**专业实践重点专题实践周报（2023.07.03-2023.07.07）**

杜国豪，面壁智能

1. 本周工作总结及收获

本周我主要进行了一家金融客户相关网站的调研，并成功爬取了航运公司的网站数据。通过数据清洗后，我进行了大语言模型的训练。在这个过程中，我掌握了调研方法和爬虫技术，并对数据清洗和语言模型训练有了更深入的了解。通过这次实践，我对金融行业和航运行业有了更全面的认识，并且提升了数据处理和机器学习的技能。

1. 本周工作存在的问题

在本周的实践过程中，我遇到了一些问题。首先，在爬取航运公司网站数据时，遇到了一些反爬虫机制，需要花费一些时间和精力去绕过。其次，在数据清洗过程中，我发现了一些噪声数据和缺失值，需要进行进一步处理和填充。这些问题对我的工作进度造成了一定的影响

1. 下一步工作计划

在接下来的工作中，我计划解决本周存在的问题并继续推进实践项目。具体工作计划包括：

对航运公司网站的反爬虫机制进行进一步研究，找到更好的方法来绕过限制。

继续进行数据清洗工作，处理噪声数据和填充缺失值，确保训练数据的质量。

探索更多的大语言模型训练方法和技巧，提高模型的性能和准确性。

分析训练得到的模型结果，与金融客户讨论并收集反馈，进行进一步的改进和优化。

1. 其他（可另起标题）

本周我还参加了一场关于金融科技的线上研讨会，与行业专家和同行进行了深入的交流和讨论。这次研讨会对我理解金融科技的发展趋势和应用前景有了很大的帮助。我将把这些新的见解和知识应用到我的实践项目中，进一步提升实践的质量和价值。