Centre for Artificial Intelligence

Islamic University of Science & Technology

Master of Technology
Artificial Intelligence

Curriculum

for

Master of Technology Artificial Intelligence
Two Year Programme
2024 Onwards

Programme Description

The Master of Technology (M.Tech) Artificial Intelligence (AI) at the Centre for AI, Islamic University of Science & Technology, is a two-year program designed to provide students with comprehensive knowledge and practical skills in the field of AI. This program encompasses a wide range of topics including mathematics, statistics, programming, machine learning, deep learning, and their applications. Emphasis is placed on both theoretical foundations and hands-on experience to prepare students for careers in AI which includes AI in industry and research.

Program Educational Objectives (PEOs)

PEO 1

Focuses on building foundational skills in core areas such as mathematics, statistics, and programming languages.

PEO 2

Aims to provide deep understanding and hands-on skills in Data Science, Machine Learning, Computer Vision, and Natural Language Processing.

PEO 3

Concentrates on the use of state-of-the-art tools and high-performance computing platforms for problem-solving in AI.

PEO 4

Empowers students to tackle AI challenges with a rigorous and analytical mindset, promoting creativity and interdisciplinary collaboration, aligned with NEP 2020.

PEO 5

Inculcates professionalism and ethical foundations, encouraging contributions to societal challenges through AI in line with NEP 2020.

PEO 6

Equips students with industry-relevant skills and expertise to meet the evolving demands of the professional industry.

Program Structure

The proposed M. Tech program would span for two years. The course curriculum is developed in accordance with the AICTE Model Curriculum in consultation with industry experts. It shall be regularly updated to align with the latest advancements and industry requirements. It comprises core concepts, practical skills, and hands-on experience in a balanced blend of theoretical coursework, laboratory exercises, and industry-oriented projects.

A. Definition of Credit

Credit is one of the primary methods used to determine and document that student has met academic requirements. Credits are awarded upon completing and passing a course.

1 hr. Lecture (L) per week	1 credit
1 hr. Tutorial (T) per week	1 credit
2 hrs. Practical (P) per week	1 credit

B. Range and distribution of Credits

Credits earned in the range of 77 and above shall be required for a student to be eligible to get MTech Artificial Intelligence.

Distribution of Credits		
S No	**Semester**	**Total Credits**
1	I	23
2	II	20
3	III	20
4	IV	14
Total Credits		**77**

C. Course Code and Definition

All courses (except Open Electives) are denoted by a seven-digit alphanumeric code (XXXXXXX), three alphabets followed by three numerals, followed by one alphabet.

1. The first three alphabets designate the department teaching the course, i.e., the discipline to which the course belongs, e.g., CAI for Centre for Artificial Intelligence.
2. The first numeral following the three alphabets indicate the level of the course, 1 to 4 for undergraduate 1st to 4th year; 5 to 7 for postgraduate 1st to 3rd year, and 8, 9 for PhD.
3. The next two numerals are the unique identification numbers for the course. Courses running in odd semesters are labeled from 01 to 49 and courses running in even semesters are labeled from 50 to 99.
4. The last alphabet indicates the nature of the course. It is one amongst four choices, C (Core Course), E (Elective (Discipline Centric)), G (Elective (Generic)), F (Foundation Course). Since, open electives are identified by a zero in place of the level numeral (at fourth digit), therefore the last digit does not have significance in their course code, and hence will not be used for the definition of the same.
5. Open Electives have a zero in place of the above level numeral and thereby six digits only.

Course Outline

Semester I						
S NO	Course Title	Course Code	Hours Per Week			Credits
			L	T	P	
1	Introduction to Artificial Intelligence	CAI501C	3	1	0	4
2	Machine Learning	CAI502C	3	1	0	4
3	Mathematical Foundations for Machine Learning	CAI503C	0	0	4	2
4	Machine Learning Lab	CAI504C	3	1	0	4
5	Python Programming Language	CAI505C	3	1	0	4
6	Python Programming Language Lab	CAI506C	0	0	4	2
7	Elective I/ MOOCs*		3	0	0	3
Total Credits			**15**	**4**	**4**	**23**

Semester II						
S NO	Course Title	Course Code	Hours Per Week			Credits
			L	T	P	
1	Optimization Techniques for Machine Learning	CAI550C	3	1	0	4
2	Deep Learning	CAI551C	3	1	0	4
3	Deep Learning Lab	CAI552C	0	0	4	2
4	Applied Statistics And Probability	CAI553C	3	1	0	4
5	Exploratory Data Analysis Lab	CAI554C	3	0	0	3
6	Elective II/ MOOCs*		3	0	0	3
Total Credits			**15**	**3**	**2**	**20**

Semester III						
S NO	Course Title	Course Code	Hours Per Week			Credits
			L	T	P	
1	Deep Reinforcement Learning	CAI601C	3	1	0	4
2	Advanced Deep Learning	CAI602C	3	1	0	4
3	Research Methodology	CAI603C	3	0	0	3
4	Dissertation Work-Phase I	CAI604C	0	0	6	3
5	Elective III/ MOOCs*		3	0	0	3
7	Elective IV/ MOOCs*		3	0	0	3
Total Credits			**12**	**2**	**6**	**20**

Semester IV					
S NO	Course Title	Course Code	Hours Per Week		Credits
			L	P	
1	Dissertation Work-Phase II	CAI655C	0	28	14
Total Credits			**0**	**14**	**14**

* *Subject to proper approval from the departmental monitoring committee (DMC), students will have the option to earn credits from the MOOCs platforms (SWAYAM/NPTEL) in addition to choosing courses from the proposed elective baskets. The MOOCs courses are detached from the semester structure, eliminating the necessity to complete them within the same semester they're introduced.*

List of Electives for the First Year (Semester I)						
S No	Course Title	Course Code	Hours Per Week			Credits
			L	T	P	
1	Soft Computing Techniques	CAI501E	3	0	0	3
2	Big Data Analytics	CAI502E	3	0	0	3
3	Intelligent Information Retrieval	CAI503E	3	0	0	3
4	Pattern Recognition	CAI504E	3	0	0	3
5	Advanced Algorithms and Analysis	CAI505E	3	0	0	3
List of Electives for the First Year (Semester II)						
6	Machine Learning for Signal Processing	CAI550E	3	0	0	3
7	Electronic Design Automation	CAI551E	3	0	0	3
8	Computer Vision	CAI552E	3	0	0	3
9	Fuzzy Logic and its Applications	CAI553E	3	0	0	3
10	Bio-Inspired Computing	CAI554E	3	0	0	3
List of Electives for the Second Year (Semester III)						
11	Statistical Modelling for Computer Sciences	CAI601E	3	0	0	3
12	Data Engineering	CAI602E	3	0	0	3
13	Cognitive Systems	CAI603E	3	0	0	3
14	Digital Imaging Techniques and Analysis	CAI604E	3	0	0	3
15	Quantum Artificial Intelligence	CAI605E	3	0	0	3
16	Natural Language Computing	CAI606E	3	0	0	3
17	MLOps	CAI607E	3	0	0	3
18	Federated Learning	CAI608E	3	0	0	3

Detailed Syllabus

Core Courses

Introduction to Artificial Intelligence

Semester I

Course Code
CAI501C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Acquire foundational knowledge of Artificial Intelligence.
- ✔ Develop proficiency in various AI problem-solving techniques
- ✔ Acquire advanced knowledge in AI representation methods
- ✔ Understand and apply ethical principles in AI design and deployment

Course Content

UNIT I

AI vs. Human Intelligence, Definition of Artificial Intelligence, Narrow AI vs. General AI, Subfields of AI, History and Evolution of AI, Applications and limitations of current AI systems, Ethical and societal challenges of AI

(6 hours)

UNIT II

Intelligent agents: Agents and Environments, the concept of rationality, the nature of environments, structure of agents, problem solving agents, problem formulation.

Search: informed search , uninformed search , Local search, adversarial search, Constraint Satisfaction Problems

(8 hours)

UNIT III

Knowledge representation, Logical representation including predicate and first-order logic, Limitations of logical representation, Semantic networks: structure and usage, Frames: attributes, slots, values, and inheritance, Scripts for event sequences, Rule-based deduction systems: forward and backward chaining, resolution, Conceptual graphs: basics and construction, Reasoning under uncertainty: review of probability, Bayesian networks, Hidden Markov Models (HMM), Challenges in knowledge representation.

(10 hours)

UNIT IV

Expert systems:- Introduction, basic concepts, structure of expert systems, the human element in expert systems how expert systems works, problem areas addressed by expert systems, expert systems success factors, types of expert systems, expert systems and the internet interacts web, knowledge engineering, scope of knowledge, difficulties, in knowledge acquisition methods of knowledge acquisition, selecting an appropriate knowledge acquisition method, societal impacts reasoning in artificial intelligence, inference with rules, model based reasoning, case based reasoning, explanation & meta knowledge inference with uncertainty representing uncertainty

(8 hours)

UNIT V

Introduction to AI ethics, key ethical principles in AI: fairness, accountability, transparency, integrity, sustainability, control, democracy, interoperability, privacy concerns and data protection in AI systems, bias and discrimination in algorithmic decision-making, AI in surveillance and its implications for human rights, governance frameworks for AI (national and international perspectives), role of AI ethics in autonomous and semi-autonomous systems, regulatory and legal challenges in AI deployment, ethical AI design and implementation practices, case studies of ethical dilemmas and resolutions in AI applications.

(10 hours)

Text Books

1. Russell, S. J., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach (3rd ed.). Chapters on Knowledge Representation.

Reference Books

1. Ertel, W. (2018). Introduction to artificial intelligence. Springer.
2. Coeckelbergh, M. (2020). AI ethics. The MIT press essential knowledge series.
3. Wooldridge, M. (2021). A brief history of artificial intelligence: what it is, where we are, and where we are going. Flatiron Books.

Web References

1. SWAYAM NPTEL
   An Introduction to Artificial Intelligence
   By Prof. Mausam
2. MIT OpenCourseWare
   Artificial Intelligence
   By Patrick Henry Winston

Machine Learning

Semester I

Course Code
CAI502C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Acquire foundational knowledge in machine learning principles.
- ✔ Develop skills in decision tree algorithms, instance-based learning, and feature selection techniques.
- ✔ Master probabilistic models, logistic regression and support vector machines.
- ✔ Implement and analyze artificial neural networks, focusing on perceptrons and multilayer architectures.
- ✔ Employ ensemble methods and clustering techniques such as bagging, boosting, random forests, and k-means.

Course Content

UNIT I

Introduction to Machine Learning, Definition of learning systems, Goals and applications of machine learning, Different types of learning paradigms, Hypothesis space and inductive bias, Aspects of developing a learning system: training data, concept representation, function approximation, Basic Machine learning pipeline – training, testing and validation. Evaluation Metrics – accuracy, precision, recall, ROC curve. Cross validation, Overfitting, Under fitting, bias-variance tradeoff, Regularization Theory.

