Online Harmonizing Gradient Descent for Imbalanced Data Streams One-Pass Classification

Han Zhou[1], Hongpeng Yin[1*], Xuanhong Deng[1] & hyy5476@163.com[1]

[1]*The School of Automation, Chongqing University, Chongqing, China.*

The code is available at https://github.com/Kan9594/OHGD.

Appendix A Theoretical Justification

In this section, we provide the theoretical analysis that gives the sublinear regret bound achieved by OHGD. Recalling the following assumptions and corollary.

Assumption 1. Let $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_T, y_T)$ be the data stream $\mathcal{D}$, where $y_i \in \{-1, +1\}$ and $\|\mathbf{x}_i\| \leqslant 1$ for all t. The imbalance ratio is stationary $\rho = \frac{N_n^T}{N_p^T} \geqslant 1$.

Assumption 2. The decision domain $\mathcal{W}$ contains the origin $\mathbf{0}$, and its diameter is bounded by D, i.e.,

$$\max_{\mathbf{w_1}, \mathbf{w_2} \in \mathcal{W}} \|\mathbf{w_1} - \mathbf{w_2}\|_2 \leqslant D. \tag{A1}$$

Corollary 1. Under these assumptions, for the hinge loss $\mathcal{L}(\mathbf{w}_i) = \max(0, 1 - y_i \mathbf{w}_i \mathbf{x}_i)$, we have $0 \leqslant \mathcal{L}(\mathbf{w}_i) \leqslant (D+1)$.

Appendix A.1 The proof of Lemma.1.

We provide the following lemma that would facilitate the theoretical analysis. To ease our discussion, we denote $\mathcal{M}$ as the prediction error indexes set: $\mathcal{M} = \{t | \mathcal{L}(\mathbf{w}_t) > 0\}$. Similarly, we denote M_p^t and M_p^t are the number of misclassified positive and negative instances before t-th round, respectively. $M^T = |M| = M_p^T + M_n^T$ is the total number of misclassified instances in data stream $\mathcal{D}$.

Lemma 1. The sum of the re-weighting parameter α_t has an upper bound:

$$\sum_{t=1}^{T} \alpha_t \leqslant 2(D+1)(\rho M_p^T + M_n^T). \tag{A2}$$

Proof. Recalling the definition of α_t:

$$\alpha_t = 2 \frac{\rho_t [\sum_{i \in \mathbb{I}_{(-)}} R_i(\mathbf{x}_i, y_i)] \mathbb{I}_{(+)} + [\sum_{i \in \mathbb{I}_{(+)}} R_i(\mathbf{x}_i, y_i)] \mathbb{I}_{(-)}}{\sum_{i \in \mathbb{I}_{(-)}} R_i(\mathbf{x}_i, y_i) + \sum_{i \in \mathbb{I}_{(+)}} R_i(\mathbf{x}_i, y_i)} \mathcal{L}_t(\mathbf{w}_t | \mathbf{x}_t, y_t). \tag{A3}$$

When the truth is positive, we have:

$$\alpha_t = 2 \frac{\rho_t [\sum_{i \in \mathbb{I}_{(-)}} R_i(\mathbf{x}_i, y_i)] \mathbb{I}_{(+)}}{\sum_{i \in \mathbb{I}_{(-)}} R_i(\mathbf{x}_i, y_i) + \sum_{i \in \mathbb{I}_{(+)}} R_i(\mathbf{x}_i, y_i)} \mathcal{L}_t(\mathbf{w}_t | \mathbf{x}_t, y_t) \leqslant 2\rho_t(D+1). \tag{A4}$$

When the truth is negative, we have:

$$\alpha_t = 2 \frac{[\sum_{i \in \mathbb{I}_{(+)}} R_i(\mathbf{x}_i, y_i)] \mathbb{I}_{(-)}}{\sum_{i \in \mathbb{I}_{(-)}} R_i(\mathbf{x}_i, y_i) + \sum_{i \in \mathbb{I}_{(+)}} R_i(\mathbf{x}_i, y_i)} \mathcal{L}_t(\mathbf{w}_t | \mathbf{x}_t, y_t) \leqslant 2(D+1). \tag{A5}$$

* Corresponding author (email: yinhongpeng@cqu.edu.cn)

Summing α_t over T leads to an upper bound:

$$\sum_{t=1}^{T} \alpha_t \leqslant 2(D+1)[\sum_{t\in\mathbb{I}_{(y=+1)}}^{T} \rho + \sum_{i\in\mathbb{I}_{(y=-1)}}^{T} 1]. \tag{A6}$$

Recalling the definition of M_p^t and M_n^t, we thus have:

$$\sum_{t=1}^{T} \alpha_t \leqslant 2(D+1)(\rho M_p^T + M_n^T). \tag{A7}$$

This completes the proof. ■

Appendix A.2 The proof of Theorem.1.

Theorem 1. By dynamically setting $\lambda_t = \sqrt{\frac{D+1}{\rho M_n^t + M_p^t}}$, for any $\mathbf{w}^* \in \Re^d$, the following regret bound holds for the proposed OHGD on the data stream $\mathcal{D}$:

$$R(T) \leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(\frac{2\epsilon+1}{\epsilon}\sqrt{T} - 1). \tag{A8}$$

where the ϵ denotes the minimum of α_t.

Proof. Due to the convexity of the loss function, the following inequality holds for any $\mathbf{w}$:

$$\mathcal{L}_t(\mathbf{w}_t) - \mathcal{L}_t(\mathbf{w}^*) \leqslant \nabla\mathcal{L}_t(\mathbf{w}_t)(\mathbf{w}_t - \mathbf{w}^*). \tag{A9}$$

Since we defined the harmonized gradient descent updating $\mathbf{w}_{t+1} = \mathbf{w}_t - \alpha_t\lambda_t\nabla\mathcal{L}_t(\mathbf{w}_t)$ in Eq.**??**, then we have:

$$\|\mathbf{w}_{t+1} - \mathbf{w}^*\|^2 = \|\mathbf{w}_t - \mathbf{w}^*\|^2 + \alpha_t^2\lambda_t^2\|\nabla\mathcal{L}_t(\mathbf{w}_t)\|^2 - 2\alpha_t\lambda_t\nabla\mathcal{L}_t(\mathbf{w}_t)(\mathbf{w}_t - \mathbf{w}^*). \tag{A10}$$

Accordingly,

$$\nabla\mathcal{L}_t(\mathbf{w}_t)(\mathbf{w}_t - \mathbf{w}^*) = \frac{\|\mathbf{w}_t - \mathbf{w}^*\|^2 - \|\mathbf{w}_{t+1} - \mathbf{w}^*\|^2}{2\alpha_t\lambda_t} + \frac{\alpha_t\lambda_t}{2}\|\nabla\mathcal{L}_t\|^2. \tag{A11}$$

By summing, we can get:

$$\begin{aligned} R(T) &= \sum_{t=1}^{T}[\mathcal{L}_t(\mathbf{w}_t) - \mathcal{L}_t(\mathbf{w}^*)] \leqslant \sum_{t=1}^{T}\nabla\mathcal{L}_t(\mathbf{w}_t)(\mathbf{w}_t - \mathbf{w}^*) \\ &\leqslant \frac{\|\mathbf{w}_1 - \mathbf{w}^*\|^2 - \|\mathbf{w}_2 - \mathbf{w}^*\|^2}{2\alpha_1\lambda_1} + \cdots + \frac{\|\mathbf{w}_T - \mathbf{w}^*\|^2 - \|\mathbf{w}_{T+1} - \mathbf{w}^*\|^2}{2\alpha_T\lambda_T} + \sum_{t=1}^{T}\frac{\alpha_t\lambda_t}{2}\|\nabla\mathcal{L}_t\|^2. \end{aligned} \tag{A12}$$

Let ϵ denotes the minimum of α_t. Since we have $\|\nabla\mathcal{L}_t\| \leqslant 1$ and $\|\mathcal{L}_t\| \leqslant D+1$:

