: Distributed Query Processing , Optimization, Transactions Management** **2.1** Concept, objective, and phases of distributed query processing, Translation of global queries to fragment queries, Query optimization in centralized databases, **2.2** join queries, general queries, TRANSLATION OF GLOBAL QUERIES TO FRAGMENT QUERIES: Equivalence Transformations For Queries, Transforming Global Queries into Fragment Queries, Distributed Grouping and Aggregate Function Evaluation, Parametric Queries, **2.3** THE MANAGEMENT OF DISTRIBUTED TRANSACTIONS: A Framework for Transaction Management, Supporting Atomicity of Distributed Transactions, **2.4** Concurrency Control for Distributed Transactions, Architectural Aspects of Distributed Transactions, Transaction Schedules in Distributed databases.	15
Module -3:	**Module -3: Concurrency Control in DDBMS** **3.1** Concurrency Control Based on Timestamps, **3.2** Optimistic Methods for Distributed Concurrency Control, **3.3** Introduction to Deadlock, Distributed Deadlock prevention, **3.4** Avoidance detection and recovery, Two-Phase and Three-Phase Commit Protocol.	15
Module -4:	**Module -4: Heterogeneous Database** **4.1** Architecture of Heterogeneous Database, **4.2** Interface Standards for Relational Database : ODBC ODBC architecture, **4.3** functionality and usage of ODBC Database Integration:- Schema Translation and schema Integration,	15

	4.4 Query processing issues in Heterogeneous database.	

Reference Books:-

1) Database Systems: A Practical Approach to Design, Implementation and Management- Thomas Connolly, Carolyn Begg, Pearson Publisher, 4th Edition.
2) Database Management Systems - Raghu Ramakrishnan and Johannes Gehrke, McGraw-Hill Education publisher, illustrated Edition,2003.
3) Carlo Zaniolo, Stefano Ceri, Christos Faloustsos, R.T. Snodgrass, V.S. Subrahmanian, "Advanced Database Systems", Morgan Kaufman, 1997.
4) Distributed Databases principles & systems by Stefano Ceri, Giuseppe Pelagatti, 2nd edition, McGraw-Hill, New York, 1985, ISBN 0-07-010829-3.
5) N. TamerOzsu, Patrick Valduriez, "Principles of Distributed Database Systems", 2nd , IllustratedEdition, Prentice Hall International Inc., 1999.

Evaluation Pattern:

Total Marks: 50

Internal Continuous Evaluation (20 Marks):	**End Semester Examination (30 Marks):**
• CCE - I : 10 Marks: Objective • CCE - II: 10 Marks: Objective • Mid Semester Exam: 20 Marks: Subjective (20 Marks converted to 10 marks) • Activity: 10 marks	• Question -1: Solve the following questions (5 questions of 2 Marks) • Attempt any two questions from Q. 2 to Q. 4 (three questions of 10 marks)

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: E II	**Marks:** 30
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDST- 414 E II: IOT for Data Science I

Course Objectives:

1) Explain basic functioning of sensors and display units.
2) Familiarize the concepts of signal processing and converting elements.
3) Understand basic operation in digital systems.
4) Acquire the knowledge of microcontrollers and communication Systems.

Course Outcomes:

1) Explain the operation/working of Transducer and sensor.
2) Describe Signal conditioning and Operational amplifier.
3) Explain number system and design and construct logic gates.
4) Discuss basics of 8051microcontroller, wireless technology and communication systems.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Sensors** **1.1** Sensors and Transducers – Classification, Potentiometer, **1.2** Strain Gauge, Piezoelectric Sensor, **1.3** Linear Variable Differential Transformer (LVDT), Resistance temperature detectors (RTD), **1.4** Thermocouples, Displays - LCD, Light Emitting Diode (including OLED) displays.	15
Module -2:	**Module -2: Signal Conditioning And Data Logging** **2.1** Operational Amplifiers- Inverting, **2.2** Non-Inverting, Instrumentation Amplifier, Active filters: - Low pass, High pass filter, **2.3** Analog to Digital Converter – Successive Approximation, **2.4** Digital to Analog Converter – R- 2R ladder, Data Logging.	15
Module -3:	**Module -3: Digital Systems And Data Processing Devices** **3.1** Analog and Digital signals, Number systems - Decimal, **3.2** Binary, Hexadecimal. Logic gates-AND, **3.3** OR, NOT, NOR, NAND, EX-OR. Comparison of Microcontroller & Microprocessor, 8-Bit, **3.4** 16-Bit Microcontrollers, Study of 8051 and its Family (89C51, DS5000, 89C51VRD2),Architecture of 8051, Memories, I/O Ports.	15
Module -4:	**Module -4: Data Communication System** **4.1** Wireless Technology: Bluetooth, **4.2** Wi-Fi, Wi-Max. Communication System: Global Positioning System, **4.3** Mobile Communication, **4.4** Satellite Communication (Qualitative Idea).	15

Reference Books:-

1) The 8051 Microcontroller and Embedded Systems, Muhammad A. Mazidi, J.G. Mazidi, R.D. Mckinlay, Pearson India Education Services Pvt. Ltd., Seventeenth Edition, 2017

2) Communication Systems, Simon Heykin, Wiley Pvt. Ltd., Third Edition, 2016.

