

## 制造技术

蒋佳磊, 廖付, 郝贤伟, 等. 基于多方法优势组合的烟叶配方模块组合分类[J]. 中国烟草学报, 2024, 30(6). JIANG Jialei, LIAO Fu, HAO Xianwei, et al. Classification of tobacco leaf formula modules based on a combination of multiple methodological advantages [J]. Acta Tabacaria Sinica, 2024, 30(6). doi:10.16472/j.chinatobacco.2023.T0105

# 基于多方法优势组合的烟叶配方模块组合分类

蒋佳磊, 廖付, 郝贤伟, 汤晓东, 陈晓水, 朱书秀, 赵振杰\*

浙江中烟工业有限责任公司技术中心, 杭州市西湖区科海路 118 号 310024

**摘要:** 为深入研究卷烟品牌烟叶配方分组方法和原则, 基于 457 个烟叶样品的 12 种感官评吸指标, 比较了 4 种判别分析和 4 种机器学习方法对 4 种烟叶分类的建模集正确率 ( $R$ )、验证集正确率 ( $r$ ) 和平均正确率 ( $m$ ) 的影响, 并基于分类方法选择和权重分配构建了一种高精度的组合分类方法。结果表明: ①与判别分析相比, 机器学习的  $R$  显著提高, 而  $r$  显著下降, 且 LS-SVM 的  $R$  最高 (92.8%), FDA 和 F-BDA 的  $r$  最高 (80.2%), 但  $m$  无显著性差异; ②优化选择 M-BDA、FDA、ANN 和 KNN 四种方法, 按精度赋权建立的组合分类方法同时提高了  $R$  (95.3%) 和  $r$  (89.0%), 且  $m$  由低于 84% 提高到 92.2%, 并通过理论计算和实际结果验证了组合分类方法的普遍有效性; ③组合分类方法 Kappa 系数均大于 0.8, 方法可靠, 一致性程度高, 验证集  $m$ -F1 度量显著提升 21.2%, 模型泛化能力大为增强; ④优雅感、杂气、余味、润感和清晰度 5 项指标对分类起主要作用, 符合利群品牌的风格特征; ⑤误判样品 (6.5%) 指标评分与其模块真实类别的不匹配归因于对库存、成本和质量的平衡, 基本符合烟叶配方的调整空间。

**关键词:** 烟叶; 组合分类; 判别分析; 机器学习; 感官评吸; 配方模块

烟叶质量主要包括外观质量、物理特性、化学成分及感官质量 4 部分, 他们从不同方面反映烟叶品质特征<sup>[1]</sup>。其中化学成分和感官质量主要反映烟叶的内在质量, 是决定卷烟产品的品质与风格的主要因素。《中国烟叶质量白皮书》提出, 感官质量与化学成分按 3:1 的权重约占烟叶质量的 90%<sup>[2]</sup>; 与化学成分相比, 感官质量不仅是烟叶内在品质的直观体现, 也是目前卷烟进行产品配方设计的主要依据<sup>[3]</sup>。烟叶感官质量是指通过烟叶燃烧后产生烟气的特征经由人的抽吸感受进行鉴别, 按照 YC/T 138 包括香型、香气质、香气量、浓度、劲头、杂气、刺激性、余味等指标, 此外, YC/T 530 还包括透发性、细腻感、柔和性、润感等指标, 中烟企业会根据自身品牌的风格特征作一定程度的调整。

在实际配方设计工作中, 中烟企业一般会根据烟叶自身的品质特点和卷烟产品的定位, 将其分为不同

类型、不同用途或不同档次的功能模块, 以充分利用不同烟叶的优势进行优化组合。随着二次配方、模块化加工和分组加工等技术工艺的不断发展和应用, 按照烟叶在叶组配方中的功能对其进行定向模块化分类, 在配方前移、烟叶替代和价值提升中均具有积极意义<sup>[4]</sup>。烟叶感官质量评价与其模块化功能划分研究属于多指标判别分类范畴, 常用的判别分析方法主要有 Mahalanobis 距离判别 (MDA)、Fisher 判别 (FDA) 和 Bayes 判别 (BDA), 在烟草行业各类研究中应用也较为广泛, 如毕淑峰等<sup>[5]</sup>采用 MDA 法建立了烟叶常规化学成分与其品质间的对应关系, 张灵帅等<sup>[6]</sup>采用主成分 (PCA)-MDA 法用于卷烟的真假烟鉴别; 潘玲等<sup>[7]</sup>采用 FDA 法建立了基于常规化学成分、致香物质和物理特性 13 项指标的烤烟香型鉴别模型, 高宪峰等<sup>[8]</sup>基于鲜烟叶颜色值指标采用 FDA 法判断其成熟度; 李超等<sup>[9]</sup>采用 BDA 法建立了中性致香成分、无机元素

**基金项目:** 浙江中烟揭榜挂帅项目: 基于优化烟叶资源使用效率及提升普一类产品竞争力的配方关键技术研究与应用 (ZJZY2022A001); “利群”品牌主要烟叶产区烤烟化学成分特征与可用性评价 (ZJZY2019C003)

**作者简介:** 蒋佳磊 (1987—), 硕士, 工程师, 主要从事烟草数理统计和仿真建模工作, Email: jiangjialei@zjtobacco.com

**通讯作者:** 赵振杰 (1992—), 硕士, 工程师, 主要从事烟叶配方研究工作, Email: zhaozhenjie@zjtobacco.com

**收稿日期:** 2023-06-09; **网络出版日期:** 2024-05-11

和多酚与烟叶香型风格的判别模型, 吴雨露<sup>[10]</sup>基于烟叶的物化和感官指标采用 BDA 法建立了卷烟配方规则提取和配方维护模型。

随着大数据和人工智能的不断发展, 支持向量机 (SVM)<sup>[11]</sup>、最小二乘支持向量机 (LS-SVM)<sup>[12]</sup>、K 近邻算法 (KNN)<sup>[13]</sup>、神经网络 (ANN)<sup>[14]</sup>等基于不同数学原理的机器学习算法为烟草行业判别分类提供了不少新的思路, 但每种分类方法由于自身数学处理方式上的片面性而产生特定的分类模式, 从而对分类结果的精度和有效性造成不利的影 响。因此, 可以参考集成学习<sup>[15]</sup>的思路, 通过筛选并结合多个单一分类方法的优势形成全新的组合分类方法, 从而进一步提高分类方法的性能。本文在单一分类基础上结合分类方法选择和权重分配构建了一种更高精度的烟叶功能模块组合分类方法, 为烟叶质量评价、配方前移应用和企业数字化转型等方面提供理论基础和技术支撑。

