

山西大学

学号：202301001124

《中文信息处理技术》期末考核任务

姓 名                     罗一恺                      
院 系           计算机与信息技术学院            
专 业           计算机科学与技术            
班 级           计算机科学与技术 2301          

2025 年 6 月 1 日

# 面向 CCKS2020 金融领域篇章级事件主体抽取的理论与方法实现研究

计算机与信息技术学院 计算机科学与技术 2301 罗一恺 202301001124

(Prompt: 下面是我《中文信息处理技术》课程论文的主体部分, 请帮我根据论文主体提炼出摘要与关键词部分)

**摘要:** 随着金融领域非结构化文本数据的爆炸式增长, 从海量信息中高效、精准地抽取出结构化事件信息, 已成为辅助决策和风险评估的关键挑战。事件抽取作为核心技术, 其在金融领域的应用, 如识别财务造假、高管变动等, 对维护金融市场稳定具有不可替代的作用。本研究聚焦于 CCKS2020 “金融领域篇章级事件主体抽取” 评测任务, 该任务要求从复杂的长文本中定位事件类型并识别核心实体。为应对长文本和事件信息非局部性带来的挑战, 本研究使用一种新颖且高效的方法, 将事件主体抽取任务巧妙地转化为机器阅读理解 (MRC) 中的问答 (QA) 范式。通过为每种事件类型设计特定的自然语言问题模板, 并集成滑动窗口机制处理篇章级长文本, 我们成功利用预训练语言模型强大的语义理解能力进行事件主体抽取。实验在 CCKS2020 官方数据集上进行, 对比了 hfl/chinese-macbert-base 、 hfl/chinese-roberta-wwm-ext 以及轻量级 uer/chinese\_roberta\_L-2\_H-128 等模型。结果显示, hfl/chinese-macbert-base 模型在验证集上取得了 0.6739 的 F1 值, 0.5790 的精确率和 0.8060 的召回率, 验证了问答范式在篇章级事件主体抽取任务上的有效性与高召回率特性, 后者对金融风控场景尤为价值。同时, 实验也强调了模型容量对复杂任务性能的关键影响。本研究为金融领域事件抽取提供了一个实用且高效的解决方案, 并为未来的研究方向, 如自动化问题生成和更精细的长文本处理策略, 提供了深入见解。

**关键词:** 事件抽取; 金融领域; 问答范式; 预训练语言模型; 篇章级; CCKS2020

# 1 引言

(Prompt: 根据金融领域非结构化文本处理的挑战与事件抽取的重要性、CCKS2020 金融领域篇章级事件主体抽取任务的背景, 结合本研究的动机与具体目标 (设计问答范式方法、实验对比、优势不足剖析), 为我生成课程论文的引言部分。)

随着信息技术的飞速发展, 全球数据量呈现爆炸式增长。特别是在金融领域, 每日产生的海量非结构化文本数据, 如新闻报道、上市公司公告、社交媒体讨论等, 蕴含着大量对市场动态、企业运营和风险评估至关重要的信息。如何从这些纷繁复杂的数据中快速、准确地提取出有价值的结构化信息, 已成为金融机构和研究者面临的重大挑战。事件抽取 (Event Extraction, EE) 作为从这些文本中提取结构化信息的关键技术, 其在金融领域的应用——如自动识别财务造假、高管变动、股权质押等事件——对于提升金融分析的效率和深度、辅助智能化决策、乃至维护金融市场稳定都具有不可替代的作用。可以说, 在信息驱动的现代金融市场中, 高效、精准地从海量文本中捕获和解析事件信息, 已经从一种技术追求转变为金融机构保持竞争力、防范风险、把握机遇的基础能力。

中国知识图谱与语义计算大会 (CCKS) 举办的评测任务一直是推动中文信息处理技术发展的重要平台, 其举办的“CCKS2020 金融领域篇章级事件主体抽取”评测任务, 正是针对这一现实需求设置的典型挑战。该任务要求从篇章级金融文本中, 提取出预定义的事件类型及相应的“事件主体”(即发生事件的核心实体)。其“篇章级”的特性意味着输入文本可能很长, 与传统的句子级事件抽取任务有显著区别, 在真实世界的金融文本中, 一个复杂事件的描述往往散布在多个句子或段落中。这种信息的非局部性对模型的上下文理解能力和长距离依赖建模能力提出了更高的要求。核心挑战在于, 如何在复杂的、可能跨越多个句子的篇章级文本中, 精准地定位事件类型, 并找出与之关联的核心行动者或承担者。

