

## 基于加权集成的序贯三支决策情感分类

帅常朗,钱进\*,周川鹏

(华东交通大学 软件学院,江西 南昌 330013)

**摘要:**为了提高情感分类的性能,本文提出了一种基于加权集成的序贯三支决策情感分类模型。该模型首先对评论数据集的边界域使用不同的分类器获取各自的预测概率,再根据历史分类性能对不同分类器的预测概率进行加权集成,然后根据阈值和代价损失分别进行三支决策,将评论划分为正类、负类和边界域。对于边界域进行序贯的集成概率预测,并根据概率和阈值进一步划分为新的正类、负类和边界域。直至最细粒度上的边界域,最终通过集成二支决策得到最终的分类结果。研究表明,该模型在中文计算机评论、酒店评论和服装评论数据集上性能优于现有方法,其中在酒店评论数据集上分类准确率达到86.75%,相比于基于硬投票集成的序贯三支决策情感分类提高了3.6%。

**关键词:**多粒度分类;机器学习;集成学习;文本粒化;粗糙集

**中图分类号:**TP18 **文献标志码:**A **文章编号:**0253-2395(2025)01-0066-11

## Sentiment Classification Using Sequential Three-way Decisions Based on Weighted Ensemble

SHUAI Changlang, QIAN Jin\*, ZHOU Chuanpeng

(School of Software, East China Jiaotong University, Nanchang 330013, China)

**Abstract:** To improve the performance of sentiment classification, this paper proposes a sentiment classification model based on weighted ensemble sequential three-way decision. The model firstly uses different classifiers to obtain their respective prediction probabilities for the boundary domain of the review dataset. Then, based on historical classification performance, the prediction probabilities of different classifiers are weighted and integrated. According to the threshold and cost loss, three-way decisions are made to classify the reviews into positive, negative, and boundary domains. The boundary domain is sequentially subjected to integrated probability prediction and further classified into new positive, negative, and boundary domains according to the probabilities and thresholds. Until the finest granularity of the boundary domain is reached, a final classification result is obtained through integrated two-way decision. The research results show that this model outperforms existing methods on Chinese computer reviews, hotel reviews and clothing reviews datasets. Among them, the classification accuracy on the hotel review data set reached 86.75%, which was improved by 3.6% compared with the sequential three-branch decision emotion classification based on hard voting integration.

**Key words:** multi-granular classification; machine learning; integrated learning; text granulation; rough set

收稿日期:2024-06-30;接受日期:2024-10-24

基金项目:国家自然科学基金(62066014,62466017);江西省自然科学基金项目(20232ACB202013)

作者简介:帅常朗(1999-),男,四川达州人,硕士研究生,研究方向为三支决策。E-mail:13036216086@163.com

\* 通信作者:钱进(QIAN Jin),E-mail:qjqjlyf@163.com

引文格式:帅常朗,钱进,周川鹏.基于加权集成的序贯三支决策情感分类[J].山西大学学报(自然科学版),2025,48(1):66-76. DOI:10.13451/j.sxu.ns.2024149.

## 0 引言

随着互联网的迅速发展普及,我们的生活已经离不开网络。评论作为一种新型文本语言,如产品评论、电影评论和书籍评论等,已经深入影响着我们的日常生活。消费者常常会受到这些评论的影响,正面的评论可能会增强他们的购买意愿,而负面的评论则可能会导致他们取消购买计划。对于商家来说,好的评论可以加强他们的信心,而不好的评论则可能促使他们改进服务质量。因此,对大量评论数据进行快速、高效且准确的分类处理已成为自然语言处理领域的研究热点。情感分析<sup>[1]</sup>也被称为观点挖掘,旨在分析和挖掘隐藏在文本数据中的情感极性。情感分类作为情感分析的关键组成部分,旨在将文本数据划分为不同的情感类别,通常包括正面和负面情感。其目标是通过计算机自动识别和理解评论文本中包含的情感极性,以便从大规模文本数据中提取有用信息。常见的情感分类方法主要有基于情感词典<sup>[2]</sup>、基于特征提取的机器学习方法<sup>[3-5]</sup>和深度学习<sup>[6-8]</sup>方法。单一的使用情感词典无法涵盖所有情感表达方式和领域,因为它只包含有限的情感词汇,会导致无法准确捕捉所有情感;而单一使用机器学习的方法通常需要大量的标签数据来训练模型,成本较高;深度学习在情感分类方面取得了较大进展,但是存在数据需求量大且对硬件要求高的缺点。这些问题揭示了当前情感分类领域中的科学挑战,亟需新的方法来提升分类效果和降低成本。

在此背景下,引入三支决策<sup>[9]</sup>模型,其是在二支决策(接受与拒绝)的基础上进行的改进,引入了第三种决策选项,即延迟决策,更符合人类的认知过程,因为在某些情况下,决策者可能无法立即做出确定的决策,而需要进一步获取信息或者延迟做出决策。这样应用包括医疗诊断<sup>[10]</sup>、人脸识别<sup>[11]</sup>、隐私保护<sup>[12]</sup>等。序贯三支决策就是将三支决策作为由粗粒度到细粒度的一个动态决策过程。Zhou等<sup>[13]</sup>将动态词典结合到了三支决策中来进行情感分析。Zhang等<sup>[14]</sup>使用N-gram模型构建了一个多粒度空间,以此来进行序贯三支决策情感分类。Yang等<sup>[15]</sup>从时间和空间的角度出发,提出了基

于时空多粒度的序贯三支决策情感分析模型。Wang等<sup>[16]</sup>引入集成学习的硬投票方法结合到序贯三支决策模型进行情感分类。Chen等<sup>[17]</sup>从特征提取方面出发,提出了用于情感分析的分类特征表示三支决策模型。Su等<sup>[18]</sup>为了更好地处理复杂环境下不确定的数据,提出了基于知识融合的序贯三支决策模型社交媒体情感分析。

然而,如何提升模型的分类性能和泛化能力仍然是一个挑战。所以这里引入加权集成<sup>[19]</sup>的概念,加权集成的原理在于对多个独立分类器的历史性能赋予不同权重,通过综合各个分类器的预测结果来增强最终的分类效果。这种方法已在许多领域取得成功,尤其是在文本分类和情感分析中,研究表明加权集成能够显著提高模型的准确性和鲁棒性。加权集成的优点包括:首先,它能有效整合不同模型的优势,减轻单一模型的不足;其次,通过加权方式,能够针对不同任务动态调整模型组合,提高适应性。然而,加权集成也存在一定的缺点,如模型选择和权重分配的复杂性,这可能导致实施上的挑战。

