Bytes of Intelligence: Set - 04

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) Basic Input/Output and Data Types (5 Marks)

Write a Python program that:

- Prompts the user to input a list of numbers separated by commas.
- Converts the input into a list of integers.
- Calculates and prints the average, maximum, and minimum of the list.

Example:

```
Enter numbers separated by commas: 10, 20, 30, 40, 50
Average: 30.0
```

```
Maximum: 50
Minimum: 10
```

(b) Functions and Error Handling (10 Marks)

Define a function `divide_numbers(a, b)` that:

- Takes two arguments `a` and `b`.
- Returns the result of dividing `a` by `b`.
- Handles the `ZeroDivisionError` exception and prints an appropriate message if `b` is zero.

Demonstrate the use of this function with at least two examples, one of which should involve division by zero.

Question 2: Data Structures in Python (20 Marks)

(a) List and Dictionary Comprehensions (10 Marks)

Given a list of words:

```
words = ["data", "science", "python", "numpy", "pandas", "visualization"]
```

Perform the following tasks:

- Create a new list containing the lengths of each word using a list comprehension. *(3 Marks)*
- Create a dictionary where each word is a key and its length is the value using a dictionary comprehension. *(4 Marks)*
- Filter the original list to include only words that have more than 5 letters using a list comprehension. *(3 Marks)*

(b) Tuple and Set Operations (10 Marks)

You have two tuples representing registered users on two different platforms:

```
platform_a_users = ("alice", "bob", "charlie", "diana")
platform_b_users = ("charlie", "diana", "edward", "fiona")
```

Perform the following tasks:

- Convert the tuples to sets and find the users who are registered on both platforms. *(3 Marks)*

- Find the users who are only registered on Platform A. *(3 Marks)*
- Create a combined tuple of all unique users from both platforms. *(4 Marks)*

Question 3: NumPy Array Operations (15 Marks)

(a) Array Creation and Manipulation (10 Marks)

- Create a NumPy array `arr` of shape `(4, 4)` with random floating-point numbers between 0 and 1. *(2 Marks)*
- Multiply every element in the array by 10 and convert the array elements to integers. *(3 Marks)*
- Calculate the row-wise and column-wise sums of the array. *(5 Marks)*

(b) Advanced Indexing and Slicing (5 Marks)

Given the array `arr` from part (a):

- Replace the elements in the second row with zeros. *(2 Marks)*
- Extract the subarray consisting of the last two rows and last two columns. *(3 Marks)*

Question 4: Data Analysis with Pandas (25 Marks)

You have a CSV file `movies.csv` containing the following columns: `'MovieID'`, `'Title'`, `'Genre'`, `'Director'`, `'ReleaseYear'`, `'Rating'`.

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

- Load the dataset into a Pandas DataFrame. *(2 Marks)*
- Display the first 10 rows of the DataFrame. *(1 Mark)*
- Get a summary of the DataFrame, including data types and missing values. *(3 Marks)*
- Identify and handle any missing values in the `'Rating'` column by filling them with the average rating. *(4 Marks)*

(b) Data Filtering and Grouping (10 Marks)

- Filter the DataFrame to include only movies released after the year 2000 with a rating above 8.0. *(3 Marks)*
- Group the DataFrame by `'Genre'` and calculate the average `'Rating'` for each genre. *(4 Marks)*
- Identify the top 3 directors with the highest average movie ratings. *(3 Marks)*

(c) Data Visualization (5 Marks)

- Create a bar chart using Matplotlib or Seaborn to display the number of movies released each year from 2000 to 2023. Include appropriate labels and title.

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

(a) Matplotlib Visualization (7 Marks)

Using the filtered DataFrame from **Question 4(b)**:

- Create a scatter plot showing the relationship between `'ReleaseYear'` and `'Rating'`.
- Customize the plot with different colors for different genres.
- Include a legend, title, and axis labels.

(b) Seaborn Visualization (8 Marks)

- Using Seaborn, create a **facet grid** of histograms showing the distribution of movie ratings for each genre.
- Interpret any differences in rating distributions across genres.

Question 6: Hands-On Project (15 Marks)

You are provided with a dataset `ecommerce_data.csv` containing the following columns: `'OrderID'`, `'CustomerID'`, `'ProductID'`, `'Category'`, `'Quantity'`, `'UnitPrice'`, `'OrderDate'`.

Tasks:

- Load the dataset into a Pandas DataFrame and parse the `'OrderDate'` column as datetime. *(3 Marks)*
- Calculate the total revenue for each order and add it as a new column `'TotalRevenue'`. *(3 Marks)*
- Find the top 5 customers based on total revenue and visualize their contributions using a pie chart. *(5 Marks)*
- Using Seaborn, create a line plot showing monthly total revenue over time. Include appropriate labels and title. *(4 Marks)*

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

Data Processing and Time Series Analysis

Write a Python script that:

- Reads a CSV file `'stock_prices.csv'` containing daily stock prices with columns `'Date'` and `'ClosePrice'`. *(1 Mark)*
- Converts the `'Date'` column to datetime format and sets it as the index. *(1 Mark)*
- Calculates the **7-day moving average** of the closing prices. *(2 Marks)*
- Detects any outliers in the closing prices using the Z-score method and replaces them with the median closing price. *(4 Marks)*
- Plots the original closing prices and the moving average on the same graph using Matplotlib. Include appropriate labels, title, and legend. *(2 Marks)*