3) The 8051 Microcontroller, Kenneth Ayala, Cengage Learning India Pvt. Ltd., Third Edition, 2014.

4) Electronic Instrumentation, H.C. Kalsi, McGraw Hill (India) Pvt. Ltd. New Delhi, Twelfth Edition,2014.

5) OP-AMP and Linear Integrated Circuits, Ramakant A. Gayakwad, PHI Learning Pvt. Ltd. Delhi, Fourth Edition, 2014.	
Evaluation Pattern:	
Total Marks: 50	
Internal Continuous Evaluation (20 Marks): • CCE - I : 10 Marks: Objective • CCE - II: 10 Marks: Objective • Mid Semester Exam: 20 Marks: Subjective (20 Marks converted to 10 marks) • Activity: 10 marks	**End Semester Examination (30 Marks):** • Question -1: Solve the following questions (5 questions of 2 Marks) • Attempt any two questions from Q. 2 to Q. 4 (three questions of 10 marks)

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: Major	**Marks:** 60
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDST- 415: Research Methodology.

Course Objectives:

1) Understand the fundamentals of research and Develop research skills.
2) Explore different research methodologies and data collection and analysis techniques.
3) Enhance critical thinking and problem-solving skills.
4) Apply research methodologies in practical settings.

Course Outcomes:

1) Explain principles and concepts of research.
2) Apply different research methodologies.
3) Collect and analyze data and Interpret research findings.
4) Evaluate research studies and Apply research skills in practical settings.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Foundations of Research** **1.1** Meaning, Objectives, Motivation, Utility, **1.2** Concept of theory, empiricism, deductive and inductive theory, **1.3** Characteristics of scientific method – Understanding the language of research – Concept, **1.4** Construct, Definition, Variable. Research Process.	15
Module -2:	**Module -2: Problem Identification & Formulation** **2.1** Research Question, Investigation Question, **2.2** Measurement Issues, Hypothesis, **2.3** Qualities of a good Hypothesis, Null Hypothesis & Alternative Hypothesis, **2.4** Hypothesis Testing ,Logic & Importance.	15
Module -3:	**Module -3: Research Design** **3.1** Concept and Importance in Research, **3.2** Features of a good research design, Exploratory Research Design, **3.3** concept, types and uses, Descriptive Research Designs, concept, **3.4** types and uses. Experimental Design: Concept of Independent & Dependent variables.	15
Module -4:	**Module -4: Qualitative and Quantitative Research** **4.1** Qualitative research, Quantitative research, **4.2** Concept of measurement, causality, generalization, replication, **4.3** Merging the two approaches, Measurement, Sampling, **4.4** Data analysis.	15

Reference Books:-

1) Donald Cooper & Pamela Schindler, "Business Research Methods", TMGH, 9th edition.
2) Alan Bryman & Emma Bell ,"Business Research Methods" , Oxford University Press.
3) Kothari C.R. Research Methodology.
4) Research Methodology Dr. Shalu Katyal, 1 April 2024
5) Research Design: Qualitative, Quantitative ,J. David Creswell ,22 November 2022.

Evaluation Pattern:	
Total Marks: 100/60	
Internal Continuous Evaluation: • CCE-I – 10 Marks • CCE-II – 10 Marks • Mid Semester – 20 Marks • Activity – 10 Marks	**End Semester Examination:** • Question -1 (02 Marks = 2 X 6=12 Marks) • Question -2 (06 Marks = 6 X 2=12 Marks) • Question -3 (06 Marks = 6 X 2=12 Marks) • Question -4 (06 Marks = 6 X 2=12 Marks) • Question -5 (04 Marks = 4 X 3=12 Marks) • Question -6 (04 Marks = 4 X 3=12 Marks) • Question -7 (04 Marks = 4 X 3=12 Marks)

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: Major	**Marks:** 50
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDSP- 416: Lab Course MDST411, MDST412, MDST413

Course Objectives:

1) Master the use of the R and R Studio interactive environment.
2) Study correlation and regression, and applications probability theory.
3) To learn statistical methods and machine learning algorithms required for Data Science.
4) To visualize data and use for communicating stories from data.

Course Outcomes:

1) Install, Code and Use R Programming Language in R Studio IDE to perform basic tasks on Vectors, Matrices and Data frames.
2) Describe key terminologies, concepts and techniques employed in Statistical Analysis.
3) Conduct and interpret a variety of Hypothesis Tests to aid Decision Making.
4) Define, Calculate, Implement Probability and Probability Distributions to solve a problem.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Lab I Course MDST411, MDST412, MDST413** **1.** Diagrammatic and Graphical Representation **2.** Measures of central tendency and measures of dispersion **3.** Correlation and Regression **4.** Applications of probability **5.** Fitting of discrete and continuous distribution **6.** Creating and displaying Data. **7.** Matrix manipulations , Creating and manipulating a List and an Array **8.** Creating a Data Frame and Matrix-like Operations on a Data Frame **9.** Merging two Data Frames and Applying functions to Data Frames **10.** Using Functions with Factors, Accessing the Internet and String Manipulations **11.** Histograms and Density Charts **12.** Visualization Effects, Plotting with Layers **13.** Overriding Aesthetics and Histograms and Density Charts **14.** Simple Linear Regression – Fitting, Evaluation **15.** Simple Linear Regression – Fitting, Evaluation **16.** Case Study On Big Data **17.** Case Study on Applications of Data Science **18.** Case Studies on Supervised **19.** Case Studies Unsupervised learnings **20.** Case Study on Data Visualization	8

