

大语言模型的实际应用

摘要

大语言模型（LLMs）的快速发展为精算行业带来了重大机遇。本文旨在为精算师提供一条切实可行的路径，帮助精算师们能够高效且负责任地将 LLMs 融入到工作实践中，帮助精算师在复杂的 LLM 领域中找到方向，并提供关于 LLM 的应用、评估、部署及治理方面的指导。本文倡导以审慎且负责任的态度，将这一变革性技术应用于精算行业。

1. 引言

1.1 背景

随着生成式人工智能的快速发展，大语言模型（LLMs）应运而生，并为多个行业带来了创新机遇，其中便包括精算领域。LLMs 已在多个领域产生了深远影响，例如医学、教育、法律，以及与精算师密切相关的金融服务和软件工程。《ActuaryGPT：大语言模型在保险与精算工作的应用》（Balona, 2023）为 LLM 在精算领域应用奠定了基础。然而，LLM 生态系统正在迅速发展：

- 目前，已有超过 20 家公司和研究机构推出了 100 多个 LLMs；
- 评估与排名 LLMs 的方法不断涌现；
- LLM 的使用与部署方式不断演进；
- 其在精算行业的应用场景持续拓展。

Caveat and Disclaimer

The opinions expressed and conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries Research Institute, the Society of Actuaries or its members. The Society of Actuaries Research Institute makes no representation or warranty to the accuracy of the information.

对于没有受过正式生成式 AI 或自然语言处理训练的精算师而言，面对如此庞杂的技术领域可能会感到困惑。而即便是相关领域的研究人员，如果 LLM 不是其核心研究方向，也难以完全跟上最新动态。因此，精算行业迫切需要一本实用指南，以帮助精算师高效且负责任地将这些强大的工具投入实际应用。本文系统性地介绍 LLM 在精算领域的应用，满足了这一需求。

1.2 目标

本文旨在为精算师提供一些实用的指导和建议，使其能够在工作中负责任且高效地应用 LLMs。其主要目标包括：

- 提供当前 LLM 生态的概览（截至本文编写时）；
- 介绍 LLM 的比较与测试方法；
- 建立 LLM 在精算实践中的风险管理与治理框架；
- 提供 LLM 使用的高层指导，并推荐相关深度学习资源；
- 讨论 LLM 的部署策略；
- 探讨在精算应用场景下，开源与闭源 LLM 的优缺点权衡。

本文重点关注 LLM 在精算工作中的实际应用。尽管生成式 AI 还包括图像、音频、视频生成等形式，并可能在保险行业的其他领域产生影响，但这些内容不在本文的讨论范围内。本文依章节展开，将深入探讨 LLM 及其供应商，了解当前的 LLM 生态环境，并逐步深入 LLM 的应用，最后讨论 LLM 相关的风险管理问题。每章开头都会提供一张概览图，以引导读者理解本章核心内容，这些概览图已总结在图 1 中。

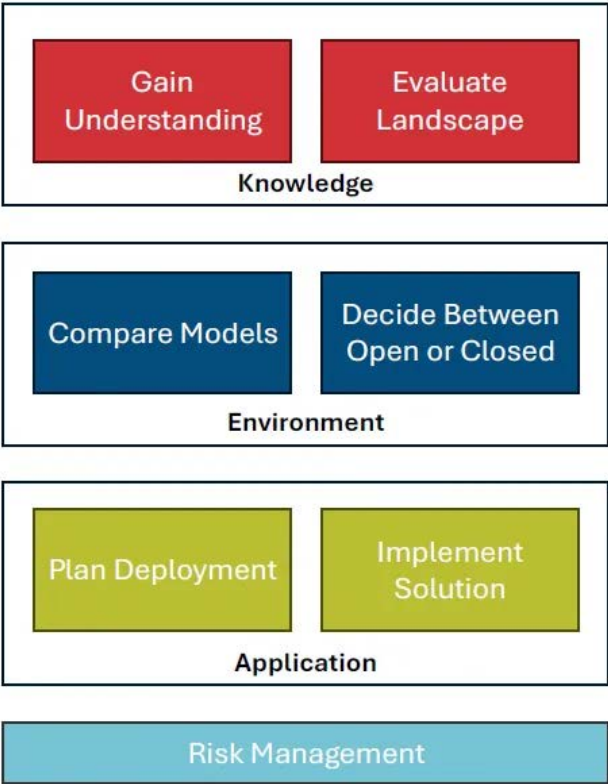
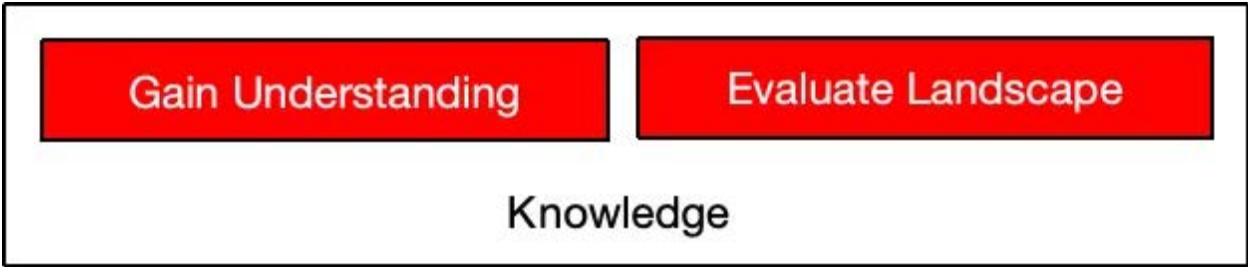


