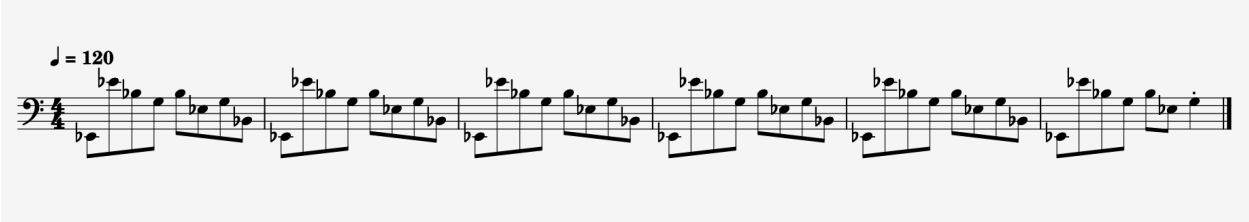
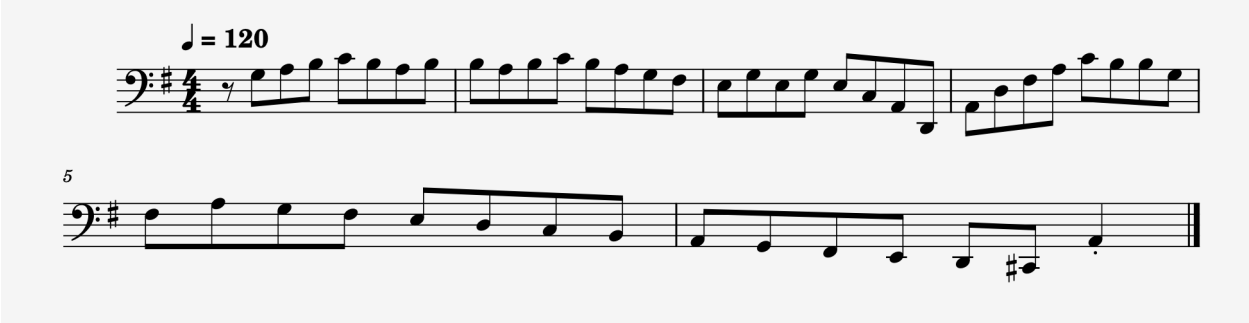
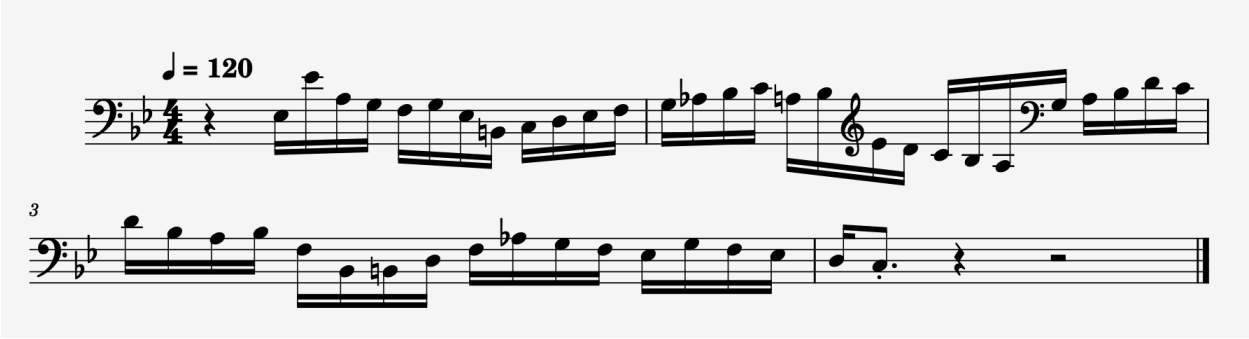
Sarah Groark
Generative AI – Assignment 6
29 November 2024

Summary of Findings

Three outputs were produced by the model with the goal of music being generated following the style of Bach's work. The following three arrangements were the results:



