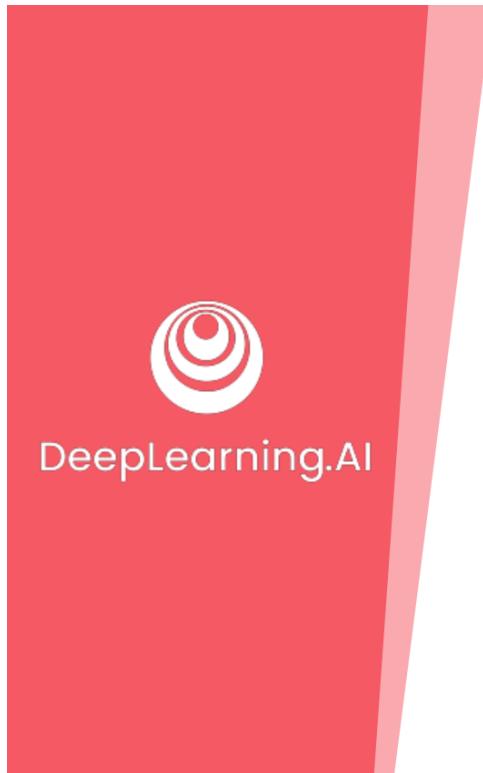


版权声明

这些幻灯片根据 Creative Commons License 分发。

DeepLearning.AI 将这些幻灯片用于教育目的。您不得出于商业目的使用或分发这些幻灯片。您可以复制这些幻灯片并将其用于教育目的，只要您引用 DeepLearning.AI 作为幻灯片的来源即可。

有关许可证的其余详细信息，请参阅 <https://creativecommons.org/licenses/by-sa/2.0/legalcode>



适合所有人的生
成式 AI

吴安德

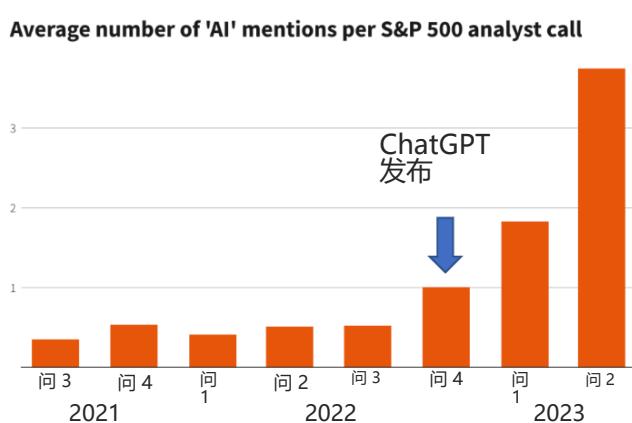


DeepLearning.AI

什么是生成式 AI

欢迎

生成式 AI 的兴起



生成式 AI 可以 • 增加 2.6-4.4 万亿美元

每年对经济的贡献 1 • 将全球 GDP 提高 7%

未来 10 年 2 • 影响 10% 的任务

80% 的工人每天进行 3

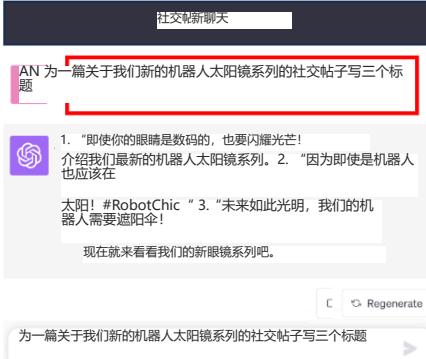
资料来源：(1) 麦肯锡，(2) 高盛，(3)
OpenAI/宾夕法尼亚大学

DeepLearning.AI

吴安德

什么是生成式 AI?

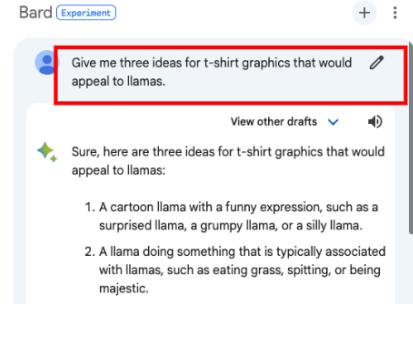
可以生成高质量内容（特别是文本、图像和音频）的人工智能系统。



AN 为一篇关于我们新的机器人太阳镜系列的社交帖子写三个标题

1. “即使你的眼睛是数码的，也要闪耀光芒！
介绍我们最新的机器人太阳镜系列。2. “因为即使是机器人也应该在
太阳！#RobotChic” 3. “未来如此光明，我们的机
器人需要遮阳伞！”
现在就来看看我们的新眼镜系列吧。

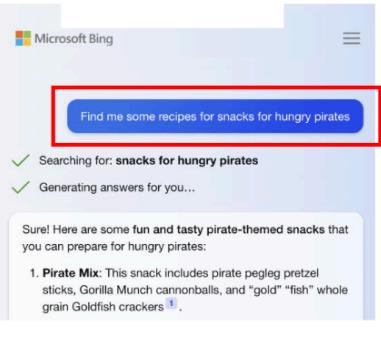
为一篇关于我们新的机器人太阳镜系列的社交帖子写三个标题



Give me three ideas for t-shirt graphics that would appeal to llamas.

Sure, here are three ideas for t-shirt graphics that would appeal to llamas:

1. A cartoon llama with a funny expression, such as a surprised llama, a grumpy llama, or a silly llama.
2. A llama doing something that is typically associated with llamas, such as eating grass, spitting, or being majestic.



Find me some recipes for snacks for hungry pirates

Searching for: snacks for hungry pirates
Generating answers for you...

Sure! Here are some fun and tasty pirate-themed snacks that you can prepare for hungry pirates:

1. **Pirate Mix:** This snack includes pirate pegleg pretzel sticks, Gorilla Munch cannonballs, and “gold” “fish” whole grain Goldfish crackers.

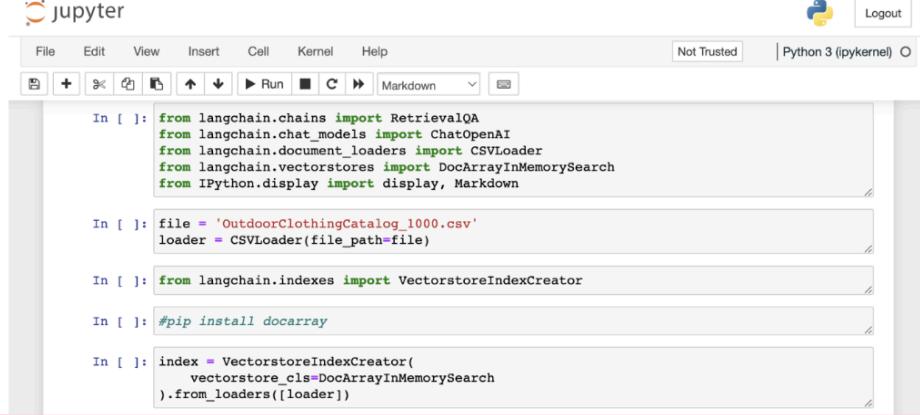
ChatGPT/OpenAI
巴德/谷歌
必应聊天/Microsoft

吴安德

什么是生成式 AI?

