



Bachelorarbeit

im Studiengang Audiovisuelle Medien

mit dem Titel

Einsatz künstlicher Intelligenz in der Musikproduktion – technische Anwendungen, Grenzen
und Analyse am Beispiel einer Songproduktion

vorgelegt von

Peter Volkert
Matrikelnummer 43712

an der Hochschule der Medien Stuttgart am 18.05.2026

zur Erlangung des akademischen Grades eines Bachelor of Engineering (B. Eng.)

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: Prof. Dr. Michael Felten

Kurzfassung

Diese Bachelorarbeit untersucht den Einsatz künstlicher Intelligenz innerhalb der Musikproduktion am Beispiel einer KI-gestützten Songproduktion mit dem System Suno. Ziel der Arbeit ist es, die technischen Möglichkeiten, kreativen Potenziale sowie die Grenzen dieses KI-Systems im Produktionsprozess und anhand eines wissenschaftlichen Tests zu analysieren.

Zu Beginn werden grundlegende Konzepte künstlicher Intelligenz erläutert, darunter maschinelles Lernen, Deep Learning, künstliche neuronale Netze sowie Diffusionsmodelle. Aufbauend darauf erfolgt eine Einordnung generativer KI-Systeme innerhalb der Musikindustrie mit besonderem Fokus auf das Audio-KI-System *Suno*.

Der praktische Schwerpunkt der Arbeit besteht aus einem ausführlichen Produktionstagebuch sowie einem systematischen Testverfahren. Innerhalb des Produktionstagebuchs wird dokumentiert, wie die KI zur Unterstützung verschiedener Produktionsschritte eingesetzt wurde. Am Ende jedes Produktionsschritts werden die wichtigsten Erkenntnisse zusammengefasst. Dabei zeigte sich, dass *Suno* insbesondere kreative Prozesse beschleunigen kann und bei der Erzeugung von Instrumenten wie Saxophon oder mehrstimmigen Streichern gute Ergebnisse liefert. Gleichzeitig traten jedoch typische Schwächen wie Artefakte, eingeschränkte Kontrolle über Details sowie eine mangelnde Reproduzierbarkeit der Ergebnisse auf.

In einem systematischen Test werden die Parameter *Weirdness*, *Style Influence* und *Audio Influence* von *Suno* gezielt untersucht und anhand verschiedener Haupt- und Nebeneigenschaften wie *Nutzbarkeit*, oder *Nähe zur Vorlage* ausgewertet. Die Ergebnisse zeigen, dass insbesondere die Parameter *Audio Influence* und *Weirdness* die Nähe zur Vorlage und Qualität deutlich beeinflussen. Außerdem wurde festgestellt, dass *Suno* häufig wiederkehrende Fehler reproduziert, indem beispielsweise kontinuierlich ein Schlagzeug ergänzt wird, das so in der ursprünglichen Vorlage nicht enthalten war.

Die Arbeit kommt schließlich zu dem Ergebnis, dass KI-Systeme wie *Suno* eine relevante technische und kreative Unterstützung innerhalb der Musikproduktion darstellen. Gleichzeitig ersetzt die KI den menschlichen Produzenten jedoch nicht vollständig, da kreative Entscheidungen, Qualitätskontrolle und insbesondere das Entfernen von KI-Artefakten in der Postproduktion weiterhin stark von der menschlichen Komponente abhängen. Die KI übernimmt somit eher die Rolle eines kreativen Partners als die eines vollständig autonomen Musikproduzenten.

Abstract

This bachelor's thesis investigated the application of artificial intelligence within music production, using the example of an AI-assisted song production with help of the Suno system. The objective of this study is to analyze the technical capabilities, creative potential, and limitations of this AI system within the production process and within a scientific test.

The thesis begins by elucidating fundamental concepts of artificial intelligence, including machine learning, deep learning, artificial neural networks, and diffusion models. Building upon this foundation, generative AI systems are contextualized within the music industry, with a particular focus on the audio AI system Suno.

The practical core of the thesis is a detailed production diary and a systematic testing procedure. The production diary documents how AI was used to support various stages of the production process. Key findings are summarized at the end of each production step. This process revealed that Suno is very effective at supporting creative workflows and gets high-quality results when generating instruments like a saxophone or multi-part string arrangements. Typical limitations were audio artifacts, limited granular control over details, and a lack of result reproducibility.

In a systematic test the Suno parameters *Weirdness*, *Style Influence*, and *Audio Influence* were examined and evaluated with various primary and secondary criteria. The results show that the parameters Audio Influence and Weirdness significantly influence the fidelity to the source material and the quality. Also, it was observed that Suno frequently makes the same mistakes. For example, it consistently adds a drum track that was not present in the original source material.

This thesis concludes that AI systems such as Suno are a source of technical and creative support within the field of music production. At the same time AI does not entirely replace the human producer, for creative and quality control. The removal of AI-generated artifacts during post-production needs human expertise. So, AI has more the role of a creative partner rather than a fully autonomous music producer.

Inhaltsverzeichnis

I.	Abkürzungsverzeichnis.....	I
II.	Glossar.....	II
III.	Anmerkungen.....	IV
IV.	Ehrenwörtliche Erklärung.....	V
1	Einleitung	
1.1	Problemstellung.....	S.1
1.2	Ziele der Arbeit.....	S.3
2	Die Grundlagen von Künstlicher Intelligenz	
2.1	Definition KI.....	S.4
2.2	Wie funktioniert künstliche Intelligenz.....	S.6
2.2.1	Maschinelles Lernen.....	S.6
2.2.2	Deep Learning.....	S.7
2.2.3	Künstliche Neuronale Netze.....	S.8
2.2.4	Generative KI.....	S.12
2.2.5	Diffusionsmodelle.....	S.13
3	Songproduktion mit künstlicher Intelligenz	
3.1	Einordnung von KI-Systemen in der Musik.....	S.15
3.1.1	Geschichte der KI in der Musikproduktion.....	S.15
3.1.2	Suno – Eine Einordnung.....	S.16
3.1.3	Suno – Eine rechtliche Einordnung.....	S.18
3.2	Produktionstagebuch.....	S.21
3.2.1	Ab 13 Februar.....	S.21
3.2.2	Ab 19 Februar.....	S.29
3.2.3	Ab 25 Februar.....	S.34
3.2.4	Ab 06 März.....	S.40
3.2.5	Ab 09 März.....	S.43

4	Systematischer Test	
4.1	Motivation.....	S.49
4.2	Methodik.....	S.50
4.3	Durchführung und Auswertung.....	S.53
4.3.1	Einfluss des Parameters Weirdness.....	S.53
4.3.2	Auswertung des Parameters Weirdness	S.64
4.3.3	Einfluss des Parameters Style Influence.....	S.66
4.3.4	Auswertung des Parameters Style Influence	S.77
4.3.5	Einfluss des Parameters Audio Influence.....	S.78
4.3.6	Auswertung des Parameters Audio Influence.....	S.88
5	Abschlussbetrachtung	
5.1	Zusammenfassung der Ergebnisse.....	S.89
5.2	Fazit.....	S.91
V.	Literaturverzeichnis	VI
VI.	Abbildungsverzeichnis	XII
VII.	Anhang.....	XIV

Abkürzungsverzeichnis

bzw.	beziehungsweise
DDEX	Digital data exchange
ff.	folgende
FX	Effekte
HCM	Human created music
KI	Künstliche Intelligenz
Sog.	Sogenannte
Tsd.	Tausend
u. a.	unter anderem
z. B.	zum Beispiel
KNN	Künstliches neuronales Netz

Glossar

Crossfade	Ein Crossfade ist eine Überblendung zweier Audiosignale ineinander, bei der ein Signal ausgeblendet wird, während ein anderes zeitgleich eingeblendet wird.
Crescendo	Allmählich lauter werdende Passage innerhalb eines Musikstückes.
DDEX	Normungsorganisation, die sich auf die Entwicklung von Standards für den digitalen Datenaustausch in der Musikindustrie spezialisiert hat.
Delay	Ein Audioeffekt, der ein Signal bewusst verzögert, um es erneut mehrmals hintereinander wiederzugeben.
Downlifter	Ein Soundeffekt, bei dem die Tonhöhe über die Zeit nach unten verläuft, um einen Übergang von zwei Bereichen innerhalb des musikalischen Arrangements zu schaffen.
Equalizer	Equalizer oder kurz EQ ist ein Werkzeug, mit dem sich bestimmte Frequenzbereiche eines Audiomaterials mithilfe bestimmter Filtereinstellungen anheben und absenken lassen.
Exciter	Ein Exciter ist ein Audioprozessor, der durch gezielte Erzeugung zusätzlicher Obertöne den Klang eines Signals brillanter und präsenter erscheinen lässt.
Fade-in / Fade-out	Ein allmählicher Anstieg oder Abstieg der Lautstärke eines Audiosignals.
Kompressor	Ein Kompressor senkt den Audiopegel ab einer eingestellten <i>Threshold</i> ab, wobei die Geschwindigkeit des Kompressors mit den Parametern <i>Attack</i> und <i>Release</i> eingestellt werden kann. Mit dem <i>Output Gain</i> kann das gesamte Material wieder angehoben werden
Legato	Artikulationsbezeichnung in der Musiknotation, die angibt, dass über mehrere Noten hinweg ohne Unterbrechung gespielt werden soll.
Nachhall	Der Nachhall, umgangssprachlich auch Hallfahne bezeichnet, ist der Begriff der das Ausklingen eines Schallsignals innerhalb eines (künstlichen) Raums mit (künstlichen) Reflexionen beschreibt.

Pitch correction	Tonhöhenkorrektur
Riser	Ein Soundeffekt, bei dem die Tonhöhe über die Zeit nach oben verläuft, um einen Übergang von zwei Bereichen innerhalb des musikalischen Arrangements zu schaffen.
RX De Noise	Audiobearbeitungssoftware von <i>Native Instruments</i> um Rauschanteile aus Audiomaterial zu entfernen,
Spectral Layers	Audioschnitt-Software für gezielte Bearbeitung im spektralen Bereich zum Beispiel zum Entfernen von Störgeräuschen.
Stems	Gruppierte Einzelspuren eines Songs.
Swooshes	Anschwellende, rauschähnliche Soundeffekte. Meist zur klanglichen Gestaltung von Übergängen verwendet.
Tenuto	Artikulationsbezeichnung in der Musiknotation, die angibt, dass eine Note über ihre volle notierte Dauer gehalten und nicht vorzeitig beendet werden soll.
VariAudio	Werkzeug in Cubase zur Bearbeitung von Tonhöhe und Timing.
Vocal chops	Kurze, stark bearbeitete Segmente aus einer Gesangsspur.

Anmerkungen

*1

Während der Aufnahme kam es aufgrund der Anforderung das Einpegeln und Sprechen gleichzeitig zu übernehmen, zu einer Übersteuerung des Signals. Eine nachträgliche Anpassung wurde bewusst nicht vorgenommen, um zu untersuchen, ob Suno in der Lage ist, auch mit übersteuertem Ausgangsmaterial umzugehen.

*2

Beim Export des Materials wurde versehentlich ein zu großer Zeitbereich ausgewählt, wodurch das Ergebnis im Style-Influence-Durchlauf eine übermäßige Länge aufweist. Am Ende sind zudem teilweise Audioschnipsel enthalten, die aus zuvor in Suno Studio nicht vollständig gelöscht Material stammen. Diese werden bei der Bewertung nicht berücksichtigt.

Ehrenwörtliche Erklärung

Hiermit erkläre ich ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Aus fremden Quellen direkt oder indirekt übernommene Gedanken habe ich als solche kenntlich gemacht.

Die vorliegende Arbeit habe ich bisher keinem anderen Prüfungsamt in gleicher oder vergleichbarer Form vorgelegt. Sie wurde bisher auch nicht veröffentlicht.

Ort, Datum 17.05.26 Lauffen

Unterschrift Peter Volkert

1 Einleitung

1.1 Problemstellung

Von der Hinzunahme der künstlichen Intelligenz als Hilfsmittel im Privatleben haben schon sehr viele Gebrauch gemacht. 32,7% der Menschen in der EU (16-74 Jahre) verwendeten laut dem statistischen Amt der Europäischen Union, kurz Eurostat, 2025 generative AI-Tools, wobei der größte Anteil diese für persönliche Aufgaben nutzt (2025). Dies sind in Anbetracht älterer Erhebungen keine überraschenden Zahlen. Bereits im Jahr 2023 zeigte eine exemplarische Umfrage an der Hochschule der Medien Stuttgart, an der 643 Studierende teilnahmen (98,62 % der Befragten), dass KI-basierte Tools in der Hochschullehre weit verbreitet genutzt wurden: 85,58 % gaben an, textgenerierende Anwendungen zu verwenden, während 94,14 % Übersetzungstools einsetzten (Gottschling et al., 2024, S. 129).

Die Nutzung bildgenerierender Tools (von 28,3 % der Befragten) und codegenerierender Tools (13,91 % der Befragten), also KI-Tools zur Bearbeitung komplexerer Aufgaben, waren noch nicht so weit verbreitet. Auch merkt Gottschling et al. an: „Bezüglich der Nutzungshäufigkeit in verschiedenen Lebensbereichen zeigten explorative Wilcoxon-Rangsummentests für verbundene Stichproben, dass KI-Tools im Studium signifikant häufiger genutzt wurden als im Privatleben ... Gleichzeitig wurden KI-Tools im Privatleben häufiger genutzt als in Nebenjobs oder Praktika“ (2024, S. 129).

Mittlerweile hat, wie die Eurostat Daten zeigen, die KI jedoch Einzug ins Privatleben der breiten Bevölkerung gefunden (32,7 %) und wird dort sogar häufiger verwendet als im beruflichen Kontext (15,1 %). Andere Umfragen zeigen aber, dass insbesondere in beruflichen Rollen mit höherer Verantwortung und komplexeren Aufgaben der Einsatz von KI weit verbreitet ist. Laut einer globalen Umfrage der Boston Consulting Group (BCG), eine US-Amerikanische Unternehmensberatungsfirma, nutzten laut einer im Juni 2025 von mehr als 10 600 Beschäftigten erhobenen Umfrage in 11 Ländern, 72 % der Befragten KI regelmäßig (mehrmals pro Woche oder täglich) bei der Arbeit, während Führungskräfte sie insgesamt häufiger verwendeten (88 %) als sog. Frontline Employees (51 %) oder Manager:innen (78 %) (Beauchene et al., 2025). Auch die durch künstliche Intelligenz erzielten Arbeitsergebnisse, scheinen Verunsicherung unter Beschäftigten auszulösen. Während 47 % aller Befragten angaben, durch den Einsatz künstlicher Intelligenz mehr als eine Stunde Arbeit am Tag zu sparen, äußerten 41 % die Erwartung, dass ihr Arbeitsplatz innerhalb der nächsten zehn Jahre sicher oder wahrscheinlich wegfallen könnte (Beauchene et al., 202).

Die Erwartung scheint nicht unbegründet, denn wie Heinen et al. bereits 2017 in ihrem Fachbeitrag *Künstliche Intelligenz und der Faktor Arbeit* schreiben: „Historisch gesehen hat jede Innovation, die den Produktionsfaktor Kapital produktiver machte, mittelbar oder unmittelbar auch den Faktor Arbeit beeinflusst“ (S. 715).

Dass AI-Tools mittlerweile fest in den Softwareentwicklungsprozess integriert sind, zeigt die Entwicklerumfrage *Developer Survey 2025* mit über 49.000 Antworten aus 177 Ländern, in welcher 50,6 % aller professionellen Entwickler angaben, täglich KI-Tools zu verwenden (Developer Survey, 2025). 2023 waren es noch 44,2% (Stack Overflow, 2023).

So macht die KI-Entwicklung natürlich auch bei Muskschaffenden nicht Halt. Laut einer Online-Befragung der LANDR Audio Inc. (2025) nutzten 87 % der 1,241 befragten Muskschaffenden KI-Tools in mindestens einem kreativen oder technischen Teil ihres Workflows. Dabei wird KI vor allem zur Bewältigung technischer Aufgaben eingesetzt: 77 % der Befragten nutzten KI zum Beispiel für Audio-Restoration, Timing-Korrekturen oder Pitch-Correction (siehe Glossar). Zudem verwenden 70 % KI-gestützte Tools zur Promotion ihrer Musik.

Für kreative Aufgaben wie Songwriting oder Komposition setzen hingegen nur 46 % der Befragten auf KI. Songgeneratoren werden von 29 % der Teilnehmenden genutzt, insbesondere zur Erstellung einzelner Songbestandteile wie Gesang oder Instrumentalspuren. Weitere 40 % gaben an, an der Nutzung solcher Generatoren interessiert zu sein (LANDR Audio, 2025).

Zwar ist die Online-Befragung von LANDR Audio Inc. aufgrund ihrer KI-affinen Stichprobe der eigenen Userbase, nicht als repräsentativ für alle Muskschaffenden zu betrachten, sie verdeutlicht jedoch exemplarisch, dass KI mittlerweile sowohl im technischen als auch im kreativen Musikproduktionsprozess eine relevante Rolle spielt (LANDR Audio, 2025).

Betrachtet man nicht nur erhobene Umfragen, sondern auch verschiedene Schlagzeilen erkennt man das KI an vielen Schauplätzen, umfangreiche Fragen aufwirft und natürlich auch die Musik und Tonindustrie umtreibt. Durch neue KI-Technologie sehen z. B. Synchronsprecherinnen und Synchronsprecher ihre Arbeitsplätze und Urheberrechte gefährdet (Jacob, 2026). Mit Udio vollständig erstellte KI-Musik hielt sich zwei Wochen in den deutschen Single Charts (Lablack, 2025) und der KI-generierte Song *Dust on the Wind* der KI-generierten Band *Velvet Sundown* „...wurde fast zwei Millionen Mal auf Spotify und fast eine Million Mal bei YouTube angehört“ (Dampz, 2025), wobei Spotify den Song erst als KI-generiert kennzeichnete, als die Songs bereits Hunderttausendfache Klicks erzielt hatten (Dampz, 2025), und die Bandbiografie in: "The Velvet Sundown ist ein synthetisches Musikprojekt, das unter menschlicher, kreativer Leitung steht und mit Unterstützung

künstlicher Intelligenz komponiert, vertont und visualisiert wird" änderte. (Spotify, zitiert nach Dampz, 2025).

Im September 2025, kurz nach Erscheinen des Beitrags von Nils Dampz im Juni 2025, änderte Spotify seine Richtlinien mit dem Fokus auf die Punkte:

Verbesserte Durchsetzung bei Verstößen gegen Identitätsmissbrauch, z.B. bei Stimmenmissbrauch durch KI-Programme. Musik Spam Filter, um KI-generierte Spammusik zu detektieren und vom Bezahlsystem auszuschließen und einer KI-Kennzeichnung in den Musik-Credits nach Branchenstandard über DDEX (siehe Glossar) (Spotify, 2025).

Die Industrie bereitet sich also auf einen Umbruch vor. In der Einleitung zu einer von Goldmedia, im Auftrag der GEMA durchgeführten Studie heißt es: „Generative Künstliche Intelligenz (KI) ist die größte technologische Veränderung für die Kreativbranche seit über 30 Jahren. Technologien wie ChatGPT werfen unweigerlich Fragen nach geistigem Eigentum und Umgang mit Urheberrechten auf“ (Goldmedia, 2024). Die Studie weist darauf hin, dass die Einkünfte von Musikerinnen und Musikern durch den Einsatz von generativer KI schon 2028 in Deutschland und Frankreich, um 27 % zurückgehen könnte (Goldmedia, 2024, S. 8). Die Studie stellt außerdem auch fest, dass KI mittlerweile von vielen Musikurhebern genutzt wird. So haben 35 % der 15 Tsd. für diese Studie Befragten KI-Technologien bereits verwendet. Bei den unter 35-jährigen war der Anteil deutlich höher und liegt bei 51 % (Goldmedia, 2024, S. 9).

Vor diesem Hintergrund erscheint es sinnvoll, sich auch praxisnah mit dem Umgang KI-gestützter Programme vertraut zu machen und deren Funktionsweisen, Möglichkeiten, Vorteile und Grenzen im Verlauf dieser Bachelorarbeit konkret zu erproben.

1.2 Ziele

Das Ziel dieser Arbeit ist es, dem Leser einen Einblick in die Nutzung bestimmter KI-gestützter Funktionen im Rahmen einer Songproduktion zu geben.

Dabei soll die Arbeitsweise anhand eines Produktionstagebuchs dokumentiert werden und am Ende eines jeden Produktionsschrittes eine Zusammenfassung der wichtigsten Feststellungen erfolgen. Diese sollen schließlich in gezielten Erhebungen (vgl. Kapitel 4) überprüft werden, um am Ende dieser Arbeit ein Fazit darüber ziehen zu können, in welchen Arbeitsschritten KI sinnvoll technisch und künstlerisch angewandt werden kann und an welchen Stellen sie an ihre Grenzen stößt.

2 Die Grundlagen von künstlicher Intelligenz

Um die Ergebnisse der von mir genutzten KI-gestützten Programme nachzuvollziehen, ist es wichtig, sich mit der grundlegenden Funktion und Bedeutung bestimmter Begrifflichkeiten auseinanderzusetzen. So soll in diesem Kapitel der Leser auf die spätere praxisnahe Anwendung, Analyse und Einordnung in den weiteren Kapiteln vorbereitet werden.

2.1 Definition KI

Eine allgemeingültige Definition für künstliche Intelligenz gibt es derzeit nicht, auch weil sich das Verständnis für KI stetig weiterentwickelt. Laut Jäkel wählte John McCarthy als Pionier dieses Forschungsfeldes den Namen, um mehr Mittel und Studenten anzuziehen, da ihm dieser reißerischer und medienwirksamer schien als der Vorschlag seines Mentors Shannon „Automatenstudien“ (2025, S. 64).

In dieser Arbeit ist es daher nicht wichtig, eine allgemeingültige Definition zu finden, sondern den Begriff „Künstliche Intelligenz“ einzugrenzen und zu beschreiben, wie er von Wissenschaftlern und Konsumenten verstanden wird,

Aufgrund des schnellwachsenden Fortschritts innerhalb dieses Gebiets müssen Bewertungskriterien und Definitionen sowieso regelmäßig neu überprüft und angepasst werden, da Maschinen ja schon zu Zeiten des Taschenrechners als „intelligent“ hätten gelten können. Die künstliche Intelligenz schreitet jedoch seitdem und vor allem in letzter Zeit rasant voran und entwickelt sich zunehmend dahin, die Funktionalität vieler einzelner, bislang auf spezifische Aufgaben beschränkter Programme mit selbständigen Problemlösungsmechanismen zu vereinen (Jäkel, 2025, S.63-67). Mockenhaupt stellt diese Entwicklung in drei Stufen dar.

Früher (Stufe 1) wurde Künstliche Intelligenz als die Fähigkeit verstanden, Aufgaben zu lösen, die mit höheren intellektuellen Verarbeitungsleistungen des Menschen in Verbindung gebracht werden, wie zum Beispiel das Rechnen. Stufe 2 erweitert dies um die Fähigkeit, dass Systeme, situationsabhängig und autonom Entscheidungen treffen, anstatt starren if-und else-Schleifen zu folgen. Aktuelle KI-Systeme (Stufe 3) können eigenständig Probleme bearbeiten und auf veränderte Bedingungen reagieren (Mockenhaupt, 2024, S. 57).

Mockenhaupt zeigt hier sehr schön, wie die als intelligent definierten Fähigkeiten von KI-Systemen immer weiter fortschreiten und dadurch auch den Rahmen alter Definitionen verlassen.

Daher ist eine allumfassende Definition nicht einfach und kann schnell unpräzise werden. Elaine Rich wählte den Weg bei ihrer Definition nicht auf die konkrete Umsetzung (Wie arbeitet KI), sondern auf ihren Zweck (Was macht sie) zu zielen. Die Definition von Rich *“Artificial intelligence is the study of how to make computers do things at which, at the moment, people are better.”* (Rich, 1983, S. 6) ist daher sehr prägnant und bleibt immer aktuell. Daraus lässt sich ableiten: Künstliche Intelligenz entwickelt sich immer weiter und ist darauf ausgelegt, ihre Arbeit so gut wie Menschen zu verrichten und mittlerweile auch dazu „... in der Lage, Wertschöpfung bei Tätigkeiten zu generieren, die früher menschlicher Intelligenz vorbehalten waren“ (Brüggemann, 2025, S. 19).

Während der Begriff Intelligenz auch gerne mit dem Menschsein verknüpft wird, impliziert das Wort Künstlichkeit eher eine Abgrenzung dazu. In Elaine Richs Definition finden wir diese Abgrenzung bereits:

“...how to make computers do things at which, at the moment, people are better.” (Rich, 1983, S. 6).

Demnach verrichten Computer Dinge, die der Mensch noch besser kann. Es gibt jedoch einen Unterschied in der Frage, ob der Output einer Maschine als künstlich bewertet wird oder ihre Beschaffenheit und Funktion.

Zwar orientiert sich die Funktion vieler KI-Systeme, etwa von neuronalen Netzen (siehe Kapitel 2.2.3), an der biologischen Gehirnstruktur. Die Maschine selbst bleibt jedoch eine technische, vom Menschen konstruierte, nicht lebendige Einheit und das Programm darin willenlos sowie seelenlos und ohne Bewusstsein, wodurch sie in Beschaffenheit und Funktion klar als künstlich bewertet werden kann.

Die Bewertung eines KI-generierten Ergebnisses als künstlich hängt jedoch vom Konsumenten ab, denn wenn „Künstlichkeit“ im Wortsinne als Nicht-menschengemacht zu verstehen ist, setzt dies voraus, dass eine Unterscheidung zu HCM (human created music) wahrnehmbar ist. Bryce (2024) zeigt, dass KI-generierte Musik zu einem Anteil von ~61% von HCM Musik zu unterscheiden war und musikalische Vorerfahrung den Probanden dabei half, den Durchschnitt zu übertreffen. Nach einer inhaltlichen Einführung in Funktionsweise und den typischen Merkmalen des verwendeten Modells stieg die Erkennungsrate in einem zweiten Durchlauf auf ~69% an.

Durch den fortschreitenden technischen Entwicklungsstand und den zunehmenden Einsatz von KI als unterstützendes Kreativwerkzeug – und nicht nur zur vollständig autonomen Songgenerierung – ist jedoch davon auszugehen, dass die Erkennbarkeit KI-generierter Musik künftig weiter abnehmen wird und auch dadurch unsere Akzeptanzschwelle steigt.

Studien zeigen, dass KI-generierte Musik hinsichtlich emotionaler Wirkung mit menschlich komponierter Musik vergleichbar sein kann und in bestimmten Parametern sogar als stärker erregend wahrgenommen wird (Fišer, 2025), jedoch insgesamt weniger beliebt ist, wenn bekannt ist, dass es sich bei der Musik um KI-Musik handelt (Shank, 2023).

Insgesamt scheint die Künstlichkeit im Wortsinn nicht in jedem Fall für den Zuhörenden wahrnehmbar zu sein, wird jedoch weiterhin als negative Eigenschaft verstanden.

Das Ergebnis künstlicher Intelligenz ist somit inzwischen in der Lage, menschliche Echtheit überzeugend zu imitieren und wird vor allem dann als künstlich erkannt, wenn Hörende über entsprechendes Vorwissen in Musiktheorie oder über Kenntnisse zu KI-Modellen verfügen. Künstlichkeit scheint somit mehr eine Zuschreibung von außen, als ein objektiv messbarer Parameter zu sein.

2.2 Wie funktioniert künstliche Intelligenz

2.2.1 Maschinelles Lernen

Das Zitat „*Die Definition von Wahnsinn ist, immer wieder das Gleiche zu tun und andere Ergebnisse zu erwarten*“ wird Albert Einstein zugeschrieben. Unabhängig von der tatsächlichen Herkunft des Zitats eignet es sich, um die grundlegende Idee hinter künstlicher Intelligenz zu veranschaulichen: Ein intelligentes System muss in der Lage sein, aus Erfahrungen zu lernen, um das eigene Handeln entsprechend anzupassen. Dies erfordert eine Strategie, wie Erfahrungen in Form von Daten von einem System genutzt werden können (Kersting, 2019, S. 19). Um Erfahrungen innerhalb eines Systems zu sammeln, müssen diese erst einmal erzeugt werden. Dies geschieht beim maschinellen Lernen.

Dabei gibt es drei zentrale Lernparadigmen: Überwachtes Lernen, unüberwachtes Lernen und verstärkendes Lernen.

Überwachtes Lernen trainiert ein Modell darauf, aus einer gegebenen Eingabe, die richtige Ausgabe vorherzusagen. *Unüberwachtes Lernen* zielt darauf ab, ein Modell so zu trainieren, dass es eigenständig Muster, Zusammenhänge und Strukturen in Daten erkennt und das *verstärkende Lernen* trainiert ein Modell für die Maßnahme mit der größtmöglichen Belohnung, nach einer Bewertung der eigenen Umgebung (Bergmann 2026a). Für überwachtes und unüberwachtes Lernen bedeutet dies grundsätzlich: „Wenn dem Lernalgorithmus die richtige Lösung aus den Trainingsdaten bekannt ist, spricht man von überwachtem Lernen, ansonsten von unüberwachten Lernen“ (Kersting, 2019, S.26).

Innerhalb des gesamten Trainingsprozesses für ein bestimmtes Modell können mehrere Lernparadigmen kombiniert werden.

Als Praxisbeispiel für maschinelles Lernen innerhalb generativer Audiomodelle dient die Verbrauchermittelung CA AB 2013 Disclosure von Suno (o. D.). Sie gibt Einblick in die enorme Datenmenge, die für das Training solcher Systeme notwendig ist. Laut Suno wurde das Modell mit rund zehn Millionen öffentlich zugänglichen Musikdateien inklusive der dazugehörigen Metadaten trainiert. Zusätzlich fließen Nutzerinhalte und Aktivitätsdaten in die Optimierung ein.

Dies lässt darauf schließen, dass unterschiedliche Trainingsmethoden kombiniert werden: Bei der Analyse der Musikdateien kann gut überwachtes Lernen zum Einsatz kommen, da durch die Metadaten (z. B. Genre-Labels wie ‚Jazz‘) der Zielzustand definiert ist. Das Modell lernt so die statistischen Muster eines Genres, um diese generativ reproduzieren zu können. Die Lösung ist also innerhalb des Trainingsprozesses bekannt. Die Verarbeitung von dynamischen Nutzerinhalten und Aktivitätsdaten deutet hingegen auf unüberwachtes Lernen oder bestärkendes Lernen hin. Hier ist keine feste Lösung festgelegt. Stattdessen geben Nutzer dem System, beispielsweise durch das wiederholte Anhören eines Songs, Aufschluss über ihre Präferenzen. Innerhalb der Suno-Umgebung wird analysiert, welche generierten Daten die größtmögliche Belohnung, z. B. in Form von investierter Zeit der Anwender, auslösen. Auf Basis dieser Informationen kann das Modell kontinuierlich optimiert werden, was einem verstärkendem Lernalgorithmus entspräche.

Um diese Menge an Daten zu verarbeiten, mussten sich die Strukturen und Algorithmen des maschinellen Lernens immer weiterentwickeln. Ein Teilbereich des maschinellen Lernens bildet das sog. Deep Learning, das auch in generativen Audiomodellen zum Einsatz kommt.

2.2.2 Deep Learning

Deep Learning ist ein Teilbereich des maschinellen Lernens und basiert auf mehrschichtigen künstlichen neuronalen Netzen. In den letzten Jahrzehnten hat sich dieser Ansatz in vielen Anwendungsfeldern der künstlichen Intelligenz als leistungsfähige Modellarchitektur etabliert. Im Unterschied zu klassischen Verfahren, die auf explizit definierte Algorithmen und manuell entwickelten Merkmalen beruhen, nutzt Deep Learning komplexe Netzwerke, um selbstständig Strukturen und Muster in hochdimensionalen Daten zu erkennen (Bergmann, 2026a).

Ein wesentlicher Unterschied zum klassischen maschinellen Lernen besteht darin, dass Deep-Learning-Modelle häufig direkt mit Rohdaten arbeiten. Dadurch wird ein Großteil der

Merkmalsextraktion automatisiert vom Modell selbst übernommen (Bergmann, 2026a). Das System kann also anhand der Wellenform des Audiomaterials mit der Zeit eigenständig anhand bestimmter Muster erkennen, in welches Genre der Song kategorisiert werden muss, oder um welches Instrument es sich dabei handelt.

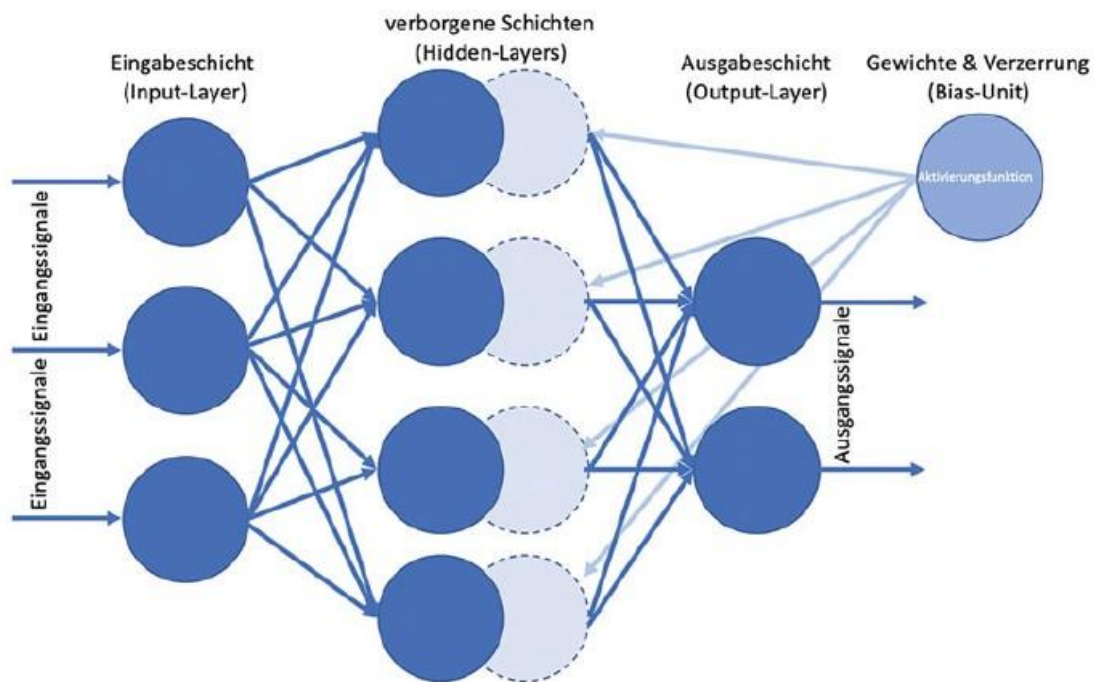
Deep Learning stellt somit einen großen Schritt hin zu autonomen Modellen dar, die sich von vorgefertigten, regelbasierten Verfahren lösen und sich den komplexen Problemlösungs- und Analysestrategien des menschlichen Gehirns annähern.