## 1 材料与方法

### 1.1 材料

根据已进入配方中使用的 457 个烟叶样品 (2016—2019 年全国 11 个主要产烟省份的上部、中部和下部片烟, 由浙江中烟工业有限责任公司提供) 作用特点和烟草工业企业长期积累的配方经验数据, 将其分为增醇提质 (A)、增浓提劲 (B)、平衡烟气 (C) 和填充辅助 (D) 4 种功能模块, 分别为 A1~A163、B1~B179、C1~C88、D1~D27, 采用分层采样原则按比例随机抽取 70% 数据作为建模集用于分类模型构建, 余下 30% 数据作为验证集用于模型可靠性验证。

### 1.2 方法

#### 1.2.1 感官评吸

参考 YC/T530—2015 并作一定调整对烟叶样品进行感官评吸。结合 4 种功能模块的使用特点设计感官评吸指标, 包括优雅感、细腻感、甜香、透发性、清晰度、香气量、成团性、柔和性、杂气、润感、刺激性和余味 12 项, 其中优雅感来自于近红外模型评价得分, 是“利群”品牌卷烟特有的兼具感官和化学组分协调性的重要评价指标<sup>[16]</sup>, 其余 11 项指标由 7 位评吸人员对上述 457 个烟叶样品进行评吸和打分, 分值均为 5~8 分, 且为正向指标。为保证结果的可靠性, 参考竞技体育打分规则, 每个指标 (除优雅感外) 各去掉 2 个最高分和 2 个最低分, 取剩余 3 人均值 (文中

所有显著性检验均采用均值)。

#### 1.2.2 单一分类方法计算

MDA、FDA、BDA 均采用 Matlab 2022b 编写相应的函数, 其中 BDA 的密度函数采用 MDA 和 FDA 两种方法分别获得, 分别形成 M-BDA 和 F-BDA 2 种方法<sup>[17]</sup>; SVM 和 LS-SVM 分别采用 Matlab 外源 LibSVM 3.23 和 LS-SVM 1.8 工具箱编写, 二者均采用网络搜索算法和 10 次 10 折交叉验证优化规则化参数和核函数参数; KNN 和 ANN 分别调用 Matlab 2022b 软件中 KNN 工具箱和 ANN 工具箱 (建立含 10 个节点的单隐层神经网络) 计算。

#### 1.2.3 组合分类方法计算

(1) 组合分类原理。“投票法”是最简单的区分分类方法好坏的标准, 即若某种分类方法将一特定样品分为某一类, 就为此类记一票, 经多种分类方法投票后计算哪一类票数最多, 就将该样品分为那一类。但“投票法”不考虑分类方法间的关系和分类精度, 未能体现出各个分类方法相似性及其在综合分类中所占的权重。因此在选择分类方法时, 方法的多样性和方法间的相关度是 2 个重要的参考因素。同时, 为使预测精度较高的方法发挥更强的作用, 从而减少精度相对较差的方法对最终结果的影响, 要增加高精度方法在综合方法中的“投票权”, 即赋予更高的权重。

(2) 分类方法的选择。采用 Yule<sup>[18]</sup>提出的  $Q$  相关系数评价两种分类方法相似度:

$$Q_{ij} = \frac{N_{ij}^{11}N_{ij}^{00} - N_{ij}^{10}N_{ij}^{01}}{N_{ij}^{11}N_{ij}^{00} + N_{ij}^{10}N_{ij}^{01}} \quad (1)$$

其中:  $i, j=1, 2, 3, \dots, m$ , 为分类方法编号;  $N_{ij}^{11}$  为同时被  $i$  和  $j$  都正确分类的样品个数,  $N_{ij}^{00}$  为同时被  $i$  和  $j$  都错误分类的样品个数,  $N_{ij}^{10}$  为被  $i$  正确分类而被  $j$  错误分类的样品个数,  $N_{ij}^{01}$  为被  $i$  错误分类而被  $j$  正确分类的样品个数。

分类方法间需要有互补的机制才能提高组合方法的分类效果, 即  $Q$  值越小越好; 但对于分类精度明显较高的方法来说, 与其它各方法间的  $Q$  值必然较大, 因此分类方法的选择需要基于以下四原则: ①平均分类精度大于 95% 的方法直接入选, 小于 75% 则舍去;

②通过计算综合  $Q$  值  $TQ_i$ , 即  $TQ_i = \sum_{j=1: j \neq i}^m Q_{ij}$ , 按照  $TQ_i$  由小到大排列依次选择; ③对于  $Q$  值大于 99.0%

的两种分类方法只取其一；④入选分类方法数应与样品类别数一致或其整数倍。

(3) 入选分类方法的赋权组合。设用  $M_k$  分类方法对样本  $s(s_1, s_2, s_3, \dots, s_n)'$  进行分类, 得到分类结果  $M_k(s) = (M_k(s_1), M_k(s_2), M_k(s_3), \dots, M_k(s_n))'$ , 另设样本的真实类别为  $T_k(s) = (T_k(s_1), T_k(s_2), T_k(s_3), \dots, T_k(s_n))'$ , 通过比较  $M_k(s)$  和  $T_k(s)$  两个类别向量, 得比较矩阵  $[E_{ij}^k]$  (其中  $i, j = A, B, C, D$ , 为 4 种样本类别;  $k = 1, 2, 3, \dots, m$ , 为分类方法),  $E_{ij}^k$  表示真实类别为  $i$  类而被  $M_k(s)$  判别为  $j$  类的样品总个数, 当  $i=j$  时, 表示  $M_k$  完成了正确的分类。则某待判样品  $s_r (r=1, 2, 3, \dots, n)$ , 其真实类别为  $i$  类而被  $M_k$  判别为  $j$  类的概率  $p_{ij}^k$  为:

$$p_{ij}^k = P(T(s_r) = i | M_k(s_r) = j) = E_{ij}^k / \sum_{i=A}^D E_{ij}^k \quad (2)$$

式中  $\sum_{i=A}^D E_{ij}^k$  为  $M_k$  判别为  $j$  类的样本总数, 记:

$$V_j^k(s_r) = \begin{cases} 1 & j = M_k(s_r) \\ 0 & j \neq M_k(s_r) \end{cases} \quad (3)$$

则  $s_r$  在  $M_k$  分类结果下, 真实地属于  $i$  类的概率:

$$w^k(s_r) = \sum_{j=A}^D p_{ij}^k V_j^k(s_r) \quad (4)$$

则在不同分类方法组合下, 得到组合分类方法  $TM$  对  $s_r$  的判定结果为:

$$TM(s_r) = \sum_{k=1}^m w^k(s_r) \quad (5)$$

$\max_{i=A}^D (TM(s_r))$  所在的类别即为样品  $s_r$  的分类。

#### 1.2.4 方法评价

实际分类问题中, 各类别的样本比例往往不太平衡, 分类方法很容易偏向大类别而放弃小类别, 导致总体分类正确率无法真实反映方法的分类能力, 本研究引入一种 Cohen 等<sup>[19-21]</sup>提出的 Kappa 系数作为样本比例“偏向性”的惩罚指标, 以检验分类方法在不同模块间的一致性程度:

$$\text{Kappa} = \frac{\text{tr}(M) - r^T M^2 r}{1 - r^T M^2 r} \quad (6)$$

其中:  $\text{Kappa} \in [0, 1]$ ;  $M$  为通过分类结果得到的

混淆矩阵<sup>[15]</sup>;  $\text{tr}(M)$  为矩阵  $M$  的迹;  $r$  为  $M$  同维度的 1 向量;  $r^T$  为  $r$  的转置。

