开题报告

Capstone Proposal

Hua Lee

2018 年 5 月 1 日

Proposal

Kaggle Competition

项目背景

深蓝在 1997 年在国际象棋比赛中击败卡斯帕罗夫。沃森在 2011 年击败了 Jeopardy 最聪明的琐事。

Web 服务通常受到人们解决这个难题的挑战，但这对计算机来说很困难。这样的挑战通常被称为 CAPTCHA（完全自动公开的图灵测试来告诉计算机和人类）或 HIP（人类交互证明）。HIP 用于多种用途，例如减少电子邮件和博客垃圾邮件，防止对网站密码进行暴力攻击。Asirra（一个 CAPTCHA，要求用户从一组 12 张照片中识别出猫和狗）是一项 HIP，通过询问用户识别猫和狗的照片而工作。这项任务对于计算机来说很难，但研究表明人们可以快速准确地完成任务。

问题描述

这个项目的目标是编写一个算法来分类图像是否包含狗或猫，是一个分类问题，我将使用深度学习对其图像进行分类。

对原始图像中的猫和狗的训练集运用深度学习算法进行训练，将数据分为训练集和验证集，然后将训练出来的模型运用于测试集进行猫狗分类，将分类的结果上传 kaggle，看分数排名。

数据或输入

数据集的链接为：https://www.kaggle.com/c/dogs-vs-cats/data

打开链接，可以直接点下载，这个过程中需要验证手机，会发一个验证码，验证完成后，统一竞赛规则，则就可以下载数据集。

数据包括 3 部分，一个 sampleSubmission.csv 的提交的样本格式，一个 test1.zip 的测试文件，一个 train.zip 的训练文件。将测试文件和训练文件解压，训练集有 25000 个狗和猫的图像，在此文件上进行训练模型，然后用模型在 test1.zip 测试集中预测标签（1 =狗，0 =猫）。

训练集中 train.zip 中有 25000 张图片，猫狗各占一半，猫狗的图片下标是从 0 到 12499，但是图片没有按照下标顺序存放。测试集 12500 张图片，没有猫狗标签。

使用 keras 需要对训练数据进行分类到不同的子目录，**ImageDataGenerator** 对数据进行增强（倾斜，旋转，缩放等产生更多不同的图片），按照 8:2 分配，80%的训练集，20%的验证集。

解决方法描述

解决猫狗分类问题，首先将将训练数据输入卷积神经网络层，运用 train.zip 集进行训练与验证，最后得到模型，将模型应用于 test1.zip 集中。对于训练模型，我会使用 keras 搭建 CNN，后端使用 tensorflow。

评估标准

性能是根据正确标记的图像的百分比进行评估的。为了确定你打破 Asirra CAPTCHA 的几率，请将你的比例提高到 12 个点。

最终评估提交到 kaggle，看分数，看排名。

基准模型

基准模型是 logloss。

$$logloss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\log(p_{ij}),$$

其中 N 是测试集中图像的数量，M 是图像类标签的数量。

项目设计

- 将 train.zip 数据分为 20%的验证集和 80%的训练集，利用 **ImageDataGenerator** 对训练集数据增强，验证集不用增强。
- 使用预训练的 **Xception**
- 模型构建使用 **dropout**
- 模型训练可以直接 **model.fit** 进行，参数为批大小，循环的轮数，验证集分割比例
- 模型应用于测试集，根据测试集准确率来评估模型好坏
- 利用 Digraph 将模型进行可视化

参考

[1] kaggle 官网：https://www.kaggle.com/c/dogs-vs-cats

[2] keras 中文文档：http://keras-cn.readthedocs.io/en/latest/

[3] tensorflow 官网： http://www.tensorflow.com/