Capstone 2: Yelp Dataset Image Classification
Project Proposal

Nils Madsen

Goal and Utility
When photos are uploaded to Yelp, not all information will necessarily be filled in by the user. Certain fields of information for the photo, such as the label, will likely be left empty for many pictures. Yet, having missing data in the database can make it difficult to run analyses, so there is a motivation to fill in data holes wherever possible. Having labels for all images also helps visitors to the website find information they want about a business.

The sheer number of pictures uploaded to Yelp on a daily basis would make it uneconomical to hire people to classify photos manually. However, modern machine learning algorithms, especially convolutional neural networks (CNN), are powerful, accurate, high-throughput tools for image classification, and therefore are well suited to solving this problem. The goal of this project will be to create a CNN model that can accurately classify the photographs in the Yelp Dataset into broad categories.

Data
Yelp has published the Yelp Dataset for academic uses. The larger dataset includes a variety of fields for each rated business, including location, name, type, and star rating, as well as full-text reviews. For each user, there is information on join date, number of reviews written, and number of compliments received. Each photo entry has a corresponding label, caption, and business it is tied to.

This project will focus entirely on the 'label' field for each photograph. This is a broad classification, with only five possible entries ('food', 'drink', 'inside', 'outside', and 'menu'). There are a total of 280,992 photographs.

Round 12 of the Yelp Dataset was downloaded from the Yelp website:
https://www.yelp.com/dataset/challenge

Approach
CNN architectures are very time consuming to design, and I have limited computing hardware. For this reason, the approach taken for this project will be to fine-tune a premade CNN model for transfer learning rather than attempt to design and train an architecture from scratch. The VGG neural network architecture designed by the Visual Geometry Group at Oxford scored second in the 2014 ImageNet classification challenge, and is available as a template in the Keras API for Tensorflow, with or without pretrained weights. This makes it a convenient model to use for this task.

The VGG16 model will be tuned and trained for this task, and may also be compared to other models, such as ResNet, if my hardware is capable of running these larger and more complex models.

Deliverables

The deliverables for this project will include code, a final report, and a slide deck.