引用格式： 张勇，王益谊，于钢，等. 结合对比学习和多任务学习的大语言模型分类技术在标准检索系统中的应用[J]. 标准化学报，2026(8)：63-71.
ZHANG Yong, WANG Yiyi, YU Gang, et al. Application of Large Language Model Classification Technology Combining Contrastive Learning and Multi-task Learning in Standard Retrieval System [J]. Journal of Standardization, 2026(8): 63-71.

结合对比学习和多任务学习的大语言模型分类技术在标准检索系统中的应用

张勇　王益谊　于钢　李娟　张阳

（中国标准化研究院）

摘　要：【目的】基于关键词的标准全文检索系统在处理用户查询时，需要从全量数据中进行检索，会导致检索系统性的召回率降低。【方法】为了解决此类问题，本文提出了一种结合对比学习与多任务学习的分类算法，将用户查询映射到中国标准分类号（CCS）分类体系，从而减小检索范围。利用层次感知的对比学习策略和知识注入的多任务框架提升该方法对CCS分类体系的细粒度类别的分类能力。【结果】该方法在用户查询分类场景的测试指标Micro-F1为89.23%；在应用线上标准检索场景后，Recall@5和MRR分别提升10.3%和9.5%。【结论】通过本文的多种实验，证明了该方法在提升基于全文检索的标准检索系统的有效性。

关键词：分类；对比学习；多任务学习；CCS

DOI编码：10.3969/j.issn.2097-857X.2026.08.008

Application of Large Language Model Classification Technology Combining Contrastive Learning and Multi-task Learning in Standard Retrieval System

ZHANG Yong　WANG Yiyi　YU Gang　LI Juan　ZHANG Yang

(China National Institute of Standardization)

Abstract: [Objective] When processing user queries, a standard keyword-based full-text retrieval system needs to perform retrieval over the full dataset. [Methods] To address this problem, this paper proposes a classification approach that combines contrastive learning and multi-task learning to map user queries into the Chinese Classification for Standards (CCS) taxonomy, thereby reducing the retrieval scope. A hierarchy-aware contrastive learning strategy and a knowledge-injected multi-task learning framework are further introduced to enhance the proposed method's ability to classify fine-grained categories within the CCS taxonomy. [Results] The proposed

基金项目：本文受中国标准化研究院基本科研业务费项目“面向大模型的结构化XML标准文档可读可理解工具研发”（项目编号：252025Y-12532）资助。

作者简介：张勇，硕士，工程师，研究方向为标准智能化。
王益谊，博士，研究员，研究方向为标准数字化、标准战略与理论。
于钢，硕士，副研究馆员，研究方向为标准信息化。
李娟，硕士，高级工程师，研究方向为标准信息化与智能化。
张阳，博士，助理研究员，研究方向为标准智能化。

method achieves a Micro-F1 of 89.23% on the user query classification task; after being deployed in an online standard retrieval setting, Recall@5 and MRR increase by 10.3% and 9.5%, respectively. [Conclusion] Extensive experiments demonstrate that the proposed method effectively improves standard full-text retrieval systems.

Keywords: classification; contrastive learning; multi-task learning; CCS

0　引言

标准作为国家技术体系和社会经济活动的基础性技术文件，是连接科技创新与产业应用的关键桥梁。据统计，我国现行有效的国家标准已逾数万项，覆盖了从基础科学到尖端技术、从工业制造到社会服务的各个领域。如何从这个庞大、复杂的知识库中快速、准确地检索到所需标准，已成为科研人员、工程师及政策制定者面临的迫切需求。一个高效、智能的标准检索系统，不仅能够加速知识的传播与应用，更能为技术创新和产业升级提供坚实的数据支撑。

当前，标准检索系统的主流技术方案多采用Solr、Elasticsearch等传统全文检索引擎。这些引擎的核心机制是倒排索引和关键词匹配。虽然技术成熟、响应迅速，但在处理标准检索这类专业任务时，检索范围过大，会引入大量来自不相关领域的噪声文献，干扰用户筛选结果，降低检索精度（低准确率）。近年来，以BERT为代表的预训练语言模型被引入检索领域。然而，直接将其应用于全库范围的语义排序，会产生大量的计算开销。因此，先对用户查询进行精准的标准领域预分类，之后再进入真正的检索过程，是优化检索系统性能的有效路径。

