

从“法学硕士”到“AI搜索”，LLM如何驱动更智能的搜索？


原创 我是哥飞 哥飞 2024-04-08 22:00:22 广东

大家好，我是哥飞。

可能部分朋友看到标题里的“法学硕士”会有些懵，这是啥？

哥飞给大家解释一下，“法学硕士”是一个梗，只要是用翻译软件——尤其是谷歌翻译——看多了AI相关的资讯，就会时不时地遇到这种情况，“LLM”被翻译为了“法学硕士”。

而在AI领域，其实“LLM”是“Large Language Model”的缩写，即“大语言模型”的意思。



Jeffrey

confident-ai.com

@jeffr_yyy

关注

It is "an" LLM, not "a" LLM

由 Google 翻译自英语

它是“一个”法学硕士，而不是“一个”法学硕士

如上图所示，在社交媒体上，有人会用这个梗来写冷笑话，谷歌翻译出来的依然还是错误的。

为什么会出现这种错误呢？

为什么“LLM”会被翻译为“法学硕士”

回答 | 深入

27

“LLM”被翻译为“法学硕士”的原因主要是因为机器翻译系统在处理大量文本数据时，如果大部分情境下“LLM”都是与法律学术相关的，则系统更容易学习到“LLM”应当被译为“法学硕士”^②。此外，“LLM”在教育领域是“Legum Magister”或“Master of Laws”的缩写，意为“法学硕士”，这进一步证实了其在学术界的普遍认知^⑩。尽管“LLM”也指代当前备受瞩目的“大语言模型 Large Language Model”，但在实际应用中，由于“大语言模型”在绝对数量级上处于下风，机器翻译系统还是默认将“LLM”翻译为“法学硕士”^④。这种现象并不是因为“大语言模型 LLM (Large Language Model)”真的和法学有强关联，而是只是一个缩写词的歧义解释而已^⑧。因此，尽管AI大语言模型的热度很高，但在机器翻译中，“LLM”仍然经常被错误地翻译成“法学硕士”。

延伸阅读

分享 深度研究 追问

大纲

缩写误解

- LLM在机器翻译中被误解
- 翻译系统依赖上下文学习特定词汇

法学硕士的定义

- Master of Law，一年制北美法学进阶课程
- 拉丁语来源，代表法学硕士Master of Laws

AI与法律学术相关性

- 大量文本数据训练导致对法律学术相关性的偏好
- 法律学术上下文中LLM出现频率高

社会认知差异

- AI热潮下，LLM与法学关联的认知差异
- 法学硕士头衔可能因AI而改变

教育领域应用

- LLM作为教育层次与内容的普遍认知
- JD、J.S.D.等其他学位的中文名称调整，反映教育层次变化

生成演示文稿

相关事件

事件名称	事件时间	事件概述
JD证书中文名称调整及学历层次明确表述 ^⑫	2019年7月2日	教育政策更新 中华人民共和国教育部留学服务中心将JD证书的中文名称统一调整为“职业法律博士文凭”，并明确美国和澳大利亚JD证书的学历层次为硕士层次。

我们拿这个问题去问一下秘塔AI搜索，得到的回答如上所示，是我们期望的正确答案，不仅提到了“Master of Laws”，也提到了“Large Language Model”。

NS

You

为什么“LLM”会被翻译为“法学硕士”

**ChatGPT**

"LLM"是英文中的缩写，代表法学硕士（Master of Laws）。"LLM"中的每个字母分别代表着不同的词，"L"代表法律（Law），"M"代表硕士（Master）。因此，将"LLM"翻译为“法学硕士”是合理的，因为它准确地传达了这个学位的含义。

NS

You

“Large Language Model”是什么

**ChatGPT**

“Large Language Model”（LLM）是指大型语言模型，它是一种基于人工智能的自然语言处理技术，旨在理解和生成自然语言文本。LLM能够处理大量的文本数据，从而学习语言的模式、结构和含义，并且能够生成具有语法正确性和上下文连贯性的文本。这种类型的模型通常使用深度学习技术，如循环神经网络（RNN）或变换器（Transformer），并且通过大规模的预训练和微调来提高性能。OpenAI的GPT系列就是一种LLM，其中的GPT-3就是目前为止最大的一个。

NS

You

为什么有时候在AI领域的“LLM”会被翻译为“法学硕士”

