

机器学习算法学术论文分析报告

Machine Learning algorithm academic paper analysis report

方建勇¹（余姚，浙江 315400）

摘要：通过超星发现系统，我们大致了解到机器学习算法学术论文所涉及的相关领域，可以明确地判断出哪些高校院所发表的学位论文较多，为我们从事机器学习算法产线的快速原型构造做了比较好的指引。

关键词：机器学习 算法 学术论文 分析报告

Abstract: Through the superstar discovery system, we generally understand the relevant fields involved in machine learning algorithm academic papers. We can clearly determine which college academic papers have published more dissertations and have done a rapid prototyping structure for our machine learning algorithm production line. Better guidance.

Key words: Machine Learning;algorithm;academic paper;analysis report

机器学习(Machine Learning, ML)是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。它是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域，它主要使用归纳、综合而不是演绎。²

一、机器学习算法学术发展趋势

表 1 机器学习算法学术发展趋势

¹ 方建勇，男，1978 年-，美国电气电子工程师学会 IEEE 会员，美国计算机学会 ACM 会员，中国工业与应用数学学会会员，中国计算机学会会员，中国中文信息学会会员，中国物流学会会员，浙江大学数学与应用数学专业毕业，MachineCoastline（机器海岸线）CEO&Founder。

² 参考百度机器学习词条。

机器学习 算法-各类型学术发展趋势								
序号	年份	图书(数量)	期刊(数量)	学位论文(数量)	会议论文(数量)	专利(数量)	报纸(数量)	科技成果(数量)
1	1965	0	1	0	0	0	0	0
2	1966	0	0	0	0	0	0	0
3	1967	0	0	0	0	0	0	0
4	1968	0	0	0	0	0	0	0
5	1969	0	0	0	0	0	0	0
6	1970	0	0	0	0	0	0	0
7	1971	0	0	0	0	0	0	0
8	1972	0	0	0	0	0	0	0
9	1973	0	0	0	0	0	0	0
10	1974	0	0	0	0	0	0	0
11	1975	0	0	0	0	0	0	0
12	1976	0	0	0	0	0	0	0
13	1977	0	0	0	0	0	0	0
14	1978	0	0	0	0	0	0	0
15	1979	0	1	0	0	0	0	0
16	1980	0	0	0	0	0	0	0
17	1981	0	1	1	0	0	0	0
18	1982	0	2	0	0	0	0	0
19	1983	0	1	0	0	0	0	0
20	1984	0	0	2	1	0	0	0
21	1985	0	0	0	0	0	0	0
22	1986	0	0	0	0	0	0	0
23	1987	0	0	1	0	0	0	0
24	1988	0	2	0	0	0	0	0
25	1989	0	1	0	7	0	0	0
26	1990	0	10	0	4	0	0	0
27	1991	0	6	0	22	0	0	0
28	1992	0	12	2	1	0	0	0
29	1993	1	12	2	6	0	0	0
30	1994	0	13	2	4	0	0	0
31	1995	0	18	5	6	0	0	0
32	1996	0	16	7	5	1	0	1
33	1997	0	27	18	2	0	0	0
34	1998	1	35	15	45	0	0	1

35	1999	0	45	15	4	0	0	2
36	2000	0	63	26	13	0	0	1
37	2001	2	62	35	9	1	0	6
38	2002	1	92	72	10	0	0	6
39	2003	4	109	130	24	3	0	5
40	2004	6	155	220	16	3	0	4
41	2005	4	209	347	35	9	0	8
42	2006	2	214	581	48	7	0	10
43	2007	4	276	642	72	9	0	11
44	2008	12	299	647	44	11	0	24
45	2009	10	307	873	61	14	0	10
46	2010	12	297	806	33	24	0	15
47	2011	9	302	962	39	23	0	12
48	2012	7	317	1016	35	40	0	8
49	2013	9	337	1141	33	40	0	7
50	2014	13	370	1282	30	62	2	4
51	2015	23	466	1300	29	121	18	3
52	2016	21	705	1133	23	225	11	1
53	2017	43	916	694	19	326	6	0
54	2018	55	1067	520	17	423	0	0

二、机器学习算法成果统计³

1、关键词

关键词涉及机器学习(3867⁴)、支持向量机(2695)、数据挖掘(1012)、神经网络(842)、遗传算法(782)、特征选择(568)、特征提取(567)、分类(424)、聚类(395)、决策树(388)、人工智能(382)、集成学习(364)、文本分类(360)、模式识别(343)、粗糙集(342)、半监督学习(334)、强化学习(324)、核函数(280)、统计学习理论(280)、故障诊断(277)、人脸识别(266)、学习算法(256)、流形学习(244)、机器人(227)、移动机器人(226)、SVM(226)、入侵检测(216)、Bp 神经网络(206)、聚类分析(202)、算法(189)、机器视觉(188)、属性约简(188)、人工神经网络(186)、路径规划(186)、分类器(179)、分类算法(178)、图像处理(172)、关联规则(165)、主动学习(165)、预测(147)、图像分割(144)、Adaboost(141)、图像分类(138)、目标跟踪(136)、稀