针对上述挑战，本文将用户查询的领域分类问题，建模为一个基于CCS分类体系的层次化多标签文本分类任务，创新性地利用大语言模型（LLM）强大的理解和生成能力，结合对比学习与多任务学习，实现对用户查询的精准识别。本文的主要贡献如下：

（1）首次将用户查询分类算法应用于标准检索领域。本文创新性地将提升检索性能任务转化为基于CCS分类体系的层次化文本分类任务，验证了分类算法在此类场景中的有效性。

（2）设计了针对层次分类任务的对比学习采样策略。为了解决CCS体系中类别语义高度重叠的难点，设计了“锚点—正例—困难负例”的层次感知批次采样算法，将相似但不同的样本聚合在一个batch中，提升了模型对CCS分类体系细粒度类别的理解能力。

（3）利用大模型的特性构建知识注入与多任务学习框架。通过将单纯的分类数据，扩展为标准标题生成、分类号知识问答等多种数据，将单一的分类模型训练转向对标准领域知识多维度问答模型的训练，提升训练后模型的泛化能力和鲁棒性。

（4）通过完整的数据实验验证本文方法的有效性。本文的训练数据部分采用真实用户日志和合成数据组成的混合训练集，测试数据采用用户真实日志，通过分类系统的F1值、检索系统的召回率和MRR，证明在用户查询分类以及应用到实际检索系统的有效性。

1　现有研究综述

目前，标准领域检索系统的主流技术方案是将全量数据入库，每次检索都是从全量数据中检索。全文检索引擎[1-2]是较为通用的方式，如Solr，Elasticsearch。基于BERT[3]的深度学习也开始逐渐应用在检索领域[4]，全文检索引擎侧重于词语层面的完全匹配。“先意图分类、后检索”的模式是大型检索系统的关键组成部分。在通用网页搜索中，系统需要判断用户查询的具体类型，从而调用不同的服务模块。比如在电商搜索中，对查询进行

精准的类目预测更是核心环节，将查询“苹果”正确分类到“手机”而非“水果”类目，是保证召回率和用户体验的基础。这些研究表明，检索前置分类对缩小检索范围、提升检索效率与精度至关重要。然而，上述场景的分类任务多集中在通用或商品领域，其分类体系相对扁平，类别间的语义界限也较为清晰。对于标准检索这类高度专业化，以及分类体系呈严格层级结构且类别间存在大量专业术语重叠的领域，相关的研究尚不充分。直接应用通用领域的分类方法往往难以捕捉到细微的语义差异，导致分类精度不足。因此，如何设计一种能感知层次结构并深度理解专业领域知识的分类模型，是该技术在标准检索领域落地的核心挑战。

对比学习是深度学习中自监督学习的范式，其核心思路是“拉近相似的，推远不相似的”，最初在视觉领域应用较为广泛。SimCLR[5]通过简单对图像进行旋转、裁剪、颜色变换等操作构造正样本，将不同的图像作为负样本训练模型。SwAV[6]通过在不同视图之间交换聚类分配来进行对比学习，依托不同视图的表示实现相似数据的自动聚类。SimCSE[7]则是应用在自然语言处理领域，利用同一输入经过2次不同dropout操作来获取正样本，将不同的样本作为负样本。引入困难负例有助于模型更好地区分语义相近但实际矛盾的句子，从而学习到更具判别力的表示。Zhang等[8]提出了一个基于对比学习的框架，该框架从角度对比和三元组对比2个维度出发，优化特征表示。

在BERT模型中，多任务学习指的是模型在训练阶段同时设置多个目标函数，比如阿里巴巴在推荐领域开展的工作[9]，谷歌基于MOE架构提出的多任务学习[10]，腾讯也开展了相关的工作[11]。生成式大模型实现多任务学习的最主流、最有效的方式是在指令微调阶段构建一个混合的多任务指令数据集，为每种任务设计一个指令，将所有的数据按照各自的指令格式化，组成一个多任务的数据集。

2 本文方法

本文将通过训练一个CCS分类模型，对用户查询进行CCS编码分类，将检索文本划分到具体的CCS编码的领域中，在部分CCS编码领域进行检索，从而缩小全文引擎的检索范围，提升检索的精度。本文将利用对比学习技术及多种类型任务的模型训练数据，增强分类模型对CCS数据的语义理解能力，使其能够更精准地捕捉用户查询文本与CCS编码之间的关联。

