



1. 引言

1.1 分类

- 机器学习中**最古老、最重要、研究最多**的问题之一
- 目前已有数十种分类算法
- 大部分现有学习算法只能在**特定的任务、单一的评价维度**上取得好的性能
- 提升学习算法的**综合性能**是机器学习领域不懈追求的目标

1.2 概念

- 由**外延**和**内涵**组成
- 外延和内涵可以**相互诱导**
- 是组成知识的**基本单元**

1.3 分类本质上是概念认知过程

论证：

- 概念认知**：对概念形成认知，即得到概念的外延或内涵
- 类别空间中的每个元素对应一个概念**
- 分类**：建立从输入空间到类别空间的映射，即得到类别空间中每个概念的外延
- 综合a—c，**得证**

1.4 现状分析

- 现有学习算法**在设计时很少完全地考虑概念认知
- 现有学习算法**很难适应用不同任务、兼顾多个评价维度
- 人类**的分类过程是概念认知的过程
- 人类**在不同任务、多个评价维度上都有出色的表现

启发

将概念认知原理引入到学习算法的设计中，有望获得像人类一样可以适应不同任务、兼顾多个评价维度的学习算法

2. 模糊学习机的一般形式

2.1 分类问题可基于相似性求解

- 认知科学依据：相似性是概念表示的基础**
- 数学工具：集合论、二元关系**
- 论证：**
 - 任意分类问题（**定义1**）都可以转化为一个与之等价的等价关系问题（**命题1**）
 - 等价关系是一种特殊的相似性综合①—②，**得证**
- 分类问题被建模为**等价关系（ER）问题（定义2）**

2.2 捕获概念的模糊性

- 认知科学依据：概念具有模糊性**



- ‘猫’和‘狗’两个视觉概念具有**模糊性**
- 模糊视角**：第3幅图片属于猫的程度是0.6，属于狗的程度0.5
- 随机视角**：第3幅图片是猫概率是0.6，是狗的概率0.4
- 随机视角是不合理的，因为第3幅图片不可能有时是猫，有时是狗

- 数学工具：模糊集、模糊二元关系**

- 由Lotfi A. Zadeh于1956年提出
- 是处理模糊性的有力工具

- 分类问题被进一步建模为**模糊等价关系（FER）问题（定义3）**

2.3 基于样例理论的概念表示

- 认知科学依据：样例理论**

核心思想：人类通过记住一些狗的对象来表示概念‘狗’

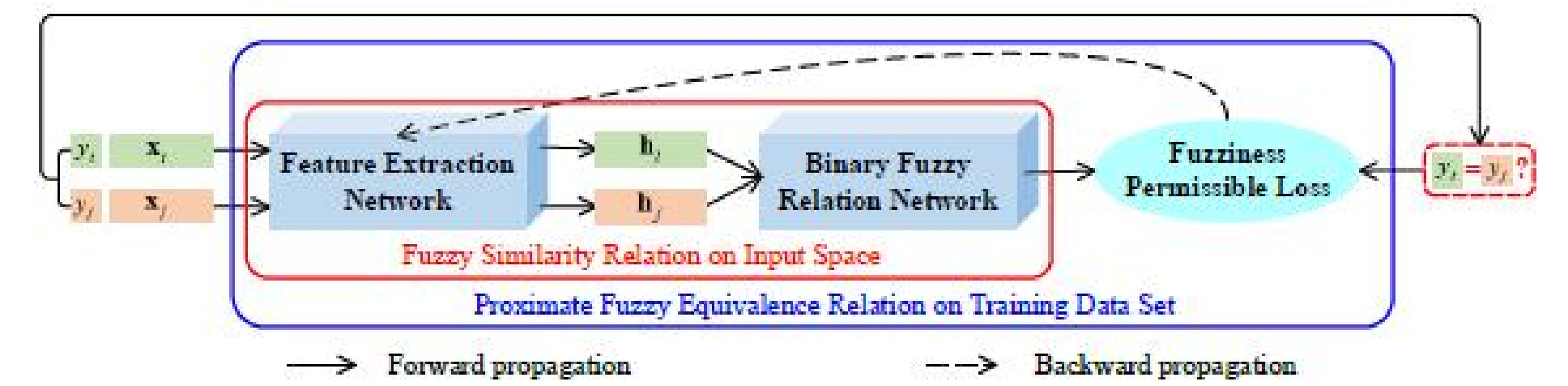
- 样例选择方法（**定义4**）

原理：一个狗的样本与其他狗的样本越相似，该样本越能代表概念‘狗’

