

基于 RAG 的煤矿安全智能问答模型

洪亮¹, 郭瑶¹, 刘兴丽², 李宗雨²

(1. 黑龙江科技大学 管理学院, 哈尔滨 150022; 2. 黑龙江科技大学 计算机与信息工程学院, 哈尔滨 150022)

摘要:面向煤矿安全智能辅助预警决策需求,以瓦斯超限煤矿安全隐患知识为数据源,基于 RAG 检索增强生成,建立了一种煤矿安全智能问答模型。应用 BGE-M3-Embedding、FAISS 开源库等方法,构建瓦斯超限煤矿安全文本块的向量数据库,以及检索和生成模块,采用 BGE-M3-Embedding 模型完成 Query 文本块向量的检索,召回 top-k 向量检索的上下文,实现煤矿安全瓦斯超限的相关提问,构建 Prompt 微调提示词,增强大语言模型生成答案。结果表明,Ragas 自动评测 Rag 检索增强三种大语言模型效果,Baichuan2-13BAR、F 忠实度及上下文相关指标最优,分别为 0.91、0.83 和 0.87。搭建煤矿安全辅助决策的智能问答原型,验证了触发瓦斯超限煤矿安全隐患智能辅助决策的实效性、可靠性及迁移性。

关键词:煤矿安全; RAG 检索增强生成; 智能问答; 智能辅助决策支持

doi:10.3969/j.issn.2095-7262.2024.03.024

中图分类号:TD79; TP391

文章编号:2095-7262(2024)03-0487-06 **文献标志码:**A

Intelligent Q & A model of coal mine safety based on RAG

Hong Liang¹, Guo Yao¹, Liu Xingli², Li Zongyu²

(1. School of Management, Heilongjiang University of Science & Technology, Harbin 150022, China;

2. School of Computer & Information Engineering, Heilongjiang University of Science & Technology, Harbin 150022, China)

Abstract: This paper aims to explore a coal mine safety intelligent question answering model in response to the demand for intelligent early warning decision-making for coal mine safety by based on RAG retrieval and enhanced generation by using the knowledge of safety hazards in coal mines with gas exceeding limits as the data source. The study consists of using the methods of BGE-M3 Embedding and FAISS open-source libraries to construct a vector database of safety text blocks for coal mines with gas exceeding limits, and a retrieval and generation module; using the BGE-M3 Embedding model to retrieve the query text block vector; recalling the context of the top-k vector retrieval; implementing the relevant questions about coal mine safety with gas exceeding limits; building a Prompt fine-tuning clue words to enhance the generation of answers in the large language model. The results showed that Ragas automatically evaluates the effectiveness of Rag retrieval in enhancing three major language models, and the optimal indicators are Baichuan2-13BAR, F, and CR indicators, reaching 0.91, 0.83 and 0.87, respectively. Building an intelligent Q & A prototype for coal mine safety assistance decision-making has verified the effectiveness, reliability, and transferability of intelligent assistance decision-making for triggering gas exceeding coal mine safety hazards.

Key words: coal mine safety; RAG retrieval-augmented generation; intelligent question answering; intelligent decision support

收稿日期: 2024-04-07

项目基金: 黑龙江省极薄煤层智能开采关键技术攻关与示范项目(2021ZXJ02A04); 黑龙江省省属高等学校基本科研业务费项目(2022-KYYWF-0569)

第一作者简介: 洪亮(1978-),男,辽宁省铁岭人,教授,博士,研究方向:煤矿安全智能应急管理,E-mail:45070955@qq.com。

0 引 言

自 2020 年 3 月,国家发展改革委等在数智时代背景下,技术赋能对煤矿安全能源局等八部委联合印发了《关于加快煤矿智能化发展的指导意见》(发改能源[2020]283 号)以来,煤矿安全化生产水平已显著提升^[1]。然而,仍有因煤矿工作人员因安全认知水平局限,决策信息不全面引发了安全隐患出现,此时,融合了自然语言处理和人工智能技术的智能问答,通过人机实时交互方式可有效辅助安全原因分析与处理措施等决策过程。近年来,基于知识图谱的领域智能问答技术成果显著。例如,文献[2-5]等验证了领域知识图谱智能问答技术可行性。然而,当下问答技术领域迁移与扩展大大受限于知识图谱自动更新能力。

