



UNIVERSITÄT  
DES  
SAARLANDES



**Fraunhofer**  
IZFP



Universität des Saarlandes  
Naturwissenschaftlich-Technische Fakultät  
Fachbereich: Materialwissenschaft und Werkstofftechnik

Masterarbeit zur Erlangung des akademischen Grades  
Master of Science

---

Methodik zur praxisnahen Klassifikation von  
akustischen Signalen auf Basis von überwachten  
maschinellen Lernverfahren am Beispiel eines  
Schweinemastbetriebs

---

Katharina Elisabeth Bollmann  
Matrikelnummer: 2565349

November 2023

Erstkorrektur: Prof. Dr.-Ing. Ute Rabe  
Zweitkorrektur: Prof. Dr. Andreas Schütze



## DANKSAGUNG

---

Mit der Abgabe dieser Arbeit endet meine Zeit als Studierende des Fachs Materialwissenschaft und Werkstofftechnik an der Universität des Saarlandes. Während dieser Zeit habe ich viel lernen dürfen: nicht nur als Studierende in den Vorlesungen, sondern auch als Hiwi an der Universität und bei Fraunhofer und in der Gremienarbeit für den Fachbereich MatWerk in der Deutschen Gesellschaft für Materialkunde. All den Menschen, die mich auf diesem Weg geprägt haben, bin ich sehr dankbar. Drei von ihnen möchte ich hier besonders hervorheben.

Danke Mama – ohne dich hätte ich die Materialwissenschaft vielleicht nicht als Studienfach gewählt. Schade, dass du nicht mehr bei uns bist.

Danke Michel – dass du so lange schon bei mir und immer für mich da bist.

Danke Michael – fürs Weiterbringen in der jDGM, als Hiwi-Betreuer und als Freund.

Diese Masterarbeit hat die letzte Phase meines Studiums eingenommen und hier möchte ich weiteren Personen danken. Danke an Sarah und das gesamte SmartPigHome-Team für die Möglichkeit, im Rahmen dieses Projekts meine Arbeit zu verfassen und mich neuen Herausforderungen zu stellen. Danke an Florian und Thies für die Einblicke in die Welt der Tierärzte und Schweine. Des Weiteren gilt mein Dank Frau Rabe und Herrn Schütze, dieses interdisziplinäre Thema wissenschaftlich zu betreuen.

Ein besonderer Dank geht an meine *Familie* sowie dem Team *Essen und Messen* für alle großen und kleinen Denkanstöße und ihre Fähigkeit, mich wieder einzufangen und zu stabilisieren.



## EIGENSTÄNDIGKEITSERKLÄRUNG

---

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbstständig und nur durch unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Stellen sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

---

Katharina Bollmann



## KURZZUSAMMENFASSUNG

---

In der Produktion steigt der Bedarf an Assistenzsystemen aufgrund der sinkenden Anzahl an Facharbeitern, der hohen technischen Anforderungen an die Produkte und des gleichzeitig hohen wirtschaftlichen Drucks. In der Fleischproduktion stehen Landwirte ebenso vor der Herausforderung, wirtschaftlich rentabel zu produzieren und gleichzeitig für das Wohl ihrer Tiere zu sorgen. Im BMEL geförderten Projekt *SmartPigHome* soll ein Assistenzsystem entwickelt werden, das sowohl mittels optischen als auch akustischen Aufnahmen in der Lage ist, kritische, tierwohlgefährdende Situationen zu erkennen.

Zur Unterscheidung von akustischen Signalen eines Schweinestalls wurde in dieser Masterarbeit eine Methodik entwickelt, die bedenkliche Schweinelaute von allen anderen Geräuschen im Stall trennt. Der Zufallswald-Klassifikator sowie der naive Bayes-Klassifikator in Kombination mit der linearen Diskriminanzanalyse erzielten von den getesteten überwachten Lernverfahren die besten Werte für Präzision sowie Sensitivität der Vorhersage. Diese beiden Methodiken wurden mittels statistischer Absicherung optimiert, sodass für Präzision und Sensitivität Werte von über 90 % erreicht wurden. Eine Überführung der entwickelten Methodik in ein Assistenzsystem wird im Rahmen des Projekts fortgeführt.



# INHALTSVERZEICHNIS

---

Danksagung.....	i
Eigenständigkeitserklärung.....	iii
Kurzzusammenfassung.....	v
Inhaltsverzeichnis.....	vii
1 Motivation und Problemstellung.....	1
2 Theoretische Grundlagen.....	3
2.1 Signalverarbeitung.....	3
2.2 Akustik.....	5
2.3 Statistik.....	7
2.4 Maschinelles Lernen.....	7
3 Stand der Technik.....	13
3.1 Passive akustische Überwachung.....	13
3.1.1 Passive akustische Überwachung in der Industrie.....	13
3.1.2 Passive akustische Überwachung in Natur und Umwelt.....	13
3.2 Klassifikation von Schweinelauten.....	14
3.2.1 Übersicht Literatur.....	14
3.2.2 Paper von Briefer et al. (38).....	15
3.2.3 Zusammenfassung Literatur zur Klassifikation von Schweinelauten.....	17
4 Materialien und Methoden.....	19
4.1 Datenaufnahme.....	19
4.1.1 Hardware.....	19
4.1.2 Software.....	20
4.1.3 Vorgehensweise.....	20
4.2 Datenanalyse.....	20
4.2.1 Software.....	20
4.2.2 Vorgehensweise.....	20
5 Ergebnisse und Diskussion.....	23
5.1 Ausarbeitung der Methodik.....	23
5.1.1 Auswahl der Klassen.....	23
5.1.2 Auswahl der Fensterbreite.....	24
5.1.3 Vorverarbeitung des akustischen Signals.....	25
5.1.4 Merkmale.....	25
5.1.5 Vorverarbeitung der Merkmale.....	27
5.1.6 Auswahl eines Modells.....	28

## Kapitel 0: Inhaltsverzeichnis

5.1.7	Bewertung der Vorhersage .....	28
5.2	Analyse der Methodik anhand der Trainingsdaten .....	29
5.2.1	Festlegung optimale Fensterbreite .....	29
5.2.2	Analyse des Geräuschanteils eines Fensters .....	29
5.2.3	Analyse der Vorhersagewahrscheinlichkeit .....	30
5.2.4	Korrelation der Merkmale .....	32
5.2.5	Reduktion der Merkmale .....	33
5.3	Übertragung der Methodik auf kontinuierliche akustische Signale .....	41
5.4	Diskussion .....	43
6	Zusammenfassung und Ausblick .....	47
7	Literatur .....	48
8	Abbildungsverzeichnis .....	51
9	Tabellenverzeichnis .....	52
10	Anhang .....	53
10.1	Theorie und Merkmalsberechnung .....	53
10.2	Skizze Maststall .....	55
10.3	Datenblätter .....	57

# 1 MOTIVATION UND PROBLEMSTELLUNG

---

In der heutigen, hochtechnologisierten Zeit sind bereits viele Prozesse in der Produktion automatisiert. Der fortschreitende Einsatz von automatisierten Abläufen geht in vielen Fällen mit einer Effizienzerhöhung sowie einer Verringerung der Anzahl der benötigten Mitarbeitenden einher. Diese Umstrukturierung der Abläufe wirkt dem ansteigenden Fachkräftemangel entgegen, doch werden weitere Verbesserungen benötigt, die Expertise bei sinkender Facharbeiter-Anzahl aufrecht zu erhalten, um Einbußen in Qualität sowie Wirtschaftlichkeit zu vermeiden. Das maschinelle Lernen liefert die Grundlage, die Erfahrung eines Menschen auf ein computergesteuertes System zu übertragen. Optische Methoden sind sowohl in der Sachgüterproduktion als auch in der Qualitätssicherung der Lebensmittelindustrie im Einsatz. (1, 2) Auch akustische Methoden beispielsweise in Form passiver akustischer Überwachung einzelner Schneidwerkzeuge werden im Zerspanungsprozess eingesetzt. (3) Je komplexer die akustisch zu überwachenden Bereiche werden, desto höher sind die Anforderungen an die Überwachungsalgorithmik, da die Geräusche vieler Quellen sich überlagern und teilweise an Oberflächen reflektiert werden. Der Mensch besitzt die Fähigkeit irrelevante Hintergrundgeräusche auszublenden und durch Erfahrung charakteristische Laute aus einer Vielzahl von Geräuschen zu erkennen, die ein Assistenzsystem erst „erlernen“ muss.

Neben der Sachgüterproduktion steht auch die Fleischproduktion unter wirtschaftlichem Druck. Die Produktionsmasse, sprich die Anzahl der Tiere, muss erhöht werden, was zur Folge hat, dass der Landwirt weniger Zeit für das einzelne Tier aufbringen kann. Auftretender Stress erhöht das Verletzungsrisiko der Tiere und somit die Wahrscheinlichkeit wirtschaftlicher Verluste, analog zu einem Maschinenstillstand in einer Produktionslinie. Daher bietet auch im Bereich der Landwirtschaft die passive akustische Überwachung kombiniert mit maschinellem Lernen sowohl wirtschaftlich als auch ethisch Potential.

Im Rahmen dieser Arbeit wird eine Methodik entwickelt, die reale, rauschbehaftete akustische Signale eines Schweinestalls mit vielen Geräuschquellen voneinander trennen und negative Stresslaute identifizieren kann. Die Methodik soll akustische Signale aufnehmen, mittels maschinellen Lernens negative Schweinelaute, die beispielsweise bei Stress oder Kämpfen mit Verletzungsrisiko auftreten können, erkennen und den Landwirt darüber informieren, und somit als ein Assistenzsystem agieren. Dieses wird im Rahmen des vom BMEL geförderten Forschungsprojekts *SmartPigHome* im Stall zum Einsatz kommen. Im Unterschied zu anderen wissenschaftlichen Untersuchungen zur Unterscheidung von Schweinelaute mittels maschinellen Lernens werden in der Erstellung sowie Ausarbeitung der Methodik bereits reale akustische Daten aus einem Stall verwendet, um die Praxisnähe zu gewährleisten.



## 2 THEORETISCHE GRUNDLAGEN

Im ersten Teil der für diese Arbeit relevanten Grundlagen wird ein Einblick in die Signalverarbeitung von Schwingungssignalen gegeben. Die Grundlagen im Bereich Akustik sowie Schallausbreitung sind für das Verständnis der Problemstellung wichtig. Im Abschnitt zur Statistik finden sich die Grundlagen zu einigen der verwendeten Merkmale der akustischen Signale und im Abschnitt zum maschinellen Lernen wird ein Überblick über die Grundbegriffe der verwendeten Methoden gegeben.

### 2.1 SIGNALVERARBEITUNG

#### *Kontinuierliche und diskrete Zeitsignale*

„Ein Signal ist eine zeitlich und örtlich veränderliche Größe“. (4) Unterschieden wird dabei neben orts- und zeitabhängigen Signalen auch zwischen kontinuierlichen und diskreten Signalen. Ist der Definitionsbereich kontinuierlich, also kann ein Signalverlauf mit einer stetigen Funktion beschrieben werden, so wird das Signal als kontinuierlich bezeichnet. Als Beispiel für eine kontinuierliche Funktion ist in Gleichung (1) eine Sinusfunktion gezeigt, die zu den harmonischen Signalen zählt. Diese Signale können über die Frequenz  $f = \omega/(2\pi)$  mit Kreisfrequenz  $\omega$  sowie die Amplitude  $A$  beschrieben werden.

$$f(t) = A \sin(\omega t) \quad (1)$$

In Abbildung 1 ist ein solches Sinussignal mit einer Amplitude  $A = 1$  sowie einer Kreisfrequenz von  $\omega = 0,5$  Hz gezeigt. Eingezeichnet ist neben der Amplitude die Schwingungs- oder Periodendauer  $T$ , die über  $T = 1/f = 2\pi/\omega$  von der Frequenz abhängt.

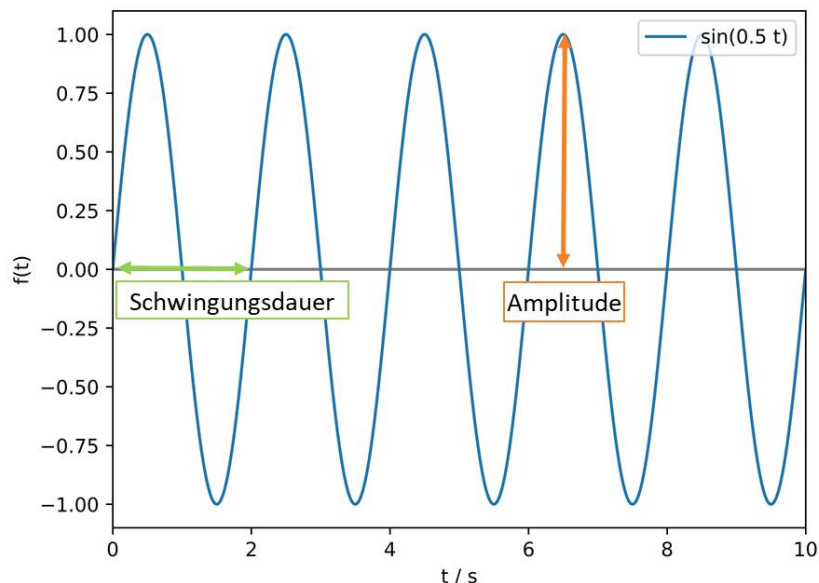


Abbildung 1: Beispiel für ein kontinuierliches Zeitsignal.

Alle mittels Sensoren aufgenommenen Signale liegen zeitdiskret vor, da der Sensor mit einer festgelegten Abtastrate die physikalische Messgröße in ein elektrisches Signal umwandelt. In Abbildung 2 ist das Sinussignal aus Abbildung 1 mit einer Abtastrate  $f_A$  von 6,9 Hz dargestellt.

## Kapitel 2: Theoretische Grundlagen

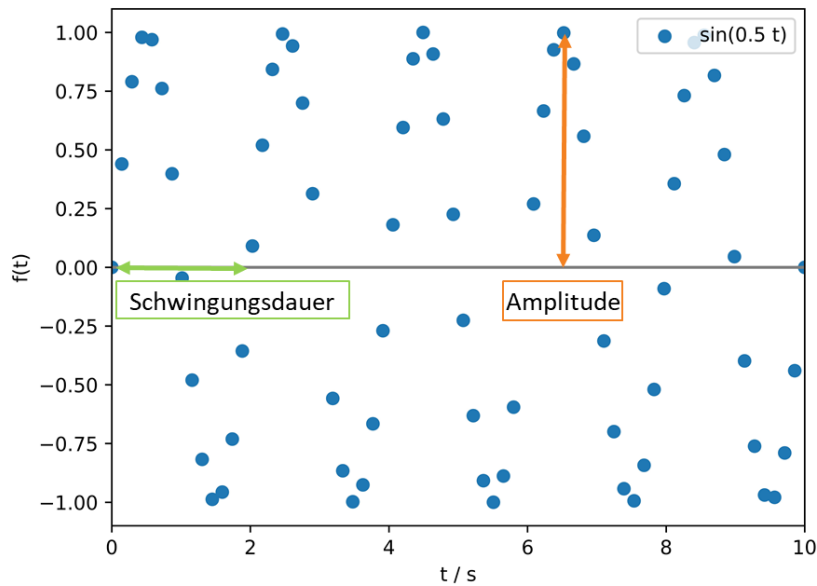


Abbildung 2: Beispiel für ein diskretes Zeitsignal.

Die allgemeine Form eines kontinuierlichen zeitabhängigen Signals ist in Gleichung (2) zu sehen.

$$f(t) = Ae^{i\omega t} \quad (2)$$

Die des diskreten zeitabhängigen Signals ist in Gleichung (3) dargestellt, die als Fourier-Reihe bezeichnet wird. Die komplexen Koeffizienten  $c_k$  werden als Fourier-Koeffizienten und  $\omega_0$  wird als Grundkreisfrequenz bezeichnet. (4, 5)

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{jk\omega_0 t} \quad (3)$$

### Fourier-Transformation

Ein Zeitsignal  $f(t)$  kann mittels der Fourier-Transformation in den Frequenzbereich überführt werden. Dabei wird analog zum Zeitbereich zwischen kontinuierlichen und diskreten Signalen unterschieden. In Gleichung (4) ist das Signal des Spektralbereichs  $F(\omega)$  für ein kontinuierliches und in Gleichung (5) für ein diskretes Zeitsignal gezeigt, wobei  $T_A$  als Abtastperiode bezeichnet wird. Diese hängt reziprok über  $T_A = 1/f_A$  von der Abtastfrequenz  $f_A$  ab.

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (4)$$

$$F(\omega) = \sum_{k=-\infty}^{\infty} c_k e^{-j\omega k T_A} \quad (5)$$

Nach dem Abtasttheorem muss die Abtastfrequenz doppelt so groß wie die höchste abzubildende Frequenz sein. Diese Frequenz wird auch als Nyquist-Frequenz bezeichnet. Die Frequenzauflösung ist reziprok von der Beobachtungsdauer abhängig und kann als Quotient der Abtastfrequenz durch die Anzahl der Samples berechnet werden. (4, 5)

## 2.2 AKUSTIK

Als Akustik wird die Lehre des Schalls bezeichnet. Schallwellen können sowohl transversal, d.h. senkrecht zur Ausbreitungsrichtung, oder longitudinal, d.h. in Ausbreitungsrichtung, schwingen. In dieser Arbeit sind vornehmlich die Longitudinalwellen von Interesse, da in Luft aufgrund der nur schwachen Wechselwirkung der Moleküle keine Scherkräfte und somit Transversalwellen übertragen werden können. Die Ausbreitung des Schalls lässt sich für den eindimensionalen Fall mit der Wellengleichung (Gleichung (6)) beschreiben, wobei  $c$  als Schallgeschwindigkeit und  $p$  als Schalldruck bezeichnet wird, dessen räumliche ( $x$ ) sowie zeitliche ( $t$ ) Änderung betrachtet wird.

$$\frac{\partial^2 p}{\partial t^2} = c^2 \frac{\partial^2 p}{\partial x^2} \quad (6)$$

In Abbildung 3 ist am Beispiel eines akustischen Strahlers dargestellt, wie der Schalldruck  $p$  reziprok zum Abstand  $r$  von der Schallquelle abnimmt.

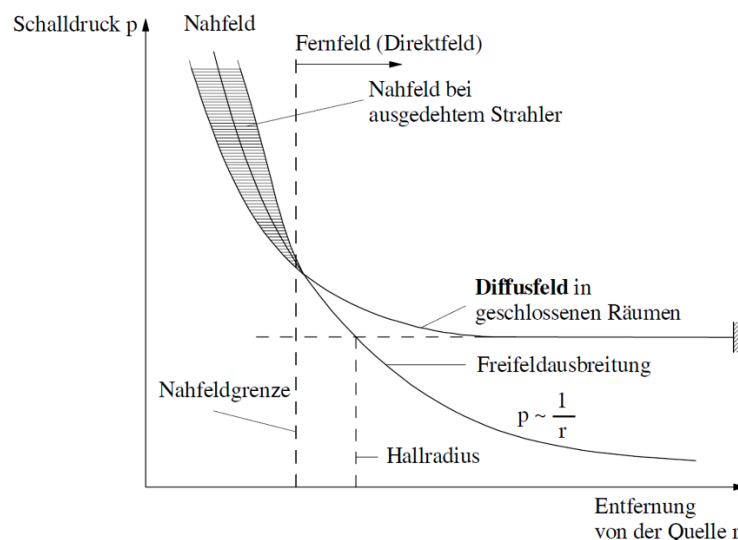


Abbildung 3: Schalldruckverhältnisse (schematisiert) in der Umgebung eines akustischen Strahlers; in Abwandlung von (6), Seite 495.

Dieser Zusammenhang lässt sich ebenfalls im Zusammenhang der Energiedichte  $E$  aufzeigen. Diese nimmt im Freien in Abhängigkeit der Schallleistung  $P$  sowie der Schallgeschwindigkeit  $c$  mit der in Gleichung (7) gezeigten Zusammenhang quadratisch mit dem Abstand  $r$  von der Schallquelle ab.

$$E = \frac{P}{4\pi cr^2} \quad (7)$$

In einem geschlossenen Raum stellt sich ein Gleichgewicht zwischen der abgestrahlten und der von den Wänden absorbierten Leistung ein, das sogenannte Diffusfeld. Die Energiedichte nimmt im Abstand des Hallradius'  $r_H$  den Wert  $E_0$  an. Der Hallradius ist wie in Gleichung (8) gezeigt von der absorbierenden Fläche  $A$  sowie deren mittleren Absorptionsgrad  $\bar{\alpha}$  abhängig. Diese Konvergenz gilt wie in Abbildung 3 zu sehen ebenso für den Schalldruck.

$$r_H = \sqrt{\frac{\bar{\alpha} A}{16\pi}} \quad (8)$$

## Kapitel 2: Theoretische Grundlagen

In sogenannten Direktfeld, das auch als Fernfeld bekannt ist, tritt freie Ausbreitung ohne Interferenzerscheinungen auf. Bei den akustischen Daten dieser Arbeit ist davon auszugehen, dass Reflexionen an den Stall- sowie Buchtenwänden auftreten und somit ein Diffusfeld vorliegt. Eine Skizze des Stalls kann Anhang 10.2 entnommen werden.

Die Schallwellen werden an den vorhandenen glatten Flächen nach dem Reflexionsgesetz reflektiert. Geht von der Quelle  $Q$  Schall aus und wird an einer Wand reflektiert, so scheint der Schall von der sogenannten Spiegelschallquelle  $Q'$  auszugehen. Bei Mehrfachreflexionen wird zwischen Spiegelquellen erster ( $Q'$ ) und zweiter ( $Q''$ ) Ordnung unterschieden. Diese sind für eine Raumecke in Abbildung 4 dargestellt.

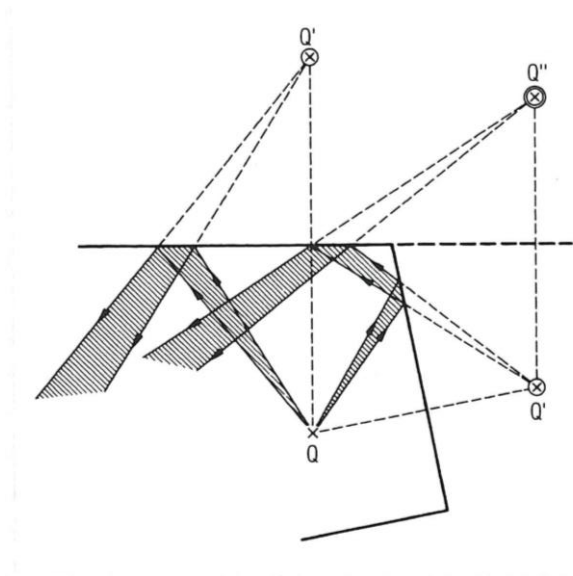


Abbildung 4: Schematische Darstellung einer Schallquelle  $Q$  in einer Raumecke und ihren Spiegelschallquellen  $Q'$  und  $Q''$ ; entnommen aus (7), Seite 334.

Neben der Reflexion an ebenen Flächen kann auch Reflexion an diffusen Flächen auftreten, die sich im Gegensatz zu den Spiegelquellen nicht mehr mit einer geschlossenen Lösung beschreiben lässt. Darüber hinaus interferieren die Schallwellen mehrerer Quellen miteinander, was zu einer Verstärkung oder einer Abschwächung des einzelnen Schallsignals führen kann. Bei Schallausbreitung über größere Entfernungen sind zudem Dämpfungseffekte wie Schallabsorption aufgrund von innerer Reibung, Wärmeleitung sowie molekularer Absorption zu berücksichtigen. Alle diese genannten Effekte haben zur Folge, dass ein einzelnes Schallsignal aufgrund der Interferenzen sowie Dämpfungen und dem damit einhergehenden Verlust an Informationen bzgl. der Ausgangsamplitude sowie  $-$ phase nicht durch Entfaltungen aus einer Überlagerung von Schallsignalen herausgestellt werden kann. (6, 7)

## 2.3 STATISTIK

### Zufallsverteilung

Die aufgenommenen akustischen Signale können als zeitabhängige Zufallssignale betrachtet werden. Die Zufallsgrößen unterliegen einer Verteilungsfunktion, der Wahrscheinlichkeitsfunktion  $p(x)$ , die mittels der zentralen Momente  $z_k$  beschrieben werden kann. Für eine Verteilungsfunktion diskreter Werte  $p(x_i)$  werden diese mit Gleichung 9 berechnet, wobei  $m_1$  der arithmetische Mittelwert, auch als Erwartungswert und gewöhnliches Mittel 1. Ordnung bezeichnet, ist.

$$z_k = \sum_{i=0}^{\infty} (x_i - m_1)^k p(x_i) \quad (9)$$

Die zentralen Momente höherer Ordnung werden als Varianz oder Dispersion ( $z_2$ ), Schiefe oder Asymmetrie ( $z_3$ ) und Wölbung oder Exzess ( $z_4$ ) bezeichnet. (4)

### Korrelierte Zufallsvariablen

Liegen einer Zufallsverteilung  $F(X,Y)$  mehrere Zufallsvariablen  $(X,Y)$  zugrunde, kann deren Abhängigkeit unter anderem mittels der Kovarianz  $\text{cov}(X,Y)$  (siehe Gleichung 10) sowie der Korrelation  $\text{cor}(X,Y)$  (siehe Gleichung 11) ausgedrückt werden, wobei  $E(X,Y)$  der Erwartungswert der jeweiligen Zufallsvariablen ist. (4)

$$\text{cov}(X,Y) = E(XY) - E(X)E(Y) \quad (10)$$

$$\text{cor}(X,Y) = \frac{\text{cov}(X,Y)}{\sqrt{z_2(X)z_2(Y)}} \quad (11)$$

Der Pearsonschen Korrelationskoeffizient  $r(X,Y)$  beschreibt als Normierung der Kovarianz, wie in Gleichung 12 zu sehen, die „Güte der linearen Vorhersagbarkeit von Y durch X.“ Je höher der Betrag des Koeffizienten im Wertebereich -1 bis +1 wird, desto stärker sind X und Y linear korreliert. (8)

$$r(X,Y) = \frac{\text{cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\text{cov}(X,Y)}{\sqrt{z_2(X)z_2(Y)}} \quad (12)$$

In der Signalverarbeitung ist eine Verringerung der Korrelation wünschenswert, die durch die Drehung der Koordinaten-Achsen in Richtung der Hauptachsen erreicht werden kann, in deren Richtung die Signale die höchste Varianz aufweisen. Diese wird als Hauptachsentransformation bezeichnet. (4)

## 2.4 MASCHINELLES LERNEN

Das maschinelle Lernen (engl. machine learning) wird in der Regel in die Bereiche des überwachten Lernens (engl. supervised learning), des unüberwachten Lernens (engl. unsupervised learning) sowie des bestärkenden Lernens (engl. reinforcement learning) unterteilt. Beim überwachten Lernen werden Eingangsgrößen, auch Merkmale (engl. feature) genannt, auf Zielgrößen (engl. target) abgebildet und das Modell erlernt diese Abbildungsfunktion. Liegen die Zielgrößen als Kategorien vor, so ist von Klassifikation bzw. Mustererkennung die Rede, liegen sie als skalare Werte vor, wird von Regression gesprochen. Beim unüberwachten Lernen werden in den Eingangsgrößen Muster erkannt, ohne dass Zielgrößen oder Arten von Mustern vorgegeben werden. Das bestärkende Lernen wird seltener als die anderen beiden verwendet und beruht auf einem Rückmeldungs-Prinzip durch seine Umwelt, bei der das Modell belohnende oder bestrafende Signale erhält. In dieser Arbeit wird sich auf die Implementierung des überwachten Lernens beschränkt.

