

# Dog Breed Image Classification

*Ishaaq Razack, Nathan Williams*

## Background

Our model uses the [dog breeds dataset](#) hosted on Kaggle. The data is broken up into train and test sets. Labels were provided for the train set; however, none were provided for the test set. We believe this was done intentionally since the data is part of a competition, so evaluation of groups' final model would be done separately.

Due to limitations of the test data, we split the training set into train, validation and test. We believe the size of the dataset is satisfactory so each group has sufficient data to perform its respective function. The splits are thus 80/20 for train and test, and we split train further into train and validation using another 80/20 split.

## Model

We used the MobileNet V2 as our pretrained model. The base model was left in its current state with the addition of another deep layer to handle the classification of our dog breeds. To perform hyperparameter tuning on our model, we ran it on our train, validation and test sets repeatedly while changing hyperparameters. We recorded the model accuracy for each instance and generated figures showing the change in accuracy and loss for each epoch. Our findings are shown below:

	Parameters				Train/Test Accuracy		
	Optimizer	Dropout Rate	Learning Rate	Epoch	Train	Validation	Test
1	Adam	0.2	0.01	10	0.9338	0.655	0.6659
2	Adam	0.2	0.1	10	0.9069	0.655	0.6639
3	Adam	0.1	0.01	10	0.9954	0.6966	0.7065
4	Adam	0.1	0.1	10	0.9655	0.6538	0.6649
5	RMSprop	0.1	0.01	10	0.9943	0.7028	0.7079
6	RMSprop	0.1	0.1	10	0.9919	0.6948	0.6952
7	RMSprop	0.2	0.01	10	0.9956	0.7003	0.7128
8	RMSprop	0.2	0.1	10	0.9846	0.6869	0.7045
9	RMSprop	0.2	0.01	12	0.9962	0.6924	0.7133
10	RMSprop	0.1	0.01	14	0.9991	0.6948	0.7143

## Fine-Tuning

Taking our best model from our hyperparameter tuning step, we unfreeze some of the top layers from our base model to do some fine-tuning on our base model. Our results after doing so slightly lowers our test accuracy to 0.6580. Therefore, we decided not to make any changes to our base model.

## Predictions

We have also provided the predictions of our best model on 30 observations from the test dataset:

Predictions VS Actual Labels