F1 度量是一种用于二分类问题的通用性能评价参数<sup>[22]</sup>, F1 度量越大则模型的分类精度越高。参考其计算方法建立多分类体系的性能评价参数 m-F1:

$$\text{m-F1} = \frac{2 \times \text{m-P} \times \text{m-R}}{\text{m-P} + \text{m-R}} \quad (7)$$

其中:  $\text{m-F1} \in [0, 1]$ ;  $\text{m-P} = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FP_i}$ ;

$\text{m-R} = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FN_i}$ ;  $n$  为类别数;  $\text{m-P}$  为多分类查

准率;  $\text{m-R}$  为多分类查全率;  $TP$ 、 $FP$ 、 $FN$  分别为通过混淆矩阵  $M$  得到的真正例、假正例、假反例。

#### 1.2.5 其它数值计算

各模块的协方差矩阵、建模集和验证集的随机分配分别采用 matlab 2022b 软件中 cov 函数、randperm 函数计算; 模块间欧式平方距离由 minitab 19 中的 MDA 法给出; 各感官指标的独立分类能力检验采用 SPSS 19 判别分析中的 Wilk's  $\lambda$  统计量; 多重比较计算、协方差相等性检验采用 matlab 2022b 自行编写代码。

## 2 结果与分析

### 2.1 感官评价指标描述性统计分析

457 个烟叶样品 4 个功能模块 12 种感官评吸得分的描述性统计分析见表 1。由表 1 可知, A、B 模块占比超过 70%, 而 D 模块比例仅 6.2%, 4 个模块比例差异较大, 需通过 Kappa 系数评价一致性程度; 从各指标均值上看, 4 个模块品质依次下降, 且 A、B 模块的标准偏差和极差均明显小于 C、D, 表明后者分布范围更加分散; 除 D 模块中杂气和润感符合正态分布 ( $P > 0.05$ ) 外, 其余均不符合正态分布 ( $P \leq 0.05$ ), 表明感官评价指标具有一定的主观随意性, 导致数据分布往往出现偏态和拖尾等非正态化表现, 模块间耦合程度增加, 也进一步提高了其分离难度。从多重比较计算结果上看, 相邻模块间 (AB、BC 和 CD) 各指标均存在极显著性差异, 表明分类在一定程度上具有可行性。从各模块整体上看, AB、BC 和 CD 的平均欧式距离分别为 3.9、10.4 和 11.4, 表明 AB 间距离远小于 BC 和 CD, 因此 AB 模块的区分可能相对困难。

表 1 烟叶感官评吸得分的描述性统计分析<sup>①</sup>  
Tab. 1 Descriptive statistical analysis of tobacco leaf sensory evaluation scores

模块	比例	指标	优雅感	细腻感	甜香	透发性	清晰度	香气量	成团性	柔和性	杂气	润感	刺激性	余味
A	35.5%	均值	7.2	7.3	7.1	7.2	7.2	7.1	7.2	7.3	7.3	7.2	7.3	7.2
		标准偏差	0.2	0.2	0.2	0.2	0.1	0.3	0.2	0.2	0.2	0.2	0.2	0.1
		极差	0.9	1.1	1.0	0.8	0.7	1.0	0.9	1.3	0.8	0.8	1.0	0.9
		<i>P</i> 值	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
B	38.9%	均值	6.8	7	6.8	7	7	6.8	7	7.1	6.9	6.8	7	6.9
		标准偏差	0.2	0.3	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.2	0.2	0.2
		极差	0.7	1.0	0.9	1.0	1.2	2.0	1.0	0.9	0.9	0.9	1.0	1.1
		<i>P</i> 值	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
C	19.3%	均值	6.2	6.5	6.3	6.7	6.4	6.4	6.4	6.5	6.2	6.2	6.4	6.2
		标准偏差	0.3	0.5	0.4	0.4	0.3	0.4	0.4	0.5	0.4	0.4	0.5	0.4
		极差	1.5	2.5	2.0	1.5	2.2	2.4	2.0	2.3	2.0	1.8	2.3	2.2
		<i>P</i> 值	<0.01	0.03	<0.01	<0.01	<0.01	0.02	<0.01	<0.01	<0.01	<0.01	<0.01	0.01
D	6.2%	均值	5.6	5.8	5.7	6.6	5.8	6.1	6	5.9	5.7	5.6	5.7	5.7
		标准偏差	0.3	0.5	0.4	0.4	0.4	0.5	0.4	0.6	0.5	0.5	0.5	0.5
		极差	1.4	2.0	1.3	1.7	1.6	2.0	1.6	2.0	1.5	1.5	1.8	2.0
		<i>P</i> 值	0.03	0.05	0.04	0.02	<0.01	0.05	0.02	0.05	0.06	0.07	0.05	0.05
AB		0.4**	0.2**	0.3**	0.2**	0.3**	0.3**	0.2**	0.2**	0.3**	0.3**	0.3**	0.3**	
BC	多重比较	0.6**	0.6**	0.5**	0.3**	0.6**	0.4**	0.6**	0.6**	0.7**	0.6**	0.6**	0.7**	
CD		0.5**	0.6**	0.6**	0.2**	0.6**	0.3**	0.4**	0.6**	0.5**	0.5**	0.7**	0.5**	

①注：若  $P>0.05$ ，则服从正态分布；反之，则不符合正态分布；“\*\*”表示差异极显著。

2.2 单一分类方法结果

按文献[17]方法对 4 种判别方法进行协方差矩阵相等性检验，结果表明各模块间协方差不全相等，判别式为“二次型”。4 种判别分析和 4 种机器学习方法

的建模集正确率 ( $R$ )，验证集正确率 ( $r$ )、平均正确率 ( $m$ ) 结果见表 2 (FDA 和 F-BDA 中判别函数取第一判别式，贡献度达到 94.3%)。

表 2 不同方法的正确率结果  
Fig. 2 The accuracy results of different methods

方法	判别分析			方法	机器学习		
	$R$	$r$	$m$		$R$	$r$	$m$
MDA	90.3%	77.2%	83.8%	SVM	91.3%	75.7%	83.5%
FDA	87.2%	80.2%	83.7%	LS-SVM	92.8%	74.3%	83.6%
M-BDA	90.3%	77.2%	83.8%	KNN	89.4%	77.2%	83.3%
F-BDA	87.2%	80.2%	83.7%	ANN	90.3%	75.7%	83.0%
<i>P</i> 值 <sup>①</sup>	0.05	0.02	0.35				

