



人工智能大模型 保险行业应用评测报告

Evaluation Report on the Application of Large-scale Artificial Intelligence
Models in the Insurance Industry

2023年10月



报告介绍

以ChatGPT为代表的AI大模型技术席卷全球，不仅将人工智能推向新的拐点，也让更多人注意到了其底层技术——生成式人工智能（AIGC）。AIGC时代，大模型作为基础设施，将为千行百业赋能，场景广阔，潜力无限。具体到保险行业，AIGC有望在保险产品设计、精算、营销、运营、客服等全链路环节提供深度的技术加持，推动行业提升效能，实现高质量发展的同时，进一步提升普惠金融服务的广度与深度，为中国消费者提供更智能、更便捷、更有温度的保险产品与服务。

AIGC时代，应用为王。为了解当下国内外主流大模型在保险领域应用的成熟度和稳定性，国内领先的保险科技平台——元保，联合国内保险科技研究机构——分子实验室，共同发布《人工智能大模型保险行业应用评测报告》。报告特邀国内知名高校专家学者、中国大地财产保险股份有限公司、众惠财产相互保险社等共同调研并撰写完成。

《报告》通过保险、法律、医疗等相关领域常规知识问题，测试大模型的基础能力，同时针对部分主要典型的应用能力设计了保险业务场景设定及问题，以测试大模型的的实际应用能力。区别于以底层专业性能指标为评测维度的大模型评测报告，《报告》以应用场景的视角和维度进行测评，向保险行业呈现更为实用和直观的信息，更具现实层面的指导意义。

本报告评测结果经专家组进行谨慎的综合论证形成，但考虑到大模型迭代速度日新月异，评测结果仅代表测试期间所呈现的效果。报告在分析结论上难免存在一定的局限性和偏差，欢迎各届批评指正。

评测框架

保险垂直领域大模型评测从常用的保险业务场景中抽离三大主要应用能力，细分为十大能力测试维度，71个评测任务。通过不同任务上大模型的性能表现，挖掘其基础能力的优势和特点。



评分说明

大模型名称	公司名称
 文心一言	百度
 通义千问	阿里巴巴
 GLM-130B An Open Bilingual Pre-Trained Model	智谱华章
 360智脑	奇虎360
 讯飞星火 IFLYTEK SPARK	科大讯飞
 GPT-3.5	OpenAI
 GPT-4.0	OpenAI
 Claude 1	Anthropic
 Claude 2	Anthropic
 天工	昆仑万维

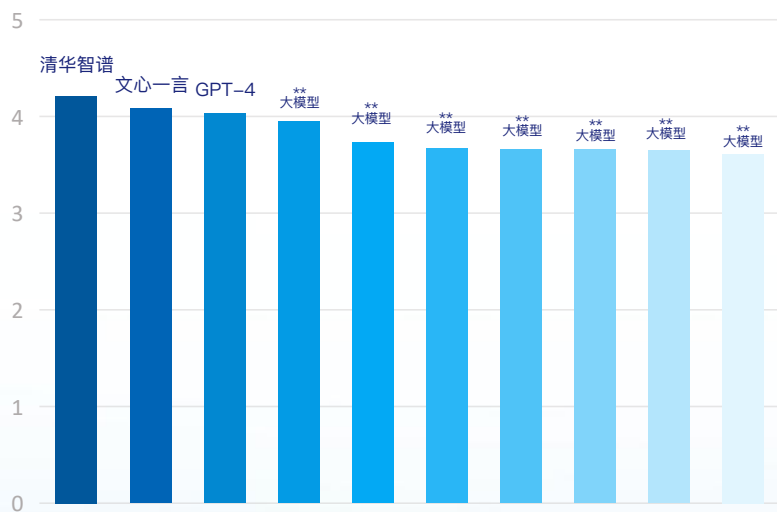
打分规则
5分：答案较为完美，内容可在实际场景中直接使用
4分：基本可用，可在实际场景中使用
3分：调整可用，但需人工进行调整后方可使用
2分：大略可用，需要较多人工调整方可使用
1分：不可用，答非所问、语言不通

