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。

作者使用了NB KNN SVM CNN LSTM 和BILSTM 每个都介绍一下 nb的参数是a=0.05 knn是= n = 29. svm没有说 cnn是The parameter that we

used were three one dimensional CNN layer with filters

= 128, kernel size = 5 and activation function = ReLU.

We used three max pooling layer with pool size = 5 and

output layer with activation function = Softmax. lstm和bilstm都是Parameter Value

Dimension 100

Number of Bi-LSTM Hidden Units 100

Maximum Sequence 1000

Batch Size 64

Activation Function Softmax

Optimizer ADAMpoch = 20 and learning rate

= 0.0006 每个方法总结成一段

朴素贝叶斯(Naive Bayes)是一种基于贝叶斯定理的基本分类技术，以其高效和有效而闻名，尤其是在处理大型数据集时。它假设预测变量之间相互独立，并以其处理分类数据的性能而闻名。朴素贝叶斯分类器的本质在于，根据贝叶斯定理，给定一组独立的预测变量，计算某个类别出现的概率。通过选择具有两个概率最高积的类来计算最可能的类c:类P(c)的先验概率和文档P(d|c)的似然: cˆ = argmax c∈C likelihood z }| { P(f1, f2,...., fn|c) prior z}|{ P(c)

在这篇论文中，我们使用了朴素贝叶斯的一种特殊变体，称为多项朴素贝叶斯(Multinomial Naive Bayes)，它特别适合使用单词计数等离散特征进行分类。多项式朴素贝叶斯的计算公式如下:

\begin{equation}

P(C\_k | \mathbf{x}) = \frac{P(C\_k) \times \prod\_{i=1}^{n} P(x\_i | C\_k)^{a\_i}}{P(\mathbf{x})}

\end{equation}

其中\( \mathbf{x} \)是一个特征向量，\( a\_i \)表示特征\( i \)在待分类实例中的出现频率。分子\( P(C\_k) \times \prod\_{i=1}^{n} P(x\_i | C\_k)^{a\_i} \)计算观察类\( C\_k \)中特征的可能性，而分母\( P(\mathbf{x}) \)用作归一化因子。在复现的论文中，平滑参数\( \alpha \)被设置为0.5，以调整在学习样本中不存在的特征，并防止进一步计算中出现零概率。