分类号：
学　号：20222108043

密　　级：公开
单位代码：10759

石河子大学

硕士学位论文

基于深层语言分析的小学作文辅助评分研究与系统实现

学位申请人　李俊杰
指导教师　于宝华 教授
申请学位类别　专业硕士
专业名称　电子信息
研究领域　计算机技术
所在学院　信息科学与技术学院

中国·新疆·石河子
2025 年 6 月

分类号：　　密　　级：公开
学　号：20222108043　　单位代码：10759

石河子大学

硕 士 学 位 论 文

基于深层语言分析的小学作文辅助评分研究与系统实现

学 位 申 请 人　李俊杰
指 导 教 师　于宝华 教授
申 请 学 位 类 别　专业硕士
专 业 名 称　电子信息
研 究 领 域　计算机技术
所 在 学 院　信息科学与技术学院

中国·新疆·石河子
2025 年 6 月

Research and System Implementation of Primary School Composition Assistant Scoring Based on Deep Language Analysis

A Dissertation Submitted to

Shihezi University

In Partial Fulfillment of the Requirements

for the Degree of

Master of Engineering

By

Li Jun-jie

(Electronic Information)

Dissertation Supervisor:Prof. Yu Bao-hua

June , 2025

摘要

语言文字是人类社会重要的交流工具和信息载体，语言文字的教育教学对于增强中华民族文化认同、维护民族团结和国家统一具有特殊意义。小学语文教学是夯实基础和良好习惯养成的重要阶段，其中写作能力反映了学生的语言表达与逻辑思维，对语言文字能力的培养尤为重要。传统的小学语文写作评阅方式存在工作强度大、效率低、评改反馈周期长、缺乏个性化指导意见等问题。基于此，本研究以新课标对学生“表达与交流”的能力要求为理论分析框架，构建多维度特征融合的作文辅助评分模型。该模型通过融合字词规范度、句式多样性、修辞手法运用及写作连贯性等评价维度，旨在开发智能化的作文辅助评分系统，为突破传统评阅模式的局限性提供技术解决方案。本研究的主要工作如下：

（1）针对目前 AES 的研究方法未能兼顾浅层语言学特征和深层语义特征的问题，本文提出了一种浅层语言学特征与修辞特征融合的句子质量分类模型。首先采集并处理了小学句子写作数据集与小学修辞手法数据集，设计并提取了 112 个句级特征作为句子写作的浅层语言学特征表达。其次，为解决主流文本分类模型修辞手法识别能力弱的问题，本文选取循环卷积神经网络（TextRCNN）提取修辞特征，在网络中引入基于预训练的动态词向量嵌入层，进一步加强其对上下文语义的理解能力，从而提升模型对句子深层次语义特征的感知，进而判断句子是否使用修辞。最后，将句级浅层语言学特征与修辞特征相融合，提取出与句子质量高度相关的特征并用于句子质量分类模型的训练。实验结果表明，所提出的改进模型在修辞手法识别任务上的微平均 F1 值达到 91.54%，加权平均 F1 值达到 90.40%，句子质量分类模型准确率达到 74.88%，验证了其在句子质量分类任务中的有效性。

（2）为解决文本连贯性分析模型对主题信息不敏感的问题，本文提出了基于迁移学习的写作连贯性分析模型。该模型采用改进的负样本生成策略，以解决小学作文数据集数据不平衡问题，通过对轻量化预训练模型（ALBERT）微调的方式扩展任务目标，使其具备主题一致性与语义连贯性综合评估的能力。结合上文研究成果，参照新课标中小学阶段语文能力的要求，构建了多粒度特征融合的作文辅助评分模型，该模型设计了基础语言能力、修辞手法运用、写作连贯性三个能力评价参数，运用专家调查法得出与本地区教学水平相匹配的参数权重。实验结果表明，经过微调的写作连贯性分析模型能很好地适应任务目标，在自建小学作文数据集上 ACC 为 91.98%，AUC 为 96.61%，在公开的小学作文数据集上 ACC 为 92.99%，AUC 为 96.63%，均优于现有同类模型。经过对比专业教师评分，作文辅助评分模型能给出清晰合理的建议评分，验证了该模型的有效性。