本研究与实现的动机在于探索一种简洁而有效的方法来应对这一挑战。我们提出将复杂的事件抽取任务转化为一个直观的问答任务, 利用在机器阅读理解领域已取得巨大成功的预训练语言模型来解决此问题。本次研究的目标是:

1. 设计并实现一种基于问答范式的事件主体抽取方法, 并集成滑动窗口机制以处理篇章级长文本。
2. 在 CCKS2020 官方数据集上对该方法进行实验, 并对比不同预训

练语言模型的性能。

3. 对所选方法的优势和不足进行细致剖析，并通过量化结果和错误分析，为未来的改进和提供方向。

## 2 研究背景与文献综述

(Prompt: 根据我提供的几篇学术论文的核心内容, 结合我项目中的代码实现, 为金融领域篇章级事件主体抽取任务生成课程论文的研究背景与文献综述)

### 2.1 事件抽取概览

事件抽取旨在识别文本中描述的特定类型的"事件"及其相关的"论元"(Arguments)。一个"事件"通常被理解为一个具体的事情或动作。事件抽取任务通常被分解为两个主要的子任务: 事件检测(识别事件类型)和事件论元抽取(识别事件参与者及其角色)。CCKS2020 的"事件主体抽取"任务, 可以看作是事件抽取的一个简化或特定形式, 它要求同时识别事件类型并抽取一个特定的论元——"事件主体"。

### 2.2 事件抽取的机器学习方法

事件抽取的方法经历了从早期基于规则和模式匹配, 到传统机器学习, 再到当前主流的深度学习的演进。以 BERT 为代表的预训练语言模型(Pre-trained Language Models, PLMs)通过在下游任务的标注数据上进行微调(fine-tuning), 已成为当前 NLP 领域的主流范式, 并在事件抽取任务上取得了显著效果。在处理包含多个子任务的复杂 NLP 问题时, 通常有流水线方法(Pipeline Methods)和联合学习方法(Joint Learning Methods)两种策略。流水线方法模块化但存在错误传播问题, 而联合学习方法通过在一个统一模型中学习所有子任务来缓解此问题。

### 2.3 基于问答(QA)的事件抽取方法

近年来, 将事件抽取任务转化为问答(Question Answering, QA)形式成为一个新的研究趋势。其核心思想是为每个待抽取的事件元素(如事件类型、特定角色的论元)设计一个或多个自然语言问题, 然后利用 QA 模型从文本中寻找答案。例如, 要抽取"股权质押"事件的"质押方", 可以提问: "在股权质押事件中,

谁是质押方？”。

这种方法具有以下优点：

1. 直观性：将复杂的结构化预测问题转化为人类易于理解的问答形式。
2. 利用现有模型：能够直接利用在机器阅读理解任务上已经训练得非常强大的预训练模型，如 BERT、RoBERTa 等。
3. 灵活性：通过设计不同的问题模板，可以灵活地抽取不同类型的事件信息，而无需设计复杂的、针对特定任务的输出层。

### 3 研究方法

(Prompt: 根据我项目中的具体代码实现，结合 README.md 文件为我创建本项目研究方法的初稿，要求详细阐述项目中的使用 研究方法，要求行文流畅，符合学术论文规范，避免使用项目符号或编号列表，主要包括 QA、滑动窗口、模型选择部分)

#### 3.1 基于问答范式的事件主体抽取

为应对 CCKS2020 金融领域篇章级事件主体抽取的挑战，本研究采用了一种创新的解决范式，其核心思想是将传统的序列抽取任务转化为一个机器阅读理解或问答任务。此方法不仅巧妙地利用了预训练语言模型在语义理解和答案定位方面的强大能力，也为处理金融文本中复杂的事件与主体关系提供了一个清晰且高效的框架。传统的事件抽取方法通常将任务分解为事件检测和论元抽取等多个阶段，这不仅容易导致错误在流水线中传播，而且难以显式地建立特定事件类型与其对应主体之间的直接关联。为了克服这些局限，本研究将整个事件主体抽取过程重构为一个问答模型。在此范式下，模型的输入和目标被重新定义：

