`

BMS INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE New Delhi)

Yelahanka, Bengaluru 560119

Bachelor of Engineering

Department of Artificial Intelligence & Machine Learning

VII Semester Scheme & Syllabus 2022
Effective from the AY 2025-26

Vision and Mission of the Department

Vision

To emerge as a leading department in AI and ML by preparing skilled, responsible, and eco-friendly professionals who use technology to improve society.

Mission

M1. To equip students with in-depth knowledge in Artificial Intelligence and Machine Learning, built upon a robust foundation in Computer Science and Engineering through industry focused curriculum with practical learning.

M2. To encourage research and innovation through industry partnerships and sustainable technology practices.

M3. To inspire students toward ethical leadership and entrepreneurship through innovative and collaborative student-led activities.

Program Educational Objectives (PEOs)

PEOs	
PEO1	Graduates will thrive as IT professionals, applying AI and ML to build creative and lasting solutions.
PEO2	Graduates will pursue further studies, research, and entrepreneurship, keeping up with technological advancements in various fields.
PEO3	Graduates will demonstrate ethics, integrity, leadership, teamwork, and a dedication to continuous learning in their careers

Program Specific Outcomes (PSOs)

PSOs	
PSO-1	Apply foundational knowledge of Computer Science along with AI and ML techniques to design effective, real-world solutions across diverse application domains.
PSO-2	Develop AI-driven innovations with a commitment to ethics, social responsibility, and sustainability

Scheme of VII Semester

BMS INSTITUTE OF TECHNOLOGY &MANAGEMENT

(Autonomous Institution Affiliated to VTU, Belagavi)

B. E. in Artificial Intelligence & Machine Learning

Scheme of Teaching and Examinations – 2022 Scheme

Outcome-Based Education (OBE) and Choice Based Credit System (CBCS)
(Effective from the academic year 2025-26 onwards)

VII Semester

Sl. No.	Course Category	Course Code	Course Title	Teaching Department (TD) and Question Paper Setting Board (PSB)	Credits Distribution					Examination				Contact Hours/week
					L	T	P	S	Total	CIE Marks	SEE Marks	Total Marks	SEE Duration (H)	
1	IPCC	BAI701	Machine Learning Operations (MLOps)	TD: CSE PSB: CSE/ISE	3	0	2		4	50	50	100	3	5
2	PCC	BAI702	High-performance computing		3	0	0		3	50	50	100	3	3
3	PEC	BAI703X	Professional Elective Course III		3	0	0		3	50	50	100	3	3
4	OEC	BAI704X	Open Elective Course II		3	0	0		3	50	50	100	3	3
5	PW	BAI705	Major Project Phase II		0	0	14		7	100	100	200	3	14
6	PCCL	BAIL706	Data Science Lab		0	0	2		1	50	50	100	3	2
TOTAL									**21**	**350**	**350**	**700**	-	

IPCC: Integrated Professional Core Course, **PCC**: Professional Core Courses, **PCCL**: Professional Core Course laboratory, **UHV**: Universal Human Value Course, **NCMC**: Non Credit Mandatory Course, **ESC**: Engineering Science Course **AEC**: Ability Enhancement Course, **SEC**: Skill Enhancement Course, **L:** Lecture, **T**: Tutorial, **P**: Practical, **CIE**: Continuous Internal Evaluation, **SEE**: Semester End Evaluation.

Integrated Professional Core Course (IPCC): Refers to Professional Core Course Theory Integrated with practical's of the same course. Credit for IPCC can be 04 and its Teaching–Learning hours (L: T: P) can be considered as (3: 0: 2) or (2: 2: 2). The theory part of the IPCC shall be evaluated both by CIE and SEE. The practical part shall be evaluated by only CIE (no SEE). However, questions from the practical part of IPCC shall be included in the SEE question paper.

Professional Elective Courses (PEC): A professional elective (PEC) course is intended to enhance the depth and breadth of educational experience in the Engineering and Technology curriculum. Multidisciplinary courses that are added supplement the latest trend and advanced technology in the selected stream of engineering. Each group will provide an option to select one course. The minimum number of students' strengths for offering a professional elective is 10. However, this condition shall not be applicable to cases where the admission to the program is less than 10.

Professional Elective Course III		**Open Elective Course II**	
Course Code	**Course Name**	**Course Code**	**Course Name**
BAI703A	Application Of AI in IoT	BAI704A	Machine Learning and its Applications
BAI703B	Augmented Reality	BAI704B	Computer Vision
BAI703C	Social Network Analysis	BAI704C	Cloud Computing
BAI703D	Quantum Computing	BAI704D	Data Mining and Data Warehousing

National Service Scheme /Physical Education/Yoga/NCC/Music: All students have to register for any one of the courses namely National Service Scheme (NSS), Physical Education (PE) (Sports and Athletics), and Yoga (YOG) with the concerned coordinator of the course during the first week of III semesters. Activities shall be carried out between III semester to the VI semester (for 4 semesters). Successful completion of the registered course and requisite CIE score is mandatory for the award of the degree. The events shall be appropriately scheduled by the colleges and the same shall be reflected in the calendar prepared for the NSS, PE, and Yoga activities. These courses shall not be considered for vertical progression as well as for the calculation of SGPA and CGPA, but completion of the course is mandatory for the award of degree.

Open Elective Courses (OEC): Students belonging to a particular stream of Engineering and Technology are not entitled to the open electives offered by their parent Department. However, they can opt for an elective offered by other Departments, provided they satisfy the prerequisite condition if any. Registration to open electives shall be documented under the guidance of the Program Coordinator/ Advisor/Mentor.

Selection of an open elective shall not be allowed if,

- The candidate has studied the same course during the previous semesters of the program.
- The syllabus content of open electives is similar to that of the Departmental core courses or professional electives.
- A similar course, under any category, is prescribed in the higher semesters of the program.
- The minimum students' strength for offering open electives is 10. However, this condition shall not be applicable to cases where the admission to the program is less than 10.

Major Project Phase II: The objective of the Project work is (i) To encourage independent learning and the innovative attitude of the students. (ii) To develop interactive attitude, communication skills, organization, time management, and presentation skills. (iii) To impart flexibility and adaptability. (iv) To inspire teamwork. (v) To expand intellectual capacity, credibility, judgment and intuition. (vi) To adhere to punctuality, setting and meeting deadlines. (vii) To instill responsibilities to oneself and others. (viii)To train students to present the topic of project work in a seminar without any fear, face the audience confidently, enhance communication skills, involve in group discussion to present and exchange ideas.

DEPARTMENT OF ARTIFICAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

Machine Learning Operations (MLOps) (3:0:1:0) 4
(Effective from the academic year 2025 -26)

Course Code	**BAI701**	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:2:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:
This course will enable students to:

- Understand the fundamentals of MLOps and set up a collaborative and version-controlled environment for managing machine learning workflows.
- Experiment tracking and pipeline reproducibility.
- Containerize and deploy ML models. Automate workflows with CI/CD.
- Scale, monitor, and automate the entire MLOps lifecycle.

Preamble:
In today's era of artificial intelligence and data-driven decision-making, machine learning (ML) has transitioned from research labs to real-world applications across industries. However, deploying and managing ML models in production at scale poses unique challenges, including reproducibility, scalability, collaboration, and monitoring. MLOps—Machine Learning Operations—emerges as a critical discipline that bridges the gap between data science and operations, ensuring the seamless deployment, management, and maintenance of ML systems.