①注：采用  $t$  检验；95%置信度下， $P\leq 0.05$  表示具有高显著性差异；下同。

由表 2 可知，不同方法  $R$  最大的是 LS-SVM，达到 92.8%； $r$  最大的是 FDA 和 F-BDA，达到 80.2%； $m$  无明显差异，均在 83%~84%。与判别分析结果相比，机器学习方法的  $R$  高显著性升高 ( $P = 0.05$ )， $r$

则高显著性下降 ( $P = 0.02$ )，而  $m$  则无显著性差异 ( $P = 0.35$ )。不少研究表明，机器学习方法虽然在一定程度上可以提高建模准确率，但容易出现“过拟合”现象<sup>[23-25]</sup>，从而造成对建模集拟合度足够高，而对验证

集精度则明显下降。

由分类结果可得各方法的混淆矩阵  $M$  (见表 3、4), 由表 2~4 可知, MDA 与 M-BDA(FDA 与 F-BDA) 虽然  $R$  和  $r$  均相同, 但混淆矩阵存在一定差异; 不同

方法中错误分类均主要发生在 AB 两类, 这与上述 AB 类间距离远小于 BC 和 CD 的结果一致; 不同方法的验证集中 CD 两类的识别上均不够稳定, 也对分类的准确率产生了一定影响。

表 3 不同判别方法混淆矩阵  
Tab. 3 The confusion matrix of different discriminant analysis

类型	真实类	MDA				FDA				M-BDA				F-BDA			
		A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
建模集	A	109	5	0	0	102	12	0	0	108	6	0	0	101	13	0	0
	B	15	107	3	0	14	107	4	0	15	109	1	0	14	108	3	0
	C	0	4	55	3	0	0	55	7	0	7	54	1	0	5	55	2
	D	0	0	1	19	0	0	4	16	0	0	1	19	0	0	4	16
验证集	A	38	10	1	0	38	11	0	0	38	10	1	0	38	11	0	0
	B	6	44	4	0	5	47	2	0	7	44	3	0	5	49	0	0
	C	0	1	19	6	0	1	18	7	0	1	20	5	0	2	18	6
	D	0	0	3	4	0	0	1	6	0	0	4	3	0	0	3	4

表 4 不同机器学习方法混淆矩阵  
Fig. 4 The confusion matrix of different machine learning

类型	真实类	SVM				LS-SVM				KNN				ANN			
		A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
建模集	A	107	7	0	0	110	4	0	0	107	7	0	0	106	8	0	0
	B	13	111	1	0	12	111	2	0	13	112	0	0	14	110	1	0
	C	0	4	57	1	0	4	58	0	0	7	55	0	0	4	56	2
	D	0	0	2	18	0	0	1	19	0	1	6	13	0	0	2	18
验证集	A	40	9	0	0	39	10	0	0	37	12	0	0	38	11	0	0
	B	9	44	1	0	9	43	2	0	10	44	0	0	12	41	1	0
	C	0	5	16	5	0	5	17	4	0	4	20	2	0	1	21	4
	D	0	0	4	3	0	0	5	2	0	0	3	4	0	0	4	3

2.3 组合分类方法结果

采用 1.2.3(2)方法计算各方法间的相似度  $Q$  值和总相似度  $TQ$  值 (表 5)。由表 2 可知, 不同分类方法的平均正确率差异不大, 且  $TQ$  值由小到大排序为:  $KNN < ANN < FDA < M-BDA < MDA < F-BDA < SVM < LS-$

$SVM$ ; 由表 5 可知, MDA 与 M-BDA, FDA 与 F-BDA 两组的相似度均超过 99%, 故须各取其一; 本研究中共分为 4 个功能模块, 故选取 M-BDA、FDA、ANN 和 KNN 4 种分类方法作为组合分类的入选方法。

表 5 各分类方法相似度  $Q$  值和总相似度  $TQ$  值  
Fig. 5 The similarity and total similarity of different classifications

	MDA	FDA	M-BDA	F-BDA	SVM	LS-SVM	KNN	ANN
FDA	0.901							
M-BDA	0.996	0.862						
F-BDA	0.901	0.993	0.901					
SVM	0.948	0.970	0.948	0.970				
LS-SVM	0.974	0.939	0.974	0.956	0.980			
KNN	0.805	0.708	0.835	0.758	0.781	0.900		
ANN	0.851	0.858	0.851	0.909	0.877	0.925	0.760	
$TQ$	6.376	6.231	6.363	6.389	6.474	6.648	5.547	6.031

采用 1.2.3(3)的方法计算 M-BDA、FDA、ANN 和 KNN 四种方法的  $TM$  值、比较矩阵  $E$  和混淆矩阵  $M$ ，并计算  $R$ 、 $r$  和  $m$ （表 6）。

表 6 组合分类方法混淆矩阵

Fig. 6 The confusion matrix of combination classification

真实类	建模集				验证集			
	A	B	C	D	A	B	C	D
A	110	4	0	0	44	5	0	0
B	9	116	0	0	5	49	0	0
C	0	0	62	0	0	0	22	4
D	0	0	2	18	0	0	1	6
正确率	95.3%				89.0%			

由表 6 可知，建模集和验证集均有 15 个样品误判；与单一分类方法相比，组合分类的误判仅发生在 AB 和 CD 间，BC 间达到完全正确分类，且组合分类的  $R$  和  $r$  均明显提高，分别达到 95.3%和 89.0%， $m$  也由低于 84%提高到 92.2%。与本研究中单一分类结果相比，

同类通过感官指标对烟叶功能模块进行分类的研究<sup>[4, 26]</sup>其平均分类正确率也在 80%左右（其中文献[4]为 76.8%，文献[26]为 80.8%），而本研究提出的组合分类方法则远高于此。

### 2.4 方法评价

按 1.2.4 方法计算 Kappa 和 m-F1 评价系数，结果见表 7。一般认为 Kappa 系数达到 0.6 时分类模型有意义，超过 0.8 时则表明达到较高的分类一致性<sup>[27]</sup>。由表 7 可知，Kappa 系数均大于 0.6，但所有单一分类的验证集均小于 0.8，表明单一分类方法虽然有意义，但模型可靠性不强；而组合分类的 Kappa 系数均大于 0.8，且相比提升显著，表明组合分类结果一致性程度高，受模块比例差异的影响小，模型可靠性强。从 m-F1 度量结果看，单一分类建模集精度远高于验证集，表明单一分类方法普遍存在模型泛化能力差的问题；与前者均值相比，组合分类建模集虽然性能提升不大（6.6%），但验证集则显著提升（21.2%），表明组合分类方法大幅度增强了模型的泛化能力，从而显著提升了模型在实际烟叶功能模块分类中的可用性。

