Curriculum Structure for Multi-Disciplinary Minor (MDM) Degree Offered by Different Programs

(In light of NEP 2020)

(NEP-Version II)

For students admitted in 2023-24 onwards

Implemented year-2024-25

Government College of Engineering, Amravati

(An Autonomous Institute of Government of Maharashtra)

Near Kathora Naka, Amravati, Maharashtra

PIN 444604

Curriculum Structure for Multi-Disciplinary Minor (MDM) Degree in Data Science

(In light of NEP 2020)

(NEP_Version II)

Offered By

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

For students admitted in 2023-24 onwards

Government College of Engineering, Amravati
(An Autonomous Institute of Government of Maharashtra)
Near Kathora Naka, Amravati, Maharashtra
PIN 444604
www.gcoea.ac.in

Program Specific Outcome (PSO's)

PSO 1: The ability to understand, analyze and demonstrate the knowledge of human cognition, Artificial Intelligence, Machine Learning and data engineering in terms of real world problems to meet the challenges of the future.

PSO 2: The ability to develop computational knowledge and project development skills using innovative tools and techniques to solve problems in the areas related to Deep Learning,

PSO 3: Ability to contribute to problem identification, analysis, design, and development of systems using principles and concepts of Artificial Intelligence and Data Science.

A. Preamble:

The preamble of data science sets the foundational principles and objectives for the field. It typically emphasizes the interdisciplinary nature of data science, its reliance on various techniques from statistics, computer science, and domain-specific knowledge, and its overarching goal of extracting insights and knowledge from data to inform decision-making and solve complex problems. This preamble often underscores the importance of data collection, data processing, analysis, interpretation, and communication of findings. Additionally, it may highlight ethical considerations, such as privacy, fairness, and transparency, as well as the need for continuous learning and adaptation in response to evolving technologies and data landscapes.

Data Science being a multidisciplinary field that has its roots in statistics, math and computer science, posits immense value to students across disciplines in their multi-fold application areas. The various skill sets required in this domain is illustrated below:

B. Structure of the MDM course:

CSE Department offer Multidisciplinary Minor Basket , Track-1 (Data Science)														
Category	**Course Code**	**Name of the Course**	**Teaching Scheme**				**Evaluation Scheme**							**Credits**
							Theory				**Practical**		**Total**	
			TH	TU	PR	Total	CT1	**CT2**	TA	ESE	ICA	ESE		
MM1	CS1315	Fundamentals of data science	3			3	15	15	10	60			100	3
MM2	CS1415	Computational Data Analytics	3			3	15	15	10	60			100	3
MM3	CS1515	Introduction to Database System	3			3	15	15	10	60			100	3
MM4	CS1615	Machine Learning	3			3	15	15	10	60			100	3
MM5	CS1715	Advance Machine Learning and Deep Learning	2			2	15	15	10	60			100	2
Total			**14**	**0**	**0**	**14**	**75**	**75**	**50**	**300**	**0**	**0**	**500**	**14**

C. **Eligibility criteria:** Students enrolled in B. Tech program other than Computer Science & Engineering are eligible. The allotment of minor degree Program will be as per the policy of the Institute.

D. Intake: Minimum 15,

E. Detailed syllabus:

SEMESTER – III

Course Code	**CS1315**							**Course category**			**MM1**
Course Name	**FUNDAMENTALS OF DATA SCIENCE**										
Teaching Scheme				Examination Scheme							Credits
				Theory					Practical		
Th	Tu	Pr	Total	CT 1	CT 2	TA	ESE	ESE Duration	ICA	ESE	Total
03	00	00	03	15	15	10	60	**2**hr **3**0 min	00	00	100	03

Course Objectives:

To make the students aware and understand:

1. Describe the significance of data science and understand the Data Science process.
2. Explain how data is collected, managed and stored for data science.
3. Analyse Basic tools of EDA, Data science process with case studies and Different algorithms.
4. Analyse Data using various Visualization techniques.
5. Explore Feature Generation and Feature Selection.

Course Contents:

Introduction To Data Science: Definition, Big Data and Data Science Hype, Datafication , Data Science Profile, Meta-Definition, Data Scientist, Statistical Inference, Populations and Samples, Populations and Samples of Big Data, Big Data Can Mean Big Assumptions, Modeling, Philosophy of Exploratory Data Analysis, The Data Science Process.