2.4 基于概念表示分类

- 原理**：一个样本应该属于与它最相似的概念

3. 模糊学习机的神经网络实现



- 数学工具：传递闭包、图论**

- 特征提取网络：捕获问题的非线性因素（**定理1**）
- 模糊二元关系网络：保证FER的规范性、自反性和对称性
- 模糊容忍损失：保留概念的模糊性，间接实现FER的传递性
- 随机梯度下降优化器：保证学习过程高效进行（理论保证：在模糊容忍损失意义下模糊相似关系可以有效逼近FER（**定理2**））

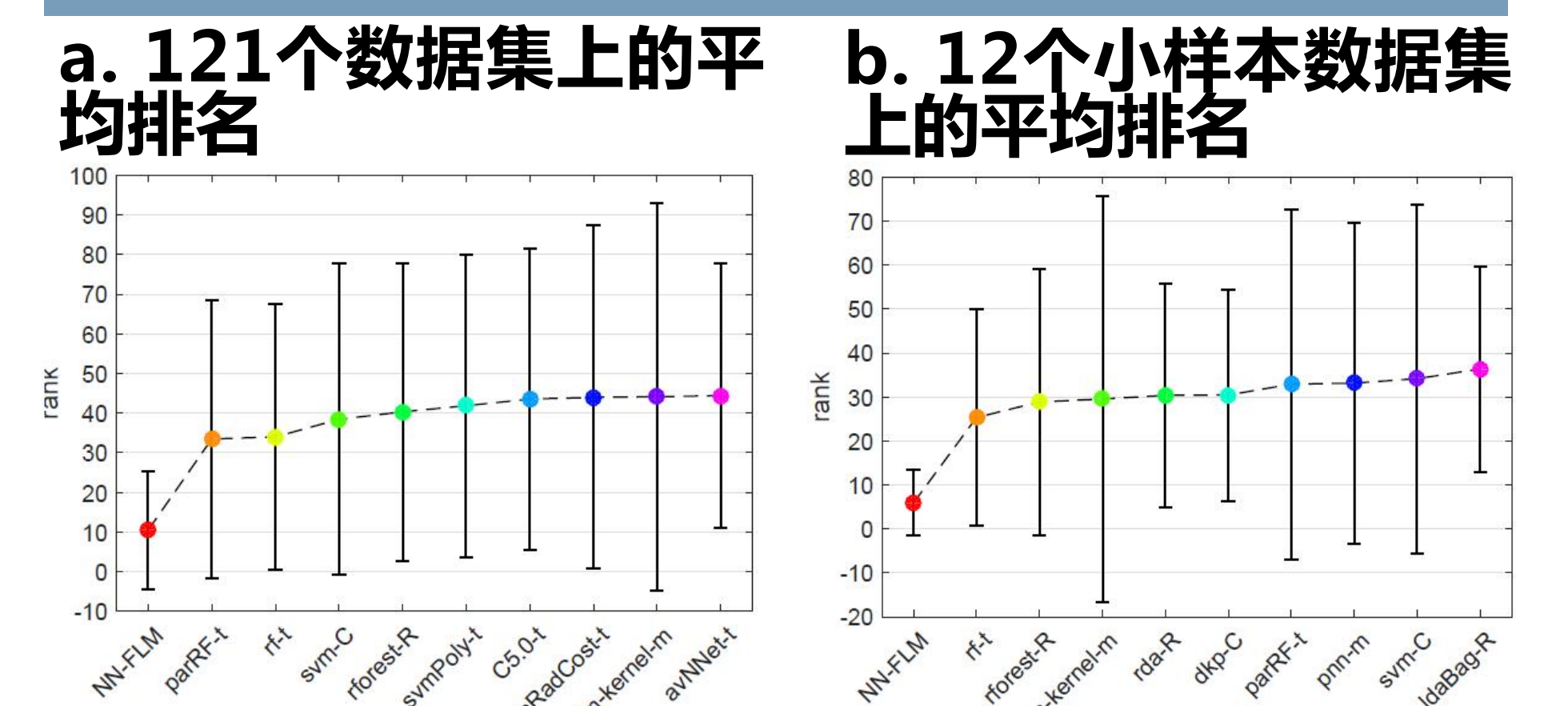
4. 实验

4.1 MNIST数据集实验



- 模糊学习机学到的相似性、选出的样例，预测的候选类标签都比较符合人类的认知，具有较强的**可解释性**
- 模糊学习机对样本的扰动具有一定的**鲁棒性**

4.2 与来自17个家族的179种分类器对比



- 模糊学习机平均排名**第1**，排名的**标准差最小**，**显著优于第2的名方法**，训练样本较少时优势更显著
- 模糊学习机具有较强的**泛化性能**

5. 结论

模糊学习机

认知科学依据

数学基础

泛化性能

可解释性

鲁棒性

高效性

...

致谢

本工作受到国家重点研发计划项目（2020AAA0106100）的支持。

