

预训练语言模型在日文文本难易度自动分类中的应用

刘君

广西大学外国语学院, 广西 南宁

收稿日期: 2025年10月29日; 录用日期: 2025年11月27日; 发布日期: 2025年12月4日

摘要

在日语教学过程中,选择难易度合适日文文本作为教学材料有利于提高日语学习者的学习兴趣及效率。日语具有词汇量大、语法复杂等特征,对文本难易度分类提出了挑战。本文尝试采用多种基于神经网络的日语预训练语言模型,通过收集历年日本语能力测试真题以及模拟题作为数据集以训练日文文本难易度自动分类模型。实验结果表明,预训练语言模型在日文文本难易度自动分类任务上能够表现出较好的性能。基于预训练语言模型的日文文本难易度自动分类方法将为计算机辅助日语学习系统以及电子化教材开发等提供有力的技术保障。

关键词

预训练语言模型, BERT, 难易度分类, 日语教育

Application of Pre-Trained Language Models in Automatic Difficulty Classification of Japanese Texts

Jun Liu

School of Foreign Languages and Literatures, Guangxi University, Nanning Guangxi

Received: October 29, 2025; accepted: November 27, 2025; published: December 4, 2025

Abstract

In the process of Japanese teaching and learning, it is important to select Japanese texts with appropriate difficulty levels as educational materials. This can help to enhance Japanese language learners' interest and efficiency. Japanese language is characterized by its large vocabulary and complex

grammar, which makes it challenging to automatically classify the difficulty levels of Japanese texts. This paper attempts to apply several neural network-based Japanese pre-trained language models in training automatic difficulty classification models of Japanese texts, by collecting historical Japanese Language Proficiency Test questions and simulated questions as datasets. Experimental results show that pre-trained language models have good performance in automatically classifying the difficulty levels of Japanese texts. The pre-trained language models-based approach for automatic difficulty classification of Japanese texts will offer significant technical support for the development of computer-assisted Japanese learning systems and electronic textbooks.

Keywords

Pre-Trained Language Model, BERT, Difficulty Classification, Japanese Language Education

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在全球化的今天，学习第二语言已成为许多人提升竞争力的重要途径。随着日本与世界其他国家和地区在经济、教育、文化等方面的交流不断加强，日语吸引了越来越多的学习者。根据日本国际交流基金 2023 年 3 月 31 日公布的《2021 年度海外日本语教育机构调查报告书》的调查结果显示：世界各国及地区的日语学习者人数达到了 3,794,714 人[1]。

在日语个性化教学过程中，基于难易度评估的文本分类在教学资源适配方面起到非常重要的作用，也是急需解决的关键技术问题。文本难易度评估的主要任务是通过对给定文本的分析，预测出该文本的难度值或难度级别，以判断该文本符合哪种层级能力水平的读者[2]。目前，先行研究多集中在英文[3]-[5]、中文[6]-[8]等语种的难易度评估上，日文难易度评估相关的研究相对较少[9]。对于日语教育领域中将日语作为第二语言的学习者而言，有效的难易度评估模型能够帮助其适配个性化学习资源，以提高其学习效果。例如，Hasebe 与 Lee (2015)通过设计日语表记、词汇等语言学特征，提出了日文文本的难易度评估公式[10]；Wang 与 Andersen (2016)以日语语法句型作为模版特征，构建了基于支持向量机的机器学习模型，以预测日文文本的难易度分类[11]；王淑一等(2022)从文字、词汇和句法三个维度构建了日语文本特征，借助 SVM 算法预测国内日语专业本科阶段所用教材的难易度分类[12]；中町礼文等(2022)通过构建 LightGBM 模型以预测日文文本的难易度[13]。从日文文本难易度评估方法来看，上述先行研究主要采用基于公式或传统机器学习的方法。这类方法极大依赖具备日语语言知识的背景的教学或研究人员事先设计用于难易度评估的语言学特征体系。但是，构建的语言学特征体系可能存在完整性、主观性等问题。

当前，人工智能(Artificial Intelligence, AI)技术迅猛发展，为社会经济的发展注入了新的活力。在 AI 时代背景下，将以大语言模型(Large Language Models, LLMs)为代表的 AI 技术运用到外语教学研究已成为未来外语教育的发展趋势[14]。“AI + 外语”的深度融合，为新时代外语学科人才培养和学科建设打开了新局面[15] [16]。