Module – I

MLOps: What and Why: Defining MLOps and its challenges, MLOPs to Mitigate Risk, MLOPs for Scale

People of MLOps: Subject Matter experts, Data Scientists, Data Engineers, Software Engineers, DevOps, Model Risk Manager / Auditor, Machine Learning Architect
Textbook 1: Chapter 1 & 2 **(08 Hours)**

Module – II

Key MLOPs Features: A Primer on ML, Model Development, Productionalization and Deployment, Monitoring, Iteration and Life Cycle, Governance
Textbook 1: Chapter 3 **(08 Hours)**

Module – III

MLOps: How – Developing Models: What is a ML model?, Data Exploration, Feature Engineering & Selection, Experimentation, Evaluating & Comparing Models, Version management & Responsibility
Preparing for Production: Runtime environments, Model Risk Evaluation, Quality Assurance for ML, Key Testing Considerations, Reproducibility & Auditability, ML Security, Model Risk Mitigation
Textbook 1: Chapter 4 & 5 **(08 Hours)**

Module – IV

Deploying to Production: CI/CD Pipelines, Building ML Artifacts, Deployment Strategies, Containerization, Scaling Deployments, Requirements & Challenges.
Monitoring & Feedback Loop: How Often Should Models be Retrained?, Understanding Model Degradation, Drift Detection in Practice, The Feedback Loop.
Textbook 1: Chapters 6 & 7 **(08 Hours)**

Module – V

MLOPs: Real – World Examples: Consumer Credit Risk Management, Marketing recommendation Engines, Consumption Forecast
Textbook 1: Chapters 9, 10 & 11 **(08 Hours)**

Lab Components

1.Introduction to MLOps:

- Setting up a Python environment with required libraries.
- Introduction to Git for version control.
- Creating a simple ML pipeline using Jupyter Notebook.

2.Building and Deploying a Simple ML Model

- Train a regression/classification model using scikit-learn.
- Save the model using joblib or pickle.
- Create a simple REST API for model inference using Flask or FastAPI.

3.Automating ML Workflows with CI/CD

- Write and automate unit tests for model code using pytest.
- Integrate GitHub Actions for CI/CD.
- Deploy a trained model to a cloud platform (e.g., AWS, GCP, or Azure).

4.Experiment Tracking with MLflow

- Install and configure MLflow.
- Train multiple models and track hyperparameters, metrics, and artifacts.
- Compare model performance using MLflow UI.

5.Model Monitoring and Logging

- Set up basic logging using Python's `logging` library.
- Log prediction metrics like accuracy and latency.
- • Create alerts for performance degradation using Prometheus and Grafana.

Course Outcomes:

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

CO1: Understand MLOps fundamentals and identify the key challenges in implementing MLOps workflows.

CO2: Design, deploy, monitor, and manage the end-to-end life cycle of machine learning models within an MLOps framework, ensuring model performance, governance, and continuous improvement throughout production.

CO3: Develop, evaluate, and prepare machine learning models for production, including data exploration, feature engineering, experimentation, model comparison, version management, and ensuring quality, security, reproducibility, and risk mitigation in production environments.

CO4: Design and implement CI/CD pipelines to automate the deployment of machine learning models.

CO5: Analyze and understand real-world applications of MLOps in domains such as consumer credit risk management, marketing recommendation engines, and consumption forecasting.

Textbooks

1. **"Introducing MLOps – How to Scale machine Learning in the Enterprise"** Mark Treveil & the Dataiku Team, O'Reilly Media Inc., Released November 2020, ISBN: 9781492083290

Reference Books

1. "**Machine Learning in Production: Master the art of delivering robust ML solutions with MLOps**", Suhas Pote, First Edition 2023, BPB Publications, India. ISBN: 978-93-55518-101.
2. **"Engineering MLOps"**, Emmanual Raj, Packt Publishing, 2021, ISBN 978-1-80056-288-2.

DEPARTMENT OF ARTIFICAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

High performance computing (3:0:0:0) 3

(Effective from the academic year 2025 -26)

Course Code:	BAI702	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:0:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:

This course will enable students to:

- Understand the basic concepts of Modern processors and Optimization Technique.
- Explain the different algorithm.
- Introduce Basics of parallelization and OpenMP
- Understand the parallel programming with modern C, C++ and new version of FORTRAN
- Explain the features of Distributed-memory parallel programming with MPI

Preamble: High performance computing (HPC) is the practice of aggregating computing resources to gain performance greater than that of a single workstation, server, or computer. HPC can be run on-premises, in the cloud, or as a hybrid of both.

Module – I

Modern processors: Stored-program computer Architecture, General purpose cache based microprocessor Architecture, Memory hierarchies, Multi core Processors, Multithreaded Processors, Vector processors.
Basic optimization techniques for serial code: Scalar Profiling, Common sense Optimizations, Simple measures, large impact, The role of compilers, C++ optimizations.

Textbook 1: Chapter 1 and 2 (Sections 1.1 to 2.5) **(08 Hours)**

Module – II

Data access optimization: Balance analysis and light speed Estimates-Storage Order, Algorithm classification and access optimizations, The Jacobi algorithm, Algorithm classification and access optimizations, Sparse matrix-vector multiply.
Parallel computers: Taxonomy of parallel computing Paradigms, Shared-memory Computers, Distributed memory computers, Hierarchical systems-Networks.

Textbook 1: Chapter 3 and 4 (Sections 3.1 to 4.5) **(08 Hours)**

Module – III

Basics of parallelization: Why parallelize, Parallelism, Parallel scalability

Shared-memory parallel programming with OpenMP: Short Introduction to OpenMP, Case study OpenMP Jacobi algorithm, Advanced OpenMP: Wavefront parallelization.

Textbook 1: Chapter 5 and 6 (Sections 5.1 to 6.3) **(08 Hours)**

Module – IV

Efficient OpenMP programming: Profiling OpenMP program, Performance pitfalls, Parallel sparse matrixvector multiply.

Locality optimizations on ccNUMA architectures: Locality of access on ccNUMA, ccNUMA optimization of sparse MVM, Placement pitfall, ccNUMA issues with C++.

Textbook 1: Chapter 7 and 8 (Sections 7.1 to 8.4) (08 Hours)

Module – V

Distributed-memory parallel programming with MPI: Message passing, A short introduction to MP,MPI parallelization of a Jacobi solver.

Efficient MPI programming: MPI performance tools, Communication parameters, Synchronization, serialization, contention, Reducing communication overhead, Understanding intranode point-to-point communication.

Textbook 1: Chapter 9 and 10 (Sections 9.1 to 10.5) (08 Hours)

Course outcomes:

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

CO1: To understand the architecture, memory hierarchy and compilers.
CO2: To describe the algorithms in optimization and parallel systems .
CO3: To apply parallelization in shared memory OpenMP.
CO4: To analyze the openMP programs and locality optimization.
CO5: To analyze the MPI and MPI performance tools

Textbooks

1. Georg Hager, Gerhard Wellein "Introduction to High Performance Computing for Scientists and Engineers", CRC Press, 2011.