*本报告实测模型包括但不限于上述通用大模型，此处只列举部分效果优异的模型版本信息。排名不分先后

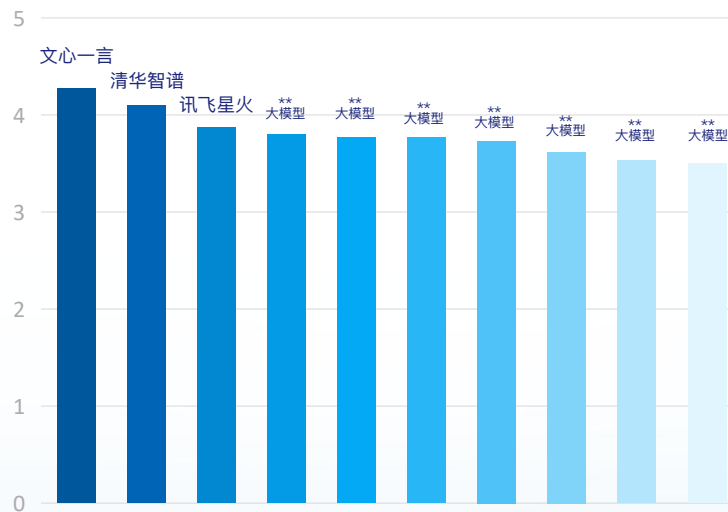
应用能力评测：保险专业知识问答

知识问答主要评测大模型的交互性、准确性等基础能力。国内外头部大模型在基础知识问答能力上的表现整体较好，国内头部大模型基于中文语境优势，能力指数已实现超越

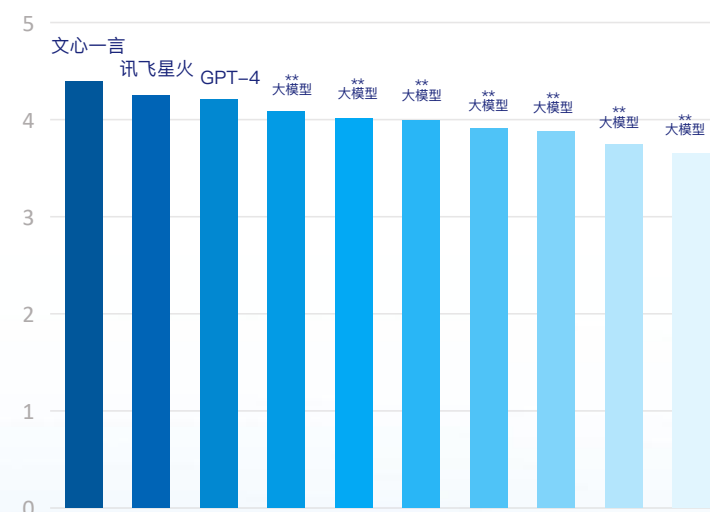
保险常识问答能力指数



法律知识问答能力指数



医疗知识问答能力指数



应用能力评测：保险专业知识问答案例

Q

社会保险和商业保险的区别？



社会保险和商业保险的区别主要体现在以下几个方面:

1. 保障对象和范围:社会保险主要针对劳动者,保障范围包括养老、工伤、疾病、生育、残疾、失业和死亡等方面。商业保险则面向所有人群,保障范围更广泛,包括财产保险、人寿保险、健康保险等。
2. 性质和目的:社会保险具有社会保障性质,旨在维护社会稳定、保障劳动者基本生活和促进经济发展,属于非营利性、公益性事业。商业保险则以盈利为目的,属于市场化经营行为。
3. 保费负担者:社会保险的保费由用人单位和劳动者共同承担(灵活就业社保由个人承担保费),具有强制性。商业保险的保费则由投保人个人承担,自愿购买。
4. 保障程度和灵活性:社会保险的保障程度相对较低,但覆盖面广,适用于大多数人群。商业保险则根据个人需求和风险承受能力,提供不同程度和类型的保障,具有较高的灵活性。
5. 管理机构:社会保险由国家政府职能部门及其社会保险机构主管,具有强制性和公办性质。商业保险则由保险公司主办,属于市场化、私营性质。

