MINISTRY OF EDUCATION AND RESEARCH OF THE REPUBLIC OF MOLDOVA

Technical University of Moldova

Faculty of Computers, Informatics, and Microelectronics

Department of Software Engineering and Automation

Approved for defense

Department head:

Ion FIODOROV, phd, associate professor

„___" _____ 2025

DEVELOPMENT OF GENERATIVE SONG PATTERNS. CREATING DATASET USING METADATA FROM MUSIC FILES

Master's project

Student: _____ **Mazur Vsevolod, IS-231M**

Coordinator: _____ **Mihai Gaidau, university assistant**

Consultant: _____ **Cojocaru Svetlana, university assistant**

Chisinau, 2025

ANNOTATION

For the diploma: "Development of generative song patterns. Creating dataset using metadata from music files",

developed by Vsevolod Mazur, Chisinau, 2024.

Keywords: audio, neural networks, data learning, metadata, frequencies, chords, dataset, spectrogram

The purpose of this work is to study the metadata of audio files, both identification and music, from the perspective of their use for training analytical and generative neural networks. Creation of a program for extracting audio metadata and generating a dataset based on them.

Tools used: Python programming language, VS Code, Music files

Explanatory note contains: introduction, 4 chapters, conclusions, bibliography with 18 titles, 27 figures, 1 table, 2 annexes.

Chapter 1: It contains an analysis of the research area, identifies possible issues and benefits of the existing solutions and defines potential solutions.

Chapter 2: Describes the principles of creating a system designed to obtain a set of data, fulfilling the set conditions and not violating the copyright of musical artists.

Chapter 3: Describes the algorithm of operation of the developed system and also demonstrates the results obtained.

Chapter 4: Describes all management aspects of project implementation from planning to product sales and profitability.

ADNOTARE

Pentru diploma: " Dezvoltare de modele inteligente generative pentru melodii. Crearea unui set de date folosind metadatele din fișiere muzicale ",

dezvoltat de Vsevolod Mazur, Chișinău, 2024.

Cuvinte-cheie: audio, rețele neuronale, data learning, metadate, frecvenţă, acorduri, set de date, spectrogramă

Scopul acestei lucrări este de a studia metadatele fișierelor audio, atât de identificare, cât și muzicale, din perspectiva utilizării acestora pentru antrenarea rețelelor neuronale analitice și generative. Crearea unui program pentru extragerea metadatelor audio și generarea unui set de date pe baza acestora.

Instrumente utilizate: Limbajul de programare Python, VS Code, fișiere muzicale

Nota explicativă conține: introducere, 4 capitole, concluzii, bibliografi cu 18 titluri, 27 figure, 1 tabel, 2 apendixuri.

Capitolul 1: Conține o analiză a domeniului de cercetare, identifică posibilele probleme și beneficii ale soluțiilor existente și definește soluții potențiale.

Capitolul 2: Descrie principiile creării unui sistem menit să obțină un set de date, îndeplinind condițiile stabilite și fără încălcarea drepturilor de autor ale artiștilor muzicali.

Capitolul 3: Descrie algoritmul de funcționare al sistemului dezvoltat și, de asemenea, demonstrează rezultatele obținute.

Capitolul 4: Descrie toate aspectele manageriale ale implementării unui proiect, de la planificare până la vânzarea produsului și obținerea profitului.

CONTENTS

LIST OF ABBRIVIATIONS AND DEFINITIONS

1. Bach chorales **-** vocal music for a choir consisting of four voices: soprano, alto, tenor, bass (SATB).
2. BPM – Beat per minute, in musical terminology, tempo - which is measured in bits per minute.
3. Musical scale - the tonal basis of music. It is a set of tones from which can build melodies and harmonies. The tones of a scale are ordered according to their pitch.
4. Arrangement - musical adaptation of an existing composition, adding differences from the original composition.
5. Guitar TABs - is an easy-to-read notation system that allows guitarists to learn how to play practically anything by themselves.
6. DAC - digital-to-analog converter is a system that converts a digital signal into an analog signal.
7. Daw – is a program that can be used to turn a regular computer into a sound recording center. This software provides all the processes associated with creating music. It can be used for various tasks: sound recording, mixing, mastering, etc.
8. VST (Virtual Studio Technology) – a format for runtime-dependent (native) real-time plugins that plug into audio editors.
9. BGM – (Background music) is a combination of musical and sound accompaniments that is actively used to create an atmosphere against the background of other activities.