2022 年 11 月 30 日,OpenAI 推出的一款人工智能技术驱动的自然语言处理工具 ChatGPT 生成式 LLM(Large language model)大语言模型出现,因其出色的文本理解和生成能力,大大推动了人工智能技术与科学研究深度结合,帮助促进发现新知识和解决科学难题的新型科学研究范式^[6]。显然,为井下生产设备植入煤矿安全“大脑”知识,辅助煤矿安全问题高效决策提供了可能^[7]。2023 年煤炭科技十大新闻中,中国煤炭科工集团西安研究院发布国内首个地质垂直领域大模型 GeoGPT,并已应用于智能矿山透明地质保障工作,有望解决海量多源数据交叉利用率低,历史资料难以有效查询等难题^[8]。然而,LLM 大语言模型无法完成超出其训练数据范围的内容生成^[9],所以,仅通过 LLM 大语言模型支持人机交互智能问答,仍然存在大量的幻觉性问题。

针对上述不足,笔者基于检索增强生成 RAG 方法,以选自瓦斯超限煤矿安全隐患知识作为数据源,构建一套煤矿安全智能问答模型,支持煤矿安全智能辅助预警决策需求。以瓦斯超限为例,探索了一种向量检索、Prompt 增强及 LLM 大语言模型生成的 RAG(Retrieval augmented generation)智能问答融技术路径,实现煤矿安全智能辅助决策支持 CMS_IDSS(Intelligent decision support system for coal mine safety)的技术可行性。以瓦斯超限为例,隐患条件、原因及措施的预警决策场景,搭建并验证智能问答的实效性,支持其他煤矿安全隐患预警场景需求的高效迁移性。

1 RAG 检索增强生成模型

文中按照 Bonczek^[10] 基于知识的“三系统”体系结构,煤矿安全辅助决策 RAG 智能问答模型构建划分为数据、模型和对话三个部件组成,如图 1 所示。

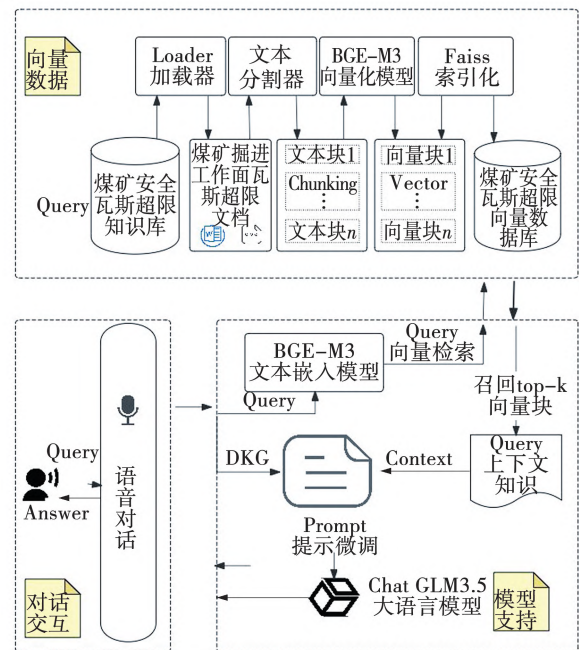


图 1 RAG 煤矿安全辅助决策智能问答模型
Fig. 1 Coal mine safety assistant decision-making RAG intelligent question answering

总体架构可划分为三部分:一是数据部件,指外挂煤矿安全瓦斯超限知识库演化的向量数据库;二是模型部件,RAG 检索增强生成模型,支持按照煤矿安全瓦斯超限类型下,瓦斯超限条件、原因及管控措施等问题检索与答案生成;三是对话部件,决策支持系统的人机接口界面。

