

文章编号: 1000-5641(2019)05-0066-08

## 基于模糊聚类和支持向量回归的成绩预测

申航杰, 琚生根, 孙界平

(四川大学 计算机学院, 成都 610065)

**摘要:** 现有的成绩预测模型往往过度使用不同类型的属性, 导致过于复杂的分数预测方法, 或是需要人工参与. 为提高学生成绩预测的准确率和可解释性, 提出了一种融合模糊聚类和支持向量回归的成绩预测方法. 首先引入模糊逻辑来计算隶属度矩阵, 根据学生的历史成绩进行聚类, 随后对每个聚类簇利用支持向量回归理论对成绩轨迹进行拟合建模. 此外, 结合学生学习行为等相关属性, 对最终的预测结果做调整. 在多个基准数据集上进行了实验测试, 验证了该方法的有效性.

**关键词:** 模糊聚类; 支持向量回归; 预测; 教育数据挖掘

**中图分类号:** TP311      **文献标志码:** A      **DOI:** 10.3969/j.issn.1000-5641.2019.05.005

## Performance prediction based on fuzzy clustering and support vector regression

SHEN Hang-jie, JU Sheng-gen, SUN Jie-ping

(College of Computer Science, Sichuan University, Chengdu 610065, China)

**Abstract:** Existing performance prediction models tend to overuse different types of attributes, leading to either overly complex prediction methods or models that require manual participation. To improve the accuracy and interpretation of student performance prediction, a method based on fuzzy clustering and support vector regression is proposed. Firstly, fuzzy logic is introduced to calculate the membership matrix, and students are clustered according to their past performance. Then, we use Support Vector Regression (SVR) theory to fit and model performance trajectory for each cluster. Lastly, the final prediction results are adjusted in combination with the students' learning behavior and other related attributes. Experimental results on several baseline datasets demonstrate the validity of the proposed approach.

**Keywords:** fuzzy clustering; support vector regression; prediction; educational data mining

收稿日期: 2019-07-28

基金项目: 四川省科技厅重点研发项目(2018GZ0182); 四川大学未来教育研究专项(SCUFEB2019004)

第一作者: 申航杰, 女, 硕士研究生, 研究方向为数据科学. E-mail: 13693499219@163.com.

通信作者: 孙界平, 男, 副教授, 研究方向为数据科学. E-mail: sunjieping@scu.edu.cn.

## 0 引言

近年来,计算机技术在教育领域已经被广泛应用.学生学业成绩预测一直是教育和学习中的重要环节.例如,当前挂科现象在学生中非常普遍,然而挂科可能会造成学生无法按时毕业或者无法找到心仪工作的后果<sup>[1]</sup>.如果能提前发现学生的学习异常,通过引导和干预就有可能阻止这些事情的发生<sup>[1]</sup>.再者教育机构也可以通过预测学生表现,为学生定制合适的辅导手段,以适应他们的需求.因此,准确地预测学生的成绩在教育管理中就显得十分重要.

目前许多数据挖掘和机器学习的方法已经应用于教育领域.现有的成绩预测方法大致可以分为两类,一是使用人工神经网络,二是建立概率统计模型. Borkar 等介绍了一种用神经网络做成绩预测并对学生进行分类的方法<sup>[2]</sup>.一些概率模型,例如多变量线性回归模型、稀疏因子分析模型<sup>[3]</sup>等都被用于预测学生成绩.张嘉等提出了基于规则的概率软逻辑模型,通过半自动的构建规则辅以决策树算法进行建模<sup>[4]</sup>.这些工作都大量使用了两种属性,一是学生以前的课程成绩,用历史成绩预测他们在后续课程中的表现.除历史成绩外,第二种属性是学生的基础行为属性,包括学生的年龄、性别、课堂出勤率、自习频率、图书馆门禁记录、座位选择情况和社会关系等等.但是过度使用不同类型的属性可能导致过于复杂的分数预测方法,其结果往往会难以解释.

