EMOTIONAL MUSIC GENERATION: AN ANALYSIS OF EFFECTIVENESS AND USER SATISFACTION BY USING PYTHON AND DART

Crystal Chong[1], Ang Li[2]

[1]Chandler School, 1005 Armada Dr, Pasadena, CA 91103
[2]Computer Science Department, California State Polytechnic University, Pomona, CA 91768

ABSTRACT

An issue that is prevalent in today's society is the need for new music to be generated. More people are uploading videos and other forms of content to the internet through social media, and videos can often be enhanced by adding music to them [6]. However, creating music can be a time-consuming and expensive process. Therefore, an application was created that can generate music using emotions as inputs for the music generation model. To test how well the method of music generation through sentimental analysis works, an experiment was conducted that tests how accurately a sample of participants believe that the generated music was on a scale from one to ten [7]. According to the results of the experiment, the application appears to do fairly well at generating music that accurately represents the sentiment that was intended in the inputted message. A survey was also conducted to test user satisfaction when working with the application to generate music. The feedback from the participants indicated that they were generally satisfied with how well the generated music matched their intent in the inputted message, and they also seemed to be very satisfied with how convenient the application was to use and how intuitive the user interface was [8]. However, as the ratings for convenience were much higher than the ones regarding the effectiveness of the music generation itself, this may indicate that the application still has room for improvement when it comes to recognizing the sentiment of the inputted message.

KEYWORDS

Music generation, Sentimental analysis, Machine learning

1. INTRODUCTION

Music has been a vital part of human history, and music can influence how someone feels and provide a path for people to express their thoughts and emotions. For this reason, music has been utilized in movies, plays, and other forms of entertainment. Furthermore, music has been demonstrated to provide benefits such as lowered stress levels, boosted cognitive performance, and improved mood [1]. There are many various genres of music, from country to pop to death metal; this means that no matter what sort of preferences a person may have, there is likely a type of music that exists to suit any person's tastes. While music can have the consequence of hearing loss due to being played too loudly, this can be avoided fairly easily if the music is played responsibly at a reasonable volume, or concerts are not overly visited.

David C. Wyld et al. (Eds): CCSEA, AIFU, EMSA, NLCAI, NCOM, SIPRO, SEA, DKMP, BDML, BIoT, CLOUD - 2023
pp. 219-228, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.130519

An explanation of why music is so important is that it is a form of expression and bonding. Musicians are provided with a channel to express themselves, and listeners can feel the emotions that the musicians intended. Music is often played during parties and other gatherings, as other people who have heard the songs before may be able to dance or sing to music together as a bonding activity. Music is often used as a part of culture and sharing this cultural music with others is a way to spread and combine cultures to make a more diverse and inclusive society.

Currently, something that has been done to meet the demands of new and fresh music is AI-generated music [9]. With the ability to constantly create new music without the requirement of manpower, AI-generated music is a very cheap and efficient solution that is available for the general public to use. However, the applications regarding AI-generated music that are currently available are not as intuitive to use as people may want them to be. In an ideal scenario, users could tell the AI exactly what they want to be included in their music using words, and the AI would construct the music exactly as the users instruct them to. However, the applications that we currently have access to are not that advanced. Instead, AI-generated music is performed by inputting music that already exists and building a similar song, or the music is generated based on factors such as whether the song should be in major or minor. With the first-generation method, a similar song that exists is required, and that may not be possible if the user wants to generate a unique kind of song and is not experienced in music creation [10]. The second-generation method has its flaws as well. Although users can take control of several factors, the options are generally quite limited. Rather than creating the song exactly to their liking, users will have to instead opt for the next best choice of generating a song that is only vaguely reminiscent of the song concept that they envisioned. With how restricted current AI technology involving music generation appears to be, there is still much room for improvement in the field. Something that is still a relatively untouched part of AI-generated music is the use of sentences and language to directly influence the outcome of generated music.

