

高等学校自然科学奖推荐项目公示

项目名称	动态不确定性知识发现理论与方法
推荐单位	西南交通大学
项目简介	
<p>如何从大数据中高效实时地发现有价值的知识具有重要的科学意义和应用价值，而大数据所呈现的体量巨大、多源异构、动态性和不确定性等特点使得知识发现理论与方法面临重大挑战，亟待发展新一代大数据知识发现技术，以突破现有成果的局限性。项目组在多个国家自然科学基金项目的持续资助下，历经十余年协同攻关，以粒计算理论为基础，以典型粗糙集模型为对象，以增量学习技术为手段，以云计算并行框架为支撑平台，在动态不确定性知识发现的理论与方法上取得了重大突破，获得了如下的重要科学发现。</p> <ol style="list-style-type: none">1. 构建了基于粒计算的增量学习理论框架和方法体系，揭示了不确定性知识在多层次粒度空间中的单调性演化规律和动态关联机理，提出了不依赖于全量数据的不确定性知识发现的增量计算理论与方法，降低了数据快速增长时学习模型及算法对时间和空间的需求。2. 设计了面向异构信息系统的粒化分析技术与近似计算模型，提出了异类特征下样本间的混合距离集成度量方法，发现了面向异构数据的融合粒化与层次分析机理，刻画了通过矩阵方法实现粒化表达与运算的内在关系，建立了异构数据的模糊粒化结构生成模型，突破了传统近似计算的局限性。3. 提出了数据体量大限制下的多粒度并行大规模知识获取方法，刻画了基于粒计算的数据分治、模型分治和知识融合机理，设计了高维数据特征选择的数据-模型并行计算框架，提出了基于云计算的大规模特征选择算法及其并行优化技术，构建了面向海量复杂数据的不确定性知识并行粗糙计算模型，实现了海量动态数据的多粒度高效知识发现。 <p>10 篇代表作发表在 IEEE TKDE, IEEE TFS 等本领域权威期刊上, SCI 他引 324 次, 总他引 479 次, 施引刊物包括 IEEE TKDE, IEEE TC, IEEE TFS, IEEE TSMC-S 等。本项目成果被加拿大皇家科学院 W. Pedrycz 院士 (IEEE TSMC-A 主编), 欧洲科学院 R. Slowinski 院士 (EUR J OPER RES 主编), 美国 X. Wu 教授 (IEEE TKDE 原主编, KIS 主编), 中国台湾 C.T. Lin 教授 (IEEE TFS 原主编), 中国澳门 C.L.P. Chen 教授 (IEEE TSMC-S 主编), 西班牙 F. Herrera 教授 (IF 主编), 日本 F. Hamido 教授 (KBS 主编) 和德国 J. Fürnkranz 教授 (DMKD 主编) 以及 AAAS/IEEE/IFSA/IRSS Fellow 等学者的大量正面引用和采用。第一完成人应邀赴新加坡、波兰、中国台湾和澳门等地举办的国际学术会议上做大会报告 17 次, 指导学生获得 2015 年天池大数据“新浪微博互动预测大赛”冠军 (奖金 20 万元), 形成重要国际影响。</p>	

主要完成人情况表

1. 姓名：李天瑞，排名：1，技术职称：教授，工作单位：西南交通大学，完成单位：西南交通大学，曾获科技奖励：无。
对本项目技术创造性贡献：1) 作为项目负责人和主要完成人，制定项目总体研究方案。2) 揭示了不确定性知识在多层次粒度空间中的演化规律和动态关联机理，构建了基于粒计算的增量学习理论框架和方法体系。3) 提出了异类特征下样本间的混合距离集成度量方法，设计了面向异构信息系统的粒化分析技术与近似计算模型。4) 刻画了基于粒计算的数据分治、模型分治和知识融合机理，创建了面向海量复杂数据的不确定性知识并行粗糙计算模型。对本项目重要科学发现中的第1、2、3点做出了创新性贡献，是所有10篇代表性论文的作者，9篇论文的通讯作者。本项目研究工作量占本人工作量的90%。
2. 姓名：陈红梅，排名：2，技术职称：教授，工作单位：西南交通大学，完成单位：西南交通大学，曾获科技奖励：无。
对本项目技术创造性贡献：创建了属性值层次化变化不确定性知识动态更新模型，发现了属性和对象同时变化知识粒度的动态演化规律，提出了不同粒度变化下不依赖于全量数据的增量学习算法，建立了增量式不确定知识高效更新方法。发现了面向异构数据的融合粒化与层次分析机理，构建了面向异构数据的复合关系粗糙集模型。对本成果重要科学发现中第1、2点做出了创新性贡献，是代表性论文2、4、7、8、9的作者。本研究工作量占本人研究工作工作量的90%。
3. 姓名：张钧波，排名：3，技术职称：副研究员，工作单位：北京京东金融科技控股有限公司，完成单位：西南交通大学，曾获科技奖励：无。
对本项目技术创造性贡献：阐明了数据动态更新时属性、样本之间关系变化的内在机理，发现了不同粒度层次之间分类知识的数学转换关系，创建了面向异构数据的增量式知识更新模型与方法。刻画了海量数据中信息粒的并行计算机理，设计了高维数据特征选择的数据-模型并行计算框架，提出了基于云计算的大规模特征选择算法及其并行优化技术。对本成果重要科学发现中第2、3点做出了创新性贡献，是代表性论文2、5、6、10的第一作者。本项目研究工作量占本人工作量的90%。
4. 姓名：罗川，排名：4，技术职称：副教授，工作单位：四川大学，完成单位：西南交通大学，曾获科技奖励：无。
对本项目技术创造性贡献：揭示了集值有序分类问题中数据动态变化时样本之间优劣关系的动态更新规律，发现了不同粒度层次之间有序分类信息粒的单调性演化规律，提出了面向动态多准则集值系统的增量式不确定性有序分类模型与方法。对本成果重要科学发现中第1点做出了创新性贡献，是代表性论文7、8、9的作者。本项目研究工作量占本人工作量的90%。