Mathematical Preliminaries: Probability, Descriptive Statistics, Correlation Analysis.
Probability Distribution: Conditional Probability and Bayes' Theorem, Random variable and probability distribution. Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of a Continuous Random variable. Various probability distribution (Binomial, Poisson, Geometric, Uniform, Exponential, Normal, Chi-Square, Student's t-distribution, F-distribution)

Inferential Statistics: Sampling and its various techniques , Estimation, Sampling distribution of mean and proportion, Normal distribution and z — statistic, Central limit theorem, Confidence Interval estimation for mean and proportion, sample size estimation, estimation of parameters.

Exploratory Data Analysis and the Data Science Process: Basic tools (plots, graphs and summary statistics) of EDA, Philosophy of EDA, The Data Science Process, Case Study: RealDirect (online real estate firm). Algorithms, machine Learning Algorithms, Three Basic Algorithms: Linear Regression, k-Nearest Neighbours (kNN), k-means, R Programs for the algorithms

Text Books:

1. Steven S. Skiena, "The Data Science Design Manual", Springer 2017.
2. Rachel Schutt & O'neil, "Doing Data Science", Straight Talk from The Frontline O'REILLY, ISBN:978-1-449-35865-5, 1st edition, October 2013.

Reference Books :

1. Cathy O Neil, Rachel Schutt, 2014, "Doing Data Science-Straight Talk from the Frontline", Orielly
2. Joel Grus," Data Science from Scratch" First Edition, April 2015

Course Outcomes:

At the end of the course the student will be able to:

CS1315.1 Understand the concepts of Data collection and management.

CS1315.2 Understand types of data Visualization techniques.

CS1315.3 Analyze the Feature Selection algorithms and Recommendation Systems.

CS1315.4 Explain and programme Data Science, Big data and fitting model.

CS1315.5 Design Map Reduce Solutions.

CO – PO – PSO Mapping:

Course Outcomes	Program Outcomes														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CS1315.1	0	3	3	2	2	2	0	2	3	3	3	1	1	2	3
CS1315.2	1	2	2	1	3	1	2	1	0	1	2	2	1	1	3
CS1315.3	2	1	3	1	0	2	2		2	1	0	2	2	0	0
CS1315.4	3	2	3	1	2	0	1	2	1	3	1	2	0	3	2
CS1315.5	2	3	1	3	1	3	1	0	2	3	3	3	2	0	0

0-Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlate

SEMESTER – IV

Course Code	**CS1415**	**Course category**	**MM2**
Course Name	**COMPUTATIONAL DATA ANALYTICS**		

Teaching Scheme				Examination Scheme								
				Theory					Practical			Credits
Th	Tu	Pr	Total	CT 1	CT 2	TA	ESE	ESE Duration	ICA	ESE	Total	
03	00	00	03	15	15	10	60	2hr 30 min	00	00	100	03

Course Objectives:

1. To explore data for understanding data characteristics.
2. To explore the statistical analysis techniques for analyzing data.
3. Tolearntheprobabilitydistributionsanddensityestimationstoperformanalysisof various kinds of data.
4. To analyze the data generated in real life problems with the help of important statistical tools and techniques.
5. To understand the application of data science in various sectors.

Course Contents:

Data analytics importance and overview: Data analytics benefits, Terminologies in data analytics, Data categorization (constant and variable; discrete and continuous; Qualitative and Quantitative; structure, semi-structured and unstructured, cross-sectional, time-series and panel), data measurement scale. Types of Analytics (Descriptive, predictive, prescriptive, diagnostic) Descriptive Analytics: Measures of Central Tendency, Measures of Variation, Measures of Shape and symmetry.

Data Munging: Properties of Data, Languages for Data Science, Collecting Data, Cleaning Data, Crowdsourcing.

Visualizing Data: Exploratory Data Analysis, Developing a Visualization Aesthetic, Chart Types, Great Visualizations Mathematical Models: Philosophies of Modelling, A Taxonomy of Models, Baseline Models, Evaluating Models, Evaluation Environment.

Feature Generation and Feature Selection (Extracting Meaning from Data): Motivating application: user (customer) retention. Feature Generation (brainstorming, role of domain expertise, and place for imagination), Feature Selection algorithms. Filters; Wrappers; Decision Trees; Random Forests. Recommendation Systems: Building a User- Facing Data Product, Algorithmic ingredients of a Recommendation Engine, Dimensionality Reduction, Singular Value Decomposition, Principal Component Analysis, Exercise: build your own recommendation system.

Data Engineering: Map reduce, Word Frequency Problem, Map Reduce Solution, Other Examples of Map Reduce, Mining Social Network Graphs: Social networks as graphs, clustering of graphs, direct discovery of communities in graphs, Partitioning 2 of graphs

Data analytics in different sectors: How Google, LinkedIn, Amazon, Netflix uses analytics Data analytics in media and entertainment industry, education, government, weather forecasting.