表 7 Kappa 和 m-F1 评价系数  
Tab. 7 Kappa and m-F1 evaluation coefficients

评价指标	样本集	MDA	FDA	M-BDA	F-BDA	SVM	LS-SVM	KNN	ANN	组合分类
Kappa	建模集	0.858	0.813	0.858	0.812	0.872	0.895	0.843	0.858	0.931
	验证集	0.665	0.708	0.664	0.704	0.637	0.614	0.658	0.640	0.837
m-F1	建模集	0.906	0.848	0.916	0.868	0.918	0.940	0.878	0.906	0.957
	验证集	0.706	0.767	0.690	0.732	0.673	0.641	0.746	0.693	0.856

### 3 讨论

通过单一分类方法优势组合形成组合分类方法的普遍有效性值得探讨。设各单一分类的正确率接近且均为  $x$ ，当采用普通“投票法”时，对于一个样本的正确分类，4 种单一分类方法必须保证有至少 3 种方法均达到正确分类，设该样本通过组合分类的正确率为  $f(x)$ ，则  $f(x) \sim x$  满足如下二项分布关系式：

$$f(x) = \sum_{i=3}^4 C_4^i x^i (1-x)^{4-i} \quad (8)$$

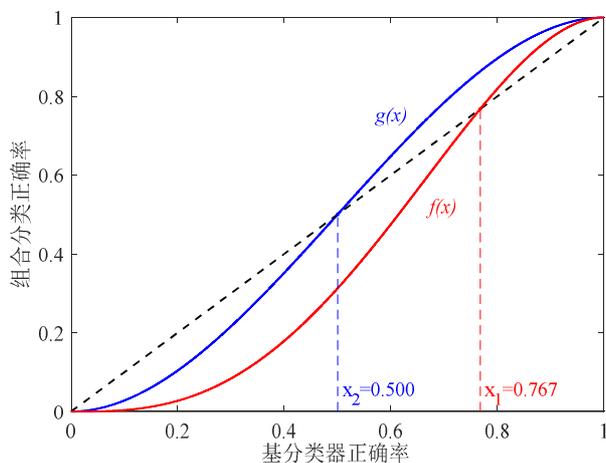
两者的函数图像如图 1。由图 1 可知，当  $x > x_1$  时， $f(x) > x$ ，表明组合分类的效果优于单一分类，通过求解  $f(x_1) = x_1$ ，得  $x_1 = 0.767$ ，即仅当单一分类的正确率大于 76.7%时，组合分类才能发挥作用。当按本研

究方法对各单一分类结果按分类精度大小进行赋权后，对于一个样本而言，有两种分类方法甚至仅一种方法完成正确分类，也有可能组合分类中形成正确的分类结果。设本研究中组合分类的正确率为  $g(x)$ ，不妨设有两种分类方法正确分类情况中，有一半的概率可以通过组合方法完成正确分类，则：

$$g(x) = C_4^2 x^2 (1-x)^2 / 2 + \sum_{i=3}^4 C_4^i x^i (1-x)^{4-i} \quad (9)$$

$g(x) \sim x$  函数图像如图 1。按同样方法可求解  $g(x_2) = x_2$ ，得  $x_2 = 0.500$ ，即理论上当单一分类的正确率大于 50.0%时，组合分类就能发挥作用。从本研究结果看，各单一分类平均正确率接近且均在 83%~84%，不仅远超 50.0%，且明显大于 76.7%，表明组合分类的有效性在理论上是可以肯定的。按上述原理采

用 4 种单一分类的真实正确率计算组合分类建模集、验证集和两者均值的理论正确率分别为：96.8%、87.2%和 92.7%，对比表 6，理论值与实际值较为相符。综上所述，通过理论分析和实际结果验证均表明通过单一分类方法优势组合形成组合分类方法是普遍有效的。

图 1  $f(x)$  和  $g(x)$  函数图像Fig. 1 Images of  $f(x)$  and  $g(x)$  functions

采用 SPSS 19 判别分析中的 Wilk's  $\lambda$  统计量检验各指标的独立分类能力，采用  $F$  检验判断其显著性，结果见表 8。Wilk's  $\lambda$  值越小表明组间差异越大，即分类能力越强<sup>[28]</sup>。由表 8 可知，12 种指标对分类结果均有极显著性影响 ( $P < 0.01$ )，其中优雅感独立分类能力最强，贡献度达到 18.0%，这可能是本研究中优雅感本身是一项通过近红外基于多年感官和化学数据得到的综合指标，既拥有满足人为需求的主观性，又具备识别数据特征的客观性，分类效果较好；杂气、余味、润感和清晰度 4 个指标次之，透发性和香气量两个指标最弱，这与利群品牌的风格定位相符。中式卷烟传统以香型划分，大部分品牌都以香气为核心进行产品风格打造<sup>[29-31]</sup>，但消费者在品烟过程中，对于香气的感觉仅是其中的一小部分，更重要的是留在口腔中的味，利群品牌卷烟追求清淡醇和的吸味风格，在评吸环节淡化香气类指标，并通过配方设计减少杂气和增强余味。

表 8 12 个指标分类能力检验

Tab. 8 Classification ability test of 12 indicators

指标	Wilk's $\lambda$	F	贡献度	$P$	指标	Wilk's $\lambda$	F	贡献度	$P$
优雅感	0.159	475.725	18.0%	<0.01	成团性	0.310	201.158	7.6%	<0.01
杂气	0.226	308.882	11.7%	<0.01	甜香	0.312	199.005	7.5%	<0.01
余味	0.234	295.111	11.2%	<0.01	柔和性	0.387	143.15	5.4%	<0.01
润感	0.244	278.859	10.5%	<0.01	细腻感	0.397	137.072	5.2%	<0.01
清晰度	0.258	258.97	9.8%	<0.01	透发性	0.553	72.808	2.8%	<0.01
刺激性	0.306	204.83	7.7%	<0.01	香气量	0.560	70.919	2.7%	<0.01

由各指标分值分布图 (图 2) 可进一步计算各模块间的分离度  $f$  (表 9)。一般认为  $f$  小于 1 时两峰难以分离， $f$  大于 1.5 时两峰完全分离， $f$  在两者之间时两峰有一定分离能力<sup>[32]</sup>。由表 9 可知，分离度均小于 1.5，其中优雅感均大于 1，杂气和润感均有两项大于 1，余味和清晰度各有一项大于 1，其余指标均小于 1，进一步表明优雅感、杂气、余味、润感和清晰度 5 项指标对分类起主要作用。另一方面， $f_{BC}$  有 5 项指标大于 1，

