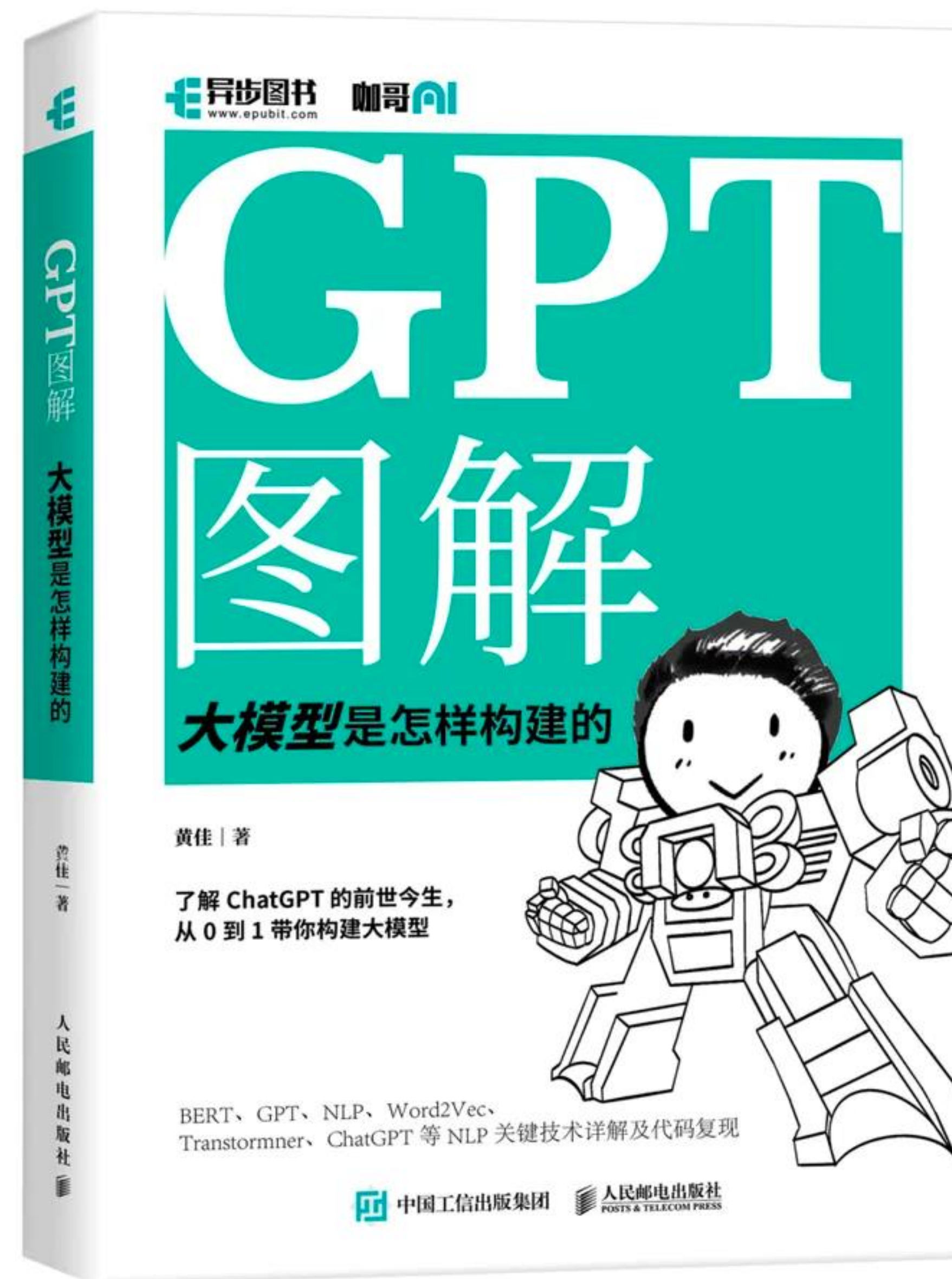


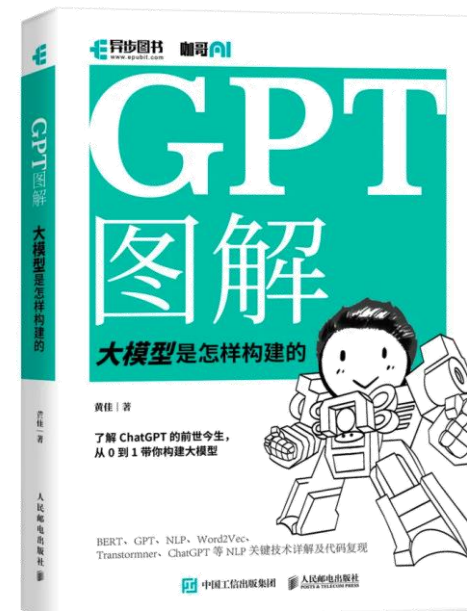
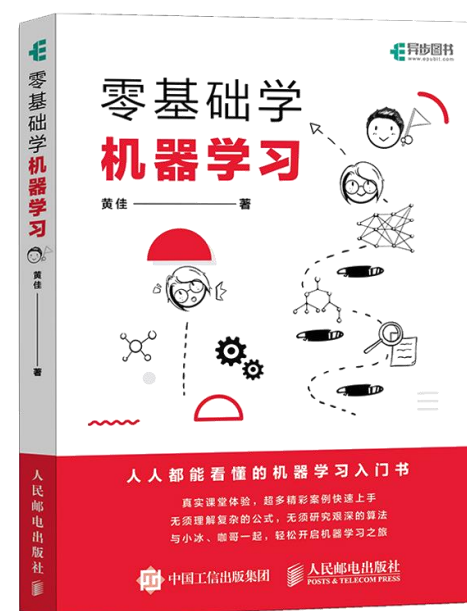
GPT 图解

黄佳 新加坡科技局 AI 研究员





- 新加坡科技研究局 资深研究员
- 极客时间专栏《零基础实战机器学习》作者
- 畅销书《零基础学机器学习》作者
- 畅销书《数据分析咖哥十话》作者
- 前埃森哲新加坡公司资深顾问



前言



在一个遥远的未来世界，科技的力量无处不在。在这个世界里，聪明的人工智能助手们为人类带来了无尽的便利。这个世界里，有两位闻名遐迩的人工智能研究员：咖哥和小冰。他们是一对师徒，携手在人工智能领域研究进步。为了让更多人了解他们的故事，咖哥决定将他们的故事写成一本书。

故事从11月底那场ChatGPT发布会开始。那天，全世界都在关注这场发布会。在那个激动人心的时刻，咖哥感受到了人工智能的力量和希望。然而，他也清楚地看到了人工智能所带来的挑战和风险。

就在这时，一个陨石从天而降，砸在了发布会现场，瞬间引起了一片混乱。人们四处奔逃，尖叫声此起彼伏。而陨石中，竟然有一个神秘的宝盒。咖哥好奇地走上前去，小心翼翼地打开了宝盒。

宝盒里，竟然是一封古老的羊皮卷。咖哥展开羊皮卷，发现上面写着一段预言：只有通过历练和传承，才能将人工智能的力量驾驭得更加完美。咖哥意识到，他的使命就是要传授给小冰和其它NLP小白ChatGPT的奥义！……