³ 数据来源于超星发现系统。

⁴ 括号内数字为出现频次，下同。

疏表示(136)、增量学习(134)、最小二乘支持向量机(132)、自然语言处理(129)、聚类算法(128)等。

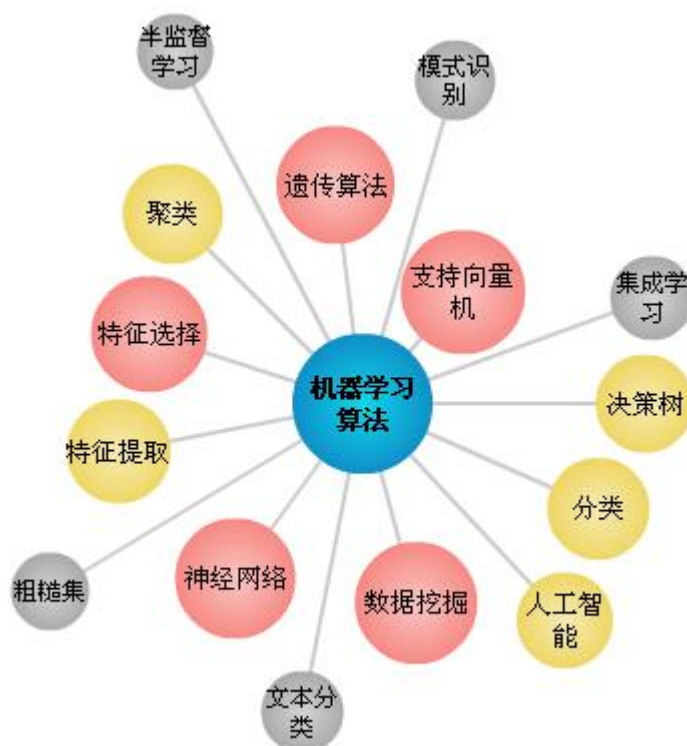


图 1 机器学习算法学术论文关键词频次泡型图

2、学术论文所在高校院所分布

学术论文所在高校院所依次为哈尔滨工业大学(586)、上海交通大学(407)、浙江大学(400)、中国科学院(397)、西安电子科技大学(389)、华南理工大学(372)、吉林大学(321)、电子科技大学(310)、大连理工大学(292)、北京航空航天大学(287)、北京邮电大学(287)、华中科技大学(273)、重庆大学(271)、中国科学技术大学(268)、中山大学(254)、北京交通大学(252)、东北大学(246)、南京理工大学(240)、西安交通大学(228)、南京大学(226)、清华大学(218)、武汉大学(213)、北京工业大学(211)、哈尔滨工程大学(208)、东南大学(208)、山东大学(185)、北京理工大学(184)、南京航空航天大学(181)、中南大学(178)、河北大学(172)、合肥工业大学(171)、厦门大学(169)、苏州大学(167)、南京邮电大学(162)、江南大学(161)、国防科学技术大学(159)、安徽大学(157)、湖南大学(157)、天津大学(156)、上海大学(151)、北京大学(150)、复旦大学(147)、北京科技大学(136)、广东工业大学(135)、太原

理工大学(127)、西北工业大学(126)、中国科学院大学(125)、山东师范大学(124)、四川大学(113)等。

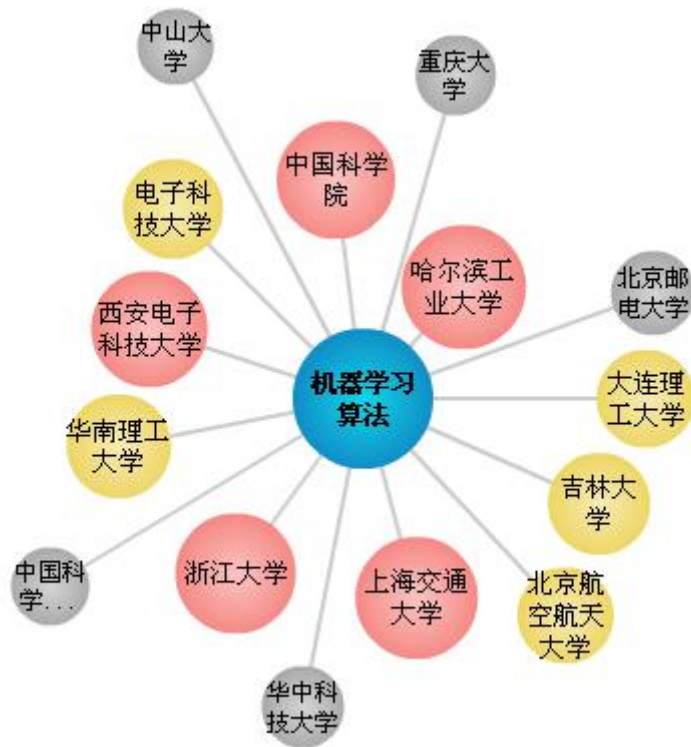


图 2 机器学习算法学术论文所在高校院所频次泡型图

3、高被引学术论文成果（2012-2018 年）

表 2 高被引学术论文成果（被引用次数超过 20 次）

[期刊] 深度学习的昨天、今天和明天 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(303)
作者：余凯，贾磊，陈雨强，徐伟（百度）
出处：计算机研究与发展 2013 第 9 期 P1799-1804 1000-1239
关键词：机器学习；深度学习；语音识别；图像识别；自然语言处理；在线广告

[期刊] 深度学习研究综述 [中文核心期刊(北大)] [统计源期刊(中信所)]

[CSCD 中国科学引文库(中科院)] 引证(287)

作者：孙志军，薛磊，许阳明，王正（电子工程学院；安徽省电子制约技术重点实验室）

出处：计算机应用研究 2012 第 8 期 P2806-2810 1001-3695

关键词：深度学习；分布式表示；深信度网络；卷积神经网络；深凸网络

[学位论文] BP 神经网络的研究分析及改进应用 引证(149)

作者：李友坤（安徽理工大学）

学位名称：硕士

出处：安徽理工大学 2012

关键词：人工神经网络；最速下降法；标准 BP 算法缺点；鲁棒性

[期刊] 基于机器学习的中文微博情感分类实证研究 [统计源期刊(中信所)]

[CSCD 中国科学引文库(中科院)] 引证(141)

作者：刘志明，刘鲁（北京航空航天大学经济管理学院）

出处：计算机工程与应用 2012 第 1 期 P1-4 1002-8331

关键词：微博；情感分类；机器学习；特征选取；特征项权重

[期刊] 大数据下的机器学习算法综述 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(128)

作者：何清，李宁，罗文娟，史忠植（中国科学院计算技术研究所智能信息处理重点实验室；中国科学院大学；河北大学数学与计算机学院）

出处：模式识别与人工智能 2014 第 4 期 P327-336 1003-6059

关键词：大数据；机器学习；分类；聚类；并行算法

[期刊] 供应链金融视角下的中小企业信用风险评估研究：基于 SVM 与 BP 神经网络的比较研究 [CSSCI 中文社科引文索引(南大)] [中文核心期刊(北大)] [统计源期刊(中信所)] 引证(119)