$f_{AB}$  有 3 项指标大于 1， $f_{CD}$  有 2 项指标大于 1，且在平均分离度上  $f_{BC}=0.95 > f_{AB}=0.75 > f_{CD}=0.67$ ，因此，从分离度上看，B、C 模块更容易分离，且与组合分类的实际结果相符合；C 和 D 两模块平均欧式距离最大，但分离度却最小，这是由于 C、D 两模块峰形分布扁平，虽然多数指标中心位置差异明显，但极差大，存在大范围交叉，从而不利于正确分类。

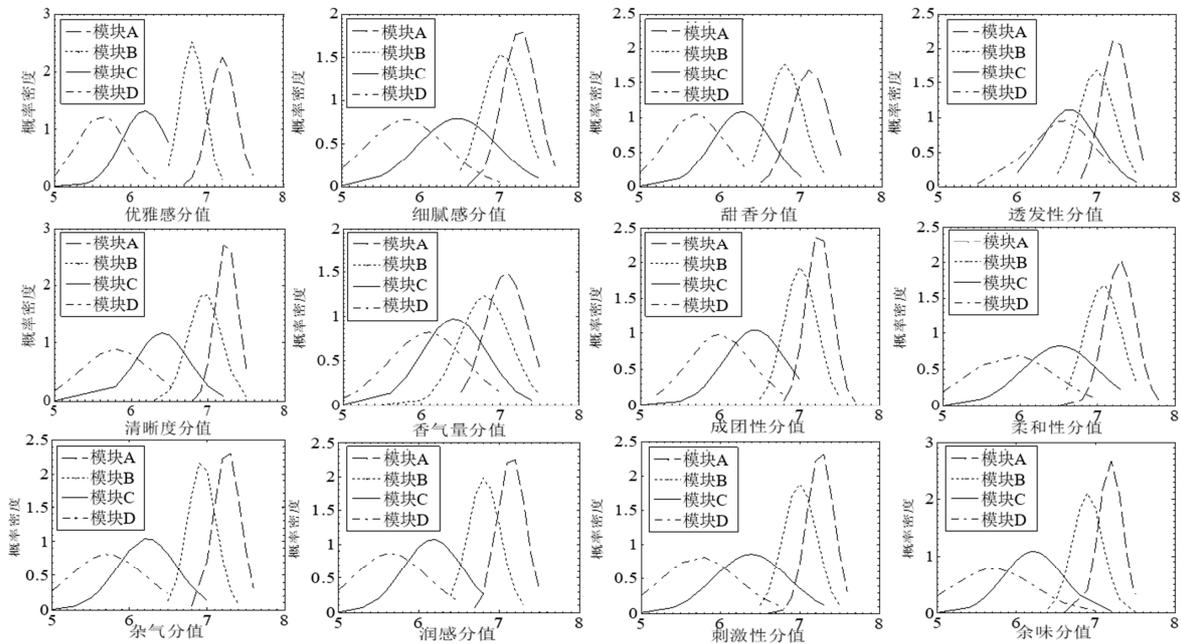


图 2 12 项指标分值分布图

Tab. 2 The distribution of sensory evaluation scores

表 9 不同模块间分离度  $f^{\text{①}}$

Fig. 9 The resolution between different modules

$f$	优雅感	细腻感	甜香	透发性	清晰度	香气量	成团性	柔和性	杂气	润感	刺激性	余味	均值
$f_{AB}$	1.05	0.59	0.73	0.74	0.59	0.51	0.66	0.59	1.03	1.03	0.66	0.79	0.75
$f_{BC}$	1.18	0.74	0.85	0.69	1.02	0.59	0.93	0.87	1.18	1.18	0.94	1.10	0.94
$f_{CD}$	1.03	0.72	0.95	0.07	0.79	0.35	0.56	0.96	0.54	0.71	0.71	0.59	0.67

注：①分离度  $f=1.18 \times$  相邻峰中心点分值之差/相邻峰半峰宽之和。

从模块功能和表 1 结果上看，总体上 A、B、C、D 4 个模块品质依次下降，但从图 2 看各指标的分布交叉耦合严重，分类方法的目的即在于建立合适的算法求解这种耦合关系。若用  $F_{AB}$ 、 $F_{BA}$ 、 $F_{CD}$ 、 $F_{DC}$  分别表示 30 个误判样品类型（具体见表 10），通过配对 t 检验发现， $F_{AB}$  极显著性小于  $F_{BA}$  ( $P<0.01$ )， $F_{CD}$  极显著性小于  $F_{DC}$  ( $P<0.01$ )，即真实模块为 A/C 的样品指标分值极显著性小于 B/D，表明这些误判样品的指标间并不存在耦合关系，即指标评分与其模块真实类

别并不相符。对工业企业来说，原烟配方在满足消费者需求，自身混配目标和产品质量指标的情况下，还需考虑库存数量、产地来源、采购成本，因此需要寻求一个库存、成本和质量的相对平衡点。基于现实情况对一些烟叶的模块类别作小范围微调（利群品牌一般在 5%左右）能帮助企业更加充分地利用库存烟叶资源，本研究中误判样品比例为  $30/457=6.56\%$ ，基本符合烟叶配方的调整空间。

表 10 误判样品比较<sup>①</sup>

Fig. 10 The comparison of miscalculation samples

类型	优雅感	细腻感	甜香	透发性	清晰度	香气量	成团性	柔和性	杂气	润感	刺激性	余味
$F_{AB}$	6.9±0.1	7.0±0.3	7.0±0.2	7.1±0.3	7.1±0.2	7.0±0.3	7.1±0.1	7.0±0.3	7.0±0.1	6.9±0.1	7.1±0.2	7.0±0.1
$F_{BA}$	7.1±0.1	7.3±0.1	7.0±0.2	7.2±0.2	7.2±0.1	7.0±0.3	7.2±0.1	7.3±0.1	7.2±0.1	7.1±0.1	7.2±0.1	7.1±0.1
$F_{CD}$	5.5±0.4	5.5±0.3	5.6±0.5	6.4±0.3	5.7±0.4	5.9±0.6	5.8±0.6	5.6±0.4	5.4±0.3	5.4±0.4	5.5±0.4	5.4±0.3
$F_{DC}$	6.1±0.2	6.5±0.6	6.2±0.2	6.8±0.2	6.4±0.3	6.3±0.3	6.6±0.2	6.5±0.3	6.3±0.1	6.3±0.2	6.4±0.4	6.3±0.4

