

重庆大学本科生毕业论文（设计）

面向多编程语言的代码检索系统设计与实现



学 生：夏劲
学 号：20214966
指导教师：刘超
专 业：软件工程

重庆大学大数据与软件学院

2025 年 6 月

Undergraduate Thesis (Design) of Chongqing University

**Design and Implementation of a
Multi-Programming Language Code
Retrieval System**



**By
XIAJin**

**Supervised by
Prof. LIU CHAO**

**Software Engineering
School of Big Data and Software Engineering
Chongqing University**

June, 2025

摘 要

随着编程语言的多样化和开源项目的迅速增长，开发者在寻找特定代码片段时面临着越来越大的挑战。现有的代码托管平台主要依赖关键词匹配和特定搜索条件进行检索，效率较低且使用门槛较高。大语言模型逐渐成为科研和工业界的热点，利用大语言模型可以轻松将用户的自然语言转译为具有查询条件的条件，并结合具体信息拓展补充用户的查询条件，从而实现更高效的查询体验。本项目旨在利用大语言模型实现一款面向多编程语言的代码检索系统。

本文首先介绍了大语言模型的发展情况，并论证了使用大语言模型进行查询条件重写的可行性，强调了提示词工程在这一过程中的重要性。接着简述了前端框架 Vue、Visual Studio Code 平台的插件开发、Elasticsearch 和 FastAPI 等后台开发框架。根据系统的应用目标，本文进行了系统需求分析，设计了系统架构、功能和数据库结构。在此基础上，本项目开发了基于 FastAPI 的后台管理逻辑服务、基于 Elasticsearch 的代码检索服务以及基于 DeepSeek-R1 的查询重写服务。通过 Visual Studio Code 平台提供的接口，结合 Vue 框架实现了 Visual Studio Code 插件的前端开发，完成了快捷键绑定、项目搜索等一系列功能。

本项目开发的面向多编程语言的代码检索系统具有界面美观、操作友好、查询效率高、查询准确等诸多优点。基于 FastAPI 开发的后台管理系统易于管理和维护，帮助开发人员实现高效的多编程语言项目代码搜索。

关键词：代码搜索；Elasticsearch；大模型提示工程

ABSTRACT

As programming languages diversify and open-source projects rapidly grow, developers face increasing challenges in finding specific code snippets. Existing code hosting platforms primarily rely on keyword matching and specific search conditions, resulting in low efficiency and a high usage threshold. Large language models have become a hot topic in both research and industry. By leveraging these models, users' natural language can be easily translated into functional query conditions, which can be further expanded with specific information to enhance the query experience. This project aims to implement a code retrieval system for multiple programming languages using large language models.

This paper first introduces the development of large language models and demonstrates the feasibility of using them for query condition rewriting, highlighting the importance of prompt engineering in this process. It then briefly describes the front-end framework Vue, plugin development for the Visual Studio Code platform, and back-end development frameworks such as Elasticsearch and FastAPI. Based on the system's application goals, a system requirements analysis was conducted, and the system architecture, functionality, and database structure were designed. On this basis, the project developed a back-end management logic service based on FastAPI, a code retrieval service using Elasticsearch, and a query rewriting service using DeepSeek-R1. The front-end development of the Visual Studio Code plugin was achieved through the platform's provided interfaces, combined with the Vue framework, implementing features such as shortcut key bindings and project search.

The code retrieval system developed in this project for multiple programming languages offers numerous advantages, including an aesthetically pleasing interface, user-friendly operation, high query efficiency, and accuracy. The back-end management system developed with FastAPI is easy to manage and maintain, aiding developers in efficiently searching for code across multi-language projects.

