Vidyalankar Institute of Technology

An Autonomous Institute affiliated to University of Mumbai

Honours/Minor Degree Programme

for

Bachelor of Technology

in

Electronics and Computer Science

(R-2022 Curriculum)

(As per AICTE guidelines, with effect from the Academic Year 2024-25)

Preamble

To meet the challenge of ensuring excellence in engineering education, the issue of quality needs to be addressed, debated, and taken forward in a systematic manner. Therefore, autonomy for Vidyalankar Institute of Technology is not merely a transition from pre-cooked syllabi to self-designed curriculum. The autonomous curriculum of the Institute offers required academic flexibility with emphasis on industry requirements and market trends, employability and problem-solving approach which leads to improving competency level of learners with diverse strengths. In line with this, the curriculum framework designed is **Choice Based Credit and Grading System (CBCGS)**. The number of credits for each category of courses learnt by learners, internships and projects is finalized considering the scope of study and the ability that a learner should gain through the programme. The overall credits and approach of curriculum proposed is in line with AICTE model curriculum.

The curriculum comprises courses from various categories like basic sciences, humanities and social sciences, engineering sciences, general education and branch specific courses including professional electives and open electives. The curriculum has core courses of branch of engineering positioned and sequenced to achieve sequential and integral learning of the entire breadth of the specific branch. These courses are completed by the third year of the engineering programme that enables learners to prepare for higher education during their final year. Professional elective courses, that begin from third year of programme, offer flexibility and diversity to learners to choose specialization from a basket of recent developments in their field of technology. The selection of unique professional elective courses based on industrial requirements and organizing them into tracks is a salient feature of this curricula ensuring employability. Open Elective courses cover multi-disciplinary, special skill development, project management and similar knowledge that make learners capable of working in an industrial environment.

For holistic development of learners, apart from technical courses, Humanities and Social Science courses develop the required soft-skills and attitude amongst learners. Our curriculum also introduces Social Service Internship and Internship with institutes abroad along with courses like Design Thinking, Wellness – Body, Mind & Spirit, Indian Traditional Knowledge system under General Education category. These general education courses aim to create balance in brain hemispheres and hence improve learners' clarity in thoughts and responses.

Additionally, curriculum provides add-on Honours/Minor degree that involves field/ domain study. Learners can avail themselves of this degree by completing the requirement of additional 18 credits.

Thus, the academic plan of VIT envisages a shift from summative to formative and competency-based learning system which will enhance learner's ability towards higher education, employability and entrepreneurship.

Chairperson, Board of Studies
Department of Electronics and Computer Science
Vidyalankar Institute of Technology

Chairperson, Academic Council
Vidyalankar Institute of Technology

[A] Guidelines for Award of Honours/ Minor Degree Programme

Honours/Minor Degree programme is introduced in order to facilitate learners to enhance the depth of knowledge, diversity, breadth and skills in emerging fields. An Honours or Minor Degree typically refers to a higher level of academic achievement either for research orientation or for improving employability. Learners can select any Honours or Minor Degree programme as per his/her choice.

In our curriculum, learners can choose to avail Honours/ Minor Degree programme by completing requirements of 18 credits, which will be over and above the minimum credits required for B.Tech. degree i.e. credit requirement for the award of degree programme and Honours/ Minor degree programme are required to be explicitly carried out. Learners shall opt for Honours or Minor specialisations during the break of Semester 5 and Semester 6. **Learners may complete the B.Tech. Degree programme without opting for Honours/Minor Degree programme** i.e. opting for Honours/ Minor Degree programme is not mandatory as a part of B.Tech. Degree programme

For Honours/Minor Degree, the learner shall select an Honours/Minor programme offered by his/her own home department.

Eligibility Criteria

- Basic eligibility for opting for Honours/Minor shall be minimum CGPA of 6.75 at the end of 4th semester and earned 81 credits from Sem 1 to Sem 4 (41 credits for DSY students).
- If student has already completed any course(s) that is listed in the chosen Honours/ Minor Degree programme, as additional learning course(s), then the transfer credits for such course(s) can be carried out towards Honours/ Minor Degree programme.
- For a student to get Honours/ Minor Degree, it is mandatory that the student completes the relevant courses before graduating.

Syllabus Scheme Template

Course		Head of Learning	Preferred Semester	Credits	Assessment Guidelines (Marks)			Total marks (Passing@40% of total marks)
Code	Name				ISA	MSE	ESE	
HM01	Industry Interaction	Theory	Break of Sem5 and Sem6	1	25	-	-	025
HMXX	Honours / Minor Degree Course 1	Theory	6	2	15	20	40	075
HMXX	Honours / Minor Degree Course 1 Lab	Practical	6	1	25	-	25	050
HM02	Survey Report/ Paper	Theory	Break of Sem6 and Sem7	2	25	-	25	050
HMXX	Honours / Minor Degree Course 2	Theory	7	2	15	20	40	075
HMXX	Honours / Minor Degree Course 2 Lab	Practical	7	1	25	-	25	050
HM03	Seminar	Theory	Break of Sem7 and Sem8	2	25	-	25	050
HMXX	Honours / Minor Degree Course 3	Theory	8	2	15	20	40	075
HMXX	Honours / Minor Degree Course 3 Lab	Practical	8	1	25	-	25	050
HM04	Capstone Project	Practical	8	4	75	-	50	125
			Total	**18**				

[B] Honours/ Minor Degree Programmes offered for B.Tech. Electronics and Computer Science

The Institute offers the listed Honours Degree Programme for learners of B.Tech. Electronics and Computer Science.

Honours/ Minor Degree Programmes Offered

Sr.No.	Honours/Minor Degree Programme	Department offering Honours	Honours applicable for	Minors applicable for
1	Next-Gen Artificial Intelligence and Machine Learning (Next-Gen AI&ML)	Electronics and Computer Science	B.Tech. Electronics and Computer Science students who have opted for AIML specialization track	None
2	Next-Gen Data Science (Next-Gen DA)	Electronics and Computer Science	B.Tech. Electronics and Computer Science students who have opted for DA specialization track	None
3	Next-Gen Internet of Things	Electronics and Telecommunication Engineering	B.Tech. Electronics and Computer Science students who have opted for IOT specialization track	As stated in Honours/Minor Degree document of Electronics and Telecommunication Engineering
4	UI/ UX	Information Technology	All B.Tech. Electronics and Computer Science students	As stated in Honours/Minor Degree document of Information Technology
5	Blockchain	Information Technology	All B.Tech. Electronics and Computer Science students	As stated in Honours/Minor Degree document of Information Technology

List of courses under each Honours/ Minor Degree Programme:

1. Next-Gen Artificial Intelligence and Machine Learning

Semester	Course Code	Course Name
VI	HMEC01T	Ethics in AI
VI	HMEC01P	Ethics in AI Lab
VII	HMEC02T	Scalable ML and BDA
VII	HMEC02P	Scalable ML and BDA Lab
VIII	HMEC03T	Generative AI models
VIII	HMEC03P	Generative AI models Lab

2. Next-Gen Data Science

Semester	Course Code	Course Name
VI	HMEC04T	Data Visualization Using R-Programming
VI	HMEC04P	Data Visualization Using R-Programming Lab
VII	HMEC05T	Time Series and Forecasting
VII	HMEC05P	Time Series and Forecasting Lab
VIII	HMEC06T	Cloud Native and BDA
VIII	HMEC06P	Cloud Native and BDA Lab

3. Next-Gen Internet of Things

Semester	Course Code*	Course Name
VI	HMET01T	Embedded Linux System
VI	HMET01P	Embedded Linux System Lab
VII	HMET02T	IoT and Data Analytic
VII	HMET02P	IoT and Data Analytic Lab
VIII	HMET03T	IoT Applications and Web Development
VIII	HMET03P	IoT Applications and Web Development Lab

* Detailed Syllabus of these courses can be obtained from the Honours/ Minor Degree document of Electronics and Telecommunication department applicable for R-2022.

