

DOI: 10.13652/j.issn.1003-5788.2020.03.031

基于 BP 神经网络的烟叶醇化感官质量仿真模拟

Simulation of sensory quality of tobacco leaf alcoholization based on
BP neural network

邓羽翔 罗 诚 李东亮 杨 杰

DENG Yu-xiang LUO Cheng LI Dong-liang YANG Jie

周 东 杜 薇 陈思昂

ZHOU Dong DU wei CHEN Si-ang

(四川中烟工业有限责任公司技术中心, 四川 成都 610066)

(Technology Center, China Tobacco Sichuan Industrial Co., Ltd., Chengdu, Sichuan 610066, China)

摘要:通过建立烟叶醇化感官质量的预测模型,用于预测烟叶醇化过程中的感官变化,突破烟叶养护工作传统以经验指导烟叶库存周转和配方使用。选取 4 个仓库中 18 个不同样品烟叶在间隔 6 个月,共 36 个月自然醇化后的醇化感官评分作为初始样本。通过因子分析,划分出烟叶醇化度等级,再进行 BP 神经网络训练,经检验表明此模型应用于预测烟叶醇化感官质量性能较好,均方误差最大值(MSE)为 $1.00E-06$,预测可知第 42 个月烟叶醇化度等级仅凉山会理 C3F 不合格。

关键词:烟叶醇化;感官评价;因子分析;BP 神经网络

Abstract: Establish a prediction model for sensory quality of tobacco leaf alcoholization, which is used to predict sensory changes in the process of tobacco leaf alcoholization process, breaking through the traditional empirical method to guide tobacco leaf inventory turnover and formulation use. In this paper, 18 alcoholic sensory scores of the 18 different samples in 4 warehouses were taken as initial samples at intervals of 6 months, totally 36 months. Firstly, Through factor analysis, the tobacco alcoholization level is divided, and then to perform BP neural network test and prediction, the results show that this model is feasible and effective for predicting the sensory quality of tobacco leaf alcoholization, and Mean Squared Error(MSE) is $1.00E-06$. It is predicted that the degree of alcoholification of tobacco leaves in the 42th month can only be seen as a failure of Liangshan Huili C3F.

Keywords: tobacco leaf alcoholization; sensory evaluation; factor

analysis; BP neural network

烟叶发酵是卷烟加工中极为重要的环节,良好的发酵工艺可以改善烟叶品质,降低烟叶青杂气和刺激性,使烟草特征香气显露,显著提高其可用性^[1]。打叶复烤后烟叶需自然醇化,醇化会使烟叶内主要化合物和致香成分发生急剧的变化^[2-3],经相关性分析和回归分析发现主要化学成分与感官质量存在直接或间接的关系^[4-7]。杨波等^[8]对四川凉山会理的烟叶进行了跟踪仓储试验,采用了描述统计和回归分析拟合了烟叶醇化品质变化的曲线方程,得到了烟叶适宜的醇化时间。刘红光等^[9]发现复烤片烟感官品质和化学成分可用性指数(CCUI)随醇化时间的延长均呈现先上升后下降的变化趋势。对于烟叶醇化感官评价研究,大多数研究仅限于传统的数理统计方法,包括一般线性相关性分析、主成分分析、多元回归分析、曲线回归分析、逻辑回归分析、因子分析等,这类统计方法在定性研究中能够快速找到规律性,但目前已知影响烟叶醇化质量的主要因子是微生物群落在一定温度、相对湿度、氧气浓度下发生的生化、物理反应,而烟叶醇化质量不能仅以感官质量中的杂气、刺激性或主要化合物和致香成分含量所决定,选用何种因子和因子采用何种表达式是一种推测,影响了因子的多样性和某些因子的不可测性,使得回归分析在某些情况下受到限制,且无法达到有效预测效果。

神经网络的优点在于它具有模拟多变量而不需要对输入变量做复杂的相关假设的能力^[10]。杨淑娥等^[11]研究表明,经过 BP(Back Propagation)神经网络训练可避免重复的因子表达,在指标间相关度较高、呈非线性变化,或数据缺漏不全的情况下仍可得到比较满意的结果。试

基金项目:四川中烟工业有限责任公司科研项目(编号:JL/CY-ZYGSJ003-04)