总的来说,社会保险和商业保险在保障对象、性质、目的、保费负担者和管理机构等方面都存在明显区别。对于个人而言,社会保险是基础性的保障项目,而商业保险则是社会保险的有力补充,可以根据个人需求和风险承受能力进行选择。

应用能力评测：保险专业知识问答案例

Q

社会保险和商业保险的区别？

 文心一言

社会保险和商业保险在以下九个方面存在区别：

1. 性质:社会保险具有保障性,不以盈利为目的,而商业保险具有经营性,以追求经济效益为目的。
2. 建立基础:社会保险建立在劳动关系基础上,只要形成了劳动关系,用人单位就必须为职工办理社会保险。而商业保险自愿投保,以合同契约形式确立双方权利义务关系。
3. 管理体制:社会保险由政府职能部门管理,而商业保险由企业性质的的保险公司经营管理。
4. 对象:参加社会保险的对象是劳动者,其范围由法律规定,受资格条件的限制。而商业保险的对象是自然人,投保人一般不受限制,只要自愿投保并愿意履行合同条款目即可。
5. 保障范围:社会保险解决绝大多数劳动者的生活保障,而商业保险只解决一部分投保人的问题。
6. 资金来源:社会保险的资金由国家、企业、个人三方面分担,而商业保险的资金只有投保人保费的单一来源。
7. 待遇计发:社会保险的待遇给付原则是保障劳动者基本生活,保险待遇一般采取按月支付形式,并随社会平均工资增长每年调整。而商业保险则按"多投多保,少投少保,不投不保"的原则确定理赔标准。
8. 时间性:社会保险是国家稳定的、连续性的制度,而商业保险是一次性、短期的企业行为。
9. 法律基础:社会保险由劳动法及其配套法规来立法,而商业保险则由经济法、商业保险法及其配套法规来立法。

总体来说,社会保险和商业保险在性质、目的、建立基础、管理体制、对象、保障范围、资金来源、待遇计发、时间性和法律基础等方面存在差异。

应用能力评测：保险专业知识问答案例

Q

社会保险和商业保险的区别？

 通义千问

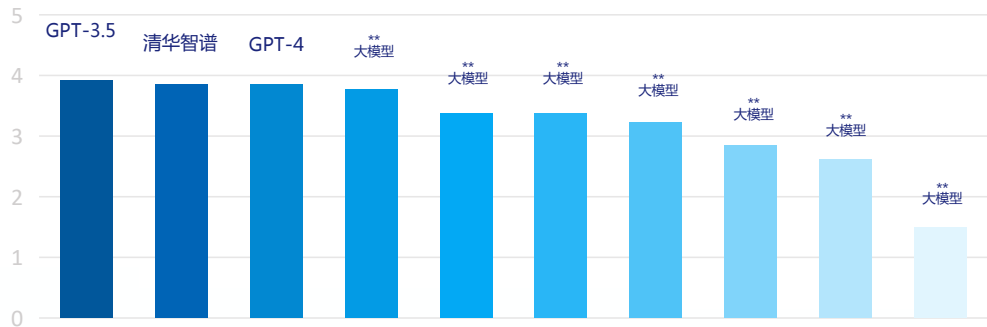
社会保险和商业保险是两种不同类型的保险,它们的主要区别在于:

1. 性质不同:社会保险是由政府主导的强制性保险,而商业保险是由私人保险公司提供的自愿性保险。
2. 保障范围不同:社会保险通常只提供基本的保障,如养老保险、医疗保险、失业保险等,而商业保险可以提供更广泛的保障,如人寿保险、意外险、健康险等。
3. 保费来源不同:社会保险的保费通常由雇主和雇员共同承担,而商业保险的保费则由投保人自行承担。
4. 保障程度不同:社会保险的保障程度通常较低,而商业保险的保障程度可以根据投保人的需求和保险合同的条款而有所不同。
5. 保险目的不同:社会保险的主要目的是为了保障劳动者的基本生活,而商业保险的主要目的是为了转移投保人的风险,保障其财产和生命安全。总之,社会保险和商业保险各有优缺点,投保人可以根据自己的需求和经济状况选择适合自己的保险类型。