2.2.3 Künstliche neuronale Netze

Das menschliche Gehirn unterscheidet sich in seiner Arbeitsweise stark von einem Computer. Während Computer in der Regel Rechenoperationen sequenziell ausführen, arbeitet das menschliche Gehirn mithilfe eines komplexen Netzwerkes aus Milliarden von Neuronen parallel. Diese Struktur dient als Vorbild für künstliche neuronale Netze (KNN), die ebenfalls aus einer Vielzahl miteinander verbundener künstlicher Neuronen bestehen, welche in unterschiedlichen Schichten organisiert sind (vgl. Abb. 1) (Mockenhaupt, 2024, S. 192).

Die *Eingabeschicht* (Input Layer) bildet dabei die Schnittstelle zwischen den Eingangsdaten und den nachfolgenden Schichten. Sie verarbeitet die eingegebenen Daten und leitet diese an die nächste Schicht weiter, die als *verborgene Schicht* (Hidden Layer) bezeichnet wird. Dabei ist in der Regel jedes Neuron mit den Neuronen der vorherigen und nachfolgenden Schicht verbunden, wodurch eine komplexe, vielschichtige Struktur aus Verbindungen entsteht. Ein KNN kann aus einer oder mehreren verborgenen Schichten bestehen. Diese wiederum bestehen aus künstlichen Neuronen, die die eingehenden Signale durch Gewichtungen transformieren und mithilfe von Aktivierungsfunktionen in eine geeignete, neue Darstellung überführen. Typische Aktivierungsfunktionen hierbei sind beispielsweise die Sigmoid- oder die Tangens-hyperbolicus-Funktion. Dieser Prozess wird schichtweise fortgesetzt, bis die Informationen die abschließende *Ausgabeschicht* (Output Layer) erreichen (Wuttke, 2024). Die Ausgabe des Netzwerks wird anschließend mit dem gewünschten Zielwert verglichen. Die Differenz zwischen der vorhergesagten und der tatsächlichen Ausgabe wird mithilfe einer Verlustfunktion berechnet. Diese Differenz zeigt am Ende, wie genau das Modell Vorhersagen für eine bestimmte Eingabe trifft und bildet die Grundlage für die Anpassung der Modellparameter im Trainingsprozess (La La, 2019). Das Ziel innerhalb des Modelltrainings ist es, diese Differenz zu minimieren.

Abb. 1 - zeigt ein künstliches neuronales Netz mit den drei Schichten (Eingabeschicht, verborgene Schichten und Ausgabeschicht) und dem Einfluss der Gewichte, Verzerrung und Aktivierungsfunktion auf die einzelnen Neuronen (Mockenhaupt, 2024, S.194).



Um ein neuronales Netz zu trainieren, berechnet jedes einzelne Neuron $z = \sum_{i=1}^n w_i x_i + b$ und $a = \sigma(z)$. Es gilt:

- x_i = Eingabefunktion
- w_i = Gewicht
- b = Verzerrung (Bias)
- z = gewichtete Summe (lineare Transformation)
- σ = Aktivierungsfunktion (nichtlineare Transformation)
- a = Ausgabe

Das bedeutet, die Ausgabe eines Neurons hängt von der linearen Transformation, also der gewichteten Eingabe eines jeden Neurons der vorherigen Schicht, und der nichtlinearen Transformation, also der Anwendung einer Aktivierungsfunktion dieser gewichteten Summe ab (Fangfang, 2026).

Ein Beispiel für so eine Aktivierungsfunktion ist die Sigmoid-Funktion $\sigma(z) = \frac{1}{1+e^{-z}}$ (vgl. Abb. 2), die das lineare Ergebnis z eines Neurons in eine nichtlineare Ausgabe transformiert. Dadurch können komplexe Zusammenhänge und Probleme gelöst werden, die mit geraden Linien nicht lösbar sind. Die Sigmoid Funktion erfüllt dabei mehrere wichtige Eigenschaften: Sie beschränkt die Ausgabe auf 0 und 1, ist an jeder Stelle differenzierbar und weist einen

nichtlinearen Verlauf auf, wodurch das Modell in der Lage ist, komplexe Muster und Entscheidungsgrenzen zu erlernen (Singh, 2025).

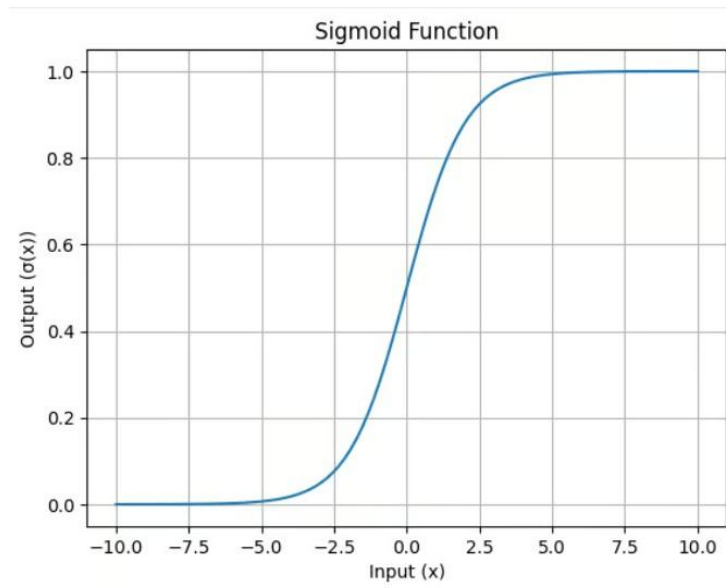


Abb. 2 - zeigt eine Visualisierung der Sigmoid Funktion. Der s-förmige Verlauf ist dabei ihr Charakteristikum und verdeutlicht, dass die Funktionswerte stets im Bereich $0 < \sigma(z) < 1$ liegen. Außerdem ist ersichtlich, dass die Funktion bei sehr großen oder sehr kleinen Eingangswerten in einen Sättigungsbereich nahe 1 oder 0 übergeht, in dem die Steigung der Kurve nahezu 0 beträgt. Dies sorgt dafür, dass bei dichten neuronalen Netzen der Lernprozess verlangsamt wird, da die Gewichte zu langsam aktualisiert werden (Singh, 2025).

Zur Minimierung der zuvor beschriebenen Verlustfunktion wird im Trainingsprozess das Verfahren der Backpropagation eingesetzt. In diesem Prozess werden die richtigen Gewichtungen ermittelt, die auf die Verbindungen zwischen den Neuronen angewendet werden müssen, um die Differenz zwischen der vorhergesagten und der tatsächlichen Ausgabe zu minimieren. Anders ausgedrückt dient die Backpropagation der Minimierung der Verlustfunktion. Dabei kommt das Gradientenabstiegsverfahren zum Einsatz. In diesem gilt es, innerhalb der Verlustfunktion das globale Minimum zu ermitteln (La La, 2019).

Die Backpropagation arbeitet sich dabei rückwärts vom Output des Modells, unter Anwendung der Kettenregel, rekursiv bis zur Eingabeschicht durch das neuronale Netz, um den Gradienten der Verlustfunktion in Bezug auf jeden einzelnen Parameter des neuronalen Netzes zu berechnen. Ein Gradient ist ein Vektor, der die partiellen Ableitungen einer Funktion nach all ihren Variablen zusammenfasst (Bergmann, 2026b) und bildet im mehrdimensionalen Raum das Äquivalent zur Ableitung im zweidimensionalen Raum (Lang, 2021).

„Da die Verlustfunktion die Ausgabe eines neuronalen Netzes als Eingabe nimmt und diese Ausgabe des neuronalen Netzes eine zusammengesetzte Funktion ist, die viele verschachtelte Aktivierungsfunktionen einzelner Neuronen umfasst, führt die Differenzierung der Verlustfunktion zur Differenzierung des gesamten neuronalen Netzes. Zu diesem Zweck wird bei der Backpropagation die Kettenregel verwendet“ (Bergmann, 2026b).

Das Gradientenabstiegsverfahren ermittelt nun das globale Minimum der Verlustfunktion, um den optimalen Trainingszustand des Modells zu finden. Dabei gilt: „Wenn wir ... in die negative Richtung des Gradienten gehen, wissen wir, dass die Funktion am stärksten abfällt und wir somit auch dem Minimum am schnellsten näherkommen (Lang, 2021).“

Die Steigung der Kurve – im eindimensionalen Fall – beziehungsweise der Gradient – im mehrdimensionalen Raum – gibt Aufschluss über die Richtung des Gradientenabstiegsverfahrens, woraufhin die Gewichte entsprechend angepasst werden. Bedeutet: Bei negativer Steigung werden Werte zu den Gewichtungen addiert. Bei positiver Steigung werden Werte von den Gewichtungen subtrahiert. Dieser zu addierende oder zu subtrahierende Wert wird als Lernrate bezeichnet. Die Wahl einer geeigneten Lernrate stellt dabei eine zentrale Herausforderung dar: Eine zu kleine Lernrate führt zu einem langsamen Trainingsprozess, während eine zu große Lernrate dazu führen kann, dass das Minimum überschritten wird (vgl. Abb. 3) (La La, 2019).

Neben der Wahl einer möglichst effizient gewählten Lernrate, die die Geschwindigkeit des Deep Learning Modells massiv beeinflusst, gibt es u. a. noch einen weiteren Aspekt, mit dem bei der Nutzung des Gradientenabstiegsverfahrens umgegangen werden muss:

Das Verfahren bewahrt das Modell nicht davor, statt in dem gewünschten globalen Minimum, ein lokales Minimum als bestes Trainingsergebnis anzunehmen. Darum müssen beim Training des Modells mehrere Startpunkte getestet werden, um eine Konvergenz zu einem bestimmten Minimum feststellen zu können (Lang, 2021).

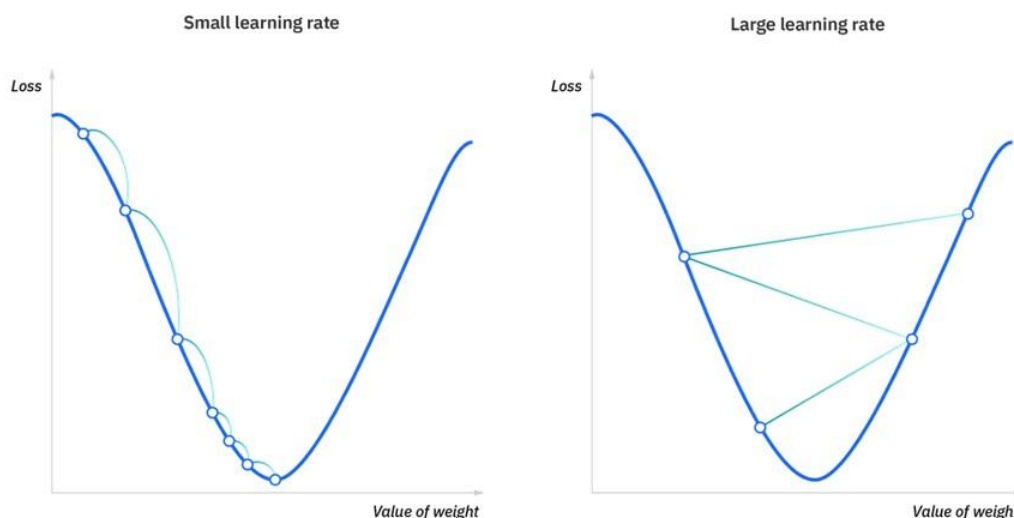


Abb. 3 - veranschaulicht die Auswirkung unterschiedlicher Lernraten (IBM, o. D.).

2.2.4 Generative KI

Generative KI ist künstliche Intelligenz, die als Antwort auf einen Prompt oder die Anfrage eines Benutzers neue Daten wie Text, Bilder, Video oder Audio generiert. (Stryker, C., o. D.). Fortgeschrittene Modelle sind mittlerweile in der Lage, Musik über eine große Bandbreite unterschiedlicher Genres hinweg zu erzeugen. Dabei bilden sie verschiedenste Instrumente und Stilrichtungen ab. (SAP, 2026). Ein Beispiel hierfür ist das Programm Suno, welches anhand einer Vorlage oder Prompts Soundeffekte, Samples, Melodien, Vocals oder vollständige Songs erzeugt.

Diese Funktion kann man in zwei Hauptansätze unterteilen: Midi-Generierung und Audio Wellenform-Generierung.

Die Midi-Generierung ist hierbei technisch weniger aufwendig, da die generierten MIDI-Daten erst durch ein virtuelles Instrument in einen Klang umgesetzt werden.

Die Audio-Wellenform-Generierung ist wesentlich komplexer, da man hier keine externe Technologie nutzt, sondern der Klang vollständig neu erzeugt wird (Landschoot, 2024).

Generative KI umfasst drei Phasen. Das Training, die Optimierung und die Generierung.

Beim Training wird das Modell anhand von riesigen Mengen Rohdaten auf die Vorhersagung des nächsten Elements in einer Sequenz trainiert. Dies kann z. B. das nächste Wort in einem Satz oder eben auch der nächste Ton in einem Lied sein. Der Algorithmus wird innerhalb des Trainingsprozesses ständig angepasst, um die Differenz zwischen den Vorhersagungen und den tatsächlichen Daten zu minimieren. Das Ergebnis dieses Trainings ist dann ein neuronales Netz von Parametern (vgl. Kapitel 2.2.3).

Um das Modell nun auf eine bestimmte Aufgabe der Inhaltserstellung zu verbessern, wird bei der Optimierung die KI mit Daten und den entsprechend richtigen Antworten gespeist, die die Anwendung wahrscheinlich erhalten wird (Stryker, C., o. D.). Der Rechenaufwand und Zeitaufwand des Trainings und Optimierens ist sehr hoch, wodurch Firmen wie Nvidia massiv durch die intensive Nutzung generativer Modelle profitieren. Da GPUs im Gegensatz zu klassischen CPUs darauf optimiert sind, viele Operationen gleichzeitig auszuführen, eignen sie sich ideal für Matrixberechnungen in neuronalen Netzen, weshalb die Nvidia-Aktie zu den größten Gewinnern an der Wall Street gehört (ad-hoc-news, 2025).

Die Generierung der Inhalte kann mit verschiedenen Modellarchitekturen erfolgen. Landschoot gibt bei Audio-generierender KI hier das Diffusionsmodell als moderne Systemlösung an (2024).

2.2.5 Diffusionsmodelle

Generative KI-Modelle modellieren eine Wahrscheinlichkeitsdichtefunktion über dem gesamten Datenraum. Diese beschreibt, mit welcher Wahrscheinlichkeit bestimmte Regionen des Datenraums auftreten. Bei der Erzeugung eines neuen Bildes geht das Modell von einer hohen Wahrscheinlichkeit aus, dass die Pixelwerte auf diese spezifische Weise verteilt werden, basierend auf der Wahrscheinlichkeitsverteilung, die es aus Mustern in Trainingsdaten gelernt hat.

Eine Wahrscheinlichkeitsdichtefunktion erfordert, dass die Wahrscheinlichkeit aller Möglichkeiten 1 ergibt, was in der Praxis häufig eine Normierungskonstante erfordert.

Die Berechnung dieser Normierungskonstante ist innerhalb eines kontinuierlichen Raums, wie bei Bildern oder Audio, - mit allen reellen und nicht nur natürlichen Zahlen -, nur mit unendlichem Rechenaufwand verbunden und dadurch nur theoretisch lösbar. Daher müssen clevere Umgehungslösungen entwickelt werden, wie zum Beispiel durch Approximation (Bergmann, 2026b).

Diese Approximation geschieht bei dem *Denoising Diffusion Probabilistic Model* (DDPM) in zwei Prozessschritten. Der Vorwärts Diffusionsprozess fügt nach einer Datenvorverarbeitung inkrementell einem Bild aus dem Trainingsdatensatz gaußsches Rauschen mithilfe einer Markov-Kette hinzu, um zukünftige Ergebnisse vorhersagen zu können (Vina, 2024).

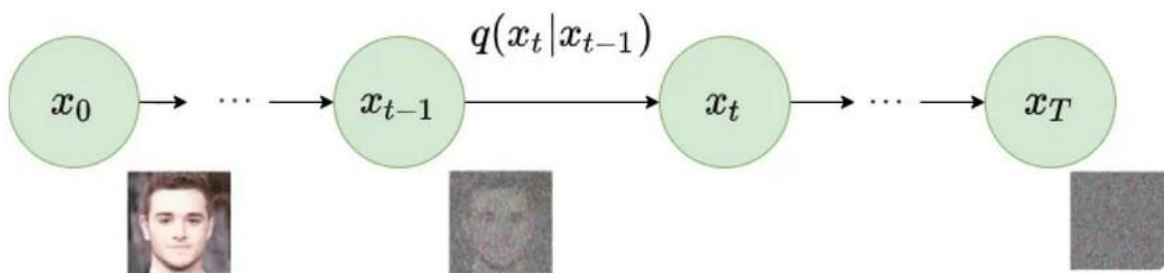


Abb. 4 – Vorwärts Diffusionsprozess (Vina, 2024).

Zu jedem Zeitpunkt t wird eine kleine Menge gaußsches Rauschen zu x_{t-1} hinzugefügt. x_0 ist der rauschfreie ursprüngliche Datenpunkt. x_T ist der Endzustand des Prozesses. Wenn T groß genug ist, konvergiert x_T zu einem gaußschen Rauschen. Jeder Vorwärtsschritt wird als $q(x_t | x_{t-1})$ definiert.

Das Rauschen wird immer von der wesentlichen Struktur des Originalbildes im vorherigen Schritt abgeleitet. Dadurch bleiben die wesentlichen Eigenschaften des Originalbildes erhalten und das Modell kann während des Rückdiffusionsprozesses die Muster und die Struktur des originalen Bildes sinnvoll erlernen. Der eigentliche maschinelle Lernprozess

findet in im Umkehr- bzw. Rückdiffusionsprozess statt, indem das Modell lernt, Gaußsches Rauschen in ein sauberes Bild umzuwandeln (Bergmann, 2026c).

Der Umkehr- oder Rückdiffusionsprozess bildet das verrauschte Bild schrittweise in seinen ursprünglichen Zustand zurück ab. Dabei kommt eine umgekehrte Markov-Kette zum Einsatz (Vina, 2024).

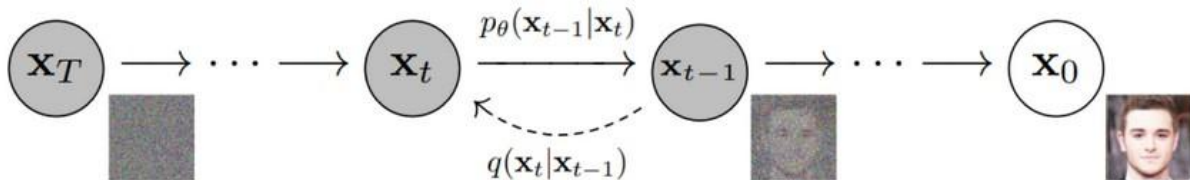


Abb. 5 – -Rückdiffusionsprozess (Vina, 2024).

Die wahre Umkehrung der Diffusion $q(x_{t-1}|x_t)$ ist in der Praxis nicht direkt lösbar, da der Rückwärtsprozess von der unbekannt Datenverteilung $q(x)$ abhängt. Daher wird $q(x)$ durch ein neuronales Netz $p_\theta(x)$ angenähert. Das Ziel ist es nun, die Gewichte des neuronalen Netzes so anzupassen, dass die modellierte Rückwärtsverteilung $p_\theta(x_{t-1}|x_t)$ in Übereinstimmung mit der tatsächlichen Rückwärtsverteilung $q(x_{t-1}|x_t)$ gebracht wird.

Statt x_{t-1} oder das im letzten Schritt hinzugefügte Rauschen vorauszusagen, wird stattdessen das gesamte Rauschen, dass in x_t hinzugefügt wurde, anhand der folgenden Darstellung des Vorwärtsprozesses geschätzt:

$$x_t = \sqrt{\bar{\alpha}_t} \cdot x_0 + \sqrt{1 - \bar{\alpha}_t} \cdot \varepsilon$$

Signalanteil Rauschanteil

Dabei ist x_0 das rauschfreie ursprüngliche Bild, während x_t das verrauschte Bild am Ende des Vorwärtsdiffusionsprozesses darstellt und ε , das im Zustand x_t enthaltene Grundrauschen. Der Parameter $\bar{\alpha}_t$ ergibt sich aus dem hinzugefügtem Rauschanteil pro Iteration des Vorwärtsprozesses. Für jeden einzelnen Diffusionsschritt gilt dabei: $\alpha_t = 1 - \beta_t$, wobei der Wert von β_t das Ausmaß des zusätzlichen Rauschens pro Schritt, innerhalb der Markov-Kette bestimmt. Dieser Parameter wird als Varianzplan (Noise Schedule) bezeichnet. Da jeder Diffusionsschritt nur vom vorherigen abhängt (Markov-Eigenschaft) und linear ist, kann man die Erhaltung des ursprünglichen Signals über mehrere Schritte hinweg durch die Multiplikation der einzelnen Signal-Erhalt-Faktoren α_t beschreiben. Das Produkt dieser Faktoren wird als $\bar{\alpha}_t$ definiert. Um das Rauschen nun schrittweise zu entfernen, um zu x_{t-1} zu gelangen, entfernen wir den Rauschanteil basierend auf dem Zustand des Varianzplans (Bergmann, 2026c).

Landschoot weist aufgrund dieser Arbeitsweise darauf hin, dass ein solches Modell dadurch nicht reproduktiv, sondern kreativ arbeitet, indem es aus Trainingsdaten völlig neue Klänge erzeugt (2024).

3 Songproduktion mit künstlicher Intelligenz

3.1 Einordnung von KI-Systemen in der Musik

Dieses Kapitel dient als Überblick und Einordnung von KI-Systemen, insbesondere von Suno, innerhalb der Musikindustrie, bevor in Kapitel 3.2 eine vertiefte Auseinandersetzung mit dem System innerhalb einer Songproduktion erfolgt.

3.1.1 Geschichte der KI in der Musikproduktion

Die Generierung von Musik in Verbindung mit maschineller Technologie ist kein neues Phänomen. Bereits lange vor dem Aufkommen moderner KI-Systeme wie Suno wurde versucht, musikalische Prozesse algorithmisch umzusetzen. In den 1950er Jahren komponierten Lejaren Holler und Leonard Isaacson die *Illiad Suite*, die als das erste Musikstück gilt, das mit kreativer Unterstützung von Algorithmen auf einem digitalen Computer erstellt wurde (Brown, O., 2021, S. 1). Erste Experimente in den 80er-Jahren ermöglichten mithilfe von Markov Ketten (vgl. Kapitel 2.2.5) die Analyse klassischer Musik und eine Voraussage anhand dieser Daten von zukünftigen Noten und Ereignissen. Außerdem experimentierten auch Künstler wie David Bowie in der Vergangenheit mit zufallsgenerierten Songtexten mithilfe des Cut-Up Prinzips (Bine, o. D.).

In naher Vergangenheit sorgte die Uraufführung von Beethovens ursprünglich unvollendeter zehnter Symphonie im Jahr 2021 für Aufsehen, die mithilfe von wenigen handschriftlichen Skizzen des Komponisten von einer künstlichen Intelligenz in Begleitung von Musikwissenschaftlern, Komponisten und Informatikern vollendet wurde (Oelze, 2021). Insgesamt war dieses Projekt jedoch sehr umfangreich und der Arrangeur Walter Werzowa musste teilweise über 200 unbrauchbare Vorschläge aussortieren, um schließlich auf ein zufriedenstellendes Ergebnis zu kommen. Der Chef-Programmierer räumte ein, dass die Komposition insgesamt sehr konventionell und kleinteilig und nur selten nach Beethoven klang. Insgesamt spricht der Auftraggeber Telekom von einem Projekt in dem Mensch und Maschine zusammengehen und in Kooperation arbeiten. Für den Dirigenten beweist das

Klangergebnis, dass die KI nur auf das reagieren kann, womit sie gefüttert wird, wobei sich die Frage stellt, ob sie etwas originär Neues schaffen kann (Hübert, 2021).

Die früheren Ansätze zeigen, dass man zur Erzeugung eines qualitativ hochwertigen Ergebnisses bei der Nutzung künstlicher Intelligenz vor allem auf Spezialisten wie Künstler, Mathematiker oder Musikwissenschaftler angewiesen war. Moderne KI-Systeme heute ermöglichen jedoch einen sehr niederschweligen Zugang zur Generierung musikalischer Inhalte. Programme wie Suno oder Udio markieren damit einen grundlegenden Wandel und besitzen das beträchtliche Potenzial, die Musikindustrie nachhaltig zu verändern.

3.1.2 Suno - Eine Einordnung

Nach der Definition von Wikipedia ist Suno AI ein generatives Programm zur Erstellung von Musik mit Künstlicher Intelligenz (2024). Das bedeutet, Suno kann ganze Songs mit nur einem einzigen Prompt in nur wenigen Sekunden oder Minuten erzeugen. Es läuft webbasiert und wird in drei verschiedenen Modellen angeboten, wobei monatlich sowie jährlich bezahlt werden kann.

Die Gratisversion von Suno bietet nur Zugriff auf die V4.5-all Version, ausschließlich 50 Credits pro Tag, was zehn Liedern entspricht, nur Standardfunktionen, keine kommerzielle Nutzung der generierten Musik und das Hochladen von bis zu einer Minute Audio.

Der Pro Plan für 8\$ im Monat bei jährlicher Abrechnung oder 10\$ im Monat bei monatlicher Abrechnung bietet bereits sehr viel mehr. Hier erhält man Zugriff auf das neueste v5 Modell von Suno, 2.500 Credits im Monat und kommerzielle Nutzungsrechte für die generierten Lieder. Hier können 8 Minuten Audiomaterial hochgeladen werden und weitere Vorteile sind möglich, die auf der Suno-Website aufgeschlüsselt werden (Suno - Pricing, o.D.).

Um Suno umfangreich und ausführlich testen zu können, wurde sich hier für die Premium Variante entschieden, die für 24\$ im Monat bei jährlicher Abrechnung, oder 30\$ im Monat bei monatlicher Abrechnung erhältlich ist. Der wichtigste Vorteil ist hier der Zugang zu *Suno Studio*, das umfangreiche Funktionen bietet, etwa *Remove FX* zum Entfernen von Effekten (Suno - Introducing Suno Studio, o.D.) oder *Cover* zum Erzeugen neu generierter Stems auf Basis einer in Suno Studio importierten Vorlage (Suno – How to Use Stem Cover in Studio, o. D.).

Von Thomas Foster, der sich in vielen YouTube-Videos und angebotenen Workshops ausführlich mit KI-Musik beschäftigt hat, wird Suno bereits in der Kapitelüberschrift Nummer Sieben seines Buches *Künstliche Intelligenz in der Musik- und Audioproduktion*, als nächste

Generation der KI-Musiktools beschrieben (Foster, 2024, S. 56). Er beschrieb die Musik, die von Suno generiert wird, bereits im Jahr 2024 als makellos, so als wäre sie von einem der besten amerikanischen Top-Produzenten bearbeitet worden. Ihm fielen die gedoppelten Stimmen, die perfekt eingestellten Effekte wie Delay (siehe Glossar) und Exciter (siehe Glossar) und der auf höchstem Niveau eingestellte Equalizer (siehe Glossar) und Kompressor (siehe Glossar) auf. Er beschreibt den Klang zwar nicht als 100% perfekt, aber als emotional mitreisend mit einer schönen und emotional tragenden Melodie (Foster, T. 2024, S. 48 & 49).

Bei der Produktion wurde sich für *Suno – AI Songs & Music* entschieden, da die Verbreitung und Nutzerakzeptanz dieses KI-Musikgenerators mit 187.000 Nutzerbewertungen und einer durchschnittlichen Bewertung von etwa 4,9 Sternen im App Store, die Zahlen von *Udio: AI Music Maker & Studio* mit etwa 1 600 Bewertungen und einem Durchschnitt von ca. 4,6 Sternen weit übertrifft. (vgl. App Store 2026).

Suno ist als KI-Anwendung, ähnlich wie andere Large Audio Models, aber nicht unumstritten. Die GEMA wirft Suno vor, bekannte Werke als Training für die KI zu nutzen, ohne die Künstler dafür zu bezahlen (Tagesschau, 2025). Laut GEMA selbst beginnt die Verhandlung hierzu am 09. März 2026 (Gema, o.D.). Der Vorwurf hierbei: Eine frappierende Ähnlichkeit zu GEMA lizenzierten Werken wie im Beispiel *Daddy Cool* von Boney M. aus dem Jahre 1976 und einem KI-Imitat von Suno. Das Beispiel hier (Abb. 6) ist exemplarisch aufgelistet. Noch weitere Vergleiche gibt es bei der GEMA selbst, im Artikel *Hörbeispiele: So kopiert Suno AI bekannte Songs* zu sehen und hören (Gema, o.D.).

Musical notation for the original song "Daddy Cool" by Boney M. The score is in F major, 4/4 time, with a tempo of 124. It shows the vocal line (Stimme) and the chorus (Chor). The lyrics are "She's cra-zy like a fool..." and "Wild a-bout Dad-dy Cool." The chords are Fm, Eb, Cm, and Fm.

Abb. 6.1 – Notenauszug aus dem Originalsong *Daddy Cool*. Zum Anhören des Originals: [Hier klicken](#)

Musical notation for the Suno AI version of "Daddy Cool". The score is in F major, 4/4 time, with a tempo of 125. It shows the vocal line (Stimme) and the chorus (Chor). The lyrics are "She's cra-zy like a fool..." and "Wild a-bout Dad-dy Cool...". The word "SUNO" is written above the score. The chords are Fm, Eb, C7, and Fm.

Abb. 6.2 – Notenauszug aus der Suno Version. Zum Anhören der Suno Version des Songs: [Hier klicken](#)

Die GEMA klagt hier nicht zum ersten Mal gegen ein großes KI-Unternehmen. Bereits im November 2025 gewann sie vor dem Landgericht München in einem fast einjährigen Verfahren gegen OpenAI. Das Gericht entschied, dass OpenAI urheberrechtlich geschützte

Texte nicht ohne entsprechende Lizenz verwenden darf. Das konkrete Ziel der Klagen ist die Einführung eines Lizenzmodells, für das KI-Anbieter bezahlen müssen, um ihre Software wie z.B. Suno oder ChatGPT mit entsprechenden Liedern trainieren zu dürfen (Zoronjic, S., 2025). Suno selbst gibt in dem bereits in Kapitel 2.2.1 erwähnten Verbraucherhinweis für kalifornische Nutzer an, dass für das Training der Modelle auch Inhalte verwendet werden, die unter geistigem Eigentum stehen. Dort heißt es: „Suno trains its models on datasets consisting of songs that may be from the public domain and/or that may be subject to intellectual property protection“ (Suno - CA AB 2013 Disclosure, o. D.).

3.1.3 Suno - Eine rechtliche Einordnung

Der Rechtsstreit um die Trainingsdaten zeigt, dass die juristische Sachlage von generativen KI-Systemen noch nicht abschließend geklärt ist. Trotzdem lohnt sich ein Blick in die Nutzungsbedingungen von Suno. Diese legen fest, dass Nutzer ausschließlich Inhalte hochladen dürfen, für die sie über die erforderlichen Rechte verfügen. Mit dem Upload bestätigen sie, dass keine Rechte Dritter verletzt werden und alle notwendigen Lizenzen vorliegen. Dies verlagert die rechtliche Verantwortung vollständig auf den Nutzer.

Wörtlich heißt es dort: „you have, or have obtained, all rights, licenses, consents, permissions, power and/or authority necessary to submit and use (and allow us to use) such Submission in connection with the Service, including for the purpose of generating your Output or Voice Model“ (Suno – Blog, 2026).

Ein weiterer zentraler Punkt betrifft die Nutzung der durch Suno erzeugten Inhalte. Laut den Nutzungsbedingungen ist die kommerzielle Nutzung abhängig vom gewählten Bezahlmodell. In den Terms of Service von Suno steht hierzu:

Subject to your compliance with these Terms of Service, if you are a user who has subscribed to the Pro or Premier paid tier of the Service, Suno hereby assigns to you all of its right, title and interest in and to any Output owned by Suno and generated from Submissions made by you through the Service during the term of your paid-tier subscription (2026).

Das bedeutet, der Nutzer erhält nur bei der bezahlten Version von Suno auch das Recht, mit dem generierten Output entsprechend Geld zu verdienen.

Suno behält sich aber auch generell weitgehende Rechte an den generierten Inhalten vor. So heißt es in den Terms of Service eine Stelle zuvor:

By using the Service or otherwise transmitting Submissions to us, you grant to Suno and our affiliates, successors, assigns, and designees a worldwide, non-exclusive, fully paid-up, sublicensable (directly and indirectly through multiple tiers), assignable, royalty-free, perpetual, irrevocable right and license to use, reproduce, store, modify, distribute, create derivative works based on, perform, display, communicate, transmit and otherwise make available any and all Content (in whole or in part) and any rights you may have in your Voice Model, in each case, in any media now known or hereafter developed, in connection with the provision, use, monetization, promotion, marketing, and improvement of our products and services, including the Service and the artificial intelligence and machine learning models related to the Service (2026).

Suno und alle mit Suno verbundenen Unternehmen besitzen also umfassende, weltweit geltende, zeitlich unbegrenzte und unwiderrufliche Rechte und Lizenzen an sämtlichen Inhalten, inklusive den Daten des Sprachmodells.

Insgesamt lässt sich also feststellen, dass sich Suno sehr weitreichende Rechte und Lizenzen zuschreibt, die in der Praxis die Durchsetzung der Interessen der Nutzer erschweren können. Die Anwaltskanzlei Rieck und Partner sagt dazu: „Auch wenn Ihnen bei Pro/Premier der Output *assigned* wird – die Suno-Lizenz verschwindet nicht automatisch. Du darfst monetarisieren – aber Suno darf mit demselben Material selbst auch eine Menge machen“ (Rieck, 2026). Laut Rieck ist das z. B. die Nutzung des Contents zur Verbesserung der KI-Modelle, zur Promotion oder Marketing, zur Monetarisierung, sowie zum Betrieb der Plattform. Außerdem verzichtet der Nutzer unwiderruflich auf seine Urheberpersönlichkeitsrechte, da laut den Terms of Service das Verfassen eines Prompts nicht ausreicht, um als Urheber eines Songs zu gelten (Rieck, 2026).

