

中南财经政法大学

本科生毕业论文（设计）



论文题目	:	基于语言模型的古琴减字谱打谱研究
姓名	:	石锴文
学号	:	201821130123
班级	:	信管 1801 班
年级	:	2018 级
专业	:	信息管理与信息系统
学院	:	信息与安全工程学院
指导教师	:	刘勘
完成时间	:	2022 年 5 月 7 日

作者声明

本毕业论文（设计）是在导师的指导下由本人独立撰写完成的，没有剽窃、抄袭、造假等违反道德、学术规范和其他侵权行为。对本论文（设计）的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。因本毕业论文（设计）引起的法律结果完全由本人承担。

毕业论文（设计）成果归中南财经政法大学所有。

特此声明。

作者专业 : **信息管理与信息系统**

作者学号 : **201821130123**

作者签名 :

年 月 日

基于语言模型的古琴减字谱打谱研究

石锴文

*Research on Guqin reduced notation transcription
based on language model*

Shi, Kai wen

2022 年 5 月 7 日

摘 要

古琴是我国优秀的非物质文化遗产，距今已有两千多年的传承历史。随着 2003 年成功申报联合国“人类口头和非物质遗产代表作”，有关古琴的传承和保护受到广泛关注。在这一背景下，如何利用现代化的智能信息技术，为弘扬和传承古琴艺术进行技术性的研究工作，成为一个重要的课题。

减字谱是古琴独有的记谱方法，琴谱数字化有利于曲风分析、智能辅助打谱等研究工作。现存古琴谱一百四十余种，3365 个不同版本的传谱，658 支不同的传曲。借助计算机高效的存储及运算功能，可以极大推动古琴艺术相关的研究、保护和推广工作。减字谱出来表达乐曲弹奏的指法以外，还能够表示一定的乐句以及大致的节奏，但是每个音的音长和节奏都不确定，需要经过长时间的分析考究，从而把定出节奏。这个过程就是“打谱”，打谱的过程少则几个月，多则两三年，非常消耗打谱者的精力。如果有一种智能打谱的算法能够辅助打谱，挖掘琴谱中难以被人眼发现的信息，那么古琴艺术将会迎来极大的发展。

本文在现有研究工作的基础上，进一步对古琴数字化智能化打谱技术进行研究。对计算机辅助打谱的定义进行讨论，将其定义为序列标注任务：确定每个减字所对应的音符类型。另外依据古琴减字的特点，提出了基于分词解构减字组合的方法，将减字拆分为左手指法、右手指法、徽位和弦位四个部分。该方法能够有效降低词表的大小，在古琴打谱低资源的数据场景下，可以减少预训练语言模型对数据量的需求。

本文的核心工作是在 BERT 模型的基础上，依据古琴减字的特性为其设计了语言模型——QinBERT，并且基于该预训练模型对下游任务进行微调，每个实验组中最优模型 F1 值均达到 0.4 以上。除此之外，本文为不同流派的打谱设计了动态打谱机制，通过接受一首琴曲中一定数量的标注，完成后续打谱任务。实验结果表明，动态打谱机制能够在一定程度上延续打谱者的流派风格。

关键词： 古琴；减字谱；打谱；语言模型；序列标注

Abstract

Guqin is an excellent intangible cultural heritage in my country, with a history of more than 2,000 years. With the successful application to the United Nations in 2003 for the "Masterpiece of the Oral and Intangible Heritage of Humanity", the inheritance and protection of the guqin has received widespread attention. In this context, how to use modern intelligent information technology to carry out technical research work for the promotion and inheritance of Guqin art has become an important topic.

The jianzi notation is a unique notation method for the guqin, and the digitization of the notation is conducive to research work such as genre analysis and intelligently assisted notation. There are more than 140 kinds of guqin scores, 3,365 different versions of biography, and 658 different biography. With the efficient storage and computing functions of computers, the research, protection and promotion of the art of guqin can be greatly promoted. In addition to expressing the fingering of the music, the minus-character notation can also express certain phrases and general rhythm, but the length and rhythm of each note are uncertain, and it takes a long time to analyze and study, so as to determine the rhythm. This process is "playing notation", and the process of notation can be as little as a few months, or as long as two or three years, which consumes a lot of the energy of the scorer. If there is an algorithm for intelligent notation that can assist in notation and dig out the information that is difficult to be found by the human, then the art of guqin will usher in great development.

Based on the existing research work, this paper further studies the digital and intelligent notation technology of guqin. Discuss the definition of computer-aided notation, and define it as a sequence labeling task: determine the note type corresponding to each minus character. In addition, according to the characteristics of guqin jianzi, a method of deconstructing jianzi based on word segmentation is proposed, and the jianzi are divided into four parts: left fingering, right fingering, emblem position and chord sequence. This method can effectively reduce the size of the vocabulary, and in the low-resource data scenario of guqin playing, it can reduce the data volume requirement of the pre-trained language model.

The core work of this paper is to design a language model-QinBERT based on the characteristics of guqin jianzi based on the BERT model, and fine-tune the downstream tasks based on the pre-training model. The F1 value of the optimal model in each experimental group was above 0.4. In addition, this paper designs a dynamic notation

mechanism for different genres of notation. By accepting a certain number of annotations in a song, the subsequent notation tasks are completed. The experimental results show that the dynamic transcribing mechanism can continue the genre style of transcribers to a certain extent.

Key words: Guqin; Jianzipu; Jianzipu transcribing; Language model; Sequence labeling

目录

一、绪论	- 1 -
(一) 研究背景与意义	- 1 -
(二) 相关工作概述	- 2 -
(三) 打谱任务定义	- 5 -
(四) 本文组织结构	- 6 -
二、相关理论基础	- 7 -
(一) 自然语言表示学习	- 7 -
(二) 预训练语言模型	- 8 -
三、减字谱数据集构建	- 9 -
(一) 数据来源及录入方法	- 9 -
(二) 数据集描述及复核	- 10 -
四、预训练减字语言模型	- 12 -
(一) 减字字符编码	- 12 -
(二) 语言模型结构	- 13 -
(三) 遮罩策略设计	- 13 -
五、基于预训练语言模型的打谱模型	- 14 -
(一) 整体打谱流程	- 14 -
(二) 打谱模型结构	- 15 -
(三) 动态打谱模型	- 16 -
六、实验及结果	- 18 -
(一) 实验设置	- 18 -
(二) 实验过程	- 18 -
(三) 实验结果与分析	- 20 -
七、总结与展望	- 21 -
(一) 工作总结	- 21 -
(二) 展望	- 21 -
主要参考文献	- 23 -

一、绪论

（一）研究背景与意义

“窈窕淑女，琴瑟友之”，中国古琴文化源远流长三千年。琴是我国优秀的非物质文化遗产，自古以来就是人们精神高洁的象征。1977 年，由我国著名古琴家管平湖先生演奏的古琴曲《流水》被刻录在一张喷金铜唱片上，随美国“旅行者”号发射到太空。古琴对于中国文化乃至世界人类文明的重要性不言而喻。随着时代演化，在文化强国的今天，古琴音乐历久弥新，其文化遗产和保护受到广泛关注。在这样的背景下，运用信息技术保护、传承和弘扬古琴文化成为非物质文化遗产保护的一个重要课题。

在众多中国传统民族乐器中，古琴的记谱法较为独特。最早的琴曲是完全由文字记录下来的，文字谱中包含演奏者弹奏所使用的手指、弹奏的弦位、徽位和演奏者的神情。直到唐代末年，琴家曹柔创立了减字谱，即由文字谱减化而来的乐谱，主要记录弹奏时的指法和对应的弦位及徽位。这种记谱法方便记录和传阅，具有“字简而义尽，文约而意赅”的特点，因此沿用千年。

虽然减字谱广为流传，仍然有不少乐谱在传播过程中遗失或谬误。另外，古琴谱只记录了演奏方式和大致节奏，并没有记录每个音符的时值。由于古琴传承依靠老琴家口授相传，即使知道乐曲的减字谱，人们依然难以还原其本身风貌。在世界文化交流频繁的今天，减字谱具有教学讲演的优势，但是并不如简谱或者五线谱那样方便传播和交流。

因此，众多古琴家着手琴曲的“翻译”工作——打谱。古琴曲打谱需要对乐句节奏进行确定，对曲谱中对残缺和谬误进行补充和修正，最后用现代音乐的记谱方式准确描述乐曲的节奏和音高。这项工作费时费力，需要打谱者广泛搜集资料，往往消耗人们大量的时间和精力。故琴界有“小曲打三月，大曲打三年”的说法，许多古琴家打谱 30 余年，费尽心血仅完成 20 首琴曲的打谱工作。然而，先人流传下来的曲谱三千有余，能够被当代琴家演奏的不过百余首。