Reference Books:-

1) Ken Black, 2013, *Business Statistics*, New Delhi, Wiley.
2) Lee, Cheng. et al., 2013, *Statistics for Business and Financial Economics*, New York: Heidelberg Dordrecht.
3) Anderson, David R., Thomas A. Williams and Dennis J. Sweeney, 2012, *Statistics for Business and Economics*, New Delhi: South Western.
4) Waller, Derek, 2008, *Statistics for Business*, London: BH Publications.
5) Zaki, M. J., & Meira Jr., W., Data Mining and Analysis: Fundamental Concepts and Algorithms (Cambridge, UK: Cambridge University Press), (2014).

Evaluation Pattern:	
Total Marks: 50	
	End Semester Examination: • Question -1 (10 Marks) • Question -2 (10 Marks) • Question -3 (10 Marks) • Activity 20 Marks.

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: Major	**Marks:** 30
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDSP 417: Lab Course II MDST 414 E1

Course Objectives:

1) Understand the fundamental concepts of data.
2) Master the use of the SQL and interactive environment.
3) Understand principles of databases.
4) Discuss the concept of procedure oriented, object-oriented programming languages, Database Management.

Course Outcomes:

1) Apply the basics of data, information, system and Database.
2) Evaluate basics of different database models for software development.
3) Design the basics of Relational algebra operations and Relational Calculus.
4) Demonstrate SQL basics and write queries to perform different operations on real world data.

Module	Title and Contents	Hrs
Module -1:	**Module -1: MDSP 416: Lab II Course MDST414E-I** **1.** Installation of PG SQL **2.** To perform operation on DDL **3.** To perform operation on DML **4.** To perform operation on DCL **5.** To perform operation on TCL **6.** To perform SQL Keywords **7.** To perform the query on Operator **8.** To perform queries on clauses **9.** To perform Aggregate Functions **10.** To perform operation Joins **11.** To perform Order by query **12.** To perform Group by statement **13.** To perform operation on Subquery **14.** To perform the operation on Indexing **15.** To perform the operation on View **16.** To perform the operation on LIKE, IN, Rollback in sql **17.** To perform the operation on Trigger **18.** To perform the operation on Cursor **19.** To perform operation NULL operator and NOT NULL **20.** To perform operation on CHECK constraint .	8

Reference Books:-

1) R. Elmasri, S.B. Navathe, (2010), Fundamentals of Database Systems 6th Edition,Pearson Education.
2) R. Ramakrishanan, J. Gehrke, (2002), Database Management Systems 3rd Edition,McGraw-Hill.
3) Silberschatz, H.F. Korth, S. Sudarshan, (2010), Database System Concepts 6th Edition,McGraw Hill
4) Silberschatz Abraham, Database System Concepts, Mc Graw Hill, 2021
5) Teorey Toby, Database Modeling and Design: Logical Design, Morgan Kaufmann, 2010.

Evaluation Pattern:	
Total Marks: 50	
	End Semester Examination: • Question -1 (10 Marks) • Question -2 (10 Marks) • Question -3 (10 Marks) • Activity 20 Marks.

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - I
Type: Major	**Marks:** 50
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDSP 418: Lab III Based on MDST 415

Course Objectives:

1) Understand the fundamentals of research and Develop research skills.
2) Explore different research methodologies and data collection and analysis techniques.
3) Enhance critical thinking and problem-solving skills.
4) Apply research methodologies in practical settings.

Course Outcomes:

1) Explain principles and concepts of research.
2) Apply different research methodologies.
3) Collect and analyze data and Interpret research findings.
4) Evaluate research studies and Apply research skills in practical settings.

Module	Title and Contents	Hrs
Module -1:	**Module -1: MDSP 416: Lab III Based on MDST 415** **1.** Identify a real-life problem and describe how it can be converted into a research problem. **2.** Write a brief research proposal including *title*, *objective*, and *research motivation*. **3.** Choose a social issue and explain whether you would apply *deductive* or *inductive* theory. Justify your answer. **4.** From a research article, identify and define one *construct*. Explain how it was measured. **5.** Create a flowchart of the research process using a topic of your choice (e.g., mobile usage and academic performance). **6.** Formulate a *research question* and *investigation questions* for studying the effects of online learning on mental health. **7.** Design a questionnaire (5–7 questions) to measure student satisfaction in online courses. Highlight any measurement issues you faced. **8.** Identify a topic and formulate both *null* and *alternative* hypotheses. **9.** Choose any hypothesis and outline the steps to test it using real or dummy data. **10.** Observe a real-world scenario and develop a *testable* hypothesis. Explain the logic behind it. **11.** Select a research topic and design an *exploratory research plan* using secondary data and interviews. **12.** Choose a subject (e.g., "effects of diet on weight loss") and create a *descriptive research design* framework. **13.** Identify the *independent* and *dependent* variables in a hypothetical study: *Example: "Effect of social media usage on exam performance."* **14.** Prepare a table showing variables, type of study, and expected data collection method.	8

	15. Critically evaluate an existing research article's design and suggest improvements. **16.** Conduct 3 short interviews on any topic (e.g., study habits) and summarize key qualitative insights. **17.** Collect simple numerical data from 10 people (e.g., number of study hours/day) and represent it using a bar chart. **18.** Use both qualitative (interview) and quantitative (survey) data to study one issue (e.g., smartphone addiction). Combine the results. **19.** Choose a topic and decide which sampling technique (random, stratified, convenience) is most suitable. Justify your choice. **20.** Calculate basic statistical analysis (mean, median, SD) using dummy data you collect from your surroundings.	

