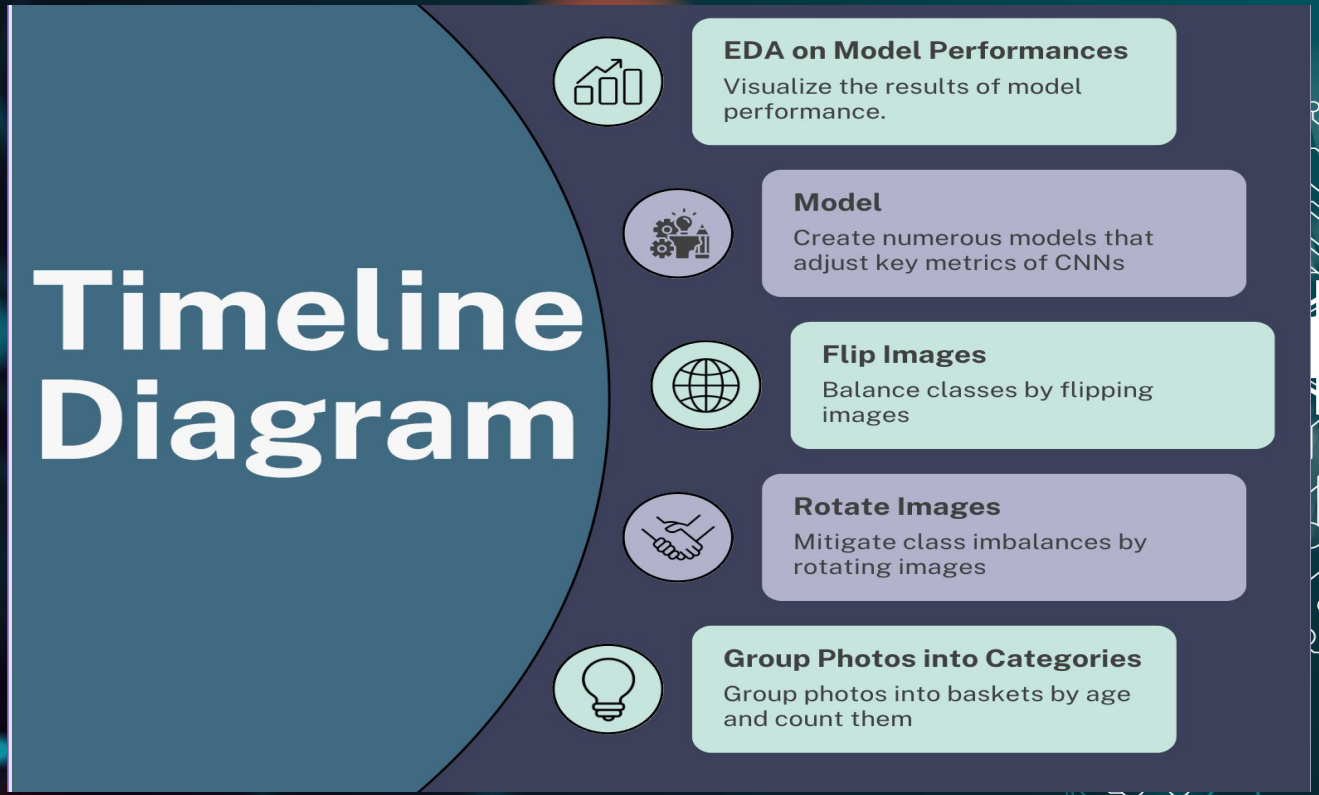
Facial Age Classifier

Nicholas Sanso

Objective

Create a model that can correctly identify the age of an individual in a picture

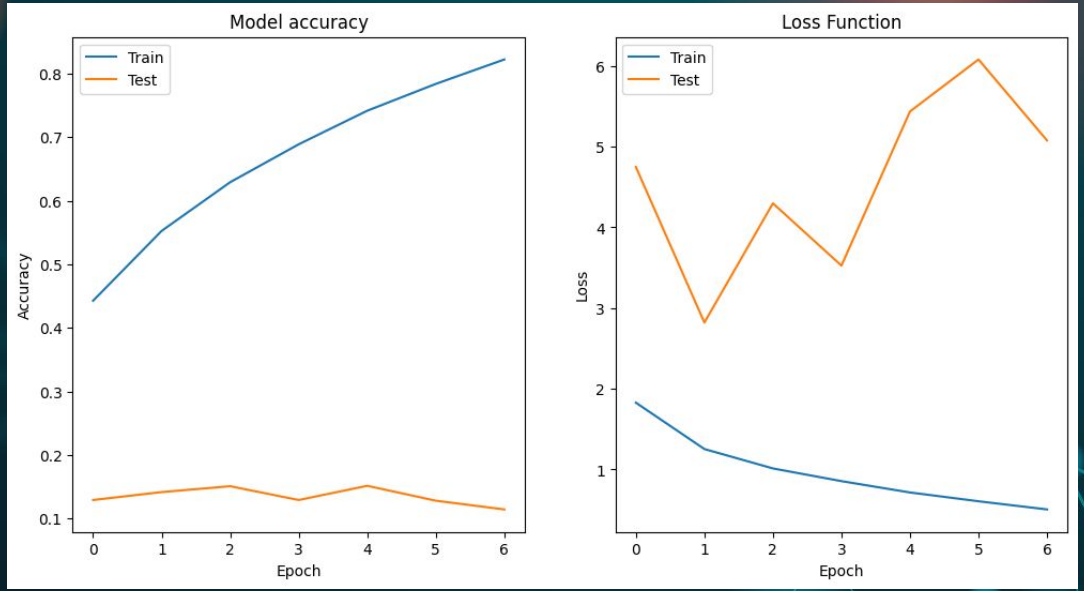
Workflow


Timeline Diagram
EDA on Model Performances
Visualize the results of model performance.
Model
Create numerous models that adjust key metrics of CNNs