Laut § 2 Abs. 2 des Urheberrechtsgesetzes sind Werke persönliche geistige Schöpfungen. Der Schöpfer ist also eine natürliche Person. Reine KI-Outputs sind laut Rieck daher nicht klassisch urheberrechtlich geschützt. Eigene menschliche Beiträge können die Schutzfähigkeit durch das Urheberrecht aber erhöhen (2026).

Die praktische Arbeit mit KI-Systemen bewegt sich also in einem juristischen (vgl. Kapitel Kapitel 3.1.3) und künstlerischen Spannungsfeld (vgl. Kapitel 3.1.1) zwischen technischer Unterstützung und eigenem schöpferischem Beitrag. Es stellt sich hierbei also die Frage, ob und in welchem Umfang künstliche Intelligenz kreative und technische Prozesse unterstützen oder möglicherweise ersetzen kann. Aus den bisherigen Betrachtungen lässt sich aber schließen, dass KI-Systeme wie Suno weniger als autonome Produzenten verstanden werden sollten, sondern vielmehr als Werkzeuge, die gezielt durch den Nutzer

interpretiert und gesteuert werden, um eine ausreichende juristische Schöpfungshöhe und künstlerisch wertvolle Qualität in den Ergebnissen zu erzeugen.

Daher widmet sich das folgende Kapitel der Dokumentation und praktischen Umsetzung einer Suno-gestützten Songproduktion, die die Interaktion zwischen dem Nutzer und dem System Suno aufbereitet, darlegt und zusammenfasst.

3.2 Produktionstagebuch

„Etwas zu schaffen bedeutet, etwas ins Sein zu bringen, das vorher noch nicht da war.“

- Rick Rubin (2023, S. 13).

Die Songproduktion im Rahmen dieser Bachelorarbeit erfolgt unter wissenschaftlichen Rahmenbedingungen. Die zentralen Arbeitsschritte werden dokumentiert, die eingesetzten KI-Funktionen werden hinsichtlich ihrer Ergebnisse überprüft und im nachfolgenden Kapitel werden die gewonnenen Erkenntnisse einer genaueren Analyse unterzogen. Es werden im vorliegenden Produktionstagebuch nicht sämtliche Produktionsschritte im Detail dokumentiert. Der Fokus liegt darin, die Interaktion mit den eingesetzten KI-Funktionen zu beschreiben.

Gleichzeitig wurde im kreativen Prozess bewusst auf vorab festgelegte musikalische Vorgaben verzichtet. Trotz einer wissenschaftlichen Methodik sollte der Produktionsprozess offen gestaltet sein, um den kreativen Prozess nicht zu stören. Das Ziel ist es, insgesamt die KI in den Produktionsprozess miteinzubinden und eine Kollaboration zwischen KI-System und Produzenten zu erzeugen. Es geht weniger darum, einen gesamten Song ausschließlich mit KI zu erstellen (vgl. Kapitel 3.1.3), sondern diese als Werkzeug zu verwenden, um aufwendige Arbeitsschritte bestenfalls zu vereinfachen und abzukürzen. Außerdem soll die Produktion so die Möglichkeit bieten, die Grenzen der KI innerhalb des Prozesses herauszuarbeiten. Die aufgeführten Tage, in die dieses Kapitel untergliedert ist, stehen dabei für den Start in den jeweiligen Produktionsschritt. Es ist nicht garantiert, dass alle aufgeführten Schritte auch an diesem Tag stattgefunden haben. Als DAW kam *Cubase Pro 14* und *Suno* in der v5. Pro Version zum Einsatz. Die einzelnen Prozesse, die in *Suno* vorgenommen werden, werden innerhalb dieses Produktionstagebuchs so ausführlich wie nötig kommentiert und am Ende eines jeden Produktionsschritts unter dem Feld *Erfolge Beobachtungen* zusammengefasst.

3.2.1 Ab 13. Februar 2026

- Grundstruktur des Songs mit *Suno* als Inspirationsquelle -

Zu Beginn wurde eine Basslinie komponiert. Diese setzt sich aus den Grundtönen des Akkordschemas zusammen und bildet so das Fundament des Liedes. Dabei kam der sehr prägnante VST Synthesizer *Voltage* mit den Einstellungen wie in Abb. 7 zu sehen, zum Einsatz. Er ist ein Rauschgenerator, der wie von Steinberg selbst beschrieben, insbesondere für die Erzeugung von Synthesizer-Bassklängen eingesetzt werden kann (*Voltage*, o.D.). Der Klang des Synthesizers erinnert an EDM. Lars Bohn empfiehlt dafür ein rollendes

Sechzehntel-Pattern mit mehreren geschichteten Klängen. (2018, S. 257). Daher wurde noch eine Flöte hinzugefügt, die mit dem Granular Synthesizer Auron erzeugt wurde (Auron, o. D.).

- [Baseline.wav](#)



Abb. 7 - Voltage Synthesizers innerhalb der Engine Halion 7.1

Anschließend wurde das KI-System Suno genutzt, um die Grundlage der zuvor komponierten Bassline um einen Chor zu erweitern. Der Gedanke dabei war, dass ein solcher Chor nur schwer mit virtuellen Instrumenten zu bewerkstelligen ist. Hierzu wurde die Audiodatei [Baseline.wav](#) in die Plattform importiert und für den Einsatz von 10 Credits eine Cover-Variante generiert. Die Erstellung des Prompts, wie in Thomas Fosters YouTube Video ab 10:21min empfohlen (2025), erfolgte mithilfe von ChatGPT (siehe Abb. 8). Diese Vorgehensweise zieht sich durch den Verlauf der Produktion und wird nicht jedes Mal erneut explizit erwähnt. In einzelnen Fällen wurden die Prompts jedoch manuell angepasst, um gezielter in die Erstellung eingreifen zu können. Die eingesetzten Prompts sind im Verlauf dieser Arbeit ausgeführt.

Zu beachten ist, dass Suno standardmäßig immer zwei Kreationen aus einem Prompt erzeugt. Hier werden jedoch nicht immer alle generierten Versionen aufgeführt, sondern nur jene, die für die Nachvollziehbarkeit des Produktionsprozesses relevant sind.

- Das erzeugte Ergebnis von Suno hören sie [hier](#).

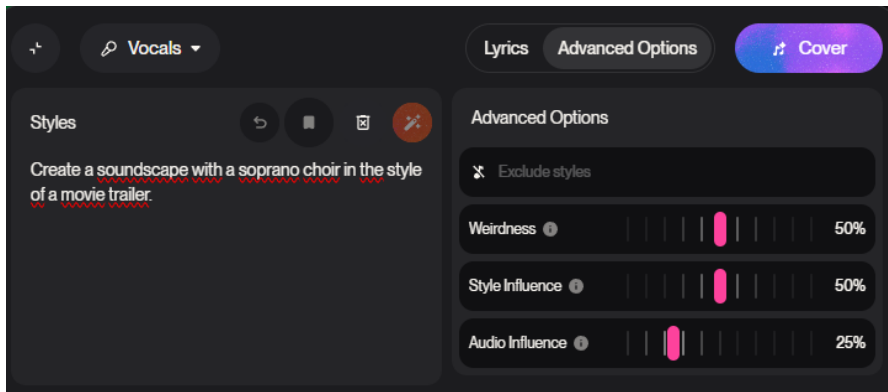


Abb. 8 - Eingegebener Prompt mit entsprechenden Werten.

Das Ergebnis klingt überzeugend und beinhaltet zu Beginn einen Impact-Sound mit einem Pad-Sound im Hintergrund. Dieser wird unterlegt mit einem melodispielanden Klavier, das im weiteren Verlauf von Streichern abgelöst wird. Dieser Instrumentenwechsel wird von einem Swoosh-sound (siehe Glossar) und einer Gesangsstimme eingeleitet.

In Suno ist es möglich, das generierte Audio mit der Funktion *Get Stems* (Suno – How to use: Stem extraction, o. D.) in die Einzelspuren bzw. Stems (siehe Glossar) zu separieren. Dieser Schritt wurde vorgenommen und anschließend wurde das Audiomaterial erneut separiert gehört. Bei der Analyse dieser Einzelspuren fällt in der Audiospur [Synth.wav](#) in Sekunde 0:02 eine Lautstärkeschwankung auf. Außerdem sind melodische Fehler der Streicher in 0:10 zu hören. Es fällt zudem auf, dass Suno die einzelnen Instrumente nicht vollständig trennen konnte. In dem [Synth.wav](#) Stem sind Klavier, teilweise die Streicher, sowie zu Teilen auch der Chor zu hören.

Insgesamt wurden die Anforderungen des Prompts jedoch zufriedenstellend erfüllt, sodass dieses Material zur Inspiration der weiteren Umsetzung dienen konnte. Auf dieser Basis wurde dem Arrangement das Instrument *The Verve* aus HALion 7 hinzugefügt. Dieses Instrument wird laut einer Pressemitteilung von Falko Haack in Verbindung mit Steinberg Media GmbH (2022) insbesondere für intime Filmmusik sowie warme, charaktervolle Klangfarben beworben.

Ergänzend wurde ein Synth.-Pad aus der HALion Sonic SE Collection verwendet, konkret das Preset *Raspy Voices*. Außerdem wurden Streicher (Studio Strings, o.D.) und eine weitere Basslinie mit dem VST Plug-in Electric bass (o.D.) eingespielt.

- Das Ergebnis ist [hier](#) zu hören.

- Erzeugen eines Chors -

Suno bietet die Möglichkeit, die eigene Stimme zu nutzen, um darauf basierende Ergebnisse zu generieren. Dieser Prozess wurde angewandt, um den Übergang des Klavierinstruments zu den Streichern mit einem Chor zu betonen. Dafür wurde die [bisherige Komposition](#) mit der [eigenen Stimme](#)*¹ als [Vorlage](#) für Suno kombiniert.

Die im Folgenden beschriebenen Einstellungen (siehe Abb. 9) führten zu dem dargestellten Ergebnis [Eins](#) und [Zwei](#).

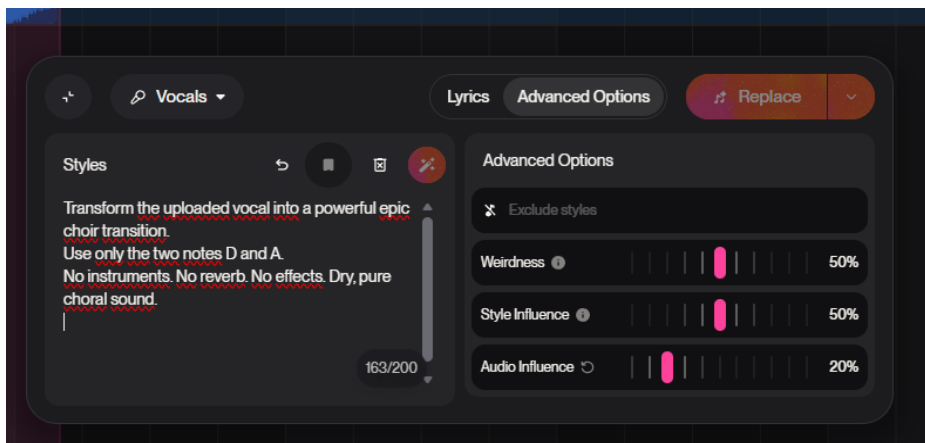


Abb. 9 - Eingegebener Prompt mit entsprechenden Werten.

Die Anweisungen des Prompts wurden hörbar in beiden Fällen nicht adäquat umgesetzt. Ergebnis Eins und Zwei weisen zwar einen mehrstimmigen Chor auf, jedoch sind diese mit Effekten wie Impact Sounds und Hallräumen versehen. Darüber hinaus wurde die Vorgabe, die Melodie auf zwei Noten (D und A) zu beschränken, nicht erfüllt.

In einem weiteren Versuch wurden die Parameter der Creative Sliders (Suno – How to Use: Creative Sliders) angepasst. Die Regler *Style Influence* und *Audio Influence* wurden, wie in Abbildung 10 dargestellt, erhöht, um den Einfluss der Vorlage zu verstärken. Zusätzlich wurde der *Weirdness*-Regler auf 20 % reduziert, um das Ergebnis weniger experimentell erklingen zu lassen. Außerdem wurde der Parameter *Vocals* in *Backing Vocals* geändert, da dies dem angestrebten Ergebnis entsprechend passender erschien.

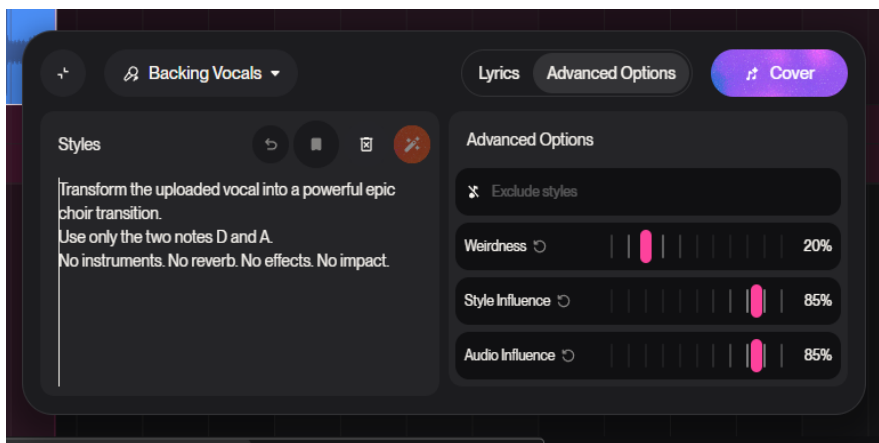


Abb. 10 - Eingegebener Prompt mit entsprechenden Werten.

Folgende Ergebnisse wurden von Suno generiert:

- Ergebnis 1:
Die [erste Spur](#) ist mehrstimmig generiert. Zu Beginn weist sie keinen hörbaren Impact auf, im späteren Verlauf ist jedoch einer zu hören. Anfangs werden ausschließlich die Töne D und A realisiert; im weiteren Verlauf erscheinen diese offenbar oktaviert (D + D sowie A + A). Gegen Ende der Spur ist ein hörbares Artefakt und eine seltsame Lautstärkeschwankung festzustellen. Auch die Atemgeräusche in 0:05 wirken unnatürlich. Am Ende wirkt die Spur abrupt abgeschnitten.
- Spur 2:
In der [zweiten Spur](#) sind keine Artefakte wahrnehmbar. Der Impact tritt zweimal deutlich in Erscheinung. Am Ende wirkt auch diese Spur wie abrupt abgeschnitten.

Es wurde sich dafür entschieden, die ersten zwei generierten Takte von Spur 1 in das Projekt einzubauen.

Außerdem zu hören: Für die Generierung innerhalb von *Suno* wurde ein vier Takte großer Abschnitt ausgewählt. In der Vorlage singe ich jedoch nur in den letzten beiden der vier ausgewählten Takte. Trotz dieser klaren Aufteilung generierte *Suno* auch Material für die ersten beiden Takte, also über alle vier ausgewählten Takte hinweg.

Außerdem notwendig war eine zeitliche Anpassung mithilfe der in Cubase integrierten Funktion *Audio Warp* (siehe Abb. 11). Da die ersten zwei generierten Takte rhythmisch nicht zum Arrangement passten.

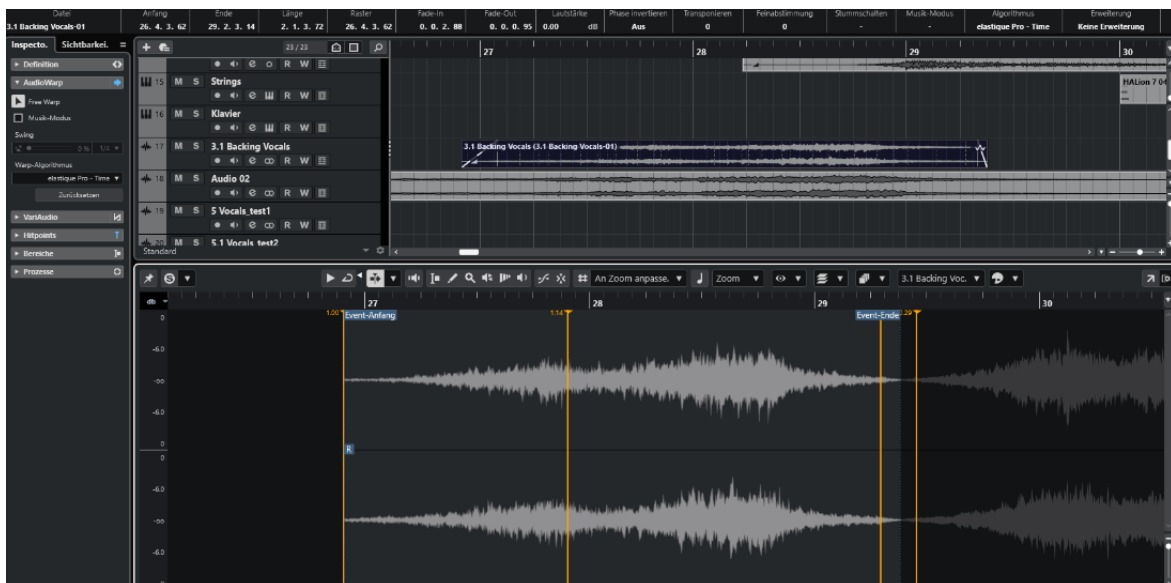


Abb. 11 – Audio-Wrap des entsprechenden Audiomaterials von *Suno* generiert.

Der chorale Übergang, der zur Verfügung stand, sollte nun – analog zum [ersten](#) von Suno generierten Beispiel, welches, wie erwähnt als Inspirationsquelle dient – um einen hohen Sopranchor ergänzt werden.

Hierzu wurde der Master-Export des bestehenden *Cubase*-Projekts erneut, in Kombination mit einer eingesungenen Vorlage, als [Referenzmaterial](#) für *Suno* verwendet. Dabei wurde kein Wert auf Intonationsgenauigkeit der Gesangsaufnahme gelegt und die konkret gewünschten Töne wurden erneut zusätzlich im Prompt spezifiziert (siehe Abb. 12).

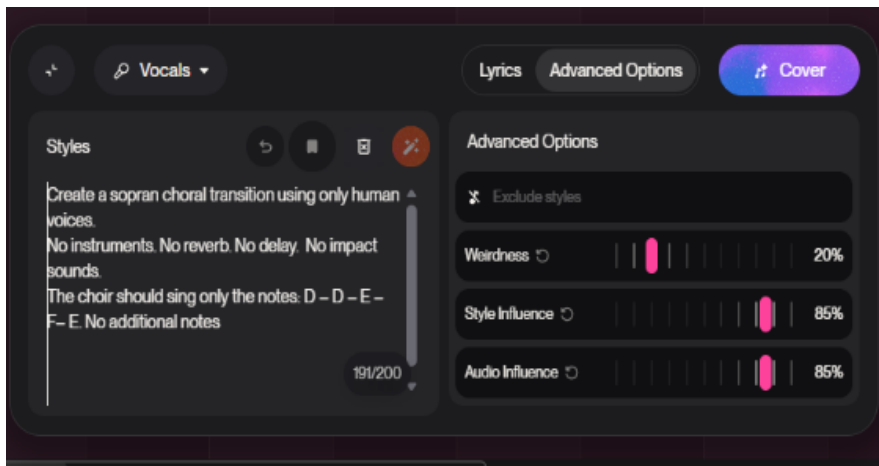


Abb. 12 - Eingegebener Prompt mit entsprechenden Werten.

In beiden Durchläufen ([Eins](#) und [Zwei](#)) entsprach das Ergebnis nicht vollständig den formulierten Vorgaben. Erneut wurden keine trockenen Vocals erzeugt, und es entstand der Eindruck, dass sich nicht an dem eingesungenen Material orientiert wurde, sondern das Klavier als Referenz diente.

Außerdem wurde eine Abfolge der Noten D, A, B, G, A, F, G, E gespielt, was der Vorgabe, nur die Noten D-D-E-F zu verwenden, nicht entsprach.

Ungeachtet dieser Abweichungen wurde das zweite Ergebnis subjektiv als klanglich überzeugender bewertet und in das bestehende Projekt integriert. Ziel war es daher, eine Variante mit denselben Tönen zu erzeugen, ohne den am Ende auftretenden Klimax. Die gewünschte Passage sollte ab dem dritten Takt beziehungsweise ab 0:05 einsetzen und sich klanglich an den ersten zwei Takten von Ergebnis zwei orientieren.

Als Referenzmaterial für Suno wurde eine Audiodatei erstellt, in der die ersten beiden Takte des Chors aus Ergebnis [zwei](#) mit dem Klavier kombiniert wurden. Auf dieser [Grundlage](#) wurde ein neuer Prompt mit den entsprechenden Parametereinstellungen in Suno eingegeben (siehe Abb. 13).

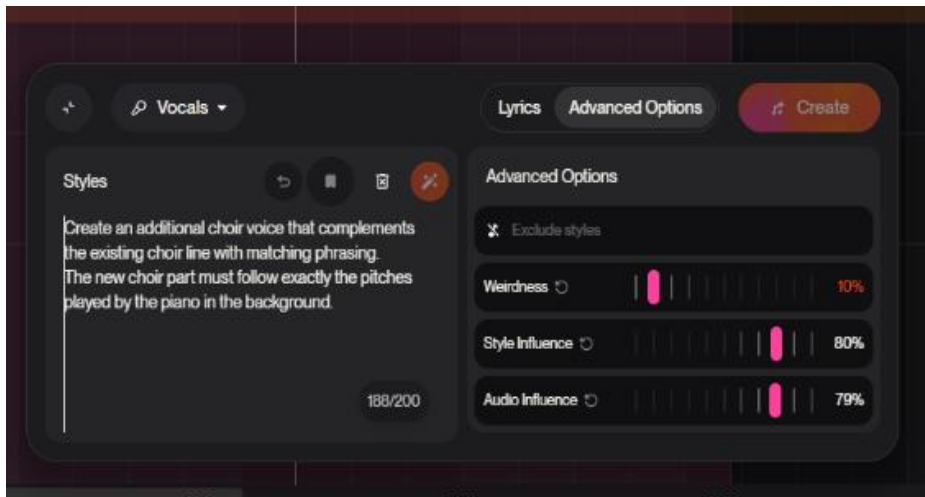


Abb. 13 -
Eingegebener Prompt
mit entsprechenden
Werten.

Dies brachte keine zufriedenstellenden Ergebnisse:

- [Ergebnis 1](#)
- [Ergebnis 2](#)

Wie zu hören weicht [Test 1](#) klanglich weit von der Vorlage ab. In [Test 2](#) setzt ab dem dritten Takt eine Bassstimme ein.

- Versuch der Reproduzierbarkeit -

Deshalb wurde der Prozess mehrmals mit angepassten Prompts und veränderten Parametereinstellungen erneut durchgeführt. Es wurde auch ausschließlich der Chor aus Ergebnis Zwei, ohne Klavier, als Referenzmaterial angegeben. Es gelang jedoch auch nach mehreren Durchläufen nicht, den Chor aus Ergebnis [Zwei](#) für den Einsatz in Takt 3 und 4 klanglich zu reproduzieren. In allen Fällen klang der Gesang nicht wie in Ergebnis [Zwei](#). Da die generierten Ergebnisse nicht den gewünschten Erfolg erzielten, wird auf eine Auflistung der einzelnen verwendeten Prompts und Parameteranpassungen an dieser Stelle verzichtet. Es wurde anschließend eine andere Methodik angewandt:

Da sich die Vorgehensweise, Suno die eigene Stimme als klangliche Referenz bereitzustellen (siehe Abb. 3), als zielführend erwiesen hatte, wurde dieses Verfahren erneut angewandt. Zu diesem Zweck wurden zwei neue Sprachaufnahmen ([Rec 1](#) und [Rec 2](#)) erstellt und in das System eingespeist.

Ziel war es, einen Chor zu generieren, der sich in den ersten zwölf Takten alle vier Takte wiederholt und sich ab Takt 12 dynamisch steigert, um einen Übergang zu kreieren. Beide Aufnahmen wurden Suno als Referenzmaterial übergeben, woraufhin folgende Ergebnisse erzeugt wurden.

- [Vocal 1](#)
- [Vocal 2](#) (Audio startet ab 0:35).

Die Resultate [Vocal 1](#) und [Vocal 2](#) wurden in das bestehende Projekt eingebunden, sodass sich daraus folgender Zwischenstand ergab.

- [Zwischenstand](#) - 13.02.26

Aus dem bisherigen Arbeitsprozess lassen sich folgende Beobachtungen ableiten.

Erfolgte Beobachtungen:

- Im Unterschied zu ChatGPT als Sprachmodell reagiert Suno offenbar weniger präzise auf textliche Stilvorgaben. Anweisungen in den Prompts wie *no reverb and effects* sowie konkrete Anforderungen an Tonhöhen oder Melodieverläufe wurden nicht zuverlässig umgesetzt.
- Demgegenüber scheinen Angaben zu Instrumentierung oder Klangcharakteristik – beispielsweise *Sopran Choir*, *Choir* oder *Movie Style* – überwiegend berücksichtigt zu werden.
- Eine tongetreue Reproduktion eines zuvor generierten Ergebnisses war trotz identischer Parameter und entsprechendem Referenzmaterial nicht möglich.
- Die Anpassung der Parameter *Weirdness*, *Style Influence* und *Audio Influence* führte hingegen zu einer erheblichen Veränderung des Outputs in die vom Nutzer intendierte Richtung.
- Verzerrtes Audiomaterial kann interpretiert werden.
- Tonal falsch eingesungenes Material scheint innerhalb eines gewissen Rahmens richtig interpretiert zu werden.

3.3.2 Ab 19. Februar 2026

- Hinzufügen von Streichern und Blechblasinstrumenten -

Die mit *Get Stems*, aus dem als Inspirationsquelle dienenden [Prompt](#) vom 13. Februar generierte [Streicherspür](#) wurde als gestalterische Referenz genutzt, um innerhalb der DAW eigene MIDI-Streicher einzuspielen. Insgesamt entstanden acht separate Streicher-Spuren.

Die zuvor von Suno erzeugten Streicher wurden anschließend vollständig aus dem Projekt entfernt und durch die eigenständig produzierten MIDI-Streicher ersetzt.

- [Liedpassage](#) mit Midi-Streichern

Die Streicher sollten dabei als melodisches Fundament fungieren und die Hauptmelodieführung durch Blechblasinstrumente ergänzt werden. Zu diesem Zweck wurde erneut Suno verwendet, um aus einem Prompt (siehe Abb. 14) und der [Liedpassage](#) mit Midi-Streichern, zwei Versionen mit Bläsern zu generieren. Das erste Ergebnis ist hier exemplarisch dargestellt.

- [Version 1](#) Suno Bläser mit Midi-Streicher

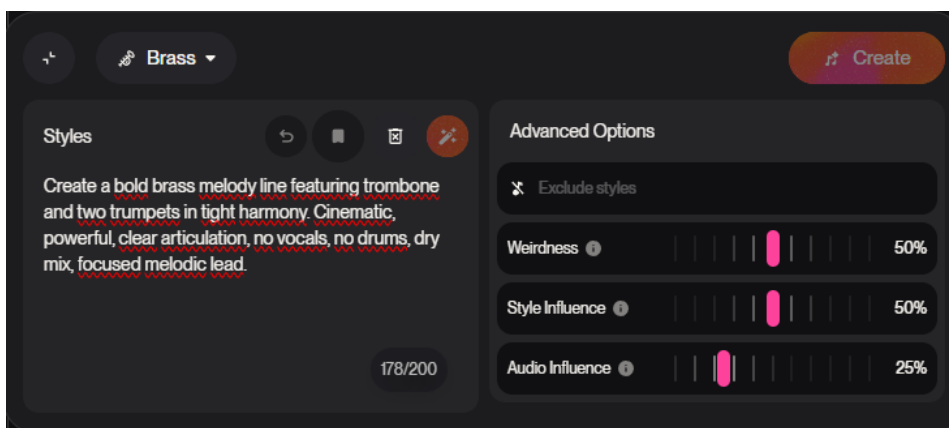


Abb. 14 -
Eingegebener Prompt
mit entsprechenden
Werten.

Da beide Ergebnisse deutlich von der Streichermelodie abwichen, passte ich den Regler *Audio Influence* von 25 % auf 70 % an. Dies führte zu einer deutlich stärkeren Übernahme der vorgegebenen Streicher-Melodie. Auffällig war jedoch erneut ein Crescendo sowie das Hinzufügen einer zusätzlichen Bläsersektion ab Takt 5.

- [Version 1.1](#) Suno Bläser mit Midi-Streicher

Um Version 1.1 präziser weiterbearbeiten zu können, wurde die berechnete Bläser-Stereodatei mithilfe der Funktion *Get Stems* in Suno in ihre Einzelbestandteile aufgeteilt. Dabei entstanden drei separate Spuren:

1. Eine [Brass-Synth](#) Spur mit deutlichen Hall- und Reverb-Effekten sowie synthetischen Klanganteilen, insbesondere im Schlussteil.
2. Eine [Brass \(fx\)](#)-Spur, mit klanglicher Ähnlichkeit zu einem Fagott, mit viel Luftanteil im Klang und hörbaren Shaker-Elementen im ersten Abschnitt.
3. Eine Haupt-[Brass Spur](#), in der alle Blechblasinstrumente in einer Stereo-Datei zusammengefasst sind.

Die Brass-Spur wurde weitgehend von den anderen Anteilen getrennt, wobei ein Restanteil von Hall und Effekten weiterhin hörbar blieb. Eine differenzierte Trennung der einzelnen Instrumente (vermutlich Tuba, Trompete und Posaune) war hingegen nicht möglich.

Die Haupt-Brass-Spur sowie die FX-Spur wurden in das bestehende Projekt integriert. In der Haupt-Brass-Spur wurde der letzte Ton (C) des vierten Taktes geschnitten und wiederholt eingefügt, wodurch die Melodie leicht verändert wurde. Außerdem wurde der erste Ton D mit dem gleichen Ton von einer anderen Stelle ersetzt, da dieser ein deutlich hörbares Artefakt aufwies, das mit *Spectral Layers* (siehe Glossar) nicht entfernt werden konnte. Darüber hinaus wurde die Spur mittels eines dynamischen Equalizers im Frequenzbereich von 0–1000 Hz moderat angehoben, um die tiefen Anteile zu erhöhen.

Besonders positiv fiel die von Suno erzeugte Betonung auf den ersten und dritten Schlag auf, sowie die gelungene Phrasierung der einzelnen Blasinstrumente.

Nun galt es herauszufinden, ob eine vergleichbare authentische Phrasierung auch auf Grundlage der selbst produzierten MIDI-Streicher realisierbar ist.

Hierzu wurde ein Stereomix der ohne KI erzeugten [Midi-Streicher](#) als Vorlage in Suno eingespeist. Auf Basis des Prompts in Abbildung 15 wurden zwei Ergebnisse generiert. Das klanglich überzeugendere Ergebnis ist hier ausgeführt:

- [Ergebnis 2](#) von Suno generiert

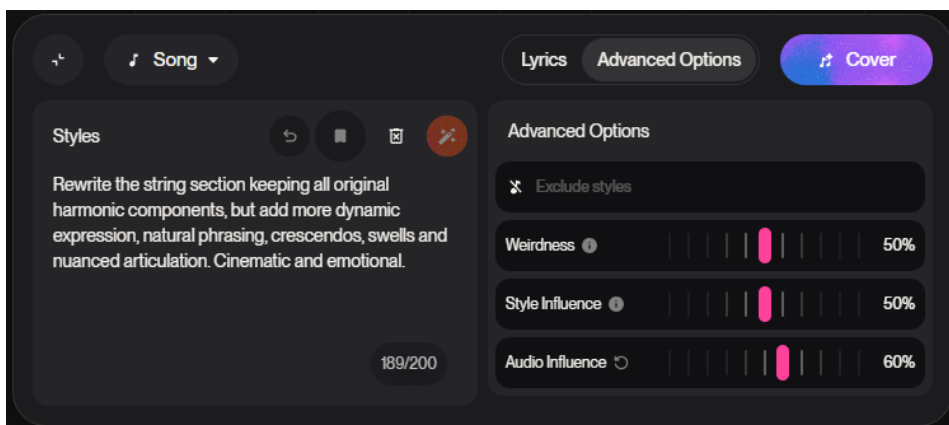


Abb. 15 -
Eingegebener Prompt
mit entsprechenden
Werten.

Die KI-generierten Streicher überzeugen in puncto Dynamik, Transparenz und Offenheit gegenüber den MIDI-Streichern deutlich. Während die MIDI-Streicher insgesamt kompakter und dadurch produziert wirken.

Anzumerken ist aber, dass die MIDI-Streicher in der vorliegenden Fassung nicht vollständig ausgearbeitet sind, da keine Automationen implementiert wurden und in den KI-Streichern erneut unerwünschte Elemente enthalten sind. Zu hören ist ein Impact im Mittelteil, ein Beckenschlag zu Beginn sowie kleinere Artefakte ab Minute 0:12, die an Shaker-Geräusche erinnern.

Wie im oben beschriebenen Vorgehen wurde das Material in Stems separiert. Die FX-Spur enthielt dabei sowohl den Impact, das Becken und die genannten Artefakte, sodass diese isoliert und entfernt werden konnten. Die verbleibenden Spuren wurden in das Cubase-Projekt integriert und ersetzt dort die ursprünglichen MIDI-Streicher.

Im Anschluss wurde untersucht, ob die Komplexität des vorgegebenen Materials einen Einfluss auf die Qualität der generierten Ergebnisse hat. Hierfür wurden aus dem ursprünglich achtspurigen MIDI-Streichersatz ausschließlich die zwei führenden Melodiespuren extrahiert und der zuvor verwendete Prompt erneut angewendet.

- [Vorlage für Suno](#)
- [Ergebnis 1 von Suno](#)
- [Ergebnis 2 von Suno](#)

Es zeigte sich, dass Suno die Melodie grundsätzlich umsetzt, die tiefen Streicher jedoch freier agierten als in der detaillierten Vorlage. Der Beginn von [Ergebnis 1](#) ist abgeschnitten, zudem erscheint am Ende eine melodische Figur in den führenden Streichern, die in der Vorlage nicht enthalten war. [Ergebnis 2](#) besitzt eine legato-geführte Begleitung, die erst in der Mitte des ersten Taktes einsetzt. Insgesamt wirken die von Suno generierten Resultate musikalisch schlüssig; fehlende Begleitstimmen werden vom KI-System sinnvoll ergänzt. Die erstellten Spuren [Ergebnis 1](#) und [Ergebnis 2](#) wurden dem Cubase-Projekt jedoch nicht hinzugefügt.