作者简介:邓羽翔,男,学士。

通信作者:李东亮(1973—),男,四川中烟工业有限责任公司研究员,博士。E-mail:360188288@qq.com

收稿日期:2019-12-19

实验通过对醇化周期内烟叶感官质量评分进行 BP 神经网络训练,对烟叶醇化过程中感官质量变化的预测方法进行的研究,旨在建立 BP 神经网络模型,为仓储环节烟叶周转、配方设计使用提供指导。

1 材料和方法

1.1 材料

醇化样品:2012 年广东南雄、云南曲靖、福建三明、河南洛阳、重庆丰都、四川凉山 6 个产区的打叶复烤后的片烟,每个产区均包含上(B2F)、中(C3F)、下(X2F) 3 个部位等级,共计 18 种片烟样品,四川中烟工业有限责任公司。每种样品各取同一批次的成品片烟 24 件,分别存放于四川中烟工业有限责任公司的成都多宝寺仓库、成都多宝寺气调仓库、什邡李子园仓库、西昌新厂仓库进行自然醇化,每种样品每个仓库存放 6 件。

1.2 试验设备

卷烟机:ZB25 型,上海烟草机械厂。

1.3 方法

1.3.1 醇化后样品取样及卷烟样品制备 从复烤片烟

入库起,以每 6 个月为周期对醇化的片烟样品取样,试验周期 36 个月,共取样 6 次。抽样方法按 GB/T 19616—2004,每次随机取 1 箱(下次取样时,排除前期已取样的烟箱),从上、中、下 3 个取样点,每个取样点取 5 kg 左右。剔除烟叶中残留的烟梗后,将烟片切成 1.0 mm 左右的烟丝。切丝前要求烟叶的水分在 12% 左右,若烟叶较湿,采用自然晾干的方法平衡水分,若烟叶较干,采用蒸锅来增大水分。切丝后,将烟丝混合均匀,采用四分法保留 500 g,用于评吸卷烟样品的卷制。在同一机台,按照“娇子(时代阳光)”工艺标准,进行卷烟样品的卷制。

1.3.2 感官评吸方法 采用“九分值”行业单料烟感官评价方法^[12],从香气特性、烟气特性、口感特性三方面进行评价,其中:香气特性包括香气质、香气量、丰满程度及杂气,烟气特性包括浓度、劲头、细腻程度及成团性,口感特性包括刺激性、干燥感、干净程度、甜度及余味。

1.3.3 训练样本集的准备 6 位感官评吸专业人员按“九分值”单料烟感官评吸方法每隔 6 个月,共 6 次感官评吸,结果见表 1。

表 1 样本的感官评吸总分描述性统计

Table 1 Descriptive statistics of total sensory scores of samples

样本来源	样本数	极小值	极大值	均值	标准差	全距	偏度	峰度
成都多宝库	108	57.00	80.00	69.075 0	4.768 67	23	-0.464	0.029
成都多宝气调库	108	59.00	78.00	70.653 7	4.005 61	19	-0.682	0.497
什邡李子园库	108	53.50	79.00	68.847 2	5.412 20	25	-0.567	0.159
西昌新厂库	108	52.00	78.00	68.175 9	5.393 11	26	-0.509	0.126

2 综合评分指数的构建与等级划分

基于主成分分析分别对 4 个仓库在 6 个时间节点年的感官评吸结果进行因子分析、指数构建与等级划分。以成都多宝仓库为例。

2.1 综合指数的构建

根据主成分分析方法的原理,以因子方差累计贡献率 > 80% 为准则来提取主成分因子^[13]。由表 2 可知,3 个因子解释了 18 个指标 91.64% 的方差。因此提取前 3 个主成分可以代替整体信息。

表 2 主成分分析结果

Table 2 Principal component analysis results

成分	初始特征值			提取平方和载入		
	合计	方差/%	累积/%	合计	方差/%	累积/%
1	12.300	68.334	68.334	12.300	68.334	68.334
2	2.566	14.255	82.589	2.566	14.255	82.589
3	1.525	8.474	91.064	1.525	8.474	91.064
4	0.927	5.152	96.216			
5	0.681	3.784	100.000			

