VARDHAMAN
COLLEGE OF ENGINEERING

CURRICULUM
For
Bachelor of Technology

Computer Science and Engineering(AI&ML) (Honors)

Under

Choice Based Credit System (CBCS)

B. Tech. - Honors Program

(For batches admitted from the Academic Year 2025 - 2026)

August 2025

VARDHAMAN COLLEGE OF ENGINEERING
(Autonomous)
Affiliated to **JNTUH,** Approved by **AICTE,** Accredited by **NAAC** with **A++** Grade
Kacharam, Shamshabad, Hyderabad- 501 218, Telangana, India
www.vardhaman.org, info@vardhaman.org

III B.Tech. I Semester													
#	Course Code	Title of the Course	Category	Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
				CI		LI	TW + SL	H	C	CIE	SEE	Total
				L	T	P	SL					
Theory Courses												
1	H2701	Data Warehousing and Business Intelligence	PC	45	-	-	45	90	3	40	60	100
	H2702	Information Retrieval system										
2	H2703	Ethics and AI	PC	45	-	-	45	90	3	40	60	100
	H2704	Reinforcement Learning & Decision Making										
Total				90	0	0	90	180	6	80	120	200

III B.Tech. II Semester													
#	Course Code	Title of the Course	Category	Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
				CI		LI	TW + SL	H	C	CIE	SEE	Total
				L	T	P	SL					
Theory Courses												
1	H2705	Data Stream Mining	PC	45	-	-	45	90	3	40	60	100
	H2706	Mining Massive Datasets										
2	H2707	Spatial And Multimedia Databases	PC	45	-	-	45	90	3	40	60	100
	H2708	Exploratory Data Analysis										
Total				90	0	0	90	180	6	80	120	200

IV B.Tech. I Semester													
#	Course Code	Title of the Course	Category	Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
				CI		LI	TW + SL	H	C	CIE	SEE	Total
				L	T	P	SL					
Theory Courses												
1	H2709	Text Mining and Time Series Analysis	PC	45	-	-	45	90	3	40	60	100
	H2710	Medical Image Data Processing										
2	H2711	Optimization Techniques in Machine Learning	PC	45	-	-	45	90	3	40	60	100
	H2712	Knowledge Representation and Reasoning										
Total				90	0	0	90	180	6	80	120	200

IV B.Tech. II Semester													
#	Course Code	Title of the Course	Category	Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
				CI		LI	TW + SL	H	C	CIE	SEE	Total
				L	T	P	SL					
Experiential Learning Course													
1	H2041	Technical Report Writing	PW	-	-	-	90	90	2	40	60	100
Total				0	0	0	90	90	2	40	60	100

III B.Tech. I Semester

H2701 Data Warehousing and Business Intelligence

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course covers the fundamentals of Data Warehousing and Business Intelligence, including data warehouse design, OLAP, BI architecture, and performance metrics. It emphasizes data-driven decision-making, explores advanced BI trends like Big Data and mobile BI, and introduces practical tools such as Pentaho and KNIME, along with ethical and integration considerations.

Course Pre/Co-requisites

A9509- Database Management Systems (DBMS)

Course Outcomes

After the completion of the course, the student will be able to:

H2701.1. Understand architecture of data warehouse and OLAP operations.

H2701.2. Understand Fundamental concepts of BI and Analytics.

H2701.3. Application of BI Key Performance indicators

H2701.4. Design of Dashboards, Implementation of Web Analytics.

H2701.5. Understand and implement advanced Business Intelligence tools, techniques, and ethical practices for effective data-driven decision-making.

Course Syllabus

Unit-I:

Introduction:

DATA WAREHOUSE: Data Warehouse-Data Warehouse Architecture- Multidimensional Data Model-Data cube and OLAP Technology-Data Warehouse Implementation -DBMS schemas for Decision support - Efficient methods for Data cube computation.

Unit-II:

Adversarial Search

Business Intelligence: Introduction – Definition, Leveraging Data and Knowledge for BI, BI Components, BI Dimensions, Information Hierarchy, Business Intelligence and Business Analytics. BI Life Cycle. Data for BI - Data Issues and Data Quality for BI.

Unit-III:

Propositional Logic:

BI Implementation - Key Drivers, Key Performance Indicators and Performance Metrics, BI Architecture/Framework, Best Practices, Business Decision Making, Styles of BI-vent-Driven alerts – A cyclic process of Intelligence Creation. The value of Business Intelligence-Value driven & Information use.

Unit-IV:

Planning:

Advanced BI – Big Data and BI, Social Networks, Mobile BI, emerging trends, Description of different BI-Tools (Pentaho, KNIME).

Unit-V:

Reinforcement learning:

Business intelligence implementation-Business Intelligence and integration implementation connecting in BI systems- Issues of legality- Privacy and ethics- Social networking and BI.

Books and Materials

Text Books:

1. JIAWEI HAN & MICHELINE KAMBER, Elsevier. *Data Mining – Concepts and Techniques* , 0^{rd} ed., Pearson Education 2009.
2. Rajiv Sabherwal. *Business Intelligence* , 0^{rd} ed., Wiley Publications, 2012.