INTRODUCTION

Currently, there is a wide distribution of applications built on neural networks throughout the world. In turn, such applications have begun to be actively used both by us, in everyday life, and by developers, artists, scriptwriters, and many other professions.

In new smartphones, the main innovations concern "intelligent assistants" built on the basis of neural networks and helping in everyday life. If earlier each subsequent update of the phone model was accompanied by significant hardware improvements, but today, for example, almost all improvements in the quality of photos on a smartphone are associated with improved artificial intelligence technologies aimed at post-processing the received photos by changing the color parameters, as well as improving the quality of the video shot by drawing additional frames, thereby increasing the number of frames per second, thereby improving the softness of the picture.

Modern video cards work on a similar technique. Currently, the processor market is trending towards reducing the process technology, i.e. reducing the size of resistors and increasing their number per unit area. Thus, processors are now manufactured using a 3 nm process technology, and research is underway to introduce a 2 nm process technology, therefore, to increase productivity during this period of time, both in the CPU and GPU markets, most of the increase in productivity directly depends on AI cores, due to which productivity with the same process technology can be higher. Modern real-time visualization technologies render images using such AI cores, thereby increasing the number of displayed frames, which in turn adds smoothness to the picture in games and movies, also due to the additional rendering of intermediate frames.

Quite often, they began to use image generation to create concepts or find inspiration. Or use text assistants to speed up systematic work that does not require human creativity. Artificial intelligence is actively used in cars as cruise control, allowing cars to move in traffic, adhering to markings, braking in traffic jams based on the braking of cars in front, and also analyzing speed limits and traffic lights around, using cameras, radars and lidars. All these technologies allow increasing the user's productivity, or correcting his actions, assisting him. Thus, actively participating in a person's life.

The implementation of artificial intelligence technologies and generative neural networks is a modern method of optimizing most of the technical processes in companies, thereby allowing to save development budgets by saving the budget requested for outsource, as well as adding chats based on gpt for a chat consultant who can direct the user to the desired page taking into account their requests. Although this technology is currently considered new and fashionable, it has the potential that some often overestimate.

For the most part, "creative" neural Models are used by people and companies that are not carriers of the analogue profession. For example, when Stable diffusion became publicly available, “non-artists” began actively polluting various platforms for artists, including Pinterest, DeviantArt, ArtStation and

others, with generated images created using neural networks, which caused a very negative reaction among people working or actively engaged in artistic creativity, not because they "take away their work" but because often in such works there were mistakes that a "non-artist" would not see, from composition to anatomy, and those who use such platforms for improving their own skill, starting to use generated works as a reference, they get bad habits, which slows down their improvement. Or, for example, there were cases when companies entrusted customer service to a neural network and suffered losses.

It should be understood that this technology is not a replacement for current professions, but only a convenient tool for work that should be implemented in the work cycle, but not replace the worker himself. Understanding this, can not only speed up the work but also raise the level of the employee, due to his memorization of different solutions, which the neural model demonstrates to him, which will increase his knowledge base and teach him to use them. The same applies not only to work, but also to training, hobbies. With the help of neural networks, can learn languages and engage in those types of hobbies that seemed distant due to the complexity and lack of understanding of how everything works. Such types of hobbies include creating music.