Linear regression, Simple and Multiple Linear regression, Polynomial regression, evaluating regression model.

(10 hours)

UNIT II

Discriminative Learning, Logistic regression, Multi-class Regression, SoftMax Regression, Maximum margin classifiers and Support Vector Machines, Hard and soft margin, Higher dimensional space and Kernel trick. Decision trees – Concept of pure and impure nodes and the measure of impurity using entropy and Gini index.

(8 hours)

UNIT III

Generative Learning: Bayes Classifier and Naïve Bayes Classifier. Maximum Likelihood Estimation, Maximum a Posteriori Estimation. Non-Parameterized Density Estimation: Parzen window and KNN Density estimation.

Clustering: Agglomerative clustering, K-means, Gaussian Mixture models and Expectation Maximization algorithm.

(8 hours)

UNIT IV

Sequence Modelling and Ensemble Methods: Hidden Markov Models, Conditional Random Fields. Bagging & boosting and its impact on bias and variance C5.0 boosting, Random forest, AdaBoost, Gradient Boosting Machines and XGBoost

(6 hours)

UNIT V

Computational learning theory, Introduction to Dimensionality Reduction, The Curse of Dimensionality, Principal Component Analysis (PCA), Singular Value Decomposition (SVD), Linear Discriminant Analysis (LDA), Kernel PCA, t-Distributed Stochastic Neighbor Embedding (t-SNE).

(8 hours)

Text Books

2. Machine Learning. Tom Mitchell. McGraw- Hill, 2010.
3. Aurelien Geron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, Oreilly, March 2017.

Reference Books

4. Alpaydin, Ethem. Introduction to machine learning. MIT press, 2020
5. Kevin P. Murphy, “Machine Learning: A Probabilistic Perspective”, MIT Press
6. Christopher Bishop, “Pattern Recognition and Machine Learning” Springer, 2007.

Web References

3. Coursera DeepLearning.AI
   Machine Learning Specialization
   By Andrew Ng, Geoff Ladwig,
   Aarti Bagul
4. Google for Developers
   Google Machine Learning Education
   By Google

Mathematical Foundations for Machine Learning

Semester I

Course Code
CAI503C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Understand fundamental concepts of linear algebra.
- ✔ Master dimensionality reduction techniques.
- ✔ Acquire advanced skills in calculus and optimization.
- ✔ Understand basic probability concepts and statistical measures.

Course Content

UNIT I

Systems of equations, Linear Dependence and Independence, Vector Spaces and Subspaces, Norm of a Vector, Operations with Vectors (Sum, Difference, Scalar Multiplication), Dot Product, Matrix Operations (Multiplication and Inverse), Determinants, Eigenvalues and Eigenvectors, Special Matrices and Their Properties, Least Squares and Minimum Norm Solutions, Matrix Decomposition Algorithms.

(10 hours)

UNIT II

Derivatives: Basic concepts of calculus, Derivative of common functions, Meaning of e and the derivative of e^x, Derivative of log x, Existence of derivatives, Properties of derivative, Partial derivatives, gradient, directional derivatives, convex function and its properties,chain rule of differentiation.

(6 hours)

UNIT III

Unconstrained and Constrained optimization, Numerical optimization techniques for constrained and unconstrained optimization, Optimization using slope method, Optimization using gradient descent, Adapting Newton's method for optimization, Penalty function method, Lagrange Multiplier Method.

(8 hours)

UNIT IV

Basic concepts of probability, conditional probability, independence, Bayes theorem, measures of central tendency (mean, median, mode), measures of dispersion (variance, standard deviation), quantiles, box-plots, populations and samples

(8 hours)

UNIT V

Random variables, discrete distributions, continuous distributions, joint distributions, marginal and conditional distributions, multivariate normal distribution, Central Limit Theorem, unbiased versus biased estimates, maximum likelihood estimation, confidence intervals, hypothesis testing.

(8 hours)

Text Books

1. M.P. Diesenroth, A. Aldo Faisal, Cheng Soon Ong, Mathematics for Machine Learning, Cambridge University Press, 2020.

2. W. Cheney, Analysis for Applied Mathematics. New York: Springer Science + Business Medias, 2001.

Reference Books

1. W. Cheney, Analysis for Applied Mathematics. New York: Springer Science + Business Medias, 2001.
2. S. Axler, Linear Algebra Done Right (Third Edition). Springer International Publishing, 2015.
3. J. Nocedal and S.J. Wright, Numerical Optimization. New York: Springer Science + Business Media, 2006.
4. J.S. Rosenthal, A First Look at Rigorous Probability Theory (Second Edition). Singapore: World Scientific Publishing, 2006.

Web References

1. SWAYAM NPTEL
   Essential Mathematics for Machine Learning
   By Prof. Sanjeev Kumar, Prof. S. K. Gupta | IIT Roorkee

2. Coursera DeepLearning.AI
   Mathematics for Machine Learning and Data Science Specialization
   By Luis Serrano

Machine Learning Lab

Semester I

Course Code
CAI504C

2 Credits

L	T	P
0	0	4

Course Outcomes

- ✔ Develop expertise in data preprocessing techniques essential for cleaning and preparing datasets for machine learning models.
- ✔ Master feature engineering to enhance model performance by creating, selecting, and transforming features effectively.
- ✔ Apply machine learning algorithms to accurately classify and predict outcomes across diverse datasets.
- ✔ Evaluate and fine-tune machine learning models using appropriate metrics.

LIST OF PRACTICAL EXPERIMENTS

1. **Iris Species Classification**
   Objective: Identify the species of iris plants (setosa, versicolor, or virginica) based on the measurements of their petals and sepals.
   Dataset: The Iris dataset includes 150 records of iris plants, with four features: the lengths and the widths of the sepals and petals.

2. **Titanic Survival Prediction**
   Objective: Predict whether a passenger survived the Titanic disaster based on features such as age, gender, and ticket class.
   Dataset: The Titanic dataset includes passenger data like name, age, gender, ticket class, and survival status.

3. **Boston Housing Price Prediction**
   Objective: Predict the median value of homes in various Boston districts using features like crime rate, property tax rate, and pupil-teacher ratio.
   Dataset: The Boston Housing dataset contains information on various housing attributes along with the median value of homes in various areas of Boston.

4. **Diabetes Progression Prediction**
   Objective: Predict the progression of diabetes in patients one year after baseline using various diagnostic measurements.
   Dataset: The dataset includes measurements such as body mass index, blood sugar levels, and age, collected from a study on diabetes progression.

5. **Spam Email Detection**
   Objective: Classify emails as spam or not spam by analyzing their text content.
   Dataset: The dataset typically consists of a collection of email texts labeled as 'spam' or 'not spam'.

6. **MNIST Handwritten Digit Classification**
   Objective: Recognize handwritten digits (0-9).
   Dataset: The MNIST dataset consists of 70,000 grayscale images, each 28x28 pixels, of handwritten digits. It's divided into a training set of 60,000 images and a test set of 10,000 images.

7. **Movie Recommendation System**
   Objective: Recommend movies to users based on their past viewing history and preferences.
   Dataset: Datasets for this task often include user ratings for various movies, which can be used to learn preference profiles for individual users.

8. **Stock Prices Prediction**
   Objective: Predict future stock prices based on historical price and volume data.
   Dataset: Historical stock prices data, which includes daily opening, closing, highest, and lowest prices, and volume of stocks traded.

9. **Sentiment Analysis of Text**
   Objective: Determine the sentiment (positive, negative, neutral) of a piece of text, such as a tweet or a product review.
   Dataset: This involves datasets containing text data with corresponding sentiment labels.

10. **Fashion-MNIST Classification**

    Objective: Classify images of clothing items into 10 categories (e.g., T-shirts, trousers, shoes).

    Dataset: Fashion-MNIST is a dataset comprising 70,000 grayscale images of 10 fashion categories, with each image being 28x28 pixels.

Text Books

1. Aurelien Geron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, Oreilly, March 2017.

Reference Books

1. Dr. M Gopal, Applied Machine Learning, 1st Edition, McGraw-Hill,2018
2. Aurelien Geron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, Oreilly, March 2017.

Web References

1. freeCodeCamp
   Machine Learning for Everybody
   By Kylie Ying
2. codebasics Youtube
   Machine Learning Tutorial Python
   By Dhaval Patel

Python Programming Language

Semester I

Course Code
CAI505C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Develop a strong foundation in the basic Python programming language.
- ✔ Master object-oriented programming concepts within Python.
- ✔ Gain expertise in data analysis using NumPy and Pandas.
- ✔ Acquire skills to visualize data effectively using Matplotlib and Seaborn.
- ✔ Gain hands-on experience with advanced topics such as design patterns and Python web frameworks.

Course Content

UNIT I

Introduction to Development Environments: Familiarization with Jupyter Notebooks and Python IDEs like PyCharm and Visual Studio Code. Python Basics: Syntax, Variables, Data Types, and Control Structures. Functions and Modules: Defining Functions, Scope, and Importing Modules. Data Structures: Lists, Tuples, Sets, Dictionaries, and Comprehensions. File Handling: Techniques for Reading from and Writing to Files.

(8 hours)

UNIT II

Introduction to OOP Concepts, Classes and Objects, Attributes and Methods, Constructors (__init__ method), Encapsulation, Inheritance, Polymorphism, Special Methods (like __str__ and __repr__), Class and Static Methods, Property Decorators, Composition vs Inheritance

(8 hours)

UNIT III

NumPy: Understanding ndarrays, Data Types and Attributes, Array Creation and Properties, Indexing and Slicing, Array Mathematics (Addition, Subtraction, Scalar Multiplication, Division), Aggregation and Statistical Functions, Array Manipulation (Reshape, Concatenate, Split).

Pandas: Series and DataFrames, Data Importing and Exporting, Data Cleaning and Preparation, Data Manipulation (Indexing, Selection, Filtering), Working with Missing Data, GroupBy Operations, Merging and Joining DataFrames, Reshaping and Pivoting.

(8 hours)

UNIT IV

Basic Plotting with Matplotlib (Line Graphs, Bar Charts, Histograms, Scatter Plots), Customizing Plots (Colors, Labels, Legends), Advanced Plotting Techniques (Subplots, 3D Plots, Interactive

Visualizations), Introduction to Seaborn, Seaborn's Built-In Datasets, Statistical Plotting with Seaborn (Distribution Plots, Categorical Plots, Pair Plots, Heatmaps)

(8 hours)

UNIT V

Iterators and Generators, Decorators, Context Managers, Regular Expressions, Testing and Debugging (using unittest), Virtual Environments, Introduction to Python Web Frameworks (Flask), Design Patterns

(8 hours)

Text Books

1. Downey, A. (2012). Think python. " O'Reilly Media, Inc.".

Reference Books

1. Shaw, Z. A. (2024). Learn Python The Hard Way. Addison-Wesley Professional.
2. Sweigart, A. (2016). Invent your own computer games with python.
3. Barry, P. (2016). Head first Python: A brain-friendly guide. " O'Reilly Media, Inc.".
4. Matthes, E. (2023). Python crash course: A hands-on, project-based introduction to programming.