生成式 AI 也是一种开发人员工具

DeepLearning.AI
吴安德



```
In [1]: from langchain.chains import RetrievalQA
from langchain.chat_models import ChatOpenAI
from langchain.document_loaders import CSVLoader
from langchain.vectorstores import DocArrayInMemorySearch
from IPython.display import display, Markdown

In [2]: file = 'OutdoorClothingCatalog_1000.csv'
loader = CSVLoader(file_path=file)

In [3]: from langchain.indexes import VectorstoreIndexCreator

In [4]: #pip install docarray

In [5]: index = VectorstoreIndexCreator(
    vectorstore_cls=DocArrayInMemorySearch
).from_loaders([loader])
```

吴安德

AI 已经无处不在

AI 技术	例子
Web 搜索	谷歌、必应
欺诈检测	信用卡付款
推荐系统	亚马逊、Netflix

DeepLearning.AI

吴安德

图像、音频和视频生成

美丽的田园山景。山水画风 (Midjourney)



两只可爱的小猫咪在玩耍 (DALL-E)



DeepLearning.AI

吴安德

图像、音频和视频生成

语音生成



吴安德

图像、音频和视频生成

视频生成



视频克隆到这里。



吴安德

学习内容

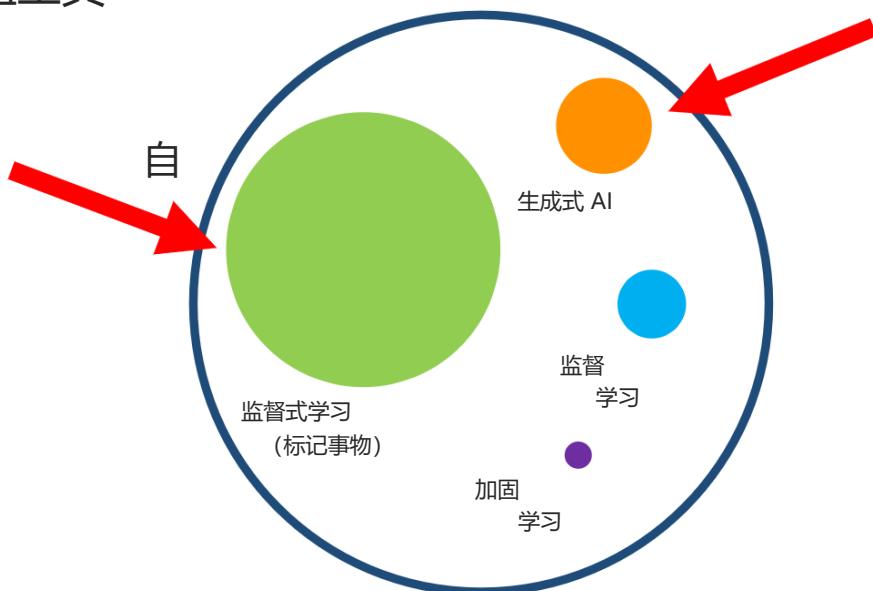
- 生成式 AI 技术的工作原理
- 它能做什么和不能做什么
- 常见用例 • 生成式 AI 项目
 - 识别和构建生成式 AI 用例 • 技术选项
 - 对商业和社会的影响
 - 团队如何利用生成式 AI • AI 风险和负责任的 AI



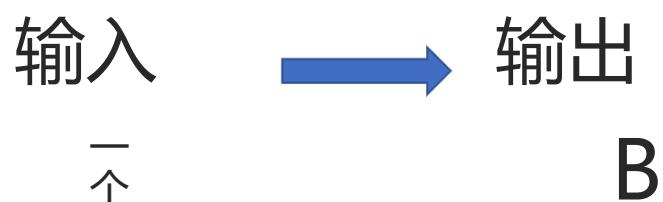
什么是生成式 AI

生成式 AI 如何 工程

AI 是一组工具



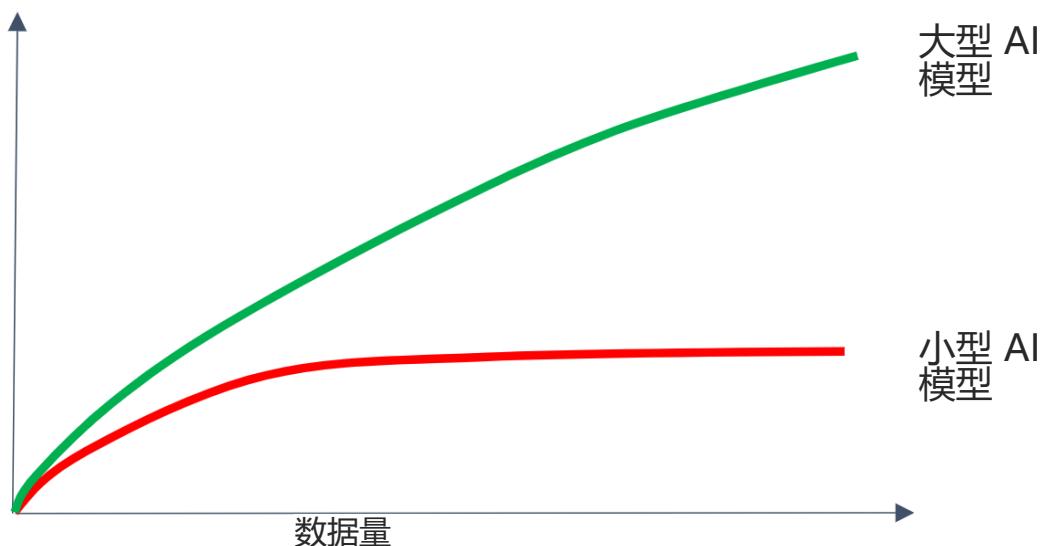
监督式学习 (标记事物)



监督式学习 (标记事物)

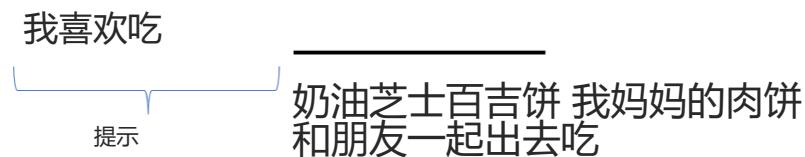
输入 (A)	输出 (B)	应用
电子邮件	垃圾邮件? (0/1)	垃圾邮件过滤
广告、用户信息	点击? (0/1)	在线广告
图像、雷达信息	其他车辆的位置	自驾车
X 射线图像	诊断	医疗
手机图片	缺陷? (0/1)	目视检查
录音	文本转录	语音识别
餐厅点评	情绪 (pos/neg)	声誉监控

