01

IMAGE CAPTION GENERATOR

IT WORKSHOP

Group -15

INTRODUCTION

→ In this project, we have presented a deep learning model to describe images and generate captions using computer vision and **machine translation**. This project aims to detect different objects found in an image, recognize the relationships between those objects and generate captions. The dataset used is **Flickr8k** and the programming language used was Python3, and an ML technique called Transfer Learning will be implemented with the help of the **CNN model**, to demonstrate the proposed experiment.

→ Image Caption Generator or Photo Descriptions is one of the Applications of Deep Learning In Which we have to pass the image to the model and the model does some processing and generating captions or descriptions as per its training.

IMAGE CAPTIONING TECHNIQUE

The task of image captioning can be divided into two modules logically :

1.Image based model — Extracts the features of our image.
CNN is a profound learning calculation that takes in the info picture, allocates significance to various components/protests in the picture, and recognizes it from each other.

2.Language based model — which translates the features and objects extracted by our image based model to a natural sentence.
LSTM is used to store the input data, as well as supply predictions about the subsequent datasets through its own.

Image Caption Generator Model

For our image based model we are using - CNN (Convolutional Neural Network) which is a Deep Learning algorithm which takes in an input image, and then helps it differentiate one image from the other.

One of the most popular applications of this architecture is image classification .It is basically used for identifying if an image is a bird, a plane or Superman, etc.

For language based models — we are relying on LSTM.

LSTM- Long Short Term Memory networks are a type of Recurrent Neural Network (RNN) which is well suited for sequence prediction problems. Based on the previous text, we can predict what the next word will be.

dataset for Training the model

In this project, we are using the ***Flickr8k_dataset****.*
The dataset contains two directories:

- *Flickr8k_Dataset: Contains 8091 photographs in JPEG format.*

- *Flickr8k_text: Contains a number of files containing different sources of descriptions for the photographs.*

.

Libraries Required

Following are the necesarry libraries required for our project :

1. *Tensorflow*
2. *Keras*
3. *NumPy*
4. *Pandas*
5. *Matplotlib*
6. *OS*
7. *Pil*
8. *SkLearn*
9. *Nltk*