Reference Books

1. Peter Pacheco-An Introduction to Parallel Programming-Morgan Kaufmann (2011)

2. Michael W. Berry, Kyle A. Gallivan, Efstratios Gallopoulos, Ananth Grama, Bernard Philippe, Yousef Saad, Faisal Saied, "High-performance scientific computing: algorithms and applications", Springer, 2012.

3. Victor Eijkhout, "Introduction to High Performance Scientific Computing", MIT Press, 2011.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Choice Based Credit System (CBCS)

SEMESTER – VII

APPLICATION OF AI IN IOT (3:0:0:0) 3

(Effective from the academic year 2025 -26)

Course Code	**BAI703A**	Course Category	**PEC**
Course Credits	**3**	CIE Marks	**50**
Teaching Hours/Week (L:T:P:S)	**3:0:0**	SEE Marks	**50**
Total Number of Contact Hours	**40**	Exam Hours	**3 Hours**

Course Objectives:

This course will enable students to:

1. Foster a comprehensive understanding of how Artificial Intelligence (AI) enhances the design and operation of Internet of Things (IoT) systems, creating intelligent, autonomous ecosystems.
2. Equip students with the technical competence to design and implement adaptive IoT applications leveraging AI models, efficient architectures, and real-time data analytics.
3. Enable learners to apply AI-driven methodologies and IoT technologies in addressing complex, real-world problems and conducting innovative research in the field.

Preamble:

This course explores the synergy between Artificial Intelligence and IoT. It covers fundamental IoT architectures, sensor interfacing, AI-based data analytics, and real-world smart city use cases. Students will gain hands-on exposure to sensor-actuator interfacing and learn how AI algorithms enhance IoT applications.

Module – I

Introduction to Artificial Intelligence: Definition of AI, Learning systems, Role of AI in IoT, Goals and applications of AI in smart systems.

Aspects of developing intelligent IoT systems: Training data, Concept representation, Role of machine intelligence in enhancing IoT platforms.

Introduction to IoT: Definition of IoT, Characteristics of IoT systems, Physical Design (devices, sensors, actuators), Logical Design (architecture, functional blocks).

Functional blocks of IoT: Sensing layer, Networking layer, Service layer, and Application layer. **Communication models & APIs:** RESTful APIs, MQTT Protocol, Publish-Subscribe and Request-Response communication models.

Textbook 1: Chapters: 1, 2 **(08 Hours)**

Module – II

IoT Data Analytics and Communication Technologies

IoT Analytics: Data flow and processing pipeline in IoT applications.

Challenges in IoT Analytics: Data heterogeneity, real-time processing, latency, and scalability. Data Acquisition: Techniques and protocols for data acquisition from IoT devices. Data Exploration and Pre-processing: Exploratory analysis, Noise reduction, Feature extraction, and transformation.

IoT Technologies and Architecture: End-to-end IoT architecture and enabling technologies. Networking Protocols: MQTT, CoAP, HTTP, LoRaWAN, and NB-IoT. Devices and Gateways: Role of devices, communication modules, and gateways in IoT.

Textbook 1: Chapters 4,5,6 **(08 Hours)**

Module – III

IoT Data Analytics and Communication Technologies

IoT Analytics: Data flow and processing pipeline in IoT applications. Challenges in IoT Analytics: Data heterogeneity, real-time processing, latency, and scalability. Data Acquisition: Techniques and protocols for data acquisition from IoT devices. Data Exploration and Pre-processing: Exploratory analysis, Noise reduction, Feature extraction, and transformation.

IoT Technologies and Architecture: End-to-end IoT architecture and enabling technologies. Networking

Protocols: MQTT, CoAP, HTTP, LoRaWAN, and NB-IoT. Devices and Gateways: Role of devices, communication modules, and gateways in IoT.

Textbook 1: Chapters 7, 8 **(08 Hours)**

Module – IV

Distributed Computing Frameworks for IoT

Distributed Computing Models: Fog computing, Edge computing, and Cloud computing architectures. Cloud Computing: Characteristics of cloud platforms for IoT – scalability, elasticity, and multi-tenancy. Service Models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS). Deployment Models: Public, Private, Community, and Hybrid Cloud.

Networking in Distributed Systems: Role of protocols and web services (REST, SOAP) in IoT-cloud interaction.

Case Study: Comparative study of Fog and Cloud computing in latency-sensitive IoT applications.

Textbook 1: Chapter 9,10 **(08 Hours)**

Module – V

Applications of AI in Smart IoT Systems

Smart City Applications: Smart Energy (predictive consumption), Smart Mobility (traffic forecasting), Smart Citizen services (public safety and waste management), Urban Planning (real-time monitoring). Industrial IoT Applications: Predictive maintenance, quality monitoring, and real-time process optimization. Smart City Data Characteristics: Big data aspects – Volume, Velocity, Veracity in smart environments.

Application of AI Models in IoT: Using Decision Trees, SVM, and KNN on IoT-generated data for decision-making. Case Study Analysis: Real-world examples of AI in Smart City and Industrial IoT applications.

Textbook 2: Chapter 7,8 **(08 Hours)**

Course Outcomes:

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

CO1: Understand the role of Artificial Intelligence in enhancing IoT solutions

CO2: Analyse and design IoT architectures and networking protocols.

CO3: Implement real-world sensor and actuator integrations in IoT platforms.

CO4: Deploy distributed computing frameworks to enable scalable and intelligent IoT solutions.

CO5: Apply AI techniques to address challenges in smart cities and industrial IoT environments.

Textbooks

1. Vijay Madisetti, Arshdeep Bahga, "Internet of Things: A Hands-on Approach," University Press, 2014.
2. J. Watt, R. Borhani, A. K. Katsaggelos, "Machine Learning Refined: Foundations, Algorithms and Applications," Cambridge University Press, 2016.

Reference Books

1. Michael Stanley & Jongmin Lee, "Sensor Analysis for the Internet of Things," Morgan & Claypool Publishers, 2018.
2. Raj Kamal, "Internet of Things: Architecture and Design," McGraw Hill, 2017.
3. Andrew Minteer, "Analytics for the Internet of Things," Packt Publishing, 2017.

e – Resources / MOOCs / Other Web Links related to the course:

1. NPTEL: "Internet of Things," Prof. Sudip Misra, IIT Kharagpur. https://nptel.ac.in/courses/106/105/106105195/
2. Coursera: "AI for Everyone" by Andrew Ng. https://www.coursera.org/learn/ai-for-everyone
3. Stanford Online: "Machine Learning" by Andrew Ng. https://online.stanford.edu/courses/cs229-machine-learning

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

AUGMENTED REALITY (3:0:0:0) 3
(Effective from the academic year 2025 -26)

Course Code	BAI703B	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:0:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:
This course will enable students to:

1. Gain a foundational understanding of Augmented Reality (AR) and Virtual Reality (VR).
2. Acquire essential technical skills needed for AR and VR development, including programming, 3D modelling, and using game development engines like Unity.
3. Understand how AR systems track objects and estimate their positions.
4. Gain knowledge of 3D computer graphics fundamentals, rendering techniques, and the tools and libraries available for 3D modelling and rendering in AR.
5. Learn design principles and interaction techniques specific to AR, including software architecture, design patterns, and user interface design.