$$
\begin{aligned}
R(T) &\leqslant \frac{\|\mathbf{w}_1-\mathbf{w}^*\|^2-\|\mathbf{w}_{t+1}-\mathbf{w}^*\|^2}{2\epsilon\lambda_T}+\sum_{t=1}^{T}\frac{\alpha_t\lambda_t}{2}\|\nabla\mathcal{L}_t\|^2 \\
&\leqslant \frac{D^2}{2\epsilon\lambda_T}+\sum_{t=1}^{T}\frac{\alpha_t\lambda_t}{2}\|\nabla\mathcal{L}_t\|^2 \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho M_n^T+M_p^T}}{2\epsilon}+\frac{1}{2}(\sum_{t\in\mathbb{I}_{(y=-1)}}^{T}\frac{\rho M_n^t\|\mathcal{L}_t\|}{M_n^t+M_p^t}\cdot\sqrt{\frac{D+1}{\rho M_n^t+M_p^t}}+\sum_{t\in\mathbb{I}_{(y=+1)}}^{T}\frac{M_p^t\|\mathcal{L}_t\|}{M_n^t+M_p^t}\cdot\sqrt{\frac{D+1}{\rho M_n^t+M_p^t}}) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho M_n^T+M_p^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}}{2}(\sum_{t\in\mathbb{I}_{(y=-1)}}^{T}\frac{\rho M_n^t+M_p^t}{M_n^t+M_p^t}\cdot\frac{1}{\sqrt{\rho M_n^t+M_p^t}}+\sum_{t\in\mathbb{I}_{(y=+1)}}^{T}\frac{\rho M_n^t+M_p^t}{M_n^t+M_p^t}\cdot\frac{1}{\sqrt{\rho M_n^t+M_p^t}}) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho M_n^T+M_p^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}}{2}(\sum_{t\in\mathbb{I}_{(y=-1)}}^{T}\frac{\sqrt{\rho M_n^t+M_p^t}}{M_n^t+M_p^t}+\sum_{t\in\mathbb{I}_{(y=+1)}}^{T}\frac{\sqrt{\rho M_n^t+M_p^t}}{M_n^t+M_p^t}) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho M_n^T+M_p^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(\sum_{t\in\mathbb{I}_{(y=-1)}}^{T}\frac{1}{\sqrt{M_n^t+M_p^t}}+\sum_{t\in\mathbb{I}_{(y=+1)}}^{T}\frac{1}{\sqrt{M_n^t+M_p^t}}) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho M_n^T+M_p^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(\sum_{t\in\mathbb{I}_{(y=-1)}}^{T}\frac{1}{\sqrt{M^t}}+\sum_{t\in\mathbb{I}_{(y=+1)}}^{T}\frac{1}{\sqrt{M^t}}) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho M_n^T+M_p^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(\sum_{t=1}^{M^T}\frac{1}{\sqrt{M^t}}) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}\sqrt{M^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(1+\int_1^{M^T}\frac{1}{\sqrt{M^t}}dt) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}\sqrt{M^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(1+[2\sqrt{M^t}]_1^{M^T}) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}\sqrt{M^T}}{2\epsilon}+\frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(2\sqrt{M^T}-1) \\
&\leqslant \frac{(D+1)^{\frac{3}{2}}\sqrt{\rho}}{2}(\frac{2\epsilon+1}{\epsilon}\sqrt{T}-1).
\end{aligned}
\tag{A13}
$$

In this setting, OHGD achieves an $O(\sqrt{T})$ regret bound. This completes the proof. ■

Appendix A.3 The proof of Theorem.2.

Theorem 2. By dynamically setting $\lambda_t = 1/\sqrt{t}$, for any $\mathbf{w}^* \in \Re^d$, the following regret bound holds for the proposed OHGD on the data stream $\mathcal{D}$:

$$R(T) \leqslant (\frac{D^2}{2\epsilon}+\rho D+\rho)\sqrt{T}-\frac{\rho D+\rho}{2}, \tag{A14}$$

where the ϵ denotes the minimum of α_t.

Proof. Rewriting Eq.A13 as:

$$
\begin{aligned}
R(T) &\leqslant \frac{\|\mathbf{w}_1-\mathbf{w}^*\|^2-\|\mathbf{w}_{t+1}-\mathbf{w}^*\|^2}{2\epsilon\lambda_T}+\sum_{t=1}^{T}\frac{\alpha_t\lambda_t}{2}\|\nabla\mathcal{L}_t\|^2 \\
&\leqslant \frac{D^2}{2\epsilon\lambda_T}+\sum_{t=1}^{T}\frac{\alpha_t\lambda_t}{2}\|\nabla\mathcal{L}_t\|^2 \\
&\leqslant \frac{D^2\sqrt{T}}{2\epsilon}+\frac{1}{2}\Big(\sum_{t\in\mathbb{I}_{(y=-1)}}^{T}\frac{\rho M_n^t\|\mathcal{L}_t\|}{M_n^t+M_p^t}\cdot\frac{1}{\sqrt{t}}+\sum_{t\in\mathbb{I}_{(y=+1)}}^{T}\frac{M_p^t\|\mathcal{L}_t\|}{M_n^t+M_p^t}\cdot\frac{1}{\sqrt{t}}\Big) \\
&\leqslant \frac{D^2\sqrt{T}}{2\epsilon}+\frac{\rho(D+1)}{2}\Big(\sum_{t=1}^{M^T}\frac{1}{\sqrt{t}}\Big) \\
&\leqslant \frac{D^2\sqrt{T}}{2\epsilon}+\frac{\rho(D+1)}{2}(2\sqrt{M^T}-1) \\
&\leqslant \Big(\frac{D^2}{2\epsilon}+\rho D+\rho\Big)\sqrt{T}-\frac{\rho D+\rho}{2}.
\end{aligned}
\tag{A15}
$$

In this setting, OHGD achieves an $O(\sqrt{T})$ regret bound. This completes the proof. ∎

Appendix A.4 The proof of Theorem.3.

Theorem 3. If the learning rate λ_t is fixed as λ, for any $\mathbf{w}^*\in\Re^d$, the upper regret bound of OHGD is:

$$
R(T)\leqslant\frac{D^2}{2\epsilon\lambda}+(D+1)(\rho M_p^T+M_n^T)\lambda, \tag{A16}
$$

where the ϵ denotes the minimum of α_t.

Proof. When the learning rate is fixed, the defined harmonized gradient descent is $\mathbf{w}_{t+1}=\mathbf{w}_t-\alpha_t\lambda\nabla\mathcal{L}_t(\mathbf{w}_t)$. Then, we have:

$$
\begin{aligned}
\|\mathbf{w}_{t+1}-\mathbf{w}^*\|^2 &= \|\mathbf{w}_t-\alpha_t\lambda\nabla\mathcal{L}_t(\mathbf{w}_t)-\mathbf{w}^*\|^2 \\
&= \|\mathbf{w}_t-\mathbf{w}^*\|^2+\alpha_t^2\lambda^2\|\nabla\mathcal{L}_t(\mathbf{w}_t)\|^2 \\
&\quad -2\alpha_t\lambda\nabla\mathcal{L}_t(\mathbf{w}_t)(\mathbf{w}_t-\mathbf{w}^*).
\end{aligned}
\tag{A17}
$$

Accordingly,

$$
\nabla\mathcal{L}_t(\mathbf{w}_t)(\mathbf{w}_t-\mathbf{w}^*)=\frac{\|\mathbf{w}_t-\mathbf{w}^*\|^2-\|\mathbf{w}_{t+1}-\mathbf{w}^*\|^2}{2\alpha_t\lambda}+\frac{\alpha_t\lambda}{2}\|\nabla\mathcal{L}_t\|^2. \tag{A18}
$$

By summing over T and using Lemma 1, we can get:

$$
\begin{aligned}
R(T) &= \sum_{t=1}^{T}[\mathcal{L}_t(\mathbf{w}_t)-\mathcal{L}_t(\mathbf{w}^*)]\leqslant\sum_{t=1}^{T}\nabla\mathcal{L}_t(\mathbf{w}_t)(\mathbf{w}_t-\mathbf{w}^*) \\
&\leqslant \frac{\|\mathbf{w}_1-\mathbf{w}^*\|^2-\|\mathbf{w}_2-\mathbf{w}^*\|^2}{2\alpha_1\lambda}+\cdots+\frac{\|\mathbf{w}_T-\mathbf{w}^*\|^2-\|\mathbf{w}_{T+1}-\mathbf{w}^*\|^2}{2\alpha_T\lambda}+\sum_{t=1}^{T}\frac{\alpha_t\lambda}{2}\|\nabla\mathcal{L}_t\|^2 \\
&\leqslant \frac{\|\mathbf{w}_1-\mathbf{w}^*\|^2-\|\mathbf{w}_{t+1}-\mathbf{w}^*\|^2}{2\epsilon\lambda}+\frac{\lambda}{2}\sum_{t=1}^{T}\alpha_t \\
&\leqslant \frac{D^2}{2\epsilon\lambda}+(D+1)(\rho M_p^T+M_n^T)\lambda.
\end{aligned}
\tag{A19}
$$

By setting $\lambda=\sqrt{\frac{D+1}{\rho M_n^T+M_p^T}}$, we may achieve $\mathcal{O}(\sqrt{T})$ regret:

$$
R(T)\leqslant\frac{(D+1)^{\frac{3}{2}}}{2}\Big(\frac{2\epsilon+1}{\epsilon}\sqrt{\rho M_p^T+M_n^T}\Big). \tag{A20}
$$

This setting turns out that for appropriate choices of λ, the OHGD achieves sub-linear regret bound with the time horizon for the hinge loss. Proof completes. ∎

Table B1 The descriptions of one-pass methods and their parameter settings.