如果能将智能信息技术应用到辅助打谱工作当中，帮助打谱者完成乐谱的节奏划分和指法修订，那么古琴艺术会迎来崭新的发展。

（二）相关工作概述

1. 计算机辅助打谱综述

有关古琴计算机打谱的研究主要集中在国内，研究的进程还停留在琴谱编码和内容分析上，其余的工作也主要集中在琴曲的音乐性研究上。我国著名琴家，中国科学院研究员陈长林先生早在 1989 年就开始尝试将计算机技术应用到古琴研究当中^①。主要完成了琴谱的编码设置，以及计算机的输出现实工作。陈长林先生在论文中提到古琴的打谱：“研究整理古琴音乐遗产（包括打谱），历来都是靠‘人脑’进行，能否用‘电脑’进行辅助是值得研究的”。可见从一开始，计算机辅助古琴打谱的方向就已经存在。1993 年，喻辉对减字谱音高的提取进行研究，首先设计出一套新的编码规则，之后将不同指法对应的音高计算出来^②。

本世纪以来，计算机技术发展迅速，国内有越来越多的科研人员将信息技术运用到古琴音乐的研究当中。2008 年，厦门大学的周昌乐等人使用分析树的方法，对指法序列进行依存分析，从而构建出减字谱数字化的方案^③。并且设计减字谱字体，完成了减字字库的构建，使古琴谱独有的“文字”能够在计算机中编辑。2009 年，顾亚丽提出基于遗传算法的古琴减字谱自动生成，通过设计出的新编码规则，将琴谱建模为编码序列，使用遗传算法对序列进行续写，完成琴曲的自动创作^④。在随后的 2010 年，周昌乐等人继续深入古琴减字谱的研究：构造了一个减字谱语料库，并使用两个关联规则挖掘算法，对古琴曲的指法共现单元进行挖掘分析^⑤；提出了将现代乐谱转化为减字谱的翻译方法，将琴曲分为虚音和主干部分，使用最大匹配法对虚音部分进行翻译，使用统计方法、模拟退火算法及 Viterbi 算法对主干部分进行处理，用基于规则的方法对半虚半实的部分进行翻译^⑥；提出提取古琴减字谱音高的计算方法^⑦。同年，刘洋等人也提出了一种减字谱的描述方法，对减字谱字符的各个部分进行拆解，使用 XML 格式进行描述^⑧。胡丹在 2014 年详细介绍了上海印刷技术研究所，有关古琴数字化应用的研究，主要包括减字谱的标准编码和计算机输入法，有效地提高了琴谱数字化的效率^⑨。吕沁沁就古琴音乐的传播难点，讨论了数字化打谱标准以及核心数字技术研发对整个古琴艺术的意义，阐述了中国传统音乐内容传播的互联网创新产业模式^⑩。2018 年，张英俐等人也提出了一种基于遗传算法的古琴曲生成方法，提出一种

^① 陈长林. 电脑在古琴音乐研究中的初步应用[J]. 计算机学报, 1989, (07): 525-533.

^② 喻辉. 古琴减字谱音高信息电脑自动处理原理[J]. 交响. 西安音乐学院学报, 1993, (03): 56-58.

^③ 丁晓君, 叶婷婷, 周昌乐. 古琴减字谱的编码与编辑方法[J]. 中国音乐学, 2008, (02): 93-96.

^④ 顾亚丽. 基于遗传算法的古琴减字谱曲段自动生成[D]. 厦门大学, 2009.

^⑤ 高宋倮, 周昌乐. 现代乐谱的古琴减字谱自动翻译研究[J]. 心智与计算, 2010, 4(01): 45-54.

^⑥ 庄雄雄, 周昌乐, 齐京峰, 高宋倮, 吕兰兰. 提取古琴减字音高的计算方法及其软件实现[J]. 计算机应用与软件, 2010, 27(04): 10-12+37.

^⑦ 顾亚丽, 周昌乐. 古琴曲语料库的构建与规则提取[J]. 心智与计算, 2010, 4(02): 128-138.

^⑧ 刘洋, 唐英敏. 一种古琴减字谱符号的结构化描述方法[J]. 北京信息科技大学学报(自然科学版), 2010, 25(S2): 61-65.

^⑨ 胡丹. 古琴减字谱的数字化应用[J]. 印刷杂志, 2014, (05): 59-62.

^⑩ 吕沁沁. 试论中国传统音乐内容传播的互联网创新产业模式——以古琴数字化为例[J]. 编辑学刊, 2017, (05): 67-71.

编码方式记录每个指法的音色和时值等参数，再利用遗传算法进行迭代生成新的曲谱^①。陈根方从记谱法、音乐标注、风格分析、智能打谱和算法作曲等 24 个方面简述了古琴艺术的数字化保护内容，并利用无监督聚类算法对古琴泛音和按音的音律进行分类^②。到 2020 年，李圣辰等人构建了一个包含 71 首琴曲的古琴音乐数据集，但是其中并没有录入减字谱的指法信息^③。

总体来看，国内外有关减字谱的计算机研究非常有限：减字谱的编码规则并不统一，几乎每一项工作的编码都会有差异；另外，使用的研究方法都是基于统计模型加以改进的，没有充分挖掘减字谱当中的语义信息，几乎没有机器学习和深度学习在古琴数字化当中的应用。

^① 张英俐, 刘弘. 一种基于遗传算法的古琴曲生成方法[J]. 信息技术与信息化, 2018, (09):28-30.

^② 陈根方. 古琴艺术数字化保护概述与琴律智能分析[J]. 复旦学报(自然科学版), 2019, 58(03):335-342.

^③ 李圣辰, 吴雨松. 1 个中国古琴曲的符号化音乐数据集介绍及其应用实例[J]. 复旦学报(自然科学版), 2020, 59(03):276-285.

2. 音乐语言模型综述

受自然语言处理技术中的 Word2Vec 词向量模型的启发，早期关于符号音乐的研究通过上下文音乐符号预测来学习音乐符号的嵌入表达。Huang 等人和 Madjiheurem 等人将和弦视为单词，并使用 Word2Vec 模型学习和弦表示^{①②}。Herremans 等人、Chuan 等人和 Liang 等人将音乐片段划分为具有固定持续时间的非重叠音乐片段，并为每个片段训练嵌入^{③④⑤}。Tatsunori 等人[24]将音符分组，并将这些音符小组视为表示学习的单词^⑥。然而，上面提到的基于 Word2Vec 的方法只使用了的规模较小网络模型，并且只使用了 4 到 5 个上下文音乐 token 作为输入，并没有使用大规模深度预训练语言模型。

Tsai 和 Ji 的工作代表了使用预训练模型进行符号音乐分类的首次尝试^⑦。其实验表明，对于 9 类作曲家分类任务，基于 RoBERTa 的 Transformer 编码器预训练模型和使用 GPT2 的 Transformer 编码器预训练模型优于未经预训练的基本方法。预训练将 GPT2 模型的分准确率从 46% 提高到 70%。然而，他们工作中考虑的符号数据格式是“乐谱图像”，这种数据格式较 MIDI 格式使用频率不高。Zeng 等人提出了 MusicBERT，使用 MIDI 格式的数据进行预训练^⑧。使用一百多万多个多轨 MIDI 音乐作品进行预训练后，MusicBERT 被应用于两个生成音乐任务，即旋律补全和伴奏建议，以及两个序列级别的判别任务，即流派分类和风格分类，其实验结果表明预训练后的模型较基本方法有显著的性能优势。Chou 等人同样依据 MIDI 数据提出了 MidiBERT，并开展了音乐语言模型预训练的研究工作^⑨。与 MusicBERT 不同的是，MidiBERT 预训练使用的语料库只有公开的 4166 个钢琴 MIDI 数据文件。MusicBERT 能够处理多轨道的 MIDI 数据，而 MidiBERT 专为单轨的钢琴 MIDI 数据设计。有关音乐语言模型的研究基本集中在国外学者的工作当中，整体的技术演变非常符合自然语言处理技术的演进：从基本的静态词向量技术发展到了预训练语言模型。这也说明了音乐这一特殊的语言，是可以从自然语言处理的角度进行研究的。

^① Huang C, Duvenaud D, Gajos K Z. ChordRipple: Recommending Chords to Help Novice Composers Go Beyond the Ordinary[C]// the 21st International Conference. ACM, 2016: 241-250.

^② Madjiheurem S, Qu L, Walder C. Chord2Vec: Learning Musical Chord Embeddings[C]//In Proceedings of the constructive machine learning workshop at 30th conference on neural information processing systems (NIPS2016), 2016.

^③ Herremans D, Chuan C H. Modeling Musical Context with Word2Vec[C]// First International Workshop on Deep Learning and Music, 2017.

^④ Chuan C H, Agres K, Herremans D. From context to concept: exploring semantic relationships in music with Word2Vec[J]. Neural Computing and Applications, 2020, 32(1).

^⑤ Liang H, Lei W, Chan P, et al. PiRhDy: Learning Pitch-, Rhythm-, and Dynamics-aware Embeddings for Symbolic Music[J]. ACM, 2020:574-582.