5.姓名：刘盾，排名：5，技术职称：教授，工作单位：西南交通大学，完成单位：西南交通大学，曾获科技奖励：无。

对本项目技术创造性贡献：揭示了动态数据环境下不同层次数据之间的泛化关系和信息粒度变化表示形式，刻画了动态数据边界样本的分类不确定性，提出了不依赖于全量数据的不确定性分类模型的增量计算方法。对本成果重要科学发现中第1点做出了创新性贡献，是代表性论文3、6、7的作者。本项目研究工作量占本人工作量的60%。

代表性论文专著目录

序号	论文专著名称/刊名/作者	影响因子	年卷页码 (xx年xx卷xx页)	发表时间 年月日	通讯作者	第一作者	SCI他引次数	他引总次数	是否国内完成
1	A rough sets based characteristic relation approach for dynamic attribute generalization in data mining/ KNOWLEDGE-BASED SYSTEMS/Li, Tianrui; Ruan, Da; Geert, Wets; Song, Jing; Xu, Yang	4.396	2007 (20): 485-494	2007年6月	李天瑞	李天瑞	78	108	是
2	Composite rough sets for dynamic data mining/ INFORMATION SCIENCES/Zhang, Junbo; Li, Tianrui; Chen, Hongmei	4.305	2014 (257): 81-100	2014年2月	李天瑞	张钧波	44	60	是
3	Incorporating logistic regression to decision-theoretic rough sets for classifications/ INTERNATIONAL JOURNAL OF APPROXIMATE REASONING/Liu, Dun; Li, Tianrui; Liang, Decui	1.766	2014 (55): 197-210	2014年1月	刘盾	刘盾	40	53	是
4	A rough-set-based incremental approach for updating approximations under dynamic maintenance environments/ IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING/Chen, Hongmei; Li, Tianrui; Ruan, Da; Lin, Jianhui; Hu, Chengxiang	2.775	2013 (25): 274-284	2013年2月	李天瑞	陈红梅	38	52	是
5	A parallel method for computing rough set approximations/ INFORMATION SCIENCES/Zhang, Junbo; Li, Tianrui; Ruan, Da	4.305	2012 (194): 209-223	2012年7月	李天瑞	张钧波	29	57	是
6	Rough sets based matrix approaches with dynamic attribute variation in set-valued information systems/ INTERNATIONAL JOURNAL OF APPROXIMATE REASONING/Zhang, Junbo; Li, Tianrui; Ruan, Da; Liu, Dun	1.766	2012 (53): 620-635	2012年6月	李天瑞	张钧波	26	53	是
7	Incremental approaches for updating approximations in set-valued ordered information systems/ KNOWLEDGE-BASED SYSTEMS/Luo, Chuan; Li, Tianrui; Chen, Hongmei; Liu, Dun	4.396	2013 (50): 218-233	2013年9月	李天瑞	罗川	22	31	是
8	A decision-theoretic rough set approach for dynamic data mining/ IEEE TRANSACTIONS ON FUZZY SYSTEMS/ Chen, Hongmei; Li, Tianrui; Luo, Chuan; Horng, Shi-Jinn; Wang, Guoyin	8.415	2015 (23): 1958-1970	2015年12月	李天瑞	陈红梅	19	24	是
9	A rough set-based method for updating decision rules on	2.775	2014 (26): 2886-2899	2014年12月	李天瑞	陈红梅	17	23	是

	attribute values' coarsening and refining/IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING/Chen, Hongmei; Li, Tianrui; Luo, Chuan; Horng, Shi-Jinn; Wang, Guoyin								
10	A parallel matrix-based method for computing approximations in incomplete information systems/IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING/Zhang, Junbo; Wong, Jian-Syuan; Pan, Yi; Li, Tianrui	2.775	2015 (27): 326-339	2015 年 2 月	李天瑞	张钧波	11	18	是