模型训练时，采用对比学习方法，通过构建正负样本对，让模型学习到不同CCS编码下文本的差异特征，从而提升分类的准确性。同时，为了进一步丰富模型的训练数据，提高其泛化能力，还将生成多种类型任务的模型训练数据，对模型进行多任务联合训练。通过这2个阶段的大模型技术应用，本文旨在提出一个高效、准确的基于CCS编码的分类算法，为标准检索系统提供更精准的领域划分，从而减小检索数据的领域范围，提升检索系统的性能。

在推理阶段，本文的结构见图1。

2.1 对比学习方法应用

为增强模型在CCS层次化标签体系中的特征表示能力，本文借鉴对比学习的原理，对传统的随机批次训练范式进行优化。标准随机采样策略在构建批次时，无法保证包含能有效提升模型对细粒度类别辨别能力的样本组合。本文设计并实现了一种混合批次采样策略。该机制将训练批次分为2类：

（1）随机批次。遵循传统深度学习的训练方式，从训练集中完全随机地抽取样本构成批次，以保证模型的探索能力和泛化性。

（2）层次感知批次。通过“锚点—正例—困难负例”的方式构建批次数据。随机选择一个锚点数据，从与锚点同属一个最细粒度类别的样本中采样，构成正例。最后，关键一步是从与锚点共享二级或三级父节点的其他类别中进行采样，构

图1 本文方法结构图

成困难负例。在损失函数中最小化锚点与正例的距离，同时最大化其与困难负例的距离，该策略能有效提升模型的类内聚性和类间可分性，尤其是在区分语义高度重叠的“兄弟”类别方面效果显著。

在训练流程中，以各占50%的比例混合使用这2种批次，以平衡模型的全局泛化与局部辨别能力。Loss函数见式（1）~式（3）：

$$L_{\text{total}}=(1-\alpha)*L_{\text{SFT}}+\alpha*L_{\text{SCL}} \tag{1}$$

$$L_{\text{SFT}}=-(1/(L-L_{\text{prompt}}))*\sum_{i=L_{\text{prompt}}+1}^{L}logP(x_i|x_{<i}) \tag{2}$$

$$L_{\text{SCL}}=-\log\frac{e^{\text{sim}(h_i,h_i^+)/\tau}}{\sum_{j=1}^{N}(e^{\text{sim}(h_i,h_i^+)/\tau}+e^{\text{sim}(h_i,h_j^-)/\tau})} \tag{3}$$

如公式（1）所示，Loss分成了2个部分：L_{SFT}表示大模型指令微调时的交叉熵损失，L_{SCL}用于计算对比学习中的向量特征表示的差异性。α是一个超参数，用于平衡2个损失函数的权重。

公式（2）为一般微调阶段的交叉熵损失函数。因为模型在指令微调时要屏蔽掉指令部分，因此计算损失从第L_{prompt}个Token开始。

公式（3）中h_i^+是句子x_i的dropout数据，构成正

样本对。h_j^-是句子x_j的不相关数据，作为难负例。sim(·)表示向量相似度计算函数，τ是温度控制参数，可以调节对负样本的关注程度。

2.2 多任务学习的知识注入应用

为了增强模型的鲁棒性和泛化能力，模型不仅要能做出正确分类，更要理解分类决策所依据的领域知识。为此，引入了知识注入式多任务学习框架。通过将主分类任务与多种探究领域知识不同侧面的辅助任务联合训练，引导大语言模型从一个单纯的分类判别器，完成标准领域和CCS编码知识注入，转换为具备领域知识理解与推理能力的知识处理器。

本文将所有辅助任务的数据全部构造成“指令—回答”的数据对。在训练阶段，这些来自辅助任务的指令数据被混合在一起，构成一个统一的训练集。模型在训练时，接收一条指令，并被要求生成期望的回答，其优化目标统一为序列到序列的交叉熵损失（L_{SFT}）。本文的方法可以实现参数效率与知识共享，所有任务共享同一套模型参数，使得模型在参数空间中寻找一个能够同时满足所有任务需求的通用表示。本文设计的辅助任务旨在从不同维度向模型注入关于标准领域的核心知识，具体设计了4种辅助任务。

（1）标准文本的CCS分类。通过标准的摘要、标题、内容等字段，建立标准文本内容与其所属类别之间的直接映射，注入关于类别定义和范围的知识。

（2）标准标题生成。基于标准摘要、内容，训练模型的主题概括与抽象能力，并且能够让模型理解标准内容的核心要点，注入关于标准核心主题识别的知识。

（3）CCS分类号问答。直接给出分类号，让模型输出分类号的中文名称。使模型直接学习CCS分类体系本身的结构化知识和层级关系，该任务向模型注入关于分类法本体的知识。