^⑥ Tatsunori, Hirai, Shun, et al. Melody2Vec: Distributed Representations of Melodic Phrases based on Melody Segmentation[J]. Journal of Information Processing, 2019, 27:278-286.

^⑦ Tsai T, Ji K. Composer Style Classification of Piano Sheet Music Images Using Language Model Pretraining[J]. 2020:176-183.

^⑧ Zeng M, Tan X, Wang R, et al. MusicBERT: Symbolic Music Understanding with Large-Scale Pre-Training[J]. 2021.

^⑨ Chou Y H, Chen I C, Chang C J, et al. MidiBERT-Piano: Large-scale Pre-training for Symbolic Music Understanding[J]. 2021.

3.文献综述小结

综合前文的现状描述，可以发现国内有关音乐语言模型的研究几乎是空白，辅助打谱的研究进程仍然停留在减字编码和浅显的生成和风格分类上。表 1 初步分析了两个领域研究角度的特点和不足。

表 1 本文所涉及的主要研究领域对比

研究领域	研究角度	目标	局限性
计算机辅助打谱	减字编码	利用数字化技术对古琴谱进行存储和表示	没有形成统一的编码规则
	音高提取	将减字谱翻译为现代音乐的表示	缺乏对音符时值的提取
	曲风分析	对乐曲流派以及指法关联进行分析	没有充分利用减字谱的上下文信息
音乐语言模型	静态词向量	获取音乐符号的嵌入表达	训练语料规模小，网络规模小
	预训练语言模型	使用大规模语料学习音乐符号当中的语言信息	预训练任务的设计没有结合音乐理论

从表 1 可以看出，辅助打谱这一领域还停留在基本的字符级别的分析，没有使用自然语言处理技术对乐谱进行挖掘。而音乐语言模型领域的研究，已经开始将大规模预训练技术应用到音乐挖掘的下游任务当中。究其原因有二：其一，古琴减字谱本身存在多种记录方式，不同的琴家有不一样的记谱习惯；其二，减字编码规则统一，不同的研究工作有不同的编码规则。因此，计算机打谱研究从根本上很难在研究步调上统一。相比之下，音乐语言模型研究所使用的数据逐渐向 MIDI 格式统一，从而能够衍化出更加深层次的技术应用。

要想从根本上改善计算机辅助打谱的研究现状，就需要有官方、权威、统一的记谱和编码规则。另外从技术演进的角度来说，计算机辅助打谱的研究中还没有引入自然语言处理技术，而这些技术已经开始在音乐语言模型领域的研究大放光彩。故将自然语言处理技术应用到辅助打谱研究当中，是十分有必要的。

（三）打谱任务定义

打谱是指将古琴的减字谱转录为通用的简谱或者五线谱，打谱人员通过确定每个音符的音高、时值和各个小节的节奏。从相关工作的调查来看，计算机辅助打谱并没有确定的工作范式。相关工作中，已经有研究人员设计了通过减字提取音高的算法，那么当前工作的重点则落在时值和节奏问题上。古琴节奏自由多变，一首琴曲中可能会出现多种节奏类型的小节。如果把每个音的时值确定下来，通过一定的推断，就可以把小节划分出来，节

奏问题就迎刃而解了。关于时值的确定，可以通过判断每个音的音符类型来确定，如：“全音符”“四分音符”“八分音符”等。

古琴的减字谱通过减字组合的方式描述演奏方法。通常来说，一个或多个减字会对应一个或多个左手和右手的动作，所以并不是每个减字都对应一个音符，那么古琴打谱的过程可以描述为序列标注或者序列生成。从辅助打谱的角度来说：序列生成的方法不便于减字和音符的对齐，打谱者难以对成谱进行识别和排版；如果使用序列标注方法，只需要引入一个新的类别标签即可解决多个减字对应一个音符的问题。

综上所述，本文将计算机辅助打谱任务定义为序列标注任务：确定每个减字所对应的音符类型。

（四）本文组织结构

论文的第一章为绪论，主要介绍研究背景、论文的创新点和论文结构。并且阐述国内外对古琴打谱的研究现状，并补充阐述语言模型算法和自然语言处理技术在类似应用场景的研究工作；

第二章为相关理论基础，介绍本文所使用到的主要方法和技术，包括表示学习理论基础、预训练语言模型、序列标注；

第三章为减字谱数据集构建，介绍数据来源和数据的录入方法，并对数据集进行描述性的统计分析；

第四章为预训练减字语言模型，主要介绍语言模型在减字谱数据集上的预训练策略和过程；

第五章为基于预训练语言模型的打谱模型，详细介绍打谱模型的结构，并对其中的优化过程进行分析；

第六章为实验及结果，描述对比实验的设置和实验参数的设置，并对实验结果进行分析；

第七章为总结与展望，总结研究过程中的亮点与不足，为今后的研究工作提出新的方向。

二、相关理论基础

（一）自然语言表示学习

自然语言是天然的非结构化数据，将其存储在计算机当中最直接的方法是通过符号表示，利用符号系统对每个字符进行编码。面向算法模型进行建模时，则将其转换为独热编码向量，向量的每个维度表示某个字符或者单词是否出现。显然，独热编码天然具有稀疏性，而且向量的每个维度都正交。在词表较大的时候，这样的表示向量并不能很好地表达自然语言的非结构化特征，模型也难以从稀疏向量中获取有效信息。因此，对于自然语言的表示学习研究一直是自然语言处理的重要课题。图 1 梳理了自然语言表示学习研究中的重要技术节点^①。

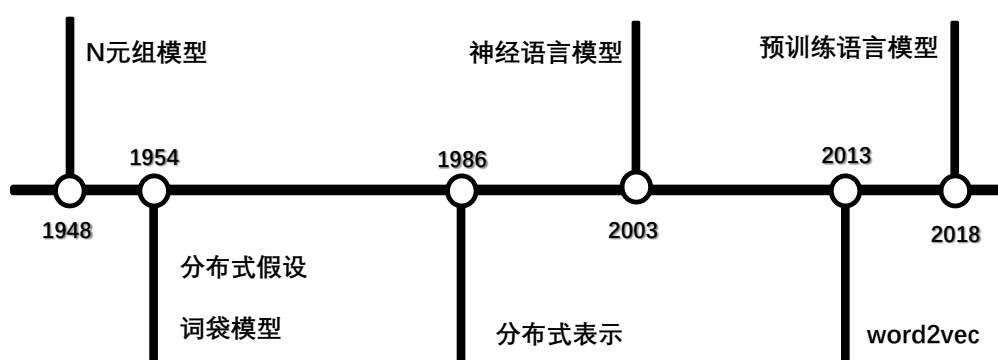


图 1 自然语言表示学习理论发展时间线

最早的词表示模型是 N 元组模型：在预测语言序列中的下一个词的时候，通常会关注前 N 个词^②。语料规模非常大时，可以计算出非常良好的条件概率分布，这个概率分布能够很好地表达词的含义。随后而来的是词袋模型：将一篇文档看作其内容的单词集合，而不关注单词的顺序^③。N 元组模型和词袋模型揭示了自然语言处理当中的非常重要的分布式假设：具有相似分布的语义单元具有相似的含义。分布式假设奠定了后续众多语言模型的理论基础。

依照分布式假设，分布式表示向量的每一个维度都为连续的实数，这些表示向量张成的向量空间即为被表示实体的语义向量空间。随着算力增长和深度学习的不断发展，分布式表示已经成为表示学习中最常用的表示方法。

分布式表示的最典型研究是神经语言模型：首先用随机初始化的向量表示每一个词，将其作为神经网络的权重，之后依据语料计算语言序列中每个词的联合概率分布，并预测下一个词^④。通过这种方法训练神经网络获取词表示的方法被成为词嵌入方法，即将词嵌入到一个连续的语义向量空间当中。最后得到的向量在每个维度上并不具有良好的解释性，但是这些向量的确能够表达词的语义信息。受神经语言模型的影响，Word2Vec、GloVe、

^① Liu Z, Lin Y, Sun M. Representation Learning for Natural Language Processing. 2021: 4-7.

^② Claude E Shannon. A mathematical theory of communication. Bell system technical journal, 27(3):379-423, 1948.

^③ Zellig S Harris. Distributional structure. Word, 10(2-3):146-162, 1954.