Preamble:
Augmented reality is an interactive experience in which a real-world environment is enhanced with computer-generated visual elements, sounds, and other stimuli. It can provide a user with a heightened, more immersive experience than they would experience otherwise, which adds to the user's enjoyment or understanding. virtual Reality (VR) is a computer-generated environment with scenes and objects that appear to be real, making the user feel they are immersed in their surroundings. AR blends virtual content with the real world, enhancing the user's perception of reality in the physical world. VR completely immerses users in a simulated environment, totally disconnecting them from the physical world. The course also focuses on Calibration and Registration, Understanding Unity, Pose Estimation and Tracking, Computer Vision for AR and Designing AR Systems.

Module – 1

Introduction to Augmented Reality: History of AR, AR Scenarios, the future of AR, Applications of AR.
Virtually Everything for Everyone: What is Virtual reality?, Differences between virtual reality and Augmented reality, Applications versus games, Types of VR experiences, types of HMD, How virtual reality works?, Technical Skills that are important to VR.
Textbook 1: Chapter 1 and Textbook 2: Chapter 1 **(08 Hours)**

Module – 2

Understanding Unity, Content and scale: Technical requirements, Installing Unity, Getting started with unity, Creating a simple diorama.
Setting up your project for VR: Introducing the Unity XR platform, Choosing your target VR Platform and toolkits, Enabling virtual reality for your platform, Building and run your project, Building for Oculus Quest
Textbook 2: Chapter 2, Chapter 3 **(08 Hours)**

Module – 3

Pose Estimation and Tracking: Pose Tracking in AR, Classifications of Tracking, Stationary Tracking System, Mobile Sensor-Based Tracking, Optical Tracking, Hybrid Tracking, Marker-Based Tracking and AR, Diminished Reality, Marker-less Tracking and AR.
Textbook 1: Chapter 3 (Sections 3.2 to 3.10) **(08 Hours)**

Module – 4

3D Graphics in AR: Basics of 3D Computer Graphics, 3D Rendering, 3D Model Importers/Loaders, 3D modeling software's, Available Graphics libraries.
Textbook 1: Chapter 5 **(08 Hours)**

Module – 5

Designing AR Systems: Design principles for AR, Designing interactions for AR, Software Architecture and Design Patterns for AR, Designing AR interfaces, Examples of AR Interfaces.
Textbook 1: Chapter 6 **(08 Hours)**

Course Outcomes:

The students will be able to: **(List the COs as per the course requirements)**

CO1: Understand the AR/VR applications and usage, HMDs, and necessary technical skills.
CO2: Demonstrate understanding of technical requirements and VR development basics with oculus integration
CO3: Analyze various pose tracking techniques, 3D graphics and design interactions and principles in AR.
CO4: Build and deploy AR and VR projects on target platforms, such as Oculus Quest, using appropriate toolkits and development practices

Textbooks

1. Chetankumar G Shetty, "Augmented Reality: Theory, Design and Development", McGrawHill Publications 2020.
2. Jonathan Linowes -"Unity 2020 Virtual Reality Projects" Third Edition Packt> Paperback – 30 July 2020

Reference Books

1. Jonathan Linowes, Krystian Babilinski, "Augmented Reality for Developers: Build practical augmented reality applications with Unity, ARCore, ARKit and Vuforia", Paperback – Import,Packt Publishing Limited, 9 October 2017.
2. Schmalstieg/Hollerer, "Augmented Reality: Principles & Practice", Paperback–12, Pearson Education India, October 2016.
3. Chitra Lele, "Artificial Intelligence Meets Augmented Reality: Redefining Regular Reality", Paperback – 1, BPB Publications, January 2019.

Alternate Assessment Tools (AATs) suggested:

- **Design and development of AR/VR model / Case study report**
- **MOOC course**

Web links / e – resources:

- NPTEL Course on Virtual Reality by Prof Steven LaValle, IIT Madras,https://nptel.ac.in/courses/106/106/106106138/
- NPTEL Course on Virtual Reality Engineering by Prof.M Manivannan, IIT Madras, https://nptel.ac.in/courses/121/106/121106013/
- NPTEL Course on Introduction to Computer Graphics by Prof. Prem K Kalra, IIT Delhi, https://nptel.ac.in/courses/106/102/106102065/
- NPTEL Course on Computer Graphics by Prof. Sukhendu Das, IIT Madras, https://nptel.ac.in/courses/106/106/106106090/
- NPTEL Course on Computer Graphics by Prof. Samit Bhattacharya, IIT Guwahati, https://nptel.ac.in/courses/106/103/106103224/

DEPARTMENT OF ARTIFICAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

SOCIAL NETWORK ANALYSIS (3:0:0:0) 3
(Effective from the academic year 2025 -26)

Course Code	BAI703C	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:0:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:
This course will enable students to:

- Understand the basics of Social Network levels and network measures.
- Distinguish various network growth models and rank methods
- Evaluate community structures and link prediction models.
- Find out the cascade actions in networks.
- Discover the usage of graph representation learning with different case studies.

Preamble: This course will focus on the analysis of massive networks which provide many computational, algorithmic, and modeling challenges. The course will cover research on the structure and analysis of such large networks and on models and algorithms that abstract their basic properties. This course also provides the insight on how to practically analyze large-scale network data and through models for network structure and evolution. Topics covered in this course are how information spreads through society; robustness and fragility of networks; algorithms for the World Wide Web; prediction and recommendation in online social networks; representation learning for large networks; etc.

Module – I

Introduction: Introduction, Applications,Preliminaries, Three Levels of SNA, Historical Development, Graph Visualization Tools
Network Measures: Network Basics, Node Centrality, Assortativity, Transitive and Reciprocity, Similarity, Degeneracy.
Textbook 1: Chapter 1,2 **(08 Hours)**

Module – II

Network Growth Models: Properties of real-world networks, Random network model, Ring Lattice Network Model, Watts Strogatz model. Preferential Attachment model, Price's model, Local -world network growth model,
Link Analysis: Applications, Signed networks, Strong and weak Ties, Link analysis and algorithms, Page Rank, Personalized Page Rank, Divrank, Simrank, PathSim.
Textbook 1: Chapter 3,4 **(08 Hours)**

Module – III

Community Structure in Networks: Applications, Types of communities, Community detection methods, Disjoint community detection, overlapping community detection, local community detection, community detection vs community search, evaluation.
Link Prediction: Applications, temporal changes in a network, Problem definition, Evaluating link prediction networks, Heuristic Models, probabilistic Models, Supervised Random walk, Information-theoretic model.
Textbook 1: Chapter 5,6 **(08 Hours)**

Module – IV

Cascade Behavior and Network Effects: Preliminaries, Cascade model, case study, Probabilistic cascades, epidemic models, independent cascade models, Cascade prediction.

Anomaly Detection in Static Networks: Outliers vs Network-based anomalies, challenges, Anomaly detection in static and dynamic networks

Textbook 1: Chapter 7,8 **(08 Hours)**

Module – V

Graph Representation Learning: Machine learning pipelines, Intuition behind representation learning, benefits, criteria of GRL, GRL pipelines, representation learning methods.

Application and Case Studies: Malicious activities on OSNs, Sock puppets in OSNs, Modeling the spread of COVID 19, Recommendation System.