Course Outcomes:

At the end of the course the student will be able to:

CS1415.1 Understand and describe the role of data science and its tools.

CS1415.2 Apply mathematical and statistical principles to the analysis of data.

CS1415.3 Apply correlations, distributions and hypothesis tests for inference.

CS1415.4 Apply basic analysis techniques for real time problems.

CS1415.5 Implement data analytics in different sectors.

CO – PO – PSO Mapping:

Course Outcomes	Program Outcomes														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CS1415.1	2	1	2	2	3	2	0	1	0	1	0	1	2	2	1
CS1415.2	2	1	2	2	3	2	0	1	0	1	0	2	2	3	2
CS1415.3	3	3	3	3	3	0	0	2	1	1	0	2	2	2	2
CS1415.4	3	3	3	3	3	0	0	2	1	1	0	2	2	3	2
CS1415.5	3	3	3	3	3	0	0	2	1	1	0	1	2	3	2

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlate

SEMESTER – V

Course Code	CS1515								Course category			MM
Course Name	INTRODUCTION TO DATABASE SYSTEM											
Teaching Scheme				Examination Scheme								
Th	Tu	Pr	Total	Theory					Practical		Total	Credits
				CT1	CT2	TA	ESE	ESE Duration	ICA	ESE		
03	00	00	03	15	15	10	60	**2**hr 30 min	00	00	100	03

Course Objectives:

I. To Understand the basic concepts of DBMS
II. To learn data models, conceptualize and depict a database system using ER diagram.
III. To understand the different database models and language queries to access databases
IV. To understand the normalization forms in building an effective database tables

Course Content

Introduction: Database System Applications, Purpose of Database System, Views of data, data models, Database Languages, database architecture and components of DBMS, Database System Applications, Database Users and Administrators, Database System Structure, History of Database Systems. Entity-Relationship Model, Basic Concepts, Design Issues, Entity- Relationship Diagram, ER Model, notations, examples.

Relational Model: Relational Data Model, Concept of relations, schema-instance distinction, referential integrity constraints, keys, referential integrity and foreign keys, relational algebra operators, Extended Relational-Algebra Operations

SQL: Introduction, data definition in SQL, table, key and foreign key definitions, update behaviors. Querying in SQL, notion of aggregation, aggregation functions group by and having clauses, embedded SQL. Domain Constraints, Referential Integrity, Assertions, Triggers, Security and Authorization, Authorization in SQL, Encryption and Authentication.

Transactions: Transaction processing and Error recovery - concepts of transaction processing, ACID properties, concurrency control, locking based protocols for CC, error recovery and logging, undo, redo, undo-redo logging and recovery methods.

Current Issues: Rules, Knowledge Bases, Active and Deductive Databases, Multimedia Databases Multimedia Data Structures, Multimedia Query languages, Spatial Databases.

Text Book

1. Silberschatz, Abraham, Korth, Henry F., and Sudharshan, S., *"Database System Concepts"*, 6thEdition, Tata McGraw Hill, 2017.

Reference Books

1. Elmasri, R. and Navathe, S. B. "Fundamentals of Database Systems", 5th Edition, Pearson/Addison Wesley, 2017.
2. Connolly, Thomas and Begg, Carlolyn, "Database Systems: A Practical Approach to Design, Implementation, and Management", 5th Edition, Addison-Wesley, 2014.

Course Outcomes

After completion of this course student will be able to

CS1515.1 Comprehend the complex query processing techniques

CS1515.2 Design and implement multimedia databases and writing query structure

CS1515.3 Install, configure and interact with a relational database management system.

CS1515.4 Master the basics of SQL and construct queries using SQL.

CS1515.5 Ability to develop skill set in Multimedia Databases, of transaction processing, Transaction management, and database administration techniques

CO – PO – PSO Mapping:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CS1515.1	3	3	3	1	1	0	0	0	0	0	2	2
CS1515.2	3	3	3	1	1	0	0	0	0	0	2	2
CS1515.3	3	3	2	1	1	0	0	0	0	0	2	2
CS1515.4	3	2	2	1	1	0	0	0	0	1	2	2
CS1515.5	3	2	2	1	1	0	1	0	0	1	2	2

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated

CO – PO – PSO Mapping: As per the new NBA Guidelines w.e.f. 01/01/2025

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CS1515.1	3	3	3	1	1	0	0	0	0	0	2
CS1515.2	3	3	3	1	1	0	0	0	0	0	2
CS1515.3	3	3	2	1	1	0	0	0	0	0	2
CS1515.4	3	2	2	1	1	0	0	0	0	1	2
CS1515.5	3	2	2	1	1	0	1	0	0	1	2

