

新的 Hopfield 神经网络分类器 在葡萄酒质量评价中的应用

Application of New Hopfield Neural Network Classifier in the Quality Evaluation of Grape Wine

刘延玲 Liu Yanling

(天津大学系统工程研究所, 天津 300072)

(Institute of Systems Engineering of Tianjin University, Tianjin 300072, China)

摘要: 建立一种新的 Hopfield 神经网络分类器模型, 该模型通过训练单层前向神经网络来设计, 数据兼容性强, 可以直接处理来自 UCI 数据库的葡萄酒的理化性质测试指标数据和专家的感官评价等级数据, 实现葡萄酒质量分类。仿真结果表明, 该分类器设计简单, 耗时短, 分类效果明显。

Abstract: A model of the quality evaluation of grape wine based on Hopfield neural network classifier is proposed in this paper. The design of model is achieved by training a single-layer feedforward network. The model is compatible with data from UCI dataset. The data are composed of physicochemical properties and sensory test results of wine. The simulation results show that the design of classifier model is simple and it achieves effective class outcome.

关键词: Hopfield 神经网络; 分类器; 质量评价; 感官评价

Key words: Hopfield neural network; classifier; quality evaluation; sensory test

中图分类号: TP183

文献标识码: A

文章编号: 1006-4311(2012)02-0181-02

0 引言

现在, 葡萄酒因其特殊的营养价值和较好的保健效果, 越来越受到中国广大消费者的欢迎。在此形势下, 葡萄酒认证和质量评价得到关注。葡萄酒认证保证了市场中酒的质量, 同时保护了消费者的利益。葡萄酒的认证包括理化性质分析和感官评价两种手段。质量评价是认证中的重要阶段, 它有益于提高葡萄酒的酿造工艺, 同时为市场定位的提供决策信息。葡萄酒的质量分类既考虑理化性质分析(酒精度、柠檬酸和 PH 值等指标), 又要考虑品酒师的感官评

价, 是比较复杂的任务^[1]。

国外对葡萄酒质量评价的研究成果比较丰富, 其中 Cortez^[1] 用支持向量机建立了葡萄酒质量分类模型, 分级效果较好; Moreno^[2] 基于红酒的矿物质元素含量, 用概率神经网络成功将 54 个样本分为两类。国内对此课题的研究多集中在统计学方法分析和专家的感官评价, 文献[3]用分析软件 SPSS 对 40 个样本进行了聚类分析, 文献[4]中论述了感官评价在葡萄酒加工工艺、成分分析和评价结果统计分析等方面的研究进展, 文献[5]提出了一种可视化的葡萄酒质量评价方法。本文将构建一类新的 Hopfield[6]神经网络分类器模型, 利用葡萄酒的理化性质指标和感官评价结果将其质量进行分类。该分类器设计简单, 分类耗时比较短, 能取得好的分类效果。

作者简介: 刘延玲(1986-), 女, 山东潍坊人, 硕士生, 天津大学系统工程研究所, 研究方向为神经网络设计 and 应用。

同功能的节点组成监控网络, 这大大增加了节点部署问题的复杂性。因此, 有必要对不同种类不同功能的节点进行分类, 并对其成本、物理特性进行描述, 建立节点类型和功能数据库。在满足给定的监控任务下, 以系统建设成本和网络服务质量为目标, 研究具体的节点选择方案。

移动感知节点的部署问题是研究重点, 将多种形式的移动节点部署问题归纳为多目标组合优化问题或节点移动的行为决策问题, 包括纯移动节点的部署问题和混合网络节点的部署问题, 针对不同要求, 给出具体的部署问题框架和模型及具体的部署算法。主要包括: ①障碍物模型的建立及避障算法进行研究; ②将多种形式的移动节点部署问题归纳为多目标组合优化问题, 建立以节点移动距离、节点剩余能量等为目标, 以网络连通和区域覆盖为约束条件的部署问题数学模型, 研究集中式部署策略; ③将多种形式的移动节点部署问题归纳为节点移动的行为决策问题, 在满足多级连通和多级覆盖的条件下, 每个节点通过设计的规则, 决定自身的移动策略, 进而研究其分布式部署策略; ④由于监控环境的复杂多变, 网络往往很难达到稳态, 为了防止节点的振动, 研究带阈值的移动节点部署策略; ⑤为了实现多级连通和多级覆盖, 监控区域边界处需要部署多个节点, 增加监控区域边界效应的研究。