按照三部分的设计思路:首先,完成支持索引的向量数据库构建模块;通过文本加载、文本块切分、基于 BGE - M3 - Embedding、FAISS 开源库等方法,构建文本块的向量数据库。接着,问题检索和答案生成模块构建:针对煤矿安全瓦斯超限相关提问,采用 BGE - M3 - Embedding 模型完成 Query 文本块向量的检索,召回 top - k 向量检索的上下文;再结合提问构建 Prompt 微调提示工,增强大语言模型生成答案。最后,选择科大讯飞语音接口实现人机语言交互。

1.1 向量数据库构建

本阶段是一个离线的数据准备阶段,将选自“煤矿安全知识”公众号的《煤矿50项瓦斯超限常见原因及管控措施》文本向量化后,构建索引并存入数据库的过程。主要包括文档加载模块、文档分割模块、文档向量化模块及向量数据入库模块等。

文档加载模块:指定煤矿安全瓦斯超限知识库,通过 Loader 加载器完成煤巷掘进工作面、采煤工作面、岩巷掘进工作面等瓦斯超限文档加载,可兼容 txt、pdf 等文档类型。

文档分块模块:利用分割器对加载文档按“句”颗粒切分,以保证语义的完整性,同时需要考虑 Tokens 限制,一般 512 个 tokens。

向量化模块:基于多语言、多颗粒、多功能的 BGE-M3-Embedding^[11] 向量模型。该模型支持 100 多种语言;且支持短句到长文,最高可处理 8 192 长度文本输入;同时,支持包括 Dense retrieval 稠密检索,lexical retrieval 稀疏检索,Multi-Vec retrieval 多向量检索的三种检索手段融合,进行多路找回响应文档,再基于三种相似度得分平均求和对找回结果进行排序,如图 2 所示。

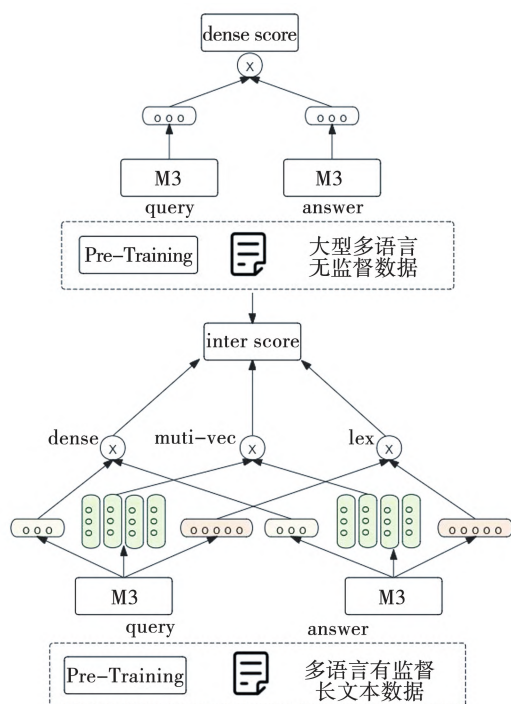


图2 BGE-M3-embedding 模型

Fig. 2 BGE-M3-embedding model

BGE-M3-Embedding 模型训练过程涉及了对比学习和知识蒸馏主要阶段,通过这种结合对比学

习损失和蒸馏损失的训练,BGE-M3-Embedding 模型旨在学习到有效的 query 和文档表示,旨在信息检索、问答系统等相关任务中取得良好的性能。

1.1.1 对比学习损失

对比学习旨在通过比较不同样本间的差异来提高模型的表现。在 BGE-M3-Embedding 模型,对比学习的目标是拉近 query 与相关文档之间的距离,同时推开 query 与不相关文档之间的距离,通过 InfoNCE 损失函数来实现的,其中,NCE 代表噪声对比估计,公式为

$$L_{\text{ln}} = -E_x \ln \left[\frac{f_k(s_{t+k}, c_t)}{\sum_{s_j \in S} f_k(s_j, c_t)} \right], \quad (1)$$

式中: c ——上下文;

s ——语句;