本文引入模糊逻辑来计算隶属度矩阵,提出一种融合模糊聚类和支持向量回归的成绩预测方法.首先利用模糊 C 均值算法(Fuzzy C-means Clustering, FCM)对学生历史成绩属性做聚类,接着对每个聚类簇单独进行支持向量回归(Support Vector Regression, SVR),并根据学生行为对最终预测结果做出调整.本方法结合了学生历史成绩和行为习惯等属性进行综合预测,在3个数据集上进行实验,结果表明,本方法在成绩预测精度方面有明显性能提升.

## 1 相关工作

### 1.1 成绩预测相关研究

学生成绩预测是心理学、教育学和计算机科学的交叉领域.目前的成绩预测方法可分为基于神经网络训练和概率统计模型两类. Borkar等提出了一种用神经网络做成绩预测并对学生进行分类的方法<sup>[2]</sup>.张嘉等提出了基于规则的概率软逻辑模型,通过半自动的构建规则辅以决策树算法进行建模<sup>[4]</sup>.张麒增等使用学生缺课次数、离讲台距离、上课迁移座位次数、历史成绩来对学生成绩进行预测,采用过采样技术配合神经网络提升模型预测能力<sup>[8]</sup>.孙毅等介绍了基于多元线性回归的预测模型<sup>[9]</sup>.陈岷等提出了融合因子分析和神经网络的体育成绩预测模型<sup>[10]</sup>.Bunkar等使用学生的班级考试成绩、研讨班成绩、作业成绩、课堂出勤率和实验室工作情况等,来预测学生学期结束时的分数<sup>[12]</sup>.利用神经网络的黑箱来预测学生成绩,模型会缺乏可解释性,不够直观且不利于理解.利用回归模型直接进行预测往往需要人为地进行特征选择.通过构建规则进行成绩预测的方法也需要人工参与,人为地给出常识知识作为规则确立的先决条件,其中包含的不正确信息可能会影响预测结果.

### 1.2 模糊 C 均值聚类

模糊 C 均值算法是模糊聚类算法中最重要的聚类算法,在模式分类、机器学习和数据挖掘等领域中有着广泛应用<sup>[5]</sup>.模糊 C 均值算法相对于其他聚类算法有着许多的优势,如模型数学表达易于理解符合实际、优化求解方法多样、收敛理论严谨、算法易于借助计算机编程实现、模糊聚类效果表现优良等<sup>[6]</sup>.

模糊 C 均值聚类采用模糊划分,为每个样本确定其属于各个组的隶属度,用 0 到 1 之间

的数表示. 该算法的优化准则为:

$$\begin{cases} JFCM(U, P) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - \theta_i\|^2, \\ \sum_{i=1}^c u_{ij} = 1, \quad 0 \leq u_{ij} \leq 1, \quad j = 1, 2, \dots, n. \end{cases} \quad (1)$$

其中  $P = (\theta_i)_{c \times d}$  为聚类中心矩阵,  $U = (u_{ij})_{c \times n}$  表示模糊隶属度矩阵,  $d_{ij} = \|x_j - \theta_i\|$  表示样本与聚类中心的距离,  $m(m > 1)$  表示模糊指标. 本文采用模糊 C 均值聚类算法, 在处理不确定问题方面有着天然优势.

### 1.3 支持向量回归

考虑样本集  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , 引入敏感度  $\varepsilon$  的损失函数为:

$$L\varepsilon(f(x_i) - y_i) = \begin{cases} 0, & |f(x_i) - y_i| < \varepsilon \\ |f(x_i) - y_i| - \varepsilon, & |f(x_i) - y_i| \geq \varepsilon \end{cases} \quad (\varepsilon \text{ 为正数}). \quad (2)$$

本文使用支持向量回归, 将其用于成绩预测中, 希望能找到参数使得函数  $f(x_i)$  和学生真实成绩之间有尽量小的偏差.