Flip Images
Balance classes by flipping images
Rotate Images
Mitigate class imbalances by rotating images
Group Photos into Categories
Group photos into baskets by age and count them

Base Model Results

- The Google Efficient Net model performed poorly

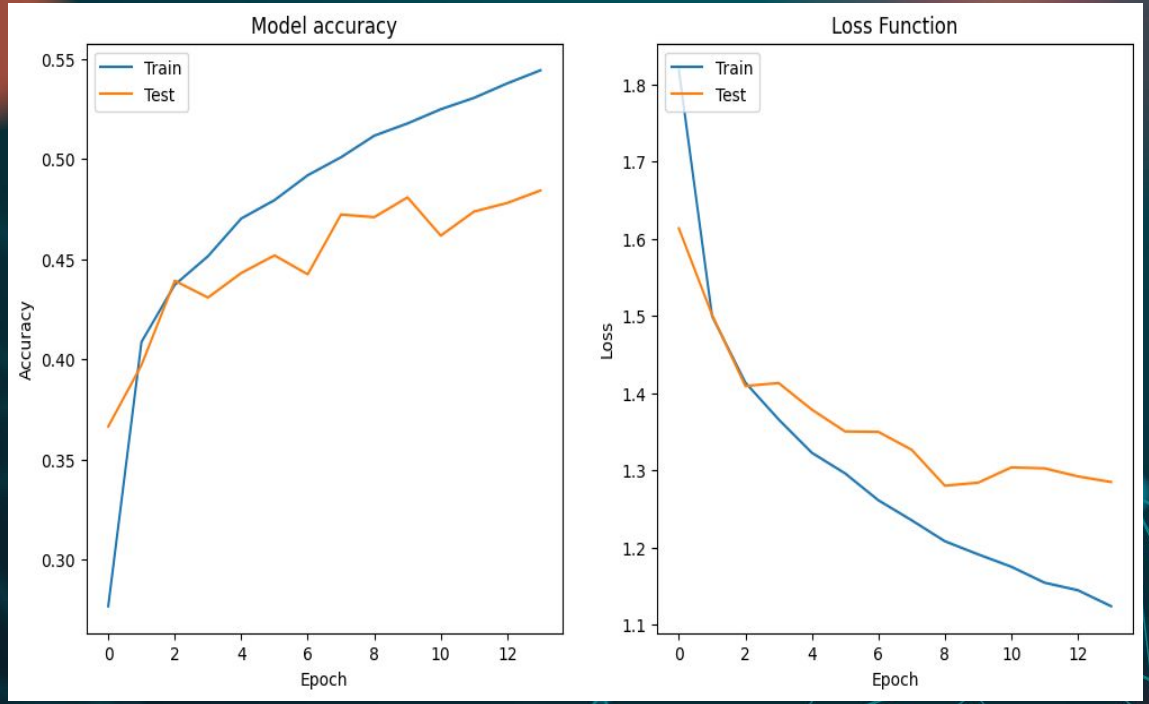
It had low accuracy, high cross entropy, and quickly became overfit.



