

# 基于群决策的 P2P 借贷信用风险评估<sup>①</sup>



姜雪莹, 秦 进

(中国科学技术大学 管理学院, 合肥 230026)

**摘 要:** 运用群决策方法, 对随机森林、神经网络、梯度提升树三种算法所生成的个体学习器进行集成, 构建基于群决策的 P2P 借贷信用风险评估模型. 选取人人贷、拍拍贷的数据进行实验研究, 结果显示, 集成模型的风险评估效果较个体学习器有所提升, 且优于传统的逻辑回归方法.

**关键词:** 群决策; P2P 借贷; 信用风险评估; 机器学习; 集成

引用格式: 姜雪莹, 秦进. 基于群决策的 P2P 借贷信用风险评估. 计算机系统应用, 2019, 28(5): 226-231. <http://www.c-s-a.org.cn/1003-3254/6901.html>

## Group Decision-Making Method for Credit Risk Assessment in P2P Lending

JIANG Xue-Ying, QIN Jin

(School of Management, University of Science and Technology of China, Hefei 230026, China)

**Abstract:** In this study, we propose a combination approach based on group decision-making method, using random forest, neural network and GBDT as individual learners, to assess credit risk of borrowers in P2P lending. To validate the proposed method, two real-world datasets from PPDai.com and renrendai.com are examined. The results show that, compared with the individual learners, the proposed method has made a better performance.

**Key words:** group decision-making; P2P lending; credit risk assessment; machine learning; ensemble

### 1 引言

P2P 借贷是指个人用户之间借助专业的互联网借贷平台进行的小额借贷交易. 近年来, P2P 借贷行业在中国发展迅猛, 网贷之家的数据显示, 2017 年国内 P2P 借贷交易额达 2.8 万亿元, 较 2016 年增长超过 40%, 活跃投资人数达 440 万人. 为维持行业健康发展, 需进行有效的风险控制.

借款人信用风险是指借款人未在约定期限内还本付息的风险, 即违约风险. 由于 P2P 借贷不经由传统金融机构, 个体投资者在面临信息不对称问题的同时缺乏规避风险的专业能力, 因此, 如何合理评估借款人信用风险是一个重要的研究问题<sup>[1]</sup>. 信用风险评估问题旨在预测借款项目的违约概率. 对于已到达还款期限的项目, 称未按时还本付息的项目为违约项目, 违约概率

记为 1, 按时还本付息的项目为不违约项目, 违约概率记为 0. 将项目违约概率作为被解释变量, 从到达还款期限的项目数据中训练得到回归模型, 可预测新发布的借款项目的违约概率, 从而为平台管理与投资者投资提供参考.

为解决 P2P 借贷的信用风险评估问题, 诸多机器学习算法被用于预测借款项目的违约概率, 研究表明随机森林 (Random Forest)<sup>[2]</sup>, 神经网络 (Neural Network)<sup>[3]</sup>、梯度提升树 (GBDT)<sup>[4]</sup>等算法具有较高的预测准确度. 由于 P2P 借贷信用风险评估问题涉及的借款项目特征复杂, 且呈非线性关系, 在不同的数据环境下, 运用单一算法会受到不同数据特征的干扰<sup>[5]</sup>. 而多种算法之间分析数据特征的角度存在差异, 通过将多种算法进行集成, 能够达成互补, 使集成效果优于单

<sup>①</sup> 基金项目: 国家自然科学基金面上项目 (71571175)

Foundation item: General Program of National Natural Science Foundation of China (71571175)

收稿时间: 2018-11-25; 修改时间: 2018-12-12, 2018-12-20; 采用时间: 2018-12-26; csa 在线出版时间: 2019-05-01

一算法<sup>[6]</sup>. Xia等提出,将优质算法引入集成模型中,有助于提升模型预测的准确度<sup>[5]</sup>.因此,考虑对多种优质信用风险评估模型进行集成.群决策是使群体中个体观点达成一致的决策过程<sup>[7]</sup>,在诸多领域中得到广泛应用<sup>[8,9]</sup>.本文基于群决策思想,选取在P2P借贷信用风险评估问题中表现优异的随机森林、神经网络、梯度提升树算法,分别构建信用风险评估模型,并运用群决策方法对以上三个模型进行集成,通过各模型间的优势互补,为P2P借贷信用风险评估提供更为准确的方法.为验证模型效果,选取人人贷、拍拍贷数据进行实证研究.