^④ James L McClelland, David E Rumelhart, PDP Research Group, et al. Parallel distributed processing. Explorations in the Microstructure of Cognition, 2:216-271, 1986.

fastText 等技术应运而生。在通用或者特定语料上训练词向量，再将其应用到下游任务中，成为了各个自然语言处理任务的范式。

（二）预训练语言模型

随着 ELMo 和 BERT 一类预训练语言模型的提出，自然语言处理迎来了新的飞跃。与之前 Word2Vec 一类的词表示模型不同，预训练语言模型使用规模更大的语料、模型的参数和结构更加巨大繁复。而且预训练语言模型并不为每个词指定一个固定的词表达，而是依据上下文为每个词动态计算词向量。因此，这种动态的语言表示对文本语义的表达能力更强。以迁移学习的形式，在下游任务上对大规模预训练模型进行微调则成为当今自然语言处理的最常用范式。

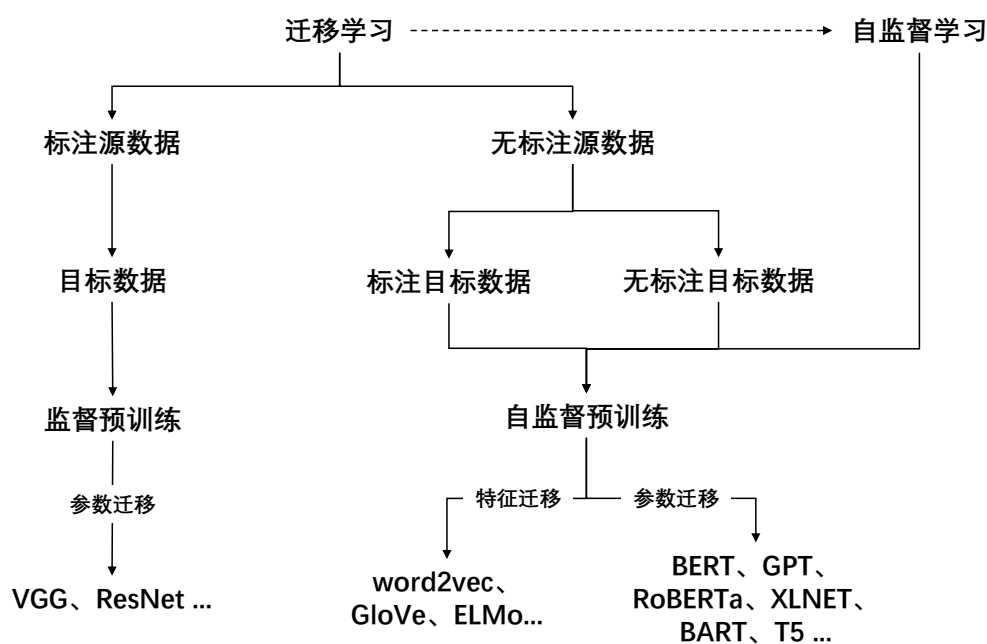


图 2 预训练方法谱系图

预训练模型在下游任务上进行微调本质上属于迁移学习，图 2 展示了不同预训练方法的训练过程及对应示例。在自然语言处理领域中的预训练均基于自监督学习，即源数据本身并没有标签。预训练时，利用语料的上下文顺序关系构造伪数据标签，随后进行自监督预训练。

以 BERT 模型的预训练为例，BERT 的预训练任务包括遮罩预测和下一句话预测。遮罩预测任务为：在原始的语料中，使用 “[MASK]” 随机替换部分 token，再用模型预测这些位置原本的 token。下一句话预测任务为：将原始语料中不同的语句使用 “[SEP]” 分隔，通过 “[CLS]” 的表示向量预测这些语句是否存在先后顺序的关系。

三、减字谱数据集构建

（一）数据来源及录入方法

现存的有关古琴音乐的数据集均不包含减字信息，故笔者选择手工录入减字谱数据。由于李圣辰等人在带有简谱的古琴曲集上构建了古琴音高的数据集，这些曲谱是目前广泛被接受的版本，是古琴教学及演奏的标准版本^①。为了便于后续工作的进行，本文在这些古琴曲集上选择部分琴曲进行减字及对应时值的数据开展录入工作^{②③}。

风 求 凰

1 = C
正调定弦: 1 2 4 5 6 1 2

据《梅庵琴谱》(1931年)
李 祥 霆订指法

慢板

1 2 3 | 5 5 6 1 1 | 5 6 2 1 2 - | 1 1 1 6 1 6 5 |

荀 瑟 荀 瑟 上七 已 荀 荀五 七 勾 七 正 荀 荀 荀 瑟 瑟 瑟

5 1 | 3 5 3 2 1 1 - | 1 2 3 5 | 6 5 5 5 - |

瑟 瑟 荀 瑟 荀 荀 瑟 瑟 瑟 瑟 瑟 瑟

图 3 古琴曲《风求凰》曲谱节选

图 3 展示了《风求凰》的部分选段，以其中的第三和第四小节为例，将如下数据保存到文本文件中：

“食指七徽勾四弦，八分音符\ n 勾五弦，八分音符挑七弦，八分音符*\ n 勾六弦，十六分音符\ n 挑七弦，四分音符-\ n 泛止，其他\ n\ n 散勾六弦，八分音符\ n 大指七徽六分勾七弦，四分音符\ n 剔七弦，八分音符\ n 大指九徽抹七弦，八分音符\ n 上七徽六分，八分音符\ n 大指九徽挑七弦，八分音符\ n 无名指十徽掐起\ n\ n”。

文本文件的每一行记录一对减字和对应的音符类型，减字和音符类型用全角逗号隔开，不同小节用换行符隔开，一首曲子保存在一个文件中。

^① 李圣辰,吴雨松. 1 个中国古琴曲的符号化音乐数据集介绍及其应用实例[J]. 复旦学报(自然科学版),2020,59(03):276-285.

^② 朴东生, 张殿英. 古琴考级曲集 1[M]. 北京: 人民音乐出版社, 2010.

^③ 李祥霆, 龚一. 古琴考级曲集 2[M]. 北京: 人民音乐出版社, 2010.

图 4 列举了数据录入过程中遇到的 3 类特殊减字情况，按照一个动作对应一个音符的原则，对不同类别的减字组合情况采取不同的录入策略。

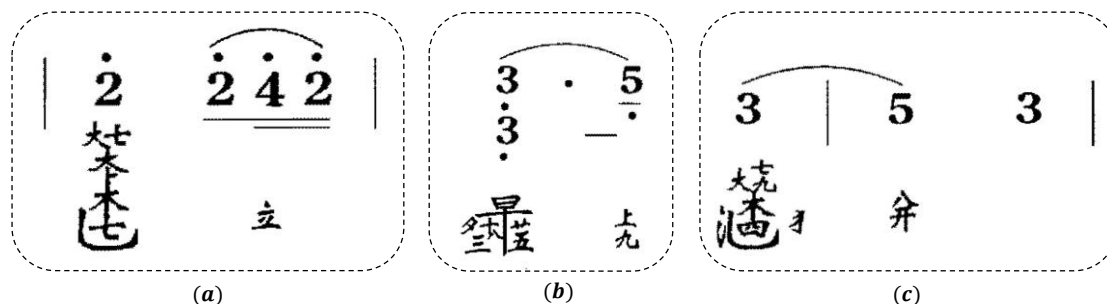


图 4 三类特殊减字示例

其中(a)录入为“大指七徽抹七弦，四分音符\|n 挑七弦，八分音符\|n 进，十六分音符\|n 复，十六分音符\|n\|n”。对于“抹挑”指法的组合减字进行替换，将其替换为“抹”和“挑”的单独减字；对于“撞”减字，将其拆分为“进”和“复”减字，使得一个减字能够对应一个音符。

示例(b)则录入为“撮无名指十徽八分三弦散五弦，四分音符*\|n 上九徽，八分音符”。在这个例子当中，两个减字对应了两组音符类型，如果将“撮”记录为“四分音符-”，那么“上九徽”则不对应任何音符。故将“撮”记录为“四分音符*”，那么两个减字都对应一个音符，这样记录也符合演奏过程中的动作逻辑。

示例(c)录入为“大指七徽九分抹四弦，四分音符\|n 猱，其他\|n\|n 上七徽，四分音符\|n 挑七徽九分四弦，四分音符\|n\|n”。“分开”指法在古琴演奏中通常与“抹挑”或者“勾剔”组合，并且对应三个动作。故按照演奏逻辑将其拆解为三个减字即可。