4. UI/ UX

Semester	Course Code*	Course Name
VI	HMIT01T	Foundation of UI/UX
VI	HMIT01P	Foundation of UI/UX Lab
VII	HMIT02T	UX Design, Evaluation and ARVR
VII	HMIT02P	UX Design, Evaluation and ARVR Lab
VIII	HMIT03T	Use cases in UI/UX
VIII	HMIT03P	Use cases in UI/UX Lab

* Detailed Syllabus of these courses can be obtained from the Honours/ Minor Degree document of Information Technology department applicable for R-2022.

5. Blockchain

Semester	Course Code*	Course Name
VI	HMIT04T	Blockchain Technology
VI	HMIT04P	Blockchain Technology Lab
VII	HMIT05T	Smart Contract and Crypto Currencies
VII	HMIT05P	Smart Contract and Crypto Currencies Lab
VIII	HMIT06T	Decentralize & Blockchain Technologies
VIII	HMIT06P	Decentralize & Blockchain Technologies Lab

*Detailed Syllabus of these courses can be obtained from the Honours/ Minor Degree document of Information Technology department

Learners of Electronics and Computer Science Department who wish to opt for Honours/Minor Degree Programme offered by other departments can obtain details of the same from Section-B and Section C of the Honour/ Minor Degree programme document of respective department.

[C] Honours/ Minor Degree Programmes Course Syllabus

Course Name: Ethics in AI

Course Code: HMEC01T

Category: Honours (Next-Gen AIML)

Preamble:

The rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized numerous industries and daily life, introducing unprecedented opportunities and challenges. As these technologies integrate deeply into societal structures, it becomes imperative to consider their ethical, social, and environmental implications. This course aims to provide learners with a foundational understanding of ethical principles and frameworks as applied to AI/ML systems.

Pre-requisites:

Artificial Intelligence, Machine Learning

Course Objectives:

1. Understand ethical considerations in AI and ML development and deployment.
2. Explore frameworks for ethical decision-making in AI systems.
3. Assess potential biases, privacy issues, and impacts of AI on society.
4. Learn legal and policy implications related to AI and ML.
5. Develop skills to implement ethical practices in AI projects.
6. Foster critical thinking to address ethical challenges in real-world scenarios.

Course Outcomes:

Learner will be able to:

CO1: Demonstrate understanding of ethical principles in AI/ML.
CO2: Identify and mitigate bias and fairness issues in datasets and algorithms.
CO3: Apply ethical frameworks to evaluate AI systems.
CO4: Design AI systems with accountability, transparency, and fairness.
CO5: Understand societal impacts
CO6: Engage in ethical decision-making during AI system development.

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
2	-	2	-

Assessment guidelines:

Head of Learning	ISA	MSE	ESE	Total
Theory	15	20	40	75

The assessment/evaluation guidelines for the courses of different credits are mentioned in the above table. Notwithstanding the above, each course faculty shall have the choice to decide her/his assessment methodology based on the nature of the course. Faculty may propose the revised assessment methodology for his/her course. However, the revised assessment methodology shall be approved by a panel constituted at institute level and published to the learners before the commencement of the semester.

Detailed Syllabus:

Module No.	Module Name	Content	No of Hours
1	Introduction to Ethics in AI/ML	Overview of Ethics: Moral principles, ethics vs. legality. Why Ethics in AI/ML? Risks and challenges. Key Ethical Issues in AI: Bias, fairness, accountability, transparency. Case Studies: Real-world AI ethical dilemmas.	5
2	Bias and Fairness in AI/ML	Definition and Types of Bias in AI: Dataset bias, algorithmic bias. Techniques to Detect and Mitigate Bias in ML Models. Fairness Frameworks: Disparate impact, equalized odds. Ethical Data Collection and Preprocessing.	5
3	Privacy and Security Concerns	Privacy Challenges in AI: Data collection, storage, and sharing. Ethical Guidelines for User Data Protection. Security Risks in AI Systems: Deepfakes, adversarial attacks. GDPR and Other Privacy Regulations.	5
4	Accountability and Transparency	Need for Explainable AI (XAI). Strategies for Creating Transparent AI Systems. Accountability in AI Decision-Making. Ethical Implications of Autonomous Systems.	5
5	Societal Impacts of AI/ML	Impacts on Employment and Workforce. AI and Social Inequality. Misinformation and AI-Generated Content. AI in Healthcare, Education, and Governance.	5
6	Legal and Ethical Frameworks in AI	Overview of AI Ethics Guidelines (IEEE, UNESCO, etc.). AI Laws and Policies: International and regional perspectives. Intellectual Property and AI-Generated Content. Future Directions in AI Ethics.	5
Total			**30**

Textbooks:

1. AI Ethics: A Textbook by Paula Boddington - A comprehensive introduction to ethical challenges in AI systems.
2. Atlas of AI by Kate Crawford - Discusses the societal and environmental impact of AI

Reference Books:

1. The Ethical Algorithm by Aaron Roth & Michael Kearns - Explores designing socially aware algorithms
2. Human Compatible by Stuart Russell - Focuses on aligning AI with human values
3. Moral Machines: Teaching Robots Right From Wrong by Wendell Wallach and Colin Allen - Examines embedding ethics into AI systems.

Online Resources for Learning:

1. **Responsible AI Toolkit Reading List** - A curated set of academic papers and books focusing on responsible AI development and ethics Responsible AI Toolkit
2. **Oxford Academic Journals** - Offers edited volumes on AI ethics and related fields, often authored by leading experts

Course Name: Ethics in AI Lab

Course Code: HMEC01P

Category: Honours (Next-Gen AIML)

Preamble:

The rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized numerous industries and daily life, introducing unprecedented opportunities and challenges. As these technologies integrate deeply into societal structures, it becomes imperative to consider their ethical, social, and environmental implications. This lab manual is designed to equip learners with practical skills to address ethical concerns in AI and ML systems while fostering a deeper understanding of fairness, accountability, transparency, and privacy. Through guided experiments and critical discussions, students will explore the balance between technological innovation and ethical responsibility, preparing them to design and deploy AI systems that respect human values and promote societal good.

Pre-requisites:

Artificial Intelligence, Machine Learning

Lab Objectives:

- Understand ethical considerations in AI and ML development and deployment.
- Explore frameworks for ethical decision-making in AI systems.
- Assess potential biases, privacy issues, and impacts of AI on society.
- Learn legal and policy implications related to AI and ML.
- Develop skills to implement ethical practices in AI projects.
- Foster critical thinking to address ethical challenges in real-world scenarios.

