Pop Music Midi Generation using Long Short-Term Memory based Recurrent Neural Network

Don Vincent Goodrich S. Ignacio and Hans Y. Sicangco
College of Computer Studies
Angeles University Foundation
Angeles City, Philippines
ignacio.donvincentgoodrich@auf.edu.ph, sicangco.hans@auf.edu.ph

Dr. James A. Esquivel
College of Computer Studies, Faculty
Angeles University Foundation
Angeles City, Philippines
esquivel.james@auf.edu.ph

Abstract

As music composition and technology grow, the pursuit for music creation with machine learning as a trans-human collaborator is a relatively new venture in the world of artificial intelligence. Various RNN models for music genres such as classical, folk, and jazz has been made but none specifically for pop music. The study aims to fill in this gap. A pop music dataset was created from Billboards' Hot 100 Year End Charts of 2021. The dataset was used to train LSTM-based RNN models which were then evaluated through loss. This revealed that the model with the lowest loss came from the ADAM optimizer making it the best choice with 200 epochs. The results of the chosen model were then evaluated through its perplexity and a listening test. With a low perplexity score, the model was deemed confident in creating novel midi samples. The listening test then tested the created midi samples, resulting in a good rating.

Keywords
A.I., Pop, Midi, LSTM, RNN

1. Introduction

Music is a way of life that has evolved throughout centuries of culture and tradition. From the simple hum of a random tune to the grandiose bands and orchestras, it can be a fact that music is an irreplaceable part of our life. Anything that contains rhythm, melody, and harmony is music no matter the instrument or culture, these elements are present in all styles of music in every period around the world (Epperson 2022). What does music have to do with computer science then? In recent times, the creation of music has been greatly enhanced with technology. The development of artificial intelligence is of much help to that, using algorithms to help create different sounds and tunes to music. Computers have also improved immensely, giving computers the ability to acquire intelligence (Zhang and Yu 2021).

Music and artificial intelligence are not new. Multiple RNNs (Recurrent Neural Networks) have already been made to make trained models of music though mainly limited to classical music. Because of this information, the researchers asked the question, “How about pop music?” Pop music is also known for its simple and catchy hooks and repetitiveness which makes it easier to understand than other genres of music such as classical music (Hoffman Academy Team 2021).

A problem that the researchers have seen is the absence of a pop music database. This may be propagated by the fact that the pop music scene is a fickle space that changes very fast and has many developments in only mere months. This is in contrast to classical music which is very static and rigid due to most music being made in the past. With this, the researchers decided to select Billboard’s top 100 pop songs of 2021. The researchers propose to create a Recursive Neural Network trained with a pop music dataset with the top songs of 2021. This can help in making a pop music dataset for future researchers and making an RNN that is trained with pop music.

1.1 Objectives

General Objective

- to develop a trained LSTM-based RNN that generates pop music samples and implemented in a web application.

Specific Objectives

- to collect music midi files and determine their qualities to be compiled into a pop music dataset and organized into a Data Frame of pitches, durations and steps
- to design and develop an LSTM-based RNN
- to train an LSTM model with a pop music dataset
- to evaluate the model through an optimizer
- to evaluate the music samples the model generates
- to develop a web application that uses the LSTM model for pop music generation purposes

2. Literature Review

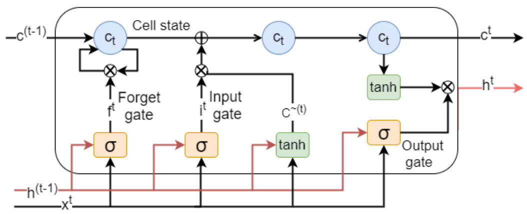
There are studies that have been made way back in 1980 about developments in music (Roads 1980). But in the field of deep learning and neural networks, it is relatively new with developments being made over the past decade. One such example comes from Magenta, an open-source project that is worked on by the AI department of Google to help further advance music and arts through neural networks and machine learning.

Deep Learning for Music Composition

Deep Learning Algorithms have been instrumental in the creation of various papers in line with musical works. Google Magenta is one that has pioneered this venture towards the arts and neural networks. It has garnered the attention of musicians and developers worldwide for their use of deep learning techniques to create new innovations in the field of music (Metz 2017; CBC Music 2021; Livneh 2018). With the music that their algorithms generate, comprehensive studies, open-source models, and interactive web apps, Google Magenta has been frontrunner in this endeavor. Multiple papers have been created from Google's endeavor. These include the Turkish makam music composition by Parlak et al. (2021), Song from Pi by Chu et al. (2017) and other music related research that take Magenta as relevant literature.