2 组网优化策略

根据监控区域性状和面积、节点的理想分布函数、节点的感知半径和通信半径及节点数目, 采用概率感知模型, 利用数学方法计算出区域 M-覆盖的概率 $P(\text{COV}_M)$, 再利用图论方法计算出网络 N-连通的概率 $P(\text{CON}_N)$ 。再分别计算出理想部署状态下: ①对于给定节点类型和监控区域的形状和面积, 在满足多级连通(M -连通, $P(\text{COV}_M) > 0.99$)和多级覆盖(N -覆盖, $P(\text{CON}_N) > 0.99$)要求的情况下, 所需节点的最少数目; ②对于给定节点数目和节点类型, 在满足 $P(\text{COV}_M) > 0.99$ 和 $P(\text{CON}_N) > 0.99$ 要求的情况下, 最多能覆盖多大的监控区域。

对不同种类不同功能的节点进行分类, 并对其成本、物理特性进行描述, 建立节点类型和功能数据库。按照不同监控任务需求, 建立以网络建设费用、CPU 响应时间和数据服务为目标, 以节点能耗、

带宽和存储能力为约束条件的节点、操作系统选择数学规划模型或多目标优化模型, 设计相应的组合优化算法统一框架, 得出最佳的网络节点和操作系统的选型, 以满足特定监控任务。

主要解决两类问题的部署优化。一是对监控区域已知, 并能部署一个强大的核心控制节点的情况下, 事先将选择的节点部署于一个较小的区域, 一旦紧急事件发生, 这些节点会主动移动至监控区域完成监控任务。这类问题的研究采用集中式部署策略, 将部署问题描述为多目标优化问题, 由控制节点执行部署算法并计算出移动感知节点的目标位置。其中多目标优化数学模型由最小化节点移动距离、最小化节点剩余能量和最小化部署时间等, 约束条件由多级连通和多级覆盖以及任一节点的邻居节点数目和邻居节点间距离组成, 求解算法采用进化计算、协同进化等方法。二是针对无法确定监控环境或无法部署一个强大的核心控制节点的情况下, 有空中随机抛洒的形式将节点部署于监控区域, 再利用节点分布式自组织算法得到精确部署方案。

3 总结与结论

主要针对面向工业生产监控中移动感知节点的部署优化及行为决策问题进行探讨, 包括系统部署框架、节点选择、适合不同场景应用的集中式和分布式部署算法、节点任务管理、网络资源管理等关键技术, 其主要的目标是用管理学方法解决其中的优化和决策问题, 通过建模、求解、仿真等过程给出最佳的部署方案, 以达到网络能耗、生命周期、可靠服务等最优化目的, 对无线移动感知监控网络的广泛应用具有重要的指导意义。

参考文献:

- [1] C. Song, M. Guizani, H. Sharif, Adaptive clustering in wireless sensor networks by mining sensor energy data, Computer Communications, 30 (14-15) (2007) 2968-2975.
- [2] 易本顺, 陈杰, 肖进胜, 无线传感器网络优化的任务管理算法研究, 电子与信息学报, 32(11) (2010) 2606-2611.
- [3] 缪强, 郑扣根, 无线传感器网络的路由协议设计研究[J], 计算机应用研究, 2004.08.

1 Hopfield 神经网络分类器

Hopfield 网络是一种反馈型神经网络,具有联想记忆的特征,可以作为一类有效的分类器^[6]。Hopfield 神经网络可用下面的常微分方程组来描述 $\dot{U} = -AU + WF(U) + \Theta$ (1)

其中 $U = [u_1, u_2, \dots, u_n]^T \in R^n$ 表示神经元状态向量, $\Theta = [\theta_1, \theta_2, \dots, \theta_n]^T \in R^n$ 是实常向量, $A = \text{diag}[a_1, a_2, \dots, a_n] \in R^{n \times n}$ 是正定对角阵, $W \in R^{n \times n}$ 表示突触连结权值矩阵, $F(U) = [f_1(u_1), f_2(u_2), \dots, f_n(u_n)]^T$ 表示神经元激活函数, $f_i(u)$ ($i=1, 2, \dots, n$)可微且满足

