3WC-GBNRS++:A novel three-way classifier with granular-ball neighborhood rough sets based on uncertainty (Supplementary materials)

Jie Yang, Zhuangzhuang Liu, Shuyin Xia, Guoyin Wang, *Senior Member, IEEE*, Qinghua Zhang, *Senior Member, IEEE*, Shuai Li, Taihua Xu

This is the supplementary file for the paper entitled 3WC-GBNRS++:A novel three-way classifier with granular-ball neighborhood rough sets based on uncertainty in IEEE Transactions on Fuzzy Systems.

S1 THE ORIGINAL GRANULAR-BALLS GENERATION METHOD

Algorithm S1: The original granular-balls generation method [1]

Input: Dataset $\mathbb{D}$, purity threshold p;
Output: The set of granular balls GB_list.
Initialize $GB\_list = \{\}$;
Function `Split_GB(`GB`)`:
  if *the purity of* $GB > p$ **then**
    Add GB to the GB_list;
  end
  else
    $GB_1, GB_2 \leftarrow GB$ is subdivided into two granular-balls by 2-means clustering algorithm;
    `Split_GB(`GB_1`)`;
    `Split_GB(`GB_2`)`;
  end
return
Function `Main(`$\mathbb{D}$`)`:
  Turn $\mathbb{D}$ to a granular ball GB;
  `Split_GB(`GB`)`;
  return GB_list;
return

In the process of generating granular-balls, Algorithm S1 introduces a purity threshold as a parameter to control the granularity. The purity threshold plays a crucial role in controlling the degree of granularity in the division of these granular-balls. The purity of a granular-ball can be determined by calculating the proportion of the majority label. If the purity of a granular-ball falls below the purity threshold, it will be divided again. Therefore, the higher the purity threshold, the higher the purity of the final generated granular-balls.

The process of threshold acquisition is shown in Algorithm S2. First, the quality index of each granular-ball in GBNRS++ is computed. Subsequently, the average fuzziness of $\mathbb{X}_{\mathbb{R}}^{\mathbb{J}}$ is determined. Next, based on the objective function, we calculate $H_{\psi(\mathbb{X}_{\mathbb{R}}^{\mathbb{J}})}$ and $\left|\bar{H}_{\psi(\mathbb{X}_{\mathbb{R}}^{\mathbb{J}})} - \bar{H}_{\mathbb{X}_{\mathbb{R}}^{\mathbb{J}}}\right|$ under varying thresholds by adjusting the step size. When the minimum value is achieved, the optimal thresholds are obtained. The main steps in Algorithm 2, from step 6 to step 14, involve searching for optimal thresholds (β^*, α^*) by exploring all possible thresholds.

Algorithm S2: The process of threshold acquisition

Input: $DBGB\_list$, step size L
Output: The decision thresholds α^* and β^*
1 **for** $GB \in DBGB\_list$ **do**
2   Compute the quality index $m(GB)$;
3 **end**
4 Computing $H_{\mathbb{X}_{\mathbb{R}}^{\mathbb{J}}} = \frac{1}{|U|}\sum_{GB\in DBGB\_list}\xi(GB)$, where $\xi(GB) = 4m(GB)(1-m(GB))$;
5 Assuming $\beta = 1-\alpha$ and $Min = H_{\mathbb{X}_{\mathbb{R}}^{\mathbb{J}}}$;
6 **for** $\beta = 0$ *to* 0.5 **do**
7   $\alpha = 1-\beta$;
8   Computing $H_{\psi(\mathbb{X}_{\mathbb{R}}^{\mathbb{J}})}$ and $\left|\bar{H}_{\psi(\mathbb{X}_{\mathbb{R}}^{\mathbb{J}})} - \bar{H}_{\mathbb{X}_{\mathbb{R}}^{\mathbb{J}}}\right|$;
9   **if** $\left|\bar{H}_{\psi(\mathbb{X}_{\mathbb{R}}^{\mathbb{J}})} - \bar{H}_{\mathbb{X}_{\mathbb{R}}^{\mathbb{J}}}\right| < Min$ **then**
10     $Min = \left|\bar{H}_{\psi(\mathbb{X}_{\mathbb{R}}^{\mathbb{J}})} - \bar{H}_{\mathbb{X}_{\mathbb{R}}^{\mathbb{J}}}\right|$;
11     $\beta^* = \beta$;
12     $\alpha^* = 1-\beta^*$;
13   **end**
14   $\beta = \beta + L$;
15 **end**
16 **return** α^* and β^*;

Algorithm S3 gives the details of choosing the turning point. To compute the set DIS, all GBs must be traversed, resulting in a time complexity of $O(N)$.To compute dV and $AllMaxDiff$ each takes $O(N)$ time since it involves traversing the elements of DIS. The time complexity of this loop over i can be considered as $O(N)$.

REFERENCES

[1] M. Ester, H. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD'96)*, 1996, p. 226–231.

Algorithm S3: Construction of adaptive granular-ball neighborhood

Input: $DBGB\_list$, parameters Local R and Global R, predicted object x_{test}.