Web References

1. Code with Mosh
Python Tutorial - Python Full Course for Beginners
By Mosh Hamedani
2. Harvard University
CS50
Introduction to Computer Science
By David J. Malan

Python Programming Language Lab

Semester I

Course Code
CAI506C

2 Credits

L	T	P
0	0	4

Course Outcomes

- ✔ Master Python basics, data manipulation with Pandas, and numerical operations with NumPy.
- ✔ Learn data visualization using Matplotlib and Seaborn.
- ✔ Understand and apply key machine learning algorithms using Scikit-learn.
- ✔ Gain skills in model evaluation, performance metrics and hyperparameter tuning.
- ✔ Get introduced to deep learning with TensorFlow and Keras, culminating in a comprehensive final project.

LIST OF PRACTICAL EXPERIMENTS

1. **Basic Syntax and Script Writing**
   Experiment: Write a simple Python script that takes user input, processes it, and outputs a result, such as a script that calculates the area of a circle given its radius.
2. **Data Types and Variables**
   Experiment: Create variables of different data types (integer, float, string, list, tuple, dictionary) and perform basic operations on them, like adding numbers or concatenating/joining strings.
3. **Control Flow**
   Experiment: Write programs that use if, elif, and else statements to make decisions, and use for and while loops to iterate over sequences or repeat actions until a condition is met.
4. **Functions and Modules**
   Experiment: Define functions to perform specific tasks. Also, learn to use Python modules by importing and using functions from the standard library.
5. **File Handling**
   Experiment: Read from and write to files in Python. Create a script that reads a text file and counts the frequency of each word in the file.
6. **Error Handling and Exceptions**
   Experiment: Write a program that handles different types of exceptions, such as handling division by zero or handling file operations when a file does not exist.
7. **Classes and Object-Oriented Programming**
   Experiment: Create a class representing a simple concept, such as a Book with attributes like title and author, and methods to display book info.

8. **List Comprehensions and Generators**
   Experiment: Use list comprehensions to create lists in a single line of code. For example, create a list of squares of the first 10 natural numbers. Also, experiment with generators to generate an infinite sequence.
9. **Decorators and Higher-Order Functions**
   Experiment: Write decorators to modify existing functions, such as a decorator that logs function calls or measures the execution time of functions.
10. **Regular Expressions**
    Experiment: Use regular expressions to perform complex string matching and extraction, such as extracting all email addresses from a large text.
11. **Web Scraping**
    Experiment: Use libraries like BeautifulSoup or Scrapy to scrape data from web pages.
12. **Web Development**
    Experiment: Build a simple web application using a framework like Flask.
13. **Database Interaction**
    Experiment: Connect to a SQL database using sqlite3 or another database library and perform CRUD operations.
14. **Data Analysis and Visualization**
    Experiment: Use pandas and matplotlib to analyze a dataset and create visualizations like histograms and scatter plots.
15. **Asynchronous Programming**
    Experiment: Write asynchronous code using asyncio to perform multiple tasks concurrently.
16. **Script Packaging and Distribution**
    Experiment: Package a Python script and distribute it as a package/installable module.

Text Books

1. Downey, A. (2012). Think python. " O'Reilly Media, Inc.".

Reference Books

1. Shaw, Z. A. (2024). Learn Python The Hard Way. Addison-Wesley Professional.
2. Sweigart, A. (2016). Invent your own computer games with python.
3. Barry, P. (2016). Head first Python: A brain-friendly guide. " O'Reilly Media, Inc.".
4. Matthes, E. (2023). Python crash course: A hands-on, project-based introduction to programming.

Optimization Techniques for Machine Learning

Semester II

Course Code
CAI550C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ To understand the theory of optimization methods and algorithms developed for solving various types of optimization problems
- ✔ To develop and promote research interest in applying optimization techniques in problems of engineering and technology
- ✔ To apply the mathematical results and numerical techniques of optimization theory to concrete engineering problems.

Course Content

UNIT I

Foundations: convex sets, functions, Taylor series and local function approximation, Numerical methods for computing derivatives, automatic differentiation, Numerical computation of eigenvalues / singular values, conjugates, subdifferentials, weak and strong duality, Introduction to optimisation problems, Optimality conditions, Nonconvex optimality & stationarity, Tractable nonconvex problems,

(10 hours)

UNIT II

Unconstrained optimization: Local search as a general optimization paradigm, Random local search and the curse of dimensionality, Steplength rules and intuition, Formal principles of random local search

(8 hours)

UNIT III

First order methods: Gradient descent - normalized and unnormalized versions, Step size selection and convergence, Conservative theoretical guarantees and linesearch, Formal principles of gradient descent, Steepest descent - norm generalization, Batch / Stochastic gradient descent, Variations on normalized stochastic gradient descent

(8 hours)

UNIT IV

Second order methods: Newton's method basics - normalized and unnormalized versions, Newton's method, stepsize, and backtracking linesearch, Honest adjustments for non-convex functions, Scaling issues with Newton's method, Newton and secant method as zero-finding algorithms, Quasi-Newton Methods, BFGS and L-BFGS algorithms

(6 hours)

UNIT V

Methods of constrained optimization: Projected / proximal gradient algorithms, Barrier and penalty methods, Interior point methods, Primal-dual methods, Duality and Lagrangian

(8 hours)

Text Books

1. Nowozin, S., Wright, S. J., & Sra, S. (Eds.). (2011). Optimization for Machine Learning. MIT Press.

Reference Books

1. Brownlee, J. (2021). Optimization for machine learning. Machine Learning Mastery.
2. Aggarwal, C. C., Aggarwal, L. F., & Lagerstrom-Fife. (2020). Linear algebra and optimization for machine learning (Vol. 156). Springer International Publishing.

Deep Learning

Semester II

Course Code
CAI551C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Able to introduce deep learning and application of modern neural networks.
- ✔ Understand the deep learning algorithms extract layered representations of data.
- ✔ Utilize Deep Neural Networks for Image Analysis.
- ✔ Explore the Foundations and Applications of Deep Learning, the Driving Force Behind Recent AI Advancements.

Course Content

UNIT I

History of Deep Learning, McCulloch Pitts Neuron, Perceptrons, Perceptron Learning Algorithm, Multilayer Perceptrons (MLPs), Representation

Power of MLPs, Sigmoid Neurons, Gradient Descent, Feedforward Neural Networks (FFNs), Representation Power of FFNs, Backpropagation

(8 hours)

UNIT II

Optimization algorithms and activation functions: Gradient Descent (GD), Momentum based GD, stochastic GD, mini-batch GD, Adagrad, RMSProp, Adam.

Initialization techniques: Xavier and He initialization.

Regularization: Bias Variance Tradeoff, L2 regularization, Early stopping, Dataset augmentation, Parameter sharing and tying, Injecting noise at input, Dropout, Batch Normalization,

(8 hours)

UNIT III

Convolutional Neural Networks (CNN): Convolution operation, filters, Padding and Stride, Sparse Connectivity and Weight Sharing, Max Pooling and NonLinearities.

Transfer Learning and pretrained CNN architectures: AlexNet, ZFNet, VGGNet, GoogleNet, ResNet. Batch Normalization, Dropout.

(8 hours)

UNIT IV

Basic Concepts in Object Detection**:** Bounding box and annotation techniques, Non-maximum suppression (NMS), R-CNN and its evolution (Fast R-CNN, Faster R-CNN), You Only Look Once (YOLO) series, Single Shot MultiBox Detector (SSD)

Semantic Segmentation, U-Net and its variants image segmentation, SegNet and its architecture, Instance and Panoptic Segmentation, Mask R-CNN for instance segmentation, Metrics for performance evaluation (mAP for detection, IoU for segmentation)

(8 hours)

UNIT V

Recurrent Neural Networks (RNN): Sequence Learning problems, Intuition behind RNN, sequence classification, sequence labeling, Model, Loss function, Learning algorithm, Evaluation. Vanishing and Exploding gradient. LSTMs and GRUs, Encoder Decoder models, Attention mechanism, Graph Neural Networks (GNNs).

(8 hours)

Text Books

1. Bengio, Yoshua, Ian J. Goodfellow, and Aaron Courville. "Deep learning." An MIT Press book in preparation. (2015) .

Reference Books

1. Bengio, Yoshua. "Learning deep architectures for AI." Foundations and trends in Machine Learning 2.1 (2009): 1127.
2. Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.".
3. Trask, A. W. (2019). Grokking deep learning. Simon and Schuster.

Web References

1. NPTEL
   Deep Learning
   By Mitesh M. Khapra
2. MIT 6.S191
   Introduction to Deep Learning
   By MIT

Deep Learning Lab

Semester II

Course Code
CAI552C

2 Credits

L	T	P
0	0	4

Course Outcomes

- ✔ To Build The Foundation Of Deep Learning.
- ✔ To understand How to Build The Neural Network.
- ✔ To enable students to develop successful machine learning concepts.

LIST OF PRACTICAL EXPERIMENTS

1. **Implementing Gradient Descent Algorithm from Scratch**

   Objective: Understand and implement the gradient descent optimization algorithm to minimize a simple cost function.

2. **Data Preprocessing**

   Objective: Load, reshape, normalize, and preprocess data for a neural network model. This includes converting labels to one-hot encoding

3. **Building and Training Neural Networks using Tesorflow**

   Objective: Build, train, validate, and infer with a neural network using Keras, and learn to save and reload the model

4. **Building and Training Neural Networks using PyTorch**

   Objective: Build, train, validate, and infer with a neural network using PyTorch, and learn to save and reload the model

5. **Binary Classification of Images**

   Objective: Create a CNN that can differentiate between cat and dog images.

   Dataset: Cats vs. Dogs Dataset (commonly found on Kaggle)

6. **Multiclass Classification of Images**

   Objective: Build a simple convolutional neural network (CNN) to classify images

   Datasets: MNIST, Fashion-MNIST and CIFAR-10

7. **Implementing Data Augmentation**

Objective: Apply data augmentation techniques to enhance the training dataset for a neural network, improving model robustness and helping prevent overfitting.

8. **Transfer Learning for Image Classification**
   Objective: Utilize a pre-trained model (like VGG16, ResNet, or MobileNet) as a feature extractor and fine-tune it to classify a new set of images.
   Dataset: Use the Oxford 102 Flowers dataset for flower classification or the Stanford Cars dataset for car classification.

