

# Musik und Emotion: Ein auf SVM-Klassifikation basierendes anwendungsorientiertes Werkzeug zur Emotionserkennung

Michael Wittland<sup>1</sup>, Vincent Raciti<sup>2</sup> und Michael Oehler<sup>3</sup>

**Abstract:** Produktmarken werden mit emotionalen Eigenschaften verbunden, um ihnen einen persönlichen Charakter zu verleihen. Musik ist ein wichtiges Medium, dem Verbraucher die emotionalen Assoziationen einer Marke zu vermitteln. Zusammen mit TRO Music Services wurde ein Tool entwickelt, welches automatisiert überprüft, ob ein Musikstück zu einer gewünschten Markenpersönlichkeit passt. Grundlage zur Beschreibung der Markenpersönlichkeit bildet das Hevnersche Emotionsmodell [He36]. Ein Datenkorpus von 1831 Musikstücken wurde hierzu von Probanden getaggt. Anschließend wurde damit ein Support Vector Classification Algorithmus trainiert, Emotionstags anhand von akustischen Merkmalen zu erkennen. Der Datenkorpus besteht größtenteils aus professionell produzierter, unveröffentlichter Musik, wie sie auch in Mainstream-Radiosendern gespielt werden könnte. Einerseits werden so die Urteile der Probanden nicht aufgrund außermusikalischer Faktoren, die mit dem Interpretieren zu tun haben, beeinflusst, andererseits ist eine größere ökologische Validität als bei Produktionsmusik gewährleistet. Bei einer Kategorisierung von Musikstücken in 8 Hevnersche Emotionsdimensionen werden gegenwärtig durchschnittlich etwa 71% der Stücke in jeder Emotionskategorie korrekt klassifiziert.

**Keywords:** Music Emotion Recognition, Support Vector Classification, Hevner Emotion Model

## 1 Einleitung

Marken und deren Produkte sind ständige Begleiter in unserem Leben. Wir umgeben uns mit ihnen im Alltag. Die Auswahl der Marken kann hierbei Ausdruck unserer Persönlichkeit und Einstellung sein. Die Form der Bindung, die hierbei entsteht, ist vergleichbar mit einer persönlichen Beziehung: Wir wählen Charaktereigenschaften und emotionale Attribute aus, welche zu uns passen. Die Marke ist neben dem rein sachlichen Wert also auch eine Verbindung menschlicher Eigenschaften, die mit der Marke assoziiert werden. Man spricht dann von einer Markenpersönlichkeit [Aa99]. In der Werbung werden neben den sachlichen Vorteilen eines Produktes, wie Preis und Güte, auch die emotionalen Eigenschaften, also die Markenpersönlichkeit, der dahinterstehenden Marke kommuniziert [La16].

---

<sup>1</sup> Hochschule Düsseldorf, Gronauer Str. 51, 51063 Köln, mja.wittland@gmail.com

<sup>2</sup> TRO Music Services, Zimmerstr. 19, 40215, vincent@tro.de

<sup>3</sup> Universität Osnabrück, Institut für Musikwissenschaft, Abteilung für Musik und Medientechnologie, Neuer Graben/Schloss, Osnabrück 49074, michael.oehler@uos.de

Musik ist in Werbung eine sehr wichtige Komponente für die Erzeugung von Emotionen sein. Die Auswahl der richtigen Musik ist daher von zentraler Bedeutung für das Erschaffen einer Markenpersönlichkeit [Br16]. Die emotionale Interpretation eines Musikstückes ist allerdings meist sehr subjektiv und variiert unter Umständen deutlich von Person zu Person [JS01]. Um Informationen über den emotionalen Gehalt eines Stückes zu gewinnen, werden häufig Probanden in Hörversuchen nach den empfundenen bzw. transportierten Emotionen des Musikstückes befragt. Dies ist jedoch mit nicht unerheblichem Kosten- und Zeitaufwand verbunden, der unter Umständen in einem anwendungsorientierten Kontext Budget und zeitlichen Rahmen überschreitet.

