Minor in AIML and Data Science

Programming and Data Science

Course Code	Course Name
MI-BT051	Programming and Data Science
Course pre-requisites	

Course Objectives
1. Learn Python programming from basics. 2. Use Python tools to work with simple datasets. 3. Understand basic statistics used in data analysis.

Course Outcomes
Upon successful completion of the course, students should be able to 1. Write Python code for simple data tasks. 2. Use NumPy and Pandas to clean and organize data. 3. Make graphs and charts with Matplotlib and Seaborn. 4. Understand basic statistics like mean and standard deviation.

Course Content		
Module No	**Details**	**Hrs**
1	**Python Programming Basics** • Variables, data types (integers, strings, lists) • Operators, expressions, and basic input/output • Conditional statements (if-else), loops (for, while)	**06**
2	**Functions and File Handling** • Defining and calling functions • Parameters, return values, and scope • Reading from and writing to text and CSV files	**06**
3	**Introduction to NumPy** • Creating arrays, indexing, and slicing • Basic array operations and mathematical functions • Array broadcasting and reshaping	**06**
4	**Data Handling with Pandas** • Creating and using Series and DataFrames	**06**

	• Reading data from CSV, Excel, and JSON files • Filtering, sorting, merging, and grouping data	
5	**Data Visualization** • Introduction to Matplotlib and Seaborn • Creating line plots, bar charts, histograms, and scatter plots • Customizing plots with labels, legends, and titles	**06**
6	**Data Cleaning and Preprocessing** • Identifying and handling missing data • Encoding categorical variables • Scaling and normalizing features	**04**
7	**Basic Statistics and Probability** • Measures of central tendency (mean, median, mode) • Measures of dispersion (range, variance, standard deviation) • Basics of probability and normal distribution	**08**

Text Books
• *Python for Data Analysis* by Wes McKinney • *Think Python* by Allen B. Downey • *Head First Statistics* by Dawn Griffiths

References
• NPTEL: "Python for Data Science" by IIT Madras • Google Colab tutorials, Kaggle beginner Python exercises

CO-PO Mapping:

CO\PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	3	2	2	1	3								1	1
CO2	3	2	2	1	3								1	1
CO3	2	2	2	1	3								1	1
CO4	3	3	2	1	2								1	1

Machine Learning and Statistical Modeling

Course Code	Course Name
MI-BT052	Machine Learning and Statistical Modeling
Course pre-requisites	Linear algebra, Probability and statistics

Course Objectives

1. Learn simple machine learning techniques.
2. Use statistics to test and improve models.
3. Understand supervised and unsupervised learning.

Course Outcomes

Upon successful completion of the course, students should be able to

1. Explain how basic ML models work.
2. Build and test simple models like linear regression and decision trees.
3. Use accuracy and confusion matrix to check model quality.
4. Create a simple ML pipeline using Python.

Course Content		
Module No	**Details**	**Hrs**
1	**Math for Machine Learning** • Basics of vectors, matrices, dot products • Understanding probability and statistics • Cost functions and how models learn	**06**
2	**Linear and Logistic Regression** • Simple and multiple linear regression • Logistic regression for classification • Using scikit-learn to build models	**06**
3	**Tree-Based Models** • Decision trees for classification and regression • Introduction to Random Forest • Basics of Support Vector Machines (SVM)	**06**
4	**Clustering and Dimensionality Reduction** • K-Means and Hierarchical Clustering • DBSCAN for density-based clustering • PCA for reducing data dimensions	**06**

	Model Evaluation	06
5	• Accuracy, precision, recall, F1-score • Confusion matrix, ROC curve, AUC • K-fold cross-validation	
6	**Feature Engineering** • Selecting and extracting features • Handling missing and categorical data • Normalizing and scaling features	06
7	**ML Pipeline and Case Studies** • Steps of an ML pipeline (preprocessing to prediction) • Case studies using open datasets • Python project with train/test split and evaluation	06

Text Books
• *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* by Aurélien Géron • *An Introduction to Statistical Learning* by Gareth James et al.

References
• NPTEL: "Machine Learning" by Prof. P. Abinash (IIT Kharagpur) • scikit-learn official documentation

CO-PO Mapping:

CO\PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	3	2	2	1	3								1	1
CO2	3	2	2	1	3								1	1
CO3	2	2	2	1	3								1	1
CO4	3	3	2	1	3								1	1

Artificial Intelligence and Deep Learning

Course Code	Course Name	
MI-BT053	Artificial Intelligence and Deep Learning	
Course pre-requisites		

Course Objectives

1. Learn what AI is and how it is used.
2. Understand how deep learning works.
3. Use tools like TensorFlow or PyTorch for projects.