Model Results

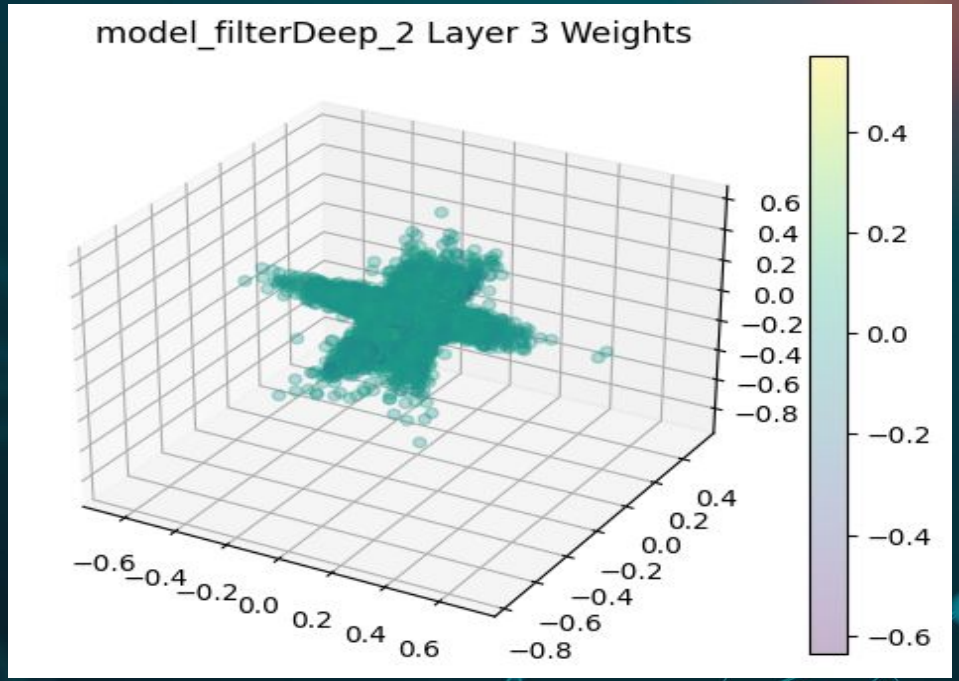
- One of the better performing models

All the models at higher number of epochs, the model became overfit



Modeling EDA

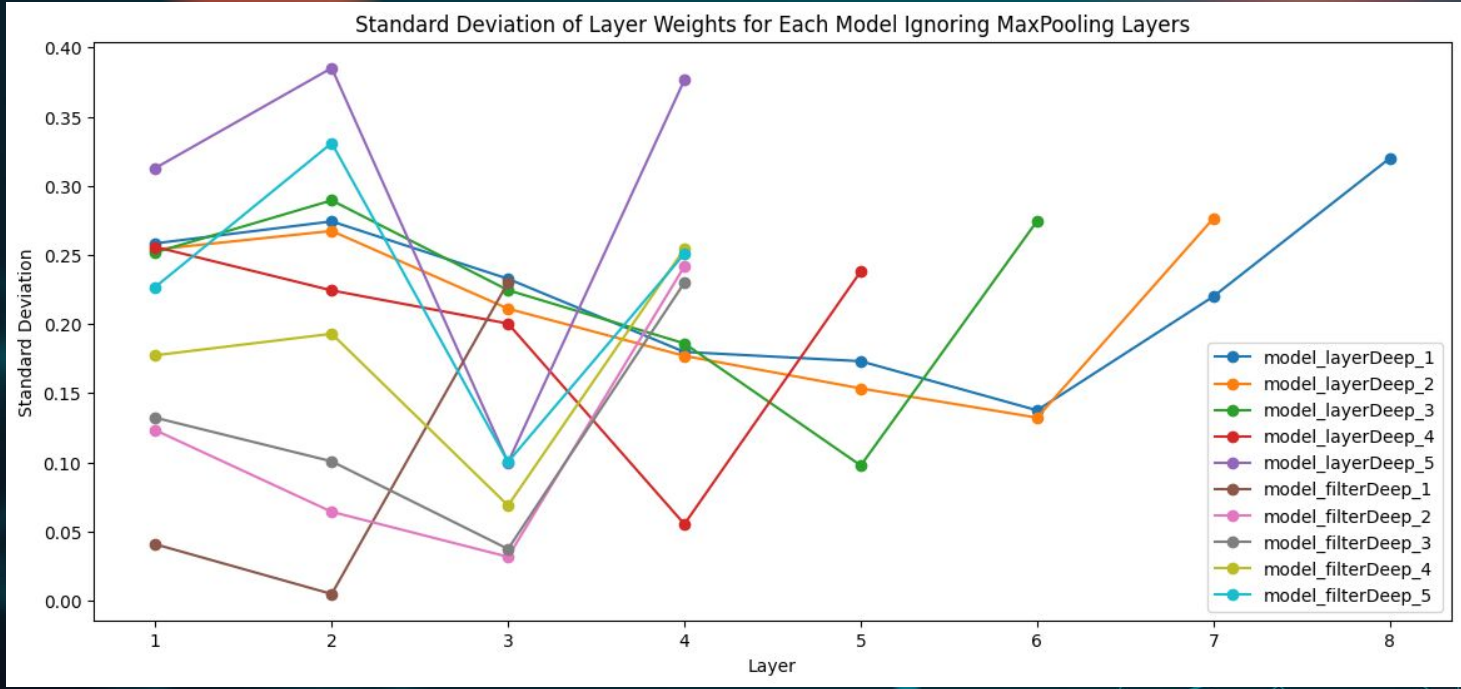
- If certain features often appear together in the input data, the weights of the CNN in these models could cluster together forming a starfish pattern



Modeling EDA

The final layer (which is always a dense layer) saw a sharp increase in the standard deviation of the layer's weights.

The second to last layer, which is also a dense layer, always saw a precipitous decline in the standard deviation of the layer's weights.



Primary Conclusions

1. There are probably a limited number of patterns that the model relies heavily on
2. The improvement in performance with the reduction in batch size indicates that the model may have been converging to a local minimum when minimizing the loss function during the (training) gradient descent process. Reducing batch size likely led to a more stochastic gradient descent, helping the model bypass local minimums.
3. Considerations for model performance must be balanced against hardware requirements. MaxPooling2D's contribution to reducing training time is indispensable for all CNN models.
4. A model that outperformed the base model was established.

Next Steps

1. Implement transfer learning by seeing if the age classifier could classify race, gender or other attributes commonly targeted by advertisers or pollsters.
2. Balance the number of modified images added to each class. Increase the number of modified images in order to still balance the classes.
3. Pull the weights of each of the neural networks to explore the standard deviation of the weights which will indicate the variety of features that the model is using to learn.
4. Run more models to isolate other parameters in the model