

KI in der Musikproduktion

Wolfgang Rup
Matrikelnummer: 35053
Ton-Seminar
Sommersemester 2020

Inhaltsverzeichnis

1. Funktionsprinzipien in der Musikproduktion angewandter KI
 - 1.1 Markov-Ketten
 - 1.2 Neural Networks/Neuronale Netze
2. Beispiele für angewandte Systeme
 - 2.1 Flow Machines
 - 2.2 AIVA
 - 2.3 Melodrive
 - 2.4 Magenta
 - 2.5 Dadabots
 - 2.6 Amper
3. Fazit und Ausblick
4. Literaturverzeichnis
5. Abbildungsverzeichnis

1. Funktionsprinzipien in der Musikproduktion angewandter KI

Um die Arbeitsweise der in der Musikproduktion angewandten Künstlichen Intelligenzen zu verstehen, ist es sinnvoll, sich zunächst mit der Definition von KI auseinanderzusetzen. Eine KI steht unter der Prämisse, menschliches intelligentes Verhalten imitieren zu können. Darunter versteht man man gemeinhin die Fähigkeit, sich auf neue Umstände einstellen und Probleme lösen zu können, sowie Lernfähigkeit. Eine KI durchläuft in der Regel folgende Arbeitsschritte:

- Lernen: Der KI wird Information eingespeist, anhand derer sie Zusammenhänge analysiert und Muster erkennt.
- Schlussfolgerung: Anhand den aus dem Lernprozess hervorgegangenen Erkenntnissen entwickelt die KI Regeln und trifft Annahmen für folgende Ereignisse.
- Selbstkorrektur: Der Fehler in der Genauigkeit der Annahmen wird bestimmt und bei weiteren Berechnungen berücksichtigt. ^[1]

Die größten Herausforderungen beim Realisieren einer KI unterscheiden sich grundlegend von den Problemen, mit denen ein Mensch kämpfen mag. Eine künstliche Intelligenz kann beispielsweise komplexe mathematische Berechnungen problemlos erledigen, während Aufgaben, die ein Mensch intuitiv bestreiten würde, weitaus schwieriger sind, wie z.B. Gesichtserkennung, Handschrift- und Spracherkennung, oder eben auch die Komposition von „ansprechend“ klingender Musik.

1.1 Markov-Ketten

Eine Technik, die vor allem bei früher KI-Musik Verwendung fand, sind die sogenannten Markov-Ketten. Dabei handelt es sich um ein mathematisches Konzept, das zu Beginn des 20. Jahrhunderts vom russischen Mathematiker Andrey Markov entworfen wurde. Eine Markov-Kette ist eine Folge von stochastischen Ereignissen, bei der jeweils die Wahrscheinlichkeit bestimmt wird, dass auf ein eingetretenes Ereignis ein bestimmtes weiteres Ereignis folgt. Bei einer Markov-Kette erster Ordnung bezieht sich die Wahrscheinlichkeit des nächsten Ereignisses immer nur auf das vorherige Ereignis, es gibt aber auch Markov-Ketten höherer Ordnung. So bezieht sich bei einer Markov-Kette n-ter Ordnung die Probabilität des folgenden Ereignisses auf n vorhergehende Ereignisse. ^[2]

Markov wandte diesen Prozess zur Veranschaulichung u.a. auf Gedichte an. Hierbei wurden die Wahrscheinlichkeiten, dass ein Vokal auftritt, und dass ein Vokal auf einen Konsonanten oder auf einen anderen Vokal folgt, bestimmt. Aus diesen Daten entstanden nun neue Gedichte. ^[2]

Ein weiteres Beispiel, wie man ein Markov-Modell visualisieren kann, ist die Wettervorhersage anhand der Wahrscheinlichkeiten aufeinanderfolgender Wetterereignisse (Abb. 1).

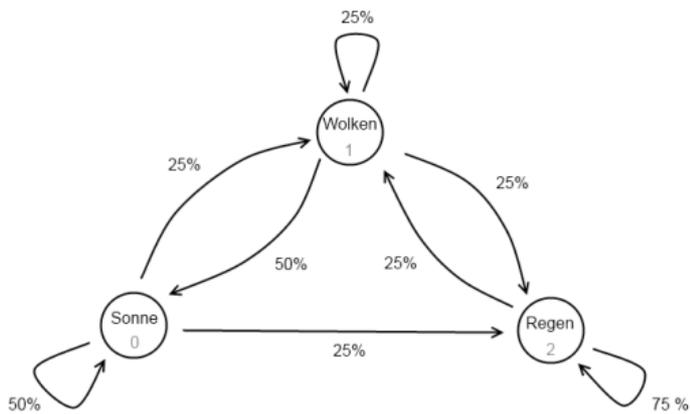


Abbildung 1: Wettervorhersage mittels Markov-Ketten

Bei der Anwendung in der Musikproduktion geht man ähnlich vor. Zunächst wird als Datensatz ein melodischer Korpus erstellt, der aus einigen, mitunter auch sehr vielen Musikstücken besteht. Dann wird der melodische Korpus analysiert in Hinsicht der Wahrscheinlichkeiten von Note zu Note oder Akkord zu Akkord. Anhand dessen erfolgt nun eine neue Aneinanderreihung der Noten/Akkorde, wodurch ein neues Musikstück entsteht. ^[3]

Dabei gilt: Ist eine Sequenz der Länge n nicht im Korpus enthalten, wird sie nicht in einer Kreation n -ter Ordnung auftreten. Umgekehrt bedeutet dies, dass Sequenzen der Länge n übernommen werden können, was bei höheren Ordnungen zu einem Problem werden kann. Sollte der Fall eintreten, dass mehrere Sequenzen der Länge n aus dem Korpus aufeinanderfolgend reproduziert werden, kommt man einem Plagiat sehr nahe. Ein Lösungsansatz ist die Funktion MAXORDER, die die Länge von in Sequenzen enthaltenen Datenblöcken auf ein festgelegtes Maximum beschränkt. ^[4]

1.2 Neural Networks/Neuronale Netze

(auch artificial neural networks, deep learning)

Neural Networks, im Deutschen auch als Künstliche Neuronale Netze bezeichnet, sind die am weitesten verbreitete Technologie von Künstlichen Intelligenzen zur Kreation von Musik. Dabei handelt es sich um Systeme, die als digitale Nachbildung des menschlichen Nervensystems operieren sollen. Um dessen komplexe Aufgaben nachahmen zu können, wird es mit Information in Form von Beispielen „gefüttert“. Diese werden vom Neural Network (NN) analysiert und zur Entwicklung von Regeln verwendet. Dabei steigt die Genauigkeit des Ergebnisses, also wie nah der Output am gewünschten Resultat liegt, mit der Anzahl der Beispiele. Hier fällt auch häufig der Begriff „deep learning“, der auf eine vielfach („tief“) verschachtelte Struktur aus Regeln verweist. ^[5]

Das Prinzip eines Neural Networks kann vereinfacht betrachtet mit dem mathematischen Modell der linearen Regression verglichen werden (vgl. Abb. 2). Zunächst werden Startwerte angegeben, diese können zufällig oder bestimmt gewählt sein. Dann werden Werte aus einem Datensatz eingespeist, aus denen durch Interpolation Funktionswerte ermittelt werden, was schließlich eine vollständige mathematische Funktion ergibt. ^[6] Nach einem Durchlauf wird der Fehler zwischen Funktion und

erwartetem/gewünschtem Ergebnis bestimmt, hierbei können auch bestimmte Werte höher gewichtet werden. ^[7] Das System wird sodann fortlaufend neue Funktionswerte suchen, bis der Fehler minimal wird. ^[6]

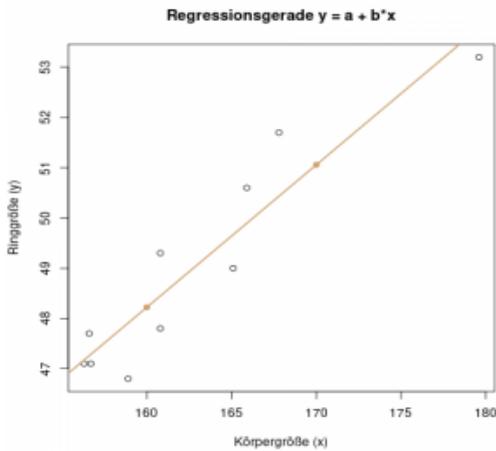


Abbildung 2: Lineare Regression

Vom weiten Feld der NNs werden in der Musikproduktion üblicherweise RNNs (Recurrent Neural Networks) benutzt. Dabei handelt es sich um ein Neural Network, welches nach dem Durchlaufen auf sich selbst zurückführen kann und so mehrmalige Durchläufe einer Instanz vornehmen. Das hat den Effekt eines Gedächtnisses, sodass bereits gelernte Regeln nicht wieder vergessen werden, sondern später wieder „in Erinnerung gerufen“ und erneut angewandt werden können. ^[8]

Eine relevante Weiterentwicklung von RNNs sind LSTMs (Long Short Term Memory Networks). Hier wird der Erinnerungsaspekt von RNNs auf Langzeit-Abhängigkeiten erweitert. Dies wird durch eine sog. Erinnerungszelle realisiert; durch „Gates“ entscheidet das System, welche Daten in die Erinnerungszelle hinein kommen, wie lange sie dort bleiben und welche davon verwendet werden. Der Vorteil, der durch LSTMs in der Musikproduktion erzielt wird, ist die Fähigkeit, Strukturen über ein Musikstück hinweg beizubehalten und wiederkehrende Elemente einzubauen, was einen wichtigen Bestandteil in vielen Genres darstellt. ^{[9][10]}

2. Beispiele für angewandte Systeme

Schon lange wird versucht, mit Hilfe von Computern überzeugende Musik zu erzeugen, in den letzten Jahren allerdings gab es, durch die Fortschritte in der KI-Technologie, einen geradezu exponentiellen Anstieg an Systemen und Werkzeugen zur Komposition mit künstlicher Intelligenz. Im Folgenden sollen einige davon, mit ihren spezifischen Funktionsweisen und Anwendungsgebieten, vorgestellt werden.

2.1 Flow Machines

Flow Machines wurde 2012 vom französischen Musiker und Wissenschaftler François Pachet im Dienst von Sony-CSL entwickelt. Dessen Flow Composer basiert auf Markov-Ketten höherer Ordnung, die um die Methode „meter“ erweitert wurden. Damit ist es möglich, die Zeit als Faktor in Markov-Modelle zur

bringen, was mehrere Verbesserungen mit sich brachte: Zum einen ermöglichte dies, die Dauer von Noten zu beeinflussen; zum anderen war es zuvor nicht möglich, Markov-Ketten mit festgelegter Länge zu generieren, nun konnte man aber auf Wunsch Musikstücke von bestimmter Länge erzeugen. Flow Machines analysierte ca. 13000 Stücke diverser Künstler als Basis, je nach gewolltem Stil werden andere Stücke in den melodischen Korpus aufgenommen. Der Output erfolgt in Lead Sheets, die das volle Arrangement einem Künstler überlassen. Darüber hinaus ist die Harmonisierung bestehender Melodien in einem frei wählbaren Stil möglich. ^{[11][12]}

2.2 AIVA

AIVA (Artificial Intelligence Virtual Artist) ist ein 2016 gegründetes Startup-Unternehmen. Der AIVA-Composer basiert auf einer Variation von RNNs/LSTMs, wie auch alle Beispiele, die im Folgenden behandelt werden. Zudem arbeitet AIVA mit Reinforcement Learning, was bedeutet, dass eine Person der KI positives Feedback für gute Ergebnisse gibt, damit diese Aspekte weiterverfolgt werden. Darüber hinaus werden sog. Genetic Algorithms verwendet, eine Optimierungsmethode, mit der sich das System evolutionsartig kontinuierlich weiterentwickeln soll. AIVA liefert MIDI-Daten als Ausgabe, der Verwendungszweck liegt im Bereich Filme und Games. Es sind verschiedene Stile implementiert, der Fokus liegt auf Klassischer Musik. ^{[13][14]}

2.3 Melodrive

Ebenfalls als Startup gegründet wurde Melodrive 2018 in Berlin. Das Unternehmen hat seinen Fokus auf Games gelegt, es bietet auch Unterstützung für die Game-Engine Unity an. Das Besondere an Melodrive ist, dass es Musik in Echtzeit generieren kann. ^[15] Dies wurde unter anderem auch in einem 24 Stunden Livestream auf Twitch demonstriert, bei dem die Zuschauer durch Schlagwörter im Chat die Musik live beeinflussen konnten, indem ein Programm den Wörtern bestimmte Emotionen zuordnete, die dann in der erzeugten Musik wiederspiegelt wurden. ^[16] Das Tool erfreute sich bei Indie-Entwicklern großer Popularität, auch weil eine kostenlose Version angeboten wurde, große Unternehmen blieben allerdings skeptisch in Anbetracht ihrer Qualitätsansprüche. ^[17]

2.4 Magenta

Unter dem Namen Magenta operieren verschiedene Forschungsprojekte von Google, seit 2016 sind einige vom Magenta-Team entworfene Tools als open source verfügbar. Diese arbeiten teils mit MIDI, teils mit rohen Audiodateien. Das Programm Magenta Studio soll als kreativer Helfer fungieren, indem es aus einer gegebenen Notensequenz eine Musiksequenz oder ein Stück kreiert, oder sogar die Ausgangssequenz anhand der Magenta Library zufallsgeneriert. ^[18] Die prominentesten Ergebnisse erzielte Studio bei der Analyse von Drum Tracks, aus der es eine abwechslungsreiche Methode zum „Humanisieren“ von Beats

entwickelte. ^[19] Ebenfalls zu Magenta gehört der Synthesizer Nsynth, der KI benutzt, um nach persönlichen Kriterien neue Sounds zu erzeugen, im Gegensatz zum reinen Layering bei vergleichbaren Alternativen. ^[20]

2.5 Dadabots

In den USA wurde 2012 mit der Arbeit an Dadabots begonnen, einem nichtkommerziellen, experimentellen Projekt mit KI-Musik. Der Algorithmus von Dadabots arbeitet im Gegensatz zu den meisten anderen RNN-basierten Systemen nicht mit MIDI oder Noten, sondern mit rohen Audiodateien. Dafür wird ein sog. SampleRNN-Algorithmus benutzt, der ursprünglich für text-to-speech Anwendungen erdacht wurde. Dieser tastet eingespeiste Musik (meist Metal-Alben, da diese ein gewisses Maß an Chaos tolerieren können) Sample für Sample ab und setzt sie daraus neu zusammen. Da mit dem Projekt kein konkretes Ziel verfolgt wird, lässt man das RNN „unsupervised learning“ betreiben, d.h. es wird kein Einfluss zur Erzielung etwaiger gewünschter Resultate genommen. Der einzigartige Ansatz, mit ganzen Audiodateien zu arbeiten, birgt jedoch auch den Nachteil, dass eine immense Rechenleistung vonnöten ist, die von der Hardware geschultert werden muss. ^{[21][22]}

2.6 Amper

Ebenfalls aus den USA stammt das Startup Amper, das 2014 gegründet wurde. Amper greift auf eine Datenbank aus über einer Million Samples zurück, die mit realen Instrumenten aufgenommen wurden. Überdies wurden keine bestehenden Musikstücke eingelesen, vielmehr sind Datensätze implementiert, die musiktheoretische Grundsätze zur Komposition anwenden. Dadurch stehen viele Parameter zur Verfügung, die manipuliert werden können, wie Stil, Stimmung oder Dauer, aber auch komplexere Attribute wie z.B. Dynamik oder Anschlagsstärke der Noten. Der Output erfolgt als Audio, welches im Ganzen oder als Stems exportiert werden kann; das Zielformat ist hauptsächlich (Hintergrund-)Musik für Videos und Games. ^{[23][24]}

3. Fazit und Ausblick

Künstliche Intelligenz hat im letzten Jahrzehnt in der Musikproduktion deutlich an Bedeutung gewonnen. Die Eigenkreationen von KIs sind überzeugender geworden, und finden bereits Einsatz z.B. als unkomplizierte Hintergrundmusik in Werbung und Videos. Vor allem aber hat sich KI als nützliches Instrument für Musiker herausgestellt, mit dem Inspirationen gewonnen oder Schreibblockaden umgangen werden können. Alles in allem ist das Feld der künstlichen Intelligenz in der Musikproduktion noch lange nicht weitläufig begangen, und ihre Zukunft ist vielversprechend.