## Kapitel 2: Theoretische Grundlagen

Der allgemeine Ablauf des überwachten Lernens sieht im ersten Schritt das Training eines Modells auf einem Teil der Gesamtdaten vor, das in einem zweiten Schritt für unbekannte Eingangsgrößen, die ebenfalls Teil der Gesamtdaten sind, die Zielgrößen vorhersagt. Mittels der wahren Zielgrößen kann die Qualität der Vorhersage beurteilt werden. (9, 10)

### Aufteilung der Gesamtdaten

Die Paare, bestehend aus Eingangs- und Zielgrößen, werden vor dem Erstellen des Modells in der Regel in ein Trainingsset, ein Testset sowie ein Validierungsset aufgeteilt (engl. split). Üblicherweise werden 80 % der Daten zu Trainings- und 20 % zu Testdaten, sofern keine Validierungsdaten zum Optimieren des Modells benötigt werden. Mittels der Kreuzvalidierung (engl. cross validation) kann die Qualität der erstellten Modelle beurteilt werden. Bei der Kreuzvalidierung werden alle Daten in  $k$  Abschnitte, auch als Faltungen bezeichnet, unterteilt und  $k$  Modelle trainiert, wobei für jedes Modell ein anderer der  $k$  Datensätze als Validierungs- und die restlichen als Trainingsdaten verwendet werden. Im Spezialfall der stratifizierten Kreuzvalidierung wird für jede der  $k$  Teilmengen eine annähernd gleiche Verteilung der Klassen sichergestellt. Für den Fall  $k = 4$  ist das Vorgehen der stratifizierten Kreuzvalidierung (engl. stratified  $k$ -fold cross validation) für zwei Klassen in Abbildung 5 dargestellt. (9)

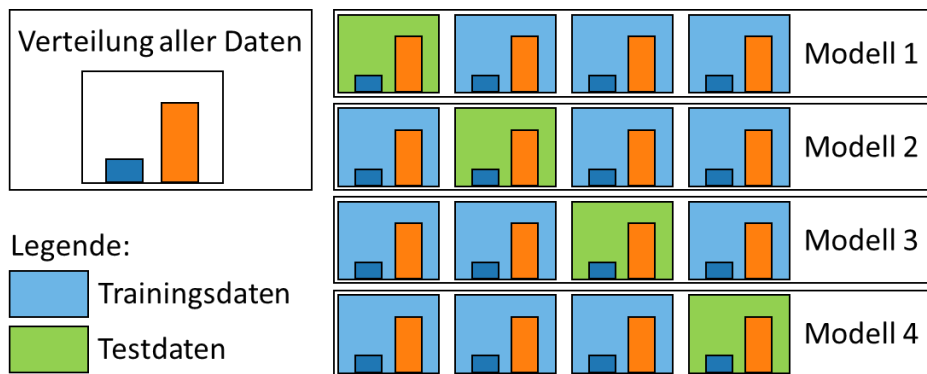


Abbildung 5: Schematische Darstellung der Auswahl an Test- und Trainingsdaten für eine stratifizierte Kreuzvalidierung für  $k=4$  Faltungen bzw. Modelle.

### Skalierung der Daten

Damit alle Daten, die unter Umständen in sehr unterschiedlichen Wertebereichen liegen, gleichwertig in das Training des Modells eingehen, werden diese in der Regel skaliert, sodass der Mittelwert null und die Standardabweichung bei eins liegt. (10)

### Dimensionsreduktion der Daten

Da nicht alle Modelle hochdimensionale Daten verarbeiten können, wird häufig vor dem Trainieren eines Modells eine Dimensionsreduktion durchgeführt. Zu diesen Algorithmen zählen die Hauptkomponenten-Analyse (engl. principal component analysis, PCA) sowie die lineare Diskriminanz-Analyse (engl. linear discriminant analysis, LDA).

Analog zur in Abschnitt 0 vorgestellten Hauptachsentransformation werden bei der Hauptkomponenten-Analyse die Achsen bestimmt, die die größte Varianz aufweisen. Die Vektoren dieser Hauptachsen werden zur Transformation der Ausgangsdaten eingesetzt. Die Hauptachsen werden als Hauptkomponenten bezeichnet und sind zueinander orthogonal. Für einen zweidimensionalen Beispiel-Datensatz sind in Abbildung 6 die Achsen mit der höchsten sowie der niedrigsten Varianz dargestellt.

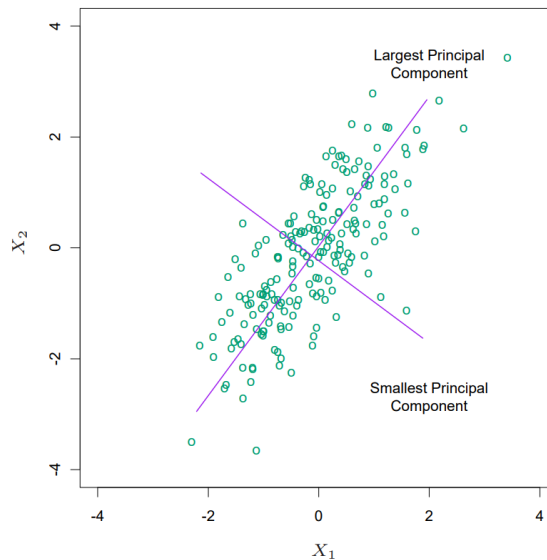


Abbildung 6: Darstellung der Achse mit der größten Varianz (largest principal component) sowie die mit der kleinsten Varianz (smallest principal component); entnommen aus (10), Seite 67.

Bei der linearen Diskriminanz-Analyse wird ein  $n$ -dimensionaler Raum auf einen  $(n-1)$ -dimensionalen Raum transformiert, wobei dieses Verfahren die Zielgrößen benötigt, also zu den überwachten Lernalgorithmen zählt. Für jede Klasse werden Mittelwert sowie Varianz berechnet. Die Transformation hat zum Ziel, den Abstand der Mittelwerte zu maximieren und die Varianz der einzelnen Klassen zu minimieren. In Abbildung 7 ist ein Beispiel für die Mittelwerte sowie Varianzen dreier Klassen dargestellt. Dieses Verfahren unterliegt der Annahme, dass die zugrundeliegenden Daten Gaussverteilt sind. (9, 10)

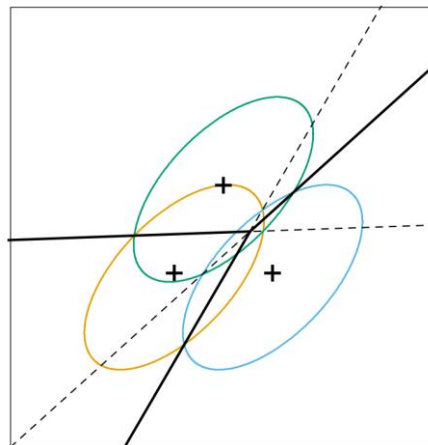


Abbildung 7: Darstellung des Mittelwerts (Kreuz) sowie der Kovarianz (Ellipse) für jede Klasse sowie die paarweisen Entscheidungsachsen der linearen Diskriminanzanalyse; entnommen aus (10), Seite 109.

### Erstellen eines Modells

Da in dieser Arbeit Kategorien unterschieden werden sollen, werden alle Modelle auf Klassifikatoren beruhen. Eine Klassifikation kann sowohl für binäre als auch für Multi-Klassen erfolgen. Beispiele für Klassifikatoren sind:

- der Bayes-Klassifikator, der auf der Annahme unabhängiger Merkmale beruht, die bei realen Problemen selten auftreten,
- der Nächste-Nachbarn-Algorithmus (engl. k-nearest-neighbor, short: KNN), der anfällig für den Fluch der Dimensionalität (engl. curse of dimensionality) ist, also bei einem hochdimensionalen Problem die benachbarten Datenpunkte nicht ausreichen, eine

## Kapitel 2: Theoretische Grundlagen

belastbare Vorhersage zu treffen, und häufig in Kombination mit Datenreduktionstechniken verwendet wird,

- die Support Vector Machine, die zwei sich nicht überlappende Klassen linear separiert, (11)
- sowie der Entscheidungsbaum.

Da der in dieser Arbeit eingesetzte Klassifikator auf Entscheidungsbäumen beruht, wird dieser ausführlich eingeführt.

Ein Entscheidungsbaum startet am sogenannten Wurzel-Knoten (engl. root node) mit allen Daten, der gleichzeitig ein Entscheidungsknoten (engl. decision node) ist. Für jeden Entscheidungsknoten wird das Kriterium gesucht, das die Daten am besten voneinander trennt und den größten Informationsgewinn liefert. Liegt danach nur noch eine Klasse vor, so wird dieser als Endknoten (engl. leaf node) bezeichnet. Als Maß für den Informationsgewinn wird üblicherweise der Gini-Index, die Entropie oder die Falschklassifizierungsrate herangezogen, deren Maximum angestrebt wird um die sogenannte Unreinheit (engl. impurity) des gesamten Entscheidungsbaums zu minimieren. Die allgemeine Berechnung des Gini-Index ist in Gleichung (13) zu sehen.  $p_i$  ist die Wahrscheinlichkeit, mit der ein Datenpunkt Klasse  $i$  zugeordnet wird. In Abbildung 8 sind alle drei Maße für eine binäre Klassifikation über die Wahrscheinlichkeit aufgetragen.

$$\text{Gini} = 1 - \sum_{i=1}^n (p_i)^2 \quad (13)$$

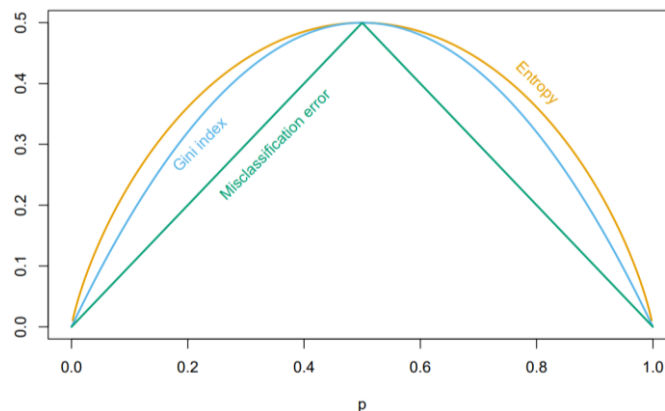


Abbildung 8: Auftragung der Entropie, des Gini-Indexes sowie der Falschklassifizierungsrate über die Wahrscheinlichkeit für eine binäre Klassifikation zur Veranschaulichung des Informationsgewinns; entnommen aus (10), Seite 309.

Da in der Regel für alle Daten mehr als ein Merkmal vorliegt, wird an jedem Entscheidungsknoten für jedes Merkmal eine binäre Klassifikation durchgeführt, bei der das sogenannte Attribut, der Trenn-Wert, so lange angepasst wird, bis der Gini-Index maximal ist. Ausgewählt wird am Ende das Attribut, das insgesamt den höchsten Informationsgewinn liefert. (9, 10) Je höher der erzielte Informationsgewinn ist, desto höher ist der Wert der Wichtigkeit des dazugehörigen Merkmals. Am Ende der Klassifizierung ergibt sich eine Rangliste der wichtigsten Merkmale des Entscheidungsbaums. (12) Bei der rekursiven Merkmalseliminierung wird rekursiv das Merkmal mit der geringsten Wichtigkeit für die Vorhersage entfernt.

Der Entscheidungsbaum ist robust gegen Ausreißer, dennoch ist ein einzelner Baum aufgrund der festen Attribute instabil. Leichte Änderungen in den Eingangsdaten haben einen großen Einfluss auf die Entscheidungen, was die Vorhersagen erheblich verschlechtern kann. Zur Stabilisierung der Ergebnisse werden die sogenannten Ensemble-Methoden, die mehr als ein einziges Ergebnis zur Vorhersage verwenden, eingesetzt.

Für den Fall des Entscheidungsbaums wird die Bagging-Methode (=bootstrap aggregating) verwendet, die in diesem speziellen Fall besser unter dem Begriff Zufallswald (engl. random forest) bekannt ist. Ein Zufallswald besteht aus vielen Entscheidungsbäumen, die wiederum nur auf einer Teilmenge aller Merkmale sowie Datensätze trainiert werden, da die Datensätze zufällig mit Zurücklegen ausgewählt werden. Im Fall des Testens liefert jeder Baum eine Vorhersage, die per Mehrheitsentscheid die Klasse vorhersagen. Die Vorhersagewahrscheinlichkeit des Zufallswalds gibt an, wie viele der Bäume für die jeweilige Klasse „gestimmt“ haben. (9, 10)

**Bewertung von Vorhersageergebnissen**

Zur Beurteilung der Qualität des verwendeten Modells werden dessen Vorhersagen in der sogenannten Wahrheitsmatrix (engl. confusion matrix), wie sie in Abbildung 9 für eine binäre Klassifikation zu sehen ist, zusammengefasst. Die Werte der einzelnen Felder werden für jede Klasse in die Kategorien *richtig positiv RP* (engl. true positive), *richtig negativ RN* (engl. true negative), *falsch positiv FP* (engl. false positive) sowie *falsch negativ FN* (engl. false negative) zugeordnet. (9) Im Fall von mehr als zwei Klassen werden für Klasse *i* der Wert (i,i) als *richtig positiv*, alle Werte der *i*-ten Spalte als *falsch negativ*, alle Werte der *i*-ten Zeile als *falsch positiv* und alle verbleibenden Werte der Wahrheitsmatrix als *richtig negativ* bezeichnet. (13) Ein Beispiel kann Abbildung 27 im Anhang entnommen werden.

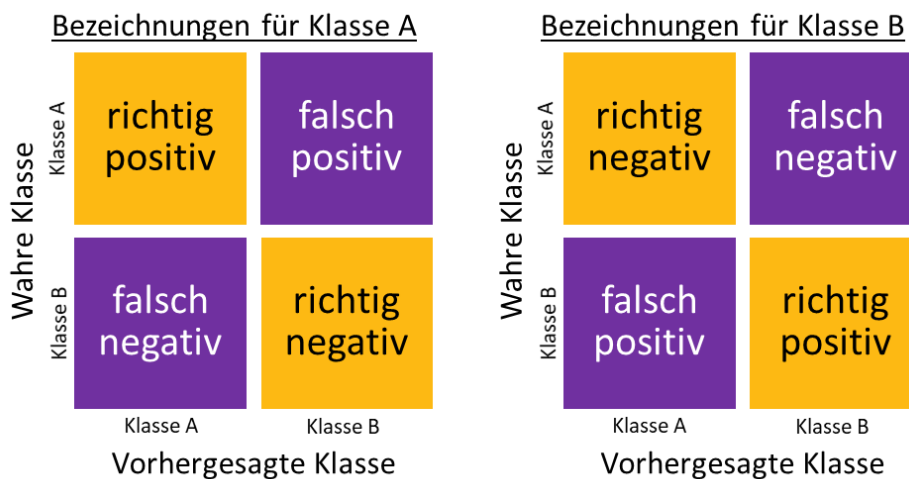


Abbildung 9: Schematische Darstellung einer binären Wahrheitsmatrix inklusive Bezeichnungen der Vorhersagen: Links für Klasse A; Rechts für Klasse B.

Zur Beurteilung der Qualität der Vorhersage werden in dieser Arbeit für jede Klasse einzeln die Präzision (engl. precision), Berechnung Gleichung 14, sowie die Sensitivität (engl. recall), Berechnung Gleichung (15, herangezogen. Darüber hinaus existieren weitere Maße zur Beurteilung der Vorhersagequalität wie beispielsweise die Spezifität (engl. specificity), die Genauigkeit (engl. accuracy) sowie das F-Maß (engl. F1-score). Da die Genauigkeit nur den Anteil aller richtigen Klassifikationen angibt und keine Rückschlüsse auf die einzelnen Klassen zulässt, wurde auf diesen Wert in dieser Arbeit verzichtet. (9)

$$\text{Präzision}_{\text{Klasse } i} = \frac{RP_{\text{Klasse } i}}{RP_{\text{Klasse } i} + FP_{\text{Klasse } i}} \tag{14}$$

$$\text{Sensitivität}_{\text{Klasse } i} = \frac{RP_{\text{Klasse } i}}{RP_{\text{Klasse } i} + FN_{\text{Klasse } i}} \tag{15}$$

### *Einfluss der Datenauswahl auf das Ergebnis*

Werden die Algorithmen des überwachten maschinellen Lernens auf reale Daten angewendet, so liegt eine erhöhte Wahrscheinlichkeit vor, dass nicht alle Klassen im selben Umfang vorliegen. Üblicherweise gelten die Daten ab einem Klassenverhältnis von 1:100 als unausgewogen (engl. imbalanced). Einige Klassifikatoren, unter anderem die auf einem Entscheidungsbaum basierenden, liefern für die kleinere Klasse ein schlechteres Vorhersageergebnis, da für diese weniger Endknoten existieren und somit die Gewichte dieser Ergebnisse eine geringere Konfidenz aufweisen.

Liegt bei den Gesamtdaten eine Unausgewogenheit der Klassen vor, existieren mehrere Möglichkeiten diese Ausgewogenheit herzustellen. Dazu zählen die Überabtastung (engl. oversampling), bei der Werte der kleineren Klasse dupliziert und dem Gesamtdatenset hinzugefügt werden, sowie die Unterabtastung (engl. undersampling), bei der Werte der größeren Klasse ausgelassen werden. Ein Beispiel für die Überabtastung ist der SMOTE-Algorithmus (synthetic minority oversampling technique; dt. synthetische Minoritätsüberabtastungstechnik), bei dem ein künstlicher Datenpunkt der kleineren Klasse als Mittelpunkt der benachbarten Punkte, unter Berücksichtigung des euklidischen Abstands, erzeugt wird. (14) In dieser Arbeit wird eine Unterabtastung durch das zufällige Auslassen der größeren Klasse erzielt.

## 3 STAND DER TECHNIK

---

In diesem Kapitel wird ein Überblick über die für diese Arbeit relevanten Themen der passiven akustischen Überwachung in der Industrie sowie in Natur und Umwelt sowie über die Klassifikation von Schweinelauten gegeben.

### 3.1 PASSIVE AKUSTISCHE ÜBERWACHUNG

Als passive akustische Überwachung (engl. passive acoustic monitoring, PAM) wird die Überwachung von Volumina mittels Schall bezeichnet, wobei nicht die Wechselwirkung von aktiv eingebrachtem Schall mit dem Volumen analysiert, sondern der im Volumen erzeugte Schall betrachtet wird. Von den vielfältigen Anwendungsfeldern der passiven akustischen Überwachung, die Volumina von Spanwerkzeugen bis Ozeane umfassen, werden einige im Folgenden exemplarisch vorgestellt.

#### 3.1.1 Passive akustische Überwachung in der Industrie

Abrupte Änderungen der Eigenspannungen eines Werkstoffs, beispielsweise durch Versetzungsbewegungen, Diffusion oder Kriechen ausgelöst, führen zu akustischer Emission (AE), die passiv überwacht werden kann. Mittels dieser Methode lässt sich der Verschleiß von Werkzeugen detektieren und überwachen, wozu charakteristische Merkmale des akustischen Signals, wie beispielsweise die Anzahl der Amplituden über einem Schwellwert oder die Signalintensität, berechnet werden. (3)

Da der detektierte Schall, wie von Chen et al. beschrieben, mehr als eine Quelle haben kann und dieser, wie von Ravindra et al. untersucht, abhängig von den Schneidparametern ist, können Schwellwerte einzelner Merkmale für Werkzeugverschleiß nicht universell festgelegt werden. (3, 15) Die Komplexität erfordert eine Kombination der Merkmale sowie bei Bedarf die Erweiterung des Merkmalsraums, um eine Trennung der Zustände zu ermöglichen. Die Wavelet-Transformation sowie die Empirische Modenzerlegung sind Beispiele für Zerlegungen bzw. Transformationen des akustischen Signals, anhand derer wieder charakteristische Merkmale ermittelt werden können. (15, 16) Neben Methoden der multivarianten Statistik, wie beispielsweise der lineare Diskriminanz-Analyse, können zur Zustandsklassifikation neuronale Netze zum Einsatz kommen. (3) Die neuronalen Netze können nicht nur mit Merkmalen der akustischen Signale oder den akustischen Daten selbst trainiert werden, sondern es ist auch möglich mittels Bildern des Zeitsignals oder der Short-Time-Fourier-Transformation Zustände zu unterscheiden. Zur Anomalie-Detektion akustischer Signale kommen häufig bildbasierte neuronale Netze zum Einsatz. (17–19)

#### 3.1.2 Passive akustische Überwachung in Natur und Umwelt

Die Bioakusik beschäftigt sich unter anderem mit der Lauterzeugung und Lautausbreitung von Tieren. (20) Dabei spielt die passive akustische Überwachung eine wichtige Rolle, da Lebensräume untersucht werden können, ohne die darin lebenden Tiere in ihrem Verhalten zu stören. Zwar ist prinzipiell auch eine optische Überwachung möglich, diese eignet sich aber häufig nicht für kleine, nachtaktive und/oder versteckt lebende Tiere. (21) Eingesetzt wird die passive akustische Überwachung beispielsweise bei Fledermäusen, Elefanten und Meeressäugetieren. Anhand der akustischen Signale kann auf die Gruppengröße geschlossen und somit der Bestand bedrohter Tierarten überwacht werden. (21–23)

### 3.2 KLASSIFIKATION VON SCHWEINELAUTEN

Die Identifikation von negativen Stresslauten sowie der Klassifikation von Schweinelauten ist Gegenstand zahlreicher Veröffentlichungen. In Abschnitt 3.2.1 wird ein Einblick in die Literatur mit Fokus auf den Arbeiten, auf die im Abschnitt 0 näher vorgestelltem Paper Bezug genommen wird, gegeben.

#### 3.2.1 Übersicht Literatur

In einer der ersten wissenschaftlichen Arbeiten von 1954 über das Verhalten von Schweinen werden die emotionalen Laute von Schweinen als Musik-Noten dargestellt. (24) Im Paper der Autoren Schrader und Hammerschmidt von 1997 werden mögliche Vorgehensweisen zur Extraktion von Merkmalen beschrieben. (25) Darauf aufbauend verfolgten zwei Autorengruppen parallel unterschiedliche Ansätze zur Stimmungsanalyse von Schweinen.

Die Autoren Schön, Puppe, Manteuffel et al. extrahierten 1998 mehrere, geordnete Werte aus dem Frequenz- sowie Amplitudenbereich als Merkmale. (26) Die Anwendung der linearen prädiktiven Kodierung (Linear predictive coding, LPC) der Lautsignale, die in der Sprachsynthese verwendet wird, ermöglichte 2001 die Reduktion der Merkmale auf 12 LPC-Koeffizienten, die zur Erkennung und Klassifikation von negativen Stresslauten mittels eines neuronalen Netzes eingesetzt werden. (27) Daraus wurde das System STREMODO zur Erkennung und Registrierung von Stresslauten entwickelt. Das LPC-Modell reagiert „empfindlich auf Veränderungen der Resonanzfrequenzen“ und somit auch indirekt auf die motorische Anspannung der Tiere. Zum Training sowie Testen wurden Laute von Ferkeln im Alter von 2 sowie 5 Wochen der Kategorien *Stress* sowie *kein Stress* sowie Laute von Sauen während der Laktation und Stallgeräuschen, beides Kategorie *kein Stress*, verwendet. Die Falschklassifizierungsrate lag bei 5 %. Die verwendete Vorgehensweise weist nach Aussage der Autoren eine hohe Unempfindlichkeit gegenüber Hintergrundgeräuschen auf. Dieses System ist auf die Erkennung hoher Töne von Ferkeln spezialisiert. (28) Das Verfahren wurde patentiert und von der Firma Big Dutchman vertrieben. (29) Das System ist auf der Website der Firma Big Dutchman in Verbindung mit dem Patentnamen nicht mehr aufzufinden. Darüber hinaus existiert kein jüngeres Paper der Autorenschaft. Das System wurde jedoch begleitend in anderen Publikationen, beispielsweise in (30–32), als Referenz eigener Untersuchungen verwendet.

Eine Übersicht anderer Merkmale von Stress- sowie Schmerzlauten veröffentlichten 2003 die Autoren Marx, Horn, von Borell et al. Diese werden mit den verschiedenen Aufnahmesituationen korreliert und die Varianzen der Werte beschrieben. Zu diesen Merkmalen gehören unter anderem die Lautlänge, die maximalen Amplituden im Zeit- sowie Frequenzbereich sowie die Energie des Lauts. Aufgenommen wurden die Laute von Ferkeln bis zu einem Alter von 19 Tagen in Situationen, in denen die Tiere mit oder ohne Anästhetikum kastriert oder fixiert wurden. Nach Aussage der Autoren können die entwickelten Lautparameter perspektivisch die Grundlage bilden, bei einer geringen Vokalisationsdichte Laute automatisiert zu klassifizieren. (33)

Die in den Folgejahren erschienenen Veröffentlichungen verweisen sehr häufig auf die Arbeit mindestens einer der acht genannten Autoren, wie beispielsweise in (34–37), in denen der Merkmalsraum, die Aufnahmesituation und/oder das Vorhersagemodell angepasst wird. Eines dieser Papers, verfasst von Briefer et al. (38), wird in folgendem Abschnitt genauer analysiert.

### 3.2.2 Paper von Briefer et al. (38)

In dem von Briefer et al. im Jahr 2022 veröffentlichten Paper wurden Merkmale identifiziert, die zur Unterscheidung von emotionalen Zuständen von Schweinen mittels eines neuronalen Netzes eingesetzt werden können. Untersucht wurden 7414 Schweinelaute von 411 Schweinen. (38)

#### **Datenerfassung**

Aufgenommen wurden die 7414 Laute in 19 unterschiedlichen Situationen eines gesamten Schweinelebens, von der Kastration über das Säugen bis hin zur Schlachtung. Die Schweine wurden einzeln bzw. in Kleingruppen einer der 19 Situationen ausgesetzt und ihre Laute aufgezeichnet. Jede der Situationen wurde von mehreren Tierärzten unabhängig nach „positiv“ oder „negativ“ klassifiziert, ohne dass das Verhalten des individuellen Tieres betrachtet wurde. Darüber hinaus sind die Laute hinsichtlich ihrer Anzahl in den 19 Klassen unausgeglichen.

#### **Datenvorverarbeitung**

Alle Laute wurden manuell aus kontinuierlichen Aufnahmen herausgeschnitten, wobei darauf geachtet wurde, Störgeräusche oder Überlappung mehrerer Laute zu vermeiden. Alle verwendeten Laute weisen eine Länge von 0,043 s bis 3,595 s auf.

#### **Merkmale der Laute**

Ein Teil der Merkmale wurde mittels der von Linguistikern verwendeten Software *Praat* ermittelt. Im Folgenden sind alle 10 verwendeten Merkmale aufgelistet und beschrieben.