（二）数据集描述及复核

本文共录入 20 首琴曲，最长的琴曲包含 310 个减字，所有琴曲中最长的小节包含 14 个减字。

由于本文采用的数据集为人工录入，数据录入过程中难免产生错误。因此本文通过两阶段数据复核，对数据进行核验，以保证数据集的正确性和可用性。复核流程如图 5 所示。

第一阶段复核为正确性核验阶段，主要检验录入过程中是否输入错误的字符，并计算每个小节的音符时值之和是否为整数拍。

第二阶段复核为逻辑性核验阶段，主要检查减字的输入是否与琴谱一致，有调整的数据段落是否符合演奏逻辑，这一次核验由人工完成。

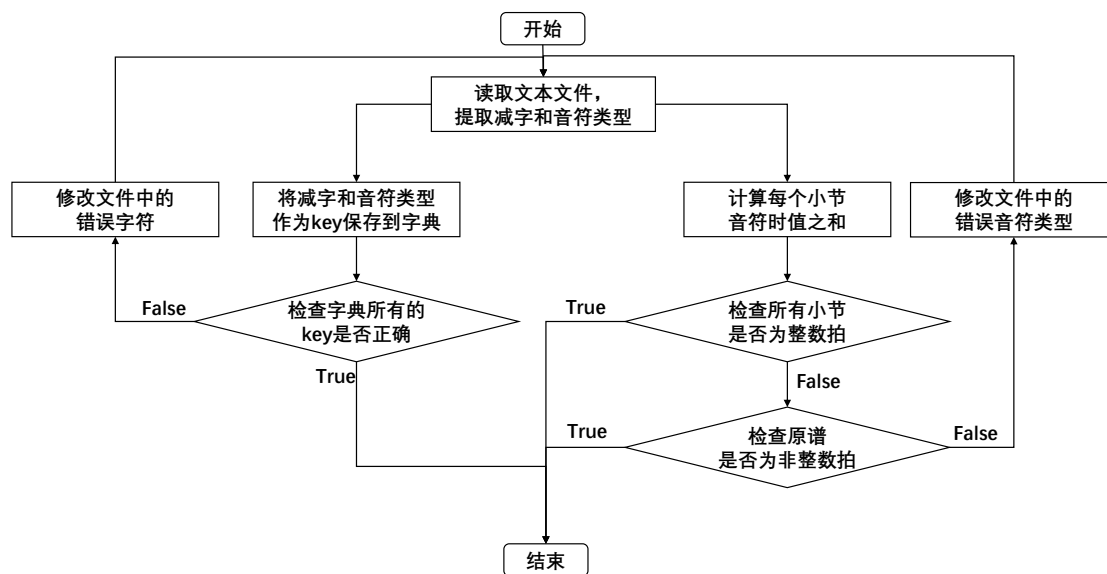


图 5 第一阶段数据复核流程图

正确性核验阶段的复核通过程序脚本来完成，对于字符检查：首先遍历所有文本文件，通过字典保存减字和音符类型，最后只需要检查字典中的键值是否正确即可。对于音符类型检查：在遍历过程中计算每个小节的音符时值之和，和为奇数时，说明该小节为非整数拍，此时需要与原谱进行核对，根据核对情况调整数据文件内容。

正确性核验可以保证人工录入数据的正确性，能够规避键盘输入的符号错误。由于古琴音乐节奏自由，每个小节不一定是标准的整数拍，因此对音符类型的检验更加复杂，在发现非整数拍的时候还需要增加对原谱的核对环节。

逻辑性核验可以检查由于连续录入误差而产生的逻辑错误，找出正确性核验阶段漏检的错误。如将“四分音符”“八分音符”输入成连续的两个“八分音符”，正确性核验阶段会将其判断为整数拍，但原谱是非整数拍。因此需要增加人工检查，以保证录入数据的逻辑正确性。

四、预训练减字语言模型

针对古琴减字的序列标注问题，引入基于自监督的预训练语言模型能够缓解下游任务标注数据少的问题。由于古琴减字的独特性，目前没有基于古琴减字语料训练的语言模型。因此，本文依据减字的特点设计了基于 BERT Encoder 的语言模型——QinBERT。

（一）减字字符编码

通常情况下，减字由四部分组成：左手指法、徽位、右手指法、弦位，示例如图 6 所示。古琴有 7 根弦，13 个徽位，几十余种左右手指法，自由排列构造出的减字非常丰富。如果直接将减字当作 token，词表的复杂度会很大，语言模型难以从有限的语料中学习各个减字之间的关系。



图 6 减字结构解析示例

本文使用一种解构减字的方法对减字进行编码，只需要将左右手指法、徽位和弦位确定，即可表达出一个动作。对于古琴减字谱中非指法类的减字，则直接填充四个部分。如泛音开始的提示“泛起”，左右手指法、徽位和弦位均填充为“泛起”。

数据录入时，没有明确指定每个减字对应的四个部分。因此，在进行编码时首先构建指法字典，再利用 jieba 分词器对减字进行分词，从而得到解构的减字编码。例如“大指七徽九分挑四弦”会被切分为“大指\七徽九分\挑\四弦”，得到的四个部分则分别对应了编码当中的左手指法、徽位、右手指法和弦位。

同样的，这种方法无法应对特殊的减字。例如“撮无名指十徽八分三弦散五弦”，分词之后会得到多个动作组合，与编码的四个部分并不一致。除了指法“撮”，类似的还有“拨”“刺”等，会出现一个动作产生多个音符的情况。为了应对这种情况，本文将这类特殊减字进行化简，由于右手指法的动作更具有代表性，故将其编码的四个部分全部填充为右手指法。在尽可能保留指法信息的情况下，让解构减字的方法能够处理所有减字。

（二）语言模型结构

如图 7 所示，QinBERT 模型的主干结构采用了 BERT 模型的编码器，Embedding 部分将左右手指法、徽位和弦位 4 个输入分别进行嵌入表达，随后将四个嵌入表示向量加在一起并输入一个全连接层。完成 token 嵌入表达之后，为了引入序列位置信息，在 token 嵌入向量当中加入相对位置嵌入向量，随后一起输入到若干个 BERT 编码器中进行特征提取。特征提取之后，将 BERT 隐藏层的输出向量分别输入到 4 个全连接层当中，对掩码部分进行预测。

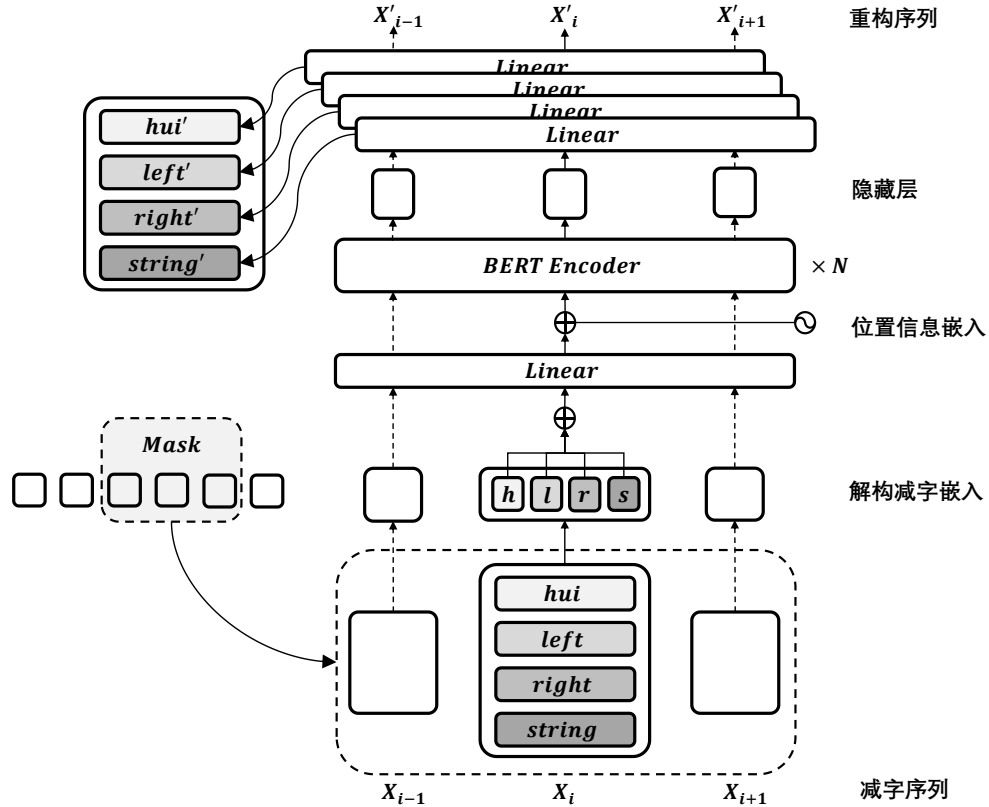


图 7 QinBERT 减字语言模型结构

（三）遮罩策略设计

受 BERT 中的掩码语言模型的启发，本文随机掩码输入标记序列中的一些元素并预测被掩码的元素。一种简单的屏蔽策略是随机屏蔽一些标记如图 4 所示。但是由于古琴减字谱存在顺延的记录方法，随机屏蔽元素会导致信息泄漏，从而无法很好地学习上下文表示。

因此将同一组减字表达视为一个单元并同时被掩蔽，这样可以避免信息泄漏，通过预训练可以学习到更好的上下文表示。每次按照对于被掩蔽的元素，其中 80% 被替换为 MASK，其中 10% 被替换为随机元素，10% 保持不变。本文参考 RoBERTa 的预训练任务，在预训练中移除了下一句预测任务。

