Dog Breed Classifier

Machine Learning Nanodegree Project Report

Jiawei Sun

**1. Definition**

**1.1 Project Overview**

**1.2 Problem Statement**

**1.3 Metrics**

The metric that we use to evaluate the performance of dog classifying model is “accuracy”. Accuracy is defined as the number of items correctly classified divided by the number of all classified items. In the equation, TP, FN, FP and TN represent the number of true positives, false negatives, false positives and true negatives, respectively.

We use Cross Entropy Loss as the loss function to optimize the machine learning algorithm. The loss is calculated on training and validation datasets, and it measures the model performance in these two data sets. Cross-entropy is a better measure than MSE for this project, because the decision boundary in a classification task is large, while MSE doesn’t punish misclassifications enough.

**2. Analysis**

2.1 Dog Images

Udacity provides 8351 dog images. Among them, 6680 images (80%) are used for training, 835 images (10%) for validation and 836 images (10%) for testing.

2.2 Human Images

There are 13,233 human images in total.

**3. Methodology**

**4. Results**

**5. Conclusions**

<https://github.com/chauhan-nitin/Udacity-IdentifyCustomerSegments-Arvato/blob/master/Identify_Customer_Segments.ipynb>

<https://github.com/MrDaubinet/Dog-Breed-Classifier/blob/master/dog_app.ipynb>

<https://github.com/littlejsun/CNN-dog-breed-classifier/blob/main/dog_app.ipynb>