## 2 基于群决策的P2P借贷信用风险评估算法及模型构建

一个群体中,每个个体对于特定事件均有自己的主观观点,使群体中个体观点达成一致的决策过程,称为群决策.在群决策过程中,每个个体在其他个体观点的基础上对自身观点进行修正,在一系列迭代修正后,群体中所有个体观点趋于一致且不再变动,从而形成群体的观点.运用群决策方法构建P2P借贷信用风险评估模型,具有以下特性:第一,交互性,群体中所有个体间通过观点的交互影响,实现优势互补;第二,动态性,群决策过程针对不同样本独立进行,从而对不同情况形成更有针对性的结论;第三,异质性,运用群决策方法,能够实现不同种类算法间的集成.

### 2.1 基于群决策的P2P借贷信用风险评估集成算法

在群决策过程中,个体的主观观点会受到其他个体的影响而发生改变,从而令群体中所有个体的观点达成一致<sup>[10]</sup>.假设现有 $N$ 个信用风险评估模型 $M_1, M_2, \dots, M_N$ ,分别运用不同的机器学习算法从训练数据中产生,称为个体学习器.第 $i$ 个个体学习器 $M_i$ 对测试集中借款项目违约概率的预测值为 $P_i$ .构建基于群决策的集成模型,关键在于衡量个体学习器 $M_i$ 的预测值 $P_i$ 受到其余个体学习器预测值的影响.

记 $P_i^*$ 为受到其余个体学习器预测值的影响后, $M_i$ 的修正预测值,取 $P_i^*$ 为所有个体学习器预测值的线性组合,即:

$$P_i^* = \sum_{j=1}^N w_{ij} P_j \quad (1)$$

其中, $w_{ij}$ 表示个体学习器 $M_j$ 对个体学习器 $M_i$ 的影响权

重, $\sum_{j=1}^N w_{ij} = 1$ .设 $W$ 为 $N^*N$ 大小的矩阵,矩阵中的元素为 $w_{ij}(i=1, \dots, N, j=1, \dots, N)$ , $P^* = (P_1^*, \dots, P_N^*)'$ , $P = (P_1, \dots, P_N)'$ ,则:

$$P^* = WP \quad (2)$$

为使所有个体学习器的修正预测值达成一致,需迭代进行式(2)所示修正过程,直至所有个体学习器的修正预测值保持一致且不再变动.由于 $W$ 是 $N^*N$ 的转移矩阵,可将其视为马尔科夫链的一步转移概率矩阵.因此,所有个体学习器的修正预测值达成一致,当且仅当存在向量 $\pi = (\pi_1, \dots, \pi_N)$ ,满足:

$$\begin{aligned} \pi W &= \pi \\ \sum_{i=1}^N \pi_i &= 1 \end{aligned} \quad (3)$$

解出向量 $\pi$ 的值,最终得到所有个体学习器的集成结果 $R$ :

$$R = \sum_{i=1}^N \pi_i P_i \quad (4)$$

集成结果 $R$ 即为集成模型对借款项目违约概率的预测值.

上述过程中,一个重要的问题是如何获取个体学习器 $M_j$ 对个体学习器 $M_i$ 的影响权重 $w_{ij}$ .参考Basir等<sup>[11]</sup>的研究,设置个体学习器的不确定度,将不确定度分为两种类型:局部不确定度和全局不确定度.其中,局部不确定度衡量的是个体学习器自身决策不确定的程度,全局不确定度衡量的是个体学习器在受到另一个个体学习器的影响时决策的不确定程度.

个体学习器 $M_i$ 的违约概率预测值为 $P_i$ ,则不违约概率预测值为 $1 - P_i$ .记个体学习器 $M_i$ 的局部不确定度为 $U_{ii}$ , $U_{ii}$ 满足:

$$U_{ii} = -P_i \log_2(P_i) - (1 - P_i) \log_2(1 - P_i) \quad (5)$$

当 $P_i$ 趋向0或1时,个体学习器 $M_i$ 对借款项目违约与否的判定清晰,此时不确定度 $U_{ii}$ 趋向于0.当 $P_i$ 趋向0.5时,个体学习器 $M_i$ 对借款项目违约与否的判定近似随机,此时不确定度 $U_{ii}$ 趋向于1.因此,局部不确定度 $U_{ii}$ 能够反映个体学习器自身决策不确定的程度.