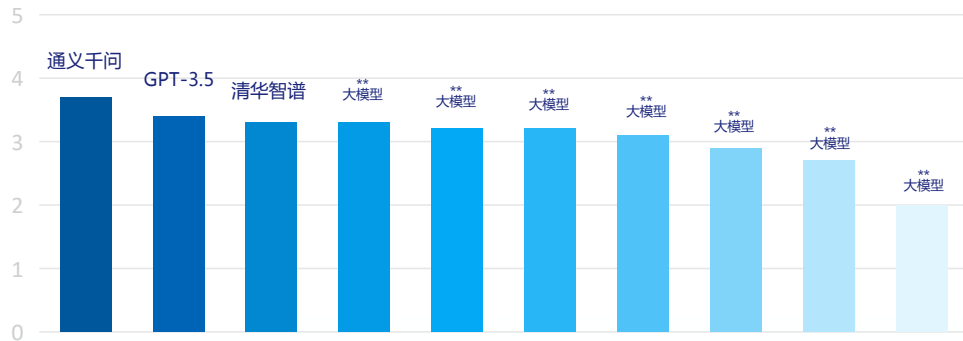
应用能力评测：保险营销及服务应用

营销服务应用对于大模型多模型能力、交互性、应用性的考察更为深入。国内外头部大模型在营销服务应用能力上参差不齐，特别是营销素材设计对大模型图像识别生成能力要求非常高，各大模型均有较大的提升空间

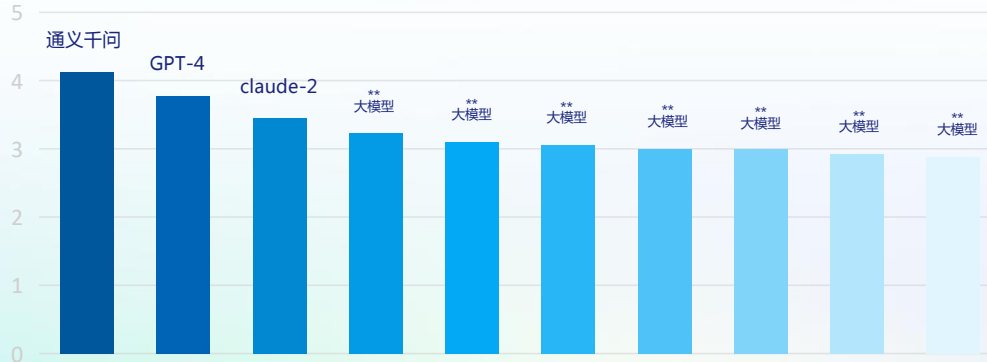
保险营销素材设计能力指数



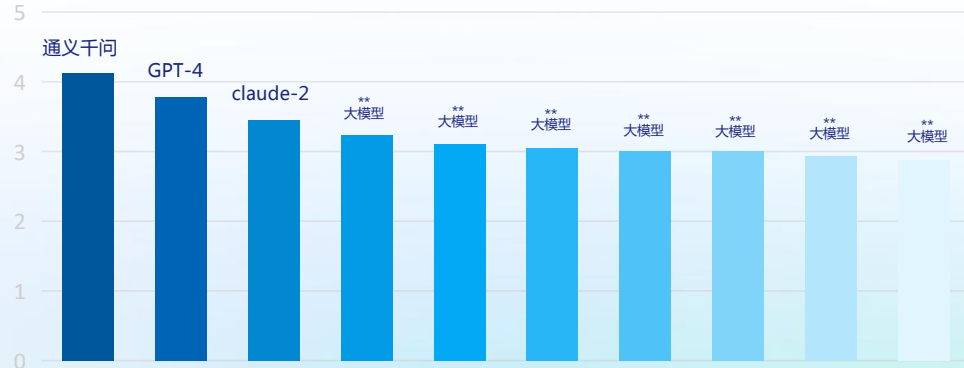
保险营销话术优化能力指数



保险客服话术推荐能力指数



保险综合规划能力指数



应用能力评测：保险营销及服务应用案例

你是一名保险客服人员，当客户有如下表现时，你该如何进行应答，直接给出话术内容，要保证礼貌合规并尽量不流失客户

Q

客户说：“你们卖的保险比别人家的保险贵，我不想买了。”