五、基于预训练语言模型的打谱模型

（一）整体打谱流程

对于输入的减字序列，图 8 描述了 QinBERT 打谱的整体流程。首先用不含标注信息的减字序列构造掩码样本，通过预测掩码的预训练任务完成 QinBERT 的训练。之后对于不同的场景可以使用不同的方法获得打谱结果：全部减字序列由模型进行打谱时，直接将其输入到模型当中，即可获得预测标签；人工和模型混合打谱时，首先由打谱者标注一段减字序列，将其与未标注序列一起输入动态打谱模型，即可获得剩余未标注序列的预测标签。

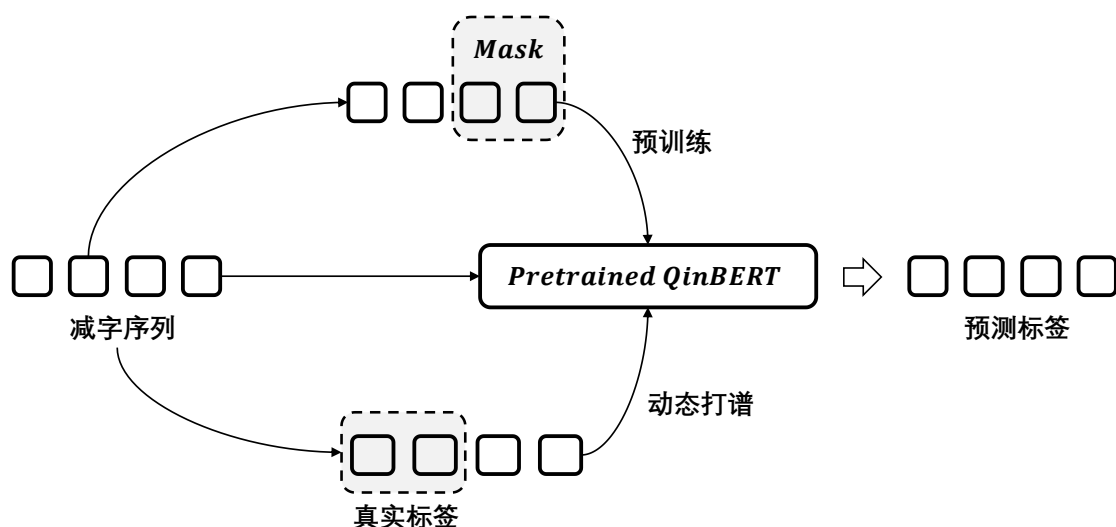


图 8 QinBERT 打谱流程

打谱者在使用模型辅助打谱时，可以选择先将待标注曲谱输入到 QinBERT 模型当中，得到初步的音符类型。经过人工调整之后，将带有部分人工校正后的标注连同剩余数据一起输入动态打谱模型。如此往复，可以完成整首曲子的打谱工作。

如果在打谱过程中遇到减字缺失或者可疑的谬误，可以使用预训练的语言模型的掩码预测功能对缺失或谬误进行预测，利用预训练模型中包含的上下文信息帮助打谱者对曲谱进行修正。

（二）打谱模型结构

本文的第一章第三节将计算机辅助打谱任务定义为序列标注任务，故将 QinBERT 隐藏层输出输入到 CRF 层当中进行解码，得到最终的序列标签，模型结构如图 9 所示。

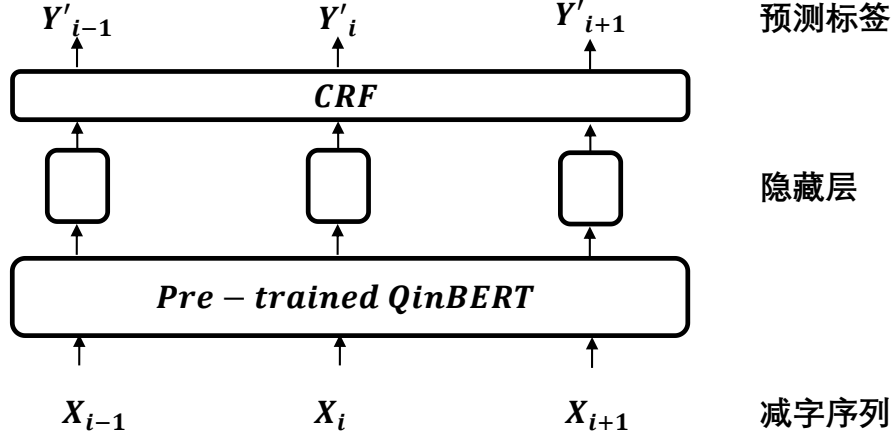


图 9 减字打谱模型结构

CRF 是一种马尔可夫随机场模型，与传统的 softmax 标签输出相比，CRF 通过概率计算能够保证输出标签的逻辑性^①。例如在 BIO 序列标注中，I 标签不会独立出现在 B 标签之前，即训练数据没有出现这样的样本，那么这一个概率路径就不会被模型选中，从而确保 I 标签在 B 标签后出现。

给定序列 $X = (X_1, X_2, \dots, X_n)$ 和标签序列 $Y = (Y_1, Y_2, \dots, Y_n)$ ，其条件概率 $P(Y|X)$ 可由公式(1)计算。

$$P(Y|X) = \frac{1}{Z(X)} \exp \left(\sum_{i,k} \lambda_k t_k(Y_{i-1}, Y_i, X, i) + \sum_{i,l} \mu_l s_l(Y_i, X, i) \right) \quad (1)$$

$$Z(X) = \sum_Y \exp \left(\sum_{i,k} \lambda_k t_k(Y_{i-1}, Y_i, X, i) + \sum_{i,l} \mu_l s_l(Y_i, X, i) \right)$$

其中 t_k 和 s_l 为特征函数， λ_k 和 μ_l 为对应权重， $Z(X)$ 为规范化因子。计算条件概率 $P(Y|X)$ 的过程为概率最大的最优路径求解问题，通过基于动态规划的维特比算法可以得出解码序列 $(Y'_1, Y'_2, \dots, Y'_n)$ 。

^① 李航,统计学习方法,第二版.北京:清华大学出版社,2019: 218-220.

（三）动态打谱模型

1. 动态打谱机制

由于古琴流派众多，各个流派的演奏风格不同，对相同乐句的处理方式不同。因此在进行打谱时，打谱者会根据流派的习惯和风格进行调整。为了满足不同流派的打谱需求，本文提出一种动态打谱模型，通过已标注的序列段落对未标注的序列段落进行调整。打谱者可以由人工完成一定量的打谱工作，再将标注好的数据与之后未标注的数据一同输入动态打谱模型，模型根据已经完成的段落，调整标签输出的概率分布，以达到风格延续的效果。动态打谱过程如图 10 所示。

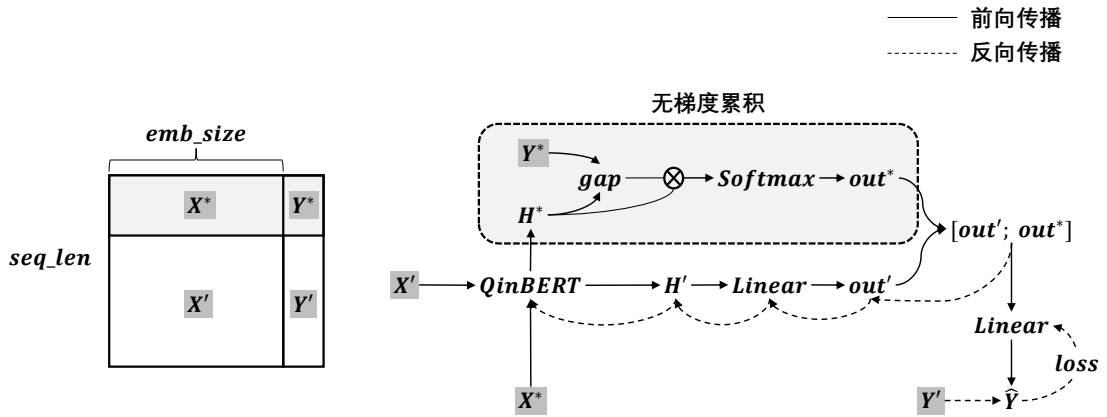


图 10 动态打谱机制

对于已经标注好的数据，将减字序列和对应的标签一同输入 QinBERT 模型。使用模型的编码器对标注好的减字序列进行编码，得到隐藏层输出之后计算与真实标签之间的交叉熵，以此度量模型与真实值之间的差距。随后将交叉熵与激活值相乘，再经过 Softmax 函数归一化，得到标签差距概率分布向量 out' 。 out' 的计算过程如公式(2)所示。

$$\begin{aligned} gap &= - \sum Y' \cdot \log(H') \\ out' &= \text{softmax}(gap \cdot H') \end{aligned} \quad (2)$$

对于未标注的数据，在得到 QinBERT 编码的隐藏层激活输出之后，将标签差距概率分布向量与未标注输出的概率分布向量进行拼接融合。得到融合向量之后，将其输入到全连接层，得到最终的标注输出。