记已知个体学习器 $M_j$ 的违约概率预测值的情况下,个体学习器 $M_i$ 的违约概率预测值为 $P_{ij}$ ,则不违约概率预测值为 $1 - P_{ij}$ .记个体学习器 $M_i$ 的全局不确定度为 $U_{ij}$ , $U_{ij}$ 满足:

$$U_{ij} = -P_{ij} \log_2(P_{ij}) - (1 - P_{ij}) \log_2(1 - P_{ij}) \quad (6)$$

与局部不确定度类似,全局不确定度 $U_{ij}$ 能够反映个体学习器在受到另一个个体学习器的影响时决策的不确定程度。

$P_{ij}$ 表示的是个体学习器 $M_i$ 在个体学习器 $M_j$ 影响下的违约概率预测值,取 $P_{ij}$ 为 $P_i$ 与 $P_j$ 的线性组合,即:

$$P_{ij} = P_j I_{ij} + P_i (1 - I_{ij}) \quad (7)$$

取 $I_{ij}$ 为 sigmoid 函数,即:

$$I_{ij} = \frac{1}{1 + e^{-(Acc_j - Acc_i)}} \quad (8)$$

其中, $Acc_i$ 为个体学习器 $M_i$ 的预测准确率。

求得局部不确定度与全局不确定度后,希望给不确定度低的个体学习器分配较高的权重,给不确定度高的个体学习器分配较低的权重.因此,设置规划问题,

$$\begin{cases} \min z_i = \sum_{j=1}^N w_{ij}^2 U_{ij}^2 \\ \sum_{j=1}^N w_{ij} = 1 \end{cases} \quad (9)$$

其中, $w_{ij}$ 是个体学习器 $M_j$ 对个体学习器 $M_i$ 的影响权重.采用拉格朗日乘子法,令:

$$L_i = \sum_{j=1}^N w_{ij}^2 U_{ij}^2 - \rho (\sum_{j=1}^N w_{ij} - 1) \quad (10)$$

求 $L_i$ 对 $w_{ij}$ 的偏导并令结果等于0,结合 $\sum_{j=1}^N w_{ij} = 1$ ,得到 $w_{ij}$ 的表达式:

$$w_{ij} = \frac{1}{U_{ij}^2 \sum_{k=1}^N U_{ik}^{-2}} \quad (11)$$

由此获取个体学习器 $M_j$ 对个体学习器 $M_i$ 的影响权重 $w_{ij}$ ,随后根据式(3)、(4),即可得到所有个体学习器的集成结果 $R$ .

## 2.2 基于群决策的P2P借贷信用风险评估模型构建过程

假设现有 $N$ 种机器学习算法,基于群决策的P2P借贷信用风险评估模型的构建过程,描述如下:

(1) 分别运用 $N$ 种机器学习算法,在训练数据中训练出个体学习器 $M_1, M_2, \dots, M_N$ ,并得到个体学习器的预测准确率 $Acc_1, \dots, Acc_N$ .

(2) 应用个体学习器 $M_1, M_2, \dots, M_N$ ,对测试集中借款项目的违约概率进行预测,预测值为 $P_1, \dots, P_N$ .

(3) 运用式(5)求得个体学习器的局部不确定度,运用式(6)–(8)求得个体学习器的全局不确定度。

(4) 运用式(11)求得权重 $w_{ij}(i=1, \dots, N, j=1, \dots, N)$ .

(5) 将 $w_{ij}$ 代入式(3),解得向量 $\pi$ 的值。

(6) 运用式(4),最终得到所有个体学习器的集成

结果 $R$ .

集成结果 $R$ 表示集成模型对借款项目违约概率的预测,若 $R > 0.5$ ,预测借款项目违约,若 $R \leq 0.5$ ,预测借款项目不违约。

## 2.3 个体学习器描述

选取随机森林(Random Forest)、神经网络(Neural Network)、梯度提升树(GBDT)三种机器学习算法,分别构建回归模型,作为群决策集成模型的个体学习器.在P2P借贷信用风险评估问题中,以P2P借贷借款项目数据集作为个体学习器的训练数据集,选取借款项目及借款人的信息作为回归模型的解释变量,借款项目违约概率作为回归模型的被解释变量,由此构建的回归模型,能够用于新发布的借款项目的违约概率预测.以下是对随机森林、神经网络、梯度提升树三种算法的简要介绍。

