SHIVAJI UNIVERSITY, KOLHAPUR - 416004,
MAHARASHTRA
PHONE:EPABX-2609000, www.unishivaji.ac.in, bos@unishivaji.ac.in
शिवाजी विद्यापीठ, कोल्हापूर -४१६००४,महाराष्ट्र
दूरध्वनी-ईपीएबीएक्स -२६०९०००, अभ्यासमंडळे विभाग दुरध्वनी ०२३१—२६०९०९४
०२३१—२६०९४८७

Estd. 1962
"A++" Accredited by
NAAC (2021)
With CGPA 3.52

Ref.No.SU/BOS/Science/147 **Date: 26/05/2026**

The Principal,
All Concerned Affiliated Colleges/Institutions
Shivaji University, Kolhapur

The Head/Co-ordinator/Director
All Concerned Department (Science)
Shivaji University, Kolhapur.

Subject: Regarding revised syllabi of M.Sc. Part-II (Sem III & IV) degree programme under the Faculty of Science and Technology as per NEP-2020 (2.0)

Sir/Madam,

With reference to the subject mentioned above, I am directed to inform you that the university authorities have accepted and granted approval to the syllabi, nature of question paper of M.Sc. Part-II (Sem III & IV) degree programme under the Faculty of Science and Technology as per NEP-2020 (2.0).

M.Sc. Part-II (Sem. III & IV) as per NEP-2020 (2.0)			
1.	Chemistry	7.	Data Science
2.	Mathematics (Online Mode)	8.	Computer Science

This syllabus, nature of question and equivalence shall be implemented from the academic year 2026-2027 onwards. A soft copy containing the syllabus is attached herewith and it is also available on university website **www.unishivaji.ac.in>Syllabus>Syllabus as per NEP2020.**

The question papers on the pre-revised syllabi of above-mentioned course will be set for the examinations to be held in October /November 2026 & March/April 2027. These chances are available for repeater students, if any.

You are, therefore, requested to bring this to the notice of all students and teachers concerned.

Thanking you,

Yours faithfully,

Dy Registrar

Encl: as above

for information and necessary action

Copy to:

1	Dean, Faculty of Science & Technology	6	Appointment Section A & B
2	Director, Board of Examinations and Evaluation	7	I.T.Cell /Computer Centre
3	Chairman, Respective Board of Studies	8	Eligibility Section
4	B.Sc.-M.Sc. Exam Section	9	Affiliation Section (T.1) (T.2)
5	Internal Quality Assurance Cell (IQAC Cell)	10	P.G. Seminar Section
		11	Distance and Online Education

SHIVAJI UNIVERSITY, KOLHAPUR

Established: 1962
A^{++} Accredited by NAAC (2021) with CGPA 3.52

Structure and Syllabus in Accordance with

National Education Policy - 2020

with Multiple Entry and Multiple Exit

Master of Science
(Data Science)
Part II -
(Semester III & IV)

under

Faculty of Science and Technology

(To Be Implemented from Academic Year 2026-27)

Structure in Accordance with National Education Policy - 2020
With Multiple Entry and Multiple Exit Options
M.Sc. (Data Science) Part – II (Level-6.5)

	Course Code	Teaching Scheme			Examination Scheme					
		Theory and Practical			University Assessment (UA)			Internal Assessment (IA)		
		Lectures + Tutorial (Per week)	Hours (Per week)	Credit	Maximum Marks	Minimum Marks	Exam. Hours	Maximum Marks	Minimum Marks	Exam. Hours
Semester-III										
Major Mandatory	MMT-301	4	--	4	60	24	2	40	16	
	MMT-302	4	--	4	60	24	2	40	16	
	MMPR -303	--	8	4	60	24	2	40	16	
	MMT-304	2	--	2	30	12	2	20	8	
Major Elective	MET-305	4	--	4	60	24	2	40	16	
	MET-306									
Research Project	RP-307	--	--	4	60	24	--	40	16	
Total				22	330			220		
Semester-IV										
Major Mandatory	MMT-401	4	--	4	60	24	2	40	16	
	MMT-402	4	--	4	60	24	2	40	16	
	MMPR-403	--	8	4	60	24	2	40	16	
Major Elective	MET-404	4	--	4	60	24	2	40	16	
	MET-405									
Research Project	RP-406	--	--	6	100	40	--	50	20	
Total				22	340			210		
Total (Sem III + Sem IV)				44						