本文在序贯三支决策的基础上,结合加权集成和N-gram语言模型,构建了一个多粒度的加权集成序贯三支决策情感分类研究模型,该模型将多个独立分类器的效果根据它们的历史分类性能赋予不同的权重,然后对多个分类器的结果进行加权集成,从而得到最终的分类结果。通过在中文计算机、酒店和服装评论数据集上应用加权集成多种独立分类器的效果来提高模型的性能。最后的结果表明,通过加权集成后的序贯三支决策模型能够显著提高情感分类的性能,并增强模型的泛化能力。

## 1 基本原理

### 1.1 粗糙集

由Pawlak<sup>[20]</sup>在1982年提出,是一种处理不精确、不确定与不完全数据的新的数学方法。在粗糙集理论中特征选择的研究对象主要以符号型数值的数据为主,可以表示为一个四元组 $S=(U, A, V, f)$ ,其中 $U$ 为非空有限对象集合,即为论域, $A$ 为非空有限特征集合, $V_a \in A, V_a$

表示特征  $a$  的值域。 $f$ : 为  $U \times A \rightarrow V$  是一个信息函数,  $\forall x \in U, a \in A$ , 定义  $f(x, a)$  表示  $x$  在特征  $a$  上的取值, 则有  $f(x, a) \in V_a$ 。若有  $A = C \cup D$ , 且  $C \cap D = \emptyset$ , 其中  $C$  为条件属性集合,  $D$  为决策属性集合, 此信息系统又被称为决策系统。其等价关系的定义如下所示,

$$E_A = \{(x, y) \in U \times U \mid \forall a \in A \subseteq C \subseteq At\}。 (1)$$

其中  $(U, E_A)$  是非空有限特征集合上的近似空间,  $U/E_A$  是对  $U$  的划分,  $x$  的等价类如下所示,

$$[x]_{E_A} = [x]_A = [x] = \{y \in U \mid (x, y) \in E_A\}。 (2)$$

状态集为  $\Omega = (X, X^c)$ , 分别表示该评论属于积极评论还是消极评论, 其中属于  $X$  的条件概率使用公式(3)计算。

$$p(X \mid [X]) = \frac{|X \cap [x]|}{|[x]|}。 (3)$$

不属于  $X$  的条件概率为  $1 - p(X \mid [x])$ 。动作集为  $A = \{a_P, a_B, a_N\}$ , 三种动作依次对应接受, 延迟决策和拒绝。损失函数表如表 1 所示。

表 1 三支决策损失函数

Table 1 Loss functions of three-way decisions

决策规划	$X$	$X^c$
$P$	$\lambda_{PP}$	$\lambda_{PN}$
$B$	$\lambda_{BP}$	$\lambda_{BN}$
$N$	$\lambda_{NP}$	$\lambda_{NN}$

基于上面的损失函数表可以给出阈值的计算公式如下。

$$\alpha = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}, (4)$$

$$\beta = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}, (5)$$

$$\gamma = \frac{(\lambda_{PN} - \lambda_{NN})}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}。 (6)$$

其中的  $\alpha$  表示最小接受阈值,  $\beta$  表示最小拒绝阈值,  $P$ 、 $B$  和  $N$  代表接受、延迟分类、和拒绝的三种决策规则。 $\lambda_{PP}$  表示文本被正确划分为正类的成本,  $\lambda_{NN}$  表示文本被正确划分为负类的成本,  $\lambda_{BP}$  表示文本被划分为边界类但实际为正类的成本,  $\lambda_{BN}$  表示文本被划分为边界类但实际为负类的成本,  $\lambda_{PN}$  表示文本被划分为正类但实际为负类的成本,  $\lambda_{NP}$  表示文本被划分为负类但实际为正类的成本。

## 1.2 序贯三支决策

三支决策理论是 Yao<sup>[9]</sup>提出的一种相比于二元思维更符合人类日常思维的一种决策模式, 如果无法根据现有的信息和提示来进行决策, 则进入延迟决策, 提取更多的信息后再来进行决策, 三支决策规则旨在为概率粗糙集的三个邻域提供明确的语义解释, 并在不确定或信息不完整的情况下提供一种决策方法。这三个邻域分别对应概率粗糙集的构建, 包括接受、拒绝和不承诺规则。在实际应用中, 通过调整不同的阈值, 可以产生不同的决策结果, 因此选择适当的阈值至关重要。通过引入贝叶斯决策过程到概率粗糙集模型中, 根据不同的决策成本选择使总体风险最小化的决策代价作为损失函数。从三支决策方法提出以来, 已经有许多的专家学者根据不同的需求改进提出了新的三支决策。例如, Chen 等<sup>[21]</sup>将三支决策和多属性决策相结合提出了 3WD-gmcr 模型, 用于处理冲突分析问题, Qian 等<sup>[22]</sup>引入多层次决策表, 构建了三支决策的分层决策模型, Wang 等<sup>[23]</sup>将犹豫模糊信息与三支决策方法相结合, 在排序和分类方面表现出良好的性能, Qian 等<sup>[24]</sup>对多粒度三支决策的现状进行了研究。

然而, 现有的三支决策方法未考虑在决策过程中获取和使用信息的成本。在实际问题中, 获得解决问题的有效信息往往是一个逐步迭代的过程, 初始获得的信息可能不足以直接支持做出三支决策。基于这一考虑, 提出了序贯三支决策方法, 将三支决策作为序贯决策的一个中间过程, 以最小化信息成本的方式获得更好的决策效果。序贯三支决策的核心思想是按照从粗到细的顺序, 在每一粒度中, 当现有信息不足以支持接受或拒绝决策时, 选择延迟决策, 等待下一粒度补充信息后再进行三支决策判断。对于常见的二支决策问题, 该过程会在合适的粒度下根据信息量增加和决策成本的考虑获得二支决策结果。

三支决策理论传统二支决策理论的扩展。在解决实际问题时, 二支决策适用于信息充足或信息获取成本较低的情况, 但实际上获得的初始信息往往不足以直接支持决策者做出适当的决策, 对于文本情感分类也是这样。其

中的阈值设置为  $0 < \beta_i \leq \alpha_i < 1$ , 阈值也会随粒度的变化而变化, 粗粒度具有更大的  $\alpha$  和更小的  $\beta$ , 细粒度则相反。假设  $i = 1, 2, 3, \dots, i$  为粒度层次, 则阈值具有以下规律:  $0 < \beta_1 \leq \beta_2 \leq \beta_3 \leq \dots \leq \beta_i < \alpha_i \leq \dots \leq \alpha_3 \leq \alpha_2 \leq \alpha_1 < 1$ 。