**ChatGPT**

在AI领域，“LLM”通常被翻译为“Large Language Model”（大型语言模型），因为它是指一种特定的人工智能模型类型。然而，在某些特定的上下文中，尤其是在非专业人士或非英语母语人士的交流中，可能会出现误解或混淆，导致将“LLM”误译为“法学硕士”。

这种误译可能是因为人们不熟悉“Large Language Model”这一术语，而将其与熟悉的“LLM”（法学硕士）混淆。此外，有些机器翻译工具可能会根据上下文错误地将“LLM”翻译为“法学硕士”。

因此，在沟通和翻译中，尤其是涉及到专业术语的时候，确保清晰地表达和理解术语的含义非常重要。



而如果拿同样的问题去问GPT3.5，就会发现得到的回答只提到了“Master of Laws”，于是哥飞不得不手动补充背景信息来提示GPT，终于也得到了正确答案。

从上面的过程中，我们会发现，GPT等大语言模型基于历史信息去训练，只要被训练出来了，知识库就已经过期了。



You

你的最新知识库截止到哪一天？



ChatGPT

我的最新知识截止到2022年1月份。



虽然时不时会拿最新的知识库去训练，但也追不上最新的动态，如GPT最早的知识库是2021年9月，后来更新到了22年1月。

但是现在已经2024年4月了，也就是说，GPT3.5的知识库已经落后现实世界26个月了。