（4）标准文本同类判别。给出2个标准的内容、标题、摘要，判断2个标准是否属于同一CCS类别。该任务强化模型在细粒度上的类别边界感知能力是对比学习的补充。

4种辅助任务的数据示例见表1。

同时，为了确保合成数据的质量、多样性以及与真实标准内容的强相关性，本文设计了一套以标准内容为输入的提示模板框架。该框架旨在让教师模型根据一份具体的标准文本，逆向生成用户可能会使用的查询。本文设计了三类模板，以模拟不同认知水平和查询习惯的用户，具体见表2。

2.3 推理阶段的输出方式设计

在推理阶段，模型的输出方式直接影响其在

表1 辅助任务

任务类型	输入	输出
标准文本的CCS分类	请对以下标准摘要进行CCS分类。\n摘要：本标准适用于饮用天然矿泉水水源水及其直接提供给消费者的所有预包装产品的测定	X51–饮料制品
标准标题生成	请为以下标准内容生成一个简洁的标题。\n内容：本部分规定了GB 150适用范围内的钢制压力容器的制造、检验与验收要求；其他材料制压力容器的制造、检验与验收要求按相关标准。本部分适用的压力容器结构形式为单层焊接压力容器、锻焊压力容器和多层压力容器（包括多层筒节包扎、多层整体包扎、钢带错绕和套合容器）	铸件X射线底片评定方法
CCS分类号问答	请解释CCS分类号“A50”代表什么	计量综合
标准文本同类判别	判断标准A和标准B是否属于同一个最细粒度的CCS类别？\n标准A：《焊接接头金相检验方法》\n标准B：《压力容器焊接工艺评定》\n回答：是	是

表2 用户查询生成提示模板

类型	目标	提示词
专家问题	模拟相关领域专家进行检索查询时的输入	角色：你是一个标准检索领域的专家。 Context: 我将提供一项国家标准以及对应的【CCS编码】。 Task: 请你站在专业工程师或科研人员的角度，根据这份标准的核心内容，生成5条高度专业化的检索查询。 Requirements: 查询必须精确、简短。 查询应包含标准摘要中的关键技术术语、参数或标准方法。 查询风格应符合专业人士的搜索习惯
新手问题	模拟普通用户进行标准检索时的输入	角色: 你是一位善于模拟不同用户角色的AI助手。 Context: 我将提供一项国家标准以及对应的【CCS编码】。 Task: 请你模拟一个初学者或非专业人士（例如：学生、一线操作员），根据这份标准所解决的实际问题，生成5条口语化的检索查询。 Requirements: 查询应使用日常用语，避免使用专业术语。 查询可以是一个问题，描述一个场景，或一个模糊的需求
困难问题	同时选取2篇“兄弟”类别的标准文献，生成语义上容易混淆的困难查询，强化模型的相似类别的区分能力	角色：你是一位精通分类法、善于制造困难样本的AI专家。 Context: 我将提供2项标准，包括标准内容和对应的CCS编码，这2项标准的CCS类别为“兄弟类别”（类别含义有较强的关联性和相似性）。 Task: 请你根据2项标准的内容，生成“困难”检索查询。 Requirements: 结合2项标准，查询中必须巧妙地包含一些可能与【干扰项CCS编码】相关的词汇或概念，使其容易被误分类

实际应用中的可用性和与下游检索系统的集成效率。根据CCS分类体系存在的层级结构，以及大语言模型输出格式的多样性，本文设计了以下多种输出方式，并通过实验对比其优劣。

第一种是纯分类号输出方式，例如：“A29”。指令模型仅输出预测的最细粒度CCS分类编码，例如直接返回“A29”。这种方式输出字符串极短，且分类号作为唯一标识符完全避免了自然语言的歧义，便于下游系统进行快速、低开销的解析。然而，这种方式缺少人类可读的语义信息和分类体系的层级上下文，使得结果的直观性和可调试性变得极差，出现分类错误，人工分析错误的代价也较大。

第二种是编码与中文混合输出方式。通过指令模型生成一个由特定分隔符连接的半结构化文本，如“A29—材料防护”。这种格式参考思维链的方式，同时输出中英文，帮助模型在输出时思考。但是在做输出结果的解析时，需要对输出进行中英文匹配校验，并加入错误处理机制。