作者：胡海青，张琅，张道宏（西安理工大学经济与管理学院）

出处：管理评论 2012 第 11 期 P70-80 1003-1952

关键词：供应链金融；中小企业；信用风险评估；SVM；BP 神经网络

[期刊] 随机森林模型在分类与回归分析中的应用 [统计源期刊(中信所)] 引证(103)

作者：李欣海（中国科学院动物研究所）

出处：应用昆虫学报 2013 第 4 期 P1190-1197 2095-1353

关键词：随机森林；分类树；判别分析；回归；机器学习

[学位论文] 电力变压器状态评估及故障诊断方法研究 引证(101)

作者：郑含博（重庆大学）

学位名称：博士

出处：重庆大学 2012

关键词：电力变压器；状态评估；油中溶解气体分析；故障诊断；故障预测

[学位论文] 推荐系统关键技术研究 引证(98)

作者：任磊（华东师范大学）

学位名称：博士

出处：华东师范大学 2012

关键词：推荐系统；基于项目的协作推荐；混合推荐；用户建模；稀疏性；概念漂移

[学位论文] 基于协同过滤的推荐算法研究 引证(91)

作者：刘青文（中国科学技术大学）

学位名称：博士

出处：中国科学技术大学 2013

关键词：推荐系统；协同过滤；社交网络；数据稀疏问题；冷启动问题；用户相似度模型；模型融合；堆叠回归；基于用户的方法；基于项目的方法

[期刊] 迁移学习研究进展 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(87)

作者: 庄福振, 罗平, 何清, 史忠植 (中国科学院智能信息处理重点实验室 (中国科学院计算技术研究所))

出处: 软件学报 2015 第 1 期 P26-39 1000-9825

关键词: 迁移学习; 相关领域; 独立同分布; 生成模型; 概念学习

[期刊] 微博客中转发行为的预测研究 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(87)

作者: 张旸, 路荣, 杨青 (中国科学院自动化研究所模式识别国家重点实验室)

出处: 中文信息学报 2012 第 4 期 P109-114, 121 1003-0077

关键词: 微博客; 转发; 特征加权模型

[学位论文] 基于支持向量机的聚类及文本分类研究 引证(76)

作者: 平源 (北京邮电大学)

学位名称: 博士

出处: 北京邮电大学 2012

关键词: 内容安全; 支持向量机; 聚类分析; 支持向量聚类; 文本分类

[期刊] 玻尔兹曼机研究进展 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(64)

作者: 刘建伟, 刘媛, 罗雄麟 (中国石油大学自动化研究所)

出处: 计算机研究与发展 2014 第 1 期 P1-16 1000-1239

关键词: 玻尔兹曼机; 可见单元; 隐单元; 概率分布; 期望值; 模拟退火; 吉布斯采样; 马尔可夫链

[学位论文] 基于聚类分析的图像分割算法研究 引证(62)

作者: 许晓丽 (哈尔滨工程大学)

学位名称：博士

出处：哈尔滨工程大学 2012

关键词：图像分割；聚类分析；谱聚类；mean；shift；近邻传播；数据点竞争

[学位论文] 聚类分析中的相似性度量及其应用研究 引证(61)

作者：白雪（北京交通大学）

学位名称：博士

出处：北京交通大学 2012

关键词：聚类分析；相似性度量；信息瓶颈；基于信息的聚类；信息学习理论；联合互信息；图像处理

[期刊] 差分隐私保护及其应用 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(60)

作者：熊平，朱天清，王晓峰（中南财经政法大学信息与安全工程学院；澳大利亚迪肯大学信息技术学院；武汉轻工大学数学与计算机学院；中国科学院计算技术研究所无线传感网络实验室）

出处：计算机学报 2014 第 1 期 P101-122 0254-4164

关键词：差分隐私；数据发布；数据挖掘；机器学习；统计查询；隐私保护

[图书] 这就是搜索引擎 核心技术详解 引证(60)

作者：张俊林著

出处：北京：电子工业出版社 2012 300 页

ISBN: 978-7-121-14865-1

主题词：互联网络-情报检查

[学位论文] 大规模数据集高效数据挖掘算法研究 引证(54)

作者：申彦（江苏大学）

学位名称：博士

出处：江苏大学 2013

关键词：大规模数据集；挖掘算法；关联规则；磁盘 FPTREE；半监督学习；EM 算法；概率聚类

[期刊] 受限波尔兹曼机 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(54)

作者：张春霞，姬楠楠，王冠伟（西安交通大学数学与统计学院；长安大学理学院；西安工业大学机电工程学院）

出处：工程数学学报 2015 第 2 期 P159-173 1005-3085

关键词：机器学习；深度学习；受限波尔兹曼机；对比散度；Gibbs 采样

[学位论文] 深度学习神经网络在语音识别中的应用研究 引证(54)

作者：陈硕（华南理工大学）

学位名称：硕士

出处：华南理工大学 2013

关键词：神经网络；语音识别；反向传播算法(BP)；限制玻尔兹曼机(RBM)；深度信念网络(DBN)

[学位论文] 基于语义统计分析的网络舆情挖掘技术研究 引证(51)

作者：万源（武汉理工大学）

学位名称：博士

出处：武汉理工大学 2012

关键词：网络舆情挖掘；Fisher 判别准则；突发事件分类；局部潜在语义分析；GARCH 类模型

[学位论文] 基于深度学习的人脸识别研究 引证(50)

作者：林妙真（大连理工大学）

学位名称：硕士

出处：大连理工大学 2013

关键词：深度学习；人脸识别；人脸姿态；深度信念网络

[期刊] 基于采样策略的主动学习算法研究进展 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(49)

作者：吴伟宁，刘扬，郭茂祖，刘晓燕（哈尔滨工业大学计算机科学与技术学院）

出处：计算机研究与发展 2012 第 6 期 P1162-1173 1000-1239

关键词：机器学习；主动学习；采样策略；标记代价；样例选择

[期刊] 深度学习结构和算法比较分析 [中文核心期刊(北大)] [统计源期刊(中信所)] [CA 化学文摘(美)] 引证(47)