Um dennoch schnell einen möglichst objektiven Blick auf die Emotionen eines Stückes gewinnen zu können, wurde in Zusammenarbeit mit der Musikagentur TRO Music Services ein Tool zur automatisierten Erkennung von Emotionen in Musikstücken mittels maschinellen Lernens entwickelt, welches bezüglich der verwendeten Methoden einerseits den aktuellen Stand der Forschung berücksichtigt ([FR10, BF13]), andererseits bekannte Probleme, wie z.B. die Grenzen der Verallgemeinerbarkeit von Ergebnissen in Abhängigkeit der in den Studien verwendeten Musikkorpora, adressiert. Gerade in Bezug auf die intendierte Anwendungsorientiertheit, d.h. den praktischen Einsatz im Agenturalltag, ist dies ein sehr wichtiger Aspekt.

Die automatisierte Erkennung von Emotionen in Musik ist bereits seit langem ein Bereich des Music Information Retrieval. Es besteht allerdings die Schwierigkeit, dass sich die Güte der Algorithmen kaum vergleichen lässt, da unterschiedliche Emotionsmodelle oder Musikdatenbanken mit unterschiedlichen Genreschwerpunkten verwendet werden. Aus diesem Grund wurde 2007 die *MIREX Mood Classification Task* ins Leben gerufen. Hier werden jährlich neue Emotionsklassifizierungsalgorithmen an einer vorgegebenen Produktionsmusikdatenbank getestet und verglichen. Ähnlich wie in der vorliegenden Arbeit beruht das Emotionsmodell der MIREX Mood Classification Task auf dem Tagging von Emotionsgruppen (Cluster). Tab. 1 zeigt eine Auswahl der besonders gut klassifizierenden Algorithmen der MIREX Task von 2011 bis 2016. Es wurde bei der Auswahl darauf geachtet, dass die Algorithmen nicht nur in einem Bereich gut funktionieren, sondern über alle Emotionskategorien gut klassifizieren. Existieren von einem Autor mehrere Einreichungen, die einen ähnlichen Ansatz verfolgen, so wurde derjenige mit der besten Klassifikationsgenauigkeit oder - bei vergleichbaren Klassifikationsgenauigkeiten - der mit dem neusten Erscheinungsdatum gewählt.

Um eine ungefähre Einschätzung über die Güte des vorliegenden Algorithmus zu erhalten, wurde für dessen Evaluation ebenso wie in der MIREX Mood Task die Klassifizierungsgenauigkeit als Evaluationsmaß verwendet. Da jedoch ein anderes Emotionsmodell und eine andere Musikdatenbank verwendet wurden, ist ein Vergleich mit der MIREX Mood Classification Task nur bedingt möglich. Außerdem durfte bei der MIREX Task nur eine Emotion mit einem Musikstück assoziiert werden, während in dieser Arbeit ein Stück auch mit mehreren Emotionstags verbunden werden kann. Dies erscheint in diesem Kontext sinnvoll um die emotionale Komplexität eines Musikstückes vollständig zu beschreiben.

Emotionskategorien	[SS15]	[Wu13]	[PR13]	[RW15]	[TC11]
Emotionsgruppe 1	55.83	60.00	59.17	62.50	58.33
Emotionsgruppe 2	57.50	57.50	55.83	60.83	55.00
Emotionsgruppe 3	74.17	80.00	85.83	80.83	74.17
Emotionsgruppe 4	55.83	66.67	55.83	63.33	62.50
Emotionsgruppe 5	82.50	77.50	81.67	80.00	85.83
<b>Mittlere Klassifikationsgenauigkeit</b>	<b>65.17</b>	<b>68.33</b>	<b>67.67</b>	<b>69.50</b>	<b>67.17</b>

Tab. 1: Die besten Ergebnisse der MIREX Mood Classification Task.