表 1：问答范式的的相关表述以及在本项目中的相关定义

上下文	原始的、未经切分的金融领域篇章文本。
问题	针对每一个预定义的事件类型，我们设计了一个固定的、指向性的自然语言问题模板。例如，对于“股权质押”这一事件类型，对应的问题是：“在事件‘股权质押’中，事件主体是谁？”。
答案	在给定上下文和问题的条件下，模型需要定位并抽取的“事件类型”“事件主体”为该问题的答案。若当前上下文中并未发生该

	问题所指代的事件，则该问题被视为无答案（No-Answer）。
输入	一段金融领域的文本 $T$ ，长度不定。
输出	(事件类型, 事件主体) 二元组的集合。
事件类型	任务预定义了一系列金融领域的事件类型，例如：财务造假、股权质押、增持、减持等。
事件主体	指发生特定事件类型的核心实体，通常是公司名称、人名或机构名称。

通过这种转换，事件主体抽取的本质从在文本中进行开放式搜索，转变为在特定问题（即事件类型）的引导下，对上下文（即原始文本）进行答案（即事件主体）的精准定位。这种方法极大地明确了模型的任务目标，将对事件类型的识别内隐地融合到对问题的理解之中，使得模型能够集中其注意力于寻找与特定事件语义相关的核心实体，从而有效提升了抽取的准确性。

### 3.2 面向长文本的滑动窗口处理机制

金融公告等篇章级文本的长度常常远超如 BERT 等主流预训练模型所能处理的 512 个 token 的上限。若采用简单的截断策略，极有可能丢失包含关键事件信息的核心段落。为了在不牺牲信息完整性的前提下应用强大的预训练模型，本研究设计并实现了一套完备的滑动窗口（Sliding Window）机制，将其深度应用于数据处理的模型推理的全过程。

在模型训练阶段，该机制被用于长文本的预处理。我们使用一个固定尺寸的窗口（例如 512 个 token）以设定的步长（例如 128 个 token）在原始长文本上进行滑动，从而将一篇长文档切分为多个存在内容重叠的短文本块（chunks）。对于每一个文本块，如果它完整地包含了某个事件主体的文本跨度（span），那么该文本块与对应的事件类型问题便构成了一个正样本。由于窗口之间存在重叠，一个真实的事件主体可能作为答案出现在多个不同的文本块中，这种设计天然地实现了对正样本的数据增强（Data Augmentation），有助于模型学习在不同上下文中识别同一实体。反之，若一个文本块不包含特定事件问题的答案，则构成负样本。

在模型预测阶段，面对一篇待分析的长文章，滑动窗口机制同样发挥着核心作用。文章首先被切分为若干个重叠的文本块，模型对每一个文本块进行独立的推理，预测所有可能存在的答案，并为每个预测结果计算一个置信度分数。在所

有文本块的预测完成后，我们设计了一个结果聚合（**Result Aggregation**）策略，该策略会综合比较所有从同一篇文章中预测出的答案，最终选择置信度得分最高的预测结果作为该事件类型在整个篇章中的唯一事件主体。通过这一机制，本研究成功地将短文本设计的预训练模型的能力扩展至篇章级任务，确保了模型能够接触并利用完整的上下文信息。

### 3.3 高效的负采样与模型选型

在将任务转换为问答范式后，一个直接的挑战是数据不平衡问题。若为每一段文本都匹配所有未在其中发生的事件类型作为负样本，将会产生海量的“无答案”样本，这不仅会急剧增加训练数据量和计算开销，还可能导致模型过度倾向于预测“无答案”，从而影响其对真实事件的识别能力。为解决此问题，本研究引入了负采样（**Negative Sampling**）策略。具体而言，对于每一个正样本（即文本中真实发生的事件），我们仅从所有未发生的事件类型中随机挑选固定数量（在本研究中设置为 2 个，过多的负样本给模型训练造成了极大的性能上的压力）来生成对应的“无答案”负样本。该策略在保证模型能够学习区分不同事件类型的能力的同时，将训练数据量控制在了一个合理的范围内，显著提升了训练效率与模型的收敛效果。