$f_k(s_{t+k}, c_t)$ ——得分函数。

1.1.2 知识蒸馏损失函数

知识蒸馏损失函数通过 teacher 模型的知识转移到 student 模型中,使得 student 模型具有优质的性能。其损失函数通常由两部分组成:一部分是常规的损失函数,如交叉熵;另一部分是蒸馏损失,公式为

$$L = (1 - T) * L_{\text{CE}} + T * L_{\text{KD}}, \quad (2)$$

式中: L_{CE} ——常规的损失函数;

L_{KD} ——蒸馏损失,表示两个模型输出的概率分布之间的距离;

T ——温度参数,用于平衡两部分的损失的重要性。

在 BGE-M3-Embedding 模型中,将三种不同方式的相似度得分进行加权求和,得到一个 teacher 分数。然后,student 模型的三种相似度得分将学习这个 teacher 分数的信息,从而得到三个蒸馏损失,蒸馏损失的计算使用均方误差(MSE)。

向量数据索引库构建:文中采用 Facebook AI Research 开发的 Faiss 开源库,支持 RAG 的 query 向量相似文本表示向量的快速检索。

1.2 问题检索与答案生成模块

当用户提问后,关键环节包括数据检索模块和注入 Prompt 微调模块以及大模型生成答案。

数据检索模块:用户提出问题“煤巷掘进工作面瓦斯超限的常见原因有哪些?”,则系统首先选择 BGE-M3-Embedding 模型,从煤巷掘进工作面瓦斯超限向量数据库中检索出与问题最相关 top-k 个文本库作为问题上下文。

注入 Prompt 微调提示工程:Prompt 提示作为大模型生成过程的一部分,其设计对于输出答案的准确性和相关性具有显著影响^[12]。文中煤矿安全领

域专家角色描述、瓦斯超限等任务描述、瓦斯超限的原因及管控措施等作为示例,并规范输出格式等支持大模型文本生成。

文中选择 ChatGLM3 的 LLM 大语言模型生成^[13],该模型基于 Transformer 神经网络架构的解码部分^[14],在模型的预训练部分通过上文的词来预测下一个词,属于无监督的预训练。如,给定一个无监督的语料 $U = \{u_1, u_2, \dots, u_n\}$,用 U 表示每一个 token(词),当设置窗口长度为 k ,预测句中的第 i 个词时,则使用第 i 个词之前的 k 个词,同时也根据超参数 θ ,来预测第 i 个词,公式为

$$L_1(U) = \sum_i \ln P(u_i | u_{i-k}, u_{i-k-1}, \dots, u_{i-1}; \theta)。(3)$$

2 实验结果与分析

2.1 数据集

实验数据集包括煤矿安全瓦斯超限知识库和评测数据集。其中:煤矿安全瓦斯超限管理制度、原因分析及管控措施等相关文档组成知识库;而评估数据集主要包括问题(question)、问题正确答案(ground truths)、根据问题检索到的上下文信息生成的答案(answer)、根据问题检索得到的上下文相关提示信息(contexts)等构成用 OpenAI 接口生成,评测 Dataset。

2.2 评价指标

文中选择 Ragas^[15] 自动化评估方法对模型对结果评估,并以答案相关性公式为

$$A_R = \frac{1}{n} \sum_{i=1}^n \text{sim}(\mathbf{q}, \mathbf{q}_i), \quad (4)$$

式中, \mathbf{q}_i ——查询样本 i 的编码向量。

该指标衡量生成的答案与查询之间的相关性。分数越高,相关性越好。忠实度的公式为

$$F = \frac{|V|}{|S|}, \quad (5)$$

式中: S ——生成的 statement 数量;

V ——可被支撑的 S 的数量;

该指标保证答案是基于给定上下文信息所生成的,评估模型生成答案与提供给模型的上下文提示信息中所表达事实的准确性,得到同样在 0 到 1 之间,得分越高表示答案与提示相关性越高。CR 上下文相关度公式为

$$C_R = \frac{n}{m}, \quad (6)$$

式中: n ——对答案有用的句子数量;

m ——上下文中全部句子的数量。

这是一个衡量检索质量的指标,主要评估检索到的上下文对查询的支持程度。

2.3 实验环境

本实验选择 RTX 4090 24GB 显存 GPU、RTX 3090 24GB 显存 GPU,开发选择 python3.9 版本, RAGAS0.0.22 版本、Langchain0.0.325 版本。