Reference Books:-

1) Donald Cooper & Pamela Schindler, "Business Research Methods", TMGH, 9th edition.
2) Alan Bryman & Emma Bell ,"Business Research Methods" , Oxford University Press.
3) Kothari C.R. Research Methodology.
4) Research Methodology Dr. Shalu Katyal, 1 April 2024
5) Research Design: Qualitative, Quantitative ,J. David Creswell ,22 November 2022.

Evaluation Pattern:	
Total Marks: 50	
	End Semester Examination: • Question -1 (10 Marks) • Question -2 (10 Marks) • Question -3 (10 Marks) • Activity 20 Marks.

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - II
Type: Major	**Marks:** 60
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 421: Foundations for Data Science - II

Course Objectives:

1) Understand the various aspects in district mathematics in data science.
2) Understand Data Analysis and probability theory.
3) Study the concept of Linear Algebra and Calculus.
4) Understand how to use Regression model.

Course Outcomes:

1) Design discrete mathematics for any real-world application.
2) Handle data analysis and probability theory.
3) Apply concept of linear algebra and calculus.
4) Manage regression model.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Discrete mathematics for Data Science:** **1.1** Concept of set, cardinality of set, **1.2** finite, infinite and uncountable infinite sets, **1.3** Basic set operations, **1.4** Principal of inclusion Exclusion, Trees: Basic terminologies, search tree: Binary & M-ary tree.	15
Module -2:	**Module -2: Data Analysis & Probability Theory** **2.1** Data Representation, Average, Spread, Experiments, **2.2** Outcomes, Events, Probability, **2.3** Permutations and Combinations, Random Variables, **2.4** Probability Distributions, Mean and Variance of a Distribution, Binomial, Poisson, and Hyper geometric Distributions.	15
Module -3:	**Module -3: Linear Algebra and Calculus** **3.1** Linear Algebra: Matrix and vector algebra, **3.2** systems of linear equations using matrices, **3.3** linear independence, **3.4** Matrix factorization concept/LU decomposition, Understanding of calculus: concept of function and derivative.	15
Module -4:	**Module -4: Graph Theory for Data Science** **4.1** Connectivity: Definition and Simple Properties, **4.2** Bridges, Cut Vertices and Connectivity, **4.3** Euler Tours: Euler Graphs, Properties of Euler Graph, **4.4** the Chinese Postman Problem.	15

Reference Books:-

1) Applied Statistics and Probability for Engineers, Douglas C. Montgomery, George C. Runger, 2018
2) Goodfellow, I., Bengio, Y., & Courville, A., Deep Learning (Cambridge, MA: MIT Press), (2016).
3) Vanderbei, R. J., Linear Programming: Foundations and Extensions (New York: Springer), (2014).

<table>
<tr><td colspan="2">4) The R Software-Fundamentals of Programming and Statistical Analysis -Pierre Lafaye de Micheaux, RémyDrouilhet, Benoit Liquet, Springer 2013
5) Wasserman, L. All of Statistics: A Concise Course in Statistical Inference (New York: Springer), (2013).</td></tr>
<tr><td colspan="2">Evaluation Pattern:</td></tr>
<tr><td colspan="2">Total Marks: 100/60</td></tr>
<tr><td>Internal Continuous Evaluation:
• CCE-I – 10 Marks
• CCE-II – 10 Marks
• Mid Semester – 20 Marks
• Activity – 10 Marks</td><td>End Semester Examination:
• Question -1 (02 Marks = 2 X 6=12 Marks)
• Question -2 (06 Marks = 6 X 2=12 Marks)
• Question -3 (06 Marks = 6 X 2=12 Marks)
• Question -4 (06 Marks = 6 X 2=12 Marks)
• Question -5 (04 Marks = 4 X 3=12 Marks)
• Question -6 (04 Marks = 4 X 3=12 Marks)
• Question -7 (04 Marks = 4 X 3=12 Marks)</td></tr>
</table>

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - II
Type: Major	**Marks:** 60
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 422: Python Programming

Course Objectives:

1) Understand python programming.
2) Imbibe build and package python modules for reusability.
3) Understand suitable knowledge about their implementation.
4) Understand various file handling techniques and database interactions.

Course Outcomes:

1) Explain principles python programming.
2) Implement clear and effective python code.
3) Discuss Data Analysis using python libraries.
4) Try to understand Error handling.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Introduction To Python** **1.1** Introduction, Various IDEs,Numeric data types: int, float, complex, String, **1.2** list and list slicing, Tuples, Control Flow Conditional blocks using if, **1.3** Else and elif Simple for and while loops in python For loop using ranges, **1.4** string, list and dictionaries Loop manipulation using pass, continue, break and else.	15
Module -2:	**Module -2: Functions and Packages** **2.1** Functions Arguments, **2.2** Lambda Expressions, Function Annotations, **2.3** Modules Organizing python projects into modules Importing own module as well as external modules, **2.4** Packages, Programming using functions, modules and external packages.	15
Module -3:	**Module -3: Data Structures and Python File Operations** **3.1** Lists as Stacks, Queues, Comprehensions, Tuples and sequences, Sets, **3.2** Dictionaries, reading config files in python, Writing log files in python, **3.3** Understanding read functions, read (), redline () and readlines (), Understanding write functions, write() and writelines(), **3.4** Manipulating file pointer using seek, Programming using file operations.	15
Module -4:	**Module -4: Database Interaction SQL** **4.1** Database connection using python, **4.2** Creating and searching tables, **4.3** Reading and storing config information on database, **4.4** Programming using database connections.	15

Reference Books:-

1) McKinney, W., Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (Sebastopol, CA: O'Reilly Media), (2017).
2) Dusty Phillips, "Python 3 Object-oriented Programming Second Edition", Pack Publishing, Greg, 2015.
3) VanderPlas, J." Python Data Science Handbook: Essential Tools for Working with Data (Sebastopol, CA: O'Reilly Media)", (2016).
4) Charles Dierbach," Introduction to Computer Science Using Python: A Computational Problem-Solving Focus", John Wiley & Sons,2013.
5) Jan van Eijck , Christina Unger, "Computational Semantics with Functional Programming", Cambridge University Press, 2012 .

Evaluation Pattern:

Total Marks: 60

Internal Continuous Evaluation:	**End Semester Examination:**
• CCE-I – 10 Marks • CCE-II – 10 Marks • Mid Semester – 20 Marks • Activity – 10 Marks	• Question -1 (02 Marks = 2 X 6=12 Marks) • Question -2 (06 Marks = 6 X 2=12 Marks) • Question -3 (06 Marks = 6 X 2=12 Marks) • Question -4 (06 Marks = 6 X 2=12 Marks) • Question -5 (04 Marks = 4 X 3=12 Marks) • Question -6 (04 Marks = 4 X 3=12 Marks) • Question -7 (04 Marks = 4 X 3=12 Marks)

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - II
Type: Major	**Marks:** 60
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 423: Data Preparation Analysis

Course Objectives:

1) understand the importance of data and data preprocessing.
2) understand data cleaning and conditioning.
3) understand an ETL – Extract, Transform and Load – process and ETL tools.
4) get acquainted with data visualization techniques for exploratory analysis.

Course Outcomes:

1) Design Data Gathering and Data Discovery.
2) Handle Cleaning and Conditioning Data.
3) Manage Exploratory Analysis.
4) Apply Data Visualization and Advanced Tools for Data Preparation.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Data Gathering and Data Discovery** **1.1** Identifying potential data sources, **1.2** Gathering data, Data discovery- understanding the data, assessing data, **1.3** data formats, Parsing, Selecting features, **1.4** Transformation, Scalability and real-time issues.	15
Module -2:	**Module -2: Cleaning and Conditioning Data** **2.1** Data Preparation Basic Models: Data Integration, **2.2** Data Cleaning, Data Normalization, Min-Max Normalization, Z-score Normalization, **2.3** Decimal Scaling Normalization, Consistency checking, Heterogeneous and missing data, **2.4** Dealing with missing values, Duplicate, values, Noise, Inconsistent data, Outliers.	15
Module -3:	**Module -3: Exploratory Analysis** **3.1** Formulating Hypothesis, Data Terminology, **3.2** Data Exploration, Data Exploration through Summary Statistics, **3.3** Data Exploration through Plots, Feature Engineering, Feature selection, **3.4** Feature transformation, Dimensionality reduction	15
Module -4:	**Module -4: Data Visualization and Advanced Tools for Data Preparation** **4.1** Visualization techniques, Different types of plots, Designing visualizations, Time series, **4.2** Geolocated data, Correlations and connections, **4.3** Hierarchies and networks, Interactivity. Web scraping, **4.4** Data from social networks, Open- source tools for data preparation: Open Refine, R/Python libraries for data preparation and visualization.	15

Reference Books:-

1) Bateman, S., & D'Ignazio, C., Data Feminism (Cambridge, MA: MIT Press), (2020).

2) Wickham, H., & Grolemund, G., R for Data Science: Import, Tidy, Transform, Visualize, and Model Data (Sebastopol, CA: O'Reilly Media), (2016).