序贯三支决策的模型图如图 1 所示。

图 1 展示了序贯三支决策的具体实现步骤: 先将论域 (数据集) 划分为 POS<sub>1</sub> (正域)、BND<sub>1</sub> (边界域) 和 NEG<sub>1</sub> (负域), 按照从粗粒度 1—细粒度  $i$  的顺序依次对得到的 BND (边界域) 进行三个域的划分, 直至最细粒度上划分为两个域。

### 1.3 集成学习

集成学习利用多个基本模型的集体智慧, 通过整合各个模型的预测结果, 从而产生比单个模型更为准确和稳健的预测。这种方法能够弥补单个模型的局限性, 并在面对复杂数据或任务时提供更可靠的解决方案。集成学习的核心思想是通过结合多个模型的优势以及合理的整合策略, 使得集成模型具有更好的泛化能力和预测性能, 从而在实际应用中取得更好的效果。根据其实现方式和策略的不同可以分为 Bagging<sup>[25]</sup>, Boosting<sup>[26]</sup>, Stacking<sup>[27]</sup> 和 Voting<sup>[28]</sup>。将集成学习融入其他领域来提高性能

已经成为一个趋势<sup>[29-30]</sup>, Huang 等<sup>[30]</sup>将反向传播神经网络 (Back Propagation Neural Network, BPNN) 模型与集成三支决策相结合, 提高了其预测性能; Jiang 等<sup>[31]</sup>将集成三支聚类和阴影集相结合提出了 S-M3WCE 模型, 对比其他集成聚类算法获得了更高的性能; Qian 等<sup>[32]</sup>使用四种具有不同决策标准的三支决策模型通过集成算法处理分类问题, 相比于其他传统三支决策模型可以获得更高的分类精度和更低的延迟率。

## 2 基于加权集成学习的序贯三支决策文本情感分类

### 2.1 粒化文本

文本粒化是将文本数据划分成更小的片段或单位的过程<sup>[33]</sup>。在文本情感分类中, 通常将粒度层次分为词语、句子、段落和篇章级。研究者往往专注于某一特定粒度层次, 但文本情感分析存在较多模糊性。仅在粗粒度进行特征提取与分类虽能节省成本, 但可能忽略关键情感信息, 导致分类性能下降; 而在细粒度上进行特征提取与分类则可提高分类准确率, 但成本较高, 效率低下。因此, 本文提出根据每条评论中情感信息含量作为粒度划分的依据, 逐渐由粗到细划分粒度层次, 以求解这一问题。

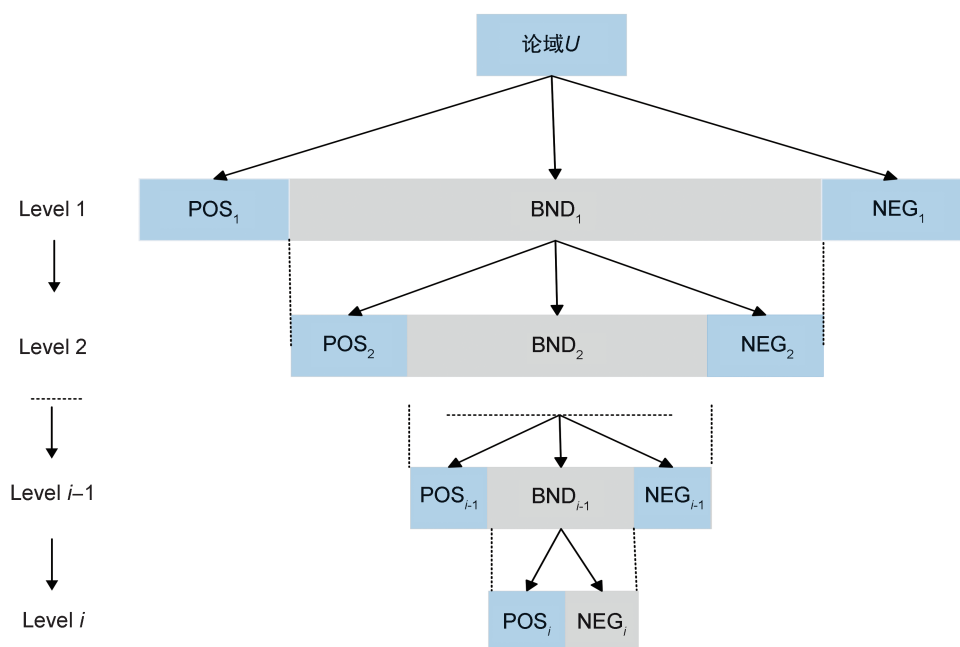


图 1 序贯三支决策模型图

Fig. 1 Diagram of sequential three-way decisions model

2.2 加权集成下的序贯三支决策文本情感分类

首先在较粗粒度进行三支决策,将划入到边界域的部分根据下一粒度层次中的情感信息特征继续进行三支决策,最后获得分类结果。为了增强模型的泛化能力和分类性能,在上面的基础上引入了集成学习,它是将多个弱分类器通过策略进行融合依次变成一个强分类器获得更好的分类结果。具体的思路是:从粗到细的多个粒度下,分别先使用不同的分类器独立地对评论对象进行概率预测,再对得到的不同概率根据经验和历史分类性能进行权重分配和加权集成,再将得到的加权集成概率根据三支决策的损失和阈值设置得到正、负和边界三个域,再序贯执行上面的步骤,直到最细粒度,再对最细粒度的边界域进行二支决策,将每个粒度层次得到的正样本和负样本相加得到最终的分类结果。如图2所示。

详细的描述如算法1所示。

**算法1** 基于加权集成的序贯三支决策情感分类方法

输入: 中文评论数据集  $X = \{x_1, x_2, \dots, x_n\}$ , 不同粒度的阈值对  $(\alpha_i, \beta_i, \gamma_i)$ , 加权集成分类器集合  $\{model_1, model_2, \dots, model_k\}$ , 以及每个分类器的权重  $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ 。