Lab Outcomes:

Learner will be able to:

CO1: Demonstrate understanding of ethical principles in AI/ML.
CO2: Identify and mitigate bias and fairness issues in datasets and algorithms.
CO3: Apply ethical frameworks to evaluate AI systems.
CO4: Design AI systems with accountability, transparency, and fairness.
CO5: Understand societal impacts
CO6: Engage in ethical decision-making during AI system development.

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
-	2	-	1

Assessment guidelines:

Head of Learning	ISA	MSE	ESE	Total
Practical	25	-	25	50

The assessment/evaluation guidelines for the courses of different credits are mentioned in the above table. Notwithstanding the above, each course faculty shall have the choice to decide her/his assessment methodology based on the nature of the course. Faculty may propose the revised assessment methodology for his/her course. However, the revised assessment methodology shall be approved by a panel constituted at institute level and published to the learners before the commencement of the semester.

List of Experiments

1. Bias Detection and Mitigation in ML Models
2. Privacy Preservation Using Differential Privacy
3. Explainable AI (XAI): Interpreting Black-Box Models
4. Adversarial Attack and Defense Strategies
5. Fairness in AI-Powered Recommendation Systems
6. Simulating Ethical Dilemmas in Autonomous Systems
7. Energy Efficiency Analysis of ML Models
8. Ethical Concerns in NLP Models for Sentiment Analysis
9. Misinformation Detection Using AI
10. Guidelines for Ethical Data Collection and Labeling

Course Name: Scalable Machine Learning & Big Data Analytics

Course Code: HMEC02T

Category: Honours (Next-Gen AIML)

Preamble:

This course covers distributed machine learning and large-scale data processing frameworks. Students will learn to design scalable ML pipelines for massive datasets using cloud platforms and parallel computing. Applications include recommendation systems, IoT analytics, and real-time prediction systems.

Pre-requisites:
• Machine Learning Fundamentals
• Python/Java/Scala Programming
• Database Systems

Course Objectives:

After completing this course, you will be able to learn:
1. Understand distributed computing fundamentals for big data processing
2. Implement efficient data storage and processing in distributed systems
3. Design and implement scalable machine learning algorithms
4. Develop real-time analytics solutions using stream processing
5. Apply advanced analytics techniques to complex data types
6. Deploy and monitor end-to-end scalable ML systems

Course Outcomes:
CO1: Explain distributed computing architectures and their applications
CO2: Build optimized data pipelines using distributed storage systems
CO3: Implement distributed ML algorithms for large datasets
CO4: Create real-time analytics solutions for streaming data
CO5: Apply graph, time series, and geospatial analytics at scale
CO6: Deploy and maintain production-ready ML systems

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
2	-	2	-

Assessment Guidelines:

Head of Learning	ISA	MSE	ESE	Total
Theory	15	20	40	75

The assessment guidelines for the courses of different credits are mentioned above. Notwithstanding the above, each course faculty shall have the choice to propose her/his assessment methodology based on the nature of the course. However, the proposed assessment methodology shall be approved by a panel constituted at Institute level and published to the learners before the commencement of the semester.

Detailed Syllabus:

Module No.	Module Name	Content	No. of Hours
1	Big Data Fundamentals	• Introduction to distributed systems • Hadoop ecosystem overview • Spark architecture and RDDs • Cloud platforms comparison (AWS/GCP) • Hands-on: Setting up Spark cluster	4
2	Distributed Data Processing	• HDFS architecture and operations • NoSQL databases (MongoDB/Cassandra) • Kafka streaming fundamentals • Data partitioning strategies • Lab: Building ETL pipeline with Spark+Kafka	6
3	Scalable Machine Learning	• Spark MLlib pipeline components • Distributed algorithms (K-means, Random Forest) • Federated learning concepts • Lab: Implementing collaborative filtering at scale	6
4	Real-Time Analytics	• Stream processing with Flink/Spark Streaming • Window functions and state management • Fraud detection system design • Lab: Real-time anomaly detection	6
5	Advanced Applications	• Graph analytics with GraphX • NLP for big data (Spark NLP) • Geospatial processing techniques • Lab: Social network analysis	6
6	Deployment & Scaling	• Containerization with Docker • Kubernetes orchestration • ML model monitoring • Capstone project presentation	6
Total			**30**

Textbooks:

1. Learning Spark: Lightning-Fast Data Analytics – Jules Damji, Brooke Wenig, Tathagata Das, Denny Lee
2. Designing Data-Intensive Applications – Martin Kleppmann
3. Machine Learning Engineering – Andriy Burkov
4. Streaming Systems – Tyler Akidau, Slava Chernyak, Reuven Lax

Reference Books:

1. Hadoop: The Definitive Guide – Tom White
2. Kubernetes in Action – Marko Luksa
3. Federated Learning – Qiang Yang, Yang Liu, Yong Cheng

Course Name: Scalable Machine Learning & Big Data Analytics Lab

Course Code: HMEC02P

Category: Honours (Next-Gen AIML)

Preamble:

This hands-on lab course develops practical skills in distributed data processing and large-scale machine learning using industry-standard tools. Through experimental learning, students will work with big data technologies to implement scalable ML pipelines, process streaming data, and deploy models in cloud environments. The lab focuses on applied techniques using Spark, Hadoop, Kafka, and cloud platforms in a project-based environment.

Pre-requisites:

- Programming Fundamentals (Python/Java/Scala)
- Machine Learning Basics
- Database Concepts

Course Objectives:

After completing this lab course, students will be able to:

1. Implement distributed data processing pipelines
2. Develop and optimize scalable machine learning models
3. Deploy ML solutions in cloud environments

Course Outcomes:

Upon completion, learners will be able to:

CO1: Execute big data processing workflows using Spark/Hadoop

CO2: Build distributed ML models with Spark MLlib

CO3: Deploy scalable ML solutions using cloud platforms

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
-	2	-	1

Assessment Guidelines:

Head of Learning	ISA	MSE	ESE	Total
Theory	25	--	25	50

The assessment guidelines for the courses of different credits are mentioned above. Notwithstanding the above, each course faculty shall have the choice to propose her/his assessment methodology based on the nature of the course. However, the proposed assessment methodology shall be approved by a panel constituted at Institute level and published to the learners before the commencement of the semester.