作者：李海峰，李纯果（河北大学教务处；河北大学数学与计算机学院）

出处：河北大学学报(自然科学版) 2012 第 5 期 P538-544 1000-1565

关键词：深度机器学习；无监督贪婪学习算法；DBNs；RBMs

[期刊] 一种新的不平衡数据学习算法 PCBoost [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(45)

作者：李雄飞，李军，董元方，屈成伟（吉林大学符号计算与知识工程教育部重点实验室；长春理工大学应用数学系；长春理工大学经济管理学院）

出处：计算机学报 2012 第 2 期 P2202-2209 0254-4164

关键词：数据挖掘；不平衡数据；集成学习；提升；扰动

[期刊] 基于粒计算的 K-medoids 聚类算法 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(45)

作者：马箐，谢娟英（陕西师范大学计算机科学学院）

出处：计算机应用 2012 第 7 期 P1973-1977 1001-9081

关键词：传统 K-medoids 聚类算法；快速 K-medoids 聚类算法；粒计算；等价关系；聚类

[期刊] 随机森林理论浅析 引证(45)

作者: 董师师, 黄哲学 (深圳市高性能数据挖掘重点实验室 深圳
(518055); 中国科学院深圳先进技术研究院 深圳 (518055))

出处: 集成技术 2013 第 1 期 2095-3135

关键词: 随机森林; 数据挖掘; 机器学习

[期刊] 基于确定学习的机器人任务空间自适应神经网络控制 [EI 工程索引
(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)]
引证(44)

作者: 吴玉香, 王聪 (华南理工大学自动化科学与工程学院)

出处: 自动化学报 2013 第 6 期 P806-815 0254-4156

关键词: 确定学习; 径向基函数神经网络; 自适应神经网络控制; 机器人

[学位论文] K-Means 聚类算法的研究与改进 引证(44)

作者: 崔丹丹 (安徽大学)

学位名称: 硕士

出处: 安徽大学 2012

关键词: 聚类分析; 数据挖掘; K-Means 聚类算法

[图书] 支持向量机的算法设计与分析 引证(44)

作者: 杨晓伟; 郝志峰

出处: 北京: 科学出版社 2013 206 页

ISBN: 978-7-03-037869-9

主题词: 向量计算机-算法分析-算法设计

摘要: 支持向量机的研究是近十余年来机器学习和模式识别领域中研究的热点, 受到了计算数学.....; 然后沿着先直观后严谨的思路介绍了二分类模型和回归模型; 最后对与分解算法、最小二乘支持向量机、多分类、模糊支持向量机.

[学位论文] 贝叶斯网络结构学习与推理研究 引证(43)

作者：朱明敏（西安电子科技大学）

学位名称：博士

出处：西安电子科技大学 2013

关键词：贝叶斯网络；结构学习；条件独立测试；联接树；条件线性高斯网络最大主子图分解；人工蜂群优化；Markov 边界；Markov 等价；推理计算

[图书] 深入理解大数据 大数据处理与编程实践 引证(43)

作者：黄宜华主编

出处：北京：机械工业出版社 2014 488 页

ISBN：978-7-111-47325-1

主题词：数据管理-高等学校-教材

[期刊] 基于样本空间分布密度的初始聚类中心优化 K-均值算法 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(42)

作者：谢娟英，郭文娟，谢维信，高新波（陕西师范大学计算机科学学院；西安电子科技大学电子工程学院；深圳大学信息工程学院 ATR 国家重点实验室）

出处：计算机应用研究 2012 第 3 期 1001-3695

关键词：聚类；K-均值聚类；初始中心；邻域；样本分布密度

[期刊] 一种结合主动学习的半监督文档聚类算法 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(40)

作者：赵卫中，马慧芳，李志清，史忠植（湘潭大学信息工程学院；中国科学院计算技术研究所智能信息处理重点实验室；西北师范大学数学与信息科学学院）

出处：软件学报 2012 第 6 期 P1486-1499 1000-9825

关键词：半监督聚类；文档聚类；主动学习；成对约束

[期刊] 混合粒子群算法求解多目标柔性作业车间调调度度问题 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)]

院)] 引证(40)

作者：张静，王万良，徐新黎，介婧（浙江工业大学计算机科学与技术学院；浙江工业大学信息工程学院）

出处：控制理论与应用 2012 第 6 期 P715-722 1000-8152

关键词：粒子群；多目标优化；柔性作业车间调度问题；Baldwinian 学习策略

[学位论文] 电力变压器故障诊断方法研究 引证(39)

作者：武中利（华北电力大学）

学位名称：博士

出处：华北电力大学 2013

关键词：变压器；故障诊断；振动信号；因子分析；基因表达式；物元模型；局部放电

[学位论文] 基于 Hadoop 的数据挖掘算法的分析与研究 引证(38)

作者：张明辉（昆明理工大学）

学位名称：硕士

出处：昆明理工大学 2012

关键词：数据挖掘；Hadoop；Mahout；K-Means

[学位论文] 基于 Spark 的机器学习平台设计与实现 引证(37)

作者：唐振坤（厦门大学）

学位名称：硕士

出处：厦门大学 2014

关键词：Spark；机器学习；海量数据挖掘

[学位论文] 基于学习算法的 WLAN 室内定位技术研究 引证(37)

作者：邓志安（哈尔滨工业大学）

学位名称：博士

出处：哈尔滨工业大学 2012

关键词：室内定位；接收信号强度；核函数；流形学习；接入点选择；聚类分块

[期刊] 基于外观模型学习的视频目标跟踪方法综述 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(37)

作者：张焕龙，胡士强，杨国胜（上海交通大学航空航天学院；洛阳理工学院计算机与信息工程系；中央民族大学信息工程学院）

出处：计算机研究与发展 2015 第 52 卷 第 1 期 P177-190 1000-1239

关键词：外观模型学习；稀疏表示；度量学习；视频跟踪；半监督学习

[期刊] 基于模糊聚类分析的云计算负载平衡策略 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(37)

作者：姚婧，何聚厚（陕西师范大学计算机科学学院）

出处：计算机应用 2012 第 1 期 P213-217 1001-9081

关键词：云计算；模糊聚类；负载平衡；粒子群优化算法

[学位论文] 视觉显著性应用研究 引证(37)

