**“数学与人工智能应用”企业联合研究项目**

**资助课题选题指南**

课题1. 文档知识的高效机器检索与数学

[背景] 在当今的技术环境中，通用大模型正在各个领域中展现出其强大的潜力。然而，当我们将通用模型进行商业化应用时，需要对其进行领域化，以便为精准聚焦的业务场景带来实际的价值。为了进行有效的领域化，我们需要对领域性甚至专业性的文档进行高效准确的理解。业界已经在这方面应用中取得初步成效，然而这个过程仍然充满挑战，需要借助于数学理论和方法攻克难题。本方向课题旨在研究和开发一种更智能的文档理解与问答方法，能够快速理解含有复杂逻辑和抽象概念的多模态文档（文档中包含文本、图片、表格等），并能在此基础上基于文档知识准确回答专业群体的问题。

[问题]（可以从以下问题中选择一个或几个进行研究）：

1. 对于多模态文档建立高效的向量化表征算法，使得文本、图片、表格等多形态数据可以在统一的向量空间中进行表征。

2. 构建高效的向量相似度及匹配算法，可针对大规模文档快速精准地进行向量搜索。

课题2. 生物医药与数学

[背景] 在药物发现和设计过程中，分子动力学可以提供有关药物分子和生物分子之间相互作用的详细信息。它可以模拟药物分子在生物体内的行为，如药物的结合方式、药物与蛋白质的相互作用以及药物的稳定性等。

分子动力学模拟可以提供许多有价值的信息。它可以揭示药物分子与蛋白质之间的结合方式，从而帮助研究人员了解药物的作用机制。此外，它还可以预测药物分子的构象变化、溶解度、药物释放速率等性质，有助于药物分子的优化设计。

分子动力学模拟基于牛顿力学原理，通过数值方法来求解分子系统的运动方程。它将分子系统建模为一组粒子（通常是原子或分子），并使用经典力场来描述粒子之间的相互作用。通过迭代求解粒子的运动方程，可以模拟出分子系统在一定时间内的运动轨迹。通常，在药物发现领域涉及到的分子系统具有变量个数多，变量之间存在两两交互作用等特点，利用经典方法直接进行数值求解效率较低。

[问题]：

1. 针对药物发现和设计过程中涉及到的分子动力学方程，研究利用RBM等新方法进一步提升计算效率，并结合具体场景，利用新方法得到药物发现和设计过程中1-2个新结果。

课题3. 稀土材料与数学

[背景] 稀土材料是由稀土元素制成的材料，它们在许多高科技应用中起着至关重要的作用。近些年的研究表明，通过第一性原理计算、分子动力学模拟、人工智能等方法，可以有效预测材料性质，高效设计新材料，在电子元件、医药等领域有着广泛应用。

在稀土材料仿真预测中，通常会采用微观，介观，宏观等方法。微观层面会从电子或原子的角度构建随机微分方程组进行仿真，构建的随机微分方程组通常具有变量个数多、求解速度慢等问题。同时，近年来，也会采用将微观，介观，宏观等模拟方法结合起来的多尺度模拟方法对稀土材料进行仿真，多尺度模拟方法可以更好地理解材料的性质，优化材料的性能。此外，也会利用人工智能的方法对材料的性质、结构进行预测，人工智能的方法为材料预测带来了新思路，但同时也面临无法给出预测结果的物理解释等问题

[问题]（可以从以下问题中选择一个或几个进行研究）：

1. 构建一套高效通用的稀土材料微观模拟仿真算法，在保证精度的前提下，解决已有方法中求解速度慢的问题。

2. 利用数学方法（如微分方程、概率统计、优化理论等）与人工智能相结合的方式，构建一套高效精准的多尺度模拟方法。

3. 利用数学与统计方法（如概率因果推断等），提高材料预测结果的可解释性。

课题4. 非线性期望计算及金融应用

[背景] 自概率论公理化体系建立后，概率相关理论在各个领域得到广泛的应用与发展，尤其体现在对市场的不确定因素问题的度量与分析中，例如度量金融风险的重要指标风险价值（Value at Risk，简称VaR）就是基于公理化概率论体系所构建的。1994年J．P Morgen投资银行在其风险评价系统中引入VaR的概念，之后，巴塞尔银行委员会将VaR纳入银行业风险管理指标，随着《巴塞尔II协议》的全球化使用，加速推动了VaR的进一步发展，VaR成为了衡量市场风险的首选方法。随着VaR的推广，越来越多的学者关注到其局限性。1997年，Taleb指出VaR无法估计罕见事件的风险，2008年，Einhorn和Brown指出VaR忽略了损失的尾部风险，只关注其分布中心附近的可控风险，进而导致风险被低估。除此之外，VaR不具有次可加性，即投资组合的VaR值可能会高于单个投资VaR值之和，高估了投资组合的风险，与分散投资降低风险的普遍规律相悖。

为了解决这个问题，中国科学院院士彭实戈教授跳出一般概率理论的框架，建立非线性期望空间的理论框架，提出一系列非线性期望框架下的基础性质、定义、定理等，引入次线性期望理论并提出了G-正态分布、G-布朗运动等概念。研究表明，非线性G-期望框架下的风险度量可以有效刻画数据中的模糊性，同时对极端情形具有较好的鲁棒性。基于G-期望所构建的G-VaR在理论上被证明满足次可加性，与分散投资降低风险的普遍规律所一致。

[问题]：

1. 结合某个领域（银行、证券、保险等），基于G-VaR构建情景分析与压力测试高效计算方法。

课题5. 人工智能赋能数学教育

在AI时代到来的背景下，数学教育正面临前所未有的发展机遇与挑战。为了适应新时代教育的需求，探索数智人、大语言模型等人工智能技术在数学教育中的创新应用。请自行提出一个具体的相关问题并进行研究，例如AI助教、虚拟教师、数学定理自动证明等。