（3）小学作文辅助评分系统的实现。本文设计并实现了一个 Web 端与移动端结合的小学作文辅助评分系统，该系统实现了包括系统管理、作业管理、作业提交、辅助批改、评阅结果查看等功能。通过系统试运行表明本系统能够有效提高小学作文写作批改效率，弥补了传统批改模式下缺乏个性化指导的短板。

关键词：作文辅助评分；修辞手法识别；写作连贯性；多粒度特征融合

Abstract

Language is an important communication tool and information carrier of human society. Language education and teaching is of special significance for strengthening the cultural identity of the Chinese nation, maintaining national unity and national unity. Chinese teaching in primary schools is an important stage to consolidate the foundation and cultivate good habits. Writing ability reflects students' language expression and logical thinking, which is particularly important for the cultivation of language and writing ability. The traditional way of evaluating primary school Chinese writing has many problems, such as high work intensity, low efficiency, long feedback period of evaluation, and lack of personalized guidance. Based on this, this research takes the new curriculum standard's requirements for students' ability of "expression and communication" as the theoretical analysis framework, and constructs a multi-dimensional feature fusion composition auxiliary scoring model. This model aims to develop an intelligent composition auxiliary scoring system by integrating evaluation dimensions such as word standardization, sentence diversity, rhetorical devices and writing coherence, and to provide a technical solution to break through the limitations of the traditional scoring model. The main work of this study is as follows:

(1) In view of the problem that the current research methods of AES fail to take into account the shallow linguistic features and deep semantic features, this thesis proposes a sentence quality classification model that combines the shallow linguistic features and rhetorical features. This thesis first collected and processed the primary school sentence writing data set and the primary school rhetorical device data set, designed and extracted 112 sentence level features as shallow linguistic features of sentence writing. Secondly, in order to solve the problem that the mainstream text classification model has a weak ability to recognize rhetorical devices, this thesis selects a circular convolutional neural network (TextRCNN) to extract rhetorical features, and introduces a dynamic word vector embedding layer based on pre training in the network to enhance the model's perception of the deep semantic features of sentences, and then judge whether sentences use rhetoric. Finally, the sentence level shallow linguistic features and rhetorical features are combined to extract features highly related to sentence quality and used for the training of sentence quality classification model. The experimental results show that the micro average F1 value of the proposed improved model reaches 91.54%, the weighted average F1 value reaches 90.40%, and the accuracy of the sentence quality classification model reaches 74.88%, which verifies its effectiveness in the sentence quality classification task.

(2) To solve the problem that the text coherence analysis model is not sensitive to topic information, this thesis proposes a writing coherence analysis model based on transfer learning. The model uses an improved negative sample generation strategy to solve the problem of data imbalance in primary school composition data sets, and expands the task objectives by fine-tuning the lightweight pre training model

(ALBERT), so that it has the ability to comprehensively evaluate the theme consistency and semantic coherence. Combined with the above research results and referring to the requirements of the language ability of primary and secondary schools in the new curriculum standard, this thesis constructs a multi granularity feature fusion composition auxiliary scoring model. This model designs three ability evaluation parameters, namely, basic language ability, rhetoric use, and writing coherence. The expert survey method is used to obtain the parameter weights that match the local teaching level. The experimental results show that the finely tuned model can well adapt to the task objectives. ACC is 91.98% and AUC is 96.61% on the self built primary school composition data set, and ACC is 92.99% and AUC is 96.63% on the open primary school composition data set AICFE, all of which are better than existing similar models. At the same time, by comparing the scores of professional teachers, the essay assisted scoring model can give clear and reasonable suggestions and scores, which verifies the effectiveness of the model.

