

My thought-processes during many stages of this project are documented as follows:

1. Exploratory Data Analysis

- After a quick look of the dataset, I found some inconsistencies such as dark edges, varying resolutions in the image files
- In order to mitigate these issues, I took the following steps in the preprocessing stage.

2. Preprocessing

- The images were resized to a standard resolution of 256*256 which is suitable for the model that is being used.
- Even though the processed image was blurred, the improvement in the contrast ratio was the desired goal. Because the classification of the images is dependent on how well the machine learns the structure and format of the document rather than focusing on the words.
- This strategy could be hindered due to the presence of dark edges around the image. So, their removal was essential.

3. Model Architecture

- VGG16 model architecture was chosen as it was effective, simple and easily fine-tunable as compared to other models.
- Since VGG16 model is also good in transfer learning, we can partition an image into 4 parts and train the model that way and stack all the 5 models to obtain an even accurate and precise model.
- Since the custom vgg16 model on whole image produced desirable outcomes after some fine-tuning, I decided not to proceed further with the transfer learning concept.