Textbook 1: Chapter 9,10 **(08 Hours)**

Course outcomes:

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

CO1: Describe the levels of SNA and network measures.
CO2: Outline various network growth models and different rank methods.
CO3: Experiment With different community structures and link prediction models.
CO4: Evaluate cascade behaviour in networks.
CO5: Analyse the usage of graph representation learning through case studies.

Textbooks

1. Social Network Analysis, Tanmoy Chakraborty, Wiley, 2021

Reference Books

1. Network Science, Albert-Lazzlo Barabasi
2. Social Network Analysis: Methods and Appl ications, Stanley Wasserman, Katherine Faus

DEPARTMENT OF ARTIFICAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

Quantum Computing (3:0:0:0) 3
(Effective from the academic year 2025 -26)

Course Code	BAI703D	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:0:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:
This course will enable students to:

- Understand the fundamental concepts of quantum computing.
- Learn the principles of quantum mechanics as they apply to quantum computing.
- Develop the ability to implement key quantum algorithms.
- Gain practical skills in quantum programming using tools like Qiskit.
- Explore advanced quantum protocols and their implications.
-

Preamble: The field of quantum computing merges principles of quantum mechanics with information science, promising to revolutionize computation and communication. This syllabus is designed to provide a comprehensive introduction to quantum computing, starting with fundamental concepts and progressing to advanced topics, including quantum programming and protocols. Through these modules, students will gain a solid foundation in both theoretical and practical aspects of quantum computing, preparing them for further research or careers in this emerging field.

Module – I

Introduction and Overview: Global perspectives, Quantum bits, Quantum computation: Single qubit gates, Multiple qubit gates, Measurements in bases other than the computational basis, Quantum circuits, Qubit copying circuit?, Example: Bell states, Example: quantum teleportation, Quantum algorithms: Classical computations on a quantum computer, Quantum parallelism, Deutsch's algorithm, The Deutsch–Jozsa algorithm, Quantum algorithms summarized.

Textbook 1: Chapter 1 **(08 Hours)**

Module – II

Introduction to Quantum Mechanics: Linear algebra, Bases and linear independence, Linear operators and matrices, The Pauli matrices, Inner products, Eigenvectors and eigenvalues, Adjoints and Hermitian operators, Tensor products, Operator functions, The commutator and anti-commutator, The polar and singular value decompositions.
The Postulates of Quantum Mechanics: State space, Evolution, Quantum measurement, Distinguishing quantum states, Projective measurements, POVM measurements, Phase, Composite systems, Quantum mechanics: a global view.
Application: superdense coding

Textbook 1: Chapter 2 **(08 Hours)**

Module – III

Basic Linear Algebra: Hilbert Spaces, Products and Tensor Products, Matrices, Complex Spaces and Inner Products, Matrices, Graphs and Sums Over Paths.

Quantum Algorithms Overview: Phils Algorithm: The Algorithm, The Analysis, An Example, A Two-Qubit Example, Phil Measure Up, Quantum Mazes versus Circuits versus Matrices, Problems, Summary and Notes. Deutch's Algorithm: The Algorithm, The Analysis, Superdense Coding and Teleportation, Problems, Summary Notes, Deutch-Jozsa Algorithm: The Algorithm, The Analysis, Problems, Summary and Notes**.**

Textbook 2: Chapter 3, 7, 8 & 9 **(08 Hours)**

Module – IV

Quantum Programming: IBM Quantum: Services, Quantum Composer, Quantum Processor, Simulator, Quantum Assembly Language: OpenQASM, Quantum Experience Standard Header, Open QASM in IBM Quantum, Quantum Adder, Qiskit: Quantum Composer, Quantum Lab, Simulator, Quantum Processor, Other Quantum Programming Langauages

Textbook 3: Chapter 5 **(08 Hours)**

Module – V

Entanglement and Quantum Protocols: Measurements: Product States, Maximally Entangled States, Partially Entangled States, Bell Inequalities: EPR Paradox and Local Hidden Variables, Bell Inequalities and the CHSH Inequality, Quantum, Processor, Experiment, Other, Experiments, No-Signaling Principle, Other, Theories, Monogamy of Entanglement: Classical, Correlations, Quantum Entanglement, Superdense Coding: The Problem, Classical Solution, Quantum Solution, Quantum Teleportation.

Textbook 3: Chapter 6 **(08 Hours)**

Course Outcomes:

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

CO1: Demonstrate a solid understanding of quantum computing basics.
CO2: Apply linear algebra and quantum mechanics principles to quantum computing problems.
CO3: Implement and analyze quantum algorithms effectively.
CO4: Design and simulate quantum circuits using quantum programming tools.
CO5: Understand and explain advanced topics in quantum entanglement and quantum protocols.

Textbooks

1. "Quantum Computation and Quantum Information" by Michael A. Nielsen and Isaac L. Chuang
2. "Quantum Algorithms via Linear Algebra: A Primer" by Richard J. Lipton and Kenneth W. Regan
3. "Introduction to Classical and Quantum Computing" by Thomas Wong

Reference Books

1. "Quantum Computing for Computer Scientists" by Noson S. Yanofsky and Mirco A. Mannucci
2. "An Introduction to Quantum Computing" by Phillip Kaye, Raymond Laflamme, and Michele Mosca

DEPARTMENT OF ARTIFICAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

Machine Learning and its Applications (3:0:0:0) 3
(Effective from the academic year 2025 -26)

Course Code	**BAI704A**	CIE Marks	**50**
Teaching Hours/Week (L:T:P:S)	**3:0:0:0**	SEE Marks	**50**
Total Number of Contact Hours	**40**	Exam Hours	**3 Hours**

Course Objectives:
This course will enable students to:

- Provide an understanding of how machine learning is transforming the healthcare industry.
- Analyze key technological advancements and ethical considerations in retail ML applications.
- Explore the impact of ML on the finance sector.
- Understand the lifecycle of ML models and the challenges of deploying them in production environments.
- Explore post-production activities, including model security, A/B testing, and future trends in MLOps.

Preamble:
As we stand on the cusp of a technological revolution driven by artificial intelligence and machine learning, this course aims to equip the learners with the foundational knowledge necessary to harness the power of these transformative technologies. In this course, learners will explore the machine learning applications in sectors like Healthcare, Retail & Finance. However, understanding machine learning models is only part of the equation. To truly leverage these technologies in real-world applications, it is essential to understand the principles and practices of MLOps (Machine Learning Operations). MLOps is a set of practices that aims to deploy and maintain machine learning models in production reliably and efficiently. The course covers key concepts such as model deployment, monitoring, scaling, and lifecycle management, ensuring that learners are well-versed in the end-to-end process of bringing machine learning models from development to deployment.

Module – I

Overview of Machine Learning in Healthcare – How ML is transforming Healthcare.

Key Technological advancements in Healthcare – Narrow vs. Broad ML, Importance of ML in Healthcare.

Case Studies in Healthcare – Lab Coordinator Problem, Hospital Food Wastage Problem

Textbook 1: Chapters 1, 2 & 4 **(08 Hours)**

Module – II

Overview of Machine Learning in Retail – Retail Segments, Retail Value Proposition, The Process of Technology Adoption in the Retail Sector, The current state of Analytics in the Retail Sector.

