  

Boot Camp on Artificial Intelligence



**Practical Assignment – 5**

**Group Members:**  **PRN:**

Mihir Patil 122B1B213

Siddhesh Patil 122B1B218

Mihir Pande 122B1B200

Rohan Mengal 122B1B175

Shrawani Pande 122B1B201

Yash Mulay 122B1B178

Mubin Tamboli 22TCOMP2071

Shailesh Galande 22TCOMP2024

Dhanashree Patil 22TCOMP2093

Shivani Pandey 22TCOMP2109

**Date of Submission: Maximum Marks:**

# Lab Assignment: Text Preprocessing with Python Objective:

To clean and preprocess raw text data, preparing it for further analysis or input into machine learning models. Students will use Python libraries such as **NLTK**, **re**, and **Pandas** to complete this task.

# Tasks:

You will be working with a dataset containing product reviews. Your task is to:

1. Load the dataset.
2. Perform the following text preprocessing tasks:
   * Convert all text to lowercase.
   * Remove punctuation, numbers, and special characters.
   * Tokenize the text into individual words.
   * Remove stopwords.
   * Perform stemming and lemmatization.
   * Calculate the **Term Frequency (TF)** of words in each review

# Dataset:

Use the CSV file named product\_reviews.csv