### 2.3.1 随机森林

随机森林(Random Forest)是解决回归问题常用的算法,它以一系列未剪枝的决策树作为基学习器,在构建 bagging 集成的基础上,于决策树的训练过程中引入变量的随机选择思想<sup>[12]</sup>.决策树的训练过程是特征空间的递归分区过程,在每个非叶节点,决策树从当前节点的变量集合中选取最优变量对数据集进行划分.而在随机森林中,决策树每个节点的最优变量是从变量集合的一个随机子集中选取的.随机森林中的决策树,建立在用 bootstrap 采样法从训练集中选取的样本集合上.在基于不同采样构建一系列决策树后,随机森林采用平均法对决策树输出的结果进行整合.随机森林能够快速并行地处理高维度的数据,在P2P借贷信用风险评估中,借款项目及借款人的特征数量较多,运用随机森林能够减轻特征选择的负担,对特征进行综合分析.同时,随机森林对于缺失数据具有较强的抗干扰能力,在P2P借贷中,存在借款人信息不全的情况,而随机森林能够在数据缺失的环境下得出较为准确的结果。

### 2.3.2 神经网络

神经网络(Neural Network)是模拟生物神经网络构建的并行交互系统,它具有多种学习算法,其中使用误差逆传播算法训练的单隐层前馈神经网络(简称BP神经网络),是最具有代表性的神经网络算法之一. BP神经网络分为三层:输入层、隐层、输出层,每一层由若干神经元构成,且每层神经元与下一层所有神



经元相互连接. 输入层神经元接收外界输入, 并将信号传递至隐层. 隐层神经元接收到上层神经元输出值的加权平均, 将其与阈值相比较, 并通过激活函数产生神经元输出, 传递至输出层. 输出层同样进行信号加工, 并生成最终的输出值. 神经网络训练的过程, 就是确定神经元间连接权重以及神经元阈值的过程. 根据神经网络输出值的误差, 基于梯度下降策略, 可以对参数进行调整, 并重新进行加权平均, 如此进行迭代学习, 直至误差处于允许范围之内, 即可得到最终的神经网络模型<sup>[13]</sup>. 神经网络具有较强的非线性映射能力, 能够拟合 P2P 借贷中借款项目及借款人的信息与借款项目违约概率间的非线性关系. 同时, 神经网络具有自学习与自适应的能力, 在 P2P 借贷借款项目信息多样化的情况下, 神经网络对新产生的借款项目具有较好的解释能力.

### 2.3.3 梯度提升树

梯度提升树 (GBDT) 是一种重要的 boosting 算法, 以分类回归树 (CART) 作为弱学习器, 通过迭代将若干弱学习器集成为一个强学习器. GBDT 采用了前向分布算法, 其核心思想是, 每一轮迭代中构建的弱学习器, 建立在上一轮迭代集成的强学习器的损失函数的梯度下降方向. 构建 GBDT 模型, 首先要根据初始训练集训练一个基学习器  $f_1(x)$ , 随后进行以下迭代过程. 假设前一轮迭代得到的强学习器为  $f_{i-1}(x)$ , 损失函数为  $L(y, f_{i-1}(x))$ , 本轮迭代的目标是寻找弱学习器  $h_i(x)$ , 使得本轮迭代得到的强学习器  $f_i(x) = f_{i-1}(x) + h_i(x)$  满足

损失函数  $L(y, f_i(x))$  最小化. 由于损失函数的负梯度可作为损失的近似值, 因此通过构造决策树对损失函数的负梯度进行拟合即可得到弱学习器  $h_i(x)$ . GBDT 可以灵活处理多种类型的数据, 对于异常值的鲁棒性较强, 在 P2P 借贷较为复杂的数据环境下, GBDT 具有较高的预测准确率.

## 3 实验分析

### 3.1 实验数据及变量描述

选取中国具有代表性的两家 P2P 平台人人贷与拍拍贷的数据, 对基于群决策的 P2P 借贷信用风险评估模型的效果进行实验验证. 其中, 从人人贷选取 2012 年 1 月至 2015 年 12 月间发布的借款项目 15 912 个, 从拍拍贷选取 2015 年 1 月到 2017 年 1 月间发布的借款项目 50 167 个. 将借款项目的违约概率作为被解释变量, 两个平台借款项目的违约状况分布如表 1 所示.