In creating a music generating model, RNN is a good choice to start with. RNNs are perfect for sequential data such as speech, music, video, and other time series data (Donges 2021). LSTMs (Long Short-Term Memory) were introduced by Hochreiter and Schmidhuber (1997), as an improvement to RNNs. Unlike RNNs which have 1 channel, LSTMs have 2 channels that pass through 3 gates that help allow it to 'forget' useless information and 'remember' important information. The top channel is the previous cell state, and the bottom channel is what changes the top channel as it moves toward the output for the cell also known as the hidden state. The bottom channel data passes through a sigmoid activation layer and a tanh layer. In the forget gate, it is determined what data is to be kept and what is forgotten. It uses a sigmoid activation function where 1 means to keep and 0 means to forget. Then it passes through the input gate where after passing through a sigmoid and tanh layer, it is multiplied with each other for the sigmoid output to choose what is important in the tanh output. The top channel, with its information is pointwise multiplied and pointwise added to update the previous cell state into a new cell state.

Lastly, it passes through the last gate that decides what will be the output. The new cell state passes through a tanh later and the hidden state passes through a sigmoid layer which is then multiplied and becomes the new hidden layer. The output of both channels is moved on to the following steps. The explanation of the LSTM is in line with Phi's (2018) article. This process can be seen in Figure 1 from Yin's (2018) paper that talks about LSTMs. Conner et al. (2022) state that LSTMs are useful in making predictions for producing outputs in a series. Others are on the same page as LSTMs are also used in other papers such as BachProp by Colombo & Gertsner (2018), and other works concerning music generation. In line with this, Parlak et al. (2021), Chu et al. (2017), and Simon & Oore's (2017) papers use piano music as the baseline instrument in their midi files. The midi files as well vary between papers. Some range from 50 to as many as 130000 midi files depending on the scope of said research.



σ = sigmoid, = pointwise multiplication, = pointwise addition, = previous cell state,
 = previous hidden state ct = new cell state ht = new hidden state

Figure 1. LSTM cell

Music Midi Datasets

POP909 is a study by Wang et al. (2020) where a series of data sets were compared based on multiple factors. These factors include Polyphony, Lead Melody, Audio Time-alignment, Beat, Key, Chord, Modality, and their content. Their dataset, also named POP909, was compared to the following: MAESTRO Dataset, E-piano Dataset, Lakh MIDI, Nottingham Database, RWC POP, CrestMuse and JSB-Chorale. These datasets generally consist of classical piano performances and music from different genres, except for their POP909 dataset. After closer inspection, it appeared to only utilize Chinese pop music, which was not in the researcher's scope. Various datasets have also been created for music research. The Bach Doodle Dataset (Google Magenta 2019), Expanded Groove MIDI Dataset (Google Magenta 2020), Groove MIDI Dataset (Google Magenta 2019), The Largest MIDI Collection on the Internet (Midiman 2016), and many others. None met the pop midi dataset requirements that the researchers were looking for. The researchers decided to create their own dataset for pop music.

3. Methodology

The study's researchers chose the iterative waterfall as the model of development for this project. This software development methodology is a combination of the incremental development approach of the iterative model, where developers could jump back and forth between steps when an improvement could be made, and the rigid structure of the waterfall model which gives structure to the workflow. As seen in figure 2, a workflow aligned with the iterative waterfall development model was chosen. This is A Real-Life Data Science Development Workflow (Byrne 2017) with some modifications on the pre-processing part of the workflow.

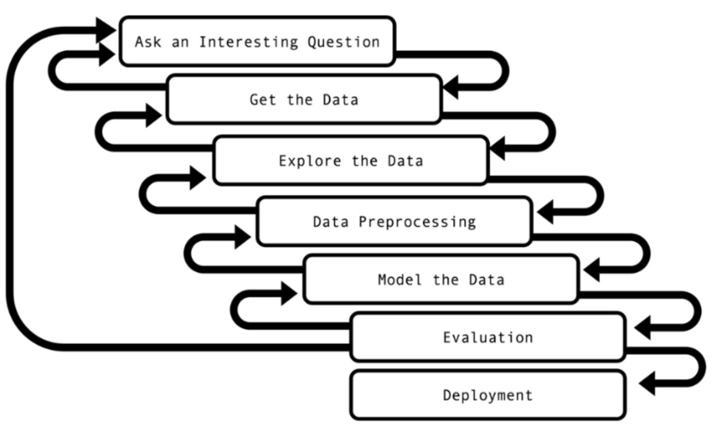


Figure 2. A Real-Life Data Science Development Workflow

Upon examining different studies and datasets available online, the researchers noticed a lack of AI implementations on western pop music. This led them to ask the question “How about pop music?” In order to obtain the data set, a list of songs must first be determined. As the prevailing source for studies concerning the top list of popular songs in each year (Napier and Shamir 2018), the list of songs will be coming from the Billboard's Year End Hot 100 Songs. The list of the most popular songs in 2021 (Billboard 2021) can be seen in Table 1.