那么要怎么样才能让AI能够回答最新的问题呢？

答案是把大模型和传统搜索结合，开发一个AI智能搜索引擎出来。

这就是今天哥飞要给大家介绍的主角，秘塔AI搜索。

这是一个崭新的产品，最大的特点就是完全无广告，免登录注册即可使用。



没有广告，直达结果

点击输入想了解的问题

● 全网 ▾



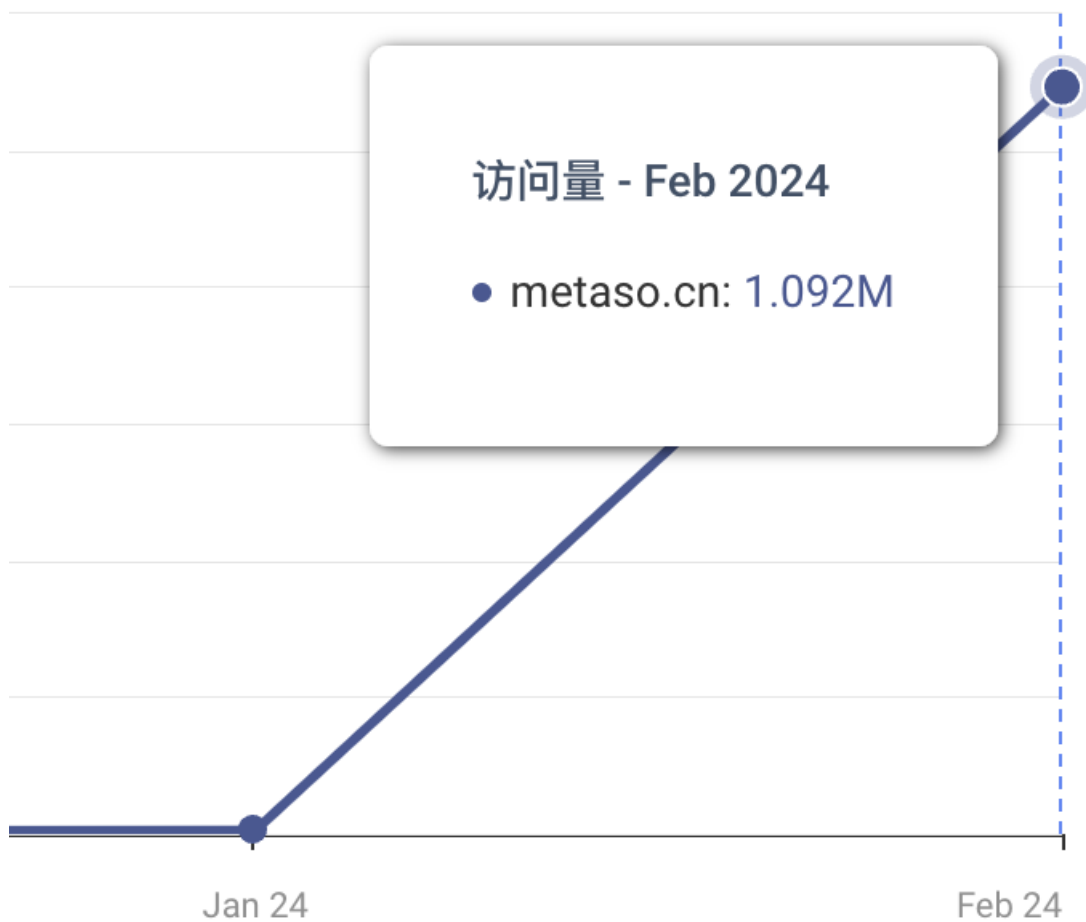
简洁

深入

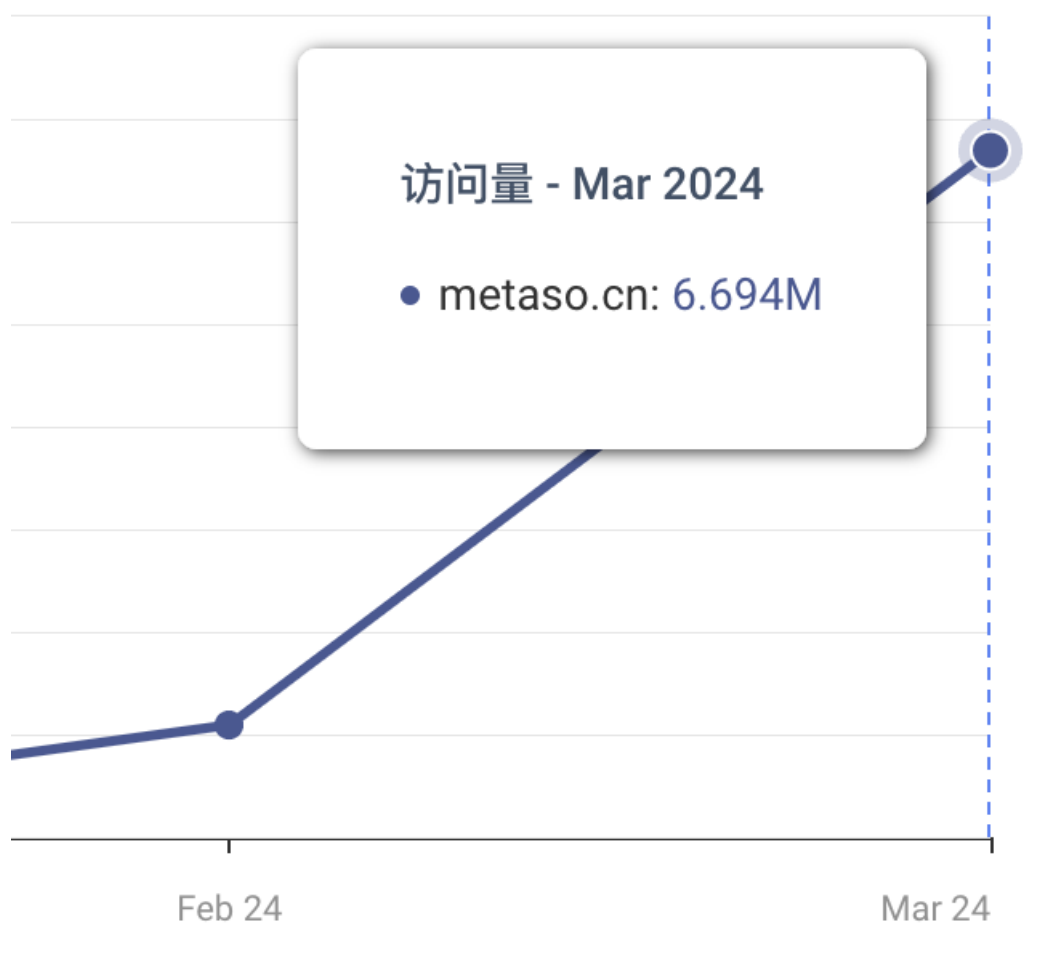


研究

用 Similarweb 看下这个网站的流量数据，2月访问量109万，可谓上线就爆发。



再看3月流量，简直坐火箭，从109万直接涨到了669万。



为什么？
为什么流量能坐火箭？
因为完全免费且真的好用，大家用了都喜欢用。

哥飞是谁？

回答 | 深入

5

哥飞是一个多面的人物，根据我搜索到的资料，我们可以看到他在不同领域有着不同的身份和成就。

哥飞在商业、SEO技巧和AI赚钱方面有深入的分享和实践价值，表明他在这个领域具有敏锐的商业嗅觉和丰富的知识^①。此外，他在即刻App上的个人主页显示，他已经写了225篇教程，其中220篇是公开免费的原创内容，这进一步证明他在分享知识和经验方面的乐于助人和专业性^②。

