Sheth NKTT College of Commerce and Sheth JTT College of Arts, Thane
(Autonomous)
Syllabus Post Graduate Programme As per NEP-2020
(w.e.f. 2025-26)
M.Sc. (Data Science) Part II Sem III

Category	Code	Course	Credit
Mandatory	MDN301	Next Generation Database	**4**
	MDNP302	Next Generation Database Practical	**2**
	MDT303	Text Mining and Natural Language Processing	**4**
	MDTP304	Text Mining and Natural Language Processing Practical	**2**
	MDD305	Data Compliance and auditing	**2**
Elective (Any One)	MDI306	Image and Video Analytics	**2Th+2 P**
	MDIP307	Image and Video Analytics Practical	
	MDD308	Data Science for Agriculture	
	MDDP309	Data Science for Agriculture Practical	
	MDE310	Econometrics and Finance	
	MDEP311	Econometrics and Finance Practical	
RP	MDRP312	Research Project	**4**
		Total	**22**

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Mandatory	
Name of the Dept: **Science and Technology**	
Course Title: **Next Generation Databases**	
Course Code: MDN301	Course Level:6.5
Type: **Theory**	
Course Credit: 4 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 100 Marks	

Course Objectives:

1. Understand the fundamentals of Big Data, its applications, and the Big Data ecosystem.
2. Differentiate between Database systems and Data Warehouses, and understand RDBMS models.
3. Distinguish between ACID and BASE properties and understand distributed file systems.
4. Gain proficiency in various NoSQL database types including document, columnar, and graph databases.

Course Outcomes:

Upon completing this course, the student will be able to:

CO1. Understand the fundamentals of Big Data, its applications in various domains and the technologies in the Big Data ecosystem.

CO2. Understand the difference between Database systems and data warehouses, the architecture of Enterprise Data warehouses, RDBMS models, SQL queries, and ACID properties.

CO3. Differentiate between ACID and BASE, and understand HDFS File Structure and the MapReduce Framework.

CO4. Understand the different types of NoSQL databases like document, columnar, and graph databases.

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	1. Getting Started, Big Data, Facts About Big Data, Big Data Sources, Types of bigdata (structured, semi structured, unstructured), Usage of Big Data (Banking and Finance, Media, and communications etc.), Data governance Policies and Procedures, Legacy Systems and Big Data, Data Storage, Data Processing, Big Data Technologies. 2. Relational Database, Database Design Data Storage, Data warehouse and Data Mining, Information Retrieval, OLTP vs OLAP, Advantages of warehousing approach, Components of a Datawarehouse, types of schemas star schema and snowflake schema, SQL statements, DDL, DML queries, advanced subqueries.	15
II	1. A Brief History of NoSQL, ACID vs. BASE, CAP Theorem, The BASE, Comparison of SQL and NoSQL, Advantages and Disadvantages, Categories of NoSQL Databases (key-value, columnar, document, graph database) 2. History, MongoDB Design Philosophy, Non-Relational Approach, JSON-Based Document Store, Performance vs. Features, Running the Database Anywhere, SQL Comparison. The Data Model, JSON and BSON, The Identifier (_id), Basic Querying, Create and Insert, Explicitly Creating Collections, Inserting Documents Using Loop, Inserting by Explicitly Specifying _id, Update, Delete, Read, Using Indexes, Using Conditional Operators, Regular Expressions, MapReduce, aggregate (), MongoDB Document Data Model Approach.	15
III	1. Concept of distributed systems, transparencies in Distributed system, concept of middleware, Introduction to HDFS, Core components of Hadoop, Hadoop web UI, differences between Hadoop 1,2 and 3 architectures. HDFS shell commands Introduction to MapReduce concept, Pig interface, queries in Pig Latin, data types of Pig Latin. Pig grunt shell. 2. Architecture of Hive, Configuration and starting hive shell, features of Hive, components of hive: user interface, meta store, HiveQL process engine, execution engine, query execution in Hive, data types in Hive, types of tables in Hive, create table, load data in Hive from HDFS.	15