Key Technological Advancements in Retail – Narrow vs. Broad ML in Retail, Importance of ML in Retail, Research Design Overview: Data Collection Methods, Data Analysis, Ethical Considerations.

Case Studies in Retail AI – Recommendation Engine Creation for Online Retail Mart.

Textbook 1: Chapters 7, 8 &10 **(10 Hours)**

Module – III

Overview of Machine Learning in Finance – Financial Segments, Finance value Proposition, The Process of Technology Adoption in the Finance Sector.

Key Technological Advancements in Finance – Narrow vs. Broad ML in Finance, Importance of ML in Finance, Research Design Overview: Data Collection Methods, Data Analysis, Ethical Considerations.

Case Studies in Finance AI – Stock Market Movement Prediction, Detecting Financial Statements Fraud
Textbook 1: Chapters 13, 14 & 16 **(10 Hours)**

Module – IV

Challenges in ML Model Deployment – ML Life Cycle, Types of model deployment, Challenges in deploying models in the production environment, MLOps, Benefits of MLOps.

CI / CD for ML – CI / CD pipeline for ML, Continuous Integration (CI), Continuous Delivery / Deployment (CD), Continuous Training (CT).
Textbook 2: Chapters 3 & 9 **(06 Hours)**

Module – V

Monitoring & Debugging – Importance of Monitoring, Fundamentals of ML monitoring, Metrics for monitoring your ML system, Drift in ML.

Post – Productionizing ML models – Bridging the gap between the ML model and the creation of business value, Model security, A/B Testing, MLOps in the future.

Textbook 2: Chapters 14 & 15 **(06 Hours)**

Course Outcomes:
At the end of the course the student will be able to:
CO1: Analyze practical healthcare case studies to understand real-world applications of ML.
CO2: Describe the different retail segments and the value proposition of ML in retail.
CO3: Identify the financial segments and apply ML techniques to solve financial problems.
CO4: Identify challenges in deploying ML models in production and understand the benefits of MLOps.
CO5: Apply post – production techniques to enhance business value, ensure model security, and conduct testing.

Textbooks

1. **"Machine Learning Applications using Python: Case Studies from Healthcare, Retail and Finance"** Puneet Mathur, Apress, 2019, ISBN-13 (pbk): 978-1-4842-3786-1, ISBN-13 (electronic): 978-1-4842-3787-8.
2. **"Machine Learning in Production: Master the art of delivering robust ML solutions with MLOps",** Suhas Pote, First Edition 2023, BPB Publications, India. ISBN: 978-93-55518-101.

Reference Books

1. **"Engineering MLOps"**, Emmanual Raj, Packt Publishing, 2021, ISBN 978-1-80056-288-2.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

Computer Vision (3:0:0:0) 3
(Effective from the academic year 2025 -26)

Course Code:	BAI704B	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:0:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:
This course will enable students to:

- Learn basic principles of image formation, image processing algorithms and different algorithms for recognition from single and multiple images
- Understand the core vision tasks - Learn methods for segmenting images into meaningful regions.
- Understand the basic concepts of deep learning and its significance in the field of computer vision
-

Preamble: Computer Vision(CV) is a field of Artificial Intelligence (AI) that enables computers to interpret and make decisions based on visual data from the world. This course will introduce students to the fundamental principles and techniques in computer vision, including image processing, feature extraction. The field of deep learning has brought about significant advancements in computer vision, enabling machines to understand and interpret visual data with unprecedented accuracy.

Module – I

Introduction and Image formation: What is computer vision? A brief history, Geometric primitives and transformations, Photometric image formation, The digital Camera, The Pinhole Perspective, Weak Perspective, Cameras with lenses, The Human Eye, Intrinsic and extrinsic parameters, Geometric Camera Calibration.

Textbook 1: Chapter 1 - 1.1 & 1.2, Chapter 2 - 2.1 to 2.3
Textbook 2: Chapter 1-1.1 to 1.3 **(08 Hours)**

Module – II

Early Vision: Linear filters and convolution, Shift invariant linear systems, Spatial frequency and Fourier transforms, Sampling and Aliasing, Filters as Templates, Local Image Feature: Computing and representing the image gradient.

Textbook 2: Chapter 4.1-4.5, 5.1-5.2 **(08 Hours)**

Module – III

Mid level Vision: Segmentation by clustering: Important Applications, Image segmentation by clustering Pixels, Segmentation, Clustering and graphs. Grouping and model fitting: The hough transform, Fitting lines and planes.

Textbook 2: Chapter 9.1-9.4, 10.1-10.2 **(08 Hours)**

Module – IV

Why Deep Learning: What is AI and Deep Learning**,** Motivation of deep architecture, Applications, Future potential and challenges.

Getting yourself ready for Deep learning: Deep learning with GPU, Deep learning software frameworks, setting up deep learning on AWS.

Textbook 3: Chapter 1 and 2 **(08 Hours)**

Module – V

Deep learning in Computer Vision: Origin of CNNs, Fine-tuning CNNs, Popular CNN architectures.

Textbook 3: Chapter 4 **(08 Hours)**

Course outcomes :

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

CO1: Implement fundamental image processing techniques required for computer vision.
CO2: Learn methods for segmenting images into meaningful regions.
CO3: Understand image formation process**.**
CO4: Demonstrate a solid understanding of deep learning fundamentals and its significance in computer vision.

Textbooks

1. Computer Vision: Algorithms and Applications (CVAA), Richard Szeliski, Springer 2nd edition, 2020.
2. Computer Vision - A modern Approach, by D. Forsyth and J.Ponce, Prentice Hall, 2nd edition, 2012
3. Wei Di, Anurag Bhardwaj, Jianing Wei, " Deep Learning Essentials" , Packt Publisher, 2018

Reference Books

1. Ben Shneiderman, Designing the User Interface: Strategies for Effective Human-Computer Interaction, Pearson, 5th Edition.
2. Donald A. Norman, The Design of Everyday Things, Basic Books, Revised Edition.
3. Neural Networks and Deep learning, Charu C. Aggarwal, Springer International Publishing, 2018

DEPARTMENT OF ARTIFICAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

Cloud Computing **(3:0:0:0) 3**
(Effective from the academic year 2025 -26)

Course Code:	BAI704C	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:0:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:
This course will enable students to:

- Understand the fundamentals of cloud computing.
- Describe the Architecture of cloud computing
- Illustrate the cloud application programming and aneka platform
- Contrast different cloud platforms used in industry.

Preamble: Cloud computing can also be thought of as utility computing or on-demand computing. The name cloud computing was inspired by the cloud symbol that's often used to represent the internet in flowcharts and diagrams. Cloud computing is the on-demand delivery of computing services such as servers, storage, databases, networking, software, and analytics. Rather than keeping files on a proprietary hard drive or local storage device, cloud-based storage makes it possible to save remotely.