作者：敖欢欢（中国科学技术大学）

学位名称：博士

出处：中国科学技术大学 2013

关键词：视觉显著性；显著性计算模型；显著性边缘提取；视觉显著性表示方法；遥感图像弱小目标检测

[期刊] 利用 word2vec 对中文词进行聚类研究 引证(36)

作者：郑文超，徐鹏（北京邮电大学网络技术研究院）

出处：软件 2013 第 34 卷 第 12 期 P160-162 1003-6970

关键词：数据挖掘；聚类；分词；词向量；神经网络

[期刊] 遗传算法和粒子群优化算法的性能对比分析 [统计源期刊(中信所)]
[中文核心期刊(北大)] [CSCD 中国科学引文库(中科院)] 引证(35)

作者: 张鑫源, 胡晓敏, 林盈 (中山大学电子与通信工程系; 中山大学公共卫生学院卫生信息研究中心 广东省卫生信息学重点实验室; 中山大学心理学系)

出处: 计算机科学与探索 2014 第 8 卷 第 1 期 P90-102 1673-9418

关键词: 遗传算法; 粒子群优化算法; 单峰; 多峰; 性能对比

[期刊] 基于特征子集区分度与支持向量机的特征选择算法 [EI 工程索引(美)]
[中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(35)

作者: 谢娟英, 谢维信 (陕西师范大学计算机科学学院; 深圳大学信息工程学院 ATR 国家重点实验室)

出处: 计算机学报 2014 第 8 期 P1704-1718 0254-4164

关键词: 特征选择; 支持向量机; 相关性; 特征子集区分度; 特征区分度

[学位论文] 基于分布式平台 Spark 和 YARN 的数据挖掘算法的并行化研究 引证(35)

作者: 梁彦 (中山大学)

学位名称: 硕士

出处: 中山大学 2014

关键词: 云计算; 数据挖掘; 聚类分析; k-means; canopy; 并行算法; Spark; YARN

[期刊] 稀疏子空间聚类综述 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(34)

作者: 王卫卫, 李小平, 冯象初, 王斯琪 (西安电子科技大学数学与统计学院)

出处: 自动化学报 2015 第 41 卷 第 8 期 P1373-1384 0254-4156

关键词: 高维数据; 子空间聚类; 稀疏表示; 低秩表示

[学位论文] 工业机器人的高速高精度控制方法研究 引证(34)

作者：刘海涛（华南理工大学）

学位名称：博士

出处：华南理工大学 2012

关键词：工业机器人；参数辨识；有限时间控制；鲁棒 H^∞ ；时延估计

[期刊] 基于表示学习的中文分词算法探索 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(33)

作者：来斯惟，徐立恒，陈玉博，刘康，赵军（中国科学院自动化研究所 模式识别国家重点实验室）

出处：中文信息学报 2013 第 5 期 P8-14 1003-0077

关键词：表示学习；中文分词

[期刊] 一种新型多标记懒惰学习算法 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(32)

作者：张敏灵（东南大学计算机科学与工程学院）

出处：计算机研究与发展 2012 第 11 期 P2271-2282 1000-1239

关键词：机器学习；多标记学习；懒惰学习；功能基因组学；图像分类

[期刊] 贝叶斯网络研究综述 引证(31)

作者：胡春玲（合肥学院网络与智能信息处理重点实验室）

出处：合肥学院学报(自然科学版) 2013 第 1 期 P33-40 1673-162X

关键词：贝叶斯网络；推理；学习

[期刊] 大数据与宏观经济分析研究综述 [CSSCI 中文社科引文索引(南大)] [中文核心期刊(北大)] 引证(30)

作者：刘涛雄，徐晓飞（清华大学经济学研究所）

出处：国外理论动态 2015 第 1 期 P57-64 1674-1277

关键词：大数据；宏观经济；数据挖掘；机器学习算法；预测

[期刊] 一种基于多特征和机器学习的分级行人检测方法 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(29)

作者：种衍文，匡湖林，李清泉（武汉大学测绘遥感信息工程国家重点实验室）

出处：自动化学报 2012 第 3 期 P375-381 0254-4156

关键词：四方向特征；熵梯度直方图；自适应增强算法；GAB 级联；支持向量机；两级检测

[期刊] 基于 HOG 和 SVM 的公交乘客人流量统计算法 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CA 化学文摘(美)] [CSCD 中国科学引文库(中科院)] 引证(28)

作者：徐超，高梦珠，查宇锋，曹利民（安徽大学智能与信号处理教育部重点实验室；安徽微纳电子器件与集成电路设计省级实验室）

出处：仪器仪表学报 2015 第 2 期 P446-452 0254-3087

关键词：HOG；SVM；数据关联；Camshift 跟踪算法

[期刊] 不平衡数据的集成分类算法综述 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(28)

作者：李勇，刘战东，张海军（新疆师范大学网络信息安全与舆情分析重点实验室；南京航空航天大学计算机科学与技术学院）

出处：计算机应用研究 2014 第 5 期 P1287-1291 1001-3695

关键词：不平衡数据；集成学习；分类；代价敏感；数据采样

[学位论文] 视觉地形分类和四足机器人步态规划方法研究与应用 引证(28)

作者：李彬（山东大学）

学位名称：博士

出处：山东大学 2012

关键词：视觉地形分类；极端学习机；四足机器人；步态规划；中枢模式发

生器

[期刊] 分类不平衡协议流的机器学习算法评估与比较 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(26)

作者: 张宏莉, 鲁刚 (哈尔滨工业大学 计算机科学与技术学院 计算机网络与信息安全技术研究中心)

出处: 软件学报 2012 第 6 期 P1500-1516 1000-9825

关键词: 不平衡; 特征选择; 流量分类; 集成学习; 单因子实验

[期刊] 基于机器学习的网络流量分类研究进展 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(26)

作者: 王涛, 余顺争 (广东工业大学网络与信息化工程系; 中山大学电子与通信工程系)

出处: 小型微型计算机系统 2012 第 5 期 P1034-1040 1000-1220

关键词: 机器学习; 网络流; 网络流量分类; 统计特征

[期刊] 最小方差优化初始聚类中心的 K-means 算法 [统计源期刊(中信所)] [中文核心期刊(北大)] [CSCD 中国科学引文库(中科院)] 引证(26)