Reference Books:

1. Efraim Turban, Ramesh Sharda, Jay Aronson, David King *Decision Support and Business Intelligence Systems* , 9^{th} ed., Pearson Education, 2009.
2. David Loshin *Business Intelligence - The Savy Manager's Guide Getting Onboard with Emerging IT,* 0^{0} ed., Morgan Kaufmann Publishers, 2009.
3. Philo Janus, Stacia Misner *Building Integrated Business Intelligence Solutions with SQL Server,* 0^{0} ed., TMH, 2011.
4. Carlo Verellis *Business Intelligence Data Mining and Optimization for decision making* , 0^{0} ed., Publication: (Wiley).
5. Alex Berson and Stephen J. Smith- Tata McGraw- Hill *Data Warehousing, Data Mining & OLAP* , 10^{th} ed., reprint 2007.
6. W. H. Inmon *Building the Data Warehouse* , 0^{0} ed., Wiley Dreamtech India Pvt. Ltd.
7. MARGARET H DUNHAM *Data Mining Introductory and Advanced topics* , 0^{0} ed., PEA.

H2702- Information Retrieval System

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course offers an overview of Information Retrieval Systems, focusing on techniques for organizing, indexing, and retrieving textual and multimedia data. It covers search algorithms, data structures, clustering, and visualization methods to enhance information access. Students gain insights into modern retrieval approaches, including web and multimedia search, enabling efficient and intelligent information management.

Course Pre/Co-requisites

A9503-Data Structures.

Course Outcomes

After the completion of the course, the student will be able to:

H2702.1. Ability to apply IR principles to locate relevant information large collections of data .

H2702.2. Ability to design different document clustering algorithms.

H2702.3. Implement retrieval systems for web search tasks.

H2702.4. Design an Information Retrieval System for web search tasks.

H2702.5. Ability to evaluate the performance and effectiveness of information retrieval systems using suitable metrics and techniques.

Course Syllabus

Unit-I:

Introduction:

Introduction to Information Retrieval Systems: Definition of Information Retrieval System, Objectives of Information Retrieval Systems, Functional Overview, Relationship to Database Management Systems, Digital Libraries and Data Warehouses Information Retrieval System Capabilities: Search Capabilities, Browse Capabilities, Miscellaneous Capabilities

Unit-II:

Adversarial Search

Cataloging and Indexing: History and Objectives of Indexing, Indexing Process, Automatic Indexing, Information Extraction Data Structure: Introduction to Data Structure, Stemming Algorithms, Inverted File Structure, N-Gram Data Structures, PAT Data Structure, Signature File Structure, Hypertext and XML Data Structures, Hidden Markov Models.

Unit-III:

Propositional Logic:

Automatic Indexing: Classes of Automatic Indexing, Statistical Indexing, Natural Language, Concept Indexing, Hypertext Linkages. Document and Term Clustering: Introduction to Clustering, Thesaurus Generation, Item Clustering, Hierarchy of Clusters.

Unit-IV:

Planning:

User Search Techniques: Search Statements and Binding, Similarity Measures and Ranking, Relevance Feedback, Selective Dissemination of Information Search, Weighted Searches of Boolean Systems, Searching the INTERNET and Hypertext Information Visualization: Introduction to Information Visualization, Cognition and Perception, Information Visualization Technologies.

Unit-V:

Reinforcement learning:

Text Search Algorithms: Introduction to Text Search Techniques, Software Text Search Algorithms, Hardware Text Search Systems Multimedia Information Retrieval: Spoken Language Audio Retrieval, Non-Speech Audio Retrieval, Graph Retrieval, Imagery Retrieval, Video Retrieval.

Books and Materials

Text Books:

1. Gerald J. Kowalski, Mark T. Maybury. *Information Storage and Retrieval Systems – Theory and Implementation* , 2nd ed., Springer.

Reference Books:

1. Frakes, W.B., Ricardo Baeza-Yates *Information Retrieval Data Structures and Algorithms* , Prentice Hall, 1992.
2. Robert Korfhage – John Wiley & Sons *Information Storage & Retrieval.*
3. Yates and Neto Pearson Education. *Modern Information Retrieval.*

H2703 - Ethics and AI

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course provides an overview of the ethical and societal dimensions of Artificial Intelligence, examining its influence on individuals, institutions, and the environment. It explores international ethical frameworks, algorithmic transparency, data privacy, and roboethics, along with real-world case studies in healthcare, automation, and industry. The course aims to develop an understanding of responsible AI design, fostering trust, accountability, and sustainability in technological advancements.

Course Pre/Co-requisites

A9702-Machine Learning

Course Outcomes

After the completion of the course, the student will be able to:

H2703.1. Learn about morality and ethics in AI.

H2703.2. Acquire the knowledge of real time application ethics, issues and its challenges.

H2703.3. Understand the ethical harms and ethical initiatives in AI.

