知识与智能

汇报人: 项一卓 2024.04

CONTENTS 目录

1. 知识与智能的关系

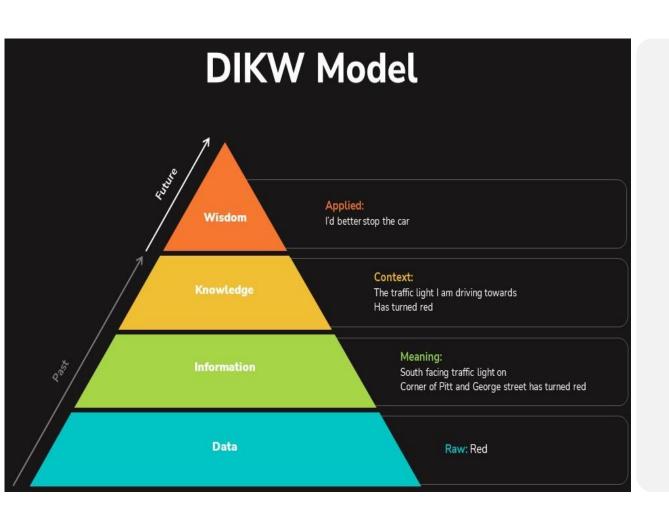
2. 知识在现有人工智能方法中的应用

3. 知识在大模型中的表现形式

4. 知识图谱与知识问答系统

知识和智能的关系

DIKW (Data - Information - Knowledge - Wisdom)



数据 ----> 信息 ----> 知识 ----> 智慧

知识是智能行为的基础

智能是应用知识来解决问题的能力

知识与智能

$\bigcap 1$

知识:

是通过经验、教育和学习获得的信息和理解 提供了理解世界和解决问题的基础 个人发展的支持, 社会进步的驱动

02

智能:

计算 感知 认知

通过计算机模拟,让机器获得和人类相似的智慧,解决智能时代下的实际问题

实例



学习语言

词汇、语法规则、语境……

流利地运用 → 智能



解决数学问题

分析问题 ──知识库中选择解决方法

正确答案

AI模拟人类智能

通过机器学习模型

从大量的数据中学习知识,并将知识应用到解决问题的过程中

如:通过训练数据学习语言模式,并能够生成合乎语法和语境的文本

在人工智能方法中的应用

1、专家系统



使用专家系统解决问题

临床症状——>医学知识库——>提供诊断与治疗建议

通过使用规则库和推理引擎,专家系统能够模拟医生的诊断过程,实现智能的医疗决策



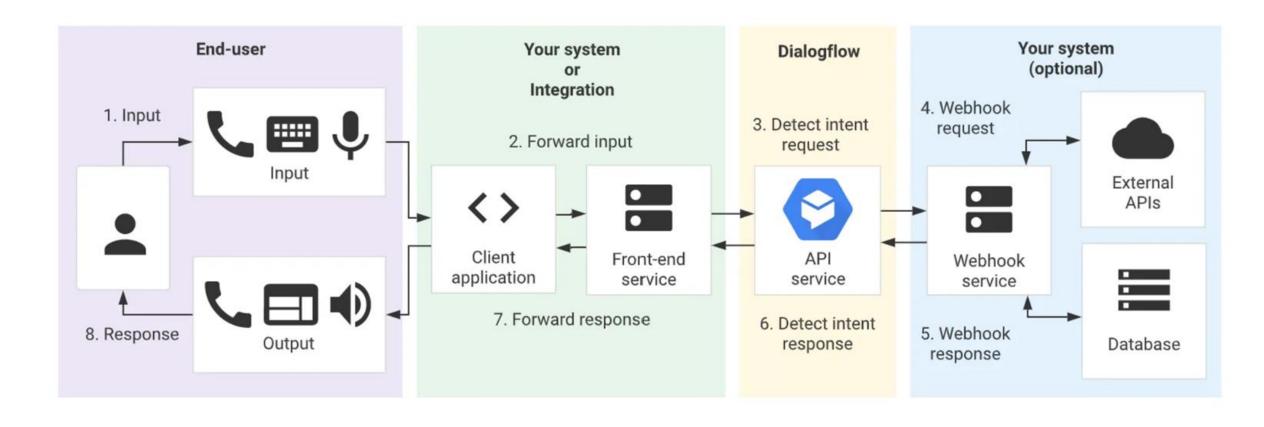
2. 自然语言处理



聊天机器人 (Chatbot)

