# README

# GOOGLE DESCRIBE FUNCTION

The Machine Learning project discussed here is capable of describing images/videos using deep learning techniques( CNN and LSTM).

I have attached all the required files in the folder on google drive including the model I trained. You can directly run the model I trained and generate descriptions for images/videos. I have mentioned the steps below.

I am also attaching “modeltraining” jupyter notebook in case anyone wishes to optimise the code or retrain the model for any reason.

I have trained the model on flickr 30k dataset available here https://www.kaggle.com/datasets/adityajn105/flickr30k

The captions for image training are already uploaded in the folder as a txt file under the name “flickrcaption”.

To retrain the model you would have to download the image dataset from the website.

Steps to run the pretrained model:

- Save this folder on your local machine.
- For videos, open the “videodescribe” jupyter notebook and enter the video address in the filename and run the notebook. It will describe the given video frame by frame.
- For images, follow the same steps in “imagedescribe” jupyter notebook

Or

- Run the “WebPortal for imagedescribe” jupyter notebook after which a local link would be provided in the notebook itself where you can experience this functionality on a web portal.

**Note:**

The machine learning model employed in this project has demonstrated a considerable level of accuracy, considering the dataset's size (25k images) during its training phase. While the model performs well within the confines of its training data, it's essential to note that its performance may vary when exposed to images or contexts outside its training scope. The model's accuracy and efficacy could be influenced by the dataset's scale and diversity Continuous refinement and expansion of the dataset may further enhance the model's accuracy and applicability across broader scenarios.

## Some Testing Results:

(descriptions generated by model followed by image)

```
In [61]: filename = r"C:\Users\KRIDAY PARMAR\Flickr\flickr30k_images\241347067.jpg"
         print(get_image_caption(filename,modelx))
         display(Image(filename))

1/1 [==============================] - 1s 512ms/step
 Football Player In Red And White Uniform Is Trying To Catch The Ball
```

```
In [62]: filename = r"C:\Users\KRIDAY PARMAR\Flickr\flickr30k_images\4749181324.jpg"
         print(get_image_caption(filename,modelx))
         display(Image(filename))

1/1 [==============================] - 1s 516ms/step
 Man In Blue Shirt Is Standing In Front Of Building
```

```
1/1 [==============================] - 1s 551ms/step
 Dog Is Running Through The Grass
```

```
1/1 [==============================] - 1s 523ms/step
 Little Girl Is Playing With Toy
```

```
1/1 [==============================] - 1s 922ms/step
 Man In Black Shirt And Black Pants Is Playing The Guitar
```

```
1/1 [==============================] - 1s 550ms/step
 Young Boy Is Playing In The Water
```

Thank you,

KRIDAY PARMAR