图 1：LLM 精算实践概览图

通过本文，精算师将会扎实掌握将 LLM 融入工作的方法，理解相关风险与收益，并在这一快速发展的领域中保持领先。

1.3 什么是大型语言模型（LLM）？



要理解大型语言模型（Large Language Model, LLM），不妨将其名称拆解为三个部分：大型（Large）、语言（Language）和模型（Model）。为了更清晰地理解其概念，也可以调整顺序，从模型（Model）开始，再到语言（Language），最后是大型（Large）（即 MLL）。

需要注意的是，以下解释将仅保持在概括性层面。LLM 的使用门槛非常低，不需要深入掌握其技术细节即可上手，但对于希望深入了解的读者，本文也会提供进一步的阅读资源。

1.3.1 模型 (Model)：人工智能的核心部分

精算师对数学模型的概念并不陌生。从本质上讲，模型是对现实世界过程的数学表征。LLM 的核心也是某种现实世界过程的数学表征，这一点将在下一节详细讨论这一点。不同于物理学中的直观数学模型，LLM 的底层架构完全是计算机模型。它由一个庞大的神经网络组成，这些神经元按照特定的结构相互连接，使模型能够通过统计方法学习数据的表示形式。这种学习的能力是 LLM 的关键特性。LLM 归类为人工智能 (Artificial Intelligence, AI) 的一种形式。人工智能是一个广义术语，可以有多种解释，但在本文语境下，它指的是让计算机执行特定任务的能力。图 2 展示了 AI 作为最广泛的概念，其中包括了各种用于训练机器完成任务的方法。

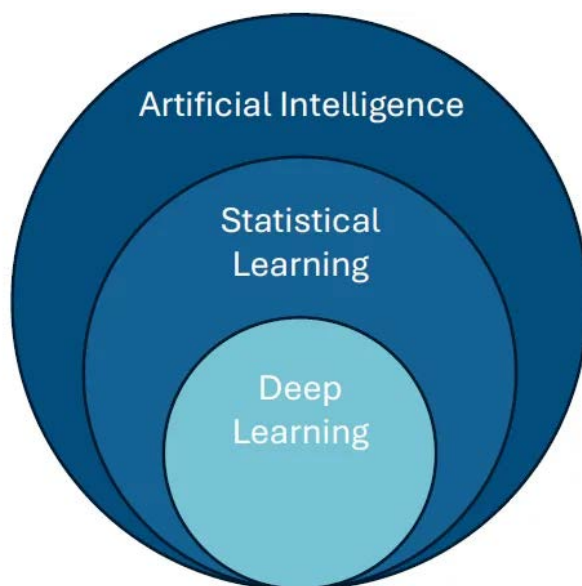


图 2：人工智能 (AI) 作为机器学习任务的最广泛上级集合

通过图 2 可以看到统计学习 (Statistical Learning) 和深度学习 (Deep Learning) 的概念：

- **统计学习 (Statistical Learning)**：又称机器学习 (Machine Learning)，是一个研究领域，该领域专注于通过数据训练模型，使其能够学习数据的结构；
- **深度学习 (Deep Learning)**：LLM 属于深度学习的范畴，深度学习是统计学习的一个子集，它采用神经网络 (Neural Networks) 来进行计算。神经网络本质上是一系列连接的权重 (Weights，也称为神经元) 所组成的网络，这些神经元相互连接以执行计算，深度学习指的是包含多个隐藏层的神经网络。数学上，它们可以被视为从 R_n 到 R_m 的映射函数，图 3 展示了一个单层神经网络的示例，图 4 则展示了一个包含四个隐藏层 (Hidden Layers) 的深度神经网络。