为了全面评估本研究方法的有效性并探索模型能力对任务性能的影响，我们选取了三个具有代表性的中文预训练语言模型进行对比实验：

- **hfl/chinese-macbert-base**: 作为本研究的主要基线模型。**MacBERT** 通过将掩码语言模型（MLM）任务改进为“**MLM as Correction**”任务，增强了模型对文本中存在的拼写错误和用词不当等噪声的鲁棒性，这一特性使其非常适合处理来源于真实世界、可能存在格式转换错误的金融文本。
- **hfl/chinese-roberta-wwm-ext**: 作为一个高性能的对比模型。**RoBERTa** 通过采用全词掩码（**Whole Word Masking**）策略和在更大规模、更多样化的中文语料上进行预训练，通常在各类下游任务中展现出更强的性能。引入此模型旨在评估更先进的预训练策略对本任务的性能提升上限。
- **uer/chinese\_roberta\_L-2\_H-128**: 作为一个轻量级的对比模型。该模型层数较少，参数量远小于前两者。引入此模型旨在进行一种反向的消融研究，用以探索在处理复杂的金融事件抽取任务时，模型容量的下限在何处，从而反向证明任务的复杂性对模型能力提出的内在要求。

## 4 数据集分析

(Prompt: 请根据以下提供的详细信息,为我润色论文的第4部分 数据集分析。要求行文流畅,符合学术论文规范,避免过多使用项目符号或编号列表,将其组织成一个连续的段落。内容需涵盖数据集的来源介绍、本研究中各文件(训练、内部验证、批量预测、定性分析)的具体用途,并明确提及用于定量评估的内部验证集以及用于批量预测和定性分析的文件、具体脚本或指令。)

### 4.1 数据集描述

本研究采用的实验数据源于 CCKS2020 “面向金融领域的篇章级事件主体与要素抽取”评测任务的官方数据集。该数据集由蚂蚁集团与中国科学院自动化研究所联合提供,其文本语料主要采集自互联网公开的新闻报道以及上市公司的公告文件。值得注意的是,部分源自 PDF 格式的公告文本已被转换为无结构的纯文本,这意味着数据中可能包含一定的格式噪声,这对模型的鲁棒性提出了挑战。

官方数据集包含三个核心文件: `event_entity_train_data_label.csv` 作为训练集,包含 72,515 条已标注样本; `event_entity_dev_data.csv` 作为开发集 (A 榜测试集),包含 900 条样本; 以及 `event_entity_test_data.csv` 作为最终测试集 (B 榜测试集),包含 299,395 条样本。

数据的组织形式体现了任务的特性。在训练集中,每条样本均由“文本 ID”、“文本内容”、“事件类型”和“事件主体”四个字段构成,以制表符分隔。一个关键特征是,同一篇文本若包含多个不同的事件,或同一事件类型下有多个主体,会在数据中以多条独立样本的形式出现。这揭示了任务中存在的多事件共存和元素重叠的复杂性。相比之下,开发集与测试集仅提供“文本 ID”和“文本内容”,需要模型进行端到端的预测。

### 4.2 本研究中的数据使用方案

鉴于官方提供的开发集与测试集均未附带标签,无法直接用于模型训练过程中的性能监控和超参数调优。为此,本研究设计了如下一套完整且严谨的数据使用方案,确保实验的可靠性与可复现性。

模型训练主要依赖于官方提供的 `event_entity_train_data_label.csv` 文件。在数据预处理阶段,我们利用此文件中丰富的标注信息,将其转换为问答范式所

需的大量（上下文、问题、答案）三元组，作为模型学习的基础。

定量评估是衡量模型性能的关键。为了获得客观、可对比的性能指标，我们从原始训练集中随机抽样并固定了一个包含 500 条样本的子集，作为本研究的内部验证集。此验证集在模型训练过程中被完全隔离，仅用于最终的性能评测。我们编写了自动化评估脚本 `src/analysis.py` (`python -m src.analysis`)，该脚本加载各模型的最优权重，对该验证集进行预测，并依据官方标准计算精确率（Precision）、召回率（Recall）和 F1 值。

开发集批量预测是模拟真实竞赛场景的重要环节。本研究实现了对官方 `event_entity_dev_data.csv` 开发集的完整预测功能。通过 `python -m src.batch_predict` 指令，可以调用任一训练好的模型，对开发集中的全部 900 条数据逐一进行事件主体抽取。预测结果将按照官方要求的格式（文本 ID、事件类型、事件主体）保存至 `output/` 目录下，并以模型名称加以区分。此功能不仅生成了可用于手动评估或提交至在线评测系统的结果文件，也验证了本研究方案在处理未知数据时的端到端能力。

