Report on Different Approaches and Results

Roll Number: 21f2000120

1. Generic Decisions:

- The ***train data*** has two gender classes with 75% and 25% proportion for the classes, hence they are ***imbalanced***, therefore ***train test split with stratify*** has been used.
- The ***sizes*** of the ***input images*** are in the range 25 X 25 to 501 X 501, the ***mean*** value is around ***255*** and since the ***ResNet*** models were trained on input size of ***224*** and our train dataset images mean value of size is closer to it, input size of ***224 X 224*** has been ***chosen***.
- The ***input pixels*** are ***normalized*** using mean and standard deviation (RGB) of the train set as it ***decreases the dependency on weight initialization***. Data Augmentation is performed to improve the generalization. The Age value is also normalized with mean and standard deviation.
- A very small ***validation set (0.05%)*** is defined so as to ***validate after each epoch***. A batch size of ***64*** has been chosen to implement the idea of ***mini batch*** gradient descent to utilize the ***GPU*** for training and reduce the computation time.

2. Pre-trained Model :

- ***Efficient net b3*** has been chosen (12.2 million parameters) in order to leverage ***computational cost*** and not compromise on ***accuracy***.
- The ***head of the pre-trained*** model is made as ***Identity*** and ***2 heads are added*** which combines a sequential network of ***linear*** layers, ***ReLU*** and ***Dropout***.
- A ***learning rate of 5e-5*** is used as the recommended is 2e-5 to 5e-5 for fine tuning.
- ***5 Epochs*** has been chosen since it's a pre-trained model which has learnt features.
- There has been ***41% percent reduction in training loss*** from epoch 1 to 5 and a ***5% increase in validation score*** (from 76% to 81%). The notebook gives a score of 81% in the test set.

3. Custom-CNN :

- Custom CNN with ***5 convolution layers*** with channels, kernel size as 32,7 -> 64,7 -> 128,5 -> 256,3 -> 512,3 respectively with batch normalization, relu activation, max pooling layers for the first 4 layers and a global average pooling layer in the 5^{th} layer. 2 Heads were used similar to the pre trained model implementation. The model has ***1.92 million parameters***.
- A ***learning rate of 2e-5*** had been chosen as higher learning rates did not lead to stable increase in validation score.
- ***15 Epochs*** had been chosen considering it is a custom model.
- There has been ***35% reduction in training loss*** from epoch 1 to 15 and a ***15% increase in validation score*** (from 56% to 71%). The notebook gives a score of 73% in the test set.

4. Learnings and Future Work :

- Fine tuning a pre-trained model performs well even with very less epochs while the computational cost is higher. Custom model improves slowly as the number of epochs increase and the computational cost also increases with the increase in number of epochs.
- With respect to custom CNN, model architectures (with higher kernel sizes or increase in number of layers keeping a constant kernel size), different learning rates (with decay) and average pooling layers in some layers can be tried to improve the model performance.