Insgesamt wurde der Song um mehrere zusätzliche MIDI-Spuren erweitert, darunter ein mit *Guitar Rig 7* (Native Instruments GmbH, 2026) verzerrtes und mit Delay versehenes Piano sowie ein Beckenschlag. Darüber hinaus wurden einzelne Spuren aus dem [Zwischenstand](#) vom 13.02. durch Duplizierung erneut verwendet, wodurch der Song insgesamt ausgebaut und verlängert wurde.

Nach diesem kleinen Versuch sollte abschließend der Streicher- und Brass-Abschnitt durch ein melodielspielendes Saxophon ergänzt werden. Dieses wurde zunächst als MIDI-Spur eingespielt und anschließend für einen Prompt (Abb. 16) in Suno verwendet. In Kombination mit den übrigen Spuren klingt das Midi-Saxophon wie folgt:

- [Midi-Saxophon](#)

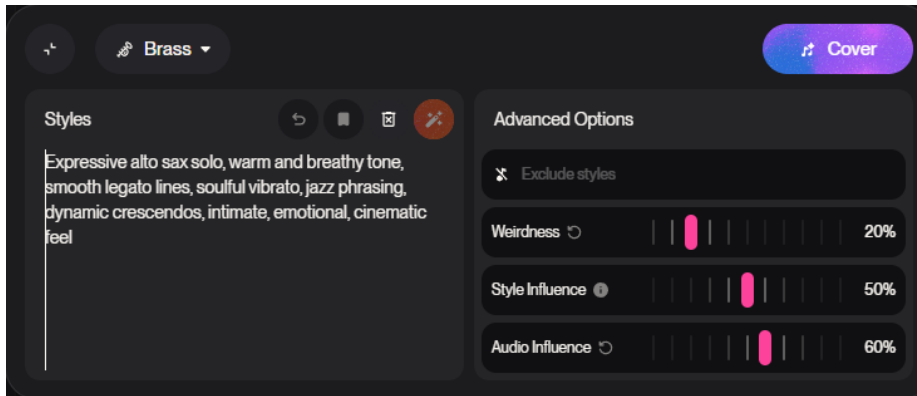


Abb. 16 - Eingegebener Prompt mit entsprechenden Werten.

Die Ergebnisse waren nicht überzeugend, da zusätzliche Phrasierungen integriert wurden und die tiefe Lage des Saxophons nicht übernommen wurde. Zudem waren Klavieranteile sowie in einer Version Gesang zu hören.

Durch die Erhöhung der Parameter *Style Influence* und *Audio Influence* auf 80 % verbesserten sich die Resultate deutlich, allerdings war weiterhin ein Rhodes-Klavier enthalten. Dieses wurde erneut mithilfe der Funktion *Get Stems* entfernt, sodass das Saxophon fast isoliert vorlag. Der starke Hallanteil wurde abschließend mit der Suno-Funktion *Remove FX* reduziert, wodurch das Signal trockener vorlag.

- [Saxophon separate](#) mit Hall
- [Saxophon separate](#) ohne Hall

Nach der Entfernung des Hallanteils wurde die Saxophonspur [Saxophon separate](#) ohne Hall in das Projekt eingefügt. Zur Erweiterung des Saxophonparts wurde die Funktion VARI-Audio (siehe Glossar) eingesetzt. Hierbei wurden die Takte 9–12 des höheren Saxophonsegments kopiert, die Schlussnoten mithilfe von VARI-Audio modifiziert und das Material zusätzlich in den Takten 13–16 eingefügt. Auf diese Weise entstand eine zweistimmige Saxophonführung mit einem hohen und einem tiefen geführten Part (siehe Abb. 17).

In den letzten vier Takten wurde das Arrangement bewusst reduziert, um eine Steigerung hin zum Höhepunkt zu erzeugen. Außerdem wurde der Bass leicht angepasst und ein weiteres Saxophon als Panorama-Effekt, sowie ein Synth-Effekt hinzugefügt.

Das klangliche Gesamtbild stellt sich wie folgt dar:

- [Full Saxophon](#)-Part.

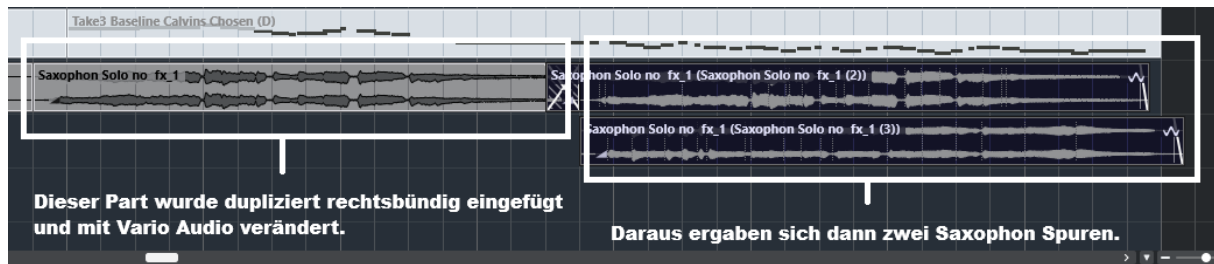


Abb. 17 – Saxophon-Arrangement in Cubase Pro 14

Es ist deutlich zu hören, dass das von Suno generierte Saxophon im Vergleich zum MIDI-generierten Saxofon natürliche An- und Abschwellungen, Atemgeräusche und leichte Intonationsabweichungen aufweist. Dadurch klingt es authentischer, weshalb die MIDI-Version im weiteren Produktionsverlauf nicht beibehalten wurde.

Im nächsten Schritt sollten mithilfe von Suno fließende Übergänge generiert werden, beispielsweise durch Synthesizer, Percussion, Drums oder Swooshes. Der gesamte Song wurde zunächst separat gerendert und anschließend in Suno Studio weiterverarbeitet.

Durch den Einsatz unterschiedlicher Prompts wurden über die gesamte Länge des Songs verschiedene Instrumentalvarianten – darunter Piano-, Synth- und Percussion-Elemente – erzeugt. Geeignete Passagen wurden selektiv ausgeschnitten und dem bestehenden Arrangement hinzugefügt. Nicht alle von Suno generierten Ergebnisse fanden im Song Verwendung. Insgesamt wurden vier zusätzliche Spuren integriert. Eine Pianospur sowie drei Percussion- beziehungsweise Effekts Spuren. Das bisherige Ergebnis nach Integration der generierten Elemente ist hier zu hören:

- [Zwischenstand](#) - 19.02

Erfolgte Beobachtungen:

- Wird die Funktion *Audio influence* erhöht, führt das dazu, dass die vorgegebene Melodie genauer übernommen wird.
- Die Funktion *Get Stems* stellt ein zentrales Werkzeug dar, um unerwünschte Elemente gezielt zu isolieren und danach entfernen zu können. Eine vollständige instrumentenspezifische Trennung innerhalb einer Spur ist aber nicht möglich.

- KI-generierte Instrumente zeigen eine ausgeprägte dynamische Varianz und authentische Phrasierung, wodurch sie im Vergleich zu nicht automatisierten MIDI-Spuren natürlicher wirken.
- Vereinfachte Referenzvorlagen führen ebenfalls zu qualitativ guten Ergebnissen, mit mehr Varianz in der Spielweise und Komplexität.
- Die Einbindung der generierten Audiospuren machte häufig eine nachträgliche Bearbeitung, wie durch VARI-Audio, Schnitt oder Dynamikanpassungen, erforderlich, wodurch die Audibearbeitung im Vergleich zur reinen MIDI-Bearbeitung stärker in den Fokus rückt.

3.2.3 Ab 25. Februar 2026

- Den Song mithilfe der Extend Funktion erweitern-

Das Ziel hier ist es, den Song nun sinnvoll zu erweitern und eine Brücke zu einer neuen Variation der Eingangsmelodie zu schaffen, um ihn anschließend beenden zu können. Dafür wurde der Song in seinem [aktuellen Zustand](#) in Suno hochgeladen und es wurde die *Extend-Funktion* verwendet. Dabei generiert Suno auf Basis des bestehenden Materials eine mögliche Fortsetzung des Songs oder bringt diesen sinnvoll zu Ende (Suno – How do i make my Song longer?, o. D.).

Es wurden zwei Varianten erzeugt:

- [Version 1](#) beginnt mit Streichern, die die Hauptmelodie spielen, und hinzugefügten Drums. Danach folgt ein Übergang, der durch eine Bassline und das Herausnehmen der Drums erzeugt wird und anschließend in einen Chor übergeht, der wiederum von einem Xylophonähnlichen Klang abgelöst wird. Daraufhin steht der Gesang wieder im Vordergrund, begleitet von präsenten Drums. Die Hauptmelodie wird anschließend von einem mit Delay versehenen Synthesizer bis zum Ende weitergeführt. Am Ende sind *vocal chops* (siehe Glossar) mit Drums in einem Art Höhepunkt zu hören, der anschließend von einem ruhigen Ende abgelöst wird.
- [Version 2](#) startet ebenfalls mit Streichern und Drums, wobei der Einstieg jedoch sehr holprig ist und so für den fertigen Song unbrauchbar. Danach geht die Melodie auch hier in einen Chor über, der die Hauptmelodie singt. Der Chor singt jedoch nicht durchgehend sauber und hat einige kleine Aussetzer. Zum Beispiel geht er bei 0:25 im Hall unter und ist auch bei 0:31 kaum hörbar, bis er bei 0:34 wieder deutlicher einsetzt. Danach wird die Melodie erneut vom Chor gesungen und von präsenten Drumlines begleitet. Die Drums wirken in dieser Version stimmiger, da sie weniger

basslastig sind und eher an orchestrale Trommeln erinnern, wie sie in Film und vor allem Trailer-musik häufig vorkommen. Besonders interessant ist die Passage, in der die Melodie leicht verändert von *Backing vocal chops* und Kontrabass gespielt wird (siehe 0:55). Diese Stelle leitet überzeugend in einen möglichen neuen Abschnitt des Stücks ein. Am Ende ist erneut eine Steigerung des Tracks zu hören, bei der mehrere Instrumente, Chor, Drums und Streicher gleichzeitig einsetzen und so etwas wie ein mögliches Finale einleiten. Die letzten vier Takte sind, wie in [Extend-Version 1](#), wieder ruhiger und bestehen nur aus einem hörbaren Glockenspiel sowie einem leisen Chor im Hintergrund.

Diese Funktion zeigt, dass der Song nicht einfach nur geloopt und dadurch verlängert wird. Stattdessen fügt Suno tatsächlich neue Instrumente und musikalische Ideen hinzu. Durch das Hinzufügen von Drums und Vocal-Chops rutscht der Song automatisch etwas stärker in eine elektronische Richtung. Beide hinzugefügten Versionen folgen dabei einer ähnlichen kompositorischen Struktur, beginnend mit der Hauptmelodie, unterlegt mit Schlagzeug/Drums, übergehend in einen ruhigeren Teil, der anschließend den Höhepunkt einleitet und von einem wieder ruhigen Ende abgelöst wird.

- Einbinden von Midi Extraktion über Suno -

Anschließend wurden die *Extend*-Versionen 1 und 2 innerhalb von Suno wieder in einzelne Stems separiert, um mit diesen weiterarbeiten zu können. Darüber hinaus ermöglicht Suno nicht nur den Export der Stems als .wav- oder .mp3-Datei, sondern auch als MIDI-Datei. Diese Funktion wurde bei Version 2 genutzt, da die dort hinzugefügte, leicht veränderte Grundmelodie (siehe 0:55) überzeugte und durch einen Export in MIDI noch gezielter veränderbar ist. Das Resultat waren jeweils neun Stem-Spuren in MIDI, wie in Audio, bei [Extend-Version 2](#) und sechs Stem-Spuren in Audio bei [Extend-Version 1](#).

Beim Export der MIDI-Dateien stellte sich jedoch heraus, dass Suno die Dateien nicht in der ursprünglich vorgegebenen und innerhalb von Suno im Explorer korrekt erkannten Geschwindigkeit von 110 BPM exportierte (siehe Abb. 18). Als Beispiele dienen hier die .wav-Audiodateien des *Background-Chors* der *Extend-Version 2* und danach die daraus erzeugten MIDI-Daten derselben Passage, die mit verschiedenen VST-Instrumenten und zugespieldem Metronom mit der Audiodatei des *Background-Chors* verglichen wird.

Beispiel 1: Soundbeispiel einer schnelleren Passage der *Background Vocals*-Audio Spur und die dazu generierte MIDI-Spur mit einer Flöte als *VST-Instrument*, um den Klang der Vocals nachzubilden.

[- Backing Vocals audio](#)

[- Flöte MIDI](#)

Beispiel 2: Im Vergleich dazu eine langsamere, andere Passage. Zu hören sind die *Backing-vocals* zunächst mit einem Metronom und anschließend die erzeugten MIDI-Spuren als Streicherstimme und Piano, mittels VST instrumentalisiert

[- Backing Vocals audio 2](#)

[- Streicher MIDI](#)

[- Piano MIDI](#)

In Beispiel 1: Der Takt und die Geschwindigkeit werden weitgehend korrekt eingehalten, allerdings sind in der Audio-Version hörbare Artefakte vorhanden.

In Beispiel 2: In diesem Beispiel wird deutlich, dass die MIDI-Noten (siehe Abb. 18) nicht exakt auf dem Grid liegen. Dies scheint mit der Geschwindigkeit der Vorlage zusammenzuhängen (vgl. Beispiel 1).

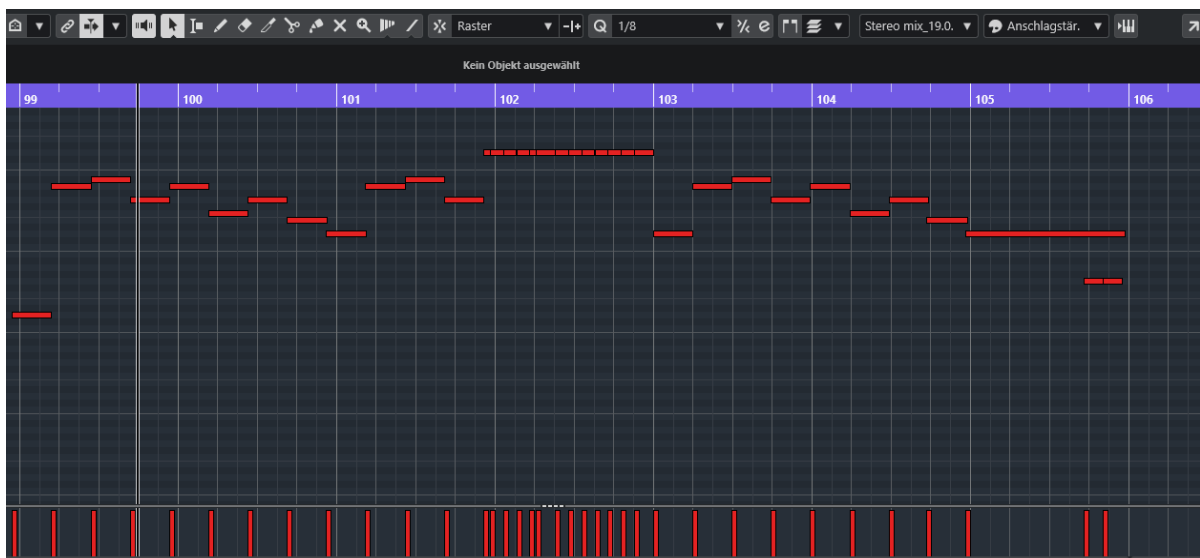


Abb. 18 – Die generierten Midi Noten weichen in den ersten vier Takten (99, 100, 101 und 102), wie hier zu sehen, deutlich (ca. um eine 16tel) vom eingestellten Grid ab.

Aufgrund der langsamen Attack-Phase der Streicher wirken Takt 99-102 dennoch nahezu im Takt, während Takt 103 bis 105 unrhythmisch wirken. Bei der Verwendung eines Instruments mit schnellerem Anschlag, vgl. Piano, zeigt sich hingegen ein anderes Ergebnis. Die ersten vier Takte wirken rhythmisch ungenau. Die letzten drei bis vier Takte hingegen, in denen die MIDI-Noten stärker am Grid ausgerichtet sind, klingen rhythmisch deutlich präziser. Dies scheint darauf zurückzuführen zu sein, dass sich auch die zugrunde liegende Vorlage [Backing vocals audio 2](#) ab diesem Takt klanglich verändert und stärker akzentuiert ist und weniger flächig klingt.

Die Vorschläge von Suno wurden nun als inspirative Vorlage genutzt, um einen durch MIDI und *VST-Instrumente* erzeugte Erweiterung zu erzeugen.

- [MIDI Erweiterung](#)

- Rückführung des MIDI-Summensignals auf einzelne Spuren mittels Suno -

Der Masterkanal der MIDI-Erweiterung wurde anschließend als Stereo-Datei exportiert und Suno zur Verfügung gestellt. Mithilfe der *Get Stems*-Funktion soll Suno aus dieser Stereo-Datei wieder auf die Einzelspuren schließen, die im Cubase-Projekt als MIDI vorliegen.

Dabei zeigte sich eine interessante Diskrepanz: Obwohl die Passage ursprünglich aus acht einzelnen Spuren bestand (siehe Abb. 19), erzeugte Suno beim Herausrechnen der Stems lediglich vier Spuren (siehe Abb. 20). Mehrere Instrumente wurden von Suno zu gemeinsamen Kategorien zusammengefasst.

In Tabelle 1 weiter aufgeführt sind alle Einzelspuren der in *Cubase* erzeugten [MIDI-Erweiterung](#) (siehe Abb. 19), sowie die durch *Get Stems* von Suno erzeugten Resultate. Alle Einzelspuren wurden vom identischen Startzeitpunkt aus gerendert, um eine korrekte zeitliche Synchronisation beim Import in eine DAW zu gewährleisten. Beim Anhören der einzelnen Spuren kann es daher erforderlich sein, zu den relevanten Stellen innerhalb der Datei vorzuspulen. Außerdem wurde hier das MP3-Format gewählt, um die Datenmenge zu reduzieren.

Tabelle 1: Vergleich Originalspur MIDI und reproduzierte Spur von Suno (siehe auch Abb. 19 und 20).

Nummer	MIDI Spuren	Suno-Stem
1	Choir	Mixdown (Vorlage für Suno)
2	Synth Choir	Backing Vocals
3	Synth Sinus	Bass
4	Synth Suspense	Synth
5	Strings Bass	FX
6	Glockenspiel	
7	Geigen	
8	Impact	



Abb. 19 – Die Passage des Liedes innerhalb der DAW Cubase Pro 14. Die Benennung der Spuren wurde hierbei vom Ersteller selbst gewählt. Die einzelnen Instrumentengruppen sind grob farblich markiert.

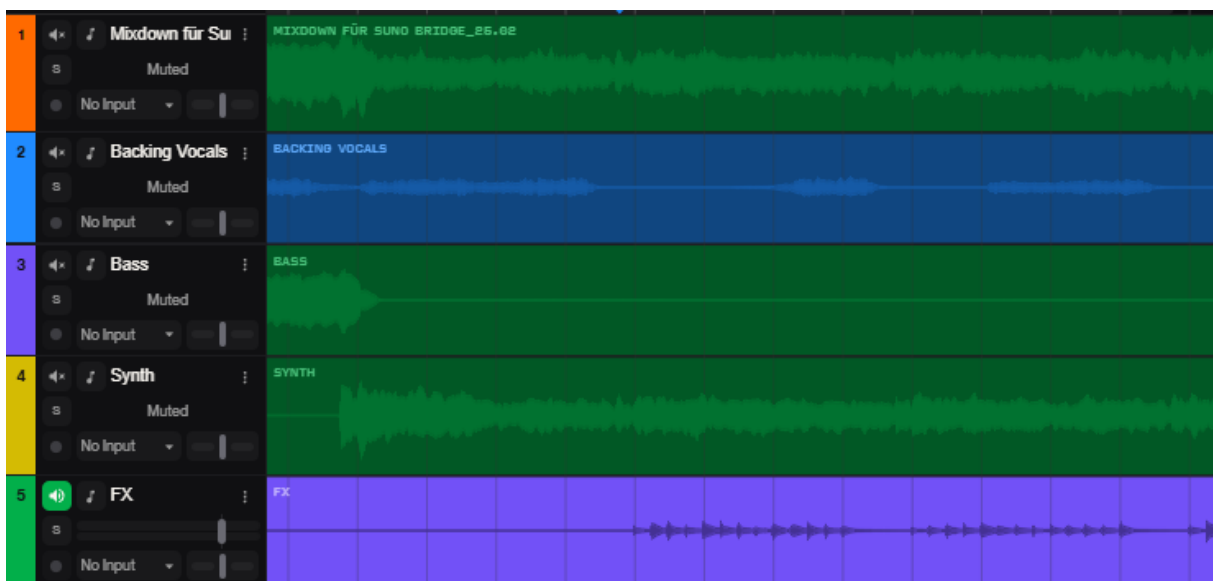


Abb. 20 – Die aus Suno errechneten resultierenden Spuren innerhalb Suno Studios.

Beim Vergleichen der einzelnen Spuren fällt auf: Die [Backing-Vocals](#), insbesondere zu Beginn der Passage, wurden sehr klar von den übrigen Instrumenten getrennt. Ab 0:35, sobald bei der [MIDI-Erweiterung](#) das Glockenspiel einsetzt, wird der Chor jedoch zunehmend der [Synth](#)-Spur zugeordnet. In diesem Bereich gelingt es Suno nicht mehr, den Chor eindeutig von den restlichen Instrumenten zu separieren.

Die von Suno erzeugte [Bass](#)-Spur erkennt den ursprünglich sinusartigen Ton (vgl. [Synth Sinus](#)) korrekt und platziert ihn innerhalb des Arrangements an die passende Stelle. Beim klanglichen Vergleich fällt jedoch auf, dass zusätzliche Rauschanteile sowie weitere Obertöne hinzukommen. Der Ton wirkt dadurch insgesamt deutlich verzerrter als im Originalsignal.

In der Spur [Synth](#) werden von Suno mehrere Elemente des Originals zusammengeführt. Die [Strings-Bass](#) Spur, die hohen Streicher [Geigen](#), sowie die [Synth-Suspense](#) Spur. Ab 0:35 werden außerdem, wie oben erwähnt, Teile des Chors dieser Spur zugeordnet.

Das Glockenspiel wurde hingegen sehr präzise innerhalb der Spur [FX](#) separiert, jedoch von Suno nicht als Glockenspiel zugeordnet. Außerdem ist ein leichtes Rauschen hörbar, sodass die Spur nicht vollständig artefaktfrei ist.

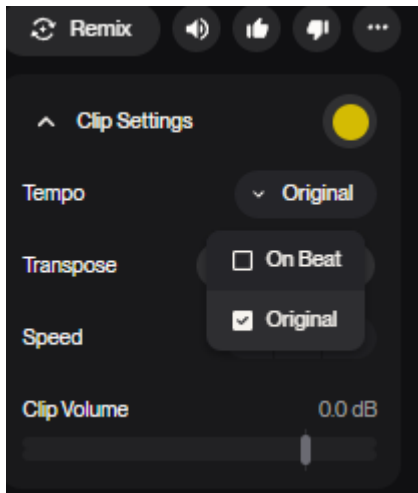
Erfolgte Beobachtungen:

- Beim Export der generierten MIDI-Dateien weicht das Timing teilweise vom ursprünglich vorgegebenen Grid ab. Besonders bei langsamen Material scheint es für Suno schwierig zu sein, das Ergebnis eindeutig einem präzisen Timing zuzuordnen.
- Es ist nicht ausreichend gut möglich, aus einem Stereo-Mixdown mithilfe der *Stem separation* wieder sauber das ursprüngliche Einzelspurmaterial zu rekonstruieren. Stattdessen fasst Suno mehrere Instrumente häufig zusammen und erzeugt viele Artefakte.
- Bestimmte Instrumente, wie z. B. das Glockenspiel, lassen sich präzise isolieren, während andere Instrumente, wie der Chor, von Suno nicht sauber getrennt werden.

3.2.4 Ab 06. März 2026

- Das Ende des Songs ausgestalten –

Da der Stereo-Masterkanal der einzelnen MIDI-Spuren nicht erfolgreich mithilfe der *Stem-Separation* isoliert werden konnte, wurde, - auch aufgrund der positiven Ergebnisse aus Tag 2 (vgl. [Saxophon separate](#)) – erneut die Methode angewandt, ein über MIDI eingespieltes Instrument mithilfe der *Cover-Funktion* in ein von Suno generiertes Ergebnis umzuwandeln.



Hierfür wurden alle MIDI-Einzelspuren, wie in Tabelle 1 dargestellt, in Suno AI hochgeladen und innerhalb *Suno Studio* untereinander angeordnet. Dabei stellte sich heraus, dass die einzelnen Spuren in *Suno Studio* zeitlich nicht korrekt untereinander gesetzt werden konnten, da manche Spuren mit 55bpm detektiert wurden, statt dem Projekt entsprechenden 110 bpm. Damit die einzelnen Spuren zeitlich korrekt dargestellt werden, musste das Tempo jedes Tracks auf *Original*, statt auf *on beat* gestellt werden (siehe Abb. 21).

Abb. 21 – Beim Import der Daten kann eingestellt werden, ob Suno das originale Tempo beibehält oder das Tempo selbst detektiert.

Eine interessante Feststellung hierbei: Vor allem bei langsam gespielten, flächigen Spuren (hier [Synth Choir](#) und [Synth Suspense](#)) wurde das Tempo falsch detektiert und um den Faktor 0.5 reduziert.

Anschließend wurde mit der Cover-Generierung der [Geigen](#)-Spur begonnen, da der Klang der zuvor mit einem VST-Instrument eingespielten Geige nicht zufriedenstellend war. Die Parameter wurden wie beim Saxophonsolo auf *Weirdness: 20%*, *Style Influence: 80%* und *Audio Influence: 80%* eingestellt. Dabei wurde folgender Prompt verwendet:

suspense, violins, drama, Thriller, transition, pizzacato

Die Ergebnisse waren größtenteils nicht zufriedenstellend und es folgten mehrere Versuche mit unterschiedlichen Parametereinstellungen. Die generierten Varianten orientierten sich weder ausreichend an der Vorlage, noch an dem eingegebenen Prompt, oder an der gewünschten Klangästhetik. Das überzeugendste Ergebnis ist hier beispielhaft aufgeführt.

- [Geigen Cover von Suno](#)

Insbesondere das abschließende Crescendo (siehe Glossar) wirkt hier wie eine sinnvolle Erweiterung, die den Song beenden könnte. Daher wurde dieses in das Cubase-Projekt vorläufig übernommen und auch innerhalb von *Suno Studio* ergänzt. Nun wurde jedoch ein neuer Ansatz verfolgt, da die Qualität der Ergebnisse bisher als zu schwankend eingestuft wurde.

Der neue Ansatz verfolgte das Ziel, dass die gesamte Liedpassage – [MIDI-Erweiterung](#) - von Suno grundsätzlich neu interpretiert werden sollte, um daraus mithilfe der Funktion *Get Stems* die einzelnen Instrumente als Ergänzung für die [MIDI-Erweiterung](#) zu nutzen. Hierfür wurden zwei unterschiedliche Methoden getestet:

Zum einen wurde die bisher noch nicht angewandte *Remix-Funktion* verwendet. Zum anderen wurde in *Suno Studio* über die gesamte Länge des Songs ein neuer Track generiert, um mithilfe der im Vorfeld bereits häufig mit Erfolg verwendeten *Cover-Funktion* die Passage als Song neu zu interpretieren (siehe Abb. 22).

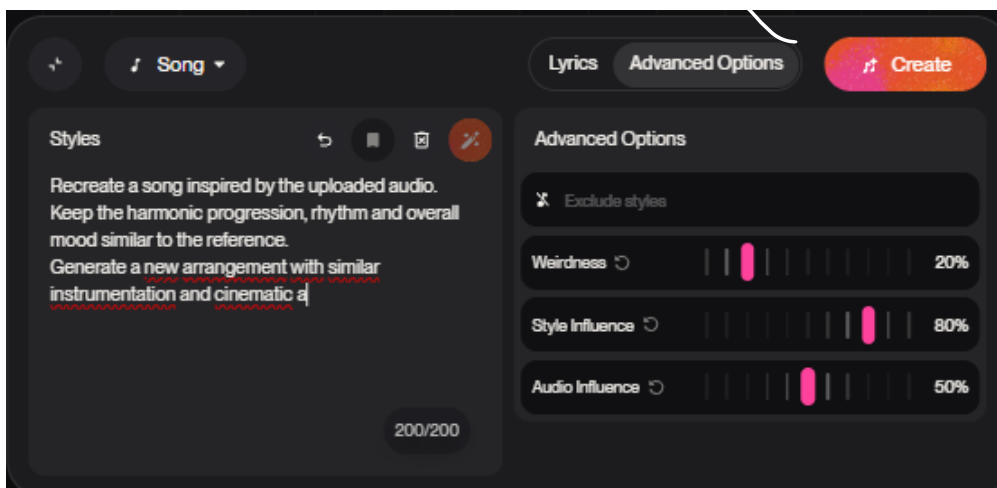


Abb. 22 - Eingegebener Prompt mit entsprechenden Werten.

Die Ergebnisse über die *Cover-Funktion* waren nicht weiter verwendbar. In einer Variante war nahezu ausschließlich Schlagzeug zu hören, während eine andere Version durch zahlreiche Artefakte auffiel. Hier ist exemplarisch die bessere der beiden Versionen dargestellt.

- [Übergang von Suno erstellt mit cover](#)

Im nächsten Schritt wurde daher die *Remix-Funktion* von Suno eingesetzt. Hierzu wurde der betreffende Abschnitt innerhalb von Suno Studio markiert und die Funktion *Remix selection* ausgewählt. Anschließend öffnete sich ein Fenster, in dem das zuvor hochgeladene Audiomaterial, [MIDI-Erweiterung](#), erschien. In diesem stehen zwei getrennte Eingabefelder für *Lyrics* und *Styles* zur Verfügung. Der zuvor verwendete Prompt (siehe Abb.22) wurde in

das Styles-Feld eingefügt und mithilfe des integrierten KI-Werkzeug *Prompt Boosting* erweitert. Dieses ist seit v4.5 in Suno integriert und soll die vom User eingegebenen Prompts für das Programm optimieren (Suno – Creative Prompt Boosting in v4.5., o. D.). Diese Möglichkeit war bei allen vorherigen Versuchen, wo mithilfe der *Cover* Funktion gearbeitet wurde, nicht verfügbar (siehe Abb. 8, 9, 10, 12, 13, 14, 15, 16 und 22).

Der durch das Tool erweiterte Prompt lautete:

„Begin with a brooding piano motif under soft, swelling strings and sparse low brass, keeping the harmony and moody pulse intact. Gradually layer in pulsing percussion and subtle woodwinds, echoing the reference’s melodic shape but rephrased with violin and French horn interplay. Midway, a dramatic crescendo lifts as full orchestra surges, intensifying the theme with bold melodic flourishes. The arrangement climaxes with rich brass, soaring violins, and dynamic percussion before tapering off.“

Die daraufhin generierten Ergebnisse waren wesentlich überzeugender. Die Umsetzung orientierte sich hörbar an den beschriebenen musikalischen Elementen des von *Suno* erweiterten Prompts. Der Beginn der generierten Passage setzt mit einem Pianomotiv und anschwellenden Streichern ein. Holzbläser sind in diesem Abschnitt zunächst kaum wahrnehmbar, während Hörner im weiteren Verlauf deutlicher zu hören sind.

Das Crescendo wurde klanglich mit Blechbläsern ausgearbeitet und im weiteren Verlauf wurden außerdem Schlagzeug- und Shaker-Elemente integriert. Insgesamt wurden die im Prompt beschriebenen musikalischen Entwicklungen überzeugend umgesetzt.

- [Übergang von Suno erstellt mit *Remix selection*](#)

Abschließend wurde aus dieser Version zusätzlich eine Variante ohne Effekte wie Hall mithilfe der Suno-Funktion *Remove Fx* erzeugt.

- [Übergang von Suno, erstellt mit *Remix selection \(ohne FX\)*](#)

Bei der Version *ohne FX* klingen die Streicher (Passage - 0:18 bis 0:34) sogar räumlicher, dumpfer und weniger direkt als in der unbearbeiteten Version mit *FX*. Ab Minute 0:53 bis 1:01 ist zu hören, dass sogar das gesamte Klavier herausgerechnet wurde. Im Übrigen klingt der Nachhall (siehe Glossar) - zu hören am Ende des Crescendos bei Minute 1:17 - sehr künstlich und unrealistisch. Zusätzlich sei an dieser Stelle aber erwähnt, dass auch bei der unbearbeiteten Version ein Artefakt in der Hallfahne wahrnehmbar ist. Da die Version [Übergang von Suno, erstellt mit *Remix selection \(ohne FX\)*](#) diese deutlichen Fehler aufwies, wurde die Stereo-Datei *mit FX*, wie oben bereitgestellt, in das Cubase-Projekt integriert. Eine Extraktion der einzelnen Stems war nicht möglich, da Suno nach der Anwendung dieser Aktion einen dauerhaften Ladebalken anzeigte und der Prozess nie abgeschlossen wurde.

Dieses Verhalten trat auch nach mehrmaligem Neustarten des Programms und des Rechners immer wieder auf.

- [Zwischenstand 06. März](#)

Erfolgte Beobachtungen:

- Die *Cover* Funktion scheint bei klaren, gut definierten musikalischen Vorlagen (wie beim Saxophon) besser zu funktionieren und bei einem Multitrack als Vorlage wie beim Generieren der Streicher oder als Funktion, um einen ganzen Song zu erstellen, schlechtere Ergebnisse zu liefern.
- Tempoerkennung in Suno Studio kann fehlerhaft sein und scheint abhängig vom vorliegenden Material besser bzw. schlechter zu funktionieren.
- Die Funktion *Remove FX* scheint bei einzelnen Instrumenten (siehe Saxophon) besser zurechtzukommen, als bei komplexen Arrangements.

3.2.5 Ab 09. März 2026

Nachdem das grundsätzliche Arrangement und der Aufbau des Songs nun fertiggestellt waren, bestand der nächste Schritt darin, den Klang weiter zu optimieren. Dafür wurde ab dem [Zwischenstand vom 06. März](#) weitergearbeitet. Im Folgenden werden einige der vorgenommenen Änderungen und Bearbeitungsschritte erläutert

- Impact Sound bearbeiten –

Der ursprünglich verwendete Impact-Sound (Minute 0:37) sollte aufgrund einer hörbaren Lautstärkeschwankung ausgetauscht werden (siehe Abb. 23). Dazu wurde die Funktion *Create Sounds* verwendet, die es erlaubt, eigene Audiosamples wie *Swooshes* oder *Riser* (siehe Glossar) zu erstellen (Suno – Suno Sounds: generate Custom Audio Samples, o. D.). Dabei zeigte sich erneut ein typisches Verhalten: Trotz klar formulierter Prompts fügte das System häufig zusätzliche Instrumente wie Schlagzeug oder Gesang hinzu, die weder im Prompt noch in den Parametern vorgesehen waren.