Als Klassifikationsmethode für die vorliegende Arbeit wurde ein auf Support Vector Machine (SVM) Klassifikatoren basierendes Modell gewählt. Dieses Modell hat sich in der Musikemotionserkennung bewährt und findet auch in drei der in Tab. 1 aufgeführten Ansätze Verwendung ([PR13, RW15, TC11]). Das akustische Merkmalsset basiert für das Training der SVM in weiten Teilen auf der Studie von Baume et al. [BF14].

## 1.1 Musikdatenbank

Die Auswahl der Musikdatenbank, mit der die SVM trainiert wird, ist von zentraler Bedeutung. In bisherigen Studien wurde oft entweder ein Korpus populärer Musikstücke (häufig eine Charts-Auswahl), eine Form von Produktionsmusik oder Kombinationen davon verwendet (z.B. [HD08, EL09, SE11, PP12, BF14, AY17]). Während populäre, in den Medien häufig präsente Musik das Problem mit sich bringt, dass die Urteile der Versuchspersonen nicht nur die Musik an sich, sondern unter Umständen auch Eigenschaften des Künstlers/der Band etc. betreffen können, die Bewertung also verzerrt wird, hat Produktionsmusik in vielen Fällen eine nur beschränkte Aussagekraft. Dies liegt u.a. an der speziellen Ausrichtung von Produktionsmusik als Hintergrundmusik in TV-, Film- und Radioproduktionen. Es handelt sich dabei häufig um instrumentale Musik [BF14], d.h. sie entspricht nicht dem überwiegenden Anteil der heute über Radio oder digitale Distributionskanäle konsumierten Musik. Andere Datenkorpora beschränken sich wiederum auf ein Genre [HR09].

Bei der Datenbank von TRO Music Services handelt es sich um ein Archiv von überwiegend in den eigenen Studios bzw. in Lizenz produzierter Musik, die für Film, TV und Werbung verwendet wird. Entgegen des im Zusammenhang mit der MIREX Mood Classification Task verwendeten APM Datenkorpus [HD08], wurde das in der Studie verwendete Datenset noch nie im Bereich der Emotionsklassifikation eingesetzt, die enthaltene Musik ist relativ aktuell. Die Datenbank umfasst ein breites Spektrum an Genres und Verwendungsszenarien und enthält sowohl instrumentale Musik als auch

Musik mit Gesang. Die Länge der Audiofiles variiert dabei zwischen 00:00:03 und 00:22:39 (hh:mm:ss). Das kürzeste Audiofile ist ein Soundlogo, das längste ist ein klassisches Musikstück. Die durchschnittliche Länge beträgt 00:02:50. Entsprechend der anwendungsorientierten Ausrichtung auf eine kommerzielle Verwendung (häufig für Werbung) ist der intendierte Spannungsverlauf jedoch meist eher klein, so dass sich eine relativ unveränderliche Grundstimmung über den ganzen Song ermitteln lässt. Die Datenbank eignet sich daher gut zum Tagging und Training eines Klassifikations-Algorithmus.

## 1.2 Support Vector Machine

Die Support Vector Machine (SVM) ist ein mathematisches Modell zur Mustererkennung in Datensätzen ([CV95, FC05]). Das Ziel einer SVM ist es, für ein Beobachtungsobjekt (z. B. ein Musikstück) mit Hilfe einer Datenmenge  $X$ , welche das Objekt beschreibt, ein gewünschtes Label  $Y$  zu erkennen. Support Vector Machines nutzen hierzu die Methode des *überwachten Lernens*: Der Algorithmus wird zunächst mit Daten trainiert, bei denen das Label  $Y$  bereits bekannt ist. Anschließend soll das trainierte Modell in der Lage sein, das Label  $Y$  bei noch unbekanntem Datensätzen nur anhand der beschreibenden Menge  $X$  auszugeben. Grundsätzlich gibt es zwei Arten der SVM: Klassifikatoren und Regressoren. Klassifikatoren können lediglich unterscheiden, ob ein Objekt einer gesuchten Klasse angehört oder nicht. Das Label  $Y$  kann also nur die Werte 1 und 0 annehmen. Bei Regressoren dagegen wird die Zugehörigkeit eines Objektes zu einer Klasse durch kontinuierliche Werte beschrieben.