定性分析与效果展示则用于直观地考察模型在真实长文本场景下的表现。我们在 `src/predict.py` 脚本中内置了一段典型的金融领域长篇公告文本。通过直接运行此脚本，可以调用训练好的模型，对其进行完整的篇章级事件主体抽取，并观察模型在处理跨句、多事件等复杂情况下的实际输出。这为后续的错误分析和案例研究提供了直观的素材。

### 4.3 评价指标与协议

本研究完全遵循 CCKS2020 官方评测任务所制定的评价协议。模型性能通过三个核心指标进行衡量：精确率（Precision, P）、召回率（Recall, R）和 F1 值（F1-Score）。为了综合评估模型在所有事件类型上的整体表现，所有指标均采用微平均（Micro-average）的方式进行计算。

评估的标准极为严格：一个预测的（事件类型，事件主体）二元组被判定为正确（即一个真正例，True Positive），当且仅当其预测的事件类型和预测的事件主体字符串，与标准答案完全一致。任何一项的偏差，无论是事件类型分类错误，还是主体抽取的边界存在出入（例如，输出了“XX 公司”而非“XX 有限公司”），都将被视为一次错误的预测。这一严格的评价标准要求模型不仅要在事件识别和主体抽取两个子任务上都具备高水平的性能，更要能够精准地建立两者之间的正确关联。

## 5 实验结果

(Prompt: 请根据我提供的论文原稿以及最新的实验数据(已修改), 对我的论文的第五部分实验结果、第六部分进行重写, 重新组织一下这两部分的逻辑结构以及语言, 使用较为流畅的语言, 不要使用项目符号以及标号, 使之符合论文的格式)

本研究为系统评估所提方法的有效性, 选取了三个具有代表性的中文预训练模型进行对比实验, 分别是 `hfl/chinese-macbert-base`、`hfl/chinese-roberta-wwm-ext` 以及轻量级的 `uer/chinese_roberta_L-2_H-128`。为了更充分地探索各模型的性能, 模型均经过了不同周期的训练(因为个人电脑算力有限), 并在一个由 500 条样本构成的固定验证集上进行评测, 其最佳性能表现详见下表。

表 2: 各模型在验证集上的性能表现

Model	Epoch	F1	Precision	Recall
<code>hfl/chinese-macbert-base</code>	5	0.6739	0.5790	0.8060
<code>hfl/chinese-roberta-wwm-ext</code>	1	0.5674	0.4533	0.7580
<code>uer/chinese_roberta_L-2_H-128</code>	10	0.1912	0.1703	0.2180

上表的量化结果清晰地展示了本研究方法的有效性。首先, `macbert-base` 模型在经过 5 个周期的训练后, F1 值达到了显著的 0.6739, 这强有力地证明了“问答范式”是解决篇章级事件主体抽取任务的一条成功且高效的技术路径。其次, 实验结果凸显了模型“高召回、中等精确”的特性, 表现最佳的 `macbert-base` 模型其召回率超过了 80%, 精确率也接近 58%。这一特性在金融风控、舆情监控等注重全面性、宁可错杀也不可放过应用场景中, 具有非常显著的实用价值。最后, 实验结果再次验证了模型容量是决定任务成败的关键因素。尽管经过了多达 10 个周期的训练, 极度轻量级的 `uer-mini` 模型其 F1 值仍未超过 0.2, 性能远不及 `base` 级别的模型, 这清晰地表明, 对于此类复杂的语义理解任务, 一个足够大的模型容量是取得良好性能的基础前提。

为更深入地理解模型的行为模式与潜在缺陷, 我们进行了定性的错误分析。分析发现, 主要的错误类型体现在三个方面。一是主体边界界定不准, 模型能够定位事件主体的大致位置, 但抽取的实体边界不够精确, 例如将“XX 科技有限公司”错误地识别为“XX 科技”。二是同类型事件混淆, 当文本中包含多个角色相似的实体时, 模型可能在判断事件的真正主体时发生混淆, 例如将股权转让的