H2703.4. Learn about AI standards and Regulations like AI Agent, Safe Design of Autonomous and Semi-Autonomous Systems.

H2703.5. Understand the concepts of Roboethics and Morality with professional responsibilities.

Course Syllabus

Unit-I:

Introduction:

Definition of morality and ethics in AI, Impact on society, Impact on human psychology, Impact on the legal system, Impact on the environment and the planet, Impact on trust.

Unit-II:

Adversarial Search

International ethical initiatives, Ethical harms and concerns, Case study: healthcare robots, Autonomous Vehicles, Warfare and weaponization.

Unit-III:

Propositional Logic:

Model Process for Addressing Ethical Concerns During System Design, Transparency of Autonomous Systems, Data Privacy Process, Algorithmic Bias Considerations, Ontological Standard for Ethically Driven Robotics and Automation Systems .

Unit-IV:

Planning:

Robot-Roboethics, Ethics and Morality, Moral Theories-Ethics in Science and Technology - Ethical Issues in an ICT Society- Harmonization of Principles- Ethics and Professional Responsibility Roboethics Taxonomy.

Unit-V:

Reinforcement learning:

Challenges Opportunities- ethical issues in artificial intelligence- Societal Issues Concerning the Application of Artificial Intelligence in Medicine- decision-making role in industries-National and International Strategies on AI.

Books and Materials

Text Books:

1. Y. Eleanor Bird, Jasmin Fox-Skelly, Nicola Jenner, Ruth Larbey, Emma Weitkamp and Alan Winfield. *The ethics of artificial intelligence: Issues and initiatives* , EPRS European Parliamentary Research Service Scientific Foresight Unit (STOA) PE 634.452 - March 2020.
2. Patrick Lin, Keith Abney, George A Bekey. *Robot Ethics: The Ethical and Social Implications of Robotics* ,The MIT Press- January 2014.

Reference Books :

1. Paula Boddington *Towards a Code of Ethics for Artificial Intelligence (Artificial Intelligence: Foundations, Theory, and Algorithms)*, November 2017.
2. Mark Coeckelbergh *AI Ethics*, The MIT Press Essential Knowledge series, April 2020.

Web Link:

1. https://sci-hub.mksa.top/10.1007/978-3-540-30301-5 65
2. https://www.scu.edu/ethics/all-about-ethics/artificial-intelligence-and-ethics-sixteen challenges -and opportunities/
3. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/
4. https://sci-hub.mksa.top/10.1159/000492428

H2704 Reinforcement Learning & Decision Making

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

Reinforcement Learning & Decision Making is an advanced course that provides a comprehensive foundation in sequential decision making under uncertainty, focusing on how intelligent agents learn optimal behaviour through interaction with an environment.This course covers core concepts such as agent–environment interaction, reward modelling, return and discounting, Markov Decision Processes, and Bellman equations, and progresses through key solution approaches including dynamic programming, Monte Carlo methods, and temporal-difference learning. Students gain in-depth exposure to policy evaluation, control methods, function approximation, and the distinction between on-policy and off-policy learning. This course further explores modern policy gradient and actor-critic algorithms, including REINFORCE and PPO, and extends to deep reinforcement learning techniques such as DQN, DDPG, A3C, and SAC. Emphasis is placed on practical applications in robotics, game AI, and autonomous systems, along with considerations of exploration strategies, safety, ethics, and fairness in intelligent decision-making.

Course Pre/Co-requisites

A9702-Machine Learning.

Course Outcomes

After the completion of the course, the student will be able to:

H2704.1. Understand the fundamentals of reinforcement learning, including agent environment interaction, types of RL, and solving decision-making problems using Markov Decision Processes and Bellman equations .

H2704.2. Apply dynamic programming and Monte Carlo methods to perform policy evaluation, policy improvement, and control in model-based RL settings.

H2704.3. Implement temporal-difference learning algorithms like TD(0), Sarsa, and Q-learning and extend them using eligibility traces and function approximation techniques.

H2704.4. Develop and analyze policy gradient and actor-critic methods, including REINFORCE and PPO, to optimize policies in continuous and high-dimensional action spaces.

H2704.5. Employ deep reinforcement learning techniques (DQN, DDPG, A3C, SAC) and exploration strategies to solve complex tasks in robotics, games, and autonomous systems, considering safety and ethical decision-making.

Course Syllabus

Unit-I:

Introduction:

Introduction to Reinforcement Learning & MDPs Foundations of RL: Agent-Environment Interaction, Types of RL: Model-based vs. Model-free, Reward Signals, Return, and Discounting, Markov Decision Processes (MDPs), Bellman Equations and Optimality

Unit-II:

Adversarial Search

Dynamic Programming & Monte Carlo Methods: Policy Evaluation and Policy Improvement, Value Iteration and Policy Iteration, Monte Carlo Prediction and Control, First-visit and Every-visit Methods, Limitations of DP and MC Approaches.