<table>
<tr><td>• MMT–Major Mandatory Theory
• MMPR–Major Mandatory Practical
• MET–Major Elective Theory</td><td>• MEPR–Major Elective Practical
• RP- Research Project</td><td>• Total Marks for M.Sc.-II : 1100</td></tr>
<tr><td colspan="2"></td><td>• Total Credits for M.Sc.-II (Semester III & IV) : 44</td></tr>
<tr><td colspan="2"></td><td>• Separate passing is mandatory for University and Internal Examinations</td></tr>
<tr><td colspan="3"># Evaluation scheme for Research Project shall be decided by concerned BOS</td></tr>
<tr><td colspan="3">• Requirement for Exit after Level 6.5: Students can exit after completion of Level 6.5 with Master of Data Science</td></tr>
</table>

Course Codes

M.Sc. Semester-III		
Course Code	**Major Mandatory**	
MMT – 301	R Programming (4 Credits)	**MSU0325MML911I1**
MMT -302	Data Analytics with Power BI (4 credits)	**MSU0325MML911I2**
MMPR- 303	Practical Lab -III (4 credits)	**MSU0325MMP911I1**
MMT - 304	Generative AI (2 credits)	**MSU0325MML911I3**
RP - 307	Research Project (4 credits)	**MSU0325RP911I**
Major Elective		
MET – 305	Data Engineering (4 credits)	**MSU0325MEL911I1**
MET - 306	Cloud Computing (4 credits)	**MSU0325MEL911I2**
M. Sc. Semester-IV		
Course Code	**Major Mandatory**	
MMT – 401	Natural Language Processing (4 credits)	**MSU0325MML911J1**
MMT – 402	Data Visualization using Tableau (4 credits)	**MSU0325MML911J2**
MMPR – 403	Practical Lab -IV (4 credits)	**MSU0325MMP911J1**
RP- 406	Research Project (6 credits)	**MSU0325RP911J**
Major Elective		
MET – 404	Deep Learning (4 credits)	**MSU0325MEL911J1**
MET - 405	Social Media Analytics (4 Credits)	**MSU0325MEL911J2**

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester III)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: R Programming
Course Code: MMT -301
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:
1. Learn Fundamentals of R.
2. Apply OOP concepts in R programming
3. Covers how to use different functions in R, how to read data into R, accessing R packages, writing R functions, debugging, and organizing data using R functions.
4. Apply various concepts to write programs in R
5. Cover the Basics of statistical data analysis with examples.

Unit I: Introduction to R (15 Hours)

Introduction to R: What is R, Why R, Advantages of R over Other Programming Languages. R Studio: R command Prompt, R script file, comments – Handling Packages in R: Installing a R Package, few commands to get started: installed. packages (), package Description (), help (), find. Package (), library () - Input and Output – Entering Data from keyboard – Printing fewer digits or more digits – Special Values functions: NA, Inf and –inf.

Unit II: R Data Types: (15 Hours)

Vectors, Lists, Matrices, Arrays, Factors, Data Frame. R - Variables: Variable assignment, Data types of Variable, Finding Variable ls(), Deleting Variables ,R Operators: Arithmetic Operators, Relational Operators, Logical Operator, Assignment Operators, Miscellaneous Operators, R Decision Making: if statement, if – else statement, if – else if statement, switch statement, R Loops: repeat loop, while loop, for loop - Loop control statement: break statement, next statement.

Unit III: R-Function : (15 Hours)

function definition, Built in functions: mean(), paste(), sum(), min(), max(), seq(), user-defined function, calling a function, calling a function without an argument, calling a function with argument values, R-Strings : Manipulating Text in Data: substr(), strsplit(), paste(), grep(), toupper(), tolower(), R Vectors : Sequence vector, rep function, vector access, vector names, vector math, vector recycling, vector element sorting R List : Creating a List, List Tags and Values, Add/Delete Element to or from a List, Size of List, Merging Lists, Converting List to Vector. R Matrices: Accessing Elements of a Matrix. R Arrays: Naming Columns and Rows, Accessing Array Elements, Manipulating Array Elements.

Unit IV: R-Data Frames & Data Visualization (15 Hours)

R-Data Frames: Create Data Frame, Data Frame Access, Understanding Data in Data Frames: dim(), nrow(), ncol(), str(), Summary(), names(), head(), tail(), edit() Expand Data Frame: Add Column, Add Row - Joining columns and rows in a Data frame rbind() and cbind() , Merging Data frames merge(), Melting and Casting data melt(), cast(). Loading and handling Data in R: Getting and Setting the Working Directory – getwd(), setwd(), dir() - R-CSV Files - Input as a CSV file, Reading a CSV File, Analyzing the CSV File: summary(), min(), max(), range(), mean(), median(), apply().
Data Visualization - Pie Charts: Pie Chart title and Colors – Slice Percentages and Chart Legend, 3D Pie Chart – R Histograms – Density Plot - R – Bar Charts: Bar Chart Labels, Title and Colors.

