Utilizing Neural Networks To Categorize Songs

Max Bermont

Agenda

Define the Problem

The Data

The Model

The Results

Introduction

- Spotify is one of the largest audio streaming services in the world. They have 345 million active users and over 70 million songs.

Defining the Problem

I have been tasked by Spotify to create a model that will assign genres to songs that are being added to the platform.
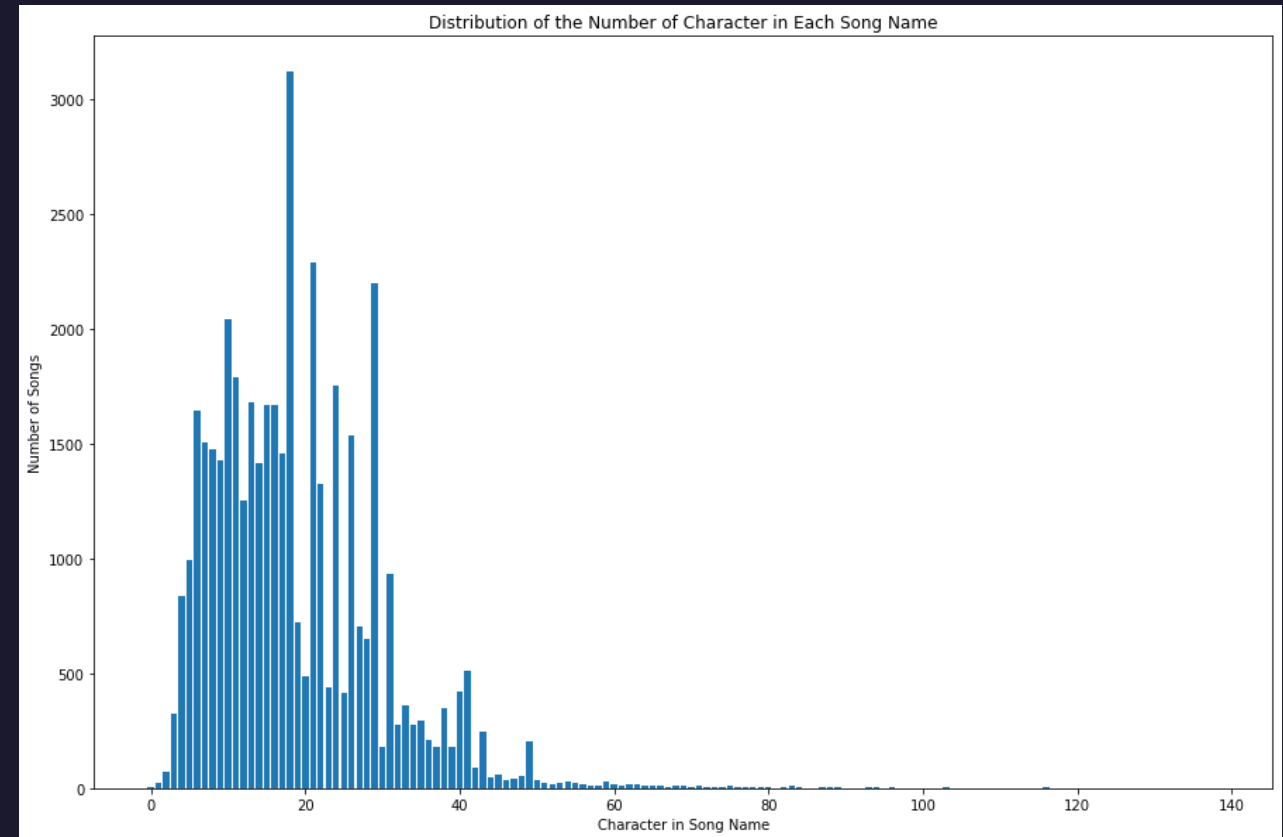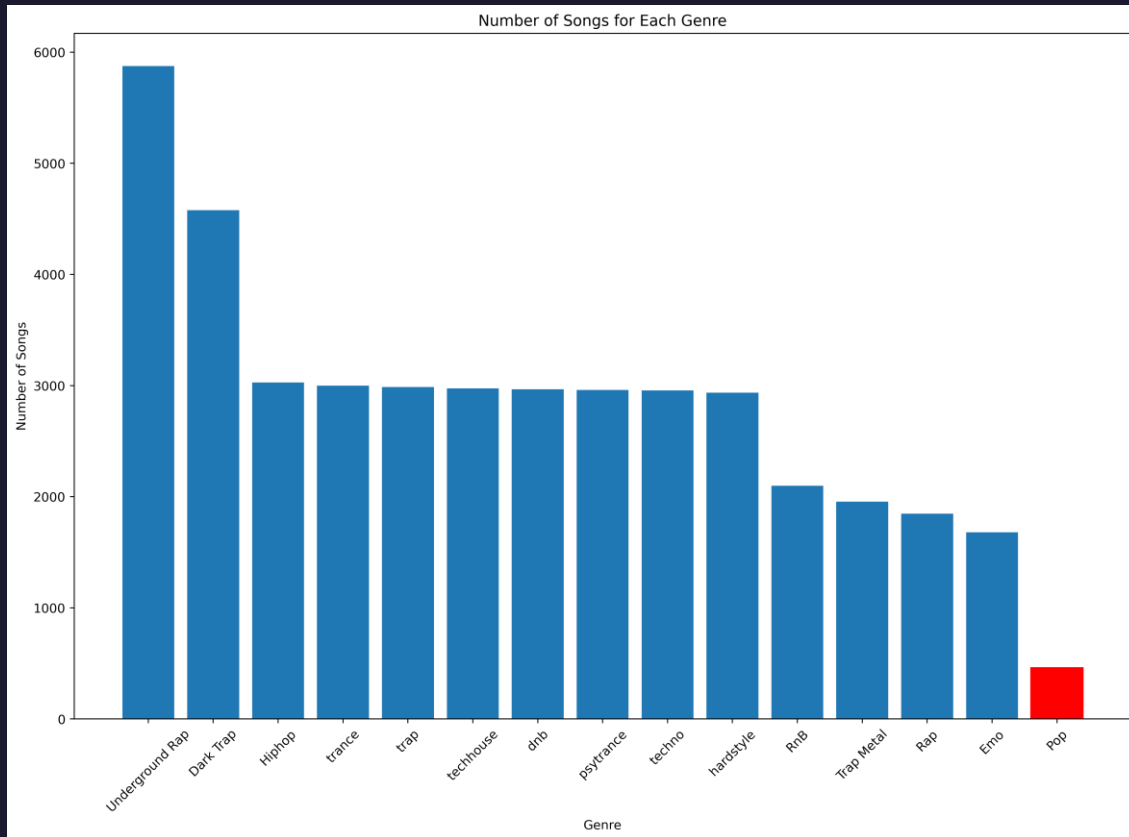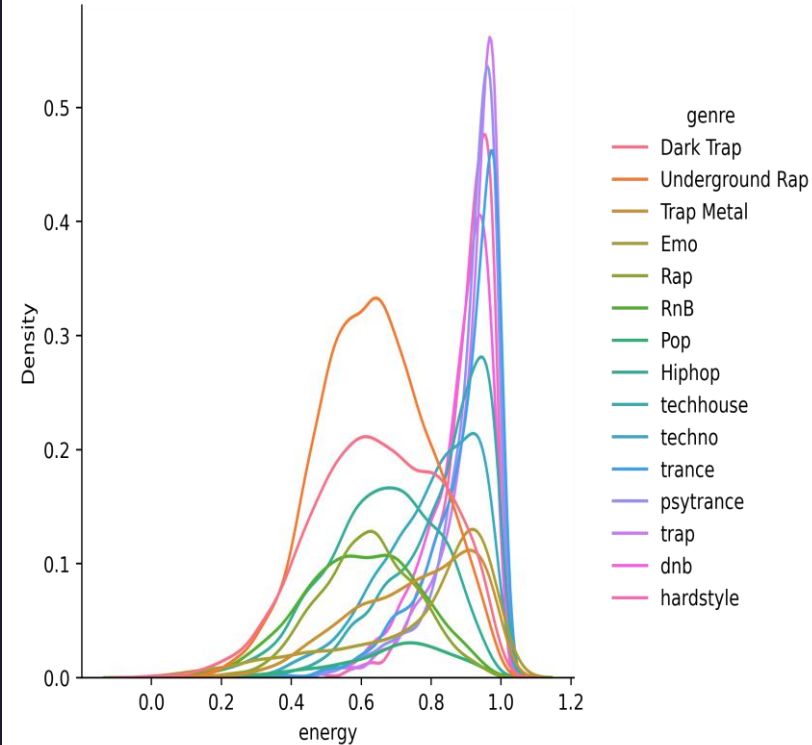
The Data