基于上述背景分析，本文以日语能力考试的难易度等级为参考，通过收集历年日语能力考试真题以及模拟题中的阅读题型文本，分别采用多种日语预训练语言模型构建日文文本难易度自动分类模型，并对比不同模型的性能差异，最终获得适用于日文文本难易度自动分类的最优模型。

2. 数据收集与预处理

2.1. 数据收集

本文参考日本语能力测试(Japanese Language Proficiency Test, JLPT)划分的难易度等级。JLPT 由日本国际交流基金会与日本国际教育支援协会设立，是目前全球规模最大的用于评价母语为非日语的学习者能力的考试。JLPT 的难易度等级，分为 N1 到 N5 共计五个级别。其中，N1 为最高级别难度，N5 为最基础难度。

首先，从现有市面上已发行的日语能力考试真题(N1~N3 级别) [17]-[19] 以及面向 JLPT 的日语教辅书(N1~N5 级别) [20]-[29] 收集了阅读题型的文本；其次，采用光学字符识别(Optical Character Recognition, OCR)技术将纸质文本转换为电子化文本；最后，删除掉了以图片、表格等形式出现的阅读题型文本。本文收集的日语文本数量统计说明如表 1 所示：

Table 1. Data statistics of the Japanese texts
表 1. 日语文本数据统计说明

序号	难易度	文本数量
1	JLPT-N1	335
2	JLPT-N2	320
3	JLPT-N3	190
4	JLPT-N4	145
5	JLPT-N5	90
总计		1080

2.2. 数据预处理

本文根据实际教学过程中大学日语专业学习者的学习进阶情况以及日文文本难易度自动分类实验设计需求，进一步对 JLPT 的难易度进行了四种不同分类，并对对应 JLPT 不同难度的文本进行了相应标注。具体分类及标注说明如表 2 所示：

Table 2. Difficulty classification-based data annotation of the Japanese texts
表 2. 基于难易度分类的日语文本数据标注

序号	分类法	标注	难易度
1	二分类	0	JLPT-N1、N2
		1	JLPT-N3、N4、N5
		0	JLPT-N1、N2
2	三分类	1	JLPT-N3、N4
		2	JLPT-N5
		0	JLPT-N1
3	四分类	1	JLPT-N2
		2	JLPT-N3
		3	JLPT-N4、N5

续表

4	五分类	0	JLPT-N1
		1	JLPT-N2
		2	JLPT-N3
		3	JLPT-N4
		4	JLPT-N5

此外，为了有效训练日语文本难易度分类模型，本文将标注后的数据集按照 8:2 的比例划分为训练集和测试集。

3. 实验设计

3.1. 模型选择

本文分别选取了由日本京都大学、日本情报通信研究机构以及日本东北大学发布的日语预训练语言模型，共计 14 种。这些模型采用双向 Transformer 编码器表征(Bidirectional Encoder Representations from Transformers, BERT) [30] 算法训练而成。日语预训练语言模型信息如表 3 所示：

Table 3. Information of the pre-trained language models
表 3. 预训练语言模型信息

序号	模型名称	发布单位
1	kyoto-L12	日本京都大学
2	kyoto-L24	
3	NICT-32K	日本情报通信研究机构
4	NICT-100K	
5	tohoku-bert-base-japanese	
6	tohoku-bert-base-japanese-v2	
7	tohoku-bert-base-japanese-v3	
8	tohoku-bert-base-japanese-char	
9	tohoku-bert-base-japanese-char-v2	日本东北大学
10	tohoku-bert-base-japanese-char-v3	
11	tohoku-bert-large-japanese	
12	tohoku-bert-large-japanese-v2	
13	tohoku-bert-large-japanese-char	
14	tohoku-bert-large-japanese-char-v2	

3.2. 评价指标

为了有效评估日语文本难易度分类模型性能，本文采用以下三个指标：

$$\text{精确率}(\text{Precision}) = \frac{\text{被模型预测为正类文本中真正是正类的文本数}}{\text{被模型预测为正类的文本数}}$$

$$\text{召回率(Recall)} = \frac{\text{被模型预测为正类文本中真正是正类的文本数}}{\text{测试集中所有真正的正类文本数}}$$

$$\text{F1值(F1 Score)} = \frac{2 * \text{精确率} * \text{召回率}}{\text{精确率} + \text{召回率}}$$