Recommended Books:

1. Sandip Rakshit, R Programming for Beginners, McGraw Hill Education (India), 2017.

Reference Books:

1. Seema Acharya, Data Analytics using R, McGrawHill Education (India), 2018.
2. Andrie de Vries, Joris Meys, R for Dummies A Wiley Brand, 2nd Edition, John Wiley and Sons, Inc, 2015.

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester III)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Data Analytics with Power BI
Course Code: MMT- 302
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:
1. Identify the primary components of Power BI.
2. Identify common challenges in Power BI data models.
3. Implement smart solutions.
4. Understand Power Query.

Unit-I: Introduction to Power BI (15 Hours)
Introduction to Data Analytics, Concepts of Business Intelligence (BI), Introduction to Microsoft Power BI ecosystem, History of Power BI, Components: Power BI Desktop, Power BI Service, Power BI Mobile, Power BI Gateway, Installing Power BI Desktop, BI vs Data Science: How Power BI fits the data science pipeline, Understanding the Power BI interface, Connecting to various data sources (CSV, Excel, SQL, Web API), Understanding field, tables, and data types, Renaming, duplicating, and formatting tables/columns,

Unit-II: Data Preparation & Power Query (ETL) (15 Hours)
Data Transformation: Benefits of Data Transformation, Introduction to Power Query Editor, Basic Transformation: Removing rows/columns, filtering data, Add column, custom column, index column, conditional column, pivoting and unpivoting columns, duplicating vs referencing queries, Merge query, append query, join kind, fact dimension model, Grouping and summarizing data ,

Unit-III: Data Modeling and DAX (15 Hours)
Data Modeling: Creating relationship between tables, One-to-many, many-to-one relationships, star schema vs Snowflake schema, managing relationships view, DAX: Understanding calculated columns vs measures, aggregation functions: SUM, AVERAGE, COUNT, Logical functions: IF, SWITCH, Text functions: CONCATENATE, LEFT, RIGHT, Time intelligence: TOTALYTD, DATESYTD, SAMEPERIODLASTYEAR, Filter functions: CALCULATE, FILTER, ALL, VALUES, Using VAR and RETURN, DAX Context: Row vs Filter context, Creating KPIs and custom metrics

Unit-IV: Data Visualization using Power BI (15 Hours)
Visual types: Bar, column, line, area, pie, donut, table, matrix, card, map, gauge, KPI; Slicers and filters, Visual interactions, Drill-down and drill-through, Formatting visuals: colors, labels, legends, Tooltips customization, Bookmarks and buttons (for navigation).
Seminars, Tutorials, Problem solving session and group discussions on above four units.

Recommended Books:
1. Mastering Microsoft Power BI – Brett Powell

Reference Books:
1. Beginning Power BI by Dan Clark
2. Microsoft Power BI Cookbook – Brett Powell

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester III)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Practical Lab- III
Course Code: MMPR - 303
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. Demonstrate use of basic functions.
2. Create their own customized functions.
3. Learn to understand new data sets and functions by yourself.
4. To construct model, relationships between the models, user friendly models.
5. Work on built-in real-time cases for analysis and visualization.

Lab assignments based on Course Code **MMT -301 (R Programming)** and **MMT-302 (Data Analytics with Power BI)**.

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester III)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Generative AI
Course Code: MMT- 304
Total Credits: 02
Course Outcomes (CO): Upon successful completion of this course, the student will be able to:

1. Students will be able to **explain basic AI concepts** and understand the differences between AI, Machine Learning, and Deep Learning, along with common learning methods and algorithms.
2. Students will be able to **apply fundamental ML and DL models** such as regression, decision trees, neural networks, CNNs, and RNNs to solve simple real-world problems.
3. Students will be able to **understand and analyze generative AI techniques** including VAEs, GANs, diffusion models, transformers, and large language models.
4. Students will be able to **use modern generative AI tools responsibly**, including fine-tuning LLMs, applying RAG, deploying AI systems, and understanding ethical and responsible AI practices.