(3) The realization of the auxiliary scoring system for primary school composition. This thesis designs and implements an auxiliary scoring system for primary school composition, which combines the Web end with the mobile end. The system implements functions such as system management, homework management, homework submission, auxiliary correction, and review results. The trial operation of the system shows that the system can effectively improve the efficiency of primary school composition writing correction, and make up for the lack of personalized guidance under the traditional correction mode.

Key words: Composition assisted scoring; Rhetoric recognition; Writing coherence; Multi granularity feature fusion

石河子大学学位论文独创性声明及使用授权声明

学位论文独创性声明

本人所呈交的学位论文是在我导师的指导下进行的研究工作及取得的研究成果。据我所知，除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确的说明并表示谢意。

研究生签名：  时间： 2025 年 05 月 26 日

使用授权声明

本人完全了解石河子大学有关保留、使用学位论文的规定，学校有权保留学位论文并向国家主管部门或指定机构送交论文的电子版和纸质版。有权将学位论文在学校图书馆保存并允许被查阅。有权自行或许可他人将学位论文编入有关数据库提供检索服务。有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

研究生签名：  时间： 2025 年 05 月 26 日

导师签名：  时间： 2025 年 05 月 26 日

目录

第 1 章 绪论

1.1 研究背景及意义

语言文字的运用是我国义务教育阶段语文课程的主要组成部分，其中写作更是教学的核心与重点。教语用〔2017〕1 号《教育部国家语委关于进一步加强学校语言文字工作的意见》指出“提高语言文字应用能力是学校培养高素质人才的基本内容。语言文字应用能力的培养要从小抓起，良好的口语、书面语表达水平和语言综合运用能力，是国民综合素质的重要构成要素，在个人成长成才过程中具有不可替代的作用。”[1]，在《义务教育语文课程标准》[2]（2022 年版）在课程目标中对不同学段都提出了发展书面语言运用能力的要求，以上政策的颁布表明国家十分重视语言文字的教育教学工作，尤其是小学阶段的写作教学对于学生语言运用能力培养至关重要。

写作是语言应用能力培养的核心，其能力提升需要经过大量训练，同时也离不开教师的悉心指导。然而传统的教学模式需要教师投入大量的精力进行作文评阅，存在效率低、评改反馈滞后和评价标准不统一等现实问题。作文自动评分（Automated Essay Scoring, AES）是指利用自然语言处理（Natural Language Processing, NLP）技术对学生的作文进行自动化评估和打分的过程[3]。AES 研究经历从基于机器学习的回归与分类方法[4]到基于深度学习的神经网络方法[5]的发展，现阶段针对 AES 系统的研究主要是通过提取作文的语言学特征与深度语义特征，利用深度学习方法对作文进行建模，最终得到作文的预测得分。AES 技术能克服传统批改方式的局限性，挖掘文本深层次语义信息，从而提高 AES 系统的整体效能。随着研究的不断深入基于深度学习的 AES 系统在评分一致率、平均分差、相关度等指标方面都已取得较大进展。

中文作为一种表意文字，不仅拥有庞大的字符集和复杂的字词关系，还包含着丰富的成语、俗语、惯用语等文化元素。此外，在写作手法上中文往往蕴含深厚的文化底蕴和情感色彩，写作时常用到比喻、拟人、排比等修辞手法。这种深层次的语言表达方式增加了 AES 系统对中文语义理解的难度，也影响了作文评分的准确性。从语言建模的角度来看，中文相较于英文在文字结构、语义特征、语法规则、表达方式上均存在明显的差异，使得已有的研究成果难以直接应用到中文作文的评分环节，因此，如何构建合理且有效的中文作文评分方法，捕获作文中具有区分度的特征，并利用深度学习方法提取作文中深层语义信息，从不同维度综合评估作文是解决中文作文自动批改的重点和难点。

