

基于 FNN 模型的决策算法研究

侯庆山, 邢进生

(山西师范大学 数学与计算机科学学院, 山西 临汾 041000)

摘要: 鉴于证据理论对样本分类和决策过程的复杂性以及不稳定性, 提出了一种基于神经网络模型和模糊集理论的样本决策算法。为了降低样本分类和决策过程的复杂性, 增强算法的稳定性和适用性, 在该算法中, 设计并提出了一种新的隶属度函数。应用提出的隶属函数对相关数据集样本进行模糊化处理, 得到数据集的模糊化矩阵, 其中输入样本数据与不同样本类别的隶属度相关联。根据隶属度矩阵, 并结合性能较好的激活函数 Swish-B, 通过神经网络分类器, 样本将被归属于特定的类。基于鸢尾花数据集对其进行可视化分析, 将该方法与传统的证据理论及相关改进算法进行比较, 验证了所设计的隶属度函数具有良好性能, 同时实验结果证明了该算法的合理性与有效性, 算法过程更为简单, 鸢尾花数据集的分类准确率高达 98%。

关键词: 复杂性; 稳定性; 证据理论; 样本分类; 神经网络; 模糊集理论; 隶属度

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2020)12-0092-07

doi:10.3969/j.issn.1673-629X.2020.12.017

Research on Decision Algorithm Based on Fuzzy Neural Network

HOU Qing-shan, XING Jin-sheng

(School of Mathematics and Computer Science, Shanxi Normal University, Linfen 041000, China)

Abstract: In view of the complexity and instability of evidence theory in sample classification and decision-making, a sample decision algorithm based on neural network model and fuzzy set theory is proposed. In order to reduce the complexity of sample classification and decision-making process and enhance the stability and applicability of the algorithm, a new membership function is designed and proposed to blur the relevant data set samples to obtain the fuzzy matrix of the data set. The input sample data is associated with the membership of different sample categories. According to the membership matrix, the samples will be assigned to a specific class by the neural network classifier. Based on iris data set and visual analysis, this method is compared with traditional evidence theory and related improved algorithms. It is proved that the new membership function has excellent performance. Experiment shows that the proposed algorithm is reasonable and effective with simpler process. The classification accuracy of the iris data set is as high as 98%.

Key words: complexity; stability; evidence theory; sample classification; neural networks; fuzzy set theory; membership

0 引言

数据挖掘技术的不断发展和完善,吸引了相关信息产业和大量研究人员的注意。随着计算机相关技术的不断成熟和进步,数据量变得越来越庞大,更多的数据可以被利用。利用相关的数据挖掘技术可以将这些数据转化为更有意义的信息和知识。收集到的信息可以应用到市场调查、生产控制、进化分析和科学探索^[1-2]等领域。样本分类是一种重要的数据挖掘技术,涉及到从现实世界的大型数据库中提取出有效的知识模式。与有效知识数据挖掘的相关研究是非常重要的,决策过程的复杂程度以及决策结果的好坏往往

取决于所用的分类方法是否有效。

由于输入数据的不精确性、模糊性、类别间边界的重叠性以及特征定义的不明确性,在对数据样本进行分类时的任何阶段都可能存在不确定性。传统的样本分类算法中证据理论是模糊决策的有效工具,在对样本分类时,往往需要对样本的特征进行筛选、降维以及为样本分配权重等处理。这些处理过程往往需要人工的参与,因此传统的样本分类算法通常是比较复杂的。首次出现的 Dempster-Shafer 证据理论^[3]被认为是智能系统中推理不确定性信息的有效模型,但当高度冲突的证据被组合起来时,样本数据的分类结果往往是

收稿日期:2020-01-18

修回日期:2020-05-19

基金项目:山西省软科学基金资助项目(2011041033-03)

作者简介:侯庆山(1995-),男,硕士研究生,研究方向为数据挖掘、计算智能;邢进生,教授,博士,研究方向为计算智能、数据挖掘、计算机视觉。

错误的^[4-6]。Murph等人^[7]提出了一种基于计算所有证据平均值的证据组方法,但并没有考虑证据权重对于融合结果的影响,将所有的权重设定为同一值。Han等人^[8]提出了一种改进的基于证据距离测量的平均组方法,每个证据主体的权重都被考虑在内。但其组合规则不能有效地处理高度冲突的证据组合。Deng等人^[9]提出了一种新的加权平均证据组方法,但随着样本特征增多,利用不确定性度量来确定证据主体的权重变得困难。