根据 SPSS 算出成分载荷矩阵,再用载荷矩阵中的指标系数除以对应各个主因子的特征根,得出各个原始变量的系数,进一步得到因子得分系数矩阵,主成分因子表示:

$$F_1 = 0.774 x_1 + 0.661 x_2 + 0.114 x_3 + 0.669 x_4 + 0.873 x_5 + 0.859 x_6 + 0.503 x_7 + 0.901 x_8 + 0.930 x_9 + 0.837 x_{10} + 0.913 x_{11} + 0.863 x_{12} + 0.959 x_{13} + 0.968 x_{14} - 0.797 x_{15} + 0.971 x_{16} + 0.971 x_{17} + 0.846 x_{18}, \quad (1)$$

$$F_2 = -0.144 x_1 + 0.272 x_2 + 0.758 x_3 + 0.279 x_4 + 0.246 x_5 + 0.473 x_6 - 0.843 x_7 - 0.410 x_8 - 0.034 x_9 + 0.465 x_{10} - 0.176 x_{11} + 0.187 x_{12} - 0.023 x_{13} - 0.133 x_{14} + 0.456 x_{15} - 0.234 x_{16} - 0.039 x_{17} + 0.299 x_{18}, \quad (2)$$

$$F_3 = 0.399 x_1 - 0.268 x_2 - 0.492 x_3 + 0.559 x_4 - 0.367 x_5 - 0.194 x_6 - 0.182 x_7 - 0.137 x_8 - 0.148 x_9 + 0.277 x_{10} + 0.288 x_{11} - 0.202 x_{12} + 0.130 x_{13} - 0.162 x_{14} + 0.363 x_{15} + 0.005 x_{16} + 0.117 x_{17} + 0.322 x_{18}. \quad (3)$$

根据因子的贡献率,将每个变量的原数据代入,即可得到每个变量的综合评价得分,计算公式为:

$$Y = \frac{\lambda_1 F_1 + \lambda_2 F_2 + \lambda_3 F_3}{F} \quad (4)$$

综合评分结果如表 3 所示,评分越大则等级越高,评分越小则等级越低。

表 3 各个仓库烟叶醇化度综合评分

Table 3 Comprehensive score of tobacco leaf alcoholization degree in each warehouse

样本来源	第 6 个月	第 12 个月	第 18 个月	第 24 个月	第 30 个月	第 36 个月
成都多宝仓库	0.288	0.511	0.78	0.122	-0.376	-1.346
成都多宝仓库(气调)	0.134	-0.544	-1.530	-1.590	-2.049	2.777
什邡李子园空调库	-0.425	1.347	-1.519	2.454	3.364	3.638
西昌新厂库	0.592	0.605	0.935	1.617	-2.031	-2.888
平均醇化评分	0.147	0.480	-0.334	0.651	-0.273	0.545

2.2 评分等级的划分

取各个地区各个时间节点的综合指数的前 3 个四分位数,得到 3 个等级临界值,将醇化得分指数划分为 4 个

区间,即对各个仓库评分进行等级划分。划分区间即可确定每个仓库对应时间节点相应的烟叶醇化等级,见表 4。

表 4 烟叶醇化度等级评分

Table 4 Tobacco leaf alcohol degree grade

样本来源	不合格	合格	醇化良	醇化优
成都多宝仓库	$(-\infty, -0.252]$	$(-0.252, 0.205]$	$(0.205, 0.455]$	$(0.455, +\infty)$
成都多宝仓库(气调处理样单独抽)	$(-\infty, -1.575]$	$(-1.575, -1.037]$	$(-1.037, -0.036]$	$(-0.036, +\infty)$
什邡李子园空调库	$(-\infty, 0.018]$	$(0.018, 1.901]$	$(1.901, 3.140]$	$(3.140, +\infty)$
西昌新厂库	$(-\infty, -1.375]$	$(-1.375, 0.599]$	$(0.599, 0.852]$	$(0.852, +\infty)$

3 BP 神经网络在模拟应用中的训练与应用

BP 神经网络具有任意复杂的模式分类能力和优良的多维函数映射能力,解决了简单感知器不能解决的异或(Exclusive OR, XOR)和一些其他问题。从结构上讲, BP 网络具有输入层、隐藏层和输出层见图 1。从本质上讲, BP 算法就是以网络误差平方为目标函数、采用梯度下降法来计算目标函数的最小值^[14]。