IV	1. Columnar database: HBASE What are columnar databases, Architecture of HBASE, HMaster, Region server, Zookeeper components, Advantages, and disadvantages of HBase, HBASE shell commands, DDL commands, DML commands, security commands, queries. 2. Graph database: Importance of graph database, Graph model, learning about node and relationships, creating a node, creating relationship between nodes, analyzing similarity between nodes, case study using Neo4j	15
	Total Hours	60

References:

1. SQL & NoSQL Databases: Models, Languages, Consistency Options and Architectures for Big Data Management, by Andreas Meier (Author), Michael Kaufmann (Author)
2. Hadoop: The Definitive Guide, Tom White, O'Reilly Media, June 2009
3. Big Data Analytics: Concepts, Techniques, Tools and Technologies, by G. Sudha, M. Thangaraj, S. Suguna (Author), PHI, 30 June 2022
4. Professional NoSQL, Shashank Tiwari, ISBN: 978-0-470-94224-6, Wiley ,September 2011

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Major Mandatory	
Name of the Dept: **Science and Technology**	
Course Title: **Next Generation Databases Practical**	
Course Code: MDNP302	Course Level:6.5
Type: **Practical**	
Course Credit: 2 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. Perform DDL, DML and TCL SQL queries on any RDBMS.
2. Gain thorough knowledge about the structural and conceptual difference between different types of NoSQL systems and their implementations.

Course Outcomes:

Upon completing this course, the student will be able to:

CO1. Perform DDL, DML and TCL SQL queries on any RDBMS.

CO2. Perform ETL operations on any public cloud platform.

CO3. Gain knowledge about bigdata ecosystem and various bigdata technologies.

CO4. Gain thorough knowledge about the structural and conceptual difference between different type of NoSQL systems and their implementations.

Syllabus: NEP 2020 w.e.f 2026-27

Note: - MySQL/ PostgreSQL, Snowflake, Hadoop, MongoDB, Neo4j

Sr. No.	Name of Practical	Hours
1	Perform DDL, DML and TCL queries in any RDBMS	
2	Access any Datawarehouse in cloud platform like Snowflake, Bigquery and perform ETL, data exploration including visualization	
3	Hadoop shell commands	
4	Wordcount implementation using Pig	
5	Data wrangling using Pig.	
6	Create managed and external table in HIVE, view in Hadoop Web UI.	
7	Create a namespace in HBASE with tables. Perform DML queries on the tables	
8	Create a document and collection in MongoDB, performing advanced queries with conditional operators and aggregate functions.	
9	Perform aggregation using Map() and Reduce() function in MongoDB	
10	Creating nodes and relationships in graph database using cypher. Perform exploration queries	
	Total Hours	**60**

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Major Mandatory	
Name of the Dept: **Science and Technology**	
Course Title: **Text Mining and Natural Language Processing**	
Course Code: MDT303	Course Level:6.5
Type: **Theory**	
Course Credit: 4 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 100 Marks	

Course Objectives:

1. Develop the ability to pre-process, clean, and represent text data effectively.
2. Implement and evaluate text classification and topic modeling techniques.
3. Understand and apply advanced NLP methods.
4. Gain proficiency in using Python libraries like NLTK and spaCy for NLP tasks.

Course Outcomes:

Upon completing this course, the student will be able to:

CO1. Pre-process, clean, and represent text data effectively.
CO2. Implement and evaluate text classification and topic modeling techniques.
CO3. Understand and apply advanced NLP methods.
CO4. Gain proficiency in using Python libraries such as NLTK and spaCy.