总体上看,数据样本分类^[10-11]就是发现区分数据类型的相关模型算法,以便模型能够预测具有未知类标签值的样本。人工神经网络^[12-14]是一种流行的数据建模工具,可以执行类似于人脑的智能任务。多层感知器反向传播网络^[15-16]算法是神经网络领域中进行数据分类的可靠方法之一。径向基函数神经网络^[17-18]是另一种利用径向基函数作为激活函数的强大神经网络模型。此外,针对输入样本数据的不确定性,基于经典集合论的模糊集合论^[19-21]在处理现实生活中的不确定性方面是非常灵活的。在模糊系统中,特征与不同样本类别的隶属度相关联。

该文提出的算法与传统的证据理论算法相比,利用神经网络模型并结合模糊理论对数据集进行有效的分类,避免了人工参与样本特征分配权重、特征选择等

问题,过程更为简单,效果更加良好。神经网络处理数值和定量数据,模糊系统处理符号和定性数据。该算法有效地结合了这两种方法的优势,利用基于模糊集理论的方法将模糊系统的类人逻辑推理与神经网络的学习相结合,样本分类的算法效率大幅提升。

1 基础理论

1.1 模糊理论

1.1.1 模糊因素集和模糊因素权重集

模糊因素集被定义为 $\Omega = \{b_1, b_2, \dots, b_m\}$, 其中 $b_i, i \in [1, m]$ 表示第 i 个决定分类结果的模糊因素。 $W = \{w_1, w_2, \dots, w_m\}$, 其中 $w_i, i \in [1, m]$ 表示第 i 个影响因素对分类结果的影响程度。

1.1.2 隶属度函数

模糊决策的结果与隶属度函数存在紧密的联系,隶属度函数定义如下:对于模糊数据集 Ω 中的任意元素 b_i , 都有一个数 $F(b_i) \in [0, 1]$ 与其相对应,当 b_i 在 Ω 中变动时, $F(b_i)$ 称作隶属度函数。隶属度的大小表明了输入样本属于该类可能性的多少。

1.2 前馈神经网络模型

1.2.1 结构组成

前馈神经网络由多层连续的非线性函数构成,包含输入层、隐含层、输出层三部分。结构如图1所示。

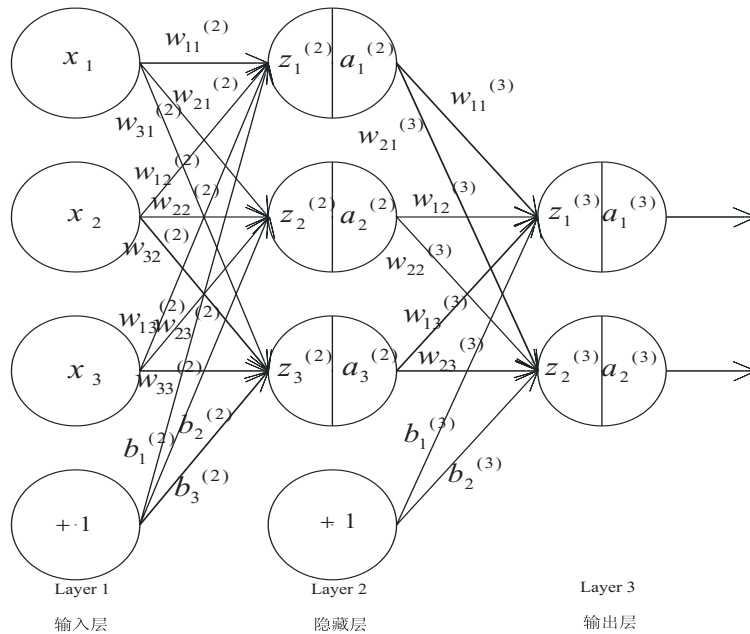


图1 前馈神经网络结构

1.2.2 前向传播

前馈神经网络可通过式(1)进行传播。通过逐层的网络传递,由输出层将最终结果输出。网络的传递及输出过程可由式(2)表达。

$$\begin{cases} Z^{(l)} = W^{(l)} * a^{(l-1)} + b^{(l)} \\ a^{(l)} = f_i(z^{(l)}) \end{cases} \quad (1)$$