支持向量机应用于回归预测等问题则称为支持向量回归, 与传统机器学习方法相比有较好的学习性能, 克服了传统方法的大样本要求, 能很好地解决小样本、非线性、维数灾难及局部极小等实际问题<sup>[7]</sup>.

## 2 基于模糊聚类和支持向量回归的成绩预测

### 2.1 FCSVR 模型框架

本文建立基于模糊聚类和支持向量回归的成绩预测模型(performance prediction based on fuzzy clustering and support vector regression, FCSVR), 主要分为 3 步, 首先对样本进行聚类、随后用回归算法拟合成绩、最后对成绩做出些许调整, 形成学生最终成绩预测机制. 其实现流程如下(见图 1).

- (1) 预处理原始数据并划分训练集和测试集;
- (2) 对训练集和测试集的所有学生历史成绩(不包括其最终成绩)进行模糊聚类, 通过相似性将他们分为不同类别, 并计算每个样本关于不同类别的模糊隶属度值;
- (3) 基于隶属度最大的原则进行聚类, 对每个聚类簇使用支持向量回归算法建模, 确定最终成绩与历史成绩记录之间的关系;
- (4) 根据测试集学生自身的学习行为属性, 找到具有相同或者相似行为的学生, 对目标学生最终成绩做出调整;
- (5) 通过在步骤(2)中获得的模糊隶属度矩阵, 在步骤(3)中确定的关系以及步骤(4)中对成绩的调整, 来获得目标学生的最终成绩.

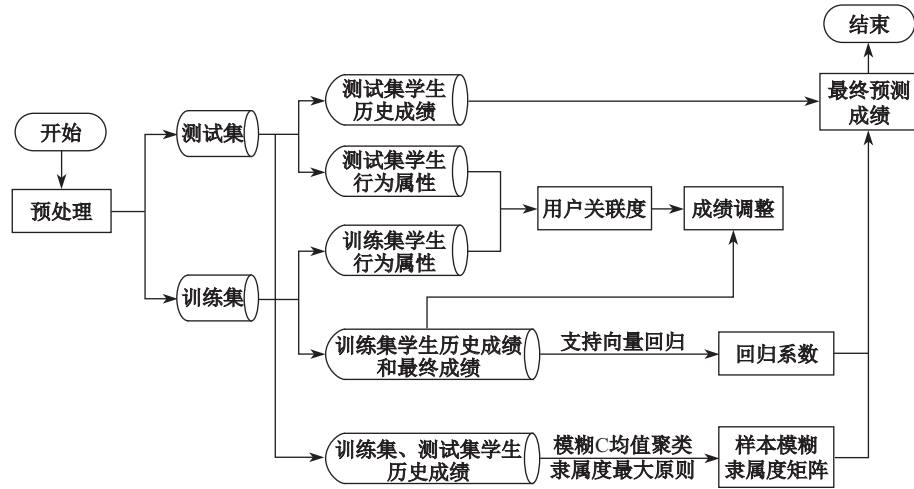


图 1 FCSVR 模型流程图

Fig. 1 Flowchart of the FCSVR Model

本文提出的基于模糊聚类和支持向量回归的成绩预测模型, 首先对所有样本进行模糊聚类, 将样本空间划分为若干个簇(子空间), 使同一个簇中的对象有很大相似性, 不同簇间的对象有很大差异性. 传统的聚类分析是一种硬划分, 将样本严格划分到某个类中, 类之间的界限是明确的. 在成绩预测问题中, 样本没有严格的类属划分, 某些样本可能具有“亦此亦彼”的性质, 因此本文采用模糊 C 均值聚类, 使系统有更好的灵活性. 然后, 针对每一个簇构造 SVR 模型来拟合学生成绩, 经过划分后的样本空间数据量减少, SVR 能很好地适用于小样本问题. 同时, 聚类之后使用 SVR 可以改善 SVR 训练时间过长的不足, 而且能提高模型预测精度.