Table 1. Billboard's Year End Hot 100 Songs

Billboard's Year End Hot 100 Songs		
Rank	Song Name	Artist
1	Levitating	Dua Lipa
2	Save Your Tears	The Weeknd & Ariana Grande
3	Blinding Lights	The Weeknd
4	Mood	24kGoldn Featuring iann dior
5	Good 4 U	Olivia Rodrigo
6	Kiss Me More	Doja Cat Featuring SZA
7	Leave The Door Open	Silk Sonic (Bruno Mars & Anderson .Paak)
8	Drivers License	Olivia Rodrigo
9	Montero (Call Me By Your Name)	Lil Nas X
10	Peaches	Justin Bieber Featuring Daniel Caesar & Giveon
11	Butter	BTS
12	Stay	The Kid LAROI & Justin Bieber
13	Deja Vu	Olivia Rodrigo
14	Positions	Ariana Grande
15	Bad Habits	Ed Sheeran
16	Heat Waves	Glass Animals
17	Without You	The Kid LAROI
18	Forever After All	Luke Combs
19	Go Crazy	Chris Brown & Young Thug
20	Astronaut In The Ocean	Masked Wolf
21	34+35	Ariana Grande Feat. Doja Cat & Megan Thee Stallion
22	What You Know Bout Love	Pop Smoke
23	My Ex's Best Friend	Machine Gun Kelly X blackbear
24	Industry Baby	Lil Nas X & Jack Harlow
25	Therefore I Am	Billie Eilish
26	Up	Cardi B
27	Fancy Like	Walker Hayes
28	Dakiti	Bad Bunny & Jhay Cortez
29	Best Friend	Saweetie Featuring Doja Cat
30	Rapstar	Polo G
31	Heartbreak Anniversary	Giveon
32	For The Night	Pop Smoke Featuring Lil Baby & DaBaby
33	Calling My Phone	Lil Tjay Featuring 6LACK

34	Beautiful Mistakes	Maroon 5 Featuring Megan Thee Stallion
35	Holy	Justin Bieber Featuring Chance The Rapper
36	On Me	Lil Baby
37	You Broke Me First.	Tate McRae
38	Traitor	Olivia Rodrigo
39	Back In Blood	Pooh Shiesty Featuring Lil Durk
40	I Hope	Gabby Barrett Featuring Charlie Puth
41	Dynamite	BTS
42	Wockesha	Moneybagg Yo
43	You Right	Doja Cat & The Weeknd
44	Beat Box	SpotemGottem Featuring Pooh Shiesty Or DaBaby
45	Laugh Now Cry Later	Drake Featuring Lil Durk
46	Need To Know	Doja Cat
47	Wants And Needs	Drake Featuring Lil Baby
48	Way 2 Sexy	Drake Featuring Future & Young Thug
49	Telepatia	Kali Uchis
50	Whoopty	CJ
51	Lemonade	Internet Money & Gunna Featuring Don Toliver & NAV
52	Good Days	SZA
53	Starting Over	Chris Stapleton
54	Body	Megan Thee Stallion
55	Willow	Taylor Swift
56	Bang!	AJR
57	Better Together	Luke Combs
58	You’re Mines Still	Yung Bleu Featuring Drake
59	Every Chance I Get	DJ Khaled Featuring Lil Baby & Lil Durk
60	Essence	Wizkid Featuring Justin Bieber & Tems
61	Chasing After You	Ryan Hurd With Maren Morris
62	The Good Ones	Gabby Barrett
63	Leave Before You Love Me	Marshmello X Jonas Brothers
64	Glad You Exist	Dan + Shay
65	Lonely	Justin Bieber & benny blanco
66	Beggin’	Maneskin
67	Streets	Doja Cat
68	What’s Next	Drake
69	Famous Friends	Chris Young + Kane Brown
70	Lil Bit	Nelly & Florida Georgia Line
71	Thot Shit	Megan Thee Stallion
72	Late At Night	Roddy Ricch

