

doi: 10.7690/bgzdh.2021.02.019

# 基于 XGboost 算法的面向任务携行航材品种确定方法

宋传洲<sup>1</sup>, 陈育良<sup>1</sup>, 杨宜霖<sup>2</sup>

(1. 海军航空大学岸防兵学院, 山东 烟台 264001; 2. 中国人民解放军 94679 部队, 南京 210000)

**摘要:** 为解决传统选取航材品种方法对人的经验依赖较多、准确率不高、运算效率低等问题, 提出一种采用 XGboost 算法的携行航材品种选择方法。建立分类特征体系, 使用 XGboost 算法对特征进行重要性排序、分析和筛选, 构建精简版分类特征体系, 使用 K 折交叉验证法和经验调参对样本数据分组和训练, 并与 GBDT、RF、Adaboost 等分类算法的结果比析。结果表明: XGboost 算法可减少人为因素干预, 在携行航材品种确定应用中具有高效性、科学性和优越性。

**关键词:** 携行航材; 品种确定; XGboost; 分类; 面向任务

中图分类号: E075; TP202 文献标志码: A

## A Method to Determine Type of Spare Parts for Mission-oriented Aircraft Based on Xgboost

Song Chuanzhou<sup>1</sup>, Chen Yuliang<sup>1</sup>, Yang Yilin<sup>2</sup>(1. School of Coast Guard, Navy Aviation University, Yantai 264001, China;  
2. No. 94679 Unit of PLA, Nanjing 210000, China)

**Abstract:** In order to solve these problems of traditional selecting aircraft spare parts method, such as more relying on human experience, low accuracy and low operational efficiency, a new method of selecting aircraft spare parts by using XGboost algorithm was proposed. The classification feature system was established. The XGboost algorithm was used to sort, analyze and screen the features, and the simplified version of the classification feature system was constructed. The k-fold cross validation method and empirical reference were used to group and train the sample data, and the results were compared with GBDT, RF, Adaboost and other classification algorithms. The results show that the XGboost algorithm can reduce the interference of human factors, and it is efficient, scientific and optimal.

**Keywords:** portable aircraft spare parts; variety determination; XGboost; classification; mission oriented

## 0 引言

航材携行保障是航材保障工作的重要内容之一。航材携行保障可分为舰载机携行保障、专项任务携行保障、转场携行保障、战时携行保障等<sup>[1]</sup>。各机种在执行任务期间, 航材的消耗情况受任务等级、周围环境、工作强度、储存空间、维修能力等很多因素影响。为此, 需针对某机种的消耗特点进行研究, 并结合任务情况来确定需携带航材的品种和数量。

在以往把品种确定视作分类方法的研究中, 侧重于品种确定的方法有 AHP 评价法<sup>[2]</sup>、模糊综合评判法<sup>[3]</sup>、灰色局势决策法<sup>[4]</sup>、粗糙集理论<sup>[5 6]</sup>、主成分分析法<sup>[7]</sup>、Topsis 评价法<sup>[8]</sup>等。这些方法都是分析影响器材消耗的各种因素, 把不容易量化的部分数值化, 合并降维的同时建立指标评价体系并赋予相

应权重, 从而确定备件品种。还有一些组合方法如模糊灰度关联法<sup>[9]</sup>、FAHP<sup>[10]</sup>、DHGF<sup>[11]</sup>等。这些方法对人的经验要求较高, 各影响因素间的相互作用容易被忽略, 随着航材种类增多, 模型运算效率较低, 工作量剧增, 且准确性不高。

XGboost 算法是以 Boosting 的集成学习为基本思想提出的一种提升算法, 近年来在许多领域的应用中都获得良好成效<sup>[12]</sup>。与传统算法相比, XGboost 算法先进性表现在预测准确度、运算效率、算法修正的容错性等诸多方面, 可根据实际情况给出的不同条件有针对性地进行算法调优, 特别是对其超参数的优化。笔者在对模型参数进行优化的同时, 使用 K 折交叉验证法对可靠样本数据进行分组和训练, 并与 GBDT、RF、Adaboost、BP、KNN、SVM 等机器学习分类算法进行对比分析, 验证了该算法的实用性、高效性、科学性和优越性。