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated

SEMESTER - VI

Course Code	**CS1615**	**Course category**	MM
Course Name	**MACHINE LEARNING**		

Teaching Scheme				Examination Scheme								
				Theory					Practical		Total	Credits
Th	Tu	Pr	Total	CT1	CT2	TA	ESE	ESE Duration	ICA	ESE		
03	00	00	03	15	15	10	60	**2**hr 30 min	00	00	100	03

Course Objectives:

Aim of course is

I. To use machine learning library (e.g., Python with Scikit-learn) to build and train models on real data
II. To explore supervised, unsupervised and reinforcement learning paradigms of machine learning.
III. To design and implement machine learning solutions to classification, regression, and clustering problems.
IV. To learn to apply machine learning algorithms on real world problems.

Course Content

Introduction: Machine Learning , Problems Machine Learning Can Solve ,Knowing Your Task and Knowing Your Data , Necessity of Python ,scikit-learn, Installing scikit-learn , Essential Libraries and Tools , Jupyter Notebook ,NumPy, SciPy ,matplotlib , pandas , mglearn , Classifying Iris Species ,Meet the Data, Measuring Success: Training and Testing Data ,Building Your First Model: k-Nearest Neighbors ,Making Predictions , Evaluating the Model

Supervised Learning: Classification and Regression, Generalization, Overfitting, and Underfitting, Relation of Model Complexity to Dataset Size, Supervised Machine Learning Algorithms, Some Sample Datasets, k-Nearest Neighbors, Linear Models, Naive Bayes Classifiers, Decision Trees, Ensembles of Decision Trees, Kernelized Support Vector Machines, Neural Networks (Deep Learning), Uncertainty Estimates from Classifiers, The Decision Function, Predicting Probabilities, Uncertainty in Multiclass Classification

Unsupervised Learning and Preprocessing: Types of Unsupervised Learning, Challenges in Unsupervised Learning, Preprocessing and Scaling, Different Kinds of Preprocessing, Applying Data Transformations, Scaling Training and Test Data the Same Way, The Effect of Preprocessing on Supervised Learning, Dimensionality Reduction, Feature Extraction, and Manifold Learning, Principal Component Analysis (PCA), Clustering, k-Means Clustering, Agglomerative Clustering, Comparing and Evaluating Clustering Algorithm

Representing Data and Engineering Features: Categorical Variables, One-Hot-Encoding (Dummy Variables), Numbers Can Encode Categoricals, Binning, Discretization, Linear Models, and Trees, Interactions and Polynomials, Univariate Nonlinear Transformations, Automatic Feature Selection, Univariate Statistics, Model-Based Feature Selection, Iterative Feature Selection, Utilizing Expert Knowledge

Model Evaluation and Improvement: Cross-Validation, Cross-Validation in scikit-learn, Benefits of Cross-Validation, Stratified k-Fold Cross-Validation and Other Strategies, Grid Search, Simple Grid Search, The Danger of Overfitting the Parameters and the Validation Set , Grid Search with Cross-Validation, Evaluation Metrics and Scoring, Keep the End Goal in Mind, Metrics for Binary Classification, Metrics for Multiclass Classification, Regression Metrics, Using Evaluation Metrics in Model Selection

Text Book:

3. "Introduction to Machine Learning with Python", Andreas C. Müller & Sarah Guido, O'Reilly Media, Incorporated, 2018
4. "Introduction to Statistical Machine Learning with Applications in R", Hastie, Tibshirani, Friedman, Springer, 2nd Edition-2012.

Reference Books:

2. "Introduction to Machine Learning", Ethem Alpaydin PHI 2nd Edition-2013.

Course Outcomes:

After completion of this course student will be able to

CS1615.1 Understand fundamental machine learning concepts, essential libraries and tools
CS1615.2 Identify suitable algorithms for different types of problems.
CS1615.3 Create unsupervised learning models for handling unknown pattern
CS1615.4 Understand the importance of representing data in a way that is suitable for the machine learning algorithm
CS1515.5 Validate and evaluate the performance of machine learning models using metrics like accuracy, precision, recall, F1-score, and understand the limitations of different metrics.