FLOWCHART

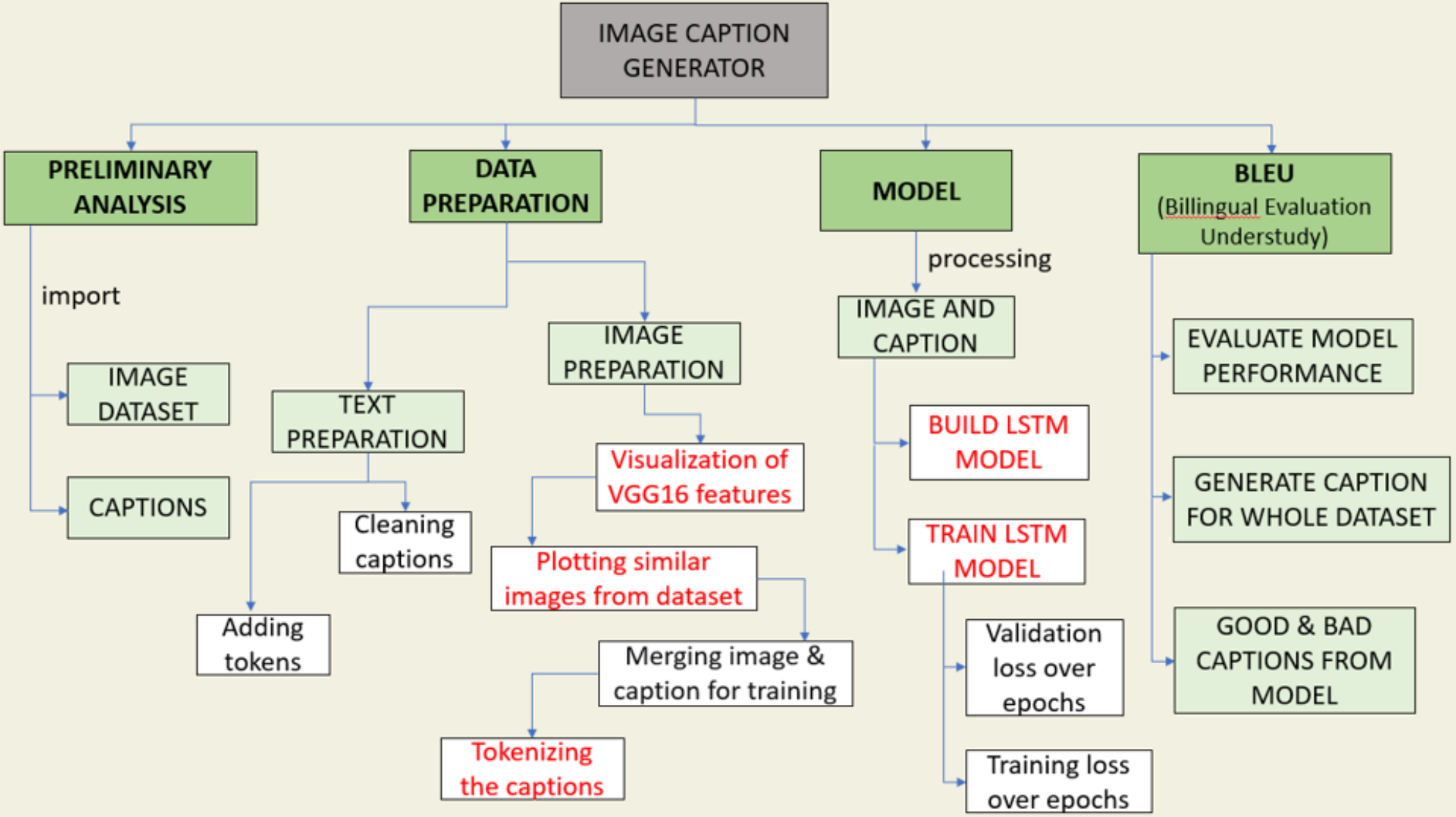
IMAGE CAPTION GENERATOR
PRELIMINARY ANALYSIS
DATA PREPARATION
MODEL
BLEU (Billingual Evaluation Understudy)
import
IMAGE DATASET
CAPTIONS
TEXT PREPARATION
Cleaning captions
Adding tokens
IMAGE PREPARATION
Visualization of VGG16 features
Plotting similar images from dataset
Merging image & caption for training
Tokenizing the captions
processing
IMAGE AND CAPTION
BUILD LSTM MODEL
TRAIN LSTM MODEL
Validation loss over epochs
Training loss over epochs
EVALUATE MODEL PERFORMANCE
GENERATE CAPTION FOR WHOLE DATASET
GOOD & BAD CAPTIONS FROM MODEL

RESULT

imagecaption.ipynb

C: > Users > sohal > OneDrive > Desktop > MLAI > imagecaption.ipynb > import matplotlib.pyplot as plt

+ Code + Markdown ···

··· Bad Caption

true: child in pink dress is climbing up set of stairs in an entry way
pred: boy in red shirt is standing on the air
BLEU: 7.176794039009363e-232

true: black dog and spotted dog are fighting
pred: black and white dog is running in the grass
BLEU: 1.384292958842266e-231

true: little girl covered in paint sits in front of painted rainbow with her hands in bowl
pred: man in red shirt is jumping over the air
BLEU: 4.832402486973385e-232

true: man lays on bench while his dog sits by him
pred: black and white dog is running on the grass
BLEU: 1.1193096620723278e-231

true: man in an orange hat starring at something
pred: man in black shirt is standing in front of the camera
BLEU: 5.477489369001354e-155

imagecaption.ipynb

C: > Users > sohal > OneDrive > Desktop > MLAI > imagecaption.ipynb > import matplotlib.pyplot as plt

+ Code + Markdown ···

Good Caption

true: brown dog is running in the sand

pred: brown dog is running in the water

BLEU: 0.8091067115702212

Group members

Shrujal Bansal

02101172020

Sohali Baisla

02501172020

Rishika Pal

04301172020

Mehak Aggarwal

03401172020