收稿日期: 2020-08-30; 修回日期: 2020-10-16

作者简介: 宋传洲(1989—), 男, 山东人, 硕士, 从事航空器材、仓储勤务技术研究。E-mail: 792868502@qq.com。

## 1 基本理论

### 1.1 XGboost 的基本理论

XGboost 算法<sup>[13]</sup>是集成多个不同弱分类器的增强分类器，通过对其损失函数进行 2 阶泰勒展开。与通常只采取 1 阶导数进行优化不同，该算法优化过程将同时保留 1 阶导数和 2 阶导数的综合信息，以缩减模型在训练集上的收敛时间。此外，引入列抽样和在损失函数中加入可优化的正则项，既能有效控制模型复杂度，又可避免出现过拟合。其原理如下：

设  $D = \{(x_i, y_i)\} (|D|=n, x_i \in R^d, y_i \in R)$ 。其中， $n$  为航材样本数，每件航材包含特征数为  $d$ ， $x_i$  为第  $i$  个航材的取值。这个树的集成模型是通过持续添加新的决策树，利用残差多次迭代拟合得到终值，即树的数目为  $K$  的相加函数就是最终预测的结果：

$$\left. \begin{aligned} \hat{y}_i &= \phi(x_i) = \sum_{k=1}^K f_k(x_i), \quad f_k \in F \\ F &= \left\{ f(x) = \omega_{q(x)} \right\} (s: R^d \rightarrow T, \quad \omega \in R^T) \end{aligned} \right\}. \quad (1)$$

式中： $\phi$  为样本  $x_i$  和预测值  $\hat{y}_i$  的函数关系； $F$  为所有分类树的集合； $f_k$  为第  $k$  棵树的函数模型； $s$  为  $x_i$  相应叶节点序列所标记的映射； $T$  为叶子节点的数目； $R_T$  为叶子节点权重  $\omega$  的空间，每一个独立  $\omega_{q(x)}$  映射得到一个  $f(x)$ ，而每一个  $f_k$  都有且只有一个对应的  $q$ （树结构）和  $\omega$ 。

为了学习函数集，XGboost 模型中在目标函数中加入正则化后得到损失函数：

$$\left. \begin{aligned} L(\phi) &= \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \\ \Omega(f) &= \gamma T + \lambda \|\omega\|^2 / 2 \end{aligned} \right\}. \quad (2)$$

式中： $l$  为可微凸损失函数，用以表示  $\hat{y}_i$  和  $y_i$  之间的差异； $\gamma$  和  $\lambda$  分别为叶子数目和叶子权重的正则化参数， $\gamma$  的取值限制节点继续分裂；令  $\Omega(f)$  成为惩罚项，控制模型复杂度，避免过拟合发生。通过加法运算，将式(1)预测值  $\hat{y}_i$  递推展开，若第  $t$  次迭代后，用  $\hat{y}_i^t$  表示第  $i$  个样本点的预测值，可得到前  $t-1$  棵树的预测值是前  $t-1$  棵树的输出值之和  $\hat{y}_i^{(t-1)}$ ，设  $f_t$  为第  $t$  棵树的函数， $f_t(x_i)$  为第  $t$  棵树的输出结果， $\hat{y}_i^{(t-1)} + f_t(x_i)$  相加，得到新的预测值，从而得到最小化损失函数  $L^{(t)}$ ，一个新的第  $t$  棵树的目标函数：

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t). \quad (3)$$

式中  $\hat{y}_i^{(t-1)}$  处用泰勒公式 1 阶 2 阶展开得到：

$$L^{(t)} \approx \sum_{i=1}^n \left[ l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t). \quad (4)$$

式中， $g_i = \partial l(y_i, \hat{y}_i^{(t-1)}) / \partial \hat{y}_i^{(t-1)}$ ， $h_i = \partial^2 l(y_i, \hat{y}_i^{(t-1)}) / \partial^2 \hat{y}_i^{(t-1)}$ ，去掉常数项得到简化的损失函数：

$$L^{(t)} = \sum_{i=1}^n \left[ g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t). \quad (5)$$