## 1.3 Emotionsmodell

Ein gebräuchliches Modell zur Beschreibung von Emotionen ist das *Circumplex Model of Affect* [Ru80] von James Russell, von dem eine Reihe von Abwandlungen existieren. Bei diesem Modell sind die Emotionen in ein zweidimensionales Koordinatensystem mit den Achsen Valenz (horizontal) und Aktivierung (vertikal) eingeordnet. Die Valenz beschreibt, ob eine Emotion positiv oder negativ belegt ist. Für Werbemusik hat diese Art der Klassifikation allerdings einen entscheidenden Nachteil: Eine Marke wird sich auf der Valenz-Achse nachvollziehbarerweise immer im positiven Bereich positionieren, wodurch diese Dimension ihre Aussagekraft verliert.

Für Musik kommt bei der Einteilung nach Valenz erschwerend hinzu, dass beispielsweise traurige Musik durchaus angenehm, also positiv, ist, während die Emotion Traurigkeit allgemein jedoch dem negativen Bereich der Valenzachse zugeordnet wird [JS08, S.572].

Aus diesen Gründen basiert das Tagging der TRO Datenbank auf dem Emotionsmodell von Kate Hevner. Hevner ließ in einer Studie eine Anzahl von Probanden 66 Adjektive kreisförmig so anordnen, dass Gruppen entstehen, die jeweils einen fließenden Übergang zu den im Kreis benachbarten Gruppen haben. Das so entwickelte Modell setzt sich aus

acht Gruppen kreisförmig angeordneter Adjektive zusammen [He36]. Die Musikagentur TRO verwendet eine reduzierte Variante des Hevnerschen Emotionsmodells mit insgesamt 40 Adjektiven, wobei jede der acht Gruppen fünf Adjektive enthält (siehe Abb. 1). Zudem wurden die Adjektive ins Deutsche übersetzt. Wenn ein Adjektiv mehrere Übersetzungen hatte, wurde darauf geachtet, dasjenige zu wählen, welches der Bedeutungskonnotation der Gruppe am meisten entspricht. Dass die Gruppen nicht vollständig trennscharf sind, zeigt das Adjektiv *lebhaft* welches sowohl in Gruppe 6 als auch in Gruppe 5 auftaucht. Das Adjektiv wurde dennoch beibehalten, um den Emotionsraum der Gruppe adäquat zu beschreiben.

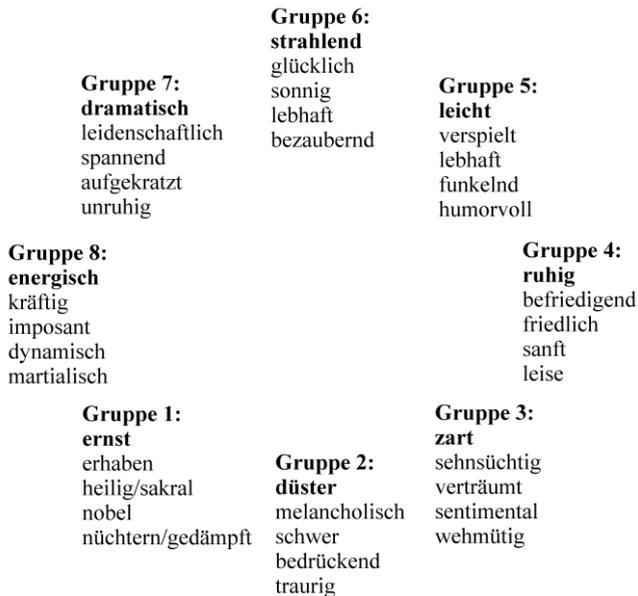


Abb. 1: Das verwendete Emotionsmodell nach [He36].