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	**Introduction to Text Mining and NLP**: Overview of text mining and NLP, Applications of NLP in various domains, Key challenges in text mining (e.g., handling unstructured data, language ambiguity), Introduction to Python libraries for NLP (NLTK, spaCy)	**15**
II	**Text Pre-processing and Representation, Techniques:** Tokenization, stemming, lemmatization, Stopwords removal and text normalization, Text cleaning: handling misspellings, special characters, and case normalization, Methods: Bag-of-Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), Word, Document, and sentence embeddings	**15**
III	**Text Classification:** Overview of text classification tasks, Naive Bayes, SVM, and Logistic Regression for text classification, Evaluation metrics for classification, Introduction to topic modeling, Latent Dirichlet Allocation (LDA), Sentiment analysis, Named Entity recognition, Part of Speech tagging	**15**
IV	**Advanced NLP Techniques and Transformer Models:** Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), Implementation using Keras and TensorFlow, Introduction to transformers (BERT, GPT), Pre-trained models and fine-tuning.	**15**
	Total Hours	**60**

References:

1. Speech and Language Processing by Daniel Jurafsky and James H. Martin,2008, Pearson Prentice Hall
2. Natural Language Processing with Python by Steven Bird, Ewan Klein, and Edward Loper, 2009, O'Reilly Media
3. Neural Network Methods for Natural Language Processing by Yoav Goldberg, 2017, Morgan & Claypool Publishers
4. Foundations of Statistical Natural Language Processing by Christopher D. Manning and Hinrich Schütze, 1999, MIT Press

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Major Mandatory	
Name of the Dept: **Science and Technology**	
Course Title: **Text mining and Natural Language Processing Practical**	
Course Code: MDTP304	Course Level:6.5
Type: **Practical**	
Course Credit: 2 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	
Course Objectives: 1. Understand fundamental concepts of NLP and its processing. 2. Apply NLP techniques in various applications of NLP.	
Course Outcomes: Upon completing this course, the student will be able to: CO1. Understand fundamental concepts of NLP and its processing. CO2. Apply NLP techniques in various application of NLP	

Syllabus: NEP 2020 w.e.f 2026-27

Sr. No.	Name of Practical	Hours
1	Pre-processing methods for text data (tokenization, stop-words etc…)	
2	Implement Stemming	
3	Implement Morphological Analysis	
4	Implement N-gram Model	
5	Implement Part-of-Speech tagging	
6	Implement Chucking	
7	Implement Text Summarization	
8	Implement Named Entity Recognition	
9	Implement Sentiment Analysis	
10	Implement One real time NLP Application on a dataset available.	
	Total Hours	**60**

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Major Mandatory	
Name of the Dept: **Science and Technology**	
Course Title: **Data Compliance and Auditing**	
Course Code: MDD305	Course Level:6.5
Type: **Theory**	
Course Credit: 2	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. Provide a comprehensive overview of major data protection and privacy regulations like GDPR.
2. Equip learners with methodologies for planning, conducting, and documenting data audits.

Course Outcomes:

Upon completing this course, the student will be able to:

CO1. Gain a comprehensive understanding of major data protection and privacy regulations such as GDPR, and their implications for businesses.

CO2. Develop the ability to plan, conduct, and document data audits using industry-standard methodologies and tools to ensure data accuracy and compliance.