73	Kings & Queens	Ava Max
74	Anyone	Justin Bieber
75	Track Star	Mooski
76	Time Today	Moneybagg Yo
77	Cry Baby	Megan Thee Stallion Featuring DaBaby
78	All I Want For Christmas Is You	Mariah Carey
79	No More Parties	Coi Leray Featuring Lil Durk
80	What's Your Country Song	Thomas Rhett
81	One Too Many	Keith Urban Duet With P!nk
82	Arcade	Duncan Laurence
83	Yonaguni	Bad Bunny
84	Good Time	Niko Moon
85	If I Didn't Love You	Jason Aldean & Carrie Underwood
86	Knife Talk	Drake Featuring 21 Savage & Project Pat
87	pov	Ariana Grande
88	Just The Way	Parmalee x Blanco Brown
89	Take My Breath	The Weeknd
90	We're Good	Dua Lipa
91	Hell Of A View	Eric Church
92	Rockin' Around The Christmas Tree	Brenda Lee
93	Put Your Records On	Ritt Momney
94	Happier Than Ever	Billie Eilish
95	Single Saturday Night	Cole Swindell
96	Things A Man Oughta Know	Lainey Wilson
97	Throat Baby (Go Baby)	BRS Kash
98	Tombstone	Rod Wave
99	Drinkin' Beer. Talkin' God. Amen.	Chase Rice Featuring Florida Georgia Line
100	Todo de Ti	Rauw Alejandro

After determining the songs to be included, midi files will be required, to obtain the data that will be used in training the model. Musescore is the site where all of the midi files will be procured. Musescore's library of music contains more than a million scores, made by approximately 200,000 musicians, making it the largest library online (Lin and Yang 2021; Musescore, n.d.). With its high-quality scores being produced and curated by official means and user ratings, this allows for the best midi file choices. The researchers plan to use this website as the source of midi files for the dataset.

The data of each song will be kept track on a table in order to determine whether the said midi file can be used for data processing. Data attributes will be aligned with how POP909 (Wang et al 2020) has organized their song list and information. Other information, deemed necessary in this study, will also be added by the researchers. To create the table, Power BI will be used for its visualization features aimed for exploring data. Data attributes included are defined in Table 2 below.

Table 2. Song Attributes

Song Attributes				
Field Name	Data Type	Data Format	Description	Example
Song Title	Text		Title or Name of the song	Diwa ng Pasko
Rank	Integer		Numerical position in the Billboard's Year End Hot 100 Songs List	5
Release Date	Date / Time	DD/MM/YY	Day, Month and Year the song was released	10/10/2022
Artist Name	Text		Name of the song's artist/s	Justin Bieber
Key	Text			C#
Number of Instruments	Integer		Integer corresponding to the number of instruments included inside the midi file	5
Presence of Vocals	Boolean		Determines if vocals are present inside the midi file	True

After implementing the necessary changes, based on the requirements, the midi files will be saved for organizing. Along with the midi files, a text file containing the notes of the song will be produced that will include the pitches, durations and steps of each note (Wang et al. 2020) using the pretty midi library.

In order to use these midi files in training the model, each midi file must first be converted into a DataFrame of pitches, durations and steps. In this study, the notes will be ranging from C3 to C6, giving the note vocabulary a size of 22. In order to normalize the data, the pitches must be converted from its pitch letter string value into a corresponding numerical value. C3 will be considered the 1st index and C6 will be considered the 22nd index. The normalized value can be computed by dividing the note's index by the vocabulary size (Tensorflow 2022).

$$pitch = \frac{pitch\ index}{vocabulary\ size}$$

Lastly, each song will be divided into features and labels. The features will contain a list of 25 notes pertaining to the sequence of notes that will precede the label note. The LSTM-based RNN model will be created with the help of Tensorflow and Keras. The input layer will receive the note sequence which contains each note's pitch, step and duration. As stated by Mauthes (2018), the optimal number of hidden units for an LSTM layer will be 256 for music generation purposes. On the other hand, 3 Dense Layers will be used which corresponds to the outputs: pitch, step and duration (Verma 2021). A loss function will be aiding the optimizer in solving for the loss between the actual note and predicted note. The mean squared error, as suggested by Mangal et al. (2019), will be used. In training the model, it is suggested by Mauthes (2018) to train between 50 and 200 epochs when developing an LSTM model for music generation. The researchers tested varying epochs within a range of 50 to 200, increasing its step by 50.

Multiple optimizers must be tested to find out the best choice that will fit the data (Conner et al. 2022). With this, Giordano (2022) has suggested trying out 2 optimizers for testing, namely, the Adam Optimizer and the Stochastic gradient descent. The optimizer that produces the lowest loss will be chosen. A quantitative approach will be used through a listening test, as previously done by Zou et al. (2021), which will be the basis for the following questions:

- **Tonality** - the relationship between notes revolving around a center tone. Does the music revolve around a certain home key? Do the notes sound like they belong together?
- **Rhythm** - the element of time in music. Did the music stay a tempo?
- **Form** - the structure and organization of a musical composition. Does the music have structure?
- **Melody** - a series of notes played in an order that is memorable and recognizable. Does the melody sound pleasant or predictable?