根据 BP 神经网络模型原理,需将采集到烟叶醇化得分指标数据划分为训练集、检验集和预测集。具体而

言,将 4 个仓库中的前 4 个时间节点的样本数据作为网络模型的训练集,以后第 5 个时间节点(间隔 30 个月)的数据作为检测集,以最后时间节点(间隔 36 个月)作为预测集,对 4 个仓库的烟叶醇化度进行检验与预测。

3.1 BP 神经网络模型建立

采用 Matlab 软件对烟叶醇化得分模型进行训练。调用 Trainlm 函数网络对输入数据和期望输出值进行神经网络训练,训练策略采用调节训练次数和期望误差进行对比试验,使实际输出值和期望输出间的误差达到最小。当达到较合适的期望误差值,则网络的训练结束。确定较合适的值是通过不断调整隐含层神经元个数来确定的,一般认为增加隐含层可以降低网络误差,提高精度,但也会使网络复杂化,增加了网络的顺联时间和出现“过拟合”,通过试验可知增加隐含层神经元个数来获得较低的误差,其训练效果要比增加隐含层层数更容易实现^[15]。其输入数据为 4 个仓库中的前 4 个时间节点的 72 组数据,建立的 BP 神经网络模型的输入层节点有 11 个,试验对比确认网络误差最小时对应隐含层节点为 10;对应的期望输出值则是根据综合指数和烟叶醇化等级划分区间,该仓库不同种类的烟叶醇化等级,模型输出层节点有 4 个。当网络性能达到期望误差要求时训练结束,网络训练学习率为 0.01,将期望误差设置为 1.00E-

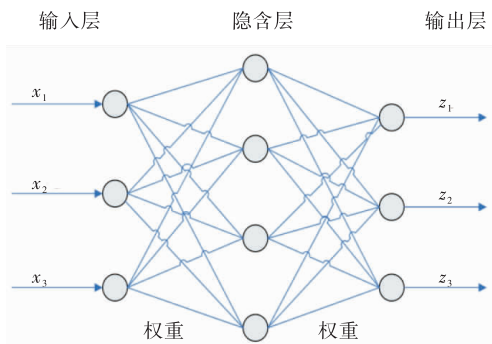


图 1 BP 神经网络映射结构图

Figure 1 BP neural network mapping diagram

06. 网络训练的误差拟合图见图 2, BP 神经网络训练界面见图 3。

3.2 BP 神经网络模型的检验

期望目标输出由主成分因子分析法得出的烟叶醇化综合指数转化而来, BP 神经网络模型检验通过比较实际输出和期望目标输出大小判断, 若实际输出小于期望目标输出, 判断模型准确性较高, 反之较低。采用样本数据中第 5 个时间节点(间隔 30 个月)数据作为检验集, 将其烟叶醇化等级状况与期望的目标输出进行检验, 误差较小则表示模型通过检测。由表 5 中模型的检验输出烟