##### **Lautlänge – Dur (s)**

Die Lautlänge beschreibt die Länge eines Lauts, die manuell mithilfe des Zeitsignals sowie des Spektrogramms ausgewählt wird

##### **Amplituden-Variation – AmpVar (dB/s), Amplituden-Modulations-Rate – AmpModRate (s<sup>-1</sup>), Amplituden-Modulations-Ausmaß – AmpModExtent (dB)**

Alle drei Parameter wurden auf Grundlage der Intensitätskontur, die mit *Praat* erzeugt wird, bestimmt. Eine genauere Beschreibung ist dem Paper von Charlton et al. (39) zu entnehmen. Die drei Merkmale werden wie folgt berechnet:

- AmpModRate = Anzahl der Amplitudenmodulationen / Lautlänge
- AmpModExtent = Mittelwert der einzelnen Peak-to-Peak-Ausmaße der einzelnen Amplitudenmodulationen
- AmpVar = kumulative Amplitudenmodulationen (=Summe der einzelnen Peak-to-Peak-Ausmaße der einzelnen Amplitudenmodulationen) / Lautlänge

##### **Erstes Quantil – Q25% (Hz), Zweites Quantil – Q50% (Hz), Drittes Quantil – Q75% (Hz)**

Die Quantile werden vom Spektrum des Lauts bestimmt und geben an, bei welcher Frequenz 25%-50%-75% der Gesamtamplitude des Spektrums erreicht werden. Das Quantil von 50% entspricht dem Schwerpunkt des Spektrums.

##### **Frequenz mit der höchsten Amplitude – FPeak (Hz)**

Das Merkmal FPeak gibt die Frequenz mit der größten Amplitude des cepstral-smoothed Spektrums, das mittels *Praat* berechnet wird, an. (35)

##### **Harmonizität – Harmonicity**

Die Berechnung der Harmonizität erfolgte mit einem Befehl aus *Praat*. Die Harmonizität, im Englischen auch unter Harmonic-to-Noise-Ratio bekannt, ist ein Maß für die Periodizität eines akustischen Signals. (40)

### Wiener Entropie – WienEntropy

Die Wiener Entropie ist ein für *Praat* veröffentlichtes Skript, das ein Maß für die Ebenheit der spektralen Leistungsdichte darstellt. Dieser Wert entspricht dem Quotienten aus dem geometrischen und des arithmetischen Mittelwerts und wird mittels einer logarithmischen Skala ausgedrückt. (41)

### Modellbildung und –bewertung

Im Rahmen dieses Papers wurden zwei Vorhersagemodelle getestet. Zum einen eine permutierte Diskriminanzanalyse auf der Grundlage der vier ausgewählten Merkmale Lautlänge, Amplituden-Modulations-Rate, Zweites Quantil (50 %) sowie Wiener Entropie, zum anderen eine Klassifikation mittels eines neuronalen Netzes, das Bilder der Spektrogramme der Laute verarbeitet. Im Fall der Diskriminanzanalyse waren 85,2 % der Vorhersagen in die beiden Kategorien „positiv“ und „negativ“ richtig. (38)

### Übertragung einiger Merkmale auf ein akustisches Signal eines realen Schweinelauts

Zur Visualisierung der Merkmale des Papers werden ausgewählte Merkmale des Zeit- sowie des Amplitudenbereichs anhand eines akustischen Signals eines Schweinelauts erläutert. In Abbildung 10 ist das Zeitsignal eines Schweinelauts mit linearer Skala sowohl komplett als auch ausschnittsweise gezeigt.

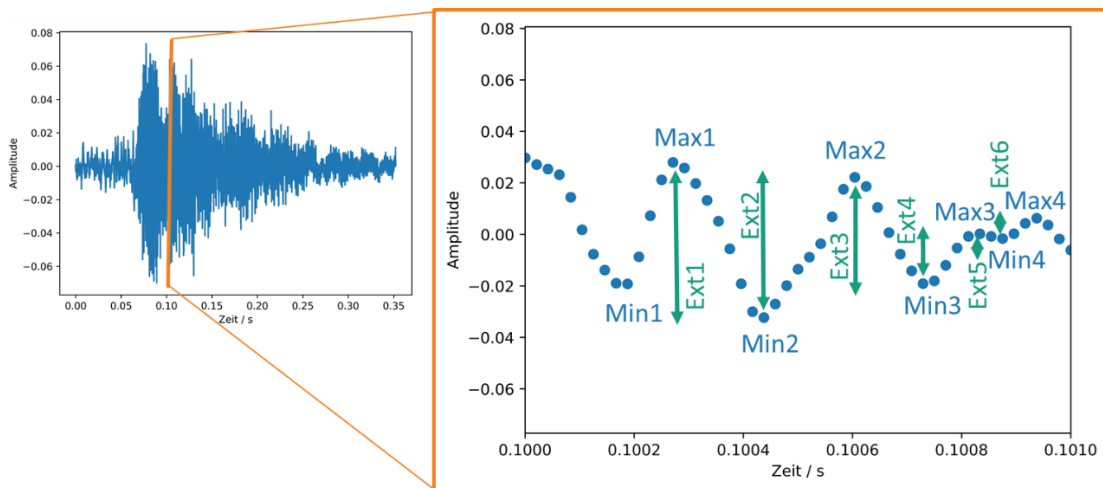


Abbildung 10: Zeitsignal eines Schweinelauts sowie ein Ausschnitt daraus inklusive Beschriftung von charakteristischen Eigenschaften, die zur Bestimmung der Merkmale relevant sind.

Zur Bestimmung der Merkmale AmpModRate, AmpVar sowie AmpModExtent werden alle lokalen Maxima (Max) und Minima (Min) des Signals bestimmt. Für die Amplitudenmodulationsrate wird die Anzahl aller Extrema durch die dazugehörige Signallänge dividiert. Für die Merkmale AmpVar sowie AmpModExtent werden die Beträge der Differenzen zwischen zwei aufeinanderfolgenden Extrema (Ext) benötigt. Die Amplitudenvariation gibt den Mittelwert der Beträge an, das Amplituden-Modulations-Ausmaß die Summe der Beträge, normiert auf die dazugehörige Signallänge.

Zur Bestimmung des Spektrums wird eine Fourier-Transformation des akustischen Signals durchgeführt. Dieses ist in Abbildung 11 für den Schweinelaut aus Abbildung 10 dargestellt.

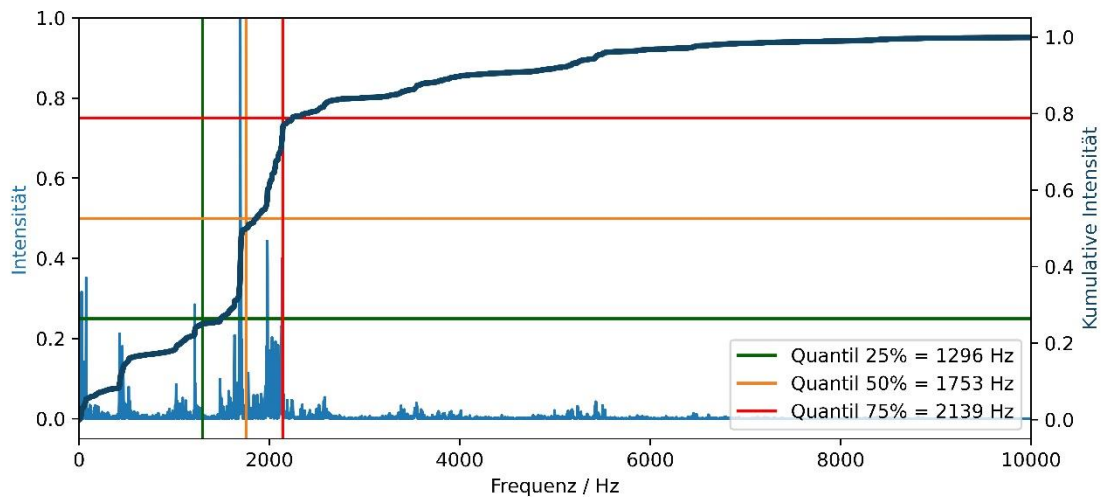


Abbildung 11: Spektrum eines Schweinelauts inklusive der kumulativen Intensität zur Bestimmung der Frequenzquantile.

Zur Bestimmung der Frequenzquantile werden alle Intensitäten kumulativ aufsummiert. Die Frequenzquantile geben die Frequenz an, bei der 25 % / 50 % / 75 % der kumulativen Intensität erreicht werden.

### 3.2.3 Zusammenfassung Literatur zur Klassifikation von Schweinelauten

In der Literatur finden sich einige Ansätze zur Klassifikation von Schweinevokalisationen. Die dort untersuchten Laute wurden jedoch in der Regel in gezielt herbeigeführten Situationen aufgenommen, indem Tiere separiert oder gestresst wurden. In den Papern von Schön, Puppe, Manteuffel et al. sowie Marx, Horn, von Borell et al. wurden die Laute von Ferkeln untersucht, während Briever et al. alle Lebensabschnitte eines Schweins betrachten. Da bei Briever et al. die Situationen und nicht das Tierverhalten bei der Klassifikation der Laute berücksichtigt wurden, erweist sich die Übertragbarkeit auf ein Assistenzsystem schwierig, zumal auch nur die Laute ohne Umgebungsgeräusche in das Training der verwendeten Modelle eingeflossen sind. Das patentierte System STREMODO ist unempfindlich gegenüber Hintergrundgeräuschen und wurde bereits von der Firma BigDutchman vertrieben. Ein aktueller Vertrieb kann nicht mehr nachgewiesen werden.

Da ein Assistenzsystem zur Erkennung negativer bzw. bedenklicher Schweinelaute nicht oder nicht mehr existiert, wird aufbauend auf den Ergebnissen der vorher genannten Autoren eine Methodik für den späteren Einsatz in einem Assistenzsystem für einen Maststall mit Umgebungsgeräuschen entwickelt. Abgrenzend zu den vorgestellten Papern wird dabei auf die Separierung in einen abgetrennten Raum sowie die Manipulation einzelner Tiere zur Erzeugung von Trainingsdaten verzichtet. Zur Modellerstellung werden stattdessen Geräusche eines realen Maststalls mit etwa 150 Tieren verwendet, die neben einer gelabelten Schweinevokalisation sowohl überlagernde Laute mehrerer Tiere als auch Umgebungsgeräusche enthalten können.



## 4 MATERIALIEN UND METHODEN

### 4.1 DATENAUFNAHME

#### 4.1.1 Hardware

Zur Erzeugung der Audiodateien wurde die in Abbildung 12 schematisch gezeigte, mobile Audiobox verwendet.

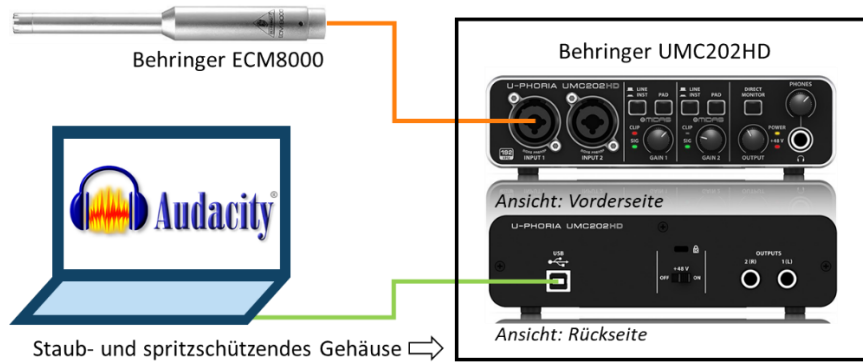


Abbildung 12: Schematischer Aufbau der mobilen Audiobox; mit Bildern aus (42–44).

Über das omnidirektionale Kondensator-Mikrofon (Typ ECM8000 von Behringer) mit einem Frequenzbereich von 20 Hz bis 20 kHz werden akustische Signale aufgenommen und mittels des Analog-Digital-Wandlers (Behringer U-Phoria UMC202HD) digitalisiert. Die Datenblätter der verwendeten Geräte können dem Anhang 10.2 entnommen werden. Die 48 V Spannungsversorgung des Mikrofons erfolgt genauso wie die Datenverbindung zum Mess-Computer über eine USB-Schnittstelle. Die analoge Vorverstärkung des akustischen Signals wird mittels eines Drehreglers am AD-Wandler eingestellt. Zum Schutz der elektronischen Anschlüsse vor Staub, Spritzwasser sowie anderen Umwelteinflüssen im Schweinestall befindet sich der AD-Wandler in einem darauf ausgelegten Gehäuse. Zusätzlich wurden nur staubdichte Anschlüsse gewählt. Die originale Audiobox ist in Abbildung 13 zu sehen.



Abbildung 13: Foto der mobilen Audiobox; Links: Außenansicht. Rechts: Innenansicht.

## Kapitel 4: Materialien und Methoden

### 4.1.2 Software

Zur Audioaufnahme wurde die frei verfügbare Software *Audacity* (Version 3.2.3) unter Windows mit einer Abtastfrequenz von 48 kHz und einer Amplitudenauflösung von 24-Bit, umgesetzt durch Puls-Code-Modulation, verwendet.

### 4.1.3 Vorgehensweise

Alle Audioaufnahmen wurden von einem Tierarzt der Stiftung Tierärztliche Hochschule Hannover durchgeführt, da neben dem Erstellen der Aufnahmen auch das Tierverhalten beobachtet und bewertet werden muss. Die Aufnahme erfolgte in einem Mast-Stall mit etwa 150 Schweinen an je einem Tag zweier unterschiedlicher Mastdurchgänge. Das Mikrofon hing zentral zwischen zwei Buchten auf einer Höhe von etwa 2,20 m. Eine Skizze des Stalls kann in Anhang 10.2 eingesehen werden. Alle Schweine befanden sich im selben Raum und können daher auf den Aufnahmen zu hören sein. Aufgrund des Ziels, Geräusche eines realen Stalls zur Erstellung einer Klassifikationsmethodik zu verwenden, erfolgte keine Separation einzelner Schweine. Eine Separation von einzelnen Schweinen wäre aus tierschutzrechtlichen Gründen nicht möglich gewesen, da die Separation bereits als Tierversuch eingestuft wird.

Mit *Audacity* wurden Zeiträume von 30 bis 45 Minuten aufgenommen, während der Tierarzt parallel das Tierverhalten, das zu markanten Lauten zuzuordnen war, inklusive Zeitstempel notiert hat. Für eine Variation der Schweinelaute wurde der Aufnahmeort für jede Aufnahme verändert. Aufgrund der hohen Schweineanzahl konnten nicht alle in diesem Zeitraum auftretenden Geräusche klassifiziert werden. Die genaue Identifikation der Start- und Endsamplerepunkte der einzelnen Geräusche erfolgte im Rahmen dieser Arbeit durch das Anhören der einzelnen Geräusche sowie der Betrachtung der dazugehörigen Amplitude mittels *Audacity*. Diese kontinuierlichen Aufnahmen inklusive Labels des Schweineverhaltens dienen als Datengrundlage für diese Arbeit. Insgesamt liegen dieser Arbeit 1041 abgegrenzte Laute zugrunde.

## 4.2 DATENANALYSE

### 4.2.1 Software

Die Datenanalyse erfolgt mit Python (Version 3.9.10). Die wichtigsten Bibliotheken inklusive Version können folgender Liste entnommen werden:

- joblib, Version 1.3.1
- matplotlib, Version 3.7.2
- numpy, Version 1.24.4
- pandas, Version 1.3.5
- scipy, Version 1.11.1
- scikit-learn, Version 1.3.0
- seaborn, Version 0.12.2
- Signal-Analysis, Version 0.1.26
- soundfile, Version 0.12.1

### 4.2.2 Vorgehensweise

Das Erstellen eines Modells, das bestimmte auftretende Arten eines akustischen Signals voneinander trennen soll, erfolgt prinzipiell nach dem in Abbildung 14 gezeigten Ablauf. In Kapitel 5.1 wird diese Methodik mit für den Anwendungsfall konkreten Werten versehen.

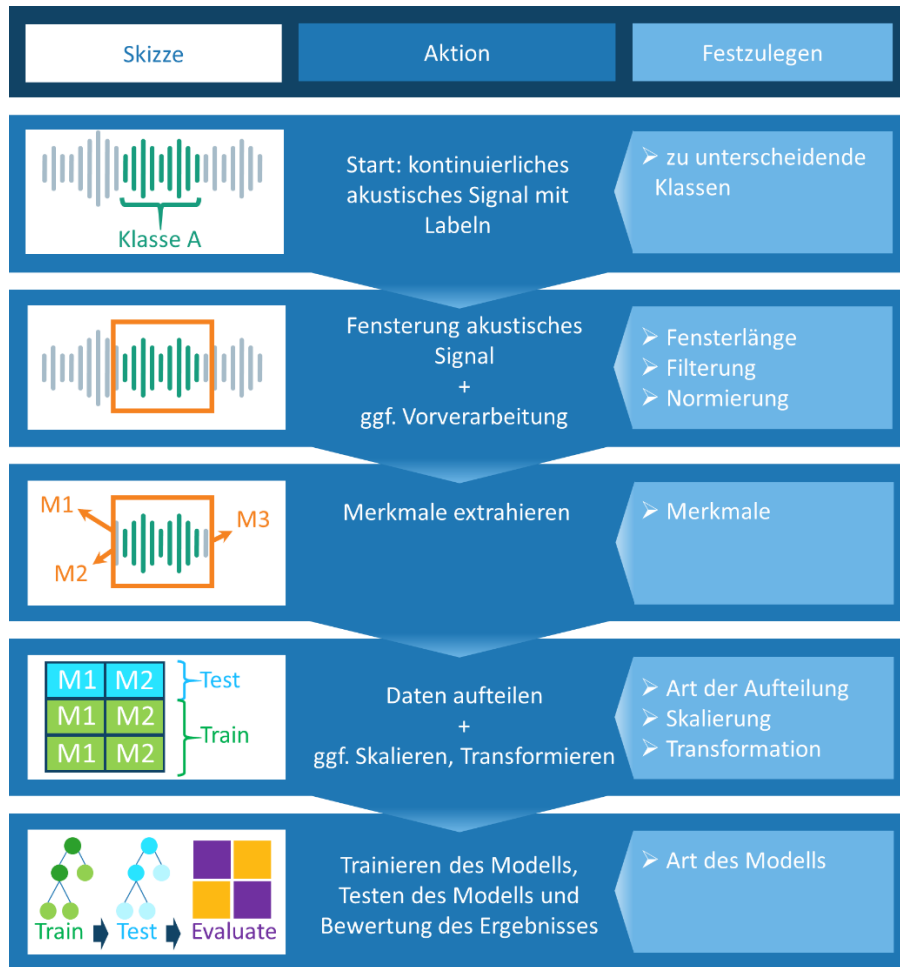


Abbildung 14: Allgemeiner Ablauf der Datenauswertung. Links: schematische Darstellung der Vorgehensweise. Mitte: Beschreibung der Vorgehensweise. Rechts: Übersicht der festzulegenden Parameter.

Voraussetzung für das Erstellen eines Modells sind kontinuierliche akustische Signale, die komplett oder abschnittsweise gelabelt sind. Das Labeln erfolgt wie in Abschnitt 4.1.3 beschrieben, wobei ein möglichst präzises Label pro Laut vergeben wird. Da dadurch eine hohe Anzahl unterschiedlicher Labels zustande kommt, muss eine Entscheidung getroffen werden, welche Labels voneinander getrennt werden sollen. Zur besseren Übersicht werden diese in Klassen zusammengefasst. Die abschnittsweise gelabelten akustischen Signale werden im nächsten Schritt als Fenster aus dem kontinuierlichen Signal extrahiert und einzeln analysiert. Jedes der Fenster wird in dieser Arbeit als *Geräusch* bezeichnet. Sofern das Geräusch von einem Schwein erzeugt wurde, wird die Bezeichnung *Laut* bzw. *Schweinelaut* verwendet. Auf jedes dieser Fenster kann sowohl eine Filterung als auch eine Normierung angewendet werden. Anschließend werden Charakteristika des Geräuschs ermittelt, die in Merkmalen zusammengefasst werden. Dazu können Eigenschaften wie Lautstärke, Geräuschlänge oder dominierende Frequenzen gezählt werden. Diese Merkmale werden zum Trainieren eines Klassifikationsmodells sowie zur Vorhersage der Klasse herangezogen. Um die Vorhersagen des Modells bewerten zu können, werden die Merkmale in ein Trainings- sowie ein Testdatenset aufgeteilt, von denen der erste Teil zur Erstellung und der zweite zur Evaluierung des Modells herangezogen wird. Das Modell setzt sich aus Skalierung, Transformation und Klassifikation der Merkmale zusammen, wobei die ersten beiden Schritte optional sind. Die Evaluierung des Modells erfolgt anhand der Wahrheitsmatrix, in der die richtigen und falschen Klassifizierungen dargestellt sind. Anhand der Komponenten dieser Matrix können weitere Werte zur Bewertung der Vorhersage abgeleitet werden.



## 5 ERGEBNISSE UND DISKUSSION

Im Hinblick auf die Anwendung, kritische, tierwohlgefährdende Situationen anhand passiver akustischer Analyse zuverlässig zu erkennen, wird eine Methodik benötigt, die die kontinuierlich aufgenommenen akustischen Signale auswertet. In Abbildung 14 ist das Konzept der Methodik schematisch dargestellt, das im nachfolgenden Schritt ausgearbeitet wurde. Dabei wurden Parameter festgelegt und zur Optimierung der Methodik bei Bedarf angepasst. Die Evaluierung der Methodik an unbekanntem akustischen Signalen sowie die Integration in ein (Prototypen-) System mit kontinuierlicher Verarbeitung ist nicht mehr Teil dieser Arbeit, sondern wird im Projekt *SmartPigHome* weitergeführt.

### 5.1 AUSARBEITUNG DER METHODIK

Die Ausarbeitung der Methodik zur Detektion kritischer Situationen im Schweinestall erfordert eine erste Festlegung von denen in Abbildung 14 vorgestellten Parametern. Zu diesen zählen die Geräuschklassen nach denen mittels des Modells unterschieden werden soll, die Vorverarbeitung des akustischen Signals, die Merkmale, die die Grundlage des Modells bilden sollen, deren Vorverarbeitung sowie die Auswahl des Modells selbst.

#### 5.1.1 Auswahl der Klassen

##### *Hierarchische Struktur der Geräuschklassen*

Im Konsortium des Projekts *SmartPigHome* wurde die in Abbildung 15 gezeigte hierarchische Übersicht der Geräusche im Schweinestall aufgestellt. Diese erhebt keinerlei Anspruch auf Vollständigkeit. Ziel der Aufstellung der Hierarchie war das Ordnen der auftretenden, beobachteten Schweinelaute (Ebene 5) mit dem Anspruch, so wenig wie möglich das Tierverhalten zu vermenschlichen, jedoch auch den Grad des potentiellen Stressniveaus herauszustellen. Da nur eine Beobachtung der Tiere und kein Tierversuch, wie beispielsweise das Messen von Hirnströmen oder das gezielte Provozieren von Stress, durchgeführt wurde, unterliegen sowohl die Klassen als auch die Einteilung der akustischen Signale in diese Klassen einer menschlichen Unsicherheit.

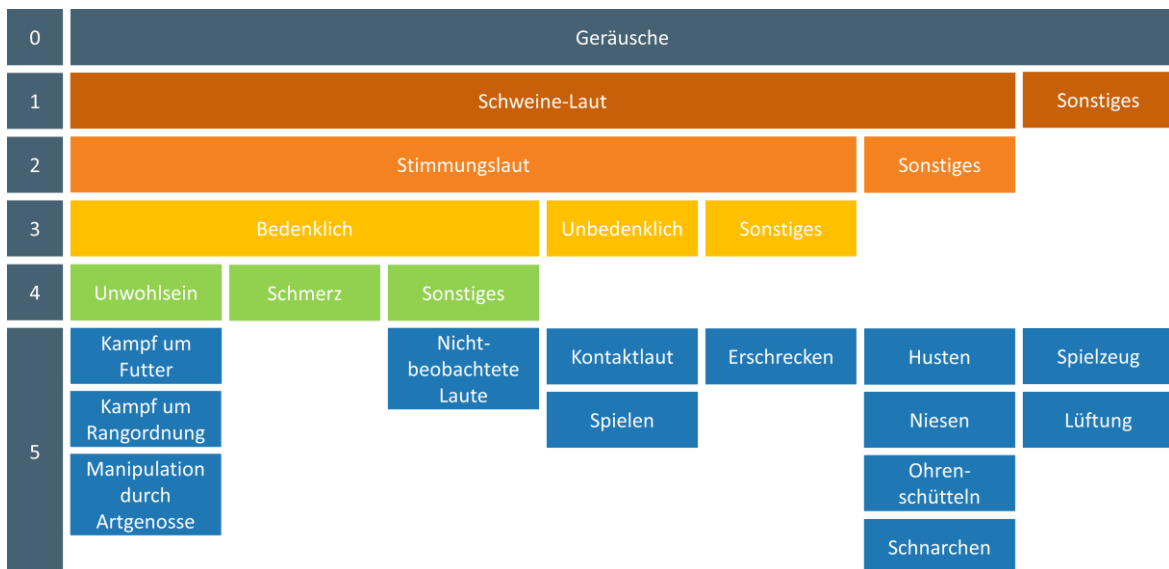


Abbildung 15: Hierarchische Anordnung der Geräuschklassen.

### Einteilung der Geräuschklassen in zwei Hauptklassen

Aus der Sicht des Schweinehalters ist eine automatische Detektion vieler der in Ebene 5 aufgezeigten Geräusche bzw. Laute erstrebenswert. Aus technischer Sicht ist ein Vorhersagemodell sowohl bei binären als auch bei Multiklassen möglich. Da die Gesamtanzahl der Geräusche einiger kleiner Klassen bei der Anzahl von 1041 Testgeräuschen zu gering ist und somit keine Gleichverteilung der in Ebene 5 vorliegenden Geräusche existiert, beschränkt sich diese Arbeit auf ein Zwei-Klassen-Modell. Die zwei Klassen müssen so gewählt werden, dass in beiden Klassen ausreichend akustische Daten vorliegen. Für das Projekt ist das Erkennen von kritischen, tierwohlgefährdenden Situationen am notwendigsten, daher wird zwischen den Klassen „Bedenkliche Laute“ und „Restliche Geräusche“ unterschieden.

Darüber hinaus ist eine ausgewogene Anzahl an Geräuschen in beiden Klassen notwendig, da abhängig vom Modell, die überwiegende Klasse eines unausgewogenen Datensets bevorzugt werden könnte. Zur Vermeidung der Einflüsse auf das Vorhersageergebnis wurde aus den vorhandenen Geräuschen ein ausgewogenes Datenset ausgewählt.

#### 5.1.2 Auswahl der Fensterbreite

Zur Auswahl einer geeigneten Fensterbreite kann die in Abbildung 16 gezeigte Verteilung der Länge aller 933 Schweinelaute der 1041 Geräusche herangezogen werden. Zur besseren Übersicht wurde in dieser Darstellung auf die langen Geräuschabschnitte ohne Schweinelaut sowie die vereinzelt negativen Laute mit einer Länge größer als sieben Sekunden verzichtet. Ein Großteil der Schweinelaute weist eine Länge zwischen 0,1 s und 3 s auf.