其中, $f_i(\cdot)$ 表示第 l 层网络节点的激活函数, $W^{(l)} \in R^{m^{(l)} \times m^{(l-1)}}$ 表示第 $l-1$ 层和第 l 层之间的权值矩阵, 其中 $m^{(l)}$ 表示第 l 层的节点个数。 $b^{(l)} \in R^{m^{(l)}}$ 表示第 $l-1$ 层和第 l 层之间的偏置值。 $z^{(l)} \in R^{m^{(l)}}$ 表示第 l 层网络节点的输入值, $a^{(l)} \in R^{m^{(l)}}$ 表示第 l 层网络节点的输出值。

$$\begin{aligned} x &= a^{(0)} \rightarrow z^{(1)} \rightarrow a^{(1)} \rightarrow z^{(2)} \rightarrow \cdots \rightarrow a^{(L-1)} \rightarrow \\ &z^{(L)} \rightarrow a^{(L)} \end{aligned} \quad (2)$$

其中, L 表示网络的层数。

1.2.3 反向传播

对于输入样本 (X, y) , 网络损失采用交叉熵损失, 损失定义如式(3)所示。

$$F_{\text{loss}}(y, \hat{y}) = -y^T \log \hat{y} \quad (3)$$

假定训练集 $\Omega = \{(X^{(n)}, y^{(n)})\}_{n=1}^N$, 将训练集数据输入到网络中得到网络输出 $\hat{y}^{(n)}$, 在训练集 Ω 上的结构化风险函数可由式(4)定义。

$$R(W, b) = \frac{1}{N} \sum_{n=1}^N F_{\text{loss}}(y^{(n)}, \hat{y}^{(n)}) + \frac{1}{2} \lambda \|W\|_F^2 \quad (4)$$

$$\|W\|_F^2 = \sum_{l=1}^L \sum_{i=1}^{m^{(l)}} \sum_{j=1}^{m^{(l-1)}} (W_{ij}^{(l)})^2 \quad (5)$$

其中, W 和 b 分别表示网络中的权重矩阵与偏置向量, $\|W\|_F^2$ 表示用于防止过拟合的正则化项, λ 为可调节超参数。

通过随机梯度下降法对网络参数进行学习和调整, 在网络的迭代训练过程中, 第 l 层权重和偏置参数的更新可由式(6)、式(7)表达。

$$W^{(l)} \leftarrow W^{(l)} - \alpha \frac{\partial R(W, b)}{\partial W^{(l)}} = \quad (6)$$

$$W^{(l)} - \alpha \left(\frac{1}{N} \sum_{n=1}^N \frac{\partial F_{\text{loss}}(y^{(n)}, \hat{y}^{(n)})}{\partial W^{(l)}} + \lambda W^{(l)} \right)$$

$$b^{(l)} \leftarrow b^{(l)} - \alpha \frac{\partial R(W, b)}{\partial b^{(l)}} = \quad (7)$$

$$b^{(l)} - \alpha \left(\frac{1}{N} \sum_{n=1}^N \frac{\partial F_{\text{loss}}(y^{(n)}, \hat{y}^{(n)})}{\partial b^{(l)}} \right)$$

其中, α 表示学习率。

2 算法设计与实现

2.1 隶属度函数设计

考虑 N 个输入样本, 每个输入样本包含 k 个属性, 输入样本共属于 M 个类别。定义输入样本集 $S = \{s_1, s_2, \dots, s_n\}$, 样本属性集 $A = \{a_1, a_2, \dots, a_k\}$, 样本类别集 $C = \{c_1, c_2, \dots, c_M\}$ 。定义的样本属性隶属度函数可由式(8)表达。

$$f(a_{ji}; e, l, z) = \frac{1}{1 + |\frac{a_{ji} - z}{e}|^{2l}} \quad (8)$$

$$i \in [1, k], j \in [1, n]$$

式中, 参数 z 决定了隶属度函数的中心位置, 参数 e, l 控制交界点处的斜率。隶属度函数的形状及位置由参数 e, l, z 决定, 通过调控相关参数得到满足需求的属性隶属度函数。输入样本某一类别的隶属度可由式(9)表达。

$$f_{jm}(s_j; a_{j1}, a_{j2}, \dots, a_{jk}) = \sum_{i=1}^k f(a_{ji}; e, l, z) \quad (9)$$

$$j \in [1, n], i \in [1, k], m \in [1, M]$$

对输入样本的类别隶属度进行归一化操作, 可由式(10)表达。

$$f_{jm}(s_j; a_{j1}, a_{j2}, \dots, a_{jk}) = \frac{f_{jm}(s_j; a_{j1}, a_{j2}, \dots, a_{jk})}{\sum_{m=1}^M f_{jm}(s_j; a_{j1}, a_{j2}, \dots, a_{jk})} \quad (10)$$