4. 实验结果与分析

4.1. 二分类实验结果

从二分类实验结果(如表 4 所示)来看,除了模型 4 (NICT-100K)之外,其他模型取得的 F1 值均超过了 0.8。其中,模型 12 (tohoku-bert-large-japanese-v2)在精确率、召回率以及 F1 值三项指标上均获得最高值,其分值均为 0.926。实验结果表明,该模型在将日文文本分类为高级难度文本(JLPT-N1、N2)和低级难度文本(JLPT-N3、N4、N5)两个级别上表现出了最优的性能。

Table 4. Comparison of experimental results for binary classification

表 4. 二分类实验结果对比

序号	模型名称	精确率 (Precision)	召回率 (Recall)	F1 值 (F1 Score)
1	kyoto-L12	0.860	0.857	0.858
2	kyoto-L24	0.880	0.880	0.880
3	NICT-32K	0.890	0.889	0.890
4	NICT-100K	0.759	0.759	0.759
5	tohoku-bert-base-japanese	0.896	0.894	0.895
6	tohoku-bert-base-japanese-v2	0.873	0.870	0.871
7	tohoku-bert-base-japanese-v3	0.879	0.857	0.868
8	tohoku-bert-base-japanese-char	0.874	0.861	0.867
9	tohoku-bert-base-japanese-char-v2	0.874	0.861	0.867
10	tohoku-bert-base-japanese-char-v3	0.861	0.833	0.847
11	tohoku-bert-large-japanese	0.907	0.907	0.907
12	tohoku-bert-large-japanese-v2	0.926	0.926	0.926
13	tohoku-bert-large-japanese-char	0.891	0.889	0.890
14	tohoku-bert-large-japanese-char-v2	0.879	0.880	0.880

注: 表中分值最大的数字用粗体表示。

4.2. 三分类实验结果

从三分类实验结果(如表 5 所示)来看,模型 12 (tohoku-bert-large-japanese-v2)以及模型 14 (tohoku-bert-large-japanese-char-v2)的精确率最高,分值均为 0.784; 模型 7 (tohoku-bert-base-japanese-v3)以及模型 14 (tohoku-bert-large-japanese-char-v2)的召回率最高,分值均为 0.773; 模型 14 (tohoku-bert-large-japanese-char-v2)的 F1 值最高,分值为 0.778。综合三分类结果来看,模型 14 (tohoku-bert-large-japanese-char-v2)在将日文文本分类为高级难度文本(JLPT-N1、N2)、中级难度文本(JLPT-N3、N4)和低级难度文本(JLPT-N5)三个级别上表现出了最优的性能。

Table 5. Comparison of experimental results for ternary classification
表 5. 三分类实验结果对比

序号	模型名称	精确率 (Precision)	召回率 (Recall)	F1 值 (F1 Score)
1	kyoto-L12	0.710	0.708	0.709
2	kyoto-L24	0.759	0.755	0.757
3	NICT-32K	0.734	0.732	0.733
4	NICT-100K	0.583	0.583	0.583
5	tohoku-bert-base-japanese	0.771	0.764	0.767
6	tohoku-bert-base-japanese-v2	0.723	0.727	0.725
7	tohoku-bert-base-japanese-v3	0.770	0.773	0.771
8	tohoku-bert-base-japanese-char	0.719	0.704	0.711
9	tohoku-bert-base-japanese-char-v2	0.767	0.750	0.758
10	tohoku-bert-base-japanese-char-v3	0.751	0.755	0.753
11	tohoku-bert-large-japanese	0.761	0.745	0.753
12	tohoku-bert-large-japanese-v2	0.784	0.750	0.767
13	tohoku-bert-large-japanese-char	0.762	0.759	0.760
14	tohoku-bert-large-japanese-char-v2	0.784	0.773	0.778

注：表中分值最大的数字用粗体表示。

4.3. 四分类实验结果

从四分类实验结果(如表 6 所示)来看，模型 14 (tohoku-bert-large-japanese-char-v2)取得了最高的精确率，分值均为 0.761；模型 12 (tohoku-bert-large-japanese-v2)取得了最高的召回率，分值均为 0.722；模型 12 (tohoku-bert-large-japanese-v2)的 F1 值最高，分值为 0.735。综合四分类结果来看，模型 12 (tohoku-bert-large-japanese-v2)在将日文文本分类为高级上等难度文本(JLPT-N1)、高级下等难度文本(JLPT-N2)、中级难度文本(JLPT-N3、N4)和低级难度文本(JLPT-N5)四个级别上表现出了更好的性能。