第三种是结构化JSON输出方式。例如{"code": "A29", "name": "材料防护"}。JSON数据格式固定，解析方便，并且后续的扩展性较好，可以持续扩展层级输出、置信度输出等信息。但是其输出体积略大于纯文本，需要模型消耗更多的输出Token。

3　实验设置及结果分析

3.1 实验数据

借助于“国家数字标准馆”的标准检索系统，本文获取了大量的真实用户查询日志，但是真实日

志的标注成本较高，因此本文构建了一个合成数据与真实数据的混合训练集。通过多种不同的提示模板，使用DeepSeek V3.2进行训练数据合成，生成了覆盖CCS全体系约2万条模拟训练数据。为了模拟实际检索系统的真实用户，这些提升模板包含不同的难度和不同的语言风格。同时，为了保证模型的鲁棒性，本文收集了约1 500条经脱敏处理的真实用户检索日志。

本文从线上标准检索系统收集了1 249条真实用户检索日志，并且由3位标准化领域的专家进行交叉标注，构建了高质量的测试集。

3.2 实验设置

本文使用千问3系列模型进行全量参数的指令微调。设计了多种不同维度的实验来全面评估本文方法的有效性，同时在分类阶段和检索阶段通过多种指标评价本文方法带来的提升。

3.2.1 消融实验及分析

（1）对比学习和多任务学习的消融实验

为了验证对比学习和多任务学习2个核心模块的有效性，在第3.1节所述的测试集上使用QWEN3-14模型进行了消融实验。所有模型均采用JSON输出格式进行评估，结果如表3所示。

基线方法为使用3.1节中构建的训练数据进行全量参数微调训练。

从表3可以看出，本文提出的完整方法在所有分类指标上均显著优于所有对比模型，其中Micro-F1达到了89.23%，相比仅进行标准微调的基线方法提升了6.68%，证明了所提方法的有效性。

“本文对比学习”相比“基线方法”有显著提升，层次感知的对比学习策略，尤其是“困难负例”的引入，有效增强了模型在区分相似类别等语义相近、易混淆类别上的辨识能力。

“本文多任务学习”使用第2.2节提及的任务构建了2万条训练数据，再结合第3.1节的3万条合成数据，共计5万条数据进行了模型训练。相比“基线方法”也有明显提升，说明通过引入标准标题生成、分类号问答等辅助任务，模型能够学习到更深层次的领域知识，从而提升了主分类任务的性能。

“本文对比学习+多任务学习”的组合取得了最佳性能，证明了精细化的对比学习与领域知识显式注入的多任务学习之间存在显著的协同效应，二者共同推动模型性能的跃升。

（2）不同输出格式的对比实验

为了探究第2.3节中设计的3种输出格式对模型性能及下游集成的影响，对完整模型进行了输出格式的对比实验，结果见表4。

表4 输出格式对比结果

输出格式	Micro-F1	解析错误率	平均输出Token数
纯分类号	89.23%	0.13%	3.3
编码与中文混合	89.87%	1.04%	15.7
结构化 JSON	91.36%	0.16%	25.5

由以上实验可以看出。纯分类号格式虽然简洁、Token消耗少，但在F1指标上略有下降，这可能是因为缺乏语义上下文的约束，模型更容易出错。

编码与中文混合格式的F1分数有所提升，证明思维链式的输出有助于模型进行更准确的推理。但其半结构化的特点导致下游解析错误率较

表3 模型各核心模块的消融实验结果

单位：%

方法	Accuracy	Micro-Precision	Micro-Recall	Micro-F1
基线方法	82.39	82.14	82.97	82.55
本文对比学习	86.21	86.63	85.61	86.13
本文多任务学习	87.93	87.68	86.9	87.34
本文对比学习+多任务学习	89.51	88.58	89.87	89.23

高，需要设计复杂的规则进行处理。

结构化JSON格式在取得最高F1分数的同时，使解析错误率控制在0.16%。虽然Token消耗略高，但其鲁棒性、可读性和后续扩展性（如增加置信度字段）是最好的。因此，本文将其作为与下游检索系统集成的最终方案。

3.2.2 下游检索系统性能提升实验

为了最终验证本文方法对标准检索系统性能的实际提升效果，本文将表现最佳的分类模型（采用 JSON 输出格式）集成到检索流程中，形成“先分类、后检索”的二级检索系统。实验在与线上的真实数据一致的测试环境中进行，结果见表5。