1.2 国内外研究现状

在语言应用能力的培养过程中，写作占据着极为重要的地位，对写作能力的训练，不仅能促进学生思想、心理、文化素质的完善与发展，也能提高学生的表达沟通能力[6]。然而，随着学生数量的剧增，教师评阅作文的负担日趋繁重，传统的评阅模式已经难以满足需求。近年来，人工智能发展迅速，深度学习等技术逐渐成熟，使得作文自动评分 AES 已经成为自然语言处理 NLP 在教育领域最重要的应用之一[7]。本章将从基于浅层语言学特征的方法与基于深度学习的方法两个方面对作文自动评分领域的研究进行梳理。

1.2.1 基于浅层语言学特征的作文自动评分研究现状

基于浅层语言学特征的 AES 需要针对特定场景精心设计特征，通过这些特征，能直观地对作文进行评估和分析，其分析过程相对简洁，能快速地给出作文在语言层面的评估反馈，有助于教师对学生在写作方面存在的问题给出准确、有针对性的评估。但基于特征的作文评分方法往往只关注于作文表面的语言现象，对作文质量分析不够全面和深入，在某些场景下难以准确地反映作文真实水平。

（1）基于浅层语言特征的方法

最早的基于特征的 AES 为美国杜克大学 Ellis Page 教授等人研发的 PEG（Project Essay Grade）[8]，该系统设计并提取了包括单词长度、作文长度、标点数量、生僻字数量等浅层语言学特征作为评分标准，但由于缺乏文章结构、内容语义等方面的因素，很容易出现文章内容很长，毫无逻辑，却可获得较高分数的问题。受限于上个世纪算法与硬件性能的限制，早期的 AES 系统大部分基于手动提取特征。Burstein 等研发了一款面向大规模考试的作文评分系统 E-rater[9]，E-rater 通过标注词性，分析文本句法结构、文本结构等特征，同时考虑主题贴切程度、对照人工评分的标准来建立作文评分模型，虽然该系统更接近人工评分的思路，但是对深层次语义特征的探究较少，对于一些模板类文章时无法给出准确评分。

梁茂成等[10]利用基于特征的方法对作文进行建模，提取了语法错误、句法结构等特征，采用线性回归的方式对作文进行评分。梁茂成同时对国外成熟的自动评分系统 PEG 和 IEA 系统进行深入分析，提炼出二者对作文评分效果影响系数较高的语言学特征，并对国内学生写作评分进行了回归分析[11]，但该系统依然存在对深层语义特征利用不足的问题，系统整体的准确性有所欠缺。彭星源等[12]对作文中出现的词汇进行评级，并分级汇编成词汇质量等级表，并对中文论文进行了自动评分。周明等[13]从分别提取了词汇、句法和结构作文特征，使用线性回归方法构建了自动评分模型。余立清[14]针对特殊场景设计了多种人工特征，进而构建了一套作文自动评分系统。彭丽莎等[15]对传统中文作文

评分的研究做了总结和改进，设计了一套适用于中考语文作文自动评分系统，该系统分别从文章的文字、词汇、内容与语言多个维度分析，提取中考语文作文的特征项，设计并构建了中考语文中文辅助评分系统。主要用于为阅卷教师提供作文的参考评分，该系统在特定领域取得了较好的效果。