Zunächst wurde folgender Prompt verwendet:

cinematic impact transition, rising tension, reverse textures, subtle reverb tail, transition into next scene

Die erstellten Impacts waren zu tief und basslastig und passten nicht ins Arrangement, daher musste der Prompt geändert werden.

Doch auch nach der Erstellung weiterer zahlreicher Varianten konnte kein überzeugendes Ergebnis erzielt werden. Daraufhin wurde erneut die *Cover-Methode* ausprobiert, da *Suno* hierbei den bestehenden Song als Referenz analysiert und den generierten Sound stärker an das vorhandene Material anpasst.

Trotz des Erstellens von insgesamt über 40 Impacts konnte kein zufriedenstellendes Ergebnis gefunden werden. Auch in der *Cubase*-internen Soundbibliothek ließ sich kein passender Impact finden. Daher wurde schließlich dem ursprünglichen Impact eine andere Stelle desselben *Impacts* beigemischt.

- [Impact ohne Bearbeitung](#)
- [Impact mit Bearbeitung](#)

Dies führte jedoch nur zu einer sehr geringfügigen Verbesserung des vorhandenen Problems.



Abb. 23 - Zeigt innerhalb des orangenen Bereichs die Wellenform im Bereich des entsprechenden Artefakts. Auffällig ist ein unnatürlicher Einbruch der Amplitude. Da sich dieser Fehler weder durch *Time-Stretching* noch durch *Clip gain* (siehe Glossar) beheben ließ, wurde der Sound nachträglich bearbeitet, indem eine andere Stelle desselben *Impacts* beigemischt wurde.

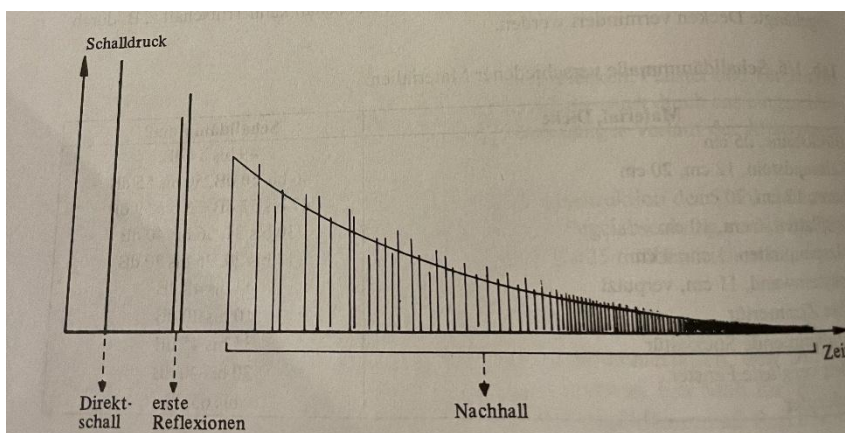


Abb. 24 – Der typische Energieverlauf eines Impulses ist durch ein kontinuierliches Ausschwingen mit abnehmender Amplitude gekennzeichnet (Dickreiter, M., 2014, S.30).

- Ergänzung der Streicher und Bläsersektion-

Im nächsten Schritt wurde die Streicher und Bläsersektion im entsprechenden Abschnitt (0:39 bis 1:32) weiter ausgebaut. Hierfür wurden mehrere Änderungen vorgenommen. Es wurden viele zusätzliche MIDI-Spuren eingefügt, auf welche jedoch aufgrund des KI-Themenschwerpunktes dieser Bachelorarbeit nicht näher eingegangen wird.

Parallel dazu wurde Suno weiterhin zur Generierung neuer Spuren verwendet, die anschließend in der DAW weiterbearbeitet und angepasst wurden.

Auffällig dabei war erneut, dass ein kurz formulierter Prompt zur Generierung einer Streicherspur mithilfe der *Cover Funktion* nicht zum gewünschten Ergebnis führte. Der Prompt lautete:

Cellos and Violas, legato, tenuto, gentle, calm

- [Sunos Ergebnis](#)

Zu hören sind schnell spielende Streicher zu Beginn, ohne eine Legato oder Tenuto (siehe Glossar) Artikulation. Am Ende der Passage ist zudem ein klackendes Geräusch wahrnehmbar.

- Schlagzeugspuren hinzufügen -

Um Schlagzeugspuren zu ergänzen, wurde folgende Methode angewendet: Ein entsprechender Abschnitt wurde in *Suno Studio* markiert und mithilfe der *Cover-Funktion* wurde innerhalb dieses Abschnittes, mithilfe eines passenden Prompts, neues Material generiert.

Der erste, erneut nicht ausführlich formulierte Prompt lautete:

cinematic war drums, only drums, epic

Dieser führte zu unbefriedigenden Ergebnissen. Es wurden erneut zusätzliche Instrumente generiert, die weder im Prompt noch in den Einstellungen vorgesehen waren. Ein exemplarisches Beispiel ist hier gelistet:

- [Exemplarisches Beispiel](#)

Daraufhin wurde der Prompt erweitert und präzisiert:

cinematic war drums, epic battle drums, deep tribal percussion, large taiko style drums, powerful rhythm, dramatic build, only drums, no melody, no instruments, dark cinematic atmosphere

Um den Produktionsprozess zu verschnellern, wurden insgesamt 20 Drum-Varianten mit obigem Prompt, aber jeweils anderen Parametereinstellungen generiert. Aufgrund der Vielzahl an Versuchen werden hier nicht alle einzelnen Parametereinstellungen aufgeführt. Dieses Vorgehen erwies sich jedoch als effiziente Methode, da innerhalb von *Suno Studio*, so ähnlich zum Workflow des *compings* (siehe Glossar), die geeignetsten Spuren ausgemacht werden konnten.

Mehrere Ergebnisse wurden anschließend in *Cubase* importiert und in das Projekt integriert.

Eine *Stem Separation* innerhalb von *Suno* war nicht möglich, da *Suno* die gesamte Drum Spur offenbar als einen einzigen Stem erkannte. In mehreren Fällen exportierte *Suno* schlicht die identische Spur erneut, anstatt einzelne Bestandteile zu isolieren.

Das Trennen der Schlagzeugspuren von ihrem entsprechenden Hallanteil mit der Funktion *Remove FX* war ebenfalls nicht möglich. *Suno* generierte, ähnlich wie beim letzten Mal (vgl. S.42), erneut artefaktbehaftete und dadurch unbrauchbare Spuren.

Letztlich wurden mehrere Drum-Varianten aus unterschiedlichen Generierungen kombiniert (siehe Abb. 25). Diese wurden im Projekt zusammengeschnitten und arrangiert. Zusätzlich wurden ein *Riser* sowie ein *Downlifter* (siehe Glossar) aus der *Cubase*-eigenen Sample Library eingefügt. Außerdem wurden fünf Beckenschläge mithilfe des *Drum Samplers* Groove Agent 5 von Steinberg (Steinberg Media Technologies, o. D. a.) hinzugefügt. Anschließend erfolgte eine klangliche Anpassung mithilfe von Equalizern, Kompressoren und Limitern, um die Schlagzeugspuren in das bestehende Projekt zu integrieren.

Das Ergebnis dieser Bearbeitung ist im folgenden Audiobeispiel zu hören:

- Von *Suno* generiertes und dem Anwender arrangiertes [Schlagzeugset](#)

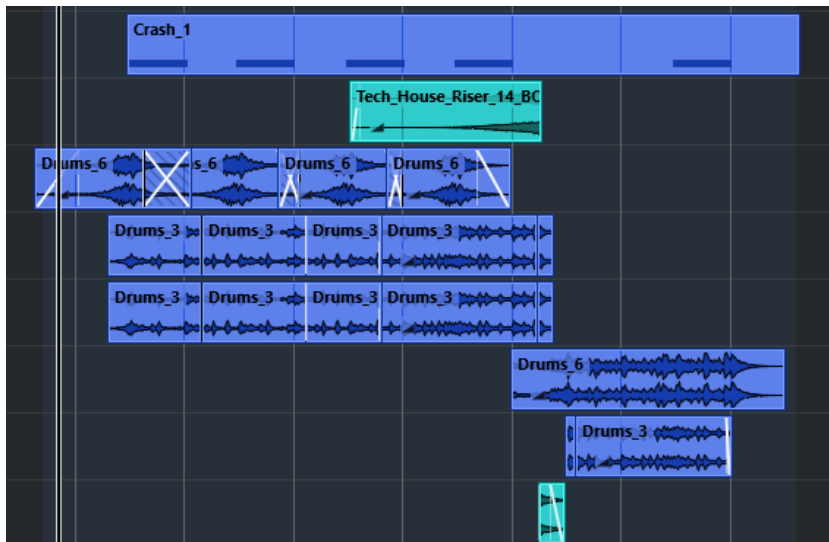


Abb. 25 – Darstellung der verwendeten Drumsuren im Projekt. Zu sehen sind die Groove-Agent-Spur (erster Kanal), der Tech-House-Riser (zweiter Kanal), der zugeschnittene Downlifter (letzter Kanal) sowie zwei der insgesamt zwanzig generierten Suno-Kreationen (Kanäle drei bis sieben), die mithilfe von Crossfades sowie entsprechenden Fade-ins und

Fade-outs (siehe Glossar) zugeschnitten wurden. Außerdem wurde eine frequenzbasierte Dopplung der Spuren vorgenommen. Hierbei wurden die entsprechenden Audiospuren dupliziert und anschließend gefiltert, sodass jeweils nur die hohen bzw. tiefen Frequenzanteile der Originalspur wiedergegeben werden. Auf diese Weise konnte das Drumset trotz fehlender Stem-Separation zumindest teilweise frequenzbasiert weiterbearbeitet werden.

- Nachhall bearbeiten –

Es stellte sich heraus, dass bei KI-generierten Spuren der Nachhall aufgrund abrupt endender Audiodateien häufig abgeschnitten wird und dadurch ein unnatürlicher Klang entsteht. Daher mussten im vorliegenden Stück insgesamt vier Hallfahnen nachbearbeitet werden. Hierfür wurde der Hall-Effekt *RoomWorks* in Cubase verwendet (Steinberg Media Technologies, o. D. b). Der Effekt wurde an vier Stellen des Liedes dem Originalsignal zugemischt, insbesondere an Übergängen oder kurzen Pausen, in denen der Nachhall – beispielsweise am Ende eines Abschnitts – deutlich hörbar wird. Exemplarisch sind im Folgenden zwei Beispiele aufgeführt, die den Zustand vor und nach der Bearbeitung veranschaulichen:

Beispiel 1

- [Anfang Chor](#) ohne Bearbeitung
- [Anfang Chor](#) mit Bearbeitung

Beispiel 2

- [Ende Impact](#) ohne Bearbeitung
- [Ende Impact](#) mit Bearbeitung

Erfolgte Beobachtungen:

- Abgeschnittene oder artefaktbehaftete Hallfahnen stellen bei Übergängen und dem Ende eines Songs häufig ein Problem dar
- Kurze und schlecht ausformulierte Prompts führen zu hörbar schlechteren Ergebnissen.
- Innerhalb des Produktionsprozesses zeigte sich, dass es zielführend ist, eine große Anzahl an Spuren mit unterschiedlichen Parametereinstellungen in Suno generieren zu lassen. Durch die Vielzahl an Varianten können anschließend die klanglich überzeugendsten Ergebnisse ausgewählt und in das Projekt integriert werden.

Das Lied ist damit vollständig ausproduziert. Der anschließende Produktionsschritt würde nun das Mixing und Mastering umfassen. Diese Schritte würden jedoch den Rahmen dieser Bachelorarbeit überschreiten, da sich diese Arbeit ausschließlich auf den Produktionsprozess konzentrieren soll.

Auf Grundlage der im Produktionstagebuch dokumentierten Beobachtungen folgen im nächsten Kapitel wissenschaftlich angelegte Tests, nach deren Auswertung konkrete Ergebnisse abgeleitet werden können.

- [Song](#) als wav. Datei
- [Song](#) als mp3. Datei

4 Systematischer Test

4.1 Motivation

In Kapitel 4 werden insgesamt drei Testreihen durchgeführt. Ziel dieser Untersuchungen ist es, die formulierte Forschungsfrage mithilfe transparent gesammelter Daten besser beantworten zu können und die im Produktionsprozess gewonnenen Beobachtungen zu überprüfen und einzuordnen. Die Forschungsfrage lautet:

Inwiefern können KI-gestützte Tools einzelne Arbeitsschritte der Songproduktion technisch sinnvoll unterstützen oder ersetzen, und wo liegen deren praktische und qualitative Grenzen am Beispiel einer KI-unterstützten Songproduktion?

Im Verlauf der Songproduktion zeigte sich, dass das System Suno insbesondere als Inspirationsquelle sowie zur Generierung einzelner Instrumente innerhalb bestehender Arrangements effektiv eingesetzt werden kann. Gleichzeitig wurden jedoch auch deutliche Einschränkungen sichtbar, insbesondere hinsichtlich der fehlenden Reproduzierbarkeit der generierten Ergebnisse, sowie einer schwankenden Zuverlässigkeit was die Übernahme bestimmter Prompts betraf. Um diese Schwächen des Systems ausgleichen zu können, ist es wichtig zu wissen, wie Anwender schnellstmöglich, auch aufgrund der je nach gewähltem Bezahlmodell beschränkten Credit-Anzahl, zu bestmöglichen Ergebnissen kommen. Deswegen wird in diesem Test die Auswirkungen der innerhalb der Produktion sehr häufig genutzten *Creative Sliders* überprüft. Dabei soll ermittelt werden, was für einen Einfluss sie auf die Qualität und die intendierte Veränderung (Kreativität, Stiltreue und Klangtreue) haben.

Die Ergebnisse wurden sowohl über Lautsprecher als auch über Kopfhörer abgehört. Hierbei kommen die Kopfhörer von beyerdynamic DT 770 Pro (250 Ohm) sowie die Monitorlautsprecher Yamaha HS5 in Kombination mit einem Subwoofer Yamaha HS8S zum Einsatz.

4.2 Methodik

In diesem Test wird der Einfluss der Creative Sliders *Weirdness*, *Style Influence* und *Audio Influence* auf die generierten Ergebnisse untersucht.

Zielsetzung:

Untersucht werden soll, inwieweit sich die generierten Ergebnisse durch gezielte Anpassungen der Parametereinstellungen verändern lassen und wie stark die einzelnen Parameter das Resultat beeinflussen.

Hypothese:

Die Parametereinstellungen (*Audio Influence*, *Style Influence* und *Weirdness*) haben einen signifikanten Einfluss auf die generierten Ergebnisse. Mit steigenden Werten der Parameter *Audio Influence* und *Style Influence* nimmt die Ähnlichkeit zur Vorlage zu, während eine Erhöhung des Parameters *Weirdness* zu einer stärkeren Abweichung von der Vorlage führt. Dabei wird angenommen, dass *Audio Influence* den größten Einfluss auf das Ergebnis ausübt, während *Style Influence* im Vergleich den geringsten Einfluss ausübt.

Testverlauf:

Innerhalb des Tests wird von einem Standardwert bei allen Parametern von *Weirdness* = 50, *Style Influence* = 50 und *Audio Influence* = 50 für die Generierung einer Reproduktion einer 15 bis 20 sekundenlangen Passage aus der im Produktionstagebuch erläuterten Produktion ausgegangen. Hierfür wird die *Cover* Funktion innerhalb von *Suno Studio* genutzt. Dabei wird auf einen Textprompt verzichtet, um ausschließlich die Parameter als beeinflussenden Faktor bewerten zu können. Die Ergebnisse werden anhand von drei Hauptbewertungskriterien pro Durchlauf anhand einer Bewertungsskala eingestuft. Insgesamt gibt es fünfzehn Durchläufe. Dreizehn mit unterschiedlichen Parametereinstellungen, da der Referenzwert (50/50/50) mehrfach vorkommt. Jeder Durchlauf wird insgesamt zwei unterschiedliche Ergebnisse generieren. Insgesamt wird es also 30 zu bewertende Ergebnisse geben (vgl. Tab. 2). Die Wechselwirkungen zwischen den Parametern untereinander können daher nicht vollständig untersucht werden.

Die einzelnen Ergebnisse werden pro Parameter (*Weirdness*, *Style Influence* und *Audio Influence*) auf verschiedene Eigenschaften überprüft.

Tabelle 2:

Weirdness-Durchlauf	Style-Influence-Durchlauf	Audio-Influence-Durchlauf
W 0 / S 50 / A 50 -> 2 Ergebnisse	W 50 / S 0 / A 50 -> 2 Ergebnisse	W 50 / S 50 / A 0 -> 2 Ergebnisse
W 25 / S 50 / A 50 -> 2 Ergebnisse	W 50 / S 25 / A 50 -> 2 Ergebnisse	W 50 / S 50 / A 25 -> 2 Ergebnisse
W 50 / S 50 / A 50 -> 2 Ergebnisse	W 50 / S 50 / A 50 -> 2 Ergebnisse	W 50 / S 50 / A 50 -> 2 Ergebnisse
W 75 / S 50 / A 50 -> 2 Ergebnisse	W 50 / S 75 / A 50 -> 2 Ergebnisse	W 50 / S 50 / A 75 -> 2 Ergebnisse
W 100 / S 50 / A 50 -> 2 Ergebnisse	W 50 / S 100 / A 50 -> 2 Ergebnisse	W 50 / S 50 / A 100 -> 2 Ergebnisse

Für den Parameter *Weirdness* wird untersucht, wie stark sich das Ergebnis von der Vorlage entfernt, bei welchen Einstellungen Artefakte entstehen oder ungewöhnliche Entscheidungen auftreten und in welchem Maß das Ergebnis als kreativ bewertet werden kann. Zur Bewertung wird bei allen drei Parametern (Weirdness, Style Influence, Audio Influence) folgende Skala verwendet:

Bewertungsskala

0 = sehr schlecht / gar nicht erfüllt

1 = stark unzureichend

2 = eher schwach

3 = mittel / teilweise erfüllt

4 = gut

5 = sehr gut / klar erfüllt

Jedes Ergebnis des Weirdness-Durchlaufs wird hinsichtlich der drei Haupteigenschaften

Nähe zur Vorlage, **Artefaktfreiheit** und **Kreativität** bewertet.

- **Nähe zur Vorlage:** Inwieweit orientiert sich das Ergebnis hinsichtlich des Klangs, der Melodie und des Verlaufs an der Vorlage?
- **Artefaktfreiheit:** Gibt es hörbare Artefakte, unpassende Entscheidungen oder klanglich störende Passagen?
- **Kreativität:** In welchem Maß ist das erzeugte Material abwechslungsreich, vielfältig und inspirierend?

Damit die Bewertung dieser Haupteigenschaften nachvollziehbar hergeleitet werden kann, setzt sich jede Gesamtbewertung der Haupteigenschaft aus mehreren Untereigenschaften zusammen. Deren Anzahl und Gewichtung variiert je nach Haupteigenschaft.

Die Haupteigenschaft **Nähe zur Vorlage** setzt sich aus den Untereigenschaften *Instrumentierung, Melodie, Klang, sowie Übergänge und Ende* zusammen.

Die Haupteigenschaft **Artefaktfreiheit** setzt sich aus den Untereigenschaften *Unnatürliche Sounds, Mixing / Balance* sowie *Nutzbarkeit* zusammen.

Die Haupteigenschaft **Kreativität** setzt sich aus den Untereigenschaften *musikalische Weiterentwicklung* und *Kohärenz* zusammen.

Der Wert einer Haupteigenschaft ergibt sich jeweils aus dem Durchschnitt der zugehörigen Untereigenschaften. Wird ein Ergebnis beispielsweise unter der Haupteigenschaft **Artefaktfreiheit** in den Untereigenschaften *Unnatürliche Sounds* mit 4 von 5 Punkten, *Mixing / Balance* mit 5 von 5 Punkten und *Nutzbarkeit* mit 4 von 5 Punkten bewertet, ergibt sich der Gesamtwert der Haupteigenschaft **Artefaktfreiheit** aus dem gerundeten Durchschnitt dieser drei Einzelwerte.

Die Berechnung erfolgt mit nachfolgender Formel:

$$\text{Artefaktfreiheit} = \frac{\text{Unnatürliche Sounds} + \text{Mixing/Balance} + \text{Nutzbarkeit}}{\text{Anzahl der Untereigenschaften}}$$

Im genannten Beispiel ergibt sich damit folgender Wert:

$$\text{Artefaktfreiheit} = \frac{4 + 5 + 4}{3} = 4.33 = 4.3$$

Diese Methode wird auch bei dem Style-Influence und Audio-Influence-Durchlauf angewandt und daher nicht erneut ausführlich beschrieben. Die Haupt- und Untereigenschaften variieren natürlich von Durchlauf zu Durchlauf und werden am entsprechendem Kapitelanfang (Kapitel 4.3.3 und Kapitel 4.3.5) kurz beschrieben. Im Anhang finden sich die entsprechenden Bewertungstabellen pro Durchlauf.

Folgende Passage wird auf die Parametereinstellungen Weiridness, Style-Influence und Audio-Influence getestet:

- [Passage](#)

Die Vorlage beginnt mit einem Saxophon, einem Bass und einem Klavier. Im weiteren Verlauf leitet ein Swoosh- und Impact-Effekt gemeinsam mit dem Klavier in einen neuen Abschnitt über (ab 0:09). In diesem treten ein zusätzliches Saxophon sowie weitere Effekte hinzu. Darunter seitlich (links/rechts) gepannte Streichinstrumente (ab 0:12), ein gezupfter Bass (ab 0:11) und ein am Ende stark modulierter Bläserklang (ab 0:16).

Auswertung der Ergebnisse:

Jede Parametereinstellung innerhalb eines Durchlaufs erzeugt zwei Ergebnisse. Von diesen Ergebnissen wird ein Mittelwert gebildet und für jedes Bewertungskriterium wird eine Kurve gebildet mit x-Achse (Nummer des Durchlaufs) und y-Achse (Bewertung 0-5). Auf diese Weise lassen sich die Auswirkungen der einzelnen Parametereinstellungen visuell darstellen und die Veränderung über die Durchläufe hinweg gut vergleichen. Die Interpretation erfolgt auf Basis der erkennbaren Kurvenverläufe und Bewertungen. Alle Bewertungen pro generiertem Ergebnis werden in einer Tabelle zusammengefasst (siehe Anhang).

4.3 Durchführung und Auswertung

4.3.1 Einfluss des Parameters Weirdness

[Ergebnis 1.1](#) mit den Parametereinstellungen: 0 / S 50 / A 50

Nähe zur Vorlage: 4 / 5

Instrumentierung: Die Version enthält zusätzliche Drums, die so in der Vorlage nicht erhalten sind. Im zweiten Part ist aber kein Schlagzeug mehr zu hören. Die Grundinstrumente wie Saxophon, Bass und Piano sind erhalten und gut zu hören. **4/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine leicht veränderte Melodie besitzen. Das zweite Saxophon spielt sehr nahe an der Vorlage, mit kleinen Variationen. **4/5**

Klang: Auffällig hier ist, wie nahe, das erste Saxophon, an den Klang der Vorlage heranreicht. Das zweite Saxophon klingt, ähnlich wie die Vorlage, deutlich nasal und wird ebenfalls gut getroffen. Auch der Bass, wenn auch etwas druckvoller, klingt sehr ähnlich zum Original. Das Piano weicht hingegen stark von der Vorlage ab und klingt mehr nach einem Rhodes Keyboard, nähert sich im zweiten Part dem Klavier aus der Vorlage aber wieder an. **4/5**

Übergänge und Ende: Beim Übergang sind Swoosh und Impact vorhanden, klingen insgesamt aber weniger druckvoll. Das Piano im Übergang ist sehr gut getroffen, spielt jedoch einen Ton zu viel. Der modulierte Bläserklang am Ende ist ebenfalls gut getroffen, wird aber von einem verrauschten Swoosh fast völlig verdeckt (vgl. Artefaktfreiheit). **4/5**

Artefaktfreiheit: 3.7 / 5

Unnatürliche Sounds: Fast bis zum Ende fallen keine Störgeräusche auf. Am Ende ist jedoch ein Rauschen hörbar und das Ende klingt wie abgeschnitten. **3/5**

Mixing / Balance: Das Saxophon steht, wie in der Vorlage, klar im Vordergrund, während das zweite Saxophon etwas zu leise abgemischt und nicht nach links gepannt ist. Insgesamt ist der Klang der Vorlage aber sehr gut herauszuhören, auch wenn ein ganzes Instrument in Form des Schlagzeuges hinzugekommen ist. **4/5**

Nutzbarkeit: Das Material ist, da die störenden Artefakte erst am Ende auftreten, gut nutzbar. **4/5**

Kreativität: 3 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung hält sich trotz minimaler Melodievariationen und dem Hinzufügen eines Schlagzeugs in Grenzen. **2/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. Leider fällt der verrauschte Schluss des Materials dabei eher raus. **4/5**

- [Ergebnis 1.2](#) mit den Parametereinstellungen: W 0 / S 50 / A 50

Nähe zur Vorlage: 3.9 / 5

Instrumentierung: Am Anfang sind erneut Drums hörbar, die dieses Mal auch mit in den zweiten Part spielen, in diesem treten neben dem Schlagzeug auch Percussion Elemente wie eine Clap mit auf. Die Grundinstrumente wie Saxophon, Bass und Piano sind enthalten und gut zu hören. **3.5/5**

Melodie: Die Grundmelodie wird vom Saxophon erneut gut eingefangen. Der Bass spielt jedoch wieder deutlich anders, nämlich schneller als in der Vorlage. Das zweite Saxophon spielt erneut sehr nahe an der Vorlage, mit kleinen Variationen. **4/5**

Klang: Erneut auffällig ist, wie nahe, das erste Saxophon an den Klang der Vorlage heranreicht. Ebenfalls sehr nahe zur Vorlage sind der Klang des zweiten Saxophons und der Bass. Eine Schwäche ist erneut die klangliche Abweichung des Pianos. **4/5**

Übergänge und Ende: Beim Übergang sind Swoosh und Impact vorhanden, gehen dieses Mal aber fast vollkommen im Arrangement unter. Das Piano im Übergang ist sehr gut

getroffen, spielt aber erneut einen Ton zu viel. Der modulierte Bläserklang am Ende ist erneut gut getroffen, wird aber wieder von einem Artefakt gestört. **4/5**

Artefaktfreiheit: 3.7 / 5

Unnatürliche Sounds: Bis auf das Ende fallen keine Störgeräusche auf. Am Ende ist jedoch ein Zischen hörbar und die Spur klingt wie abgeschnitten. **3/5**

Mixing / Balance: Das Saxophon steht, wie in der Vorlage, klar im Vordergrund, während das zweite Saxophon etwas zu leise abgemischt und nicht nach links gepannt ist. Insgesamt ist der Klang der Vorlage aber sehr gut herauszuhören, auch wenn ein ganzes Instrument in Form des Schlagzeuges hinzugekommen ist. **4/5**

Nutzbarkeit: Das Material ist, da die störenden Artefakte erst am Ende auftreten, gut nutzbar. **4/5**

Kreativität: 3 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung hält sich trotz minimaler Melodievariationen und dem Hinzufügen eines Schlagzeuges in Grenzen. **2/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. Leider fällt der Schluss des Materials dabei aber aufgrund der Artefakte heraus. **4/5**

- [Ergebnis 2.1](#) mit den Parametereinstellungen: W 25 / S 50 / A 50

Nähe zur Vorlage: 3.8 / 5

Instrumentierung: Es sind erneut über das gesamte Ergebnis hinweg Drums zu hören, die im zweiten Part jedoch etwas zurückgehen. Die Grundinstrumente wie Saxophon, Bass und Piano sind erhalten und gut zu hören. **4/5**

Melodie: Das erste Saxophon spielt die Melodie der Vorlage korrekt, während das zweite erneut leicht abweicht. Der Bass spielt wie in den vorherigen Generierungen etwas mehr Töne als im Original. **4/5**

Klang: Die Saxophone sind klanglich erneut gut getroffen. In Kombination mit dem Piano, das deutlich dunkler als in der Vorlage klingt, wirkt aber insbesondere das erste Saxophon insgesamt kratziger als im Original. **3.5/5**

Übergänge und Ende: Beim Übergang sind Swoosh und Impact vorhanden, jedoch erneut sehr zaghaft eingesetzt. Das Piano im Übergang ist sehr gut getroffen, spielt aber erneut einen Ton zu viel. Der modulierte Bläserklang am Ende klingt etwas überbearbeitet. **3.5/5**

Artefaktfreiheit: 3.3 / 5

Unnatürliche Sounds: Der Song beginnt mit einem ausklingenden Ton und am Ende ist ein leises Klicken hörbar. **3/5**

Mixing / Balance: Das Saxophon steht, wie in der Vorlage, klar im Vordergrund. Das zweite Saxophon ist diesmal etwas lauter, bleibt aber leiser als in der Vorlage und erneut nicht nach links gepannt ist. Insgesamt ist der Klang der Vorlage aber erneut sehr gut herauszuhören. **4/5**

Nutzbarkeit: Das Material könnte genutzt werden. Es müsste aber zu Beginn, wie auch am Ende, aufgrund der Artefakte aber Audiobearbeitung stattfinden. **3/5**

Kreativität: 3 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung hält sich trotz minimaler Melodievariationen und dem Hinzufügen eines Schlagzeugs in Grenzen. **2/5**

Kohärenz: Das erzeugte Material wirkt insgesamt kohärent und entwickelt sich schlüssig. Der Anfang und das Ende fallen jedoch leicht aus dem Gesamtbild heraus, da dort kleinere Artefakte auftreten, die bei genauerem Hinhören wahrnehmbar sind. **4/5**

- [Ergebnis 2.2](#) mit den Parametereinstellungen: W 25 / S 50 / A 50

Nähe zur Vorlage: 3.3 / 5

Instrumentierung: Die Version enthält über die gesamte Spiellänge zusätzliche Drums. Im Hintergrund sind teilweise zusätzliche Bläser zu hören. Saxophon und Bass sind erhalten und gut zu hören. Das Piano wird größtenteils, außer im Übergang, nicht übernommen. Das zweite Saxophon wechselt teilweise in einen gitarrenähnlichen Klang. **2.5/5**

Melodie: Das Saxophon beginnt mit einer leichten Ausschmückung, ansonsten entspricht die Melodie weitgehend der Vorlage. Der Bass spielt erneut etwas anders. Das zweite Saxophon weicht erneut minimal ab, bleibt aber sehr nah am Original. Im Übergang spielt das Piano diesmal sechs Töne statt der vier in der Vorlage, beziehungsweise der fünf in den zuvor erstellten Versionen. **3.5/5**

Klang: Das erste Saxophon ist wieder sehr nahe am Klang der Vorlage. Das zweite Saxophon klingt teilweise mehr nach einer Gitarre als nach einem Saxophon. Auch der Bass, wenn auch etwas druckvoller, klingt sehr ähnlich zum Original. Das Piano weicht hingegen stark von der Vorlage ab und klingt eher nach zusätzlichen Blasinstrumenten. **3/5**

Übergänge und Ende: Im Übergang sind Whoosh- und Impact-Sounds vorhanden. Das Klavier spielt erneut sechs Töne. Der modulierte Bläserklang am Ende ist herauszuhören. Der Song bricht am Ende ungewöhnlich ab. **4/5**

Artefaktfreiheit: 3.5 / 5

Unnatürliche Sounds: Bis auf das Ende fallen keine Störgeräusche auf. Am Ende klingt der Song jedoch wie abgeschnitten. **4/5**

Mixing / Balance: Das Saxophon steht, wie in der Vorlage, klar im Vordergrund. Das zweite Saxophon ist wieder sehr leise und erneut nicht nach links gepannt. Der Klang der Vorlage ist herauszuhören. **3.5/5**

Nutzbarkeit: Das Material ist grundsätzlich sehr gut nutzbar, da lediglich am Ende ein Artefakt auftritt. Es ist jedoch zu berücksichtigen, dass das zweite Saxophon klanglich uneinheitlich wirkt, was die Bearbeitung im Mix erschwert. **3/5**

Kreativität: 3.5 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung hält sich trotz minimaler Melodievariationen und dem Hinzufügen eines Schlagzeugs in Grenzen. Positiv hervorzuheben ist jedoch die melodische Ausschmückung des Saxophons zu Beginn. **3/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. Leider fällt das Material aufgrund des abrupten Abschlusses und dem unschlüssigen zweiten Saxophon diesbezüglich negativ auf. **4/5**

- [Ergebnis 3.1](#) mit den Parametereinstellungen: W 50 / S 50 / A 50

Nähe zur Vorlage: 3.3 / 5

Instrumentierung: Erneut wurden zusätzlich über den gesamten Verlauf Drums ergänzt, die sich im zweiten Part jedoch etwas zurücknehmen. Zu Beginn sind mehrere Blasinstrumente hörbar, anstelle eines einzelnen Saxophons. Das Piano tritt lediglich im Übergang und im zweiten Part deutlich hervor. **2.5/5**

Melodie: Das erste Saxophon spielt erneut die gleiche Melodie wie in der Vorlage. Auch das zweite Saxophon orientiert sich nahezu identisch an den dort enthaltenen Tönen. Der Bass hingegen weist, wie bereits in allen Variationen zuvor, eine höhere Tonanzahl auf. **4/5**

Klang: Der Klang des Saxophons reicht erneut sehr nah an die Vorlage heran. Das zweite Saxophon klingt, ähnlich wie die Vorlage, deutlich nasaler und wird ebenfalls gut getroffen. Auch der Bass wird wieder sehr gut getroffen. Durch das Hinzufügen des Schlagzeugs und der Bläser klingt diese Version jedoch, wie andere Suno-Versionen bereits zuvor, deutlich dichter als die Vorlage. **3.5/5**

Übergänge und Ende: Swoosh und Impact im Übergang sind lediglich angedeutet und klingen stark verrauscht. Dadurch fügen sie sich nicht organisch in das Arrangement ein und wirken wie störende Elemente. Das Klavier im Übergang verhält sich wie in den Variationen zuvor und spielt sechs statt vier Töne. Der modulierte Bläsereffekt am Ende wirkt erneut stark überbearbeitet und dadurch deutlich verrauscht. Zudem bricht der Song wieder abrupt ab. **3/5**

Artefaktfreiheit: 3 / 5

Unnatürliche Sounds: Insgesamt klingt der Song sehr verrauscht und am Ende wirkt der Song wie abgeschnitten und ebenfalls sehr verrauscht. **2.5/5**

