音乐，这种古老而普遍的艺术形式，在人类历史中发挥着重要作用。音乐可以同时包含与表达情感，有多种理论探索这种关系（Juslin，2001；Budd，2002）。这种联系得到了当代实证研究的进一步支持，该研究确定了三种类型的音乐情感：体现的、认知的和联想的（Przybysz，2013）。最近的一项研究发现，无论个人的音乐偏好如何，音乐结构都可以触发特定的情绪（Silva，2021）。这些发现共同强调了音乐对我们情感体验的强大影响。无论是在欢庆的场合还是在哀悼的时刻，音乐都以其独特的方式触动着我们的心灵。它能够激发起我们内心深处的情感，无论是喜悦、悲伤、爱情还是怀旧。在不同的文化和社会中，音乐都扮演着核心的角色，它不仅是情感表达的媒介，更是连接人与人之间情感的桥梁。而且从神经科学的观点表明音乐也可以影响复杂的神经生物学过程，并可用作各种精神疾病的替代疗法(Shuai-Ting Lin2011)。这些发现强调了音乐作为促进心理健康工具的潜力。在过去的十年中，随着大量易于获取的数字音乐库的指数式增长，对音乐及其相关数据进行有效组织和搜索的挑战也日益凸显。音乐信息检索（MIR）作为这个领域的科学领域正迅速向搜索和组织音乐及相关数据的自动化系统发展。虽然一些常见的搜索和检索类别，如歌手或流派，更容易量化为一个“正确的”(或普遍同意的)答案，但音乐本身的情感表达，它可能是高度主观的，难以量化ykim(2010)。所以其情感内容的主观性和复杂性意味着传统方法可能不足以满足日益增长的需求。而通过情感搜索音乐是用户[1]使用的主要标准之一，所以来自Spotify和Last.fm等网站的真实音乐数据库每天都会变得越来越大，这需要大量的人工工作来保持更新。其中歌词和音频作为音乐的两个最主要的组成部分，音乐艺术家通常会通过这两个方面来传达情感维度，而不同的方面可能会传达出截然不同的情感维度和强度。在今天这个数据驱动的环境下，机器学习和深度学习的快速发展为解析音乐中的复杂情感内容开辟了新的可能性。这些先进的技术允许我们自动化地处理和分析大规模的音乐数据集，揭示音乐创作中隐藏的情感层面。特别是在处理包含丰富情感表达的歌词和复杂音频特征。

\section{问题及现有解决方案}

尽管音乐情感分析取得了进展，但现有方法在全面捕捉音乐情感本质方面往往存在局限性。许多研究主要集中于抒情内容或音频功能，忽略了将两者结合起来的协同潜力。这种疏忽可能导致对音乐的情感影响的部分理解，限制了在个性化音乐推荐、治疗干预和音乐教育中的情感分析等领域应用的有效性。

\section{目标和贡献}

该项目旨在应用机器学习、深度学习和自然语言处理技术开发一个综合模型，分析歌词和音频数据，以更深入地理解音乐中的情感表达。目标是超越现有的情感分类基准，增强模型的泛化能力。该模型不仅有望为音乐情感分析提供一个新的视角，而且有望在音乐推荐系统、情感治疗、音乐教育等领域产生积极的实际影响，为音乐情感分析提供更详细、更全面的途径。

\vspace{5mm}

\noindent

复现现有研究：成功复现了现有文献中提出的模型，这是验证和基础工作的重要步骤，为后续的改进和发展奠定了基础。

单歌词模型的改进：在复现现有研究基础上对原有模型的基础上进行了关键的优化，包括嵌入技术，预处理步骤和精细模型结构和参数的调整，使单歌词模型的性能超越了原论文的基线，展现了模型改进的有效性

综合模型的开发和优化：通过融入音频特征，开发了一个更为鲁棒且泛化能力更强的综合模型，该模型同样超过了相关领域第二篇论文的基线。

创建带情绪的音乐数据集：使用优化后的模型，创建了一个涵盖近10年top100歌曲并附有情绪标签的数据集，为音乐推荐系统和情感分析提供了宝贵的数据资源。

**实际应用评估**：为了进一步验证模型的有效性，我将模型的预测结果与近年来的世界大事的真实数据进行了对比分析。这一评估过程展示了模型在识别和反映与重大事件相关的情感趋势方面的能力，增强了研究的实用性和相关性。

近年来，音乐情感识别(MER)[1]逐渐成为研究热点，已成为音乐信息检索(MIR)[2]领域的重点研究领域之一。MER致力于对音乐作品中所表达的情感和情绪进行深入洞察和准确识别，实现音乐的个性化推荐和准确分类。

在Russell的情感模型[3]的指导下，该领域得到了持续而广泛的关注和研究。[3]提供了更细致的情感分类。

音乐情感识别的深度探索不仅推动了个性化音乐系统的快速发展，而且显示出巨大的社会价值和商业应用潜力。从技术角度来看，音乐情感识别(MER)的研究主要集中在利用歌词和音频的特征来预测音乐的情感属性。在歌词分析方面，[4]提出利用歌词的风格(StyBF)、结构(StruBF)和语义(SemBF)特征来提高情感分类的准确率。另一项研究“MoodyLyrics”[5]提出了一种基于内容词情感特异性的歌词情感标注方法。同时，“LyBERT”[6]利用迁移学习和BERT模型对歌词进行详细的多类别情感分类。此外，一项研究[7]使用朴素贝叶斯对音乐的积极度进行分类，发现基于歌词的分类在积极度方面表现出很高的准确率。这些研究证明了歌词在音乐情感研究中的重要价值。另一方面，在音频特征领域，研究[8]利用MFCC和残差相位特征，采用支持向量机(SVM)对音乐情感进行分类。另一项研究[7]使用SVM对音乐的唤醒度进行分类，发现音频分类在唤醒度方面也取得了很高的准确率。此外，还有研究[9]使用向量距离计算，结合Spotify的价值和能量特征值，并参考Russell的情感模型来更准确地定义音乐的情感分类。另外，另一项研究[10]使用Spotify的音频特征评估了逻辑回归在预测歌曲情感方面的表现，得到了较高的准确率。这些研究进一步证实了音频特征在情感研究中的重要价值。虽然目前单一特征在MER中已经显示出一定的有效性，但随着研究的深入，其局限性逐渐显现。

因此，大多数研究者认为融合多种特征是提高音乐情感识别准确率的重要途径。例如，[11]和[7]的研究表明，结合音频和歌词特征可以增强情感分类，突出了从两个维度融合特征的价值。此外，集成学习作为一种有效的方法，也被证明可以显著提高分类精度。[12]中强调集成学习在提高情感分类准确率方面发挥着重要作用，而[13]探索了特征级和决策级融合方法在情感分析中的应用价值。基于这些分析，本项目旨在有效融合歌词特征、音频特征和集成学习方法，开发更全面、更精确的音乐情感识别模型。