Syllabus: NEP 2020 w.e.f 2024-25

Unit No.	Content	Hours
I	**UNIT I: Data Compliance** **General Data Protection Regulation (GDPR)** - Data Protection, Material scope of the GDPR, The building blocks of the GDPR, Compliance framework – the standards of protection **Eprivacy** - Regulating the electronic communications sector, The relationship between data protection and ePrivacy **Operational Data Protection** - The three layers of an organization, implementing data protection in the people layer, implementing data protection in the paper layer. **Core Law of Data Protection-** The principles of data protection, Lawfulness, fairness and transparency – the first principle, Purpose limitation – the second principle, Data minimization – the third principle. **The Rights of Data Subjects-** Transparency and information rights, Rights over data processing. Mechanisms to Support Operational Compliance. Programmatic Approaches for Delivering Data Protection By Design And Default.	15
II	**UNIT II: Data Auditing** **Information Security Performance Metrics and Audit:** Pre-audit checklist, Information Gathering, Vulnerability Analysis, External Security Audit, Internal Network Security Audit, Firewall Security Audit, IDS Security Auditing, Social Engineering Security Auditing, Web Application Security Auditing. **Vulnerability Management:** Information Security Vulnerabilities – Threats and Vulnerabilities, Human-based Social Engineering, Computer-based Social Engineering, Social Media Countermeasures, Vulnerability Management – Vulnerability Scanning, Testing, Threat management, Remediation etc. **Information Security Assessments:** Vulnerability Assessment, Classification, Types of Vulnerability Assessment, Vulnerability Assessment Phases, Vulnerability Analysis Stages, Characteristics of a Good Vulnerability Assessment Solutions &Considerations, Vulnerability Assessment Reports – Tools and choosing a right Tool, Information Page 10 of 37 https://e-next.in Security Risk Assessment, Risk Treatment, Residual Risk, Risk Acceptance, Risk Management Feedback Loops etc. **Configuration Reviews:** Introduction to Configuration Management, Configuration Management Requirements-Plan Control, Development of configuration Control Policies, Testing Configuration Management etc.	15
	Total Hours	30

References:

1. Data Protection and Compliance, 2nd Edition by Stewart Room, Michelle Maher, Niall O'Brien, Adam Panagiotopoulos, Shervin Nahid, Richard Hall, Tughan Thuraisingam, James Drury-Smith, Simon Davis, Mark Hendry, Jamie Taylor, Ben Johnson Released November 2021
2. Data auditing The Ultimate Step-By-Step Guide by Gerardus Blokdyk, 2018, 5STARCooks
3. Assessing Information Security (strategies, tactics, logic and framework) by A Vladimirov, K.Gavrilenko, and A.Michajlowski, Addison-Wesley Professional , 2010

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Elective	
Name of the Dept: **Science and Technology**	
Course Title: **Image and Video Analytics**	
Course Code: MDI306	Course Level:6.5
Type: **Theory**	
Course Credit: 2	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. Describe the foundational concepts and theoretical framework of digital image and video processing.
2. Develop the capability to implement algorithms for image enhancement, feature extraction, and video-based object tracking.

Course Outcomes:

Upon completing this course, the student will be able to:

CO1. Apply image processing techniques to enhance, segment, and extract features from images for diverse applications.

CO2. Implement motion analysis and utilize video analytics to perform event recognition and anomaly detection in dynamic scenes.

Syllabus: NEP 2020 w.e.f 2024-25

Unit No.	Content	Hours
I	**Unit 1: Image Analysis** **Introduction to Image Processing:** Overview of digital images and basic image processing concepts, Image acquisition and representation, Introduction to image processing tools (e.g., OpenCV, PIL) **Image Enhancement Techniques:** Histogram equalization, Filtering: smoothing, sharpening, Edge detection and image segmentation **Feature Extraction and Description:** Keypoint detection (SIFT, SURF), Feature descriptors (HOG, LBP), Object recognition and classification Advanced Image Analysis: Image transforms (Fourier, Wavelet), Morphological operations, Image stitching and panoramic images	15
II	**Unit 2: Video Analytics** **Introduction to Video Processing:** Video acquisition and representation, Basic concepts of video processing, Introduction to video analytics frameworks **Motion Analysis:** Optical flow, Motion detection and tracking, Background subtraction techniques **Object Detection and Tracking:** Algorithms for object detection (YOLO, SSD), Multi object tracking (SORT, Deep SORT), Real-time tracking systems. **Video Event and Activity Recognition:** Temporal event detection, Activity recognition and anomaly detection, Video summarization techniques.	15
	Total Hours	**30**

References:

1. Intelligent Image and Video Analytics, El-Sayed M. El-Alfy, George Bebis, Mengchu Zhou, Routledge, Taylor and Francis Group, CRC, 2023
2. Machine Learning for Audio, Image and Video Analysis:Theory and Applications, Francesco Camastra, Alessandro Vinciarelli, Springer, 2015
3. Computer Vision Using Deep Learning Neural Network Architectures with Python and Keras, Apress 2021(UNIT-III,IV and V), Vaibhav Verdhan, 2021
4. “Image Processing, Analysis, and Machine Vision”, Milan Sonka, Vaclav Hlavac, Roger Boyle, 4th edition, Thomson Learning, 2013.
5. Richard Szeliski, “Computer Vision: Algorithms and Applications”, Springer Verlag London Limited,2011.
6. “Video Analytics for Business Intelligence”, Caifeng Shan, FatihPorikli, Tao Xiang, Shaogang Gong, Springer, 2012.
7. “Computer Vision: A Modern Approach”, D. A. Forsyth, J. Ponce, Pearson Education, 2003.
8. “Computer & Machine Vision”, E. R. Davies, (2012), Fourth Edition, Academic Press.

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M.Sc (Data Science)** **Semester: III**
Course Category/Vertical: Elective
Name of the Dept: **Science and Technology**
Course Title: **Image and Video Analytics Practical**
Course Code: MDIP307 Course Level:6.5
Type: **Practical**
Course Credit: 2 credits
Hours Allotted: 60 Hours
Marks Allotted: 50 Marks
Course Objectives: 1. Capture, manipulate, and format images and video frames using OpenCV. 2. Implement and evaluate image enhancement and segmentation methods
Course Outcomes: Upon completing this course, the student will be able to: CO1. Capture, manipulate, and format images and video frames using OpenCV and implement and evaluate image enhancement and segmentation methods. CO2. Utilize algorithms to extract and match features between images and apply transformations and morphological operations for complex image analysis. CO3. Implement motion detection and object tracking in dynamic video scenes and execute face detection and event recognition for video summarization.

Syllabus: NEP 2020 w.e.f 2026-27

Note: Access to Python and relevant libraries (OpenCV, scikit-image, etc.).

Sr. No.	Content	Hours
1	Image Acquisition and Basic Processing: Acquire and manipulate images (resize, crop, format conversion) using OpenCV.	
2	Histogram Equalization: Enhance image contrast by applying histogram equalization techniques.	
3	Image Filtering and Edge Detection: Apply filters and detect edges using Sobel and Canny methods	
4	Image Segmentation: Segment images using global and adaptive thresholding and k-means clustering.	
5	Feature Extraction and Matching: Extract keypoints and match features between images using SIFT or ORB.	
6	Image Transformation and Morphology: Perform Fourier/Wavelet transforms and apply morphological operations.	
7	Image Stitching and Panorama Creation: Stitch multiple images together to create a panoramic image.	
8	Video Capture and Frame Extraction: Capture video and extract frames at regular intervals using OpenCV.	
9	Motion Detection and Optical Flow: Detect motion and analyze optical flow using background subtraction and optical flow algorithms.	
10	Object Detection in Videos: Detect objects in video sequences using YOLO or SSD models.	
11	Face Detection Using Haar Cascades: Implement face detection in images and videos using Haar cascade classifiers.	
12	Multi-Object Tracking: Track multiple objects in videos using algorithms like SORT or Deep SORT.	
13	Video Event Detection and Summarization: Detect events and create summarized videos highlighting key activities.	
	Total Hours	**60**

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Elective	
Name of the Dept: **Science and Technology**	
Course Title: **Data Science for Agriculture**	
Course Code: MDD308	Course Level:6.5
Type: **Theory**	
Course Credit: 2	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. Proficiency in data collection, preprocessing, and analysis for agricultural data.
2. Ability to build predictive models and perform EDA specific to agriculture

Course Outcomes:

Upon completing this course, the student will be able to:

CO1. Proficiency in data collection, preprocessing, and analysis for agricultural data and ability to build predictive models and perform EDA specific to agriculture.