预先构建知识库 ——> 回答问题

如:虚拟客服、对话式AI



3、图像识别

知识提取

卷积神经网络(CNN)——深度学习

通过不断调整参数来捕获图像中的特征和模式

通过知识提取实现对图像内容的理解和识别



在大模型中的表现形式

1、传统机器学习 vs 深度学习

传统机器学习方法

- ·规则和特征工程来获取知识,
- ·模型训练和预测

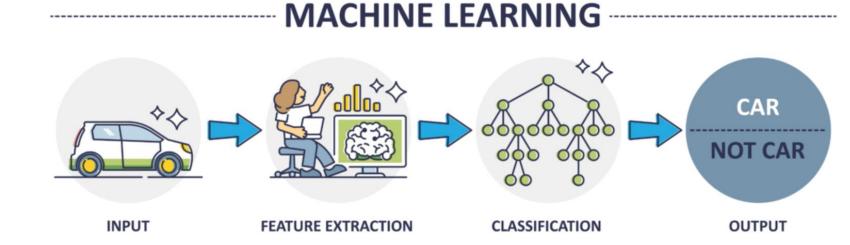
优缺点:

· 优点: 精确度高

· 缺点: 足够多的规则

足够多的数据

颗粒度





1、传统机器学习 vs 深度学习

深度学习

- · 从数据中学习特征和模式
- · 通过大量的数据和反复的试错来 自动学习特征和规律

----- DEEP LEARNING



优缺点:

·优点:学习能力强

适应性好

上限高

·缺点: 计算量大, 硬件需求高, 模型设计复杂



1、传统机器学习 vs 深度学习



2、支持向量机 (Support Vector Machine, SVM)

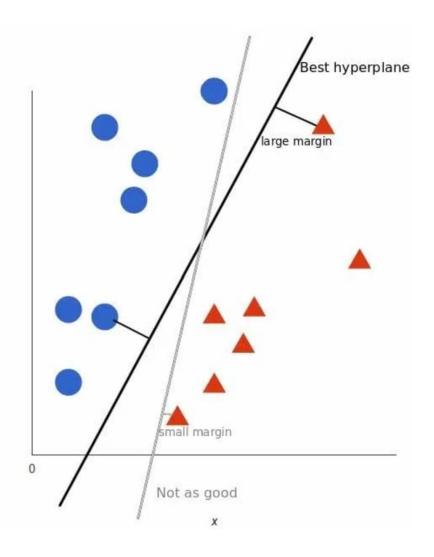
https://www.bilibili.com/video/BV1yo4y1o7A3

https://www.bilibili.com/video/BV1es4y1F72m

监督学习方法

寻找最优的超平面将不同类别的样本分隔

进行预测与分类



2、支持向量机在实际场景中的应用

文本分类

对文档或消息进行分类。 如在新闻网站上,可以使用SVM模型来自动将新闻文章分为"政治"、"体育"、"娱乐"等不同的类别垃圾邮件过滤器可能使用SVM来识别垃圾邮件和正常邮件。

图像识别

SVM

通过使用不同的核函数, SVM能够在高维空间中 找到决策边界, 从而用于图像识别任务, 如手写 数字识别或面部识别中。

生物信息学

识别基因序列模式、药物发现。

在疾病诊断中, SVM可以用于分析基因表达数据, 以识别是否存在特定疾病的风险。

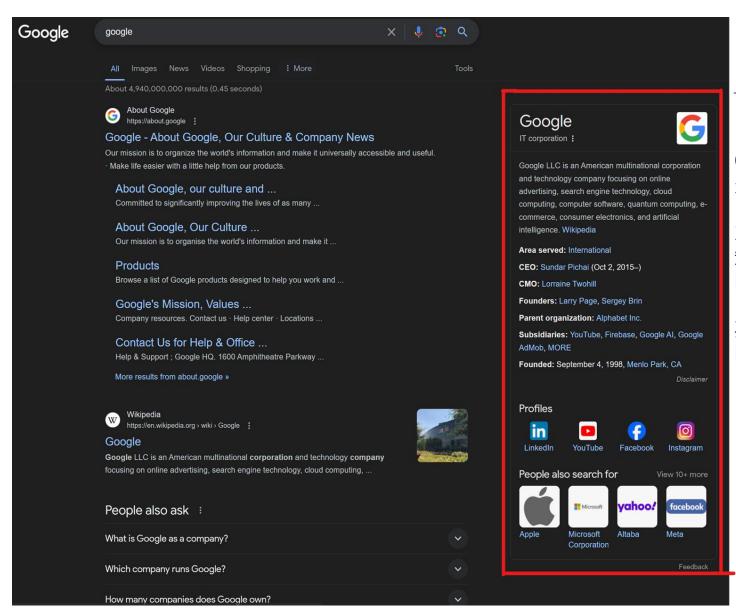
金融预测

预测股票价格的走势或者用于信用评分。

如在信用卡欺诈检测中,SVM可以用于分析消费 者的交易记录,并自动标识出可能的欺诈性交易。

知识图谱与知识表示

1、Google知识图谱(Google Knowledge Graph)