作者: 谢娟英, 王艳娥 (陕西师范大学计算机科学学院)

出处: 计算机工程 2014 第 8 期 P205-211, 223 1000-3428

关键词: 聚类; K-means 算法; 方差; 紧密度; 初始聚类中心

[期刊] 基于 Bagging 的概率神经网络集成分类算法 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(26)

作者: 蒋芸, 陈娜, 明利特, 周泽寻, 谢国城, 陈珊 (西北师范大学计算机科学与工程学院)

出处: 计算机科学 2013 第 5 期 P242-246 1002-137X

关键词: 分类; BP 神经网络; 概率神经网络; 集成学习; Bagging

[学位论文] 应用 Apriori 关联规则算法的数据挖掘技术挖掘电子商务潜在客户 引证(26)

作者: 乌文波 (浙江工业大学)

学位名称: 硕士

出处: 浙江工业大学 2012

关键词: 数据挖掘; 关联规则; Apriori 算法; 潜在客户挖掘

[期刊] 概率图模型研究进展综述 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(26)

作者: 张宏毅, 王立威, 陈瑜希 (机器感知与智能教育部重点实验室(北京大学); 北京大学信息科学技术学院智能科学系)

出处: 软件学报 2013 第 24 卷 第 11 期 P2476-2497 1000-9825

关键词: 概率图模型; 概率推理; 机器学习

[期刊] 基于 SVM 的房地产投资风险评价及应用 [CSSCI 中文社科引文索引(南大)] [中文核心期刊(北大)] 引证(26)

作者: 李毅 (重庆大学建设管理与房地产学院)

出处: 统计与决策 2012 第 1 期 P70-72 1002-6487

关键词: 房地产; 支持向量机; 风险评价

[期刊] 相关向量机分类方法的研究进展与分析 [中文核心期刊(北大)] [统计源期刊(中信所)] 引证(26)

作者: 赵春晖, 张焱 (哈尔滨工程大学信息与通信工程学院)

出处: 智能系统学报 2012 第 4 期 P294-301 1673-4785

关键词: 相关向量机; 改进型相关向量机; 高光谱图像; 分类算法

[期刊] 大数据处理模型 Apache Spark 研究 引证(26)

作者: 黎文阳 (四川大学计算机学院)

出处: 现代计算机 2015 第 8 期 P55-60 1007-1423

关键词：Spark；Hadoop；MapReduce；大数据；数据分析

[期刊] 移动机器人轨迹跟踪的模糊 PID-P 型迭代学习控制 [EI 工程索引(美)]
[中文核心期刊(北大)] [统计源期刊(中信所)] [CA 化学文摘(美)] [CSCD 中国科学引
文库(中科院)] 引证(25)

作者：刘国荣，张扬名（湘潭大学信息工程学院；湖南工程学院）

出处：电子学报 2013 第 8 期 P1536-1541 0372-2112

关键词：移动机器人；迭代学习控制；轨迹跟踪；模糊控制

[学位论文] Android 恶意软件静态检测方案的研究 引证(25)

作者：童振飞（南京邮电大学）

学位名称：硕士

出处：南京邮电大学 2012

关键词：AndroidAPK；DEX；恶意软件；静态检测；机器学习

[期刊] 短文本信息流的无监督会话抽取技术 [EI 工程索引(美)] [中文核心期
刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(25)

作者：黄九鸣，吴泉源，刘春阳，张旭，贾焰，周斌（国防科学技术大学计
算机学院；国家计算机网络应急技术处理协调中心）

出处：软件学报 2012 第 4 期 P735-747 1000-9825

关键词：会话抽取；短文本；短文本信息流；无监督；时序特征；上下文相
关度

[学位论文] WLAN 位置指纹室内定位关键技术研究 引证(25)

作者：陈丽娜（华东师范大学）

学位名称：博士

出处：华东师范大学 2014

关键词：室内定位；无线局域网；接收信号强度；聚类分块；接入点选择；
特征选择

[学位论文] 聚类分析及其在图像处理中的应用 引证(25)

作者: 肖宇 (北京交通大学)

学位名称: 博士

出处: 北京交通大学 2012

关键词: 聚类分析; 相似性度量; 本质维数; 半监督聚类; 家族相似性; 图像分割; 图像去噪

[学位论文] 基于卷积神经网络的深度学习算法与应用研究 引证(24)

作者: 陈先昌 (浙江工商大学)

学位名称: 硕士

出处: 浙江工商大学 2014

关键词: 卷积神经网络; 深度学习; 模式识别; 光学字符识别; 交通标示识别

[期刊] 基于邻域的 K 中心点聚类算法 [中文核心期刊(北大)] [统计源期刊(中信所)] [CA 化学文摘(美)] [CSCD 中国科学引文库(中科院)] 引证(24)

作者: 谢娟英, 郭文娟, 谢维信 (陕西师范大学计算机科学学院; 西安电子科技大学电子工程学院; ATR 国家重点实验室深圳大学信息工程学院)

出处: 陕西师范大学学报(自然科学版) 2012 第 4 期 P16-22 1672-4291

关键词: 邻域; K 中心点算法; 样本密度; 聚类; 样本空间分布

[期刊] 基于马氏决策过程模型的动态系统学习控制:研究前沿与展望 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(24)

作者: 徐昕, 沈栋, 高岩青, 王凯 (国防科学技术大学机电工程与自动化学院自动化研究所; 中国科学院自动化研究所; 复杂系统智能管理与控制国家重点实验室; 美国亚利桑那大学系统与工业工程学院; 国防科学技术大学军事计算实验与平行系统技术研究中心)

出处: 自动化学报 2012 第 5 期 P673-687 0254-4156

关键词：学习控制；Markov 决策过程；增强学习；近似动态规划；机器学习；自适应控制

[期刊] 基于高斯颜色模型和 SVM 的交通标志检测 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CA 化学文摘(美)] [CSCD 中国科学引文库(中科院)] 引证(24)