表 1 人人贷、拍拍贷借款项目违约状况分布

平台	项目总数	违约		不违约	
		项目数	占比 (%)	项目数	占比 (%)
人人贷	15 912	3076	19.33	12 836	80.67
拍拍贷	50 167	2558	5.10	47 609	94.90

从平台提供的信息中选取解释变量, 解释变量分为三类: 借款项目信息、借款人身份信息、借款人历史借款信息. 表 2 展示了从人人贷与拍拍贷提供的信息中筛选出的解释变量.

表 2 人人贷、拍拍贷解释变量汇总

	人人贷	拍拍贷
借款项目信息	借款金额、年利率、还款期限、信用等级、信用额度	借款金额、年利率、还款期限、信用等级
借款人身份信息	年龄、学历、婚姻状况、收入、房产、车产、房贷、车贷、公司规模、工作时间	年龄、性别、手机认证、户口认证、视频认证、学历认证、征信认证、淘宝认证
借款人历史借款信息	借款次数、逾期次数、还清次数	成功借款次数、成功借款金额、待还本金、历史正常还款期数、历史逾期还款期数

### 3.2 实验结果

构建群决策集成模型, 首先需要运用随机森林、神经网络及梯度提升树算法, 分别构建三个信用风险评估个体学习器. 在个体学习器构建过程中, 一些重要参数的设置对于个体学习器的预测准确性存在一定影响, 需要设置的参数包括随机森林模型中决策树的数量  $n$ , 随机森林中决策树节点随机选择特征的数量  $m$ ,

神经网络隐层中神经元的数量  $s$ , 神经网络的最大迭代次数  $d$ , 梯度提升树中决策树的数量  $t$ . 通过实验, 对比设置参数值不同的情况下模型的预测准确性, 可确定参数值. 根据实验结果, 取  $n=500$ ,  $m=4$ ,  $s=8$ ,  $d=200$ ,  $t=5000$ .

在构建个体学习器后, 依据 2.2 节中基于群决策的 P2P 借贷信用风险评估模型构建过程, 可以对个体

学习器进行集成. 将三种个体学习器的结果与集成结果进行比较, 同时运用信用风险评估的传统方法逻辑回归算法构建模型作为比较的基准, 验证基于群决策的 P2P 借贷信用风险评估模型的效果.

为评判信用风险评估模型的预测效果, 若模型预测的借款项目违约概率大于 0.5, 预测借款项目为违约项目, 若模型预测的借款项目违约概率小于等于 0.5, 预测借款项目为不违约项目. 选取五个指标评判模型的预测效果: 准确率、AUC、召回率、精确率、F1 值. 准确率指模型正确预测的样本数占总样本数的比例. AUC 指随机挑选一个违约项目和一个不违约项目, 违约项目预测值高于不违约项目预测值的概率. 召回率指实际不违约的项目中被正确预测为不违约的比例. 精确率指被预测为不违约的项目中预测正确的比例. F1 值是召回率和精确率的调和平均值, 用于综合召回率和精确率对模型进行评价.

表 3 采用人人贷数据构建信用风险评估模型结果

模型	准确率 (%)	AUC	召回率 (%)	精确率 (%)	F1 值
逻辑回归	89.00	0.9318	93.57	92.84	0.9320
随机森林	90.22	0.9458	95.95	92.24	0.9406
神经网络	90.76	0.9590	95.52	93.19	0.9434
梯度提升树	89.44	0.9327	98.01	98.83	0.9374
群决策集成模型	91.26	0.9613	96.81	92.70	0.9470

由于拍拍贷平台的数据存在严重的类别不平衡问题, 即不违约项目的数据量远大于违约项目的数据量. 运用类别不平衡数据构建模型对项目违约状况进行预测, 会导致大部分违约项目被错误预测为不违约项目, 且模型评判指标虚高. 因此, 在采用拍拍贷数据构建信用风险评估模型前, 运用随机欠采样方法, 使得训练集中违约项目与不违约项目的比例为 1:1, 据此可验证模型在类别平衡数据下的效果. 表 4 展示了采用拍拍贷数据构建信用风险评估模型的结果. 根据表 4 数据, 群决策集成模型的准确率、AUC、召回率均高于其余模型, 精确率在相比较的模型中也处于较高位置. 且综合精确率与召回率, 群决策集成模型的 F1 值较其他模型更高.