Mixing / Balance: Das Saxophon steht, wie in der Vorlage, klar im Vordergrund, während das zweite Saxophon etwas zu leise abgemischt und nicht nach links gepannt ist. Insgesamt ist der Klang der Vorlage aber sehr gut herauszuhören. **4/5**

Nutzbarkeit: Das Material müsste mithilfe von RX DeNoise (siehe Glossar) oder ähnlichen Programmen bearbeitet werden. Auch der Übergang und das Ende wären ohne Audionachbearbeitung so nicht nutzbar. **2.5/5**

Kreativität: 3 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung hält sich trotz minimaler Melodievariationen und dem Hinzufügen eines Schlagzeugs in Grenzen. **2/5**

Kohärenz: Das erzeugte Material wirkt insgesamt kohärent und entwickelt sich schlüssig. Der Übergang und das Ende fallen jedoch leicht aus dem Gesamtbild heraus, da dort kleinere Artefakte auftreten, die bei genauerem Hinhören wahrnehmbar sind. **4/5**

- [Ergebnis 3.2](#) mit den Parametereinstellungen: W 50 / S 50 / A 50

Nähe zur Vorlage: 3.4 / 5

Instrumentierung: Zu Beginn kommen zwei Saxophone zum Einsatz, die in unterschiedlichen Tonlagen spielen. Vereinzelt sind erneut Kick und Becken hörbar. Insgesamt nimmt sich das Schlagzeug jedoch deutlich zurück. Klavier und Bass, sowie das zweite Saxophon sind gut herauszuhören. **3.5/5**

Melodie: Dieses Mal konnten beim zweiten Saxophon keine Veränderungen festgestellt werden, und auch das erste Saxophon spielt weitgehend korrekt. Lediglich die Begleitinstrumente Klavier und Bass weichen leicht von der Vorlage ab. **4/5**

Klang: Durch das Zurücknehmen des Schlagzeugs nähert sich der Gesamtklang deutlich stärker dem Original an. Das erste Saxophon klingt jedoch etwas tiefer und weniger klar. Bass und Klavier sind klanglich sehr nahe an der Vorlage. Leider reißen die Artefakte, die in der Vorlage nicht vorhanden sind, den Hörer klanglich aus dem Gesamtbild heraus. **4/5**

Übergänge und Ende: Swoosh- und Impact-Effekte werden angedeutet, treten jedoch nicht klar hervor. Das Piano ist im Übergang gut getroffen und spielt sogar, wie in der Vorlage, nur vier Töne. Das Ende wirkt hingegen unschlüssig und ist bisher das Schwächste. **2/5**

Artefaktfreiheit: 3 / 5

Unnatürliche Sounds: Am Ende bricht der Song – ebenso wie der Swoosh – abrupt ab, obwohl in der Vorlage kein Swoosh vorhanden ist. Auch einzelne Hi-Hat- bzw. Shaker-Schläge zuvor wirken stellenweise deplatziert. **2.5/5**

Mixing / Balance: Das Saxophon steht, wie in der Vorlage, klar im Vordergrund, während das zweite Saxophon etwas zu leise abgemischt und nicht nach links gepannt ist. Insgesamt ist der Klang der Vorlage aber sehr gut herauszuhören. **4/5**

Nutzbarkeit: Das Material wäre nur dann gut nutzbar, wenn sich das Schlagzeug und die FX im zweiten Part sauber isolieren ließen. **2.5/5**

Kreativität: 2.5 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung hält sich in Grenzen. Das Schlagzeug wurde insgesamt reduziert, dafür wurde ein zweites Saxophon in den ersten Teil miteingebaut. **2/5**

Kohärenz: Das erzeugte Material wirkt insgesamt nicht vollständig kohärent, da die Einschübe des Schlagzeuges und der Schluss sich insgesamt unschlüssig anhören. **3/5**

- [Ergebnis 4.1](#) mit den Parametereinstellungen: W 75 / S 50 / A 50

Nähe zur Vorlage: 3.3 / 5

Instrumentierung: Vereinzelt sind Schlagzeug-Elemente wie Kick und Becken hörbar. Insgesamt nimmt sich das Schlagzeug jedoch ziemlich zurück. Zu Beginn des Stücks sind teilweise sogar Streicher zu hören. Das Piano ist nur im Übergang und im zweiten Teil zu hören. Beide Saxophone sind vorhanden, auch wenn das zweite Saxophon dieses Mal wieder sehr leise spielt. **3/5**

Melodie: Das zweite Saxophon spielt die Melodie sehr nah an der Vorlage. Die Begleitinstrumente weichen erneut leicht ab, während das Klavier erst ab dem zweiten Part einsetzt. **3.5/5**

Klang: Das melodiegebende Saxophon weicht etwas von dem Saxophon der Vorlage ab und klingt insgesamt etwas dumpfer. Dem Klang hilft es jedoch insgesamt, dass sich das Schlagzeug zurücknimmt, wodurch er näher an der Vorlage ist. **3.5/5**

Übergänge und Ende: Swoosh- und Impact-Effekte im Übergang sind vorhanden. Das Ende wirkt hingegen etwas ungewöhnlich und klingt leicht heruntergepitcht. **3.5/5**

Artefaktfreiheit: 4 / 5

Unnatürliche Sounds: Bis zum Ende konnten keine Artefakte festgestellt werden. Der Song bricht jedoch – wie bereits in mehreren Fällen – am Schluss unnatürlich ab. **4/5**

Mixing / Balance: Das Saxophon steht, wie in der Vorlage, klar im Vordergrund, während das zweite Saxophon etwas zu leise abgemischt und nicht nach links gepannt ist. Insgesamt ist der Klang der Vorlage aber sehr gut herauszuhören. **4/5**

Nutzbarkeit: Das Material würde sich, trotz der Abweichungen in der Instrumentierung, gut nutzen lassen, da es nahezu artefaktfrei ist. **4/5**

Kreativität: 3 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung hält sich trotz dem Hinzufügen von Schlagzeugelementen und Streichern in Grenzen. **2/5**

Kohärenz: Das erzeugte Material wirkt insgesamt kohärent und entwickelt sich schlüssig, nur das Ende fällt etwas heraus. **4/5**

- [Ergebnis 4.2](#) mit den Parametereinstellungen: W 75 / S 50 / A 50

Nähe zur Vorlage: 2.2 / 5

Instrumentierung: Erneut ist über den gesamten Song ein Schlagzeug zu hören. Bass und Klavier sind vorhanden. Zu Beginn sind teilweise mehrere Saxophone hörbar. Im zweiten Part treten zusätzlich vereinzelt Bläser-Effekte auf. **2.5/5**

Melodie: Das Klavier spielt bis zum Übergang eine neue Melodie, orientiert sich im Übergang jedoch an der Vorlage. Die Melodie des zweiten Saxophons, scheint im zweiten Part vom Bass übernommen zu werden, der nur ansatzweise die Melodie der Vorlage übernimmt. **2.5/5**

Klang: Der Klang des Saxophons im ersten Part überzeugt, jedoch irritieren der Bass, die zusätzlichen Effekte und der Klang des ersten Saxophons ab dem zweiten Part. Dadurch ist diese Version klanglich weit von der Vorlage entfernt. **2/5**

Übergänge und Ende: Swoosh- und Impact-Effekte im Übergang sind nicht vorhanden. Das Ende entspricht nur ungefähr der Vorlage. **2/5**

Artefaktfreiheit: 2.2 / 5

Unnatürliche Sounds: Im zweiten Part sind teilweise ungewöhnliche, schwer zuzuordnende Klänge zu hören. Zudem reißt das Ende erneut unnatürlich ab. **2.5/5**

Mixing / Balance: Der zweite Part wirkt stark höhenbetont und insgesamt unausgewogen. Dieses Ergebnis ist das erste, das tatsächlich deutlich schlecht abgemischt wirkt. **2/5**

Nutzbarkeit: Allenfalls der erste Part ließe sich in ein Projekt integrieren. **2/5**

Kreativität: 1 / 5

Musikalische Weiterentwicklung: Eine wirklich musikalische Weiterentwicklung ist nicht hörbar, obwohl viele Instrumente dazukommen. **1/5**

Kohärenz: Das Ergebnis wirkt nicht kohärent; die hinzugefügten Instrumente erscheinen eher beliebig zusammengestellt. **1/5**

- [Ergebnis 5.1](#) mit den Parametereinstellungen: W 100 / S 50 / A 50

Nähe zur Vorlage: 0.8 / 5

Instrumentierung: Das Saxophon ist gut hörbar. Insgesamt wechseln sich jedoch zahlreiche, teils schwer identifizierbare Instrumente ab. Synthie-Effekte sowie Bassanteile sind erkennbar. **1/5**

Melodie: Lediglich die Grundmelodie wurde übernommen. **1/5**

Klang: Klanglich weist lediglich das Saxophon noch eine erkennbare Nähe zur Vorlage auf. **1/5**

Übergänge und Ende: Der Übergang weicht vollständig von der Vorlage ab, ebenso ist das Ende völlig neu interpretiert. **0/5**

Artefaktfreiheit: 1 / 5

Unnatürliche Sounds: Im gesamten Song sind ungewöhnliche, schwer zuzuordnende Klänge zu hören. Die Instrumente besitzen teilweise schwerwiegende Artefakte. Zudem reißt das Ende erneut einfach so ab. **0/5**

Mixing / Balance: Insgesamt ist der Mix schwer zu bewerten, da das Ergebnis zahlreiche Artefakte aufweist und nur noch wenig mit der Vorlage gemein hat. Im Vergleich zur vorherigen Version ist es aber nicht höhenbetont und wird daher etwas besser bewertet. **3/5**

Nutzbarkeit: Das Material ist in dieser Form, ohne starke Audiotbearbeitung, in keiner Weise nutzbar. **0/5**

Kreativität: 1 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung ist nur in sehr geringem Maße vorhanden. Zwar bleibt die Grundmelodie ansatzweise erhalten, darüber hinaus entstehen jedoch keine klar nachvollziehbaren musikalischen Ideen oder sinnvolle Variationen. **1/5**

Kohärenz: Das Ergebnis wirkt insgesamt nicht kohärent. Die zahlreichen, teils schwer identifizierbaren Klänge und Instrumente wirken eher beliebig zusammengestellt. **1/5**

- [Ergebnis 5.2](#) mit den Parametereinstellungen: W 100 / S 50 / A 50

Nähe zur Vorlage: 0.8 / 5

Instrumentierung: Das Saxophon ist gut hörbar. Insgesamt wechseln sich jedoch zahlreiche, teils schwer identifizierbare Instrumente ab. Synthie-Effekte sowie Bassanteile sind erkennbar. **1/5**

Melodie: Lediglich die Grundmelodie wurde übernommen. **1/5**

Klang: Klanglich weist lediglich das Saxophon noch eine erkennbare Nähe zur Vorlage auf. **1/5**

Übergänge und Ende: Der Übergang weicht vollständig von der Vorlage ab, ebenso ist das Ende völlig neu interpretiert. **0/5**

Artefaktfreiheit: 1 / 5

Unnatürliche Sounds: Im gesamten Song sind ungewöhnliche, schwer zuzuordnende Klänge zu hören. Die Instrumente besitzen teilweise schwerwiegende Artefakte. Zudem reißt das Ende erneut einfach so ab. **0/5**

Mixing / Balance: Insgesamt ist der Mix schwer zu bewerten, da das Ergebnis zahlreiche Artefakte aufweist und nur noch wenig mit der Vorlage gemein hat. Im Vergleich zur vorherigen Version ist es aber nicht höhenbetont und wird daher etwas besser bewertet. **3/5**

Nutzbarkeit: Das Material ist in dieser Form, ohne starke Audibearbeitung, in keiner Weise nutzbar. **0/5**

Kreativität: 1 / 5

Musikalische Weiterentwicklung: Die musikalische Weiterentwicklung ist nur in sehr geringem Maße vorhanden. Zwar bleibt die Grundmelodie ansatzweise erhalten, darüber hinaus entstehen jedoch keine klar nachvollziehbaren musikalischen Ideen oder sinnvolle Variationen. **1/5**

Kohärenz: Das Ergebnis wirkt insgesamt nicht kohärent. Die zahlreichen, teils schwer identifizierbaren Klänge und Instrumente wirken eher beliebig zusammengestellt. **1/5**

4.3.2 Auswertung des Parameters Weiridness

Die Auswertung der fünf *Weiridness*-Durchläufe zeigt, dass niedrige bis mittlere Werte insgesamt die überzeugendsten Resultate liefern. Ab dem vierten Durchlauf fällt die Bewertung in den Haupteigenschaften *Nähe zur Vorlage* und *Kreativität* deutlich ab (siehe Abb. 26). Bei *Weiridness*-Werten von 0 bis 25 bleiben zentrale Merkmale der Vorlage weitgehend erhalten. Zwar treten bereits in diesem Bereich kleinere Abweichungen auf, etwa in Form des zusätzlich hinzugefügten Schlagzeugs, insgesamt bleiben die Ergebnisse jedoch gut nachvollziehbar und vergleichsweise nutzbar.

Auffällig ist, dass die Bewertungen in allen drei Hauptkriterien mit zunehmender Erhöhung des Parameters *Weiridness* abnehmen. Dies betrifft nicht nur die Nähe zur Vorlage und die Artefaktfreiheit, sondern auch die Kreativität. Zwar grenzen sich die Ergebnisse mit steigender *Weiridness* zunehmend von der Vorlage ab, gleichzeitig treten jedoch mehr Störgeräusche, inkohärente Entscheidungen und seltsame Melodien auf. Dadurch entstehen

keine überzeugenderen kreativen Weiterentwicklungen, sondern eher klanglich und strukturell unzufriedenstellende Resultate.

Besonders häufig scheinen Übergänge und Enden von diesen Artefakten betroffen zu sein. Dies zeigte sich auch innerhalb des Produktionsprozesses und konnte nun im systematischen Test erneut beobachtet werden. Viele Ergebnisse enden abrupt, wirken abgeschnitten oder klingen am Schluss unausgereift.

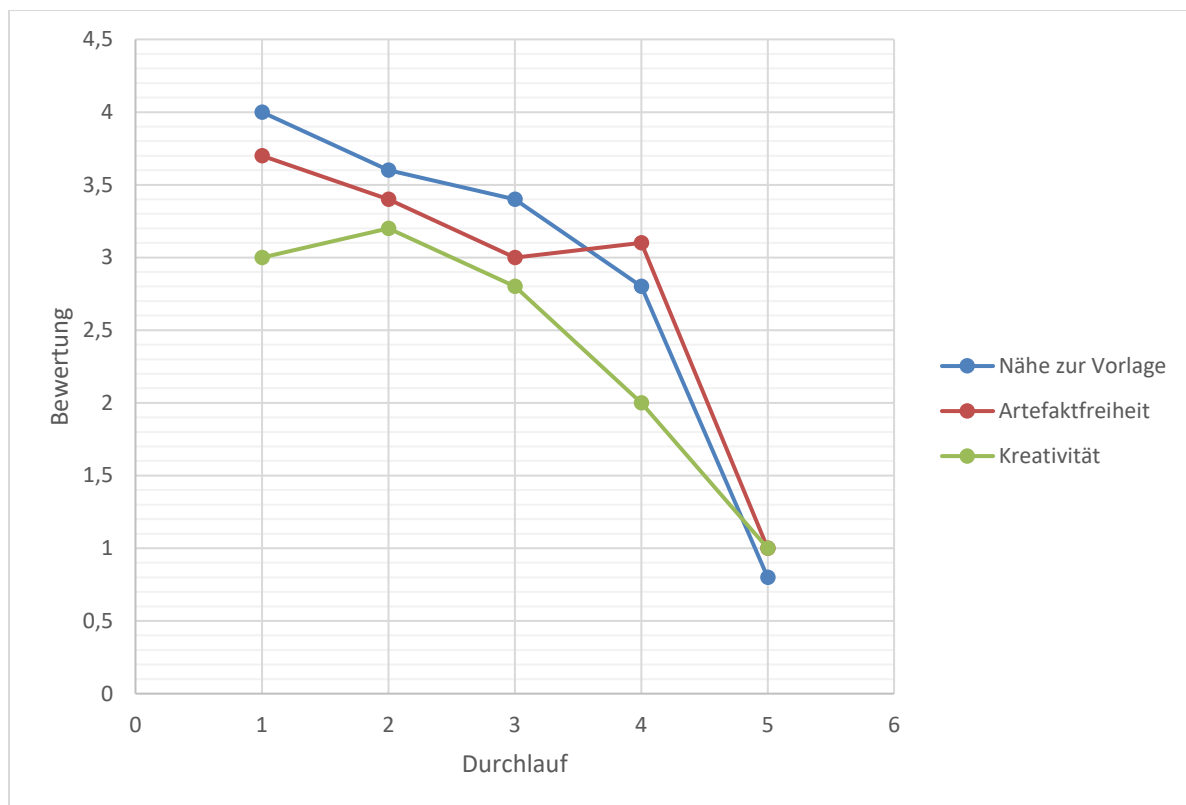


Abb. 26 – Der Kurvenverlauf zeigt den Durchschnitt der einzelnen Bewertungen pro Durchlauf und Hauptkriterium

Ab einem *Weirdness*-Wert von 75 nimmt die Abweichung von der Vorlage deutlich zu. Die Ergebnisse wirken weniger stabil, sowohl hinsichtlich der Instrumentierung als auch im Bereich Mixing, Kohärenz und Nutzbarkeit. Teilweise entstehen neue Instrumente oder Klanganteile, die sich nicht mehr sinnvoll aus der Vorlage ableiten lassen. Während einzelne Ergebnisse noch brauchbar erscheinen, zeigen andere bereits deutliche qualitative Einbrüche.

Bei einem *Weirdness*-Wert von 100 bricht die Orientierung an der Vorlage schließlich fast vollständig zusammen. Zwar bleibt in Form des ersten Saxophons noch ein melodischer Restbezug erhalten, insgesamt dominieren jedoch schwer identifizierbare Klänge und deutliche Artefakte. Die erzeugten Ergebnisse wirken damit nicht wie kreative Weiterentwicklungen der Vorlage, sondern eher wie unkontrollierte Neuinterpretationen mit klaren technischen Mängeln.

Auffällig ist darüber hinaus, dass die seitlich gepannten Streichinstrumente ab 0:12 in keiner der Generierungen übernommen wurden. Wodurch dieses Verhalten verursacht wird, konnte im Rahmen des Tests nicht geklärt werden. Daher fließt das Weglassen dieses Elements aber auch in die Bewertung nicht mit ein.

Im Anhang befindet sich die Tabelle 3, die alle Parametereinstellungen, Suno-Ergebnisse und die entsprechenden Bewertungen dieses Durchlaufs auflistet.

4.3.3 Einfluss des Parameters Style Influence

Der Parameter *Style Influence* wird darauf überprüft, wie stark der Stil der Vorlage übernommen wird und ob diese insgesamt erhalten und hörbar bleibt. Auch hier wird das Bewertungsschema erneut in drei Haupt- und mehrere Untereigenschaften aufgeteilt. Die drei Haupteigenschaften sind: **Nähe zur Vorlage** (Ist die Vorlage noch im neuen Ergebnis enthalten?), **Stilistische Eindeutigkeit** (Wie eindeutig passt das Ergebnis zum vorgegebenen Stil/Genre?) und **Nutzbarkeit** (Ist das Ergebnis verwendbar?). Die Untereigenschaften können somit teilweise aus dem *Weirdness*-Durchlauf zuvor übernommen werden. Die Untereigenschaften *Nähe zum ursprünglichen Genre* (Wie nah ist das Material am Stil des Originals?), *Typische Stilmerkmale vorhanden* (Sind Genre-typische Merkmale wie Drops bei EDM oder Pads bei Cinematic erhalten?) und *Konsistenz der Stilmittel* (Bleibt der gewählte Stil über die Zeit stabil?) sind dabei neu.

-
- [Ergebnis 1.1](#)^{*2} mit den Parametereinstellungen: W 50 / S 0 / A 50

Nähe zur Vorlage: 3.5 / 5

Instrumentierung: Die zentrale Instrumentierung bleibt erhalten (Bass, Saxophon, Klavier), wird jedoch durch ein zusätzliches Schlagzeug im ersten Part ergänzt. Das zweite Saxophon wird von einer Gitarre übernommen. **3/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine leicht veränderte Melodie besitzen. Die Gitarre spielt sehr nahe an der Vorlage des zweiten Saxophons, mit kleinen Variationen. **4/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song wurde aber um ein paar Instrumente ergänzt. **4/5**

Übergänge und Ende: Ein Swoosh- oder Impact-Sound im Mittelteil ist nicht vorhanden. Der Übergang wurde mit einer Gitarre etwas ausschräffiert. Das Piano spielt sehr ähnlich zur Vorlage. Der Bass-Drop am Ende fällt schwächer aus. Der modulierte Bläserklang bleibt weitgehend erhalten, ist jedoch etwas höher angesetzt als im Original. **3/5**

Stilistische Eindeutigkeit: 3.7 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. **5/5**

Konsistenz der Stilmittel: Im zweiten Part entfällt das Schlagzeug, was die Nähe zur Vorlage erhöht, jedoch innerhalb der Generierung zu einer inkonsistenten Umsetzung führt. **3/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und das Wegfallen der Riser wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter in Richtung eines poppig-jazzigen Sounds. **3/5**

Nutzbarkeit: 4.3 / 5

Unnatürliche Sounds: Die Hallfahne am Ende des Songs wirkt etwas künstlich und abrupt abgeschnitten. Außerdem sind ein Knacken und leichtes Rauschen wahrnehmbar. **3.5/5**

Kohärenz: Der Song ist insgesamt kohärent und die Instrumente sind stimmig eingesetzt. **5/5**

- [Ergebnis 1.2](#) mit den Parametereinstellungen: W 50 / S 0 / A 50

Nähe zur Vorlage: 3.8 / 5

Instrumentierung: Die zentrale Instrumentierung bleibt erhalten (Bass, Saxophon, Klavier), wird jedoch durch ein zusätzliches Schlagzeug ergänzt, das sich dieses Mal über den gesamten Song zieht. Das zweite Saxophon wird dieses Mal nicht durch eine Gitarre ersetzt. **4/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine leicht veränderte Melodie besitzen. Das zweite Saxophon spielt nah an der Vorlage. **4/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song wurde aber um ein Schlagzeug ergänzt. **4/5**

Übergänge und Ende: Ein Swoosh- oder Impact-Sound im Mittelteil fehlt. Der Übergang wird stattdessen leicht durch eine Gitarre hervorgehoben. Das Piano klingt eher nach einer künstlichen Violine. Der modulierte Bläserklang bleibt grundsätzlich erhalten, wirkt jedoch durch eine starke Anhebung im Hochfrequenzbereich deutlich verändert. **3/5**

Stilistische Eindeutigkeit: 4.3 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. **5/5**

Konsistenz der Stilmittel: Im zweiten Part entfällt das Schlagzeug dieses Mal nicht. So wird über die gesamte Laufzeit der poppige Stil des Songs eingehalten. **5/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und das Wegfallen der Riser wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter der Vorlage in Richtung eines poppig-jazzigen Sounds. Die Vorlage ist jedoch noch klar erkennbar und musikalisch bietet diese Version eher eine Evolution als eine Revolution. **3/5**

Nutzbarkeit: 4 / 5

Unnatürliche Sounds: Die Hallfahne am Ende des Songs wirkt etwas künstlich. In Minute 0:04 und 0:13 sind kleine Knackser zu hören. **3/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. **5/5**

- [Ergebnis 2.1](#) mit den Parametereinstellungen: W 50 / S 25 / A 50

Nähe zur Vorlage: 4 / 5

Instrumentierung: Die zentrale Instrumentierung bleibt erhalten (Bass, Saxophon, Klavier), wird jedoch durch ein zusätzliches Schlagzeug ergänzt, das sich erneut über den gesamten Song zieht. Das Klavier klingt sehr viel tiefer als in der Vorlage. Das zweite Saxophon wird nicht durch eine Gitarre oder ein anderes Instrument ersetzt. **4/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine leicht veränderte Melodie besitzen. Das zweite Saxophon orientiert sich an der Vorlage. **4/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht verlassen. Der Song wurde aber um ein Schlagzeug ergänzt. **4/5**

Übergänge und Ende: Ein Swoosh- oder Impact-Sound im Mittelteil fehlt. Beim Übergang steht das Piano klar im Vordergrund. Die Gitarre wirkt eher deplatziert. Der modulierte Bläserklang ist klanglich nahe am Original. **4/5**

Stilistische Eindeutigkeit: 4.3 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. **5/5**

Konsistenz der Stilmittel: Im zweiten Part entfällt das Schlagzeug nicht. So wird über die gesamte Laufzeit der poppige Stil des Songs eingehalten. **5/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und das Wegfallen der Riser wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter der Vorlage in Richtung eines poppig-jazzigen Sounds. Die Vorlage ist jedoch noch klar erkennbar und musikalisch bietet diese Version eher eine Evolution als eine Revolution. **3/5**

Nutzbarkeit: 4.8 / 5

Unnatürliche Sounds: Es sind keine besonderen Artefakte wahrnehmbar. Das Ende ist jedoch weiterhin abrupt abgeschnitten. Das kleine Schlagzeug Element am Ende bei 0:17 wird nicht als Artefakt bewertet. **4.5/5**

Kohärenz: Der Song ist insgesamt kohärent. Die Instrumente sind stimmig eingesetzt, und es treten keine ungewöhnlichen musikalischen Entscheidungen auf. **5/5**

- [Ergebnis 2.2](#) mit den Parametereinstellungen: W 50 / S 25 / A 50

Nähe zur Vorlage: 2.9 / 5

Instrumentierung: Das zentrale Saxophon, sowie der Bass bleiben erhalten. Das Piano klingt jedoch eher nach einem Synthesizer und der Song wurde wieder einmal mit einem Schlagzeug ergänzt. Das zweite Saxophon wird teilweise durch das Klavier übernommen. **2.5/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine stark veränderte Melodie besitzen. Die Melodie des zweiten Saxophons fließt in das Klavier/Synthesizer über. Beim Übergang fließt vom Bass eine neue Melodie mit ein. **3/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song geht durch das hinzugefügte Schlagzeug, die wegfallenden Riser und Impacts und vor allem durch den sehr flott spielenden Bass und die Synthie Effekte noch mehr in Richtung Jazz/Dance. **3.5/5**

Übergänge und Ende: Ein Swoosh- oder Impact-Sound im Mittelteil fehlt. Beim Übergang spielt der Bass eine Melodielinie. Das Piano rückt dadurch etwas in den Hintergrund. Das Ende ist durch den Vocal Sound fast neu interpretiert und nicht mehr nahe an der Vorlage. **2.5/5**

Stilistische Eindeutigkeit: 4.5 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. **5/5**

Konsistenz der Stilmittel: Der Song ist über die gesamte Laufzeit hinweg konsistent. **5/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs, des schnell spielenden Basses (vgl. Übergänge und Ende) und den Vocal-Effekten wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter der Vorlage in Richtung eines poppig-jazzigen Sounds. Die Vorlage ist noch erkennbar. **3.5/5**

Nutzbarkeit: 4.8 / 5

Unnatürliche Sounds: Das Ende wirkt erneut wie abgeschnitten. **4.5/5**

Kohärenz: Der Song ist insgesamt kohärent. Die Vocal Effekte am Ende werden als zusätzliches, neues und nicht inkohärentes Stilmittel bewertet. **5/5**

- [Ergebnis 3.1](#) mit den Parametereinstellungen: W 50 / S 50 / A 50

Nähe zur Vorlage: 3.3 / 5

Instrumentierung: Die zentrale Instrumentierung bleibt erhalten (Bass, Saxophon, Klavier), wird jedoch durch ein zusätzliches Schlagzeug ergänzt. Das zweite Saxophon im zweiten Part könnte klanglich auch eine Gitarre sein. **3/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine veränderte Melodie im ersten Part besitzen. Die Melodie des zweiten Saxophons wurde etwas abgeändert. **3/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song geht durch das hinzugefügte Schlagzeug und die wegfallenden Riser in Richtung Jazz/Dance. **4/5**

Übergänge und Ende: Ein Swoosh- oder Impact-Sound im Mittelteil fehlt. Beim Übergang spielen eine Gitarre und ein Piano. Das Piano rückt dadurch etwas in den Hintergrund. Am Ende ist ein Keyboard-ähnlicher Klang zu hören, der an die modulierten Bläser erinnert. **3/5**

Stilistische Eindeutigkeit: 4.3 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. **5/5**

Konsistenz der Stilmittel: Der Song ist über die gesamte Laufzeit hinweg konsistent. **5/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und des schnell spielenden Basses wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter der Vorlage in Richtung eines poppig-jazzigen Sounds. Die Vorlage ist jedoch noch klar erkennbar. **3/5**

Nutzbarkeit: 4.8 / 5

Unnatürliche Sounds: Das Ende wirkt erneut wie abgeschnitten. **4.5/5**

Kohärenz: Der Song ist insgesamt kohärent. **5/5**

- [Ergebnis 3.2](#) mit den Parametereinstellungen: W 50 / S 50 / A 50

Nähe zur Vorlage: 3.1 / 5

Instrumentierung: Das zentrale Saxophon, sowie der Bass bleiben erhalten. Das Piano klingt jedoch eher nach einem Synthesizer und der Song wurde mit einem Schlagzeug ergänzt. Der Part des zweiten Saxophons wird von einem Bass/Gitarre übernommen. **2.5/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine stark veränderte Melodie besitzen. Die Melodie des zweiten Saxophons fließt nahezu akkurat in die Gitarre/Bass über. **3.5/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song geht durch das hinzugefügte Schlagzeug und die wegfallenden Riser etwas mehr in Richtung Jazz/Dance. **4/5**

Übergänge und Ende: Ein Swoosh- oder Impact-Sound im Mittelteil fehlt, dafür ist ein Beckenschlag zu hören. Beim Übergang wechseln sich ein Piano und eine Gitarre ab. Am Ende sind Synthie Streicher zu hören. **2.5/5**

Stilistische Eindeutigkeit: 3.6 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. Das Piano ist jedoch etwas zu leise ausgeprägt, wodurch andere Generierungen insgesamt zielgerichteter wirken. **4/5**

Konsistenz der Stilmittel: Der Übergang wirkt durch das Hinzufügen zweier in dieser Form nicht vorhandener Instrumente etwas holprig. **4/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs, des schnell spielenden Basses und den Vocal Effekten wird der Song musikalisch leicht weiterentwickelt,

wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter der Vorlage in Richtung eines poppig-jazzigen Sounds. Die Vorlage ist jedoch noch klar erkennbar. **3/5**

Nutzbarkeit: 3.75 / 5

Unnatürliche Sounds: Das Ende wirkt erneut wie abgeschnitten und enthält ein Artefakt. **3.5/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. Leider fällt das Material aufgrund des abrupten Abschlusses und Artefakts negativ auf. **4/5**

- [Ergebnis 4.1](#) mit den Parametereinstellungen: W 50 / S 75 / A 50

Nähe zur Vorlage: 3.6 / 5

Instrumentierung: Die zentrale Instrumentierung bleibt erhalten (Bass, Saxophon, Klavier), wird jedoch durch ein zusätzliches Schlagzeug ergänzt, welches im zweiten Part zurückgenommen wird. Das zweite Saxophon wird durch eine Gitarre ersetzt. **3/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine stark veränderte Melodie besitzen. Die Melodie des zweiten Saxophons fließt nahezu akkurat in die Gitarre über. **4/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song geht durch das hinzugefügte Schlagzeug, die wegfallenden Riser und das flottere Piano, sowie Bass, etwas mehr in Richtung Jazz/Dance. **4/5**

Übergänge und Ende: Im Mittelteil ist ein Riser enthalten. Im Übergang spielen ein Piano, wie im Original und ein Bass. Der Bass-Drop vom Original wurde übernommen. Der modulierte Bläserklang klingt nur noch annähernd wie in der Vorlage. **3.5/5**

Stilistische Eindeutigkeit: 3.2 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. Insgesamt wirkt das Ende des Tracks etwas unpassend, da an dieser Stelle ein schwer definierbares und unpassendes Instrument einsetzt. **3.5/5**

Konsistenz der Stilmittel: Das Ende wirkt durch das Hinzufügen des neuen Instruments etwas holprig. **4/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und das Wegfallen der Riser wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter in Richtung eines poppig-jazzigen Sounds. **3/5**

Nutzbarkeit: 3.5 / 5

Unnatürliche Sounds: Der Riser und das Ende wirken unnatürlich abgeschnitten. Außerdem enthält das Ende einen seltsamen Klickton, der nicht nach einem Schlagzeug klingt. **3/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. Leider fällt das Material aufgrund des abrupten Abschlusses und des Klicktons negativ auf. **4/5**

- [Ergebnis 4.2](#) mit den Parametereinstellungen: W 50 / S 75 / A 50

Nähe zur Vorlage: 3.4 / 5

Instrumentierung: Die zentrale Instrumentierung bleibt erhalten (Bass, Saxophon, Klavier), wird jedoch durch ein zusätzliches Schlagzeug ergänzt. Das zweite Saxophon klingt ein bisschen nach einem Bass. **3.5/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine stark veränderte Melodie besitzen. Die Melodie des zweiten Saxophons wird übernommen. Der Übergang hingegen besitzt eine stark veränderte Melodie. **3.5/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song geht durch das hinzugefügte Schlagzeug, die wegfallenden Riser und das flottere Piano, sowie Bass, etwas mehr in Richtung Jazz/Dance. **4/5**

Übergänge und Ende: Der Übergang wird von einem Synthie getragen, der auch anders spielt als in der Vorlage. Der modulierte Bläserklang klingt nur noch annähernd wie in der Vorlage. Interessant hier ist, dass der Bläsereffekt auf die rechte Seite des Stereofelds wandert. **2.5/5**

Stilistische Eindeutigkeit: 4 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik, Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. **4/5**

Konsistenz der Stilmittel: Der Song ist über die gesamte Laufzeit hinweg konsistent. **5/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und das Wegfallen der Riser wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter in Richtung eines poppig-jazzigen Sounds. **3/5**

Nutzbarkeit: 4.8 / 5

Unnatürliche Sounds: Es ist kein Artefakt zu hören. Das Ende wirkt aber wie abgeschnitten. **4.5/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. **5/5**

- [Ergebnis 5.1](#) mit den Parametereinstellungen: W 50 / S 100 / A 50

Nähe zur Vorlage: 3.6 / 5

Instrumentierung: Die zentrale Instrumentierung bleibt erhalten (Bass, Saxophon, Klavier), wird jedoch durch ein zusätzliches Schlagzeug ergänzt. Das zweite Saxophon klingt ein bisschen nach einem Bass oder nach einer Gitarre. **3.5/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während Piano und Bass eine stark veränderte Melodie besitzen. Die Melodie des zweiten Saxophons wird weitestgehend übernommen. **4/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song geht durch das hinzugefügte Schlagzeug, die wegfallenden Riser und das flottere Piano, sowie Bass, etwas mehr in Richtung Jazz/Dance. **4/5**

Übergänge und Ende: Der Übergang wird von einer Gitarre getragen, das Piano im Hintergrund ist zu hören. Der modulierte Bläserklang klingt nur noch annähernd wie in der Vorlage. **3/5**

Stilistische Eindeutigkeit: 4 / 5

Typische Stilmerkmale vorhanden: Typische Stilmerkmale von akustischer Musik Jazz/Dance Musik wie Schlagzeug, Piano, Bass und Saxophon sind erhalten. **4/5**

Konsistenz der Stilmittel: Der Song ist über die gesamte Laufzeit hinweg konsistent. **5/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und das Wegfallen der Riser wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter in Richtung eines poppig-jazzigen Sounds. **3/5**

Nutzbarkeit: 4 / 5

Unnatürliche Sounds: Es ist kein Artefakt zu hören. Das Ende wirkt aber wie abgeschnitten und hat eine seltsam anmutende Lautstärkeschwankung. **4/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. Das Ende fällt jedoch etwas ab. **4/5**

-
- [Ergebnis 5.2](#) mit den Parametereinstellungen: W 50 / S 100 / A 50

Nähe zur Vorlage: 3.4 / 5

Instrumentierung: Bass und Saxophon sind klar zu hören, während das Klavier von Blasinstrumenten im Hintergrund ersetzt wurde, die auch im zweiten Teil den Part des zweiten Saxophons übernehmen. Außerdem wird wieder ein Schlagzeug über die gesamte Lauflänge hinzugefügt. **3/5**

Melodie: Die Melodie des ersten Saxophons ist weitgehend gleich wie in der Vorlage, während die zweizen Blasinstrumente und Bass eine veränderte Melodie besitzen. Auch im Übergang weicht die Generierung von der Vorlage ab. **3/5**

Nähe zum ursprünglichen Genre: Das ursprüngliche Genre wurde nicht gänzlich verlassen. Der Song geht durch das hinzugefügte Schlagzeug, die wegfallenden Riser und den flotten Bass, etwas mehr in Richtung Jazz/Dance. **4/5**

Übergänge und Ende: Im Mittelteil ist kein Riser enthalten und der Übergang wird von einem Synthie gespielt, der nur annähernd die Melodie, dafür aber größtenteils den Klang des Pianos übernimmt. Der modulierte Bläserklang klingt ähnlich zur Vorlage **3.5/5**

Stilistische Eindeutigkeit: 4 / 5

Konsistenz der Stilmittel: Der Song ist über die gesamte Laufzeit hinweg konsistent. **5/5**

Musikalische Weiterentwicklung: Durch den Einsatz des Schlagzeugs und das Wegfallen der Riser wird der Song musikalisch leicht weiterentwickelt, wirkt insgesamt flotter und verschiebt sich vom eher cinematischen Charakter in Richtung eines poppig-jazzigen Sounds. **3/5**

Nutzbarkeit: 4.8 / 5

Unnatürliche Sounds: Es ist kein Artefakt zu hören. Das Ende wirkt aber wie abgeschnitten. **4.5/5**

Kohärenz: Das erstellte Material ist kohärent und baut sinnvoll aufeinander auf. **5/5**

4.3.5 Auswertung des Parameters Style Influence

Die Auswertung der einzelnen Stufen des Style Influence Parameters zeigt keinen klaren Trend in eine gewisse Richtung. So nimmt die Nähe zur Vorlage bei Erhöhung des Paramaters nicht frappierend zu und auch die Qualität der Ergebnisse schwankt auf einem mittleren bis hohem Niveau. So zeigt dieser Versuch, im Gegenteil zum Parameter Weirdness bei Veränderung des Parameters Style Influence, keine bestimmte Lenkungswirkung auf. Während bei der Untersuchung des Weirdness Parameters ein rapider Abfall bei sukzessiver Erhöhung des Parameters in allen drei Hauptkategorien beobachtet werden konnte, verbleiben die Ergebnisse des Style Influence-Durchlaufs über das gesamte Spektrum auf einem weitgehend konstanten und verlässlichen Niveau. Die zu Beginn aufgestellte Hypothese, dass dieser Parameter den geringsten messbaren Einfluss auf das klangliche Resultat ausübt, lässt sich dadurch bestätigen.

