Unsupervised Image Sentiment Analysis

# Introduction

Social media has proven to be a rich field for making sense of popular opinion on a number of different topics. From companies trying to understand how their brand is being received, to political parties striving to get a hold on popular opinion, twitter mining has become an invaluable tool.

Much of this work has focused on text-based sentiment analysis. As a result, \ text-based sentiment analysis in social media is a fairly well-evolved area of machine learning.

But focusing exclusively on textual sentiment analysis misses the fact that social media is increasingly image-based. With image-centric platforms like Instagram and Snapchat becoming increasingly important, it is foolish to assume that text analysis alone can provide an accurate indicator of social media sentiment on any particular topic.

Unfortunately, much of the work in image analysis has focused on object identification, rather than sentiment determination. And much of the work that has been done on image sentiment has taken a hand-crafted, rather than machine-learning approach to sentiment ascription. Additionally, image sentiment analysis has been hampered by the fact that it is difficult to assemble sufficiently robust data image data sets that are labeled with reliable sentiments.

In this project, I will attempt to determine whether twitter text-based sentiment analysis can provide a ground-truth to effectively develop an image classifier. To do this, I propose to first classify Twitter-posted images based on textual clues. Using these ratings, I will then develop a Neural-Network based model to classify image sentiment.

The predictive ability of this model will be tested against images that have been had sentiment scores ascribed by crowd-sourcing. The model will be compared against benchmarks, and against a model derived solely from images whose sentiment has been crowd-sourced.

# Literature Review

There are a couple of different categories of literature that need to be reviewed to ensure this project is properly grounded in best practices. These are set out systematically below.

## Image Sentiment Analysis

Foo

## Text Sentiment Analysis

Foo

# Dataset

Give the description of the dataset that you are using along with the individual attributes you will or will not use in your analysis. Also mention the source of the dataset (where did you get it from). In case the data is curated and created by you please explain the details. Descriptive statistics of the attributes and datasets can also be provided here.

The data used for the bulk of this analysis is derived from captured Twitter stream data.

# Approach

Create a block diagram for the steps of your approach to clearly provide an overview. For example, if you first scrapped twitter, second applied NLP techniques to extract keywords, third labelled the tweets as positive and negative using a set of keywords, and fourth build a classifier, then you should create a box for each of the steps with arrows connecting one step to the next one. A sample block diagram is shown below.

Once this is done, explain each of the steps in detail. What are you planning to do in each step or have already done. For example, in the above case you would create subheadings for each of the steps.

## Step 1: <Name of the step>

Write details of the step 1. If there is any source code that you’d like to share then provide the link of the Github.

## Step 2: <Name of the step>

Write details of the step 2. If there is any source code that you’d like to share then provide the link of the Github.

………………………………………………………………………………………………………………………..

………………………………………………………………………………………………………………………..

## Step N: <Name of the step>

Write details of the step N. If there is any source code that you’d like to share then provide the link of the Github.

## Bibliography