BIBLIOGRAPHY

1. Ahona Rudra. “Cybersecurity In The Digital Learning Environment,” August 17, 2023. https://powerdmarc.com/cybersecurity-in-digital-learning-environment/.
2. Burgess, Richard James, and Richard James Burgess. *The Art of Music Production: The Theory and Practice*. Fourth Edition, Fourth Edition. Oxford, New York: Oxford University Press, 2013.
3. YouTube. *Composing Dark Star Thresh's Music | Behind the Scenes - League of Legends*, 2016. https://www.youtube.com/watch?v=1h9c-O3AEkE
4. AlexMatkovicMusic. “Микширование, мастеринг и производство музыки и аудио профессионального качества.” Accessed October 3, 2024. https://www.etsy.com/listing/1116907281/professional-quality-musicaudio-mixing?utm_source=OpenGraph&utm_medium=PageTools&utm_campaign=Share.
5. Mubert Inc. “Mubert.” Accessed October 5, 2024. https://mubert.com/render.
6. Hayk Martiros. “Riffusion/Riffusion-Hobby.” Python. 2022. Reprint, Riffusion, December 24, 2024. https://github.com/riffusion/riffusion-hobby.
7. “Spectrogram.” In *Wikipedia*, December 9, 2024. https://en.wikipedia.org/w/index.php?title=Spectrogram&oldid=1262037486.
8. Young-Jun Kim and Seok-Pil Lee. “A Generation of Enhanced Data by Variational Autoencoders and Diffusion Modeling.” *ResearchGate*, October 12, 2024. https://doi.org/10.3390/electronics13071314.
9. YouTube. Как Сделать Переход Между Треками ПОДРОБНО, 2022. https://www.youtube.com/watch?v=9m1TUG_7q6E.
10. “Период действия авторских прав.” In *Википедия*, December 30, 2019. https://ru.wikipedia.org/w/index.php?title=%D0%9F%D0%B5%D1%80%D0%B8%D0%BE%D0%B4_%D0%B4%D0%B5%D0%B9%D1%81%D1%82%D0%B2%D0%B8%D1%8F_%D0%B0%D0%B2%D1%82%D0%BE%D1%80%D1%81%D0%BA%D0%B8%D1%85_%D0%BF%D1%80%D0%B0%D0%B2&oldid=104247604.
11. Cano, Estefania, Derry FitzGerald, Antoine Liutkus, Mark D. Plumbley, and Fabian-Robert Stöter. “Musical Source Separation: An Introduction.” *IEEE Signal Processing Magazine* 36, no. 1 (January 2019): 31–40. https://doi.org/10.1109/MSP.2018.2874719.
12. “List of Copyright Duration by Country.” In *Wikipedia*, October 30, 2024. https://en.wikipedia.org/w/index.php?title=List_of_copyright_duration_by_country&oldid=1265058860.
13. Joe Wreschnig, Michael Urman, Lukáš Lalinský, Christoph Reiter, and Ben Ockmore. “Mutagen.” Accessed November 2, 2024. https://mutagen.readthedocs.io/.
14. McFee, Brian, Matt McVicar, Daniel Faronbi, Iran Roman, Matan Gover, Stefan Balke, Scott Seyfarth, et al. “Librosa/Librosa: 0.10.2.Post1.” Zenodo, May 14, 2024. https://doi.org/10.5281/zenodo.11192913.
15. Blendist LLC. “Find the BPM and Key for Any Song | Every Song, Every Tempo.” SongBPM. Accessed November 6, 2024. https://songbpm.com.

16. Bigo, Louis, Daniele Ghisi, Antoine Spicher, and Moreno Andreatta. “Representation of Musical Structures and Processes in Simplicial Chord Spaces.” *Computer Music Journal* 39, no. 3 (September 1, 2015): 9–24. https://doi.org/10.1162/COMJ_a_00312.

17. Imai, Yusuke, Sid C. Dellby, and Nobuaki Tanaka. “General Theory of Music by Icosahedron 1: A Bridge between ‘Artificial’ Scales and ‘Natural’ Scales, Duality between Chromatic Scale and Pythagorean Chain, and Golden Major Minor Self-Duality.” arXiv, August 6, 2022. https://doi.org/10.48550/arXiv.2103.10272.

18. Кристина Спиридонова. “8 лучших моделей разработки программного обеспечения.”Purrweb. *kristina-spiridonova-ru* (blog), October 27, 2023. https://www.purrweb.com/ru/blog/metodologii-dlya-razrabotki-po/.