<table>
<tr><th>Methods</th><th>Strategy</th><th>Classifier Type</th><th>Approaches for Imblance</th><th>Parameter</th></tr>
<tr><td rowspan="2">CSOGD</td><td>Cost_I
Cost_{II}</td><td rowspan="2">Single</td><td rowspan="2">Cost-Sensitive</td><td>$c_p = 0.95, c_n = 0.05$</td></tr>
<tr><td>Sum_I
Sum_{II}</td><td>$n_p = 0.5, n_n = 0.5$</td></tr>
<tr><td rowspan="2">onlineUnderBagging
onlineWeightedUnderBagging
onlineOverBagging
onlineWeightedOverBagging</td><td rowspan="2">–</td><td rowspan="2">Ensemble</td><td>Under-Sampling</td><td rowspan="2">$M = 10, \eta = 0.9$</td></tr>
<tr><td>Over-Sampling</td></tr>
<tr><td>onlineAdaC2
onlineCBS2</td><td>–</td><td>Ensemble</td><td>Cost-Sensitive Hybrid-Resampling</td><td>$M = 10, c_p = 1, c_n = 0.8$</td></tr>
<tr><td rowspan="2">onlineRUSBoost

onlineUnderOverBagging</td><td>1
2</td><td rowspan="2">Ensemble</td><td>Cost-Sensitive Hybrid-Resampling</td><td>$M = 10, R = 0.7$</td></tr>
<tr><td>3
–</td><td>Hybrid-Resampling</td><td>$M = 10, R = IR$</td></tr>
<tr><td>onlineEffectiveBagging</td><td>–</td><td>Ensemble</td><td>Cost-Sensitive Hybrid-Resampling</td><td>$M = 10, R = 1/IR$</td></tr>
<tr><td>OHGD</td><td>–</td><td>Single</td><td>Gradient Harmonizing</td><td>–</td></tr>
</table>

Table B2 The descriptions of chunk based methods and their parameter settings.

<table>
<tr><th>Methods</th><th>Strategy</th><th>Classifier Type</th><th>Approaches for Imbalance</th><th>Parameter</th></tr>
<tr><td>HDWE</td><td>-</td><td>Ensemble</td><td>Hellinger Distance</td><td>$M = 10$</td></tr>
<tr><td>KNORAE2</td><td>NON
ROS</td><td>Ensemble</td><td>Over-Sampling & Classifier Selection</td><td>$L = 5,\ M = 10$</td></tr>
</table>

Appendix B Experimental Analysis

This section provides some details of experimental settings and results.

Appendix B.1 Settings

Twenty-four datasets from UCI repository and KEEL with different imbalance ratio were selected as the test rigs for performance evaluation. We compared our OHGD algorithms with various online learning algorithms for imbalanced data streams, including cost-sensitive online learning [1] (CSOGD), cost-sensitive online ensemble learning [2, 3] (onlineAdaC2, onlineCSB2, onlineUnderOverBagging, onlineRUSBoost, onlineEffectiveBagging) and resampling based online ensemble learning [4, 5] (onlineUnderBagging, onlineOverBagging, onlineWeightedUnderBagging, onlineWeightedOverBagging). Table.B1 summarizes the characteristics of different methods. Their parameters are set as the original works suggested. Note that, our method does not require any parameter settings for imbalance learning. As the purpose of the experiments is to make fair comparisons among different online algorithms, we chose OGD as the base learners of ensemble learning and the number of base learners M was set as 10. The learning rate η_t was set as $0.3/\sqrt{t}$.

Although our method focuses on one-pass data stream learning, we still compared it with some chunk based methods, including HDWE [6] and KNORAE2 [7]. The number of base learners M was set as 10 and other parameters were set as the original works suggested.

Appendix B.2 Parameter Sensitivity

Fig.C1-C24 illustrate the performance variation of algorithms under different parameter settings. Since none hyper-parameters is required in OHGD, its performance curves are horizontal lines. From these figures, it can be observed that OHGD is elegantly formulated without many hyper-parameters to tune, benefiting the easy implementation in practical applications.

Appendix B.3 Performance Comparison

Applying all algorithms for classifying all class imbalanced datasets in terms of AUC, G-mean and F1score, We reported the comparison results in Table.C1-C3. Summarizing the average ranks over all three performance metrics, we found that OHGD significantly outperforms other competitors in terms of each performance measure. Similar observations also can be found in the comparisons with the chunk-based methods (Table.C4-C6).

Appendix B.4 Regret Analysis

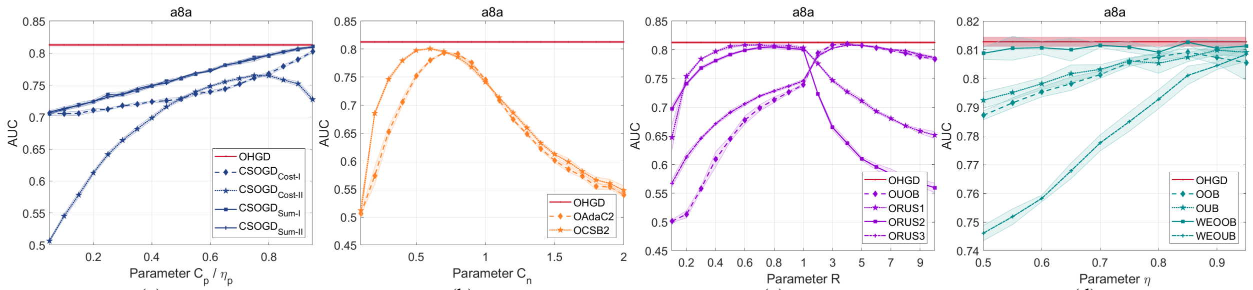
We proceed to analyze the theoretical performance of OHGD in this sub-section. Particularly, we present the *Average Accumulative Loss* to show the regret trends:

$$AvgLoss_t = \frac{\sum_{i=1}^{t} \mathcal{L}_t}{t} \tag{B1}$$

Figure.C25 illustrate the *AvgLoss* trends of OHGD on all the datasets. We observe that all the curves decrease rapidly and generally converge to a constant which show the agreements with our $O(\sqrt{T})$ regret convergence expectations in the Theorem.2.

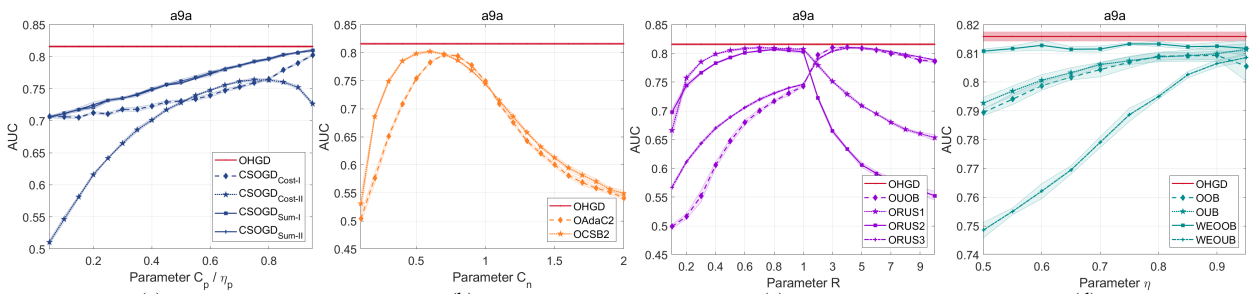
Appendix C *

References
1 Jialei Wang, Peilin Zhao, and Steven CH Hoi. Cost-sensitive online classification. *IEEE Transactions on Knowledge and Data Engineering*, 26(10):2425–2438, 2013.
2 Boyu Wang and Joelle Pineau. Online bagging and boosting for imbalanced data streams. *IEEE Transactions on Knowledge and Data Engineering*, 28(12):3353–3366, 2016.
3 Hongle Du, Yan Zhang, Ke Gang, Lin Zhang, and Yeh-Cheng Chen. Online ensemble learning algorithm for imbalanced data stream. *Applied Soft Computing*, 107:107378, 2021.
4 Shuo Wang, Leandro L Minku, and Xin Yao. A learning framework for online class imbalance learning. In *2013 IEEE Symposium on Computational Intelligence and Ensemble Learning (CIEL)*, pages 36–45. IEEE, 2013.
5 Shuo Wang, Leandro L Minku, and Xin Yao. Resampling-based ensemble methods for online class imbalance learning. *IEEE Transactions on Knowledge and Data Engineering*, 27(5):1356–1368, 2014.
6 Joanna Grzyb, Jakub Klikowski, and Michał Woźniak. Hellinger distance weighted ensemble for imbalanced data stream classification. *Journal of Computational Science*, 51:101314, 2021.
7 Paweł Zyblewski, Robert Sabourin, and Michał Woźniak. Preprocessed dynamic classifier ensemble selection for highly imbalanced drifted data streams. *Information Fusion*, 66:138–154, 2021.



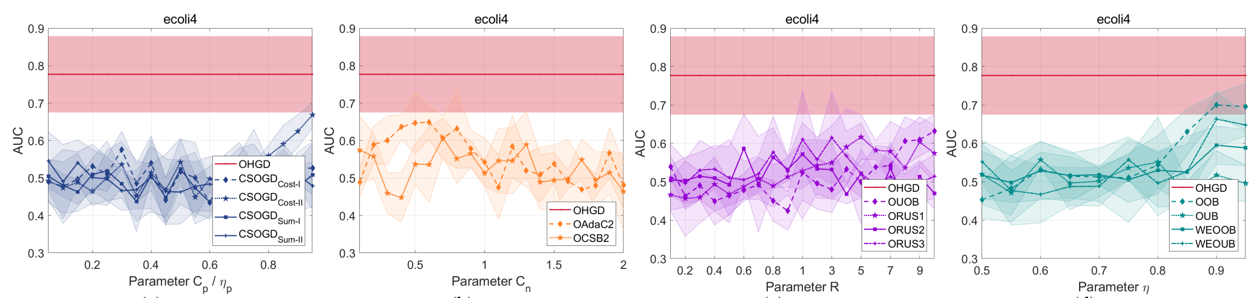
(a) (b) (c) (d)

Figure C1 Performance evaluation with varying parameters on dataset a8a, in terms of AUC.