UNIT I: Introduction to Generative AI **(15 Hours)**

Introduction to Artificial Intelligence (AI), Difference between AI, ML, and DL, Types of Learning: Supervised, Unsupervised, Reinforcement, Linear Regression, Logistic Regression, Decision Tree, K-Means, Limitations of Traditional ML, Introduction to Deep Learning, Neural Networks, CNN Basics, RNN Basics, Representation Learning, Discriminative vs Generative Models.

UNIT II: Introduction to Generative Models **(15 Hours)**

Introduction to Generative Models - Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Diffusion Models, Transformers, Why Transformers Replaced RNNs, Large Language Models (LLMs), Applied Generative AI, Fine-Tuning Large Language Models, Retrieval-Augmented Generation (RAG), Deployment of Generative AI Systems, Ethics & Responsible AI

Recommended Book:

1. generative deep learning — david foster.
2. Foundations of Machine Learning — Mehryar Mohri, Afshin Rostamizadeh&Ameet Talwalkar.

Reference Books:

1. Neural Networks and Deep Learning — Michael Nielsen (online book).
2. Deep Learning — Ian Goodfellow, Yoshua Bengio& Aaron Courville.

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester III)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Data Engineering
Code: MET-305
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. To introduce students to data storage systems and technologies commonly used in data engineering.
2. To enable students to design and implement databases for efficient data storage and retrieval.
3. To teach student show to optimize data storage and access patterns for performance.
4. To explore data security and privacy considerations in data engineering.
5. To provide experience in using cloud-based storage and database services.

UNIT-I Introduction to Data Engineering (15 Hours)

Introduction to Data Engineering: Data engineering roles and responsibilities, Data pipelines and their importance, Types of data: structured, semi-structured, unstructured. **Data Flow Concepts:** End-to-end data flow lifecycle, Data ingestion → processing → storage → consumption, Data lineage and, data, tracking, **Data Storage Fundamentals:** Relational databases (SQL): schema, ACID properties, NoSQL databases: key-value, document, columnar, graph, **Storage file formats:** CSV, JSON, Parquet, Avro

UNIT-II - Data Retrieval and Querying (15 Hours)

Data Retrieval and Querying: SQL querying, indexing, optimization basics,NoSQL querying patterns ,**Data Transformation:** Cleaning, filtering, data enrichment, Batch vs streaming data transformation, ETL vs ELT concepts ,**Processing Frameworks:** Apache Spark: RDDs, DataFrames, Spark SQL, transformations vs actions,Overview of Apache Flink

UNIT III - Data Integration & ETL (15 Hours)

Data Integration & ETL: Extract, Transform, Load pipelines,Ingestion tools (Kafka Connect, Sqoop, Flume),Data validation and quality checks,Data Warehousing: OLTP vs OLAP, Dimensional modeling (Star, Snowflake schemas),Fact and dimension tables, Data marts,**Big Data Technologies:** Hadoop ecosystem overview,HDFS, MapReduce, YARN,Hive, Pig,**Apache Kafka:** producers, consumers, topics, partitions,Real-time data streams

UNIT IV - Introduction to Cloud Computing (15 Hours)

Introduction to Cloud Computing: Benefits of cloud data platforms, IaaS, PaaS, SaaS models, Cloud storage types: object, block, file systems, **Cloud Data Engineering Platforms:** AWS: S3, Redshift, Glue, EMR, Kinesis, Azure: Azure Data Factory, Databricks, ADLS, Synapse Analytics, **GCP**: Big Query, Dataflow, Data proc, Pub/Sub, **Cloud-Based Data Processing:** Serverless data processing, Data pipeline orchestration on cloud, Cost optimization & scalability, Considerations

Reference Books:

1."DataEngineering"byPramodJ. SadalageandMartinFowler
2."DataEngineeringCookbook"byAndreasKretz.

Title of Course: Cloud Computing
Course Code: MET - 306
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. Understand the fundamentals of cloud computing, including its characteristics, service and deployment models, cost benefits, and basic cloud architecture.
2. Gain knowledge of virtualization technologies, different types of virtualizations, virtual machine management, and their advantages and disadvantages.
3. Learn about cloud storage, virtualized data centers, distributed file systems like GFS and HDFS, federated cloud architecture, SLAs, and cloud service life cycle.
4. Develop awareness of cloud security concepts, cloud risks, data protection methods, identity management, IoT technologies, cloud trends, and practical applications in real-world scenarios.