Lab Experiments:

1. **Spark Data Processing** – Implement RDD/DataFrame operations on large datasets
2. **Distributed ETL Pipelines** – Build data pipelines using Spark SQL and Hadoop
3. **Scalable ML with MLlib** – Train distributed models (Random Forest, ALS)
4. **Stream Processing** – Implement real-time analytics with Kafka+Spark Streaming
5. **Federated Learning** – Setup privacy-preserving ML with TensorFlow Federated
6. **Cloud Deployment** – Containerize models using Docker and deploy on Kubernetes
7. **Performance Optimization** – Benchmark and tune distributed ML systems
8. **Capstone Project** – End-to-end scalable ML solution for real-world problem

Tools/Platforms:

- Apache Spark
- Hadoop (HDFS/YARN)
- Apache Kafka
- Docker/Kubernetes
- AWS EMR/GCP Dataproc
- Jupyter Notebook

Text Books:

1. Learning Spark – Jules Damji et al.
2. Spark: The Definitive Guide – Bill Chambers, Matei Zaharia
3. Machine Learning Engineering – Andriy Burkov

Reference Books:

1. Hadoop: The Definitive Guide – Tom White
2. Designing Data-Intensive Applications – Martin Kleppmann
3. Kubernetes in Action – Marko Luksa

Course Name: Generative AI Models

Course Code: HMEC03T

Category: Honours (Next-Gen AIML)

Preamble:

This course provides an in-depth study of deep learning architectures, algorithms, and applications. It covers neural networks, backpropagation, CNNs, RNNs, GANs, and transformers, with hands-on implementation using TensorFlow and PyTorch. Through practical exercises, case studies, and projects, students will develop expertise in designing, training, and optimizing deep neural networks for research and industry applications.

Pre-requisites:

- Calculus, Linear Algebra,
- Probability Theory, and
- Python programming
- Machine Learning

Course Objectives:

After completing this course, students will be able to:

1. To learn the fundamentals of Neural Networks and their various types.
2. To explore Generative AI models like GANs, VAEs, and Transformers.
3. To analyze the limitations of traditional RNNs and LSTMs.
4. To discuss current trends and future directions in Generative AI research.

Course Outcomes:
Learner will be able to
CO1: Understand the evolution of AI and the significance of Deep Learning.
CO2: Apply various Neural Network architectures for tasks like image recognition and sequence modeling.
CO3: Analyze data preprocessing and training techniques for neural networks.
CO4: Design practical solutions using advanced neural networks for diverse applications

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
2	-	2	-

Assessment Guidelines:

Head of Learning	ISA	MSE	ESE	Total
Theory	15	20	40	75

The assessment guidelines for the courses of different credits are mentioned above. Notwithstanding the above, each course faculty shall have the choice to propose her/his assessment methodology based on the nature of the course. However, the proposed assessment methodology shall be approved by a panel constituted at Institute level and published to the learners before the commencement of the semester.

Detailed Syllabus:

Module No.	Module Name	Content	No. of Hours
1	Introduction to Large Language Models	Overview of Generative AI and Large Language Models. Basics of attention mechanisms and Transformer architecture. Pre-training techniques and transfer learning strategies.	5
2	GPT Models and Applications:	Study of GPT architecture and variants. Applications of GPT models in text generation and dialogue systems. GPT-based chatbot enhances E-Shop's customer support service	6
3	BERT and Advanced Techniques:	Understanding BERT architecture and pre-training objectives. Fine-tuning BERT for downstream NLP tasks.	5
4	Text Generation with Generative AI	Introduction to Text Generation, LSTM-based Text Generation, Transformer-based Text Generation, Fine-Tuning Language Models, and Text Generation Applications	5
5	Music Generation with Generative AI	Introduction to Music Generation, Music Representation, and LSTM-based Music Generation. Transformer-based Music Generation	5
6	Applications and Future Directions:	Real-world applications of large language models. Challenges and limitations of current approaches. Emerging trends and future dir ections in Generative AI.	4
Total			**30**

Textbooks:

1. Generative AI for Everyone: Understanding the Essentials and Applications of This Breakthrough Technology". Altaf Rehmani .
2. Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
3. Neural Networks and Deep Learning: A Textbook" by Charu C. Aggarwal..

Reference Books:

Generative Adversarial Networks Cookbook: Over 100 recipes to build generative models using Python, TensorFlow, and Keras" by Josh Kalin.

1. Generative AI in Software Development: Beyond the Limitations of Traditional Coding" Jesse Sprinter, 2024..
2. Douwe Osinga. "Deep Learning Cookbook" O'REILLY, SPD Publishers, Delhi.
3. Simon Haykin, Neural Network- A Comprehensive Foundation- Prentice Hall International, Inc
4. S.N.Sivanandam and S.N.Deepa, Principles of soft computing-Wiley India

Course Name: Generative AI Models Lab

Course Code: HMEC03P

Category: Honours (Next-Gen AIML)

Preamble:

This course provides an in-depth study of deep learning architectures, algorithms, and applications. It covers neural networks, backpropagation, CNNs, RNNs, GANs, and transformers, with hands-on implementation using TensorFlow and PyTorch. Through practical exercises, case studies, and projects, students will develop expertise in designing, training, and optimizing deep neural networks for research and industry applications.

Pre-requisites:

- Calculus,
- Linear Algebra,
- Probability
- Python programming

Course Objectives:

After completing this lab course, students will be able to:

1. To learn Python and TensorFlow skills for Generative AI.
2. To study techniques for cleaning and preparing data for Generative AI tasks.
3. To implement generative AI models
4. To develop innovative applications using generative AI tools and techniques.

compositions for diverse music applications.

Course Outcomes:

Upon completion, learners will be able to:

CO1:Implement Python and TensorFlow basics, including data handling and preprocessing techniques.

CO2. Implement Generative AI models such as GANs, VAEs, LSTM networks, and Transformer models for image, text, and music generation tasks.

CO3. Evaluate model performance and experiment with hyperparameters and optimization techniques to enhance Generative AI outcomes.

CO4. Develop innovative applications in image, text, and music generation, showcasing practical skills.

Course Scheme:

Contact Hours		**Credits Assigned**	
Theory	Practical	Theory	Practical
-	2	-	1

Assessment Guidelines:

Head of Learning	**ISA**	**MSE**	**ESE**	**Total**
Theory	25	--	25	50

The assessment guidelines for the courses of different credits are mentioned above. Notwithstanding the above, each course faculty shall have the choice to propose her/his assessment methodology based on the nature of the course. However, the proposed assessment methodology shall be approved by a panel constituted at Institute level and published to the learners before the commencement of the semester.

Suggested List of Practicals:

Lab Experiments:

1. Write Python scripts to implement basic operations and TensorFlow 2 tensors.
2. Preprocess and clean datasets for Generative AI applications using Python libraries such as Pandas and NumPy. Handle missing data, normalize features, and encode categorical variables.
3. Use Matplotlib or Seaborn to visualize data distributions and patterns in Generative AI datasets. Plot histograms, scatter plots, and heatmaps to analyze data characteristics.
4. Implement a Generative Adversarial Network (GAN) architecture using TensorFlow 2. Train the GAN model on a dataset such as MNIST or CIFAR-10 for image generation tasks.
5. Train a GAN model on a custom dataset for image generation. Experiment with hyperparameters, loss functions, and optimization techniques to optimize GAN training.
6. Explore advanced techniques such as Wasserstein GANs, Progressive GANs, or StyleGANs for image generation. Implement and compare these techniques for generating high-quality images. Develop applications for image and video generation using trained Generative AI models. Use the models to generate art, create deep fakes, or synthesize video content.
7. Develop applications for image and video generation using trained Generative AI models. Use the models to generate art, create deep fakes, or synthesize video content.
8. Text Generation: Implement a Long Short-Term Memory (LSTM) network using TensorFlow 2 for text generation tasks. Train the LSTM model on a dataset of text sequences and generate new text samples.
9. Text generation: Implement a Transformer-based language model (e.g., GPT) using TensorFlow 2 for text generation. Fine-tune the model on a text corpus and generate coherent and contextually relevant text.
10. Text generation: Fine-tune a pre-trained language model (e.g., GPT, BERT) using transfer learning techniques. Fine-tune the model on a domain-specific dataset and evaluate its performance for text generation tasks.
11. Text generation: Develop applications for text generation tasks such as story generation, dialogue generation, or code generation using trained Generative AI models.
12. Music Generation: Preprocess music data and represent it in a suitable format for music generation tasks. Explore MIDI or audio representations for training Generative AI models.