表 4 采用拍拍贷数据构建信用风险评估模型结果

模型	准确率 (%)	AUC	召回率 (%)	精确率 (%)	F1 值
逻辑回归	29.65	0.6410	26.43	97.90	0.4162
随机森林	72.15	0.6128	73.94	95.81	0.8347
神经网络	69.50	0.6426	70.86	95.97	0.8153
梯度提升树	70.90	0.6393	72.25	96.18	0.8252
群决策集成模型	72.89	0.6461	74.59	96.00	0.8395

## 4 结语

基于群决策方法, 对随机森林、神经网络、梯度提升树三种算法进行集成, 用于评估 P2P 借贷中借款项目的信用风险. 运用群决策方法进行集成, 能够通过不同算法评估结果间的相互影响, 实现算法间的优势互补, 从而提升评估效果. 选取人人贷、拍拍贷平台的数据进行实验验证. 考虑到拍拍贷的数据存在严重的类别不平衡问题, 对拍拍贷数据的训练集进行随机欠采样, 使得训练集达到类别平衡. 数据研究结果显示, 集成算法的信用风险评估效果相比于单一算法有所提升, 且优于传统的逻辑回归算法. 因此, 基于群决策的 P2P 借贷信用风险评估算法, 有助于平台及投资者更好的识别风险项目, 以便进一步提升风险控制能力.

## 参考文献

- Serrano-Cinca C, Gutiérrez-Nieto B, López-Palacios L. Determinants of default in P2P lending. *PLoS One*, 2015, 10(10): e0139427. [doi: 10.1371/journal.pone.0139427]
- Malekipirbazari M, Aksakalli V. Risk assessment in social lending via random forests. *Expert Systems with Applications*, 2015, 42(10): 4621–4631. [doi: 10.1016/j.eswa.2015.02.001]
- Byanjankar A, Heikkilä M, Mezei J. Predicting credit risk in peer-to-peer lending: A neural network approach. *Proceedings of 2015 IEEE Symposium Series on Computational Intelligence*. Cape Town, South Africa. 2015. 719–725.
- Xia YF, Liu CZ, Liu NN. Cost-sensitive boosted tree for loan evaluation in peer-to-peer lending. *Electronic Commerce Research and Applications*, 2017, 24: 30–49. [doi: 10.1016/j.elerap.2017.06.004]
- Xia YF, Liu CZ, Da BW, *et al.* A novel heterogeneous ensemble credit scoring model based on bstacking approach. *Expert Systems With Applications*, 2018, 93: 182–199. [doi: 10.1016/j.eswa.2017.10.022]
- Choudhury AK, Shankar R, Tiwari MK. Consensus-based intelligent group decision-making model for the selection of advanced technology. *Decision Support Systems*, 2006, 42(3): 1776–1799. [doi: 10.1016/j.dss.2005.05.001]
- Zhao M, Ma XY, Wei DW. A method considering and adjusting individual consistency and group consensus for group decision making with incomplete linguistic preference relations. *Applied Soft Computing*, 2017, 54: 322–346. [doi: 10.1016/j.asoc.2017.01.010]

- 8 贾子文, 顾煜炯, 邢月, 等. 改进专家群决策方法的风电机组故障风险评价研究. 可再生能源, 2018, 36(3): 453–460. [doi: [10.3969/j.issn.1671-5292.2018.03.020](https://doi.org/10.3969/j.issn.1671-5292.2018.03.020)]
- 9 吴胜, 李延来, 陈振颂. 群决策方法及其在供应商选择中的应用. 计算机仿真, 2018, 35(3): 184–189. [doi: [10.3969/j.issn.1006-9348.2018.03.041](https://doi.org/10.3969/j.issn.1006-9348.2018.03.041)]
- 10 Degroot MH. Reaching a consensus. Journal of the American Statistical Association, 1974, 69(345): 118–121. [doi: [10.1080/01621459.1974.10480137](https://doi.org/10.1080/01621459.1974.10480137)]
- 11 Basir OA, Shen HC. New approach for aggregating multi-sensory data. Journal of Robotic Systems, 1993, 10(8): 1075–1093. [doi: [10.1002/rob.v10:8](https://doi.org/10.1002/rob.v10:8)]
- 12 Breiman L. Random forests. Machine Learning, 2001, 45(1): 5–32. [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
- 13 Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer, 2009.

[www.c-s-a.org.cn](http://www.c-s-a.org.cn)

[www.c-s-a.org.cn](http://www.c-s-a.org.cn)