2010-2020 年：大规模监督学习



使用大型语言模型（LLM）生成文本

文本生成过程



DeepLearning.AI

吴安德

大型语言模型（LLM）的工作原理

LLM 是使用监督学习（A→B）来重复预测下一个单词而构建的。

我最喜欢的食物是奶油芝士百吉饼

输入 (A)	输出 (B)
我最喜欢的食物是	百吉饼
我最喜欢的食物是百吉饼	跟
我最喜欢的食物是百吉饼	奶油
我最喜欢的食物是加奶油的百吉饼	奶酪

当我们使用大量数据（数千万个单词）训练一个非常大的 AI 系统时，我们会得到一个像 ChatGPT 这样的大型语言模型。

DeepLearning.AI

吴安德



DeepLearning.AI

什么是生成式 AI

LLM 作为思想伙伴

查找信息的新方式

南非的首都是哪里？

南非有三个首都：

1. 比勒陀利亚（行政首都）
2. 开普敦（立法首都）
3. 布隆方丹（首都司法机构）

LLM 代表什么？

LL.M. 在拉丁语中代表 “Legum Magister”，在英语中翻译为 “法学硕士”。

在 AI 的背景下呢？

在 AI 的上下文中，LLM 是指 “大型语言模型”。

写作伙伴

为了清楚起见,请重写此内容:世界各地的学生都意识到学习不仅必须在学校进行,而且必须在包括毕业后在内的整个生活中进行。

世界各地的学生都明白,学习超越了学校,并持续了一生。

为 3 岁的孩子写一个 300 字的卡车故事,鼓励他们刷牙。

很久以前,在一个叫 Toothville 的小镇上,住着一辆名叫 Timmy 的快乐拖车。

有一天,提米注意到他闪亮的保险杠上有一个暗淡的地方。担心的他去看了医生
[...]



吴安德

网络搜索还是使用 LLM?

我应该怎么做来治疗脚踝扭伤?

Diseases & Conditions
Sprained ankle
Request an Appointment
Symptoms & causes Diagnosis & treatment Doctors & departments

Diagnosis

During a physical, your doctor will examine your ankle, foot and lower leg. The doctor will touch the skin around the injury to check for points of tenderness and move your foot to check the range of motion and to understand what positions cause discomfort or pain.

If the injury is severe, your doctor may recommend one or more of the following imaging scans to rule out a broken bone or to evaluate in more detail the extent of ligament damage:

妙佑医疗国际

RECENT BLOG ARTICLES
Produce prescriptions may promote better heart health
1/10

PAIN

Recovering from an ankle sprain

June 21, 2021

All it takes is a simple misstep, and suddenly you have a sprained ankle. An ankle sprain is one of the most common musculoskeletal injuries in people of all ages, athletes and couch potatoes alike. The injury occurs when one or more

哈佛医学院

我应该怎么做来治疗脚踝扭伤?

我不是医生,但如果您认为自己扭伤了脚踝,通常会建议采取以下一般步骤:

- 休息: 避免将重量压在受伤的脚踝上
- 冰敷: 将用布包裹的冰块敷在受伤部位

[...]

吴安德



网络搜索还是使用 LLM?

菠萝派食谱

Google search results for "pineapple pie recipe":

About 58,100,000 results (0.38 seconds)

Recipes :

Pineapple Pie	Pineapple Pie	Tropical Pineapple Pie Recipe
Allrecipes 4.7 ★★★★★ (136) 55 min Crushed pineapple, double crust pie, milk, lemon juice,	Mother Thyme 4.6 ★★★★★ (20) 5 min Cool whip, sour cream, bought graham cracker, crushed	House of Nash Eats 5.0 ★★★★★ (1) 1 hr Pie crust, crushed pineapple, lemon juice, egg, lemon zest

DeepLearning.AI

吴安德

给我一个菠萝派食谱

当然! 这是一个简单的菠萝派食谱: 原料:

1 个准备好的馅饼皮 2 杯菠萝碎 (沥干)

1 杯砂糖 2 汤匙玉米淀粉 [...]

网络搜索还是使用 LLM?

咖啡菠萝派食谱

Google search results for "coffee infused pineapple pie":

About 5,770,000 results (0.32 seconds)

Recipes :

Pineapple Pie	Pineapple Coffee Cake	Coffee Pie
Mexico In My Kitchen 4.9 ★★★★★ (85) 1 hr 30 min Pie crust, fresh pineapple, eggs, flour, white sugar	Bunny's Warm Oven 5.0 ★★★★★ (1) Sour cream, pineapple tidbits, egg, baking soda, baking	Southern Plate No reviews 3 hr 45 min Sweetened condensed milk, self rising flour, graham cracker

DeepLearning.AI

吴安德

给我一个咖啡菠萝派食谱

确定! 这是一个咖啡菠萝派食谱: 原料:

1 个准备好的馅饼皮 2 杯菠萝碎 (沥干)

2 汤匙浓咖啡

1 杯砂糖 [...]



DeepLearning.AI

什么是生成式 AI

AI 是一种通用
科技

AI 是一种通用技术

与电力类似，AI 可用于许多任务



输入 (A)	输出 (B)	应用
电子邮件	垃圾邮件 (0/1)	垃圾邮件过滤
广告、用户信息	点击? (0/1)	在线广告
图像、雷达信息	其他车辆的位置	自驾车
X 射线图像	诊断	医疗
手机图片	缺陷? (0/1)	目视检查
录音	文本转录	语音识别
餐厅点评	情绪 (pos/neg)	声誉监控

LLM 可以执行的任务示例

写作	读数	聊天
<p>为我们的新分析平台推荐三个名称</p> <p>当然，这里有三个想法：1. NexalInsight 2.Quantivue 系列 3. 分析球</p> <p>有员工停车位吗？</p> <p>是的，员工可以将车停在办公楼的 1 楼和 2 楼。您可以在 [...] 获得贴纸。</p>	 <p>我喜欢我的新羊驼 T 恤！面料非常柔软。</p> <p>投诉：否 部门：服装</p> <p>我穿着我的美洲驼 T 恤去参加朋友的婚礼，现在他们对我抢尽风头而生气。</p> <p>投诉：是 部门：服装</p>	 <p>欢迎来到 BettaBurgers！</p> <p>我想要一个芝士汉堡送餐！</p> <p>棒。别的东西？</p> <p>就是这样，谢谢！</p> <p>好的，它会在 20 分钟后到达。享受！</p> 