受让方误判为转让方。三是对于无答案样本的判断失误，这包括事件存在但模型未能抽取出主体的漏报情况，以及事件不存在但模型仍给出了错误答案的误报情况。这些错误表明，模型在处理实体边界的细微差异、理解复杂上下文中的角色关系以及学习隐晦表达方面仍有提升空间。

## 6 总结与展望

作为《中文信息处理》课程理论学习与实践探索的结合，本研究项目聚焦于金融领域篇章级事件主体抽取这一具有挑战性的真实场景，旨在将课程中所学的自然语言处理核心技术应用于解决具体问题。在本次设计中，我们成功探索并验证了一种将该复杂抽取任务高效转化为问答范式的方法。研究表明，通过将事件类型巧妙地设计成自然语言问题，可以充分利用中文预训练语言模型强大的语义理解能力进行事件主体抽取。其中，`macbert-base` 模型在验证集上取得了超过 0.67 的 F1 分数，这充分验证了该技术路线的成功与高效。实验对比分析进一步揭示，模型的容量对于处理此类复杂任务至关重要，同时，所提方法在当前实验配置下呈现出“高召回、中等精确”的突出特点，这使其在金融风控等下游场景中具备了很高的潜在应用价值。整个过程不仅是对课程知识的综合运用，更是一次宝贵的科研与工程实践，加深了我们对中文信息处理技术核心挑战与应用前景的理解。

尽管本研究取得了积极的成果，但仍存在一些局限性，这些局限性也为未来的研究指明了方向。首先，当前方法的性能在一定程度上依赖于人工设计的问题模板，未来的一个重要研究方向是探索自动化的问题生成技术，例如利用 T5 等生成模型自动优化或生成最适合事件抽取的问题，以降低人工成本并可能提升性能。其次，本研究虽然对不同模型进行了多周期的训练，但更系统性的超参数调优（如学习率、负采样比例等）仍有探索空间，这可能带来进一步的性能提升。此外，针对篇章级长文本的处理，未来应实现并评估如滑动窗口等更成熟的策略，以确保模型能够有效整合分布在整个文档中的信息。最后，可以探索将多跳推理等更先进的问-答技术引入该框架，以应对事件信息分散、逻辑关系更隐含的复杂抽取场景。通过这些方向上的持续探索，有望进一步提升基于问答范式的事件抽取技术，使其更好地服务于金融等领域的智能化信息处理需求。

## 参考文献

- [1] 王人玉, 项威, 王邦, 等. 文档级事件抽取研究综述[J]. 中文信息学报, 2023, 37(6).
- [2] 胡蓉, 万常选, 万齐智, 刘德喜, 刘喜平. 基于深度学习的篇章级事件抽取综述[J]. 计算机学报, 2025, 48(2).
- [3] MeiHan Tong, Bin Xu, Shuai Wang, 等. DocEE: A Large-Scale and Fine-grained Benchmark for Document-level Event Extraction[C/OL]. Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, 2022[2023-10-26]. <https://aclanthology.org/2022.naacl-main.291/>.
- [4] Xie J, Zhang Y, Kou H, et al. A Survey of the Application of Neural Networks to Event Extraction[J/OL]. Tsinghua Science and Technology, 2025, 30(2): 748-768[2023-10-26]. <https://doi.org/10.26599/TST.2023.9010139>.
- [5] Di Lu, Shihao Ran, Joel Tetreault, et al. Event Extraction as Question Generation and Answering[C/OL]. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Toronto: Association for Computational Linguistics, 2023: 1666-1688[2023-10-26]. <https://doi.org/10.18653/v1/2023.acl-short.143>.
- [6] Di Lu, Shihao Ran, Joel Tetreault, et al. Event Extraction as Question Generation and Answering[EB/OL]. arXiv:2307.05567v1. 2023-07-10[2023-10-26]. <https://arxiv.org/pdf/2307.05567>.
- [7] Zichuan Fu, Wentao Song, Yejing Wang, et al. Sliding Window Attention Training for Efficient Large Language Models[EB/OL]. arXiv:2502.18845v1. 2025-02-26[2023-10-26]. <https://arxiv.org/html/2502.18845v1>.
- [8] Xindi Wang, Mahsa Salmani, Parsa Omid, et al. Beyond the Limits: A Survey of Techniques to Extend the Context Length in Large Language Models[EB/OL]. arXiv:2402.02244v2. 2024-05-22[2023-10-26]. <https://arxiv.org/pdf/2402.02244v2>.