Key words: Code Search;Elasticsearch;LLM Prompt

目 录

摘要..... I

ABSTRACT..... II

1 绪论..... 1

1.1 研究目的及意义..... 1

1.2 国内外研究现状..... 1

1.3 本文研究意义..... 2

1.4 本章小结..... 2

2 正文文字格式..... 3

2.1 论文正文..... 3

2.2 字数要求..... 3

2.2.1 本科论文字数要求..... 3

2.3 本章小结..... 3

3 图表、公式格式..... 4

3.1 图表格式..... 4

3.2 公式格式..... 4

3.3 本章小结..... 5

4 结论与展望..... 6

4.1 主要结论..... 6

4.2 研究展望..... 6

参考文献..... 7

附录 A：XX 公式的推导..... 8

致谢..... 9

原创性声明和使用授权书..... 10

1 绪论

1.1 研究目的及意义

随着机器学习、深度学习等人工智能技术的快速发展，这些高新技术在解决传统领域挑战中发挥了重要作用。在搜索领域，传统搜索服务提供商通常直接采用 Elasticsearch 作为搜索引擎来完成信息检索。然而，传统搜索引擎基于倒排索引的检索方式要求用户提供极其精确的搜索词，并掌握一定的搜索技巧，这极大地限制了普通用户的搜索体验。随着人工智能技术，尤其是大语言模型（LLM）的进步，利用其自然语言友好特性来辅助搜索已成为科研界和工业界的研究热点。

Elasticsearch 是一种分布式、RESTful 风格的搜索和分析引擎，其强大的搜索能力源于其独特的倒排索引结构（Inverted Index）。这种数据结构使 Elasticsearch 具备高效的全文搜索能力、实时数据处理能力以及高可扩展性，同时能够处理结构化和非结构化数据。

大型语言模型（LLM）是拥有海量参数和卓越学习能力的高级语言模型，其核心模块是 Transformer 架构中的自注意力机制。自注意力机制作为语言建模任务的基本构建块，能够有效处理顺序数据，实现并行化计算，并捕捉文本中的远程依赖关系。LLM 的显著优势在于其能够理解自然语言并执行基于自然语言的指令，这对用户侧软件服务非常友好。然而，LLM 的另一特点是其庞大的模型参数和极高的训练成本。从头预训练一个大型模型所需的资源（如数百万高性能显卡卡时）是普通实验室或个人工作者难以承担的。因此，针对特定场景的任务，工业界普遍采用开源的大语言模型进行监督微调（SFT），或通过 Prompt 工程来实现任务目标。

本项目旨在构建一个面向多编程语言的代码检索系统，该系统将融合大语言模型的自然语言友好能力，实现对用户搜索词的重写、完善，结合 Elasticsearch 强大的多结构文本搜索能力实现用户方便快速搜索多语言项目的的能力。同时我还讲相关能力集成到 Visual Studio Code 平台，发布免费的开源插件，为所有开发者提供便捷高效的代码搜索服务。

1.2 国内外研究现状

2004 年，Shay Banon 创造了 Elasticsearch 的前身——Compass。在 Compass 的基础上，Shay Banon 进一步实现了分布式和可扩展性优化，并提供了 HTTP 接口，

使得 Java 以外的语言也能调用。2010 年，Elasticsearch 的第一个版本正式发布。

Elasticsearch 是一种成熟的搜索解决方案，一些大型搜索公司如 Google 和百度都基于 Elasticsearch 进行研发。Elasticsearch 支持 RESTful 风格的调用，客户端可以通过 HTTP 请求直接操作 Elasticsearch 服务。它还提供了多种客户端支持，包括 Java、Python、Go、PHP 等语言的 SDK，以及 JDBC、ODBC 等标准化接口。Elasticsearch 广泛应用于全文搜索、日志分析、实时数据分析、地理空间搜索以及商业智能领域的数据聚合分析。在部署方面，Elasticsearch 支持单节点模式，适用于开发和测试环境，可以通过简单的 Docker 命令快速启动。在生产环境中，通常采用分布式集群部署，通过分片（Shard）和副本（Replica）机制实现水平扩展和高可用性。集群节点可以动态扩容并自动平衡数据负载，同时支持主节点、数据节点、协调节点等角色划分，以优化资源分配。