2.4 实验结果分析

以 Baichuan2-13B、Baichuan2-13B-4bits 和 ChatGLM3-6B 三种 LLM 模型的 Ragas 评测结果,如表 1 所示。

表 1 RAG 问答系统评测数据

Table 1 RAG coal mine safety intelligent Q & A

测试对象	答案相关性	忠实度	上下文相关度
Baichuan2-13B	0.91	0.83	0.87
Baichuan2-13B-4bits	0.87	0.76	0.85
ChatGLM3-6B	0.81	0.79	0.83

对照 Ragas 评测指标, Baichuan2-13B 分别为 0.91、0.83 和 0.87,相较于另外两种模型效果更佳,因此,文中基于该 Baichuan2-13B 模型完成煤矿安全智能辅助决策支持 CMS_IDSS 核心模块,支持煤矿安全瓦斯超限智能问答原型实现。

3 RAG 煤矿安全智能问答模型应用

3.1 煤矿安全智能问答预警逻辑

为保障煤矿安全智能辅助决策支持需求实现效果,梳理煤矿安全专家认知路径,形成“煤矿安全隐患分类、触发条件、原因分析及管控措施”的煤矿安全智能问答的预警逻辑,如图 3 所示。同时,基于 RAG 煤矿安全智能问答模型,实现煤矿安全预警应用。

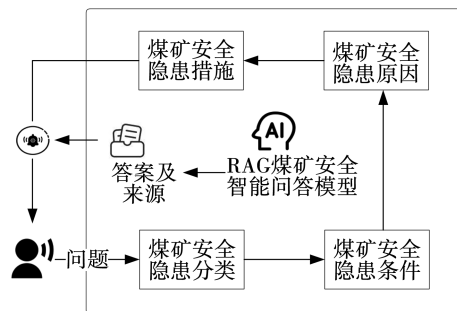


图 3 煤矿安全智能问答预警逻辑

Fig. 3 Coal mine safety intelligent Q & A early warning logic

3.2 煤矿安全智能问答系统原型验证

基于 RAG 煤矿安全智能问答模型,采用 MVC

(Model view controller) 架构,搭建了煤矿安全智能问答系统原型,及时响应煤矿安全预警问题答案,并提供答案来源,有效避免了大模型针对领域知识的幻觉,保障了 CMS_IDSS 煤矿安全决策预警问答的可靠性。

3.2.1 系统原型实效性验证

为验证 RAG 问答模型的系统原型实效性,项目组共完成 100 余次人机交互测试,测试示例如图 4 所示。当用户提问“煤巷掘进工作面瓦斯超限常见原因是什么?”,系统针对煤巷掘进工作面这类瓦斯超限隐患可能的原因。

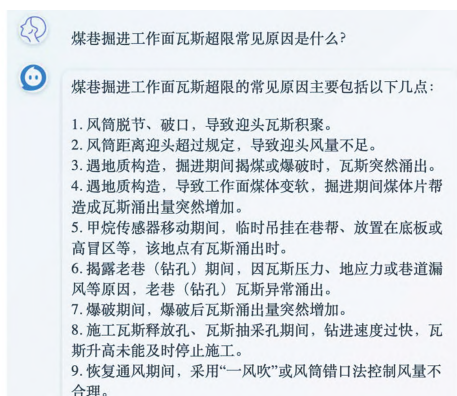


图4 智能问答原型的准确性

Fig. 4 Accuracy of intelligent Q & A prototype

3.2.2 系统原型可靠性验证

为验证 RAG 问答模型的系统原型可靠性,通过“文件来源查看”功能,提供问题答案的来源文件,如图 5 所示。同时,基于“相关推荐”模块“针对风筒脱节、破口导致超限的问题,有哪些管控措施?”,对煤矿安全隐患管理决策提供进一步的支持。