4. Data Collection

The midi files used for the dataset were sourced from musescore.com. From the 100 songs in the Billboard's Year End Hot 100 Songs, 53 were available to be downloaded by the researchers. This is due to the unavailability of the remaining 47 songs. The researchers, along with a professional musician determined the keys of each song. With the information from Billboard's Year End Hot 100 Songs, a data table was made from it with the following: Number, Song Title, Artist, Original Key, Release Date, Musescore Website, Number of Instruments, and Presence of Vocals. The release date and original key information were taken from Genius.com. The key of the songs was cross-referenced with a professional to make sure they were correct.

5. Results and Discussion

All tracks inside the midi file that was not the melody was removed. The instrument of the melody track was changed into Grand Piano and the tempo of the midi file was set to 120. After exporting the cleaned midi files, they were checked to make sure that the key of each cleaned midi file was in C Major or A minor, and that the bpm was set to 120. The midi files were shuffled and split into 3 groups: Training, Validating and Testing. The percentages and number of midi files for each group are written in Table 3 aligned with the study of Lin and Yang (2021).

Table 3. Dataset Splitting Information

	Percentage	Number of Midi Files
Training	70%	37
Validating	8%	4
Testing	22%	12

The midi files were then passed through a python code to convert them into pandas data frames. Text files were then produced, each containing the data frames composed of pitch, step, and duration. Each data frame was converted into note sequences containing 25 feature notes and 1 label note. A total of 15386 note sequences were created for training. In constructing the model, each layer was declared with the proper specifications, with the help of tensorflow and keras. The input layer received data in the shape of (25, 3) corresponding to the 25 notes in the input sequence with the 3 columns consisting of pitch, step and duration. An LSTM layer with 256 hidden units was used. Lastly, the output layer consists of 3 Dense Layers, also corresponding to the pitch, step and duration. Two optimizers were used in evaluating the model, namely, Adam Optimizer and the Stochastic Gradient Descent. Epoch values of 50, 100, 150, and 200 were used to evaluate the model. This is done to figure out the best variation for the model.

5.1 Numerical Results

From the variations made, 8 models were created for comparison and evaluation. The final loss of each model can be seen in Table 4.

Table 4. Optimizer Comparison

	Adam Optimizer	Stochastic Gradient Descent
50 epochs	0.3148689568042755	11.938611030578613
100 epochs	0.27600395679473877	11.734061241149902
150 epochs	0.25887686014175415	10.703495025634766

200 epochs	0.24806489050388336	8.736026763916016

Upon training the 8 models, the model trained with the Adam Optimizer with 200 epochs resulted with the lowest loss value. The models that have used the Adam Optimizer had the lower loss values. The models that have used the Stochastic Gradient Descent had the lowest loss value of 8.736026763916016, which had a difference of 8.487961873412133 with its Adam Optimizer counterpart. Generally, the Adam Optimizer has worked out significantly better in terms of reducing the loss of pop music generating models. With a low loss value of 0.24806489050388336, the model trained with the Adam Optimizer with 200 epochs had been chosen to be the final model to be evaluated through the listening test. Ten songs from the testing dataset were randomly picked to be included in the listening test. The listening test rated each sample in a five-point scale ranging from Very bad (1) to Excellent (5) in terms of tonality, rhythm, form, and melody. A total of 17 evaluators participated in the listening test. Table 5 contains the results of the listening test.

Table 5. Listening Test Individual Results

	Tonality	Rhythm	Form	Melody
Music Sample 1	3.41176471	4.47058824	3.82352941	3.47058824
Music Sample 2	4.11764706	4.47058824	4.05882353	4.05882353
Music Sample 3	4.17647059	3.76470588	3.17647059	3.47058824
Music Sample 4	4.58823529	4.52941176	4.35294118	4.58823529
Music Sample 5	3.70588235	4.05882353	3.41176471	3.29411765
Music Sample 6	4.17647059	4.58823529	4.17647059	4
Music Sample 7	3.58823529	4.11764706	3.29411765	3.17647059
Music Sample 8	4.35294118	4.47058824	4.05882353	4.11764706
Music Sample 9	3.94117647	3.94117647	3.76470588	3.64705882
Music Sample 10	4.52941176	4.58823529	4.11764706	4.11764706

The listening test was done to evaluate the midi samples made. Among the 10 music samples, sample 4 had the highest tonality rating with a value of 4.58823529, sample 6 and 10 had the highest rhythm rating with a value of 4.58823529, sample 4 had the highest form rating with a value of 4.35294118, and sample 4 had the highest melody rating with a value of 4.58823529. This makes sample 4 the best generated sample out of all. Meanwhile, sample 1 had the lowest tonality rating with a value of 3.41176471, sample 3 had the lowest rhythm rating with a value of 3.76470588, sample 3 had the lowest form rating with a value of 3.17647059, and sample 7 had the lowest melody rating with a value of 3.17647059. It is important to note that even if these samples had the lowest ratings, their values are still between the neutral and good values of the scale. When graded generally, Tonality had a result of 4.058823529 and Rhythm had a result of 4.3 meaning that they had a good rating. Form was 3.823529412 and Melody was 3.79411765 which means that they have a neutral, leaning into good, rating. Overall, the 10 music samples garnered the following scores for the tonality, rhythm, form and melody of its produced samples, shown in Table 6.

