Michael Siebel
Project Proposal
Group 5

The goal of my final will be to create a model that can serve as copy right detection for the HBO television series Game of Thrones. A common issue television networks face is improper distribution of their intellectual property. I intend to create a model that can correctly classify whether an image belongs to Game of Thrones or another television show. This interested me as Game of Thrones is known for having numerous characters and settings, making it more difficult to classify. The series is fantasy-based, existing in a fictional, Medieval inspired land. For a person, without knowing the characters and settings, it would be easy to classify as it is a costume-drama largely set outside or in rustic castles and ancient huts. The series further has high production values and is color treated—meaning cold locations have blue colors emphasized and warm locations have orange colors emphasized via post-production. Overall, the series appears ancient, mystical, and serious in tone. Therefore, my model likely will need to classify the attributes of the fantasy setting more than the actors' faces.

Given the difficulty of these images, the rest of my data will be made of comedic sitcoms (e.g., Seinfeld, Friends, and How I Met Your Mother). Sitcoms have closer to real world aesthetics, with characters wearing modern clothing and being in modern settings. They also have lower production values and are not color treated. Thus, the classification problem is binary, in which series tone, production values, and overall aesthetics are the main attributes to classify.

My dataset will be screenshots taken from these series. I will web scrap these images from the Internet Movie Database (IMDB.com). Images that are not directly from an episode will not be taken (e.g., behind the scenes or promotional photos). All images are pre-cropped to be 100x100 pixels. There are 3,883 Game of Thrones screenshots. I will take another ~10,000 images from at least seven different sitcom series. Therefore, the dataset will be large enough to train a deep network but will be a little imbalanced in binary classes.

I will use a convolutional neural network as it accepts matrix input without flattening, which will enable me to retain the images' RGB values. This is important as the color scheme will likely be important for training. I will start with a standard form of the network and focus on data augmentation before moving to more complex architecture in the event of classification difficulties. Currently, I am more comfortable with Keras and plan to use it to implement the network. However, as I work on Exam 2 and become more comfortable with Pytorch, I will consider moving to it.

Unfortunately, it is difficult to find useful resources to aid my project. I will mostly focus on Keras or Pytorch documentation, lecture slides and notes, and data science blogs (particularly Towards Data Science). I will use F1 scores to evaluate the network is it provides a harmonic average between precision and recall. Precision is often a more valuable metric for judging the performance of binary data in which the predicted class is under sampled. However, precision does not account for false negatives, which can still be an issue as my data is only a little imbalanced. Therefore, F1 scores provide balance in evaluation.

Below is a rough schedule outlining my completion of the project.

- Nov 22: Finish web scrapping
- Dec 4: Finish modeling
- Dec 9: Finish creating the presentation and report