**COMPUTATIONAL LINGUISTICS AND NATURAL LANGUAGE PROCESSING**

**DESCRIPTION OF THE DATASET**

**IMAGE CLASSIFICATION**

**Purpose of the project**  
The purpose of this report is to develop and evaluate a classification model capable of distinguishing between images of tiger and lion. This model leverages machine learning techniques to analyze visual features and accurately categorize input images. By exploring various algorithms and assessing their performance, this study aims to identify the most effective approach for achieving high classification accuracy. The findings of this report can, image recognition systems, and other computer vision tasks.

**Dataset Overview**

This project involves two datasets:

* Lion: Contains 1930 images.
* Tiger: Contains 1600 images.

**Data Collection and Preparation**

* Data was gathered from Google.
* Images were preprocessed by resizing, removing duplicates, and normalizing formats to ensure consistency.
* The dataset was curated to maintain a balanced distribution of tiger and lion images for accurate model training.