Machine Learning and Music Composition

Daniel Woeste
Division of Math and Sicence
University of Minnesota, Morris
Morris, Minnesota, USA 56267
woest015@morris.umn.edu

ABSTRACT

in the background section,introduce topics: Decision Trees, Over Fitting, Machine Learning, Finite Automata and others.

1. INTRODUCTION

2. BACKGROUND

We use this section to introduce important terminology for the rest of the paper.

2.1 Music

In this section we explain the terminology that is important to the music side of the paper.

2.1.1 Melody

good music definitions, need to figure out how to work them into a conherent paragraph or 2, this is super clunky to me currently

Melodic progression is the breaking down of the 8 note octave into equal parts. Either whole steps of the 12 half steps or semitones. Melody is the pitches and rhythm that are used to be aesthetically pleasing. Thus a melodic progression uses various intervals- the whole steps or half steps in combinations that are pleasing, we hope, to the listener.

The texture of the music also plays an important role in the style of music that can be generated by machine learning programs. Two different styles that are used by the later discussed programs are homophoonic and polyphonic textures. A homophinc texture is a piece of music that only contains one melody played at a single time. It may be played by more that one isntrument simultaneously, but there is not an equally important part of the music being played at that moment. On the other hand polyphonic means that there are more than one melodic line being played at the same time. Polyphoinc music is a more common style to jazz and mroe contemporary pieces, while some of the simpler classical music and pop music tend to favor homophonic due to its easier to grasp and understand style.

2.2 Machine Learning

Machine learning, a subcategory of artificial intelligence, is an area of computer sciecne that focuses on programs that are able to systematically change them selves based upon a learning period. These programs are design to alter themselves without the need for human input, these style of programs are benfical due to the fact that they are not as strictly bound to a set of instruction. Instead, during the learing period, the programs are able to change themselves to best fit however they data demands that work. This style of data drvien learning allows machine learning programs to be well suited for areas such as self driving cars, text prediction, online redommendations, and even hte complex field of music.

2.2.1 Decision Trees

Consider inserting an image about decsion trees if deemed necessary?

Decision trees are a structure that repesent the tests and outcome of an algorith. Every node on the decision tree represents a possible test or choice that an algorithm can make. The resulting branches from that come from that node represent all of the possible outcomes of that test. This can easily be repesented by flipping a coin multiple times in a row. Every time the coin is flipped, a node is created in the descision tree. In this instance, the two nodes for the coin flip are heads and tails. For every subsequent coin flip, a new set of nodes and leaves are generated. This allows the decision tree to contain all possible out comes for the Nth coin flip.

2.2.2 Overfitting

Overfitting is possible side effect that can occur with decision trees. When a decision tree is made and for an application, it can become too complex and dependant on the particular set of data that it was trained upon. This occurs when a tree gains too many parameters, making the tree too specific to the data that it was trained on, or the nodes the tree branches from are inconsequential to the actual results. A concequence of overfitting happens when the tree is then tried on data different from what is was trained on. The tree looses its accurecy at prodicting what the possible results may be.

2.3 Finite Automata

Finite automata, otherwise known as finite state machines, are representations of a an abstract computational model that can exist in only one of a finite number of states at a time. The transistion from one state to another is caused by external inputs to finite automata. Every finite automata has a start state, often represented by the head of an arrow

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/.

UMM CSci Senior Seminar Conference, December 2017 Morris, MN.

pointing the that state indicated as the start state, and an end state, often represented by a secong circle around the last state. The finite automata listen below deviates from these notations, the start state is the state with the value of negative four, and the end state is the state with the value of 4. In this automata they can be determined by the fact that the start state does not have any recuring transistions back to it, and the final state has not possible transition to a new state once it is reached.

2.4 General Framework

All of the methods described in this paper share a similar framework that would be usefull to understand. The commonality between all of the methods is that every note is generated in relation to a cache of previous notes. Meaning that every note is dependant upon a set of notes before it, whether it is only the previous note or a larger set of notes. This allows the program to determine a melodic progression for the music.

3. METHODS

In this section, we will be describing three seperate subcategories of machine learning. The three subcategories of machine learning that we will be describing are *random forests*, *markov chains*, and *neural networks*.