Unit-III:

Propositional Logic:

Temporal-Difference Learning & Function Approximation: TD(0), Sarsa, and Q-Learning Algorithms, Eligibility Traces: TD(λ), Sarsa(λ), Off-policy vs. On-policy Learning, Linear Function Approximation, Generalization in RL.

Unit-IV:

Planning:

Policy Gradient Methods and Actor-Critic Algorithms: Policy Gradient Theorem, REINFORCE Algorithm, Baselines and Variance Reduction, Actor-Critic Architectures, Trust Region and Proximal Policy Optimization (PPO).

Unit-V:

Reinforcement learning:

Deep Reinforcement Learning and Applications: Deep Q-Networks (DQN) and Experience Replay, DDPG, A3C, and SAC Algorithms, Exploration Techniques: E-greedy, UCB, Intrinsic Rewards, RL in Robotics, Game AI, and Autonomous Systems, Safety, Ethics, and Fairness in Decision Making.

Books and Materials

Text Books:

1. LanGoodfellow, YoshuaBengio, Aaron Courville. *Deep Learning* , MIT Press.

Reference Books:

1. David Silver's *RL Course Slides & Lectures – DeepMind* , University College London.
2. Marco Wiering & Martijn van Otterlo *Reinforcement Learning: State of the Art*, Springer
3. Morgan & Claypool. *CsabaSzepesvári – Algorithms for Reinforcement Learning, Yuxi Li – Deep Reinforcement Learning: An Overview*, arXiv survey.

III B.Tech. II Semester

H2705 - Data Stream Mining

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course provides an overview of data stream mining, emphasizing methods for learning from continuous and evolving data. It introduces Hoeffding Trees, ensemble models, prediction strategies, and adaptive algorithms for handling concept drift.The course also covers evaluation frameworks, change detection techniques, and practical experimental settings for real-time data streams.

Course Pre/Co-requisites

A9509-Database Management System (DBMS)

Course Outcomes

After the completion of the course, the student will be able to:

H2705.1. Understand how to formulate a knowledge extraction problem from data streams.

H2705.2. Ability to apply methods / algorithms to new data stream analysis problems.

H2705.3. Evaluate the results and understand the functioning of the methods studied.

H2705.4. Demonstrate decision tree and adaptive Hoeffding Tree concepts

H2705.5 Design and implement adaptive data stream mining solutions that handle concept drift using ensembles, sliding windows, and change detection mechanisms.

Course Syllabus

Unit-I:

MOA Stream Mining:

Assumptions, Requirements, Mining Strategies, Change Detection Strategies, MOA Experimental Settings, Previous Evaluation Practices, Evaluation Procedures for Data Streams, Testing Framework, Environments, Data Sources, Generation Speed and Data Size, Evolving Stream Experimental Setting.

Unit-II:

Hoeffding Trees:

The Hoeffding Bound for Tree Induction, The Basic Algorithm, Memory Management, Numeric Attributes, Batch Setting Approaches, Data Stream Approaches.

Unit-III:

Prediction Strategies

Majority Class, Naïve Bayes Leaves, Adaptive Hybrid, Hoeffding Tree Ensembles, Data Stream Setting, Realistic Ensemble Sizes.

Unit-IV:

Evolving Data Streams:

Algorithms for Mining with Change, A Methodology for Adaptive Stream Mining, Optimal Change Detector and Predictor, Adaptive Sliding Windows, Introduction, Maintaining Updated Windows of Varying Length.

Unit-V:

Adaptive Hoeffding Trees:

Introduction, Decision Trees on Sliding Windows, Hoeffding Adaptive Trees, Adaptive Ensemble Methods, New methods of Bagging using trees of different size, New method of bagging using ADWIN, Adaptive Hoeffding Option Trees, Method performance.

Books and Materials

Text Books:

1. Albert Bifet and Richard Kirkby *Data Stream Mining: A Practical Approach* , 9th

Reference Books:

1. Gama João *Knowledge discovery from data streams* , 3rd ed., ISBN: 978-1-4398-2611-9..
2. Albert Bifet, Ricard Gavalda *Information Storage & Retrieval* rd ed., MIT Press, 2017

H2706 - Mining Massive Datasets

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course provides a comprehensive overview of mining massive datasets using scalable data mining techniques and distributed computing frameworks. Students will study algorithms for similarity search, data stream processing, clustering, link analysis, recommendation systems, and social network mining. The course emphasizes practical approaches using MapReduce-based models and real-world case studies from web search, online advertising, recommender systems, and social media analytics.

Course Pre/Co-requisites

A9509- Database Management System(DBMS)

Course Outcomes

After the completion of the course, the student will be able to:

H2706.1. Handle massive data using MapReduce.

H2706.2. Develop and implement algorithms for massive data sets and methodologies in thecontext ofdata mining.

H2706.3. Understand the algorithms for extracting models and information from large datasets.

H2706.4. Develop recommendation systems.

H2706.5. Gain experience in matching various algorithms for particular classes of problems.

Course Syllabus

Unit-I:

Introduction:

Data Mining-Introduction-Definition of Data Mining-Statistical Limits on Data Mining,MapReduce and the New Software Stack-Distributed File Systems, MapReduce, Algorithms UsingMapReduce.