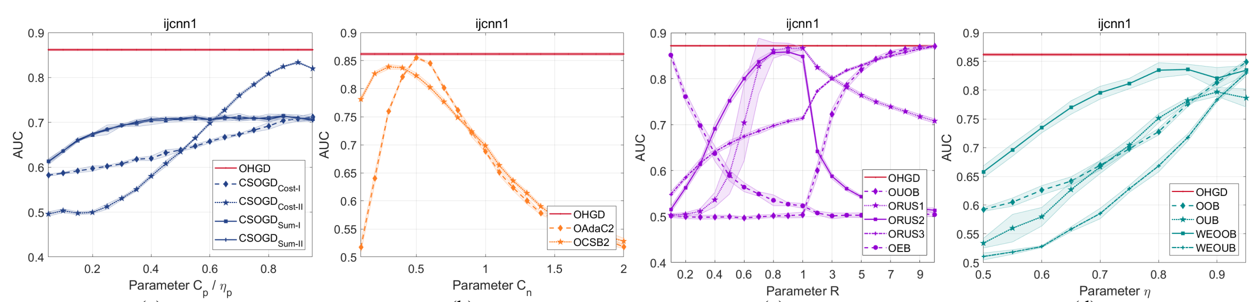
(a) (b) (c) (d)

Figure C2 Performance evaluation with varying parameters on dataset a9a, in terms of AUC.



(a) (b) (c) (d)

Figure C3 Performance evaluation with varying parameters on dataset ecoli4, in terms of AUC.



(a) (b) (c) (d)

Figure C4 Performance evaluation with varying parameters on dataset ijcnn1, in terms of AUC.

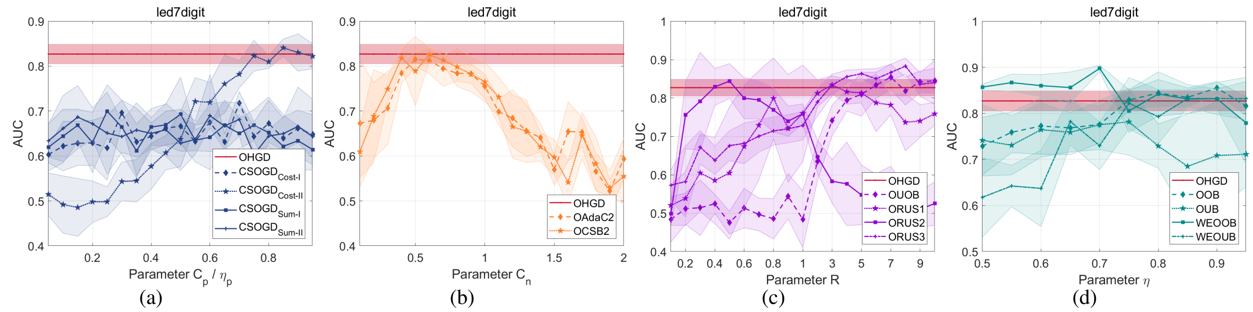


Figure C5 Performance evaluation with varying parameters on dataset led7digit, in terms of AUC.

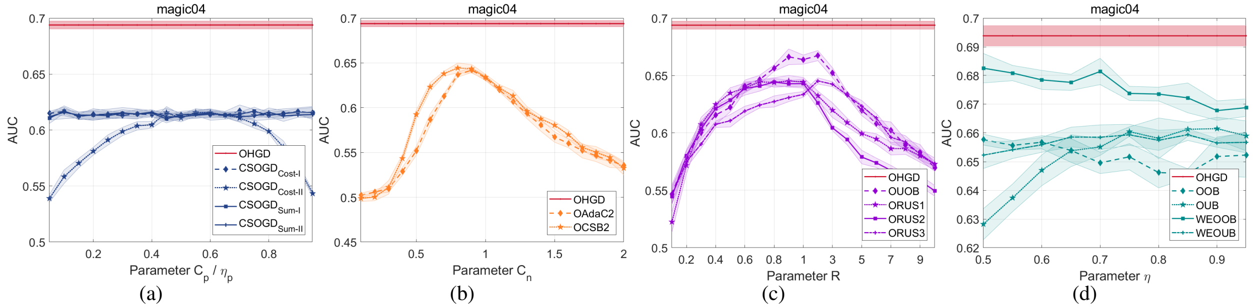


Figure C6 Performance evaluation with varying parameters on dataset magic04, in terms of AUC.

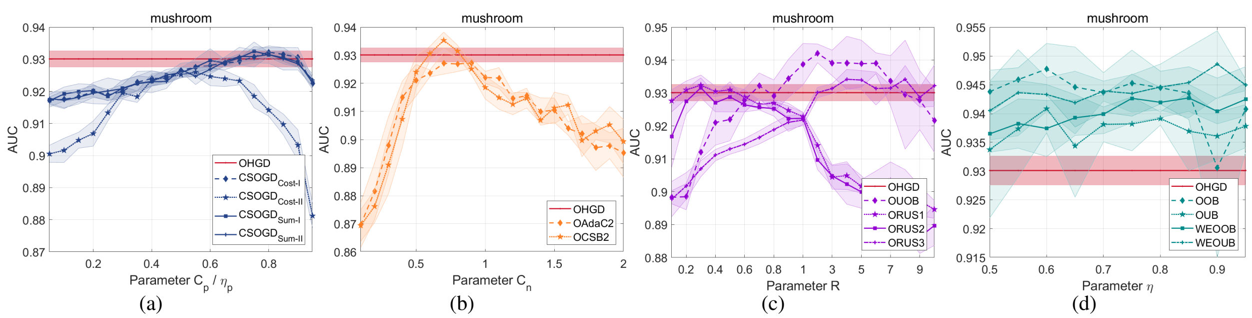


Figure C7 Performance evaluation with varying parameters on dataset mushroom, in terms of AUC.

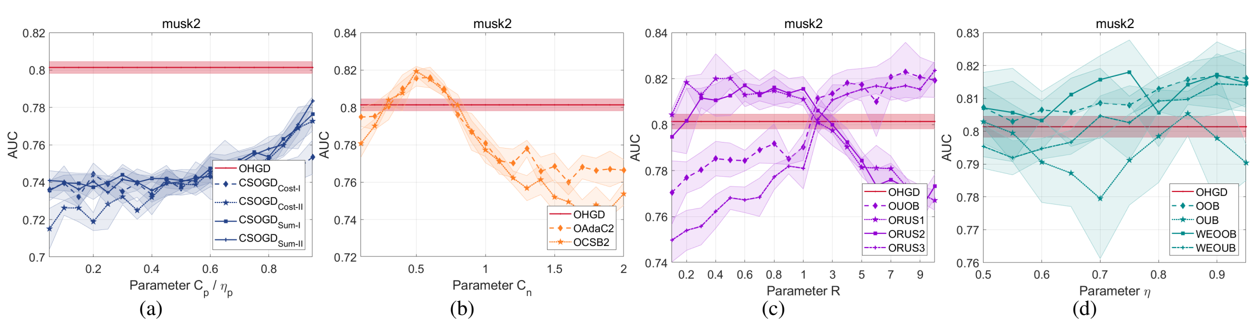


Figure C8 Performance evaluation with varying parameters on dataset musk2, in terms of AUC.

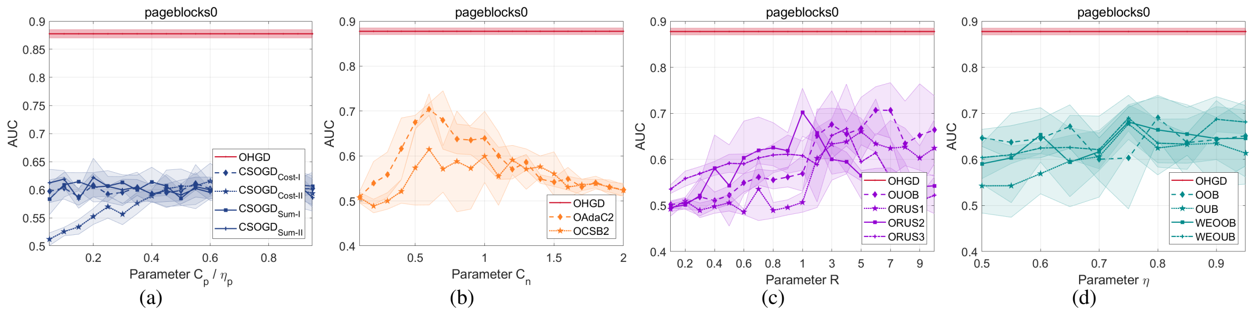


Figure C9 Performance evaluation with varying parameters on dataset pageblocks0, in terms of AUC.

