

## Magister-/Master-/Diplomarbeit

# Comparison of audio-based and language-based sentiment analysis for music recommendation

Der Ursprung der Sentiment-Analysis liegt ursprünglich im Bereich der Computerlinguistik bzw. der maschinellen Textanalyse und beschäftigt sich mit der Quantifizierung emotionalen Inhalts eines Textes. Dieser durch ihre hohe Abstraktionsebene sehr komplexe Aufgabe hat insbesondere im letzten Jahrzehnt durch einerseits starke Entwicklungen im Bereich des Machine Learnings, insbesondere im Bezug auf künstliche neuronale Netzwerke (ANNs), gleichzeitig aber auch durch den großen Zuwachs verfügbarer Textmengen im Internet, etwa in sozialen Netzwerken aber auch über den Online-Handel oder Nachrichten große Fortschritte gemacht.

Ein thematisch eng verwandtes Gebiet ist die Music Emotion Recognition (MER), welche ein Teilgebiet des Music Information Retrieval (MIR) darstellt. Diese behandelt die computerbasierte Analyse der emotionalen Wirkung von Musik basierend auf der Partitur oder dem Audiosignal. Auch in diesem Feld sind konnten vor allem in den letzten Jahren durch die starke Entwicklung im Bereich des Machine Learning beträchtliche Fortschritte verzeichnet werden.

Für kontemporäre westliche Musik, deren Ausdrucksform zum Großen Teil durch eine Kombination aus Musik und Text charakterisiert werden kann, ist demnach eine Kombination von text-basierter Sentiment-Analysis als auch Music Emotion Recognition ein interessanter Ansatz. Malheiro et al. (2013) unternahmen einen ersten Versuch, Features zu kombinieren, die auf Lyrics und Audio basieren und konnten gute Ergebnisse für Emotionserkennung zeigen. Ein weiterer Ansatz (Delbouys et al. 2018) benutzt Deep Learning für die Bewerkstelligung der gleichen Aufgabe, mit noch besseren Ergebnissen. Ziel dieser Masterarbeit ist es, eine fehlende Aufgabe in dem Bereich ausführlich anzugehen: eine statistische Gegenüberstellung der beiden Ansätze für Real-World Data zu schaffen. Dies bedeutet, dass eine Emotionserkennung sowohl auf Basis der Lyrics als auch auf Audio benutzt wird, sowie deren Kombination. Die Ergebnisse können zeigen, inwiefern solche Methoden nützlich sind und unter welchen Bedingungen.

Grundlage für diese Arbeit können dabei u.a. die APIs von Spotify, welche eine ausführliche Auflistung von Audio-Features (u.a. auch die Emotions-Parameter Valence / Arousal), sowie von Genius, über welche Lyrics einer großen Menge an kontemporärer westlicher Musik gecrawlt werden kann bieten. Schließlich könnte der Bau eines Sentiment Classifiers, der auf einer Kombination aus Text und Audio die Stimmung von Musikstücken analysiert ein finaler Teil dieser Arbeit sein.

Die bevorzugte Programmierumgebung ist Python.

## Literatur

Malheiro, R., Panda, R., Gomes, P., & Paiva, R.P. (2013). Music Emotion Recognition from Lyrics: A Comparative Study.

Delbouys, Rémi, Romain Hennequin, Francesco Piccoli, Jimena Royo-Letelier and Manuel Moussallam (2018). Music Mood Detection Based on Audio and Lyrics with Deep Neural Net. *ISMIR*

Yang, yi-hsuan & Su, Ya-Fan & Lin, Yu-Ching & Chen, Homer (2007). Music emotion recognition: The role of individuality. Proceedings of the ACM International Multimedia Conference and Exhibition.

## Voraussetzungen

Gute Kenntnisse der Signalverarbeitung und des Maschinellen Lernens (inkl. Neurone Netzwerke); Python. Erwünscht sind Kenntnisse in der qualitativen Emotionsforschung.

## Betreuung

Roman Gebhardt, [rmngebhardt@gmail.com](mailto:rmngebhardt@gmail.com)

Athanasios Lykartsis, EN107, [athanasios.lykartsis@tu-berlin.de](mailto:athanasios.lykartsis@tu-berlin.de)