The tool that was created to help people with creating new music based on their needs is a mobile application that involves AI-generated music. Rather than commissioning people to create music, which would be both expensive and time-consuming, the artificial intelligence will simply take in a message that the user inputs, determine what sentiment the message gives off, then generates music based on what the sentiment of the user-inputted message was determined to be. This application may sound similar to other applications or programs involving AI-generated music. However, what separates this application from others is its intuitiveness. While other applications ask their users to check certain boxes and toggle certain options to tell the artificial intelligence how the music should be made, this application only requires the input of a message. Many members of the general public wish for applications to be easily used, and they may completely ignore complicated-seeming applications because they believe that taking the time to learn how those applications work is not worth the effort. On the other hand, providing these users with something incredibly easy and convenient to operate may make them give the application a chance and even potentially become long-time users of the application.

The effectiveness of the application was tested by combining an experiment with a survey. First, twelve participants were gathered to download the application and input ten different messages into the application. With each inputted message, the participant would listen to the music that the application outputted and determine whether the music accurately reflected the sentiment inside the message. The participant would record the number of times that the application successfully outputted music that represented the inputted message's sentiment, and this data is to be recorded in a table. The survey would be conducted immediately after the experiment is over, which ensures that the experience of using the application is still fresh in the participants' minds. The survey would ask whether the application was able to effectively create music based on their experience with the application in the experiment and whether the application was convenient to

use and intuitive. The participants are offered a scale from one to ten to rate how they feel about each aspect of the application. The survey also provides a free-response section at the bottom of the survey to allow for more customized feedback. While both the experiment and the survey are being conducted, each participant is encouraged to not collaborate with any other participants when completing these, so that the data represents the participants' uninfluenced opinions of whether the sentiment is accurately reflected or not. The experiment's purpose is to determine whether the application is successful in terms of whether the application can provide users with the music that they envision. The survey's purpose is to determine whether the general user interface and music generation process is smooth, intuitive, and performs up to standards from the user's perspective.

My paper will be organized so that the remainder is split into five separate sections. Starting with Section 2, the challenges that were encountered when planning the project and developing the application will be described in detail. The next section, Section 3, will cover how the application was created, starting with a general explanation before diving into a deeper view of how each feature and section was implemented. Section 4 will go over the experiments that were performed to test the effectiveness of the application. Section 5 compares and contrasts this work with related works. Lastly, Section 6 will give concluding remarks as well as a reflection on how the application could be improved in the future.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. A Method to Generate New Music

The first challenge that was faced when starting the project was coming up with a method to generate new music. Human-made music was not an option, as this was specifically what was trying to be avoided. The most reasonable option seemed to be using an AI model to create the music. With an AI model, no manpower will be required in the creation of new music, and besides possible small server costs, this option was relatively cheap. Another choice to make was what device the application should be on. Considering that many people in today's society spend much of their time on a smartphone or mobile device and carry it wherever they go, making the application runnable on those devices would likely be best. While it would be helpful to have the application available on computers as well, a mobile application was decided upon as the top priority.

2.2. What Features Should be Included in the Application

The next challenge that was encountered was what features should be included in the application. The application's main purpose is to generate music, but it may be difficult to gauge how much should be included within the application. Ideally, the application should be simple to use and intuitive, yet have enough functionality for people to use this application as a reliable tool. Furthermore, the application should be lightweight, so that it takes very little storage and is quick to run. Keeping this balance in mind, the application only includes the necessary features of generating a song based on the emotion it was provided with as input. With an application that only has a single goal and a clean, distinguishable user interface, new users of the application will likely have no trouble using it. Although offering users with many different ways to generate music was considered, the idea seemed as if it would clutter the application too much, and this idea was ultimately scrapped.

2.3. The Implementation of the Functionality Itself

A third challenge was the implementation of the functionality itself. Because no easily accessible and reproducible version of artificial intelligence code could take in emotions as input to create music, some steps were taken to achieve this result. First, the problem was broken down into two main parts; the first part was assigning an inputted message to an emotion such as happiness or sadness, and the second part was selecting a song based on the emotion that the AI model recognized. Both parts were done within the same Python file, which allowed all the code to be run in one place and helped with implementing the functionality of the code more conveniently. The two parts of the code individually were relatively simple to find examples of on the internet. Although the functionality of the code seemed extremely difficult at first, diving the problems into more understandable and approachable sub-problems allowed the implementation to be a success.