## 2 Methode

### 2.1 Tagging

Die Datenbank von TRO Music Services umfasst insgesamt 1831 Stücke. Jeder Musiktitel ist mit Tags versehen, welche die Emotionen des Stückes mit Hilfe der acht Adjektive aus dem Hevnerschen Emotionsmodell (siehe Kapitel 1.3 *Emotionsmodell*) beschreiben.

Das Tagging wurde von fünf Experten durchgeführt, welche vorher in das Emotionsmodell eingeführt worden waren. Die Experten hatten beim Hören eines Songs fünf Adjektive aus jeder der acht Hevnerschen Emotionsgruppen zur Verfügung. Wurde ein Song mit mindestens zwei Adjektiven einer Emotionsgruppe versehen, so wurde der entsprechende Tag für die Emotionsgruppe vergeben. Ein Musiktitel kann mit mehreren Tags versehen sein. Tab. 2 zeigt die Verteilung der Musikstücke innerhalb der Emotionsgruppen. Durchschnittlich wurden vier Tags pro Musikstück vergeben, wobei zwölf Stücke nur einen Tag und sechs Stücke alle acht Tags erhielten.

	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5	Gruppe 6	Gruppe 7	Gruppe 8
Gruppe 1: <i>Ernst</i>	462	282	266	180	130	157	346	261
Gruppe 2: <i>Düster</i>	282	517	308	193	111	107	418	277
Gruppe 3: <i>Zart</i>	266	308	1046	728	690	676	384	320
Gruppe 4: <i>Ruhig</i>	180	193	728	851	624	605	228	151
Gruppe 5: <i>Leicht</i>	130	111	690	624	1133	966	355	495
Gruppe 6: <i>Strahlend</i>	157	107	676	605	966	1156	411	550
Gruppe 7: <i>Dramatisch</i>	346	418	384	228	355	411	872	611
Gruppe 8: <i>Energisch</i>	261	277	320	151	495	550	611	901

Tab. 2: Anzahl der Musikstücke je Emotionsgruppe.

## 2.2 Musikalische Features

Es existiert eine Vielzahl musikalischer Features zur Beschreibung von Audiodaten. Um eine für das Training relevante Vorauswahl dieser Features zu treffen, orientiert sich die Auswahl an der Studie von Baume et al. (2014) [BF14]. Dort wurde ein Support Vector Regression Algorithmus trainiert, Musikstücke anhand von musikalischen Merkmalen in ein fünfdimensionales Emotionsmodell einzuordnen. Anschließend wurde evaluiert, welche Features zur Klassifikation der Musikstücke besonders aussagekräftig waren.

Da das in dieser Studie verwendete Emotionsmodell nicht mit dem der vorliegenden Arbeit identisch ist, kann man von einer exakten Übereinstimmung der relevanten Features nicht notwendigerweise ausgehen, jedoch kann grundsätzlich eine gewisse Ähnlichkeit der relevanten Merkmale angenommen werden. Ziel war es u.a. die Anzahl der Features bei ausreichender Klassifikationsleistung relativ gering zu halten, um im anwendungsorientierten Kontext eine Zeit- und damit Ressourcenersparnis zu erzielen.

Die musikalischen Features wurden mit der MIRtoolbox unter Verwendung der Standardeinstellungen in MATLAB extrahiert [LT08]. Die nachfolgende Tabelle (Tab. 3) zeigt die Liste der letztlich genutzten Features. Der so ermittelte Trainingsvektor enthält 26 beschreibende Werte pro Song.

Kategorie	Musikalisches Feature	MIRtoolbox Funktion
Spektral	MFCC (13 Coeffs)	mirmfcc
	Flatness	mirflatness
	Spread	mirspread
	Inharmonicity	mirinharmonicity
	Brightness	mirbrightness
	Irregularity	mirregularity
	Centroid	mircentroid
	Rolloff	mirrolloff
	Harmonisch	Mode (major/minor)
Rhythmisch	Eventdensity	mireventdensity
	Tempo	mirtempo
	Pulseclarity	mirpulseclarity
Energie	RMS Energy	mirrms
Zeitlich	Zero-Crossings	mirzerocross

Tab. 3: Die verwendeten musikalischen Features.