## 2.2 模糊聚类

对训练集和测试集的所有学生进行模糊聚类, 其中的每个记录包括学生的历史成绩和最终成绩. 本文使用模糊 C 均值聚类算法, 根据学生的历史成绩, 将他们划分为不同的类别. 具有相似历史成绩的学生将被划分到同一组. 然而不同类别之间具有相关性, 某个实例难以被确切地分为某个类. 因此模糊聚类对处理这种不确定性有天然的优势. 得到的模糊隶属度矩阵将被用于最后的预测. 该算法伪代码见算法 1.

如果初始聚类中心随机选取, 容易选到噪声数据, 使算法性能变差. 所以初始聚类中心通过以下步骤生成: 首先计算每个学生历史成绩的算数平均值, 根据其算数平均值对记录排序, 使排序好的所有记录平均地分为  $K$  个簇, 其中心表示为  $C_j$  ( $1 \leq j \leq K$ ).

---

### 算法 1 历史成绩模糊聚类

---

输入: 所有样本记录  $R = \{r_1, r_2, \dots, r_n\}$ ; 初始聚类中心  $C = \{c_1, c_2, \dots, c_k\}$ ; 聚类个数  $K$ ; 模糊指标  $m$  ( $m > 1$ ); 控制迭代地最小阈值  $\varepsilon$

输出:  $K$  个聚类簇以及样本模糊隶属度矩阵

---

1: 计算初始模糊隶属度值  $u_{ij} = \frac{1}{\sum_{l=1}^K \left( \frac{\|s_i - c_j\|}{\|s_i - c_l\|} \right)^{\frac{2}{m-1}}}$  (3) ( $\|*\|$  表示欧氏距离)

2: loop

$$3: \text{ 计算聚类中心 } c_j = \frac{\sum_{i=1}^n u_{ij}^m r_i}{\sum_{i=1}^n u_{ij}^m} \quad (4)$$

4: 更新模糊隶属度矩阵

$$5: \text{ 计算目标函数 } J^i = \sum_{i=1}^n \sum_{j=1}^K u_{ij}^m \|r_i - c_j\|^2 \quad (5)$$

6: **until**  $|J^i - J^{i-1}| < \varepsilon$

### 2.3 支持向量回归

回归分析是在预测问题中常用的一种统计方法,用于估计变量之间的关系.支持向量回归的损失函数自带 L2 正则项,对数据有普适性.给出数据集  $\{y_i, x_{i1}, x_{i2}, \dots, x_{im}\} (i = 1, \dots, n)$ , 希望找到变量  $y_i$  和变量  $x_{ij} (1 \leq j \leq m)$  之间的关系.支持向量回归的基本思想就是找到一个函数  $f(x_i) = \mathbf{w}^T \mathbf{x}_i + b$ , 使其与  $y$  有最大偏差  $\varepsilon$ . 该问题可写成一个凸优化问题.

$$\begin{aligned} \min_{(\mathbf{w}, b, \xi_i)} & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l \xi_i \\ \text{满足} & \begin{cases} |\mathbf{w}^T \mathbf{x}_i + b - y_i| - \varepsilon \leq \xi_i \\ \xi_i \geq 0 \end{cases}, \end{aligned} \quad (6)$$

对于每一个聚类簇,分别计算相应的回归系数,记为  $w_i (1 \leq i \leq K)$ ,  $b_i (1 \leq i \leq K)$ .

### 2.4 最终成绩预测

一方面,学生历史成绩可以作为成绩预测的重要参考,另一方面,学生的行为等属性也会影响学生最终成绩.学生在前几个阶段有相同或相似的成绩,但在最后阶段的成绩可能也有差异.如今人们普遍认为学习行为会对学生的学业成绩产生重要影响<sup>[11]</sup>.因此,本文除历史成绩外还考虑了学生的行为习惯,用来进一步调整最终成绩,以产生更好的预测结果.