然而，需要注意的是，证据中还提到了其他几位“飞哥”，他们的背景和活动领域各不相同。例如，“飞哥简历”中的飞哥是一位CCIE国际认证程序员、IT数码极客、中国数码摄影家协会会员、抖音旗下图虫签约摄影师以及湖北省宜昌市伍家区剧协副主席^③。而“科技小飞哥”则是一位定居新加坡的科技互联网领域自媒体博主^④。另外，“有趣的飞哥”是贵州兴义人，80后小伙，入驻抖音后拥有170000粉丝，涉及三农产业^⑤。

哥飞是一个在商业、SEO技巧和AI赚钱方面有深入分享和实践的人，具有敏锐的商业嗅觉和丰富的知识^①^②。但是，由于证据中提到了多位不同的“飞哥”，他们的背景和活动领域有所不同，因此在讨论哥飞时需要明确具体指的是哪一位“飞哥”。

延伸阅读

分享

深度研究

追问

大纲

×

个人背景

- 贵州兴义人，80后小伙
- CCIE国际认证程序员
- IT数码极客
- 抖音旗下图虫签约摄影师
- 湖北省作家协会会员
- 湖北省宜昌市伍家区剧协副主席

教育与职业经历

- 高等教育背景
- IT行业多年经验

社交媒体影响力

- 抖音粉丝170000
- 入驻抖音时间：2020年6月底

商业智慧与分享

- 商业嗅觉敏锐，乐于分享
- 分享内容含金量高，有实践价值

主要成就

- 公众号原创输出225篇教程，全部公开免费可看全文
- 教程面向读者想要通过做网站出海赚钱的人群

生成演示文稿

如我们去问秘塔AI搜索“哥飞是谁”，秘塔AI搜索会把所有可能的答案都列出来，而不是只回答一个。
AI并没有自作主张的选择一个最可能的答案，也没有把多个人当作一个人，而是给我们选择，我们可以去追问。
也就是说，筛选信息，主导信息的决定权依然交回了我们用户手上。

我问的是即刻哥飞，写公众号的那个，教SEO的那个。

回答 | 深入 11

哥飞在即刻App上分享了关于SEO的详细信息和成就，具体包括：

- 哥飞通过自己的经验和理解，基于原文www.searchenginejournal.com，给大家讲解了每一个重要的SEO元标签，这些内容被分为三篇文章发布 ^①。
- 他分享了如何通过SEO和AdSense优化，将一个网站的广告收入从每月八百多美元提升到每月两千多美元的经验，这一系列内容共有122篇相关文章 ^②。
- 哥飞还提供了一系列初级入门教程，包括如何挖掘需求以及针对目标关键词优化网站SEO的一步一步指南 ^③。
- 他还推荐了搜索引擎官方出的文档作为学习SEO的最佳教程，强调了谷歌的SEO文档的重要性 ^⑦。
- 在推广方面，哥飞提到了借鉴犹太人的二八定律，即花费20%的时间创作内容，80%的时间去推广的重要性 ^⑧。
- 此外，哥飞还分享了三个月的分享大揭秘，内容涵盖商业智慧、SEO技巧、AI赚钱等方面 ^⑩。

哥飞在即刻App上关于SEO的教学不仅包括了基础理论的学习，还涉及到了实际操作经验和策略的分享，以及对SEO工具和资源的推荐。他的成就体现在成功地将个人经验和理解转化为有价值的内容，帮助他人提升SEO技能，并且通过实践案例展示了SEO优化的实际效果。

延伸阅读

分享 深度研究 追问

通过在刚才的问题下方继续追问，指定即刻哥飞之后，AI就能够确定我们想要了解的到底是哪一位，于是就能够给出准确的回答。

传统的搜索引擎要用好，还是有一定的门槛的，甚至有了“搜商”的说法。

搜商是什么？请简单介绍

回答 | 深入 19

搜商，全称为搜索智力（Search Quotient），是一种通过工具获取新知识的能力，与智商（IQ）、情商（EQ）并列为人类的三种智力因素之一 ^②。它强调的是在信息爆炸的时代，知道如何高效、快速地获取所需的知识，并对这些知识进行融合以解决实际问题的能力 ^④。搜商的核心包括搜索意识、搜索渠道和搜索技术 ^①，这表明搜商不仅仅是使用搜索引擎这么简单，而是涉及到对信息的敏感度、获取信息的渠道选择以及利用技术手段进行有效搜索的综合能力。