Table 6. Listening Test Grouped Results

Tonality	4.058823529
Rhythm	4.3
Form	3.823529412
Melody	3.79411765

In the study of Ji et. al (2020), the two most commonly used metrics for evaluating models are loss and perplexity. A low perplexity shows that the model trained is suitable for unseen data which makes it have the ability to generate more novel music. The perplexity of the model is 3.2532912837. With a low perplexity score of 3.2532912837, the model is deemed to generate more novel music confidently, as it's not perplexed by the data it has been given. This also means that the model doesn't get confused on what music note to give upon generating the music.

5.2 Graphical Results

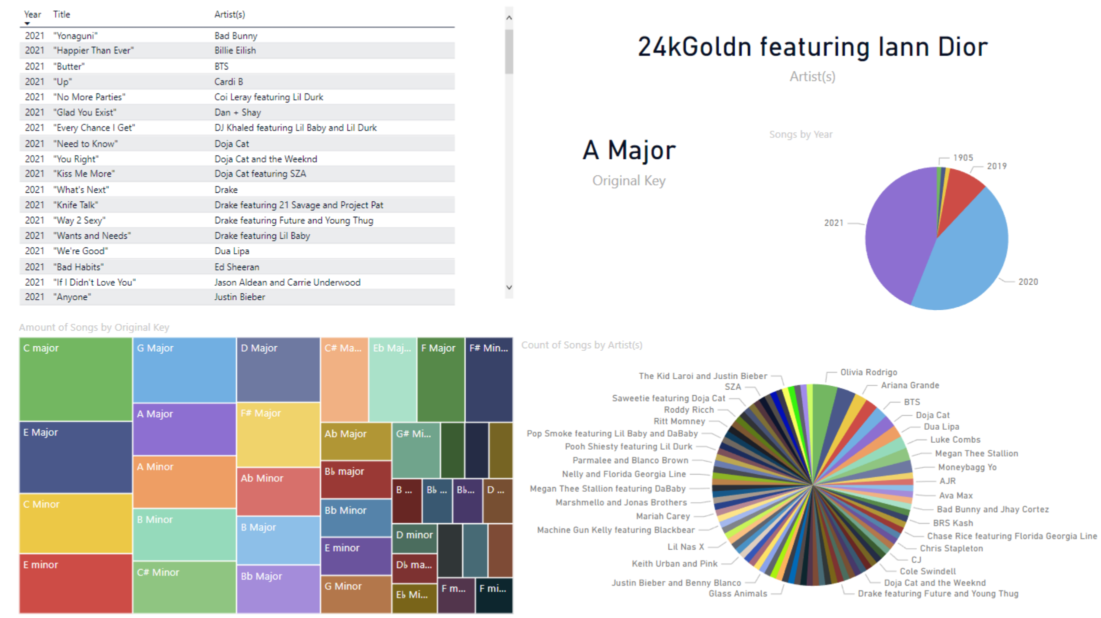


Figure 3. Power BI Visualization

According to the pie chart in figure 3, with the release dates of each song, majority of the Billboard's Hot 100 Year-End Charts came from the years 2020 and 2021. It is interesting to note that there are songs included that were released decades before 2021. These songs are Rockin Around the Christmas Tree from 1905 and All I Want for Christmas is You from 1994. Both are christmas songs which are mostly played in the holiday season. The artist which had the most songs in the database is Olivia Rodrigo with 4 songs, followed by The Weeknd with 3 songs, and everyone else having 2 to 1 song each, as seen in figure 3. In figure 3, the key that had the most songs was in C major scale with 7 songs. This is followed by E Major with 6 songs, then C minor, E minor and G major with 5 songs. It is also interesting to note that both major and minor scales of C and E were the top 4 keys in the dataset.

5.3 Proposed Improvements

An article by Tham (2021) suggests that other neural networks can be used to make similar music generating models. With this, the researchers suggest to design and develop a pop music generating Convolutional Generative Adversarial Network, aligned with the study of Yang et al. (2018), trained with the pop music dataset created in this study and determine which model works better for the dataset created. As stated in Chapter 2, some studies go as far as using 130000 midi files in training their models. The researchers suggest to widen the scope of the study, in terms of the songs included in the dataset, in order to involve more midi files in training the designed model and determine if this could create better outputs for the model. With the number of midi files involved in this study, an automated process may be needed in cleaning the midi files needed for training the model. It is also suggested to improve the existing model by determining the best number of feature notes, in terms of loss and perplexity, to be used in training the model, especially in the context of pop music.