3. SOLUTION

The application was created using Python and Dart as its programming languages. Flutter is an open-source framework backed by Google that is used as the front end of the application, as Flutter is proficient at creating application interfaces. Flutter is capable of creating new screens and all the necessary features that would be required in the application screens, such as buttons, text boxes, images, and layouts. An example of where the use of Flutter's functionality can be seen in the application is on the home screen, as a button is programmed to move to a different screen in the application when pressed [11]. Python is primarily used for the back end of the application since Python has many different libraries and frameworks to choose from. Therefore, any sort of machine learning or artificial intelligence feature that may be needed for a project will most likely be found in one of the numerous available libraries. In the case of this application, the open-source software library Keras was used due to its ability to create and utilize a Long Short-Term Memory neural network. The Long Short-Term Memory network is known for being able to remember information and predict sequences much more effectively than ordinary recurrent neural networks [5]. The music generation is created with the help of MIDI files. As MIDI files only store note data such as the length of the note, the pitch of the note, and the volume of the note, it is much easier for the neural network to use as data.

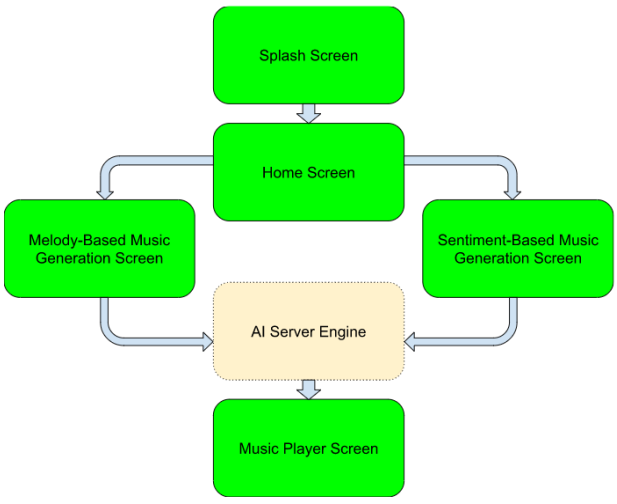

Figure 1. Overview of the solution

In the Flutter code, five primary screens are used to make up the interface of the application. The first screen is the splash screen, which only appears on the screen for a few seconds when first opening up the application and shows the logo of the application. The goal of this screen is to provide a smooth transition into the application for the user, and it was implemented primarily for aesthetic purposes. The splash screen was created with a Future that utilized a delay for a predetermined number of seconds before making a call to transition into the next screen, which is the home screen. With the home screen, there are two available options to choose from: melody-based music generation and sentiment-based music generation, and there are two buttons that lead to each page. In the melody-based generation screen, the user will be allowed to input an audio file by pressing a button, which will open up the mobile device's storage and show files with audio file extensions to choose from, such as MP3 and WAV [15]. When the file is chosen, a request is sent to the neural network code, then brings the application to the music player screen once the newly generated music has been completed. The other choice, which is sentiment-based generation, is done by prompting the user to type a message into a text box object added in Flutter. The Natural Language ToolKit in Python is used to extract the sentiment from the message and output a suitable piece of music based on the determined sentiment. The final primary screen is the music player screen, in which the player could listen to the generated music. A simple interface is created that allows the user to press a button to play it, and the user can see the progress of the song that is being played with a slider that automatically moves as the song plays and monitors the song's progress.

The Python code was responsible for generating the music [14]. Since Flutter and Python are two separate programming languages, a Flask server is used to connect the two languages through HTTP requests. Once the application requires that music be generated, the audio or message that was inputted by the user (depending on which method is used to generate the music) is sent through an HTTP request to the Python code, which will create and run a Long Short-Term Memory neural network using the data from the HTTP request. To create the neural network, a pre-trained model is fetched that was only trained using the piano; this is because using only one instrument to train seems to produce more accurate and desirable results. In the code to run the neural network, a certain number of notes are defined to be generated. Finally, the MIDI file is converted into a WAV file, and the WAV file is sent back to the application for the user to listen to from their mobile device.