令  $I_i = \{i | q(x_i) = k\}$  作为叶子节点  $k$  的样本集，代入式(5)：

$$\left. \begin{aligned} L^{(t)} &= \sum_{i=1}^n \left[ g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 = \\ &\sum_{k=1}^T \left[ (\sum_{i \in I_k} g_i) \omega_k + \frac{1}{2} (\sum_{i \in I_k} h_i + \lambda) \omega_k^2 \right] + \gamma T. \end{aligned} \right\} \quad (6)$$

由于  $q(x)$  固定，可以求得  $\omega_k$ ，进而得到最优目标函数值，但树结构  $q(x)$  未知，不可能一一穷举  $q$ ，通常使用贪心算法进行迭代，分裂叶节点，即在已有的叶子节点划分分支。

### 1.2 评价指标

笔者采用精确度(precision)、召回率(recall)、准确率(accuracy)、 $F_1$  值作为评价指标。公式如下：

$$P=T_p/(T_p+F_p), \quad R=T_p/(T_p+F_n),$$

$$A=(T_p+F_n)/(T_p+F_p+T_n+F_n), \quad F_1=2P \cdot R / (P+R).$$

其中： $T_p$  为需携行航材样本中分类正确的数量； $F_n$  为需携行航材被错误分类的数量； $F_p$  为无需携行航材被分类为需携行的数量； $T_n$  为无需携行航材被分类为无需携行的数量。

## 2 基于 XGboost 分类模型的流程

基于 XGboost 基本原理，结合某机型航材面向任务的使用情况，给出该机型的携行品种推荐基本流程：

1) 结合任务需求对不同架次的某机型航材实际消耗情况采集数据，建立可靠样本数据集，对其进行归一化处理。

2) 采取 XGboost 算法对航材样本及特征参数值进行分析，计算每个特征的重要程度并进行排序，根据实际情况取舍特征，留下对最终航材种类选取影响较大的特征。

3) 将样本分组，设置训练集和测试集，利用 K 折交叉法<sup>[14]</sup>找到最优分组比例，进行数据分组。

4) 参数调优，确定其学习深度、学习率、迭代次数等参数<sup>[15]</sup>。

携行航材品种分类模型流程如图 1 所示。

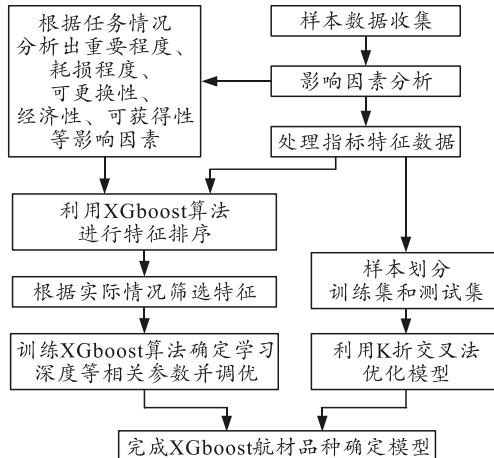


图 1 携行航材品种分类模型基本流程

### 3 建立携行品种的评价特征体系

#### 3.1 评价特征的选取

对携行航材品种的评价特征选取是建立模型的重要部分，关键在于在众多特征中选出重要的特征指标，要能全面正确地反映某次任务中航材的使用情况<sup>[16]</sup>。特征选择和构建时要符合图 2 所示原则。

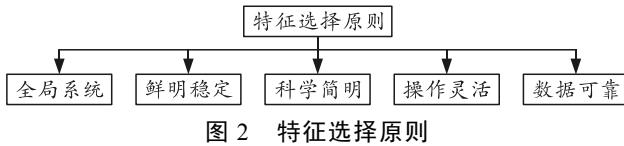


图 2 特征选择原则

按以上原则对航材消耗影响因素分析如下：

1) 通常情况下，失效性多以失效率说明，即执行某任务时某航材平均发生故障的概率，也可通过其平均故障间隔时间表征。该间隔时间指发生故障后经修理和更换零件后仍能正常工作，2 次相邻故障间的工作时间。因为这类数据很难准确获得，故使用失效性统一代替，科学区分等级说明某航材在执行某任务时发生故障的可能。该特征直接反映航材的可靠性，更是判断航材损耗程度的重要特征。