注：① $F_{AB}$  表示将 A 误判为 B 样品的指标均值； $F_{BA}$ 、 $F_{CD}$ 、 $F_{DC}$  同理。

## 4 结论

(1) 8种单一分类方法中LS-SVM的建模正确率最高(92.8%), FDA和F-BDA的验证正确率最高(80.2%), 且与判别分析相比, 机器学习法显著提高了建模正确率, 降低了验证正确率, 但平均正确率不变。(2) 择优选择M-BDA、FDA、ANN和KNN 4种方法, 并按精度赋权建立的组合分类方法同时提高了建模正确率(95.3%)和验证正确率(89.0%), 平均正确率由低于84%提高到92.2%, 远高于同类研究结果, 并通过理论计算和实际结果验证了组合分类方法的普遍有效性。(3) 组合分类方法Kappa系数均大于0.8, 方法可靠, 一致性程度高, 验证集m-F1度量显著提升21.2%, 模型泛化能力大为增强。(4) 优雅感、杂气、余味、润感和清晰度5项指标对分类起主要作用, 符合利群品牌的风格特征。(5) 误判样品(6.5%)指标评分与其模块真实类别的不匹配归因于对库存、成本和质量的平衡, 基本符合烟叶配方的调整空间。

## 参考文献

- [1] 闫克玉, 王光耀, 李春松. 我国烤烟质量分析评价研究进展[J]. 郑州轻工业学院学报: 自然科学版, 2007, 22(2/3):49-52.  
YAN Keyu, WANG Guangyao, LI Chunsong, et al. Advance in evaluation research of flue-cured tobacco quality in China[J]. Journal of Zhengzhou University of light industry (Natural Science), 2007, 22(2/3):49-52.
- [2] 郑州烟草研究院. 中国烟叶质量白皮书(2016年)[M]. 北京: 中国烟草总公司, 2016.  
Zhengzhou Tobacco Research Institute. White paper of China tobacco leaf quality[M]. Beijing: China National Tobacco Corporation, 2016. (自译)
- [3] 王强, 李孟军, 陈英武. 卷烟配方数据挖掘技术研究进展[J]. 中国烟草科学, 2007, 28(4): 14-17.  
WANG Qiang, LI Mengjun, CHEN Yingwu. Research progress in data mining technology on cigarette formulation[J]. Chinese Tobacco Science, 2007, 28(4): 14-17.
- [4] 邵惠芳, 赵昕宇, 许自成, 等. 基于Fisher判别分析的烤烟感官质量与工业应用价值的关系研究[J]. 中国烟草学报, 2011, 17(6): 13-18.  
SHAO Huifang, ZHAO Xinyu, XU Zicheng, et al. Studies on relationship between sensory quality and industrial availability of flue-cured tobacco based on Fisher discriminant analysis[J]. Acta Tabacaria Sinica, 2011, 17(6): 13-18.
- [5] 毕淑峰, 朱显灵, 马成泽. 判别分析在烤烟品质鉴定中的应用[J]. 中国农学通报, 2005, 21(1):79-79.  
BI Shufeng, ZHU Xianling, MA Chengze. Application of discriminant analysis indistinguishing flue-cured tobacco quality[J]. Chinese Agricultural Science Bulletin, 2005, 21(1): 79-79.
- [6] 张灵帅, 王卫东, 谷运红, 等. 近红外光谱的主成分分析-马氏距离聚类判于卷烟的真伪鉴别[J]. 光谱学与光谱分析, 2011, 31(5):1254-1257.  
ZHANG Lingshuai, WANG Weidong, GU Yunhong, et al. Identification of authentic and fake cigarettes using near infrared spectroscopy combined with principal component analysis-mahalanobis distance[J]. Spectroscopy and Spectral Analysis, 2011, 31(5):1254-1257.
- [7] 潘玲, 云月利, 陈振国, 等. 基于湖北烤烟综合质量的香型分析[J]. 华北农学报, 2015, 30(S1):217-224.  
PAN Ling, YUN Yueli, CHEN Zhenguo, et al. Analysis of aroma types based on comprehensive quality of flue-cured tobacco in Hubei[J]. Acta agriculture Boreali-Sinica, 2015, 30(S1):217-224.
- [8] 高宪辉, 王松峰, 孙帅帅, 等. 鲜烟成熟度颜色值指标及其判别函数研究[J]. 中国烟草学报, 2017, 23(1): 77-85.  
GAO Xianhui, WANG Songfeng, SUN Shuaishuai, et al. Study on color space data-based discriminating functions of fresh tobacco at various mature stages[J]. Acta Tabacaria Sinica, 2017, 23(1): 77-85.
- [9] 李超, 李娥贤, 张承明, 等. 基于因子分析的烤烟香型定量判别及其与产区的对应关系[J]. 中国烟草学报, 2016, 22(6):51-62.  
LI Chao, LI Exian, ZHANG Chengming, et al. Factor analysis based quantitative determination of flavor type and its corresponding relationship with growing areas in flue-cured tobacco[J]. Acta Tabacaria Sinica, 2016, 22(6):51-62.
- [10] 吴雨露. 基于贝叶斯网络的卷烟配方规则提取和配方维护[D]. 沈阳: 东北大学, 2016.  
WU Yulu. Rule extraction and formula maintenance of cigarette based on bayesian network[D]. Shenyang: Northeastern University, 2016.
- [11] 聂铭, 周冀衡, 杨荣生, 等. 基于MIV-SVM的烤烟评吸质量预测模型[J]. 中国烟草学报, 2014, 20(6):56-62.  
NIE Ming, ZHOU Jiheng, YANG Rongsheng, et al. MIV-SVM-based prediction model for smoking quality of flue-cured tobacco[J]. Acta Tabacaria Sinica, 2014, 20(6):56-62.
- [12] 章英, 贺立源, 叶颖泽, 等. 基于LS-SVM的烤烟烟叶产地判别[J]. 湖北农业科学, 2012, 51(3):583-585.  
ZHANG Ying, HE Liyuan, YE Yingze, et al. Identification of producing area of tobacco leaf based on LS-SVM[J]. Hubei Agricultural Sciences, 2012, 51(3):583-585.
- [13] 李航. 基于聚类和加权K近邻的烟叶分级研究[D]. 郑州: 郑州大学, 2017.  
LI Hang. The research on tobacco classification based on clustering and weighted KNN[J]. Zhengzhou: Zhengzhou University, 2017.
- [14] 石子健. 基于分类算法的卷烟感官评吸指标预测方法[D]. 沈阳: 东北大学, 2015.  
SHI Zijian. Prediction methods of cigarette sensory index evaluation based on classification algorithms[D]. Shenyang: Northeastern University, 2015.
- [15] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2017.  
ZHOU Zhihua. Machine learning[M]. Beijing: Tsinghua University Press, 2017.
- [16] 吴继忠, 毕一鸣, 李石头, 等. 一种基于近红外光谱与感官评吸互信息判别感官表征信息的方法: 中国, 201810023242.0[P]. 2018-08-17.  
WU Jizhong, BI Yiming, LI Shitou, et al. A method for discriminating sensory representation information based on near infrared spectroscopy and sensory assessment mutual information: China, 201810023242.0[P]. 2018-08-17. (自译)
- [17] 李柏年, 吴礼斌. Matlab数据分析方法[M]. 北京: 机械工业出版社, 2012.  
LI Bainian, WU Libin. Matlab data analysis method[M]. Beijing: China Machine Press, 2012.
- [18] Yule G U. On the association of attributes in statistics: with