Unit-II:

Adversarial Search

Similarity Search: Finding Similar Items-Applications of Near-Neighbor Search, Shingling ofDocuments, Similarity-Preserving Summaries of Sets, Distance Measures. Streaming Data: MiningData Streams-The Stream Data Model, Sampling Data in a Stream, Filtering Streams.

Unit-III:

Propositional Logic:

Link Analysis-PageRank, Efficient Computation of PageRank, Link Spam. Frequent Itemsets -Handling Larger Datasets in Main Memory, Limited-Pass Algorithms, Counting Frequent Items in aStream. Clustering-The CURE Algorithm, Clustering in Non-Euclidean Spaces, Clustering for Streamsand Parallelism.

Unit-IV:

Planning:

Advertising on the Web-Issues in On-Line Advertising, On-Line Algorithms, The Matching Problem,The Adwords Problem, Adwords Implementation. Recommendation Systems - A Model forRecommendation Systems, Content-Based Recommendations, Collaborative Filtering, DimensionalityReduction, The NetFlix Challenge.

Unit-V:

Reinforcement learning:

Mining Social-Network Graphs-Social Networks as Graphs, Clustering of Social-Network Graphs,Partitioning of Graphs, Simrank, Counting Triangles.

Books and Materials

Text Books:

1. Leskovec, AnandRajaraman, Jeff Ullman. *Mining of Massive Datasets* , 3rd ed.

Reference Books:

1. Jiawei Han & Micheline Kamber *Data Mining – Concepts and Techniques* , 3rd ed., Elsevier.
2. Margaret H Dunham *Data Mining Introductory and Advanced topics*, PEA.
3. Ian H. Witten and Eibe Frank *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann.

H2707 - Spatial and Multimedia Databases

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course provides a comprehensive overview of spatial and multimedia databases, focusing on the storage, indexing, querying, and retrieval of complex data types such as spatial objects, images, text, audio, and video. Students will study spatial data models and indexing structures alongside multimedia database architectures and retrieval techniques. Emphasis is placed on query optimization, distributed multimedia systems, and practical methods for managing and presenting rich multimedia content efficiently.

Course Pre/Co-requisites

A9503-Data Structures A9509-Database Management System (DBMS)

Course Outcomes

After the completion of the course, the student will be able to:

H2707.1. Explain the fundamentals of spatial databases, spatial data models, spatial concepts, and spatial query languages.

H2707.2. Analyze and apply spatial storage methods, indexing techniques, and multidimensional data structures for efficient query processing and optimization.3. Store and retrieve multimedia data.

H2707.3. Design and evaluate image and text/document database systems using similarity-based retrieval, indexing structures, and information retrieval techniques

H2707.4. Describe and implement database techniques for managing video, audio, and multimedia data, including segmentation, indexing, and multimedia query processing.

H2707.5. Develop and assess distributed multimedia database systems and presentations considering temporal, spatial constraints, and distributed media server architectures.

Course Syllabus

Unit-I:

Introduction:

Introduction to Spatial Databases: Overview, beneficiaries, GIA and SDBMS, users, Space taxonomy,query language, query processing, query optimization. Spatial Concepts and Data Models: Models ofSpatial information, three step database design, extending the ER model with spatial concept, objectsorienteddata modeling. Spatial Query Languages.

Unit-II:

Adversarial Search

Spatial Storage and Indexing: Storage-disks and files, spatial indexing, TR*, spatial join index. Queryprocessing and optimization – Evaluation of Spatial operations, query optimization, Analysis of Spatialindex structures, distributed and parallel spatial database system. Multidimensional Data Structures: kdTrees, Point Quad-trees, The MX-Quad-tree, R-Trees, comparison of Different Data Structures.

Unit-III:

Propositional Logic:

Image Databases: Raw Images, Compressed Image Representations, Image Processing:Segmentation, Similarity-Based Retrieval, Alternative Image DB Paradigms, Representing Image DBswith Relations, Representing Image DBs with R-Trees, Retrieving Images By Spatial Layout,Implementations. Text/Document Databases: Precision and Recall, Stop Lists, Word Stems, andFrequency Tables, Latent Semantic Indexing, TV-Trees, Other Retrieval Techniques.

Unit-IV:

Planning:

Video Databases : Organizing Content of a Single Video, Querying Content of Video Libraries, VideoSegmentation, video Standards Audio Databases : A General Model of Audio Data, Capturing AudioContent through Discrete Transformation, Indexing Audio Data Multimedia Databases : Design andArchitecture of a Multimedia Database, Organizing Multimedia Data Based on The Principle ofUniformity, Media Abstractions, Query Languages for Retrieving Multimedia Data, Indexing SMDSs withEnhanced Inverted Indices, Query Relaxation/Expansion.