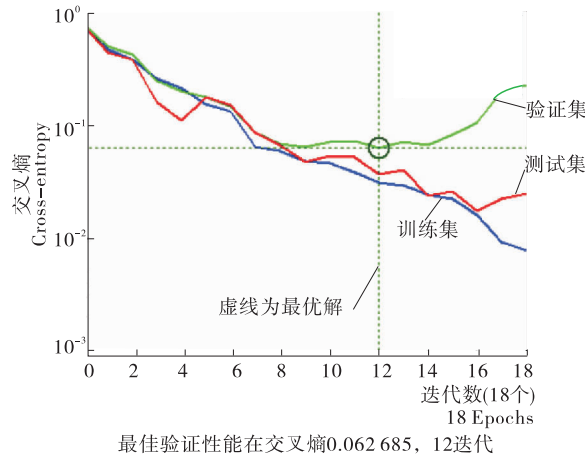


图 2 BP 神经网络误差拟合图

Figure 2 BP neural network error fitting graph

叶醇化等级状态与期望输出结果基本一致, 模型检验结果较满意, 表明网络有较好的预测效果。

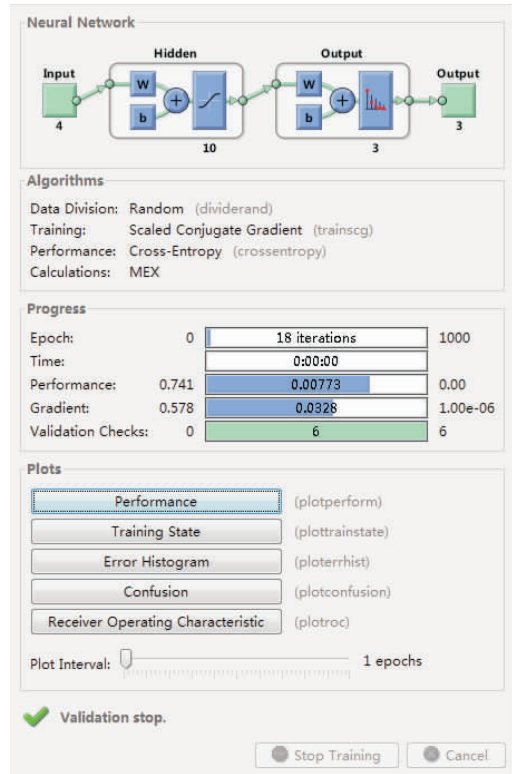


图 3 BP 神经网络训练

Figure 3 BP neural network training

表 5 BP 神经网络模型的检验输出与期望输出对比

Table 5 Comparison of test output and expected output of BP neural network model (interval of 30 months)

样本仓库	实际输出		期望目标输出转化值		
成都多宝仓库	0.910 81	0.000 21	0.039 84	0.000 73	1 000
成都多宝仓库(气调处理)	0.900 66	0.001 73	3.820 00E-05	0.001 31	1 000
什邡李子园空调库	0.990 14	0.136 34	0.000 94	0.000 12	1 000
西昌新厂库	1.000 00	0.000 82	0.000 59	3.380 00E-05	1 000

3.3 基于 BP 神经网络的预测

如神经网络原理所述, 神经网络第 n 年的输入值对应的期望目标输出值则是第 $n+1$ 个月的烟叶醇化等级状况, 这就使网络模型训练得到的输出值具有预测下一年醇化等级的功效^[16]。因此, 将间隔 36 个月数据输入模型, 则可得间隔 42 个月的 4 个仓库 18 个不同产地等级烟叶的预测结果, 预测结果以成都多宝库为例见表 6。

根据结果可知, 在下次抽检过程中, 烟叶醇化为优的产地和种类有: 福建三明 B2F、福建三明 C3F、云南罗平 B2F、C3F、重庆丰都 C3F、B2F、X2F, 可以预测这些产地不同等级烟叶在成都多宝库经 42 个月醇化水平较高, 其感官评吸预测结果也较好, 配方设计使用价值高。处于不

合格或刚合格的烟叶醇化级别的为: 福建三明 X2F、广东南雄 X2F、云南罗平 C3F 以及凉山会理 B2F, 可能是由于不同地区烟叶种植条件环境不同, 烟叶化学成分差异性较大造成, 在成都多宝库条件下醇化在这一时间节点处于低水平状态, 使用价值欠佳, 因此, 建议尽快处理, 并需进一步探讨合理的醇化环境和方式。

4 结论

烟叶醇化过程中发生复杂的生化、物理反应, 这一类问题的共同特点是: 难以达到解析解, 而神经网络具有识别复杂非线性可反映的特征, BP 神经网络能学习和储存大量输入—输出模式映射关系, 而无须事先了解描述这类映射关系的数学方程。通过 Matlab 软件, 建立烟叶

表6 成都多宝仓库烟叶醇化结果分析

Table 6 Analysis of alcoholification results of tobacco leaves in Chengdu Duobao warehouse

醇化度等级	产地种类
不合格	凉山会理 C3F
	河南洛阳 C3F
合格	云南罗平 C3F
	凉山会理 B2F
醇化良	广东南雄 X2F、C3F、B2F
	河南洛阳 B2F、X2F
	福建三明 X2F
	凉山会理 X2F
醇化度优	福建三明 B2F、C3F
	云南罗平 B2F、C3F
	重庆丰都 C3F、B2F、X2F

醇化感官质量 BP 神经网络模型,应用在时间节点质量预测方面,从检验结果可知具有准确性,为烟叶醇化质量评价和预测提供了科学的实践方法和可靠的理论依据,可指导烟叶仓储养护工作者、卷烟配方师提前采取风险预控措施,规避由醇化不当导致烟叶质量下降所带来的损失。