9. **Sentiment Analysis**
   Objective: Train a neural network to classify movie reviews from the IMDB dataset as positive or negative.
   Dataset: IMDB Movie Reviews

10. **Stock Prices Prediction**
    Objective: Build a model using RNNs to predict future stock prices based on historical price data.
    Dataset: Any stock price historical data

11. **Language Detection**
    Objective: Train a neural network to detect the language of a given text snippet.
    Dataset: WiLI-2018, a benchmark dataset for language identification

12. **Generative Adversarial Network (GAN)**
    Objective: Generate digits by training a GAN on Identify the Digits (MNIST) dataset
    Dataset: Identify the Digits (MNIST)

13. **Graph Neural Network (GNN)**
    Objective: Implement and explore basic Graph Neural Network (GNN) architectures to solve problems related to molecular data
    Dataset: MoleculeNet (Tox-21)

Text Books

1. Chollet, François. "Deep Learning with Python," Manning Publications, 2017.

Reference Books

1. Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press.
2. The Elements of Statistical Learning by T. Hastie, R. Tibshirani, and J. Friedman, Springer.
3. Probabilistic Graphical Models. Koller, and N. Friedman, MIT Press.

Applied Statistics and Probability

Semester II

Course Code
CAI553C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Understand the fundamental principles of statistics and their application in engineering, including the design and analysis of experiments and observational studies.
- ✔ Master measures Of Central tendency and dispersion to describe data distributions effectively
- ✔ Gain expertise in key probability distribution and their applications in engineering problems.
- ✔ Develop the ability to conduct and interpret hypothesis testing, including the use of t-tests, F-tests, and Chi-Square tests for data analysis.

Course Content

UNIT I

Basics of Statistics: The Role of Statistics in Engineering, Basic Principles, Retrospective Study, Observational study, Designed Experiments, Observing Processes over time, Mechanistic and Empirical Models, Probability and Probability Models

(8 hours)

UNIT II

Measures of central tendency: mean, median, and mode; Measures of dispersion, Range, Quartile Deviation, Mean Deviation, Standard Deviation, Coefficient of variance, Skewness, Kurtosis.

(8 hours)

UNIT III

Probability Distribution: Sample Spaces and Events, Interpretations and Axioms of Probability, Addition Rules, Conditional Probability, Total Probability. Random Variables, Concept of Random Variable, Bernoulli Distribution, Binomial Distribution, Poisson Distribution, Normal Distribution.

(8 hours)

UNIT IV

Correlation and Regression: Concept and types, Karl Pearson Method, Rank Spearman Method, Least Square Method, Discrete Random Variables and Probability Distributions. Continuous Random Variables and Probability Distributions. Joint Probability Distributions.

(8 hours)

UNIT V

Testing of Hypothesis: Testing of Hypothesis, Null and alternative hypothesis, level of significance, one-tailed and two-tailed tests, tests for large samples (tests for single mean, difference of means, single proportion, difference of proportions), t-test, F- test, Chi-Square Test.

(8 hours)

Text Books

1. Douglas C. Montgomery and George C. Runger, Applied Statistics and Probability for Engineers, John Wiley & Sons. 2016

Reference Books

1. Blake I., An Introduction to Applied Probability, John Wiley & Sons.
2. Yagolam A. M. and Yagolam I. M., Probability and Information, Hindustan Publishing Corporation, Delhi, 1983.
3. Statistics for Management - Richard I Levin, David S Rubin - Prentice Hall India –6thEdn, ISBN-81-203-0893-X.

Exploratory Data Analysis Lab

Semester II

Course Code
CAI554C

1 Credits

L	T	P
0	0	2

Course Outcomes

- ✔ Gain proficiency in Python-based data visualization (Matplotlib, Pandas, Seaborn) and introductory Tableau Desktop skills.
- ✔ Master data aggregation and manipulation techniques, including pivot tables and the split-apply-combine strategy.
- ✔ Acquire foundational skills in creating charts, performing data filtering and sorting in Tableau
- ✔ Learn to analyze relationships between quantitative factors using scatterplots, and regression diagnostics.

LIST OF PRACTICAL EXPERIMENTS

1. **Basic Plotting:**
   Create line plots and scatter plots using Matplotlib.
   Customize plots with labels, legends, and annotations.
2. **Advanced Graph Types:**
   Generate histograms, box plots, and bar charts to visualize distributions and comparisons.
   Explore the use of subplots to display multiple plots in one figure.
3. **Interactive Visualizations:**
   Implement interactive elements such as buttons and sliders to modify plots in real time.
4. **Statistical Data Visualization:**
   Use Seaborn to create visually appealing statistical plots like violin plots, swarm plots, and pair plots.
   Understand how to integrate Seaborn with Matplotlib for enhanced customization.
5. **Facet Grids and Categorical Data Visualization:**
   Employ FacetGrid to create a grid of plots based on a dataset's features.
   Visualize categorical data using bar plots, box plots, and count plots.
6. **Basic Interactive Charts:**
   Create interactive charts like line charts, scatter plots, and area charts using Plotly.
   Explore Plotly's capabilities to modify aspects of the chart interactively.
7. **Complex Interactive Visualizations:**
   Develop complex visualizations such as 3D plots and geographic data maps.
   Use Dash by Plotly to create web-based interactive dashboards.
8. **Integrating Plotly with Web Applications:**
   Learn how to embed Plotly visualizations into web applications.

Explore interactive features that allow users to manipulate data or change visualizations through web interfaces

9. **Summary Statistics and Correlation Analysis**
   Generate summary statistics tables and heatmaps of correlations between variables.
10. **Principal Component Analysis (PCA)**
    Apply PCA to reduce dimensions and visualize the dataset in two or three dimensions.
11. **Missing Data Analysis**
    Visualize and quantify missing data in datasets.
12. **Outlier Detection**
    Use scatter plots and box plots to identify outliers in datasets.
13. **Time Series Analysis**
    Explore autocorrelation and trends over time with line plots and rolling averages.
14. **Distribution Analysis**
    Use kernel density plots and cumulative distribution functions (CDFs) to analyze the underlying distributions of variables.
    Create Q-Q plots to assess the normality of the distributions.
15. **Bivariate Analysis**
    Generate pair grids to visualize relationships between each pair of variables.
    Use bubble charts to represent three variables on two dimensions, with the size of the bubble adding an additional dimension
16. **Multivariate Analysis**
    Conduct cluster analysis to identify natural groupings in the data.
    Visualize high-dimensional data using t-distributed Stochastic Neighbor Embedding (t-SNE).
17. **Text Data Exploration**
    Utilize word clouds to visualize the most frequent terms in textual data.
    Conduct sentiment analysis to gauge the sentiment expressed in text data and visualize the results.
18. **Calculate and visualize PDFs and CDFs to understand the probability distributions of different variables within the dataset**

Text Books

1. Tamara Munzner, Visualization Analysis & Design, CRC Press, 2014.

Reference Books

1. Scott Murray, Interactive Data Visualization for the Web, (2e), O'Reilly, 2017.

2. Wingston Chang, R graphics cookbook, O'Reilly. (2013).
3. Andy Field, Jeremy Miles, and Zoe Field, Discovering Statistics Using R, SAGE Publications Ltd. 2012.

Deep Reinforcement Learning

Semester III

Course Code
CAI601C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Understand the core principles of reinforcement learning
- ✔ Understand the model decision-making process using Markov Decision making processes
- ✔ Understand and apply exploration and exploitation strategies in reinforcement learning
- ✔ Acquire Skills in Monte Carlo methods for practical policy evaluation and control
- ✔ Understand function approximation in the context of Reinforcement Learning

Course Content

UNIT I

Overview of reinforcement learning, Difference between supervised, unsupervised, and reinforcement learning, Components of reinforcement learning (agent, environment, state, action, reward, transaction function, Discount, episode), Bandit problem.

(5 hours)

UNIT II

Understanding Markov Decision Processes (MDPs), Policy, State value function, Action value function, Bellman equation, Dynamic Programming—Policy iteration, Policy Improvement, Value iteration, and Limitations of dynamic programming.

(9 hours)

UNIT III

Exploration and exploitation strategies- Random, Greedy, Epsilon-Greedy, Softmax, UCB.
Monte Carlo Methods: First-Visit Monte Carlo, Every-Visit Monte Carlo, Monte Carlo simulation for policy evaluation.

(9 hours)

UNIT IV

Temporal Difference Learning (TD), n-step TD, Q-learning, SARSA, bootstrapping.

(8 hours)

UNIT V

Function Approximation, linear function approximation, Deep Q-networks (DQN), Double Deep Q-networks (DDQN), Dueling Q-networks (Dueling DQN), Policy gradient methods, and Actor Critic methods.

(9 hours)

Text Books

1. "Reinforcement Learning: An Introduction" by Richard S. Sutton and Andrew G. Barto.

Reference Books

1. "Grokking Deep Reinforcement Learning" by Miguel Morales, 2020
2. Alexander Zai , Brandon Brown, Deep Reinforcement Learning in Action, 2020, 1st Edition, Manning Publications.
3. Mohit Sewak, Deep Reinforcement Learning: Frontiers of Artificial Intelligence, 2019, Springer.
4. Sugiyama, Masashi, Statistical reinforcement learning: modern machine learning, 2015, Chapman and Hall

Advanced Deep Learning

Semester III

Course Code
CAI602C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Acquire knowledge to design advanced deep learning models to address novel challenges in practical applications
- ✔ Understand and apply the concept of Meta Learning
- ✔ Gain skills in implementing Transformer architectures and LLMs
- ✔ Develop the ability to design and train various types of GANs
- ✔ Learn to implement and innovate with Graph Neural Networks

Course Content

UNIT I

Meta-learning: Concepts of Few-Shot Learning, One-Shot Learning, and Zero-Shot Learning, Meta-learning Algorithms (Model-Agnostic Meta-Learning (MAML), Prototypical Networks, Matching Networks), Meta-Reinforcement Learning, Applications of Meta-learning in Real-World Scenarios, Neural Architecture Search, Hyperparameter Optimization, Meta-Learning for Fast Adaptation of Deep Networks, Continuous Adaptation via Meta-Learning in Nonstationary and Competitive Environments, Learning to Generalize: Meta-Learning for Domain Generalization

(8 hours)

UNIT II

Attention Mechanisms and Transformers: Introduction to Attention Mechanisms, Types of Attention (Scaled Dot-Product Attention, Multi-Head Attention), The Architecture of Transformers, Self-Attention Mechanism, Positional Encoding, Encoder-Decoder Structure in Transformers, Application of Transformers in Natural Language Processing (NLP), Fine-Tuning Pretrained Vision Transformers, Large Language Models (LLMs), Applications of LLMs

(8 hours)

UNIT III

Generative Adversarial Networks: Introduction to GANs, Understanding Generators and Discriminators, Theoretical Foundations of GANs, Types of GANs (DCGAN, CGAN, InfoGAN, CycleGAN), Training Stability and Techniques, Loss Functions in GANs, Evaluation Metrics for GANs, StyleGAN, BigGAN, RadialGAN, Relativistic discriminator

(8 hours)

UNIT IV

Learning on graphs: Graph Representation Techniques, Node Embedding Techniques, Graph Neural Networks (GNNs), Spectral Methods for Graph Analysis, Spatial Methods for Graph Learning, Applications of Graph Learning (Social Network Analysis, Recommendation Systems, Fraud Detection), Key Architectures in GNNs Graph Convolutional Networks, Semi-Supervised Classification with Graph Convolutional Networks, Gated Graph Sequence Neural Networks, Graph Attention Networks, GraphSAGE, Inductive Representation Learning on Large Graphs, Hierarchical Graph Representation Learning with Differentiable Pooling, GraphRNN: Generating Realistic Graphs with Deep Auto-regressive Models, Learning Graphical State Transitions

(8 hours)

UNIT V

Introduction to Explainability and Interpretability, Techniques for Model Interpretability, Feature importance methods, permutation importance, Partial dependence plots (PDPs), Local interpretable model-agnostic explanations (LIME), SHapley Additive exPlanations (SHAP), Interpretability in Computer vision (saliency maps and Grad-CAM), Interpretability in Natural language processing (attention mechanisms), Quantitative Testing with Concept Activation Vectors (TCAV), Improving Simple Models with Confidence Profiles

(8 hours)

Text Books

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

Reference Books

1. Vasilev, I. (2019). Advanced Deep Learning with Python: Design and implement advanced next-generation AI solutions using TensorFlow and PyTorch. Packt Publishing Ltd.
2. Trask, A. W. (2019). Grokking deep learning. Simon and Schuster.
3. Aurelien Geron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, Oreilly, March 2017.