### 2.3 Training des Algorithmus

Da ein Musikstück mehrere Emotionstags haben kann, wurde für jede der acht Emotionsgruppen ein eigener SVM Klassifikator trainiert.

Der gesamte Datensatz ist für manche Emotionsgruppen sehr unausgewogen, d.h. zum Beispiel, dass auf eine kleine Anzahl von Songs innerhalb einer Klasse eine große Anzahl von Songs außerhalb der Klasse kommen (siehe Tab. 1 **Fehler! Verweisquelle konnte nicht gefunden werden.**). Dadurch kann es geschehen, dass die SVM eine scheinbar gute Klassifikationsgenauigkeit erzielt, wenn der Algorithmus alle Songs ausnahmslos außerhalb der Klasse einordnet. Um eine gute Voraussetzung für das Training zu schaffen, wurde daher für jede SVM ein eigener Trainingsdatensatz erstellt. Dieser besteht jeweils zur Hälfte aus Musikstücken innerhalb und zur anderen Hälfte aus Musikstücken außerhalb der zu trainierenden Emotionsgruppe. Um beispielsweise eine SVM für Gruppe 1 des Emotionsmodells (siehe Abb. 1) zu trainieren, wurden alle zu Gruppe 1 gehörigen Songs sowie eine gleich große Anzahl zufällig gewählter Songs, welche nicht Gruppe 1 angehören, in den dafür vorgesehenen Trainingsdatensatz aufgenommen. Durch diese gleiche Verteilung wird nun eine SVM, die alle Songs des Trainingsdatensatzes außerhalb der Emotionsgruppe einordnet für diesen Datensatz durchschnittlich eine Klassifikationsgenauigkeit von nur 50% erzielen. Da die Gesamtzahl der Songs innerhalb jeder der acht Emotionsgruppen unterschiedlich groß ist, variiert auch entsprechend die Größe der Trainingsdatensätze für jede SVM.

Die für das Training der SVM in Matlab verwendete Kernelfunktion ist die gaußsche radiale Basisfunktion. Um zu vermeiden, dass Features mit großen Werten andere Features mit kleineren Werten dominieren, wurden die Werte der Features außerdem vor dem Training standardisiert.

Zur Evaluation des SVM Klassifikators wurde während des Trainings die Methode der k-fachen Kreuzvalidierung verwendet. Die Kreuzvalidierung wurde an jeder SVM mit den  $k=\{10, 8, 6, 4\}$  durchgeführt. Als Maß für die Genauigkeit der SVMs wurde die arithmetisch gemittelte Klassifizierungsgenauigkeit der vier Kreuzvalidierungen mit unterschiedlichen k-Werten gewählt (Tab. 4).

## 3 Ergebnisse

In der nachfolgenden Tabelle (Tab. 4) ist das Ergebnis des Trainings dargestellt. Spalte 3 zeigt das gemittelte Ergebnis der Kreuzvalidierungen. In der zweiten Spalte ist die Größe des Trainingssets für jede Emotionsgruppe angegeben.

<b>Emotions- gruppe</b>	<b>Größe des Trainingssets</b>	<b>mittlere Genauigkeit [%]</b>
Gruppe 1	924	71,61
Gruppe 2	1034	70,23
Gruppe 3	1570	72,06
Gruppe 4	1702	73,61
Gruppe 5	1396	69,63
Gruppe 6	1350	68,57
Gruppe 7	1744	69,21
Gruppe 8	1802	76,26

Tab. 4: Die Klassifikationsgenauigkeit der SVMs.