2. 优化过程分析

对动态打谱模型的优化过程进行分析，有利于理解动态机制在模型当中的作用。首先通过抽象函数对模型反向传播过程进行推导，随后打印出简化模型的 Pytorch 反向传播计算图进行印证。

动态打谱模型使用交叉熵损失函数作为优化目标，整个模型抽象为 f ，QinBERT 模型用 Q 来表示， θ_{Linear} 为融合向量拼接之后输入的全连接层参数。求导过程如公式(3)所示。

$$\begin{aligned}\frac{\partial Loss}{\partial \theta} &= \left(-\frac{Y'}{Q} + \frac{1-Y'}{1-Q} \right) \frac{\partial f}{\partial \theta} \\ \frac{\partial f}{\partial \theta} &= \theta_{Linear} \left[\frac{\partial Q(X')}{\partial \theta}; Q(X^*) \right]\end{aligned}\quad (3)$$

由于标注数据输入后只计算与真实值之间的差距，并不对模型进行参数优化，所有对应部分的拼接向量不产生梯度。动态打谱模型在进行反向传播时，损失信息通过最后一个全连接层传递到 QinBERT 当中，而标注序列产生的梯度不会向前传递。这就避免了训练过程中标签泄露的可能性。

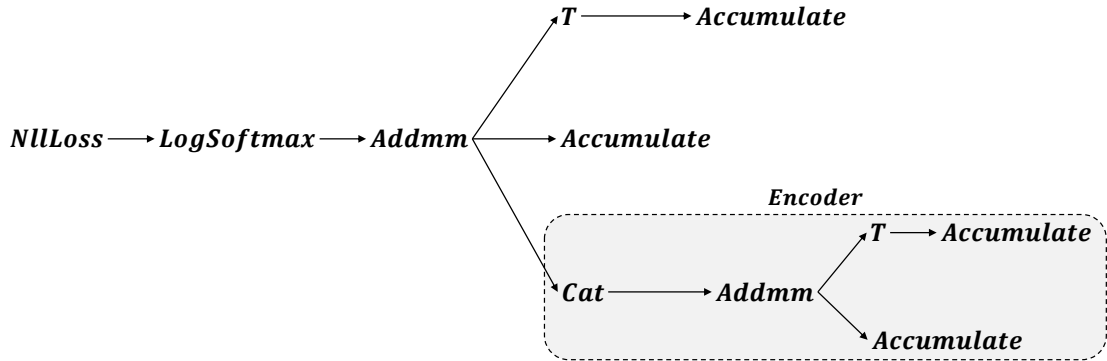


图 11 动态打谱机制反向传播计算图

由于 QinBERT 模型在反向传播过程中，涉及的模块非常多，分析起来较为复杂，故使用简单的全连接层代替 QinBERT 作为编码器，以便于分析和理解动态打谱机制。

图 11 展示了简化后动态打谱机制在 Pytorch 中的反向传播计算图。Pytorch 在进行前向计算时构建梯度的有向计算图，反向传播时从叶子节点出发进行梯度自动计算。从图 11 中可以看到，在 Cat 拼接算子之后是全连接层的优化子图。也就是说，在图 10 中 gap 的计算过程并没有被 Pytorch 的计算图记录下来。

据此可以推断出，QinBERT 在进行反向优化时，gap 的计算同样没有形成计算图节点。因此动态打谱机制的反向传播与一般的打谱模型速度接近，而且产生的计算图大小相近，运算所占用的内存也相近。

六、实验及结果

（一）实验设置

MusicBERT 的实验当中指出：更短的序列长度可以使模型一次性处理更长的音乐片段，从而能够提升模型对音乐的理解能力。因此本文设置三个对照组进行组内和组间的对照，以检验本文提出方法的性能，具体设置如下：

对照组一：探究不同序列长度对打谱效果的影响。使用不同长度的序列输入，对打谱模型进行训练。短序列为 1 个小节，长序列为多个小节及整首曲子。本组实验控制 QinBERT 层数为 12 层。

对照组二：不同的预训练语言模型对下游任务的影响。将基于静态词向量与动态词向量的语言模型进行对比，静态词向量设置 Word2Vec 和 GloVe，动态词向量设置不同参数量的 BERT。本组实验控制输入序列为长序列。

对照组三：动态打谱模型中不同 gap 计算策略对打谱效果的影响。在 gap 的计算中，除了使用交叉熵度量输出和真实值的差距，还可以使用 KL 散度或者绝对值。不同的人工标注完成度对打谱效果也会有影响，由于引入的人工标注，测试集的分布会产生变化。因此每个完成度都对应一个空白对照组，即在微调模型的基础上之间删除前段数据，以保证模型性能指标的可比性。本组实验控制人工标注比例为 15%。

（二）实验过程

实验环境为：Linux(Ubuntu)，GPU 为显存 24GB 的 NVIDIA GeForce RTX 3090，CPU 为内存 63G 的 2 核 Intel(R) Xeon(R) Platinum 8156。

首先对收集到的数据进行核验，修正其中的错误之后，利用 jieba 对减字进行分词，构造解构的减字数据集。随后训练语言模型：静态词向量训练 Word2Vec 的 skip-gram 模型和 CBOW 模型，以及 golve 模型；动态词向量训练 BERT 和 XLNet。完成语言模型的训练之后，对预训练的模型进行下游任务的微调。实验的主要参数设置如表 2 所示：

表 2 实验关键参数设置

参数	参数值
动态语言模型预训练 epoch	5
长序列下游任务微调 epoch	200
短序列下游任务微调 epoch	50
动态语言模型预训练 batch size	128
下游任务微调 batch size	64
动态语言模型预训练学习率	2e-5
下游任务微调学习率	1e-4
动态语言模型维度	768
多头自注意力头数	12

对于下游的微调任务，为了避免模型在训练过程中陷入局部最优或者局部平坦区域，使用带热重启的余弦退火算法对学习率进行动态调整。在进行大轮数模型训练时，该方法可以有效地提升模型泛化性能，尤其在后期训练当中可以通过震荡学习率使模型脱离局部最优。带热重启的余弦退火调整学习率效果如图 12 所示。

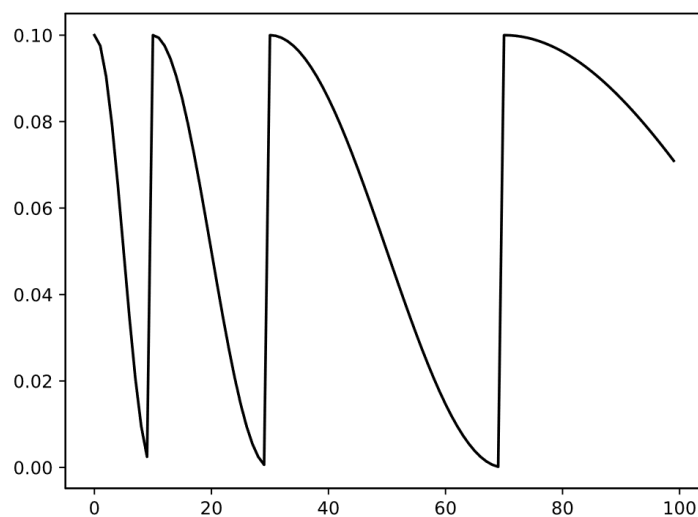


图 12 带热重启的余弦退火调整学习率效果

（三）实验结果与分析

实验当中的下游微调任务采用留一法进行数据集划分，即每次选择 1 首琴曲作为测试集，最后计算各个性能指标的平均值^①。度量打谱模型的性能主要采用宏平均的 Precision、Recall 和 F1。

表 3 各组对照实验结果

实验分组	Method	Precision	Recall	F1
对照组一	短序列 QinBERT	0.33	0.35	0.31
	长序列 QinBERT	0.43	0.44	0.40
对照组二	Word2Vec +BiLSTM	0.22	0.27	0.24
	Word2Vec+Transformer	0.27	0.39	0.30
	GloVe +BiLSTM	0.21	0.25	0.22
	Glove+Transformer	0.28	0.31	0.30
	6 层 QinBERT	0.38	0.40	0.34
	12 层 QinBERT	0.43	0.44	0.40
对照组三	Cut QinBERT	0.45	0.48	0.42
	Dynamic QinBERT (KL)	0.47	.044	0.41
	Dynamic QinBERT (CrossEntropy)	0.49	0.48	0.44

从第一组实验可以发现长序列模型的微调效果优于短序列，这与 MusicBERT 实验得出的结论不同：更短的序列长度可以使模型一次性处理更长的音乐片段，从而能够提升模型对音乐的理解能力。笔者认为原因在于古琴减字序列与 MIDI 序列存在差异，古琴减字谱存在多对一的音符对应关系，因此需要更长的序列长度使数据段落包含更多信息。

从第二组实验结果来看，使用静态词向量的模型性能普遍不佳。一方面，本文收集到的数据并不多，静态词向量技术难以捕捉到减字的上下文关系；另一方面，多层词向量的叠加再表达丰富的减字组合上能力不足。与第一组实验对照起来，动态语言模型显然学习到了更多减字序列的上下文信息，且参数量大的模型学习到更多减字序列间的语言关系。

从第三组实验结果来看，动态打谱方法在测试集上的效果普遍较好。这与需要标注段落的位置和数量有关：截断后的后半序列标签更少，信息在自注意力机制的特征抽取下更加集中。与其他两组实验对照来看，动态打谱机制能够在一定程度上延续打谱者的标注风格，并运用在后半段的标注上。

^① 周志华. 机器学习. 北京：清华大学出版社, 2016: 25-27.