#### 算法 2 最终成绩预测

输入: 训练集  $U = \{r_1, r_2, \dots, r_n\}$ ; 测试集  $T = \{r_1, r_2, \dots, r_d\}$ ; 超参数  $N, \alpha$

输出: 最终成绩的预测值

1: 根据 2.2 小节得到的模糊隶属度矩阵,针对测试样本在每一聚类簇中计算 SVR 回归的结果,表示为  $\text{temp\_score}_i = \sum_{i=1}^K u_i \cdot f(x_i)$  (7), 其中  $K$  是聚类簇的个数;

2: 计算测试样本与每个归属同一个类的训练样本之间的欧氏距离  $d_i (1 \leq i \leq |c_t|)$ ,  $|c_t|$  表示该测试样本所属类里实例数;

3: 将步骤 2 中计算的欧氏距离从小到大排序,选择前  $N$  个训练样本,记为  $S_n$ ;

4: 对于  $S_n$  中的每一个实例,使用学生的学习行为等相关属性,计算其与测试样本的余弦相似度  $\text{Sim}_t = \frac{\sum_{i=1}^m A_i^S \times A_i^T}{\sqrt{\sum_{i=1}^m (A_i^S)^2} \times \sqrt{\sum_{i=1}^m (A_i^T)^2}}$  (8), 其中  $A_i (1 \leq i \leq m)$  表示学生行为属性,  $A_i^S$  表示测试样本属性,  $A_i^T$  表示  $V_n$  中的第  $T$  个实例.

5: 引入超参数  $\alpha$  来灵活地控制学生行为属性对最终预测成绩的影响,计算对成绩的修正  $b_i = \alpha \cdot \sum_{n=1}^N (\text{Sim}_n \cdot (s_n - s_i))$  (9), 其中  $s_n$  表示第  $n$  个训练样本最终成绩,  $s_i$  表示测试集中第  $i$  个样本利用 SVR 和隶属度值计算得到的预测成绩;

6: 得到最终成绩的预测值  $\text{score}_i = \text{temp\_score}_i + b_i$  (10),  $\text{temp\_score}_i$  表示步骤 1 中支持向量回归的结果.

### 2.5 时间复杂度分析

首先使用 FCM 进行聚类, 其时间复杂度为  $O(nktd)$ , 其中  $n$  表示样本数目,  $k$  表示聚类簇数目,  $t$  表示算法迭代次数,  $d$  表示样本维度. 随后使用 SVR 对学生成绩进行拟合, 支持向量回归是一个凸二次规划问题, 其训练结果只与支持向量有关. 算法的复杂度也只与训练样本的数目有关, 与样本维度无关, 具有  $O(n^3)$  以上的时间复杂度,  $n$  表示样本数目. 最后预测最终成绩只需要线性复杂度的计算  $O(nk)$ ,  $k$  表示聚类簇数目,  $n$  表示样本数目.

## 3 实 验

### 3.1 数据集与评价指标

本文使用 3 个成绩预测数据集来对模型进行评估.

(1) 葡萄牙中学生数据集, 包含葡萄牙两所中学的 1 064 条学生数据, 属性包括学生成绩、父母受教育程度、家庭情况以及学生行为习惯等相关特征. 数据由 UCI 机器学习库提供(<http://archive.ics.uci.edu/ml/datasets/Student+Performance>), 两个文件分别提供数学(UCI-Math)和葡萄牙语(UCI-Portuguese)成绩.

(2) 中国高校学生日常数据集(Stu-Common), 包含某高校学生的 60 多万条活动记录, 属性包括图书馆进出记录、消费记录、借阅书籍记录以及学生学期排名信息等. 数据来自网站 <http://www.dcjingsai.com/common/cmpt> 提供的学生排名预测竞赛数据集.

