1. Automatic music mood recognition using Russell’s two-dimensional valence-arousal space from audio and lyrical data as classified using SVM and Naïve Bayes（2019）

音频分类在唤醒度方面取得了较高的准确率，而歌词分类在效价方面也取得了较高的准确率。此外，经过特征选择后，最佳双峰分类器几乎包含了来自最佳歌词和音频分类器的所有特征(1065个可能特征中的1057个特征)。这表明了两个维度特征的重要性

1. A MUSIC BASED MOOD REGULATION SYSTEM USING SENTIMENT ANALYSIS, RUSSELL'S CIRCUMPLEX MODEL AND VECTOR DISTANCE CALCULATION TO IMPROVE THE PRODUCTIVITY OF THE USER（2022）

该模型使用了詹姆斯·拉塞尔(James Russell)开发的罗素情感复杂模型，该模型认为情感分布在一个二维圆形空间中，包含效价和唤醒维度。然后将其作为我们之前创建的基于效价-唤醒的歌曲数据库的输入，以找到最佳匹配的歌曲为用户排队，并逐步使用它们来改善他们的精神状态并达到目标状态

1. Multi-Emotion Classification for Song Lyrics （2021）

另一方面，在歌词上训练的模型达到了与域外数据上训练的模型相当的精度，即使歌词数据集比它们的对应数据集小几个数量级，从线条到歌曲级别进行聚合，从不同的角度进行注释，并且由不同的音乐流派组成。研究结果强调了使用域内数据进行歌曲情感分类的重要性。

1. EMOTION ANALYSIS OF SONGS BASED ON LYRICAL AND AUDIO FEATURES 2015

基于提取的歌词特征和音频特征的加权组合，将歌曲划分为情感类别。这种组合是通过使用大量的统计实验来实现的，这些实验最终收敛到每个特征的所需权重。实验通过调整阈值和kNN算法考虑的邻居数，在准确性和模糊性之间取得了显著的平衡。

1. Emotionally-Relevant Features for Classification and Regression of Music Lyrics 2016

该文主要研究歌词在MER过程中的作用。我们提出了新的风格、结构和语义特征，以及一个包含180首歌词的真实数据集，并根据罗素情感模型进行了人工标注。同时，通过实验识别了特征与情感以及特征之间的关系的可解释规则。关于回归，结果表明，与针对音频的类似研究相比，我们在唤醒方面取得了类似的表现，在效价方面取得了更好的表现。

1. MoodyLyrics: A Sentiment Annotated Lyrics Dataset 2017

我们希望探索基于内容词的歌词情绪标注和基于词典的情绪标注方法在多大程度上与主观、人工或基于标签的标注相一致。我们还创建并展示了MoodyLyrics，这是一个带有情绪标注的歌词的大型极化数据集，将供公众使用。与使用用户标签和人工标注的歌词数据集进行对比，该方法的准确率达到74.25%，在音乐领域具有较好的性能。这一结果在一定程度上证明了，在没有主观反馈的情况下，音乐作品的情绪标注也是一个可以解决的问题

7.Music Mood Prediction Based on Spotify’s Audio Features Using Logistic Regression 2022

多年来，音乐流媒体平台因其音乐传播方式和音乐队列为听众量身定制而变得流行。在这些应用中，机器学习一直用于音乐推荐。本文将探索一种基于Spotify PH市场上可用的Spotify应用程序接口(API)音频组件来建模曲目情绪的创新方法 使用了矢量距离计算与spotify的valence和enrgy的音频特征和拉塞尔的情感模型来更准确的定义情感分类

7.LyBERT: Multi-class classification of lyrics using Bidirectional Encoder Representations from Transformers (BERT) 2022

通过迁移学习特征并结合预训练模型BERT。迁移学习之后，然后将BERT模型应用于数据集，以提高模型的准确性。总体精度达到92

8.Russell, J. A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161-1178.

詹姆斯·拉塞尔提出的罗素情感复杂模型，在多媒体和心理学领域已经被广泛采纳。该模型基于一个二维空间，由效价（valence）和唤醒（arousal）两个维度构成。

9.Sentiment classification: The contribution of ensemble learning

实验结果表明，集成学习方法可以获得比基础学习器更好的结果。在20种方法中，随机子空间SVM的准确率最高。

10. Sentiment analysis: Bayesian Ensemble Learning

[在本文中，我们追求集成学习](https://www.sciencedirect.com/topics/computer-science/ensemble-learning)的范式，以降低与语言歧义相关的噪声敏感性，从而提供更准确的极性预测在黄金标准数据集上的实验结果表明，所提出的方法优于传统分类和集成方法。

11. MUSIC EMOTION RECOGNITION: A STATE OF THE ART REVIEW

本文探讨了音乐情感识别的广泛研究，特别关注使用上下文文本信息(如网站、标签和歌词)的方法和基于内容的方法，以及结合多个特征域的系统。

12. [Music emotion recognition: A state of the art review](https://archives.ismir.net/ismir2010/paper/000045.pdf)

13. Fusing audio, visual and textual clues for sentiment analysis from multimodal content

我们使用特征级和决策级融合方法来合并从多种模式中提取的情感信息。达到了很高的准确率。

14. Music emotion recognition: The combined evidence of MFCC and residual phase

总结

在过去的几年里，音乐情感识别（Music Emotion Recognition, MER）[11]成为音乐信息检索（Music Information Retrieval, MIR）[12]的一个重要研究方向。MER的目标是理解和识别音乐作品中所表达的情感和情绪，从而更好地推荐和分类音乐。这一过程在拉塞尔的情感模型[8]的指导下得到了广泛的研究和应用。

在单一特征的研究中，歌词和音频特征通常被单独用来进行音乐的情感分析。近年来的文献研究显示，歌词在MER中具有重要作用。例如，[5]一项研究提出了针对歌词的新的风格特征（StyBF）、结构特征（StruBF）和语义特征（SemBF），并通过这些特征成功提高了F-measure的评估指标。"MoodyLyrics"[6]通过创建一个带有情绪标注的歌词数据集，探索了基于内容词和基于词典的歌词情绪标注方法的准确性。研究“LyBERT”[7]通过使用迁移学习特征并结合预训练模型BERT实现了歌词的多类别分类。

在音频特征此方向的研究通常关注音乐的声学属性，[14]组合了梅尔频率倒谱系数（MFCC）和残差相位（RP）特征的方法应用非线性支持向量机学习算法来获得了各种情绪之间的最佳分类。[1]这篇文章借助支持向量机（SVM）成功地对音乐的唤醒度进行了准确的分类。另外一篇文献则采用了不同的方法。[7]这个研究中使用了Spotify音频特征并结合拉塞尔的情感模型[8]和矢量距离计算来更准确地定义音乐的情感分类。

然而，尽管这两种特征在单独应用时都展示出一定的准确性，但它们各自也有局限性。[4]探讨了如何基于歌词和音频特征的加权组合将歌曲划分为情感类别。另一篇文献[1]表明，通过整合音频的唤醒度特征和歌词的效价特征，可以实现更精确的情感分类。在集成学习方面，文献如[10]强调了集成学习方法在提高情感分类准确率方面的作用。另外，[13]探讨了特征级和决策级的融合方法在情感分析中的应用。因此，我计划使用集成学习来综合这两个分开的模型，即歌词特征模型和音频特征模型，以期望进一步提升情感分类的准确性。通过整合这两种特征，我们可以补充它们各自的局限性，并期望得到一个更全面和准确的音乐情感识别模型。