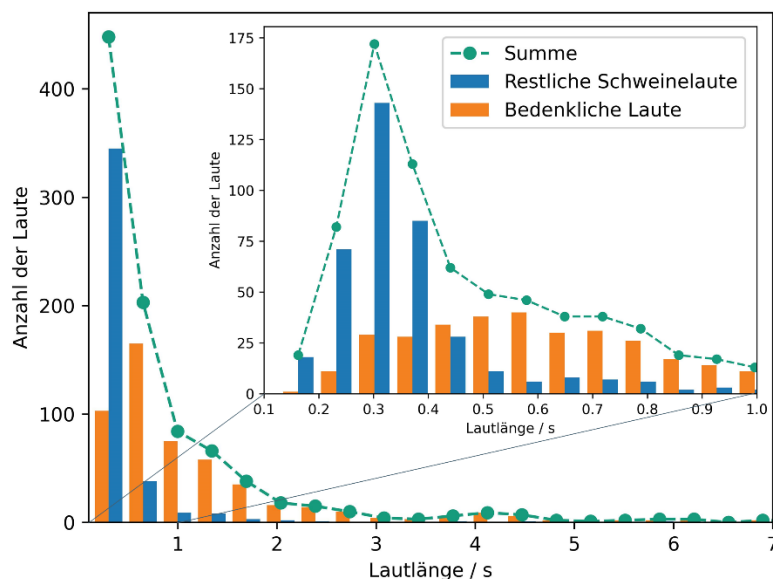


Abbildung 16: Verteilung der Längen von 933 Schweinelauten.

Unter dem Gesichtspunkt der späteren Übertragbarkeit der Ergebnisse auf ein System mit kontinuierlicher Verarbeitung, das akustische Signale mit gleichbleibender Puffer-Größe analysieren soll, wird eine feste Fenstergröße zum Trainieren des Modells verwendet. Ein gelabelter Signalabschnitt kann sowohl im Ganzen als auch in Unterabschnitte unterteilt für die Merkmalsextraktion herangezogen werden. Aufgrund der großen Varianz der Längen der Signalabschnitte werden in Abbildung 17 Möglichkeiten der Fensterung abhängig von der gewählten Fensterbreite sowie der jeweiligen Geräuschlänge gezeigt.

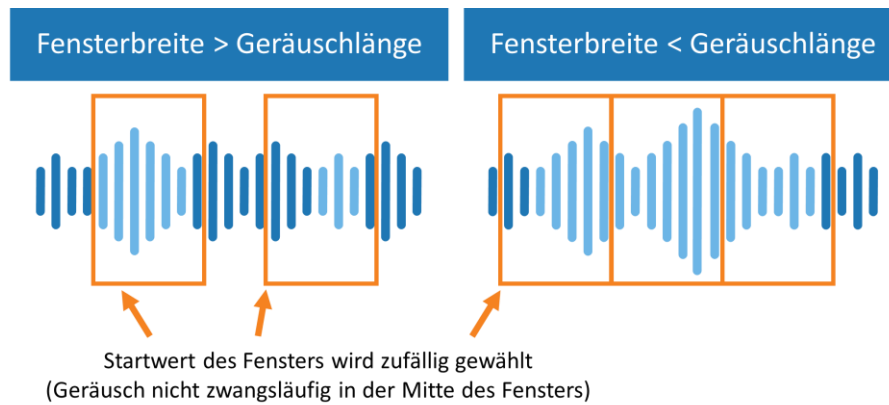


Abbildung 17: Schematische Darstellung der beiden Möglichkeiten, abhängig von der vorliegenden Geräuschlänge sowie der gewählten Fensterbreite das Fenster um das Geräusch zu legen.

Im ersten Fall, in dem die gewählte Fensterbreite größer ist als die Geräuschlänge, liegt das gesamte Geräusch innerhalb des Fensters. Zur Vermeidung einer Beeinflussung des späteren Modells, stimmen Mitte des Geräuschs und Mitte des Fensters nicht zwingend überein. Der Startwert des Fensters ergibt sich aus der Differenz des Startwerts des Geräuschs sowie einem zufälligen Wert, der Maximal der Differenz von Geräuschlänge und Fensterbreite entsprechen kann. Im zweiten Fall, in dem die Geräuschlänge größer als die gewählte Fensterbreite ist, wird zunächst die Geräuschlänge durch die Fensterbreite dividiert. Die Anzahl der Fenster entspricht diesem Quotienten addiert um eins. Der Startwert des ersten Fensters wird äquivalent zum ersten Fall ermittelt. Jedes darauffolgende Fenster beginnt mit dem Endwert des vorherigen.

Unter Berücksichtigung der Verteilung der Geräuschlängen ist eine Fensterbreite von einer Sechzehntel Sekunde (0,0625 s), was etwa der Hälfte des kürzesten Lauts entspricht, bis hin zu 4 s, also etwa eine Sekunde größer als ein Großteil der längsten Schweinelaute, möglich. Aufgrund dieser großen Varianz an möglichen Fensterbreiten erfolgt die Auswahl einer optimalen Fensterbreite in Kapitel 5.2.1.

### 5.1.3 Vorverarbeitung des akustischen Signals

#### Filterung

Aufgrund der Charakteristik des Mikrofons (Auszüge des Datenblatts im Anhang 10.3) dürfen nur Frequenzen zwischen 20 Hz und 20 kHz ausgewertet werden. Daher wurde auf jedes Fenster mit der Filterfunktion `scipy.signal.filtfilt` ein Butterworth-Hochpass (Ordnung 4) bei 20 Hz sowie ein Butterworth-Tiefpass (Ordnung 4) bei 20 kHz angewendet.

#### Normierung

Alle vorliegenden Audioaufnahmen sind in einem Amplitudenbereich zwischen -1 und +1 normiert, wobei die betragsmäßige Amplitude von eins dem Wert entspricht, bei dem das Audiosystem übersteuert.

#### 5.1.4 Merkmale

Als Merkmale werden im ersten Schritt zum einen die häufig verwendeten Charakteristika von Zeitsignalen wie

- die Extrema,
- die Energie,
- die Lautstärke,
- die vier Momente der Stochastik
- sowie die Fourierkoeffizienten der maximalen Frequenz,

## Kapitel 5: Ergebnisse und Diskussion

zum anderen die Merkmale aus dem in Abschnitt 0 beschriebenen Paper von Briefer et al. (38) herangezogen. Insgesamt ergeben sich dadurch 36 Merkmale.

### Merkmale Zeitbereich

Die Merkmale der Fenster der akustischen Signale werden einmal auf der linearen Skala zwischen -1 und +1, die in Abschnitt 5.1.3 zur Normierung näher erläutert wird, und einmal auf einer logarithmischen Dezibel-Skala bestimmt. Da Werte von exakt Null im Zeitsignal vorliegen können und der Logarithmus bei null gegen minus unendlich konvergiert, ist für die Umrechnung der linearen Skala in die logarithmische notwendig, das gesamte Signal um einen Wert, in diesem Fall  $10^{-32}$ , in der Amplitude zu verschieben. Die Umrechnung erfolgt mit Gleichung (16).

$$\text{logarithmische Amplitude} = 20 * \log_{10} (|\text{lineare Amplitude}|) \quad (16)$$

Alle Merkmale des Zeitbereichs sind in Tabelle 1 zusammengefasst. Die Implementierung dieser Merkmale in Python ist Tabelle 18 des Anhangs zu entnehmen.

*Tabelle 1: Merkmale des Zeitbereichs.  
x = Merkmal wurde auf dazugehöriger Skala bestimmt.*

Merkmals (Abkürzung)	Lineare Skala	Logarithmische Skala	Beschreibung
maximale Amplitude (maxAmp)	x	x	Betragsmaximum der Amplitude
Energie (Energie)	x	x	Fläche unter der Kurve der gleichgerichteten Amplitude
Lautstärke (Lautst)	x	x	Quadratisches Mittel der Amplitude
Mittelwert (MW_Z)	x	x	Mittelwert der gleichgerichteten Amplitude
Varianz (Var_Z)	x	x	Varianz der gleichgerichteten Amplitude
Schiefe (Skew_Z)	x	x	Schiefe (engl. skewness) der gleichgerichteten Amplitude
Wölbung (Kurt_Z)	x	x	Wölbung (engl. kurtosis) der gleichgerichteten Amplitude
Amplituden-Modulations-Rate (AmpModRate)	x		Anzahl der Amplituden-Modulationen im Zeitbereich, normiert auf die Fensterlänge
Amplituden-Modulations-Ausmaß (AmpExt)	x	x	Mittelwert der einzelnen Peak-to-Peak-Ausmaße der einzelnen Amplitudenmodulationen
Amplituden-Variation (AmpVar)	x	x	Summe der einzelnen Peak-to-Peak-Ausmaße der einzelnen Amplitudenmodulationen, normiert auf die Fensterlänge
Harmonizität (Harm)	x		Harmonizität (auch engl. Harmonic-to-noise-ratio) ist das Verhältnis aus periodischen- und nicht-periodischen Komponenten eines Sprachsignals

**Merkmale Frequenzbereich**

Für die Merkmale des Frequenzbereichs wird sowohl die Intensität mit linearer sowie logarithmischer Skala berücksichtigt. Die Umrechnung der Intensität erfolgt nach Gleichung (17).

$$\text{logarithmische Intensität} = 20 * \log_{10} (|\text{lineare Intensität}|) \quad (17)$$

Alle Merkmale des Frequenzbereichs sind analog zu denen des Zeitbereichs in Tabelle 2 hinterlegt. Die Python-Implementierung ist in Tabelle 19 im Anhang zu finden.

*Tabelle 2: Merkmale des Frequenzbereichs.  
x = Merkmal wurde auf dazugehöriger Skala bestimmt.*

Merkmale	Lineare Skala	Logarithmische Skala	Beschreibung
Mittelwert (MW_F)	x	x	Mittelwert der Amplitude
Varianz (Var_F)	x	x	Varianz der Amplitude
Schiefe (Skew_F)	x	x	Schiefe (engl. skewness) der Amplitude
Wölbung (Kurt_F)	x	x	Wölbung (engl. kurtosis) der Amplitude
Frequenz mit höchster Amplitude (maxF)	x		Frequenz, bei der die Amplitude des Spektrums maximal ist
Realer Anteil Fourier-Koeffizient (real_FC)	x		Realteil des Fourierkoeffizienten des Maximums der Fouriertransformation
Imaginärer Anteil Fourier-Koeffizient (imag_FC)	x		Imaginärteil des Fourierkoeffizienten des Maximums der Fouriertransformation
Frequenz Quantil 25% (FQ25)	x		Quantil, das beschreibt, bei welcher Frequenz 25% der gesamten Intensität des Spektrums erreicht werden
Frequenz Quantil 50% (FQ50)	x		Quantil, das beschreibt, bei welcher Frequenz 50% der gesamten Intensität des Spektrums erreicht werden
Frequenz Quantil 75% (FQ75)	x		Quantil, das beschreibt, bei welcher Frequenz 75% der gesamten Intensität des Spektrums erreicht werden
Spektrale Ebenheit = Wiener Entropie (SpE)	x	x	Quotient aus dem arithmetischen Mittel und dem geometrischen Mittel der spektralen Leistungsdichte

Der Parameter der spektralen Ebenheit ist nicht für alle Geräusche bestimmbar, da der geometrische Mittelwert zu Null wird. Dies trifft auf 0,1 % der Geräusche bei einer Fensterlänge von einer Sekunde zu. Bei der Erstellung des Modells werden diese Datensätze nicht mitberücksichtigt und im späteren praktischen Einsatz wird dieser Anteil zu vernachlässigen sein.

**5.1.5 Vorverarbeitung der Merkmale**

Für die optimale Ausnutzung aller vorhandenen Daten wird die stratifizierte k-fache Aufteilung mit dem Faktor k = 5 eingesetzt. Dabei ist immer eine Faltung Testdatensatz während die anderen vier zum Training des Modells genutzt wird, was einem Training-Test-Verhältnis von 80:20 entspricht.

## Kapitel 5: Ergebnisse und Diskussion

Durch diese Art des Aufteilens der Daten liegen am Ende des Trainings fünf Modelle vor. Aufgrund der Werteverteilung innerhalb eines Merkmals sowie der Abweichungen der Merkmale untereinander ist eine wie in Kapitel 5.1.3 beschriebene Skalierung erforderlich, damit alle Merkmale gleichwertig in das Modell eingehen. Dazu wird der *StandardScaler* verwendet. Eine Transformation der Merkmale erfolgt in diesem Kapitel nicht.

### 5.1.6 Auswahl eines Modells

Wie in Kapitel 2.4 beschrieben existiert eine große Anzahl an Modellen, die zum Vorhersagen von Klassen verwendet werden kann. Da die Zuordnung der beobachteten Laute zu den verschiedenen Klassen möglichst nachvollziehbar sowie schnell erfolgen soll, wurde der Zufallswald als Modell ausgewählt. Der Zufallswald von *scikit-learn* hat die Grundparameter: 100 Bäume, Bewertung anhand des Gini-Indexes sowie keine Begrenzung in der Tiefe. Für die Vorhersage einer Klasse eines einzelnen Geräuschs werden die Vorhersagewahrscheinlichkeiten der beiden Klassen herangezogen. Der Laut wird der Klasse mit der höheren Wahrscheinlichkeit zugeordnet. Liegt die Wahrscheinlichkeit in beiden Fällen bei 50 %, so erfolgt die Zuordnung zufällig. In Abschnitt 5.2.3 wird dies näher erläutert.

### 5.1.7 Bewertung der Vorhersage

In der Wahrheitsmatrix werden die Ergebnisse der Vorhersage des Modells graphisch zusammengestellt. Die daraus abgeleiteten Werte für Präzision und Sensitivität dienen als Bewertungsgrundlage des Modells. Da eine fünffache Aufteilung erfolgt, werden die Werte für Präzision sowie Sensitivität gemittelt sowie jeweils der minimale und maximale Wert der fünf einzelnen Modelle bestimmt. Diese Werte werden beim Vergleich von Modellen, die auf unterschiedlichen Parametern beruhen, zur Bestimmung des besten Modells wie folgt herangezogen. Im ersten Schritt werden die Mittelwerte verglichen und das Modell mit dem höchsten Mittelwert als bestes erachtet. Unterscheiden diese sich nicht, so wird das Modell als bestes erachtet, das die höchsten Minima der Präzisionen und Sensitivitäten aufweist. Weisen die Minima auch keine Unterschiede auf, so wird das Modell mit dem höchsten Maximum als bestes erachtet.

## 5.2 ANALYSE DER METHODIK ANHAND DER TRAININGSDATEN

Auf Basis der in Kapitel 5.1 angelegten Methodik inklusive der Vorhersagebewertung werden in diesem Kapitel sowohl eine für die Trainingsdaten optimale Fensterbreite vorgestellt als auch weiterführend die möglichen Einflussfaktoren auf die Falschklassifizierungs-Rate analysiert. Des Weiteren werden drei Methoden zur Reduktion der Merkmale getestet und verglichen.

### 5.2.1 Festlegung optimale Fensterbreite

Aufbauend auf den in Kapitel 5.1.2 vorgestellten Ablauf sowie den bereits festgelegten Parametern kann eine Festlegung einer optimalen Fensterbreite erfolgen. Unter Berücksichtigung der in Tabelle 3 aufgelisteten Parameter wurden für jede Fensterbreite fünf Modelle nach der k-fachen Aufteilung, basierend auf den Merkmalen, die aus Fenstern verschiedener Länge bestimmt wurden, trainiert. Die Ergebnisse der Testdaten sind in Tabelle 4 zusammengestellt, wobei die Zeile mit den höchsten Werten hervorgehoben ist.

Tabelle 3: Parametersatz Variation Fensterlänge.

<b>Klasse A</b>	Bedenkliche Laute	<b>Merkmale</b>	alle aus Kapitel 5.1.4
<b>Klasse B</b>	Restliche Geräusche	<b>Skalierung</b>	StdScaler
<b>Fensterbreite</b>	wird variiert	<b>Transformation</b>	Keine
<b>Anz. Fenster pro Klasse</b>	variiert	<b>Modell</b>	Zufallswald

Tabelle 4: Ergebnisse bei variierenden Fensterbreiten.

Fensterbreite in s	Anzahl Fenster pro Klasse	Klasse A: Bedenkliche Laute						Klasse B: Restliche Geräusche					
		Präzision			Sensitivität			Präzision			Sensitivität		
		Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.
0,0625	11646	87%	86%	87%	87%	86%	88%	87%	86%	88%	87%	86%	87%
0,125	5953	88%	86%	89%	87%	86%	88%	87%	85%	88%	88%	86%	90%
0,25	2633	90%	88%	91%	88%	87%	90%	87%	86%	89%	90%	88%	91%
0,5	1677	91%	89%	93%	89%	88%	91%	89%	88%	91%	91%	88%	94%
1	968	92%	90%	94%	90%	88%	95%	90%	88%	95%	92%	90%	94%
1,5	756	91%	87%	94%	90%	89%	93%	90%	88%	93%	91%	86%	95%
2	666	91%	88%	94%	90%	85%	94%	90%	85%	94%	91%	88%	94%
3	594	90%	88%	93%	86%	83%	88%	85%	82%	87%	90%	88%	93%
4	568	87%	85%	90%	83%	80%	85%	82%	78%	85%	88%	85%	90%

Da in diesem Vergleich das Modell der Fensterlänge einer Sekunde die insgesamt höchsten Werte aufweist, wird diese Fensterlänge verwendet.

### 5.2.2 Analyse des Geräuschanteils eines Fensters

Aufgrund der unterschiedlichen Geräuschlängen sowie der daraus folgenden Fensterung ist der Geräuschanteil nicht in jedem Fenster gleich. Mögliche Einflüsse der Falschklassifizierungen werden in diesem Abschnitt genauer untersucht.

Zur Unterscheidung der Einflüsse der Geräuschlängen und den Möglichkeiten der Fensterung werden die Geräusche mit einer kürzeren Geräusch- als Fensterlänge und jene mit einer Geräuschlänge größer oder gleich der Fensterlänge voneinander getrennt betrachtet, da ein Geräusch im ersten Fall komplett und im zweiten Fall nur teilweise im betrachteten Fenster liegt.

Mit der in Tabelle 5 zusammenfassten Parameter, sind in Tabelle 6 für beide der genannten Fälle sowohl die richtig als auch falsch klassifizierten Zuordnungen in Abhängigkeit des Geräuschanteils im betrachteten Fenster, zusammengefasst in Intervalle von 20 %, zusammengetragen.

## Kapitel 5: Ergebnisse und Diskussion

Tabelle 5: Parametersatz zur Analyse des Geräuschanteils im Fenster sowie Analyse der Vorhersagewahrscheinlichkeit.

<b>Klasse A</b>	Bedenkliche Laute	<b>Merkmale</b>	alle aus Kapitel 5.1.4
<b>Klasse B</b>	Restliche Geräusche	<b>Skalierung</b>	StdScaler
<b>Fensterbreite</b>	1 Sekunde	<b>Transformation</b>	Keine
<b>Anz. Fenster pro Klasse</b>	968	<b>Modell</b>	Zufallswald

Tabelle 6: Übersicht Anteil der richtig- bzw. falsch-zugeordneten Geräusche in Abhängigkeit des Geräuschanteils im jeweils betrachteten Fenster, unterteilt nach dem Verhältnis Fenster- zu Geräuschlänge.

Anteil Geräusch im Fenster	Geräusche, für die gilt: Fensterlänge > Geräuschlänge						Geräusche, für die gilt: Fensterlänge < Geräuschlänge					
	Gesamt klassifiziert		Richtig klassifiziert		Falsch klassifiziert		Gesamt klassifiziert		Richtig klassifiziert		Falsch klassifiziert	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
[0,0; 0,2)	16	100	12	75	4	25	16	100	12	75	4	25
[0,2; 0,4)	275	100	232	84	43	16	60	100	41	68	19	32
[0,4; 0,6)	137	100	117	85	20	15	100	100	85	85	15	15
[0,6; 0,8)	104	100	90	87	14	13	155	100	142	92	13	8
[0,8; 1,0]	52	100	48	92	4	8	1021	100	985	96	36	4
<b>Gesamt</b>	<b>584</b>	<b>100</b>	<b>499</b>	<b>85</b>	<b>85</b>	<b>15</b>	<b>1352</b>	<b>100</b>	<b>1265</b>	<b>94</b>	<b>87</b>	<b>6</b>

In beiden Fenster-zu-Geräuschlänge-Verhältnissen nimmt der Anteil der Falschklassifizierungen tendenziell ab je größer der Geräuschanteil im Fenster ist. Dennoch muss berücksichtigt werden, dass weder die Anzahl der Geräusche innerhalb derselben Geräuschlänge gleichverteilt noch zwischen den beiden Geräuschlängen die Verteilung über die Anteile hinreichend gleich ist.

### 5.2.3 Analyse der Vorhersagewahrscheinlichkeit

Wie bereits in Abschnitt 5.1.6 erwähnt, erfolgt die Zuordnung der Geräusche in die zwei Hauptklassen in die mit der größeren Vorhersagewahrscheinlichkeit. In diesem Abschnitt wird diese genauer in Korrelation mit den Falschklassifizierungen und eine Möglichkeit zur Reduzierung der Falschklassifizierungsrate untersucht.

#### **Korrelation der Vorhersagewahrscheinlichkeit mit der Falschklassifizierungsrate**

Mit dem in Tabelle 5 angegebenen Parametersatz ergeben sich die Verteilung der Vorhersagewahrscheinlichkeiten, die in Abbildung 18 für alle Geräusche, untergliedert in Klassen sowie deren Bewertung der Klassifizierung, dargestellt ist.

Bei zunehmender Vorhersagewahrscheinlichkeit nimmt die Summe der richtigen Klassifizierungen zu, während die der falschen abnimmt. Im Intervall [0,5;0,6) ist festzustellen, dass bei Klasse A die Summe der falschen Klassifizierungen höher ist als die der richtigen.

In diesem Intervall sind darüber hinaus 4 Klassifizierungen mit einer Vorhersagewahrscheinlichkeit von 0,5 enthalten, von denen zwei richtig und sechs falsch Klasse A zugeordnet wurden. Diese Zuordnung erfolgt zufällig durch die zugrundeliegenden Zufallswälder, die standardmäßig aus 100 Bäumen bestehen. Zur Vermeidung einer solchen Falschklassifizierung wäre eine Möglichkeit, eine ungerade Anzahl an Bäumen zu wählen.

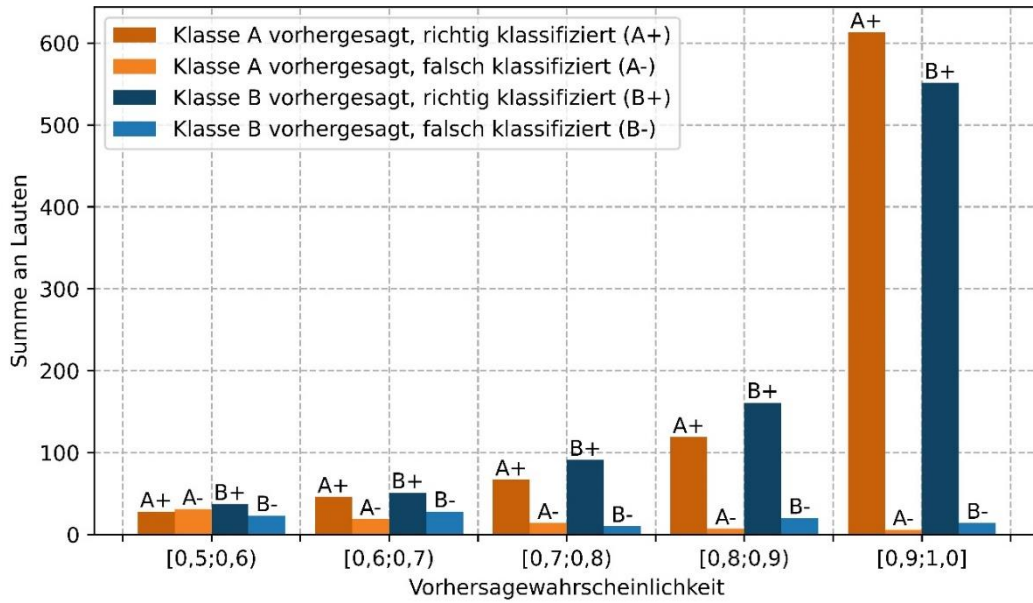


Abbildung 18: Histogramm (Bin-Größe = 0.1) für die Verteilung der Vorhersagewahrscheinlichkeiten aller vorhergesagten Geräusche, unterteilt nach Klassen sowie Klassifizierungsbewertung.

**Reduzierung der Falschklassifizierungsrate**

Diese Trainingsdaten weisen für alle Klassifizierungen ein Richtig-zu-Falsch-Klassifizierungs-Verhältnis von 55:45 im Intervall [0,5;0,6), von 67:33 im Intervall [0,6;0,7) sowie von 87:13 im Intervall [0,7;0,8) der Vorhersagewahrscheinlichkeiten auf. Diese Werte liegen unter dem Durchschnitt aller Klassifizierungen. Daher wäre eine Möglichkeit zur Reduzierung der Falschklassifizierungen alle Vorhersagen mit einer Wahrscheinlichkeit unterhalb eines Grenzwertes in eine dritte Klasse zuzuordnen.

In Abbildung 19 sind die Ergebnisse der Wahrscheinlichkeitsmatrizen sowie der daraus abgeleiteten Werte für Präzision und Sensitivität unter Berücksichtigung eines Grenzwertes zu einer dritten Klasse dargestellt. Diese dritte Klasse wird als *Uneindeutige Zuordnung* bezeichnet.

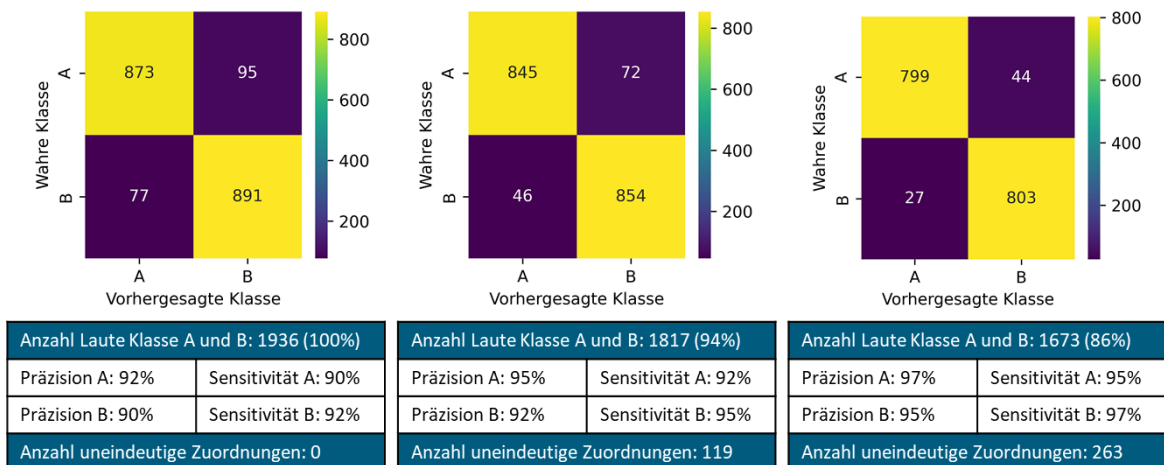


Abbildung 19: Wahrscheinlichkeitsmatrizen sowie Ergebnisse von mittlere Präzision und mittlere Sensitivität für alle Vorhersagewahrscheinlichkeiten Links: ≥ 50%, Mitte: ≥ 60% sowie Rechts: ≥ 70%.