输出: 决策结果, 即 POS 和 NEG

1. initialize  $U_1 = U, U_n = \emptyset, i = 1, POS =$

$\emptyset, NEG = \emptyset, BND = \emptyset$

2. for  $i=1, 2, \dots, n-1$  do

3.  $prob\_x = 0;$

4. for each classifier model\_j in ensemble classifiers do

5.  $prob\_x += \lambda_j * model\_j.predict\_proba(x_i)[1];$

6. end for

7.  $POS_i = \{x \in U_i | prob\_x \geq \alpha_i\}$

8.  $BND_i = \{x \in U_i | \beta_i < prob\_x < \alpha_i\}$

9.  $NEG_i = \{x \in U_i | prob\_x \leq \beta_i\}$

10.  $POS = POS \cup POS_i$

11.  $NEG = NEG \cup NEG_i$

12.  $U_{i+1} = BND_i$

13.  $i = i + 1$

14. end for

15. if  $U_n \neq \emptyset$  then

16.  $prob\_x = 0$

17. for each classifier model\_j in ensemble classifiers do

18.  $prob\_x += \lambda_j * model\_j.predict\_proba(x_i)[1]$

19. end for

20.  $POS_n = \{x \in U_n | prob\_x \geq \gamma_n\}$

21.  $NEG_n = \{x \in U_n | prob\_x < \gamma_n\}$

22.  $POS = POS \cup POS_n$

23.  $NEG = NEG \cup NEG_n$

24. end if。

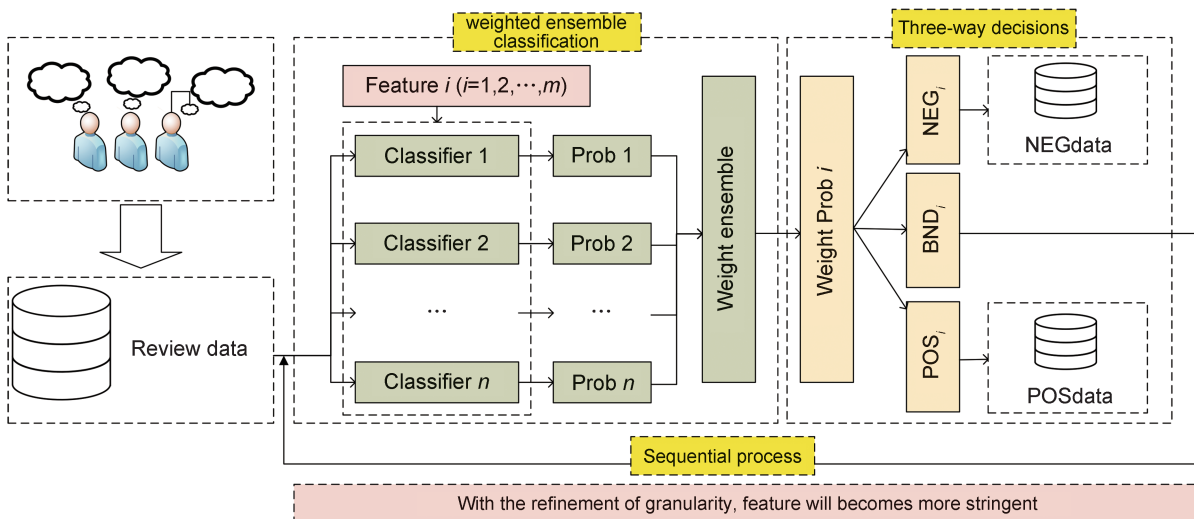


图2 基于加权集成的序贯三支决策情感分类模型

Fig. 2 A sequential three-way decisions sentiment classification model based on weighted ensemble

算法1首先使用多个分类器模型分别对文本进行概率预测再根据设置的权重 $\lambda$ 进行加权概率的集成,最后将得到的概率与设置的两个阈值 $\alpha$ 和 $\beta$ 进行比较,通过三支决策划分为正、负和边界域,再对边界域通过不同分类器的重新概率预测,进行序贯的集成概率计算,最后对最细粒度的边界域使用集成概率进行二支决策得到正域和负域,再将上面步骤得到的正域和负域相加即为最终分类结果。包含两个嵌套的循环,外层遍历数据集中的每条文本,内层循环遍历每个分类器,上面文本数量为 $n$ ,分类器数量为 $k$ ,得出时间复杂度为 $O(n \cdot k)$ 。这意味着算法的执行时间会随着文本数量和分类器数量的增加而线性增长,展现了算法在处理大规模数据集和复杂分类任务时的效率。算法1能够在提高分类精度的同时,有效处理分类不确定性,为文本情感分类问题提供了高效且准确的解决方案。

### 3 实验结果与分析

#### 3.1 实验数据集

为了验证本文所提出的方法有效性和优越性,采用中文计算机评论、酒店评论和服装评论三个数据集作为本文的实验数据集,其中中文计算机评论数据集中有4000条评论,正负评论各2000条;酒店评论数据集总共有6000条评论,正负评论各占3000条,从中随机选取正负评论各2000条;服装评论数据集共有10000条,正负评论各5000条,从中随机选取正负评论各2000条。数据集信息如表2所示。

##### 3.1.1 预处理

文本预处理是自然语言中处理的重要步骤,一般分为文本清洗、分词、移除停用词和词形还原等,因为英文文本具有文本空格隔断的特征,所以英文评论数据集不用进行分词,而对于中文评论数据集来说需要先进行分词,再进行文本清洗和移除停用词的操作,表3是预

处理后的部分评论文本。

表2 数据集构成

Table 2 Composition of dataset

类别	计算机评论	酒店评论	服装评论
正样本	2 000	2 000	2 000
负样本	2 000	2 000	2 000
总评论数	4 000	4 000	4 000

##### 3.1.2 文本特征提取

N-gram模型是一种用于建模序列数据的统计语言模型,在自然语言处理任务中具有广泛的应用,如语言建模、文本生成和文本分类。该模型基于马尔可夫假设,即当前词的出现只与前面 $N-1$ 个词有关。具体而言,N-gram模型通过统计文本中连续 $N$ 个词(或字符)的出现频率和分布情况,从而推断出文本的语言规律和结构特征。借助N-gram模型,能够计算在给定前 $N-1$ 个词的情况下,下一个词(或字符)出现的概率,从而实现自动建模和预测。具体的概率预测公式(7)所示:

$$P(\omega_i | \omega_1, \omega_2, \dots, \omega_{i-1}) = P(\omega_i | \omega_{i-N+1}, \omega_{i-N+2}, \dots, \omega_{i-1}) \quad (7)$$

一般分为unigram、bigram和trigram,相对应的分别为一元模型、二元模型和三元模型。其中unigram只考虑单个词语的出现频率和分布情况,不考虑词语之间的顺序关系;bigram是指在N-gram模型中考虑两个连续词语的组合,每个词语的出现概率与其前一个词语的出现有关,即用前一个词语来预测下一个词语的出现概率;trigram中每个词语的出现概率与其前面两个词语的出现有关,用前两个词语组合来预测下一个词语的出现概率。

