

**Vidya Pratishthan's Kamalnayan Bajaj Institute of
Engineering and Technology, Baramati.
(An Autonomous Institute)**



Faculty of Science and Technology

Board of Studies

Department of Artificial Intelligence and Data Science

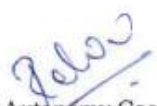
Syllabus
Multidisciplinary Minor (MDM)

(Pattern 2024)
(w.e.f. AY: 2025-26)

Syllabus: Multidisciplinary Minor w. e. f. AY: 2025- 2026

Multidisciplinary Minor in Artificial Intelligence and Data Science

Course Code	Courses Name	Teaching Scheme			Examination Scheme and Marks							Credits			
		TH	PR	TUT	Activity	ISE	ESE	TW	PR	OR	Total	TH	PR	TUT	Total
AI24051	Data Processing and Analysis	2	2	-	10	-	60	30	-	-	100	2	1	-	3
AI24052	Fundamentals of Programming Language	2	2	-	10	-	60	30	-	-	100	2	1	-	3


Dept. Autonomy Coordinator
Mrs. R. S. NaiK


Dept. Academic Coordinator
Mr. P.N. Shendage


HOD, AI&DS
Dr. C.S. Kulkarni


Dean Autonomy
Dr. C. B. Nayak


Dean Academics
Dr. S. M. Bhosle


Principal
Dr. S. B. Lande
Vidya Pratishthan's
Kamainsyan Bajaj Institute of
Engineering & Technology, Baramati
Vidyanagari, Baramati-413133





Vidya Pratishthan's
Kamalnayan Bajaj Institute of Engineering and Technology, Baramati
(Autonomous Institute)

AI24051- Data Processing and Analysis

Teaching Scheme:

Theory: 2 Hours/Week

Practical: 2 Hour/Week

Credits

03

Examination Scheme:

Activity:10 Marks

ESE: 60 Marks

Term Work: 30 Marks

Prerequisites: Python Programming

Course Objectives:

- To understand the need of Data Science
- To understand computational statistics in Data Science
- To provide a comprehensive knowledge of data science using Python.
- To learn the essential concepts of data analytics and data visualization.

Course Outcomes (COs): The students will be able to learn:

CO1: Apply basic data manipulation techniques using Pandas.

CO2: Identify and apply the need and importance of pre-processing techniques.

CO3: Implement data visualization using visualization tools in Python programming.

CO4: Apply various algorithms and evaluate their performance using metrics.

Course Contents

Unit I Introduction to Data Science and Pandas Basics (06 Hours)

Data science: Definition and importance of data science in various industries. Overview of the data science process and the role of a data scientist. Datafication, and data science lifecycle. Ethical considerations and challenges in data science.

Getting Started with Pandas: Overview of Pandas library and its architecture. Introduction to data structures: Series and DataFrames. Indexing, selection, and filtering data. Basic operations: sorting, ranking, reindexing, and handling missing data.

Unit II Statistical Inference & Data Wrangling (6 Hours)

Statistical Inference: Measures of central tendency: Mean, median, mode, and their application in data analysis. Measures of dispersion: Variance, standard deviation, and range. Introduction to Bayes' Theorem and its relevance to data science. Pearson Correlation and its role in understanding relationships between variables.

Data Wrangling: Combining and merging datasets using Pandas. Techniques for reshaping data (pivoting, melting). Handling overlapping data, removing duplicates, and replacing values. Transforming data for analysis (scaling, encoding, and feature extraction).

Unit III Plotting, Visualization & Exploratory Data Analysis (EDA) (6 Hours)

Plotting & Visualization: Importance of data visualization in data science. Overview of different types of data visualizations: Line plots, bar charts, histograms, scatter plots. Introduction to **Matplotlib** and **Seaborn** for data visualization in Python. Plot customization: Titles, labels, colors, legends, and annotations.

Exploratory Data Analysis (EDA): Visualizing distributions, relationships, and trends in the data. Using heatmaps, pairplots, and correlation matrices for EDA. Identifying patterns, outliers, and potential

data issues through visualization.

Unit IV Machine Learning Basics & Model Evaluation (6 Hours)

Introduction to Machine Learning: Overview of machine learning: Types of learning (supervised vs unsupervised). Introduction to classification and regression tasks. Clustering algorithms: K-Means and hierarchical clustering. Evaluation metrics: Accuracy, precision, recall, F1-score.

Model Evaluation & Selection: Cross-validation and its importance in model evaluation. Overfitting and underfitting: Understanding and mitigating these issues. Introduction to Ridge regression for regularization.