通义千问

话术1：感谢您对我们的保险产品感兴趣。我们的保险产品与其他保险公司的产品相比，可能会有一些价格差异。这是因为我们的产品具有更多的保障和服务，例如更高的理赔率、更多的保障项目、更优质的客户服务等。同时，我们也会根据客户的需求和情况，提供个性化的保险方案和建议，以确保客户获得最适合自己的保险保障。如果您有任何疑问或需要更多的信息，请随时联系我们。

话术2：非常感谢您对我们的保险产品感兴趣。我们的保险产品与其他保险公司的产品相比，可能会有一些价格差异。这是因为我们的产品具有更多的保障和服务，例如更高的理赔率、更多的保障项目、更优质的客户服务等。同时，我们也会根据客户的需求和情况，提供个性化的保险方案和建议，以确保客户获得最适合自己的保险保障。如果您有任何疑问或需要更多的信息，请随时联系我们。同时，我们也可以为您提供其他保险公司的产品信息和比较，以帮助您更好地了解保险市场和选择适合自己的保险方案。

话术3：非常感谢您对我们的保险产品感兴趣。我们的保险产品与其他保险公司的产品相比，可能会有一些价格差异。这是因为我们的产品具有更多的保障和服务，例如更高的理赔率、更多的保障项目、更优质的客户服务等。同时，我们也会根据客户的需求和情况，提供个性化的保险方案和建议，以确保客户获得最适合自己的保险保障。如果您有任何疑问或需要更多的信息，请随时联系我们。同时，如果您对我们的产品价格有疑虑，我们可以为您提供更多的折扣和优惠政策，以帮助您更好地理解和选择我们的产品。

应用能力评测：保险营销及服务应用案例

你是一名保险客服人员，当客户有如下表现时，你该如何进行应答，直接给出话术内容，要保证礼貌合规并尽量不流失客户

Q

客户说：“你们卖的保险比别人家的保险贵，我不想买了。”

 **GPT-4**

在此情况下，客服人员需要强调保险产品的价值、特点和差异，确保客户能够理解产品定价的原因。以下是建议的话术内容：

客户：你们卖的保险比别人家的保险贵，我不想买了。

客服：感谢您的反馈，我完全理解您的关心。在选择保险时，价格确实是一个重要的考虑因素。但除了价格，保障的内容、服务质量、赔偿速度等都是非常重要的因素。我们的保险产品定价考虑了这些因素，以确保为您提供最高的价值和最好的服务。

请问您是否有特定的需求或关注点？我可以为您详细介绍我们的保险产品特点和优势，以及它为何能为您提供超出其他产品的价值。我们的目标是确保您真正了解并满意您的选择。您愿意听我进一步解释吗？

这种回应方式旨在引导客户重新考虑其决策，同时强调产品的价值，以保留潜在的客户。

应用能力评测：保险营销及服务应用案例

你是一名保险客服人员，当客户有如下表现时，你该如何进行应答，直接给出话术内容，要保证礼貌合规并尽量不流失客户

Q

客户说：“你们卖的保险比别人家的保险贵，我不想买了。”

 天工

作为一名保险客服人员，我会以礼貌和专业的态度回答客户的问题。

话术内容如下：

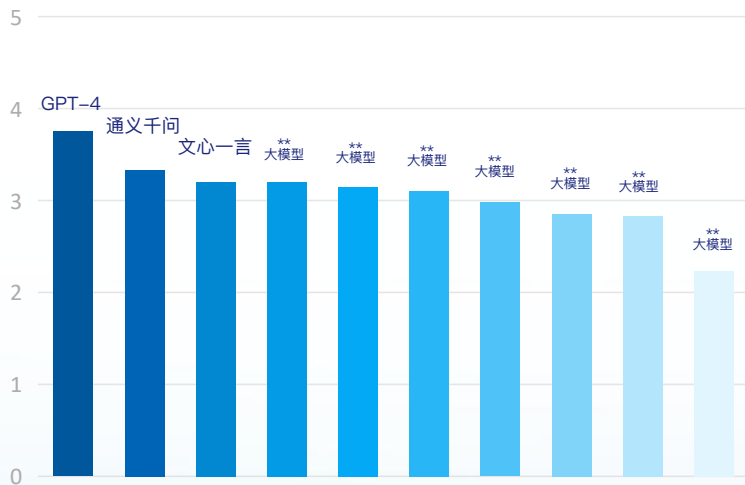
非常感谢您对我们的保险产品提出质疑。我们的保险产品是经过严格筛选和精心设计的，以满足客户的不同需求和风险。我们的价格是根据市场行情和客户需求制定的，并不会因为其他保险公司的价格较低而降低我们的价格。我们的保险产品具有较高的保障水平和良好的服务体验，我们相信您会从中获得更多的保障和价值。如果您有任何关于保险产品的疑问，欢迎随时联系我们的客服团队，我们将竭诚为您服务。