Unit-V:

Reinforcement learning:

Creating Distributed Multimedia Presentations: Objects in Multimedia Presentations, SpecifyingMultimedia Documents with Temporal Constraints, Efficient Solution of Temporal PresentationConstraints, Spatial Constraints. Distributed Media Servers: Distributed multimedia server architecture,distributed retrieval plans, optimal distributed retrieval plans.

Books and Materials

Text Books:

1. Shashi Shekhar, Sanjiv Chawla. *Spatial Databases-A Tour* , Pearson Education.
2. V.S. Subrahmanian. *Principles of Multimedia Database Systems* , Morgan Kauffman.

Reference Books:

1. Lynne Dunckley *Multimedia Databases: An object relational approach* , Pearson Education.
2. Prabhakaran *Multimedia Database Systems*, Prabhakaran, Springer.

H2708 - Exploratory Data Analysis

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course provides a comprehensive introduction to Exploratory Data Analysis (EDA), data preprocessing, outlier detection, feature selection, and dimensionality reduction. Students will learn practical techniques to summarize, visualize, and clean data, handle missing values, and reduce data dimensionality for better analysis. The course emphasizes applying statistical and computational methods to prepare data for effective data mining, machine learning, and predictive analytics.

Course Pre/Co-requisites

A9702 - Machine Learning.

Course Outcomes

After the completion of the course, the student will be able to:

H2708.1. Handle missing data in the real-world data sets by choosing appropriate methods.

H2708.2. Summarize the data using basic statistics. Visualize the data using basic graphs and plots.

H2708.3. Identify the outliers if any in the data set.

H2708.4. Choose appropriate feature selection and dimensionality reduction.

H2708.5. Techniques for handling multi-dimensional data.

Course Syllabus

Unit-I:

Introduction to Exploratory Data Analysis:

Data Analytics lifecycle, Exploratory Data Analysis (EDA)– Definition, Motivation, Steps in data exploration, The basic data types Data Type Portability.

Unit-II:

Preprocessing - Traditional Methods and Maximum Likelihood Estimation:

Introduction to Missing data, Traditional methods for dealing with missing data, Maximum Likelihood Estimation – Basics, Missing data handling, improving the accuracy of analysis. Preprocessing Bayesian Estimation: Introduction to Bayesian Estimation, Multiple Imputation-Imputation Phase, Analysis and Pooling Phase, Practical Issues in Multiple Imputation, Models for Missing Notation Random Data.

Unit-III:

Data Summarization & Visualization:

Statistical data elaboration, 1-D Statistical data analysis, 2-D Statistical data Analysis, N-D Statistical data analysis.

Unit-IV:

Outlier Analysis:

Introduction, Extreme Value Analysis, Clustering based, Distance Based and Density Based outlier analysis, Outlier Detection in Categorical Data. Feature Subset Selection: Feature selection algorithms: filter methods,

wrapper methods and embedded methods, Forward selection backward elimination, Relief, greedy selection, genetic algorithms for features selection.

Unit-V:

Dimensionality Reduction:

Introduction, Principal Component Analysis (PCA), Kernel PCA, Canonical Correlation Analysis, Factor Analysis, Multidimensional scaling, Correspondence Analysis.

Books and Materials

Text Books:

1. Glenn J. Myatt. *Making sense of Data: A practical Guide to Exploratory Data Analysis and Data Mining* , 9th ed.

Reference Books:

1. Charu C. Aggarwal *1."Data Mining The Text book"*, 3rd ed., Springer, 2015.
2. Craig K. Enders, *"Applied Missing Data Analysis"* rd ed., The Guilford Press, 2010.

IV B.Tech. I Semester

H2709 Text Mining and Time Series Analysis

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course introduces time series similarity measures, representations, and transformations, along with mining techniques such as indexing, classification, clustering, forecasting, anomaly detection, and segmentation. It explores advanced time series analysis methods including frequency- and symbolic-based representations, trend and periodicity analysis, and similarity search.

Course Pre/Co-requisites

A9701-Artificial Intelligence A9702-Machine Learning

Course Outcomes

After the completion of the course, the student will be able to:

H2709.1 Understand the importance of time series and text data.

H2709.2. Enhance analytical skills for mining time series and text data.

H2709.3 Develop skills to leverage sequence patterns.

H2709.4 Create real-life case studies using text data.

H2709.5 Analyse sentiment, opinions, and text streams for real-world applications .

Course Syllabus

Unit-I:

Time Series Similarity Measures:

Euclidean Distance & Lp Norms, Dynamic Time Warping (DTW), Longest Common Subsequence (LCSS) Similarity, Probabilistic Similarity Methods, General Transformations of Time Series, Similarity Search in Time-Series Analysis.

Unit-II:

Time Series Representations:

Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), Singular Value Decomposition (SVD), Piecewise Linear Approximation (PLA), Piecewise Aggregate Approximation (PAA), Adaptive Piecewise Constant, Approximation (APCA), Symbolic Aggregate Approximation (SAX)

Unit-III:

Mining Time Series Data:

: Indexing (Query by Content) Time Series, Classification, Time Series Clustering, Forecasting Techniques, Time Series Summarization, Anomaly Detection Segmentation.