实验采用评价回归中常用的均方差(mean squared error, MSE)平均绝对值误差(mean absolute error, MAE)、R2 决定系数(R2 score, R2) 3 个评价指标对本文提出的方法进行考察. MSE、MAE 的值越小说明预测的精度越高, 模型被接受的可能性越大. R2 决定系数越接近于 1, 说明回归方程的拟合程度越好.

$$MSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (12)$$

$$R2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (13)$$

本文提出的方法主要用于学生成绩预测, 该方法也可以应用为分类模型. 将本文的方法用作对目标学生进行二分类, 分为合格和不合格两类. 对于 UCI-Math 和 UCI-Portuguese 两个数据集, 将分数大于 10 分(总分 20 分)的学生记为合格. 对于 Stu-Common 数据集, 将排名在前 345 名(共 538 人)的学生记为合格. 对于此二分类问题, 可以用准确率 (Precision) 来度量.

$$Precision = \frac{TP}{TP + FP}. \quad (14)$$

### 3.2 参数设定

在进行实验之前需要对数据进行预处理, 并设定超参数. 例如学生的学习时间、缺席次数等等属性值都被归一化为 0 到 1 之间的范围. 对于“YES”和“NO”的属性值分别被设置为 1 和 0.

在实验中左右的结果都取 10 次实验结果的平均. 由于学生最后阶段的成绩是一个 0 到 20 的整数, 但是预测的最终成绩是以浮点数的形式得出, 预测数据将以四舍五入的形式转换回整数, 从而计算预测结果的准确率, 也能更好地支持可解释性.

实验过程中将聚类个数  $K$  设为 5, 模糊指标  $m$  设为 2, 最大偏差  $\varepsilon$  设置为  $10^{-6}$ ,  $V_n$  中最大训练样本数  $N$  设置为 5. 在 UCI-Math、UCI-Portuguese 两个数据集中, 系数  $\alpha$  设为 0.2, 在 Stu-Common 数据集, 系数  $\alpha$  设为 0.05. 训练集和测试集的比例为 7:3.

### 3.2 实验结果

为了分析本文提出的方法, 在 3 个数据集上进行了实验. 该工作还与其他 3 种方法进行了比较, 实验配置如下.

- (1) FCSVR: 本文提出的方法.
- (2) MLR: 标准的多变量线性回归方法.
- (3) BR: 贝叶斯岭回归方法.
- (4) EN: 弹性网络回归方法.
- (5) SVR: 支持向量回归方法.

表 1 和表 2 展示了 5 种方法的均方差 MSE 和平均绝对值误差 MAE 的实验对比结果. 整体上看, 本文提出的方法(FCSVR)误差都要低于其他 3 种方法(MLR、BR、EN、SVR). 这 5 种方法在 UCI-Math、UCI-Portuguese 两个数据集上效果比较好, 而 Stu-Common 数据集由于拥有更多的数据量以及较为复杂的属性关系, 总体效果稍差. 其中, 多变量线性回归方法和贝叶斯岭回归方法效果相等, 弹性网络回归方法效果最差.

表 1 5 种方法的均方差对比

Tab. 1 Mean squared error comparison of the five methods studied

数据集	方 法				
	FCSVR	MLR	BR	EN	SVR
UCI-Math	0.21	0.30	0.29	0.62	0.36
UCI-Portuguese	0.17	0.29	0.30	1.38	0.62
Stu-Common	0.26	0.30	0.30	0.73	0.30

表 2 5 种方法的平均绝对值误差对比

Tab. 2 Mean absolute error comparison of the five methods studied

数据集	方 法				
	FCSVR	MLR	BR	EN	SVR
UCI-Math	0.32	0.31	0.30	0.53	0.37
UCI-Portuguese	0.28	0.31	0.31	0.87	0.44
Stu-Common	0.39	0.44	0.44	0.73	0.43

对比方法 1 和方法 5 这组消融实验, 只使用支持向量回归的效果远不及本文提出的 FCSVR 模型, 证明了本文方法的有效性.