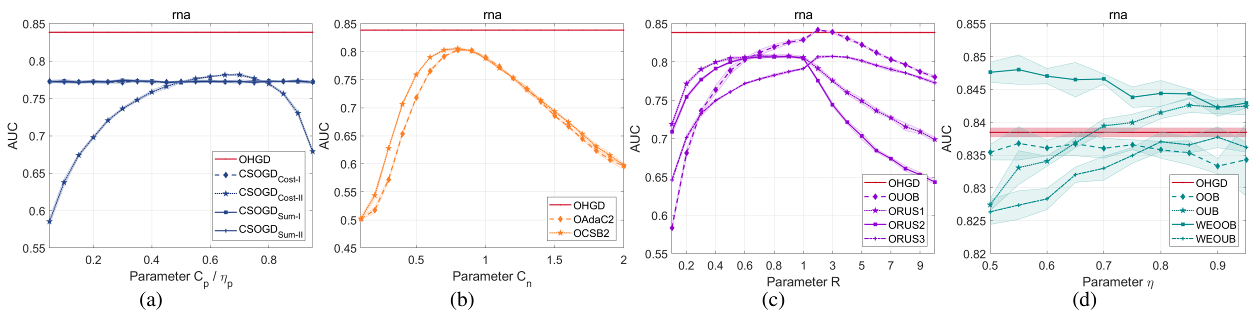


Figure C10 Performance evaluation with varying parameters on dataset rna, in terms of AUC.

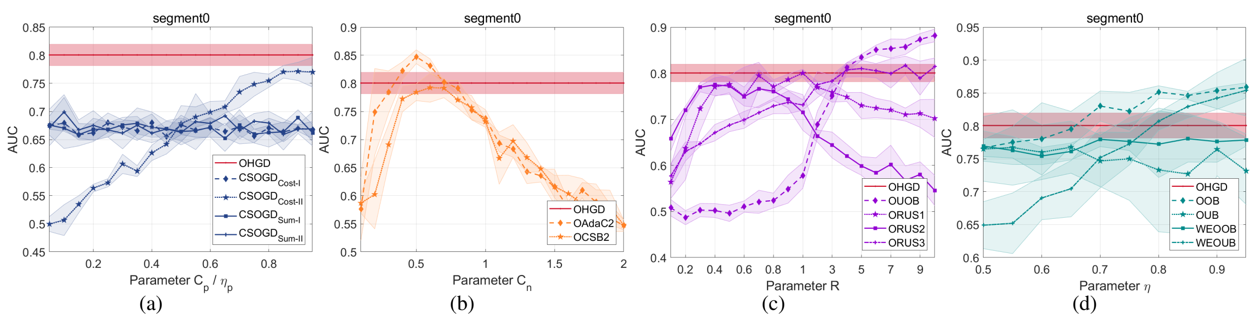


Figure C11 Performance evaluation with varying parameters on dataset segment0, in terms of AUC.

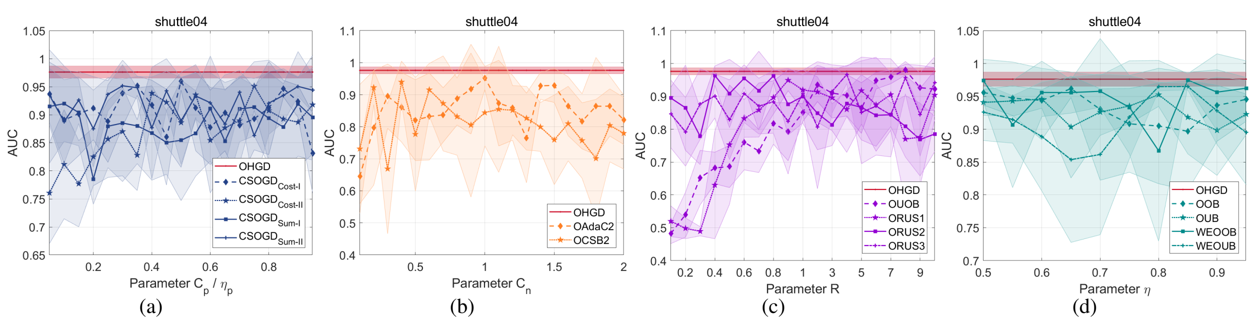


Figure C12 Performance evaluation with varying parameters on dataset shuttle04, in terms of AUC.

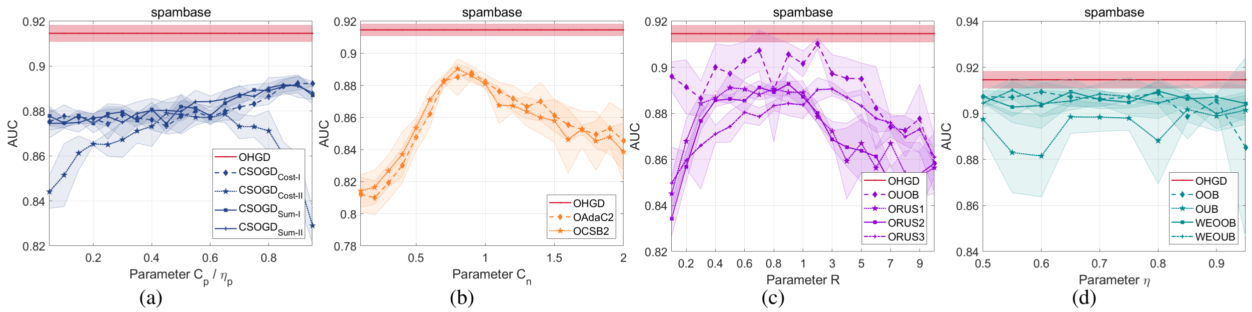


Figure C13 Performance evaluation with varying parameters on dataset spambase, in terms of AUC.

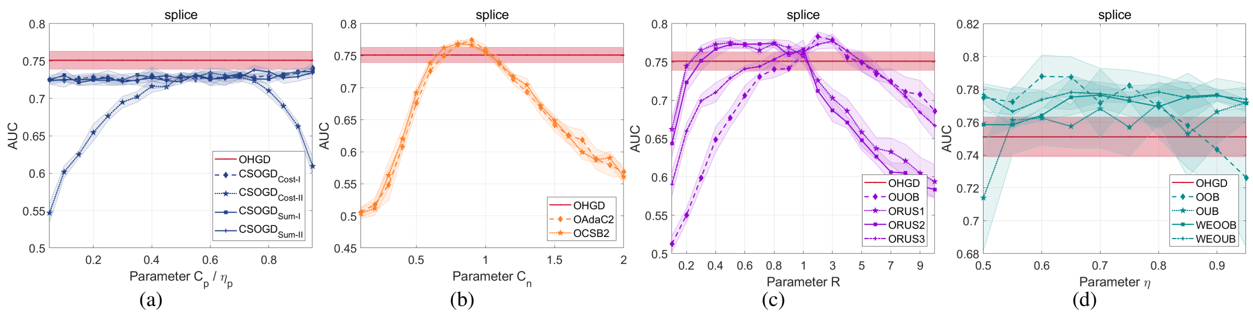


Figure C14 Performance evaluation with varying parameters on dataset splice, in terms of AUC.

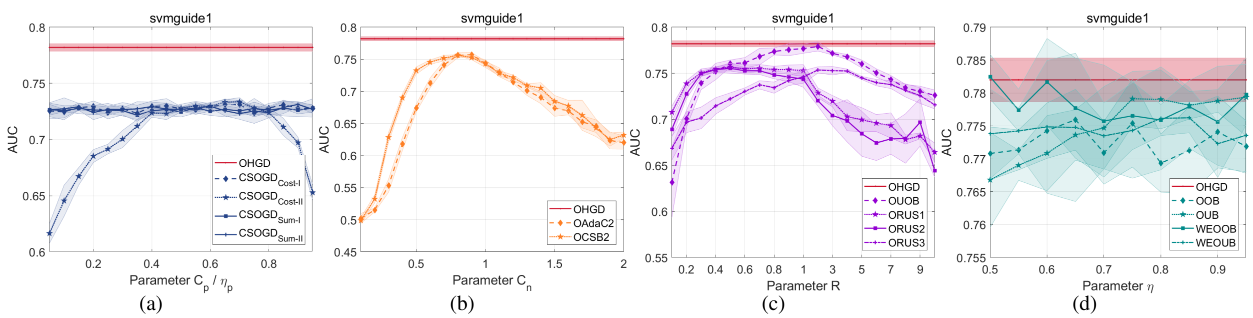


Figure C15 Performance evaluation with varying parameters on dataset svmguide1, in terms of AUC.