5.4 Validation

The paper sought to answer the question "What about pop music?". Specifically, to expand research on LSTM RNN's using pop music and help musicians by giving them pop music midi samples to use as inspiration or continuation in music making. The paper addresses this by first, creating a pop music database consisting of the Billboard's Hot 100 Year-End Charts in the form of midi files used in this study and to be used by future research. Next, by training the model with the dataset and accepting midi files consisting of 25 notes to make a midi output based on the trained model. Evaluation metrics such as mean squared error and perplexity were used to evaluate the model. A listening test was also conducted to evaluate the qualities of each music sample.

5. Conclusion

The study successfully accomplished its objectives starting with the creation of a dataset consisting of songs from the Billboard's Hot 100 Year-End Charts. An LSTM-based RNN Model was also designed and developed with the proper specifications. It was trained using the dataset created with various optimizers and epoch values. The multiple models were evaluated using the mean squared error loss function wherein the model with the lowest loss value was chosen for the final model. With its low loss value and perplexity score, the model was ready to create music samples. The music samples were then subjected to a listening test where the results showed that in terms of tonality, rhythm, form, and melody, they had a rating from 3.7 - 4.3 in the 5-point scale, which are interpreted as having a good rating. Lastly, a web application was also developed to make accessing the model much more convenient. These were all accomplished using the programming language Python, with the help of external libraries such as tensorflow, prettymidi, etc. With the accomplishments written above, all the objectives have been met. This study has filled in the gap, being the lack of pop music research in deep learning.

References

Billboard, YEAR-END CHARTS Hot 100 Songs 2021, Available: https://www.billboard.com/charts/year-end/hot-100-songs/, Accessed on May 18, 2022.

Byrne, C., Development Workflow for Data Scientists, Available: https://resources.github.com/downloads/development-workflows-data-scientists.pdf, May 25, 2022.

CBC Music, Artificial intelligence is shaping the future of music — but at what cost?. Available: https://www.cbc.ca/music/artificial-intelligence-is-shaping-the-future-of-music-but-at-what-cost-1.5952955, Accessed on March 14, 2022.

Chu, H., Urtasun, R., & Fidler, S., SONG FROM PI: A MUSICALLY PLAUSIBLE NETWORK FOR POP MUSIC GENERATION. *Arxiv,* 2017. https://doi.org/10.48550/arXiv.1611.03477

Colombo, F., Brea, J., & Gerstner, W., Learning to Generate Music with BachProp. *Proceedings of the 16th Sound and Music Computing Conference*, p. 380-386, June 12, 2019.

Conner, M., Gral, L., Adams, K., Hunger, D., Strelow, R., & Neuwirth, A. (2022, March 23). Music Generation Using an LSTM, *Midwest Instruction and Computing Symposium,* 2022.

Donges, N., A Guide to RNN: Understanding Recurrent Neural Networks and LSTM Networks. Available: https://builtin.com/data-science/recurrent-neural-networks-and-lstm, May 25, 2022.

Epperson, G., Music | Art Form, Styles, Rhythm, & History, Available: https://www.britannica.com/art/music, Accessed on March 2, 2022.

Giordano, D., 7 tips to choose the best optimizer - Towards Data Science. Available: https://towardsdatascience.com/7-tips-to-choose-the-best-optimizer-47bb9c1219e, May 24, 2022.

Google Magenta, Bach Doodle Dataset. Available: https://magenta.tensorflow.org/datasets/bach-doodle, Accessed on May 25, 2022.

Google Magenta, Datasets. Available: https://magenta.tensorflow.org/datasets/, Accessed on March 14, 2022.

Google Magenta, Groove MIDI Dataset. Available: https://magenta.tensorflow.org/datasets/groove Accessed on May 25, 2022.

Google Magenta, MusicVAE: Creating a palette for musical scores with machine learning. Available: https://magenta.tensorflow.org/music-vae, Accessed on March 14, 2022

Google Magenta, The Expanded Groove MIDI Dataset. Available: https://magenta.tensorflow.org/datasets/e-gmd, Accessed on May 25, 2022.

Google Magenta, The MAESTRO Dataset. Available: https://magenta.tensorflow.org/datasets/maestro, Accessed on May 25, 2022.

Hoffman Academy, Classical vs. Pop Music for Piano Learners. Available: https://www.hoffmanacademy.com/blog/classics-vs-pop/, Accessed on May 25, 2022.

Ji, S., Luo, J., & Yang, X., A Comprehensive Survey on Deep Music Generation: Multi-level Representations, Algorithms, Evaluations, and Future Directions. Available: https://arxiv.org/pdf/2011.06801.pdf, March 14, 2022.