在大语言模型方面，

1.3 本文研究意义

在研究意义部分，阐述你的研究对学术界和实践领域的重要性。强调你的研究如何推动学科的发展，解决问题或者对社会产生积极影响。可以涉及到研究的创新性、实用性、理论意义等方面。此部分需要突出你的研究对于学术和实际领域的价值，使读者明白为何该研究是有意义的……

1.4 本章小结

小结部分是引言的收尾，总结引言中提到的关键信息。简要回顾研究的背景、问题、目的和结构，并为读者提供对整个研究的整体印象。强调研究的重要性，为后续章节的阅读做铺垫。小结要紧凑而有力，概括引言的核心内容，使读者对整篇论文的主旨有明确的认识……

2 正文文字格式

2.1 论文正文

论文正文是主体，一般由标题、文字叙述、图、表格和公式等部分构成。一般可包括理论分析、计算方法、实验装置和测试方法，经过整理加工的实验结果分析和讨论，与理论计算结果的比较以及本研究方法与已有研究方法的比较等，因学科性质不同可有所变化。

2.2 字数要求

2.2.1 本科论文字数要求

论文主体部分字数要求：理工类专业一般不少于 1.5 万字，其他专业一般不少于 1.0 万字。

2.3 本章小结

本章介绍了……

3 图表、公式格式

3.1 图表格式

本章将主要介绍一些图表和公式的格式...

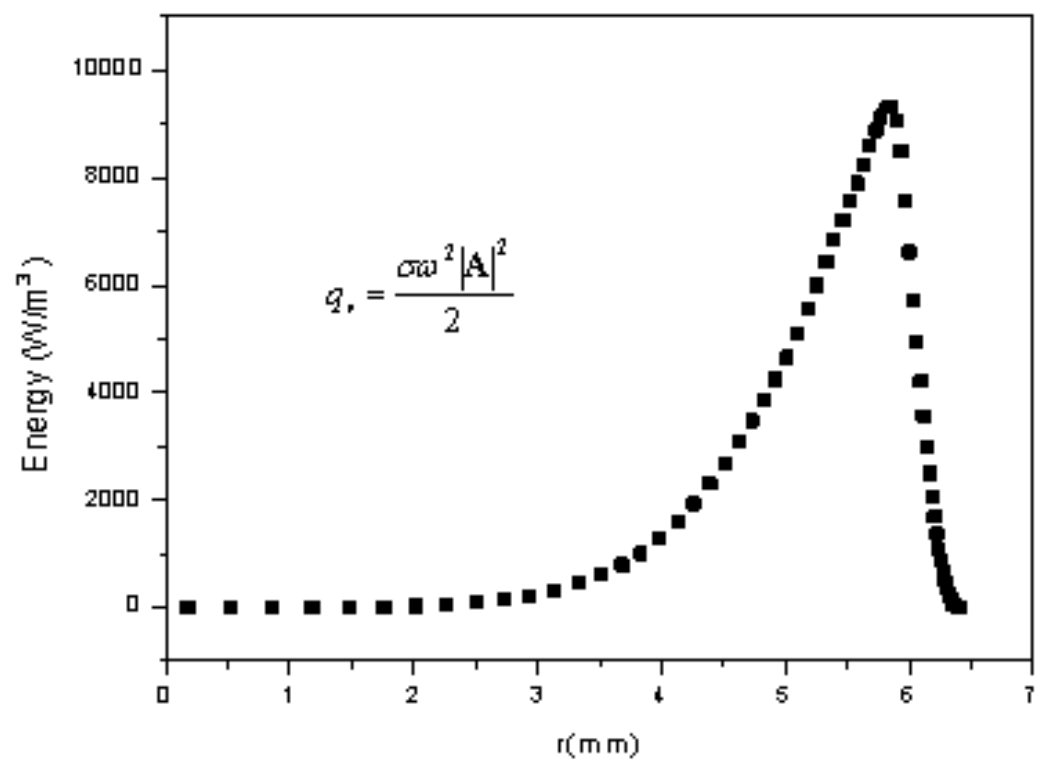


图 3.1 内热源沿径向的分布

3.2 公式格式

$$\frac{1}{\mu} \nabla^2 A - j \omega \sigma A - \nabla \left(\frac{1}{\mu} \right) \times (\nabla \times A) + J_0 = 0 \tag{3.1}$$