CO – PO – PSO Mapping:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CS1615.1	2	2	2	2	0	0	0	0	0	0	0	1
CS1615.2	2	3	3	2	2	0	0	0	0	0	1	1
CS1615.3	2	2	3	2	3	0	0	0	0	0	1	1
CS1615.4	2	3	2	3	2	0	0	0	0	0	0	1
CS1615.5	2	3	3	3	2	0	0	0	2	0	1	1

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated

CO – PO – PSO Mapping: As per the new NBA Guidelines w.e.f. 01/01/2025

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CS1615.1	2	2	2	2	0	0	0	0	0	0	1
CS1615.2	2	3	3	2	2	0	0	0	0	0	1
CS1615.3	2	2	3	2	3	0	0	0	0	0	1
CS1615.4	2	3	2	3	2	0	0	0	0	0	1
CS1615.5	2	3	3	3	2	0	0	0	2	0	1

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated

Curriculum Structure for Multi-Disciplinary Minor (MDM) Degree in Artificial Intelligence

(In light of NEP 2020)

(NEP_Version II)

Offered By

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

For students admitted in 2023-24 onwards

Government College of Engineering, Amravati
(An Autonomous Institute of Government of Maharashtra)
Near Kathora Naka, Amravati, Maharashtra
PIN 444604
www.gcoea.ac.in

A. Preamble:

It focusses on integrating AI and Machine Learning skills to all core engineering students to enhance their employability skills. The prime objective of this Programme is to develop students to excel in Machine Learning specific areas by exposing them to understand mathematical, statistical, and biologically inspired computational models with engineering and scientific principles to devise solutions for societal and business problems. As the business world witnesses the power of data-driven decision making, there is a lot of demand for professionals who can build and execute mathematical and statistical models/algorithm. To harness the power of data, effective methodologies are required for extracting hidden patterns from the data. This Programme focusses more on algorithmic aspects of AI and Machine Learning where students study various statistical and machine learning algorithms, neural network architectures, methods to build ML stack, and required technologies for building and hosting ML models. The courses in this Minor degree designed to introduce the students to core skills such as foundations of Artificial Intelligence, Machine learning, Feature Engineering for ML, Data Analytics and Deep learning architectures and its applications.

B. Structure of the MDM course:

CSE Department offer Multidisciplinary Minor Basket , Track-1 (Artificial Intelligence)														
Category	**Course Code**	**Name of the Course**	**Teaching Scheme**				**Evaluation Scheme**							**Credits**
							Theory				**Practical**		**Total**	
			TH	TU	PR	Total	CT1	**CT2**	TA	ESE	ICA	ESE		
MM1	CS1316	Introduction to Artificial Intelligence	3			3	15	15	10	60			100	3
MM2	CS1416	Data Mining	3			3	15	15	10	60			100	3
MM3	CS1516	Machine Learning	3			3	15	15	10	60			100	3
MM4	CS1616	Deep Learning	3			3	15	15	10	60			100	3
MM5	CS1716	Generative AI with Large Language Model	2			2	15	15	10	60			100	2
Total			**14**	**0**	**0**	**14**	**75**	**75**	**50**	**300**	**0**	**0**	**500**	**14**

C. **Eligibility criteria:** Students enrolled in B. Tech program other than Mechanical Engineering are eligible. The allotment of minor degree Programme will be as per the policy of the Institute.

D. Intake: Minimum 15,

E. Detailed syllabus:

SEMESTER – III

Course Code	**CS1316**						**Course category**			MM1
Course Name	**INTRODUCTION TO ARTIFICIAL INTELLIGENCE**									
Teaching Scheme				Examination Scheme						Credits
				Theory				Practical		
Th	Tu	Pr	Total	CT1	CT2	TA	ESE	ESE Duration	ICA	ESE	Total	
03	00	00	03	15	15	10	60	2hr 30 min	00	00	100	03

Course Objectives:

1. To understand the concept of Artificial Intelligence
2. To apply the methods of solving problems using Artificial Intelligence.
3. To learn the knowledge representation techniques, reasoning techniques and planning
4. To introduce the concepts of Expert Systems and machine learning.
5. To implement the classification and clustering methods for complex problems.

Course Contents:

Introduction: The Foundations of Artificial Intelligence, The History of Artificial Intelligence, The State of the Art, Agents and Environments, Good Behavior: The Concept of Rationality, The Nature of Environments, The Structure of Agents.

Problem Solving: Problem-Solving Agents, Example Problems, searching for Solutions, Uninformed Search Strategies, Informed (Heuristic) Search Strategies, Local Search Algorithms and Optimization Problems, Searching with Nondeterministic Actions

Search Algorithms: Problem solving agents, search algorithms terminologies, properties of search algorithms, types of search algorithms. Uniformed/Blind Search Algorithms: Breadth-first Search, Depth-first Search, Depth-limited Search, Iterative deepening depth-first search, Uniform cost search, Bidirectional Search.