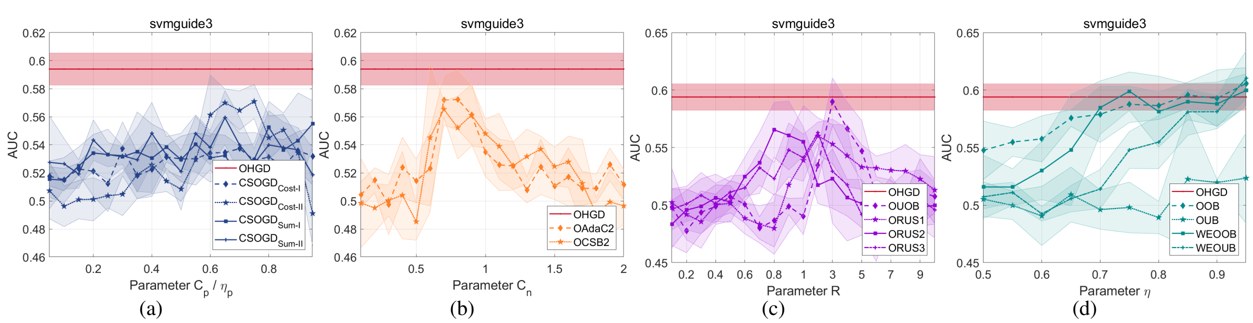
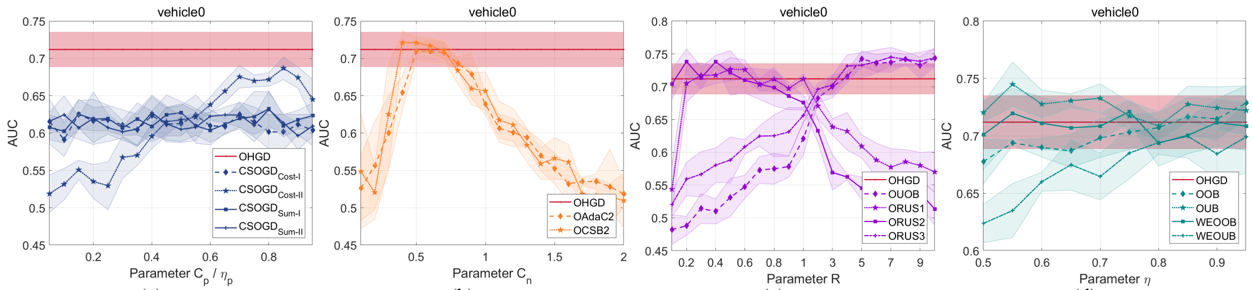
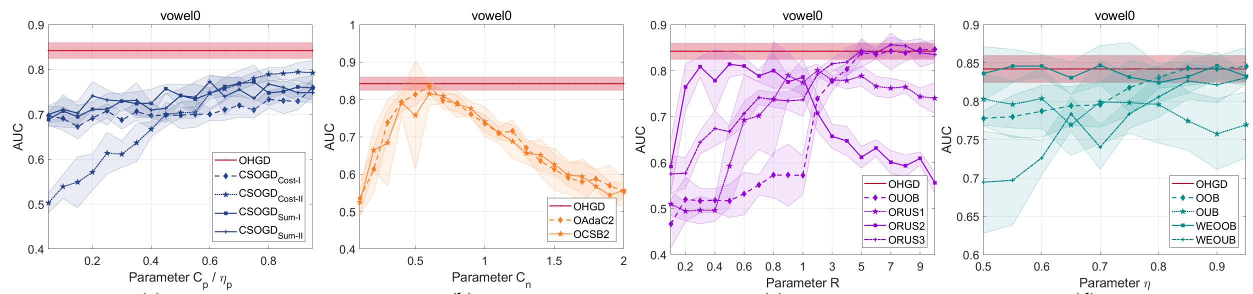


Figure C16 Performance evaluation with varying parameters on dataset svmguide3, in terms of AUC.



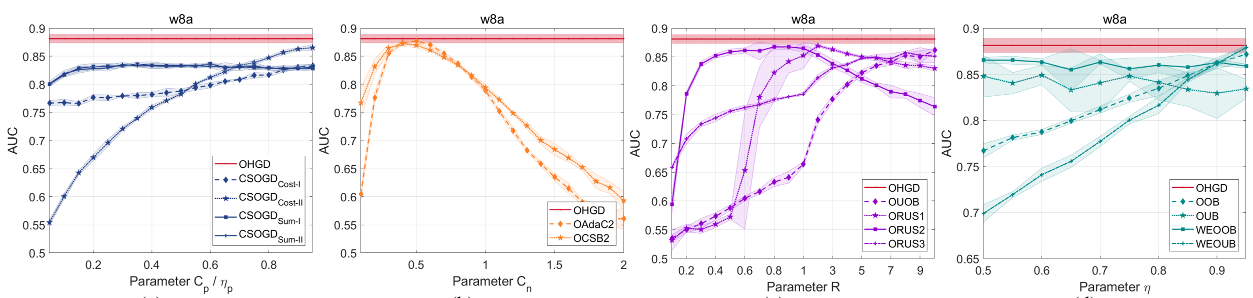
(a) (b) (c) (d)

Figure C17 Performance evaluation with varying parameters on dataset vehicle0, in terms of AUC.



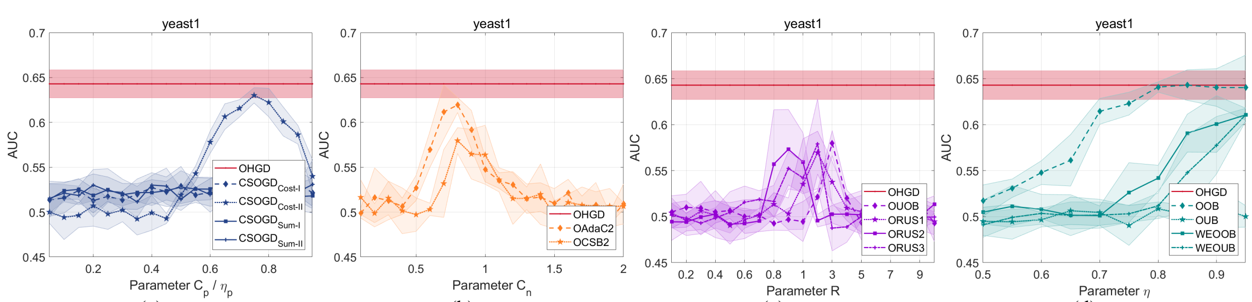
(a) (b) (c) (d)

Figure C18 Performance evaluation with varying parameters on dataset vowel0, in terms of AUC.



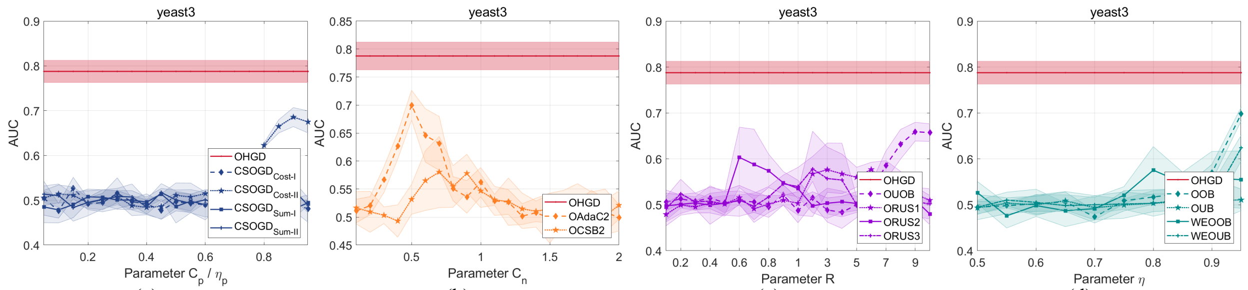
(a) (b) (c) (d)

Figure C19 Performance evaluation with varying parameters on dataset w8a, in terms of AUC.



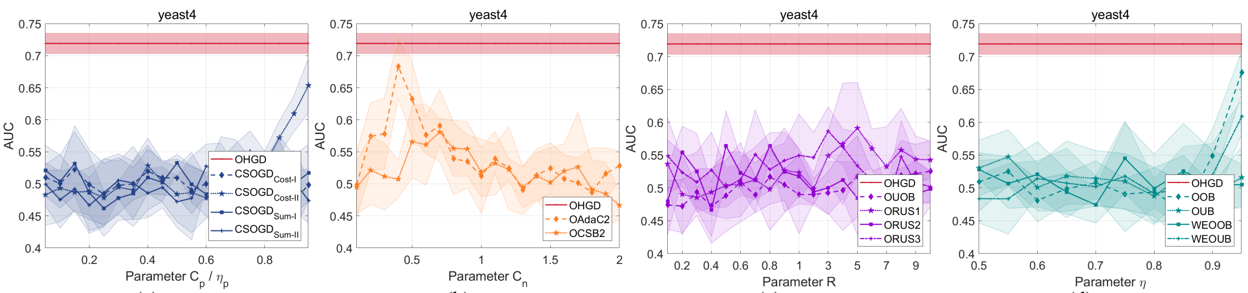
(a) (b) (c) (d)

Figure C20 Performance evaluation with varying parameters on dataset yeast1, in terms of AUC.