Lin, C., & Yang, M., *InverseMV: Composing Piano Scores with a Convolutional Video-Music Transformer*. Cornell University. 2021, https://doi.org/10.48550/arXiv.2112.15320

Livneh, N., Can AI Create Better Music Than Humans? Available: https://www.forbes.com/sites/forbestechcouncil/2018/11/26/can-ai-create-better-music-than-humans/?sh=47ebcbf11936, Accessed on March 14, 2022,.
Mangal, S., Modak, R., & Joshi, P., (2019, August 2). LSTM Based Music Generation System. *International Advanced Research Journal in Science, Engineering and Technology*, Vol. 6, Issue 5, pp. 47-54, 2019.
Mauthes, N., VGM-RNN: Recurrent Neural Networks for Video Game Music Generation. *Master's Projects*, p. 595. 2018.
Metz, C., How A.I. Is Creating Building Blocks to Reshape Music and Art. The New York Times. Available: https://www.nytimes.com/2017/08/14/arts/design/google-how-ai-creates-new-music-and-new-artists-project-magenta.html, Accessed on March 14, 2022.
Musescore, About Musescore, Available: https://musescore.com/about, Accessed on May 18, 2022.
Napier, K., & Shamir, L., Quantitative Sentiment Analysis of Lyrics in Popular Music. *Journal of Popular Music Studies*, 30(4), pp. 161–176., 2018, https://doi.org/10.1525/jpms.2018.300411.
Parlak, I., Cebi, Y., Isikhan, C., & Birant, D., Deep learning for Turkish makam music composition. *Turkish Journal of Electrical Engineering & Computer Sciences*. pp. xx – xx, May 20, 2021.
Phi, M., Illustrated Guide to LSTM's and GRU's: A step by step explanation. Available: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21 May 25, 2022.
Roads, C., Artificial Intelligence and Music. *Computer Music Journal*, vol. *4*, no. 2, pp. 13–25, 1980, https://doi.org/10.2307/3680079
Simon, I., & Oore, S., Performance RNN: Generating Music with Expressive Timing and Dynamics. Available: https://magenta.tensorflow.org/performance-rnn, March 14, 2022.
Simoni, M., Algorithmic Composition: A Gentle Introduction to Music Composition Using Common LISP and Common Music. Available: https://quod.lib.umich.edu/s/spobooks/bbv9810.0001.001, March 25, 2022.
TensorFlow, Generate music with an RNN | TensorFlow Core. Available: https://www.tensorflow.org/tutorials/audio/music_generation, Accessed on May 24, 2022.
Verma, Y., A Hands-On Guide to Automatic Music Generation using RNN. Available: https://analyticsindiamag.com/a-hands-on-guide-to-automatic-music-generation-using-rnn/, May 24, 2022.
Wang, Z., Chen, K., Jiang, J., Zhang, Y., Xu, M., Dai, S., Gu, X., & Xia, G., POP909: A POP-SONG DATASET FOR MUSIC ARRANGEMENT GENERATION. *In Proceedings of 21st International Conference on Music Information Retrieval*, pp. xx-xx, Montreal, Canada, August 17, 2020.
Yin, H., *LSTM Cell* [Illustration]. Research Gate. Available: https://www.researchgate.net/figure/Long-Short-Term-Memory-LSTM-cell-Fundamental-components-of-an-LSTM-cell-are-a-forget_fig1_329619613, May 24, 2022.
Zhang, C., & Lu, Y., Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, vol. 23, no 1, September 2021.
Zou, Y., Zou, P., Zhao, Y., Zhang, K., Zhang, R., & Wang, X. (2021, November 3). MELONS: GENERATING MELODY WITH LONG-TERM STRUCTURE USING TRANSFORMERS AND STRUCTURE GRAPH. *Arxiv*, November 3, 2021.

Biography

Hans Y. Sicangco is an Undergraduate in Angeles University Foundation taking up his Bachelor of Science in Computer Science majoring in Data Science. He is a University Scholar, a Microsoft Learn Student Ambassador, and a member of the Honor Society. He loves music and likes to make projects that revolve around music.

Don Vincent Goodrich S. Ignacio is an Undergraduate in Angeles University Foundation taking up his Bachelor of Science in Computer Science majoring in Data Science. He is a University Scholar, a member of the Honor Society and the Junior Philippine Computer Society.

James A. Esquivel is a Professor of Computer Science at Angeles University Foundation and a recipient of the Most Outstanding IT Educator award. He holds a Bachelor of Science degree in Mathematics Major in Computer Science from University of Santo Tomas, a Masters in Computer Science from De La Salle University, and a Doctorate in Information Technology from Angeles University Foundation. Has worked on Power-aware cloud computing infrastructures, green computing, and enterprise mobile apps development.