

## HELP DOCUMENT:

The code can be run on google collab after installing a few packages. However, the code accesses the dataset directly from the Kaggle repository and requires a unique API token which can be downloaded from the website with any valid user account. The API token is a .json file which needs to be uploaded to google colab before performing any operations on the dataset.

The dataset is in a zip file, using the ! unzip command the files are extracted. The rest of the code can be run using the run command.

Any packages that throws up the “ModuleNotFoundError” can be installed through !pip install <package name> and all the names of the functions are in the latest python version (3) which needs to be followed throughout.

## PACKAGES/LIBRARIES USED:

1. Tensorflow - Used for building and training of the Convolutional Neural Network (CNN) model using the dataset.
2. Keras - Acts as an interface for the Tensorflow library.
3. Scikit Learn - Provides a consistent interface for classification
4. Numpy - For storing the sizes of the images in the form of arrays (working with large arrays)
5. Pandas - Provides support for multidimensional arrays and in loading the .csv files.
6. Matplotlib - Used for plotting the accuracy and validation accuracy plot for both test and train datasets.

## **Supervisor:**

Dr. Siddhaling Urolagin,

PhD, Post Doc, (Machine Learning),

[Dr.siddhaling@gmail.com](mailto:Dr.siddhaling@gmail.com)

[www.github.com/siddhaling](https://www.github.com/siddhaling)

[www.researchreader.com](http://www.researchreader.com)