2) 飞机在执行某专项任务、处于特殊环境或维修有一定条件限制等情况下，若发生故障，未知在其维修后是否能正常使用时，需考虑该航材是否可分解为外场可更换单元级别，对其进行直接更换。在可以更换的情况下，还要考虑某一维修级别对损坏单元进行拆卸、安装的能力。此外，更换水平用来反映维修人员的技术水平、配备的维修设备、设施和级别等条件。

3) 任务需求与任务损伤和环境严酷度 2 个特征都有很强的关系，任务需求是表征某航材针对某具体任务的重要程度，执行任务的环境也直接表征该任务外部影响因素，需综合考虑温度、湿度、盐

度、风力、生物作用等因素<sup>[17]</sup>。一次任务中，由于不同航材其本身的环境适应性不同，再加上所处飞机的位置不同，其受到的环境影响也不相同。显然，环境适应性强、可靠性高、处于保护层内的航材受到环境影响较少。如果是在执行特殊的任务，比如有对抗损耗的作战任务，则通过任务损伤表征。

4) 寿命较长的航材发生故障的概率往往较低，且数据很难获得，基于对其消耗数量进行拟合研究规律的方法适用性较差；因此，需采用其生命周期规律来代替历史消耗数据作为依据，结合其他特征，尤其是重要程度，最终判断其是否需要携带。此外，倘若某航材剩余可靠寿命或者平均寿命少于任务时间，在任务中不能限时修复只能更换，那该航材也需要携带。为方便研究，将正在使用的航材区分初级期、成熟期或衰退期来表示这类情况。

5) 由于数值型数据易受量纲影响，购置费、储运费等指标取值往往远大于其他指标，若不进行规范化处理，将会大大降低模型的分类性能。为提高分类学习速率，降低取值较大属性对模型分类效果的影响，笔者对连续数值的购置费和存储保养费划分区间，转化成离散型数值。

6) 可获得性特征主要用来表现某航材失效后是否被替代。通用性使用器件是否是标准件、通用件、专用件来表征，一般特设航材都是专用器件，专用性越强，越需要携带。此外，航材在航材携行空间有限的情况下，可携带的航材在体型上会受到一定的限制，可通过专车专送、邮寄等方式进行运送，也可视为携行航材；因此，需进一步判断某航材在不损坏其性能的情况下能否被运送，即可储运性能。

7) 所有数值均采取归一化处理，使不同的特征取值可进行加权并具可比性，保证模型运行的收敛速度，从而提升模型的分类效率。

使用归一化公式，对序列  $x_1, x_2, \dots, x_n$  进行求值变换：

$$y_i = (x_i - \min_{1 \leq j \leq n} \{x_j\}) / (\max_{1 \leq j \leq n} \{x_j\} - \min_{1 \leq j \leq n} \{x_j\})$$

得到  $y_1, y_2, \dots, y_n \in [0, 1]$ ，该序列是无量纲序列。

#### 3.2 选取特征

如表 1 所示，根据问卷调查并咨询该领域专家、机务保障人员、航材股保障人员等，结合相关文献资料，最终分析得出重要程度、耗损程度、可更换性、经济性、可获得性等 5 类共 20 种评价指标特征<sup>[18]</sup>。

表 1 影响航材携行选取的特征

一级特征	二级特征	特征说明
重要程度	危害程度	发生故障后对飞机安全以及完成任务的影响程度
	任务需求	某种航材对完成某项任务的影响程度
	保障等级	在某任务中对某航材要求的保障等级
	失效性	执行某任务时某航材发生故障或失效的可能
	可维修性	出现故障后能否限时维修并修复
耗损程度	使用强度	飞机执行任务时航材的使用强度对航材损耗的影响程度
	环境	实际使用环境对航材的影响，需综合考虑温度、湿度、盐度、风力、生物作用等因素
	严酷度	严酷度
	使用水平	技术操作人员的使用情况对航材消耗的影响
	任务损伤	任务中飞机损伤的程度，作战任务中任务损伤可视为战斗损伤
可更换性	历史消耗	通过消耗历史判断航材损耗的可能
	生命周期	航材所处在的使用阶段
	故障可预见	航材发生故障的模式是否可规律地预见
	可更换单元	航材是否可分解为外场可更换单元级别
	更换水平	维修人员技术水平、维修设备、设施和级别等条件
经济性	可拆卸能力	某一维修级别对损坏单元进行拆卸、安装的能力
	购置费	购买某航材所需的费用
	储运	携行期间存储保养和运输的费用
	保养费	发生故障后是否可被替代
	通用性	航材被携带时是否会受空间限制，也就是航材的体积
可获得性	携行空间	在不损坏航材性能的情况下是否能够被单独运送或邮递
	可储运性	