Die Werte für Präzision und Sensitivität steigen bei zunehmendem Grenzwert und damit abnehmender Anzahl an Zuordnungen zu Klasse A oder B an. Wie in Tabelle 4 zu sehen variieren die Ergebnisse der fünf Klassifizierungen der k-fachen Aufteilung um bis zu 5 %. Daher bedeutet die Steigerung der Vorhersageergebnisse um 3 % bzw. 5 % durch die Anlegung eines Schwellwerts nicht pauschal eine Verbesserung der Vorhersage. Darüber hinaus muss beachtet werden, dass die

## Kapitel 5: Ergebnisse und Diskussion

Geräusche der Klasse *Uneindeutige Zuordnungen*, in diesem Beispiel 6 % bzw. 14 % der Geräusche, nicht vorhergesagt und in diesem Fenster keine bedenklichen Geräusche identifiziert werden. Eine nachgelagerte, separate Klassifizierung wäre daher notwendig.

### 5.2.4 Korrelation der Merkmale

Zur Analyse des Merkmalsraums wird die lineare Korrelation der Merkmale herangezogen. Die Beträge der Pearson-Koeffizienten sind für alle Merkmale in Abbildung 20 dargestellt.

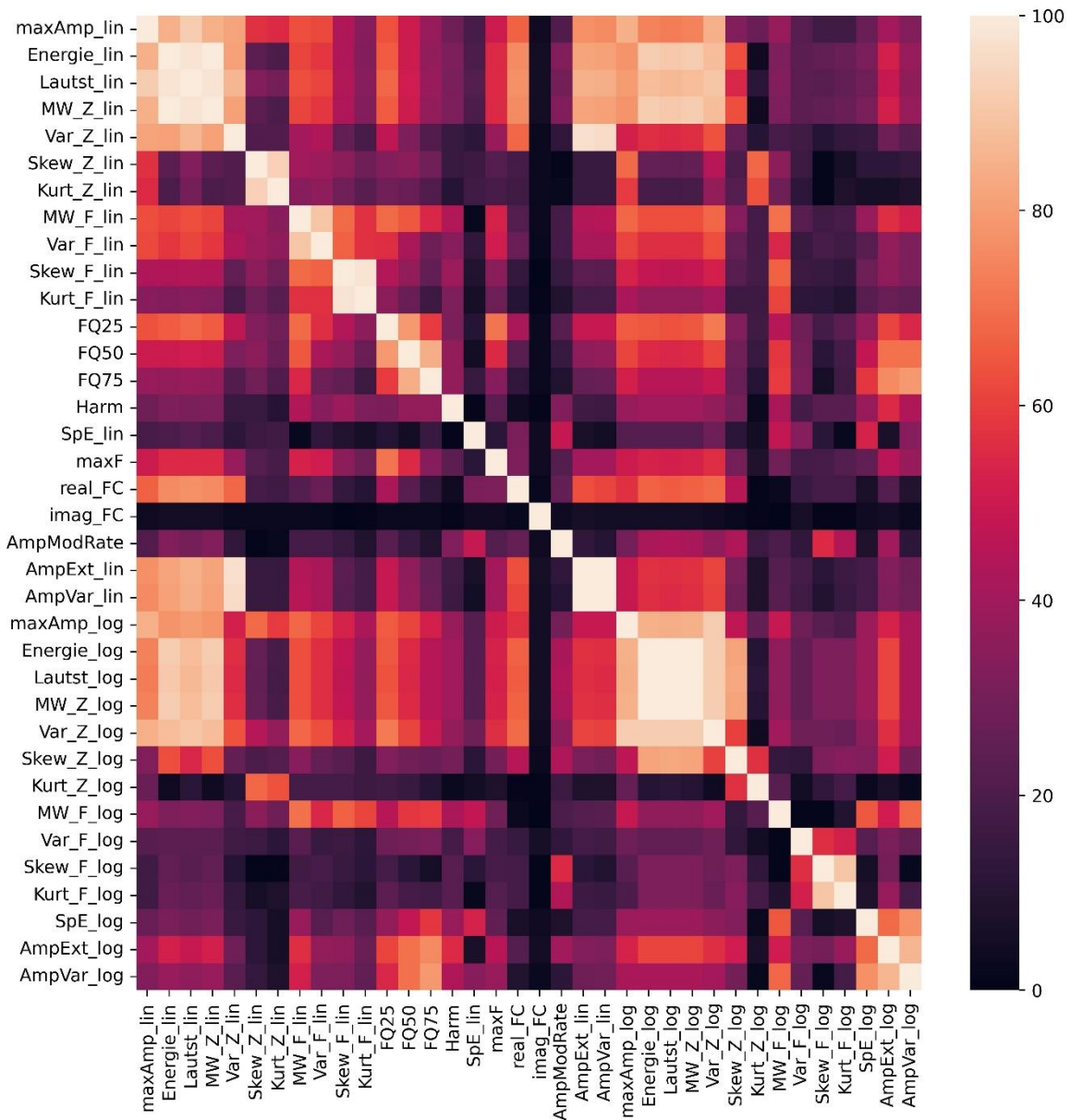


Abbildung 20: Ergebnisse der linearen Pearson-Korrelation aller 36 Merkmale als Betrag des Pearson-Koeffizienten in Prozent.

Maximale Amplitude, Energie, Lautstärke, Mittelwert sowie Varianz im Zeitbereich weisen sowohl für die lineare als auch die logarithmische Skala eine lineare Korrelation von über 80 % auf. Darüber hinaus korrelieren die Merkmale Schiefe und Wölbung des Zeit- sowie Frequenzbereichs, Mittelwert und Varianz des linearen Frequenzbereichs, und Amplituden-Modulations-Ausmaß und Amplituden-Variation auf linearer sowie logarithmischer Skala paarweise stark mit Koeffizienten von ebenfalls mindestens 80 %. Der Imaginäranteils des Fourierkoeffizienten weist als einziges Merkmal keine lineare Korrelation mit einem der anderen Merkmale auf.

### 5.2.5 Reduktion der Merkmale

Die starke Korrelation einiger Merkmale weist auf Redundanzen von Informationen zwischen den 36 Merkmalen hin. In diesem Abschnitt werden die rekursive Merkmalseliminierung, die Hauptkomponenten- sowie die lineare Diskriminanzanalyse zur Reduktion des Merkmalsraums jeweils ausgehend von den 36 Merkmalen getestet und verglichen. Ein reduzierter Merkmalsraum reduziert die Komplexität der Vorhersage.

#### Rekursive Merkmalseliminierung

Ziel des Einsatzes der rekursiven Merkmalseliminierung ist die Bestimmung einer reduzierten Anzahl an Merkmalen, auf deren Basis ebenso gute Vorhersagen beider Klassen wie unter der Berücksichtigung aller zur Verfügung stehenden 36 Merkmale getroffen werden. Im ersten Schritt wurde dazu erst die Anzahl der reduzierten Merkmale ermittelt. In einem zweiten Schritt wurden die Merkmale bestimmt, die am häufigsten ausgewählt wurden, wenn auf die vorher bestimmte Anzahl reduziert wurde.

#### Anzahl der wichtigsten Merkmale

Mit dem Parametersatz aus Tabelle 7 wurde mittels rekursiver Merkmalseliminierung die Anzahl der Merkmale reduziert und mit den verbleibenden Merkmalen mit denselben Parametern insgesamt fünf Modelle trainiert. In Abbildung 21 sind die Mittelwerte, das Minimum sowie das Maximum der Präzision sowie Sensitivität der fünf Modelle einer Merkmalsanzahl dargestellt.

Tabelle 7: Parametersatz rekursive Merkmals-Eliminierung zur Bestimmung der Merkmalsanzahl.

<b>Klasse A</b>	Bedenkliche Laute	<b>Merkmale</b>	<i>werden variiert</i>
<b>Klasse B</b>	Restliche Geräusche	<b>Skalierung</b>	StdScaler
<b>Fensterbreite</b>	1 Sekunde	<b>Transformation</b>	Keine
<b>Anz. Fenster pro Klasse</b>	968	<b>Modell</b>	Zufallswald

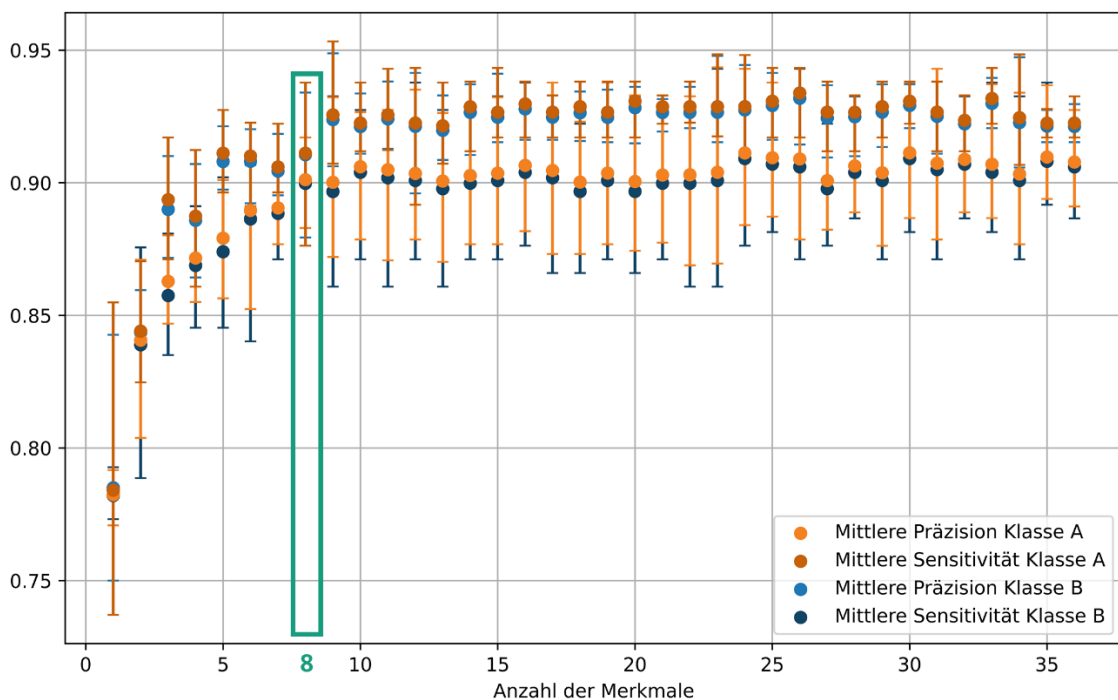


Abbildung 21: Mittelwert, Minima und Maxima der Präzisionen sowie Sensitivitäten der Ergebnisse der rekursiven Merkmalseliminierung, aufgetragen über die Anzahl der zur Modellerstellung verwendeten Merkmale.

## Kapitel 5: Ergebnisse und Diskussion

Bis zu einer Anzahl von acht Merkmalen steigen die mittlere Sensitivität von Klasse A sowie die mittlere Präzision von Klasse B von 78 % auf 90 % stetig an. Da die mittleren Präzisionen von Klasse A sowie die mittleren Sensitivitäten von Klasse B für alle Merkmale mindestens so hoch wie die jeweiligen anderen beiden Werte sind, wird zur Übersichtlichkeit nur noch die Sensitivität von Klasse A aufgrund ihrer niedrigsten Mittel- sowie Minimalwerte betrachtet.

Der Mittelwert der Sensitivität bleibt bei einer Anzahl höher als acht Merkmalen im Bereich 90 % bis 91 %. Der Minimalwert wird nie schlechter als 86 % und beträgt maximal 89 %. Die Anzahl der Merkmale kann von 36 auf acht reduziert werden, ohne dass das Vorhersageergebnis sich signifikant verschlechtert.

### Statistische Ermittlung der acht wichtigsten Merkmale

Da zur Bestimmung der wichtigsten Merkmale die k-fache Aufteilung zum Einsatz kam, liegt für jede Aufteilung aufgrund unterschiedlicher Trainingsdaten eine andere Rangliste an wichtigsten Merkmalen vor. Zur Bestimmung der Art der acht wichtigsten Merkmale wurden zur Abbildung dieser Varianz insgesamt 20 anstelle einer einzigen fünffach-Aufteilung durchgeführt. Verwendet wurde der in Tabelle 8 aufgeführte Parametersatz. In Tabelle 9 sind die durch die rekursive Merkmals-Eliminierung ausgewählten Merkmale inklusive ihrer Wahrscheinlichkeit, bei den durchgeführten 100 Eliminierungen ausgewählt zu werden, zusammengefasst.

*Tabelle 8: Parametersatz zur Bestimmung der Häufigkeit der Merkmale bei einer rekursiven Merkmalseliminierung auf acht Merkmale.*

<b>Klasse A</b>	Bedenkliche Laute	<b>Merkmale</b>	<i>werden variiert, Anzahl konstant bei 8</i>
<b>Klasse B</b>	Restliche Geräusche	<b>Skalierung</b>	StdScaler
<b>Fensterbreite</b>	1 Sekunde	<b>Transformation</b>	Keine
<b>Anz. Fenster pro Klasse</b>	968	<b>Modell</b>	Zufallswald

*Tabelle 9: Ausgewählte Merkmale sowie ihre Wahrscheinlichkeit bei einer rekursiven Merkmalseliminierung auf acht Merkmale ausgewählt zu werden.*

Nr.	Merkmalsbezeichnung	Wahrscheinlichkeit
1	AmpExt_log	100%
2	AmpExt_lin	100%
3	FQ50	100%
4	FQ75	99%
5	AmpVar_lin	99%
6	SpE_lin	98%
7	Skew_Z_log	91%
8	AmpModRate	74%
9	Kurt_F_log	42%
10	Lautst_log	41%
11	Energie_log	31%
12	MW_Z_log	20%
13	Energie_lin	2%
14	Skew_F_log	2%
15	MW_Z_lin	1 %

Bei der Reduktion auf acht Merkmale werden insgesamt 15 Merkmale von der rekursiven Merkmalseliminierung ausgewählt. Dies liegt in der Zusammensetzung der gewählten Trainingsdaten begründet. Sieben der Merkmale werden in über 90 % der Fälle ausgewählt. Das achte Merkmal wird in 74 % und das neunte in nur 42 % der Fälle ausgewählt. Da diese beiden Werte sich um mehr als 30 % unterscheiden, können die ersten acht Merkmale der Liste eindeutig als wichtigste acht Merkmale der gesamten Vorhersage definiert werden.

Mittels der acht ausgewählten Merkmale wurde mit der fünffach-Aufteilung Modelle trainiert, deren Vorhersageergebnis in Tabelle 10 zusammengestellt ist.

Tabelle 10: Ergebnisse Vergleich 36 gegenüber acht Merkmale.

Anzahl Merkmale	Klasse A: Bedenkliche Laute						Klasse B: Restliche Geräusche					
	Präzision			Sensitivität			Präzision			Sensitivität		
	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.
36	92%	90%	94%	90%	88%	95%	90%	88%	95%	92%	90%	94%
8	92%	92%	93%	91%	89%	93%	90%	88%	93%	93%	92%	93%

Die leichten Unterschiede der Präzisionen sowie Sensitivitäten zwischen den Vorhersagen ist in der Streuung der Werte der Merkmale, abhängig von der Aufteilung der Daten, begründet. Das Vorhersageergebnis bleibt gleich, die Anzahl der zugrundeliegenden Merkmale wird übersichtlicher, sodass im folgenden Abschnitt diese genauer analysiert werden können.

#### Analyse des statistischen Verhaltens der acht wichtigsten Merkmale

Von den acht wichtigsten Merkmalen werden fünf im Zeit- und drei im Frequenzbereich berechnet. In Abbildung 22 und Abbildung 23 sind die acht wichtigsten Merkmale in der Reihenfolge ihrer Wahrscheinlichkeit gegeneinander aufgetragen und farblich in die beiden Hauptklassen unterteilt. Dieser Darstellung wurden alle Daten ein ausgewogenes Datenset zugrunde gelegt. Auf der Hauptdiagonalen sind die Verteilungen der beiden Klassen für das jeweilige Merkmal dargestellt.

Die Verteilungen von Klasse A ist für die ersten fünf Merkmale erkennbar breiter als die der Klasse B. Bei der Betrachtung der graphischen Darstellungen der Merkmale gegeneinander können die beiden zu unterscheidenden Klassen meist mit einer gedachten Linie voneinander getrennt werden.

Die Merkmale AmpExt\_lin sowie AmpVar\_lin weisen eine lineare Abhängigkeit auf, da das eine Merkmal ein Maß für die Frequenz der mittleren und das andere der absoluten Änderung der Intensität ist. Die Merkmale AmpExt\_lin und AmpExt\_log sind aufgrund ihrer Berechnung logarithmisch miteinander verknüpft. Die beiden Frequenzquantile können nicht komplett unabhängig voneinander betrachtet werden, denn liegt die Frequenz von FQ50 höher so wird das auch für die Frequenz von FQ75 der Fall sein.

## Kapitel 5: Ergebnisse und Diskussion

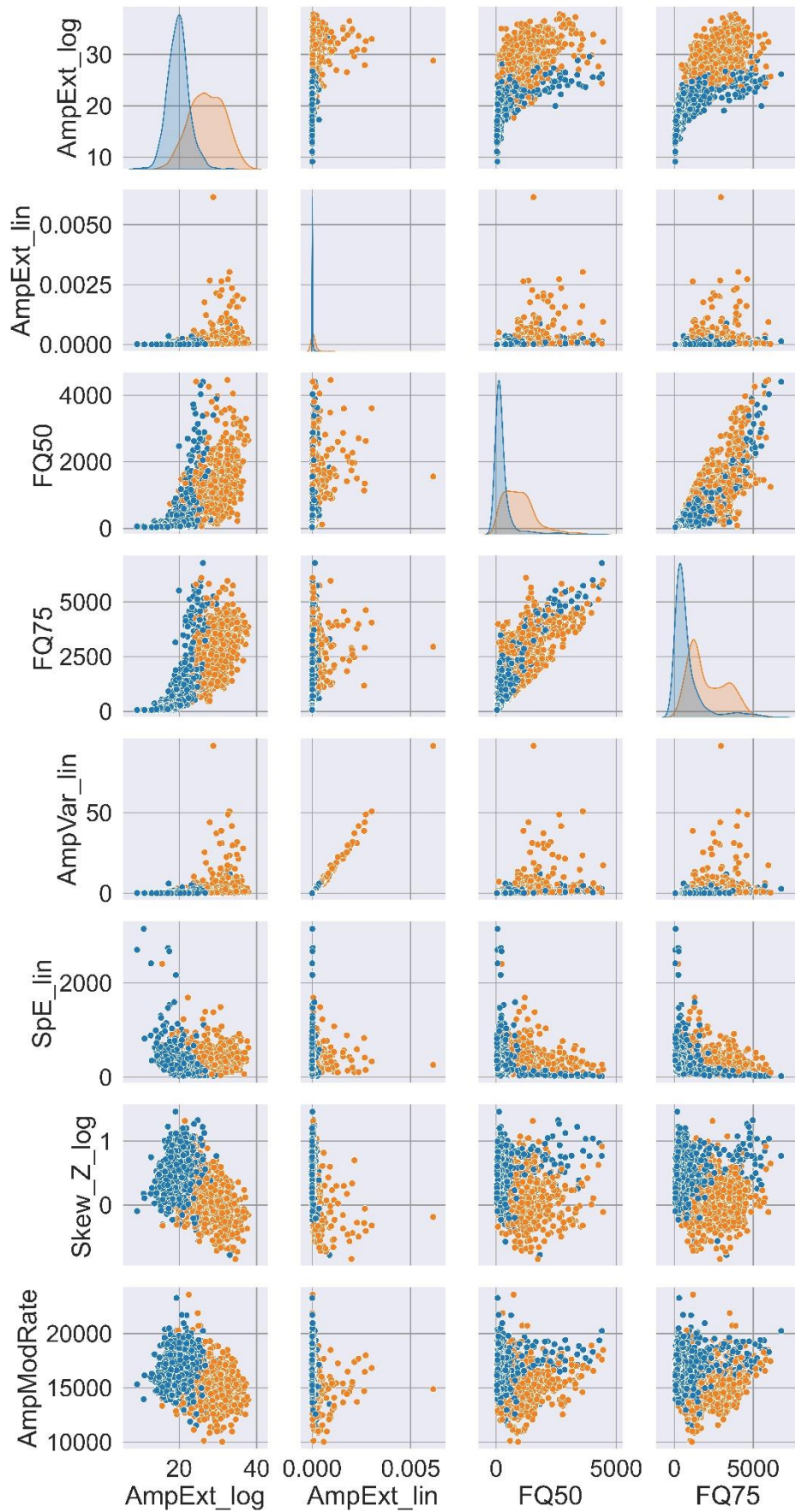


Abbildung 22: Paarweise Auftragung der acht wichtigsten Merkmale. Teil 1.  
 Orange = Klasse A (bedenkliche Laute); Blau = Klasse B (restliche Geräusche)

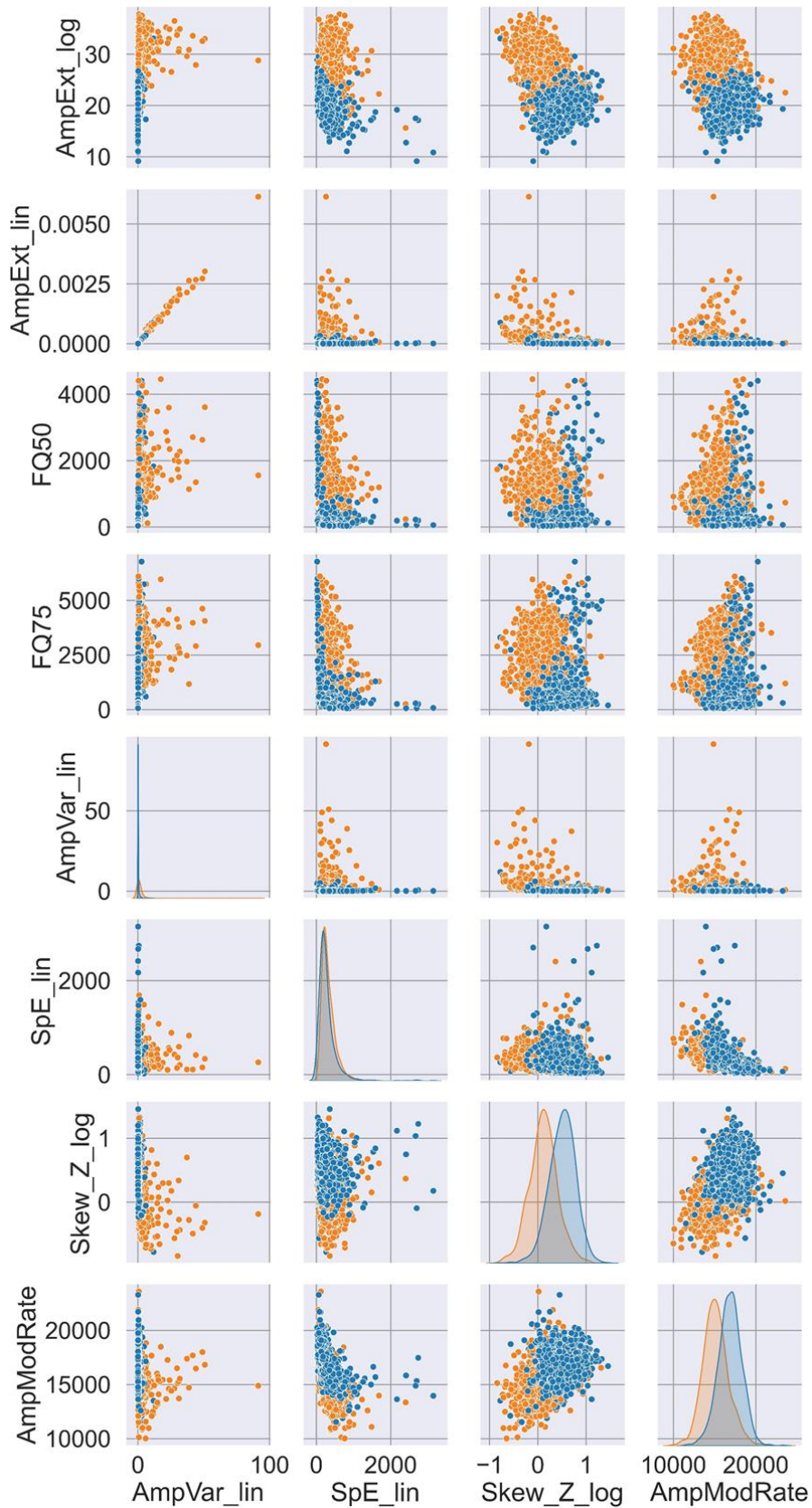


Abbildung 23: Paarweise Auftragung der acht wichtigsten Merkmale. Teil 2.  
 Orange = Klasse A (bedenkliche Laute); Blau = Klasse B (restliche Geräusche)

### Hauptkomponentenanalyse

Mit dem in Tabelle 15 zusammengefassten Parametersatz werden für alle 36 Merkmale eine Hauptkomponentenanalyse (PCA) durchgeführt. Alle 36 Hauptkomponenten werden zur

## Kapitel 5: Ergebnisse und Diskussion

Vorhersage genutzt und die Ergebnisse der Modelle in Tabelle 16 mit den Vorhersageergebnissen der untransformierten Merkmale verglichen.

Tabelle 11: Parametersatz für die Hauptkomponenten-Analyse.

<b>Klasse A</b>	Bedenkliche Laute	<b>Merkmale</b>	alle aus Kapitel 5.1.4
<b>Klasse B</b>	Restliche Geräusche	<b>Skalierung</b>	StdScaler
<b>Fensterbreite</b>	1 Sekunde	<b>Transformation</b>	PCA, alle Komponenten
<b>Anz. Fenster pro Klasse</b>	968	<b>Modell</b>	Zufallswald

Tabelle 12: Ergebnisse des Vergleichs untransformierter gegenüber PCA-transformierter Merkmale.

Transformation	Klasse A: Bedenkliche Laute						Klasse B: Restliche Geräusche					
	Präzision			Sensitivität			Präzision			Sensitivität		
	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.
Keine	92%	90%	94%	90%	88%	95%	90%	88%	95%	92%	90%	94%
PCA	91%	88%	94%	92%	87%	95%	92%	87%	95%	90%	88%	94%

Die Wertebereiche der Ergebnisse der Merkmale ohne Transformation sowie der mit PCA-Transformation berechneten weisen eine große Überlappung auf, wobei die minimalen Ergebniswerte im Fall der untransformierten Merkmale höher als die der transformierten sind.

### Anzahl der Hauptkomponenten

Zur Beurteilung des Potentials der Hauptkomponentenanalyse zur Dimensionsreduktion wurde eine rekursive Merkmals- bzw. in diesem Fall Hauptkomponentenreduktion durchgeführt. Die Ergebnisse sind in Abbildung 24 dargestellt.