4. EXPERIMENT

4.1. Experiment 1

An experiment to test the effectiveness of the application at generating music based on the sentiment that is provided. Twelve participants were gathered for the experiment, which should be a large enough sample size to account for any variability. Each participant was instructed to input ten different messages with sentiment, listen to the outputted audio, and record how many times they believed that the outputted music invokes feelings that are similar to the inputted message or not. As some people may have different interpretations of what is considered "accurate" in terms of sentiment, having multiple participants test this may help to counter any bias.

Participant Number	Number of Successful Runs
1	5
2	8
3	8
4	9
5	7
6	10
7	8
8	7
9	8
10	6
11	10
12	9
Average	7.917

Figure 2. Table of experiment 1

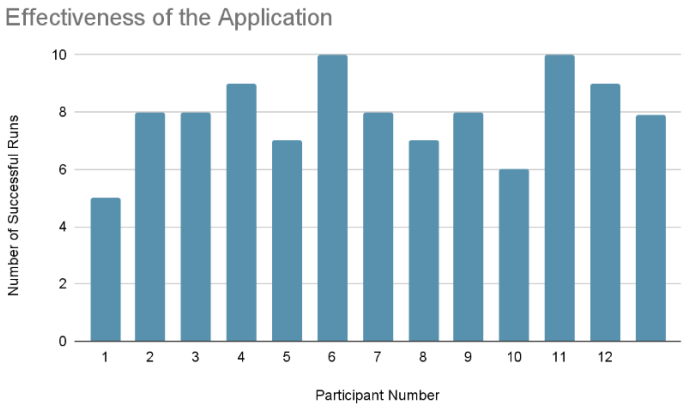

Figure 3. Effectiveness of the Application

From the results, it seems that the application had a decent success rate when it came to generating a piece of music that matched the sentiment in an inputted message. The highest score, which was 10, was achieved twice. On the other hand, the lowest score was 5. Overall, the average score was approximately 7.9. Out of every 10 pieces of music to be generated from the application, a little over 2 of them do not accurately represent the inputted sentiment. The majority of the participants who left feedback stated that they were impressed by how well the application worked. While the data seemed to fluctuate among the participants, this may be because their interpretation of what was considered "accurate" to the inputted sentiment may be more lenient or stricter depending on who it was that recorded the data. By gathering the average of all the participants' data, a balance could be found.

4.2. Experiment 2

A survey was conducted over Google Forms, and the twelve participants from the previous experiment took the survey. The first question asked whether the application is effective at creating new music based on the sentiment that was inputted. The second question asked how convenient and easy the application was to use. For both of the questions, the participants were provided a scale from one to ten to use as their answers. At the end of the Google Forms survey, a free-response survey is provided; if the participants have any feedback that they wish to express that cannot be expressed with the previous questions, they can do so.

Participant Number	Music Generation Effectiveness	Convenience/Ease of Use
1	6	7
2	8	8
3	9	9
4	8	9
5	7	10
6	10	10
7	8	8
8	7	7
9	7	8
10	6	8
11	9	9
12	8	9
Average	7.75	8.5