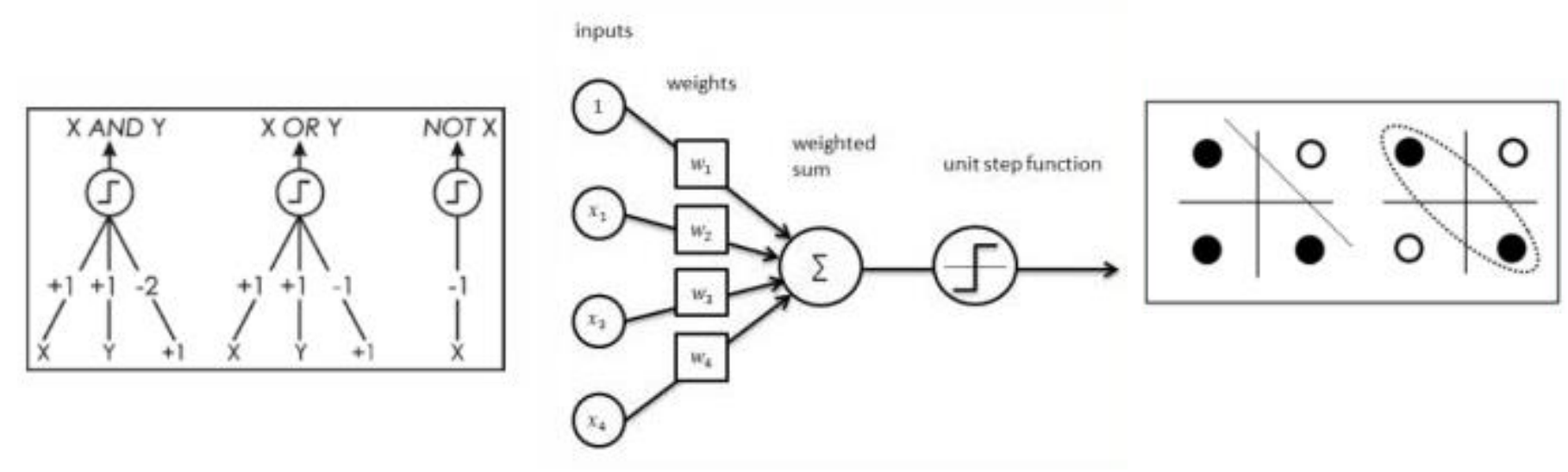
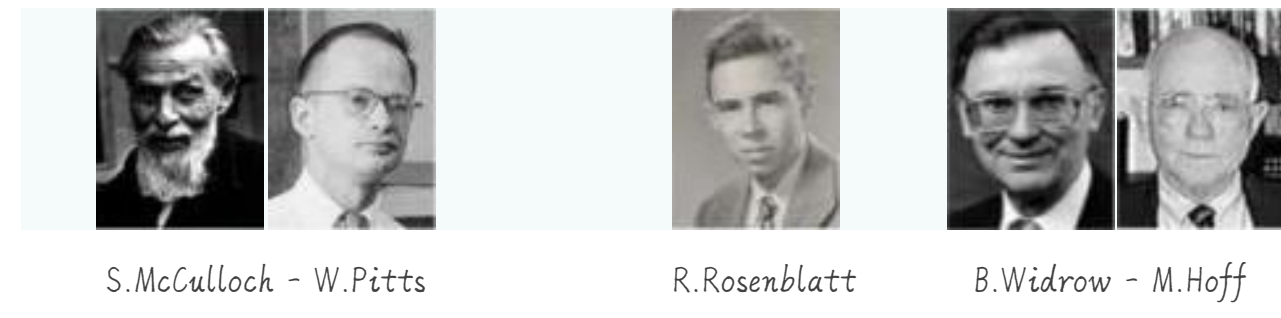
AI技术的发端时期