Tools/Platforms:

- Python (Pandas, Matplotlib, Seaborn)
- R (tidyverse, ggplot2)
- Tableau/Power BI
- Jupyter Notebook/RStudio

Textbooks:

1. Responsible AI: Implementing Ethical and Unbiased Algorithms, by Shashin Mishra and Sray Agarwal

2. Generative AI in Practice: 100+ Amazing Ways Generative Artificial Intelligence is Changing Business and Society, Bernard Marr

Reference Books:

1. Generative AI with Python and TensorFlow 2 Create images, text, and music with VAEs, GANs, LSTMs, Transformer models", Joseph Babcock and Raghav Bali
2. Generative Adversarial Networks: An Overview by Vinod Nair and Geoffrey E. Hinton.
3. Hands-On Generative Adversarial Networks with PyTorch 1.x by Stefano Bosisio and Vijayabhaskar J.

Course Name: Data visualization using R Programming

Course Code: HMEC04T

Category: Honours (Next-Gen DS)

Preamble:

It introduces fundamental concepts, advanced techniques, and best practices while introducing tools and libraries within the R ecosystem. It will train learners to interpret, design, and communicate insights effectively with real-world applications that bring into closer view the role of data visualization in decision-making.

Pre-requisites:

Skill Based Lab-Python Programming

Course Objectives:

1. Develop a good understanding of the theoretical underpinnings of data visualization, including principles of clarity, accuracy, and storytelling.
2. Learn how to use R programming and its visualization libraries such as ggplot2, plotly, and sf to create a range of charts and graphs.
3. Learn to clean, manipulate, and prepare data effectively for visualization using R's data wrangling packages like dplyr and tidyr.
4. Explore advanced visualization techniques, including geospatial mapping, interactive dashboards, and visualizations for complex datasets such as networks and time series.
5. Understand how to design and present visualizations that clearly communicate insights, supported by reproducible reports and presentations using R Markdown.
6. Demonstrate the ability to apply data visualization methods to real-world problems from a variety of domains, including business, healthcare, and environmental studies, through case studies and projects.

Course Outcome

Learner will be able to:
CO1: Understand and apply data visualization principles to enhance decision-making and avoid common pitfalls.
CO2: Demonstrates competence in using R and key libraries for visualization (ggplot2, dplyr, plotly) for data analysis.
CO3: Create and customize appropriate visualizations, such as scatter plots, bar charts, and histograms.
CO4: Prepare and maintain data for visualization through transformation and tidy data principles.
CO5: Improve advanced visualizations for multivariate, time series, and geospatial data, both statically and interactively.
CO6: Design accessible, effective visualizations using strong communication and aesthetic principles for diverse audiences.

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
2	-	2	-

Assessment guidelines:

Head of Learning	ISA	MSE	ESE	Total
Theory	15	20	40	075

The assessment/evaluation guidelines for the courses of different credits are mentioned in the above table. Notwithstanding the above, each course faculty shall decide her/his assessment methodology based on the course's nature. Faculty may propose the revised assessment methodology for his/her course. However, the revised assessment methodology shall be approved by a panel constituted at institute level and published to the learners before the commencement of the semester.

Detailed Syllabus:

Module No.	Module Name	Content	No of Hours
1	Introduction to Data Visualization and R Basics	Importance of Data Visualization: Role in data analysis and decision-making, Definition and importance of data visualization in analytics and decision-making, Principles of effective visualization (clarity, simplicity, and accuracy), how to avoid misleading visualizations, Differences between exploratory and explanatory data visualization. Components: data, visual encodings, and context, Understanding visual perception and cognitive load. Introduction to R Programming: Overview of R and RStudio, Key libraries for visualization: ggplot2, dplyr, plotly Understanding Data Structures in R: Vectors, data frames, tibbles, and lists, Loading and exploring datasets in R	5
2	Fundamentals of Data Visualization	Types of Data and Their Visualization Needs: Categorical, numerical, temporal, and geospatial data, Matching chart types to data types. Overview of Chart Types; Scatter plots, bar charts, and line charts, Histograms, density plots, boxplots, and pie charts The Grammar of Graphics (ggplot2): Understanding layers: data, aesthetics, and geometries, Customizing plots with themes, labels, and legends. Understanding data mappings and coordinates Best Practices for Chart Selection: Choosing appropriate charts for categorical, numerical, and temporal data.	5
3	Data Preparation for Visualization	Data transformation: filtering, aggregating, and reshaping. Tidy Data Principles: Importance of structured data for visualization, Using tidyr and dplyr for data preparation. Handling Large Datasets: Sampling techniques and data summarization. Efficient visualization strategies for big data	5
4	Advanced Visualization Techniques	Visualization of Multivariate Data: Techniques for high-dimensional data (e.g., scatterplot matrices, parallel coordinates), Visualizing clustering and classification results. Multi-Panel Visualizations: Faceting techniques for subset comparison, Overlaying plots and combining visualizations Time Series Visualizations: Trend lines and seasonal patterns, Temporal patterns and trends, Smoothing and seasonal decomposition	5

		Geospatial Visualizations: Mapping spatial data and geospatial patterns, Incorporating layers, heatmaps, and choropleth maps, maps with ggplots and sf	
5	Design Principles and Aesthetic Customization	Designing Visualizations for Communication: Structuring narratives for data-driven presentations, Color Theory and Accessibility: Effective use of color in data visualization, Ensuring accessibility (e.g., colorblind-friendly palettes), Customizing Visualization Elements: Titles, labels, legends, and annotations for clarity, Layout and spacing for better readability.	
6	Interactive and Specialized Visualizations	Interactive Visualization Concepts: The need for interactivity in data exploration, Overview of tools for creating interactive visualizations Specialized Visualization Types: Network visualizations (e.g., node-link diagrams), Statistical visualizations: regression plots, confidence intervals, and distributions, Representing regression results and confidence intervals, Diagnostic plots and residual analysis Tree maps and hierarchical data visualizations, Visualization for Big Data, Techniques for summarizing and aggregating large datasets, Sampling and visual encoding strategies	5
Total			**30**

Textbooks:

1. R for Data Science by Hadley Wickham and Garrett Grolemund
2. Data Visualization: A Practical Introduction by Kieran Healy

Reference books:

1. Practical Data Science with R by Nina Zumel and John Mount

Course Name: Data visualization using R Programming Lab

Course Code: HMEC04P

Category: Honours (Next-Gen DS)

Preamble:

The Data & Feature Engineering Lab provides hands-on experience in applying data preprocessing and feature engineering techniques to real-world datasets. Students will learn to collect, clean, preprocess, and transform data to prepare it for analysis and modeling. In the feature engineering process, you start with your raw data and use your own domain knowledge to create features that will make your machine learning algorithms work. In this module we explore what makes a good feature. Through practical exercises and projects, students will develop proficiency in using tools and libraries commonly employed in data engineering tasks.