- From Kaggle

- Data Process

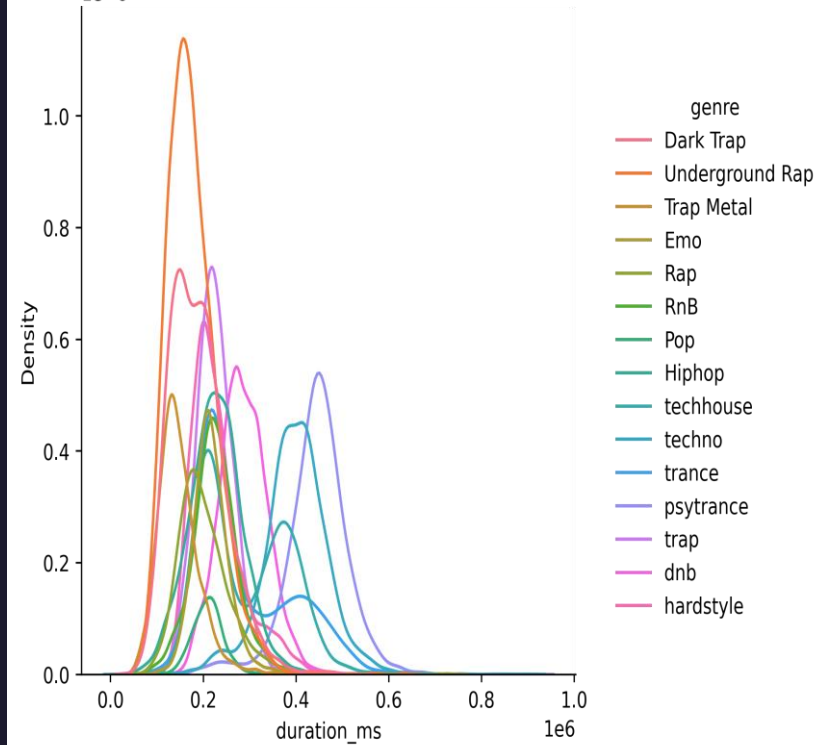- Feature Validation

- Target Values

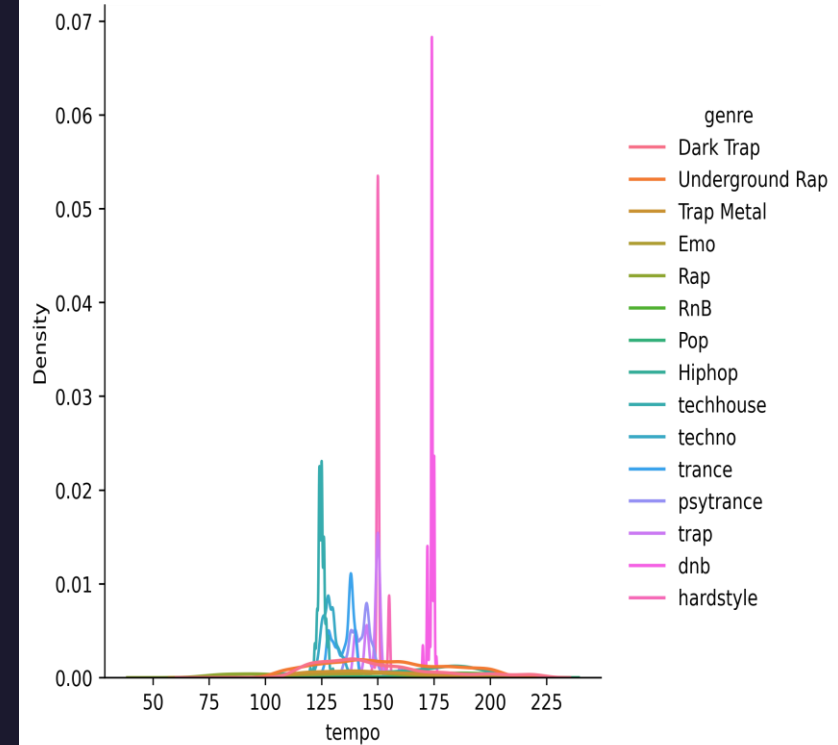# Charts

Charts (contd.)

The Distributions of Energy Percentage Values for Each Genre
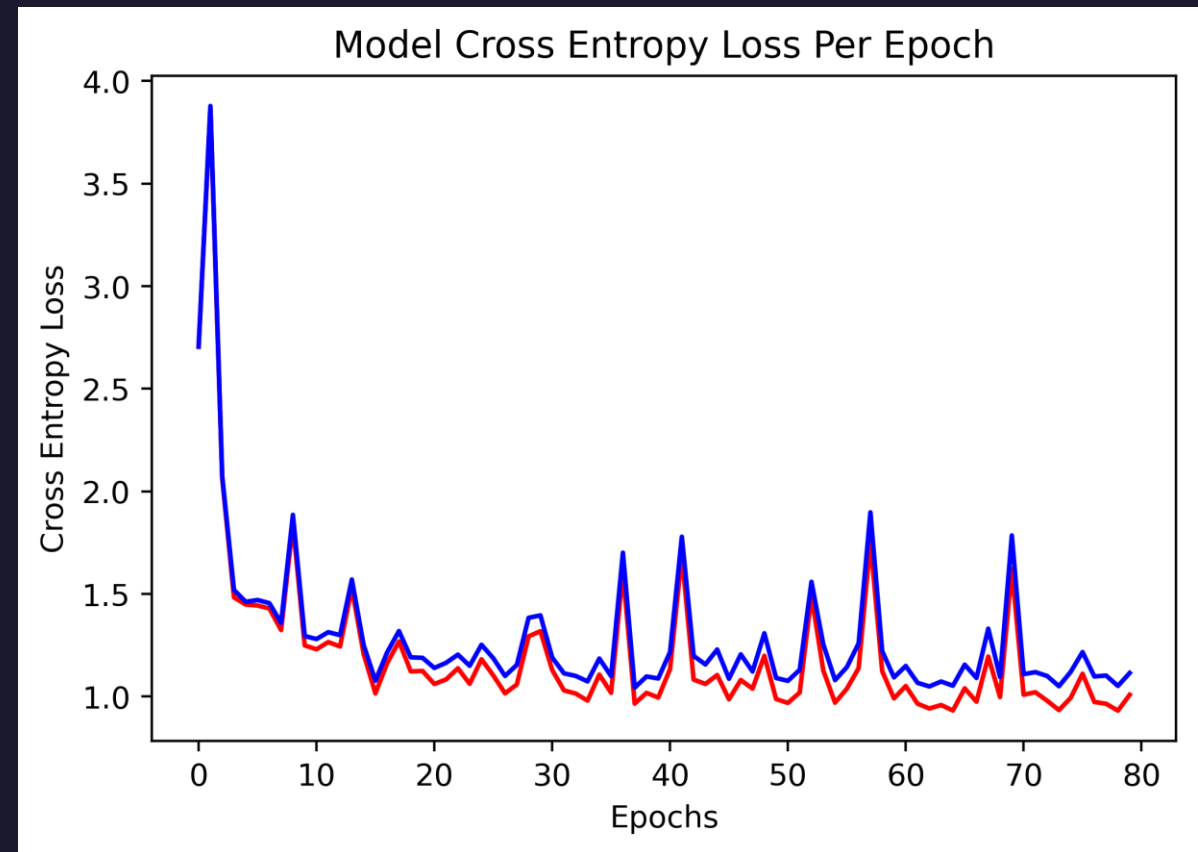

The Distribution of Song Duration for Each Genre

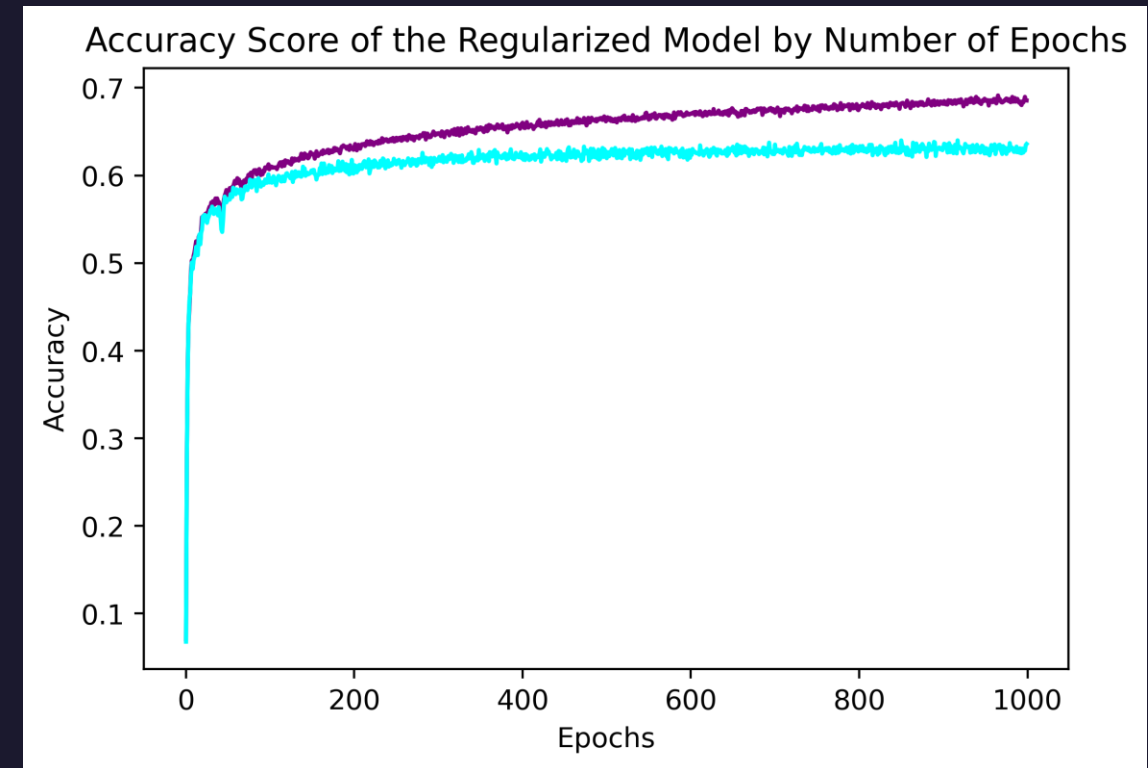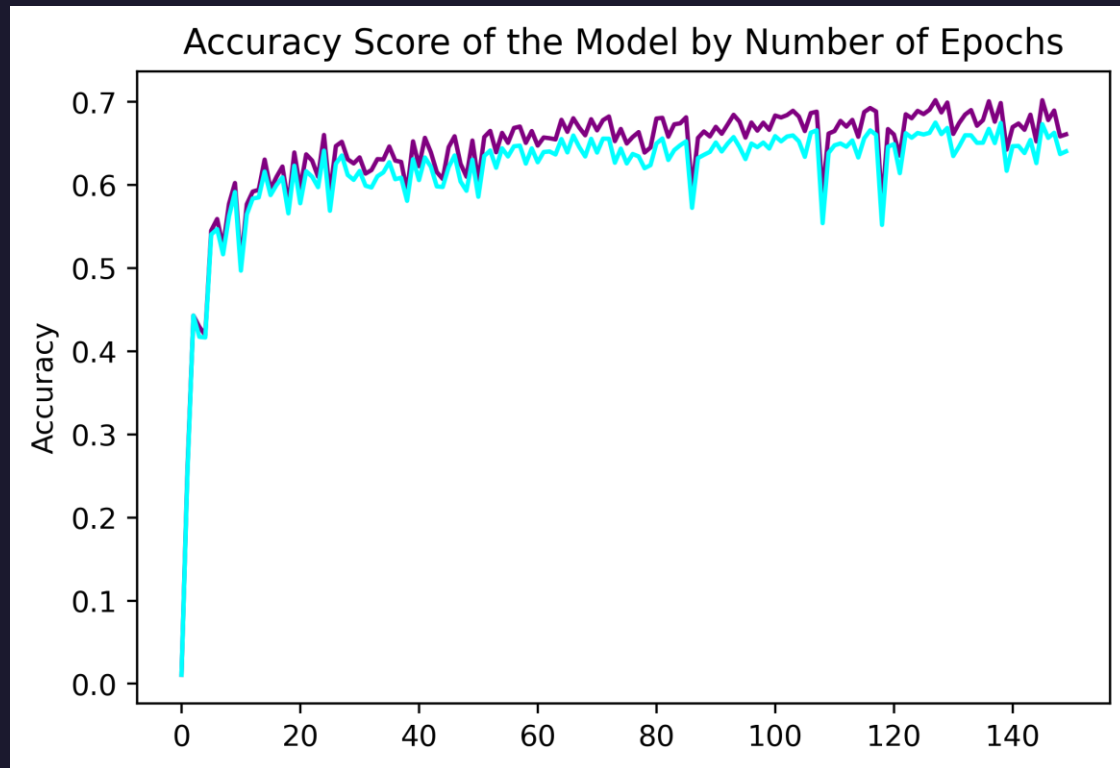The Distributions of Tempos for Each Genre

The Model:

- Multiclassification Neural Network (using PyTorch)

- Baseline Accuracy: ~6.59%

- Best Testing Accuracy of Model: 67.50%

- 6 Layers

  - 4 Hidden Layers

  - 1 Input Layer

  - 1 Output Layer

To Regularize?

Accuracy Score of the Model by Number of Epochs

Accuracy Score of the Regularized Model by Number of Epochs

Conclusion & Recommendations

- I created a neural network model to predict the genre of a song given attribute data about the song.

- This model without regularization is over ten-fold better than the baseline model.

- Utilizing Dropout Regularization did not improve the model in terms of accuracy, due to the data deficiency. It did improve the model by making the model less over fit and far less erratic. However, regardless of the model set up, I have never been able to achieve an accuracy score above 70%

Thank You

Max Bermont

mjbermont@gmail.com

https://github.com/mbermont/