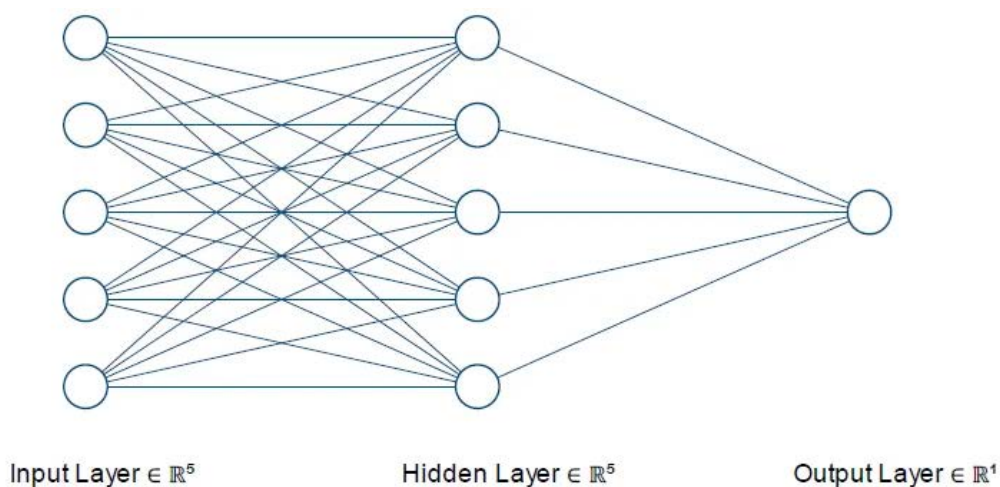


图 3：单层神经网络示例

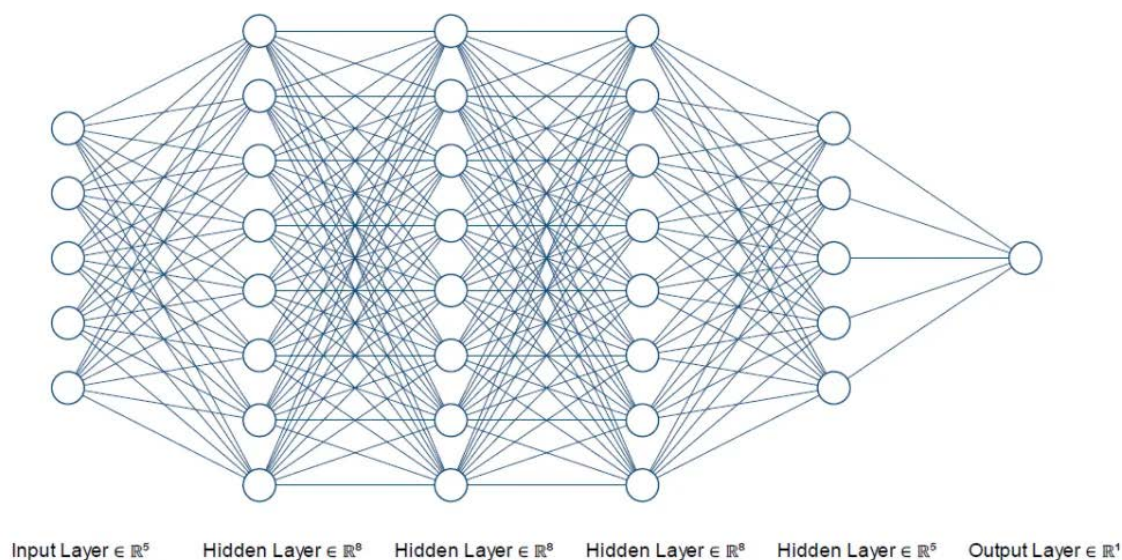


图 4：深度神经网络示例

深度学习的强大之处在于这些隐藏层，也就是模型的架构（Architecture）。这些隐藏层执行表征学习（Representation Learning），通过优化权重选择，使模型针对特定任务构建最佳的数据表示。

1.3.2 语言 (Language)

在 LLM（大型语言模型）的语境中，“语言”指的是模型在理解和生成自然语言方面的能力。自然语言处理（Natural Language Processing, NLP）是人工智能（AI）领域中专注于此任务的一个分支。要理解 NLP，需要掌握两个核心概念：分词（Tokenization）和词嵌入（Word Embedding）。