Prerequisites:

Skill Based Lab-Python Programming

Course Objectives:

1. Master R programming and manipulate various data structures effectively.
2. Perform exploratory data analysis (EDA) and analyze data patterns.
3. Match appropriate visualizations to different data types for meaningful insights.
4. Create and customize visualizations using the grammar of graphics (ggplot2).
5. Transform and prepare data for visualization using dplyr and tidyr.
6. Visualize high-dimensional and geospatial data and communicate insights clearly.

Course Outcomes:

Learner will be able to:
CO1: Apply R and RStudio tools to manipulate and explore data structures.
CO2: Perform basic exploratory data analysis and find insights in the data.
CO3: Select and apply the appropriate visualizations for categorical and numerical data.
CO4: Create and customize visualizations using ggplot2 and the grammar of graphics.
CO5: Transform and aggregate data for visualization using dplyr and tidyr.
CO6: Visualize high-dimensional and geospatial data and effectively communicate insights.

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
-	2	-	1

Assessment guidelines:

Head of Learning	ISA	MSE	ESE	Total
Practical	25	-	25	50

Suggested list of experiments:

Sr. No.	List of experiments
1	Introduction to R and RStudio
2	Create, access, and modify various data structures in R
3	Loading and Exploring Datasets and perform basic exploratory data analysis (EDA).
4	Visualizing Categorical and Numerical Data and Match the right chart type to different data types
5	Understand the grammar of graphics and create plots like scatter plots, line charts, and box plots using ggplot2
6	Data transformation techniques for data visualization (filtering, aggregating, reshaping).Using dplyr and tidyr to filter and aggregate a dataset
7	Implement data sampling techniques, perform summarization
8	Visualize high-dimensional data using techniques like scatterplot matrices and parallel coordinates.
9	Visualize geospatial data using ggplot2 and sf using choropleth maps
10	Design effective visualizations for storytelling and communication.

Textbooks:

1. R for Data Science by Hadley Wickham and Garrett Grolemund
2. Data Visualization: A Practical Introduction by Kieran Healy

Reference books:

1. Practical Data Science with R by Nina Zumel and John Mount

Course Name: Time Series and Forecasting

Course Code: HMEC05T

Category: Honours (Next-Gen DS)

Preamble:

This course provides a comprehensive introduction to time series analysis and forecasting techniques. It covers fundamental concepts such as time series decomposition, trend analysis, smoothing techniques, and machine learning-based forecasting models.

Pre-requisite:

- Probability Theory
- Basic Statistics
- Linear Algebra

Course Objectives:

1. Understand the fundamental components of time series data.
2. Apply different time series models for forecasting.
3. Learn statistical and machine learning approaches for time series forecasting.
4. Evaluate forecast models using error metrics.
5. Apply forecasting techniques to real-world datasets.

Course Outcomes:

After successful completion of the course students will be able to:

CO1: Identify trends, seasonality, and cyclic patterns in time series data.

CO2: Implement and compare different smoothing techniques like Moving Averages and Exponential Smoothing.

CO3: Apply ARIMA, SARIMA, and GARCH models for time series forecasting.

CO4: Use deep learning models such as LSTMs for advanced forecasting.

CO5: Evaluate forecast models and optimize parameters for better predictions.

Course Scheme:

Contact Hours		**Credits Assigned**	
Theory	Practical	Theory	Practical
2	-	2	-

Assessment Guidelines:

Head of Learning	**ISA**	**MSE**	**ESE**	**Total**
Theory	15	20	40	75

The assessment/evaluation guidelines for the courses of different credits are mentioned in the above table. Notwithstanding the above, each course faculty shall have the choice to decide her/his assessment methodology based on the nature of the course. Faculty may propose the revised assessment methodology for his/her course. However, the revised assessment methodology shall be approved by a panel constituted at institute level and published to the learners before the commencement of the semester.

Detailed Syllabus:

Module No.	Module Name	Content	No. of Hours
1	Introduction to Time Series Data	Describe the concept of time series data, including its key components: Trend, Seasonality, Noise, and Cyclicity. Explore the concepts of stationarity and non-stationarity. Discuss how time series analysis is applied across different industries, such as finance, economics, and retail. Additionally, highlight the importance of time series plots and visualization in analyzing the data.	4
2	Time Series Decomposition and Stationarity	Time series decomposition: Additive and Multiplicative models. Decomposing time series into trend, seasonal, and residual components. Identifying stationarity: Augmented Dickey-Fuller (ADF) test. Methods for transforming non-stationary time series into stationary series (e.g., differencing, log transformations). Exploratory Data Analysis (EDA) for time series.	6
3	Classical Time Series Forecasting Methods	Moving Averages: Simple, weighted, and exponential moving averages. Exponential Smoothing: Simple, Double, and Triple Exponential Smoothing. Autoregressive Integrated Moving Average (ARIMA) Models: Understanding AR, MA, and ARMA models. Model building: Identification, Estimation, and Diagnostic checking. Autocorrelation function (ACF) and partial autocorrelation function (PACF). ARIMA vs SARIMA (Seasonal ARIMA)	8
4	Advanced Forecasting Model	Seasonal Decomposition of Time Series (STL). Vector Autoregression (VAR): Multivariate time series analysis. GARCH Models: Forecasting volatility in financial markets. Machine Learning Approaches: Regression-based models for forecasting (e.g., linear regression with lag variables). Neural networks for time series forecasting. Ensemble methods for improved forecasting accuracy.	6
5	Model Evaluation and Selection	Evaluation metrics for forecasting accuracy: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE). Cross-validation techniques for time series (e.g., rolling forecasting origin).	4

		Model diagnostics: Residual analysis, checking for autocorrelation and heteroscedasticity. Model comparison and selection based on performance metrics.	
6	Advanced Topics and Emerging Trends in Time Series Forecasting	Deep learning for time series forecasting (e.g., LSTM, GRU networks). Transfer learning in time series forecasting. Forecasting in the presence of outliers and missing values. Explainable AI in time series forecasting.	2
		Total	**30**

Textbooks:

Sr. No	Textbook Titles	Author/s	Publisher	Edition	Module Nos.
1	Time Series Analysis: Forecasting and Control	G.E.P. Box, G.M. Jenkins, G.C. Reinsel, G.M. Ljung	Wiley	5th Edition, 2015	2, 3, 5
2	Forecasting: Principles and Practice	Rob J. Hyndman, George Athanasopoulos	OTexts (Free Online), Monash Univ.	3rd Edition, 2021	1, 2, 3, 4, 5
3	Time Series Analysis and Its Applications: With R Examples	R.H. Shumway, D.S. Stoffer	Springer	4th Edition, 2017	2, 3, 4, 5
4	The Analysis of Time Series: An Introduction	Chris Chatfield	CRC Press	6th Edition, 2003	1, 2, 3

Reference Books:

Sr. No	Textbook Titles	Author/s	Publisher	Edition	Module Nos.
1	Analysis of Financial Time Series	Ruey S. Tsay	Wiley	3rd Edition, 2010	3, 4, 5
2	Forecasting: Methods and Applications	Spyros G. Makridakis, Steven C. Wheelwright, R.J. Hyndman	Wiley	3rd Edition, 1998	1, 3, 5
3	Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model (Journal)	G. Zhang	International Journal of Forecasting	Vol. 19, 2003	4, 6
4	Deep Learning for Time Series Forecasting (Research/Articles)	Various Authors	IEEE / Springer / arXiv	--	6

Course Name: Time Series and Forecasting LAB

Course Code: HMEC05P

Category: Honours (Next-Gen DS)

Preamble:

The Time Series and Forecasting Lab equips students with practical skills to analyze and model time-dependent data using statistical and machine learning techniques. It provides hands-on experience with tools like Python and R to explore trends, seasonality, and forecasting models such as ARIMA and neural networks. The lab enhances analytical thinking by applying theory to real-world datasets across various domains.