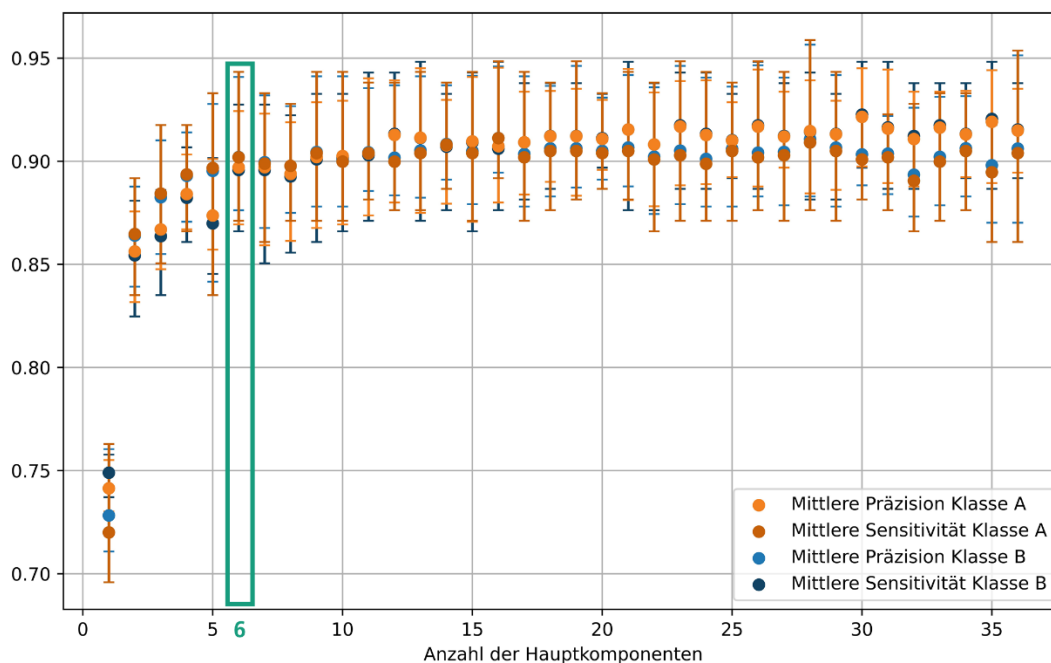


Abbildung 24: Mittelwert, Minima und Maxima der Präzisionen sowie Sensitivitäten der Ergebnisse der Vorhersagen, aufgetragen über die Anzahl der zur Modellerstellung verwendeten Hauptkomponenten.

Ab einer Anzahl von sechs Hauptkomponenten beträgt der kleinste der vier mittleren Werte mindestens 89 % und maximal 91 %. Eine Steigerung des Vorhersageergebnisses wird durch die Reduktion von Hauptkomponenten nicht erreicht.

Identifikation der Hauptkomponenten

Analog zum Abschnitt *Identifikation der acht wichtigsten Merkmale* des Abschnitts zur rekursiven Merkmalseliminierung wurden die in zusammengetragenen wichtigsten Hauptkomponenten bestimmt.

Tabelle 13: Ausgewählte Hauptkomponenten sowie ihre Wahrscheinlichkeit bei einer rekursiven Hauptkomponenteneliminierung auf sechs Hauptkomponenten ausgewählt zu werden

Nr.	1	2	3	4	5	6	7	8	9	10
Hauptkomponente	1	3	12	7	5	31	2	17	19	20
Wahrscheinlichkeit	100%	100%	100%	97%	75%	56%	29%	28%	25%	24%

Nr.	11	12	13	14	15	16	17	18	19	20
Hauptkomponente	16	32	35	30	34	18	6	10	9	33
Wahrscheinlichkeit	15%	14%	11%	10%	7%	3%	2%	2%	1%	1%

Die ersten sechs Hauptkomponenten der Hauptkomponentenanalyse sind nicht identisch mit denen, die die rekursive Komponentenreduktion ausgewählt hat. Der Vollständigkeit halber sind in Tabelle 14 die Ergebnisse für die Merkmale zusammengetragen, die zum einen auf die ersten sechs Hauptkomponenten transformiert wurden und zum anderen auf die, die auf die Hauptkomponenten 1, 3, 12, 7, 5 und 31 transformiert wurden.

Tabelle 14: Ergebnisse Vergleich 36 Merkmale gegenüber auf sechs Hauptkomponenten reduzierte transformierte Merkmale.

Art der Merkmale	Klasse A: Bedenkliche Laute						Klasse B: Restliche Geräusche					
	Präzision			Sensitivität			Präzision			Sensitivität		
	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.
36, untransformiert	92%	90%	94%	90%	88%	95%	90%	88%	95%	92%	90%	94%
6, (1-6), PCA-transformiert	90%	85%	94%	88%	86%	90%	88%	85%	90%	90%	84%	95%
6, (1, 3, 12, 7, 5, 31), PCA-transformiert	90%	87%	92%	89%	85%	94%	89%	85%	94%	90%	87%	93%

Beide Ergebnisse für die mittels Hauptkomponentenanalyse transformierten Merkmale weisen im Mittel schlechtere Vorhersageergebnisse als die 36 untransformierten Merkmale auf.

**Lineare Diskriminanzanalyse**

Analog zur Hauptkomponentenanalyse wird eine lineare Diskriminanzanalyse (LDA) für alle 36 Merkmale durchgeführt. Da diese für den binären Fall nur eine Komponente ergibt, wurde der naive Bayes-Klassifikator für die Vorhersage eingesetzt. Alle zur Vorhersage verwendeten Parameter sind in Tabelle 15 und die Ergebnisse, verglichen mit der der untransformierten Merkmale in Tabelle 16 aufgeführt.

Tabelle 15: Parametersatz für die lineare Diskriminanzanalyse.

<b>Klasse A</b>	Bedenkliche Laute	<b>Merkmale</b>	alle aus Kapitel 5.1.4
<b>Klasse B</b>	Restliche Geräusche	<b>Skalierung</b>	StdScaler
<b>Fensterbreite</b>	1 Sekunde	<b>Transformation</b>	LDA
<b>Anz. Fenster pro Klasse</b>	968	<b>Modell</b>	naiver Bayes-Klassifikator

## Kapitel 5: Ergebnisse und Diskussion

Tabelle 16: Ergebnisse Vergleich untransformierter gegenüber LDA-transformierter Merkmale.

Transformation	Klasse A: Bedenkliche Laute						Klasse B: Restliche Geräusche					
	Präzision			Sensitivität			Präzision			Sensitivität		
	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.	Mittl.	Min.	Max.
Keine	92%	90%	94%	90%	88%	95%	90%	88%	95%	92%	90%	94%
LDA	92%	91%	93%	91%	90%	92%	91%	90%	92%	92%	91%	93%

Die Präzisionen sowie Sensitivitäten der Vorhersagen der LDA-transformierten Merkmale liegen im Bereich von 90 % bis 93 % während die der untransformierten Merkmale eine größere Spanne von 88 % bis hin zu 95 % aufweisen. Beide Vorhersagen sind hinreichend gleich.

Aufgrund der Reduzierung der Merkmale auf eine Komponente durch die lineare Diskriminanzanalyse kann in Abbildung 25 die Trennbarkeit der Klassen in einer zweidimensionalen Grafik gezeigt werden. Dargestellt sind die Werte der linearen Diskriminanzanalyse für eins der fünf Testdatensets sowie der Schwellwert, an dem der naiven Bayes-Klassifikator die Klassen trennt.

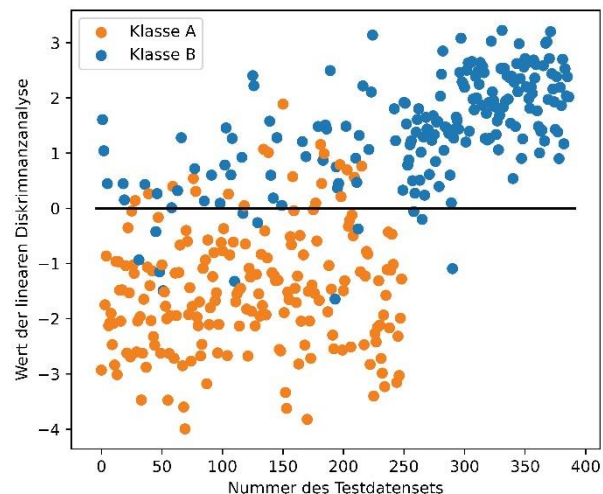


Abbildung 25: Werte der linearen Diskriminanzanalyse sowie Schwellwert des naiven Bayes-Klassifikators für ein Testdatenset.

### Vergleich der Reduktionsmethoden

Die Wertebereiche der Ergebnisse für Präzisionen und Sensitivitäten sind sowohl für die 36 Merkmale als auch für die acht Merkmale sowie die Merkmale der linearen Diskriminanzanalyse hinreichend gleich. Eine Dimensionsreduktion mittels der Hauptkomponentenanalyse erzielt bei den Mittelwerten bis zu 2 % und bei den Extremwerten bis zu 5 % schlechtere Ergebnisse, weswegen diese Reduktionsmethode nicht weiterverwendet wird. Durch die rekursive Merkmalseliminierung kann die Anzahl der zu berechnenden Merkmale auf acht reduziert werden. Bei der linearen Diskriminanzanalyse entfällt aufgrund eines einzigen Schwellwerts des naiven Bayes-Klassifikators die rechenaufwändigere Klassifikation des Zufallswalds. Die letzteren beiden Methoden werden im folgenden Kapitel auf den unbekanntenen kontinuierlichen akustischen Signalen getestet.

### 5.3 ÜBERTRAGUNG DER METHODIK AUF KONTINUIERLICHE AKUSTISCHE SIGNALE

Nach der Entwicklung der Methodik sowie deren Anpassung in den vorangegangenen Kapiteln muss die Methodik an für die beiden Modelle unbekanntem kontinuierlichen akustischen Signalen evaluiert werden. Die daraus abgeleiteten Ergebnisse können zur Verbesserung der Methodik genutzt werden. Dazu wurde im Projekt entwickelter Produktivcode verwendet, der auf den Ergebnissen der vorangegangenen Kapitel basiert. In Tabelle 17 sind die in den vorherigen Kapiteln entwickelten Parameter der Methodik zusammengefasst.

Tabelle 17: Parameter der in dieser Arbeit entwickelten Methodik.

<b>Klasse A</b>	Bedenkliche Laute	<b>Modell 1</b>	Zufallswald, reduzierte Merkmale
<b>Klasse B</b>	Restliche Geräusche	<b>Reduzierte Merkmale</b>	AmpExt_log
<b>Fensterbreite</b>	1 Sekunde		AmpExt_lin
<b>Filter-Frequenzen</b>	20 Hz, 20 kHz		FQ50
<b>Skalierung</b>	StdScaler		FQ75
<b>Transformation</b>	keine		AmpVar_lin
<b>Schwellwert</b>	60 %		SpE_lin
<b>Uneindeutige Zuordnungen</b>			Skew_Z_log
<b>Fensterüberlappung</b>	0,5 Sekunden	<b>Modell 2</b>	AmpModRate
			LDA

Die zur Vorhersage benötigten Merkmale des kontinuierlichen akustischen Signals werden auf Fenstern mit einer Sekunde Länge bestimmt. Für jedes dieser Fenster erfolgt die Filterung, die Extraktion der Merkmale und deren Skalierung sowie Klassifikation. In Abhängigkeit des Schwellwerts wird dieses Fenster dann Klasse A, Klasse B oder Klasse *Uneindeutige Zuordnungen* zugeordnet. Zur Vermeidung von Fehlklassifizierungen aufgrund eines zu geringen Geräuschanteils wird das Fenster mit einer Überlappung von 0,5 s verschoben. Die jeweils eingesetzten Modelle wurden im Unterschied zu den bisherigen Modellen mit allen der 968 Geräusche je Klasse mit einer Länge von einer Sekunde trainiert.

In Abbildung 26 ist als erstes ein Ausschnitt eines unbekanntem kontinuierlichen akustischen Signals, gefolgt von den Labels des Tierarztes und im Anschluss die Vorhersagen beider Modelle dargestellt. Im Gegensatz zu den in dieser Arbeit verwendeten Trainingsdaten wurden die exakten Start- und Endwerte der einzelnen vom Tierarzt gelabelten Geräusche dieses kontinuierlichen Signals nicht ermittelt. Dies kann unter Umständen zu einem Zeitversatz von mehreren Sekunden zwischen Label und tatsächlichem Auftreten des Geräuschs führen.

## Kapitel 5: Ergebnisse und Diskussion

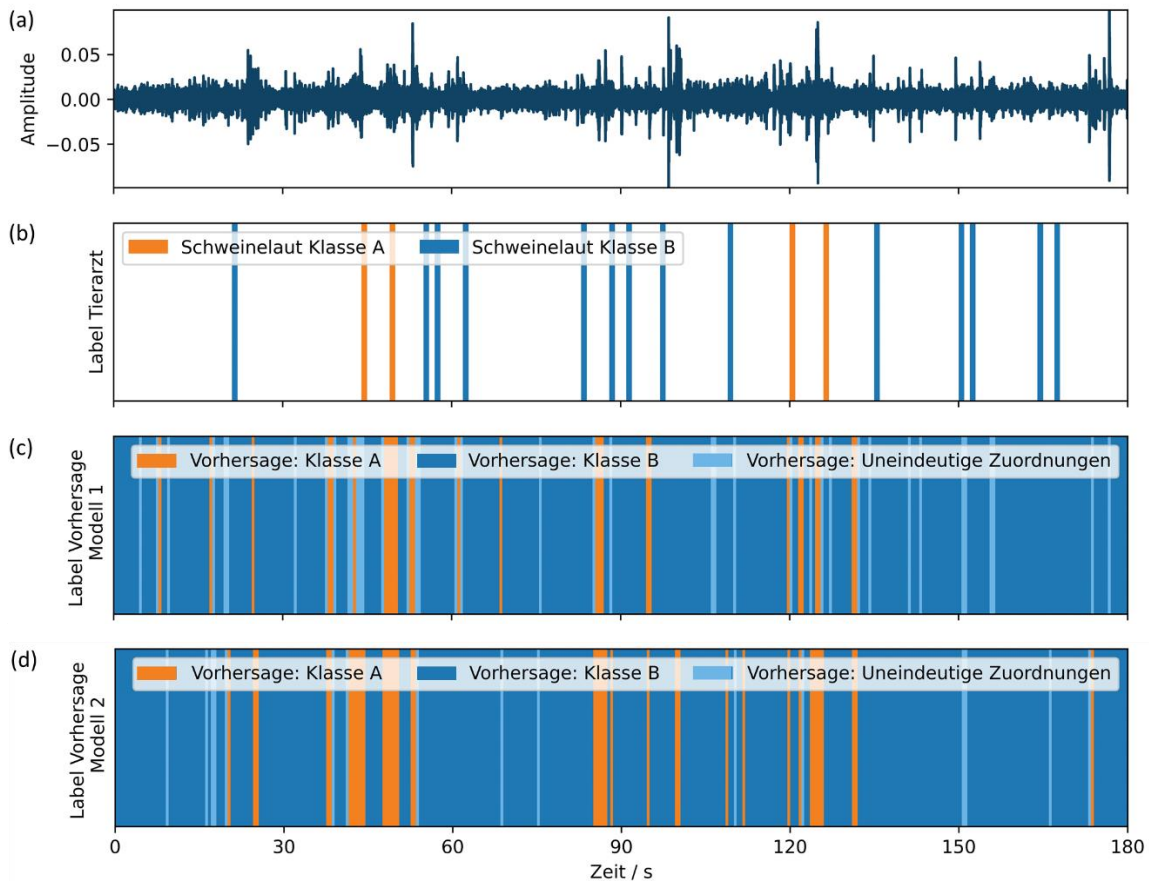


Abbildung 26: (a) unbekanntes kontinuierliches akustisches Signal; (b) Labels von Schweinelauten; (c) vorhergesagte Labels Modell 1; (d) vorhergesagte Labels Modell 2.

Eine Auswertung der Präzisionen sowie Sensitivitäten der vorhergesagten Ereignisse kann aufgrund der fehlenden Startwerte nur qualitativ erfolgen. In den Zeitbereichen, in denen der Tierarzt bedenkliche Laute erkannt hat, werden diese auch mit einer Abweichung von maximal zwei Sekunden von beiden Modellen erkannt. Darüber hinaus werden weitere Fenster mit bedenklichen Lauten vorhergesagt, zu denen keine tierärztliche Einschätzung vorliegt. Aufgrund des hohen Anteils der Vorhersagen der restlichen Geräusche kann über die Zuordnung der Klasse B keine Aussage getroffen werden, da in dieser Klasse sowohl Stallgeräusche als auch andere Schweinelaute enthalten sind. Qualitativ ist kein signifikanter Unterschied zwischen den beiden getesteten Modellen feststellbar.

## 5.4 DISKUSSION

Mit Ausnahme der Fensterlänge wurden in Kapitel 5.1 für alle festzulegenden Parameter der in Abbildung 14 konzipierten Methodik Werte beziehungsweise Verfahren unter Verwendung der Trainingsdaten ausgewählt. In Kapitel 5.2 wurden die Trainingsdaten weiterhin genutzt, eine Fensterlänge festzulegen sowie Optimierungen der Vorhersagen durchzuführen.

### *Aufnahme und Labeln der Daten*

Alle in dieser Arbeit verwendeten Trainingsdaten wurden aus kontinuierlichen Aufnahmen, die mit einem Mikrofon an verschiedenen Stellen im selben Stall aufgenommen wurden, extrahiert. Dabei liegt nicht für jedes aufgenommene Sample ein Label vor, da nur spezifische Geräusche durch den Tierarzt erfasst wurden. Die Ursprungssignale sind daher in drei Kategorien, die zwei Hauptklassen sowie dem Anteil der *uneindeutigen Zuordnungen*, unterteilt.

Die Aufnahmezeit sollte aufgrund des Abtasttheorems mindestens 40 kHz betragen, da Menschen Frequenzen zwischen 20 Hz und 20 kHz hören können und dieser Bereich abgedeckt werden sollte. Handelsübliche Mikrofone im musikalischen Bereich, die in dieser Arbeit verwendet wurden, sind ebenfalls für diesen Frequenzbereich ausgelegt und können höhere Frequenzbereiche nicht abdecken. Da Schweine Frequenzen zwischen 40 Hz und 40 kHz wahrnehmen können, könnten mit einem Mikrofon mit erweitertem Frequenzbereich der Einfluss des größeren Frequenzbereichs auf die Merkmale untersucht werden. (45)

### *Hierarchische Struktur an Geräuschklassen und deren Bewertung*

Während in den Untersuchungen von (37, 38) die Situationen in positiv und negativ unterschieden werden, ohne das einzelne Tierverhalten zu analysieren, erfolgt die Zuordnung in unbedenkliche (positive) sowie bedenkliche (negative) Laute im Projekt *SmartPigHome* in Kombination mit der Beobachtung des Schweineverhaltens. Dabei wurden speziell Laute herausgearbeitet, die die Tiere zur Kommunikation miteinander verwenden, in Abgrenzung zu zufälligen Lauten wie Niesen oder Schnarchen. Gefühlszustände wie *Freude* oder *Angst* wurden bewusst nicht als Kategorie ausgewählt, um eine Überinterpretation des Tierverhaltens im Sinne der Vermenschlichung zu vermeiden.

### *Hauptklassen*

Aufgrund der unzureichenden Anzahl der Geräusche einiger Unterklassen der Ebene 5 der Abbildung 15 ist nur eine Unterscheidung zwischen größeren Klassen höherer Ebenen möglich, weswegen sich in dieser Arbeit auch nur auf zwei Hauptklassen beschränkt wurde. Prinzipiell ist die in dieser Arbeit erstellte Methodik übertragbar auf andere sowie mehr zu unterscheidende Geräuschklassen. Beispielsweise könnte das Erkennen gesundheitsrelevanter Laute wie Husten und Niesen für den Landwirt von Interesse sein, um deren Ursache frühzeitig auf den Grund gehen zu können. Dazu ist eine ausreichend große Datenbasis notwendig. Das Labeln der Daten erfordert noch einen großen Zeitaufwand, da zum Training eines Modells möglichst präzise Start- sowie Endzeitpunkte des Geräuschs herausgearbeitet werden müssen. Durch die in dieser Arbeit entwickelte Algorithmik ist eine schnellere Erfassung potentieller Trainingsdaten zum Nachtrainieren denkbar, da die unbekannteren akustischen Signale bereits in die zwei Hauptklassen unterteilt werden.

Durch die Berücksichtigung der Vorhersagewahrscheinlichkeiten der zwei Hauptklassen, wurde in Abschnitt 5.2.3 eine weitere Klasse, die der *Uneindeutigen Zuordnungen*, eingeführt. Wie in diesem Abschnitt gezeigt, können bei einer geringen Vorhersagewahrscheinlichkeit von unter 60 % häufiger Falschklassifizierungen auftreten. Die Wahl der Schwelle, bis zu welcher Vorhersagewahrscheinlichkeit die Fenster der dritten Kategorie zugeordnet werden, wurde in diesem Fall unter Berücksichtigung aller Falschklassifizierungen einer fünffachen Aufteilung getroffen. Je höher dieser Schwellwert liegt, desto weniger Falschklassifizierungen treten auf und

## Kapitel 5: Ergebnisse und Diskussion

die Werte für Präzision und Sensitivität verbessern sich. Jedoch sollte dieser Anteil nicht unabhängig von dem Anteil der Geräusche betrachtet werden, die aufgrund des Schwellwerts nicht klassifiziert werden. Für die Anwendung wäre ein möglichst kleiner Anteil der dritten Kategorie erstrebenswert, der durch das Nachtrainieren des ursprünglichen Modells oder auch durch das Entwickeln weiterer Merkmale umgesetzt werden könnte.

### *Ausgewogenes Datenset*

Zum Trainieren eines Modells, das möglichst gut beide Klassen vorhersagen soll, ist ein ausgewogenes Datenset notwendig. Bei einem unausgewogenen Datenset würde die überwiegende Klasse stärker trainiert werden und die kleinere Klasse mit einer höheren Wahrscheinlichkeit im Test schlechter abschneiden. Da eine Überabtastung durch exaktes Kopieren der Werte der kleineren Klasse zu unerwünschtem Overfitting führen kann (46), kann der SMOTE-Algorithmus verwendet werden. Dieser weist jedoch auch Nachteile wie das Auftreten von Überschneidungen von Samples sowie Anfälligkeiten für Rauschen auf. (47) Daher wurde die Methode der Unterabtastung gewählt. Zum Nachtrainieren des Modells für die Praxis wäre eine Überabtastung in einem nächsten Schritt denkbar.

### *Fensterung*

Beim Fenster der relevanten Geräusche aus dem kontinuierlichen Signal wurde sich, konträr zum Paper von Briefer et al. (38), gegen eine exakte Fensterung der Geräusche entschieden, da in der Praxis ein Puffer, der der Fensterung entspricht, vorgesehen ist. Aufgrund der unterschiedlichen Lautstärken sowie Geräuschlängen der Geräusche beider Hauptklassen stellt sich das automatisierte Ermitteln der exakten Start- und Endzeitpunkte als Herausforderung dar, die in dieser Arbeit nicht fokussiert wird, da der erwartende Rechenaufwand als hoch und der Gewinn in der automatischen Klassifizierung als gering einzuschätzen ist.

Die Startwerte des ersten Fensters der Geräuschfensterung wurden bei den Trainingsdaten zufällig gewählt, da der Puffer der kontinuierlichen Verarbeitung auch die Geräusche zufällig unterteilt. Allerdings werden im Training Geräusche, die kürzer als die Fensterlänge sind, nicht zufällig halbiert, sondern im Ganzen für das Training verwendet. Da nicht das gesamte kontinuierliche Signal zum Training gelabelt wurde, können keine Rückschlüsse auf die Geräuschklasse der Bereiche des Fensters gezogen werden, die kein gelabeltes Geräusch enthalten. In den meisten Fällen kann davon ausgegangen werden, dass die jeweils andere Klasse im ungelabelten Fensterbereich vorliegt. Aufgrund der großen Anzahl an Tieren im Stall können auch zwei Geräusche derselben Klasse sehr schnell aufeinander folgen, von denen nur eines gelabelt ist. Dies könnte einer der Gründe sein, dass in Abschnitt 5.2.2 die Fenster mit einem geringeren Geräuschanteil prozentual häufiger falsch klassifiziert wurden, da dort die andere Geräuschklasse überwiegt. In der späteren Anwendung bietet sich daher ein Überlapp der Fenster an, um präziser Geräusche, die durch die Fensterung ungünstig geteilt werden, dennoch richtig zuordnen zu können.

Die Fensterlänge wurde in Abschnitt 5.2.1 auf eine Sekunde festgelegt, da für diese Fensterlänge bei den gegebenen Daten das Modell mit den besten Ergebnissen für Präzision sowie Sensitivität trainiert wurde. Alle über fünffache Aufteilung ermittelten Präzisionen und Sensitivitäten überlappen sich für die angrenzenden Fensterlängen. Daher wird in der späteren Anwendung das Modell mit der Fensterlänge auszuwählen sein, das die besten auf den kontinuierlichen Daten ermittelten Ergebnisse liefert. Da die Trainingsdaten die Realität aufgrund unzureichender Mengen aller möglichen Geräusche nur bedingt abbilden, kann auch ein Modell für eine von einer Sekunde abweichenden Fensterlänge die besten Ergebnisse liefern.

### *Vorverarbeitung der akustischen Daten*

Aufgrund der niederfrequenten Störgeräusche, verursacht durch elektrisches Rauschen bei 50 Hz sowie beispielsweise durch die Lüftung des Stalls, war in der Entwicklung der Methodik ein Hochpass vorgesehen. Zur Vermeidung einer Beeinflussung der verwendeten, frequenzabhängigen Merkmale wurde dieser Ansatz zurückgestellt und zunächst die untere Auflösungsgrenze des Mikrofons herangezogen. Da auch auf dem logarithmierten Spektrum Merkmale bestimmt wurde, war eine obere Grenzfrequenz notwendig, die sich ebenfalls an der Grenzfrequenz des Mikrofons orientiert. Diese Implementierung der Pässe wird für die nach der Arbeit geplanten Tests durch einen Bandpass von 20 Hz bis 20 kHz ersetzt werden. Auf eine Verkleinerung des Frequenzbereichs sowie eine Änderung der Normierung wurde zugunsten einer übersichtlichen Versuchs-Matrix verzichtet. Zur Verbesserung der Merkmale ist dennoch eine gezieltere Filterung denkbar, damit eine sicherere Trennung der Merkmale mittels des Klassifikators erfolgen kann.

### *Merkmale*

In dieser Arbeit wurden 36 Merkmale verwendet, von denen 14 sowohl auf einer linearen als auch auf einer logarithmischen Skala berechnet wurden. Da das zu entwickelnde Modell die Geräusch-Differenzierung eines Menschen nachbilden soll und das menschliche Gehör sowohl bei Tonhöhen als auch Lautstärke logarithmisch differenziert, bietet sich die logarithmische Skalierung für die Merkmalsextraktion an. Einige der Merkmale weisen wie in Abschnitt 5.2.4 eine lineare Korrelation auf. Eine Dekorrelation ist für den Zufallswald aufgrund seiner Eigenschaften nicht zwingend notwendig, weswegen die Ergebnisse, die mit transformierten Merkmalen ermittelt wurden, entweder genauso gut oder etwas schwächer ausgefallen sind.

Da die Auswahl der 36 Merkmale zu Beginn empirisch erfolgte, wurden mehrere Methoden zur Dimensionsreduktion getestet. Hinsichtlich der Ergebnisse von Präzision und Sensitivität kann die Anzahl der verwendeten Merkmale mittels der rekursiven Merkmalseliminierung auf acht Merkmale reduziert werden, da sich die Ergebnisse bei einer höheren Anzahl an Merkmalen nicht mehr signifikant verbessern. Von diesen acht Merkmalen werden drei im Frequenz- und fünf im Zeitbereich berechnet. Die Frequenzquantile sind ein Maß für die Tonhöhe des Geräuschs und die spektrale Ebenheit ein Maß für das Auftreten einer spezifischen Frequenz bzw. eine Gleichverteilung der Frequenzen. Die Merkmale des Amplituden-Modulations-Ausmaßes, die Amplituden-Modulation sowie der Amplituden-Variation sind ein Maß für die Lautstärkenänderung innerhalb eines Fensters. Die Schiefe des Signals im Zeitbereich ist ebenso ein Maß für den zeitlichen Verlauf der Lautstärke.