Text Books:

1. David Dietrich, Barry Hiller, "Data Science and Big Data Analytics", EMC Education services, Wiley publication, 2012, ISBN0-07-120413-X.
2. Wes McKinney, "Python for Data Analysis", O'REILLY, ISBN:978-1-449-31979-3, 1st edition, October 2012.
3. Rachel Schutt & O'neil, "Doing Data Science", O'REILLY, ISBN:978-1-449-35865-5, 1st edition, October 2013.

Reference Books:

1. Joel Grus, "Data Science from Scratch: First Principles with Python", O'Reilly Media, 2015
2. Matt Harrison, "Learning the Pandas Library: Python Tools for Data Munging, Analysis, and Visualization, O'Reilly, 2016.
3. Chirag Shah, "A Hands-On Introduction to Data Science", Cambridge University Press, (2020), ISBN: 978-1-108-47244-9.
4. Wes McKinney, "Python for Data Analysis", O'Reilly media, ISBN: 978-1-449-31979-3.
5. Trent Haunk, "Scikit-learn Cookbook", Packt Publishing, ISBN: 9781787286382

E-Resources:

1. https://onlinecourses.nptel.ac.in/noc21_cs69/preview
2. <https://nptel.ac.in/courses/106106179>
3. https://onlinecourses.swayam2.ac.in/imb23_mg64/preview

List of Assignments

1. Perform the following operations using Python on any open source dataset (e.g., data.csv)
 - Import all the required Python Libraries.
 - Locate open source data from the web (e.g., <https://www.kaggle.com>). Provide a clear description of the data and its source (i.e., URL of the web site).
 - Load the Dataset into pandas dataframe.
 - Data Preprocessing: check for missing values in the data using pandas isnull(), describe() function to get some initial statistics. Provide variable descriptions. Types of variables etc. Check the dimensions of the data frame.
2. Load the open source dataset () and apply statistical inference.

- Calculate and interpret measures of central tendency (mean, median, mode) and dispersion (variance, standard deviation, range) to summarize data.
- Explore Pearson correlation to assess relationships between variables.
- Apply Bayes' Theorem to solve probability-based problems in data analysis. Load the open source dataset ()

3. Load the open source dataset (e.g.csv) and perform data wrangling with pandas.

- Combine and merge multiple datasets using various Pandas functions.
- Reshape data using pivoting and melting techniques to restructure data for analysis.
- Clean data by handling duplicates, overlapping data, and performing value replacements.
- Scale numerical data and encode categorical variables to prepare datasets for statistical analysis or machine learning.

4. Load the any open source dataset (e.g. csv) and with Matplotlib and seaborn libraries in python.

- Create a various plot such as line plots, bar charts, histograms, and scatter plots.
- Customize plots by adding titles, labels, colors, legends, and annotations to enhance readability and understanding.
- Use visualizations to identify patterns, outliers, and potential data issues in datasets, facilitating deeper insights into the data for decision-making.
- Find the correlation between variables using suitable plot.

5. Implement a simple unsupervised machine learning model (e.g., K-Means clustering and evaluate its performance using cross-validation.

6. Implement a supervised machine learning model (e.g., Logistic Regression or Linear Regression).

- Train the model and evaluate its performance using accuracy or R-squared (for regression).

7. Implement k-fold cross-validation on a regression or classification model (e.g., Linear Regression or Logistic Regression).

- Compare the results of cross-validation with a single train-test split.

8. Choose a classification dataset (e.g., Titanic dataset or Breast Cancer dataset).

- Implement model evaluation metrics for classification (k-means or logistic regression) .
- Evaluate a classification model using accuracy, precision, recall, and F1-score.



Vidya Pratishthan's
Kamalnayan Bajaj Institute of Engineering and Technology, Baramati
(Autonomous Institute)

AI24052: Fundamentals of Programming Language

Teaching Scheme:

TH: 02 Hrs./Week

PR: 02 Hrs./Week

Credit: 03

Examination Scheme:

Activity: 10 Marks

ESE: 60 Marks

Term Work: 30 Marks

Prerequisite: Basic Knowledge of Computers.

Course Objective:

- Learn the structural components of a C Program.
- Develop Problem-Solving Skills Using C.
- Learn data structures like arrays and structures to obtain solutions to solve the problems.
- Learn concepts of modular programming to design the solutions to the problems.

Course Outcomes:

1. Develop C programs utilizing variables, operators and expressions effectively
2. Implement C programs using decision-making constructs, and looping mechanisms to solve computational problems efficiently.
3. Utilize arrays, strings, and structures in C programming to develop efficient and structured solutions.
4. Apply modular programming using function.

Course Contents

Mapping of Course Outcomes for Unit I

CO1

UNIT I

INTRODUCTION TO C PROGRAMMING

06 Hours

Overview of C: History and importance of C, Structure of C program, executing a C program, Algorithms and flowcharts

Constants, Variable and Data Types: Keywords and Identifiers, Constants, Variables, Data types, Declaration of variables, Assigning values to variables, Defining symbolic constants.