Table 6. Comparison of experimental results for four-class classification

表 6. 四分类实验结果对比

序号	模型名称	精确率 (Precision)	召回率 (Recall)	F1 值 (F1 Score)
1	kyoto-L12	0.626	0.602	0.614
2	kyoto-L24	0.709	0.676	0.692
3	NICT-32K	0.681	0.671	0.676
4	NICT-100K	0.489	0.444	0.465
5	tohoku-bert-base-japanese	0.704	0.681	0.692
6	tohoku-bert-base-japanese-v2	0.706	0.662	0.683
7	tohoku-bert-base-japanese-v3	0.646	0.648	0.647
8	tohoku-bert-base-japanese-char	0.630	0.630	0.630

续表

9	tohoku-bert-base-japanese-char-v2	0.645	0.639	0.642
10	tohoku-bert-base-japanese-char-v3	0.656	0.593	0.623
11	tohoku-bert-large-japanese	0.715	0.694	0.704
12	tohoku-bert-large-japanese-v2	0.748	0.722	0.735
13	tohoku-bert-large-japanese-char	0.734	0.648	0.688
14	tohoku-bert-large-japanese-char-v2	0.761	0.690	0.724

注：表中分值最大的数字用粗体表示。

4.4. 五分类实验结果

从五分类实验结果(如表 7 所示)来看，模型 13 (tohoku-bert-large-japanese-char)取得了最高的精确率，分值均为 0.73；模型 7 (tohoku-bert-base-japanese-v3)取得了最高的召回率，分值均为 0.681；模型 11 (tohoku-bert-large-japanese)的 F1 值最高，分值为 0.683，其次为模型 7，分值为 0.682。综合四分类结果来看，模型 7 (tohoku-bert-base-japanese-v3)在将日文文本分类为高级上等难度文本(JLPT-N1)、高级下等难度文本(JLPT-N2)、中级上等难度文本(JLPT-N3)、中级下等难度文本(JLPT-N4)和低级难度文本(JLPT-N5)五个级别上表现出了更好的性能。

Table 7. Comparison of experimental results for five-class classification**表 7. 五分类实验结果对比**

序号	模型名称	精确率 (Precision)	召回率 (Recall)	F1 值 (F1 Score)
1	kyoto-L12	0.589	0.579	0.584
2	kyoto-L24	0.653	0.644	0.648
3	NICT-32K	0.634	0.630	0.632
4	NICT-100K	0.410	0.412	0.411
5	tohoku-bert-base-japanese	0.686	0.671	0.678
6	tohoku-bert-base-japanese-v2	0.649	0.625	0.637
7	tohoku-bert-base-japanese-v3	0.683	0.681	0.682
8	tohoku-bert-base-japanese-char	0.602	0.593	0.597
9	tohoku-bert-base-japanese-char-v2	0.652	0.625	0.638
10	tohoku-bert-base-japanese-char-v3	0.660	0.593	0.625
11	tohoku-bert-large-japanese	0.690	0.676	0.683
12	tohoku-bert-large-japanese-v2	0.678	0.676	0.677
13	tohoku-bert-large-japanese-char	0.730	0.667	0.697
14	tohoku-bert-large-japanese-char-v2	0.692	0.625	0.657

注：表中分值最大的数字用粗体表示。

4.5. 实验结果分析

通过对不同模型的表现，可以发现基于 BERT 预训练语言模型的方法在日文文本难易度自动分类任务上具有一定的有效性。实验结果表明了 BERT 预训练语言模型强大的泛化能力。这主要得益于 BERT