这里构造特征的原则是第一粒度使用unigram筛选名词,第二粒度为bigram筛选形容词+名词,在第三粒度考虑到数据规模避免数据稀疏性和计算复杂度的问题,所以不使用trigram,而是使用unigram+bigram来筛选副词+形容词+名词作为特征。这种按照情感信息量

表3 预处理评论文本

Table 3 Pre-processed comment text

编号	原文本	预处理后文本	Sentiment
1	服务很差!!! 服务员的态度很不好! 环境也一般	服务 很差 服务员 态度 很 不好 环境 也 一般	Negative
2	环境非常好,服务非常好,交通不方便,价格有点贵	环境 非常 好 服务 非常 好 交通 不 方便 价格 有点 贵	Positive

不同的划分具有层层递进更加准确高效地捕捉其中的情感信息,比如:不错的服务,在第一个粒度层次“服务”无法分辨情感倾向,放在边界域进行第二粒度“错服务”,更偏向消极类情感,初定为消极类,第三粒度“不错服务”显然为积极类,这里可以得到积极类的倾向大于消极类的倾向,也就推翻了前面的初定义为消极类。这样层层递进的特征构造方法能够更好地反映文本中的情感信息,提高情感分类的准确性和效率。具体的如表4所示。

表4 特征选取

Table 4 Characteristic selection

粒度层次	待选特征			
1	酒店	地毯	服务	交通
2	好酒店	脏地毯	好服务	方便交通
3	最好酒店	太脏地毯	不好服务	很方便交通

### 3.2 评价指标

为了衡量本文方法的性能,使用了 Accuracy ( $A$ )、Precision ( $P$ )、Recall ( $R$ )和  $F1$  四种常见的性能指标。根据分类结果的混淆矩阵来进行计算,其中  $TP$  表示真正例数量(分类器将正例正确分类为正例的数量),  $TN$  表示真负例数量(分类器将负例正确分类为负例的数量),  $FP$  表示假正例数量(分类器将负例错误分类为正例的数量),  $FN$  表示假负例数量(分类器将正例错误分类为负例的数量)。具体的计算公式如下:

$$A = \frac{TP + TN}{TP + FN + FP + TN}, \quad (8)$$

$$P = \frac{TP}{TP + FP}, \quad (9)$$

$$R = \frac{TP}{TP + FN}, \quad (10)$$

$$F1 = \frac{2 \times P \times R}{P + R}. \quad (11)$$

### 3.3 实验设置

#### 3.3.1 损失及阈值设置

这里设置每个粒度下的损失,一般来说是根据自己的经验和实际情况来进行综合设置。损失其实就是对于分类错误的代价,因为本文研究的是情感分类,最终的结果应该为积极类和消极类两类,所以就是将积极类和消极类分为正、负和不确定域所付出的代价,根据常识及原则来看,从粗粒度到细粒度,错误分类的

代价也在提高,这里设置为从粗粒度开始,每个粒度损失依次增加;而将积极类分到负域和不确定域的代价应该比将消极类分到正域和不确定域的代价更高,正确分类的损失都设置为0。考虑到分类代价的问题,所以这里只选择三个粒度进行三支决策。这样的设置不仅能够考虑情感信息的重要性,还能合理地控制决策的成本,从而实现更加有效的情感分类。根据上面阈值的计算公式和给出的定义以及综合考虑分类效果,将计算机评论的第一粒度层次的阈值对设置为  $\alpha_1 = 0.74, \beta_1 = 0.30$ ; 酒店评论的阈值对设置为  $\alpha_1 = 0.73, \beta_1 = 0.27$ ; 服装评论的阈值对设置为  $\alpha_1 = 0.71, \beta_1 = 0.32$ 。

#### 3.3.2 加权集成权重设置

设置逻辑回归(Logistic Regression, LR)、朴素贝叶斯(Naive Bayes, NB)和支持向量机(Support Vector Machine, SVM)三个基分类器的权重比在0—10以内,三者之和为10。在第一粒度下通过遍历权重输出错误分类的数量确定最佳权重比。根据遍历结果可知在计算机评论数据集中 LR:NB=6:4 时错误样本数量最少,设置为最佳权重比。而在酒店评论数据集中得出的最佳权重组合为 SVM:LR:NB=7:2:1。在服装评论数据集中最佳权重组合为 SVM:LR:NB=5:3:2。

#### 3.4 实验结果及分析

为了验证阈值设置的合理性和有效性,这里遍历阈值对不同组合第一粒度下对于酒店评论、计算机评论和服装评论数据集的分类准确率形成的散点,其中阈值  $\alpha$ , 阈值  $\beta$  和准确率(Accuracy)为坐标。注:准确率的取值范围为0到1,0表示0%而1表示100%。例如,准确率为0.85时,表示模型的正确率为85%。在后续的分析 and 图表中,将始终以0到1之间的小数形式表示准确率、查准率、查全率和  $F1$  值。

从图3可以看出,在计算机评论数据集中,当第一粒度的阈值对为  $\alpha_1 = 0.74, \beta_1 = 0.30$  时能获得更高的准确率,在酒店评论数据集中当第一粒度的阈值对为  $\alpha_1 = 0.73, \beta_1 = 0.27$  能获得更好的实验结果,而在服装评论数据集中当第一粒度阈值对为  $\alpha_1 = 0.71, \beta_1 = 0.32$  能获得更高的准确率,证明了三个阈值对选取的有效性。

