

Bytes of Intelligence: Set - 02

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 Data Types and Input/Output (5 Marks)

Write a Python program that:

- Prompts the user to input their full name and age.
- Prints a greeting that includes their name and calculates the year they will turn 100 years old.

Example:

```
Enter your full name: John Doe
Enter your age: 20
Hello, John Doe! You will turn 100 years old in the year 2101.
```

(b) Control Flow and Functions (10 Marks)

Define a function `is_prime(n)` that checks if a given number `n` is a prime number. The function should return `True` if `n` is prime, and `False` otherwise.

- Use this function to print all prime numbers between 2 and 100.
-

Question 2: Data Structures in Python (20 Marks)

(a) List Manipulations (10 Marks)

Given the list:

```
numbers = [12, 75, 150, 180, 145, 525, 50]
```

Write a Python script that:

- Iterates through the list and prints each number.
 - Skips any number greater than 500.
 - Stops the loop if a number greater than 150 is encountered.
 - Skips numbers that are divisible by 5.
-

(b) Dictionary and Set Operations (10 Marks)

You have two sets of student names:

```
english_students = {"Emma", "Olivia", "Ava", "Isabella", "Sophia"}  
math_students = {"Olivia", "Ava", "Mia", "Charlotte", "Amelia"}
```

Perform the following tasks:

- Find the **students who are enrolled in both English and Math** courses. (3 Marks)
 - Find the **students who are enrolled in English but not in Math**. (3 Marks)
 - Create a dictionary where the keys are student names from both sets, and the values are the courses they are enrolled in. (4 Marks)
-

Question 3: NumPy Array Operations (15 Marks)

(a) Array Creation and Indexing (8 Marks)

- Create a NumPy array of shape `(5, 5)` filled with random integers between 1 and 100. (2 Marks)

- Replace all the even numbers in the array with zero. (3 Marks)
 - Calculate the sum of all the numbers in each column and store the results in a one-dimensional array. (3 Marks)
-

(b) Array Reshaping and Mathematical Operations (7 Marks)

Given the array:

```
array = np.arange(1, 13)
```

- Reshape the array into a (4, 3) matrix. (2 Marks)
 - Compute the natural logarithm of each element in the reshaped matrix. (3 Marks)
 - Flatten the matrix back into a one-dimensional array. (2 Marks)
-

Question 4: Data Analysis with Pandas (25 Marks)

You are provided with a CSV file `sales.csv` containing the following columns: `'OrderID'`, `'Product'`, `'Category'`, `'Quantity'`, `'UnitPrice'`, `'OrderDate'`.

(a) Data Loading and Cleaning (10 Marks)

- Load the dataset into a Pandas DataFrame. (2 Marks)
 - Check for and handle any missing values in the `'Quantity'` and `'UnitPrice'` columns. Explain your method. (5 Marks)
 - Convert the `'OrderDate'` column to datetime format and extract the year into a new column `'Year'`. (3 Marks)
-

(b) Data Manipulation and Analysis (10 Marks)

- Calculate the **total sales** (`Quantity * UnitPrice`) for each order and add it as a new column `'TotalSales'`. (3 Marks)
 - Find the **top 3 products** by total sales. (4 Marks)
 - Group the data by `'Category'` and calculate the **average unit price** for each category. (3 Marks)
-

(c) Advanced Filtering (5 Marks)

- Filter the DataFrame to show all orders made in the year 2022 for products in the 'Electronics' category with a quantity greater than 5.
-

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

(a) Matplotlib Visualization (7 Marks)

Using the DataFrame from Question 4:

- Create a **pie chart** showing the proportion of total sales by product category.
 - Include a legend, title, and ensure that the percentages are displayed on the pie chart.
-

(b) Seaborn Visualization (8 Marks)

- Using Seaborn, create a **heatmap** to show the correlation between `'Quantity'`, `'UnitPrice'`, and `'TotalSales'`.
 - Interpret the heatmap and comment on any strong correlations.
-

Question 6: Hands-On Project (15 Marks)

You are given a dataset `population_data.csv` containing the following columns: `'Country'`, `'Year'`, `'Population'`, `'Area'`, `'GDP'`.

Tasks:

- Load the dataset into a Pandas DataFrame. *(2 Marks)*
 - For the year 2020, calculate the **population density** (Population divided by Area) for each country and add it as a new column `'PopulationDensity'`. *(4 Marks)*
 - Identify the **top 5 countries** with the highest population density in 2020. *(3 Marks)*
 - Using Matplotlib, plot a **bar chart** of these top 5 countries showing their population densities. Include appropriate labels and title. *(3 Marks)*
 - Save the plot as an image file named `'population_density.png'`. *(3 Marks)*
-

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

Data Processing and Visualization

Write a Python script that:

- Reads a text file `'article.txt'`. *(1 Mark)*
- Cleans the text by removing punctuation and converting it to lowercase. *(2 Marks)*
- Counts the frequency of each unique word. *(3 Marks)*
- Creates a Pandas DataFrame with words and their corresponding frequencies. *(2 Marks)*

- Plots a **horizontal bar chart** of the top 15 most frequent words using Seaborn. Include appropriate labels and title. *(2 Marks)*