（2）联合神经网络模型与浅层语言学特征的方法

近些年的研究表明联合使用神经网络模型与语言学特征能有效提升 AES 系统的性能。刘浩坤[16]提出了扣题度等特征，并融合了基于词嵌入表示的语义特征，提出了一种多模型融合方法进行自动作文评分。崔建鹏[11]构建了基于图神经网络的作文评分模型，将语句通顺度、文本匹配度融合到神经网络中来提取作文特征。Cozma 等[4]将字符级特征与词嵌入表示特征融合从而提取语义特征，通过实验验证，有效提升了模型性能。Liu 等[17]提出了一种两阶段模型（Two-Stage Learning Framework, TSLF），该方法利用神经网络技术，分析和提取文章的语义、流畅度以及关联性特征，并将这些自动提取的特征与手动识别的特征相结合，以实现对文章自动评分。Farag 等[18]将句子之间的连贯性特征与深度学习模型相融合，有效提升了作文评分模型的性能。周险兵等[19]采用 Doc2Vec 模型对主题信息进行提取，同时结合句法特征等人工特征以及卷积神经网络（Convolutional Neural Network, CNN）和长短期记忆网络（Long Short-Term Memory, LSTM）提取的语义信息，对作文进行评分。

综上所述，基于浅层语言学特征的 AES 系统主要依赖于人工设计的特征，特征选取的角度与数量直接影响到最终的评分效果，容易出现“低质量高分”的极端情况。另外，此类方法未将深层次语义特征融入作文评分中，虽然在特定场景下具有一定的使用价值，但是离大规模应用还有大差距。目前，基于特征的 AES 系统通过结合神经网络模型在特殊场景下的作文评分任务中具有明显优势，联合使用神经网络模型与浅层语言学特征已经成为当前研究的热点。

1.2.2 基于深度学习的作文自动评分研究现状

基于浅层语言学特征的作文自动评分方法通常只提取作文的浅层语言学特征，难以捕获作文的深层次语义特征。近年来，随着硬件设备算力的快速提升，深度学习方法逐渐成为主流，基于深度学习的作文自动评分技术也取得了许多研究成果。

（1）基于优美表达的方法

在小学阶段的语文作文评分过程中，修辞手法等优美表达是作文的加分点之一。黄凯[20]提取主题特征并利用卷积神经网络 CNN 捕捉句子优美性特征，综合二者并对作文评分。黄志娥[21]从字、词、语法与优美程度等角度中选取了 100 多个作文特征，对 HSK 作文语料库的作文进行研究，得到了 19 个相关度较高的作文特征，解释了作文长短、

词汇量、语言优美程度对作文评分的影响力。刘明杨等[22]在作文自动评分中融入文采特征，通过构建修辞分类体系、设计启发式规则实现了修辞识别，实验表明评分性能有显著提升。付瑞吉[23]等使用深度学习的方法对中学生中文作文优美句进行了识别研究，并将优美句识别相关特征加入作文自动评分任务中，降低了计算机评分与人工评分之间的误差。穆婉青等[24]使用卷积神经网络 CNN 融合结构相似度算法对文学作品以及高考阅读材料中的排比句进行了识别。文治等[25]将情感判别与反问句的修辞手法结合起来用以识别中英文中的修辞手法。赵晓妮等[26]将修辞手法识别技术与情感分析的研究相结合，实验证明修辞识别能有效提升模型对情感的判别准确度。石昀东等[27]运用深度学习方法、建立引用库等方法对修辞手法进行特征提取，并运用于小学作文的自动分类任务中。

（2）基于文本连贯性的方法

在作文评分过程中，连贯性也是关键的评估指标之一。基于深度学习的作文自动评分方法不仅考虑文章的语义表征，文章的主题相关性[28]与连贯性[29]等都是该领域中的常用特征。Tay 等[29]为增强长短期记忆网络对句子向量的编码能力，提出了 skipflow 机制，利用句子的向量相似度对整篇文章进行建模，进而获取文章的连贯性特征来辅助评分。针对传统文档表征忽视篇章连贯的问题，Mim 等[30]提出基于文本自监督学习的无监督框架，通过建模句子间语义连贯性自动捕捉文章“衔接－连贯”特征。Farag 和 Yannakoudakis[18]提出了一种基于多任务学习的多层级神经网络，在底层网络的子任务中学习单词语法角色，在顶层网络中预测篇章级别的连贯性得分。通过引入子任务，底层的单词级别任务显著提高了模型预测连贯性得分的准确性。Aralikatte 等[31]搭建了 Transformer 编码层对输入文本编码，构建文档浅层语义网络图，采用强化学习与奖励机制，有效提升模型对文本的全局连贯性预测准确率。Jwalapuram 等[32]引入自监督学习机制，区分连贯文本和非连贯文本，通过自适应抽样策略挖掘非连贯样本，并使用动态编码器组建编码层提取文本特征。苏娜[33]构建了文本语义结构树分析模块，将文本连贯性问题转化成对语义结构树的分析，其次逐层抽取特征，再通过最大熵分类模型对篇章结构和语义关系进行识别，验证了该技术的有效性。贺亚琼等人[34]以篇章连贯性为评估汉语议论文质量的重要因素，提出分别从语义表达、组织结构和整体逻辑来评估文本的连贯性方法。