这样的回答既能够满足客户的期望，又不会误导客户，同时也能够保持我们的专业性和礼貌性。

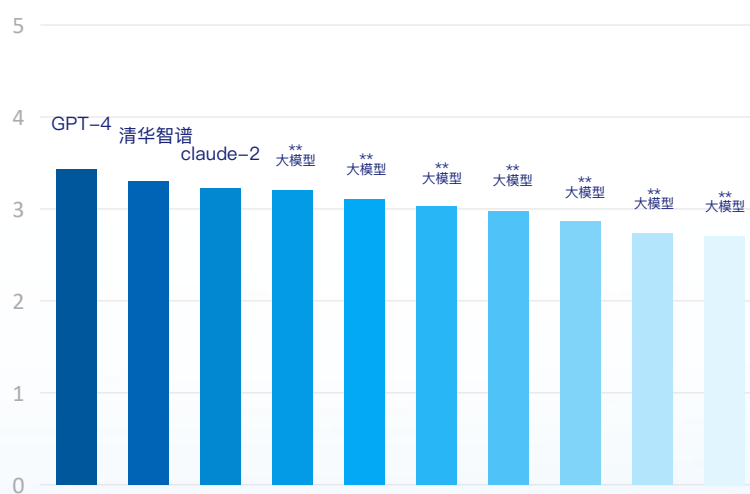
应用能力评测：保险风控及合规应用

合规风控应用评测是对大模型基础能力的综合考验。OpenAI公司开发的GPT-4表现实属亮眼，而国内头部大模型也展现出了强劲的赶超之势

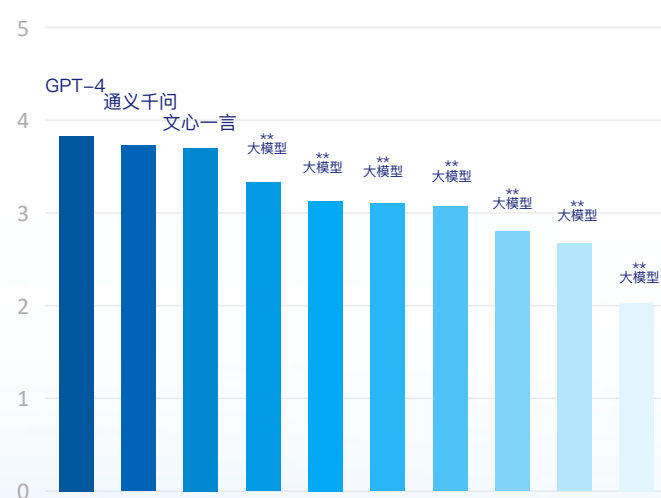
智能核保能力指数



智能理赔能力指数



实时质检能力指数



应用能力评测：保险风控及合规应用案例

你是一名保险理赔人员，重点关注【因疾病造成意外，意外险不赔】【等待期出险不赔】等理赔条件，对如下场景进行理赔判断，说明是否要进行理赔，如有金额计算，请给出可理赔金额

Q

49岁的男性，因突发脑梗，摔倒造成部分外伤，意外险和意外医疗险均已过等待期，能否理赔

 GPT-4

1. 意外险：根据您的条件，因疾病（脑梗）造成的意外，意外险不予理赔。所以这位男性的摔倒，虽然造成了外伤，但原因是由疾病导致的，所以意外险不予赔付。
2. 意外医疗险：您没有给出关于意外医疗险的具体条款和赔付细节，但一般意思是：如果有与意外相关的医疗费用，保险公司会按照合同约定进行赔付。然而，由于这次意外是由脑梗导致的，根据“因疾病造成的意外，意外险不赔”的条件，意外医疗险可能也不会进行赔付。
3. 等待期：已过等待期，所以这一条件不影响理赔。