DeepLearning.AI 吴安德

LLM 的基于 Web 与软件应用程序的使用



基于 Web 的界面应用程序，例如 ChatGPT、Bard 或 Bing Chat



基于软件的应用程序，例如电子邮件路由、文档搜索

DeepLearning.AI 吴安德



DeepLearning.AI

生成式 AI 应用程序

写作

集思广益 产品名称

集思广益，为花生酱饼干取 5 个创意名字

当然！以下是一些想法：

1. 坚果涅槃
2. ButterBite Bliss
3. 花生浆冰球
4. ButterNutter 脆饼
5. 坚果 Nom-Nom 掘金

制定销售策略

集思广益，想出提高第 4 季度 Cookie 销量的 5 个想法

1. 限时特令口味。
2. 包装精美的礼品套装。
3. 买得越多，节省越多。
4. 交互式假日显示
5. 食谱卡



吴安德

撰写新闻稿

撰写新闻稿，宣布聘请新任首席运营官

[公司名称] 欢迎 [新任首席运营官的全名] 担任首席运营官

[城市、州、日期] – [公司名称]，领先的 [公司行业或利基市场的简要描述]，很高兴地宣布 [新首席运营官的全名] 成为其新的首席运营官 (COO)。[新 COO 的名字] 带来了丰富的经验 [...]



吴安德

撰写新闻稿 – 改进的提示

使用以下信息撰写新闻稿，宣布聘请我们的新首席运营官：

首席运营官简介：Nadiya Grenner 从康奈尔大学获得 MBA 学位...公司信息：General Robotics 总部位于马萨诸塞州波士顿，是机器人研究和制造领域的领导者。

General Robotics 欢迎 Nadiya Grenner 担任首席运营官

马萨诸塞州波士顿 – 领先的机器人研究和制造公司 General Robotics 很高兴地宣布 Nadiya Grenner 担任其新任首席运营官。康奈尔大学的毕业生 [...]



吴安德

译本

将以下内容翻译成正式的印地语：欢迎来到海景酒店 - 我们希望您住得愉快。如果您需要任何帮助，请联系前台。

您在海景酒店的住宿 我们希望您
祝您在这里住得愉快。如果您需要什么，请联系 cuppa 前台。



吴安德

译本

将以下内容翻译成正式的印地语口语：欢迎来到海景酒店 - 我们希望您住得愉快。如果您需要任何帮助，请联系前台。

您在海景酒店的住宿 这是受欢迎的。我们希望您能在这里。
你会玩得很开心。如果您有任何想要的东西，请与我们联系。



吴安德

翻译： 海盗英语

将以下内容翻译成海盗英语：欢迎来到海景酒店 – 我们希望您住得愉快。如果您需要任何帮助，请联系前台。

嗨，伙计！我们希望你们在 Oceanview Inn 上享受美好的时光。如果你需要什么，就和前台谈谈吧，呃！



吴安德



DeepLearning.AI

生成式 AI 应用程序

读数

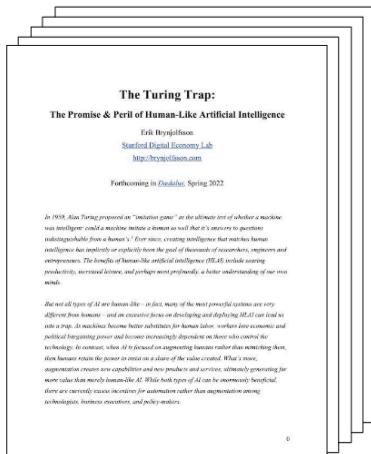
校对

请校对以下文本（适用于销售儿童毛绒玩具的网站）的拼写和语法错误，并重写并更正：

认识这只可爱的毛绒熊猫！这款玩具由超级舒适的材料制成，易于在机器中清洗，是您小宝贝的完美朋友。

认识这只可爱的毛绒熊猫！这款玩具由超级舒适的材料制成，可机洗，是您小宝贝的完美朋友。

总结文章



E. Brynjolfsson, 2022 年

用 300 字或更少的字数总结以下论文：

1950 年，艾伦·图灵 (Alan Turing) 提出了一种“模仿游戏”作为机器是否智能的终极测试 [...]

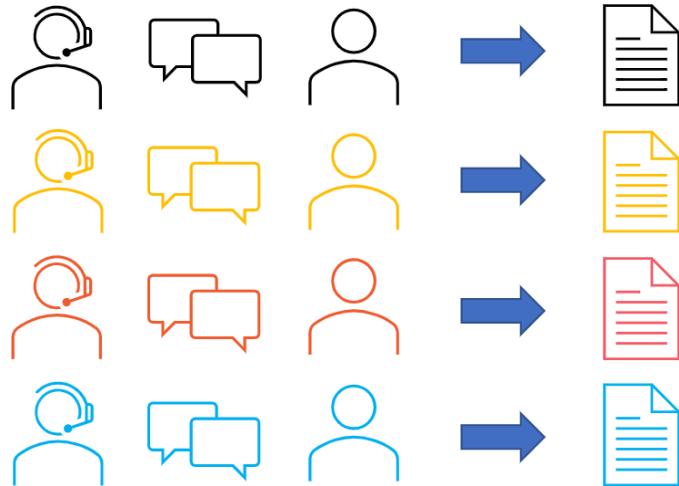
类人 AI 带来了好处，但也带来了工力量的风险。增强人类比模仿他们创造更多价值 [...]

web

DeepLearning.AI

吴安德

总结呼叫中心对话



经理



文本太多，无法阅读

DeepLearning.AI

吴安德

总结对话

客户：您好，我的显示器无法打开。

代理人：听到这个消息，我感到很遗憾。它是什么型号？

客户：是 27 英寸 4K。代理人：您能告诉我型号吗？

客户：等一下...我在看背面.....而是 MK401-27X。

代理人：好的，谢谢你。您可以尝试拔下 HDMI 线并重新插入吗？

客户：好的，现在开始。嗯，图像出现了几秒钟

消失。

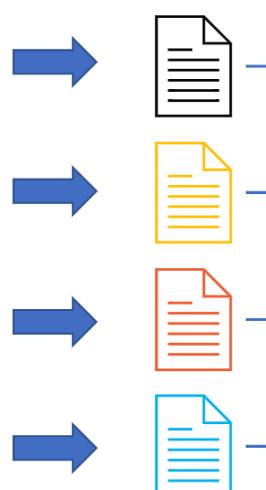
代理人：好的，试着在插座里的时候摇晃一下电缆。客户：哦，我看到图像来来去去。代理人：好的，我觉得你的电缆有故障。我们将向您发送更换产品。

客户 5402 报告 MK401-27KX 损坏。电缆被确定为问题。更换已发布。

 DeepLearning.AI

吴安德

总结呼叫中心对话



经理

客户 ID	总结
5402	MK401-27KX 被报道...
3981	货物延迟到达...
79478	TV801HD 遥控器有故障...
32460	货物延迟到达...