Module – I

Introduction ,Cloud Computing at a Glance, The Vision of Cloud Computing, Defining a Cloud, A Closer Look, Cloud Computing Reference Model, Characteristics and Benefits, Challenges Ahead, Historical Developments, Distributed Systems, Virtualization, Web 2.0, Service-Oriented Computing, Utility-Oriented Computing. Computing platforms and technologies

Virtualization, Introduction, Characteristics of Virtualized, Environments Taxonomy of Virtualization Techniques, Execution Virtualization, Other Types of Virtualization, Virtualization and Cloud Computing, Pros and Cons of Virtualization, Technology Examples Xen: Paravirtualization, VMware: Full Virtualization, Microsoft Hyper-V

Textbook 1: Chapter 1,3 **(08 Hours)**

Module – II

Cloud Computing Architecture, Introduction, Cloud Reference Model, Architecture, Infrastructure / Hardware as a Service, Platform as a Service, Software as a Service, Types of Clouds, Public Clouds, Private Clouds, Hybrid Clouds, Community Clouds, Economics of the Cloud, Open Challenges, Cloud Definition, Cloud Interoperability and Standards Scalability and Fault Tolerance Security, Trust, and Privacy Organizational Aspects.
Aneka: Cloud Application Platform, Framework Overview, Anatomy of the Aneka Container, From the Ground Up: Platform Abstraction Layer, Fabric Services, foundation Services, Application Services.

Textbook 1: Chapter 4,5 **(08 Hours)**

Module – III

Concurrent Computing: Thread Programming, Introducing Parallelism for Single Machine Computation, Programming Applications with Threads, What is a Thread? , Thread APIs, Introducing the Thread Programming Model, Aneka Thread vs. Common Threads.

High-Throughput Computing: Task Programming, Task Computing, characterizing a Task, Computing Categories, Frameworks for Task Computing, Task-based Application Models, Embarrassingly Parallel Applications, Parameter Sweep Applications, MPI Applications, Workflow Applications with Task Dependencies.

Textbook 1: Chapter 6,7 **(08 Hours)**

Module – IV

Data Intensive Computing: Map-Reduce Programming, What is Data-Intensive Computing?, Characterizin Data-Intensive Computations, Challenges Ahead, Historical Perspective, Technologies for Data-Intensiv Computing, Storage Systems, Programming Platforms, Aneka MapReduce Programming, Introducing th MapReduce Programming Model, Example Application

Textbook 1: Chapter 8 **(08 Hours)**

Module – V

Cloud Platforms in Industry: Amazon Web Services, Compute Services, Storage Services, Communication Services, Additional Services, Google AppEngine, Architecture and Core Concepts, Application Life-Cycle, Cost Model, Observations, Microsoft Azure, Azure Core Concepts, SQL Azure, Windows Azure Platform Appliance.

Cloud Applications Scientific Applications, Healthcare: ECG Analysis in the Cloud, Biology: Protein Structure Prediction, Biology: Gene Expression Data Analysis for Cancer Diagnosis, Geoscience: Satellite Image Processing, Business and Consumer Applications, CRM and ERP, Productivity, Social Networking, Media Applications, Multiplayer Online Gaming.

Textbook 1: Chapter 9,10 **(08 Hours)**

Course outcomes:

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

CO1: To Understand the fundamentals of cloud computing , virtualization

CO2: Explain the services offered by the of cloud computing

CO3: Illustrate architecture and programming in cloud

CO4: Describe the platforms for development of cloud applications and List the. application of cloud

Textbooks

1. Rajkumar Buyya, Christian Vecchiola, and Thamarai Selvi Mastering Cloud. Computing McGraw Hill Education, 2013.

Reference Books

1. Dan C. Marinescu, Cloud Computing Theory and Practice, Morgan Kaufmann, Elsevier 2013.

.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING
Choice Based Credit System (CBCS)
SEMESTER – VII

Data Mining and Data Warehousing (3:0:0:0) 3
(Effective from the academic year 2025 -26)

Course Code	BAI704D	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	3:0:0:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:
This course will enable students:

- To understand the principles of Data warehousing and Data Mining.
- To be familiar with the Data warehouse architecture and its Implementation.
- To know the Architecture of a Data Mining system.
- To understand the various Data preprocessing Methods.
- To perform classification and prediction of data.

Preamble: The rapid growth of data in various domains has made it essential to extract meaningful insights from vast datasets. This course introduces students to the foundational concepts of Data Warehousing and Data Mining, focusing on their architectures, functionalities, and applications. It equips learners with techniques for data preprocessing, classification, clustering, and association rule mining. Additionally, the course explores advanced topics such as mining spatial, multimedia, and web data, preparing students to handle real-world data analytics challenges effectively.

Module – I

Data Warehousing and Business Analysis: - Data warehousing Components –Building a Data warehouse – Data Warehouse Architecture – DBMS Schemas for Decision Support – Data Extraction, Cleanup, and Transformation Tools –Metadata – reporting – Query tools and Applications – Online Analytical Processing (OLAP) – OLAP and Multidimensional Data Analysis.

Text Book 2: Chapter 8.2.3,8.4,9.6,9.9 Text Book 1: 4.1.4 ,4.1.6,4.2,4.3.4 (08 Hours)

Module – II

Data Mining: - Data Mining Functionalities – Data Preprocessing – Data Cleaning – Data Integration and Transformation – Data Reduction – Data Discretization and Concept Hierarchy Generation- Architecture of A Typical Data Mining Systems- Classification Of Data Mining Systems.

Association Rule Mining: - Efficient and Scalable Frequent Item Set Mining Methods – Mining Various Kinds of Association Rules – Association Mining to Correlation Analysis – Constraint-Based Association Mining.

Text Book 1: Chapter 1.4 ,3.1,3.2,3.3,3.4,3.5,3.5.6,4.1.4, 6.1.2,6.2,6.3.2 (08 Hours)

Module – III

Classification and Prediction: - Issues Regarding Classification and Prediction – Classification by Decision Tree Introduction – Bayesian Classification – Rule Based Classification – Classification by Backpropagation – Support Vector Machines – Associative Classification – Lazy Learners – Other Classification Methods – Prediction – Accuracy and Error Measures – Evaluating the Accuracy of a Classifier or Predictor – Ensemble Methods – Model Section.

Text Book 2: 4.2,4.3,4.5,4.8,4.10 Textbook 1: 9.2,9.3,9.4,9.5,8.6 (08 Hours)

Module – IV

Cluster Analysis: - Types of Data in Cluster Analysis – A Categorization of Major Clustering Methods – Partitioning Methods – Hierarchical Methods – Density-Based Methods – Grid-Based Methods – Model-Based Clustering Methods – Clustering High-Dimensional Data – Constraint-Based Cluster Analysis – Outlier Analysis.

Text Book 1:10.110.2,10.3, 10.4,10.5,11.1,11.2 **(08 Hours)**

Module – V

Mining Object, Spatial, Multimedia, Text, and Web Data: Multidimensional Analysis and Descriptive Mining of Complex Data Objects – Spatial Data Mining – Multimedia Data Mining – Text Mining – Mining the World Wide Web.

Text Book 1: 13.1,13.1.2,13.1.3 **(08 Hours)**

Course outcomes:

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

CO1: Understand the principles of Data warehousing and Data Mining.
CO2: Understand the Association rule mining concept.
CO3: Explain the concept of classification and prediction.
CO4: Categorize various clustering methods
CO5: Understand the Multidimensional Analysis and Descriptive Mining of Complex Data Objects

Textbooks

1. Jiawei Han, Micheline Kamber and Jian Pei"Data Mining Concepts and Techniques", Third Edition, Elsevier, 2011.
2. Alex Berson and Stephen J. Smith "Data Warehousing, Data Mining & OLAP", Tata McGraw – Hill Edition, Tenth Reprint 2007.