$$0 < \frac{f_i(x) - f_i(y)}{x - y} \leq k, \forall x, y \in R。$$

Zheng 等在文献 [7] 中提出了一种基于单层前向神经网络的 Hopfield 记忆存储器设计方法。该单层前向神经网络只有一个输入层和一个输出层,其中 $p = [p_1, p_2, \dots, p_{m_1}]^T$ 是输入向量, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_{m_2}]^T$ 是输出向量, $b = [b_1, b_2, \dots, b_{m_2}]^T$ 是输出层的阈值, $w \in R^{m_1 \times m_2}$ 是权值矩阵,输出层的传输函数定义为显性函数 $\psi(x) = x$ 。通过训练上述定义单层前向神经网络,将给定的样本数据作为平衡点被网络(1)存储。详细的规则如下^[7] ①选择训练数据对 $T = \{(P^j, AU^j)\}_{j=1}^m$, 其中 U^j 是要记忆存储的数据, $P^j = F(U^j)$ 是输入向量, AU^j 是目标向量。②选取 A 和 F 。③用函数 newff 构建单层前向神经网络。④训练③中构建的网络,得出 w 和 b 。⑤设置 $W = w$ 和 $\Theta = b$, 即 $AU^j = wP^j + b = WF(U^j) + \Theta$ 成立。

2 分类器应用实例

2.1 葡萄酒质量分类数据 葡萄酒质量分类的数据来源于葡萄酒的认证过程,包括红葡萄酒和白葡萄酒两种,选取 UCI 数据库^①。这两种酒的理化性质指标均为 11 个,用 β_i ($i=1, 2, \dots, 11$)分别表示,详细为固定酸度(酒石酸)、挥发性酸(乙酸)、柠檬酸、残糖、氯化物、游离二氧化硫、总二氧化硫、密度、PH 值、硫酸盐、酒精浓度^[5]。第 12 项为专家的感官评价结果,即葡萄酒的质量等级,表示为 $\hat{\beta}$,分为 0 级(质量最差)到 10 级(质量最好)。数据库中的红葡萄酒分为 6 个等级,质量评价是 3, 4, \dots , 8; 白葡萄酒分为 7 个等级,质量评价是 3, 4, \dots , 9。

2.2 分类过程和结果 首先将葡萄酒数据进行预处理,将红葡萄酒和白葡萄酒的理化性质指标 $\beta_5, \beta_6, \beta_7$ 分别做如下处理: $\beta_5^* = 100 \times \beta_5$, $\beta_6^* = 0.1 \times \beta_6$, $\beta_7^* = 0.01 \times \beta_7$, 其中 β_i^* , $i=5, 6, 7$ 是处理后的数据。进一步的,分类的依据是感官评价结果 $\hat{\beta}$ 不小于 5 的质量等级定义为好,反之为一,分别用 2, -2 表示。其次选取红葡萄酒和白葡萄酒数据分别作为单层前向神经网络的训练样本,其详细情况为表 1 和表 2 所示。网络参数具体设置如下: $A = \text{diag}[1, 1, \dots, 1]$, $f_i(u) = \tanh(u)$, $i=1, 2, \dots, 12$, 生成训练样本对 $T = \{(P^j, AU^j)\}_{j=1}^2$, 其中 U^j 分别是表 1 和表 2 处理后需要记忆存储的数据, $P^j = F(U^j)$ 是输入向量, AU^j 是目标向量。最后是执行网络的训练以及选择,即执行设计规则③~⑤。

为测试分类器的效果,分别将数据库中的其它数据作为分类测试样本,即作为网络的初始状态输入,观察是否收敛到训练样本数据。分类结果表明,测试样本数据实现了正确的分类。图 1 表明质量等级 $\hat{\beta}=8$ 的红葡萄酒数据收敛到表 1 中质量等级 $\hat{\beta}=7$ 的训练样本数据上,即质量分类为好;图 2 表明质量等级 $\hat{\beta}=4$ 的红葡萄酒数据收敛到表 1 中质量等级 $\hat{\beta}=3$ 的训练样本数据上,即质量分类为一般;同样,图 3 表明质量等级 $\hat{\beta}=9$ 的白葡萄酒数据收敛到表 2 中质量等级 $\hat{\beta}=7$ 的训练样本数据上,即质量分类为好;图 4 表明质量等级 $\hat{\beta}=4$ 的红葡萄酒数据收敛到表 2 中质量等级 $\hat{\beta}=3$ 的训练样本数据上,即质量分类为一般。进一步的,图 1~图 4 均表明收敛时间短,在第 10 步时已经实现。