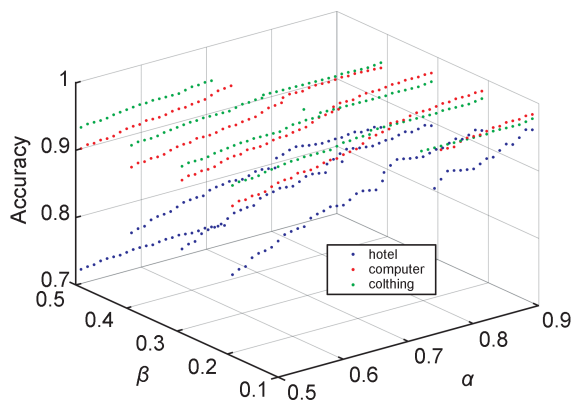


图3 第一粒度下不同阈值组合的准确率

Fig. 3 Accuracy of the combination without thresholds at the first granularity

为验证本文提出的粒度构建方法的有效性,以酒店评论数据集为例,使用NB进行分析,计算三个粒度下的分类准确率,如图4所示。

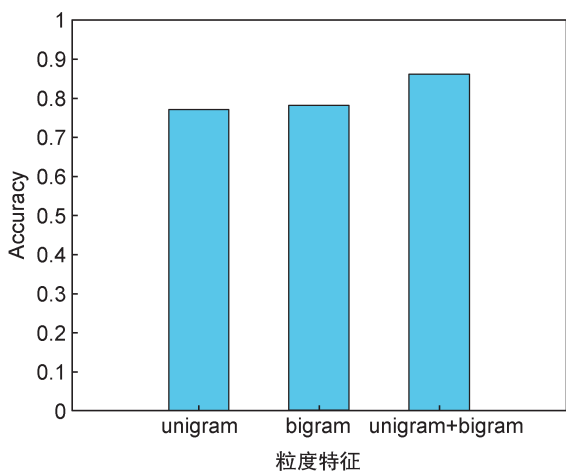


图4 朴素贝叶斯在三个粒度特征下对酒店评论数据集的分类准确率

Fig. 4 Classification accuracy of naive bayes on the hotel review dataset with three granularity features

从图4中可以看出使用unigram提取名词作为特征时,准确率较低,使用bigram提取形容词+名词作为特征时提升了1个百分点,使用unigram+bigram提取副词+形容词+名词作为特征时,有大幅提升,这是因为形容词和副词往往包含大量的情感信息,这个结果也验证了多粒度构造的有效性。

本文以LR、NB和SVM这三个二支决策方法作为基分类器,以文献[14]中的基于N-gram语言模型的多粒度序贯三支情感分类模型(N-gram-based Multi-granularity Sequential Three-

way Sentiment Classification, NSTWSC)和文献[16]中的结合集成学习方法的序贯三支情感分类模型(Ensemble Learning for Sequential Three-way Sentiment Classification, ESTWSC)的对比基准,首先用三种三支决策方法在两个数据集上对比三个二支决策方法的效果,再使用本文提出的基于加权集成的序贯三支决策情感分类模型(Weighted Ensemble-based Sequential Three-way Decision Sentiment Classification, WETWSC)对比NSTWSC和ESTWSC的结果。分别在中文酒店评论、计算机评论和服装评论数据集上进行实验,来验证本文提出的基于加权集成的序贯三支决策的情感分类模型WETWSC的有效性。

通过分析表5中的查准率P,可以发现在三种二元分类方法中,针对中文计算机评论数据集,负样本方面,LR表现最为出色,而SVM在正样本上表现最佳。相反,在中文酒店评论数据集中,NB在负样本上表现最佳,而LR则在正样本方面表现优异。这表明了每种二元分类方法各有其优劣之处,呈现出一定的平衡性。对于三种三支决策分类方法,结果显示在三个数据集的正、负样本上,均优于二元分类方法的查全率效果。这表明相较于二元分类方法,三元分类方法在情感分类中的查准率更为有效。

进一步对比三种三元分类方法在三个数据集上的表现,可以发现在计算机评论数据集中,本文提出的WETWSC方法在负样本上表

表5 查准率P的对比结果

Table 5 Comparing results of Precision (P)

方法	类别	计算机评论	酒店评论	服装评论
LR	Negative	0.863	0.863	0.829
	Positive	0.890	0.867	0.858
SVM	Negative	0.859	0.811	0.823
	Positive	0.910	0.835	0.824
NB	Negative	0.755	0.881	0.825
	Positive	0.830	0.817	0.814
NSTWSC	Negative	0.888	0.855	0.852
	Positive	0.904	<b>0.889</b>	0.914
ESTWSC	Negative	0.910	0.839	<b>0.947</b>
	Positive	<b>0.931</b>	0.822	0.922
WETWSC	Negative	<b>0.918</b>	<b>0.887</b>	0.921
	Positive	0.898	0.878	<b>0.929</b>

现最佳,相较于NSTWSC和ESWSC分别提升了3个百分点和0.8个百分点。而在正样本方面,ESTWSC表现最佳。而在酒店评论数据集上,NSTWSC方法在正样本方面表现最佳,而在负样本方面,WETWSC方法表现最优。在服装评论中,ESTWSC在正样本上表现最佳,本文提出的WETWSC在负样本最佳,但是ESTWSC存在正、负样本查准率相差较大的缺陷。综合来看,在三个数据集上的查准率上,本文的WETWSC均取得了不错的结果。

表6 查全率R的对比结果

Table 6 Comparing results of recall (R)

方法	类别	计算机评论	酒店评论	服装评论
LR	Negative	0.911	0.884	0.844
	Positive	0.861	0.844	0.844
SVM	Negative	<b>0.916</b>	0.863	0.799
	Positive	0.850	0.775	0.846
NB	Negative	0.852	0.825	0.783
	Positive	0.724	<b>0.876</b>	0.851
NSTWSC	Negative	0.906	<b>0.903</b>	0.881
	Positive	0.886	0.834	0.886
ESTWSC	Negative	0.914	0.841	0.910
	Positive	<b>0.926</b>	0.820	<b>0.954</b>
WETWSC	Negative	0.907	0.901	<b>0.921</b>
	Positive	0.910	0.831	0.928

通过分析表6查全率R的对比结果,我们可以得出以下结论:

针对计算机评论数据集的负样本方面,SVM方法的查全率最高,达到了0.916。而在正样本方面,ESTWSC方法的表现最佳。对于酒店评论数据集的负样本方面,NSTWSC方法的查全率最高,达到了0.903,略高于本文的WETWSC。而在正样本方面,NB方法的表现最佳,但是同样存在正负样本上表现差异较大不平衡的结果。而在服装评论数据集上,ESTWSC在正样本上表现最佳,本文的WETWSC在负样本上表现最佳。

总体来说,ESTWSC和WETWSC两种方法在计算机评论和服装评论数据集的查全率上均取得了较好结果,而由于酒店数据集中评论的特点导致查全率一般。

通过对表7中的F1值对比结果的深入分析,我们可以得出如下结论:在计算机评论和

服装评论数据集上ESTWSC表现较好,但是除计算机评论的正样本的F1值领先本文方法较大之外,其他样本上都相差不大,相比于其他的方法,两种集成方法均展现了卓越的性能,F1值处于领先地位。而在酒店评论上,本文的WETWSC在F1值上大幅领先,其中在负样本上相比于NSTWSC和ESTWSC分别提高了2个百分点和5个百分点。