Research Methodology

Semester III

Course Code
CAI603C

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Develop a foundational understanding of scientific research
- ✔ Master scientific writing and effective communication
- ✔ Acquire technical skills in scientific documentation
- ✔ Understand the use of databases and research metrics in scientific research

Course Content

UNIT I

Science and Scientific Research: Knowledge and the epistemology of knowledge, deductive and inductive inference, a brief history of scientific ideas, important thinkers and scientific advancements, principles of effective research, self-development in research, the creative process, roles of the problem-solver and problem-creator, identifying and solving new problems, literature survey, developing a research plan, writing research proposals.

(8 hours)

UNIT II

Scientific Writing and Communication: Steps in writing a research report, layout of the research report, writing references and bibliography, importance and planning of effective presentations, how to write good scientific papers, models of the paper writing process, benefits of targeting reputable journals, peer review process, responding to reviewer comments, reviewing papers, identifying publication misconduct, dealing with complaints and appeals with publishers and journals.

(8 hours)

UNIT III

Technical and Scientific Documentation Using LaTeX: Introduction to LaTeX, short history and main attractions of LaTeX, automatic styling according to journal requirements, cross-referencing, writing complex mathematical expressions, typical LaTeX input files, the edit/format/preview process, embedding references, bibliography management using BibTeX, creating presentations using Beamer, introduction to Overleaf, practical sessions on LaTeX usage.

(8 hours)

UNIT IV

Databases: Overview of indexing databases, citation databases like Web of Science and Scopus.

Research Metrics : Understanding metrics including Journal Impact Factor, SNIP, R-Index, CiteScore, h-index, g-index, i10-index, altmetrics.

(8 hours)

Text Books

1. Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches. Sage Publications.

Reference Books

1. Booth, W. C., Colomb, G. G., & Williams, J. M. (2009). The craft of research. University of Chicago press.
2. Schimel, J. (2012). Writing Science: How to Write Papers That Get Cited and Proposals That Get Funded. Oxford University Press.

Detailed Syllabus

Electives

Soft Computing Techniques

Course Code
CAI501E

Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Know about the concepts of Fuzzy logic, crisp logic, fuzzy relation, fuzzy implication rule
- ✔ Know about optimization theory, genetic computing and evolutionary computing.
- ✔ Know about the concepts of the neural network, Perceptron, implementation and training

Course Content

UNIT I

Introduction of soft computing: What is Soft Computing, soft computing vs. hard computing, soft computing paradigms, and applications of soft computing. Basics of Machine Learning. Dealing with Imprecision and Uncertainty- Probabilistic Reasoning- Bayesian network, Pearl's Scheme for Evidential Reasoning, Dempster-Shafer Theory for Uncertainty Management, Certainty Factor based Reasoning

(7 hours)

UNIT II

Neural Networks: Basics of Neural Networks- Neural Network Structure and Function of a single neuron: Biological neuron, artificial neuron, definition of ANN, Taxonomy of neural net, characteristics and applications of ANN, McCulloch Pitt model, different activation functions, Supervised Learning algorithms- Perceptron (Single Layer, Multi-layer), Linear separability, ADALINE, MADALINE, RBF networks , Widrow Hoff, learning rule, Delta learning rule, Back Propagation algorithm, Un-Supervised Learning algorithms- Hebbian Learning, Winner take all.

(9 hours)

UNIT III

Fuzzy Logic: Fuzzy set theory, Fuzzy set versus crisp set, Crisp relation & fuzzy relations, Fuzzy systems: crisp logic, fuzzy logic, introduction & features of membership functions, Fuzzy rule base system: fuzzy propositions, formation, decomposition & aggregation of fuzzy rules, fuzzy reasoning, fuzzy inference systems, Mamdani Fuzzy Models – Sugeno Fuzzy Models, Adaptive Neuro-Fuzzy Inference Systems Architecture.

(6 hours)

UNIT IV

Optimization: Derivative-based Optimization – Descent Methods – The Method of Steepest Descent – Classical Newton's Method, Simulated Annealing, Random Search, Downhill Simplex Search Derivative-free Optimization- Genetic algorithm Fundamentals, basic concepts, working principle, encoding, fitness function, reproduction, Genetic modeling: Inheritance operator, cross over, mutation operator, Generational Cycle, Convergence of GA, Applications & advances in GA, Differences & similarities between GA & other traditional methods.

(8 hours)

Text Books

1. S, Rajasekaran & G.A. Vijayalakshmi Pai, "Neural Networks, Fuzzy Logic & Genetic Algorithms, Synthesis & Applications", PHI Publication.
2. Jyh-Shing Roger Jang, Chuen-Tsai Sun, EijiMizutani, "Neuro-Fuzzy and Soft Computing", Prentice-Hall of India. 1996

Reference Books

1. SAndries P Engelbrecht, Computational Intelligence: An Introduction, Wiley Publications.2007
2. S.N. Sivanandam & S.N. Deepa, "Principles of Soft Computing", Wiley Publications.2018

Big Data Analytics

Course Code
CAI502E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ To work with big data platforms and explore big data analytic techniques for business applications.
- ✔ To design efficient algorithms for mining the data from large volumes.
- ✔ To Analyze the HADOOP and Map Reduce technologies associated with big data analytics.
- ✔ To explore Big Data Applications Using Pig and Hive.

Course Content

UNIT I

Introduction to Big Data Platform – Challenges of Conventional Systems - Intelligent data analysis – Nature of Data - Analytic Processes and Tools - Analysis vs Reporting.

(5 hours)

UNIT II

Introduction To Streams Concepts – Stream Data Model and Architecture - Stream Computing - Sampling Data in a Stream – Filtering Streams – Counting Distinct Elements in a Stream – Estimating Moments – Counting Oneness in a Window – Decaying Window - Real time Analytics Platform (RTAP) Applications - Case Studies - Real Time Sentiment Analysis- Stock Market Predictions.

(10 hours)

UNIT III

History of Hadoop- the Hadoop Distributed File System – Components of Hadoop Analyzing the Data with Hadoop- Scaling Out- Hadoop Streaming- Design of HDFS-Java interfaces to HDFS Basics- Developing a Map Reduce Application-How Map Reduce Works-Anatomy of a Map Reduce Job run-Failures-Job Scheduling-Shuffle and Sort – Task execution - Map Reduce Types and Formats- Map Reduce Features- Hadoop environment.

(10 hours)

UNIT IV

Applications on Big Data Using Pig and Hive – Data processing operators in Pig – Hive services – HiveQL – Querying Data in Hive - fundamentals of HBase and ZooKeeper - IBM InfoSphere BigInsights and Streams.

(5 hours)

Text Books

1. Tom White "Hadoop: The Definitive Guide" Third Edition, O'reilly Media, 2012.
2. Seema Acharya, Subhasini Chellappan,Big Data Analytics, Wiley 2015.

Reference Books

1. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007.
2. Chris Eaton, Dirk DeRoos, Tom Deutsch, George Lapis, Paul Zikopoulos, "Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data", McGrawHill Publishing, 2012.
3. Anand Rajaraman and Jeffrey David Ullman, "Mining of Massive Datasets", CUP, 2012.
4. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", John Wiley& sons, 2012.
5. Glenn J. Myatt, "Making Sense of Data", John Wiley & Sons, 2007.
6. Pete Warden, "Big Data Glossary", O'Reilly, 2011.

Intelligent Information retrieval

Course Code
CAI503E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Describe the genesis and variety of information retrieval situations.
- ✔ Construct a variety of information retrieval models and techniques.
- ✔ Execute methods and principles of information retrieval system.
- ✔ Develop Methods for implementing information retrieval systems.
- ✔ Evaluate the emerging information retrieval practices in library services and on the Web.

Course Content

UNIT I

Fundamentals of IR Systems, Models and Indexing: Overview of IR Systems, Information retrieval using the Boolean model, the dictionary and postings lists, Tolerant retrieval, Automatic Indexing, Index construction and compression, Scoring, Vector space model and term weighting.

Document Representation and Analysis: Statistical Characteristics of Text, Regular Expressions, Text Normalization, Edit Distance, N-Gram Language Models, Naive Bayes and Sentiment Classification-Logistic Regression for Document Analysis

(10 hours)

UNIT II

Query Processing and Evaluation: Basic Query Processing, Data Structure and File Organization for IR, Evaluation in information retrieval-Relevance feedback, User Profiles, Collaborative Filtering and query expansion.

(7 hours)

UNIT III

Retrieval Models: Similarity Measures and Ranking, Boolean Matching, Vector Space Models, Probabilistic Models, XML Retrieval, Language models for information retrieval.

(6 hours)

UNIT IV

Web Search Analysis: Web search basics, web characteristics, index size and estimation, near duplicates and shingling, web crawling, distributing indexes, connectivity servers, link analysis, web as a graph, PageRank, Hubs and authoritative pages, summarization, question answering

(7 hours)

Text Books

1. C. D. Manning, P. Raghavan, and H. Schutze, Introduction to Information Retrieval, Cambridge University Press (2008).

Reference Books

1. Ricardo Baezce Yates, Berthier Ribeiro-Neto, Modern Information Retrieval: The Concepts and Technology behind Search (2ndEd, 2010).
2. Mikhail Klassen, Matthew A. Russell, Mining the Social Web, O'Reilly Media, Inc., 3rd Edition (2019)

Advanced Algorithms and Analysis

Semester I

Course Code
CAI504C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Understand and apply key algorithmic strategies with basic programming skills.
- ✔ Differentiate between computational complexity classes and tackle NP-Complete problems.
- ✔ Employ probabilistic and amortized analysis to evaluate algorithm performance.
- ✔ Design approximation algorithms for NP-hard challenges, including basic parallel computing techniques,

Course Content

UNIT I

Defining Key Terms: Algorithm complexity, Greedy method, Dynamic Programming, Backtracking, Branch-and-bound Techniques; Examples for understanding above techniques; Memory model, linked lists and basic programming skills.