Regression: Linear Regression- Model representation for single variable, single variable Cost Function, Gradient Decent for Linear Regression, Gradient Decent in practice.
Logistic Regression- Classification, Hypothesis Representation, Decision Boundary, Cost function, Advanced Optimization, Multi-classification (One vs All), Problem of Overfitting.

Clustering: clustering as a machine learning task, different types of clustering techniques, partitioning methods, k-medoids, hierarchical clustering. Use-cases centered on classification and clustering.

TEXT BOOKS:

1. Rich & Knight, Artificial Intelligence, second edition, Tata McGraw Hill
2. Stuart Russell, Peter Norvig, "Artificial Intelligence: A Modern Approach", 3rd edition, Prentice Hall, 2009.
3. Machine Learning, Saikar Dutt, Subramanian Chandramouli, Amit Kumar Das, Pearson India

REFERENCES:

1. Practical Workbook Artificial Intelligence and Soft Computing for Beginners, Anindita Das Bhattacharjee, and Shroff Publisher-X team Publisher
2. Machine Learning, Tom Mitchell, McGraw Hill, 2017
3. Pattern Recognition and Machine Learning, Christopher M. Bishop, Springer, 2011
4. The Elements of Statistical Learning, T. Hastie, R. Tibshirani, J. Friedman, 2 nd Edition, 2011

Course Outcomes:

CS1316.1 Illustrate the scope of Artificial Intelligence in the real world
CS1316.2 Demonstrate various machine learning algorithms and its preliminaries
CS1316.3 Summarize and learn various supervised learning algorithms
CS1316.4 Describe and apply the concepts of classification and regression
CS1316.5 Summarize and learn various unsupervised learning algorithms

CO – PO – PSO Mapping:

Course Outcomes	Program Outcomes														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CS1316.1	0	3	3	2	2	2	0	2	3	3	3	1	1	2	3
CS1316.2	1	2	2	1	3	1	2	1	0	1	2	2	1	1	3
CS1316.3	2	1	3	1	0	2	2		2	1	0	2	2	0	0
CS1316.4	3	2	3	1	2	0	1	2	1	3	1	2	0	3	2
CS1316.5	2	3	1	3	1	3	1	0	2	3	3	3	2	0	0

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlate

SEMESTER – IV

Course Code	**CS1416**								**Course category**			MM2
Course Name	**DATA MINING**											
Teaching Scheme				Examination Scheme								Credits
				Theory					Practical		Total	
Th	Tu	Pr	Total	CT 1	CT 2	TA	ESE	ESE Duration	ICA	ESE		
03	00	00	03	15	15	10	60	3hr 00 min	00	00	100	03

Course Objectives:

1. To learn about data mining Concepts.
2. To study the different data mining techniques.
3. To have knowledge in Data mining concepts.
4. To apply Data mining concepts in different fields.
5. To apply association rules to measure the quality of data.

Course Contents:

Basic Data Mining Tasks – Data Mining Versus Knowledge Discovery in Data Bases – Data Mining Issues – Data Mining Matrices – Social Implications of Data Mining – Data Mining from Data Base Perspective.

Data Mining Techniques – a Statistical Perspective on data mining – Similarity Measures – Decision Trees – Neural Networks – Genetic Algorithms.

Classification: Introduction – Statistical – Based Algorithms – Distance Based Algorithms – Decision.

Clustering Tree – Based Algorithms – Neural Network Based Algorithms – Rule Based Algorithms – Combining Techniques: Introduction – Similarity and Distance Measures – Outliers – Hierarchical Algorithms. Partitioned Algorithms.

Association Rules: Introduction - Large Item Sets – Basic Algorithms – Parallel & Distributed Algorithms – Comparing Approaches – Incremental Rules – Advanced Association Rules Techniques – Measuring the Quality of Rules.

TEXT BOOK:

1. Jiawei Han & Micheline Kamber, "Data Mining Concepts & Techniques", 2011, 3 rd Edition.

REFERENCE BOOK:

1. Margaret H.Dunbam, "Data Mining Introductory and Advanced Topics", Pearson Education 2003.

WEB REFERENCES:

NPTEL & MOOC courses titled Data Mining ¬ https://nptel.ac.in/courses/106105174/

CO – PO – PSO Mapping:

Course Outcomes	Program Outcomes														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CS1316.1	0	3	3	2	2	2	0	2	3	3	3	1	1	2	3
CS1316.2	1	2	2	1	3	1	2	1	0	1	2	2	1	1	3
CS1316.3	2	1	3	1	0	2	2		2	1	0	2	2	0	0
CS1316.4	3	2	3	1	2	0	1	2	1	3	1	2	0	3	2
CS1316.5	2	3	1	3	1	3	1	0	2	3	3	3	2	0	0

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlate

SEMESTER - V

Course Code	CS1516								Course category			MM
Course Name	MACHINE LEARNING											
Teaching Scheme				Examination Scheme								Credits
Th	Tu	Pr	Total	Theory					Practical		Total	
				CT1	CT2	TA	ESE	ESE Duration	ICA	ESE		
03	00	00	03	15	15	10	60	2hr 30 min	00	00	100	03

Course Objectives:

Aim of course is

I. To use machine learning library (e.g., Python with Scikit-learn) to build and train models on real data
II. To explore supervised, unsupervised and reinforcement learning paradigms of machine learning.
III. To design and implement machine learning solutions to classification, regression, and clustering problems.
IV. To learn to apply machine learning algorithms on real world problems.

Course Content

Introduction: Machine Learning , Problems Machine Learning Can Solve ,Knowing Your Task and Knowing Your Data , Necessity of Python ,scikit-learn, Installing scikit-learn , Essential Libraries and Tools , Jupyter Notebook ,NumPy, SciPy ,matplotlib , pandas , mglearn , Classifying Iris Species ,Meet the Data, Measuring Success: Training and Testing Data ,Building Your First Model: k-Nearest Neighbors ,Making Predictions , Evaluating the Model

Supervised Learning: Classification and Regression, Generalization, Overfitting, and Underfitting, Relation of Model Complexity to Dataset Size, Supervised Machine Learning Algorithms, Some Sample Datasets, k-Nearest Neighbors, Linear Models, Naive Bayes Classifiers, Decision Trees, Ensembles of Decision Trees, Kernelized Support Vector Machines, Neural Networks (Deep Learning), Uncertainty Estimates from Classifiers, The Decision Function, Predicting Probabilities, Uncertainty in Multiclass Classification

Unsupervised Learning and Preprocessing: Types of Unsupervised Learning, Challenges in Unsupervised Learning, Preprocessing and Scaling, Different Kinds of Preprocessing, Applying Data Transformations, Scaling Training and Test Data the Same Way, The Effect of Preprocessing on Supervised Learning, Dimensionality Reduction, Feature Extraction, and Manifold Learning, Principal Component Analysis (PCA), Clustering, k-Means Clustering, Agglomerative Clustering, Comparing and Evaluating Clustering Algorithm

Representing Data and Engineering Features: Categorical Variables, One-Hot-Encoding (Dummy Variables), Numbers Can Encode Categoricals, Binning, Discretization, Linear Models, and Trees, Interactions and Polynomials, Univariate Nonlinear Transformations, Automatic Feature Selection, Univariate Statistics, Model-Based Feature Selection, Iterative Feature Selection, Utilizing Expert Knowledge

Model Evaluation and Improvement: Cross-Validation, Cross-Validation in scikit-learn, Benefits of Cross-Validation, Stratified k-Fold Cross-Validation and Other Strategies, Grid Search, Simple Grid Search, The Danger of Overfitting the Parameters and the Validation Set , Grid Search with Cross-Validation, Evaluation Metrics and Scoring, Keep the End Goal in Mind, Metrics for Binary Classification, Metrics for Multiclass Classification, Regression Metrics, Using Evaluation Metrics in Model Selection

Text Book:

1. "Introduction to Machine Learning with Python", Andreas C. Müller & Sarah Guido, O'Reilly Media, Incorporated, 2018
2. "Introduction to Statistical Machine Learning with Applications in R", Hastie, Tibshirani, Friedman, Springer, 2nd Edition-2012.

Reference Books:

1. "Introduction to Machine Learning", Ethem Alpaydin PHI 2nd Edition-2013.

Course Outcomes:

After completion of this course student will be able to

CS1516.1 Understand fundamental machine learning concepts, essential libraries and tools
CS1516.2 Identify suitable algorithms for different types of problems.
CS1516.3 Create unsupervised learning models for handling unknown pattern
CS1516.4 Understand the importance of representing data in a way that is suitable for the machine learning algorithm
CS1516.5 Validate and evaluate the performance of machine learning models using metrics like accuracy, precision, recall, F1-score, and understand the limitations of different metrics.