表7 F1值的对比结果

Table 7 Comparing results of F1 values

方法	类别	计算机评论	酒店评论	服装评论
LR	Negative	0.877	0.874	0.836
	Positive	0.875	<b>0.855</b>	0.851
SVM	Negative	0.887	0.836	0.811
	Positive	0.879	0.804	0.835
NB	Negative	0.801	0.852	0.803
	Positive	0.774	0.846	0.832
NSTWSC	Negative	0.897	0.875	0.866
	Positive	0.895	0.851	0.900
ESTWSC	Negative	<b>0.912</b>	0.840	<b>0.928</b>
	Positive	<b>0.928</b>	0.821	<b>0.937</b>
WETWSC	Negative	0.910	<b>0.894</b>	0.921
	Positive	0.904	0.854	0.928

最后在两个数据集上对两种不同的三支决策方法和二支分类效果最好的SVM准确率进行比较,如图5所示。

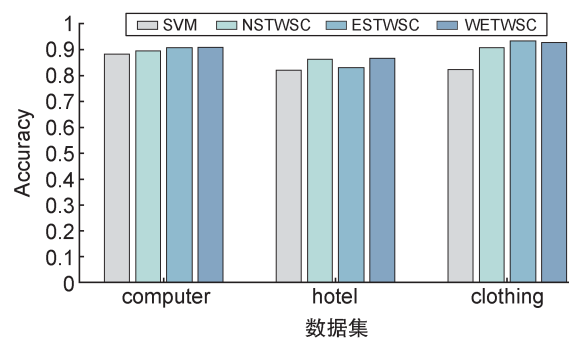


图5 SVM、NSTWSC、ESTWSC和WETWSC在中文计算机、酒店和服装评论数据集上的分类准确率比较

Fig. 5 Comparison of classification accuracy of SVM, NSTWSC, ESTWSC, and WETWSC on chinese computer, hotel, and clothing review datasets

从图5中可以看出在计算机数据集上,本文提出的WETWSC相比于NSTWSC和SVM的准确率分别提高了1.2个百分点和2.5个百分点,稍高于ESTWSC。在酒店数据集上,WET-

WSC 相比于 ESTWSC、NSTWSC 和 SVM 的准确率分别提高了 3.6 个百分点、0.3 个百分点和 4.6 个百分点。而在服装评论数据集上,本文的 WETWSC 高于 SVM 和 NSTWSC,但是略低于 ESTWSC。值得注意的是上面的方法在酒店评论数据集上的表现均不如在服装评论和计算机评论数据集上的表现,这是由于酒店评论中往往存在观点表述不清,难以判断情感倾向的评论,加大了情感分类难度。

综合来说,本文提出的 WETWSC 方法在计算机、酒店和服装评论数据集的性能表现优异,其能够更加准确地识别负样本和正样本,为情感分类任务提供了更为可靠的解决方案。

#### 4 结论与展望

在已有的三支决策与情感分类结合的研究基础上,本文引入了加权集成学习,提出了一种基于加权集成的序贯三支决策情感分类模型。该模型利用 N-gram 构建了一个多粒度结构,在每个粒度层次上采用加权集成的方法对预测概率进行加权,旨在提高模型的泛化性能和分类准确度。实验结果表明将加权集成学习融入三支决策模型中能够显著提高情感分类的性能。相对于传统的简单三支决策方法硬投票集成的序贯三支决策方法,本文提出的加权集成方法具有更高的性能指标和更强的鲁棒性。

然而,本文中使用的权重是基于历史分类性能和经验设定的,可能存在一定的局限性。未来的研究方向包括优化权重的确定方式以及深入探索更有效的特征提取方法。这些工作将进一步提升模型的性能,并丰富情感分类领域的研究内容。

#### 参考文献:

- [1] BING L. Sentiment Analysis and Opinion Mining (Synthesis Lectures on Human Language Technologies)[M]. Berlin: Springer, 2012.
- [2] TABOADA M, BROOKE J, TOFILOSKI M, et al. Lexicon-based Methods for Sentiment Analysis[J]. *Comput Linguist*, 2011, **37**(2): 267-307. DOI: 10.1162/coli\_a\_00049.
- [3] ZOU H, TANG X H, XIE B, et al. Sentiment Classification Using Machine Learning Techniques with Syntax Features[C]//2015 International Conference on Computational Science and Computational Intelligence (CSCI). New York: IEEE, 2015: 175-179. DOI: 10.1109/CSCI.2015.44.
- [4] AGARWAL B, MITTAL N. Machine Learning Approach for Sentiment Analysis[M]// Prominent Feature Extraction for Sentiment Analysis. Cham: Springer, 2016: 21-45. DOI: 10.1007/978-3-319-25343-5\_3.
- [5] TRIPATHY A, AGRAWAL A, RATH S K. Classification of Sentiment Reviews Using N-gram Machine Learning Approach[J]. *Expert Syst Appl*, 2016, **57**: 117-126. DOI: 10.1016/j.eswa.2016.03.028.
- [6] DANYAL M M, KHAN S S, KHAN M, et al. Proposing Sentiment Analysis Model Based on BERT and XLNet for Movie Reviews[J]. *Multimed Tools Appl*, 2024, **83**(24): 64315-64339. DOI: 10.1007/s11042-024-18156-5.
- [7] ZGORNI G, QUSSAY A, ELGENDY Z, et al. Deep Learning Approaches for Sentiment Analysis: Comparative Results[C]//2024 6th International Conference on Computing and Informatics (ICCI). New York: IEEE, 2024: 59-64. DOI: 10.1109/ICCI61671.2024.10485058.
- [8] REDDY R, NAOMAN A A, CHARAN G V S, et al. Sentiment Analysis of Steam Reviews Using Transformer Models[M]//Lecture Notes in Electrical Engineering. Singapore: Springer Nature Singapore, 2024: 719-727. DOI: 10.1007/978-981-99-7137-4\_70.
- [9] YAO Y Y. An Outline of a Theory of Three-way Decisions[C]//International Conference on Rough Sets and Current Trends in Computing. Berlin, Heidelberg: Springer, 2012: 1-17. DOI: 10.1007/978-3-642-32115-3\_1.
- [10] YE J, SUN B Z, BAI J C, et al. A Preference-approval Structure-based Non-additive Three-way Group Consensus Decision-making Approach for Medical Diagnosis[J]. *Inf Fusion*, 2024, **101**: 102008. DOI: 10.1016/j.inffus.2023.102008.
- [11] LI H X, ZHANG L B, HUANG B, et al. Sequential Three-way Decision and Granulation for Cost-sensitive Face Recognition[J]. *Knowl Based Syst*, 2016, **91**: 241-251. DOI: 10.1016/j.knosys.2015.07.040.
- [12] QIAN J, JIANG H Y, YU Y, et al. Multi-level Personalized K-anonymity Privacy-preserving Model Based on Sequential Three-way Decisions[J]. *Expert Syst Appl*, 2024, **239**: 122343. DOI: 10.1016/j.eswa.2023.122343.
- [13] 周哲, 商琳. 一种基于动态词典和三支决策的情感分析方法[J]. *山东大学学报(工学版)*, 2015, **45**(1): 19-23. DOI: 10.6040/j.issn.1672-3961.1.2014.250.  
ZHOU Z, SHANG L. A Sentiment Analysis Method Based on Dynamic Lexicon and Three-way Decision