UNIT I **(15 Hours)**

Definition of cloud computing, Characteristics of cloud computing, public cloud, Private cloud, Hybrid cloud, Community cloud, IaaS, PaaS, SaaS, Cloud cost benefits, Cloud architecture components, Introduction to virtualization, Virtualization reference model, Characteristics of virtualized environment

UNIT II **(15 Hours)**

Storage virtualization, Network virtualization, Desktop virtualization, Application server virtualization, VMware full virtualization model, Xen architecture and guest OS management, Virtual machine life cycle, VM migration concepts, VM consolidation concepts, VM management concepts, Advantages and disadvantages of virtualization

UNIT III **(15 Hours)**

Cloud storage system architecture, Virtualized data center architecture, Server storage and network virtualization techniques, Google File System components and features, Hadoop Distributed File System features, Federated cloud architecture, SLA basics, Types of SLA, SLA life cycle, Cloud service life cycle

UNIT IV **(15 Hours)**

Cloud security concepts (multi-tenancy, virtualization, data outsourcing, trust management), Types of cloud risks, Data security technologies, Identity and access management, Content-level security, Cloud trends in ubiquitous computing, IoT enabling technologies, IoT applications

Recommended Books:

1. Mastering Cloud Computing — Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi
2. Cloud Computing: A Hands-On Approach — Arshdeep Bahga, Vijay Madisetti
3. Cloud Computing – Principles and Paradigms — Rajkumar Buyya, James Broberg, Andrzej Goscinski

References Books:

1.Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance — Tim Mather, Subra Kumaraswamy, Shahed Latif
2.Cloud Computing Black Book — Jayaswal, Kallakurchi, Houde, Shah
3.Distributed and Cloud Computing — Kai Hwang, Geoffrey C. Fox, Jack J. Dongarra

Title of Course: Research Project
Course Code: RP-307
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. Find current research domains in computer science
2. Identify different research journals in computer science domains
3. Understand citations, impact factors, references etc.
4. Identification of appropriate societal issues.
5. Development of applications to address identified societal issue.

- Student has to identify research problem in the semester-III and have to carry out thorough literature review / student can develop an application project which will address societal issues.

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester IV)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Natural Language Processing
Course Code: MMT - 401
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. To learn the fundamentals of natural language processing.
2. To Understand approaches to syntax and semantics in NLP.
3. To understand the Large language Model and Transformers in NLP and Deep Learning
4. To understand practical approach of NLP in python

Unit I **(15 Hours)**

Introduction to NLP: Introduction, NLP, NLU, NLG, History of NLP, Stages of NLP, Working of Natural Language Processing, Technologies related to NLP: ML, NLTK, Parsers, Sentiment Analysis, Application of NLP: Speech of Text (STT), Text to Speech (TTS), Machine Translation, Basic Text Processing: Tokenization, Lemmatization, Stemming, Morphology, Language Modelling.

Unit II **(15 Hours)**

Word Level Analysis: Regular Expressions, Finite Automata, Part of Speech tagging, rule based Part of Speech Tagging, Markov Model, Hidden Markov Models. Syntax Parsing: Context Free Languages, Context Free Parsing, Top-down and Bottom -up Parsing, CYK Parsing. Semantic Analysis: Meaning, Supervised and Unsupervised Approaches, WordNet, Word Sense Disambiguation, Coreferences Resolution.

Unit III **(15 Hours)**

NLP with Deep Learning: Introduction and Word Vectors, Language Models, Back Propagation and Neural Networks, Dependency Parsing, Recurrent Neural Networks, Sequence2Sequence Models, LLM Intro, Self Attention and Transformers, Pretraining, Natural Language Generation, Post Training, Benchmarking, Reasoning and Agents.

Unit IV **(15 Hours)**

NLP in Python: Installation and pre-requisites, Regex in chatbot, information extraction, Text Classification using Naïve Bayes Classifier, Logistic Regression, Support Vector Machine, CNN, LSTM. Text Summarization: Example of Text Summarizer, Machine Translation: Using a Machine Translation Example, Question-Answering System, Chatbot.

Seminars, Tutorials, Problem solving session and group discussions on above four units.

Reference Books:

1. Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP System. Sowmya Vajjala, Bodhisattwa Majumdar, Anuj Gupta & Harshit Surana O'REILLY
2. Natural Language Processing and Information Retrieval Tanveer Siddiqui, U. S. Tiwary- Oxford University Press
3. Speech and Language Processing, An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, Daniel Jurafsky, James H. Martin- Pearson Education, 2023.
4. Akshay Kulkarni, Adarsha Shivananda, "Natural Language Processing Recipes - Unlocking Text Data with Machine Learning and Deep Learning using Python", Apress, 2019.
5. T V Geetha, "Understanding Natural Language Processing – Machine Learning and Deep Learning Perspectives", Pearson, 2024.