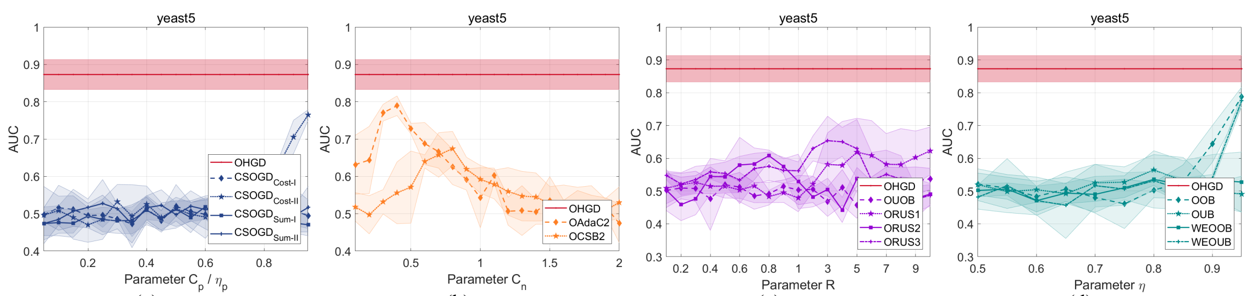
(a) (b) (c) (d)

Figure C21 Performance evaluation with varying parameters on dataset yeast3, in terms of AUC.



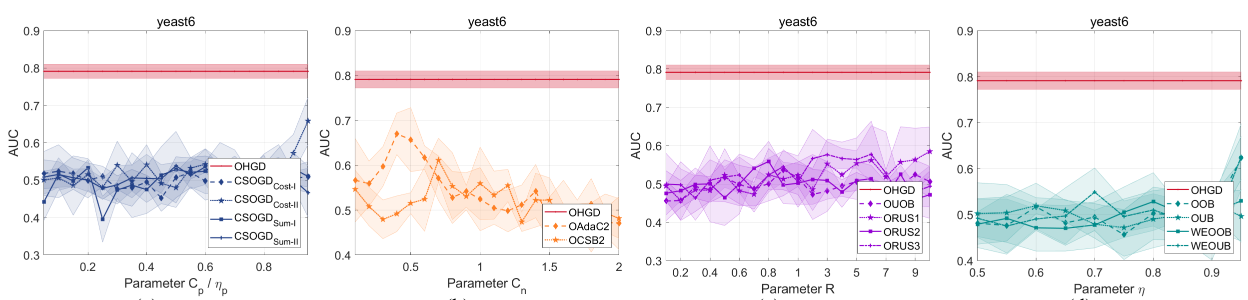
(a) (b) (c) (d)

Figure C22 Performance evaluation with varying parameters on dataset yeast4, in terms of AUC.



(a) (b) (c) (d)

Figure C23 Performance evaluation with varying parameters on dataset yeast5, in terms of AUC.



(a) (b) (c) (d)

Figure C24 Performance evaluation with varying parameters on dataset yeast6, in terms of AUC.

Table C1 The results of different methods in terms of AUC.

AUC	**OHGD**	**CSOGD**				**AdaC2**	**CSB2**	**UOB**	**RUSB**			**OOB**	**OUB**	**WOOB**	**WOUB**	**OEB**
		Cost$_I$	**Cost**$_{II}$	**Sum**$_I$	**Sum**$_{II}$				**1**	**2**	**3**					
a8a	0.816 ±0.003	0.791 ±0.002	0.737 ±0.002	0.782 ±0.002	0.805 ±0.002	0.802 ±0.002	0.814 ±0.002	0.812 ±0.003	0.814 ±0.005	0.805 ±0.001	0.815 ±0.002	0.814 ±0.002	0.798 ±0.003	0.808 ±0.002	0.81 ±0.003	0.807 ±0.004
a9a	0.815 ±0.002	0.797 ±0.001	0.735 ±0.002	0.786 ±0.002	0.804 ±0.001	0.803 ±0.002	0.815 ±0.001	0.814 ±0.002	0.815 ±0.002	0.806 ±0.001	0.816 ±0.002	0.815 ±0.001	0.801 ±0.002	0.81 ±0.003	0.811 ±0.001	0.809 ±0.002
ecoli4	0.791 ±0.094	0.505 ±0.059	0.612 ±0.062	0.498 ±0.056	0.636 ±0.047	0.543 ±0.084	0.568 ±0.042	0.667 ±0.036	0.498 ±0.056	0.568 ±0.099	0.482 ±0.056	0.544 ±0.035	0.471 ±0.042	0.479 ±0.063	0.573 ±0.081	0.536 ±0.006
ijcnn1	0.871 ±0.003	0.752 ±0.005	0.841 ±0.002	0.75 ±0.005	0.851 ±0.002	0.859 ±0.001	0.825 ±0.003	0.869 ±0.001	0.787 ±0.069	0.836 ±0.007	0.864 ±0.003	0.853 ±0.001	0.84 ±0.006	0.837 ±0.004	0.858 ±0.006	0.851 ±0.007
led7digit	0.834 ±0.038	0.533 ±0.046	0.83 ±0.02	0.548 ±0.058	0.841 ±0.034	0.806 ±0.036	0.816 ±0.016	0.838 ±0.029	0.5 ±0.064	0.802 ±0.031	0.695 ±0.129	0.85 ±0.029	0.685 ±0.098	0.822 ±0.029	0.856 ±0.036	0.732 ±0.088
magic	0.757 ±0.002	0.716 ±0.003	0.566 ±0.007	0.746 ±0.003	0.747 ±0.003	0.635 ±0.006	0.739 ±0.01	0.746 ±0.004	0.625 ±0.098	0.734 ±0.003	0.755 ±0.002	0.758 ±0.002	0.742 ±0.003	0.75 ±0.003	0.749 ±0.002	0.744 ±0.003
mushroom	0.931 ±0.002	0.889 ±0.002	0.888 ±0.005	0.937 ±0.002	0.937 ±0.002	0.923 ±0.006	0.919 ±0.003	0.899 ±0.007	0.93 ±0.002	0.928 ±0.002	0.931 ±0.003	0.929 ±0.004	0.918 ±0.011	0.921 ±0.007	0.903 ±0.005	0.911 ±0.006
musk2	0.798 ±0.007	0.769 ±0.006	0.796 ±0.007	0.784 ±0.008	0.813 ±0.004	0.823 ±0.006	0.78 ±0.005	0.783 ±0.006	0.75 ±0.012	0.758 ±0.007	0.769 ±0.009	0.794 ±0.005	0.709 ±0.012	0.791 ±0.008	0.711 ±0.009	0.709 ±0.008
pageblocks0	0.883 ±0.009	0.707 ±0.01	0.849 ±0.007	0.716 ±0.007	0.869 ±0.007	0.864 ±0.009	0.843 ±0.007	0.85 ±0.009	0.573 ±0.152	0.823 ±0.017	0.503 ±0.008	0.851 ±0.011	0.731 ±0.038	0.829 ±0.009	0.74 ±0.046	0.714 ±0.043
rna	0.91 ±0.003	0.816 ±0.001	0.866 ±0.002	0.907 ±0.001	0.922 ±0.001	0.902 ±0.002	0.504 ±0.003	0.822 ±0.004	0.501 ±0.002	0.879 ±0.003	0.862 ±0.01	0.897 ±0.002	0.692 ±0.011	0.84 ±0.003	0.81 ±0.003	0.800 ±0.006
segment0	0.97 ±0.011	0.752 ±0.009	0.939 ±0.008	0.752 ±0.013	0.962 ±0.012	0.947 ±0.012	0.918 ±0.012	0.927 ±0.015	0.554 ±0.061	0.835 ±0.056	0.869 ±0.02	0.958 ±0.008	0.791 ±0.032	0.945 ±0.005	0.845 ±0.011	0.834 ±0.049
shuttle04	0.982 ±0.007	0.877 ±0.011	0.982 ±0.005	0.874 ±0.018	0.971 ±0.008	0.967 ±0.005	0.965 ±0.012	0.98 ±0.006	0.693 ±0.22	0.94 ±0.035	0.904 ±0.036	0.976 ±0.007	0.873 ±0.043	0.968 ±0.01	0.932 ±0.028	0.886 ±0.048
spambase	0.911 ±0.004	0.892 ±0.003	0.825 ±0.006	0.915 ±0.003	0.915 ±0.004	0.879 ±0.006	0.895 ±0.005	0.906 ±0.003	0.91 ±0.003	0.905 ±0.003	0.911 ±0.003	0.906 ±0.009	0.899 ±0.006	0.911 ±0.004	0.906 ±0.005	0.905 ±0.004
splice	0.769 ±0.01	0.577 ±0.012	0.554 ±0.01	0.735 ±0.014	0.752 ±0.015	0.573 ±0.017	0.498 ±0.012	0.625 ±0.028	0.496 ±0.008	0.502 ±0.013	0.756 ±0.014	0.516 ±0.013	0.504 ±0.008	0.716 ±0.05	0.664 ±0.041	0.689 ±0.028
svmguide1	0.799 ±0.004	0.717 ±0.003	0.581 ±0.015	0.789 ±0.003	0.787 ±0.003	0.699 ±0.005	0.756 ±0.004	0.779 ±0.005	0.782 ±0.002	0.781 ±0.003	0.79 ±0.003	0.776 ±0.005	0.757 ±0.004	0.795 ±0.003	0.786 ±0.004	0.789 ±0.005
svmguide3	0.632 ±0.017	0.53 ±0.018	0.524 ±0.016	0.532 ±0.025	0.616 ±0.016	0.552 ±0.017	0.568 ±0.059	0.605 ±0.01	0.495 ±0.021	0.541 ±0.04	0.508 ±0.041	0.619 ±0.019	0.558 ±0.027	0.585 ±0.031	0.6 ±0.02	0.585 ±0.021
vehicle0	0.718 ±0.026	0.501 ±0.024	0.587 ±0.015	0.504 ±0.019	0.666 ±0.026	0.638 ±0.037	0.505 ±0.042	0.608 ±0.024	0.517 ±0.029	0.487 ±0.034	0.501 ±0.017	0.689 ±0.026	0.499 ±0.021	0.517 ±0.024	0.544 ±0.041	0.540 ±0.026
vowel0	0.863 ±0.02	0.56 ±0.034	0.839 ±0.011	0.565 ±0.028	0.835 ±0.024	0.834 ±0.017	0.806 ±0.027	0.84 ±0.019	0.674 ±0.122	0.805 ±0.021	0.805 ±0.052	0.837 ±0.013	0.667 ±0.038	0.822 ±0.03	0.823 ±0.016	0.518 ±0.062
w8a	0.901 ±0.003	0.78 ±0.005	0.894 ±0.002	0.779 ±0.004	0.891 ±0.001	0.89 ±0.003	0.87 ±0.002	0.899 ±0.002	0.709 ±0.164	0.891 ±0.001	0.894 ±0.004	0.873 ±0.005	0.849 ±0.011	0.867 ±0.005	0.888 ±0.004	0.873 ±0.012
yeast1	0.662 ±0.018	0.525 ±0.017	0.543 ±0.01	0.526 ±0.011	0.636 ±0.014	0.568 ±0.016	0.49 ±0.012	0.619 ±0.016	0.501 ±0.019	0.499 ±0.014	0.495 ±0.016	0.65 ±0.015	0.498 ±0.012	0.568 ±0.02	0.567 ±0.039	0.571 ±0.034
yeast3	0.743 ±0.024	0.513 ±0.027	0.657 ±0.012	0.498 ±0.023	0.683 ±0.022	0.646 ±0.058	0.527 ±0.054	0.614 ±0.028	0.484 ±0.036	0.569 ±0.056	0.512 ±0.028	0.522 ±0.02	0.524 ±0.021	0.493 ±0.016	0.516 ±0.039	0.515 ±0.037
yeast5	0.715 ±0.031	0.51 ±0.026	0.633 ±0.016	0.51 ±0.037	0.622 ±0.047	0.565 ±0.055	0.58 ±0.03	0.65 ±0.038	0.486 ±0.044	0.537 ±0.03	0.49 ±0.054	0.487 ±0.055	0.5 ±0.035	0.499 ±0.041	0.568 ±0.066	0.492 ±0.059
yeast6	0.856 ±0.048	0.502 ±0.036	0.743 ±0.038	0.512 ±0.056	0.753 ±0.04	0.654 ±0.048	0.621 ±0.092	0.785 ±0.048	0.509 ±0.041	0.546 ±0.051	0.524 ±0.044	0.511 ±0.046	0.493 ±0.03	0.499 ±0.041	0.593 ±0.056	0.527 ±0.034
yest4	0.779 ±0.028	0.495 ±0.04	0.644 ±0.036	0.51 ±0.056	0.648 ±0.053	0.558 ±0.082	0.56 ±0.039	0.664 ±0.027	0.487 ±0.042	0.54 ±0.056	0.5 ±0.079	0.502 ±0.041	0.49 ±0.057	0.507 ±0.048	0.557 ±0.03	0.505 ±0.039