Die lineare Diskriminanzanalyse weist genauso gute Ergebnisse auf wie die 36 untransformierten Merkmale. Da aufgrund der binären Klassifikation alle Merkmale auf eine Komponente reduziert werden können, ist nur der mittels der naiven Bayes-Klassifikation ermittelte Schwellwert auf unbekanntem Daten notwendig. Auf eine aufwendigere Klassifikation wie dem Zufallswald kann in diesem Fall verzichtet werden.

Des Weiteren wurde die Hauptkomponentenanalyse zur Dimensionsreduktion analog zur rekursiven Merkmalseliminierung getestet. Die Ergebnisse für alle 36 Hauptkomponenten war nahezu identisch zu denen der 36 untransformierten Merkmale. Da die sechs Komponenten, die mittels der rekursiven Komponenteneeliminierung ausgewählt wurden, wider Erwarten nicht den sechs ersten Hauptkomponenten entsprachen, wurden mit beiden Arten der sechs Hauptkomponenten Vorhersagen durchgeführt. Beide waren im Mittelwert um bis zu 2 % und in den Extremwerten bis zu 5 % schlechter als die untransformierten Merkmale. Da die beiden voran genannten Transformationen im Vergleich die besseren Ergebnisse erzielt haben, wurde die Hauptkomponentenanalyse nicht mehr auf unbekanntem kontinuierlichen akustischen Signalen getestet.

### *Auswahl und Bewertung der Modelle*

Trotz der großen Vielfalt an existierenden Klassifikatoren wurde sich in dieser auf überwachten Lernalgorithmen fokussierten Arbeit auf den Zufallswald sowie eine Bayes-Klassifikation der zwei Komponenten der linearen Diskriminanzanalyse beschränkt. Die Zuordnung zu unterscheidenden Klassen erfolgte durch einen Menschen, der während des Labels Entscheidungen ähnlich einem Entscheidungsbaum getroffen hat. Daher wurde der Zufallswald als Ensemble-Methode des Entscheidungsbaums im ersten Schritt als Klassifikator ausgewählt. Durch den Einsatz einer anderen Ensemble-Methode könnten die Ergebnisse der einzelnen Bäume anders gewichtet und somit eine Verbesserung der Ergebnisse erzielt werden. Aufgrund des zeitlich begrenzten Rahmens wurde in dieser Arbeit darauf verzichtet, weitere Entscheidungsbaum-basierten Klassifikatoren zu testen und diese zu tunen. Durch den Einsatz der linearen Diskriminanzanalyse der zur Verfügung stehenden Merkmale kann zur Klassifikation der verbleibenden zwei Komponenten der Bayes-Klassifikation, der weniger komplex als der Zufallswald ist, eingesetzt werden.

Die Bewertung der jeweiligen Methodiken erfolgte anhand der Ergebnisse für Präzisionen sowie Sensitivitäten beider Klassen. Diese Werte wurden ausgewählt, da sie zum einen die Verteilung der Klassen innerhalb der vorhergesagten Klasse und zum anderen die Verteilung der Klassen innerhalb der wahren Klasse widerspiegeln. Je höher die Sensitivität ist, desto mehr Geräusche der dazugehörigen Klasse wurden richtig zugeordnet. Bei einem unausgewogenen Datenset spielt dieser Wert eine wichtigere Rolle als die Präzision. Da das vorliegende Datenset ausgewogen ist, sollten sowohl die Werte für Präzision als auch für Sensitivität möglichst hoch sein sowie im idealen Fall denselben Wert aufweisen. Alle durchgeführten Vorhersagen weisen eine maximale Differenz von 3% innerhalb der vier Werte für Präzision und Sensitivität auf. Beim Vergleich der Vorhersageergebnisse der beiden Methodiken, die auf unbekannte kontinuierliche akustische Signale angewendet wurden, liegen signifikante Überlappungen der Werte vor, weswegen keine Methodik eindeutig zu bevorzugen ist.

### *Übertragung der Methodik auf unbekannte kontinuierliche akustische Signale*

Qualitativ kann am Beispiel in Abbildung 26 gezeigt werden, dass eine Übertragung der Methodik zur Vorhersage der beiden Geräuschklassen auf unbekannte kontinuierliche akustische Signale möglich ist. Aufgrund der fehlenden präzisen Start- und Endwerte der Geräusche können die Anteile der Richtig- und Falschklassifizierungen nicht wie bei den vorherigen Untersuchungen bestimmt werden. Da dies Teil der nächsten Schritte ist, wird diese Diskussion im Kapitel Zusammenfassung und Ausblick weitergeführt.

Dieses kontinuierliche akustische Signal weist weniger bedenkliche Laute als andere Geräusche in der Vorhersage auf, was zu erwarten war. Ab welcher Quantität an bedenklichen Lauten die Gesamtsituation als alarmierend im Sinne eines Assistenzsystems einzustufen ist, muss zu einem späteren Zeitpunkt des Projekts in Kooperation mit den Tierärzten erarbeitet werden.

## 6 ZUSAMMENFASSUNG UND AUSBLICK

---

In dieser Arbeit wurde gezeigt, dass eine Trennung von rauschbehafteten Signalen eines Schweinestalls in den Klassen *Bedenkliche Laute* und *Restliche Geräusche* sowohl mittels eines Zufallswalds als auch mittels einer naiven Bayes-Klassifikation in Kombination mit einer linearen Diskriminanzanalyse möglich ist. Für jedes der 968 akustischen Signalen je Klasse mit einer Länge von einer Sekunde wurden 36 Merkmale extrahiert, mit denen fünf Modelle im Sinne der stratifizierten Kreuzvalidierung mit dem Klassifikator des Zufallswalds erstellt wurden. Die Tests dieser Modelle erzielten eine mittlere Präzision sowie eine mittlere Sensitivität von 91 %. Mittels der rekursiven Merkmalseliminierung konnte die Anzahl der benötigten Merkmale auf acht reduziert werden, ohne eine Verschlechterung der Ergebnisse hervorzurufen. Dieselben Ergebnisse konnten durch eine lineare Diskriminanzanalyse der 36 Merkmale und einer naiven Bayes-Klassifikation erzielt werden. Der Falschklassifizierungsanteil der Vorhersagen kann unter Berücksichtigung der Vorhersagewahrscheinlichkeiten reduziert werden, indem alle unterhalb 60 % zu *uneindeutigen Zuordnungen* gezählt werden.

Des Weiteren wurde gezeigt, dass die Methodik und die auf den Trainingsdaten erstellten Modelle auch auf kontinuierliche akustische Signale übertragen werden konnten. Auch hier ließ sich qualitativ kein Unterschied zwischen beiden Reduktionsmethoden feststellen. Zur Quantifizierung der Vorhersagegenauigkeit müssen im Nachgang dieser Arbeit die gelabelten Bereiche des akustischen Signals präzise herausgearbeitet werden. Unter Umständen erreicht eine Reduktionstechnik dann signifikant bessere Ergebnisse. Aufbauend auf die Vorhersagegenauigkeit können dann gezielt Fenster zum Nach-Labeln ausgewählt werden, um den Anteil der Falschklassifizierungen und den der uneindeutigen Zuordnungen zu verringern.

Im Rahmen des Projekts sollte das nachtrainierte Modell dann auf die akustischen Signale der stationären Mikrofone desselben Maststalls angewendet werden. Jedoch können jegliche Änderung der Randbedingungen, wie zum Beispiel eine veränderte Hard- oder Software, eine abweichende Messposition bis hin zu einem anderen Stall inklusive der individuellen Lüftungs-, Futter- sowie Trinkanlagen, einen Einfluss auf die Merkmale haben und somit die Vorhersageergebnisse des Zufallswalds erheblich verschlechtern, da dieser aufgrund seiner starren Grenzwerte jeder Entscheidung Veränderungen der Werte nicht abfangen kann. Daher sollten die Randbedingungen schrittweise variiert und ihre Einflüsse auf die Merkmale sowie das Vorhersageergebnis analysiert werden. Unter Umständen muss für jedes Aufnahmeszenario ein neues Modell trainiert werden. Zudem ist eine Anpassung der Fensterbreite, der Merkmale, des Klassifikators sowie des Schwellwerts der uneindeutigen Zuordnungen in Abhängigkeit des Aufnahmeszenarios denkbar.

Durch die Erhebung weiterer Trainingsdaten kann mit der in dieser Arbeit beschriebenen Vorgehensweise eine Methodik entwickelt werden, die weitere Schweinelaute erkennt. Aufgrund der relativ geringen Anzahl der Geräusche in mancher der untersten Klassen ist dies mit der dieser Arbeit zugrundeliegenden Datenbasis noch nicht möglich. Beispielsweise ist für den Landwirt von Interesse, über Anzeichen von Krankheiten frühzeitig informiert zu werden. Eine Ausweitung auf die Laute Niesen und Husten wäre somit denkbar. Ein System, das die Huster von Schweinen überwacht, ist bereits unter dem Namen Soundtalks der Firma Boehringer-Ingelheim erwerbbar. Das sichere Erkennen und Unterscheiden mehrerer Klassen wäre somit von Interesse und zurzeit nicht kommerziell verfügbar.

Die in dieser Arbeit entwickelten Vorgehensweise zur Erstellung einer Methodik ist nicht auf den Schweinestall begrenzt und kann ebenso auf andere Systeme, die akustisch überwacht werden sollen, angewendet werden.

# 7 LITERATUR

---

1. Arora A, Gupta R. A Comparative Study on Application of Artificial Intelligence for Quality Assurance in Manufacturing. In: 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA): IEEE; 2022. S. 1200–6.
2. Wieme J, Mollazade K, Malounas I, Zude-Sasse M, Zhao M, Gowen A et al. Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review. *Biosystems Engineering* 2022; 222:156–76. doi: 10.1016/j.biosystemseng.2022.07.013.
3. H.V. Ravindra, Y.G. Srinivasa, R. Krishnamurthy. Acoustic emission for tool condition monitoring in metal cutting. *Wear* 1997; 212(1):78–84. Verfügbar unter: <https://www.sciencedirect.com/science/article/pii/S0043164897001373>.
4. Meffert B, Hochmuth O. *Werkzeuge der Signalverarbeitung: Grundlagen, Anwendungsbeispiele, Übungsaufgaben*. München, Boston, Harlow: Pearson Studium; 2004. (i - Informatik technische Informatik).
5. Papula L. *Mathematik für Ingenieure und Naturwissenschaftler*. 13., durchges. Aufl. Erscheinungsort nicht ermittelbar; 2012. (Mathematik für Ingenieure und Naturwissenschaftler).
6. Lerch R, Sessler G, Wolf D. *Technische Akustik*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009.
7. Müller G, Hrsg. *Taschenbuch der technischen Akustik: Mit 119 Tabellen*. 3., erw. und überarb. Aufl. Berlin, Heidelberg: Springer; 2004. (engineering online library).
8. Henze N. *Stochastik für Einsteiger: Eine Einführung in die faszinierende Welt des Zufalls*. 9., erweiterte Auflage 2012. Wiesbaden: Vieweg+Teubner Verlag / GWV Fachverlage GmbH Wiesbaden; 2011. (Für Einsteiger).
9. Murphy KP. *Machine learning: A probabilistic perspective*. 4. print. (fixed many typos). Cambridge, Mass.: MIT Press; 2013. (Adaptive computation and machine learning series). Verfügbar unter: <http://www.cs.ubc.ca/%7Emurphyk/MLbook/>.
10. Hastie T, Tibshirani R, Friedman JH. *The elements of statistical learning: Data mining, inference, and prediction*. Second edition, corrected at 12th printing 2017. New York, NY: Springer; 2017. (Springer series in statistics).
11. Kubat M. *An Introduction to Machine Learning*. 3rd ed. 2021. Cham: Springer International Publishing; Imprint Springer; 2021. (Springer eBook Collection).
12. Saarela M, Jauhiainen S. Comparison of feature importance measures as explanations for classification models. *SN Appl. Sci.* 2021; 3(2). doi: 10.1007/s42452-021-04148-9.
13. Krüger F. *Activity, Context, and Plan Recognition with Computational Causal Behaviour Models [Dissertation]*. Rostock: Universität Rostock; 2016.
14. He H, Garcia EA. Learning from Imbalanced Data. *IEEE Trans. Knowl. Data Eng.* 2009; 21(9):1263–84. doi: 10.1109/TKDE.2008.239.
15. Chen X, Li B. Acoustic emission method for tool condition monitoring based on wavelet analysis. *The International Journal of Advanced Manufacturing Technology* 2007.
16. Li R, He D. Rotational Machine Health Monitoring and Fault Detection Using EMD-Based Acoustic Emission Feature Quantification. *IEEE Transactions on Instrumentation and Measurement* 2012; 61(4):990–1001. doi: 10.1109/TIM.2011.2179819.
17. Tagawa Y, Maskeliūnas R, Damaševičius R. Acoustic Anomaly Detection of Mechanical Failures in Noisy Real-Life Factory Environments. *Electronics* 2021; 10(19). Verfügbar unter: <https://www.mdpi.com/2079-9292/10/19/2329>.
18. Robert Müller, Fabian Ritz, Steffen Illium, Claudia Linnhoff-Popien. Acoustic Anomaly Detection for Machine Sounds based on Image Transfer Learning. In: *Proceedings of the 13th International Conference on Agents and Artificial Intelligence: SCITEPRESS - Science and Technology Publications*; 2021.
19. Becker P, Roth C, Roennau A, Dillmann R. Acoustic Anomaly Detection in Additive Manufacturing with Long Short-Term Memory Neural Networks. In: *2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA)*; 2020. S. 921–6.
20. Herman I, Impett L, Wollner PKA, Blackwell AF. Augmenting Bioacoustic Cognition with Tangible User Interfaces. In: *Foundations of Augmented Cognition*. Cham: Springer International Publishing; 2015. S. 437–48.
21. Rory Gibb, Ella Browning, Paul Glover-Kapfer, Kate E. Jones, Luca Börger. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution* 2018; 10(2):169–85. doi: 10.1111/2041-210x.13101.

22. Johan Bjorck, Brendan H. Rappazzo, Di Chen, Richard Bernstein, Peter H. Wrege, Carla P. Gomes. Automatic Detection and Compression for Passive Acoustic Monitoring of the African Forest Elephant; 2019.
23. Len Thomas, Armando Jaramillo-Legorreta, Gustavo Cardenas-Hinojosa, Edwyna Nieto-Garcia, Lorenzo Rojas-Bracho, Jay M. Ver Hoef et al. Last call: Passive acoustic monitoring shows continued rapid decline of critically endangered vaquita. *The Journal of the Acoustical Society of America* 2017; 142(5):EL512-EL517. doi: 10.1121/1.5011673.
24. Grauvogl A. Über das Verhalten des Hausschweines unter besonderer Berücksichtigung des Fortpflanzungsverhaltens [Dissertation]; 1958.
25. SCHRADER L, Hammerschmidt K. COMPUTER-AIDED ANALYSIS OF ACOUSTIC PARAMETERS IN ANIMAL VOCALISATIONS: A MULTI-PARAMETRIC APPROACH. *Bioacoustics* 1997; 7(4):247–65. doi: 10.1080/09524622.1997.9753338.
26. Peter-Christian Schön, Birger Puppe, Tatjana Gromyko and Gerhard. Common Features and Individual Differences in Nurse Grunting of Domestic Pigs (*Sus scrofa*): A Multi-Parametric Analysis. *Behaviour* 1999; (Vol. 136, No. 1 (Jan., 1999)):49–66.
27. Schön PC, Puppe B, Manteuffel G. Linear prediction coding analysis and self-organizing feature map as tools to classify stress calls of domestic pigs (*Sus scrofa*). *The Journal of the Acoustical Society of America* 2001; 110(3 Pt 1):1425–31. doi: 10.1121/1.1388003.
28. Manteuffel G, Schön PC. STREMODODO, ein innovatives Verfahren zur kontinuierlichen Erfassung der Stressbelastung von Schweinen bei Haltung und Transport. *Arch. Anim. Breed.* 2004; 47(2):173–81. doi: 10.5194/aab-47-173-2004.
29. Puppe B, Schön PC, Tuchscherer A, Manteuffel G. Castration-induced vocalisation in domestic piglets, *Sus scrofa*: Complex and specific alterations of the vocal quality. *Applied Animal Behaviour Science* 2005; 95(1-2):67–78. doi: 10.1016/j.applanim.2005.05.001.
30. Leidig MS, Hertrampf B, Failing K, Schumann A, Reiner G. Pain and discomfort in male piglets during surgical castration with and without local anaesthesia as determined by vocalisation and defence behaviour. *Applied Animal Behaviour Science* 2009; 116(2-4):174–8. doi: 10.1016/j.applanim.2008.10.004.
31. Vandermeulen J, Bahr C, Tullo E, Fontana I, Ott S, Kashiha M et al. Discerning pig screams in production environments. *PLoS One* 2015; 10(4):e0123111. doi: 10.1371/journal.pone.0123111.
32. Wirthgen E, Kunze M, Goumon S, Walz C, Höflich C, Spitschak M et al. Interference of stress with the somatotropic axis in pigs - lights on new biomarkers. *Sci Rep* 2017; 7(1):12055. doi: 10.1038/s41598-017-11521-5.
33. Marx G, Horn T, Thielebein J, Knubel B, Borell E von. Analysis of pain-related vocalization in young pigs. *Journal of Sound and Vibration* 2003; 266(3):687–98. doi: 10.1016/S0022-460X(03)00594-7.
34. Garcia M, Gingras B, Bowling DL, Herbst CT, Boeckle M, Locatelli Y et al. Structural Classification of Wild Boar (*Sus scrofa*) Vocalizations. *Ethology* 2016; 122(4):329–42. doi: 10.1111/eth.12472.
35. Tallet C, Linhart P, Policht R, Hammerschmidt K, Šimeček P, Kratinova P et al. Encoding of situations in the vocal repertoire of piglets (*Sus scrofa*): a comparison of discrete and graded classifications. *PLoS One* 2013; 8(8):e71841. doi: 10.1371/journal.pone.0071841.
36. Friel M, Kunc HP, Griffin K, Asher L, Collins LM. Positive and negative contexts predict duration of pig vocalisations. *Sci Rep* 2019; 9(1):2062. doi: 10.1038/s41598-019-38514-w.
37. Briefer EF, Vizier E, Gygax L, Hillmann E. Expression of emotional valence in pig closed-mouth grunts: Involvement of both source- and filter-related parameters. *The Journal of the Acoustical Society of America* 2019; 145(5):2895. doi: 10.1121/1.5100612.
38. Briefer EF, Sypherd CC-R, Linhart P, Leliveld LMC, La Padilla de Torre M, Read ER et al. Classification of pig calls produced from birth to slaughter according to their emotional valence and context of production. *Sci Rep* 2022; 12(1):3409. doi: 10.1038/s41598-022-07174-8.
39. Charlton BD, Zhihe Z, Snyder RJ. Vocal cues to identity and relatedness in giant pandas (*Ailuropoda melanoleuca*). *The Journal of the Acoustical Society of America* 2009; 126(5):2721–32. doi: 10.1121/1.3224720.
40. Paul Boersma and David Weenink. Praat Manual; 16.08.23, 15:30. Verfügbar unter: <https://www.fon.hum.uva.nl/praat/manual/Harmonicity.html>.
41. Gabriël J.L. Beckers. Praat scripts [Stand: 16.08.23, 15:30]. Verfügbar unter: <https://gbeckers.nl/pages/phonetics.html>.

## Kapitel 7: Literatur

42. Audacity. Audacity [Stand: 01.03.23; 17:00]. Verfügbar unter: [https://www.hamm.de/fileadmin/user\\_upload/Medienarchiv\\_neu/Bilder/Grafik/Logo/Sonstige/Audacity\\_Logo\\_With\\_Name.png](https://www.hamm.de/fileadmin/user_upload/Medienarchiv_neu/Bilder/Grafik/Logo/Sonstige/Audacity_Logo_With_Name.png).
43. Behringer. Behringer EMC8000 [Stand: 01.03.23; 17:00]. Verfügbar unter: [https://live.staticflickr.com/3338/3642549573\\_f9bfe89abf\\_z.jpg](https://live.staticflickr.com/3338/3642549573_f9bfe89abf_z.jpg).
44. Behringer. Behringer U-Phoria UMC202HD. Verfügbar unter: <https://www.atprosound.com/wp-content/uploads/2021/03/8f94a32876820667014c9eb9517f80b4.jpeg> [Stand: 01.03.23; 14:30].
45. Heffner HE. Auditory awareness. *Applied Animal Behaviour Science* 1998; 57(3-4):259–68. doi: 10.1016/S0168-1591(98)00101-4.
46. Zhou Y, Lentz E, Michelson H, Kim C, Baylis K. Machine learning for food security: Principles for transparency and usability. *Applied Eco Perspectives Pol* 2022; 44(2):893–910. doi: 10.1002/aep.13214.
47. Jiang Z, Pan T, Zhang C, Yang J. A New Oversampling Method Based on the Classification Contribution Degree. *Symmetry* 2021; 13(2):194. doi: 10.3390/sym13020194.

## 8 ABBILDUNGSVERZEICHNIS

Abbildung 1: Beispiel für ein kontinuierliches Zeitsignal. _____	3
Abbildung 2: Beispiel für ein diskretes Zeitsignal. _____	4
Abbildung 3: Schalldruckverhältnisse (schematisiert) in der Umgebung eines akustischen Strahlers; in Abwandlung von (6), Seite 495. _____	5
Abbildung 4: Schematische Darstellung einer Schallquelle Q in einer Raumecke und ihren Spiegelschallquellen Q' und Q"; entnommen aus (7), Seite 334. _____	6
Abbildung 5: Schematische Darstellung der Auswahl an Test- und Trainingsdaten für eine stratifizierte Kreuzvalidierung für k=4 Faltungen bzw. Modelle. _____	8
Abbildung 6: Darstellung der Achse mit der größten Varianz (largest principal component) sowie die mit der kleinsten Varianz (smallest principal component); entnommen aus (10), Seite 67. _____	9
Abbildung 7: Darstellung des Mittelwerts (Kreuz) sowie der Kovarianz (Ellipse) für jede Klasse sowie die paarweisen Entscheidungsachsen der linearen Diskriminanzanalyse; entnommen aus (10), Seite 109. _____	9
Abbildung 8: Auftragung der Entropie, des Gini-Indexes sowie der Falschklassifizierungsrate über die Wahrscheinlichkeit für eine binäre Klassifikation zur Veranschaulichung des Informationsgewinns; entnommen aus (10), Seite 309. _____	10
Abbildung 9: Schematische Darstellung einer binären Wahrheitsmatrix inklusive Bezeichnungen der Vorhersagen: <u>Links</u> für Klasse A; <u>Rechts</u> für Klasse B. _____	11
Abbildung 10: Zeitsignal eines Schweinelauts sowie ein Ausschnitt daraus inklusive Beschriftung von charakteristischen Eigenschaften, die zur Bestimmung der Merkmale relevant sind. _____	16
Abbildung 11: Spektrum eines Schweinelauts inklusive der kumulativen Intensität zur Bestimmung der Frequenzquantile. _____	17
Abbildung 12: Schematischer Aufbau der mobilen Audiobox; mit Bildern aus (42–44). _____	19
Abbildung 13: Foto der mobilen Audiobox; <u>Links</u> : Außenansicht. <u>Rechts</u> : Innenansicht. _____	19
Abbildung 14: Allgemeiner Ablauf der Datenauswertung. <u>Links</u> : schematische Darstellung der Vorgehensweise. <u>Mitte</u> : Beschreibung der Vorgehensweise. <u>Rechts</u> : Übersicht der festzulegenden Parameter. _____	21
Abbildung 15: Hierarchische Anordnung der Geräuschklassen. _____	23
Abbildung 16: Verteilung der Längen von 933 Schweinelauten. _____	24
Abbildung 17: Schematische Darstellung der beiden Möglichkeiten, abhängig von der vorliegenden Geräuschlänge sowie der gewählten Fensterbreite das Fenster um das Geräusch zu legen. _____	25
Abbildung 18: Histogramm (Bin-Größe = 0.1) für die Verteilung der Vorhersagewahrscheinlichkeiten aller vorhergesagten Geräusche, unterteilt nach Klassen sowie Klassifizierungsbewertung. _____	31
Abbildung 19: Wahrscheinlichkeitsmatrizen sowie Ergebnisse von mittlere Präzision und mittlere Sensitivität für alle Vorhersagewahrscheinlichkeiten <u>Links</u> : $\geq 50\%$ , <u>Mitte</u> : $\geq 60\%$ sowie <u>Rechts</u> : $\geq 70\%$ . _____	31
Abbildung 20: Ergebnisse der linearen Pearson-Korrelation aller 36 Merkmale als Betrag des Pearson- Koeffizienten in Prozent. _____	32
Abbildung 21: Mittelwert, Minima und Maxima der Präzisionen sowie Sensitivitäten der Ergebnisse der rekursiven Merkmalseliminierung, aufgetragen über die Anzahl der zur Modellerstellung verwendeten Merkmale. _____	33
Abbildung 22: Paarweise Auftragung der acht wichtigsten Merkmale. Teil 1. Orange = Klasse A (bedenkliche Laute); Blau = Klasse B (restliche Geräusche) _____	36
Abbildung 23: Paarweise Auftragung der acht wichtigsten Merkmale. Teil 2. Orange = Klasse A (bedenkliche Laute); Blau = Klasse B (restliche Geräusche) _____	37
Abbildung 24: Mittelwert, Minima und Maxima der Präzisionen sowie Sensitivitäten der Ergebnisse der Vorhersagen, aufgetragen über die Anzahl der zur Modellerstellung verwendeten Hauptkomponenten. _____	38
Abbildung 25: Werte der linearen Diskriminanzanalyse sowie Schwellwert des naiven Bayes-Klassifikators für ein Testdatenset. _____	40
Abbildung 26: (a) unbekanntes kontinuierliches akustisches Signal; (b) Labels von Schweinelauten; (c) vorhergesagte Labels Modell 1; (d) vorhergesagte Labels Modell 2. _____	42
Abbildung 27: Wahrheitsmatrix für Multi-Klassen-Problem. Entnommen aus (13). _____	53
Abbildung 28: Skizze Maststall. _____	55

## 9 TABELLENVERZEICHNIS

<i>Tabelle 1: Merkmale des Zeitbereichs. <math>x =</math> Merkmal wurde auf dazugehöriger Skala bestimmt.</i>	26
<i>Tabelle 2: Merkmale des Frequenzbereichs. <math>x =</math> Merkmal wurde auf dazugehöriger Skala bestimmt.</i>	27
<i>Tabelle 3: Parametersatz Variation Fensterlänge.</i>	29
<i>Tabelle 4: Ergebnisse bei variierenden Fensterbreiten.</i>	29
<i>Tabelle 5: Parametersatz zur Analyse des Geräuschanteils im Fenster sowie Analyse der Vorhersagewahrscheinlichkeit.</i>	30
<i>Tabelle 6: Übersicht Anteil der richtig- bzw. falsch-zugeordneten Geräusche in Abhängigkeit des Geräuschanteils im jeweils betrachteten Fenster, unterteilt nach dem Verhältnis Fenster- zu Geräuschlänge.</i>	30
<i>Tabelle 7: Parametersatz rekursive Merkmals-Eliminierung zur Bestimmung der Merkmalsanzahl.</i>	33
<i>Tabelle 8: Parametersatz zur Bestimmung der Häufigkeit der Merkmale bei einer rekursiven Merkmalseliminierung auf acht Merkmale.</i>	34
<i>Tabelle 9: Ausgewählte Merkmale sowie ihre Wahrscheinlichkeit bei einer rekursiven Merkmalseliminierung auf acht Merkmale ausgewählt zu werden.</i>	34
<i>Tabelle 10: Ergebnisse Vergleich 36 gegenüber acht Merkmale.</i>	35
<i>Tabelle 11: Parametersatz für die Hauptkomponenten-Analyse.</i>	38
<i>Tabelle 12: Ergebnisse des Vergleichs untransformierter gegenüber PCA-transformierter Merkmale.</i>	38
<i>Tabelle 13: Ausgewählte Hauptkomponenten sowie ihre Wahrscheinlichkeit bei einer rekursiven Hauptkomponenteneliminierung auf sechs Hauptkomponenten ausgewählt zu werden</i>	39
<i>Tabelle 14: Ergebnisse Vergleich 36 Merkmale gegenüber auf sechs Hauptkomponenten reduzierte transformierte Merkmale.</i>	39
<i>Tabelle 15: Parametersatz für die lineare Diskriminanzanalyse.</i>	39
<i>Tabelle 16: Ergebnisse Vergleich untransformierter gegenüber LDA-transformierter Merkmale.</i>	40
<i>Tabelle 17: Parameter der in dieser Arbeit entwickelten Methodik.</i>	41
<i>Tabelle 18: Python-Implementierung der Merkmale des Zeitbereichs.</i>	53
<i>Tabelle 19: Python-Implementierung der Merkmale des Frequenzbereichs.</i>	54

# 10 ANHANG

## 10.1 THEORIE UND MERKMALSBERECHNUNG

### Wahrheitsmatrix für Multi-Klassen-Problem

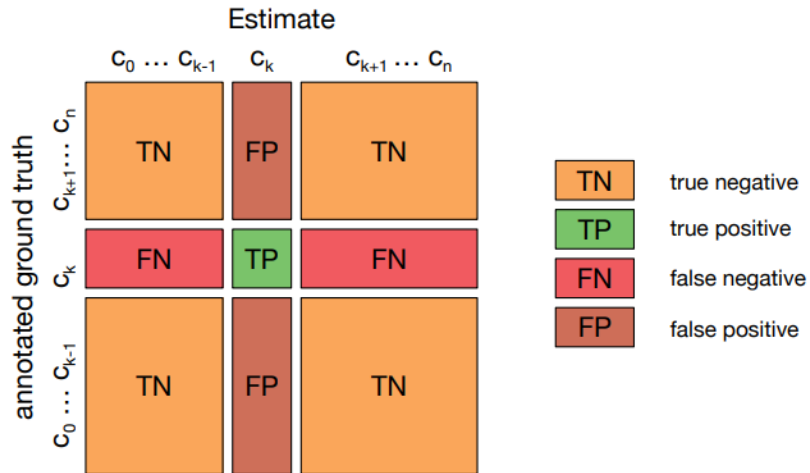


Abbildung 27: Wahrheitsmatrix für Multi-Klassen-Problem. Entnommen aus (13).