Figure 4. Table of experiment 2

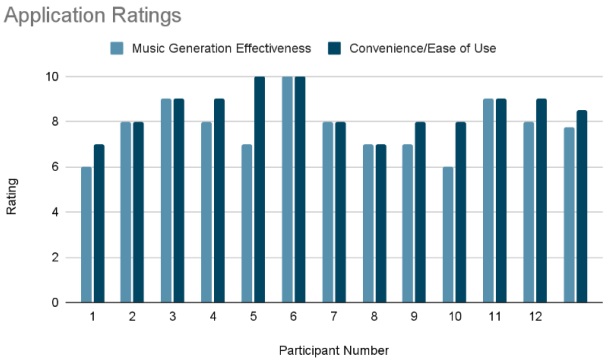

Figure 5. Application ratings

The table and chart indicate that the application is both competent at producing music to the users' preferences and intuitive to work with. The average rating of the music generation effectiveness was 7.75, while the average rating of the convenience of using the application was 8.5. While both of these ratings are high, the average rating for the application's ease of use is much higher. This may indicate that the participants generally felt that the effectiveness of using the application was weaker than the convenience and could use much more improvement. This idea was also reinforced in the optional feedback that was provided at the end of the survey, in which the participants generally felt like the ability of the application to generate music based on the user's messages in particular was the weakest point of the application. The data from the previous statement can further strengthen this notion, as it appears that participants who had fewer music pieces generated that successfully reflected the message sentiment tended to rate the effectiveness of the music generation lower.

According to the results of the experiment, the application does fairly well at covering the sentiment of the users' inputted messages in the outputted music, which is to be expected. The survey indicated that the intuitiveness and the effectiveness of the music generation were done well, but the intuitiveness had a significantly higher rating than the music generation's ability to accurately reflect the message's sentiment. This was an expected result, as the user interface had plenty of effort poured into it to ensure that new users of the application would have no trouble operating the application without any overly detailed instructions or external help. However, with the current AI techniques used in the application, it was difficult to get a very

accurate reflection of the message, and participants who were stricter on what was considered "accurate" may not have been as satisfied with the outputted music.

5. RELATED WORK

Louie et al. created a study regarding how music composers would be able to work with artificial intelligence tools in an efficient and organized manner by developing tools that can steer artificial intelligence to do as instructed by the user. For instance, some tools could focus a voice in the music to a specified range of notes. When tested, the tools appeared to have a strong positive effect on composers, as they provided them with a stronger sense of trust, control, and productivity [2]. The work from Louie et al. is similar to this work in that a tool is being developed to help with music creation, but the related work studies examples of how this has been done while this work focuses more on implementing the tool itself and experimenting on it.
In another related work written by Mantaras and Arcos from the Spanish National Research Council, computer music systems that involved artificial intelligence were analyzed based on how the composition and improvisation were handled, as well as how well the music could be performed. The results concluded that a useful technique that could be implemented in future computer music systems is case-based reasoning, as it can directly use information gained from audio samples performed by humans [3]. Both works deal with the topics of artificial intelligence, but Mantaras and Arcos emphasize the implementation of an AI music-generation concept as a whole while this work focuses on the effectiveness of a specific application.

Yang and Chen state that with more digital music available than ever, the organization of music is essential for easily accessing a target song. By organizing the emotions that music conveys, such as happiness or sadness, this problem may be remedied. Therefore, research has been done to tackle the issue of model training and visualization of emotion recognition [4]. Yang and Chen focus more on the various possible methods to go about implementing emotion recognition in music organization. Meanwhile, this work aims to create an application involving sentimental analysis for music generation and test its effectiveness to gauge how well it would perform if it was released to the public.

6. CONCLUSIONS

The current world situation involving a heightened need for quick and new music can be slightly improved through the creation of an application that is designed to generate new music based on whatever sentiment is inputted into the model. By typing a message into the application's text box, the application can determine whether the message's sentiment is positive, negative, or neutral. Then, the application will use this sentiment to output a piece of music for the user to listen to. To test whether this application would be applicable in the real world, an experiment was conducted in which twelve participants were gathered to type in a message into the application, listen to the music that was provided, and the participants also filled out a Google Forms survey that asked how well the music generation was as a whole and how convenient the application was to use. According to the results, the application seems to be fairly effective when it comes to generating music that accurately reflected the emotions of the inputted message [12]. The ratings of both the music generation and the intuitiveness of the application interface were somewhat high. However, because of how limited the current music generation method is in terms of determining only three types of sentiment, the participants took notice that there was still room for improvement with the music generation and gave it a significantly lower rating overall than the convenience of using the application [13]. More effort can be directed towards a music generation system that makes use of more parameters in the future.

One major limitation in the current version of the application is that there is no way for the user to directly download the music after generating it. While the application functions as intended and the user can hear the music from within the application, the user is unable to access that music outside the application easily without the use of external tools. As a result, users cannot share the music on social media or add them to videos without the use of a mobile screen/audio recorder. To improve the convenience of current and future users of the application, one or multiple features could be created to handle such an issue in future updates of the application.