(8 hours)

UNIT II

Overview - Class P - Class NP - NP Hardness - NP Completeness - Cook Levine Theorem - Important NP Complete Problems. Heuristic and Randomized algorithms.

(8 hours)

UNIT III

Use of probabilistic inequalities in analysis, Amortized Analysis - Aggregate Method - Accounting Method - Potential Method, competitive analysis, applications using examples.

(8 hours)

UNIT IV

Point location, Convex hulls and Voronoi diagrams, Arrangements, graph connectivity, Network Flow and Matching: Flow Algorithms - Maximum Flow – Cuts - Maximum Bipartite Matching - Graph partitioning via multi-commodity flow, Karger's Min Cut Algorithm, String matching and document processing algorithms.

(8 hours)

UNIT V

Approximation algorithms for known NP hard problems - Analysis of Approximation Algorithms Use of Linear programming and primal dual; local search heuristics; Parallel algorithms: Basic techniques for sorting, searching, merging, list ranking in PRAMs and Interconnection.

(8 hours)

Text Books

1. Michael T Goodric and Roberto Tamassia, "Algorithm Design: Foundations, Analysis and Internet Examples", John Wiley and Sons, 2002.
2. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein, "Introduction to Algorithms", Third Edition, The MIT Press, 2009.

Reference Books

1. Allan Borodin and Ran El-Yaniv: Online Computation and Competitive Analysis, Cambridge University Press, 2005.
2. Sanjoy Dasgupta, Christos Papadimitriou and Umesh Vazirani, "Algorithms", Tata McGraw-Hill, 2009.
3. RK Ahuja, TL Magnanti and JB Orlin, "Network flows: Theory, Algorithms, and Applications", Prentice Hall Englewood Cliffs, NJ 1993.

Web References

1. freeCodeCamp
   Data Structures and Algorithms with Visualizations – Full Course (Java)
   By Dinesh Varyani
2. Harvard University
   CS 224
   Advanced Algorithms
   By Harvard

Pattern Recognition

Course Code
CAI504E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Recognize the characteristics of machine learning strategies.
- ✔ Apply various supervised learning methods to appropriate problems.
- ✔ Identify and integrate more than one technique to enhance the performance of learning.
- ✔ Create probabilistic and unsupervised learning models handling unknown patterns.
- ✔ Analyze the co-occurrences of the data to find frequent patterns

Course Content

UNIT I

Pattern Classifier: Overview of pattern recognition - Discriminant functions - Supervised learning - Parametric estimation - Maximum likelihood estimation - Bayesian parameter estimation-Perceptron algorithm - LMSE algorithm - Problems with Bayes approach - Pattern classification by distance functions - Minimum distance pattern classifier.

(11 hours)

UNIT II

Unsupervised Classification: Clustering for unsupervised learning and classification - Clustering concept - C-means algorithm – Hierarchical clustering procedures - Graph theoretic approach to pattern clustering - Validity of clustering solutions.

(9 hours)

UNIT III

Structural Pattern Recognition Elements of formal grammars - String generation as pattern description - Recognition of syntactic description - Parsing -Stochastic grammars and applications - Graph based structural representation.

(5 hours)

UNIT IV

Feature Extraction and Selection: Entropy minimization - Karhunen - Loeve transformation - Feature selection through functions approximation - Binary feature selection.

(5 hours)

Text Books

1. Robert Schalkoff, Pattern Recognition: Statistical Structural and Neural Approaches Wiley–India, 2009
2. Theodoridis, S. and K. Koutroumbas, “Pattern Recognition”, Fourth Edition, San Diego, CA: Academic Press, 2009.

Reference Books

1. Robert J.Schalkoff, Pattern Recognition: Statistical, Structural and Neural Approaches, John Wiley & Sons Inc., New York, 1992.
2. Tou and Gonzales, Pattern Recognition Principles, Wesley Publication Company, London, 1974.
3. Duda R.O., and Hart.P.E., Pattern Classification and Scene Analysis, Wiley, New York, 1973.
4. Morton Nadier and Eric Smith P., Pattern Recognition Engineering, John Wiley & Sons, New York, 1993.
5. Christopher M Bishop, Pattern Recognition and Machine Learning. Springer. 2011.

Machine Learning for Signal processing

Course Code
CAI550E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Apply linear algebra and probability theory to solve problems in audio signal processing.
- ✔ Understand and apply machine learning algorithms to tasks in music information retrieval and speech recognition.
- ✔ Develop and evaluate models for various audio processing applications like genre classification, event detection, and speaker diarization.

Course Content

UNIT I

Linear Algebra Refresher: Matrices, vectors, eigenvalues, and eigenvectors. Programming Basics: Python for data manipulation, bash scripting for task automation. Digital Signal Processing for Audio: Time-frequency analysis, filtering, and feature extraction.

(8 hours)

UNIT II

Probability Theory Refresher: Bayesian inference, random variables, distributions. Machine Learning Basics: Supervised and unsupervised learning paradigms, evaluation metrics.

(6 hours)

UNIT III

Music Information Retrieval: Feature extraction from music, similarity measures, retrieval systems. Classification and Tagging: Genre classification, mood detection, auto-tagging systems.

(8 hours)

UNIT IV

Speech Recognition: Feature extraction from speech, hidden Markov models, deep learning approaches. Other Audio Processing Applications: Acoustic event detection, speaker diarization, query by humming, melody estimation.

(8 hours)

Text Books

1. “Automatic Speech Recognition: A Deep Learning Approach”, D. Yu and L. Deng, Springer, 2016.
2. “Pattern Recognition and Machine Learning”, C.M. Bishop, 2nd Edition, Springer, 2011.

Reference Books

1. "Deep Learning", I. Goodfellow, Y, Bengio, A. Courville, MIT Press, 2016.
2. "An Introduction to Audio Content Analysis", A. Lerch, Wiley-IEEE Press, 2012.
3. "Speech and audio signal processing: processing and perception of speech and music", B. Gold, N. Morgan, D. Ellis, Wiley, 2011
4. "Signal Processing Methods for Music Transcription", A. Klapuri and M. Davy, Springer, 2007.

Electronic Design Automation

Course Code
CAI551E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Understand the Foundations of EDA.
- ✔ Apply General-Purpose Methods for Combinatorial Optimization.
- ✔ Model and Simulate Electronic Designs
- ✔ Conduct High-Level Synthesis and Understand FPGA Physical Design Automation

Course Content

UNIT I

PRELIMINARIES: Introduction to design methodologies, Design Automation tools, Algorithmic Graph theory, Computational complexity, tractable and intractable problems.

(4 hours)

UNIT II

GENERAL PURPOSE METHODS FOR COMBINATORIAL OPTIMIZATION: Backtracking branch and bound, Dynamic programming, Integer linear programming, Local search, simulated annealing, tabu search, Genetic Algorithms.

Layout compaction, Placement, Floor planning and Routing problems, Concepts and algorithms.

(8 hours)

UNIT III

MODELING AND SIMULATION: Gate level modeling and simulation, Switch level modeling and simulations.

LOGIC SYNTHESIS AND VERIFICATION: Basic issues and terminology, Binary decision diagrams, two–level logic synthesis.

(8 hours)

UNIT IV

HIGH–LEVEL SYNTHESIS: Hardware models, Internal representation of the input algorithm, Allocation assignment and scheduling, some Scheduling Algorithms, some suspects of Assignment problem, High level transformations.

PHYSICAL DESIGN AUTOMATION OF FPGA TECHNOLOGIES, Physical design cycle for FPGAs, Partitioning and Routing for segmented and staggered models.

(10 hours)

Text Books

1. Algorithms for VLSI design, S.H Gerez, WILEY student edition, John Wiley & Sons (Asia) Pvt. Ltd.,1999.

Reference Books

1. Algorithms for VLSI physical design automation, 3rd edition, Naveed sherwani, springer international edition, 2005
2. Computer aided logical design with emphasis on VLSI-Hill & Peterson, Wiley.
3. M.J.S.Smith, “Application Specific Integrated Circuits”,Pearson, 2008.

Computer Vision

Course Code
CAI552E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Understand foundational principles of computer vision
- ✔ Apply techniques in image segmentation and recognition, extracting meaningful patterns from visual data through shape and texture analysis.
- ✔ Utilize computer vision methods for analyzing and managing real-world video data, employing generative models for innovative visual content creation.

Course Content

UNIT I

Introduction to Computer Vision, The Four Rs of Computer Vision, Low-level vs High-level processing, Two View Geometry, Binocular Stereopsis, Camera and Epipolar Geometry Image Formation: Planar Scenes and Homography, Depth Estimation, Robust Correspondence Image Representation: Fourier Transform, Feature Detection, Edge Detection, Local Binary Patterns Recognition: Pyramid Matching, Part-based recognition models.

(8 hours)

UNIT II

Evolution of CNN Architectures: AlexNet, MobileNet, InceptionNets, ResNets, DenseNets Advanced Neural Networks: SIFT & Single Object Recognition, Dense Neural Networks Image Enhancements: Image Quality Enhancement, Image Restoration, Super-resolution Machine Learning Techniques: Unsupervised Learning, Reinforcement Learning in Vision, Salient Detection.

(8 hours)

UNIT III

Image Segmentation: Supervised Segmentation, Agglomerative Clustering, UNet, FCN Recognition and Description: Dense Descriptors, Optical Flow & Tracking, Visual Matching: Bag-of-words Shape and Texture Analysis: Shape from Texture, Color, Motion and Edges, Face Detection.

(6 hours)

UNIT IV

Video Analytics: Crowd Analysis, Video Surveillance, Traffic Monitoring Deep Generative Models: GANs, VAEs, Zero-shot, One-shot Learning Recognition and Retrieval: Content-based Image Retrieval, Instance Recognition Anomaly Detection and Recognition: Anomalous Action Recognition, Post Estimation.

(6 hours)

Text Books

1. Computer Vision: Algorithms and Applications, 2nd ed.Richard Szeliski, The University of Washington,2022

Reference Books

1. Computer Vision: A Modern Approach (Second Edition) by David Forsyth and Jean Ponce
2. D. A. Forsyth and J. Ponce, Computer Vision, A Modern Approach, Pearson Education, 2003.

