

Machine Learning Engineer Nanodegree

“猫狗大战”开题报告

叶硕杨

06月11号，2018

开题报告

Domain Background

通过“猫狗大战”这个毕业项目，对机器学习在图形识别、图形分类上有更加系统的训练和学习。目前，机器学习在无人汽车、美颜相机、图片应用和短视频应用等等领域有越来越多的应用，帮助提升产品的“智能”水平，提高服务水平、预测帮助等，比如一些图形识别的研究在进行，比如“[AI App Identifies Plants and Animals In Seconds](#)”等等一些报道。无论是现在很火热的无人汽车，还是现在无处不在的监控系统，或者机器人领域，等等需要“视觉”的系统，对于图形的识别都具有研究的前景。

Problem Statement

在给定的已经做了“标签”的猫和狗的图片库中进行训练，通过深度学习训练识别模型，从而达到给定其他猫狗图片正确判定动物，使得最后的每张图片的判断达到90%以上。

Datasets and Inputs

数据集是kaggle的 Dogs vs. Cats Redux: Kernels Edition (<https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition/data>) 网站上下载的数据集，分为两个文件夹，训练数据集train和测试数据集test。

train 里面有25000个已经标记cat/dog的图片，各一半；test 里面有12500个未标记的图片。

Solution Statement

识别的图片只要求进行cat和dog的判断，是二元分类问题，可以使用决策树、支持向量机SVM等等算法进行。根据课程的学习，决定采用基于tensorflow的深度学习对train里面的图片进行学习，然后使用学习的模型进行测试。

Benchmark Model

使用卷积神经网络搭建神经网络模型，获得特征向量。

Evaluation Metrics

kaggle网页上说明，本问题采用的 log loss 进行训练结果的评估：

$$LogLoss = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

其中：

- n 是test测试集的图片数量
- \hat{y}_i 是可能是dog狗的预测可能性
- y_i 如果图片是狗则为1，如果是猫则为0
- $\log ()$ 是自然底数对数函数

Project Design

先对train数据进行分类处理，然后对数据进行标准化处理。

依次创建卷积层、最大池化层、扁平层和全连层，最后输出学习的卷积模型。

将模型应用于测试模型。