作者：常发亮，黄翠，刘成云，赵永国，马传峰（山东大学控制科学与工程学院；山东省机器人与制造自动化技术重点实验室 山东省科学院自动化研究所）

出处：仪器仪表学报 2014 第 1 期 P43-49 0254-3087

关键词：交通标志检测；高斯颜色模型；HOG 描述子；支持向量机

[学位论文] 基于无线自组织网络的智能路由协议的研究 引证(24)

作者：戴超（北京邮电大学）

学位名称：博士

出处：北京邮电大学 2012

关键词：智能路由；无线自组织网络；OODA；Q-学习；实验平台

[学位论文] 复杂背景下运动目标检测和识别关键技术研究 引证(24)

作者：解晓萌（华南理工大学）

学位名称：博士

出处：华南理工大学 2012

关键词：运动目标检测；目标识别；特征聚类；特征单词；Adaboost；无监督学习；高阶局部特征；特征树

[期刊] 基于流形学习算法的马铃薯机械损伤机器视觉检测方法 [CA 化学文摘(美)] [CSCD 中国科学引文库(中科院)] 引证(23)

作者：汪成龙，李小昱，武振中，周竹，冯耀泽（华中农业大学工学院；浙江农林大学信息工程学院）

出处：农业工程学报 2014 第 30 卷 第 1 期 P245-252 1002-6819

关键词：机器视觉；识别；算法；流形学习；机械损伤；马铃薯

[期刊] 主动学习算法综述 [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(23)

作者：刘康，钱旭，王自强（中国矿业大学(北京)机电与信息工程学院）

出处：计算机工程与应用 2012 第 34 期 P1-4, 22 1002-8331

关键词：主动学习；模式识别；机器学习

[学位论文] 基于目标函数的模糊聚类新算法及其应用研究 引证(23)

作者：汪庆淼（江苏大学）

学位名称：博士

出处：江苏大学 2014

关键词：模糊聚类；目标函数；模糊 c 均值聚类(FCM)；可能性 c 均值聚类(PCM)；不平衡数据集；聚类算法均衡化；模糊控制指标；m 值广义化；粒子群算法(PSO)；粒子群广义化；自适应 m 值

[学位论文] K 均值算法研究及其应用 引证(23)

作者：朱建宇（大连理工大学）

学位名称：硕士

出处：大连理工大学 2013

关键词：K 均值；离群点；交通事故；候选中心点；簇最大半径

[学位论文] 面向推荐系统的关键问题研究及应用 引证(23)

作者：刘士琛（中国科学技术大学）

学位名称：博士

出处：中国科学技术大学 2014

关键词：推荐系统；协同过滤；用户相似度模型；时间回溯模型；全局优化算法；用户属性预测模型；演化博弈

[学位论文] 基于极限学习的系统辨识方法及应用研究 引证(23)

作者: 杨易旻 (湖南大学)

学位名称: 博士

出处: 湖南大学 2013

关键词: 系统辨识; 极限学习机; 前馈神经网络; 渐进学习机; 机器人系统
辨识

[学位论文] 文本情感分类及观点摘要关键问题研究 引证(23)

作者: 张冬梅 (山东大学)

学位名称: 博士

出处: 山东大学 2012

关键词: 情感分类; 观点摘要; 不平衡数据分类; 集成学习; 语义倾向

[期刊] 概念漂移数据流分类研究综述 [中文核心期刊(北大)] [统计源期刊
(中信所)] 引证(22)

作者: 文益民, 强保华, 范志刚 (桂林电子科技大学计算机科学与工程学院;
中国科学院上海高等研究院)

出处: 智能系统学报 2013 第 2 期 P95-104 1673-4785

关键词: 大数据; 概念漂移; 增量学习; 适应学习; 数据流; 机器学习

[期刊] 数字抠像的最新研究进展 [EI 工程索引(美)] [中文核心期刊(北大)]
[统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(22)

作者: 张展鹏, 朱青松, 谢耀钦 (中国科学院深圳先进技术研究院; 中山大学)

出处: 自动化学报 2012 第 10 期 P1571-1584 0254-4156

关键词: 数字抠像; 自然图像抠像; 图像分割; 图像处理

[学位论文] SVM 核参数优化研究与应用 引证(22)

作者: 杨海 (浙江大学)

学位名称：硕士

出处：浙江大学 2014

关键词：支持向量机；核函数参数；网格搜索法；粒子群搜索法；萤火虫搜索法；空间手写识别

[学位论文] 人脸识别系统及关键技术研究 引证(22)

作者：袁正海（南京邮电大学）

学位名称：硕士

出处：南京邮电大学 2013

关键词：人脸识别；人脸检测；Adaboost 算法；局部二值模式；主成分分析；支持向量机

[期刊] 基于样例选取的属性约简算法 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(21)

作者：王熙照，王婷婷，翟俊海（河北省机器学习与计算智能重点实验室河北大学数学与计算机学院）

出处：计算机研究与发展 2012 第 11 期 P2305-2310 1000-1239

关键词：信息系统；样例选取；粗糙集；差别矩阵；属性约简

[期刊] 学习因子和时间因子随权重调整的粒子群算法 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(21)

作者：马国庆，李瑞峰，刘丽（哈尔滨工业大学机器人技术与系统国家重点实验室；长春理工大学机电工程学院）

出处：计算机应用研究 2014 第 11 期 P3291-3294 1001-3695

关键词：粒子群优化算法；学习因子；时间因子；边界限制；速度反弹

[期刊] 基于条件对数似然函数导数的贝叶斯网络分类器优化算法 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(21)

作者：王中锋，王志海（北京交通大学计算机与信息技术学院）

出处：计算机学报 2012 第 2 期 P2364-2374 0254-4164

关键词：机器学习；数据挖掘；分类器；贝叶斯网络；鉴别式训练策略

[期刊] 深度学习研究与进展 [中文核心期刊(北大)] [统计源期刊(中信所)]
[CSCD 中国科学引文库(中科院)] 引证(21)