Course Outcomes

Upon successful completion of the course, students should be able to

1. Explain search algorithms in AI.
2. Build and train a basic neural network.
3. Use CNNs for images and RNNs for sequences.
4. Use AI tools for small tasks.

Course Content		
Module No	**Details**	**Hrs**
1	**Introduction to AI** • What is Artificial Intelligence? • Examples from everyday life: navigation, voice assistants • Simple AI uses in electrical, civil and mechanical: traffic systems, smart buildings	**06**
2	**Search Algorithms** • Finding best solutions: BFS, DFS in simple terms • Applications: robot movement, path planning in warehouses	**06**
3	**Basics of Neural Networks** • How machines learn from data • Simple concepts: neurons, weights, layers • Use cases: prediction of material strength, energy load	**06**
4	**Multi-layer Perceptron (MLP)** • Training a deeper neural network • Understanding learning rate and epochs • Example: Predicting house prices or machine failure	**06**

5	**CNN for Image Data** • What is image classification? • Simple CNN layers explained • Example: crack detection in concrete images	**06**
6	**RNN for Sequence Data** • Learning from time-based data • Applications: weather forecasting, equipment usage pattern • Basics of LSTM and how memory works in AI	**06**
7	**Tools and Practice** • Using TensorFlow, Keras or PyTorch easily with examples • Mini-projects with real data • How to build a basic app using your model (Streamlit)	**06**

Text Books
• *Deep Learning* by Ian Goodfellow, Yoshua Bengio, and Aaron Courville • *Artificial Intelligence: A Modern Approach* by Stuart Russell & Peter Norvig

References
• *Neural Networks and Deep Learning* by Michael Nielsen • PyTorch & TensorFlow documentation • NPTEL: "Deep Learning" by Prof. Balakrishnan (IIT Madras)

CO-PO Mapping:

CO\PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	3	2	2	1	3								1	1
CO2	3	2	2	1	3								1	1
CO3	2	2	2	1	3								1	1
CO4	3	3	2	1	3								1	1

Advanced Data Science and Applications

Course Code	Course Name
MI-BT054	Advanced Data Science and Applications
Course pre-requisites	

Course Objectives

1. Use advanced ML tools for real-world data.
2. Learn basic big data tools like PySpark.
3. Complete a project with deployment.

Course Outcomes

Upon successful completion of the course, students should be able to

1. Use tools like XGBoost for real data problems.
2. Use PySpark to handle large data.
3. Build a simple recommender system.
4. Make and show a final project using Streamlit.

Course Content

Module No	Details	Hrs
1	**Advanced ML Algorithms** • Introduction to ensemble methods: bagging, boosting • Using XGBoost and LightGBM with scikit-learn • Use cases: defect prediction in structures, energy consumption models	**06**
2	**Big Data Tools Overview** • What is big data? Why is it important? • Basics of Hadoop and Spark • Using PySpark to handle simple large files (CSV/logs)	**06**
3	**Time-Series Forecasting** • Understanding trends and seasonal data • ARIMA and Facebook Prophet for forecasting • Example: weather, load, or demand prediction in infrastructure	**06**
4	**Recommender Systems** • What are recommendation systems? • Collaborative vs. content-based filtering • Simple example: tool or material selection recommender	**06**

5	**Natural Language Processing (NLP)** • What is text data? Simple preprocessing (cleaning, tokenizing) • Text classification: feedback analysis or issue sorting • Tools: NLTK, basic Hugging Face models	**06**
6	**Model Deployment** • Making ML models usable by others • Flask/Streamlit for web-based apps • Packaging with GitHub and Docker (simple intro)	**06**
7	**Project, Ethics and Explainability** • Small project integrating learning from all modules • Concepts of fairness, bias, and transparency in ML • Use of Explainable AI (XAI) tools like SHAP or LIME	**06**

Text Books
• *Data Science from Scratch* by Joel Grus • *Practical Statistics for Data Scientists* by Peter Bruce

References
• *Forecasting: Principles and Practice* by Rob J. Hyndman (free online) • NPTEL: "Big Data Computing" – IIT Patna • Kaggle and UCI ML Repositories (for projects)s

CO-PO Mapping:

CO\PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	3	2	2	1	3								1	1
CO2	3	2	2	1	3								1	1
CO3	2	2	2	1	3								1	1
CO4	3	3	2	1	3								1	1

Suggested Tools and Platforms

- **Languages**: Python (primary), R (optional for stats)
- **IDEs**: Jupyter Notebook, VS Code, Google Colab
- **Libraries**: Pandas, Scikit-Learn, TensorFlow, Keras, PyTorch, Matplotlib, Seaborn, NLTK, SpaCy, XGBoost
- **Datasets**: UCI, Kaggle, OpenML, real-world case data (CSV, JSON)