随着互联网尤其是移动互联网的发展，搜商成为了一个重要的职场能力。在信息数据容易获得的今天，如果缺乏高效的搜商能力，个人在信息获取上可能会比他人慢，从而影响工作效率和效果 ^⑥ ^⑩。因此，搜商被视为一种新时代的能力，对于提高个人的工作效率和适应现代社会的需求具有重要意义。

搜商的提出和重视，反映了在信息化社会中，人们对于信息获取能力的需求日益增长。它不仅仅是一种技能，更是一种能够帮助人们在海量信息中快速找到所需知识，有效解决问题的能力 ^⑨ ^⑫。此外，搜商还被认为是一种“连接式”学习的方式，需要好奇心、关键词的运用以及有效渠道的选择相结合 ^⑮。

搜商是现代社会中一个非常重要的能力，它不仅关系到个人的学习和工作效率，也是适应未来社会发展的一个关键素质。

延伸阅读

分享 深度研究 追问

大纲 生成演示文稿

SEO元标签详解

- Title、Description和Headings的重要性

网站优化经验分享

- 提升网站广告收入
- 初级入门教程

微信公众号SEO策略

- 公众号认证与关键词布局
- SEO友好前端模板的应用

SEO工具与资源

- 搜索引擎官方文档的学习

推广与流量获取

- 强化课程宣传推广
- 利用公众号SEO排名优化获取流量

商业智慧与SEO技巧

- 分享商业智慧、SEO技巧、AI赚钱方法

SEO的直接收益与品牌影响

- 订单转化与潜在客户建立

生成演示文稿

大纲 生成演示文稿

定义

- 搜商：搜索智力，获取有用信息的能力
- 缩写：SQ（SearchQuotient）

核心要素

- 搜索意识：万物皆有路径且可达的信念
- 搜索渠道：使用工具获取新知识的能力
- 搜索技术：高效、快速获取所需知识并解决问题的能力

重要性

- 被视为与智商、情商并列的人类智力因素
- 在信息爆炸的21世纪尤为重要

发展背景

- 随着互联网尤其是移动互联网的发展而诞生的新能力

实践意义

- 提高工作效率和信息获取效率
- 在必要工作中避免信息获取上的滞后

培养方法

- 培养好奇心，不满足于搜索引擎结果
- 学习和应用搜索技巧，如关键词优化、使用高级搜索命令等

生成演示文稿

而秘塔AI搜索就几乎没有使用门槛了，我们可以用自然语言把我们的疑问问出来，之后等待它的回答就好了。

秘塔AI搜索的大概原理，哥飞给大家介绍一下，用户提问之后，它会进行语义分析，然后检索互联网的信息，再整合所有搜索到的资料来回答用户的问题。

也就是在输入和输出时，都经过了大语言模型的“思考”，于是可以更准确地理解用户的问题，并且个性化地回答用户的提问。

相对于传统搜索引擎，只给出网页链接，不给总结来说，秘塔AI搜索真是个好用的产品。

不仅如此，秘塔AI搜索还有更进阶的玩法，直接丢一个链接给它，让它提炼文章重点。

读文章列出Transformer构架的优劣

<http://mtw.so/62QzLw>

阅读

Transformer 架构的优劣如下：

优点：

- 1. **解决并行计算问题：**Transformer 模型通过其独特的自注意力机制，有效地解决了传统 RNN 模型在处理长序列数据时无法有效并行计算的问题，显著提高了算法效率。
- 2. **强大的性能和广泛的应用：**Transformer 架构因其出色的性能，在自然语言处理、图像识别等多个领域得到了广泛应用，并且随着模型规模的扩大，其智能水平不断提升。
- 3. **促进 AI 技术的发展：**Transformer 的出现推动了生成式人工智能的发展，为后续的研究和应用提供了新的方向和可能性。