Fuzzy Logic and Its Applications

Course Code
CAI553E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Understand the fundamentals of Fuzzy logic and its applications
- ✔ Apply the concepts of fuzzy logic to solve real-world problems
- ✔ Design fuzzy systems for various engineering applications
- ✔ Analyze the performance of fuzzy systems

Course Content

UNIT I

Different faces of imprecision - inexactness, Ambiguity, Undecidability, Fuzzyness and certainty, Fuzzy sets and crisp sets, Probability and fuzzy logic, Fuzzy control and knowledge based systems

(7 hours)

UNIT II

Impressive concepts, Fuzzyness and imprecision, Properties of fuzzy sets, Fuzzy representation, Conventional set operations, Intersection of Fuzzy sets, Union of fuzzy sets, the complement of fuzzy sets

(7 hours)

UNIT III

Linguistic variables, Fuzzy propositions, Fuzzy compositional rules of inference-the-Min-Max rules implications and fuzzy additive rules of implication, Methods of decompositions and defuzzification -composite moments, composite maximum average of maximum values and center of maximums

(8 hours)

UNIT IV

Direct and Indirect methods with single and multiple experts, Construction from sample data - Least square method, adaptive fuzzy controllers - membership function tuning using gradient descent

(8 hours)

Text Books

1. Zimmermann H.J., ‘Fuzzy Set Theory - and its Applications’, Springer, 4th Ed., 2007
2. Timothy J. Ross, ‘Fuzzy Logic with Engineering Applications’, Wiley Publications, 4th Ed., 2016

Reference Books

1. John Yen, Reza Langari, ‘Fuzzy Logic, Intelligence, Control & Information’, Pearson Education Inc., India, 2007
2. Zdenko Kovacic, Stjepan Bogdan, ‘Fuzzy Controller Design Theory and Applications’, CRC Press, 1st Ed., 2006
3. Riza C. Berkaan, Sheldon L. Trubatch, ‘Fuzzy Systems Design Principles – Building Fuzzy IF THEN Rule Based’, IEEE Press, 1997
4. George J Klir and Bo Yuan, ‘Fuzzy Sets and Fuzzy Logic: Theory and A: Theory and Applications’, Pearson, 2015
5. M. Mitchell, ‘Introduction to Genetic Algorithms’, Indian Reprint, MIT press Cambridge, 2nd Ed, 2014

Bio-Inspired Computing

Semester II

Course Code
CAI554C

4 Credits

L	T	P
3	1	0

Course Outcomes

- ✔ Understand basic concepts of evolutionary computing.
- ✔ Understand basic features of neural and immune systems and be able to build the neural network
- ✔ Explain how complex and functional high-level phenomena emerge from low-level interactions.
- ✔ Explain the computational processes derived from neural models.
- ✔ Implement simple bio-inspired algorithms like genetic and Particle Swarm Optimization

Course Content

UNIT I

INTRODUCTION TO EVOLUTIONARY ALGORITHM: Evolutionary algorithm, components of evolutionary algorithm representation (definition of individuals), Evaluation function (Fitness function), Population, parent selection Mechanism, Variation Operators, Survivor Selection Mechanism (Replacement), Initialization, Termination Condition, evolutionary algorithm case study Cellular systems, cellular automata, modeling with cellular systems, other cellular systems, computation with cellular systems, artificial life: analysis and synthesis of cellular systems.

(6 hours)

UNIT II

NEURAL SYSTEMS: Biological nervous systems, artificial neural networks, neuron models, architecture, signal encoding, synaptic plasticity, unsupervised learning, supervised learning, reinforcement learning, evolution of neural networks, hybrid neural systems, case study.

DEVELOPMENTAL AND IMMUNE SYSTEMS: Rewriting system, synthesis of developmental system, evolutionary rewriting systems, evolutionary developmental programs, biological immune systems, lessons for artificial immune systems, algorithms and applications, shape space, negative selection algorithm, clonal selection algorithm.

(7 hours)

UNIT III

BEHAVIORAL SYSTEMS: Behavior is cognitive science, behavior in AI, behavior-based robotics, biological inspiration for robots, robots as biological models, robot learning, evolution of behavioral systems, learning in behavioral systems, co-evolution of body and control, towards self-reproduction, simulation and reality.

(6 hours)

UNIT IV

GENETIC ALGORITHMS: Representation of Individuals, Mutation, Recombination, Population Models, Parent Selection, Survivor Selection, Example Application: Solving a Job Shop Scheduling Problem

HYBRIDIZATION WITH OTHER TECHNIQUES: MEMETIC ALGORITHMS: Introduction to Local Search, Lamarckianism and the Baldwin Effect, Structure of a Memetic Algorithm, Heuristic or Intelligent Initialization, Hybridization within Variation Operators: Intelligent Crossover and Mutation, Local Search Acting on the output from Variation Operators, Hybridization During the Genotype to Phenotype Mapping, Design Issues for Memetic Algorithms.

(10 hours)

UNIT V

COLLECTIVE SYSTEMS: Biological self-organization, Particle Swarm Optimization (PSO), ant colony optimization (ACO), swarm robotics, co-evolutionary dynamics, artificial evolution of competing systems, artificial evolution of cooperation, case study.

(8 hours)

Text Books

1. D. Floreano and C. Mattiussi, “Bio-Inspired Artificial Intelligence”, MIT Press, 2008.
2. Tao Song, Pan Zheng, Mou Ling Dennis Wong, Xun Wang, “Bio-Inspired Computing Models and Algorithms”, ISBN: 978-981-3143-19-7, world scientific, 2019 F.
3. Neumann and C. Witt, “Bioinspired Computation in combinatorial optimization: Algorithms and their computational complexity”, Springer, 2010.

Reference Books

1. D. Floreano and C. Mattiussi, "Bio-Inspired Artificial Intelligence", MIT Press, 2008.
2. Tao Song, Pan Zheng, Mou Ling Dennis Wong, Xun Wang, “Bio-Inspired Computing Models and Algorithms”, ISBN: 978-981-3143-19-7, world scientific, 2019 F.
3. Neumann and C. Witt, “Bioinspired Computation in combinatorial optimization: Algorithms and their computational complexity”, Springer, 2010.
4. D. E. Goldberg, “Genetic algorithms in search, optimization, and machine learning”, Addison- Wesley, 1989.
5. Simon O. Haykin, “Neural Networks and Learning Machines”, Third Edition, Prentice Hall, 2008

Statistical Modelling for Computer Sciences

Course Code
CAI601E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ To implement statistical analysis techniques for solving practical problems.
- ✔ To perform statistical analysis on a variety of data.
- ✔ To perform appropriate statistical tests and visualize the outcome

Course Content

UNIT I

Probability Theory: Sample Spaces- Events - Axioms – Counting - Conditional Probability and Bayes' Theorem – The Binomial Theorem – Random variable and distributions: Mean and Variance of a Random Variable-Binomial-Poisson-Exponential and Normal distributions. Curve Fitting and Principles of Least Squares- Regression and correlation.

(8 hours)

UNIT II

Sampling Distributions & Descriptive Statistics: The Central Limit Theorem, distributions of the sample mean and the sample variance for a normal population, Sampling distributions (Chi-Square, t, F, z). Test of Hypothesis- Testing for Attributes – Mean of Normal Population – One-tailed and two-tailed tests, F-test and Chi-Square test - - Analysis of variance ANOVA – One way and two way classifications.

(8 hours)

UNIT III

Tabular data- Power and the computation of sample size- Advanced data handling- Multiple regression- Linear models- Logistic regression- Rates and Poisson regression- Nonlinear curve fitting.

(6 hours)

UNIT IV

Density Estimation- Recursive Partitioning- Smoothers and Generalized Additive Models- Survivals Analysis- Analyzing Longitudinal Data- Simultaneous Inference and Multiple Comparisons- Meta-Analysis- Principal Component Analysis- Multidimensional Scaling- Cluster Analysis

(8 hours)

Text Books

1. Dalgaard, Peter, "Introductory statistics with R", Springer Science & Business Media, 2008.
2. Brain S. Everitt, "A Handbook of Statistical Analysis Using R", Second Edition,LLC, 2014.

Reference Books

1. Richard Cotton, "Learning R", O'Reilly, 2013.
2. Samir Madhavan, "Mastering Python for Data Science", Packt, 2015.
3. Sheldon M. Ross,"Introduction to Probability and Statistics for Engineers and Scientists", 4th edition, Academic Press; 2009.
4. Paul Teetor, "R Cookbook, O'Reilly, 2011.

Data Engineering

Course Code
CAI602E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Master key data storage and processing technologies
- ✔ Develop expertise in designing and implementing robust ETL pipelines
- ✔ Gain proficiency in data orchestration

Course Content

UNIT I

Overview of Data Engineering: Definition, importance, and role in business and technology.

Data Systems and Architecture: Introduction to databases, Modern Data Ecosystem, data warehouses, data lakes, and data marts. Data Engineering vs. Data Science: Distinctions between the roles and how they complement each other.

(6 hours)

UNIT II

Data Storage and Processing: Types of Data, Understanding Different Types of File Formats

Relational Databases (RDBMS): Principles, design, and SQL.

NoSQL Databases: Types (Key-Value, Document, Columnar, Graph), use cases, and examples.

Foundations of Big Data: Hadoop ecosystem, Apache Spark, HDFS and Hive. Data Modeling: Concepts of normalization, denormalization, schema design, and data partitioning.

(8 hours)

UNIT III

ETL, ELT, and Data Pipelines, Data Integration and Data Integration Platforms

Data Extraction Techniques: APIs, web scraping, database queries.

Data Transformation: Data cleaning, validation, aggregation, and transformation techniques.

Data Loading: Techniques for efficient data loading into various storage systems.

(8 hours)

UNIT IV

Architecting the Data Platform, Factors for Selecting and Designing Data Stores, Importance of Data Security, How to Gather and Import Data, Data Wrangling, Tools for Data Wrangling, Querying and Analyzing Data. Data Orchestration and Monitoring: Workflow Management: Apache Airflow. Scheduling and Automation: Techniques and tools for automating and scheduling data pipelines. Monitoring and Logging: Best practices for monitoring data pipelines, logging, and error handling. (8 hours)

Text Books

1. Reis, J., & Housley, M. (2022). Fundamentals of Data Engineering. " O'Reilly Media, Inc.".

Reference Books

1. Kleppmann, M. (2017). Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems. " O'Reilly Media, Inc.".
2. Warren, J., & Marz, N. (2015). Big Data: Principles and best practices of scalable realtime data systems. Simon and Schuster.

Cognitive Systems

Course Code
CAI603E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Understand the basics Of Cognitive Computing and its differences from traditional approaches to Computing.
- ✔ Plan and use the primary tools associated with cognitive computing.
- ✔ Plan and execute a project that leverages Cognitive Computing.

Course Content

UNIT I

Cognitive science and cognitive Computing with AI, Cognitive Computing - Cognitive Psychology The Architecture of the Mind - The Nature of Cognitive Psychology – Cognitive architecture Cognitive processes – The Cognitive Modeling Paradigms - Declarative / Logic based Computational cognitive modeling – connectionist models – Bayesian models. Introduction to Knowledge-Based AI – Human Cognition on AI – Cognitive Architectures

(8 hours)

UNIT II

Decision making, Fuzzy Cognitive Maps, Learning algorithms: Nonlinear Hebbian Learning – Data driven NHL - Hybrid learning, Fuzzy Gray cognitive maps, Dynamic Random fuzzy cognitive Maps.