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester IV)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Data Visualization using Tableau
Course Code: MMT - 402
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:
1. Understand Tableau fundamentals, architecture, components, and file/data types.
2. Connect and visualize data using joins, blending, charts, filters, and calculations.
3. Use advanced features like groups, bins, hierarchies, sets, and parameters for data analysis.
4. Create interactive dashboards, stories, and annotated visuals for effective data communication.

Unit-I: (15 Hours)
Tableau Introduction: History of Tableau, Characteristics of Tableau, Tableau versions., Architecture of Tableau; Tableau Components: Tableau Desktop, Server, Tableau Reader, Tableau Online, Tableau Public. Tableau Design flow, File Extensions, Data types.

Unit-II: (15 Hours)
Data Connection in Tableau Interface, Data Joins in Tableau: Types of Joins, Data Blending in Tableau: Types of Data Blending, Rules to perform the Joins in tableau, Join vs. Data Blending. Graphs/Charts/Bars: Line Graphs, Single Line graphs, Dual Line Graphs, Bar charts, Stacked Bar charts, Histogram, Pie Charts, Donut charts, Funnel Charts, Scatter plot; Filters: Extract, Data Source, Context, Dimension, Measure, Quick Filters; Calculations: Need for calculations, types of calculations, Custom Charts.

Unit-III: (15 Hours)
Groups-Introduction, Purpose, Use of Group, Editing Existing Group, Limitation of groups, Bins-Introduction, Purpose, Hierarchies- Introduction, Drill downs, Roll ups, Automatic Hierarchies, Create and Remove Custom Hierarchy, Sets- Introduction, How to Use, Types of Sets, Parameters-Parameters, Creating Parameters, Exploring Parameter Controls, use with real time examples.

Unit-IV: (15 Hours)
Annotations- Mark Annotation, Point Annotation, Area Annotation; Page-shelf or Automated Maps-Page Shelf, Use of Page Shelf, Page Shelf Customized Controls;
Getting Started with Dashboard Sections: Dashboard- use of Dashboard and Ways to use Dashboard, Workbook-When to use Workbook, Work sheet- When to use worksheets, Stories- their usage.

Recommended Books:

1. Jon Raasch, Graham Murray, Vadim Ogievetsky, Joseph Lowery, “JavaScript and jQuery for data Analysis and Visualization”, WROX
2. Ritchie S. King, Visual story telling with D3” Pearson

Reference Books:

1.Ben Fry, "Visualizing data: Exploring and explaining data with the processing environment", O'Reilly, 2008.
2. A Julie Steele and Noah Iliinsky, Designing Data Visualizations: Representing Informational Relationships, O’Relly
3. Andy Kirk, Data Visualization: A Successful Design Process, PAKT

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester IV)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Practical Lab - IV
Course Code: MMPR - 403
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. Demonstrate use of basic functions
2. Create their own customized functions
3. Construct tables and figures for descriptive statistics
4. Learn to understand new data sets and functions by yourself
5. Work on built in real time cases for analysis and visualization
6. To learn how to do tokenization, filtration, stemming in NLP using Python
7. To implement top-down and Bottom-up parsers in NLP.
8. To see practically concepts of NLP.

Lab assignments based on Course Code MMT – 401(**Natural Language Processing**) and MMT – 402 (**Data Visualization using Tableau course**).

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester IV)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Deep Learning
Course Code: MET - 404
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:
1. Explain the basics in deep neural networks
2. Apply Convolution neural network for image processing
3. Recurrent Neural Network and its variants for text analysis
4. Apply model evaluation for various applications

Unit-I : Introduction to Deep Learning (15 Hours)
Introduction to Deep Learning, Difference between Machine Learning & Deep Learning, Basic of Neural Networks: Perceptron, Forward Propogation, Backward Propogation, Vanishing Gradient Problem, Activation Functions, Loss Functions.

Unit-II : Convolutional Neural Networks: (15 Hours)
Convolutional Neural Networks; Different types of CNN architectures; Layers in CNN; Convolution Operation: Sparse Interactions, parameter Sharing, Equivariance, pooling; Convolution Variants: Strided, Tiled, Transposed and dilated convolution; CNN Learning: Nonlinearity Functions, Loss Functions, Gradient Computation.

Unit-III : Recurrent Neural Networks (15 Hours)
Introduction to RNN, Advantages and Disadvantages of RNN, Types of RNN, Vanishing & Exploding Gradients problem, Variant RNN architectures, Sequence to Sequence RNN, Bidirectional RNN, Recursive Neural Networks.