## 4 算例分析

### 4.1 样本数据和特征取值

现有  $n$  种航材，分别记作  $H\{H_1, H_2, H_3, \dots, H_n\}$ ；

表 3 部分航材特征取值

编号	名称	H01	H02	H03	H04	H05	H06	H07	H08	H09	H10	H11	H12	H13	H14	H15
$A_1$	危害程度	6	7	6	6	6	5	5	5	5	5	3	3	2	2	1
$A_2$	任务需求	4	5	5	5	5	4	5	4	4	4	3	3	3	3	2
$A_3$	保障等级	4	4	4	4	4	4	4	4	4	4	3	3	3	3	2
$A_4$	失效性	3	7	4	4	4	5	5	5	5	5	1	1	1	1	4
$A_5$	可维修性	2	2	2	2	2	2	2	2	2	2	2	1	2	2	1
$A_6$	使用强度	4	4	4	4	4	3	3	4	4	4	4	4	4	4	4
$A_7$	环境严酷度	5	4	4	4	3	3	3	3	3	3	3	3	3	3	4
$A_8$	使用水平	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4
$A_9$	任务损伤	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2
$A_{10}$	历史消耗	1	2	2	2	2	2	2	2	2	2	1	1	1	1	3
$A_{11}$	装备寿命周期	2	2	3	3	3	1	3	2	2	2	1	1	1	2	2
$A_{12}$	故障可预见	1	1	1	1	1	2	2	2	1	1	1	1	1	1	2
$A_{13}$	可更换单元	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
$A_{14}$	更换水平	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4
$A_{15}$	可拆卸能力	3	3	3	3	3	4	4	3	3	3	2	2	3	3	3
$A_{16}$	购置费	5	4	5	4	3	3	2	4	2	5	5	3	2	2	1
$A_{17}$	存储保养费	2	1	1	1	1	2	1	1	1	1	1	3	3	2	1
$A_{18}$	通用性	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1
$A_{19}$	携行空间	1	1	1	1	1	1	2	1	2	2	2	1	2	1	1
$A_{20}$	可储运性能	1	1	1	1	1	1	1	1	1	1	2	2	2	2	1
	是否携带	是	是	是	是	是	是	是	是	是	是	否	否	否	否	否

有  $m$  种特征指标，特征值记作  $A\{A_1, A_2, A_3, \dots, A_m\}$ ，其中  $m=20$ 。特征的取值如表 2 所示。

表 2 特征的取值

代号	特征名称	取值
$A_1$	危害程度	1 一级、2 二级、3 三级、4 四级、5 五级、6 六级、7 七级
$A_2$	任务需求	1 很低、2 低、3 中等、4 高、5 很高
$A_3$	保障等级	1 低级、2 中级、3 高级、4 最高级
$A_4$	失效性	1 超极低、2 极低、3 很低、4 低、5 中等、6 高、7 很高、8 极高、9 超级高
$A_5$	可维修性	1 限时可维修、2 限时不可维修
$A_6$	使用强度	1 很低、2 低、3 中等、4 高、5 很高
$A_7$	环境严酷度	1 一级、2 二级、3 三级、4 四级、5 五级、6 六级
$A_8$	使用水平	1 合格、2 中级、3 高级、4 特级
$A_9$	任务损伤	1 轻度损伤、2 中度损伤、3 重度损伤、4 严重损伤
$A_{10}$	历史消耗	1 低、2 中、3 高、4 很高
$A_{11}$	生命周期	1 初级期、2 成熟期、3 衰退期
$A_{12}$	故障可预见	1 可预见、2 不可预见
$A_{13}$	可更换单元	1 可分解、2 不可分解
$A_{14}$	更换水平	1 一等、2 二等、3 三等、4 四等
$A_{15}$	可拆卸能力	1 很难、2 难、3 中等、4 容易、5 很容易
$A_{16}$	购置费	1 很低、2 低、3 中等、4 很高、5 超高
$A_{17}$	储运费	1 低、2 中等、3 高
$A_{18}$	通用性	1 标准件、2 通用件、3 专用件
$A_{19}$	携行空间	1 空间允许携带、2 空间不允许携带
$A_{20}$	可储运性	1 器件可储运、2 器件不可储运