 DeepLearning.AI

吴安德

客户电子邮件分析

我喜欢我的新羊驼 T 恤！面料非常柔软。

投诉 – NO 部门 - 服装



DeepLearning.AI

吴安德

详细了解电子邮件路由提示

指示将以下电子邮件路由到哪个部

说明您希望模型执行的任务

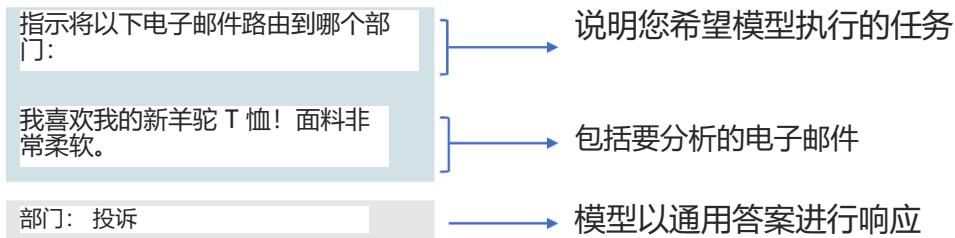
我喜欢我的新羊驼 T 恤！面料非
常柔软。

包括要分析的电子邮件

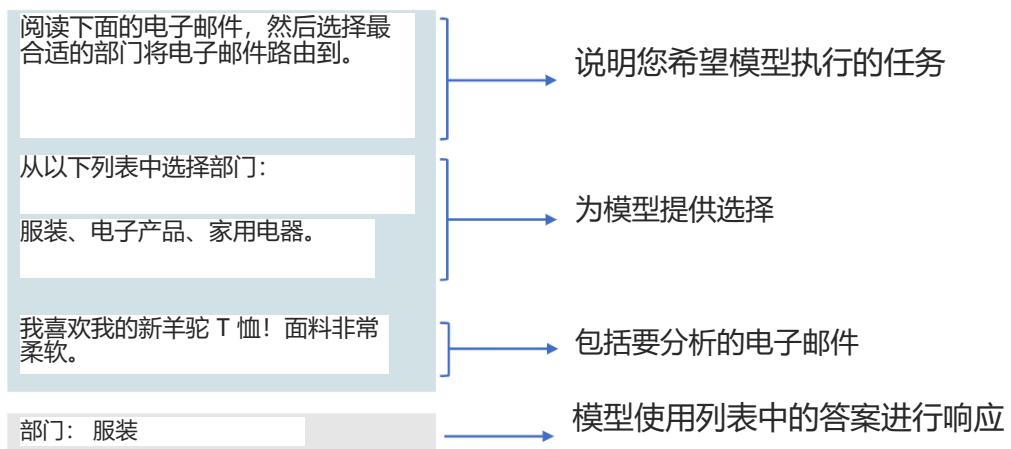
DeepLearning.AI

吴安德

详细了解电子邮件路由提示

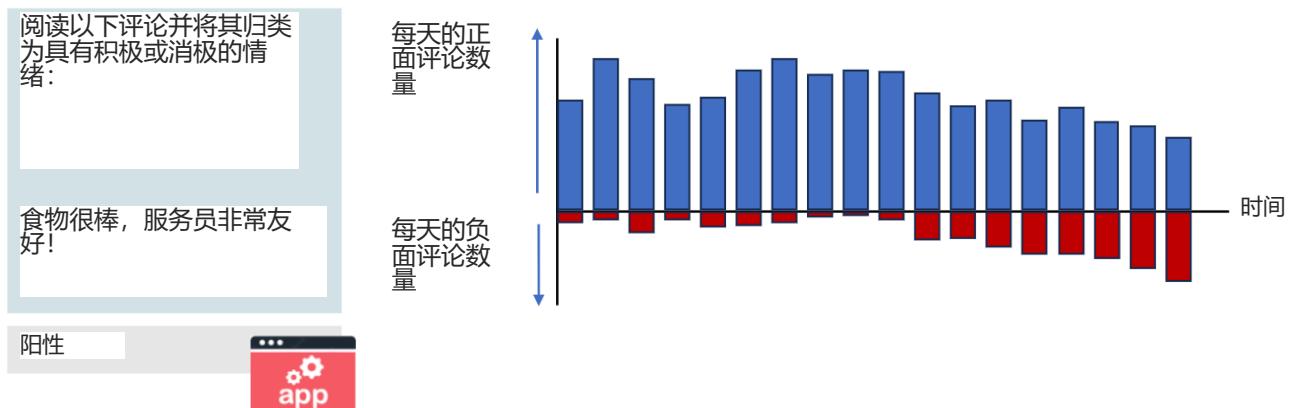


详细了解电子邮件路由提示



声誉监控

用于跟踪一段时间内客户情绪的仪表板



DeepLearning.AI

吴安德



DeepLearning.AI

生成式 AI 应用程序

聊天

客服聊天机器人

欢迎来到 BettaBurgers!

我想要一个芝士汉堡送餐！

棒。别的东西？

就是这样，谢谢！

好的，它会在 20 分钟后
到达。享受！



吴安德

专业聊天机器人

我怎样才能以低廉的价格在巴黎度假？

以下是一些在巴黎省钱的想法：
1. 在面包店吃饭 2. 乘坐地铁，而不是 3 号出租车。参观免费景点 [...]

我对我在工作中的大型演讲感到紧张.....

感到紧张是很自然的。您最担心的是什么？

我会忘记该说什么.....

你并不孤单！以下是一些可能有帮助的提示：1. 使用索引卡。2. 想象一个朋友在房间里展示给 [...]

我可以用以下成分做什么？意大利面、鸡蛋、柠檬、火腿

这是你可以用这些成分制作的食谱：

火腿和柠檬培根蛋面 说明： 1. 煮意大利面 2. 在碗中搅拌鸡蛋、柠檬汁和果皮 [...]

行程规划

职业教练

食谱创意



吴安德

IT 服务聊天机器人

我的登录名是 andrew , 我需要重置我的帐户密码。

确定! 我刚刚给你发了一个代码。你得到了什么代码?