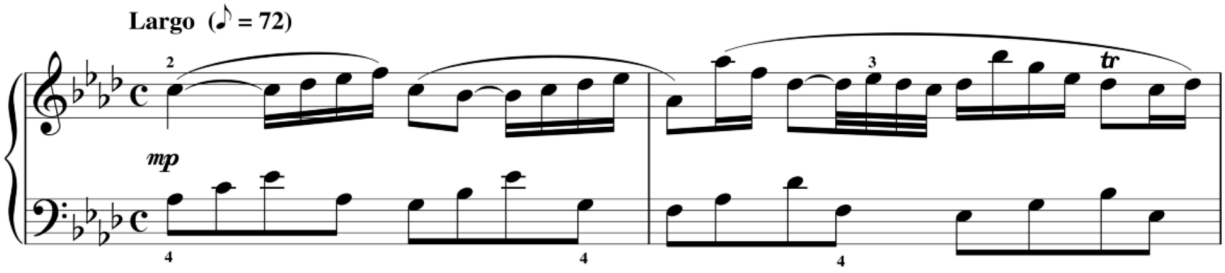




The three outputs each had a slightly different tempo, with one being much faster than the others. However, the three possess the similarity in the arrangement and range of their notes. They all exhibit a wide range of notes being played, so the sounds of each are not monotone but feature a diverse set of notes. The arrangement of these notes seem to also follow a pattern. For the most part, the notes included are part of a riff that go up and down the scale. Overall, this ascending and descending of the scale creates repeated patterns in the generated music, which result in outputs that seem ordered and realistic.

Reflection on model's ability to generate music in the style of Bach

In three different outputs of the model, the generated music produced three different pieces of Bach-inspired music that generally reflects his music style. After some research regarding Bach's specific writing style, the following generalized features are just some of the commonalities in his pieces: symmetrical arrangement, repetition, and structural balance. As mentioned above, there are several repetitions in the generated music of a chord of notes and these are present in each of the model's generated outputs, which is telling of its ability to produce progressions that reflect Bach's music style. Additionally, upon visual inspection of the generated output and comparison to his source music, there is some symmetry that comes into play in this style of composition. The following is a snippet from Bach's Piano Concerto No. 5 in F Minor:



There are some symmetrical similarities with the arrangement of the notes in each of the measures, which is somewhat similar to the generated outputs from above. Furthermore, what results from his inclusion of symmetry and repetition seems to reflect this sort of structural balance. As a result, the listener finds a pattern in the arrangement of notes that becomes familiar to their ear, creating a sort of structure for the piece. The generated outputs from the model also exhibit this sort of feature, as when listening to them, they seem to distinctly follow a pattern that is calculated and not random, making it enjoyable for the listener.

Quantitative Metrics

Due to the temporal nature of music, the approach to quantitative analysis is slightly different than the generation of other data types (ie. text, images). However, the idea remains the same — compare the generated data to the given training data. Researchers from Amazon Web Services and Carnegie Mellon University researched how to utilize a Transformer-GAN and a learned loss function. They utilized several different quantitative metrics, a few of which include: pseudo-log-likelihood (PLL), negative-likelihood (NLL), and classifier accuracy (CA). Due to the complexity of music, this variety of metrics are used to calculate accuracy of certain features. For example, pseudo-log-likelihood is responsible for computations of a local context, which take into account the dependencies between close-by notes. Negative-likelihood is determined by

computing the difference between the generated output and the actual token sequences; NLL also helps the model to fit. Classifier accuracy is used to measure the performance of the model by computing the percentage of time that the model makes accurate sequences of tokens.

Even with these, although there are plenty more that account for more complex intricacies of the music itself, there is still a level of difficulty in quantifying the success of a model's training.

Musical Quality

There are metrics that are used to determine the stylistic accuracy of generated outputs in comparison to the given data. For example, professors of computer science in Taiwan designed several metrics that they could use for their music-generating model. They used these on both the real and generated data. Two of these include EB (ratio of empty bars) and TD (tonal distance - measurement of harmonicity between tracks). These help to measure the finer details of the music to determine how similar, and in what ways, the generated output is to the given training data's style. The distribution of this data on both the real and generated data should converge over the course of training.

Additionally, many researchers seemed to have used human evaluation on their models as a form of subjective evaluation, due to the challenges that arise with quantifying a model's performance and accuracy.

Beyond Bach: Abba

I chose a second music style to analyze that I thought would be both distinct enough to generate and recognize, but also interesting to me. I've loved Mamma Mia for a long time and wanted to see what the generated outputs would be for Abba. Abba's musical pieces are compositions all laid on top of one another, so I chose the piano component to feed to the model. The generated outputs are simple and not as diverse in note ranges as Bach's outputs, but are reflective of Abba's use of piano in their music. The generated outputs feature a sort of musicality element to it that the Bach outputs did not, so I thought it was interesting to see how the model generated different pieces of music based on the input data. The source data of Abba's piano pieces in their music feature a range of approaches – long stretches of the same notes or a wide range of notes all at the same time. Since I had to train this model for a smaller amount of epochs, the outputs largely reflect the former, but still captured some stylistic elements found in Abba's music.