### Python-Implementierung der Merkmale

Tabelle 18: Python-Implementierung der Merkmale des Zeitbereichs.

Python-Implementierung der für mehrere Merkmale benötigten Berechnungen	
Berechnung der Extrema der Intensität des Zeitbereichs; benötigt für die Amplituden-Modulations-Werte	Intensität_Zeitbereich = numpy.abs(Amplitude) <sup>2</sup>
	Minima_Intensität = (numpy.diff (numpy.sign(numpy.diff (Intensität_Zeit-bereich))) > 0).nonzero()[0] + 1
	Maxima_Intensität = (numpy.diff (numpy.sign(numpy.diff (Intensität_Zeit-bereich))) < 0).nonzero()[0] + 1
	Extrema = [Minima_Intensität, Maxima_Intensität].sort()
Python-Implementierung der Merkmale	
Merkmal	Python-Implementierung
maximale Amplitude	maxAmplitude = max(numpy.abs(Amplitude))
Energie	Energie = scipy.integrate.simpson (y = numpy.abs(Amplitude), x = Zeit)
Lautstärke	Lautstärke = numpy.sqrt (sum(n*n for n in Amplitude)/len(Amplitude))
Mittelwert	Mittelwert = numpy.mean(numpy.abs(Amplitude))
Varianz	Varianz = numpy.var(numpy.abs(Amplitude))
Schiefe	Schiefe = scipy.stats.skew(numpy.abs(Amplitude))

## Kapitel 10: Anhang

Wölbung	Wölbung = <code>scipy.stats.kurtosis(numpy.abs(Amplitude))</code>
Amplituden-Modulations-Rate	<code>AmpModRate = len(Extrema) / Fensterlänge</code>
Amplituden-Modulations-Ausmaß	<code>AmpModExtent = numpy.mean(numpy.abs (numpy.diff(Extrema)))</code>
Amplituden-Variation	<code>AmpVar = numpy.sum(numpy.abs (numpy.diff(Extrema))) / Fensterlänge</code>
Harmonizität	<code>Harmonizität = Signal_Analysis.features.signal (signal = Amplitude, rate = Aufnahme- rate)</code>

Tabelle 19: Python-Implementierung der Merkmale des Frequenzbereichs.

Python-Implementierung der für mehrere Merkmale benötigten Berechnungen	
Berechnung FFT-Intensität	<code>fft = numpy.fft.fft(Amplitude)</code> <code>FFT-Intensität = numpy.real(fft * numpy.conj(fft))</code>
Berechnung der normalisierten kumulativen Intensität; benötigt zur Berechnung der Quantile	<code>Norm_Kum_Int_Freq = numpy.cumsum( Intensität_Frequenzbereich) / numpy.sum(Intensität_Frequenzbereich)</code>
Python-Implementierung der Merkmale	
Merkmal	Python-Implementierung
Mittelwert	<code>Mittelwert = numpy.mean (Intensität_Frequenzbereich[0:Aufnahmerate/2])</code>
Varianz	<code>Varianz = numpy.var (Intensität_Frequenzbereich[0:Aufnahmerate/2])</code>
Schiefe	<code>Schiefe = scipy.stats.skew (Intensität_Frequenzbereich[0:Aufnahmerate/2])</code>
Wölbung	<code>Wölbung = scipy.stats.kurtosis (Intensität_Frequenzbereich [0:Aufnahmerate/2])</code>
Frequenz mit höchster Amplitude	<code>Freq_max_Amp = Frequenz[numpy.where( numpy.fft.fft(Amplitude) == max(numpy.fft.fft(Amplitude) [0:Aufnahmerate/2]))[0]]</code>
Realer Anteil Fourier-Koeffizient	<code>FFT_Koeffizient_real = numpy.real (max(numpy.fft.fft(Amplitude) [0:Aufnahmerate/2]))</code>
Imaginärer Anteil Fourier-Koeffizient	<code>FFT_Koeffizient_imaginär = numpy.imag (max(numpy.fft.fft(Amplitude) [0:Aufnahmerate/2]))</code>
Frequenz Quantil 25%	<code>fq25% = Frequenz [np.where(Norm_Kum_Int_Freq &gt;= 0.25)[0][0]]</code>
Frequenz Quantil 50%	<code>fq50% = Frequenz [np.where(Norm_Kum_Int_Freq &gt;= 0.50)[0][0]]</code>
Frequenz Quantil 75%	<code>fq75% = Frequenz [np.where(Norm_Kum_Int_Freq &gt;= 0.75)[0][0]]</code>
Spektrale Ebenheit = Wiener Entropie	<code>(f, S) = scipy.signal.periodogram(Amplitude, fs = Aufnahme- rate, scaling='density')</code> <code>Spektrale_Ebenheit = numpy.mean(S) / scipy.stats.mstats.gmean(S)</code>

## 10.2 SKIZZE MASTSTALL

Alle Aufnahmen wurden in dem in Abbildung 28 skizzierten Maststall getätigt. Im Rahmen des Projekts wurden dort vier stationäre Mikrofone mit eigenem Aufnahmesystem installiert. Die Aufnahmen zum Training des Modells wurden mit einem mobilen Mikrofon getätigt, was über dieselben Buchten gehängt wurde. Die Position des Mikrofons wurde zwischen den Aufnahmen variiert und hing an beiden Stellen, die in der Draufsicht markiert sind. Alle Buchten sind durch metallische Wände von 1 m Höhe voneinander getrennt und verfügen über Spaltenböden. Die Außenwände sind gemauert während die Decken mit Schaumstoff verkleidet sind. Da die Decke teilweise schräg verläuft, ist auch dies in der Skizze vermerkt. Die Decke beginnt ab einer Höhe von 1,6 m bis 2,5 m bis sie die Deckenhöhe von 2,8 m erreicht hat. Futter- sowie Wasserautomaten, deren Versorgungsleitungen sowie Balken bzw. Säulen sind in der Skizze nicht vermerkt. Zur besseren Veranschaulichung sind ebenfalls zwei Fotos aus dem Bereich der akustisch überwachten Buchten hinzugefügt.

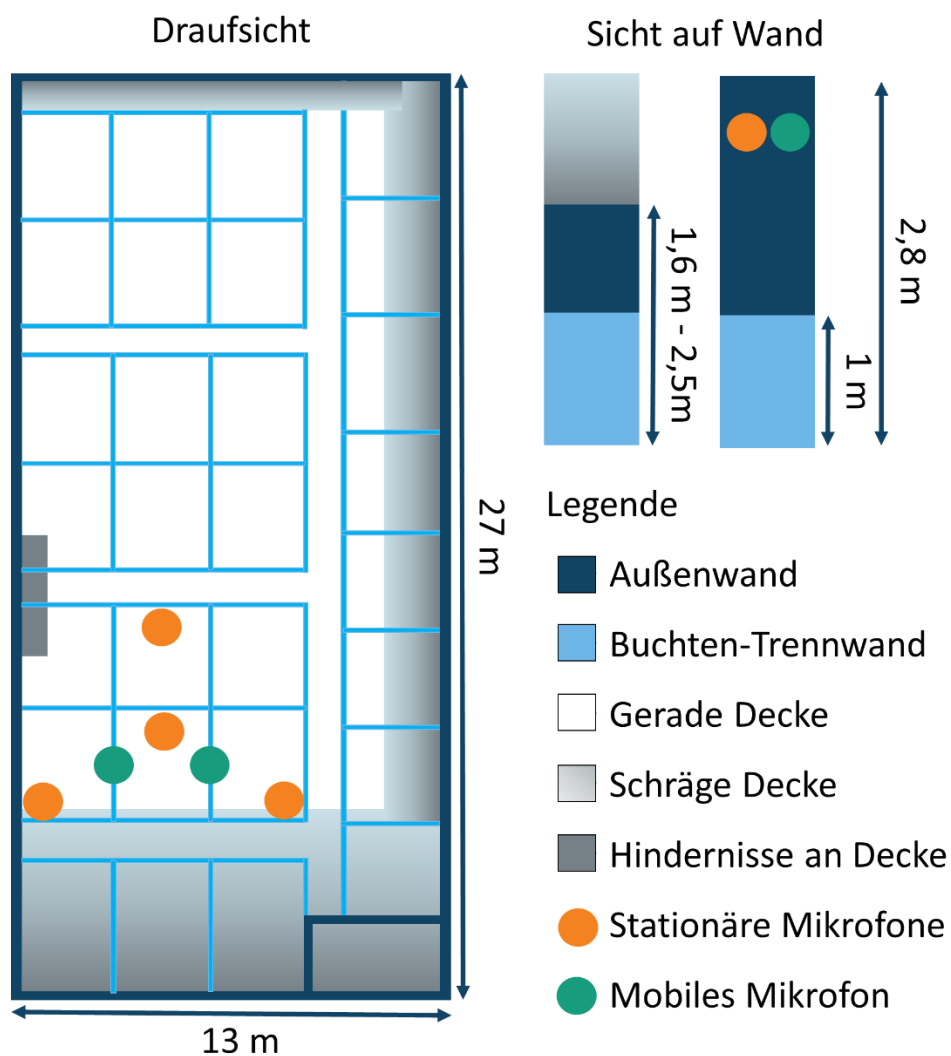


Abbildung 28: Skizze Maststall.



## 10.3 DATENBLÄTTER

*Auszüge aus dem Datenblatt zu Behringer ECM8000*

### Technical Specifications

EN

ES

FR

DE

PT

IT

NL

SE

PL

JP

CN



### MEASUREMENT CONDENSER MICROPHONE ECM8000

Ultra-Linear Measurement Condenser Microphone



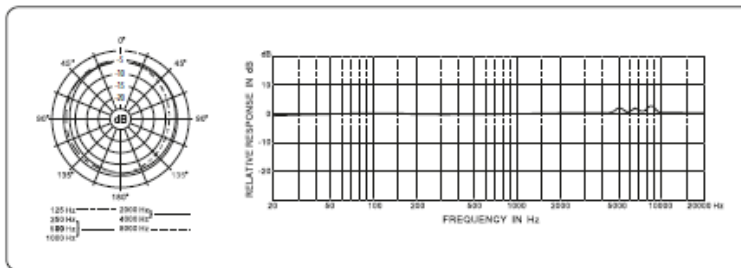
EN

# MEASUREMENT CONDENSER MICROPHONE ECM8000

## Ultra-Linear Measurement Condenser Microphone

### FEATURES

- Ultra-linear condenser microphone for measurement and recording application
- Perfectly suited for room equalization application plus high-resolution studio recording and live applications
- Exceptionally flat frequency response and ultra-high sound resolution
- Evenly weighted, true omnidirectional pattern
- Works with phantom power from +15 to +48 V
- Perfect for use with the Behringer ULTRACURVE or any other analyzer
- Ultra-low noise transformerless FET input eliminates low-frequency distortion
- Gold-plated 3-pin XLR connector for highest signal integrity
- Swivel stand mount and transport case included

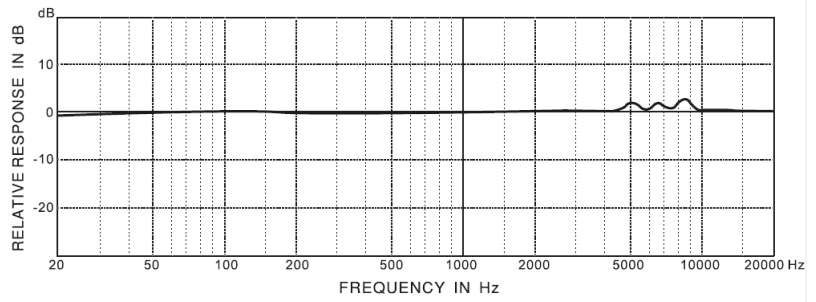
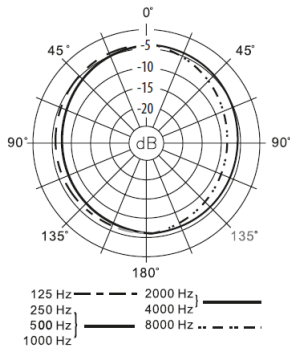


### Specifications

Type	electr. condenser
Polar Pattern	omnidirectional
Impedance	200 Ohms
Sensitivity	70 dB
Freq. Response	20-20000 Hz
Connector	gold-plated XLR
Phantom Power	+15 to +48 V
Weight	136 g

Behringer is constantly striving to maintain the highest professional standards. As a result of these efforts, modifications may be made from time to time to existing products without prior notice. Specifications and appearance may differ from those listed or illustrated.

© 2021 Music Tribe Global Brands Ltd. Technical specifications and appearance are subject to change without notice. The information contained herein is correct at the time of printing. All trademarks are the property of their respective owners. Music Tribe Global Brands Ltd. accepts no liability for any damages or loss which may be suffered by any person who relies either wholly or in part upon any description, photograph or statement contained herein.



Auszüge aus dem Datenblatt zu Behringer U-Phoria UMC202HD

Quick Start Guide



**U-PHORIA**  
**UMC404HD**  
 Audiophile 4x4, 24-Bit/192 kHz  
 USB Audio/MIDI Interface with MIDAS Mic Preamplifiers

**UMC204HD**  
 Audiophile 2x4, 24-Bit/192 kHz  
 USB Audio/MIDI Interface with MIDAS Mic Preamplifiers

**UMC202HD**  
 Audiophile 2x2, 24-Bit/192 kHz  
 USB Audio Interface with MIDAS Mic Preamplifiers

**UMC22**  
 Audiophile 2x2 USB Audio Interface with  
 MIDAS Mic Preamplifier

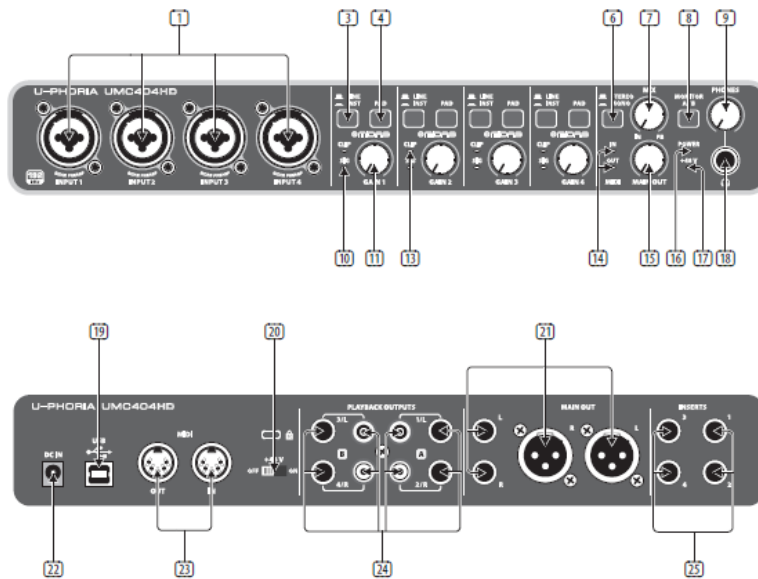
**UM2**  
 Audiophile 2x2 USB Audio Interface with  
 XENYX Mic Preamplifier

V8.0



- EN
- ES
- FR
- DE
- PT
- IT
- NL
- SE
- PL

## U-PHORIA UMC404HD/UMC204HD/UMC202HD/UMC22/UM2 Controls



## U-PHORIA UMC404HD/UMC204HD/UMC202HD/UMC22/UM2 Controls

### DE Schritt 2: Bedienelemente

- |   |  |  |
|---|--|--|
| <p><b>1</b> <b>MIC/LINE 1 (UM2 &amp; UMC22) / INPUT 1 &amp; 2 (UMC202HD &amp; UMC204HD) / INPUT 1 – 4 (UMC404HD)</b> XLR / 6,35 mm-Kombianschlüsse. Hier schließen Sie Mikrofone, Instrumente oder Audioquellen mit Line-Pegel an.</p> <p><b>2</b> <b>INST 2 (UM2 &amp; UMC22)</b> 6,35 mm-Anschluss. Hier schließen Sie Instrumente und Audioquellen an.</p> <p><b>3</b> <b>LINE / INST-Wahltaste (UMC202HD / UMC204HD / UMC404HD)</b>. Damit bestimmen Sie, ob eine Signalquelle mit Line- oder Instrumentenpegel am XLR / 6,35 mm-Kombieingang anliegt.</p> <p><b>4</b> <b>PAD-Wahltaste (UMC202HD / UMC204HD / UMC404HD)</b>. Bei Aktivierung wird der Eingangspegel der angeschlossenen Signalquelle verringert.</p> | <p><b>5</b> <b>DIRECT MONITOR-Wahltaste (UM2 / UMC22 / UMC202HD)</b>. Bei Aktivierung ist das latenzfreie Direktmonitoring der Eingangssignale (keine Verzögerung) eingeschaltet.</p> <p><b>6</b> <b>STEREO/MONO-Wahltaste (UMC204HD &amp; UMC404HD)</b>. Bei Aktivierung kann man die an INPUT 1 und INPUT 2 (UMC204HD) oder INPUTS 1 – 4 (UMC404HD) angeschlossenen Audiosignale in Mono abhören.</p> <p><b>7</b> <b>MIX-Drehregler (UMC204HD &amp; UMC404HD)</b>. Damit stellen Sie das Pegelverhältnis der Eingangssignale und Playback-Signale 1-2 am MAIN OUT-Ausgang und am  -Ausgang ein (wenn MONITOR A für den  -Ausgang gewählt ist).</p> | <p><b>8</b> <b>MONITOR A/B-Wahltaste (UMC204HD &amp; UMC404HD)</b>. Bei Aktivierung werden die Ausgangskanäle 3 &amp; 4 zum  -Ausgang geleitet.</p> <p><b>9</b> <b>PHONES-Drehregler (UMC202HD / UMC204HD / UMC404HD)</b>. Damit stellen Sie den Ausgangspegel am  (Kopfhörer)-Ausgang ein.</p> <p><b>10</b> <b>SIG LED</b> zeigt an, dass ein Audiosignal auf dem Kanal übertragen wird.</p> <p><b>11</b> <b>MIC/LINE GAIN 1 (UM2) / GAIN 1 &amp; 2 (UMC22 / UMC202HD / UMC204HD) GAIN 1 – 4 (UMC404HD)</b> Drehregler. Damit stellen Sie den Pegel für die Eingänge MIC/LINE 1 (UM2), MIC/ LINE 1 &amp; INST 2 (UMC22), INPUT 1 &amp; 2 (UMC202HD &amp; UMC204HD) oder INPUT 1 – 4 (UMC404HD) ein.</p> |
|---|--|--|

- 12 **INST GAIN 2**-Drehregler (UM2). Damit stellen Sie den Pegel am INST 2-Eingang ein.
- 13 **CLIP LED** zeigt an, dass das Audiosignal auf dem Kanal zu laut ist. Drehen Sie den entsprechenden GAIN-Regler nach links, bis die CLIP LED nicht mehr leuchtet.
- 14 **MIDI IN/OUT LEDs** (UMC204HD & UMC404HD) zeigen MIDI-Signalaktivitäten an.
- 15 **OUTPUT** (UM2 / UMC22 / UMC202HD) / **MAIN OUT** (UMC204HD & UMC404HD) Drehregler. Damit steuern Sie den Ausgangspegel an den 1(L) & 2(R) OUTPUTS (UM2 / UMC22 / UMC202HD) oder den L & R MAIN OUTS (UMC204HD & UMC404HD).
- 16 **POWER LED** zeigt an, dass das Gerät eingeschaltet ist.
- 17 **+48 V LED** zeigt an, dass die +48 V Phantomspannung aktiviert ist.
- 18 **☞-Ausgang**. Hier schließen Sie Ihre Kopfhörer zum Abspielen und Mischen des Audiomaterials an.
- 19 **USB Typ B-Anschluss**. Hier schließen Sie Ihren Computer an.
- 20 **+48 V ON/OFF-Wahlschalter** aktiviert +48 V Phantomspannung (erforderlich für professionelle Studio-Kondensatormikrofone).
- 21 **1(L) & 2(R) OUTPUTS** (UM2 / UMC22 / UMC202HD) / **L & R MAIN OUT** (UMC204HD & UMC404HD) Hier schließen Sie aktive Studiomonitore für die Wiedergabe und Abmischung an.
- 22 **DC IN** (UMC404HD) Hier schließen Sie das Gerät über den mitgelieferten Netzadapter ans Stromnetz an.
- 23 **MIDI IN/OUT** (UMC204HD & UMC404HD) Hier schließen Sie externe MIDI Controller und Module an.
- 24 **PLAYBACK OUTPUTS A 1 & 2 / B 3 & 4** (UMC204HD) / **PLAYBACK OUTPUTS A 1/L & 2/R / B 3/L & 4/R** (UMC404HD) Hier schließen Sie externe Lautsprecher für zusätzliche Monitoring-Anwendungen an.
- 25 **INSERTS 1 & 2** (UMC204HD) / **1 – 4** (UMC404HD) Hier schließen Sie externe Signalprozessoren an.

DE

60 U-PHORIA UMC404HD/UMC204HD/UMC202HD/UMC22/UM2

EN

## Specifications

Model	UMC404HD	UMC204HD	UMC202HD
<b>Input</b>			
Preamp	4 x MIDAS design	2 x MIDAS design	
Type	4 x XLR/TRS combo connector (Mic/Line/Instrument)	2 x XLR/TRS combo connector (Mic/Line/Instrument)	
Input level attenuation	4 x Pad switch	2 x Pad switch	
Frequency response	10 Hz – 50 kHz (0 / -3 dB)		
Impedance	Mic in: 3 kΩ / Instrument in: 1 MΩ		
Max. input level	Mic: -4 dBu / Line: +20 dBu / Instrument: -3 dBu		
Phantom power	+48 V, switchable		
<b>Output</b>			
Type	1 x ¼" stereo (Phones), 2 x ¼" TRS (Main Line Out), 2 x XLR (Main Line Out), 4 x ¼" TRS (Playback Line Out), 4 x RCA (Playback Line Out)	1 x ¼" stereo (Phones), 2 x ¼" TRS (Main Line Out), 4 x RCA (Playback Line Out)	1 x ¼" stereo (Phones), 2 x ¼" TRS (Line Out)
Frequency response	10 Hz – 43 kHz (0 / + 0.3 dB)		
Max. output level	+3 dBu		

Model	UMC404HD	UMC204HD	UMC202HD
<b>System Data</b>			
Dynamic range	100 dB, A-weighted	110 dB, A-weighted	
Frequency response	10 Hz – 43 kHz (0 / + 0.3 dB)		
<b>Inserts</b>			
Type	4 x ¼" TRS	2 x ¼" TRS	n/a
Impedance	100 Ω Send / 10 kΩ Return		n/a
Max. input level	+3 dBu		n/a
<b>MIDI I/O</b>			
Type	5-pin DIN; 1 in, 1 out		n/a
<b>Monitoring</b>			
Direct monitor control	Mix Balance control (input <-> playback)		Direct Monitor switch
Headphones DJ cueing	Monitor A/B switch (out 1-2 / 3-4 select)		n/a
<b>Security Lock</b>			
Kensington security lock	Yes		
<b>Digital Processing</b>			
Supported sample rates	44.1 / 48 / 88.2 / 96 / 176.4 / 192 kHz		
<b>Computer Bus Connectivity</b>			
Type	USB 2.0, type B		

62 U-PHORIA UMC404HD/UMC204HD/UMC202HD/UMC22/UM2

## Specifications

Model	UMC404HD	UMC204HD	UMC202HD
<b>Software Support</b>			
Operating systems	Mac OS X*, Windows XP* or higher*		
Drivers	Mac*: No driver required / CoreAudio supported Windows*: Available as download from behringer.com		
<b>Power</b>			
Power consumption	Max. 2.5 W		
Power supply	USB connector (bus-powered from computer) and adapter	USB connector (bus-powered from computer)	
Mains voltage	100-240 V~, 50/60 Hz	n/a	
<b>Dimensions</b>			
Dimensions (H x W x D)	45.81 x 292 x 130 mm (1.78 x 11.5 x 5.1")	46.45 x 185 x 130 mm (1.8 x 7.3 x 5.1")	46.45 x 170 x 125 mm (1.8 x 6.7 x 4.9")
Weight	1.2 kg (2.7 lbs)	0.6 kg (1.4 lbs)	0.5 kg (1.2 lbs)