Zusätzlich wurden die verwendeten musikalischen Features mit Hilfe einer Variante der *Forward Selection Wrapper Methode* [KJ97] auf ihre Relevanz für die Klassifikationsgenauigkeit des Algorithmus untersucht. Hierzu wird jedes Feature einzeln zum Training der SVM verwendet. Das Feature, welches das beste Ergebnis erzielt, wird in das noch leere Featureset mit aufgenommen. Anschließend wird das Featureset in Kombination mit jedem der verbleibenden Features getestet. Das Feature welches in Kombination mit dem Set die beste Klassifikation erzielt, wird neu in das Set aufgenommen. Dieser Vorgang wird wiederholt bis alle Features in das Set aufgenommen sind. Zeigt das Featureset ab einer gewissen Anzahl an Features keine Verbesserung im Training mehr, werden die verbleibenden Features für die Klassifikation als nicht relevant angenommen. Alle verwendeten Features trugen jedoch der Forward Selection Methode zufolge bei allen SVMs zur Verbesserung der Klassifikationsgenauigkeit bei.

Um zu überprüfen, ob größere Samplezahlen voraussichtlich eine deutliche Verbesserung der Klassifikationsgenauigkeit erzielen, wurde der Algorithmus außerdem in jeder der acht Emotionskategorien mit einer immer kleiner werdenden Anzahl von Songs trainiert. Für jede Samplegröße wurden mehrere 10-fache Kreuzvalidierungen durchgeführt.

Das Ergebnis zeigt, dass die Klassifikationsrate je nach Emotionskategorie von der maximalen Samplezahl bis zu einer Samplezahl von etwa 300-500 Songs je nach Emotionsgruppe gleichbleibend ist. Dann verschlechtern sich die Klassifikationsraten und die Ergebnisse der Kreuzvalidierungen. Bei Gruppe 7 verschlechterte sich die Klassifikation von Beginn an relativ kontinuierlich. Da das Ergebnis mit einer k-fachen Kreuzvalidierung ermittelt und somit das Training bei der Evaluierung mit  $k-1$  Subsets des ganzen Datensatzes durchgeführt wurde, sind die für das Training notwendigen Samplegrößen, um die vorliegende Klassifikationsgenauigkeit zu erzielen, eher geringer.

## 4 Fazit

Mit Hilfe von Support Vector Klassifikatoren konnte ein Algorithmus mit dem Hevnerschen Emotionsmodell trainiert werden. Es wurde ein neues, anwendungsorientiertes und ökologisch valides Datenkorpus verwendet. Die Klassifikationsgenauigkeit des trainierten Algorithmus variiert je nach Emotion zwischen rund 69% und 76% korrekter Klassifikation. Es ist anzunehmen, dass - außer in Emotionsgruppe 7 - größere Samplezahlen das Trainingsergebnis nicht maßgeblich verbessern.

## Literaturverzeichnis

- [Aa99] Aaker, J. L.: Dimensionen der Markenpersönlichkeit. In: *Moderne Marktführung*. Gabler Verlag, S. 91-102, 1999.
- [AY17] Aljanaki, A.; Yang, Y. H.; Soleymani, M.: Developing a benchmark for emotional analysis of music. *PLoS one* 12(3), 2017.
- [BF13] Barthez, M.; Fazekas, G.; Sandler, M.: Music Emotion Recognition: From Content to Context-Based Models. *From Sounds to Music and Emotions*, Springer, S. 228–252, 2013.
- [BF14] Baume, C.; Fazekas, G.; Barthez, M.; Marston, D.; Sandler, M.: Selection of Audio Features for Music Emotion Recognition Using Production Music. *AES 53<sup>rd</sup> International Conference*, London 2014.
- [Br16] Bronner, K.: Schöner die Marken nie klingen... Jingle all the Way? *Grundlagen des Audio-Branding*. Audio-Branding. Nomos, S. 50-69, 2016.
- [CV95] Cortes, C.; Vapnik, V.: Support-vector networks. *Machine learning* 20(3), S. 273-297, 1995.
- [EL09] Eerola, T.; Lartillot, O.: Prediction of multidimensional emotional ratings in music from audio using multivariate regression models. *10<sup>th</sup> International Society for Music Information Retrieval Conference*, S. 621–626, 2009.
- [FC05] Fan, R. E.; Chen, P. H.; Lin, C. J.: Working set selection using second order information for training support vector machines. *Journal of machine learning research*, 6(Dec), S. 1889-1918, 2005.
- [FR10] Fazekas, G.; Raimond, Y.; Jacobson, K.; Sandler, M.: An overview of semantic web activities in the OMRAS2 project. *Journal of New Music Research* 39(4), S. 295–311, 2010.
- [HD08] Hu, X.; Downie, J. S.; Laurier, C.; Bay, M.; Ehmann, A. F.: The 2007 MIREX audio mood classification task: Lessons learned. *9<sup>th</sup> International Society for Music Information Retrieval Conference*, S. 462–467, 2008.
- [He36] Hevner, K.: Experimental Studies of the Elements of Expression in Music. *The American Journal of Psychology* Vol. 48, S. 246-268, 1936.