CO2. Skills in remote sensing, GIS, and precision agriculture techniques and competence in developing data-driven solutions for agricultural challenges.

CO3. Enhanced understanding of sustainability and environmental impacts in agriculture.

Syllabus: NEP 2020 w.e.f 2024-25

Unit No.	Content	Hours
I	**Introduction to the role of data science in agriculture:** Key challenges and opportunities. Overview of agricultural data sources, Types of data in agriculture (soil data, weather data, crop data, etc.), Importance of big data in agriculture, Case studies of big data applications in agriculture. Data Collection Techniques: Remote sensing and satellite imagery, IoT and sensor networks in agriculture, Manual data collection methods, Data Management: Cloud storage solutions for agricultural data. Data Cleaning and Preprocessing, Exploratory data Analysis (EDA): Techniques for summarizing and visualizing data, identifying patterns and correlations, Using EDA tools and libraries.	15
II	**Machine Learning Fundamentals:** Introduction to supervised and unsupervised learning, Common algorithms (regression, classification, clustering), Model evaluation and validation. Predictive Analytics in Agriculture: Yield prediction models, Disease and pest prediction, Weather forecasting and its impact on agriculture. Advanced Machine Learning Techniques: Deep learning applications in agriculture, Computer vision for crop monitoring, Natural language processing for agricultural data, Applications of machine learning in precision farming. Case studies and real-world examples.	15
	Total Hours	30

References:

1. Data Science for Agriculture by Simon Stelling, April 14, 2020
2. Data Science for Agriculture" by David L. Bunge
3. Practical Data Science for Agriculture" by Kellee Koenig

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M.Sc (Data Science)**	**Semester: III**
Course Category/Vertical: Elective	
Name of the Dept: **Science and Technology**	
Course Title: **Data Science for Agriculture Practical**	
Course Code: MDDP309	Course Level:6.5
Type: **Practical**	
Course Credit: 2 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. Understand and apply data collection techniques for agricultural datasets.
2. Perform Exploratory Data Analysis (EDA) and develop predictive models.

Course Outcomes:

Upon completing this course, the student will be able to:

CO1. Understand and Apply Data Collection Techniques and perform Exploratory Data Analysis (EDA).

CO2. Develop Predictive Models and implement Precision Agriculture Techniques.

CO3. Utilize Remote Sensing and GIS and forecast Market Prices and Analyze Supply Chains.

Syllabus: NEP 2020 w.e.f 2026-27

Note: Python, R, data analysis libraries (Pandas, NumPy, Scikit-learn), and data visualization tools (Matplotlib, Seaborn, Tableau). 02

Sr. No.	Content	Hours
1	Web Scraping for Agricultural Data: Scrape weather data, soil data, and crop yield data from relevant websites - Tools: Python, BeautifulSoup, Scrapy. Cleaning and Preprocessing Data: Handle missing values, outliers, and normalize data for further analysis. Exploratory Analysis of Crop Yield Data: Perform descriptive statistics and visualize data distributions, trends, and correlations.	
2	Prediction Using Time Series Analysis: Predict future weather patterns using historical weather data.	
3	Crop Yield Prediction Using Machine Learning: Build a machine learning model to predict crop yields based on various factors (soil type, weather, etc.). Tools: Scikit-learn, XGBoost	
4	Analyzing Satellite Images for Crop Health Monitoring: Use satellite imagery to assess crop health and identify areas needing attention. Tools: QGIS, Python, OpenCV	
5	Land Use Classification Using Remote Sensing Data: Classify land use patterns using remote sensing data and machine learning algorithms. Tools: Python, Scikit-learn, GDAL.	
6	Soil Moisture Prediction Using Sensor Data: Predict soil moisture levels using data from soil sensors and environmental factors.	
7	Optimizing Irrigation Systems Using Data Analytics: Analyze irrigation data to optimize water usage and improve crop yield.	
8	Plant Disease Detection Using Image Processing: Detect plant diseases using image processing techniques on leaf images.	
9	Creating Yield Maps Using GPS Data: Generate yield maps to visualize spatial variability in crop production.	
10	Price Forecasting for Agricultural Products: Predict market prices for agricultural products using historical price data and other relevant factors.	
	Total Hours	**60**