预训练语言模型采用的 Transformer 框架能够有效地捕捉文本中单词和语句的语义特征，从而使得模型的分类结果更加精确。

从评估指标来看，日本东北大学发布的 BERT 预训练语言模型在各项任务中表现出了显著的优势；尤其是在二分类和四分类任务中，模型 12 (tohoku-bert-large-japanese-v2) 的多项评估指标均获得了最高分值。但是，随着分类任务的进一步细化，模型的性能也在随之下降。不同模型的表现差异可能主要源于网络结构、参数量和训练数据规模等。例如，日本东北大学发布的 BERT-large 模型在日文文本难易度自动分类任务中比 BERT-base 模型均表现更优。相比较 BERT-base 模型而言，BERT-large 模型使用了更多的 Transformer 编码器层(24 个)以及自注意力头(16 个)；其训练数据规模也更大，包括从 cc-100 和 Wikipedia 的语料，分别为 74.3 GB 和 4.9 GB。在三分类任务中，日本东北大学发布的 BERT-large-char 模型(模型 14)的表现较好。该模型在日语分词处理上采用了字符级别的嵌入，即在输入层，通过将每个字符映射为向量，然后通过神经网络来学习字符序列的表示，再输入到 BERT 中。基于字符的 BERT 模型可以更好地处理文本中的未知词问题。在五分类任务中，BERT 模型表现不佳。其原因可能在于五个难易度的日文文本数量不均衡，尤其是 JLPT-N5 的文本较少，这导致 BERT 模型能够捕捉到的文本语义特征较少。此外，导致 BERT 模型分类错误的原因可能还在于一些日文常用词对文本的难易度区分不够明显；例如，日文文本中出现频度较高的「は」、「が」等助词。

尽管 BERT 预训练语言模型在日文文本难易度自动分类任务中表现出了一定的优势，但是仍然含有一些不足以及今后可以改进的地方。首先，BERT 预训练语言模型需要大量的计算资源和数据资源，训练成本和技术门槛较高，这限制了其在语言教育等人文领域应用场景中的应用和推广。其次，BERT 预训练语言模型输出的分类结果对语言教育等人文领域的研究者和教育者来讲，在解释性方面可能存在一定的局限性。因此，在今后的研究工作中，可以探索将 BERT 预训练语言模型与语言学特征相融合，通过收集更多的个性化语言数据并采用更加有效的预训练或微调等方法，构建更加精细化的模型，以提高 BERT 预训练语言模型的表示能力、泛化性能以及可解释性，从而更加满足语言教学的实际应用需求。

5. 结论

本文提出基于预训练语言模型的方法，通过多项分类实验评估了 14 中日语 BERT 预训练语言模型在日文文本难易度自动分类任务中的表现。实验结果表明，该方法具有一定的有效性，尤其是在二分类任务上表现最优；各种模型随着分类的进一步细化，其性能也随之降低；不同机构发布的模型在难易度分类性能方面也表现出了一定的差异。综合来看，日本东北大学发布的 BERT 预训练语言模型在本次日文文本难易度分类任务中表现更优。BERT 预训练语言模型因其强大的表示能力，能够适应日文文本难易度分类任务，在日语教学领域展现出了较大的应用价值。今后，将继续探索针对日文的语言数据标注、文本分类的标准化、日语语言特征挖掘等系列问题，以构建更加精细化的模型，为智能化、个性化日语教学提供有力的技术保障。例如，可借助实时动态建模技术，模型不断输出与学习者日语能力相匹配的文本，并接受学习者的反馈，达到优化文本分类结果。

基金项目

广西哲学社会科学规划研究课题：日语文本易读性评估中融合神经网络技术的语言学特征优化组合研究(批准号：22FYY011)。

参考文献

- [1] Japan Foundation (2023) Survey on Japanese-Language Education Abroad 2021.
<https://www.jpf.go.jp/e/project/japanese/survey/result/survey21.html>