笔者选取 400 种可能携带的航材，收集 6 次任务的 2 400 个样本数据，包含样本特征共 48 000 个，其中，有 150 种是 6 次任务均携带的航材，250 种根据任务不同并不都携带。部分特征取值如表 3 所示。

#### 4.2 用 K 折交叉法对样本数据分组

笔者采用的计算机配置：Intel® Core™ i7-8750H 处理器，16 G 内存，Ge Force GTX 1060 显卡，6 G 显存，Windows 系统，python 语言，pytorch 架构。采用 K 折交叉法来划分样本数据的训练集和测试集。将 2 400 个样本集随机划分为互不相交的 K 个子集后，取  $K=2 \sim 11$  进行交叉验证。记录每次运行的结果如表 4 所示。显然取值  $K=9$  时，准确率最高，样本最终选择 9 折交叉验证法分组，测试集样本数为 267，训练集样本数为 2 133。

表 4 不同 K 取值下分类准确率对比 %

样本	$K=2$	$K=3$	$K=4$	$K=5$	$K=6$
准确率	97.17	97.85	97.66	97.97	98.03
样本	$K=7$	$K=8$	$K=9$	$K=10$	$K=11$
准确率	97.56	98.01	98.05	97.87	97.65

#### 4.3 特征重要性测试和筛选

笔者使用 XGboost 算法对 20 种特征进行重要性测试，获得如图 3 所示的特征重要性程度对比。

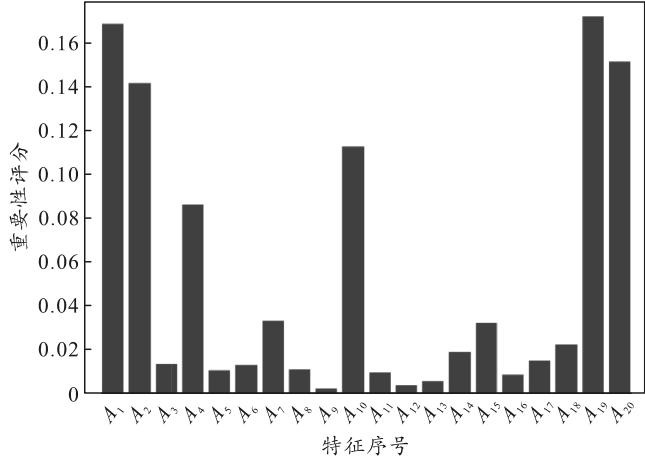


图 3 特征重要性程度对比

从图可知： $A_{19}$  携行空间、 $A_1$  危害程度、 $A_{20}$  可储运性能、 $A_2$  任务需求、 $A_{10}$  历史消耗、 $A_4$  失效率等特征指标依次排在最前， $A_{11}$  装机寿命周期、 $A_{16}$  购置费、 $A_{13}$  可更换单元、 $A_{12}$  故障可预见、 $A_9$  任务损伤等特征指标排在最后。进一步分析特征数量对分类精度的影响情况如表 5 和图 4 所示。如图 5 所示，为提高分类精度和运算速率，删除特征重要性小于 0.01 的特征，即  $A_9, A_{11}, A_{12}, A_{13}, A_{16}$  等 5 个特征。

表 5 特征数量对分类结果的影响 %

特征数量	1	2	3	4	5
准确率	76.41	82.15	90.02	95.45	96.14
特征数量	6	7	10	12	15
准确率	98.01	98.05	98.11	98.15	98.76