Thresholded Logic Unit

Perceptron

Adaline

194019501960



神经网络寒冬

AI技术的发展时期

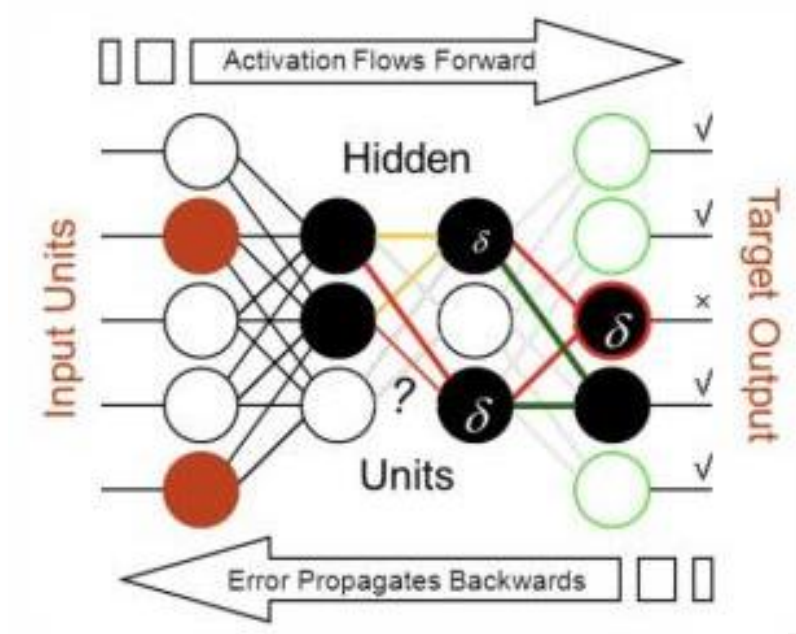
Multilayer Backprop

