description

October 31, 2018

1 Recepit Classification

1.1 Three Parts of discussion for this task:

```
1) Aim of the task
2) Dataset Collection
3) Approach exploration and finalization
```

1.1.1 1) Aim of the task:

```
As the name suggest Aim of this task is to classify a given image into a recepit or not.
Hypothesis: image satisfing the visual standard of recepit should be classified as recepit.
```

1.1.2 2) Dataset Collection:

```
Now as per the hypothesis the dataset is collected using bulk image downloader.
For this task two labels are genrated as
1: recepit
2: non-recepit

Description of download criteria:
1) Recepit: In this set of the downloaded images, search key words
that are used "grocery bill, invoice, atm-slips"
2) Non-recepit: In  this set of downloaded images, search key words
thhat are used "documents , printed office paper"

    Once the images were downloaded then the rigorious task of data cleaning started.
In this process almost all the data was visually checked and decided to make keep then
in training set or not. And since it was a itreative process, it took its own time before
satisfactory training set was created.
    The final training set had blanced class data with 600 images in each set.
```

1.1.3 3) Approach Exploration and Finalization:

```
    The exploration started with the simplest approach of using standard Computer Vision
techniques which later on was taken over by deep learnning ideas which also had its own
itreation before finalizing the used design.
```

1)Computer Vision Based technique: In this method the idea was to cash in on the most basic and dominant feature of the recepit, that is its glossy paper with to much of a grid based view. For this purpose proposed logic was to preprocess the image for noise removel and then calculating the GLCM based texture information beacuse of glossyness and also calculating the hough line based grid formation on the recepit. But the idea started failing in its preprocessing stage only because the images didn't had proper illumination. But still few images did pass through preprocessing but getting hough lines and preparing a threshold beacuse of varing size of slips was getting difficult.

2) Neural Network based approach: Since the first idea failed beacuse of above stated reasons, it was required that solution should come from more robust method and hence the neural networks.

a) Custom Architecture: As per the initial finiding varing size was a big problem and since the neural network is robust for illumination condition. So the Architecture was designed that can input the images of different size, it was inspired by the fully convolution network, with dense layer genrated using 1x1 convolution kernal size. But like any neural network which is trained from the scratch has a most important need i.e. amount of data that is required to train the model. And the collected data was total 1200 images, which is nowhere near to the fulfilling the need. But to overcome this augumentation was applied but still it was taking unpredictable time to converge. And as a result this architecture was rejected.

b) Transfer-learing based approach: As training from the scratch was getting difficult a pretrainned VGG16 model is used for tranfer-learning based approach. The reason behind choosing the VGG16 was its depth. As most of the models which are trainned on imagenet dataset are much deeper than VGG16 and deeper model will need more data and lack of data is already a constrain. So the pretrained VGG16 with 1024 fully connected layer followed by a softmax layer is added at the bottom of network to create a to-be used network.

In [ ]: