Project Proposal: Drum Generator

When you're learning to play an instrument such as the guitar, it is usually the case that a large portion of that is practicing alone. Learning scales, chords, and even other artist's songs are a great way to set goals and continue to improve. However, it may end up becoming repetitive and boring to listen to the same metronome or the same songs you're trying to master. The idea that I have come up with is to train a system on a dataset of drum tracks that vary in genre and try to generate its own tracks based on the genre you provide.

The dataset that I found for this idea is the 'Groove MIDI' dataset (https://magenta.tensorflow.org/datasets/groove) from Magenta which is comprised of 1,150 MIDI files and over 22,000 measures of drumming. As well, it features a robust set of percussion instruments to utilize as well as 18 different genres to work with including rock, blues, jazz and reggae. MIDI files were chosen for the dataset as Python would be able to easily handle note extraction, manipulation and file creation with the data in this format. As well, MIDI has its own standardized mapping between pitches and the instruments that they represent. This saves a lot of time and effort for the system to figure out what instruments are being played at any point in the track and when creating its own.

Since the focus of the capstone is 'creation', the approach for solving this problem will be the use of a deep learning model called Generative Adversarial Network or GAN. GAN models are separated into two parts, the generator and discriminator. The generator focuses on trying to create the desired output whether that be a new image or in this case a new drum track. The discriminator's job is to classify if the generator's results pass as a real track. The results of the discriminator is back-propagated to the generator, allowing it to improve. Over time the generator should be able to 'trick' the discriminator into believing that their tracks are 'real' drum tracks. The models themselves should not require the user's computer to have a gpu as this project is trying to be tailored to everyone.

The final deliverable for the capstone will be deployed as a web application. The user will have the ability to choose from a variety of genres such as rock, funk and even an option for a mixture of all of the genres combined. Once that is selected, the application will generate a track for them to listen to directly on the site. If they wish, the user can then select another genre or stay with the one they have already selected and create a new one.