图 2 显示了 FCSVR 方法、MLR 方法、BR 方法、SVR 方法和 EN 方法在 UCI-Portuguese、UCI-Math、Stu-Common 在这 3 个数据集上关于  $R^2$  决定系数的对比实验结果. 可以看到, FCSVR 方法在 3 个数据集上的  $R^2$  系数值均要高于其他 4 种算法.  $R^2$  决定系数表示拟合优度, 越接近 1 越好, 这表明了本方法的有效性.

图 3 显示了 FCSVR 方法、MLR 方法、BR 方法、SVR 方法和 EN 方法在 UCI-

Portuguese、UCI-Math、Stu-Common 在这 3 个数据集上关于分类准确率的对比实验结果. 在 5 种方法里, SVR 总体效果最差, 在 Stu-Common 数据集的准确率只有 66.7%, 与其他方法差距明显. 这说明仅仅使用 SVR 用于分类模型是不够的. FCSVR 方法在 3 个数据集上的分类准确率均高于其他 4 种方法, 说明将成绩预测应用于分类问题, FCSVR 方法也能取得很好的效果, 进一步表明了本文方法的优势与潜力.

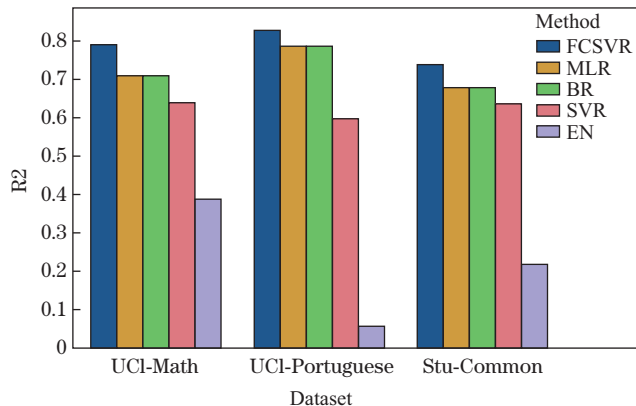


图 2 5 种方法的 R2 系数对比

Fig. 2 R2 comparison of the five methods studied

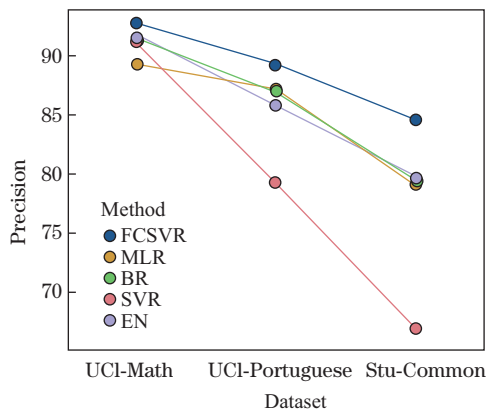


图 3 5 种方法的分类准确率对比

Fig. 3 Precision comparison of the five methods studied

## 4 结 论

本文分析了当前研究的不足, 提出了一种基于模糊聚类和支持向量回归的成绩预测方法. 该方法能有效利用学生历史成绩和行为习惯两种属性, 提高成绩预测的准确率. 在 3 个标准数据集上的对比实验展现了本文提出方法的有效性.

## [参 考 文 献]

- [1] 吕红胤, 连德富, 聂敏, 等. 大数据引领教育未来: 从成绩预测谈起[J]. 大数据, 2015, 1(4): 118-121.