CO – PO – PSO Mapping:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CS1516.1	3	3	3	1	2	0	0	0	0	0	0	2
CS1516.2	3	3	3	1	2	0	0	0	0	0	0	3
CS1516.3	3	3	3	2	2	0	0	0	0	0	0	3
CS1516.4	2	3	3	1	1	0	0	0	0	0	0	2
CS1516.5	3	3	3	2	1	0	0	0	0	0	0	3

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated

CO – PO – PSO Mapping: As per the new NBA Guidelines w.e.f. 01/01/2025

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CS1516.1	3	3	3	1	2	0	0	0	0	0	2
CS1516.2	3	3	3	1	2	0	0	0	0	0	3
CS1516.3	3	3	3	2	2	0	0	0	0	0	3
CS1516.4	2	3	3	1	1	0	0	0	0	0	2
CS1516.5	3	3	3	2	1	0	0	0	0	0	3

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated

SEMESTER -VI

Course Code	CS1616								Course category			MM
Course Name	DEEP LEARNING											
Teaching Scheme				Examination Scheme								Credits
				Theory					Practical		Total	
Th	Tu	Pr	Total	CT1	CT2	TA	ESE	ESE Duration	ICA	ESE		
03	00	00	03	15	15	10	60	2hr 30 min	00	00	100	03

Course Objectives

Aim of the course is

V. To understand the fundamental concepts of machine learning and its applications
VI. To master the concepts of classification and clustering techniques.
VII. To develop a deep understanding of convolutional neural networks (CNNs) and their architecture.
VIII. To apply deep learning techniques to large-scale datasets and real-world problems.

Course Content

Introduction: Neural Network, The Human Brain, Models of a Neuron, Neural Networks Viewed As Directed Graphs, Feedback, Network Architectures, Rosenblatt‟s Perceptron: Introduction, Perceptron, The Perceptron Convergence Theorem, Relation Between the Perceptron and Bayes Classifier for a Gaussian Environment.

Multilayer Perceptrons: Introduction, Batch Learning and On-Line Learning, The Back-Propagation Algorithm, XOR Problem, Heuristics for Making the Back- Propagation Algorithm Perform Better, Back Propagation and Differentiation.

Regularization for Deep Learning: Parameter Norm Penalties - L2 Parameter Regularization, Dataset Augmentation, Semi-Supervised Learning. Optimization for Training Deep Models: Challenges in Neural Network Optimization – Ill Conditioning, Local Minima, Plateaus, Saddle Points and Other Flat Regions.

Convolutional Neural Networks: The operation, Pooling, Convolution and Pooling as an infinitely strong prior, Variants of the basic functions, efficient algorithms, Random or Unsupervised Features, Neuroscientific Basis for Convolutional Networks.

Recurrent Neural Networks: RNN, Bidirectional RNN, Encoder-Decoder Sequence to sequence architecture, Deep Recurrent Networks, Recursive Neural Networks, The Long Short Term Memory and other Gated RNNs, Optimization for Long Term Dependencies.

Text Book:

1. Artificial Intelligence Illuminated - Ben Coppin

2. Deep Learning - Ian Goodfellow, Yoshua Bengio, Aaron Courville

Reference Books:

1. Fundamentals of Deep Learning – Nikhil Budama
2. Neural Networks and Deep Learning – Charu Aggarwal
3. Hands-on Deep Learning Algorithms with Python – Sudharsan Ravichandran

Course Outcomes:

After completion of this course student will be able to

CS1616.1 Demonstrate a comprehensive understanding of deep learning fundamentals
CS1616.2 Illustrate the learning processes and their statistical properties.
CS1616.3 Apply various deep learning architectures to solve complex problems.
CS1616.4 Design deep learning models using regularization and convolutional operations. **CS1616.5** Develop and analyze the applications

CO – PO – PSO Mapping:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CS1616.1	3	3	3	1	2	0	0	0	0	0	0	2
CS1616.2	3	3	3	1	2	0	0	0	0	0	0	2
CS1616.3	3	3	3	2	2	0	0	0	0	0	0	2
CS1616.4	2	3	3	1	1	0	0	0	0	0	0	2
CS1616.5	3	3	3	2	1	0	0	0	0	0	0	2

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated

CO – PO – PSO Mapping: As per the new NBA Guidelines w.e.f. 01/01/2025

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CS1616.1	3	3	3	1	2	0	0	0	0	0	2
CS1616.2	3	3	3	1	2	0	0	0	0	0	2
CS1616.3	3	3	3	2	2	0	0	0	0	0	2
CS1616.4	2	3	3	1	1	0	0	0	0	0	2
CS1616.5	3	3	3	2	1	0	0	0	0	0	2

0- Not correlated 1 - Weakly Correlated 2- Moderately Correlated 3- Strongly Correlated