Output: The $\delta_R^*(x_{\text{test}})$ of x_{test}.

Compute the set DIS $= \{\gamma_1, \gamma_2, \gamma_3, \ldots, \gamma_n\}$;
Len $= \text{length}(DIS)$;
$\boldsymbol{dV} = \{\gamma_i - \gamma_{i+1}\}_{i=1}^{\text{Len}-1}$;
AllMaxDiff $= \max(\boldsymbol{dV})$;
FitLen $= 20$;
for $i = Len - FitLen; i \geq 2; i = i - 1$ **do**
  $\boldsymbol{CV} = \{\gamma_j\}_{j=i+1}^{i+\text{FitLen}}$;
  $\boldsymbol{Index} = \{i+1, \ldots, i+\text{FitLen}\}$;
  $[a, b] = \text{LinearFit}(\boldsymbol{Index}, \boldsymbol{CV})$;
  PredictV $= a \times i + b$;
  CDiff $= \gamma_i -$ PredictV;
  $\boldsymbol{dCV} = \{CV_j - CV_{j+1}\}_{j=1}^{\text{FitLen}-1}$;
  CMAllMaxDiff $= \max(\boldsymbol{dCV})$;
  if $CDiff > LocalR \times CMAllMaxDiff$ ***and*** $CDiff > GlobalR \times AllMaxDiff$ **then**
    break;
  end
end
$\delta_R^*(x_{\text{test}}) = \{\gamma$ values less than $\gamma_i\}$ **return** $\delta_R^*(x_{\text{test}})$;

S2 THE BIOGRAPHIES OF AUTHORS

Jie Yang received the Ph.D. degree in computer science and technology from the Chongqing University of Posts and Telecommunications, Chongqing,China, in 2019. He is currently a Professor with Zunyi Normal University, Zunyi, China, and a masters supervisor with the Chongqing University of Posts and Telecommunications and Jiangsu University of Science and Technology. He has more than 50 publications, including IEEE TFS, Information Sciences, Knowledge-Based System, Cognitive Computation, etc. His research interests include data mining, machine learning, three-way decisions, and rough sets.

Zhuangzhuang Liu is currently pursuing a Master'sdegree in the School of Computer, at Jiangsu University of Science and Technology, Zhenjiang, China. His research interests include granular-ball computing and three-way decisions..

Shuyin Xia (Member, IEEE) received the Ph.D. degree from the College of Computer Science, Chongqing University, Chongqing, in 2015. He is currently a Professor with the College of Computer Science and Technology, Chongqing University of Posts and Telecommunications (CQUPT), Chongqing. He is also the Executive Deputy Director of the Big Data and Network Security Joint Laboratory, CQUPT. He has published more than 30 papers in journals and conferences, including IEEE TPAMI, IEEE TKDE, IEEE TNNLS, IEEE TCYB, and Information Sciences. His research interests include classifiers and granular computing.

Guoyin Wang (Senior Member, IEEE) received the B.S.,M.S., and Ph.D. degrees in computer science and technology from Xi'an Jiaotong University, Xi'an, China, in 1992, 1994, and 1996, respectively. He was a Visiting Scholar with the University of North Texas, Denton, TX, USA, and the University of Regina, Regina, SK, Canada, from 1998 to 1999. Since 1996, he has been with the Chongqing University of Posts and Telecommunications, Chongqing, China, where he is currently a Professor as well as the Vice President and the Director of the Chongqing Key Laboratory of Computational Intelligence. He was the Director of the Institute of Electronic Information Technology, Chongqing Institute of Green and Intelligent Technology, CAS, Chongqing, from 2011 to 2017. He has authored 12 books and he is the editor of dozens of proceedings of national and international conferences and has more than 200 reviewed research publications. His research interests include rough sets, granular computing, knowledge technology, data mining, neural networks, and cognitive computing.Dr. Wang is a Fellow of the International Rough Set Society, the Chinese Association for Artificial Intelligence, and the China Computer Federation.

Qinghua Zhang received an M.S. degree in computer application from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2003 and a Ph.D. degree in computer application technology from Southwest Jiaotong University, Chengdu, China, in 2010. He was at San Jose State University, USA, as a visiting scholar in 2015. Since 1998, he has been at Chongqing University of Posts and Telecommunications, where he is currently a professor and the dean of the School of Computer Science and Technology. His research interests include rough sets, fuzzy sets, granular computing and uncertain information processing.

Shuai Li received the Ph.D. degree in computer science and technology from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2022. He is currently an Associate Professor with Nanchang Hangkong University, Nanchang, China, and a masters supervisor with Nanchang Hangkong University. His research interests include machine learning, rough sets, cloud models and three-way decisions..

Taihua Xu received the Ph.D. degree in computer application technology from Southwest Jiaotong University, Chengdu, China, in 2019. He is currently an Associate Professor with Jiangsu University of Science and Technology, Zhenjiang, and a masters supervisor with the Jiangsu University of Science and Technology. His research interests include data mining, machine learning, three-way decisions, and rough sets.