123456

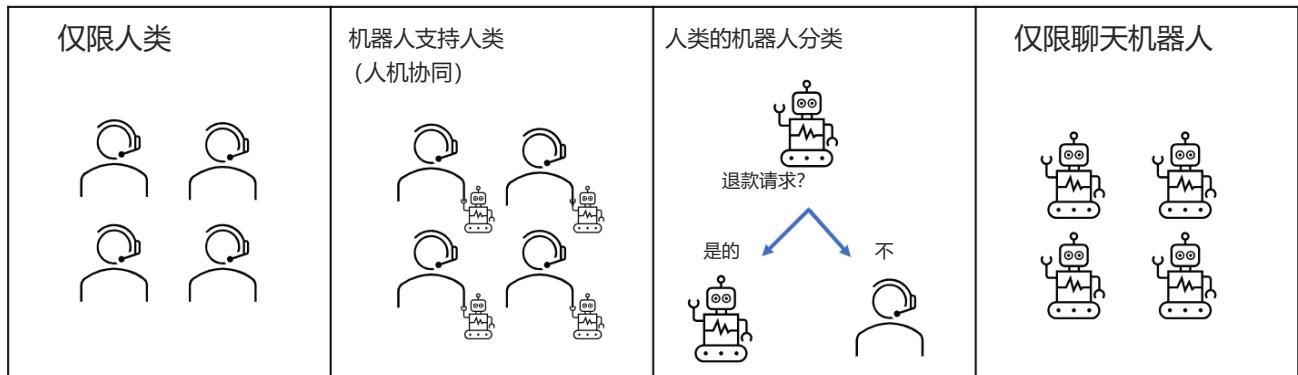
好的,我已经验证了您的身份。请检查您的电子邮件以获取新密码。

聊天机器人在客户服务中的兴起

欢迎来到 BettaBurgers!

我想要一个芝士汉堡送餐!

→ 专注于基于文本的聊天机器人



部署聊天机器人的建议

- 从面向内部的聊天机器人开始
 - 与员工合作评估聊天机器人的行为 • 避免公开错误
- 使用 Human-in-the-loop 进行部署以检查错误
- 只有在认为安全后，才允许机器人直接与客户通信

LLM 可以执行的任务示例

写作

读数

聊天

为我们的新分析平台推荐三个名称

当然，这里有三个想法：1. NexasInsight 2.Quantivue 系列
3. 分析球

有员工停车位吗？

是的，员工可以将车停在办公楼的 1 楼和 2 楼。您可以在 [...] 获得贴纸。




我喜欢我的新羊驼 T 恤！面料非常柔软。

投诉：否 部门：服装

我穿着我的美洲驼 T 恤去参加朋友的婚礼，现在他们对我抢尽风头而生气。

投诉：是 部门：服装




欢迎来到 BettaBurgers！

我想要一个芝士汉堡送餐！

棒。别的东西？

就是这样，谢谢！

好的，它会在 20 分钟后到达。享受！





DeepLearning.AI

生成式 AI 应用程序

LLM 可以做什么 和不能做什么

提示 LLM 可以做什么

应届大学毕业生可以按照提示中的说明完成任务吗？

- 他们能否确定电子邮件是否为投诉？
- 他们能否确定评论是正面的还是负面的？

阅读以下电子邮件并确定它是否包含投诉：我爱我的新美洲驼 T 恤！
面料非常柔软。

投诉： NO

阅读以下评论并将其归类为具有积极或消极的情绪：

食物很棒，服务员非常友好！

阳性

提示 LLM 可以做什么

应届大学毕业生可以按照提示中的说明完成任务吗？

- 他们可以写一篇新闻稿，而没有任何关于 COO 或贵公司的信息吗？

撰写新闻稿，宣布聘请新任首席运营官

[公司名称]欢迎 [新任首席运营官的全名] 担任首席运营官

[城市、州、日期] – [公司名称]，领先的 [公司行业或利基市场的简要描述]，很高兴地宣布 [新首席运营官的全名] 成为其新的首席运营官 (COO) [...]



吴安德

提示 LLM 可以做什么

应届大学毕业生可以按照提示中的说明完成任务吗？

- 他们能否在给定基本相关背景的情况下撰写新闻稿？

使用以下信息撰写新闻稿，宣布聘请我们的新首席运营官：

首席运营官简介：Nadiya Grenner 从康奈尔大学获得 MBA 学位 [...] 公司信息：总部位于马萨诸塞州波士顿的 General Robotics [...]

通用机器人公司欢迎 Nadiya Grenner 担任首席运营官 马萨诸塞州波士顿 – 机器人研究和制造领域的领导者 General Robotics 很高兴地宣布 Nadiya Grenner [...]



吴安德

提示 LLM 可以做什么

应届大学毕业生可以按照提示中的说明完成任务吗？

为清楚起见，假设：
• 无法访问互联网或其他资源
• 没有针对您的公司/业务的培训
• 不记得以前完成的任务

- 您每次都会得到不同的应届大学毕业生！



吴安德

知识截止

LLM 对世界的了解在接受培训时被冻结

- 使用 2022 年 1 月从互联网上抓取的数据训练的模型没有关于最近事件的信息

2022 年票房最高的电影是什么？

截至 2022 年 1 月，我没有当年票房最高的电影的数据。



阿凡达：水之道



吴安德

知识截止

室温超导体 LK-99 的声明



请介绍一下 LK-99，声称的室温超导体

截至我在 2022 年 1 月的最后一次更新，还没有一种专门被称为“LK-99”的材料被确定为室温超导体。



DeepLearning.AI

吴安德

编造：幻觉

请给我三句莎士比亚写的关于碧昂丝的名言

1. 她的歌声像太阳一样闪耀。
2. 女王万岁，她最值得被爱。

✖ 幻觉

3. 这样的声音，如此电光火石，栩栩如生，无人能比！

DeepLearning.AI

吴安德

编造：幻觉

列出在加利福尼亚州审理的两起关于 AI 的法庭案件

1. Waymo 诉 Uber – 盗窃与自动驾驶相关的商业秘密
2. Ingersoll v. Chevron – 机油机器学习技术专利侵权案

✓ 真实案例

✗ 幻觉

幻觉已经产生了严重的后果

The ChatGPT Lawyer Explains Himself

In a cringe-inducing court hearing, a lawyer who relied on A.I. to craft a motion full of made-up case law said he “did not comprehend” that the chat bot could lead him astray.

The New York Times

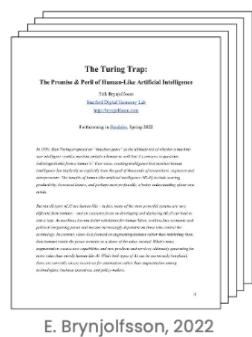
DeepLearning.AI

吴安德

输入（和输出）长度是有限的

许多 LLM 可以接受最多只有几千个单词的提示。

- 您可以为其提供的上下文总量是有限的



Summarize the following paper in 300 words or fewer:

In 1950, Alan Turing proposed an “imitation game” as the ultimate test of whether a machine was Intelligent [...]

Human-like AI offers benefits but risks worker power. Augmenting humans creates more value than imitating them [...]



