



好的，这是将您提供的英文会议纪要逐字翻译成的流畅地道中文，供中文母语者阅读。

---

## 洞察见解 (Insights)

- **职业路径指导：** Brad 就如何在不断变化的数据科学领域中找准自己的方向，以及如何在机器学习工程师和数据科学家之间做出选择寻求建议。
- **角色辨析：** Roxanne 区分了机器学习工程师（MLE）、数据科学家和数据分析师这些角色，强调了它们在工作内容模糊性上的不同程度，以及人工智能自动化对每个角色的影响。
- **技能培养：** Roxanne 建议 Brad 优先学习机器学习理论并积累动手实践经验，同时在编程技能方面重点关注 SQL 和 Python。
- **未来需求：** Roxanne 预测，由于人工智能的普及应用会带来日益增多的复杂性和问题，未来市场对数据科学家的需求将会持续存在。
- **时间分配：** Roxanne 建议将 60-70% 的学习时间分配给机器学习，30-40% 分配给 SQL、Python 和编程练习。

## 行动要点 (Action items)

- Brad 将专注于学习机器学习理论及相关的动手实践经验。
- Brad 将利用 LeetCode 等资源练习 SQL 和 Python 编程。
- Brad 将在几个月后与 Roxanne 安排下一次会谈，以讨论他的进展。
- 如果 Brad 觉得这次会谈有帮助，他将为此留下评论。

## 会议亮点 (Highlights)

- **开场与背景：**
  - Brad 为这次咖啡交流表示感谢，并对之前的延期表示歉意。
  - Roxanne 确认她可以提供导师指导，并对 Brad 使用该平台的体验表示关心。
  - Brad 提到他已安排了与另一位专攻机器学习工程领域的导师进行交流。
- **在路透社AI的角色：**
  - Brad 询问 Roxanne 在路透社AI（Reuters AI）的角色，特别是该角色更偏向于生成式AI研究还是传统的数据科学。
  - Roxanne 解释说，她的角色融合了生成式AI研究和数据科学，专注于应用人工智能

和智能体AI (agentic AI) 来处理多模态新闻数据 (文本、图片、视频)。

- 其目标是将数据变现，并提高记者和新闻制作的效率。

#### • 用例讨论：

- Brad 询问路透社AI团队使用了哪些具体的问题解决方法。
- Roxanne 举了一个高层次的例子：使用“检索增强生成” (RAG) 技术进行内容搜索，使记者能用更复杂的描述进行搜索，并利用大语言模型 (LLM) 生成文章摘要。

#### • 职业路径的困境：

- Brad 分享了他的学术背景，并表达了对数据科学领域不断发展的担忧。
- 他正处于一个十字路口，不确定是该追求机器学习工程师的职业道路，还是成为一个更偏向业务的数据科学家。
- Roxanne 建议将这个比较拆分为“机器学习工程师 vs. 数据科学家”和“侧重建模的数据科学家 vs. 侧重业务的数据科学家”两个层面来分析。

#### • 机器学习工程师 vs. 数据科学家：

- Roxanne 询问 Brad，相比于确定性和定义明确的问题，他个人是更偏好模糊性和新的可能性。
- Brad 表示，他在职业前景中看重成长空间和持续的市场需求。
- Roxanne 解释说，数据科学更多是关于反复试错，而机器学习工程师的角色则更注重稳定性和可扩展性。

#### • 技能与经验：

- Roxanne 询问 Brad 在应用开发、模型部署或模型规模化方面的动手经验。
- Brad 澄清说他仍是大一新生，正在规划课程，目标是为统计学和数学打下坚实的基础。
- 他承认自己可能缺乏软件工程技能，这或许会成为他成为机器学习工程师的一个障碍。

#### • 学业规划与黑客松 (Hackathons)：

- Roxanne 称赞 Brad 很早就开始规划自己的学业路径。
- 她建议通过参加黑客松来积累软件工程经验，但也承认多伦多大学的课业非常繁重。

#### • 实习机会：

- Roxanne 指出，与滑铁卢大学相比，多伦多大学更侧重于学术研究而非实践工作经验。
- Brad 提到了文理学院实习项目（ASIP）这个选项，但不确定其价值。
- 考虑到人工智能的自动化趋势，他寻求 Roxanne 关于数据科学与机器学习工程的前景和市场需求的看法。
- **市场需求与AI自动化：**
  - Roxanne 预测，机器学习工程的初期需求可能更高，但从长远来看更容易被AI取代。
  - 她认为，未来将需要数据科学家来解决因AI普及应用而产生的问题。
  - Roxanne 强调了数据科学工作中的模糊性以及确保AI系统可靠运行所面临的挑战。
- **区分数据科学家角色：**
  - Brad 询问如何区分数据科学家、数据分析师和机器学习工程师。
  - Roxanne 解释说，数据科学家进行实验，而数据分析师准备数据。
  - 机器学习工程师则负责将数据科学家写出的有效代码部署到生产环境中。
- **模糊的界限与公司实践：**
  - Brad 指出，数据科学家和数据分析师角色之间的界限正变得越来越模糊。
  - Roxanne 澄清说，这通常是因为公司希望以数据分析师的头衔支付较低的薪水，来做数据科学的工作。
  - 她强调，每个角色的技能要求是截然不同的。
- **后续步骤与技能发展：**
  - Brad 讨论了他的自学努力，包括HTML网页应用开发和SQL教程。
  - 他询问是应该优先学习机器学习专业知识还是SQL技能。
  - Roxanne 建议不要去考数据分析师证书，而是应该专注于SQL和机器学习。
- **SQL和机器学习的重点：**
  - Roxanne 建议在准备面试时，可以在 LeetCode 上复习一下SQL技能。
  - 她强调了理解机器学习模型如何工作的重要性，并利用AI将理论知识与实际应用结合起来。
- **优先学习机器学习理论：**
  - Brad 询问是应该优先考虑机器学习理论还是实际的编程技能。

- Roxanne 建议学习机器学习理论，并学习基础知识，而不是依赖AI来写代码。
- **区分技能：**
  - Brad 就如何在 SQL/Excel 和学习机器学习之间分配时间寻求建议。
  - Roxanne 建议不要花时间学习Excel，因为这个技能很容易上手，而且每个公司的用法都不同。
- **时间分配与动手经验：**
  - Roxanne 建议将 60-70% 的时间分配给机器学习，30-40% 分配给SQL和Python。
  - 重点应该放在机器学习的动手实践上，包括数据清洗、模型训练和评估。
- **结束语：**
  - Roxanne 表示可以提供进一步的帮助，并建议几个月后再安排一次会谈。
  - 她鼓励 Brad 为这次会谈留下评论。

### **内容概要 (TLDR)**

这次一对一的会谈聚焦于 Brad Xu 在数据科学和机器学习领域的职业抱负。Roxanne Li 作为路透社AI的一名生成式AI研究员兼数据科学家，为多伦T多大学大一新生 Brad 提供了指导和建议。他们讨论了数据科学领域不断变化的格局、机器学习工程师（MLE）和数据科学家角色之间的区别，以及 Brad 应如何培养相关技能并为其学术和职业道路做出明智决策的策略。Roxanne 强调了理解机器学习理论的重要性，并鼓励 Brad 专注于机器学习、SQL和Python的动手实践。她还建议现阶段不必追求数据分析师证书或学习Excel，因为这些技能以后更容易掌握。会谈最后，Roxanne 表示愿意提供进一步的帮助，并鼓励 Brad 为本次会谈留下评论。