Reference Books

1. K.P. Soman, Shyam Diwakar and V. Ajay "Insight into Data mining Theory and Practice", Easter Economy Edition, Prentice Hall of India, 2006.
2. G. K. Gupta "Introduction to Data Mining with Case Studies", Easter Economy Edition, Prentice Hall of India, 2006.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Choice Based Credit System (CBCS)

SEMESTER – VII

Data Science Lab (0:0:1:0) 1

(Effective from the academic year 2025 -26)

Course Code	BAIL706	CIE Marks	50
Teaching Hours/Week (L:T:P:S)	0:0:1:0	SEE Marks	50
Total Number of Contact Hours	40	Exam Hours	3 Hours

Course Objectives:

This course will enable students:

- To provide practical exposure to handling missing data and identifying/treating outliers using statistical and machine learning techniques.
- To develop proficiency in applying data transformation, normalization, and scaling methods to improve model performance and data distribution.
- To implement encoding techniques for categorical data and assess their impact on memory usage and model interpretability.
- To perform comprehensive univariate, bivariate, and correlation analysis to discover meaningful patterns and feature relationships.
- To analyze and resolve data quality issues like multicollinearity and class imbalance for building effective predictive models.

Preamble: This Data Science Laboratory is designed to provide hands-on experience in data preprocessing, transformation, visualization, and analysis using Python programming. The course emphasizes practical implementation of data imputation, encoding, scaling, anomaly detection, correlation analysis, and handling real-world data challenges such as class imbalance and multicollinearity. It also introduces students to Exploratory Data Analysis (EDA) for both tabular and time-series data, helping them build robust, clean, and insightful datasets for effective machine learning applications.

Experiments

1. **Develop a Python program to perform missing value imputation and outlier detection/treatment on a given dataset containing numeric features, where it has to perform the following subtasks:**

   a) Load a dataset with missing values and display the count of missing values in each feature.
   b) Apply imputation techniques such as mean, median, and mode to fill missing values and display the updated dataset.
   c) Implement KNN imputation for missing values and compare the results with statistical imputation methods.
   d) Detect outliers in numeric columns using box plots and scatter plots for visualization.

e) Apply statistical methods such as the Interquartile Range (IQR) to identify outliers and remove/treat them from the dataset.
f) Display the cleaned dataset and explain the changes observed after imputation and outlier treatment.

2. **Develop a Python program to perform various data transformation and scaling techniques on a given dataset containing numeric features, where it has to perform the following subtasks:**

   a) Apply normalization and standardization techniques to the dataset and display the transformed data.
   b) Apply log transformation and power transformation on skewed features and visualize the effect on their distributions.
   c) Analyse and interpret the impact of each transformation on skewness and overall data distribution.
   d) Build a simple linear regression model before and after applying transformations, and compare the model performance using suitable evaluation metrics.

3. **Develop a Python program to encode categorical variables using different encoding techniques, where it has to perform the following subtasks:**

   a) Apply one-hot encoding, label encoding, and target encoding on the categorical features of a given dataset and display the encoded outputs.
   b) Compare the memory usage and interpretability of the dataset after applying each encoding technique.
   c) Build a logistic regression classifier using the encoded features and analyse how each encoding impacts the performance of the model.

4. **Develop a Python program to perform univariate and bivariate analysis, and correlation analysis on a given dataset, where it has to perform the following subtasks:**

   a) Analyse individual features using summary statistics such as mean, median, and mode, and visualize their distributions using histograms and box plots.
   b) Calculate and interpret skewness and kurtosis to understand the shape of feature distributions.
   c) Use scatter plots, correlation heat maps, and pair plots to visually examine relationships between pairs of features.
   d) Calculate Pearson and Spearman correlation coefficients for numeric features and interpret the correlation results.

5. **Develop a Python program to detect and handle multicollinearity in a dataset containing numeric features, where it has to perform the following subtasks:**

   a) Calculate the Variance Inflation Factor (VIF) for each feature and identify features exhibiting multicollinearity.

b) Remove or combine highly correlated features based on VIF values and justify the feature selection process.
c) Build a regression or classification model before and after addressing multicollinearity, and analyse the impact on model performance.

6. **Develop a Python program to analyse and handle class imbalance in a classification dataset, where it has to perform the following subtasks:**

   a) Analyse the class distribution in the dataset and visualize the imbalance using appropriate plots.
   b) Apply resampling techniques such as oversampling, under sampling, and SMOTE to balance the dataset and display the new class distribution.
   c) Build a classifier on the original and balanced datasets, and compare the model performance using suitable evaluation metrics.

7. **Develop a Python program to perform exploratory data analysis (EDA) and detect anomalies in a given dataset, where it has to perform the following subtasks:**

   a) Detect anomalies using statistical methods such as Z-score and Interquartile Range (IQR), and visualize the anomalies using appropriate plots.
   b) Perform EDA on both the original dataset and the dataset after removing anomalies, and analyse the differences in distributions and summary statistics.
   c) Experiment with a basic anomaly detection algorithm (such as Isolation Forest or DBSCAN) and demonstrate its application on practical use cases.

8. **Develop a Python program to perform exploratory data analysis (EDA) on time series data, where it has to perform the following subtasks:**

   a) Decompose the time series data into its trend, seasonality, and residual components, and visualize each component.
   b) Plot moving averages and autocorrelation function (ACF) plots to analyse temporal patterns and dependencies in the data.
   c) Identify and interpret any periodic or seasonal trends present in the time series data.

Course outcomes:

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

CO1. Apply data cleaning techniques such as missing value imputation and outlier treatment using Python.
CO2. Implement data transformation, normalization, and scaling techniques to prepare data for modeling.
CO3. Analyze the effects of different categorical encoding methods on model performance and interpretability.
CO4. Evaluate data distributions, correlations, and anomalies through exploratory data analysis (EDA) and visualization techniques.
CO5. Design preprocessed and well-structured datasets suitable for building machine learning models in real-world applications.

Learning Sources/ Web Source Links

1. **Pandas Documentation** – Data handling and preprocessing
   https://pandas.pydata.org/docs/
2. **Scikit-learn Documentation** – Imputation, scaling, encoding, modeling
   https://scikit-learn.org/stable/documentation.html
3. **Kaggle Learn: Data Science & Python** – Hands-on tutorials and practice
   https://www.kaggle.com/learn
4. **Seaborn Documentation** – Data visualization: histograms, boxplots, heatmaps
   https://seaborn.pydata.org/
5. **Imbalanced-learn Documentation** – SMOTE and handling class imbalance
   https://imbalanced-learn.org/stable/
6. **Statsmodels Documentation** – Time series decomposition, VIF, statistical analysis
   https://www.statsmodels.org/stable/index.html

Reference Books

1. **"Python for Data Analysis" by Wes McKinney**

- Publisher: O'Reilly Media
- ISBN: 978-1491957660
- Description: A comprehensive guide to data wrangling with Pandas, NumPy, and IPython. Ideal for learning data preprocessing, EDA, and data transformation.

2. **"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron**

- Publisher: O'Reilly Media
- ISBN: 978-1492032649
- Description: Covers practical ML workflows including data preparation, model building, evaluation, and advanced techniques like handling imbalance and transformations.