参考文献

- [1] 王瑞新. 烟草化学[M]. 北京: 中国农业出版社, 2010: 201-202.
- [2] 刘红光, 胡玲, 颜亮亮, 等. 初烤烟叶自然纯化过程中内在化学成分变化研究[J]. 江西农业学报, 2015, 27(8): 71-75.
- [3] 王勇, 周浩. 醇化时间对烟叶香味成分的影响研究[J]. 安徽

- 农学通报, 2012, 18(6): 131-132.
- [4] 巩效伟, 段焰青, 黄静文, 等. 烤烟主要化学成分与烟叶等级和醇化时间的相关性研究[J]. 江西农业大学学报, 2016, 32(1): 31-34.
- [5] 王涛, 刘磊, 史绍新, 等. 复烤片烟自然醇化过程中主要雪花成分含量与感官质量的关系[J]. 河南农业科技, 2015, 44(9): 129-134.
- [6] 杜咏梅, 郭成芳, 张怀宝, 等. 水溶性糖、烟碱、总氮含量与烤烟吃味品质的研究[J]. 中国烟草科学, 2000(1): 7-10.
- [7] 闫铁军. 不同产区烤烟配方模块在醇化过程中质量的变化趋势及醇化周期研究[D]. 郑州: 河南农业大学, 2009: 30-35.
- [8] 杨波, 荔亮. 烟叶适宜醇化时间分析[J]. 现代农业科技, 2016(14): 29-30.
- [9] 刘红光, 胡玲, 赵斌, 等. 初烤、复烤烟叶协同醇化对烟叶品质的影响[J]. 烟草科技, 2017, 50(7): 31-39.
- [10] 李东亮. 数据处理方法及其在烟草质量评价中的应用[M]. 郑州: 郑州大学出版社, 2014: 187-191.
- [11] 杨淑娥, 黄礼. 基于BP神经网络的上市公司财务预警模型[J]. 系统工程理论与实践, 2005, 25(1): 12-18.
- [12] 申玉军, 邓国栋, 陈良元, 等. 一种烟草感官评价分析方法的建立及应用[J]. 烟草科技, 2011(5): 17-20.
- [13] 杨维忠. SPSS 统计分析 with 行业应用[M]. 北京: 清华大学出版社, 2011: 169-170.
- [14] 阮秀凯. 现代无线通信系统盲处理技术新进展 基于智能算法[M]. 上海: 复旦大学出版社, 2015: 28.
- [15] 吴艳. 配水管网系统中消毒副产物的研究[D]. 哈尔滨: 哈尔滨工业大学, 2006: 67-71.
- [16] 朱大奇. 神经网络原理及应用[M]. 北京: 科学出版社, 2006: 28-47.

信息窗

制茶不再凭经验 红茶“发酵”可量化

据科技日报讯记者从安徽农业大学获悉,该校茶与食品科技学院宁井铭教授等研究人员,采用近红外光谱技术结合计算机视觉系统,建立了一种简单、客观的红茶“发酵”程度综合评价方法。日前,该研究成果发表在《LWT-食品科学与技术》期刊上。

红茶是世界上最受欢迎的饮料之一,约占全球茶叶贸易量的78%。红茶是由茶叶鲜叶经萎凋、揉捻、发酵、干燥等工序加工而成,其中发酵是红茶加工中最关键的工序之一,在一定的温度、湿度条件下,揉捻后的茶叶发生一系列变化,既涉及茶叶内含物质的化学变化,又有外观的物理变化,发酵不足或过度均会影响红茶的品质。因此,在加工中正确判断红茶发酵程度至关重要。

目前,在生产加工中红茶发酵程度的判别完全依赖

人工经验,生产者通常根据叶的颜色和气味来进行主观评价,缺乏量化的评价标准,操作者的经验、心理状态、环境等因素都会影响评价结果,从而导致产品质量不统一,无法实现标准化生产。因此,精准、量化地判别红茶发酵程度是制约红茶标准化加工的技术瓶颈之一。

科研人员采用近红外光谱技术监测茶叶在发酵过程中主要内含化学成分的变化,利用计算机视觉系统监控茶叶外观颜色的变化,在信息融合的基础上,结合化学计量学方法建立了一种红茶发酵程度判别模型,该模型具有精准、量化的特点,能够在1 min之内完成判别。

据悉,该技术应用后,将有力推动茶叶加工的智能化和标准化水平,也为其他农产品标准化加工提供了借鉴。

(来源: <http://news.foodmate.net>)