Unit-IV:

Text Mining Fundamentals:

Overview of NLP Techniques & Text Representation, Word Association Mining, Text-Based Similarity Measures, Information Extraction from Text, Named Entity Recognition (NER),Text Classification, Text Clustering.

Unit-V:

Advanced NLP Models:

Contextual Embeddings: ELMo, BERT (overview),Sentence & Document Embeddings. Information Retrieval & Search Document Indexing & Ranking-IDF and Vector Space Model, Search Engine Architecture, Query Expansion & Relevance Feedback

Books and Materials

Text Books:

1. Jiawei Han & Micheline Kamber, *Data Mining Concepts and Techniques*, 2nd ed., Morgan Kaufmann Publishers 2006.
2. Michael Steinbach and Vipin-Kumar, *Introduction to Data Mining* , Pearson Education.

Reference Books:

1. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, *Mining of Massive Datasets,* , 2nd ed., Cambridge University Press & Assessment.
2. Charu C.Aggarwal, ChengXiangZhai, *MiningTextData,* Springer, 2014.

H2710 - Medical Image Data Processing

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

This course provides a comprehensive study of time series analysis and text mining, focusing on similarity measures, transformations, and representation techniques for effective pattern discovery. It covers mining tasks such as indexing, classification, clustering, forecasting, anomaly detection, and segmentation for time-dependent data.

Course Pre/Co-requisites

A9702- Machine Learning

Course Outcomes

After the completion of the course, the student will be able to:

H2710.1 Understand image processing Techniques

H2710.2 Apply restoration and segmentation technique sin medical images.

H2710.3 Understand their presentations of features and classification methods.

H2710.4 Apply Deep learning methods for medical data analysis and Generative models for generate synthetic data.

H2710.5 Use topic modeling and sentiment analysis techniques for real-world text mining applications.

Course Syllabus

Unit-I:

Time Series Similarity Measures & Transformations:

Time Series Similarity Measures, Euclidean Distance and Lp Norms, Dynamic Time Warping (DTW), Longest Common Subsequence (LCSS) Similarity, Probabilistic Similarity Methods,General Transformations of Time Series, Similarity Search in Time-Series Analysis.

Unit-II:

Time Series Representations:

Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), Singular Value Decomposition (SVD), Piecewise Linear Approximation (PLA), Piecewise Aggregate Approximation (PAA), Adaptive Piecewise Constant Approximation (APCA), Symbolic Aggregate Approximation (SAX).

Unit-III:

Mining Time Series Data:

Indexing Time Series (Query by Content), Time Series Classification, Time Series Clustering, Forecasting Techniques, Time Series Summarization, Anomaly Detection, Segmentation of Time Series, Periodicity Analysis for Time-Related Sequence Data, Trend Analysis

Unit-IV:

Text Mining Fundamentals & Similarity:

Overview of Text Mining and NLP Techniques, Text Representation Models, Word Association Mining, Text-Based Similarity Measures, Information Extraction from Text, Named Entity Recognition (NER), Text Classification, Text Clustering, Text Summarization.

Unit-V:

Advanced Text Mining & Applications:

Dimensionality Reduction and Topic Modeling, Latent Semantic Indexing (LSI), Latent Dirichlet Allocation (LDA), Short Text Understanding, Probabilistic Models for Text Mining, Mining Text Streams, Opinion Mining and Sentiment Mining, Text Mining Applications.

Books and Materials

Text Books:

1. Rafael C.Gonzalez, Richard E.Woods. *Digital Image Processing* , 9th

Reference Books:

1. Anil.K.J *"Fundamentals of Digital Image Processing"* , 3rd ed., Prentice Hall of India Pvt. Ltd., New Delhi, 1995.
2. C Bishop *"Pattern Recognition and Machine Learning",* rd ed., Springer,2006.

H2711 - Optimization Techniques in Machine Learning

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

The students will be able to understand and analyze how to deal with changing data. They will also be able to identify and interpret potential unintended effects in your project. They will understand and define procedures to operationalize and maintain your applied machine learning model. .

Course Pre/Co-requisites

A9702-Machine Learning

Course Outcomes

After the completion of the course, the student will be able to:

H2711.1. Analyze how to deal with changing data.

H2711.2. Interpret potential unintended effects in their project.

H2711.3. Comprehend and define procedures to operationalize and maintain the applied machine learning model.

H2711.4. Evaluate how to optimize the use of Machine Learning in real-life problems.

H2711.5. Evaluate the application of machine learning techniques in production and planning.

Course Syllabus

Unit-I:

Model Fitting and Error Measurement:

Optimization Using Gradient Descent, Constrained Optimization and Lagrange Multipliers, Convex Optimization, Data, Models, and Learning, Empirical Risk Minimization, Parameter Estimation, Probabilistic Modelling and Inference Directed Graphical Models.