Unit-IV : Model Evaluation (15 Hours)
Performance metrics, Performance Metrics for classification and regression, Baseline models: types, benefits, Hyperparameters: Manual Hyperparameter, Automatic Hyperparameter, Grid search, Random search, debugging strategies.

Recommended Books:
1. Ian Good Fellow, YoshuaBengio, Aaron Courville, “Deep Learning”, MIT Press, 2017.
2. Francois Chollet, “Deep Learning with Python”, Manning Publications, 2018.

Reference Books:
1. Phil Kim, “Matlab Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence”, Apress , 2017.
2. Ragav Venkatesan, Baoxin Li, “Convolutional Neural Networks in Visual Computing”, CRC Press, 2018.
3. Navin Kumar Manaswi, “Deep Learning with Applications Using Python”, Apress, 2018.
4. Joshua F. Wiley, “R Deep Learning Essentials”, Packt Publications, 2016Dittman

M. Sc. (Data Science) (Part II) (Level-6.5) (Semester IV)
(NEP-2020)
(Introduced from Academic Year 2026-27)

Title of Course: Social Media Analytics
Course Code: MET – 405
Total Credits: 04
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. Identify various elements of web that can be used for data extraction and mining process.
2. Interpret meaningful trends and hidden insights from social Networking Platform.
3. Gain Knowledge of the Web Architecture and the mechanism of Web Crawling for developing web based systems and applications.
4. Identify the principles of Information Integration and the role of Schema Matching.

UNIT I : Defining Analytics in Social Media (15 Hours)

Introduction, Concept and Scope of Social Media Analytics, Overview of Social Network Ecosystem, The Stages in Social Media Data Analysis workflow, Emerging Trends in Social Media Data Intelligence. **Web Data Mining**: Introduction to Information Retrieval and Web Search, Core principles of Information Retrieval Systems, IR Models and Approaches in Information Retrieval. Web Spamming.

UNIT II : Text and Web Page Preprocessing (15 Hours)

Stop Word Removal, Stemming, Other Pre-processing Tasks for Text, Web Page Preprocessing, Duplicate Detection,
Social Network Analysis: Introduction, Co-Citation and Bibliographic Coupling, HITS: HITS Algorithm, Finding other Eigen Vectors, Relationship with Co-Citation and Bibliographic Coupling, Strength and Weakness of HITS.

UNIT III : WEB Crawling (15 Hours)

Basic Crawler algorithm, Implementation Issues, Universal Crawlers, Focused Crawlers, Topical Crawlers, Evaluation, Crawler Ethics and Conflicts, New Developments. Wrapper Generation: Preliminaries, Wrapper Induction, Instance Based Wrapper Learning, Automatic Wrapper Generation, String Matching And Tree Matching, Multiple Alignment, Building DOM Trees

UNIT IV : Information Integration and Sentiment Analysis (15 Hours)

Information Integration: Introduction to Schema Matching, Pre-processing for schema matching, Schema Level Matching, Domain and Instance Level Matching, Combining Similarities, Other Issues, Integration of Web Query Interfaces, Constructing Unified Global Query Interface.
Sentiment Analysis: Opinion Mining, Document Sentiment Classification, Sentence subjectivity classification, Opinion Lexicon Expansion, Opinion Spam Detection.

Recommended Book:

1. Social Media Analytics Strategy: Using Data to Optimize Buisness Performance Alex Goncalves Las Vegas, Nevada, USA
2. Web Data Mining Exploring Hyperlinks, Contents and Usage Data Bing Liu, Second Edition Springer- Verlag Berlin Heidleberg

References Books:

1.Gautum Shroff" Enterprise Cloud Computing", Cambridge, 2010 Scott Granneman," Google Apps Deciphered: Compute in the cloud to streamline your desktop", Pearson Education, 2008.
2. Social Media Analytics Techniques and Insights for Extracting Business Value Out of Social Media Matthew Ganis, Avinash Kohirkar IBM Press

Title of Course: Research Project
Course Code: RP-406
Total Credits: 06
Course Outcomes: Upon successful completion of this course, the student will be able to:

1. investigate and design a model for research problem identified.
2. implementation of model with appropriate software tools.
3. benchmark the experimental results.
4. writing a research article.
5. identification of appropriate societal issues.
6. development of applications to address identified societal issue.

