Worksheet 4(b)

Student Name: Rahul Saxena   **UID:** 24MCI10204

Branch: MCA(AI&ML)   **Section/Group:** 3-B

Semester: 1st semester   **Date of Performance:** 10/09/2024

Subject Name: Python Programming   **Subject Code:** 24CAH-606

Aim/Overview of the practical:

Write an experiment import a dataset with numbers and texts keeping the text intact in python NumPy?

Task To be done:

- **Import the Dataset**: Load a CSV file containing both numbers and text using np.genfromtxt() in Python.
- **Verify Data Integrity**: Print the entire dataset to verify that both text and numeric data are loaded properly.
- **Access Specific Elements**: Access and print specific text and numeric elements from the dataset using NumPy's array indexing.

Source Code:

```
import numpy as np
np_data = np.genfromtxt('./Book1.csv', delimiter=',', dtype=object, skip_header=1)
print("Complete Dataset:\n", np_data)
print("\nText Element (Second row, second column):", np_data[0, 0])
print("Numeric Element (Third row, third column):", np_data[2, 2])
```

Output:

```
Complete Dataset:
 [[b'Rahul Saxena' b'3-B' b'204' b'754445451']
 [b'Lakshya Pratap Singh' b'3-B' b'203' b'1234567']
 [b'Atul' b'3-B' b'200' b'1234567']
 [b'Bhavishya' b'3-B' b'153' b'1234567']
 [b'Maan' b'3-B' b'784' b'1234567']]


Text Element (Second row, second column): b'Rahul Saxena'
Numeric Element (Third row, third column): b'200'
```

Learning Outcome:

- **Loading Mixed-Type Data:** Learn how to import a CSV file with both numbers and text using NumPy's genfromtxt() function, and understand the significance of using dtype=object to handle mixed data types.
- **Array Indexing:** Understand how to access specific rows and columns of data within a NumPy array, enabling manipulation of both text and numeric values.
- **Data Conversion:** Gain experience in converting a subset of the data (e.g., numeric columns) from object type to float for mathematical operations.
- **Basic Data Operations:** Perform fundamental operations like calculating the average of a numeric column within a mixed dataset, deepening understanding of how to manipulate real-world datasets in Python.