Pre-requisite:

- Proficiency in Python programming
- Basic understanding of statistics and probability
- Familiarity with data analysis libraries (Pandas, NumPy)
- Knowledge of data visualization (Matplotlib, Seaborn)

Course Objectives:

1. Preprocess and visualize time series data
2. Implement classical forecasting methods (ARIMA, Exponential Smoothing)
3. Build machine learning models (LSTMs) for time series forecasting
4. Evaluate model performance using appropriate metrics
5. Handle real-world challenges (missing data, outliers)

Course Outcomes:

After successful completion of the course students will be able to:
CO1: Identify and visualize key components of time series data such as trend, seasonality, and noise.
CO2: Apply techniques to transform non-stationary time series into stationary form for analysis.
CO3: Implement and compare classical forecasting methods including Moving Averages, Exponential Smoothing, and ARIMA.
CO4: Apply advanced forecasting models like GARCH, VAR, and machine learning algorithms for multivariate time series analysis.
CO5: Evaluate forecasting models using accuracy metrics and perform diagnostic checks to validate model performance.
CO6: Analyse real-world time series datasets to derive insights and make data-driven decisions.

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
-	2	-	1

Assessment Guidelines:

Head of Learning	ISA	MSE	ESE	Total
Practical	25	-	25	50

The assessment/evaluation guidelines for the courses of different credits are mentioned in the above table. Notwithstanding the above, each course faculty shall have the choice to decide her/his assessment methodology based on the nature of the course. Faculty may propose the revised assessment methodology for his/her course. However, the revised assessment methodology shall be approved by a panel constituted at institute level and published to the learners before the commencement of the semester.

Suggested list of Practical:

Practical No.	Module No.	Suggested list of Experiments	Topics to be highlighted	CO Map
1	1	Load and visualize time series data (e.g., stock prices, weather data). Plot trends, seasonality, and noise.	Time series components, visualization (Matplotlib/Seaborn).	CO1
2	2	Decompose time series using additive/multiplicative models (statsmodels).	Decomposition, stationarity checks (ADF test).	CO1
3	3	Implement Moving Averages and Exponential Smoothing (Holt-Winters).	Smoothing techniques, parameter tuning.	CO2
4	3	Build ARIMA/SARIMA models (auto_arima, pmdarima).	ACF/PACF plots, model diagnostics.	CO2
5	4	Forecast volatility using GARCH models (arch library).	Financial time series, heteroscedasticity.	CO2
6	4	Train an LSTM model (TensorFlow/Keras) for multi-step forecasting.	Sequence modeling, hyperparameter tuning.	CO3
7	5	Evaluate forecasts using MAE, RMSE, MAPE. Compare ARIMA vs. LSTM.	Error metrics, model selection.	CO4
8	6	Handle missing values and outliers in time series data.	Data imputation, anomaly detection.	CO1

Course Name: Cloud-Native Big Data Analytics (CN-BDA)

Course Code: HMEC06T

Category: Honours (Next-Gen DS)

Preamble:

This course is designed to equip students with the foundational and advanced skills required to leverage cloud-native technologies for scalable and efficient Big Data Analytics solutions. Through a blend of theoretical concepts and hands-on labs, students will explore cloud infrastructure, storage, processing frameworks (like Hadoop and Spark), and AI/ML services, all within modern cloud platforms such as AWS and Azure. Emphasis will be placed on practical implementation, security best practices, and integrating end-to-end analytics workflows to solve real-world business problems, ensuring proficiency in designing, deploying, and managing cloud-native Big Data applications.

Pre-requisites: Cloud Computing Lab

Course Objectives:

- To introduce students to cloud-native principles, Big Data characteristics (5Vs), and cloud service models (IaaS/PaaS/SaaS), enabling them to recall and explain their roles in modern analytics.
- To equip students with hands-on skills to deploy cloud infrastructure (AWS/Azure), configure Big Data services (Hadoop/Spark), and integrate storage/database solutions for scalable analytics.
- To develop the ability to assess the effectiveness of cloud-native AI/ML services, security frameworks (IAM), and processing tools for real-world use cases.
- To enable students to design, build, and deploy end-to-end cloud-native Big Data solutions, incorporating infrastructure, processing, AI/ML, and security best practices.

Course Outcomes:
Learner will be able to:
CO1: Students will recall cloud-native concepts, Big Data characteristics, and cloud service models.
CO2: Students will explain how cloud-native architectures support Big Data processing.
CO3: Students will deploy cloud infrastructure and configure Big Data services.
CO4: Students will evaluate storage solutions and processing frameworks for specific use cases.
CO5: Students will assess security implementations and AI/ML service effectiveness.
CO6: Students will design and implement end-to-end cloud-native Big Data solutions.

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
2	-	2	-

Assessment Guidelines:

Head of Learning	ISA	MSE	ESE	Total
Theory	15	20	40	075

The assessment guidelines for the courses of different credits are mentioned in the above table. Notwithstanding the above, each course faculty shall have the choice to propose her/his assessment methodology based on the nature of the course. However, the proposed assessment methodology shall be approved by a panel constituted at Institute level and published to the learners before the commencement of the semester.