- illustrations from the material of the childhood society, &c[J]. *Philosophical Transactions of the Royal Society of London*, 1900(194):257-319.
- [19] Cohen J A. A coefficient of agreement for nominal scales[J]. *Educational and Psychological Measurement*, 1960, 20(1): 37-46.
- [20] Su L, Gong M, Zhang P, et al. Deep learning and mapping based ternary change detection for information unbalanced images[J]. *Pattern Recognition*, 2017, 66: 213-228.
- [21] Gong M, Yang H, Zhang P, et al. Feature learning and change feature classification based on deep learning for ternary change detection in SAR images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017, 129 (jul.): 212-225.
- [22] 李航. 统计学习方法[M]. 清华大学出版社, 2012.  
LI Hang. *Statistical learning method*[M]. Beijing: Tsinghua University Press, 2012. (自译)
- [23] 赵晓芳, 万明勇, 周萍, 等. 基于 RF-WPA 混合算法优化 SVM 参数的电力工程造价预测[J]. *中国电力企业管理*, 2018, 549(36): 80-81.  
ZHAO Xiaofang, WAN Mingyong, ZHOU Ping, et al. The SVM parameters optimized based on RF-WPA for cost prediction of electric engineering[J]. *Engineering Management*, 2018, 549(36):80-81. (自译)
- [24] 黎隽男, 吕佳. 基于近邻密度和半监督 KNN 的集成自训练方法[J]. *计算机工程与应用*, 2018, 54(20):137-143.  
LI Junnan, LV Jia. Integrated self-training method based on neighborhood density and semi-supervised KNN. *Computer Engineering and Applications*, 2018, 54(20): 132-138.
- [25] 金龙, 况雪源, 黄海洪, 等. 神经网络预报模型的过拟合研究[J]. *气象学报*, 2004, 62(1):62-70.  
JIN Long, KUANG Xueyuan, HUANG Haihong, et al. Study on the overfitting of the artificial neural network for casting model[J]. *Acta Meteorologica Sinica*, 2004, 62(1):62-70.
- [26] 刘伟, 徐磊, 刘晓利, 等. 基于聚类和逐步判别分析的烤烟感官质量特征分类评价[J]. *贵州农业科学*, 2017, 45(1): 143-147.  
LIU Wei, XU Lei, LIU Xiaoli, et al. Classified evaluation of sensory quality characteristics of flue-cured tobacco based on clustering and stepwise discriminant analysis[J]. *Guizhou Agricultural Sciences*, 2017, 45(1):143-147.
- [27] Bishop C . *Pattern recognition and machine learning*[M]. Berlin: Springer, 2006.
- [28] 胡建鹏, 陈强, 黄容. 逐步贝叶斯判别分析中的变量优化方法研究[J]. *计算机工程与应用*, 2014, 50(21):63-67.  
HU Jianpeng, CHEN Qiang, HUANG Rong. Study on variable optimization method in stepwise Bayes discriminant analysis. *Computer Engineering and Applications*, 2014, 50(21): 63-67.
- [29] 于录, 阮晓明, 卢在雨, 等. 叶组配方的分组加工模块设计[J]. *烟草科技*, 2006, 39(7): 11-13, 21.  
YU Lu, RUAN Xiaoming, LU Zaiyu, et al. Method of grouping a tobacco blend for optimization of processing[J]. *Tobacco Science & Technology*, 2006, 39(7): 11-13, 21.
- [30] 马慧婷, 赵铭钦, 于海顺, 等. 基于模糊综合评判烟叶原料使用类群的初步划分[J]. *中国烟草科学*, 2015, 36(2): 1-7.  
MA Huiting, ZHAO Mingqin, YU Haishun, et al. Preliminary division of the use groups of raw tobacco leaves based on fuzzy comprehensive Evaluation[J]. *Chinese Tobacco Science*, 2015, 36(2): 1-7.
- [31] 何望望, 吴风光, 谢豪, 等. 不同分组方法对原烟配方模块质量的影响[J]. *中国烟草科学*, 2011, 32(2): 86-89.  
HE Jiewang, WU Fengguang, XIE Hao, et al. Effects of different grouping on green blend module's quality[J]. *Chinese Tobacco Science*, 2011, 32(2): 86-89.
- [32] 浙江大学. *无机及分析化学*[M]. 浙江: 浙江大学, 2006.  
Zhejiang University. *Inorganic and analytical chemistry*[M]. Zhejiang: Zhejiang University,2006.(自译)

## Classification of tobacco leaf formula modules based on a combination of multiple methodological advantages

JIANG Jialei, LIAO Fu, HAO Xianwei, TANG Xiaodong, CHEN Xiaoshui, ZHU Shuxiu, ZHAO Zhenjie\*

Technology Centre, China Tobacco Zhejiang Industrial Co., Ltd, Hangzhou 310024, China

**Abstract:** To deeply study grouping methods and principles of cigarette brand tobacco leaf formulas, based on the 457 tobacco samples of 12 kinds of sensory evaluation indexes, this study compared four discriminant analysis and four machine learning methods for modeling set accuracy (R), validation set accuracy (r), and average accuracy (m) across four tobacco leaf classifications. A high-precision composite classification method was constructed based on method selection and weight allocation. Results showed that: (1) Compared with discriminant analysis, machine learning significantly improved R, while r significantly decreased, with LS-SVM having the highest R (92.8%), and FDA and F-BDA having the highest r (80.2%), but there was no significant difference in m; (2) Optimized selection of M-BDA, FDA, ANN, and KNN methods and the composite classification method established by accuracy weighting simultaneously improved R (95.3%) and r (89.0%), and increased m from below 84% to 92.2%, validating the general effectiveness of the composite classification method through theoretical calculations and practical results; (3) The Kappa coefficient of the composite classification method was greater than 0.8, indicating reliability, high consistency, and a significant improvement in validation set m-F1 measure by 21.2%, greatly enhancing the model's generalization ability; (4) Five indicators, namely elegance, off-flavor, aftertaste, moistness, and clarity, played a major role in classification, aligning with the style characteristics of the Liqun brand; (5) Misjudged samples (6.5%) with indicator scores not matching their real module categories were attributed to the balance of stock, cost, and quality, generally conforming to the adjustment space of tobacco leaf formulas.

**Keywords:** tobacco leaf; composite classification; discriminant analysis; machine learning; sensory evaluation; formulation modules

\*Corresponding author. Email: zhaozhenjie@zjtobacco.com