Google知识图谱是Google的一个知识库,其使用语义检索从 多种来源收集信息,以提高Google搜索的质量。

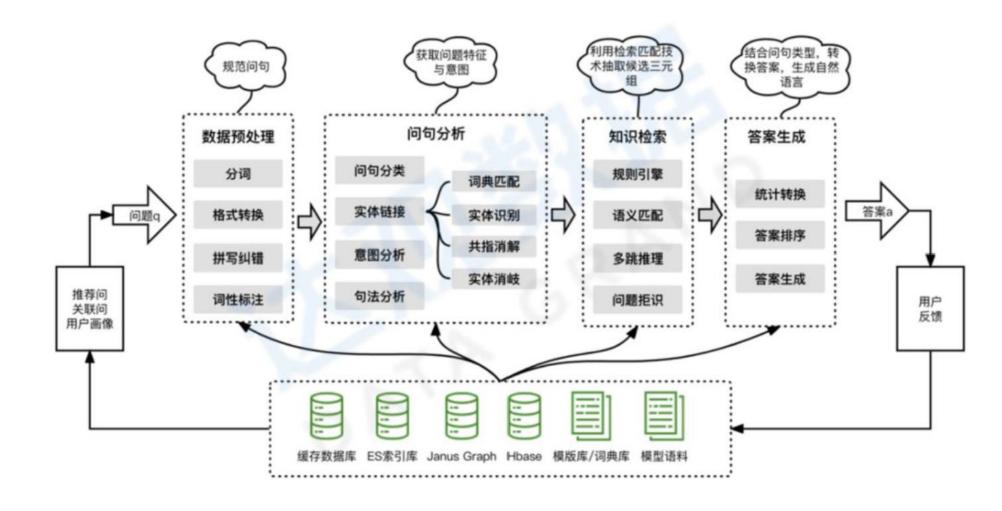
2012年加入Google搜索,2012年5月16日正式发布。Google 知识图谱除了显示其他网站的链接列表,还提供结构化及详细的关于主题的信息。

其目标是,用户将能够使用此功能提供的信息来解决他们查询的问题,而不必导航到其他网站并自己汇总信息。

2、基于知识图谱的知识问答系统

知识图谱也被应用于支持问答系统。

通过理解问题中的实体和它们的关系,知识图谱可以为问答系统提供更加准确和完整的答案。



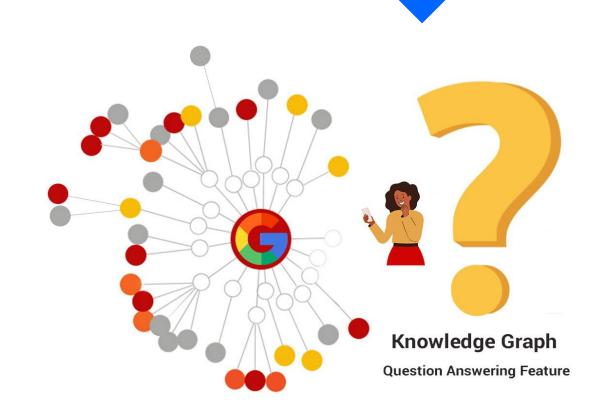
3、三种常见的问答技术路线

- 1、基于知识图谱的问答(KBQA)
- 2、基于阅读理解的问答(MRC)
- 3、常见问题问答(FAQ)

维 度	KBQA	FAQ	MRC
定义	从知识图谱中查询问题需要实体或 关系或属性或加工结果作为答案	从问题答案库中查询与问题最相 似的Q,并返回Q对应的A	根据给定的文本上下文,提取 其中的连续片段作为答案
优点	支持简单或者复杂的推理结果作为 答案	覆盖率高,类似搜索总是可以返 回最可能相关的答案	不需要对文本进行结构化
缺点	图谱构建和维护成本高	答案不一定正确,需要训练STS	需要大量问答对的标注
试用场景	解决领域场景中特定知识库的复杂 和推理问题	解决高频问题	类似搜索,解决长尾问题

4、基于知识图谱的问答系统的业务价值

- ·知识的沉淀和高效复用,知识梳理更加明确直接,实现知识的瘦身和标准化
 - ·返回的答案更能精准理解用户的意图,答案更加直接高效,避免二次推理加工
 - ·对知识管理的维护更加方便,,方便增删改查
 - ·特定业务场景下基于专家经验的复杂推理成为 可能



谢谢大家!

汇报人: 项一卓 2024.04