<table>
<tr><td colspan="2">3) Bramer, M., Principles of Data Mining (London, UK: Springer), (2016).
4) Swayne, D. F., Wickham, H., & Cook, D., Interactive and Dynamic Graphics for Data Analysis: WithR and GGobi (New York: Springer), (2015).
5) Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies (Cambridge, MA: MIT Press), (2015).</td></tr>
<tr><td colspan="2">Evaluation Pattern:</td></tr>
<tr><td colspan="2">Total Marks: 60</td></tr>
<tr><td>Internal Continuous Evaluation:
• CCE-I – 10 Marks
• CCE-II – 10 Marks
• Mid Semester – 20 Marks
• Activity – 10 Marks</td><td>End Semester Examination:
• Question -1 (02 Marks = 2 X 6=12 Marks)
• Question -2 (06 Marks = 6 X 2=12 Marks)
• Question -3 (06 Marks = 6 X 2=12 Marks)
• Question -4 (06 Marks = 6 X 2=12 Marks)
• Question -5 (04 Marks = 4 X 3=12 Marks)
• Question -6 (04 Marks = 4 X 3=12 Marks)
• Question -7 (04 Marks = 4 X 3=12 Marks)</td></tr>
</table>

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - II
Type: E I	**Marks:** 30
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 424 E I: AI for Data Science

Course Objectives:

1) Understand the various aspects in intelligent agents.
2) Understand problem solving methods.
3) Study the concept Knowledge, reasoning and planning.
4) Study the applications.

Course Outcomes:

1) Design intelligent agent.
2) Handle problem solving methods.
3) Manage Knowledge, reasoning and planning.
4) Apply the Knowledge in Applications.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Intelligent Agents Concepts** **1.1** Introduction: What is AI, Foundations History of Artificial Intelligence, **1.2** The State-of-the-Art Intelligent Agents: Agents and Environments, **1.3** Good Behavior: The Concept of Rationality, **1.4** The Nature of Environments, and The Structure of agents.	15
Module -2:	**Module -2: Problem-solving** **2.1** Solving Problems by Searching: Problem-Solving Agents, **2.2** Uninformed Search Strategies, Informed (Heuristic) Search Strategies, Heuristic Functions, **2.3** Beyond Classical Search Local Search Algorithms and Optimization Problems, Local Search in Continuous Spaces, **2.4** Searching with Nondeterministic Actions, Searching with Partial Observations, Online Search Agents and Unknown Environment.	15
Module -3:	**Module -3: Knowledge, reasoning, and planning** **3.1** Knowledge based Agents, First-Order Logic and Its Inference, **3.2** Classical Planning, Planning and Acting in the Real World, **3.3** Knowledge Representation, Quantifying Uncertainty, **3.4** Probabilistic Reasoning, Probabilistic Reasoning over Time, Making Simple Decisions, Making Complex Decisions.	15
Module -4:	**Module -4: Learning and Applications with case studies** **4.1** Learning from Examples, Knowledge in Learning, Learning Probabilistic Models, Reinforcement Learning, **4.2** AI Applications in various fields in marketing, **4.3** healthcare, banking, finance, etc. Case Studies: Credit card Fraud Analysis, Sentiment Analysis, **4.4** Recommendation Systems and Collaborative filtering, Uber Alternative Routing.	15

Reference Books:-

1) Chollet, F., Deep Learning with Python (Shelter Island, NY: Manning Publications), (2017).

2) Russell, S. J., & Norvig, P., Artificial Intelligence: A Modern Approach (New York: Pearson), (2016).

3) Goodfellow, I., Bengio, Y., & Courville, A., Deep Learning (Cambridge, MA: MIT Press), (2016).

4) Marsland, S., Machine Learning: An Algorithmic Perspective (Boca Raton, FL: CRC Press), (2014).

5) Luger G.F. and Stubblefield W.A., Artificial Intelligence: Structures and strategies for Comple roblem Solving. Addison Wesley, 6th edition, (2008).

Evaluation Pattern:

Total Marks: 50

Internal Continuous Evaluation (20 Marks):	**End Semester Examination (30 Marks):**
• CCE - I : 10 Marks: Objective • CCE - II: 10 Marks: Objective • Mid Semester Exam: 20 Marks: Subjective (20 Marks converted to 10 marks) • Activity: 10 marks	• Question -1: Solve the following questions (5 questions of 2 Marks) • Attempt any two questions from Q. 2 to Q. 4 (three questions of 10 marks)

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - II
Type: E II	**Marks:** 30
Credits: 4	**From:** A. Y. 2025-26

Name of the Course: MDST 424 E II: IOT for Data Science II

Course Objectives:

1) Gain knowledge on basic functioning of Arduino with sensors.
2) Familiarize the concepts of Arduino programming with sensors.
3) Understand basic of Raspberry Pi and to acquire the knowledge of Interfacing with Raspberry Pi.
4) Familiarize the applications of IOT.

Course Outcomes:

1) Explain the Arduino with sensors.
2) Describe concepts of Arduino with programming.
3) Explain the raspberry pi.
4) Discuss interfacing of raspberry pi.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Arduino in IOT** **1.1** Characteristics of IoT, Physical design of IoT, **1.2** Difference between IoT and M2M, **1.3** Interoperability in IoT, Introduction to Arduino Programming, **1.4** Interfacing of Temperature, Humidity, Motion, Light and Gas Sensor with Arduino.	15
Module -2:	**Module -2: Introduction of Raspberry Pi** **2.1** Introduction to Python programming, **2.2** Introduction to Raspberry Pi, **2.3** Interfacing Raspberry Pi with basic peripherals, **2.4** Implementation of IoT with Raspberry Pi.	15
Module -3:	**Module -3: Implementation of IOT** **3.1** Implementation of IoT with Raspberry Pi, **3.2** Introduction to Software defined Network (SDN), **3.3** SDN for IoT, **3.4** Data Handling and Analytics	15
Module -4:	**Module -4: Applications of IOT** **4.1** Cloud Computing, **4.2** Cloud services - SaaS, PaaS, IaaS, Sensor-Cloud, **4.3** Smart Cities and Smart Homes, Connected Vehicles, **4.4** Smart Grid, Industrial IoT Case Study: Agriculture, Healthcare, Activity Monitoring.	15

Reference Books:-

1) Antonopoulos, A., & Gillam, L., Internet of Things: Principles and Paradigms (Boca Raton, FL: CRC Press), (2017).

