Bytes of Intelligence: Set - 01

Department of AI Engineering

Semester Examination

Course Title: Deep Learning Techniques in Python for Vision and Language

Course Code: DL-101

Duration: 3 Hours

Total Marks: 100

Instructions:

- **Attempt all questions.**
- **Read each question carefully before answering.**
- **All code should be properly commented and indented.**
- **Use appropriate data structures and algorithms where necessary.**
- **Marks are indicated next to each question.**
- **Assume any necessary imports (e.g., `import numpy as np`, `import pandas as pd`).**
- **Write your code in Python programming language.**

Question 1: Python Fundamentals and Control Structures (15 Marks)

(a) Conditional Statements and Loops (8 Marks)

Write a Python script that prompts the user to enter a list of integers separated by spaces. The script should then:

- Print all the even numbers in the list.
- Calculate and print the sum of all odd numbers in the list.

Example:

```
Input: 10 15 20 25 30
Even numbers: 10 20 30
Sum of odd numbers: 40
```

(b) Functions and Scope (7 Marks)

Define a function `calculate_grade(scores)` that takes a list of student scores (integers from 0 to 100) and returns a dictionary with the count of grades according to the following criteria:

- **A**: 90 and above
- **B**: 80-89
- **C**: 70-79
- **D**: 60-69
- **F**: Below 60

Demonstrate the use of local and global variables within your function.

Question 2: Data Structures in Python (20 Marks)

(a) List and Tuple Manipulations (10 Marks)

Given the following list of tuples representing products and their prices:

```
products = [("Laptop", 1200), ("Smartphone", 800), ("Tablet", 400), ("Monitor", 300)]
```

Perform the following tasks:

- **Sort** the list based on price in descending order without using the built-in `sort()` method. *(5 Marks)*
- **Convert** the list of tuples into a dictionary where the product names are keys and prices are values. *(5 Marks)*

(b) Set and Dictionary Operations (10 Marks)

You have two dictionaries representing students enrolled in different courses:

```
course_python = {"Alice": 85, "Bob": 78, "Charlie": 92, "Diana": 88}
course_data_science = {"Eve": 91, "Frank": 76, "Alice": 89, "Charlie": 95}
```

Perform the following tasks:

- Find the **students enrolled in both courses** using set operations. *(5 Marks)*
- Create a new dictionary that contains the **average scores** of students enrolled in both courses. *(5 Marks)*

Question 3: NumPy Array Operations (15 Marks)

(a) Array Creation and Manipulation (10 Marks)

- Create a NumPy array `A` of shape `(3, 5)` containing random integers between 10 and 50. *(2 Marks)*
- Compute the **mean**, **median**, and **standard deviation** of the array. *(3 Marks)*
- **Normalize** the array so that all values are between 0 and 1. *(5 Marks)*

(b) Broadcasting and Mathematical Operations (5 Marks)

Given two NumPy arrays:

```
X = np.array([1, 2, 3])
Y = np.array([[10], [20], [30]])
```

- Use broadcasting to perform element-wise multiplication to produce a `(3, 3)` array. *(3 Marks)*
- Explain how broadcasting works in this context. *(2 Marks)*

Question 4: Data Analysis with Pandas (25 Marks)

(a) Data Loading and Basic Operations (10 Marks)

Given a CSV file `employees.csv` with the following columns: `'EmployeeID'`, `'Name'`, `'Department'`, `'Salary'`, `'JoiningDate'`.

- Load the dataset into a Pandas DataFrame. *(2 Marks)*
- Display the first five rows and get basic statistics of numerical columns. *(3 Marks)*
- Filter the DataFrame to show employees who joined after January 1, 2020, and belong to the 'Data Science' department. *(5 Marks)*

(b) Data Aggregation and Grouping (10 Marks)

Using the DataFrame from part (a):

- Calculate the **average salary** for each department. *(5 Marks)*
- Identify the department with the **highest number of employees**. *(5 Marks)*

(c) Handling Missing Data (5 Marks)

Assume the `'Salary'` column has missing values.

- Write code to **identify and count** the missing values in the `'Salary'` column. *(2 Marks)*
- **Impute** the missing salaries with the **median salary** of the respective department. *(3 Marks)*

Question 5: Data Visualization with Matplotlib and Seaborn (15 Marks)

(a) Matplotlib Visualization (7 Marks)

Using the `employees.csv` DataFrame:

- Create a **bar chart** showing the number of employees in each department. Include appropriate labels, title, and color customization.

(b) Seaborn Visualization (8 Marks)

Using Seaborn:

- Plot a **box plot** of employee salaries by department to visualize salary distribution across departments.
- Interpret any noticeable differences in salary distributions.

Question 6: Hands-On Project (15 Marks)

You are given a dataset `sales_data.csv` containing daily sales records with the following columns: `'Date'`, `'ProductID'`, `'QuantitySold'`, `'Revenue'`.

Tasks:

- Load the dataset into a Pandas DataFrame. *(2 Marks)*
- Convert the `'Date'` column to datetime objects and set it as the index. *(3 Marks)*
- Resample the data to get **monthly total revenue** and plot it using Matplotlib. *(5 Marks)*
- Identify the **top 3 products** with the highest total sales revenue and visualize the results using a Seaborn bar plot. *(5 Marks)*

Question 7: Capstone Coding Challenge (Bonus Question - Optional) (10 Marks)

Data Analysis and Visualization

Write a Python script that performs the following:

- Reads a CSV file `'weather_data.csv'` containing `'Date'`, `'Temperature'`, `'Humidity'`, `'WindSpeed'` columns. *(2 Marks)*
- Calculates the **moving average** of the temperature over a 7-day window. *(3 Marks)*
- Plots both the original temperature data and the moving average on the same graph. Include legends and labels. *(5 Marks)*