2.2 基于模糊理论的神经网络设计

(1) 求解输入样本的隶属度矩阵。

假定包含 N 个输入样本 s 以及 M 个预测类别 c , 每个样本包含了 k 个样本特征 a 。

输入样本矩阵为 $S = [s^1, s^2, \dots, s^n]^T$, 对任一输入样本 S_i 包含 k 个特征。调节隶属度函数的超参数, 根据预测类别设置 M 个具有不同参数的隶属度函数, 将输入样本分别带入所设置的 M 个隶属度函数。隶属度矩阵 $D_{f_m(s)}$ 可由式(11)表达。

$$D_{f_m(s)} = \begin{bmatrix} f_{1,1}(s_1) & f_{1,2}(s_1) & f_{1,3}(s_1) & \cdots & f_{1,M}(s_1) \\ f_{2,1}(s_2) & f_{2,2}(s_2) & f_{2,3}(s_2) & \cdots & f_{2,M}(s_2) \\ f_{3,1}(s_3) & f_{3,2}(s_3) & f_{3,3}(s_3) & \cdots & f_{3,M}(s_3) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_{N,1}(s_N) & f_{N,2}(s_N) & f_{N,3}(s_N) & \cdots & f_{N,M}(s_N) \end{bmatrix} \quad (11)$$

(2) 设置神经网络激活函数与网络损失。

在输入样本前向传播的过程中, 样本信息可由式(12)进行表达, 将输入信息 z 经过网络节点的激活函数, 得到网络节点的激活值。为简化计算, 神经网络的激活函数选取 ReLU 与 Swish-B 激活函数, 由式(13)定义表达。

$$z = \sum_{i=1}^d w_i x_i + b = W^T X + b \quad (12)$$

式中, $W = [w_1, w_2, \dots, w_n] \in R^d$ 为权重向量, b 表示偏置值。

$$\text{Swish-B} = x * \text{Sigmoid}(b * x) = \frac{x}{1 + e^{-bx}}$$

$$\text{ReLU} = \begin{cases} x & x \geq 0 \\ 0 & x \leq 0 \end{cases} = \max(0, x) \quad (13)$$

针对样本分类问题, 神经网络的损失采用交叉熵损失函数, 可由式(14)表达。

$$l(f, y) = - \sum_i^n \sum_j^m y_{ij} \log f(x_{ij}) \quad (14)$$

式中, n 表示输入样本的数量, m 表示输入样本种类, y_{ij} 表示第 i 个样本属于分类 j 的标签值, 取值为 0 或 1。 $f(x_{ij})$ 表示样本 i 被预测为 j 类别的概率。

(3) 网络参数调节。

将求解好的隶属度矩阵输入到神经网络中,利用反向传播算法对网络的参数进行不断的迭代修正。算法描述如下:

输入:样本训练集 $\Omega = \{(X^{(n)}, y^{(n)})\}_{n=1}^N$, 样本验证集 V , 网络学习率 α , 正则化项系数 λ , 网络结构层数 L , 网络节点数量 $m^{(l)}, l \in [1, L]$ 。

步骤:

- ①对网络的权值参数 W 和 b 随机初始化。
- ②对样本训练集 Ω 中的样本进行随机抽取排序。
- ③从样本训练集 Ω 中依次选取训练样本 $(X^{(n)}, y^{(n)})$ 。
- ④通过前向传播计算每一层的输入值与激活函数的值。

⑤通过反向传播算法调整每一层的误差值,更新网络的权值与偏置值。

⑥重复步骤 ② ~ ⑤ 直到网络模型在样本验证集上的错误趋于稳定。

输出:权值矩阵 W 与偏置 b 。

(4) 预测集样本的预测。

利用训练好的人工神经网络对样本预测集数据进行预测,对网络性能做出评估。

3 实验

3.1 相关数据集说明

选取鸢尾花数据集作为算法网络的实验数据集,鸢尾花数据集作为统计学习与样本分类领域的经典数据集,常被用于实验测试。数据集中记录了 150 条数据,分为 3 类鸢尾花(Iris-Setosa, Iris-Versicolour, Iris-Virginica),每类鸢尾花包含了 50 条数据。每条数据记录了鸢尾花的四个特征,即花瓣的长度和宽度,花萼的长度和宽度,通过鸢尾花的特征预测鸢尾花的品种。