图5 煤矿安全智能问答原型的可靠性

Fig. 5 Reliability of intelligent Q & A prototype

3.2.3 系统原型可迁移性验证

为验证 RAG 问答模型的系统原型的迁移性,在“煤矿安全隐患预警”知识库中加载的煤矿瓦斯超限文档即可支持上述问答,如图 6 所示。同样,通过不同领域知识库及相应文件加载即可高效实现智能问答领域应用迁移。



图6 智能问答原型的可迁移性

Fig. 6 Transferability of intelligent Q & A prototype

4 结 论

文中基于外挂煤矿安全知识库,融合检索、增强与大模型生成的 RAG 智能问答模型构建及原型应用验证,为煤矿安全生成及安全管理提供了一种效率高、可靠性好、可移植性强的 CMS_DSS 煤矿安全辅助决策支持方法。文中的模型实验及原型系统实践可知其优势显著:一是实效性强、可靠性高,基于 RAG 技术可以检索最新的信息,且直接问答答案来源,支持生成包含最新煤矿安全数据或知识的内容;二是领域迁移性好,RAG 不受限于训练数据集,可根据需要加载其他领域知识库,完成 RAG 领域智能问答模型迁移应用;三是交互友好性,支持实时的文本与语音人机交互。该研究为煤矿安全决策支持智能技术提供了有效的参考。

参考文献:

- [1] 王国法,杜毅博. 煤矿智能化标准体系框架与建设思路[J]. 煤炭科学技术, 2020, 48(1): 1-9.
- [2] 段雨希,邱芹军,田苗,等. 面向地质图的知识图谱构建及智能问答应用[J]. 地质科学, 2024, 59(2): 588-602.
- [3] 张元鸣,姬琦,徐雪松,等. 基于知识图谱关系路径的多跳智能问答模型研究[J]. 电子学报, 2023, 51(11): 3092-3099.
- [4] 范俊杰,马海群,刘兴丽. 数智时代下开源情报的军事知识图谱问答智能服务研究[J/OL]. 数据分析与知识发现, 1-15 [2024-03-31]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20231026.1305.002.html>.
- [5] 陈明,刘蓉,熊回香. 基于医疗知识图谱的智能问答系统研究[J/OL]. 情报科学, 1-9 [2024-03-31]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20231026.1305.002.html>.

- cnki.net/kcms/detail/22.1264.G2.20230915.1055.007.html.
- [6] 孙 坦, 张智雄, 周力虹, 等. 人工智能驱动的第五科研范式 (AI4S) 变革与观察[J]. 农业图书情报学报, 2023, 35(10): 4-32.
- [7] 白 雪, 胡 泊. 当 AI 模型遇见煤矿, 会发生什么? [N]. 中国煤炭报, 2023-1. 2023 年煤炭科技十大新闻[J]. 中国煤炭工业, 2024(1): 40-41.
- [8] 赵 月, 何锦雯, 朱申辰, 等. 大语言模型安全现状与挑战[J]. 计算机科学, 2024, 51(1): 68-71.
- [9] Bonczek R H. Foundations of decision support systems[M]. New York: Academic Press, 1981.
- [10] Jianlv C, Shitao X, Peitian Z. BGE M3 - Embedding: multi-lingual, multi-functionality, multi-granularity text embeddings through self-knowledge distillation[J]. arXiv, 2024, abs/2402.03216.
- [11] 时宗彬, 乐小虬. 基于本地大语言模型和提示工程的材料信息抽取方法研究[J/OL]. 数据分析与知识发现, 1-13 [2024-04-01]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20240315.1418.002.html>.
- [12] Es S, James J, Espinosa-Anke L, et al. Ragas: Automated evaluation of retrieval augmented generation[J]. arXiv preprint arXiv: 2309.15217, 2023.
- [13] 李 南, 丁益东, 江浩宇, 等. 面向大语言模型的越狱攻击综述[J/OL]. 计算机研究与发展, 1-28 [2024-04-02]. <http://kns.cnki.net/kcms/detail/11.1777.TP.20240223.1222.002.html>.
- [14] 真超奇. 基于 Graph Transformer 的知识增强预训练模型研究[D]. 北京: 北京邮电大学, 2023.
- (编辑 李德根)