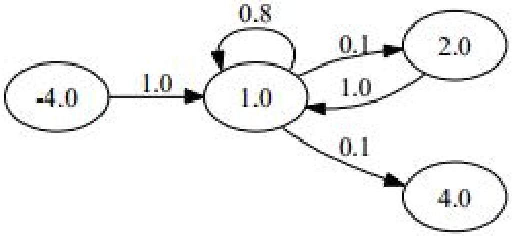

Figure 1: Finite automata detailing the Markov Generated in Automatic Composition of Music with Methods of Computational Intelligence

First, we discuss the use of random forests. Random forests operate by generating and altering a large collection of decision trees. The generated output of the random forest is the result of the average answer of all of the decision trees in the forest. As with all machine learning a set of amount of time is required to teach the algorithm to properly generate answers. For the case of random forest, each decision tree is made on only a small part of the set of data, with overlap among all of the trees. This overlap in the data that the trees are built on helps to reduce over fitting of the trees to the data. When it comes to music we discuss a program known as ALYSIA, that uses two random forests to generate music from a given set of lyrics.

Next, markov chains function fundamentally differently than random forests. Markov chains use a state machine and a statistical models to predict the next value. Figure 1 shows a simple finite automata that generates the note length, but not pitch values for a song. The values 1, 2, and 4 represent quarter, half, and whole notes respectively. The value of -4 is a starting state that holds no value musiclly, other than being a place to start the song. The automata visualizes the way the markov chain uses probabilistic determination to predict what the next note or value will be. In the terms of music this means that for each state in the automata it has a percentage chance of moving to each of the other nodes in the automata. In this case, the first note will be a quarter note with an 80% chance of having a repeated quarter note, and only a 10% chance of having a half note following.

introduce the word trained instead of generated to make sense with the neural networks paragraph

Neural networks offer yet another approach to using computers to generate music, this time being a deterministic style of algorithm. Neurals networks are a style of machine learing that is based upom the functions of the human brain. This is done by generating many highy interconnected nodes (neurons) that are working together to solve a single problem. Similar to random forests, neural networks are trained upon a given set of data to produce a network that will produce a favorable. Unlike the previous two methods, neural networks are almost always deterministic, meaning for the same input a neural network will always have the same output. This differs fundamentally from the other two methods that really heavily on elemnts of chance to determine an outcome.

4. METHODS AND MUSIC

In this section, we will discuss how the previously stated methods can be applied to the composition of music with specific examples.

4.1 ALYSIA

ALYSIA, Automated LYrical Song-writing Application, is a program that uses random forests to generate songs and melodies for a given set of pop lyrics. ALYSIA contains two different random forests that work in tandem: one to produce rhythm, the other to produce pitch values. [ALYSIA] Random forests can provide some inherent benefits over Markov chains. Even though both models work from a cache of previous notes and both predict what the next note should be. Markov chains offer only a statistical probability of what the next note should be, while random forests include a reasoning to what the next note should be.

ALYSIA is able to encode a reason for the next note because, instead of relying on entirely random chance for the next note, it uses a system known as function known as feature extraction to determine the shape of the background music that the ALYSIA is trying to compose music to. The feature extractions includes parameters such as key signature, time signature, note duration, accidental, and many others. It even includes feature extraction for the lyrics that ALYSIA is trying to match as well; such as, word frequency, and even word vowel strength (value extracted from CMU-Dict v0.07 database.) This function allows ALYSIA to examine the previous few notes of the melody, if not the entire melody, in order to predict what the best next note will be.

talk about the co creative nature of ALYSIA mentioned in the paper

Next we will examine neural networks that were used to generate an entirely different style of music, polyphonic music. Polyphonic music is a more complicated style of music that conatins multiple melody lines being played all at once.

This adds an added level of complication to what the neural network has to keep track of due to complexities of music in realtion to music patterns and harmonic realtions between the multiple lines of music. The paper by Daniel D. Johnson covers several types of neural networks including *recurrent neural networs* based upon the *long short-term memory* system.

help!, i am confused byt the lstm and how they work and the recurrent netwrok...and everything

4.2 Citations

4.3

Acknowledgments

Fix the section heards because currently they are not organized in a way that makes sense.