2) Patel, V. M., & Zhang, S, Internet of Things: A Hands-On Approach (Boca Raton, FL: CRC Press).

3) Ray, P. P. (2018). Internet of Things: A Complete Introduction (New York: Springer), (2017).

4) Banerjee, A., & Dasgupta, S., Internet of Things: Challenges and Opportunities (New York: Springer), (2015).

5) Kortuem, G., Kawsar, F., Sundramoorthy, V., & Fitton, D., Smart Objects, and the Internet of Things: AHandbook for IoT Developers (San Francisco, CA: Morgan & Claypool Publishers), (2010).

Evaluation Pattern:

Total Marks: 60

Internal Continuous Evaluation (20 Marks):	**End Semester Examination (30 Marks):**
• CCE - I : 10 Marks: Objective • CCE - II: 10 Marks: Objective • Mid Semester Exam: 20 Marks: Subjective (20 Marks converted to 10 marks) • Activity: 10 marks	• Question -1: Solve the following questions (5 questions of 2 Marks) • Attempt any two questions from Q. 2 to Q. 4 (three questions of 10 marks)

	Board of Studies in Computer Science	
	Programme: M.Sc.	**Semester** - II
	Type: Major	**Marks:** 60
	Credits: 4	**From:** A. Y. 2025-26
	Name of the Course: MDST- 425: Research Project	

Course Objectives:

1) Apply data science concepts and techniques to address a specific research problem.
2) Conduct a thorough literature review and critically analyze existing research in the field.
3) Design and implement a research plan, including data collection and preprocessing.
4) Interpret and evaluate the results of the research project.

Course Outcomes:

1) Demonstrate a deep understanding of research methodologies in data science and their application to real-world problems.
2) Formulate clear and concise research questions and objectives.
3) Conduct a comprehensive literature review and critically analyze existing research in the field.
4) Design and execute a research plan, including data collection, preprocessing, and analysis.

Module	Title and Contents	Hrs
Module -1:	**Module -1: Fundamental Research in Data Science** **1.1** Introduction to research methodologies in data science, **1.2** Identifying research problems and formulating research questions, **1.3** Reviewing and critiquing existing literature, **1.4** Ethical considerations in data science research.	15
Module -2:	**Module -2: Research Planning and Design** **2.1** Developing a research plan and timeline, **2.2** Defining research objectives and hypotheses, **2.3** Selecting appropriate research methods, **2.4** Data collection strategies and techniques.	15
Module -3:	**Module -3: Data Collection and Preprocessing, Data Analysis and Modeling** **3.1** Data acquisition and cleaning, **3.2** Exploratory data analysis Feature selection and dimensionality reduction, **3.3** Handling missing data and outliers ,Supervised and unsupervised learning techniques Regression analysis, **3.4** Classification and clustering algorithms Evaluation metrics for model performance.	15
Module -4:	**Module -4: Project Management and Professional Skills** **4.1** Time management and project organization, **4.2** Collaboration and teamwork, **4.3** Ethical considerations in data science research Intellectual property and plagiarism.	15

Reference Books:-

1) Donald Cooper & Pamela Schindler, Business Research Methods, TMGH, 9th edition.
2) Alan Bryman & Emma Bell, Business Research Methods, Oxford University Press.
3) Kothari C.R. Research Methodology.
4) Research Methodology Dr. Shalu Katyal, 1 April 2024

5) Research Design: Qualitative, Quantitative ,J. David Creswell ,22 November 2022.	
Evaluation Pattern:	
Total Marks: 60	
Internal Continuous Evaluation: • CCE-I – 10 Marks • CCE-II – 10 Marks • Mid Semester – 20 Marks • Activity – 10 Marks	**End Semester Examination:** • Question -1 (02 Marks = 2 X 6=12 Marks) • Question -2 (06 Marks = 6 X 2=12 Marks) • Question -3 (06 Marks = 6 X 2=12 Marks) • Question -4 (06 Marks = 6 X 2=12 Marks) • Question -5 (04 Marks = 4 X 3=12 Marks) • Question -6 (04 Marks = 4 X 3=12 Marks) • Question -7 (04 Marks = 4 X 3=12 Marks)

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - II
Type: Major	**Marks:** 50
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDSP 426: Lab Course I MDST 421, MDST 422, MDST 423

Course Objectives:

1) Understand the various aspects in district mathematics in data science.
2) Understand Data Analysis and probability theory. Study the concept of Linear Algebra and Calculus.
3) Understand how to use Regression model and Understand python programming.
4) To understand the importance of data and data preprocessing, data cleaning.

Course Outcomes:

1) Apply principles python programming.
2) Implement clear and effective python code.
3) Design discrete mathematics for any real-world application. Handle data analysis and probability theory.
4) Design Data Gathering and Data Discovery and Handle Cleaning and Conditioning Data.

