Adam White

CPSC 325 Final Project Write-Up

Going into this semester, as I’ve previously mentioned in my presentations, I had a lot of ambitious goals for what I wanted to accomplish this semester. While it didn’t exactly turn out how I expected it to, I ended up really learning a lot and honestly has taught me a lot about how hard it is to create accurate models using live data. Being a computer science minor, I don’t have as much experience as some when it comes to some specific parts of project development such as using APIs and Deployment, so there was a bit of a steep learning curve I didn’t fully expect on top of learning how to create an unsupervised learning model that can predict genres of images. Initially, when I was set on the idea of using the Spotify API as my API of choice for learning these new techniques, I immediately had many ideas of ways I could use it. My first idea was the audio feature analysis, in which the goal was to analyze and visualize songs’ popularity values based on their different features and genre. The hardest part of this mini-project was simply collecting the data from the API. Spotify doesn’t keep track of historical data, and when you use the “Search” endpoint, it only returns 1000 at a time per query. Essentially, this means that I could only get about 1000 songs per genre at a time unless I am able to figure out a clever way to query unique values. I didn’t end up finishing this portion of the project, but it is something I still think is possible through the Spotify API it will just take some more time, which I should have this summer! Another idea I had when I considered using the Spotify API was to create my image genre classifier using the Spotify API’s genre data. Unfortunately, this posed another problem as even though the API’s get genre endpoint says it should return a genre value, it is actually blank most of the time. However, querying by track and artist is possible, so I would essentially search for tracks, verify that the artist has that same genre listed under their information, and then I would classify that track’s album cover as that genre. It isn’t perfect, so I would manually make sure that the images were classified correctly. Also, when I was searching for album images, I made sure that there were no duplicates in each set, and that the album was not used for multiple genres. I also made sure to query out compilations as those typically ended up being very hard to classify and ended up having a lot of duplicates with remixes and such.

When it came to actually creating the image classifier, there was definitely a learning curve, but I started to really get the hang of it towards the end. I ended up using FastAI as online I had seen it being recommended as a good entry-level deep learning tool. To go along with this, watched a few lectures from their main course which is free online on how to create the vision learner! The lectures, however, didn’t always cover the material I needed so I did a lot of digging online for how to fine-tune the model more and how to measure the performance of my model. When it came to the actual performance, I was originally only getting about 38% accuracy with my first model, but after verifying the genre tag with the artists’ genre tag, reducing the batch size of the data loader, and changing the metrics to favor accuracy over error rate, I ended up getting exactly 50% accuracy with my final model. Ideally, I would like this to be a lot higher, but considering the hoops I had to jump through just to get to this point I’m still pretty happy with it. I’m definitely going to do some more research this summer on how to fine-tune the model more, but for how it works now I’m pretty happy.

Lastly, to no one’s surprise, I ended up using Streamlit as it was definitely the easiest option for deployment and I’m happy with how clean it looks! There were a few hiccups when it came to deploying it locally vs. on the cloud, namely with the requirements.txt file and the structure of the repo, but once I figured it out it worked extremely well! Overall, while I would have liked to have gotten a bit more accurate model and had finished the audio feature analysis page, I am really happy with all the new skills I learned in the process, namely learning how to read, understand, and use an API, create a deep learning model, use data visualization to clean data, and deploying using a web app. Below is my confusion matrix and classification report for my genre predictor model.