（3）基于整体评分的方法

Hussein 等[35]构建了文章整体评分神经网络模型，该模型通过组合卷积神经网络和长短期记忆网络对文章进行整体评分，并采用了多任务学习的方法分别获得文章整体评分和各维度评判标准评分，综合二者评分最后预测文章的最终得分。Kumar 等[36]使用多任务学习的方法，使用 Dong 等[37]提出的模型，首先将作文整体评分与评判标准评分相分离，分别作为评分模型的主任务与辅助任务，其次将各评判标准评分转化为特征与整体评分的文章特征表示进行拼接，最后输入线性层得到作文最终的整体评分。

Howard 等[38]提出通用语言模型微调框架，通过迁移学习显著提升文本分类任务性能。宋巍[39]等人构建三阶段训练范式。该方法通过分层知识迁移增强模型领域适应性，实验表明其能有效提升现有评分器的跨主题泛化能力。该方法避免复杂语法解析。在数据量较少的情况下 Mayfield 等[40]对 BERT 进行微调提高了模型在作文评分任务中的适用性和性能。Yang 等[3]提出融合回归和排序损失微调 BERT 模型。Xue 等[41]使用基于预训练的双向变换器模型（Bidirectional Encoder Representationfrom Transformers, BERT）模型作为特征提取器，并使用多个全连接层来对不同维度的论文特征进行评分。Sun 等[42]提出了一种将 BERT 与提示相结合的方法，通过 BERT 生成与输入文本相关的提示以及将提示与短文相匹配，帮助 BERT 获得更好的特征。Wu 等[43]提出了一种将 BERT 与主题建模相结合的方法，通过识别作文中的关键主题句，帮助 BERT 建模主题信息。Rodriguez 等[44]将比较了近几年自然语言处理领域最为先进的语言模型 BERT 和 XLNet，将其应用于作文自动评分领域并取得了较好的性能。于明诚等[45]使用 XLNet 捕获上下文语义信息，再通过计算句向量与作文主题语义相似度提取篇章主题层次特征，该方法在作文自动评分任务中表现出较好的性能。

（4）基于大语言模型的方法

大语言模型（Large Language Model, LLM）的出现显著扩展了人工智能的能力边界和应用范围，推动了人工智能大众化和平民化时代的到来[46]。2022 年发布的 ChatGPT（Chat Generative Pre-trained Transformer）是基于 Transformer 架构的生成式预训练大语言模型，能够输出复杂度较高的类人语言，适应不同自然语言处理任务并提供反馈，可用于故事续写、摘要生成、写作反馈等应用[47]。作为新兴的人工智能技术，ChatGPT 在写作评估与反馈领域的应用也备受关注[48]。相比 BERT，ChatGPT 更适用于语言生成任务[49]。然而，ChatGPT 的参数数量达到千亿级别，经过海量的文本预训练之后，ChatGPT 的语言理解能力非常强大。基于其强大的语言理解力，Mayer 等[50]对商务邮件的礼貌性进行分类研究，结果显示 ChatGPT 可以达到与人类评级相似的准确性水平。表明大语言模型用于 AES 的潜能不应被忽视[49]。