七、总结与展望

（一）工作总结

本文围绕古琴打谱的相关问题展开研究，包括计算机打谱的任务定义、数据录入、减字谱数据编码、减字语言模型设计、语言模型遮罩策略的设计、打谱模型的设计以及一种解决风格延续的动态打谱机制。

在计算机辅助打谱的任务定义中，本文讨论分析了不同定义的优劣，最终将其确认为一一对应的序列标注任务。

有关数据录入方法的部分，本文提出了保证数据质量的两阶段复核方法，通过正确性核验阶段保证人工输入的正确性，逻辑性核验阶段检查录入过程中产生的节拍逻辑错误。相应地，本文为减字谱数据设计了新的编码方式，通过分词的方法对减字组合进行解构，以到达降低词表复杂度的目的，有效提升了后续语言模型的训练效率。

由于计算机与古琴结合的研究工作还停留在表示学习的初级阶段——符号表示。为了跟进自然语言处理研究现状，本文按照自然语言表示学习的发展路线对减字语言模型进行了初步探索，提出了与解构减字契合的 QinBERT。

除了完成基本的打谱序列标注流程，本文从打谱需求的角度出发，进一步探索了动态打谱机制及其原理。通过可视化其梯度计算的动态图，分析出动态打谱各个模块梯度传播情况。

本文首次完整地提出并实现了计算机辅助打谱的工作流程，并且整理核对出一套古琴减字序列标注数据集，为后续工作的开展积累了经验。

（二）展望

作为计算机辅助打谱工作的先行者，本项目并没有任何前期工作可以参考，因此从算法层面和工程层面提出展望，为后续研究工作的开展提供思路。

从算法层面来说，本项目围绕序列标注这一核心问题开展了一系列对比实验，但仍然有不少空白需要填补。未来研究方向总结为以下三点：

从实验结果来看，模型的整体性能似乎一般。这是因为古琴打谱本身非常自由，同时也受到流派风格差异影响，打谱的结果并没有绝对的对错之分。在数据量较小的情况下，预训练语言模型能够超过基线模型，说明古琴减字序列是可以用自然语言的方法来进行建模的。在未来的研究工作中，可以用更多的减字数据训练语言模型。

关于语言模型的训练中，输入序列的长度会影响下游任务性能。实验发现，较长的序列输入需要更长的训练轮数才能达到较优的性能。笔者尝试过用长短序列交替训练，希望模型可以学习到不同长度的序列信息，但是并没有取得良好性能。后续的研究也可以关注如何有效利用不同长度的序列。

动态打谱机制是更加深入的打谱方法研究，本文梳理了该机制的数学原理，并通过计算图印证了动态机制的优化过程。在实际打谱过程中，打谱者不一定会从曲子的开头进行打谱。往后还可以探究不同位置和完成度对打谱效果的影响。

从工程层面来说，本项目的数据录入工作完全由人工完成，效率不高。后续的拓展工作思路总结为以下两点：

针对减字数据较少的问题，后续的研究工作还可以围绕古琴减字 OCR 技术展开。现有的工作只完成了古琴减字谱字符图像的锚定框选，并没有做到减字编码的输出。这与相关数据集的缺失有很大关系，在本项目基础上补充对应的图像数据即可开展后续研究。

针对减字数据录入难问题，除了需要 OCR 技术的辅助，还需要有相关的信息系统协助数据的复核及纠正工作。另外，如果 OCR 技术难以对音符类型进行判别，系统还可以引入人工判断音符类型的功能。通过按钮点击即可完成对应减字的音符类型录入。

综上两个层面所述，如果未来的研究工作能够解决这些问题，那么我国的古琴艺术会迎来新的发展生机，我国的信息智能技术在非物质文化遗产传承和保护能力上将走在世界前列，人类精神文明财富的流传则存在更多的可能性。

主要参考文献

- [1] 陈长林. 电脑在古琴音乐研究中的初步应用[J]. 计算机学报,1989,(07):525-533.
- [2] 陈根方. 古琴艺术数字化保护概述与琴律智能分析[J]. 复旦学报(自然科学版),2019,58(03):335-342.
- [3] 丁晓君,叶婷婷,周昌乐. 古琴减字谱的编码与编辑方法[J]. 中国音乐学,2008,(02):93-96.
- [4] 顾亚丽. 基于遗传算法的古琴减字谱曲段自动生成[D].厦门大学,2009.
- [5] 高宋梯,周昌乐. 现代乐谱的古琴减字谱自动翻译研究[J]. 心智与计算,2010,4(01):45-54.
- [6] 顾亚丽,周昌乐. 古琴曲语料库的构建与规则提取[J]. 心智与计算,2010,4(02):128-138.
- [7] 胡丹. 古琴减字谱的数字化应用[J]. 印刷杂志,2014,(05):59-62.
- [8] 刘洋,唐英敏. 一种古琴减字谱符号的结构化描述方法[J]. 北京信息科技大学学报(自然科学版),2010,25(S2):61-65.
- [9] 吕沁融. 试论中国传统音乐内容传播的互联网创新产业模式——以古琴数字化为例[J]. 编辑学刊,2017,(05):67-71.
- [10] 李航,统计学习方法.第二版.北京:清华大学出版社,2019: 218-220.
- [11] 李祥霆,龚一.古琴考级曲集 2. 北京:人民音乐出版社,2010.
- [12] 李圣辰,吴雨松. 1 个中国古琴曲的符号化音乐数据集介绍及其应用实例[J]. 复旦学报(自然科学版),2020,59(03):276-285.
- [13] 朴东生,张殿英.古琴考级曲集 1. 北京:人民音乐出版社,2010.
- [14] 喻辉.古琴减字谱音高信息电脑自动处理原理[J]. 交响.西安音乐学院学报,1993,(03):56-58.
- [15] 庄雄雄,周昌乐,齐京峰,高宋梯,吕兰兰. 提取古琴减字音高的计算方法及其软件实现[J]. 计算机应用与软件,2010,27(04):10-12+37.
- [16] 张英俐,刘弘. 一种基于遗传算法的古琴曲生成方法[J]. 信息技术与信息化,2018,(09):28-30.
- [17] 周志华. 机器学习. 北京:清华大学出版社,2016: 25-27.
- [18] Claude E Shannon. A mathematical theory of communication. Bell system technical journal, 27(3):379-423, 1948.
- [19] Chuan C H, Agres K, Herremans D. From context to concept: exploring semantic relationships in music with Word2Vec[J]. Neural Computing and Applications, 2020, 32(1).
- [20] Chou Y H, Chen I C, Chang C J, et al. MidiBERT-Piano: Large-scale Pre-training for Symbolic Music Understanding[J]. 2021.
- [21] Huang C, Duvenaud D, Gajos K Z. ChordRipple: Recommending Chords to Help Novice Composers Go Beyond the Ordinary[C]// the 21st International Conference. ACM, 2016: 241-250.

- [22] Herremans D, Chuan C H. Modeling Musical Context with Word2Vec[C]// First International Workshop on Deep Learning and Music. 2017.
- [23] James L McClelland, David E Rumelhart, PDP Research Group, et al. Parallel distributed processing. Explorations in the Microstructure of Cognition, 2:216–271, 1986.
- [24] Liang H, Lei W, Chan P, et al. PiRhDy: Learning Pitch-, Rhythm-, and Dynamics-aware Embeddings for Symbolic Music[J]. ACM, 2020:574-582.
- [25] Liu Z, Lin Y, Sun M. Representation Learning for Natural Language Processing. 2021: 4-7.
- [26] Madjiheurem S, Qu L, Walder C. Chord2Vec: Learning Musical Chord Embeddings[C]//In Proceedings of the constructive machine learning workshop at 30th conference on neural information processing systems (NIPS2016), 2016.
- [27] Tatsunori, Hirai, Shun, et al. Melody2Vec: Distributed Representations of Melodic Phrases based on Melody Segmentation[J]. Journal of Information Processing, 2019, 27:278-286.
- [28] Tsai T, Ji K. Composer Style Classification of Piano Sheet Music Images Using Language Model Pretraining[J]. 2020:176-183.
- [29] Zellig S Harris. Distributional structure. Word, 10(2–3):146–162, 1954.
- [30] Zeng M, Tan X, Wang R, et al. MusicBERT: Symbolic Music Understanding with Large-Scale Pre-Training[J]. 2021.