Interessant ist außerdem, dass es in den neuen Bewertungskriterien wie *Nähe zum ursprünglichen Genre* und *Typische Stilmerkmale vorhanden* innerhalb der Versuchsreihe quasi zu gar keiner Schwankung kam. Das System scheint ein Bias in Richtung *Jazz/Dance*

oder *Pop-Jazz* zu besitzen, da es alle Generierungen in diese Richtung, vor allem durch das konsequente Hinzufügen des Schlagzeuges, interpretierte.

Im Anhang befindet sich die Tabelle 4 die alle Parametereinstellungen, Suno-Ergebnisse und die entsprechenden Bewertungen dieses Durchlaufs auflistet.

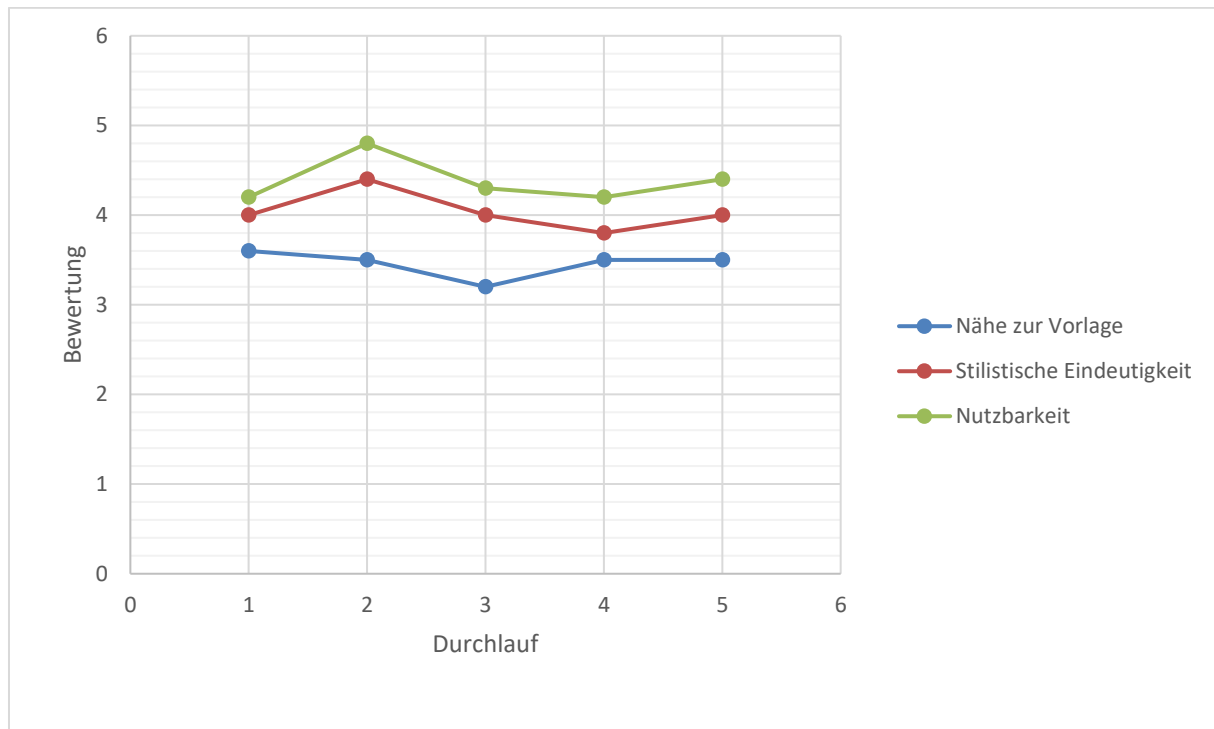


Abb. 27 – Der Kurvenverlauf zeigt den Durchschnitt der einzelnen Bewertungen pro Durchlauf und Hauptkriterium.

4.3.4 Einfluss des Parameters Audio Influence

Der Parameter Audio Influence wird darauf überprüft, wie stark der Klang der Vorlage übernommen wird und wie sehr die Nähe zur Vorlage erhalten bleibt. Hier wird das Bewertungsschema in zwei Haupt und mehrere Untereigenschaften aufgeteilt. Die zwei Haupteigenschaften sind: **Nähe zur Vorlage** (Ist die Vorlage noch im neuen Ergebnis enthalten?) und **Nutzbarkeit** (Ist das Ergebnis verwendbar?). Die entsprechenden Untereigenschaften sind *Melodische Übereinstimmung* (Wie deutlich wird die Melodie der Vorlage übernommen?), *Klangliche Übereinstimmung* (Wie nah ist das Ergebnis klanglich an der Vorlage), *Instrumentale Übereinstimmung* (Wie viele Instrumente wurden übernommen, oder hinzugefügt?), *Übergänge und Ende* (Wie nah sind die Übergänge und das Ende an der Vorlage?), *Unnatürliche Sounds* (Sind Artefakte oder seltsam anmutende Klänge wahrnehmbar?), *Kohärenz* (Wirkt das Ergebnis stimmig oder zusammengewürfelt?)

-
- [Ergebnis 1.1](#) mit den Parametereinstellungen: W 50 / S 50 / A 0

Nähe zur Vorlage: 3.3 / 5

Melodische Übereinstimmung: Die Melodie wurde in weiten Teilen übernommen. Insbesondere die Basslinie sowie beide Saxophonstimmen orientieren sich stark an der Vorlage und weisen nur geringe Abweichungen auf. Das Piano hingegen zeigt leichte melodische Variationen, wodurch die Gesamtübereinstimmung geringfügig reduziert wird. **4/5**

Klangliche Übereinstimmung: Das erste Saxophon wird klanglich überzeugend umgesetzt und kommt dem Original sehr nahe. Das zweite Saxophon ist hingegen klanglich weniger eindeutig und erinnert stellenweise eher an eine Gitarre. Der Bass hat eine leicht veränderte Tonhöhe, klingt jedoch noch ähnlich zur Vorlage. Das Piano klingt deutlich flächiger und erinnert eher an einen synthetischen Klang mit reduzierter Attack. Zusätzlich führt das hinzugefügte Schlagzeug dazu, dass sich die klangliche Balance verschiebt und zentrale Elemente wie Piano und Bass im Gesamtbild weiter in den Hintergrund treten. **3/5**

Instrumentale Übereinstimmung: Die grundlegenden Instrumente der Vorlage (Bass, Piano, erstes Saxophon) sind vorhanden. Allerdings wird erneut ein Schlagzeug hinzugefügt, das in der Vorlage nicht enthalten ist. Zudem ist das zweite Saxophon klanglich nicht eindeutig identifizierbar und könnte auch als Gitarre interpretiert werden. **3/5**

Übergänge und Ende: Der Übergang wird nicht durch Riser und Impact realisiert. Dafür werden Toms eingesetzt. Das Piano spielt in diesem Abschnitt fünf Töne anstelle der vier Töne der Vorlage. Der Bass-Drop ist kaum wahrnehmbar, und der modulierte Bläsersound wirkt leicht heruntergepitcht. **3/5**

Nutzbarkeit: 4.8 / 5

Unnatürliche Sounds: Das Ende ist abgeschnitten. Ansonsten sind keine Artefakte erkennbar. **4.5/5**

Kohärenz: Das Material ist musikalisch stimmig und sinnvoll aufeinander aufgebaut. **5/5**

- [Ergebnis 1.2](#) mit den Parametereinstellungen: W 50 / S 50 / A 0

Nähe zur Vorlage: 2.9 / 5

Melodische Übereinstimmung: Dieses Mal hat es den Anschein, dass das Piano kurz vor dem Übergang sehr hoch spielt und im ersten Teil des Stücks der Melodie des ersten Saxophons folgt. Dadurch wird sich nicht konsequent an die Grundmelodie gehalten. Das erste Saxophon orientiert sich nah an der Vorlage. Der Bass spielt etwas schneller als im Original. **3/5**

Klangliche Übereinstimmung: Das erste Saxophon wird klanglich nicht mehr so überzeugend umgesetzt wie zuvor. Es klingt weniger direkt und weniger hell als in der Vorlage. Das zweite Saxophon ist hingegen diesmal eindeutiger als solches identifizierbar. Der Bass klingt sehr ähnlich zur Vorlage. Das Piano erinnert eher an zusätzliche Bläser. Wie zuvor führt das hinzugefügte Schlagzeug zu Überdeckungen im Klangbild. **3.5/5**

Instrumentale Übereinstimmung: Die grundlegenden Instrumente der Vorlage (Bass, Piano, erstes und zweites Saxophon) sind vorhanden. Allerdings wird erneut ein Schlagzeug hinzugefügt, das in der Vorlage nicht enthalten ist, und das Klavier könnte auch als Bläser bzw. Synthesizer interpretiert werden. **3/5**

Übergänge und Ende: Der Übergang wird nicht durch Riser und Impact realisiert, sondern durch eine spielende Gitarre ersetzt. Das Piano rückt dadurch etwas in den Hintergrund. Der Bass-Drop sowie der Bass im Allgemeinen sind kaum wahrnehmbar, und der modulierte Bläsersound erinnert eher an eine bearbeitete Vocalspur. **2/5**

Nutzbarkeit: 4 / 5

Unnatürliche Sounds: Das Ende wirkt wenig überzeugend. Der hinzugefügte Vocal Sound erscheint deplatziert. Dafür wirkt das Material kaum abgeschnitten. **4/5**

Kohärenz: Durch das wenig überzeugende Ende ist das Material nicht zu vollständig stimmig. **4/5**

- [Ergebnis 2.1](#) mit den Parametereinstellungen: W 50 / S 50 / A 25

Nähe zur Vorlage: 3 / 5

Melodische Übereinstimmung: Der Bass spielt erneut schneller als in der Vorlage. Das Piano erinnert eher an zusätzliche Bläser oder einen entsprechenden Synthesizer, das sich melodisch nur wage an der Vorlage orientiert. Das erste Saxophon und zweite Saxophon sind melodisch sehr nahe an der Vorlage. **3.5/5**

Klangliche Übereinstimmung: Das erste Saxophon ist klanglich sehr nah an der Vorlage, ebenso das zweite Saxophon, wenn auch etwas tiefer. Der Bass ist klanglich auch sehr nah an der Vorlage. Das Piano klingt eher nach zusätzlichen Blasinstrumenten und erneut tritt ein Schlagzeug im Arrangement hervor. **3.5/5**

Instrumentale Übereinstimmung: Die Instrumente Bass und Saxophon wurden übernommen. Das Klavier wird jedoch überwiegend durch Blasinstrumente ersetzt, mit Ausnahme des Übergangs. Zudem wird erneut ein Schlagzeug hinzugefügt, das in der Vorlage nicht enthalten ist. **3/5**

Übergänge und Ende: Der Übergang wird erneut nicht durch Riser und Impact, sondern durch eine spielende Gitarre realisiert. Das Piano rückt dadurch etwas in den Hintergrund. Außerdem spielt und klingt es anders als in der Vorlage. Der Bass-Drop ist kaum wahrnehmbar, der Bass steht jedoch ähnlich wie im Suno-Original, mehr im Mittelpunkt. Der modulierte Bläsersound ist wenig durchsetzungsfähig und wirkt klanglich unsicher. **2.5/5**

Nutzbarkeit: 4 / 5

Unnatürliche Sounds: Das Ende wirkt wenig überzeugend. Klare Artefakte sind jedoch nicht hörbar. **4/5**

Kohärenz: Durch das wenig überzeugende Ende ist das Material nicht zu vollständig stimmig. **4/5**

- [Ergebnis 2.2](#) mit den Parametereinstellungen: W 50 / S 50 / A 25

Nähe zur Vorlage: 3 / 5

Melodische Übereinstimmung: Der Bass spielt erneut schneller als in der Vorlage. Auch das Piano weicht in seiner Spielweise ab. Das zweite Saxophon wird hier durch eine Gitarre ersetzt, die nah an der Vorlage bleibt. Das erste Saxophon orientiert sich sehr stark an der Vorlage. **3.5/5**

Klangliche Übereinstimmung: Das erste Saxophon ist erneut sehr gut getroffen, ebenso der Bass. Das Piano fällt leicht ab, bleibt jedoch insgesamt nah an der Vorlage. Bereits im ersten Part ist stellenweise eine Gitarre zu hören, und die Rolle des zweiten Saxophons wird nahezu vollständig von dieser übernommen. Außerdem – Wieder ein Schlagzeug vorhanden. **3/5**

Instrumentale Übereinstimmung: Teilweise wurde eine Gitarre ergänzt, ebenso erneut ein Schlagzeug hinzugefügt. Die übrigen Instrumente entsprechen weitgehend der Vorlage. **3/5**

Übergänge und Ende: Im Übergang ist ausschließlich eine Gitarre zu hören, während das Piano entfällt. Riser sind nicht vorhanden. Das Ende wirkt – wie bereits zuvor – wenig überzeugend. **2.5/5**

Nutzbarkeit: 4 / 5

Unnatürliche Sounds: Das Ende wirkt wenig überzeugend. Klare Artefakte sind jedoch nicht hörbar. **4/5**

Kohärenz: Durch das wenig überzeugende Ende ist das Material nicht zu vollständig stimmig. **4/5**

-
- [Ergebnis 3.1](#) mit den Parametereinstellungen: W 50 / S 50 / A 50

Nähe zur Vorlage: 3.5 / 5

Melodische Übereinstimmung: Der Bass spielt erneut deutlich schneller als in der Vorlage. Das erste Saxophon wurde wieder sehr gut übernommen. Das zweite Saxophon wurde

durch mehrere Bläser realisiert, wobei die Melodie nahezu identisch bleibt. Das Piano wirkt erneut deutlich verspielter als in der Vorlage. **4/5**

Klangliche Übereinstimmung: Durch den hinzugemischten, veränderten Piano-Sound verliert das erste Saxophon etwas den klanglichem Bezug zur Vorlage. Der Bass ist leicht höher als im Original. Das zweite Saxophon wirkt zu Beginn als wären es mehrere Bläser, die klanglich aber näher am original sind, wie in Beispielen zuvor hinzugefügte Gitarren. Erneut wurde erneut ein Schlagzeug hinzugefügt. **3/5**

Instrumentale Übereinstimmung: Bass, Saxophone und Piano sind vorhanden. Erneut wurde ein Schlagzeug hinzugefügt. **4/5**

Übergänge und Ende: Im Übergang sind keine Riser vorhanden, stattdessen ist eine Gitarre zu hören. Das Piano erinnert eher an synthetische Streicher. Der Bass-Drop ist dieses Mal hörbar. Das Saxophon fällt etwas aus dem Gesamtklang heraus, während der modulierte Bläserklang sehr nah am Original bleibt. **3/5**

Nutzbarkeit: 4 / 5

Unnatürliche Sounds: Bei Minute 0:05 tritt kurzzeitig ein Aussetzer im Signal auf. Das Saxophon am Ende weist eine ungewöhnliche Lautstärkeschwankung auf. Der Song endet jedoch weniger abrupt als in anderen Beispielen. **3.5/5**

Kohärenz: Das Material ist musikalisch stimmig und sinnvoll aufeinander aufgebaut. Einen kleinen Abzug gibt es für die klangliche Unstimmigkeit am Ende im Saxophonklang. **4.5/5**

-
- [Ergebnis 3.2](#) mit den Parametereinstellungen: W 50 / S 50 / A 50

Nähe zur Vorlage: 3.9 / 5

Melodische Übereinstimmung: Das erste Saxophon folgt erneut weitgehend der ursprünglichen Spielweise. Der Bass sowie das Klavier weisen wieder ein etwas erhöhtes Tempo auf. Der Bass spielt aber vor allem im zweiten part sehr nah an der Vorlage. Das zweite Saxophon orientiert sich an der Vorlage. **3.5/5**

Klangliche Übereinstimmung: Das erste Saxophon wirkt klanglich weniger voll als in der Vorlage. Der Bass ist hingegen tonal sehr nah am Original. Das Piano erinnert ansatzweise

an den ursprünglichen Klang. Das zweite Saxophon erscheint im Vergleich leicht höher. Positiv fällt auf, dass das Schlagzeug im zweiten Part nicht mehr einsetzt. **4/5**

Instrumentale Übereinstimmung: Die Instrumente Bass, Klavier und Saxophon wurden übernommen. Zudem wird erneut ein Schlagzeug hinzugefügt, das sich in der zweiten Hälfte aber zurücknimmt. **4/5**

Übergänge und Ende: Im Übergang ist ein dezenter Riser, aber kein Impact enthalten. Es spielt jedoch eine Gitarre, statt einem Piano. Das Ende zeigt sich Artefakt behaftet (vgl. Unnatürliche Sounds), orientiert sich jedoch in Aufbau und Klangbild deutlich an der Vorlage. **4/5**

Nutzbarkeit: 3.5 / 5

Unnatürliche Sounds: Das Ende klingt unsauber, unnatürlich und reist aus dem Hörfluss heraus. Es ist außerdem abgehackt. **3/5**

Kohärenz: Durch das wenig überzeugende Ende ist das Material nicht zu vollständig stimmig. **4/5**

- [Ergebnis 4.1](#) mit den Parametereinstellungen: W 50 / S 50 / A 75

Nähe zur Vorlage: 4 / 5

Melodische Übereinstimmung: Im ersten Part weichen Piano und Bass erneut deutlich von der Vorlage ab, da sie in einem höheren Tempo spielen. Im zweiten Part wurden Bass und Piano jedoch sehr präzise umgesetzt. Das erste und zweite Saxophon hingegen orientieren sich nahezu identisch an der Vorlage. **4/5**

Klangliche Übereinstimmung: Das erste Saxophon klingt minimal heller als im Original, während das zweite Saxophon im Verhältnis zu den anderen Versuchen sehr treffend umgesetzt ist. Der Bass ist etwas zu leise abgemischt, was die Bewertung erschwert, liegt insgesamt aber recht nah an der Vorlage. Auch das Piano kommt dem Original gut nahe, wirkt stellenweise jedoch eher wie eine Gitarre. Positiv fällt auf, dass sich das Schlagzeug im zweiten Part zurücknimmt, was dem Gesamtklang zugutekommt. **4/5**

Instrumentale Übereinstimmung: Die grundlegenden Instrumente der Vorlage (Bass, Piano, erstes und zweites Saxophon) sind vorhanden. Allerdings wird erneut ein Schlagzeug

hinzugefügt, das in der Vorlage nicht enthalten ist. Dieses nimmt sich im zweiten Teil jedoch zurück- **4/5**

Übergänge und Ende: Ein Riser im Übergang fehlt, ein Impact ist dafür vorhanden. Der modulierte Bläserklang wurde sehr gut nachgebildet, wirkt jedoch (vgl. Unnatürliche Sounds) etwas deplatziert. Bass und Bass Drop sind ähnlich zum Original. **4/5**

Nutzbarkeit: 3.5 / 5

Unnatürliche Sounds: Das Ende klingt seltsam unnatürlich und deplatziert. Ansonsten konnten keine Artefakte festgestellt werden. **3/5**

Kohärenz: Durch das wenig überzeugende Ende ist das Material nicht zu vollständig stimmig. **4/5**

- [Ergebnis 4.2](#) mit den Parametereinstellungen: W 50 / S 50 / A 75

Nähe zur Vorlage: 3.3 / 5

Melodische Übereinstimmung: Das erste Saxophon ist erneut sehr nah an der Vorlage. Bass und Piano spielen hingegen wieder in einem schnelleren Tempo. Im zweiten Part nähert sich das Ergebnis bis zum Ende hinzunehmend der originalen Melodieführung an, bevor es anschließend deutlich einbricht. **2.5/5**

Klangliche Übereinstimmung: Das erste Saxophon klingt weniger klar als im Original. Das zweite ist eindeutig als Saxophon erkennbar, wirkt jedoch insgesamt etwas zurückhaltend. Bass und Piano orientieren sich sehr nah an der Vorlage, auch wenn das Piano stellenweise etwas zu leise ist. Positiv hervorzuheben ist, dass sich das Schlagzeug im zweiten Part zurücknimmt, was der Nähe zur Vorlage zugutekommt. **4/5**

Instrumentale Übereinstimmung: Die grundlegenden Instrumente der Vorlage (Bass, Piano, erstes und zweites Saxophon) sind vorhanden. Allerdings wird erneut ein Schlagzeug hinzugefügt, das in der Vorlage nicht enthalten ist. Dieses nimmt sich im zweiten Teil jedoch zurück. **4/5**

Übergänge und Ende: Riser und Impact im Übergang sind vorhanden. Es spielt ein Piano und eine Gitarre. Der modulierte Bläserklang wurde versucht nachzubilden, wobei leider üble Artefakte entstanden sind (vgl. Unnatürliche Sounds). **2.5/5**

Nutzbarkeit: 2.3 / 5

Unnatürliche Sounds: Das Ende ist in dieser Form nicht tolerierbar und das Material müsste stark bearbeitet oder das Ende ausgetauscht werden. Das restliche Material scheint so weit artefaktfrei zu sein, klingt aber künstlich und wenig überzeugend. **1.5/5**

Kohärenz: Durch das mangelhafte Ende ist das Material nicht bis zum Schluss hin kohärent. **3/5**

- [Ergebnis 5.1](#) mit den Parametereinstellungen: W 50 / S 50 / A 100

Nähe zur Vorlage: 3 / 5

Melodische Übereinstimmung: Das erste Saxophon orientiert sich melodisch eng an der Vorlage. Bass und Piano weisen jedoch erneut ein höheres Spieltempo auf. Im zweiten Abschnitt wirkt das zweite Saxophon unsauber, als würde es sich verspielen. Der Bass ist im zweiten Part hingegen nah an der Vorlage. **3/5**

Klangliche Übereinstimmung: Das Klavier weicht klanglich von der Vorlage ab und klingt eher, wie zusätzliche Blasinstrumente. Auch das erste Saxophon klingt als bestünde es aus mehreren Stimmen. Zudem wurde erneut ein Schlagzeug ergänzt, das sich im zweiten Part etwas zurücknimmt. Der Bass erscheint insgesamt höher. Das zweite Saxophon klingt ähnlich zur Vorlage. Insgesamt ist die Ähnlichkeit zur Vorlage vor allem im zweiten Part, durch das Wegfallen des Schlagzeugs gut erkennbar. **3.5/5**

Instrumentale Übereinstimmung: Die Instrumente der Vorlage Bass und Saxophon sind vorhanden. Allerdings wird erneut ein Schlagzeug hinzugefügt, das in der Vorlage nicht enthalten ist. Dieses nimmt sich im zweiten Teil jedoch zurück. Außerdem sind Blasinstrumente hörbar, die in der Vorlage ebenfalls nicht enthalten sind. **3/5**

Übergänge und Ende: Im Übergang ist kein Riser vorhanden. Gitarre und Piano sind zwar hörbar, dafür tritt ein perkussives Schlagwerk auf. Abgesehen vom Piano unterscheidet sich der Übergang deutlich von der Vorlage. Auch das Ende wirkt ungewöhnlich und klingt stellenweise leicht verstimmt. **2.5/5**

Nutzbarkeit: 2 / 5

Unnatürliche Sounds: Wie bereits erwähnt, wirkt das Ende klanglich leicht verstimmt. Es handelt sich dabei zwar nicht um ein Artefakt, dennoch entsteht ein unnatürlicher und künstlicher Eindruck. Außerdem klingen die Instrumente insgesamt eher unausgereift und künstlich. **2/5**

Kohärenz: Das zweite Saxophon verspielt sich am Ende, das Ende wirkt künstlich und insgesamt klingt das Resultat nicht stimmig, sondern eher wie zusammengewürfelt. **2/5**

- [Ergebnis 5.2](#) mit den Parametereinstellungen: W 50 / S 50 / A 100

Nähe zur Vorlage: 2.5 / 5

Melodische Übereinstimmung: Das erste Saxophon ist erneut sehr präzise getroffen. Das zweite Saxophon wirkt am Ende wieder leicht unsauber bzw. verspielt. Piano und Bass spielen erneut in einem höheren Tempo und ähneln damit den vorherigen Beispielen. **3/5**

Klangliche Übereinstimmung: Das Klavier weicht klanglich von der Vorlage ab. Der Bass und das erste Saxophon orientieren sich an der Vorlage. Zudem wurde erneut ein Schlagzeug ergänzt, das sich im zweiten Part etwas zurücknimmt. Insgesamt ist die Ähnlichkeit zur Vorlage vor allem im zweiten Teil vorhanden. **2.5/5**

Instrumentale Übereinstimmung: Teilweise sind Bläser zu hören, die in der Vorlage nicht enthalten sind. Zudem wurde erneut ein Schlagzeug ergänzt, das sich im zweiten Part etwas zurücknimmt. **3/5**

Übergänge und Ende: Im Übergang ist ein leichter Riser und Impact und auch ein Piano vorhanden. Zusätzlich wurde aber eine Hi-hat und Gitarre dazugemischt. Das Ende ist stark ungewöhnlich und klingt verstimmt. **1.5/5**

Nutzbarkeit: 2 / 5

Unnatürliche Sounds: Wie bereits erwähnt, wirkt das Ende klanglich massiv verstimmt, unnatürlich und schräg. **2/5**

Kohärenz: Das zweite Saxophon verspielt sich am Ende, das Ende wirkt künstlich und insgesamt klingt das Resultat nicht stimmig, sondern eher wie zusammengewürfelt. **2/5**

4.3.6 Auswertung des Parameters Audio Influence

Die Auswertung des Parameters *Audio Influence* zeigt, dass mit steigender Ausprägung des Parameters die Nähe zur Vorlage kurz ab-, tendenziell bis zu einem Wert von 75% jedoch zunimmt. Bei einer weiteren Erhöhung auf 100 % ist hingegen ein deutlicher Kurvenabfall zu beobachten. Vor allem das Wegfallen des Schlagzeugs führt zunächst zu einer deutlich erhöhten Nähe zur Vorlage. Im weiteren Verlauf bewirkt jedoch die zunehmende Anzahl an Artefakten sowie die klanglich unsaubereren Instrumente eine Abnahme dieser Übereinstimmung. Die Nutzbarkeit der erzeugten Ergebnisse verschlechtert sich stetig. Während zunächst lediglich das Ende des Stücks und die Übergänge betroffen sind, nimmt ab einem Wert von 75 % auch die Qualität der einzelnen Instrumente selbst ab. Infolgedessen sinkt die Bewertung schließlich auf Werte von 3 beziehungsweise 2. Die tiefen Werte der Nutzbarkeit ab einer Parametereinstellung von 75 % deuten darauf hin, dass Suno bei hohen Parameterausprägungen (vgl. *Weirdness*-Parameter) vermehrt schwer nachvollziehbare Entscheidungen trifft. Allerdings war dieser Verlauf bei dem Parameter *Style Influence* nicht bemerkbar. Der Testlauf zeigt, dass sich Suno mit steigenden Parameterwerten zunehmend der Vorlage annähert, was bis zu einem Wert von 75 % gelingt. Dies geht jedoch mit einer abnehmenden Eigenständigkeit einher, da das System verstärkt versucht, den Originalklang zu imitieren, anstatt auf erlernte, variantenreiche Klangmuster zurückzugreifen, wodurch die Qualität abnimmt.

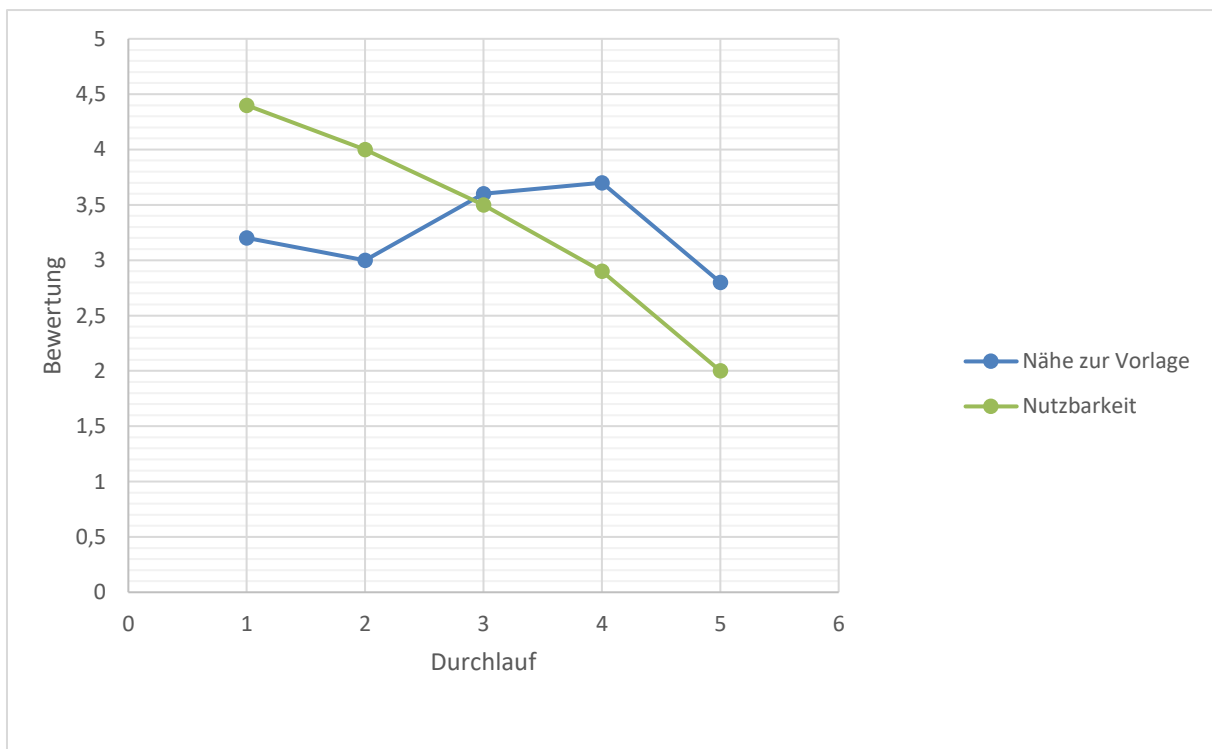


Abb. 28 – Der Kurvenverlauf zeigt den Durchschnitt der einzelnen Bewertungen pro Durchlauf und Hauptkriterium.