随后，Mizumoto 等[49]基于 TOEFL11 语料库，采用雅思写作任务二的评分标准，探究了 GPT-3 在写作评分方面的效果。研究结果发现，GPT-3 的评分与人工打分的一致性为仅为 40%左右，只能达到人工评分的中等效果。此外，将 GPT-3 生成的分数融入包含语言特征的回归模型中所带来的性能改善效果有限。Hackl 等[48]通过评估宏观经济学领域的学生作业，探索了 GPT-4 的文本评估能力。研究结果表明，GPT-4 在不同时间点和风格变化下能够一致地应用评价标准，显示出 GPT-4 可以作为 AES 工具减轻教师负担。Yancey 等[51]测试了 GPT-3.5 和 GPT-4 预测二语写作水平的能力，发现 GPT-4 可与人工打分一致性较高，能达到 85%左右，但是不同的语言指令可显著改善模型。由此可见，GPT-4 对于文本评估以及二语写作水平评估具有较高的性能。

综上所述，近年来深度学习技术与教育领域的结合日趋紧密，基于深度学习的 AES 系统能有效提高作文自动评分的性能，越来越多的研究者引入额外的特征从不同维度对作文进行建模，例如语义连贯性、主题一致性、优美表达等。另外，随着预训练语言模型的应用，大量研究利用其基于上下文的语义理解机制，在不同任务场景下对模型进行微调，都取得了不错的效果。但由于中文作文评分任务的复杂性，仍然需要通过设计语言学特征融入评分模型，以增强作文评分模型的泛化性能。此外，现阶段有关在大语言模型在教育及测评领域的应用都是大样本场景下的应用，在课堂环境下的小样本量写作评估能力如何，仍需进一步检验。因此，如何将深度学习方法与语言学特征有效结合，合理地利用大语言模型赋能作文测评，进而满足教育场景需求，构建兼具高精度、可解释性和教学实用性的 AES 系统是重要的研究方向。

1.3 主要研究内容

1.3.1 本文研究内容

本研究聚焦于小学作文辅助评分方法研究，本文先对该领域国内外研究进展进行了系统的梳理与深入分析，为解决目前小学语文作文辅助评分研究较少、针对性不强的问题，本文参照新课标中对小学语言能力培养的要求，分别从句子质量、修辞运用、写作连贯性三个维度出发，构建了多维度特征融合的作文辅助评分模型，为小学语文作文评分提供参考，具体研究内容如下：

（1）为解决目前 AES 方法存在未能兼顾浅层语言学特征与深层语义特征的问题，本文提出了一种浅层语言学特征与修辞特征融合的句子质量分类模型。首先采集并处理了小学句子写作数据集与小学修辞手法数据集，参照新课标设计了字、词、句 3 个维度共 112 个句级特征作为句子的浅层语言学特征表达。针对文本分类模型修辞手法识别能力弱的问题，通过改进 TextRCNN（Recurrent Convolutional Neural Network for Text Classification）模型的词嵌入层，引入基于预训练的动态词向量替换原有的静态词向量，增强其对上下文语义的感知能力，判别句子的修辞使用情况，进而提取修辞特征。最后融合句级浅层语言学特征与修辞特征，并对特征集合进行筛选和处理，挑选出与句子质量相关性高的特征项，输入到随机森林算法中进行分类，得到句子整体质量的评分等级。经过实验，本章设计的浅层语言学特征能有效判别句子质量，改进后的修辞手法识别模型在修辞手法识别任务上相较于主流文本分类模型有显著优势。

（2）针对常见文本连贯性分析模型对主题信息不敏感的问题。本章提出了基于迁移学习的写作连贯性分析模型，该模型在数据预处理阶段融入作文主题信息，通过改进的负样本生成策略解决数据不平衡问题，使用自建小学作文数据集对轻量化预训练语言