作者：孙志远，鲁成祥，史忠植，马刚（中国科学院计算技术研究所智能信息处理重点实验室；中国科学院大学；曲阜师范大学信息科学与工程学院）

出处：计算机科学 2016 第 2 期 P1-8 1002-137X

关键词：深度学习；机器学习；深层神经网络；图像识别；语音识别；自然语言处理

[学位论文] 基于支持向量机的金融时间序列分析预测算法研究 引证(21)

作者：鲍漪澜（大连海事大学）

学位名称：博士

出处：大连海事大学 2013

关键词：支持向量机；金融时间序列；非线性检验；混沌；预测技术

[学位论文] 面向微博短文本的情感分析研究 引证(21)

作者：刘楠（武汉大学）

学位名称：博士

出处：武汉大学 2013

关键词：自然语言处理；文本情感分析；微博短文本；机器学习

[期刊] 基于支持向量机的供应链风险评估研究 [CSSCI 中文社科引文索引(南大)] [中文核心期刊(北大)] 引证(21)

作者：舒彤，葛佳丽，陈收（湖南大学工商管理学院）

出处：经济经纬 2014 第 1 期 P130-135 1006-1096

关键词：供应链风险；风险评估；支持向量机

[学位论文] 面向图像分类和识别的视觉特征表达与学习的研究 引证(21)

作者：杨钊（华南理工大学）

学位名称：博士

出处：华南理工大学 2014

关键词：视觉特征；特征表达；特征学习；深度学习

[学位论文] 基于 SIFT 的图像配准与拼接技术研究 引证(21)

作者：袁杰（南京理工大学）

学位名称：硕士

出处：南京理工大学 2013

关键词：图像配准；SIFT；kd-tree；由粗到细；分块匹配；图像融合

[期刊] 基于灰色投影改进随机森林算法的电力系统短期负荷预测 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(20)

作者：吴潇雨，和敬涵，张沛，胡骏（北京交通大学国家能源主动配电网技术研发中心）

出处：电力系统自动化 2015 第 12 期 P50-55 1000-1026

关键词：短期负荷预测；相似日；灰色投影法；随机森林；Bagging 抽样方法；袋外估计

[学位论文] 相关向量机多分类算法的研究与应用 引证(20)

作者：柳长源（哈尔滨工程大学）

学位名称：博士

出处：哈尔滨工程大学 2013

关键词：机器学习；相关向量机；分类器；人脸识别；增量学习

[期刊] 基于卷积神经网络的图像识别算法设计与实现 引证(20)

作者：王振，高茂庭（上海海事大学信息工程学院）

出处：现代计算机 2015 第 20 期 P61-66 1007-1423

关键词：卷积神经网络；深度学习；图像识别；机器学习；神经网络

[期刊] 深度卷积神经网络在计算机视觉中的应用研究综述* [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(20)

作者：卢宏涛，张秦川（上海交通大学计算机科学与工程系）

出处：数据采集与处理 2016 第 1 期 P1-17 1004-9039

关键词：深度学习；卷积神经网络；图像识别；目标检测；计算机视觉

[期刊] 可重构卫星锂离子电池剩余寿命预测系统研究 [EI 工程索引(美)] [中文核心期刊(北大)] [统计源期刊(中信所)] [CA 化学文摘(美)] [CSCD 中国科学引文库(中科院)] 引证(20)

作者：周建宝，王少军，马丽萍，杨思远，彭宇，彭喜元（哈尔滨工业大学自动化测试与控制系；上海空间电源研究所）

出处：仪器仪表学报 2013 第 9 期 P2034-2044 0254-3087

关键词：卫星；锂离子电池；剩余寿命预测；可重构计算；故障预测和健康管理

[期刊] 多媒体技术研究：2013：面向智能视频监控的视觉感知与处理 [中文核心期刊(北大)] [统计源期刊(中信所)] [CSCD 中国科学引文库(中科院)] 引证(20)

作者：黄铁军，郑锦，李波，傅慧源，马华东，薛向阳，姜育刚，于俊清（北京大学信息科学技术学院；北京航空航天大学计算机学院；北京邮电大学计算机学院；复旦大学计算机科学技术学院；华中科技大学计算机学院）

出处：中国图象图形学报 2014 第 11 期 P1539-1562 1006-8961

关键词：视频监控；目标检测；目标跟踪；视频增强；行为识别

[期刊] 基于同义词词林的文本特征选择方法 [中文核心期刊(北大)] [统计源期刊(中信所)] [CA 化学文摘(美)] [CSCD 中国科学引文库(中科院)] 引证(20)

作者：郑艳红，张东站（厦门大学信息科学与技术学院）

出处：厦门大学学报(自然科学版) 2012 第 2 期 0438-0479

关键词：特征选择；同义词词林；遗传算法；文本分类

[学位论文] 基于半监督集成支持向量机的土地覆盖遥感分类方法研究 引证(20)

作者：刘颖（中国科学院研究生院(东北地理与农业生态研究所)）

学位名称：博士

出处：中国科学院研究生院(东北地理与农业生态研究所) 2013

关键词：支持向量机；自适应变异粒子群；半监督学习；Self-training；Gustafson-kessel；集成学习；土地覆盖；遥感分类

[学位论文] 细粒度情感分析研究 引证(20)

作者：施寒潇（苏州大学）

学位名称：博士

出处：苏州大学 2013

关键词：情感分析；意见挖掘；情感强度量化；语义角色标注；统计机器学习；自然语言处理

[学位论文] 粗糙集理论中的知识获取与约简方法的研究 引证(20)

作者：张明（南京理工大学）

学位名称：博士

出处：南京理工大学 2012

关键词：粗糙集；描述子；多粒度；决策规则；约简；知识获取；度量

[学位论文] 基于本体的自适应 Web 信息抽取方法研究 引证(20)

作者：李传席（中国科学技术大学）

学位名称：博士

出处：中国科学技术大学 2012

关键词：信息获取；信息抽取；本体学习；自适应信息抽取；领域资源发现

关系抽取；基于本体的信息抽取

三、结语

机器学习算法学术论文集中在哪些高校，各个高校的优势方向如何，纲举目张，一目了然，为我们从事机器学习算法产线的快速原型构造做了比较好的指引。

四、参考文献

- [1] 超星发现系统[EB/OL].<http://www.chaoxing.com/>
- [2] 百度词条