- [HR09] Han, B.; Rho, S.; Dannenberg, R.; Hwang, E.: SMERS: Music Emotion Recognition Using Support Vector Regression. 10<sup>th</sup> International Society for Music Information Retrieval Conference, 2009.
- [JS01] Juslin, P. N.; Sloboda, J. A.: Music and emotion: Theory and research. Oxford University Press, 2001.
- [KJ97] Kohavi, R.; John, G. H.: Wrappers for feature subset selection. Artificial Intelligence 97, S. 273-324, 1997.
- [La16] Lange, A.: Markenpersönlichkeit. Der Einfluss unbekannter Werbegesichter auf die Wahrnehmung der Markenpersönlichkeit, Springer Fachmedien Wiesbaden, S. 17-28, 2016.
- [LT08] Lartillot, O.; Toivianen, P.; Eerola, T.: A matlab toolbox for music information retrieval. Data analysis, machine learning and applications, Springer, S. 261-268, 2008.
- [JS08] Juslin, P.; Sloboda, J.: Emotional responses to music. Behavioral and Brain Sciences 31(5), S. 559-575, 2008.
- [TC11] Tardieu, D.; Charbuillet, C.; Cornu, F.; Peeters, F.: MIREX 2011: Single-Label and Multi-Label Classification Tasks <http://www.music-ir.org/mirex/abstracts/2011/TCCP4.pdf>, Stand: 02.09.2011.
- [PP12] Panda, R.; Paiva, R. P.: Music Emotion Classification: Dataset Acquisition and Comparative Analysis. 15<sup>th</sup> International Conference on Digital Audio Effects (DAFx), S. 1-7, York, UK, 2012.
- [PR13] Panda, R.; Rocha, B.; Paiva, R. P.: MIREX 2013: Mood Classification Task Submission. <http://www.music-ir.org/mirex/abstracts/2013/PRP3.pdf>. Stand: 08.09.2013
- [Ru80] Russell, J.: A Circumplex Model of Affect. Journal of Personality and Social Psychology Vol. 39 No. 6, S. 1161-1178, 1980.
- [RW15] Ren, J.; Wu, M.; Jang, J. R.: Automatic Music Mood Classification Based on Timbre and Modulation Features. IEEE 6(3), S. 236-246, 2015.
- [SE11] Saari, P.; Eerola, T.; Lartillot, O.: Generalizability and Simplicity as Criteria in Feature Selection: Application to Mood Classification in Music. IEEE Transactions on Audio, Speech and Language Processing, 19(6): S. 1802-1812, 2011.
- [SS15] Seyerlehner, K.; Schedl, M.: MIREX 2014 Optimizing the Fluctuation Pattern Extraction Process. MIREX. <http://www.music-ir.org/mirex/abstracts/2014/SS6.pdf>, Stand: 04.02.2015
- [Wu13] Wu, M: MIREX 2013: Submission for Train/Test Tasks (Draft). MIREX. <http://www.music-ir.org/mirex/abstracts/2013/JJ1.pdf>, Stand: 31.10.13.