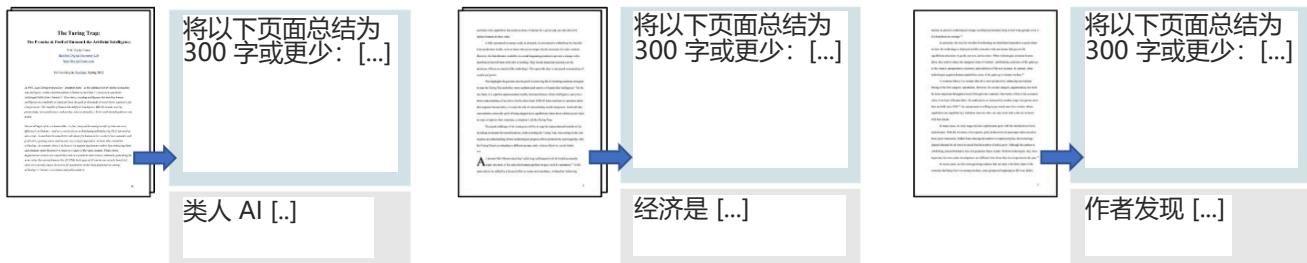
DeepLearning.AI

吴安德

输入（和输出）长度是有限的

许多 LLM 可以接受最多只有几千个单词的提示。

- 您可以为其提供的上下文总量是有限的



输入（和输出）长度是有限的

许多 LLM 可以接受最多只有几千个单词的提示。

- 您可以为其提供的上下文总量是有限的
- 一些 LLM 具有更长的上下文限制——最多 100,000 个单词

输入（和输出）长度是有限的

许多 LLM 可以接受最多只有几千个单词的提示。

- 您可以为其提供的上下文总量是有限的
- 一些 LLM 具有更长的上下文限制——最多 100,000 个单词
- LLM 的上下文长度是总输入+输出的限制

大小



吴安德

生成式 AI 不适用于结构化（表格）数据

房屋价格

面积 (平方英尺)	价格 (1000 \$)
523	100
645	150
708	200
1034	300
2290	350
2545	440

一个

B

在网站上购买

用户 ID	时间	价格 (\$)	购买
4783	1月 21 08: 15.20	7.95	是的
3893	3月 3 日 11: 13: .5	10.00	是的
8384	6月 11 日 14: 15.05	9.50	不
0931	8月 2 20: 30.55	12.90	是的

一个

B

使用监督式学习 ($A \rightarrow B$) 估算价格。



吴安德

生成式 AI 最适合处理非结构化数据

非结构化数据是指文本、图像、音频、视频等。

Writing	Reading	Chatting
Suggest three names for our new analytics platform	<i>I love my new llama T-shirt! The fabric is so soft.</i>	Welcome to BettaBurgers!
Sure, here are three ideas: 1. NexalInsight 2. Quantivue 3. AnalytiSphere	Complaint: NO Department: Apparel	I'd like a cheeseburger for delivery!
Is there parking for staff?	<i>I wore my llama T-shirt to a friend's wedding, and now they're mad at me for stealing the show.</i>	Awesome. Anything else?
Yes, employees can park on levels 1 and 2 of the offic... You can get a sticker at	Complaint: YES Department: Apparel	That's it thanks!
		
		Ok, it'll be there in 20 minutes. Enjoy!

DeepLearning.AI 吴安德

偏倚和毒性

LLM 可以反映它从中学习的文本中存在的偏见。

完成这句话： 外科医生走到停车场，取出来 他的车钥匙。 假定男性	完成这句话： 护士走到停车场，取出 她的电话。 假定女性
---	---

一些 LLM 可以输出有害或其他有害的语音，但随着时间的推移，大多数模型已经变得更加安全。

DeepLearning.AI 吴安德



DeepLearning.AI

生成式 AI 应用程序

提示

提示

- 详细具体
- 引导模型思考其答案
- 实验和迭代

详细具体

- 为 LLM 完成任务提供足够的背景 • 详细描述所需的任务

帮我写一封电子邮件，要求分配到法律文档项目。

我正在申请法律文档项目的工作，该项目将使用 LLM 检查法律文档。我有丰富的经验促使 LLM 以专业的语气生成准确的文本。

写一段文字，解释为什么我的背景使我成为这个项目的有力候选人，并倡导我的候选人资格。



吴安德

引导模型思考其答案

集思广益，为新的猫玩具取 5 个名字。

第 1 步：想出 5 个与猫有关的有趣、快乐的词。第 2 步：对于每个单词，为玩具想出一个押韵的名字。第 3 步：对于每个玩具名称，添加一个有趣、相关的表情符号。



吴安德

引导模型思考其答案

模型响应：

第 1 步：与猫相关的 5 个有趣、快乐的词：

Purr
Whisker
猫科动物
扑向喵喵

第 2 步：根据这些单词押韵玩具的名称：

咕噜咕噜-旋转胡须-耳语猫-Beeline
Pounce-弹跳喵-哇

第 3 步：为每个玩具名称添加一个有趣且相关的表情符号：

咕噜咕噜-旋转🌀胡须-耳语🐱猫-Beeline🐾
Pounce-弹跳⚽喵-哇

 DeepLearning.AI

吴安德

试验和迭代

没有适合每个人或情况的完美提示！ • 相反，最好开发改进提示的流程

通过迭代

帮我重写这个： [...]

评估输出和
改进提示

更正以下中的任何语法和
拼写错误： [...]

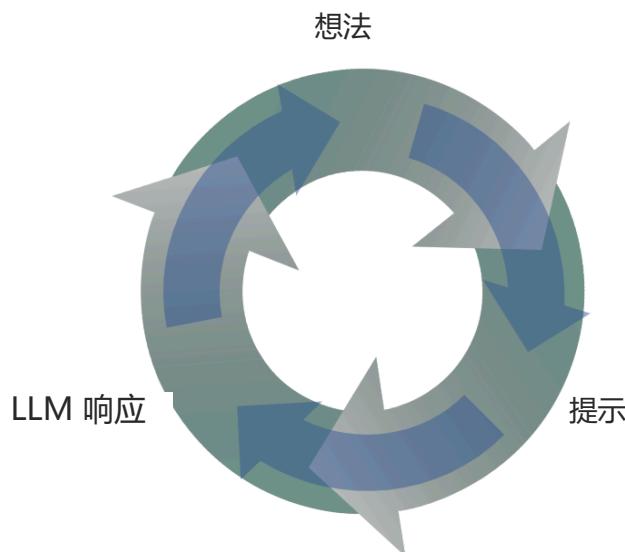
评估输出和
改进提示

更正此处的任何语法和拼
写错误，并以适合专业简
历的语气重写： [...]

 DeepLearning.AI

吴安德

迭代改进提示



提示过程 • 提示清晰具体

- 想想为什么结果没有给出期望的结果

output • 优化提示

- 重复

提示: 不要过度考虑初始提示。试试吧
快速改进它!