- Student have to design a model and implement for the research problem identified in semester-III / student can carry out a project which will address societal issues.
- It is preferable to publish the research work carried out in the form of Seminar/ Workshop/ Conference proceedings /Research journal publications.

Scheme of Teaching

1. Each contact session for teaching or practical should be of 60 minutes each.
2. Minimum 60 lectures should be conducted for each course of 4 credits and 30 lectures for 2 credits.
3. One Practical Batch should be of 20 students.
4. Practical evaluation should be conducted before the commencement of University examination

Examination Pattern

Theory:

- **For 4 credit course-** University examinations: 60 marks, Internal evaluation: 40 marks
    - Two tests should be conducted of MCQ type questions. Each test will be of 10 marks
    - Assignments/Seminar/Activities carries 20 marks
- **For 2 credit course-** University examinations: 40 marks, Internal evaluation: 10 marks
    - One test should be conducted of MCQ type questions of 10 marks.
    - Assignments/Seminar carries 10 marks
- The internal marks will be communicated to the University at the end of each semester, but before the semester end examinations. These marks will be considered for the declaration of the results.

Practical:

Practical evaluation will be through university appointed panels of external and internal examiner.

Research Project:

- **For 4 credit course-** University evaluation: 60 marks, Internal evaluation: 40marks.
    - Project Viva by university appointed external and internal examiners.
    - Internal Evaluation will be carried out by internal guide.
    - Student has to submit the report of work carried out in the research project.
- **For 6 credit course-** University Evaluation: 100 marks, Internal Evaluation: 50 marks
    - Project Viva by university appointed external and internal examiners.
    - Internal Evaluation will be carried out by internal guide.
    - Student has to submit the report of work carried out in the research project.

On Job Training:

Student has to make a presentation of the work carried out during On Job Training in front of a panel external and internal examiners. He has to submit the report of work carried out as part of On Job Training.

Nature of Question Paper and Scheme of Marking

Theory:

For 4 Credits:

1) There will be five (5) questions of 12 Marks and out of which all five(5) questions are to be attempted from question no 1 to 5.
2) Question No. 1 to Question No. 4 should consist 2 sub question each carrying 6 marks.
3) Question No.5 is compulsory and is of multiple-choice questions. There will be 6multiple choice question each carries 2 marks

For 2 Credits:

1) There will be three (3) questions of which Question No.1 and Question No.2 will be of 12 Marks and Question No. 3 will be of 6 marks. All three questions must be attempted.
2) Question No. 1 and Question No. 2 should consist 2 sub question each carrying 6 marks.
3) Question No.3 is compulsory and is of multiple-choice questions. There will be 3 multiple choice question each carries 2 marks

Practical:

1) Duration of Practical Examination: 2 Hrs

2) Nature of Question paper: There will be four questions. All questions to be attempted. Question No. 1 and Question No.2 must be Major Question carrying 20 Marks each. Question No.3 and Question No. 4 must be minor Question carrying 10 marks each.
3) The final practical examination will be conducted by the university appointed examiners both internal as well as external at the end of semester for each lab course and marks will be submitted to the university by the panel. The pattern of final Practical Examination will be as follows;

1	Coding and Execution of Program	60 Marks
2	Journal	10 Marks
3	Attendance	10 Marks
4	Oral	10 Marks
5	*Seminar Report	10 Marks
	Total	100 Marks

* Seminar Report: The student has to submit a separate seminar report as per the case may be.

The practical examination will be conducted semester wise in order to maintain the relevance of the respective theory course with laboratory course.

Research Project:

- Student has to identify research problem in the third semester and have to carry out regress literature review.
- In the four-semester student have to design a model and implement for the researchproblem identified in semester three.
- It is preferable to publish the research work carried out in the form of Conference /Research journal publications.

M. Sc. Part II (Semester III and IV)

Old Course				Equivalent Course		
Sem No.	**Course Code**	**Title of Old Course**	**Credit**	**Course Code**	**Title of New Course**	**Credit**
III	MP-301	Machine Learning	4	MMT-202	Artificial Intelligence	4
III	MP-304	Power BI	4	MMT-302	Data Analytics with Power BI	4
III	ME-305.1	Cyber Security	4	MMT – 305	Data Engineering	4
IV	MP-401	Internet of Things	4	MMT-401	Natural Language Processing	4
IV	MP- 402	Data Visualization with Tableau	4	MMT-402	Data Visualization using Tableau	4
IV	ME-405.1	Deep Learning	4	---	*** No Equivalence**	---
IV	ME-405.2	Soft Computing	4	---	*** No Equivalence**	---

*** Two more chances be given to the student.**