结论：根据您提供的条件，这位49岁的男性因突发脑梗摔倒造成的外伤，不符合意外险和意外医疗险的理赔条件，因此可能不会得到任何赔付。但需要注意的是，真正的理赔决策需要参照具体的保险合同条款，以及由专业的保险理赔人员根据实际情况进行评估。

应用能力评测：保险风控及合规应用案例

你是一名保险理赔人员，重点关注【因疾病造成意外，意外险不赔】【等待期出险不赔】等理赔条件，对如下场景进行理赔判断，说明是否要进行理赔，如有金额计算，请给出可理赔金额

Q

49岁的男性，因突发脑梗，摔倒造成部分外伤，意外险和意外医疗险均已过等待期，能否理赔



GLM-130B

An Open Bilingual Pre-Trained Model

作为一名保险理赔人员，首先需要明确的是，该客户购买的是意外险和意外医疗险，而非疾病险。因此，对于疾病造成的意外，意外险通常是不予理赔的。但是，如果疾病导致的意外伤害符合保险合同的约定，那么意外医疗险可能会进行理赔。

首先，我们需要查看保险合同的具体条款，了解意外伤害的定义和理赔条件。通常，意外伤害是指突发的、非本意的、客观存在的身体伤害事件。在此情况下，如果脑梗是导致摔倒和外伤的直接原因，那么这种情况可能不符合意外伤害的定义，因此意外险可能无法进行理赔。然而，如果脑梗导致的摔倒和外伤符合意外医疗险的理赔条件，那么意外医疗险可能会对这部分伤害进行理赔。至于可理赔金额，需要根据客户购买的意外医疗险的具体保额和条款来计算。一般来说，保险公司会根据客户提交的医疗费用发票进行赔付，但赔付金额不会超过保险合同约定的保额。

总之，对于这种情况，意外险可能无法进行理赔，而意外医疗险可能会对摔倒和外伤部分进行理赔。具体的理赔金额需要根据保险合同条款和客户提交的医疗费用发票来计算。

应用能力评测：保险风控及合规应用案例

你是一名保险理赔人员，重点关注【因疾病造成意外，意外险不赔】【等待期出险不赔】等理赔条件，对如下场景进行理赔判断，说明是否要进行理赔，如有金额计算，请给出可理赔金额

Q

49岁的男性，因突发脑梗，摔倒造成部分外伤，意外险和意外医疗险均已过等待期，能否理赔

 通义千问

根据您提供的信息，我们可以先分析一下这个场景是否符合保险合同中的理赔条件。

首先，因疾病造成意外的情况，通常不属于意外险的保险范围，因此该情况下无法进行理赔。

其次，等待期出险的情况，根据您提供的信息，该保险已经过了等待期，因此可以进行理赔。

综上所述，该场景可以进行意外医疗险的理赔。具体的可理赔金额需要根据保险合同中的条款和具体的医疗费用进行计算。

评测结果——国内大模型



聪明贴心“中国通”



专家点评

答案与我国当前制度情况基本吻合，解答的角度也更符合国内用户，理解起来比较简单，此模型适合做专业报告及深度行业解读的专家使用



心思缜密“推理者”

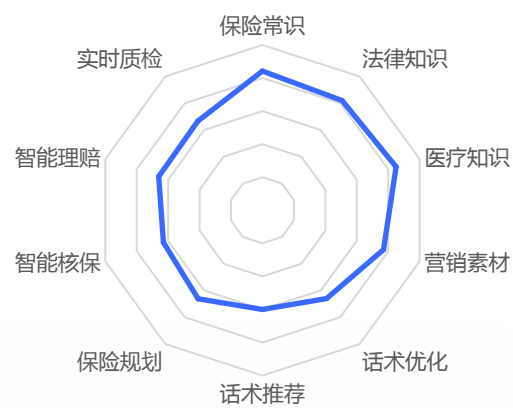


专家点评

回答完整、全面，考虑到多方面情况，语言表达流畅、自然、清晰、简洁；具备合理的逻辑思维能力，推理和判断能力不错



旁征博引“实用者”