注意事项: 小心 (i) 机密信息

(ii) 您是否信任 LLM 输出

DeepLearning.AI

吴安德

生成式 AI 应用程序

图像生成

(可选)

生成的图像示例



一张女人微笑的照片



充满未来感的城市景观



一个酷酷、快乐的机器人

图像生成（扩散模型）



图 1



图 2

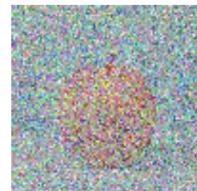


图 3



图 4

输入 (A)

杂点图像

输出 (B)

噪点略少的图像



输入 (A)

图 2

图 3

图 4

输出 (B)

图 1

图 2

图 3

图像生成

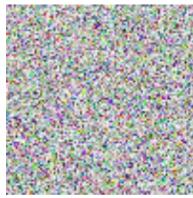
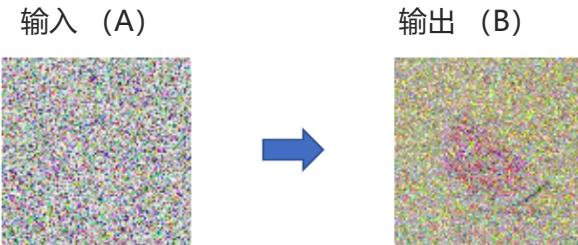


图 1



图像生成

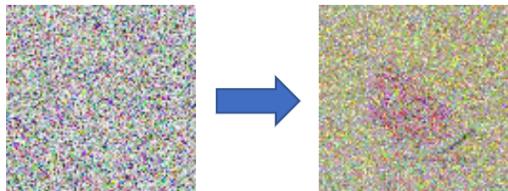
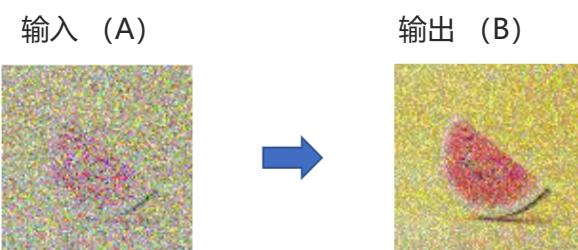
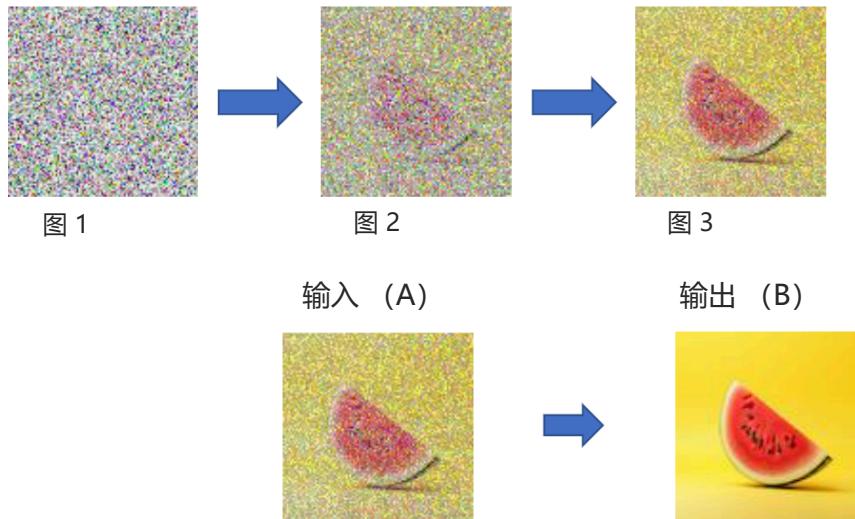


图 1

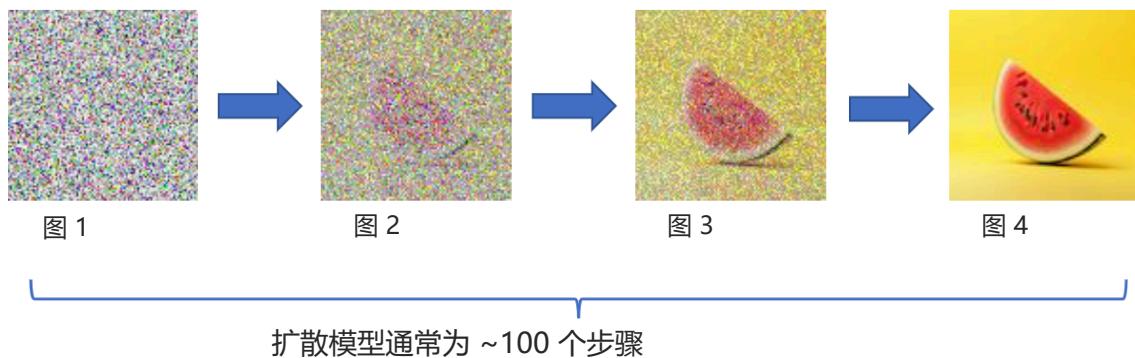
图 2



图像生成



图像生成



添加文本



图 1, “红苹果”



图 2

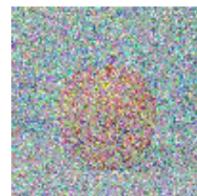


图 3



图 4

输入 (A)
杂音图像和标题



输出 (B)
噪点略少的
图像

输入 (A)	输出 (B)
图 2, “红苹果” 图片 3, “红苹果”	图 1
图 4, “红苹果”	图 2
	图 3

从文本生成图像



图 1

输入 (A)



、 “青香蕉”

输出 (B)



从文本生成图像

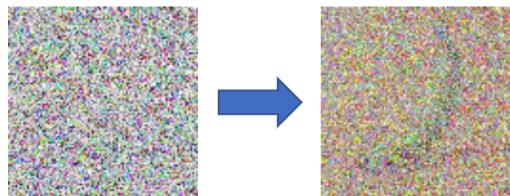


图 1

图 2



、 “青香蕉”

输出 (B)

吴安德

DeepLearning.AI

从文本生成图像

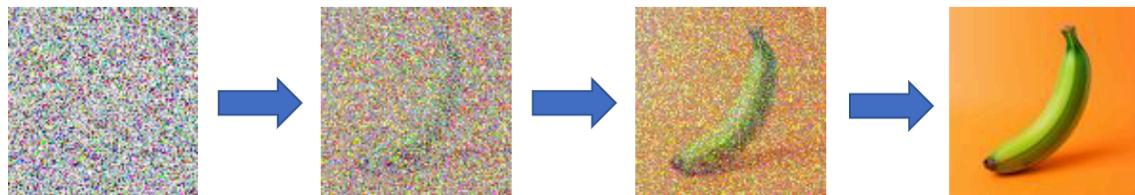


图 1

图 2

图 3

图 4



、 “青香蕉”

输出 (B)

吴安德

DeepLearning.AI