缺点：

- 1. **高计算资源需求：**尽管 Transformer 解决了并行计算的问题，但其训练和推理过程仍然需要大量的计算资源，这限制了其在资源受限环境下的应用。
- 2. **模型复杂度高：**Transformer 模型的复杂度较高，对于初学者来说，理解和掌握其工作原理和实现细节存在一定的难度。
- 3. **对数据质量和量的依赖性强：**Transformer 模型的效果很大程度上依赖于高质量和大规模的数据集。在数据不足或质量不高的情况下，模型的性能可能会受到影响。
- 4. **可能存在的过拟合风险：**在某些情况下，Transformer 模型可能会出现过拟合现象，尤其是在训练数据相对较少的情况下，需要采取相应的正则化策略来缓解这一

大纲

优点

- 解决并行计算问题
 - 提高算法效率
- 支持大规模数据处理
 - 能够从大量数据中学习，找到关系和模式
- 促进模型的自我注意和梯度下降加速
 - 自我注意机制使模型能够更好地理解输入序列
- 扩展性和通用性
 - 可以应用于多种任务，如机器翻译、图像识别等
- 成本效益高
 - 模型规模越大，成本效益越高

缺点

- 计算资源需求巨大
 - 需要大量的GPU计算资源
- 对初始架构有较高要求
 - 初始架构设计需谨慎，否则可能导致性能瓶颈
- 难以直接应用于所有领域
 - 在某些特定领域可能不如其他模型表现优异

生成演示文稿

这个功能对于经常需要读长文章、快速提炼重点的人来说简直是神器。

通过前面的介绍可以发现，秘塔AI搜索能够极大地降低小白用户的使用门槛，甚至小学生都能用上。

小学生清明放假作文要怎么写呀，我假期跟爸爸妈妈去了大梅沙玩，好多人呀，我还去吃了海鲜。

回答 | 深入 30

写一篇关于清明节假期去大梅沙玩的作文，可以从以下几个方面入手：

- 1. **开头引入**：可以以一句诗或对清明节的描述作为开头，比如“清明时节雨纷纷，路上行人欲断魂”^④，以此引出你的旅行主题和背景。
- 2. **描述大梅沙**：介绍大梅沙的位置、特色以及你对这个地方的印象。可以提到大梅沙是深圳最大的海滨浴场，有着美丽的海滩和清澈的海水^{②③}。如果有机会，还可以描述一下你在大梅沙的活动，比如游泳、沙滩排球等^{⑱ ⑲}。
- 3. **人多的体验**：描述假期期间大梅沙的人流量很大，这可能会影响游玩体验，但也增加了节日的热闹气氛^⑦。可以谈谈你对这种人多场景的感受，是觉得有趣还是有些拥挤？
- 4. **海鲜美食**：详细描述你在大梅沙吃海鲜的经历。可以提到不同种类的海鲜，如鱼类、虾类、贝类等^{⑤ ⑨ ⑩}，并分享你最喜欢的一种或几种海鲜。同时，也可以谈谈海鲜给你带来的味觉享受和营养价值的认识。
- 5. **结尾总结**：最后，可以总结这次清明节假期的意义，不仅是家庭团聚的美好时光，也是体验自然、享受生活的一次机会。同时，也可以表达对传统文化节日的理解和尊重。

在写作时，注意结合自己的实际经历和感受，使作文内容真实、生动。此外，适当运用一些修辞手法和生动的描写，可以让作文更加吸引人。希望这些建议能帮助你写出一篇优秀的作文。

延伸阅读

分享 深度研究

追问

上面是哥飞的娃问的问题，秘塔AI搜索很好地解答了这个问题。

好了，哥飞今天的文章就介绍到这里了，如果大家如果想要体验秘塔AI搜索的话，请记住域名 metaso.cn，或者点击本文左下角的“阅读原文”，一键直达。

[阅读原文](#)

大纲

×

开头

- 引入清明节的背景
- 描述清明节的意义和习俗

大梅沙之旅

- 游玩体验分享
 - 海边游玩攻略
 - 推荐路线：海滨公园→滨海文旅艺术小镇→海滨栈道→盐田海鲜街
 - 游玩项目：摩托艇、沙滩排球等
- 环境感受
 - 海景欣赏
 - 海滨公园的自然风光

海鲜美食

- 海鲜的分类和特点
- 常见海鲜种类及其食用方法

结尾

- 总结大梅沙之旅的意义，表达对亲人的思念和对自然美景的喜爱

生成演示文稿