Detailed Syllabus:

Module No	Module name	Content	No of Hours
1	Introduction to Cloud-Native Big Data Analytics	Overview of Cloud Computing concepts, Cloud-Native architectures, and principles Big Data characteristics (Volume, Velocity, Variety, Veracity, Value) Importance of integrating Big Data Analytics with Cloud-native platforms Deployment models and service models (IaaS, PaaS, SaaS)	4
2	Cloud-Native Infrastructure and Virtualization	Virtualization concepts and hosted virtualization (VirtualBox, KVM) Cloud-Native Infrastructure using AWS EC2 and Azure Compute Deploying, managing, and scaling virtual machines in the cloud Hands-on Lab: Setting up Linux/Windows VM on AWS EC2/Azure Compute	5
3	Cloud Storage and Database Services	Storage as a Service (AWS S3, Azure Storage) Database as a Service (AWS RDS, Azure SQL, MongoDB Atlas, Firebase) SQL vs NoSQL databases: CAP theorem, key-value, document, columnar, graph databases Hands-on Lab: CRUD operations in cloud-managed databases (MongoDB Atlas/AWS RDS)	5
4	Cloud-Native Big Data Processing	Hadoop ecosystem: HDFS architecture, MapReduce programming model Apache Spark architecture: RDDs, DataFrames, DAG Stream processing fundamentals: Bloom Filter, Flajolet-Martin algorithm, DGIM algorithm Hands-on Lab: Hadoop and Spark deployment in cloud environments (AWS EMR, Azure HDInsight)	6
5	Cloud-Native AI, ML, and Cognitive Services	Introduction to cloud-based AI/ML services (AWS SageMaker, Azure ML) AI-driven analytics: Predictive analytics, real-time insights Cognitive services: Computer Vision, NLP, ML APIs Hands-on Lab: Deploying and using ML models on cloud	5

Module No	Module name	Content	No of Hours
		platforms	
6	Security, Identity Management, and Capstone	Cloud Security principles and IAM (AWS IAM, Azure AD) Best practices in security, compliance, and access management for Big Data Mini Capstone Project: Develop and deploy an end-to-end cloud-native Big Data Analytics solution covering infrastructure, data storage, processing (Spark/Hadoop), AI/ML integration, security, and visualization on AWS or Azure.	5
Total			**30**

Text Books:

1. Cloud Native Data Center Networking" by Dinesh G. Dutt (O'Reilly, 2020, ISBN 978-1492045607).
2. Big Data: Principles and Best Practices of Scalable Realtime Data Systems" by Nathan Marz & James Warren (Manning, 2015, ISBN 978-1617290343).
3. Cloud Native Data Center Networking" by Dinesh G. Dutt (O'Reilly, 2020, ISBN 978-1492045607).
4. Learning Spark: Lightning-Fast Data Analytics (2nd Ed.)" by Jules S. Damji et al. (O'Reilly, 2020, ISBN 978-1492050045).

Reference Books:

1. Designing Data-Intensive Applications" by Martin Kleppmann (O'Reilly, 2017, ISBN 978-1449373320).
2. Cloud Computing for Science and Engineering" by Ian Foster & Dennis B. Gannon (MIT Press, 2017, ISBN 978-0262037242).

Course Name: Cloud-Native Big Data Analytics (CN-BDA)

Course Code: HMEC06P

Category: Honours (Next-Gen DS)

Preamble:

This lab is designed to equip students with the foundational and advanced skills required to leverage cloud-native technologies for scalable and efficient Big Data Analytics solutions. Through a blend of theoretical concepts and hands-on labs, students will explore cloud infrastructure, storage, processing frameworks (like Hadoop and Spark), and AI/ML services, all within modern cloud platforms such as AWS and Azure. Emphasis will be placed on practical implementation, security best practices, and integrating end-to-end analytics workflows to solve real-world business problems, ensuring proficiency in designing, deploying, and managing cloud-native Big Data applications.

Pre-requisites: Cloud Computing Lab

Course Objectives:

- To introduce students to cloud-native principles, Big Data characteristics (5Vs), and cloud service models (IaaS/PaaS/SaaS), enabling them to recall and explain their roles in modern analytics.
- To equip students with hands-on skills to deploy cloud infrastructure (AWS/Azure), configure Big Data services (Hadoop/Spark), and integrate storage/database solutions for scalable analytics.
- To develop the ability to assess the effectiveness of cloud-native AI/ML services, security frameworks (IAM), and processing tools for real-world use cases.
- To enable students to design, build, and deploy end-to-end cloud-native Big Data solutions, incorporating infrastructure, processing, AI/ML, and security best practices.

Course Outcomes:
Learner will be able to:
CO1: Students will recall cloud-native concepts, Big Data characteristics, and cloud service models.
CO2: Students will explain how cloud-native architectures support Big Data processing.
CO3: Students will deploy cloud infrastructure and configure Big Data services.
CO4: Students will evaluate storage solutions and processing frameworks for specific use cases.
CO5: Students will assess security implementations and AI/ML service effectiveness.
CO6: Students will design and implement end-to-end cloud-native Big Data solutions.

Course Scheme:

Contact Hours		Credits Assigned	
Theory	Practical	Theory	Practical
-	2	-	1

Assessment guidelines:

Head of Learning	ISA	MSE	ESE	Total
Practical	25	-	25	50

Suggested List of Practicals:

Sr No.	Suggested Topic(s)
1.	***Cloud VM Setup*** *Real-world Problem: Provisioning and management of virtual resources for web application deployment.* *Input Required: AWS/Azure account, Linux image.* *Suggested Outcome: Fully functional Linux VM deployed and accessible via SSH. (LO2)*
2.	***Windows Server Cloud Deployment*** *Real-world Problem: Enterprise application deployment requiring Windows infrastructure.* *Input Required: Azure account, Windows Server image.* *Suggested Outcome: Accessible and scalable Windows VM deployed on Azure. (LO2)*
3.	***Cloud Storage for Large Dataset*** *Real-world Problem: Storage solution for healthcare or financial sector's extensive data.* *Input Required: Large synthetic dataset, AWS account.* *Suggested Outcome: Efficient storage and retrieval using AWS S3.*
4.	***Mo Cloud-based Database Management*** *Real-world Problem: Customer database management for e-commerce applications.* *Input Required: MongoDB Atlas account, sample e-commerce dataset.* *Suggested Outcome: CRUD operations performed efficiently via MongoDB Atlas.*
5.	***Hadoop Cluster Deployment*** *Real-world Problem: Analyzing massive web log data.* *Input Required: AWS EMR setup, weblog dataset.* *Suggested Outcome: Operational Hadoop cluster for data analytics.*
6.	***MapReduce Data Processing*** *Real-world Problem: Analyzing textual data for insights.* *Input Required: Textual dataset.* *Suggested Outcome: Accurate word frequency count via MapReduce jobs.*
7.	***Apache Spark Data Analysis*** *Real-world Problem: Transaction analysis for fraud detection.* *Input Required: Transactional data.* *Suggested Outcome: Analytical insights from Spark DataFrame transformations.*
8	***Stream Analytics Algorithms*** *Real-world Problem: Real-time social media data processing.* *Input Required: Simulated social media data stream.* *Suggested Outcome: Effective implementation and evaluation of Bloom Filter and DGIM algorithms.*
9	***Predictive Modeling with SageMaker*** *Real-world Problem: Sales forecasting for retail businesses.* *Input Required: Historical sales dataset.* *Suggested Outcome: Predictive model deployment and evaluation using AWS SageMaker.*
10	***Sentiment Analysis using Azure Cognitive Services*** *Real-world Problem: Analyzing customer reviews.* *Input Required: Dataset of customer reviews.* *Suggested Outcome: Automated sentiment analysis with accurate classification*
11	***Cloud IAM Implementation*** *Real-world Problem: Secure access to cloud resources.* *Input Required: AWS IAM account.* *Suggested Outcome: Properly configured IAM policies and secure access management.*
12	***Capstone Cloud-Native Analytics Project*** *Real-world Problem: Comprehensive analytics solution for business intelligence.* *Input Required: Integrated dataset, cloud services.* *Suggested Outcome: End-to-end solution with infrastructure, analytics, AI, and security components.*