4. Literaturverzeichnis

1. Rouse, Margaret. (2018). „Definition Künstliche Intelligenz (KI)“.
<https://www.computerweekly.com/de/definition/Kuenstliche-Intelligenz-KI>. Zugriff: 31.08.2020.
2. Nierhaus, Gerhard. (2009). „Algorithmic Composition“. Springer-Verlag Wien.
3. Cruz, Jeffrey. (2019). „Deep Learning vs Markov Model in Music Generation“. Honors College Theses. 221.https://digitalcommons.pace.edu/honorscollege_theses/221
4. Papadopoulos, Alexandre & Roy, Pierre. (2014). „Avoiding Plagiarism in Markov Sequence Generation“.
5. Goodfellow, Ian & Bengio, Yoshua & Courville, Aaron. (2016). „Deep Learning“.
6. Briot, Jean-Pierre & Hadjeres, Gaëtan & Pachet, Francois. (2017). „Deep Learning Techniques for Music Generation - A Survey“.
7. DeepAI. „What is Weight (Artificial Neural Network)?“. <https://deepai.org/machine-learning-glossary-and-terms/weight-artificial-neural-network>. Zugriff: 10.05.2020.
8. Abolafia, Dan. (2016). „A Recurrent Neural Network Music Generation Tutorial“.
<https://magenta.tensorflow.org/2016/06/10/recurrent-neural-network-generation-tutorial>.
9. Olah, Christopher. (2015). „Understanding LSTM Networks“. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
10. Knabbe, Marina. (2017). „Erzeugung von Musiksequenzen mit LSTM-Netzwerken“.
https://edoc.sub.uni-hamburg.de/haw/volltexte/2017/3964/pdf/BA_1971279.pdf
11. Sony CSL. „Flow Machines: AI Assisted Music“. <https://www.flow-machines.com/>. Zugriff: 10.05.2020
12. Papadopoulos, Alexandre & Roy, Pierre & Pachet, Francois. (2016). „Assisted Lead Sheet Composition Using FlowComposer“. 9892. 10.1007/978-3-319-44953-1_48.
13. AIVA. „The Artificial Intelligence composing emotional soundtrack music“. <https://aiva.ai/>. Zugriff: 10.05.2020
14. Aiva Technologies. (2018). <https://medium.com/@aivatech/how-we-used-our-music-engine-to-create-the-first-ai-generated-album-of-chinese-music-9d6fa984b4e8>
15. Roth, Marcel. (2020). „Künstliche Intelligenz: Musik von Robo-Beethoven“.
<https://www.mdr.de/nachrichten/panorama/erfindungen/kuenstliche-intelligenz-musik-komponieren-100.html>
16. Dachs, Paul. (2019). „Twitch plays Melodrive – Using comments to play AI-generated music“.
<https://medium.com/the-sound-of-ai/twitch-plays-melodrive-using-comments-to-play-ai-generated-music-fde92497d414>
17. Feed Magazine. „Start-up: Melodrive, Germany, 2017“. <https://feedmagazine.tv/start-up-alley/start-up-melodrive-germany-2017/>. Zugriff: 10.05.2020.

18. Eck, Douglas. (2016). „Welcome to Magenta!“. <https://magenta.tensorflow.org/blog/2016/06/01/welcome-to-magenta/>
19. Synced. (2018). „Google AI Music Project Magenta Drops Beats Like Humans“. <https://medium.com/syncedreview/google-ai-music-project-magenta-drops-beats-like-humans-515de6e5f621>
20. Google AI. „What is NSynth Super?“. <https://nsynthsuper.withgoogle.com/>. Zugriff: 10.05.2020.
21. Kastrup, Dennis. (2019). „Wenn künstliche Intelligenz komponiert“. https://www.deutschlandfunk.de/musik-von-dadabots-wenn-kuenstliche-intelligenz-komponiert.807.de.html?dram:article_id=449591
22. Carr, CJ & Zukowski, Zack. (2018). „Generating Albums with SampleRNN to Imitate Metal, Rock, and Punk Bands“. <https://arxiv.org/pdf/1811.06633.pdf>
23. Amper music. „FAQ“. <https://www.ampermusic.com/faq/>. Zugriff: 10.05.2020.
24. Hogan, Mark. (2019). „Artificial Intelligence: Music's Next Frontier... The Drew Silverstein Interview“. <https://www.hotpress.com/music/artificial-intelligence-musics-next-frontier-drew-silverstein-interview-22786307>

5. Abbildungsverzeichnis

1. <http://www.inf.fu-berlin.de/lehre/SS11/SemAlg/eckstein-Dateien/wetter.png>. Zugriff: 10.05.2020.
2. <https://www.crashkurs-statistik.de/wp-content/uploads/2017/01/regression-gerade-300x300.png>. Zugriff: 10.05.2020.