

大模型在金融报告智能撰写中的应用实践

向俊夫 / 首席架构师

- 大模型在金融报告中的应用现状
- 大模型驱动的智能金融报告撰写
- 技术实现与挑战
- 未来展望

01

大模型在金融报告中的应用现状

目前，市场上已有多家公司开始将大模型技术应用于金融报告生成领域。主要应用方向包括：

数据处理与报告生成

- 某公司开发的系统能够处理复杂数据，将其自动输入和更新到写作材料中。
- 该系统在数据点识别和语义表格描述生成方面有所突破，有助于提高报告的准确性和一致性。

文档智能问答

- 另一家公司推出了基于大模型的文档智能问答系统。
- 用户可以上传文档并进入对话模式，系统协助用户解读文档内容，便于快速获取关键信息。
- 这种方式可以加速文档审阅过程，尤其适用于处理长篇复杂的金融报告。

财务报告审核

- 同一家公司还开发了针对上市公司年报的智能审核产品。
- 该产品主要功能包括差异对比、勾稽校验和财务错误识别等，有助于提高审计效率。
- 其目标是提高信息披露质量和审核效率，降低人为错误风险。

共同特点

- 这些应用都展现了大模型在金融文档处理领域的潜力。
- 它们提高了文档处理的效率和准确性，减少了重复性工作。
- 这些系统进行更深入的数据分析，为金融决策提供更全面的支持。

02

大模型驱动的智能金融报告撰写

功能与应用

文本生成

- 基于内部数据库和用户需求，生成高质量分析性段落
- 涵盖市场动态、发行人重大事件、财务分析等专业内容
- 支持根据用户反馈进行实时调整，提升内容的可用性和个性化

文档质量校验

- 识别多种潜在问题：
 - 数值一致性检查
 - 公式计算验证（不受格式限制）
 - 合规性审核（对照行业规范和监管要求）
- 全面文本纠错，包括错别字、标点、语法和表达准确性

数据解析与处理

- 解析用户底稿，提取关键数据和信息
- 自动清洗和标准化采集数据
- 对数据进行智能标签化，便于后续分析
- 支持定制化解析策略，适应不同需求

文本生成

- 根据内部数据库及指定文段的要求，生成分析性段落，包括市场动态、发行人重大事件概述
- 根据用户上传文本，生成段落，包括发行人财务分析、发行人走访记录等；
- 能根据用户指令，对初次生成的文本进行调整，提升内容可用性，迅速构建并调整报告模板，输出内容体现个性化视角；

文档质量校验

- 识别文档中的各类潜在问题，包括但不限于：

数值冲突——自动比对不同部分的数据是否一致，避免信息矛盾；

公式计算错误——精确校验勾稽关系，验证数据准确性，不需要依赖统一文本格式；

合规问题——比对行业规范与监管条文，确保文本表述是否有关键内容缺失或其他合规性问题；

- 文本纠错能力，识别并提示错别字、标点错误、语法错误、不严谨表达等，提升文档的专业性和可读性；
- 溯源至原始数据源或文档底稿，帮助用户迅速定位错误点，大大缩短了问题排查与修正的时间；

数据解析与处理

- 构造Prompt对各类金融公告进行智能解析，确保关键数据和信息被准确无误地提取出来；
- 自动对采集数据进行清洗和标准化处理，解决数据格式不统一、存在错误字符、字段缺失等问题；
- 即时对数据进行标签化处理，如“营收增长”、“净利润率”、“项目赛道”等标签，极大便利后续的数据分析与应用；
- 支持定制化解析策略，根据实际需求调整解析规则，以适应不同类型的金融公告或特定的数据提取需求。

03

技术实现与挑战

传统方法

1. 模板化写作

- 预先定义模板，插入变量。使用规则引擎和简单的条件判断来生成报告。

问题：

- 模板灵活性差，难以应对复杂的语境和多变的数据。
- 更新和维护模板需要大量人工干预，效率低下。
- 生成的文本较为死板，缺乏自然的语言流畅性。

2. 规则驱动的自然语言生成

- 基于预定义的规则和逻辑，通过编程实现自动化文本生成。使用语法和词汇规则，结合数据源进行写作。

问题：

- 规则的制定和维护复杂，需专家参与，成本高。
- 对于复杂的语境和非结构化数据处理能力有限。
- 难以生成高质量、多样化和富有创意的内容。

3. 统计语言模型

- 利用统计学方法，如n-gram模型，来预测文本片段。通过大量语料库训练模型，实现基本的文本生成。

问题：

- 需要大量标注数据进行训练，成本高昂。
- 模型生成的文本质量和流畅度有限。
- 无法处理长上下文和复杂的语义关系。

模型选择

挑战

能力/资源/安全的取舍平衡

- 高性能模型如GPT-4提供卓越的能力，但资源消耗巨大
- 本地化模型资源消耗相对较低，但能力有限。
- 在线API模型容易部署但数据安全性存在风险
- 本地化模型尽管成本较低，但仍需专门团队进行维护和更新。

模型选择

在线API

代表： GPT-4、Claude、文心一言、通义千问

特点：

- 规模大：** 千亿至万亿级参数。
- 能力强：** 在评测和实际体验中名列前茅，适用于复杂和高要求的任务。
- 资源占用多：** 需要高性能计算集群（如千卡、万卡集群）和专属硬件支持。

本地化部署

类型：

- 纯自研
- 基于开源模型

特点：

- 规模相对小：** 通常为十亿至百亿级参数。
- 能力一般但保持特色：** 适用于特定领域的应用，能够定制化。
- 资源占用相对合理：** 运行在单卡或多卡集群上，适合特定应用场景企业本地化部署

模型应用

挑战

研究和应用落地之间的落差：

- 研究中的先进模型和技术在实际应用中可能面临适应性和有效性的问题。

透明性和可解释性：

- 大模型的“黑箱”特性使得解释其决策过程变得困难，特别是在需要高透明度的金融行业。

专业知识和领域适应：

- 大模型需要结合专业知识进行领域适应，确保其生成的内容符合行业标准和要求。

更新维护：

- 金融行业变化快，模型需要持续更新和维护，以保持其有效性和准确性。这增加了维护成本和技术挑战。

模型应用

大模型的底层逻辑： 通过预测下一个token实现自然语言生成

使用模式：

Prompt工程： 通过设计合适的提示（prompt）来引导模型生成预期的内容。

- 数据提取：** 提供特定的文本和指令作为提示，让模型准确提取需要的字段数据。

RAG（检索增强生成）： 结合检索系统和生成模型，通过先检索相关信息再生成内容，提高生成的准确性和相关性。

- 市场分析：** 先检索最新市场数据，然后生成包含实时信息的市场分析报告。

微调： 在大模型的基础上，使用特定领域的数据进行进一步训练，以提升在特定任务上的表现。

- 规范性段落生成：** 使用金融客户的历史报告数据微调模型，稳定模型输出的风格和模式。

继续预训练： 在现有模型的基础上，继续进行预训练，使用更多的领域相关数据提高模型的整体能力。

- 领域知识增强：** 利用金融场景的海量数据继续预训练模型，增强其金融领域背景知识。

纯自研： 完全从头训练一个可用的大语言模型。

04

未来展望

未来展望

多模态数据融合

- 未来将融合文本、图像等多模态数据，生成更丰富的金融报告。

个性化与智能化

- 大语言模型将进一步支持个性化内容生成，满足用户多样化需求。

持续优化与创新

- 随着技术的发展，大语言模型将不断优化和创新，提供更高质量的服务和解决方案。

THANKS

DataFunSummit # 2024