Module	Title and Contents	Hrs
Module -1:	**Module -1: MDSP 416: Lab I Course MDST414E-I** **1.** Sets and Cardinality **2.** Set Operations **3.** Graph Traversal **4.** Matrix and Vector Algebra **5.** Systems of Linear Equations **6.** Multivariate Calculus **7.** Practical on Strings and Lists **8.** Practical on Packages **9.** Practical on Stacks **10.** Practical on Queues **11.** Practical on Tuple **12.** Practical on File Handling **13.** Import a CSV file into Python using Pandas. **14.** Identify and handle missing data **15.** Identify missing values in a dataset. **16.** Identify outliers in a dataset using statistical methods. **17.** Combine data from multiple sources **18.** Apply Principal Component Analysis (PCA) to reduce dimensionality. **19.** Summarize key data insights **20.** Calculate descriptive statistics (mean, median, standard deviation) for key variables.	8

Reference Books:-

1) Bateman, S., & D'Ignazio, C., Data Feminism (Cambridge, MA: MIT Press), (2020).
2) Wickham, H., & Grolemund, G., R for Data Science: Import, Tidy, Transform, Visualize, and ModelData (Sebastopol, CA: O'Reilly Media), (2016).
3) Bramer, M., Principles of Data Mining (London, UK: Springer), (2016).
4) McKinney, W., Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (Sebastopol, CA: O'Reilly Media), (2017).
5) Dusty Phillips, "Python 3 Object-oriented Programming Second Edition", Packt Publishing, Greg, 2015.

Evaluation Pattern:	
Total Marks: 50	
	End Semester Examination: • Question -1 (10 Marks) • Question -2 (10 Marks) • Question -3 (10 Marks) • Activity 20 Marks.

Karmaveer Bhaurao Patil University, Satara

(A State Public University Est. u/s 3(6) of MPUA 2016)

Faculty of Science and Technology

Yashavantrao Chavan Institute of Science, Satara

Board of Studies in Computer Science

Programme: M.Sc.	**Semester** - II
Type: Major	**Marks:** 50
Credits: 2	**From:** A. Y. 2025-26

Name of the Course: MDSP 427: Lab Course II MDST 424 E1

Course Objectives:

1) Understand the various aspects in intelligent agents.
2) Understand problem solving methods.
3) Study the concept Knowledge, reasoning and planning
4) Study the applications.

Course Outcomes:

1) Design intelligent agent.
2) Handle problem solving methods.
3) Manage Knowledge, reasoning and planning.
4) Apply the Knowledge in Applications.

Module	Title and Contents	Hrs
Module -1:	**Module -1: MDSP 427: Lab Course II MDST 424 E1** **1.** Practical on exploration of Python libraries such as NumPy, Pandas, Scikit- learn, and Tensor Flow **2.** Practical on implement a basic reflex agent that performs simple tasks in a predefined environment. **3.** Practical on Design an agent and evaluate its behavior in different scenarios to understand rationality **4.** Practical on Implement BFS, DFS, and UCS on a maze or graph and compare their performance **5.** Practical on Heuristic Search: A and Greedy Search* **6.** Create and compare different heuristic functions for a problem like the traveling salesman **7.** Implement and compare local search algorithms in solving optimization problems. **8.** Practical on Develop an agent that learns to navigate an unknown environment using online search strategies. **9.** Practical on Create a simple knowledge base using propositional logic and inference rules. **10.** Practical on Implement inference mechanisms in first-order logic for a simple domain like family relations. **11.** Practical on Use a planning tool (like PDDL) to design and execute a plan for a robot in a blocks world. **12.** Practical on Build a Bayesian Network to model and reason about uncertainty in a medical diagnosis problem. **13.** Practical on Implement a Hidden Markov Model (HMM) to track an object moving in a grid.	8

	14. Practical on Solve a grid-based MDP problem using dynamic programming techniques. **15.** Practical on Use a dataset to train and test a decision tree model, evaluate its performance. **16.** Practical on Apply K-means clustering on a dataset to identify natural groupings. **17.** Practical on Implement Q-learning to solve a simple grid-based navigation task. **18.** Case Study: Credit Card Fraud Detection (Develop and evaluate a machine learning model to detect fraudulent credit card transactions.) **19.** Case Study: Sentiment Analysis (Implement a sentiment analysis model on a dataset of customer reviews.) **20.** Case Study: Recommendation Systems (Create a collaborative filtering-based recommendation system for a movie dataset.).	

Reference Books:-

1) Chollet, F., Deep Learning with Python (Shelter Island, NY: Manning Publications), (2017).

2) Russell, S. J., & Norvig, P., Artificial Intelligence: A Modern Approach (New York: Pearson), (2016).

3) Goodfellow, I., Bengio, Y., & Courville, A., Deep Learning (Cambridge, MA: MIT Press), (2016).

4) Marsland, S., Machine Learning: An Algorithmic Perspective (Boca Raton, FL: CRC Press), (2014).

5) Bishop, C. M., Pattern Recognition and Machine Learning (New York: Springer), (2006).

Evaluation Pattern:	
Total Marks: 50	
	End Semester Examination: • Question -1 (10 Marks) • Question -2 (10 Marks) • Question -3 (10 Marks) • Activity 20 Marks.