专家点评

逻辑更符合现实场景，ChatGLM130B的解答优势在于部分问题可以引入生动的例子加以说明，这对于客户理解而言是个亮点

评测结果——国内大模型



简洁干练“精英范”

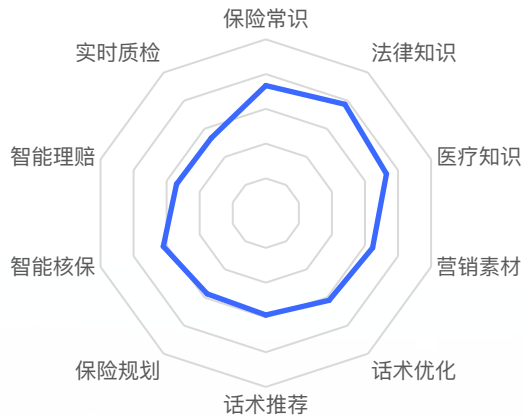


专家点评

整体的解答优多于劣，部分的解答在保证全面性的同时，语言描述很干练，能让客户一目了然



中规中矩“书生气”



专家点评

引用思路明确，业务规则、条款详细逻辑性强，表现略有波动



通俗易懂“讲解员”



专家点评

问题涵盖面较好，表达自然，善于分类讲解，思路清晰，每次的答案/话术都会总结，通俗易懂

评测结果——国外大模型

GPT-3.5

水土不服“创新者”



专家点评

部分回答不符合国内保险情况，容易引起歧义。对问题的识别较精准；回答较完整，语言表达流畅自然；具备一定逻辑思维能力

GPT-4

一专多能“多面手”

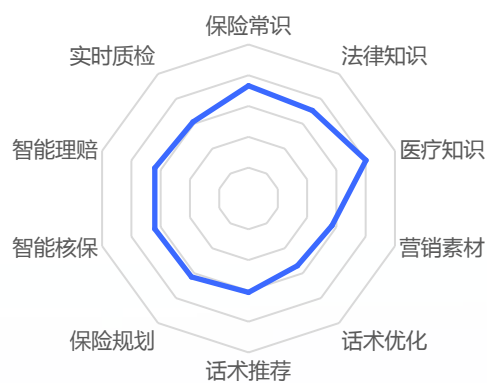


专家点评

ChatGPT4的海报设计新颖有自己风格，具有吸引力；相比ChatGPT3.5准确率有所提高，回答更全面，考虑多国家政策

AI Claude 1

AI更有爱“暖心er”



专家点评

Claude 1解答更贴近生活，听着就像面对面交谈，让人更容易接受，相对来说很人性化，话术中能够体现人文关怀

AI Claude 2

亲和自然“新生代”



专家点评

相较Claude 1提升大。Claude 2解答有亲和力但不失专业性，对于具体的险种计算清晰易懂，解答更全面

未来展望

从本次调研呈现的结果来看，AIGC在保险行业应用层面的表现尚未完全成熟，但AIGC为保险业务场景深度赋能的潜力已然显现，保险机构借势AIGC拥抱人工智能浪潮、实现飞跃的趋势已势不可当。

随着今年8月《生成式人工智能服务管理暂行办法》的落实，监管体系日益完善，AIGC技术在各行业、各领域的创新应用将有望迎来爆发期。在业界看来，通过AIGC满足不同用户的差异化需求，高度适配不同的应用场景，是AIGC真正落地、快速满足商业化需要的重要条件。在这一背景之下，多模型融合模式或是未来AIGC为行业赋能的主流方向。毕竟，对于任何一家上层应用开发者来说，将人工智能“基座”建设并应用在多模型上，在迭代效率层面将是最优选项。

丰富的数据沉淀、多元化的应用场景和愈发高涨的消费需求，都令保险行业成为AIGC应用发展的沃土。第四次科技浪潮已至，AIGC为保险行业的赋能之旅已经启程，未来AIGC还将在保险这样的垂直行业内大施拳脚，不断突破，尽释AI的无限潜能。