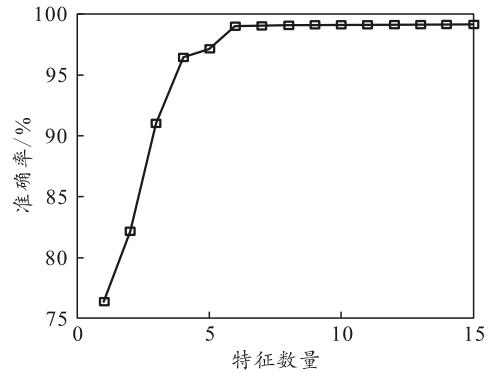


图 4 特征数量对分类准确率的影响

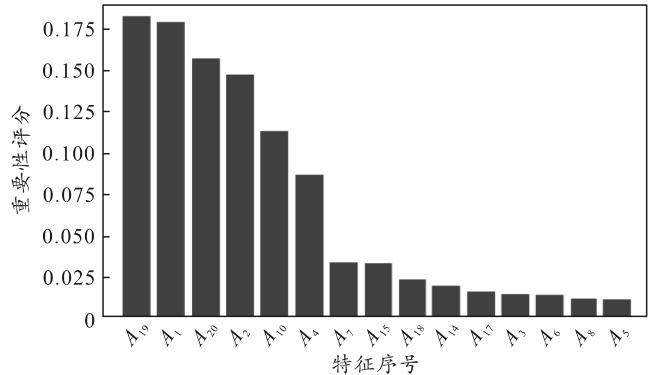


图 5 精简后的特征重要性对比

#### 4.4 XGboost 模型训练、评估和对比实验

对 XGboost 参数使用经验值进行参数优化，得到树的学习率为 0.01，深度为 6，在 1 000 次迭代下模型收敛达到最优值，利用训练好的模型与 GBDT、RF、Adaboost、BP、KNN、SVM 等机器学习分类算法进行实验对比。

经过多次实验得到不同算法结果评估对比如表 6 所示。从表中可以发现：XGboost 模型在 4 个评价指标中的结果是最好的。其中：RF 精确度与 XGboost 相近，都达到 98% 以上，但是在召回率方面与 XGboost 有近 3% 的差距，说明 RF 对于正确样本的分类遗漏较多；GBDT、Adaboost 和 SVM 等也存在这种情况；BP 与 XGboost 的召回率相近，都达到 99% 以上，但是在精确度和准确率方面与 XGboost 相比有一定的差距，说明其误检情况较多；KNN 算法各个指标都比较低；综上说明 XGboost 模型在 7 种模型中分类效果最优。

表 6 不同算法结果评估对比

算法	P/%	R/%	A/%	$F_1$
XGboost	98.56	99.42	98.76	0.989 9
GBDT	97.93	98.11	97.80	0.980 2
Adaboost	97.01	98.87	97.64	0.979 3
RF	98.16	96.55	97.12	0.973 5
BP	94.32	99.23	96.25	0.967 1
KNN	93.67	94.55	93.99	0.941 1
SVM	97.54	95.71	97.23	0.966 2

## 5 结束语

笔者通过科学选取特征评价指标,采用 K 折交叉验证法,研究了基于 XGboost 算法的携行航材品种选择问题,通过使用经验值对模型参数调优,结合航材分类效果的 4 个评价指标,与 GBDT、Adaboost、RF、BP、KNN、SVM 等模型进行对比。实验结果表明:XGboost 算法可有效避免人为因素的干预,在携行航材品种确定上的分类效果和运行效率优于其他机器学习算法。此外,在对特征筛选时,也可使用其他算法进行优化,从而进一步提高该算法的灵活性和分类效率。

## 参考文献:

- [1] 郭峰,王德心.航母舰载机航材携行品种及数量确定方法研究[J].军事运筹与系统工程,2015,29(2): 38–42, 59.
  - [2] 乃超,康锐.备件品种确定方法及决策风险分析[J].兵工自动化,2008(10): 27–30.
  - [3] 季嘉伟,孙国文,罗佳伟.基于模糊综合评判的电源车战场抢修备件品种研究[J].装备制造技术,2017(7): 214–216.
  - [4] 周一鸣,王茜,荣鹏辉.基于灰色局势决策的航空弹药保障装备备件品种确定方法[J].舰船电子对抗,2017,40(5): 56–59, 101.
  - [5] 张帅,唐全国,俞金松.基于属性的舰载机航材备件品种确定方法[J].火力与指挥控制,2015,40(7): 87–91.
  - [6] 韩俊霞,何亚群,李继军.基于混合属性值的航材库存品种确定方法[J].数学的实践与认识,2019,49(19): 139–144.
  - [7] 张闯,刘福胜,单志伟,等.一体化装备保障备件品种确定方法[J].装甲兵工程学院学报,2018,32(3): 30–34.
- 

