Christopher Graham

Project Abstract – CKME136

I propose to develop a machine learning model to predict image sentiment – one of the more interesting and developing areas in machine learning.

I will attempt to classify sentiment on social media-generated images using user-supplied texts – either keywords from Flickr image databases or from the text accompanying photos in Twitter posts.

Step 1 will be to implement a text-based sentiment analysis algorithm likely using either logistic regression or a tree-based model.

Step 2 will be to use the predictions from step 1 to train an image classifier using a neural network model.

Data sources (not exhaustive)

* Twitter stream data
* Flickr Image data from Columbia University Digital Video & Multimedia Lab (<http://www.ee.columbia.edu/ln/dvmm/vso/download/flickr_dataset.html>)
* Yahoo Labs Flickr Creative Commons 100M (http://webscope.sandbox.yahoo.com/catalog.php?datatype=i)