3.2 数据可视化分析

(1) 读取鸢尾花数据集的 csv 文件,鸢尾花数据集的部分数据如表 1 所示。每条数据包含了 Sepal. Length, Sepal. Width, Petal. Length, Petal. Width 四个属性以及品种类别 Species。数据集各类属性的均值、方差、中位数、最大值和最小值等如表 2 所示。

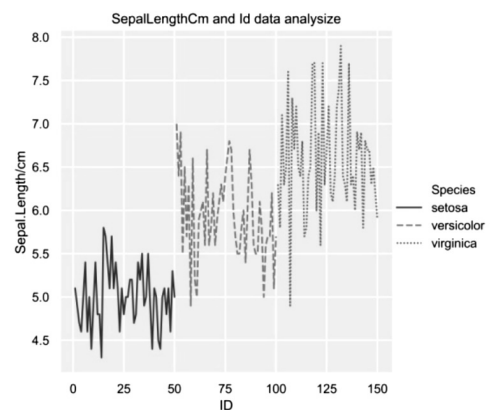
表 1 数据集部分数据

ID	Sepal. Length	Sepal. Width	Petal. Length	Petal. Width	Species
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	4.7	3.2	1.3	0.2	Setosa
4	4.6	3.1	1.5	0.2	Setosa
5	5.0	3.6	1.4	0.2	Setosa

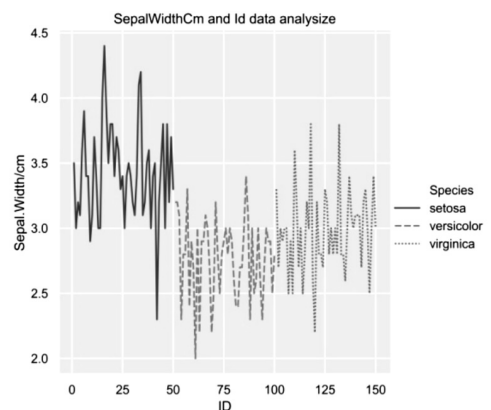
表 2 数据集属性值数据

	ID	Sepal. Length	Sepal. Width	Petal. Length	Petal. Width
Count	150	150	150	150	150
Mean	75.5	5.84	3.06	3.76	1.20
Std	43.4	0.83	0.44	1.77	0.76
Min	1	4.3	2.0	1.0	0.1
Max	150	7.9	4.4	6.9	2.5

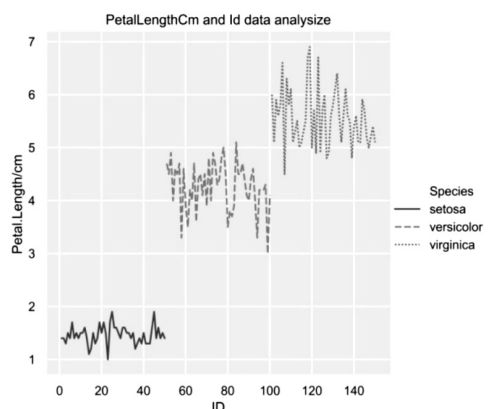
(2) 绘制 3 类鸢尾花数据各属性值与 ID 之间的关系图,如图 2 所示。



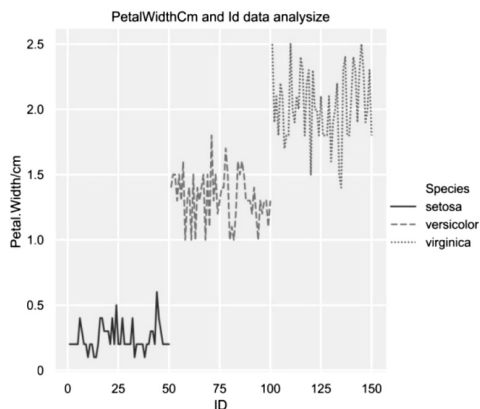
(a) 花萼长与 ID 之间关系



(b) 花萼宽与 ID 之间关系



(c) 花瓣长与 ID 之间关系



(d) 花瓣宽与 ID 之间关系

图2 各属性值与 ID 之间关系

ID 为 1 ~ 50 的为 Setosa 鸢尾花, 51 ~ 100 的为 Versicolour 鸢尾花, 剩余的为 Virginica 鸢尾花。通过关系图可以直观地分析各类鸢尾花的属性值分布范围。

3.3 训练与预测

3.3.1 数据处理

在进行相关参数设定之前, 首先要对鸢尾花数据集进行一系列的数据操作, 对数据集的操作可从以下两个方面考虑:

(1) 基于 3.2 节的数据分析, 分析鸢尾花数据集中是否存在异常点, 如果存在, 将异常点删除, 并处理某些样本属性的丢失值。丢失的样本属性值通常由该属性基于统计信息的平均值替换。

(2) 对数据集进行相关性分析, 分析样本属性之间是否存在相关性, 如果存在, 去除相关的冗余属性。

对处理过的数据进行划分, 将原始数据集划分为训练数据集、测试数据集以及验证数据集三个部分。其中, 训练数据集包含 80 组鸢尾花数据, 验证数据集包含 40 组鸢尾花数据, 验证集用于监测过拟合现象以及超参数的调整, 测试数据集包含 30 组鸢尾花数据, 用于验证模型的准确性。

3.3.2 参数设定

根据鸢尾花种类, 设定 3 种不同参数的隶属度函数, 求解样本基于不同鸢尾花种类的隶属值。基于鸢尾花数据集的可视化分析, 隶属度函数中心位置参数 Z 分别设置为: $Z_1 = 5$, $Z_2 = 6$, $Z_3 = 7$, 其中 Z_1 表示 Setosa 隶属度函数的中心, Z_2 表示 Versicolor 隶属度函数的中心, Z_3 表示 Virginica 隶属度函数的中心。 $e_1 = 1$, $e_2 = e_3 = 2$, $l_1 = l_2 = l_3 = 2$ 。

3.3.3 神经网络设计

基于对鸢尾花数据集的分析, 设计具有单隐含层的神经网络模型, 采用随机梯度下降法对网络模型的参数进行学习。神经网络输入层中的节点个数等同于隶属度函数的种类个数。输出层节点的数量与数据集中存在的类的数量相同。神经网络的输入层为数据样本对各类鸢尾花的隶属度值, 共有 3 个神经元, 输出层为类别种类, 共有 3 个输出神经元。对隐含层设置不同个数的神经元, 通过实验选择最优隐含层神经元个数, 以样本分类准确率为评价指标, 实验结果如表 3 所示。

表3 神经元数目准确率测试结果

隐藏层神经元的个数	准确率/%
8	0.93
7	0.95
6	0.96
5	0.96
4	0.98
3	0.94

根据实验结果, 最终隐藏层神经元数目设定为 4, 选取 ReLU 作为隐含层和输出层的激活函数, Swish-B 作为输入层和隐含层的激活函数, 设置网络误差为 0.000 1, 学习率为 0.01, 最大迭代次数为 100。模型训练过程的性能变化如图 3 所示。

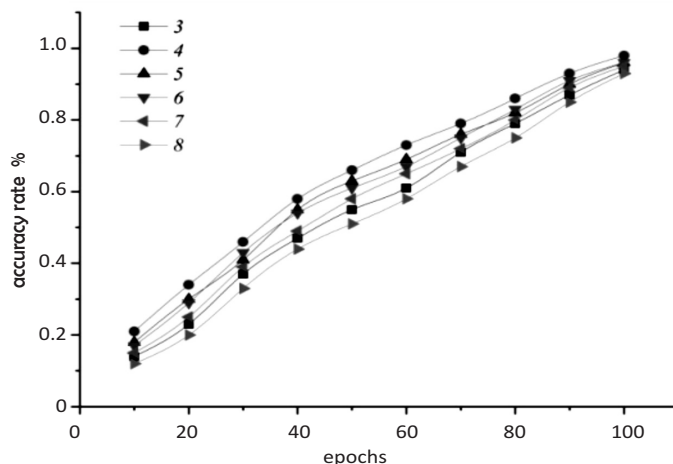


图3 不同隐层神经元数目下的训练性能

3.3.4 数据集训练和预测

在神经网络的训练阶段,首先将鸢尾花数据集分为训练集、验证集、预测集三部分,对其进行模糊化处理。把经模糊化处理后的训练集应用于提出的模糊神经网络,训练网络以建立模糊神经网络分类模型。在测试阶段,通过模糊化处理后的测试集对模型性能进行评估,根据测试集样本预测的准确性对模型进行评估。其详细步骤如下:

Step1:随机抽取鸢尾花数据集中的 80 条样本数据作为训练集,随机抽取 40 条样本数据作为验证集,剩余部分数据作预测集。

Step2:利用式(8)~式(10)求解出预测集样本和验证集样本的隶属度矩阵 D_1 、 D_2 。

$$D_1 = \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.6 \\ 0.1 & 0.5 & 0.4 \\ \dots & \dots & \dots \\ 0.7 & 0.3 & 0.0 \\ 0.2 & 0.6 & 0.2 \\ 0.8 & 0.1 & 0.1 \\ 0.6 & 0.3 & 0.1 \end{bmatrix}$$

$$D_2 = \begin{bmatrix} 0.4 & 0.2 & 0.4 \\ 0.5 & 0.3 & 0.2 \\ 0.1 & 0.3 & 0.6 \\ \dots & \dots & \dots \\ 0.4 & 0.3 & 0.3 \\ 0.7 & 0.1 & 0.2 \\ 0.2 & 0.2 & 0.6 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

Step3:设置网络误差、学习率等相关参数,利用训

练集的隶属度矩阵 D_1 对网络进行 100 次的迭代训练,利用验证集的隶属度矩阵 D_2 对神经网络的超参数进行调整。模型迭代过程中的损失变化如图 4 所示。

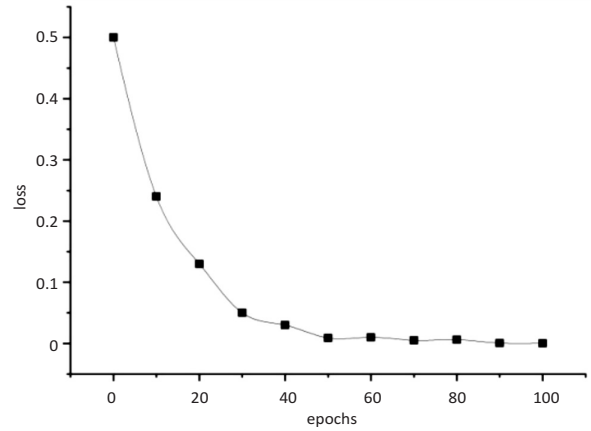


图 4 模型训练过程中的迭代损失变化

Step4:利用测试集样本预测的准确率对模型的性能进行评估。

3.4 实验结果对比

选取引言部分四种经典的基于证据理论的样本分类算法,计算传统算法基于鸢尾花数据集的分类准确率,与文中提出的样本分类方法进行比较,结果如图 5、表 4 所示。

表 4 准确率比较

Methods	20	40	60	80	100
Dempster's	0.85	0.85	0.85	0.85	0.85
Murphy's	0.88	0.88	0.88	0.88	0.88
Deng et al. 's	0.92	0.92	0.92	0.92	0.92
Han et al. 's	0.94	0.94	0.94	0.94	0.94
Ours	0.34	0.58	0.73	0.86	0.98

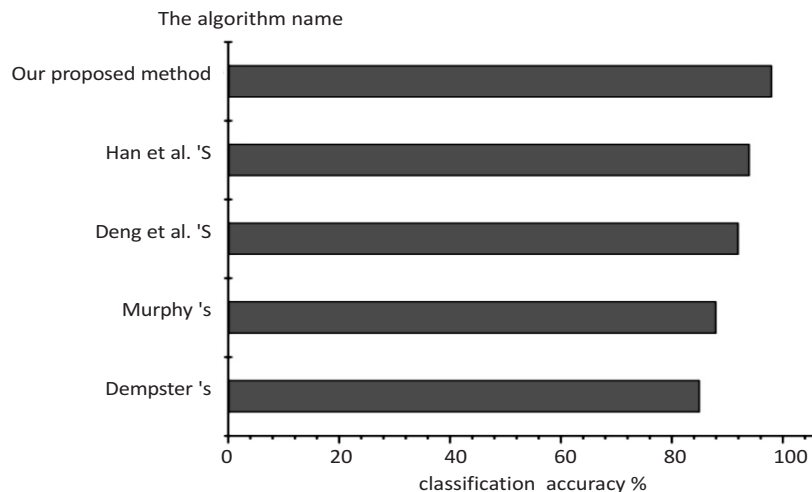


图 5 相关算法精确度比较

随着模糊神经网络训练迭代次数的增加,样本分类的准确性最终稳定到 0.98,与基于证据理论的分类算法相比,样本分类的准确率明显提升。

4 结束语

由于基于证据理论的分类算法在进行样本分类时

过于复杂,样本的某些属性间存在相关性。另外,在复杂数据集上,基于证据理论的样本分类算法往往不够理想,随着样本属性的增加,算法过程变得越来越复杂,数据集中样本的冲突变得明显。该文提出了一种基于模糊神经网络的样本分类算法,并采用鸢尾花数据集对算法进行检验。与传统的样本分类算法相比,一方面,基于模糊神经网络的样本分类算法不需要过多关注求解过程中的计算以及样本某些属性间是否存在相关性。另一方面,基于证据理论的样本分类算法中存在如何将权重合理分配给相关特征等问题,基于模糊神经网络的样本分类算法引入随机梯度下降的方法,通过多次迭代训练结合反向传播算法自动求解出合理的权重值。由实验结果可知,基于模糊神经网络的样本分类算法对分类问题的平均准确性为98%,样本分类的准确性与传统数据算法相比显著提高。