CNNs

LSTMs

SVMs

198019902000



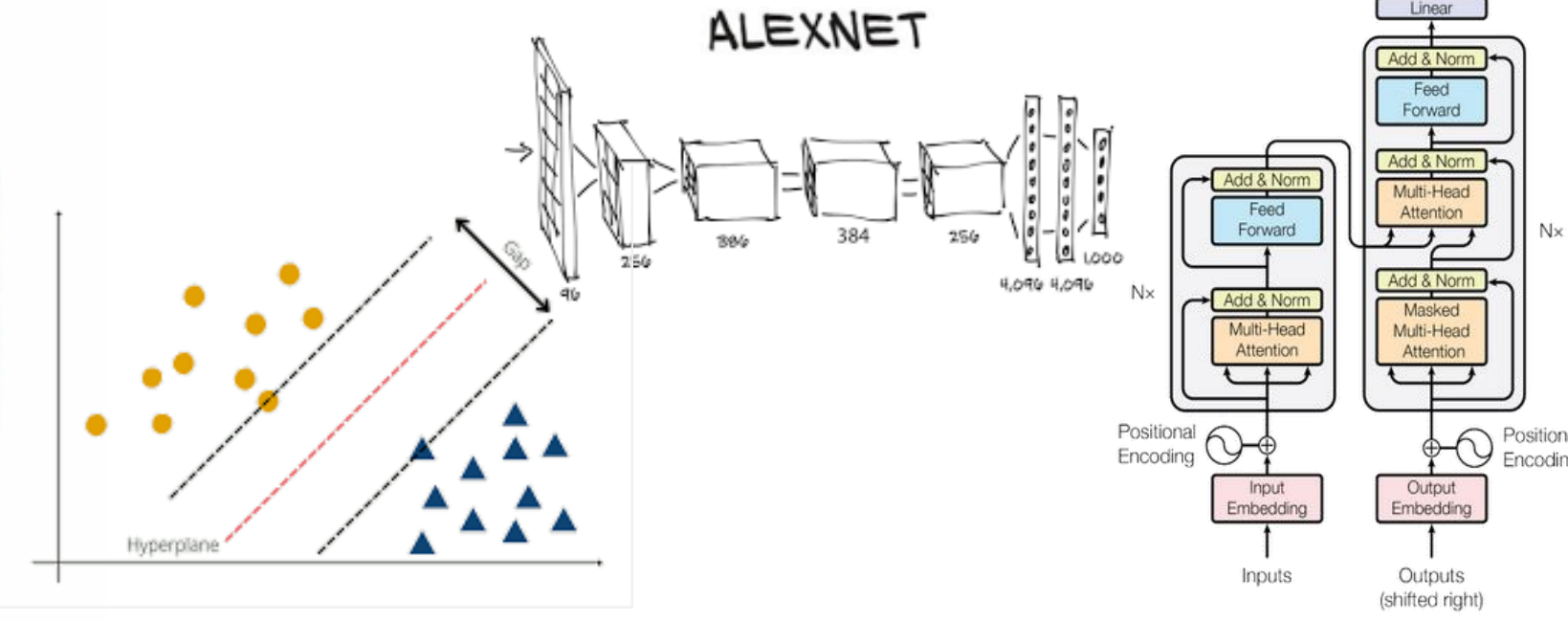
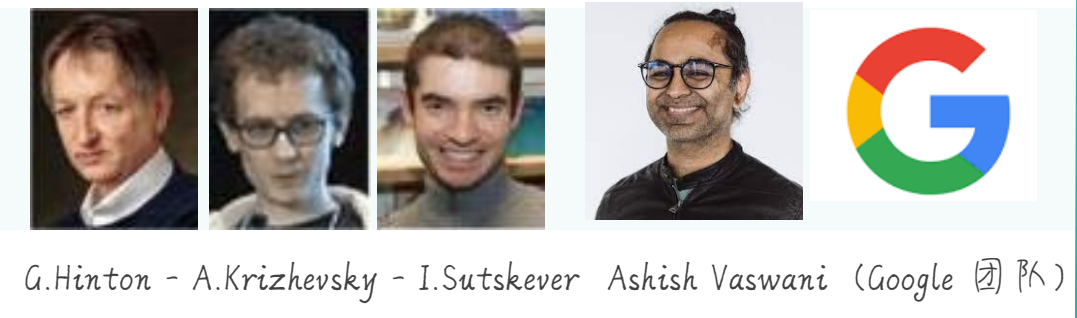
神经网络寒冬