表 3.1 高频感应加热的基本参数

| 感应频率 (KHz) | 感应发生器功率 (%×80Kw) | 工件移动速度 (mm/min) | 感应圈与零件间隙 (mm) |
|---------------|---------------------|--------------------|------------------|
| 250 | 88 | 5900 | 1.65 |
| 250 | 88 | 5900 | 1.65 |
| 250 | 88 | 5900 | 1.65 |
| 250 | 88 | 5900 | 1.65 |

续表 3.1

| 感应频率 (KHz) | 感应发生器功率 (%×80Kw) | 工件移动速度 (mm/min) | 感应圈与零件间隙 (mm) |
|---------------|---------------------|--------------------|------------------|
| 250 | 88 | 5900 | 1.65 |
| 250 | 88 | 5900 | 1.65 |

3.3 本章小结

本章介绍了……

4 结论与展望

4.1 主要结论

本文主要……

4.2 研究展望

更深入的研究……

参考文献

- [1] 杨瑞林, 李力军. 新型低合金高强韧性耐磨钢的研究 [J]. 钢铁. 1999(7): 41-45.
- [2] 于潇, 刘义, 柴跃廷, 等. 互联网药品可信交易环境中主体资质审核备案模式 [J]. 清华大学学报 (自然科学版), 2012, 52(11): 1518-1523.
- [3] Schinstock D.E., Cuttino J.F. Real time kinematic solutions of a non-contacting, three dimensional metrology frame[J]. Precision Engineering. 2000, 24(1): 70-76.
- [4] 温诗铸. 摩擦学原理 [M]. 北京: 清华大学出版社, 1990: 296-300.
- [5] 蒋有绪, 郭泉水, 马娟, 等. 中国森林群落分类及其群落学特征 [M]. 北京: 科学出版社, 1998: 5-17.
- [6] 贾名字. 工程硕士论文撰写规范 [D]. 重庆: 重庆大学, 2000: 177-178.
- [7] 张凯军. 轨道火车及高速轨道火车紧急安全制动辅助装置: 201220158825.2[P]. 2012-04-05.
- [8] 全国信息与文献标准化技术委员会. 文献著录: 第 4 部分非书资料: GB/T 3792.4-2009[S]. 北京: 中国标准出版社, 2010: 3.

附录 A：XX 公式的推导

XX 公式的推导过程是：

致 谢

致谢主要感谢导师和对论文工作有直接贡献和帮助的人士和单位。致谢言语应谦虚诚恳，实事求是。

原创性声明

郑重声明：所呈交的论文（设计）《_____》，是本人在导师的指导下，独立进行研究取得的成果。除论文（设计）中已经标注引用的内容外，本论文（设计）不包含其他人或集体已经发表或撰写过的作品成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律后果，并承诺因本声明而产生的法律结果由本人承担。

论文（设计）作者签名：_____

日期：_____

使用授权书

本论文（设计）作者完全了解学校有关保留、使用论文（设计）的规定，同意学校保留并向国家有关部门或机构送交论文（设计）复印件和电子版，允许论文（设计）被查阅和借阅。本人授权重庆大学将本论文（设计）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制方式保存和汇编本论文（设计）。

本论文（设计）属于：

保 密 ☐ 在_____年解密后适用本授权书

不保密 ☐

论文（设计）作者签名：_____ 指导教师签名：_____

日期：_____ 日期：_____