(8 hours)

UNIT III

Machine learning Techniques for cognitive decision making – Hypothesis Generation and Scoring - Natural Language Processing - Representing Knowledge - Taxonomies and Ontologies - Deep Learning.

(7 hours)

UNIT IV

Cognitive Systems in health care – Cognitive Assistant for visually impaired – AI for cancer detection, Predictive Analytics - Text Analytics - Image Analytics -Speech Analytics – IBM Watson

(7 hours)

Text Books

1. Hurwitz, Kaufman, and Bowles, “Cognitive Computing and Big Data Analytics”, Wiley, Indianapolis.
2. Neil Stillings, Steven E. Weisler, Christopher H. Chase and Mark H. Feinstein, “Cognitive Science: An Introduction”, MITPress.

Reference Books

1. Peter Fingar, Cognitive Computing: A Brief Guide for Game Changers, PHI Publication, 2015.
2. Gerardus Blokdyk ,Cognitive Computing Complete Self-Assessment Guide, 2018
3. Rob High, Tanmay Bakshi, Cognitive Computing with IBM Watson: Build smart applications using Artificial Intelligence as a service, IBM Book Series, 2019

Digital Imaging Techniques and Analysis

Course Code
CAI604E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Ascertain and describe the essentials of image processing concepts through mathematical interpretation.
- ✔ Experiment with various image segmentation and morphological operations for a meaningful partition of objects.
- ✔ Design the various basic feature extraction and selection procedures for various image applications.

Course Content

UNIT I

INTRODUCTION TO IMAGE PROCESSING: Introduction, Digital Image Fundamentals, image acquisition and display using digital devices - Human visual perception, properties – Image Formation - Image sampling and quantization-Basic relationship between pixels.

(5 hours)

UNIT II

IMAGE ENHANCEMENT: Image enhancement in the spatial domain: basic gray level transformation, Histogram Processing- Enhancement using arithmetic/Logic operations, Spatial filtering: smoothing and sharpening. Image enhancement in the frequency domain: Introduction to two-dimensional transforms- Discrete Fourier Transform, Discrete Cosine Transform, Haar Transform, Discrete Wavelet Transform - smoothing frequency domain filtering-sharpening frequency domain filtering.

(11 hours)

UNIT III

MORPHOLOGICAL IMAGE PROCESSING: Morphological Image Processing: Dilation and Erosion – Opening and Closing – Hit or Miss Transformation – Thinning – Thickening – Skeleton.

(5 hours)

UNIT IV

IMAGE SEGMENTATION: Image Segmentation: Detection of discontinuities- Object Detection Methods, Edge Linking and Boundary Detection, Thresholding Methods, Region Oriented Methods.

FEATURE EXTRACTION: Region of interest (ROI) selection - Feature extraction: Histogram based features - Intensity Features-Color, Shape Features-Local Binary Patterns (LBP), Texture descriptors- Grey Level Occurrence Matrix (GLCM).

(9 hours)

Text Books

1. Rafael C. Gonzalez, Richard E. Woods, “Digital Image Processing”, Pearson Education, Fourth Edition, 2018.

Reference Books

1. S. Sridhar, “Digital Image Processing”, Second Edition, Oxford University, 2016.
2. Anil K. Jain “Fundamentals of Digital Image Processing”, PHI, Learning Private Ltd, 2011.
3. Milan Sonka, Vaciav Hlavac, Roger Boyle, “Image Processing Analysis and Vision”, Fourth Edition, Cengage India, 2017

Quantum Artificial Intelligence

Course Code
CAI605E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Able to solve machine learning problems using quantum computations.
- ✔ Knows the use of quantum logic for data mining and applications.
- ✔ Independently understand and solve multi level problems in machine learning

Course Content

UNIT I

Introduction: Learning theory and data mining, quantum like classical computers.

(6 hours)

UNIT II

Data driven models, supervised and unsupervised learning, generalization performance, ensembles, data dependencies and examples.

(8 hours)

UNIT III

Pattern Recognition and Neural Networks, perception, Hopfield Networks, Feedforward networks, Deep learning, computational complexity.

(8 hours)

UNIT IV

Quantum Pattern Recognition, Quantum Associative Memory, Quantum Perceptron, Quantum Neural Networks, Physical Realizations.

(8 hours)

Text Books

1. Isaac Chuang, Michael Nielsen, Quantum Computation and Quantum Information, 10th Anniversary Edition, Cambridge University Press, 2011.
2. Maria Schuld, Ilya Sinayskiy, Francesco Petruccione, An introduction to quantum machine learning, 2014.

Reference Books

1. P. Wittek, Quantum Machine Learning, Elesevier, Amsterdam 2014.
2. Articles in https://arxiv.org/qu-comp
3. S. Bhattacharyya, I. Pan, A. Mani, S. De, E. Behrman, S. Chakraborti (Eds.), Quantum Machine Learning, Walter de Gruyter, Berlin, 2020

Natural Language Computing

Course Code
CAI606E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ To get introduced to language processing technologies for processing text data.
- ✔ To understand the role of Information Retrieval and Information Extraction in Text Analytics
- ✔ To acquire knowledge on text data analytics using language models.

Course Content

UNIT I

Natural Language Processing – Linguistic Background – Mathematical Foundations - Morphological Analysis-Tokenization - Stemming-Lemmatization – Boundary Determination.

(8 hours)

UNIT II

Reading unstructured data - Representing text data - Part of speech tagging – Syntactic representation -Text similarity - WordNet based similarity- Shallow parsing -Semantic representation.

(8 hours)

UNIT III

Information retrieval and Information extraction - Named Entity Recognition – Relation Identification-Template filling.

(6 hours)

UNIT IV

Language model - Probabilistic Models - n-gram language models- Hidden Markov Model- Topic Modelling - Graph Models -Feature Selection and classifiers -Rule-based Classifiers - Maximum entropy classifier – Clustering-Word and Phrase-based Clustering.

(8 hours)

Text Books

1. Christopher D. Manning and Hinrich Schutze,"Foundations of Statistical Natural Language Processing", MIT Press, 1999.

Reference Books

1. Steven Struhl, “Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence”, Kogan Page, 2015.
2. Matthew A. Russell, “Mining the Social Web”, O'Reilly Media, 2013.
3. Steven Bird, Ewan Klein and Edward Loper, “Natural Language Processing with Python”, 1st Edition, O'Reilly Media, 2009.

MLOps

Course Code
CAI607E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Understand the significance of MLOps
- ✔ Learn to deploy machine learning models using different strategies
- ✔ Gain skills in implementing CI/CD pipelines
- ✔ Develop skills in containerization and manage ML deployments

Course Content

UNIT I

Overview of MLOps and its importance, Difference between MLOps, DevOps, and DataOps, Key principles and best practices, Lifecycle stages: from model development to deployment, Managing data and model versioning, Experiment tracking and model reproducibility

(8 hours)

UNIT II

Techniques for handling large datasets, Data versioning tools like DVC, Integrating data quality checks in workflows, Automating model training processes, Scalable model training techniques, Continuous integration and delivery (CI/CD) for ML models, Evaluation metrics and validation strategies

(8 hours)

UNIT III

Deployment strategies (online, batch, streaming), Model serving frameworks (TensorFlow Serving, TorchServe), Canary deployments and A/B testing, Monitoring model performance in production, Detecting and handling model drift, Strategies for model retraining and updating

(8 hours)

UNIT IV

Containerization with Docker, Kubernetes for scalable ML deployments, Managing compute resources, Implementing feature stores for machine learning, ML workflow orchestration with tools like Kubeflow, Airflow

(8 hours)

Text Books

1. Treveil, M., Omont, N., Stenac, C., Lefevre, K., Phan, D., Zentici, J., ... & Heidmann, L. (2020). Introducing MLOps. O'Reilly Media.

Reference Books

1. Ameisen, E. (2020). Building Machine Learning Powered Applications: Going from Idea to Product. " O'Reilly Media, Inc.".
2. Gift, N., & Deza, A. (2021). Practical MLOps. " O'Reilly Media, Inc.".
3. Huyen, C. (2022). Designing machine learning systems. " O'Reilly Media, Inc.".

Federated Learning

Course Code
CAI608E

3 Credits

L	T	P
3	0	0

Course Outcomes

- ✔ Knowledge of the basic concepts, architecture, and applications of Federated Learning
- ✔ Analyze horizontal federated learning
- ✔ Understand the significance of Federated Learning for Vision, Language, and Recommendation

Course Content

UNIT I

Basics of Machine Learning :Overview of Machine Learning principles, Distributed Machine, Learning: Introduction to distributed systems, need for distributed learning, Data parallelism vs model parallelism, Fundamentals of Federated Learning: Definition and goals of Federated Learning, Key challenges and opportunities, Types of Federated Learning: Cross-device vs. Cross-silo, Vertical vs. Horizontal, Architecture and Communication Protocols: Client-server architecture, Communication protocols and efficiency, Aggregation strategies (e.g., Federated Averaging) Privacy and Security: Differential privacy, Secure multi-party computation, Privacy-preserving techniques in Federated Learning

(8 hours)

UNIT II

Federated Learning Optimization Algorithms: Federated Averaging (FedAvg), Variants of Federated Averaging, Optimization challenges and solutions, Statistical and system heterogenity, Personalization and Adaptation: Personalization Strategies in Federated Learning, Federated meta-learning, Transfer learning in Federated contexts, Robustness and Fault Tolerance: Handling non-IID data, Dealing with unreliable clients and communication failures, Robust aggregation methods, Federated Learning with Limited Resources: Resource-efficient algorithms, Model compression and quantization, Trade-offs between accuracy and efficiency

(8 hours)

UNIT III

Frameworks and Tools: Overview of Federated Learning frameworks (e.g., Flower, TensorFlow Federated, PySyft), Data Preprocessing and Management: Data partitioning and preparation for Federated Learning (IID and non-IID data), Handling heterogeneous data sources, Data privacy and anonymization techniques, Model Training and Evaluation: Training machine learning models in a Federated setup, Evaluation metrics for Federated Learning, Monitoring and debugging Federated Learning processes

(8 hours)

UNIT IV

Case study: Google Gboard, Advanced Topics: Federated transfer learning, Federated analytics, Federated LLM's, poisoning in FL, Decentralised Federated Learning

(8 hours)

Text Books

1. Yang, Q., Liu, Y., Cheng, Y., Kang, Y., Chen, T., & Yu, H. (2019). Federated Learning. Morgan & Claypool Publishers.

Reference Books

1. Jin, Y., Zhu, H., Xu, J., & Chen, Y. (2023). Federated Learning: Fundamentals and Advances. Springer Nature Singapore.

Centre for Artificial Intelligence