深度学习时代

Alex Net

Transformer

20102020

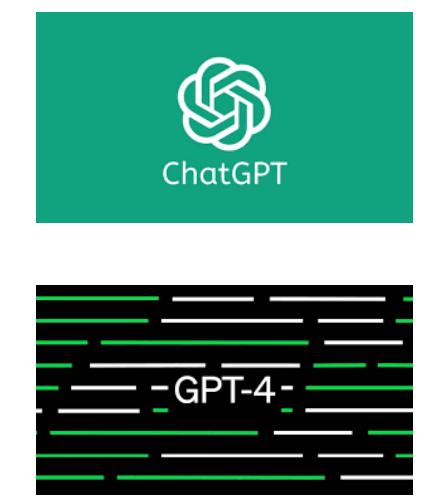


ChatGPT时代

ChatGPT

GPT4

20202023



信息/语言



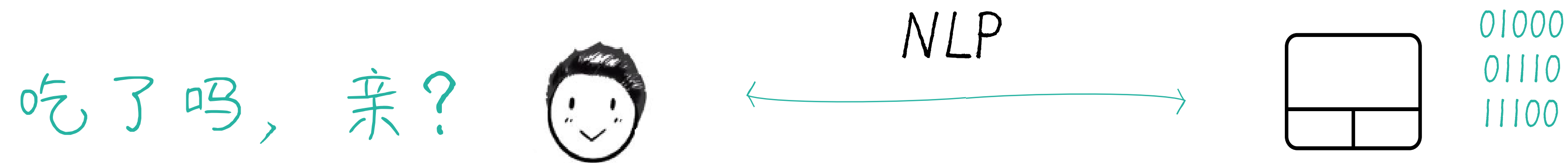
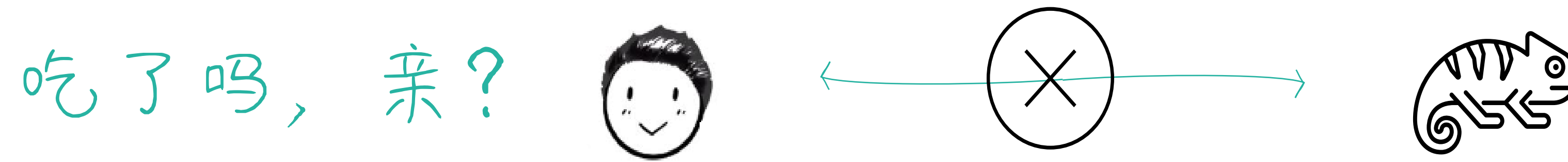
人类社会的核心就是——语言交流。语言不仅是信息的载体，也是人类情感和思想的直接表达方式。它构成了人类社会交往的基础，是文化和知识传承的关键，也承载着我们的情感。ChatGPT作为一个能够理解并使用人类语言的模型，打破了人与机器之间交流的障碍，让人们感受到了前所未有的亲密和便利。具有深远的意义，它不仅改变了人类与机器的互动方式，还可能对教育、娱乐、甚至整个社会的发展产生深刻影响。

语言模型本质上是什么？



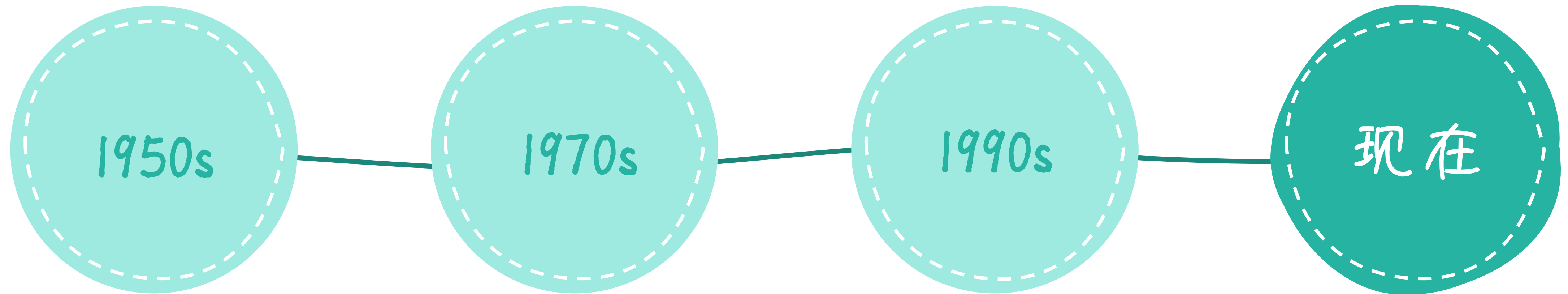
简化的通信模型

语言模型本质上是一种复杂的算法，用于对人类语言进行编码和解码。在编码过程中，它将文本或语音转换为机器可以理解的数据形式；在解码过程中，它再将这些数据转换回人类可以理解的语言形式。通过这种方式，语言模型能够捕捉和模拟人类语言的复杂结构和含义，实现有效的沟通和理解。简而言之，语言模型是信息交流的桥梁，连接了人类的语言世界和计算机的数字世界。



NLP的目标是缩小人类语言和计算机代码之间的差距，使机器能够更自然、更有效地与人类交流，提高信息处理的效率和智能化程度。

语言模型的进化



阿兰·图灵 (Alan Turing) 在论文中描述了会“思考型”机器：能够与人类自然的对话

基于规则的语言处理方法是
由语言学家开发的，用于确定计算机如何处理语言。

随着统计学和数据驱动的自然语言处理方法逐渐成为主流。

深度学习时代到来，开始出现基于深度神经网络的预训练语言模型，ChatGPT等聊天机器人。

起源于
图灵测试

基于规则

基于统计

深度学习
大数据驱动