- [2] 王蕾. 文本可读性公式研究发展阶段及特点[J]. 语言教学与研究, 2022(2): 29-40.
- [3] Chen, X. and Meurers, D. (2017) Word Frequency and Readability: Predicting the Text-Level Readability with a Lexical-level Attribute. *Journal of Research in Reading*, **41**, 486-510. <https://doi.org/10.1111/1467-9817.12121>
- [4] Arase, Y., Uchida, S. and Kajiwara, T. (2022) CEFR-Based Sentence Difficulty Annotation and Assessment. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Abu Dhabi, December 2022, 6206-6219. <https://doi.org/10.18653/v1/2022.emnlp-main.416>
- [5] Crossley, S., Heintz, A., Choi, J.S., Batchelor, J., Karimi, M. and Malatinszky, A. (2022) A Large-Scaled Corpus for Assessing Text Readability. *Behavior Research Methods*, **55**, 491-507. <https://doi.org/10.3758/s13428-022-01802-x>
- [6] Sung, Y.T., Chen, J.L., Cha, J.H., Tseng, H.C., Chang, T.H. and Chang, K.E. (2014) Constructing and Validating Readability Models: The Method of Integrating Multilevel Linguistic Features with Machine Learning. *Behavior Research Methods*, **47**, 340-354. <https://doi.org/10.3758/s13428-014-0459-x>
- [7] 吴思远, 于东, 江新. 汉语文本可读性特征体系构建和效度验证[J]. 世界汉语教学, 2020, 34(1): 81-97.
- [8] 谭可人, 兰韵诗, 张杨, 等. 基于多层级语言特征融合的中文文本可读性分级模型[J]. 中文信息学报, 2024, 38(5): 41-52.
- [9] 柴崎秀子. リーダビリティー研究と「やさしい」日本語[J]. 日本語教育, 2014(158): 49-65.
- [10] Hasebe, Y., and Lee, J.H. (2015) Introducing a Readability Evaluation System for Japanese Language Education. *Proceedings of the 6th International Conference on Computer Assisted Systems for Teaching & Learning Japanese*, Hawaii, August 2015, 19-22.
- [11] Wang, S. and Andersen, E. (2016) Grammatical Templates: Improving Text Difficulty Evaluation for Language Learners. *Proceedings of the 26th International Conference on Computational Linguistics (COLING 2016)*, Osaka, 11-16 December 2016, 1692-1702.
- [12] 王淑一, 施建军, 许培钿, 等. 面向中国日语专业本科教学的文章难度测量研究[J]. 日语学习与研究, 2022(5): 45-56.
- [13] 中町礼文, 佐藤敏紀, 西内紗恵, 等. 日本語能力試験に基づく日本語文の難易度推定[C]//言語処理学会第 28 回年次大会発表論文集. 京都: 言語処理学会事務局, 2022: 658-663.
- [14] 黄荣怀. 信息技术与教育[M]. 北京: 北京师范大学出版社, 2002.
- [15] 田臻, 彭雅清. 人工智能背景下的计算机辅助语言学习研究进展(2011-2021) [J]. 外语界, 2022(3): 53-60.
- [16] 朱晔, 王陈欣, 金慧. 智能时代计算机辅助的语言学习研究[J]. 外语教学, 2021, 42(5): 51-56.
- [17] 天明教育日语能力考试研究组. 新日本语能力考试真题大全 N1 [M]. 郑州: 河南大学出版社, 2021.
- [18] 天明教育日语能力考试研究组. 新日本语能力考试真题大全 N2 [M]. 郑州: 河南大学出版社, 2021.
- [19] 天明教育日语能力考试研究组. 新日本语能力考试真题大全 N3 [M]. 郑州: 河南大学出版社, 2021.
- [20] 刘文照, 海老原博. 新日本语能力考试 N1 读解[M]. 上海: 华东理工大学出版社, 2020.
- [21] 刘文照, 海老原博. 新日本语能力考试 N2 读解[M]. 上海: 华东理工大学出版社, 2020.
- [22] 刘文照, 海老原博. 新日本语能力考试 N3 读解[M]. 上海: 华东理工大学出版社, 2020.
- [23] 刘文照, 海老原博. 新日本语能力考试 N4 读解[M]. 上海: 华东理工大学出版社, 2020.
- [24] 刘文照, 海老原博. 新日本语能力考试 N5 读解[M]. 上海: 华东理工大学出版社, 2020.
- [25] 张鸿成. 新日语 N4 教程[M]. 上海: 上海译文出版社, 2011.
- [26] 张鸿成. 新日语 N5 教程[M]. 上海: 上海译文出版社, 2011.
- [27] 夏爱冰, 陈婧. 新日本语能力测试 N4 模拟冲刺[M]. 北京: 外文出版社, 2010.
- [28] 刘文照, 海老原博. 非凡·新日本语能力考试·N4 全真模拟试题[M]. 上海: 华东理工大学出版社, 2020.
- [29] 刘文照, 海老原博. 非凡·新日本语能力考试·N5 全真模拟试题[M]. 上海: 华东理工大学出版社, 2020.
- [30] Devlin, J., Chang, M., Lee, K. and Toutanova, K. (2019) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT 2019*, Minneapolis, 2-7 June 2019, 4171-4186.