Unit-II:

Linear Regression as an Optimization Problem:

Problem Setting, Maximum Variance Perspective, Projection Perspective, Eigenvector Computation and Low-Rank Approximations, PCA in High Dimensions, Key Steps of PCA in Practice, Latent Variable Perspective

Unit-III:

Dimensionality Reduction and Optimization:

Inference and proofs, Proof by resolution, Horn clauses and definite clauses.First-Order Logic : Syntax and Semantics of First-Order Logic, Using First Order Logic, Knowledge Engineering in First-Order Logic. Inference in First-Order Logic: Propositional vs. First-Order Inference, Unification, Forward Chaining, Backward Chaining, Resolution.

Unit-IV:

Unsupervised Learning and Expectation Maximization:

Problem Setting, Maximum Variance Perspective, Projection Perspective, Eigenvector Computation and Low-Rank Approximations, PCA in High Dimensions, Key Steps of PCA in Practice, Latent Variable Perspective

Unit-V:

Large-Margin Classification and Dual Optimization:

Separating Hyperplanes, Primal Support Vector Machine, Dual Support Vector Machine, KernelsNumerical Solution.

Books and Materials

Text Books:

1. A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for Machine Learning by Marc Peter Deisenroth*, 1st ed., Cambridge University Press, 2020.

Reference Books:

1. Stephen Boyd and Lieven Vanden berghe *Convex Optimization* , 1st ed., Cambridge University Press, 2004.
1. Patrick Henry Winston *Artificial Intelligence,* 3rd ed., Pearson Education Private Limited, India, 2001.
2. Suvrit Sra, Sebastian Nowozin and Stephen J *Optimization for Machine Learning,* MIT Press, 2011.

H2712 Knowledge Representation and Reasoning

Teaching and Learning Scheme				Hours	Credits	Assessment Marks		
CI		LI	TW+SL	H	C	CIE	SEE	Total
L	T	P	SL					
45	0	0	45	90	3	40	60	100

Course Description

Course Overview

Knowledge representation is the process of encoding knowledge in a format that can be used by computers. Reasoning is the process of using that knowledge to solve problems. There are many different ways to represent knowledge, and many different ways to reason with it.

Course Pre/Co-requisites

A9701- Artificial Intelligence.

Course Outcomes

After the completion of the course, the student will be able to:

H2712.1. Analyze theoretical and practical issues in symbolic knowledge representation and reasoning.

H2712.2. Apply the capabilities of specific knowledge representation formalisms for specific tasks.

H2712.3. Make effective use of techniques specific to specific knowledge representation problems and formalisms.

H2712.4. Make effective use of special-purpose languages for commonsense reasoning, ontologies, planning, reasoning about actions, constraint programming.

H2712.5. Handle uncertain or incomplete knowledge based problems.

Course Syllabus

Unit-I:

The Key Concepts:

Knowledge, Representation, Reasoning, Why knowledge representation and reasoning, Role of logic Logic: Historical background, representing knowledge in logic, Varieties of logic, Name, Type, Measures, Unity Amidst diversity

Unit-II:

Ontology:

Ontological categories, Philosophical background, Top-level categories, describing physical entities, Defining abstractions, Sets, Collections, Types and Categories, Space and Time

Unit-III:

Knowledge Representations:

Knowledge Engineering, Representing structure in frames, Rules and data, Object-oriented systems, Natural language Semantics, Levels of representation

Unit-IV:

Processes:

Times, Events and Situations, Classification of processes, Procedures, Processes and Histories, Concurrent processes, Computation, Constraint satisfaction, Change ConTexts: Syntax of contexts, Semantics of contexts, First-order reasoning in contexts, Modal reasoning in contexts, Encapsulating objects in contexts.

Unit-V:

Knowledge Soup:

Vagueness, Uncertainty, Randomness and Ignorance, Limitations of logic, Fuzzy logic, No monotonic Logic, Theories, Model sand the world, Semiotics Knowledge Acquisition and Sharing: Sharing Ontologies, Conceptual schema, accommodating multiple paradigms, Relating different knowledge representations, Language patterns, Tools for knowledge acquisition

Books and Materials

Text Books:

1. John F. Sowa, Thomson Learning *Knowledge Representation logical, Philosophical, and Computational Foundations* .

Reference Books:

1. Ronald J. Brachman, Hector J. Levesque, Elsevier *Knowledge Representation and Reasoning* .

Vision

To be a pioneer institute and leader in engineering education to address societal needs through education and practice.

Mission

To adopt innovative student centric learning methods.
To enhance professional and entrepreneurial skills through industry institute interaction.
To train the students to meet dynamic needs of the society.
To promote research and continuing education.

Quality Policy

We at Vardhaman College of Engineering, endeavor to uphold excellence in all spheres by adopting the best practices in effort and effect.

VCE
VARDHAMAN
COLLEGE OF ENGINEERING

VARDHAMAN COLLEGE OF ENGINEERING

(Autonomous)

Affiliated to **JNTUH,** Approved by **AICTE,** Accredited by **NAAC** with **A++** Grade
Kacharam, Shamshabad, Hyderabad- 501 218, Telangana, India
www.vardhaman.org, info@vardhaman.org