表5 检索系统性能提升实验

检索系统	Recall@5	MRR
Elasticsearch （Baseline）	72.4%	0.674
Elasticsearch+本文方法	82.7%	0.769
性能提升	+10.3%	+9.5%

接入了本文提出的查询分类模型后，检索系统的Recall@5提升了10.3%，MRR提升了9.5%。这证明通过预先缩小检索范围，可以有效过滤大量不相关领域的数据，从而提升检索结果的召回率和排序质量。同时，平均查询时延从158 ms降低到65 ms，速度提升了近60%。

3.2.3 不同尺寸模型的对比实验

为了不影响线上检索服务的并发性能，本文测试了千问多种不同尺寸的大模型在大量并发下的吞吐量以及性能表现。根据线上服务推测，模型的最大并发大约是30。测试的硬件环境为华为昇腾910B3，推理框架为VLLM-ASCEND。结果见表6。

表6 不同尺寸模型对比实验

模型	Micro-F1（分类指标）/%	Recall@5（检索系统指标）/%	平均延迟/ms
QWEN3-1.7B	88.14	79.31	160
QWEN3-4B	90.34	81.43	263
QWEN3-8B	90.78	81.75	536
QWEN3-14B	91.36	82.7	953

以上的实验结果表明模型尺寸、分类效果与系统延迟之间存在明确的多维权衡关系，这些不同尺寸的模型在“效果—延迟”二维平面上构成了一条典型的帕累托前沿（Pareto Front）。4B模型和1.7B模型相比，检索系统的Recall@5性能提升了2.12%，延迟增加了约100 ms。14B模型和4B模型相比，检索系统Recall@5提升为1.27%，但是延迟增加了2.6倍。基于上述帕累托优化分析，选择QWEN3-4B模型为部署到线上检索环境的最优选择。

4 结 语

针对基于关键词的全文检索系统因全库搜索而导致的低召回率问题，本文提出了一种融合对比学习与多任务学习的方法。核心思路是将标准检索系统的用户查询通过CCS体系的文本分类任务进行精准的领域分类，以此缩小检索范围。为了提升在CCS分类体系下的模型表现，本文设计了层次感知的对比学习策略以强化模型对相似类别的识别能力，并通过知识注入的多任务框架增强模型在标准领域的知识泛化能力。

实验显示，该方法在查询分类任务上的表现均优于基线模型。将其应用于下游检索系统后，真实场景下的Top-5召回率（Recall@5）和平均倒数排名（MRR）均有提升。通过性能和延迟的对比分析，选定QWEN3-4B作为最终模型。

本文工作证明了大模型在标准领域检索中的前景。未来的优化方向主要包括：一是数据增强，引入更多真实用户行为数据以替代合成数据；二是多标签分类，针对模糊查询输出Top-K个候选CCS分类标签并结合置信度加权，以进一步完善检索排序机制。

参考文献

[1] 于洋. 基于Solr技术的标准全文检索系统研究[J]. 电脑编程技巧与维护, 2020(5): 34-36, 58.

[2] 张绍琳, 曹平. 基于Lucene的标准全文检索技术研究与应用[J]. 航空标准化与质量, 2019(5): 53-56.

[3] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[PP/OL]. arXiv: 1810.04805.

[4] 杨旸, 蒋丽琼, 俞旻, 等. 基于两阶段排序的部分领域标准文献搜索优化方法研究[J]. 中国科技资源导刊, 2024, 56(6): 56-64.

[5] Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations[C]// International conference on machine learning. New York: PMLR, 2020: 1597-1607.

[6] Caron M, Misra I, Mairal J, et al. Unsupervised learning of visual features by contrasting cluster assignments[J]. Advances in Neural Information Processing Systems, 2020, 33: 9912-9924.

[7] Gao T, Yao X, Chen D. Simcse: Simple contrastive learning of sentence embeddings[PP/OL]. arXiv: 2104.08821.

[8] Zhang Y, Zhu H, Wang Y, et al. A contrastive framework for learning sentence representations from pairwise and triple-wise perspective in angular space[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg: ACL, 2022: 4892-4903.

[9] Ma X, Zhao L, Huang G, et al. Entire space multi-task model: An effective approach for estimating post-click conversion rate[C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. New York: ACM, 2018: 1137-1140.

[10] Ma J, Zhao Z, Yi X, et al. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2018: 1930-1939.

[11] Tang H, Liu J, Zhao M, et al. Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations[C]//Proceedings of the 14th ACM Conference on Recommender Systems. New York: ACM, 2020: 269-278.