参考文献:

- [1] LIN H Y, ZENG D, LIANG Y K, et al. GATA3 and TRPS1 are distinct biomarkers and prognostic factors in breast cancer; database mining for GATA family members in malignancies[J]. *Oncotarget*, 2017, 8(21):34750–34761.
- [2] OLTEANU D. Technical perspective: juggling functions inside a database[J]. *ACM SIGMOD Record*, 2017, 46(1):5–15.
- [3] DEMPSTER A P. Upper and lower probabilities induced by a multivalued mapping[J]. *Annals of Mathematical Statistics*, 1967, 38(2):325–339.
- [4] ZADEH L A. A simple view of the Dempster–Shafer theory of evidence and its implication for the rule of combination[J]. *AI Magazine*, 1986, 7(2):85–91.
- [5] DENG Y. Generalized evidence theory[J]. *Applied Intelligence*, 2015, 43(3):530–543.
- [6] JIANG W, WEI B, QIN X, et al. Sensor data fusion based on a new conflict measure[J]. *Mathematical Problems in Engineering*, 2016, 2016:5769061. 1–5769061. 11.
- [7] MURPHY C K. Combining belief functions when evidence conflicts[J]. *Decision Support Systems*, 2000, 29(1):1–9.
- [8] HAN D, DENG Y, LIU Q. Combining belief functions based on distance of evidence[J]. *Decision Support Systems*, 2005, 38(3):489–493.
- [9] HAN Deqiang, DENG Yong, HAN Chongzhao, et al. Weighted evidence combination based on distance of evidence and uncertainty measure[J]. *Journal of Infrared and Millimeter Waves*, 2011, 30(5):396–400.
- [10] HAN Jiawei, MICHELINE K. Data mining: concepts and techniques[M]//Data mining concepts models methods & algorithms. San Francisco, CA, United States: Morgan Kaufmann Publishers Inc., 2006:1–18.
- [11] SARVESTANI A S, SAFAVI A A, PARANDEH N M, et al. Predicting breast cancer survivability using data mining techniques[C]//Software technology and engineering (IC-STE). San Juan, PR, USA: IEEE, 2010.
- [12] MEI S, MONTANARI A, NGUYEN P M. A mean field view of the landscape of two-layers neural networks[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, 115(33):E7665–E7671.
- [13] EHLERS R. Formal verification of piece-wise linear feed-forward neural networks[C]//International symposium on automated technology for verification and analysis. Cham: Springer, 2017.
- [14] MARCJASZ G, UNIEJEWSKI B, WERON R. On the importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks[J]. *International Journal of Forecasting*, 2018, 35(4):1520–1532.
- [15] MA L, HU S, QIU M, et al. Energy consumption optimization of high sulfur natural gas purification plant based on back propagation neural network and genetic algorithms[J]. *Energy Procedia*, 2017, 105:5166–5171.
- [16] GUAN H, DAI Z, ZHAO A, et al. A novel stock forecasting model based on high-order-fuzzy-fluctuation trends and back propagation neural network[J]. *PLoS One*, 2018, 13(2):e0192366.
- [17] LIU H, LI X, WANG W, et al. Markov velocity predictor and radial basis function neural network-based real-time energy management strategy for plug-in hybrid electric vehicles[J]. *Energy*, 2018, 152:427–444.
- [18] ZARBAKSH P, ADDEH A. Breast cancer tumor type recognition using graph feature selection technique and radial basis function neural network with optimal structure[J]. *Journal of Cancer Research & Therapeutics*, 2018, 14(3):625–633.
- [19] LI W, LAI H, ZHANG D. Yoneda completeness and flat completeness of ordered fuzzy sets[J]. *Fuzzy Sets & Systems*, 2017, 313:1–24.
- [20] LIU B. Uncertainty theory: an introduction to its axiomatic foundations[M]. Berlin: Springer-Verlag, 2004.
- [21] DUBOIS D. Fuzzy sets and systems: theory and applications[M]. [s.l.]: Academic Press, 1980.