

Artificial Intelligence I: Introduction to Data Science and Machine Learning

Instructors:

Assoc. Prof. Dr. Taner Arsan

H. Fuat Alsan, PhD(c)

Sena Kılınç, PhD(c)

Week 2 Assignment

Part 1: NumPy Fundamentals

ndarrays:

1. Create a 1-dimensional NumPy array with elements [1, 2, 3, 4, 5].
2. Create a 2-dimensional NumPy array with shape (3, 4) filled with zeros.
3. Generate a random 1-dimensional array of size 7.

Vectors:

4. Perform elementwise addition of two vectors: [1, 2, 3] and [4, 5, 6].
5. Calculate the dot product of the vectors [2, 3, 1] and [1, 2, 3].

Matrices:

6. Create a 3x3 identity matrix using NumPy.
7. Multiply the matrices: $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$.
8. Transpose the matrix $C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$.

Dot Product:

9. Calculate the dot product of two vectors: [3, 7] and [7, 3].
10. Compute the dot product of vectors [2, 4] and [1, 3].

Elementwise Multiplication:

11. Multiply each element of array $A = [2, 3, 4]$ by 5.
12. Perform elementwise multiplication of arrays $X = [1, 2, 3]$ and $Y = [4, 5, 6]$.

Matrix Multiplication:

13. Multiply the matrices $D = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $E = \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix}$.
14. Perform matrix multiplication of $F = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$ and $G = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Generating ndarrays:

15. Generate a linearly and equally spaced array with 5 elements between 1 and 10.
16. Create an array of 4 evenly spaced values between 0 and 1.

Slicing ndarrays:

17. Extract the second element from the array $H = [10, 20, 30]$.
18. Slice the array $I = [1, 2, 3, 4, 5]$ to get elements from index 2 to 4.

Example Mathematical Functions:

19. Calculate the mean absolute error (MAE) between two arrays: $M = [5, 7, 9]$ and $N = [6, 8, 10]$.
20. Find the square root of each element in array $O = [4, 9, 16]$.

Part 2: Exploratory Data Analysis (EDA) with Pandas and Matplotlib

- Download **Black Friday Dataset** from Kaggle and solve the following questions accordingly.
- <https://www.kaggle.com/datasets/syedhaideralizaidi/black-friday-dataset>

1. Loading and Initial Exploration:

- a. Load the Black Friday Dataset (CSV file) into a Pandas DataFrame.
- b. Display the first 5 rows of the dataset to get a quick overview.
- c. Check the data types of each column in the DataFrame.
- d. Find the total number of rows and columns in the dataset.

2. Descriptive Statistics:

- a. Print the summary statistics (mean, median, min, max, etc.) for the 'Purchase' column.
- b. Identify and print the unique values in the 'Age' column.
- c. Calculate the average purchase amount for each gender.

3. Handling Missing Data:

- a. Identify and count the missing values in each column of the dataset.
- b. Handle missing values in the 'Product_Category_2' and 'Product_Category_3' columns by filling them with appropriate values (e.g., mean, median).

4. Feature Scaling:

- a. Scale the 'Occupation' column by applying Min-Max scaling to it.
- b. Standardize the 'Purchase' column using Standard Scaler.

5. Plotting Data:

- a. Create a histogram to visualize the distribution of the 'Age' column.
- b. Generate a box plot for the 'Purchase' column to identify outliers.
- c. Create a bar plot to compare the average purchase amount for each age group.