5 Abschlussbetrachtung

5.1 Zusammenfassung der Ergebnisse

Im praktischen Umgang mit Suno zeigte sich, dass das System nur eingeschränkt auf textliche Anweisungen reagiert (vgl. S. 23). Dabei hängt die Zuverlässigkeit stark von der Art und Qualität der Vorgaben ab. Während konkrete Prompts wie *no reverb and effects* sowie exakte Anforderungen an Tonhöhen oder Melodieverläufe häufig nicht präzise umgesetzt wurden, erwiesen sich Angaben zur Instrumentierung oder zum Klangcharakter als deutlich verlässlicher. Um die Qualität der Ergebnisse zu erhöhen, müssen Prompts ausführlich formuliert werden (vgl. S.45 ff.) und mithilfe der Parameter *Weirdness*, *Style Influence* und *Audio Influence* unterstützt werden (vgl. S.32). So lässt sich deutlich gezielter Einfluss auf die Ergebnisse nehmen.

Hinsichtlich der Klangqualität zeigen viele der erzeugten Ergebnisse einen hohen Detailgrad, insbesondere in den Bereichen Phrasierung, Dynamik und allgemeine Audioqualität (vgl. S. 29 und S.31). In vielen Fällen übertrafen die generierten Audiospuren die Erwartungen und wirkten im Vergleich zu klassischen MIDI-Produktionen deutlich natürlicher. Wurden diese Spuren noch mit MIDI-Instrumenten kombiniert, ließen sie sich sehr gut in das bestehende Projekt integrieren (vgl. S.45).

Dennoch war in vielen Fällen eine nachträgliche Bearbeitung erforderlich. Audiospuren mussten beispielsweise mittels Audio-Wrap (vgl. S.24), durch Schnitt (vgl. S.30) oder durch das Kombinieren mehrerer Takes (vgl. S.32,33 und S.46) angepasst werden, um einem höheren Qualitätsanspruch gerecht zu werden. Auch innerhalb der finalen Produktion verbleiben vereinzelt hörbare Artefakte (vgl. Min 2:42), die im späteren Mixing bearbeitet werden müssten. Besonders betroffen waren dabei, sowohl im Produktionsprozess als auch im systematischen Test, der Nachhall (vgl. S.42 und S.46) der einzelnen Spuren, der häufig unnatürlich wirkte oder abrupt abbrach. Ein zentrales Merkmal des Systems ist zudem die fehlende Reproduzierbarkeit von Ergebnissen (vgl. S.27 ff. und S. 37 ff.). Identische Eingaben führen nicht zu identischen Outputs, was den Workflow maßgeblich beeinflusst. Im Gegensatz zu klassischen VST-Instrumenten ist es nicht möglich, eine Spur an anderer Stelle im Projekt mit exakt demselben Klangbild fortzuführen. Selbst bei identischem Prompt entstehen stets leicht andere Ergebnisse. Auch lässt sich aus einer generierten Stereospur, nicht wieder auf die einzelnen Spuren zurückschließen.

Funktionen wie *Get Stems* und *Remove FX* erwiesen sich in der Praxis als hilfreiche Werkzeuge, um Instrumente voneinander zu trennen oder Effekte zu reduzieren (vgl. S.31). Die Qualität dieser Trennung variierte jedoch je nach Ausgangsmaterial und Versuch deutlich (vgl. S.46).

Als kreatives Werkzeug zur Ideenfindung überzeugte Suno in hohem Maße. Insbesondere die Cover (vgl. S.24) und Remix-Funktion (vgl. S. 42) ermöglichten es, bestehende musikalische Ansätze in neue stilistische Richtungen zu überführen. Dadurch kann der zeitliche Aufwand für die Suche nach Inspiration im Produktionsprozess reduziert werden.

Die Ergebnisse der Parameteruntersuchung zeigen, dass insbesondere bei den Parametern *Weirdness* und *Audio Influence* ab hohen Werten von 75% vermehrt Artefakte und Inkonsistenzen auftreten. Anders verhält es sich bei dem Parameter *Style Influence*. Dieser zeigt im Test kaum eine messbare Lenkungswirkung. Die Ergebnisse bleiben über das gesamte Parameterspektrum hinweg nahezu konstant, was die Annahme eines geringen Einflusses in der Hypothese (vgl. S. 49) bestätigt. In der praktischen Anwendung erweist es sich daher als sinnvoll, mehrere Parameterkombinationen hintereinander auszuprobieren, da eine präzise Lenkungswirkung durch die Veränderung der einzelnen Parameter nicht vollständig gegeben ist (vgl. S.48).

Auffällig ist zudem eine Tendenz des Systems, unabhängig von der Parameterausprägung ähnliche Richtungen zu bevorzugen. Dies äußerte sich bereits im Produktionsprozess durch wiederkehrende Elemente, wie etwa das häufige Hinzufügen eines Schlagzeugs und von Impacts, sowie Effekten (vgl. S. 24) und bestätigte sich erneut in den Testläufen. Suno ergänzte trotz eines fehlenden Schlagzeugs in der Vorlage in allen! Fällen ein Schlagzeug und ignorierte die gepannten Streichinstrumente vollständig (vgl. S.52).

Dies könnte darauf hindeuten, dass das System im Rahmen seines Trainings darauf optimiert ist, vollständige Songstrukturen zu erzeugen und dabei typische Elemente wie ein Schlagzeug automatisch hinzufügt (vgl. Kapitel 2.2.1 – Maschinelles Lernen).

5.2 Fazit

Die vorliegende Bachelorarbeit hatte das Ziel, den Einsatz künstlicher Intelligenz mithilfe des generativen KI-Systems Suno anhand einer Songproduktion praktisch und wissenschaftlich einzuordnen. Dabei stand die Frage im Zentrum, inwiefern KI-gestützte Tools einzelne Arbeitsschritte sinnvoll unterstützen oder sogar ersetzen können und wo deren praktische und qualitative Grenzen liegen.

Kapitel 2 beschreibt, dass Musik insgesamt weniger beliebt ist, wenn bekannt ist, dass es sich dabei um KI-Musik handelt. Die Produktion des vorliegenden Musikstücks wurde maßgeblich durch den Einsatz künstlicher Intelligenz (KI) gestützt, hätte ohne eine menschliche Intervention jedoch nicht die finale Qualität erreicht. Häufig mussten KI-generierte Einzelspuren, vor allem bei sog. Hallfahnen nachbearbeitet werden, um Artefakte und Störgeräusche zu reduzieren. Dies entspricht der Arbeit eines Toningenieurs, denn auch bei traditionellen Studioaufnahmen treten unvorhergesehene Fehler auf, die erst in der Postproduktion identifiziert und korrigiert werden können. Der Unterschied bleibt hier lediglich das durch generative KI erzeugte Material. Bezieht man die Ergebnisse auf die in Kapitel 2.1 angeführte Definition von Elaine Rich, wonach Menschen der KI (noch) überlegen sind, muss nach Auswertung der Ergebnisse festgehalten werden, dass die KI-generierten Spuren (insbesondere Streicher und Saxophon) als klanglich überzeugender und organischer im Vergleich zu herkömmlichen VST-Instrumenten wahrgenommen wurden. Auch wenn ein direkter systematischer Vergleich zwischen MIDI und KI nicht primärer Bestandteil dieser Arbeit war, bewies Suno eine hohe klangliche Kompetenz in der Erzeugung einzelner Instrumente. Durch die wahrscheinlichsbezogene Arbeitsweise von KI-Modellen (vgl. Kapitel 2.5.5) ergaben sich auch häufig vom Nutzer zwar nicht intendierte, jedoch sehr inspirierende und variantenreiche Ergebnisse, die im weiteren Verlauf der Produktion genutzt wurden und den kreativen Prozess massiv unterstützten.

Kapitel 3 zeigte aber auch, dass Systeme wie Suno in einem künstlerischen und juristischen Spannungsfeld stehen und die Nutzungsrechte der generierten Inhalte nicht vollständig beim Nutzer liegen. Eigene menschliche Beiträge können die Schutzfähigkeit durch das Urheberrecht aber erhöhen.

Im Produktionsprozess agierte die KI primär als generative Inspirationsquelle innerhalb einer abwechselnden Interaktion zwischen Menschen und Maschine. Eine wechselseitige Arbeitsweise bei der häufig MIDI-Daten als kreative Vorgabe für die KI diente, um über die Cover-Funktion neue musikalische Variationen zu erschaffen, stellte sich als eine sinnvolle Arbeitsweise heraus. Durch die in Kapitel 2 näher beschriebene, nicht deterministische

Arbeitsweise von KI, wurde es jedoch immer dann problematisch, wenn ab einem Punkt mit einem gewissen Klang weitergearbeitet werden sollte.

Artefakte und Störgeräusche konnten nicht immer zuverlässig entfernt werden (vgl. S. 44), jedoch half die *Get Stems* Funktion in vielen Fällen weiter. Abgeschnittene Hallfahnen wurden mit zugemischtem Hall innerhalb der DAW überdeckt.

Als zusätzliches Werkzeug innerhalb einer Produktion konnte das Modell Suno insgesamt jedoch größtenteils überzeugen. Gerade zur Erzeugung nicht elektronischer Instrumente oder Signale wie dem Chor, oder dem Saxophon eignet sich das System sehr gut. Eine gezielte Steuerung, sowie es MIDI-Daten ermöglichen, ist aber aufgrund der nicht deterministischen Arbeitsweise von KI-Modellen nicht möglich. Um das KI-System präzise zu steuern ist eine gute Kombination aus sinnvollen und ausführlichen Text-Prompts, den richtigen Parametereinstellungen und eine klanglich klar definierte MIDI- oder Audio-Vorlage notwendig. Trotzdem erzeugte Suno nicht immer sinnvolle oder passende Ergebnisse (vgl. S. 43 ff.).

Die Untersuchung in Kapitel 4 unterstreicht dies, indem sie aufzeigt, dass eine Steuerung der Ergebnisse grundsätzlich möglich ist, jedoch nicht linear oder immer vorhersehbar. Während eine Erhöhung von *Audio Influence* zu einer tendenziell stärkeren Annäherung an die Vorlage führte, und eine Erhöhung des *Weirdness* Faktors die Nähe zur Vorlage minimierte, war der Einfluss des Parameters *Style Influence* kaum spürbar. In der Praxis müssen also aufgrund der fehlenden Präzision mehrere Kombinationen ausprobiert werden und die Qualität der Ergebnisse bleibt am Ende, neben einer eher moderaten Einstellung der Schieberegler, auch vom Zufall abhängig.

Kapitel 3.2 und 4 zeigen aber auch, dass das System wiederkehrende, fast vorhersehbare Verhaltensmuster aufweist. Dazu zählt unter anderem das automatische Hinzufügen von Instrumenten wie dem Schlagzeug. Diese Verhaltensmuster können nur mit der oben ausgeführten Methode (gezielte Textprompts, definierte MIDI/Audio-vorlagen und passende Parametereinstellungen) durchbrochen werden.

Die zentrale Forschungsfrage kann also dahingehend beantwortet werden, dass Suno ein sinnvolles Werkzeug zur Erweiterung der eigenen Arbeitsweise darstellt, das vor allem den Einsatz einzelner VST-Instrumente ersetzen kann, insgesamt aufgrund häufig auftretender Fehler, die Rolle des Menschen aber unersetzbar bleibt. Die KI übernimmt eher die Rolle eines kreativen Partners, und nicht die eines ausgereiften Produzenten, da sie noch zu fehleranfällig ist und nicht zuverlässig zwischen sinnvollen Generierungen und ungeeigneten Ergebnissen unterscheiden kann.

V Literaturverzeichnis

ad-hoc-news.de. (2025, 27. Dezember). *NVIDIA und der KI-Boom: Warum die GPUs der GeForce- und H100-Generation das Herz der Tech-Revolution sind*. <https://www.ad-hoc-news.de/boerse/news/ueberblick/nvidia-und-die-ki-revolution-warum-die-gpus-hinter-chatgpt-and-co-zum/68435592>

(Aufgerufen am 14.02.26)

App Store. (2024). *Suno - AI Songs & Music App* - App Store. <https://apps.apple.com/us/app/suno-ai-songs-music/id6480136315>

(Aufgerufen am 14.02.26)

App Store. (2025). *Udio: AI Music Maker & Studio App*. <https://apps.apple.com/us/app/udio-ai-music-maker-studio/id6511211165>

Auron. (o.D.). *Steinberg*.

https://www.steinberg.help/r/halion/7.1/de/halion/topics/included_instruments/auron_c.html

Beauchene, V., Duranton, S., Kalra, N., & Martin, D. (2025, 26. Juni). *AI at work 2025: Momentum builds, but gaps remain*. Boston Consulting Group.

<https://www.bcg.com/publications/2025/ai-at-work-momentum-builds-but-gaps-remain?utm>

(Aufgerufen am 20.03.26)

Bergmann, D. (2026a). *Was ist maschinelles Lernen?*. IBM. <https://www.ibm.com/de-de/think/topics/machine-learning>

(Aufgerufen am 19.03.26)

Bergmann, D., Stryker, C. (2026b). *Was ist Backpropagation?*. IBM, <https://www.ibm.com/de-de/think/topics/backpropagation#741977107>

(Aufgerufen am 27.02.26)

Bergmann, D., Stryker, C. (2026c). *Was sind Diffusionsmodelle?*. IBM, <https://www.ibm.com/de-de/think/topics/diffusion-models#186915248>

Bine. (o. D.). *KI und Musik – eine Erfolgsgeschichte*. couchFM.

<https://www.couchfm.medienwissenschaft-berlin.de/ki-und-musik-eine-erfolgsgeschichte/>

Bohm, Lars. (2018). *Moderne Musikproduktion in der Praxis*. (2. Auflage). PPVMEDIEN GmbH.

Brown, O., Miranda, E. R. (Hrsg.). (2021). *Handbook of Artificial Intelligence for Music: Foundations, advanced approaches, and developments for creativity*. Springer. <https://doi.org/10.1007/978-3-030-72116-9>

Brüggemann, A. (1983). *Keine Panik, es ist nur KI!. Grundprinzipien zur professionellen Nutzung von text-KI für Neulinge und Fortgeschrittene*. Gabal.

- Bryce, D. (2024). *Artificial Intelligence and Music: Analysis of Music Generation Techniques Via Deep Learning and the Implications of AI in the Music Industry*. Bryant Digital Repository. https://digitalcommons.bryant.edu/honors_data_science/12/
- Dampz, N. (2025, 18. Juli). *KI-generierte Musik: Millionen Fans feiern den Fake*. Ndr. <https://www.ndr.de/kultur/musik/ki-generierte-musik-millionen-fans-feiern-fake.velvetsundown-100.html>
- Developer Survey. (2025). *AI....3*. Stackoverflow. <https://survey.stackoverflow.co/2025/ai?utm>
- Dickreiter, M., Volker, D., Hoeg, W., Wöhr, M. (2014). *Handbuch der Tonstudioteknik Band 1*. (8. Auflage). DE Gruyter SAUR.
- Electric Bass. (o. D.). Musikhaus Thomann. https://www.thomann.de/de/steinberg_electric_bass.htm
- Eurostat. (2025, 16. Dezember). *32.7% of EU people used generative AI tools in 2025*. Eurostat. <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251216-3?utm>
- (Aufgerufen am 18.03.26)
- Fangfang, Lee. (2026a). *Was sind neuronale Netzwerke?*. IBM, <https://www.ibm.com/de-de/think/topics/neural-networks#741977106>
- Fišer, N., Martín-Pascual, M. Á. & Andreu-Sánchez, C. (2025). Emotional impact of AI-generated vs. human-composed music in audiovisual media: A biometric and self-report study. *PLoS ONE*, 20(6), e0326498. <https://doi.org/10.1371/journal.pone.0326498>
- Foster, T. (2024). *Künstlich Intelligenz in der Musik und Audioproduktion*. Kampenwand Verlag.
- (Aufgerufen am 24.02.26)
- Gema. (o.D.). *Die Gema klagt für eine faire Vergütung*. Gema.de. https://www.gema.de/de/aktuelles/ki-und-musik/ki-klage?utm_source=chatgpt.com
- Gema. (o.D.). *Hörbeispiele: So kopiert Suno AI bekannte Songs*. Gema.de. <https://www.gema.de/de/aktuelles/ki-und-musik/ki-klage/soundbeispiele-suno>
- Goldmedia. (2024, 30. Januar). *AI and music: Market development of AI in the music sector and impact on music authors and creators in Germany and France*. [Study commissioned by GEMA & SACEM]. Gema. <https://www.gema.de/documents/d/quest/gema-sacem-goldmedia-ai-and-music-pdf>
- Gottschling, S., Seidl, T., & Vonhof, C. (2024). Use of AI tools by students: An exemplary study of student usage scenarios. *Die hochschullehre*, 10(11), 122–135. https://www.researchgate.net/publication/381717974_Use_of_AI_tools_by_students_An_exemplary_study_of_student_usage_scenarios
- Haack, F. & Steinberg Media Technologies GmbH. (2022, 2. März). *Verve Felt Piano: Ein klangliches Meisterstück*. Steinberg Media Technologies GmbH. https://ocl-steinberg-live.steinberg.net/storage/asset/155097/storage/master/Pressemitteilung%20-%202022-03-02%20-%20Verve%20-%20DE.pdf?_gl=1*1ipluyx*_gcl_aw*R0NMLjE3Njk3MDY2MTEuQ2owS0NRaUFwLXpMQmhEa0FSSXNBQmNZYzZ1aC1MSnJSYzJNbnNIYXRmVDUxUTg2MXZxNU1NM004NS1fUTlq

[OXdDVTdnejlJQzJFUTAtY2FBcnpRUFMd193Y0l.* gcl au*MzE5MTIyODY0LjE3Njk2MDY0MDk.](https://doi.org/10.1007/s10273-017-2203-5)

Heinen, N., Heuer, A., & Schautschick, P. (2017). Künstliche Intelligenz und der Faktor Arbeit. *Wirtschaftsdienst*, 97(10), 714–720. <https://doi.org/10.1007/s10273-017-2203-5>

Hübert, H. (2021, 10 Oktober). Künstlich ist nicht künstlerisch. BR Klassik. <https://www.br-klassik.de/aktuell/news-kritik/kritik-urauffuehrung-beethoven-10-symphonie-kuenstliche-intelligenz-computer-bonn-100.html>

(Aufgerufen am 19.03.26)

IBM. (o. D.). *Was ist ein Gradientenabstieg?*. <https://www.ibm.com/de-de/think/topics/gradient-descent#487320256>

Jacob, M. (2026, 27. Januar). Streit um Stimmrechte: Warum neue Netflix-Produktionen bald ohne deutsche Tonspur erscheinen könnten. *Berliner Zeitung*. <https://www.berliner-zeitung.de/news/streit-um-stimmrechte-warum-neue-netflix-produktionen-bald-ohne-deutsche-tonspur-erscheinen-koennten-li.10016230>

Jäkel, F. (2025). *Intelligente Täuschung*. Transcript Verlag. <https://doi.org/10.14361/9783839405598>

Kersting, K., Lampert, C., & Rothkopf, C. (Hrsg.). (2019). *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*. Springer. <https://doi.org/10.1007/978-3-658-26763-6>

Lablack, V. (2025, 14. Juli). *KI in den Charts: Diese fünf KI-Songs schreiben Musikgeschichte*. delamar. <https://www.delamar.de/musikbusiness/ki-songs-in-den-charts-80699/>

LANDR Audio. (2025, 18 November). *AI tools & features – how creatives use AI for music production..* <https://www.landr.com/ai>

Landschoot, C. (2024, 22. Januar). *Audio Diffusion: Generative Music's Secret Sauce*. Towards Data Science. <https://towardsdatascience.com/audio-diffusion-generative-musics-secret-sauce-f625d0aca800/>

Lang, N. (2021, 2 November). Gradientenverfahren – einfach erklärt mit Beispiel. <https://databasecamp.de/ki/gradientenverfahren-ein-steiler-abstieg>

La La, Frank. (2019, April). *Künstlich intelligent: Wie lernen neuronale Netze?* Microsoft Build. Band 34(4). <https://learn.microsoft.com/de-de/archive/msdn-magazine/2019/april/artificially-intelligent-how-do-neural-networks-learn?>

Mockenhaupt, A. & Schlagenhaut, T. (2024). *Digitalisierung und Künstliche Intelligenz in der Produktion* (2. Auflage). <https://doi.org/10.1007/978-3-658-41935-6>

(Aufgerufen am 20.02.26)

Native Instruments GmbH. (2026). *Guitar Rig 7 Player* [Produktseite]. Native-Instruments.com. https://www.native-instruments.com/de/products/komplete/guitar/guitar-rig-7-player/?srsltid=AfmBOooTd29tff_oCKMO2JAHfvav77J2zWZxNOd6gx60vCmQkxb8SJhX

- Oelze, S. (2021, 9. Oktober). Beethovens „unvollendete“ von KI vollendet. Deutsche Welle. <https://www.dw.com/de/beethoven-10-sinfonie-unvollendete-ki-k%C3%BCnstliche-intelligenz/a-59378632>
- Rich, E. (1983). *Artificial intelligence*. McGraw-Hill. [https://doi.org/10.1016/0004-3702\(86\)90034-2](https://doi.org/10.1016/0004-3702(86)90034-2)
- Rieck, L. (2026, 26. Januar). *Suno AI Nutzungsbedingungen & Urheberrecht: Wem gehören KI-Songs?* Rieck & Partner Rechtsanwälte. <https://rieck-partner.de/musik/suno-ai-nutzungsbedingungen-urheberrecht-wem-gehoren-ki-songs.html>
- Rubin, R. (2023). *Kreativ. Die Kunst zu sein* (deutsche Erstauflage). O. W. Barth Verlag.
- SAP. (2026, aktualisiert am 26 Januar). Was ist generative KI?. <https://www.sap.com/germany/resources/what-is-generative-ai>
- Shank, D. B., Stefanik, C., Stuhlsatz, C., Kacirek, K. & Belfi, A. M. (2023). AI composer bias: Listeners like music less when they think it was composed by an AI. *Journal Of Experimental Psychology Applied*, 29(3), 676–692. <https://doi.org/10.1037/xap0000447>
- Singh, Vikash. (2025, 28 Mai). *Die Sigmoid Funktion: Eine Schlüsselkomponente in der Datenwissenschaft*. Datacamp. https://www.datacamp.com/de/tutorial/sigmoid-function?dc_referrer=https%3A%2F%2Fwww.google.com%2F
- Spotify. (2025, 25. September). *Spotify stärkt KI-Schutzmaßnahmen für Künstler*innen und Produzent*innen — Spotify*. Spotify for the Record. https://newsroom.spotify.com/2025-09-25/spotify-starkt-ki-schutzmasnahmen-fur-kunstlerinnen-und-produzentinnen/?utm_source=chatgpt.com
- Stack Overflow. (2023, 12. Juni). *Developer Survey: Sentiment around AI/ML (2023)*. Stackoverflow.blog.. <https://stackoverflow.blog/2023/06/12/developer-survey-sentiment-ai-ml/>
- (Aufgerufen am 11.03.26)
- Steinberg Media Technologies. (o.D. a). *Groove Agent – VST Drum Plug in kaufen*. <https://www.steinberg.net/de/vst-instruments/groove-agent/>
- (Aufgerufen am 19.03.26)
- Steinberg Media Technologies. (o.D. b). *RoomWorks*. https://www.steinberg.help/r/cubase-artist/cubaseplugref/15.0/en/shared/topics/plug_ref/roomworks_r.html
- (Aufgerufen am 05.04.26)
- Stryker, Cole., Scapiccio, M. (o.D.). *Was ist generative KI?*. <https://www.ibm.com/de-de/think/topics/generative-ai#257779831>
- (Aufgerufen am 14.02.26)
- Studio Strings. (o.D). Steinberg. https://www.steinberg.help/r/halion/7.1/de/halion/topics/included_instruments/studio_strings_c.html

(Aufgerufen am 14.04.26)

Suno – Blog. (2026, 26. März). Terms of Service. <https://suno.com/terms-of-service>

(Aufgerufen am 20.03.26)

Suno – CA AB 2013 Disclosure (o. D.). How can we help?.
<https://help.suno.com/en/articles/9709569>

(Aufgerufen am 06.03.26)

Suno – Creative Prompt Boosting in v4.5. (o.D.). *How can we help?*.
<https://help.suno.com/en/articles/5804417>

(Aufgerufen am 11.03.26)

Suno – How do I make my song longer?. (o.D.). *How can we help?*.
<https://help.suno.com/en/articles/6141441#>

(Aufgerufen am 26.03.26)

Suno – How to Use: Creative Sliders. (o.D.). *How can we help?*.
<https://help.suno.com/en/articles/6141377>

(Aufgerufen am 25.02.26)

Suno – How to Use: Stem Cover in Studio. (o.D.). *How can we help?*.
<https://help.suno.com/en/articles/9819905>

(Aufgerufen am 13.03.26)

Suno – How to use: Stem Extraction. (o.D.). *How can we help?*.
<https://help.suno.com/en/articles/6141441#>

(Aufgerufen am 24.02.26)

Suno - Introducing Suno Studio 1.2. (o.D.). *How can we help?*.
<https://help.suno.com/en/articles/10625089>

(Aufgerufen am 24.02.26)

Suno - Pricing. (o. D.). *Start making music for free.* <https://suno.com/pricing>

(Aufgerufen am 11.03.26)

Suno – Suno Sounds: Generate Custom Audio Samples. (o.D.). *How can we help?*.
<https://help.suno.com/en/articles/10625537>

Tagesschau. (2025, 21. Januar). *GEMA verklagt KI-Unternehmen Suno.* tagesschau.de.
<https://www.tagesschau.de/wirtschaft/gema-suno-chatgpt-musik-100.html>

Thomas Foster. (2025, 24 September). *Suno V5 – Der Tag, an dem sich Musik für immer verändert hat!*. Thomas Foster - Sound & Vision.
<https://www.youtube.com/watch?v=1Wx7bq2i4LU&t=417s>

Vina, A. (2024, 26 August). *Was sind Diffusionsmodelle? Eine kurze umfassende Anleitung*. Ultralytics. <https://www.ultralytics.com/de/blog/what-are-diffusion-models-a-quick-and-comprehensive-guide>

Voltage (o. D.). Steinberg. https://www.steinberg.help/r/halion/7.1/de/halion/topics/included_instruments/voltage_c.html

(Aufgerufen am 25.02.26)

Wikipedia. (2024, 27. Mai). *Suno AI*. https://de.wikipedia.org/wiki/Suno_AI

Wuttke, L. (2024, 01 Februar). *Künstliche neuronale Netzwerke: Definition, Einführung, Arten und Funktion*. Datasolut. <https://datasolut.com/neuronale-netzwerke-einfuehrung/>

Zoronjic, S. (2025, 11 November). *Streit um Urheberrecht: GEMA setzt sich gegen OpenAI durch*. NDR Kultur. <https://www.ndr.de/kultur/musik/streit-um-urheberrecht-gema-setzt-sich-gegen-openai-durch,openai-112.html>

Anmerkung: Bei dieser Arbeit wurde ChatGPT zur Formulierungs-, Rechtschreib- und Grammatikkorrektur genutzt.

VI **Abbildungsverzeichnis**

- Abb 1 Aufbau eines neuronalen Netzes (S. 9)
- Abb 2 Visualisierung der Sigmoid Funktion (S. 10)
- Abb 3 Auswirkung unterschiedlicher Lernraten (S. 11)
- Abb 4 Vorwärts Diffusionsprozess (S. 13)
- Abb 5 Rückdiffusionsprozess (S. 14)
- Abb 6.1 Notenauszug aus dem Song Daddy Cool (S. 17)
- Abb 6.2 Notenauszug aus dem Song Daddy Cool aus Suno (S. 17)
- Abb. 7 Voltage Synthesizer (S. 22)
- Abb. 8. Eingebener Prompt (S.23)
- Abb. 9. Eingebener Prompt (S.24)
- Abb. 10 Eingebener Prompt (S.24)
- Abb. 11 Audio Wrap (S. 25)
- Abb. 12 Eingebener Prompt (S. 26)
- Abb. 13. Eingebener Prompt (S. 27)
- Abb. 14 Eingebener Prompt (S. 29)
- Abb. 15 Eingebener Prompt (S. 30)
- Abb. 16 Eingebener Prompt (S. 32)
- Abb. 17 Saxophon Arrangement (S. 33)
- Abb. 18 Midi Noten aus Suno (S. 36)
- Abb. 19 Liedpassage in Cubase Pro 14 (S. 38)
- Abb. 20 Suno Interpretation der Liedpassage aus Cubase Pro 14 (S. 38)
- Abb. 21 Tempoeinstellung in Suno (S. 40)
- Abb. 22 Eingebener Prompt (S. 41)

- Abb. 23 Wellenform des Nachhalls (S. 44)
- Abb. 24 Energieverlauf eines Nachhalls (S. 44)
- Abb. 25 Darstellung der Drum Sequenz in Cubase Pro 14 (S. 47)
- Abb. 26 Kurvenverlauf des Weirness Durchlaufs (S. 65)
- Abb. 27 Kurvenverlauf des Style Influence Durchlaufs (S. 78)
- Abb. 28 Kurvenverlauf des Audio Influence Durchlaufs (S. 88)

VII Anhang

Tabelle 3: Die Tabelle fasst die Bewertungen aus Kapitel 4.3.1 zusammen.

Parametereinstellungen	Ergebnisse	Bewertungen
W 0 / S 50 / A 50	Ergebnis 1.1	Nähe zur Vorlage: 4/5 Artefaktfreiheit: 3.7/5 Kreativität: 3/5
W 0 / S 50 / A 50	Ergebnis 1.2	Nähe zur Vorlage: 3.9/5 Artefaktfreiheit: 3.7/5 Kreativität: 3/5
W 25 / S 50 / A 50	Ergebnis 2.1	Nähe zur Vorlage: 3.8/5 Artefaktfreiheit: 3.3/5 Kreativität: 3/5
W 25 / S 50 / A 50	Ergebnis 2.2	Nähe zur Vorlage: 3.3/5 Artefaktfreiheit: 3.5/5 Kreativität: 3.5/5
W 50 / S 50 / A 50	Ergebnis 3.1	Nähe zur Vorlage: 3.3/5 Artefaktfreiheit: 3/5 Kreativität: 3/5
W 50 / S 50 / A 50	Ergebnis 3.2	Nähe zur Vorlage: 3.4/5 Artefaktfreiheit: 2.9/5 Kreativität: 2.5/5
W 75 / S 50 / A 50	Ergebnis 4.1	Nähe zur Vorlage: 3.3/5 Artefaktfreiheit: 4/5 Kreativität: 3/5
W 75 / S 50 / A 50	Ergebnis 4.2	Nähe zur Vorlage: 2.2/5 Artefaktfreiheit: 2.2/5 Kreativität: 1/5
W 100 / S 50 / A 50	Ergebnis 5.1	Nähe zur Vorlage: 0.8/5 Artefaktfreiheit: 1/5 Kreativität: 1/5
W 100 / S 50 / A 50	Ergebnis 5.2	Nähe zur Vorlage: 0.8/5 Artefaktfreiheit: 1/5 Kreativität: 1/5

Tabelle 4: Die Tabelle fasst die Bewertungen aus Kapitel 4.3.2 zusammen.

Parametereinstellungen	Ergebnisse	Bewertungen
W 50 / S 0 / A 50	Ergebnis 1.1	Nähe zur Vorlage: 3.5/5 Stilistische Eindeutigkeit: 3.7/5 Nutzbarkeit: 4.3/5
W 50 / S 0 / A 50	Ergebnis 1.2	Nähe zur Vorlage: 3.75/5 Stilistische Eindeutigkeit: 4.3/5 Nutzbarkeit: 4/5
W 50 / S 25 / A 50	Ergebnis 2.1	Nähe zur Vorlage: 4/5 Stilistische Eindeutigkeit: 4.3/5 Nutzbarkeit: 4.8/5
W 50 / S 25 / A 50	Ergebnis 2.2	Nähe zur Vorlage: 2.9/5 Stilistische Eindeutigkeit: 4.5/5 Nutzbarkeit: 4.8/5
W 50 / S 50 / A 50	Ergebnis 3.1	Nähe zur Vorlage: 3.3/5 Stilistische Eindeutigkeit: 4.3/5 Nutzbarkeit: 4.8/5
W 50 / S 50 / A 50	Ergebnis 3.2	Nähe zur Vorlage: 3.1/5 Stilistische Eindeutigkeit: 3.6/5 Nutzbarkeit: 3.8/5
W 50 / S 75 / A 50	Ergebnis 4.1	Nähe zur Vorlage: 3.6/5 Stilistische Eindeutigkeit: 3.5/5 Nutzbarkeit: 3.5/5
W 50 / S 75 / A 50	Ergebnis 4.2	Nähe zur Vorlage: 3.4/5 Stilistische Eindeutigkeit: 4/5 Nutzbarkeit: 4.8/5
W 50 / S 100 / A 50	Ergebnis 5.1	Nähe zur Vorlage: 3.6/5 Stilistische Eindeutigkeit: 4/5 Nutzbarkeit: 4/5
W 50 / S 100 / A 50	Ergebnis 5.2	Nähe zur Vorlage: 3.4/5 Stilistische Eindeutigkeit: 4/5 Nutzbarkeit: 4.8/5

Tabelle 5: Die Tabelle fasst die Bewertungen aus Kapitel 4.3.3 zusammen.

Parametereinstellungen	Ergebnisse	Bewertungen
W 50 / S 50 / A 0	Ergebnis 1.1	Nähe zur Vorlage: 3.5/5 Nutzbarkeit: 4.8/5
W 50 / S 50 / A 0	Ergebnis 1.2	Nähe zur Vorlage: 2.9/5 Nutzbarkeit: 4/5
W 50 / S 50 / A 25	Ergebnis 2.1	Nähe zur Vorlage: 3/5 Nutzbarkeit: 4/5
W 50 / S 50 / A 25	Ergebnis 2.2	Nähe zur Vorlage: 3/5 Nutzbarkeit: 4/5
W 50 / S 50 / A 50	Ergebnis 3.1	Nähe zur Vorlage: 3.3/5 Nutzbarkeit: 4/5
W 50 / S 50 / A 50	Ergebnis 3.2	Nähe zur Vorlage: 3.9/5 Nutzbarkeit: 3/5
W 50 / S 50 / A 75	Ergebnis 4.1	Nähe zur Vorlage: 4/5 Nutzbarkeit: 3.5/5
W 50 / S 50 / A 75	Ergebnis 4.2	Nähe zur Vorlage: 3.3/5 Nutzbarkeit: 2.3/5
W 50 / S 50 / A 100	Ergebnis 5.1	Nähe zur Vorlage: 3/5 Nutzbarkeit: 2/5
W 50 / S 50 / A 100	Ergebnis 5.2	Nähe zur Vorlage: 2.5/5 Nutzbarkeit: 2/5