- [6] PENNINGTON J, SOCHER R, MANNING C D. Glove: Global vectors for word representation[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing Association for Computational Linguistics. EMNLP, 2014: 1532-1543.
- [7] KIM Y. Convolutional neural networks for sentence classification[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). ACL, 2014: 1746-1751.
- [8] LIPTON Z C. A critical review of recurrent neural networks for sequence learning[J]. arXiv:1506.00019v1 [cs.LG].
- [9] SAK H, SENIOR A W, BEAUFAYS F. Long short term memory recurrent neural network architectures for large scale acoustic modeling[C]//Proceedings of the 2014 Conference of the International Speech Communication Association. INTERSPEECH, 2014: 338-342.
- [10] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017: 6000-6010.
- [11] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pretraining of deep bidirectional transformers for language understanding[C]//Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics Human Language Technologies. 2019:4171-4186.
- [12] YU J F, QIU M H, JIANG J, et al. Modelling domain relationships for transfer learning on retrieval-based question answering systems in e-commerce[C]// Proceedings of the 11th ACM International Conference on Web Search and Data Mining. 2018: 682-690.
- [13] NIE Y X, BANSAL M. Shortcut-stacked sentence encoders for multi-domain inference[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing Association for Computational Linguistics, EMNLP 2017, 2017: 41-45.
- [14] CONNEAU A, KIELA D, SCHWENK H, et al. Supervised learning of universal sentence representations from natural language inference data[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. ACL, 2017: 670-680.
- [15] KRATZWALD B, FEUERRIEGEL S. Putting question-answering systems into practice: Transfer learning for efficient domain customization[J]. ACM Transactions on Management Information Systems, 2019, 9(4): 15: 1-15: 20(Article No.15).
- [16] HE H, LIN J J. Pairwise word interaction modeling with deep neural networks for semantic similarity measurement[C]//Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics Human Language Technologies. 2016: 937-948.
- [17] CHEN Q, ZHU X D, LING Z H, et al. Enhanced LSTM for natural language inference[C]//Proceedings of the 2017 Annual Meeting of the Association for Computational Linguistics. ACL, 2017: 1657-1668.

(责任编辑: 李 艺)

(上接第 73 页)

- [2] BORKAR S, RAJESWARI K. Attributes selection for predicting students' academic performance using education data mining and artificial neural network[J]. International Journal of Computer Applications, 2014, 86(10): 25-29.
- [3] LAN A S, WATERS A E, STUDER C, et al. Sparse factor analysis for learning and content analytics[J]. Journal of Machine Learning Research, 2013, 15(1): 1959-2008.
- [4] 张嘉, 张晖, 赵旭剑, 等. 规则半自动学习的概率软逻辑推理模型[J]. 计算机应用, 2018, 38(11): 98-103.
- [5] 薛颖, 沙秀艳. 基于改进模糊聚类算法的灰色预测模型[J]. 统计与决策, 2017(9): 29-32.
- [6] 文传军, 詹永照. 基于样本模糊隶属度归  $n$  化约束的松弛模糊 C 均值聚类算法[J]. 科学技术与工程, 2017, 17(36): 96-104.
- [7] 赵琦, 孙泽斌, 冯文全, 等. 一种基于支持向量回归的建模方法[J]. 北京航空航天大学学报, 2017, 43(2): 352-359.
- [8] 张麒增, 戴翰波. 基于数据预处理技术的学生成绩预测模型研究[J]. 湖北大学学报(自然科学版), 2019, 41(1): 106-113.
- [9] 孙毅, 刘仁云, 王松, 等. 基于多元线性回归模型的考试成绩评价与预测[J]. 吉林大学学报(信息科学版), 2013, 31(4): 404-408.
- [10] 陈岷. 因子分析和神经网络相融合的体育成绩预测模型[J]. 现代电子技术, 2017(5): 138-141.
- [11] NÚÑEZ J C, SUÁREZ N, ROSÁRIO P, et al. Relationships between perceived parental involvement in homework, student homework behaviors, and academic achievement: Differences among elementary, junior high, and high school students[J]. Metacognition and Learning, 2015, 10(3): 375-406.
- [12] BUNKAR K, SINGH U K, PANDYA B, et al. Data mining: Prediction for performance improvement of graduate students using classification[C]//IEEE 2012 Ninth International Conference on Wireless and Optical Communications Networks (WOCN). New York: IEEE, 2012: 1-5.

(责任编辑: 张 晶)