Input and Output Operations: Input output statements, formatted input, Formatted output.

Operators and Expressions: Introduction, arithmetic, Relational, Logical, Assignment, Increment and Decrement and Bitwise operators, Arithmetic expressions, Evaluation of expressions, Precedence and Associativity of operators, Type conversions in expressions

Mapping of Course Outcomes for Unit II

CO2

UNIT II

CONTROL STRUCTURES

06 Hours

Decision Making and Branching: Introduction, Decision making with IF statement, Simple IF statement, If-Else statement, Nested if-else statements, The Switch statement, The Conditional operator, The goto statement.
Decision Making and Looping: Introduction, The for statement, The while Statement, The do-while statement, nested loops, break and continue statements

Mapping of Course Outcomes for Unit III	CO3
--	------------

UNIT III	ARRAY AND STRUCTURE	06 Hours
-----------------	----------------------------	-----------------

Characteristics of an array, One dimension and two dimensional arrays, concept of multi-dimensional arrays. Array declaration and Initialization. Operations on Arrays. Character and String input/output and String related operations. Introduction and Features of Structures, Declaration and Initialization of Structures, array of structures.

Mapping of Course Outcomes for Unit IV	CO4
---	------------

UNIT IV	FUNCTIONS	06 Hours
----------------	------------------	-----------------

Concept and need of functions. **Library functions:** Math functions, String handling functions, **User defined functions** - function definition, functions declaration, function call, scope of variables - local variables, global variables. **Function parameters:** Parameter passing- call by value & call by reference.

Books and Other Resources

Reference Books:

1. Kernighan B.W and Dennis M. Ritchie, “The C Programming Language”, 2nd Edition, 2015, Pearson Education India, ISBN: 978-93-3254-944-9.
2. Byron S. Gottfried, “Schaum’s outline of theory and problems of programming with C” 2nd Edition, McGRAW -HILL , ISBN 0-07-024035-3
3. Pradip Dey, Manas Ghosh, “Programming in C”, 2nd Edition, 2018, Oxford University Press, ISBN: 978-01-9949-147-6.
4. Yashavant P. Kanetkar, “Let Us C”, 16th Edition, 2019, BPB Publications, ISBN: 978- 93-8728-449-4.
5. Jacqueline A Jones and Keith Harrow, “Problem Solving with C”, Pearson Education. ISBN: 978-93-325-3800-9.

References:

- <http://www.studytonight.com/c/overview-of-c.php>
- <https://www.tutorialspoint.com/cprogramming>

MOOCs Courses link:

- <http://nptel.ac.in/courses/106105085/2>
- <http://nptel.ac.in/courses/106104074/1>

- <https://nptel.ac.in/courses/106/105/106105171>
- <https://nptel.ac.in/courses/106/106/106106212/>

Guidelines for Term Work Assessment:

Term work assessment will be based on overall performance of Laboratory assignments performed by a student. Each Laboratory assignment assessment will assign grade/marks based on parameters, such as timely completion, performance, efficient codes, and punctuality.

Guidelines for Term Work submission:

Problem statements will be formed based on assignments and performance will be evaluated by Internal and External Examiner. During practical assessment, maximum weightage should be given to satisfactory implementation of the problem statement. Relevant questions may be asked at the time of evaluation to test the student 's understanding of the fundamentals, effective and efficient implementation. All assignments are compulsory.

Guidelines for Laboratory Conduction:

Operating System recommended: - 64-bit Open source Linux or its derivative

Programming tools recommended: - C, Visual Studio Code

Lab Assignments

1. Write a program that takes a student's marks as input and assigns a grade based on the following criteria:

Grade: Distinction If per $\geq$ 75

Grade: A If per $\geq$ 60 and Per $<$ 75

Grade: B If per $\geq$ 55 and Per $<$ 60

Grade: Pass $\geq$ 40 and Per $<$ 55

Grade: Fail if per $<$ 40

2. Write C Program to print following patterns using loops.

```
*
* * *
* * * * *
* * * * * * *
```

3. Write a program to check whether the number is a prime number or not.
4. Write C program to find the largest and smallest element from an array.
5. Write C program to perform addition, transpose and multiplication of two 3X3 matrices using Two Dimensional Array.
6. Create a structure called "Student" with members name, age, and total marks. Write a C program to input data for five students and display the information.
7. Write C program to perform following operations without using standard string functions.
 - i) Calculate Length of given string
 - ii) Print string in the reverse order.
 - iii) Copy one string to other
 - iv) Concatenation
8. Write a function to find the factorial of the number.
9. Write a function to search an element from the array.