(上接第 60 页)

- [5] 周克良,周利锋,刘太钢,等.基于改进的 Canny 算子实时视频边缘检测系统在 FPGA 上的设计与实现[J].计算机测量与控制,2016,24(1): 219–222.
- [6] 王雪娇,陈攀.基于 FPGA 的自适应阈值边缘提取的 Canny 算法的实现[J].电子世界,2014(11): 105–106.
- [7] 乔磊.基于 FPGA 的图像边缘检测系统研究与设计[D].淮南:安徽理工大学,2017: 12–13.
- [8] PRAKASH J R. FPGA Implementation of Edge Detection using Canny Algorithm[C]//2015 Online International Conference on Green Engineering and Technologies (IC-GET 2015). IEEE, 2015: 32–35.
- [9] QIAN X, CHAKRABARTI C, KARAM L J. A distributed Canny edge detector and its implementation on

- [8] 金国栋,卢利斌,叶庆.无人机携行备件品种确定方法[J].火力与指挥控制,2008(10): 144–148.
- [9] 张强,刘士通,王开勇.模糊灰色关联分析在战储维修器材品种优化中的应用[J].军事交通学院学报,2016,18(11): 63–67.
- [10] 丁广威,霍武军,张兵.基于 FAHP 的航材库存品种的确定方法[J].四川兵工学报,2013,34(10): 59–62.
- [11] 张作刚,胡新涛,刘望.DHGF 算法在战储航材品种确定中的应用[J].四川兵工学报,2012,33(9): 36–38.
- [12] LI C, ZHENG X, YANG Z, et al. Predicting Short-Term Electricity Demand by Combining the Advantages of ARMA and XGBoost in Fog Computing Environment[J]. Wireless Communications and Mobile Computing, 2018, 6(6): 18–36.
- [13] CHEN T Q, GUESTRIN C. XGBoost: A Scalable Tree Boosting System[J]. Proceedings of the 22nd ACM SIG-KDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA, 2016, 13(2): 785–794.
- [14] TZU-TSUNG W. Parametric Methods for Comparing The Performance of Two Classification Algorithms Evaluated by K-Fold Cross Validation on Multiple Data Sets[J]. Pergamon, 2017, 65(5).
- [15] GERTZ M, GROBÉ-BUTENUTH K, JUNGE W, et al. Using The XGboost Algorithm to Classify Neck And Leg Activity Sensor Data Using on-Farm Health Recordings for Locomotor-Associated Diseases[J]. Computers and Electronics in Agriculture, 2020, 173(6).
- [16] MUCKSTADT J A. Analysis and Algorithms for Service Parts Supply Chains[M]. Beijing: National Defense Industry Press, 2016: 31–42.
- [17] 陈博,徐常凯,任佳成.基于灰色神经网络的携行航材消耗预测[J].指挥信息系统与技术,2018,9(5): 86–90.
- [18] 董晓雄,陈云翔,蔡忠义.基于不完备信息的粗糙集初始备件品种确定方法[J].系统工程与电子技术,2018,40(3): 590–594.

FPGA[C]//Digital Signal Processing Workshop & IEEE Signal Processing Education Workshop. 2011: 25–28.

- [10] 朱顺安.图像边缘检测并行算法的研究和基于 FPGA 的实现[D].武汉:中南民族大学,2013: 17–19.
- [11] 赵安才,周强.基于改进 Canny 算法的实时边缘检测系统设计与硬件实现[J].电子设计工程,2018,26(7): 189–193.
- [12] 杨康.基于 FPGA 的视频图像采集与边缘检测系统设计[D].合肥:安徽大学,2019: 35–36.
- [13] 陈梦洁,张秀梅.CANNY 边缘检测的 FPGA 实现及其改进[J].通信与广播电视台,2016(1): 14–21.
- [14] 隋旭阳,赖文娟,李健.基于 FPGA 的 DDR3 SDRAM 高速图像数据采集方法[J].兵器装备工程学报,2018,39(5): 108–111.