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M. Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Elective	
Name of the Dept: **Science and Technology**	
Course Title: **Econometrics and Finance**	
Course Code: MDE310	Course Level:6.5
Type: **Theory**	
Course Credit: 2	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	
Course Objectives: 1. Learners will understand the concepts of Econometrics. 2. Learners will be able to understand different stages in model development.	
Course Outcomes: Upon completing this course, the student will be able to: CO1. Learners will understand the concepts of Econometrics. CO2. Learners will be able to understand different stages in model development.	

Syllabus: NEP 2020 w.e.f 2024-25

Unit No.	Content	Hours
I	**Unit 1: Econometrics** Introduction to Econometrics– definitions – scope – methodology – types. Uni-variate regression model, Assumptions, Gauss Markov Theorem. Multivariate regression model and Inferential Analysis. Test of significance of parameter estimates. Test of goodness of fit and correlation Problem and Methods to detect Heteroscedasticity, Multicollinearity and Autocorrelation in the Regression Model, its causes, consequences and Remedial Measures. Dummy variables: Nature of Dummy variables – Use of Dummy Variables – Errors in Variables and its consequences.	15
II	**Unit 2: Financial Modelling** Introduction to financial modelling - objectives of financial modelling; spreadsheet features, techniques; best practices in spreadsheet design. Designing models - Model Design and structure; Building business case models; spreadsheet techniques and methods. Auditing and Testing: Essential testing and auditing techniques; Testing financial analysis model with cash flows and ratios; Debugging and checking a financial model. Macros and Security: Writing and auditing and macros; Spreadsheet security, Model Completion: Model review; Documentation; Final audit.	15
	Total Hours	**30**

References:

1. Damodar N. Gujarathi: Basic Econometrics, New Delhi: Tata McGraw Hill,5th edition, 2009
2. J.Johnston: Econometric Methods, McGraw Hill, 4th edition , 1997
3. John Hull: Futures, Options and Other Derivatives, Prentice Hall, 10th edition, 2021
4. Financial Valuation and Modeling by Sheeba Kapil, Wiely, 3rd Edition, 2021

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: **M.Sc. (Data Science)**	**Semester: III**
Course Category/Vertical: Elective	
Name of the Dept: **Science and Technology**	
Course Title: **Econometrics and Finance Practical**	
Course Code: MDEP311	Course Level:6.5
Type: **Practical**	
Course Credit: 2 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	
Course Objectives: 1. Learn the practical application of Econometrics. 2. Use Econometrics in financial modelling.	
Course Outcomes: Upon completing this course, the student will be able to: CO1. Learners will learn practical application of Econometrics. CO2. Learners will be able to use Econometrics in financial modelling	

Syllabus: NEP 2020 w.e.f 2026-27

Note: Working on - Micro soft Excel / R Programming / Python

Sr. No.	Content	Hours
1	Uni-variate Regression Model	
2	Multivariate Regression Model	
3	Methods to detect heteroscedasticity and remedial measures	
4	Methods to detect multicollinearity and remedial measures	
5	Methods to detect auto-correlation and remedial measures.	
6	Interpreting on Dummy variable	
7	Financial modelling: best practices in spreadsheet designnd Security	
8	Model Design and structure	
9	Auditing and Testing	
10	Macros and Security	
	Total Hours	**60**