深度学习模型只能处理数字，但语言大多不是数字。因此，需要一种方法将文本转换为可计算的数值形式。词嵌入（Word Embedding）的作用就是为单词转换向量，使其能够用于建模。更广义地讲，在机器学习领域，嵌入（Embedding）指的是将非数值数据映射到数值向量。在 LLM 中，分词后的离散单位（Tokens）被转换成这些数值向量，以便模型进行计算。

分词（Tokenization）是指将句子和单词拆分成更小的单位（Tokens）。Token 与单词密切相关，但通常会被进一步拆分，以揭示语言的结构。例如，子词分词（Subword Tokenization）可以将单词拆解为有意义的子部分。例如，单词 tokenization 可以拆解成 token 和 ization，因为这两个子词都有助于理解整体含义。不同的 LLM 可能采用不同的分词方式，因此理解分词过程对于掌握 LLM 的技术边界至关重要。此外，分词也是 LLM 定价的基础，模型使用的 token 数量通常决定了计算成本。例如，LLM 的使用费用通常按每 1,000 或 1,000,000 个 token 计价，1,000 个 token 大约对应 750 个英文单词。LLM 通过预测下一个 token 来生成文本，这就需要大量的文本数据进行训练，这也引出了下一个部分：“大型（Large）”。

1.3.3 大型（Large）

LLM 之所以能够达到高水平的表现，是因为它们经过了海量数据训练。底层模型必须能够以可解释的方式学习语言表征，从而可靠地完成任务，这通常需要数十亿甚至数万亿个参数（Parameters）。LLM 训练所使用的文本数据集通常来自互联网，或者是从第三方机构获取的授权数据。这些数据可能包括：

- 网站内容（如博客、新闻文章、研究论文）；
- 代码库（如 GitHub）；
- 社交媒体和论坛；
- 维基百科等。

有时，转录的音频数据也会用于训练。尽管大部分训练数据是公开可用的，但由于数据规模庞大，可能会包含部分私人信息或个人身份信息（PII）。LLM 服务商会清理数据，尽可能移除这些信息，同时过滤掉有害或不适当的内容。这些数据集的规模极为庞大，通常包含万亿级别的单词和 token。在某些情况下，还会额外添加合成数据（Synthetic Data）以增强训练效果。由于 LLM 训练涉及极大的计算量，因此其能源消耗和环境影响也受到一定批评。

LLM 训练流程

在 LLM 完成初始“预训练”（Pre-training）阶段后，模型通常会进行“微调”（Fine-tuning）以适应特定任务。例如：

- 聊天型 LLM 会在对话数据集上进行微调，使其更适合用于对话交互；
- Instruct 模型（指令型模型）会在 指令-响应数据集上微调，使其能够更好地理解并执行用户指令。

此外，LLM 还可以通过“人类反馈强化学习（Reinforcement Learning from Human Feedback, RLHF）”进一步优化。这一过程包括：

- 研究人员向 LLM 提供一组提示（Prompts）；
- 让人类标注者对模型的回答进行评分；
- 用这些评分来调整模型参数，使其生成更符合人类期望的答案。

由于预训练过程过于复杂、耗时且成本极高，对于大多数精算师来说，直接训练 LLM 并不现实。相反，微调已有模型是更可行的方式，并且能带来显著的应用价值。

LLM 的规模

目前 LLM 的规模差异很大，例如：

- Google Gemma 2B 仅有 20 亿参数；
- OpenAI GPT-4 估计拥有 1.76 万亿参数。

在存储需求方面：这些模型从 1.6GB 到超过 7TB 不等。虽然没有严格的标准，但通常认为参数数量超过 10 亿的语言模型即可称为“LLM”。目前，LLM 主要采用“Transformer”架构（但也有其他架构）。

1.3.4 总结：大型语言模型（LLM）

综合以上内容，可以定义 LLM 为一种深度学习模型，拥有大量参数，并通过大规模文本数据训练，以完成特定的语言任务。对于精算行业，LLM 的真正价值在于精算工作中选择并应用它。

About The Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying causes of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its [strategic research programs](#): aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of [topical research available](#), including an expanding collection of international and market-specific research, experience studies, models and timely research.

Society of Actuaries Research Institute
8770 W Bryn Mawr Ave, Suite 1000
Chicago, IL 60631
www.SOA.org