3 结束语

基于 Hopfield 神经网络的分类器成功实现了葡萄酒质量的评价分类,而且评价的数据是葡萄酒的理化性质测试数据和专家的感官评价结果,更有说服力。这种分类器设计简单,耗时短,可以考虑作为信息决策系统的一个模块,进一步挖掘指标变量之间的关系,

表 1 红葡萄酒训练样本数据

β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	$\hat{\beta}$
8.5	0.28	0.56	1.8	0.092	35	103	0.9969	3.3	0.75	10.5	7
11.6	0.58	0.66	2.2	0.074	10	47	1.0008	3.25	0.57	9	3

表 2 白葡萄酒训练样本数据

β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	$\hat{\beta}$
5	0.3	0.33	3.7	0.03	54	173	0.9887	3.36	0.3	13	7
6.2	0.23	0.35	0.7	0.051	24	111	0.9916	3.37	0.43	11	3

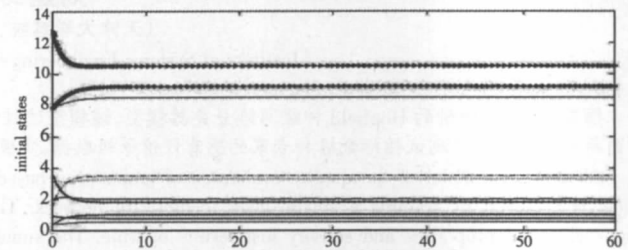


图 1 红葡萄酒测试样本分类结果

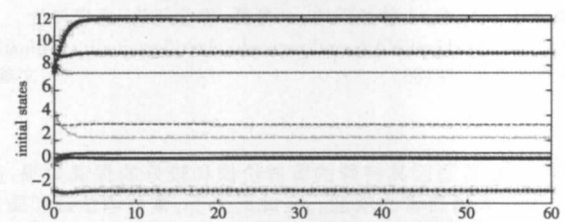


图 2 红葡萄酒测试样本分类结果

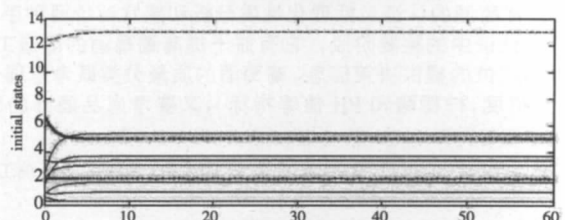


图 3 白葡萄酒测试样本分类结果

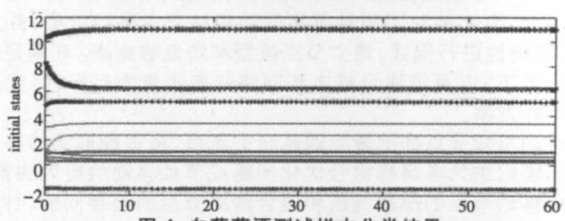


图 4 白葡萄酒测试样本分类结果

从而为葡萄酒的生产、工艺改进以及市场定位提供决策信息。

注释:①数据库链接 <http://www.ics.uci.edu/~mllearn/MLRepository.html>。

参考文献:

- [1]Cortez P, Cerdeira A, Almeida F, et al. Modeling wine preferences by data mining from physicochemical properties [J]. Decision Support Systems, 2009, 47(4): 547-553.
- [2]Moreno I, González-Weller D, Gutierrez V, et al. Differentiation of two Canary DO red wines according to their metal content from inductively coupled plasma optical emission spectrometry and graphite furnace atomic absorption spectrometry by using probabilistic neural networks [J]. Talanta, 2007, 72 (1): 263-268.
- [3]李运,李记明,姜忠军.统计分析在葡萄酒质量评价中的应用[J].酿酒科技,2009 (4):79-82.
- [4]王文静.感官评价在葡萄酒研究中的应用[J].酿酒,2007,34(4):57-59.
- [5]王金甲,尹涛,李静等.基于物理化学性质的葡萄酒质量的可视化评价研究[J].燕山大学学报,2010,34(2):133-137.
- [6]Hopfield J J. Neurons with graded response have collective computational properties like those of two-state neurons [J]. Proceedings of the National Academy of Sciences of the United States of America - Biological Sciences, 1984, 81(10):3088-3092.
- [7]Zheng P S, Zhang J X, Tang W S. Learning associative memories by error backpropagation [J]. IEEE Transactions on Neural Networks, 2011, 22(3): 347-355.