- [J]. *J Shandong Univ Eng Sci*, 2015, **45**(1): 19–23. DOI: 10.6040/j.issn.1672-3961.1.2014.250.
- [14] 张刚强, 刘群, 纪良浩. 基于序贯三支决策的多粒度情感分类方法[J]. *计算机科学*, 2018, **45**(12): 153–159. DOI: 10.11896/j.issn.1002-137X.2018.12.024. ZHANG G Q, LIU Q, JI L H. Multi-granularity Sentiment Classification Method Based on Sequential Three-way Decisions[J]. *Comput Sci*, 2018, **45**(12): 153–159. DOI: 10.11896/j.issn.1002-137X.2018.12.024.
- [15] 杨新, 刘盾, 李楸柯, 等. 基于时空多粒度的序贯三支情感分析[J]. *模式识别与人工智能*, 2020, **33**(8): 743–752. DOI: 10.16451/j.cnki.issn1003-6059.202008008. YANG X, LIU D, LI Q K, *et al.* Sequential Three-way Sentiment Analysis Based on Temporal-spatial Multi-granularity[J]. *Pattern Recognit Artif Intell*, 2020, **33**(8): 743–752. DOI: 10.16451/j.cnki.issn1003-6059.202008008.
- [16] 王琴, 刘盾. 结合集成学习的序贯三支情感分类方法研究[J]. *计算机工程与应用*, 2021, **57**(23): 211–218. DOI: 10.3778/j.issn.1002-8331.2006-0324. WANG Q, LIU D. Sequential Three-way Sentiment Classification Combined with Ensemble Learning[J]. *Comput Eng Appl*, 2021, **57**(23): 211–218. DOI: 10.3778/j.issn.1002-8331.2006-0324.
- [17] CHEN J, CHEN Y, HE Y C, *et al.* A Classified Feature Representation Three-way Decision Model for Sentiment Analysis[J]. *Appl Intell*, 2022, **52**(7): 7995–8007. DOI: 10.1007/s10489-021-02809-1.
- [18] SU J Y, LIU W J, FENG D Y, *et al.* Social Media Sentiment Analysis of Sequential Three-way Decision Model Based on Knowledge Fusion[C]//2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI). New York: IEEE, 2023: 20–24. DOI: 10.1109/ICETCI57876.2023.10176764.
- [19] XIA R, ZONG C Q, LI S S. Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification[J]. *Inf Sci*, 2011, **181**(6): 1138–1152. DOI: 10.1016/j.ins.2010.11.023.
- [20] PAWLAK Z. Rough Sets[J]. *Int J Comput Inf Sci*, 1982, **11**(5): 341–356. DOI: 10.1007/bf01001956.
- [21] CHEN L, XU H Y, PEDRYCZ W. Conflict Analysis Based on a Novel Three-way Decisions Graph Model for Conflict Resolution Method Under Hesitant Fuzzy Environment[J]. *Inf Fusion*, 2023, **100**: 101936. DOI: 10.1016/j.inffus.2023.101936.
- [22] QIAN J, TANG D W, YU Y, *et al.* Hierarchical Sequential Three-way Decision Model[J]. *Int J Approx Reason*, 2022, **140**: 156–172. DOI: 10.1016/j.ijar.2021.10.004.
- [23] WANG J J, MA X L, XU Z S, *et al.* Three-way Multi-attribute Decision Making under Hesitant Fuzzy Environments[J]. *Inf Sci*, 2021, **552**: 328–351. DOI: 10.1016/j.ins.2020.12.005.
- [24] 钱进, 郑明晨, 周川鹏等. 多粒度三支决策研究进展[J]. *数据采集与处理*, 2024, **39**(2): 361–375. DOI: 10.16337/j.1004-9037.2024.02.009. QIAN J, ZHENG M C, ZHOU C P, *et al.* Recent Advancement in Multi-granulation Three-way Decisions[J]. *J Data Acquisition and Processing*, 2024, **39**(2): 361–375. DOI: 10.16337/j.1004-9037.2024.02.009.
- [25] BÜHLMANN P, YU B. Analyzing Bagging[J]. *Ann Statist*, 2002, **30**(4): 927–961. DOI: 10.1214/aos/1031689014.
- [26] SCHAPIRE R R. A Brief Introduction to Boosting[C]//Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 1999, **99**(999): 1401–1406.
- [27] PAVLYSHENKO B. Using Stacking Approaches for Machine Learning Models[C]//2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP). New York: IEEE, 2018: 255–258. DOI: 10.1109/DSMP.2018.8478522.
- [28] PARHAMI B. Voting Algorithms[J]. *IEEE Trans Reliab*, 1994, **43**(4): 617–629. DOI: 10.1109/24.370218.
- [29] ZHANG P F, LI T R, WANG G Q, *et al.* Multi-source Information Fusion Based on Rough Set Theory: a Review[J]. *Inf Fusion*, 2021, **68**: 85–117. DOI: 10.1016/j.inffus.2020.11.004.
- [30] HUANG X F, ZHAN J M, DING W P, *et al.* An Error Correction Prediction Model Based on Three-way Decision and Ensemble Learning[J]. *Int J Approx Reason*, 2022, **146**: 21–46. DOI: 10.1016/j.ijar.2022.04.002.
- [31] JIANG C M, LI Z C, YAO J T. A Shadowed Set-based Three-way Clustering Ensemble Approach[J]. *Int J Mach Learn Cybern*, 2022, **13**(9): 2545–2558. DOI: 10.1007/s13042-022-01543-5.
- [32] QIAN J, WANG D, YU Y, *et al.* E3WD: A Three-way Decision Model Based on Ensemble Learning[J]. *Inf Sci*, 2024, **667**: 120487. DOI: 10.1016/J.INS.2024.120487.
- [33] LANG Q, PAN X J, LIU X D. A Text-granulation Clustering Approach with Semantics for E-commerce Intelligent Storage Allocation[J]. *IEEE Access*, 2020, **8**: 164282–164291. DOI: 10.1109/ACCESS.2020.3021421.