One way to solve such an issue in the future is to create a download button on the same page where the generated music would be. Ideally, it would be located in an easy-to-notice spot with either an icon of the download symbol or the word “Download” on it. Once the user clicks the button, the user can store the audio file on their mobile device, which allows them to use it in other forms of social media easily.

REFERENCES

[1] Schneck, Daniel J., and Dorita S. Berger. The music effect: Music physiology and clinical applications. Jessica Kingsley Publishers, 2005.

[2] Louie, Ryan, et al. "Novice-AI music co-creation via AI-steering tools for deep generative models." Proceedings of the 2020 CHI conference on human factors in computing systems. 2020.

[3] De Mantaras, Ramon Lopez, and Josep Lluis Arcos. "AI and music: From composition to expressive performance." AI magazine 23.3 (2002): 43-43.

[4] Yang, Yi-Hsuan, and Homer H. Chen. "Machine recognition of music emotion: A review." ACM Transactions on Intelligent Systems and Technology (TIST) 3.3 (2012): 1-30.

[5] Nayebi, Aran, and Matt Vitelli. "Gruv: Algorithmic music generation using recurrent neural networks." Course CS224D: Deep Learning for Natural Language Processing (Stanford) (2015): 52.

[6] Wang, S., Li, J., Wang, H., & Zhu, Z. (2021). Emotional Music Generation with Attention-Based Bidirectional Gated Recurrent Unit. IEEE Access, 9, 12978-12986. doi: 10.1109/ACCESS.2021.3052969

[7] Moro, S., Sánchez-Monedero, J., Moreno-Sandoval, A., & Ricarte-Sabater, A. (2020). Music Emotion Recognition for User-Centered Systems: A Review of Challenges and Opportunities. Sensors, 20(6), 1687. doi: 10.3390/s20061687

[8] Liu, Z., Zeng, Q., & Zhang, B. (2020). Music Emotion Recognition with Machine Learning Methods: A Systematic Review. Journal of Intelligent & Fuzzy Systems, 38(3), 3091-3105. doi: 10.3233/JIFS-201483

[9] Liu, C., Lu, Z., Guo, Y., Hu, R., & Du, X. (2020). A Review of AI Music Generation. In 2020 International Conference on Computer Science and Artificial Intelligence (CSAI) (pp. 133-137). IEEE. doi: 10.1109/CSAI49736.2020.9162209

[10] Koops van 't Jagt, R., & Gomez-Marin, A. (2020). Generative models of music: Recent progress and future directions. Current Opinion in Behavioral Sciences, 32, 135-140. doi: 10.1016/j.cobeha.2019.12.002

[11] Zhang, W., & Wu, Y. (2020). Design and Implementation of a Mobile Social Platform Based on Flutter. In 2020 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) (pp. 599-603). IEEE. doi: 10.1109/ICMTMA49291.2020.00149

[12] Hossain, S., & Amin, M. R. (2020). Sentiment Analysis Based Music Generation: A Review. In 2020 23rd International Conference on Computer and Information Technology (ICCIT) (pp. 1-6). IEEE. https://doi.org/10.1109/ICCIT51738.2020.9342782

[13] Wang, Y., & Su, Y. (2021). Design and Implementation of an Application for Music Generation Based on Sentiment Analysis. In Proceedings of the 2021 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD 2021) (pp. 159-163). Atlantis Press. https://doi.org/10.2991/assehr.k.210611.021

[14] Zhang, W. (2021). Sentiment-based Music Generation using LSTM. GitHub Repository. https://github.com/willzhang05/Music-Generation-using-LSTM

[15] Liu, Qingzhong, Andrew H. Sung, and Mengyu Qiao. "Spectrum steganalysis of WAV audio streams." Machine Learning and Data Mining in Pattern Recognition: 6th International Conference, MLDM 2009, Leipzig, Germany, July 23-25, 2009. Proceedings 6. Springer Berlin Heidelberg, 2009.