Table C2 The results of different methods in terms of GMEANS.

GMEANS	OHGD	CSOGD				AdaC2	CSB2	UOB	RUSB			OOB	OUB	WOOB	WOUB	OEB
		Cost$_I$	**Cost$_{II}$**	**Sum$_I$**	**Sum$_{II}$**				**1**	**2**	**3**					
a8a	0.814 ±0.002	0.801 ±0.001	0.700 ±0.003	0.786 ±0.002	0.804 ±0.002	0.789 ±0.001	0.813 ±0.002	0.808 ±0.002	0.777 ±0.025	0.786 ±0.003	0.811 ±0.004	0.809 ±0.002	0.782 ±0.003	0.810 ±0.001	0.800 ±0.003	0.804 ±0.002
a9a	0.815 ±0.001	0.803 ±0.001	0.699 ±0.002	0.787 ±0.002	0.803 ±0.001	0.789 ±0.002	0.814 ±0.001	0.810 ±0.001	0.794 ±0.018	0.790 ±0.002	0.811 ±0.002	0.811 ±0.001	0.786 ±0.004	0.811 ±0.001	0.802 ±0.002	0.805 ±0.002
ecoli4	0.797 ±0.067	0.032 ±0.100	0.601 ±0.039	0.022 ±0.071	0.632 ±0.045	0.510 ±0.089	0.531 ±0.063	0.652 ±0.023	0.042 ±0.131	0.413 ±0.226	–	0.386 ±0.084	–	0.151 ±0.135	0.320 ±0.264	–
ijcnn1	0.870 ±0.002	0.773 ±0.004	0.835 ±0.001	0.766 ±0.003	0.850 ±0.002	0.86 ±0.001	0.824 ±0.002	0.868 ±0.001	0.712 ±0.11	0.808 ±0.012	0.852 ±0.011	0.856 ±0.001	0.835 ±0.006	0.842 ±0.002	0.841 ±0.005	0.841 ±0.004
led7digit	0.837 ±0.034	0.430 ±0.075	0.827 ±0.02	0.370 ±0.094	0.843 ±0.023	0.825 ±0.027	0.819 ±0.023	0.851 ±0.016	0.097 ±0.095	0.809 ±0.04	0.505 ±0.233	0.856 ±0.014	0.705 ±0.113	0.838 ±0.018	0.838 ±0.025	0.752 ±0.083
magic	0.758 ±0.001	0.717 ±0.001	0.388 ±0.011	0.746 ±0.001	0.748 ±0.002	0.548 ±0.008	0.710 ±0.037	0.748 ±0.002	0.421 ±0.263	0.723 ±0.005	0.753 ±0.002	0.757 ±0.001	0.718 ±0.012	0.749 ±0.001	0.749 ±0.001	0.746 ±0.001
mushroom	0.933 ±0.002	0.897 ±0.001	0.881 ±0.003	0.941 ±0.001	0.941 ±0.001	0.923 ±0.004	0.919 ±0.003	0.903 ±0.008	0.932 ±0.002	0.933 ±0.002	0.933 ±0.002	0.933 ±0.004	0.924 ±0.008	0.926 ±0.005	0.911 ±0.004	0.919 ±0.005
musk2	0.778 ±0.007	0.767 ±0.003	0.775 ±0.004	0.781 ±0.005	0.792 ±0.003	0.804 ±0.004	0.768 ±0.005	0.765 ±0.006	0.743 ±0.019	0.753 ±0.012	0.761 ±0.015	0.775 ±0.005	0.695 ±0.017	0.775 ±0.003	0.705 ±0.013	0.707 ±0.0012
pageblocks0	0.886 ±0.007	0.726 ±0.011	0.841 ±0.005	0.732 ±0.008	0.876 ±0.004	0.872 ±0.006	0.853 ±0.006	0.865 ±0.007	0.19 ±0.354	0.835 ±0.015	0.009 ±0.027	0.868 ±0.008	0.747 ±0.039	0.849 ±0.005	0.738 ±0.045	0.800 ±0.006
rna	0.914 ±0.002	0.812 ±0.001	0.829 ±0.001	0.917 ±0.001	0.922 ±0.001	0.874 ±0.004	0.053 ±0.025	0.824 ±0.002	0.026 ±0.009	0.853 ±0.006	0.762 ±0.019	0.886 ±0.002	0.531 ±0.023	0.862 ±0.002	0.805 ±0.004	0.806 ±0.003
segment0	0.964 ±0.011	0.832 ±0.007	0.907 ±0.007	0.834 ±0.01	0.957 ±0.005	0.941 ±0.009	0.927 ±0.01	0.911 ±0.012	0.242 ±0.166	0.809 ±0.088	0.755 ±0.025	0.946 ±0.008	0.801 ±0.017	0.949 ±0.007	0.806 ±0.026	0.826 ±0.019
shuttle04	0.983 ±0.007	0.927 ±0.008	0.982 ±0.002	0.926 ±0.01	0.980 ±0.004	0.974 ±0.004	0.957 ±0.016	0.981 ±0.007	0.395 ±0.485	0.948 ±0.021	0.793 ±0.068	0.983 ±0.005	0.915 ±0.031	0.980 ±0.005	0.944 ±0.018	0.918 ±0.028
spambase	0.911 ±0.003	0.897 ±0.001	0.814 ±0.005	0.914 ±0.002	0.914 ±0.002	0.873 ±0.004	0.892 ±0.004	0.907 ±0.002	0.91 ±0.003	0.906 ±0.003	0.912 ±0.003	0.905 ±0.009	0.901 ±0.003	0.91 ±0.003	0.905 ±0.004	0.906 ±0.003
splice	0.768 ±0.009	0.507 ±0.022	0.283 ±0.017	0.746 ±0.005	0.75 ±0.003	0.353 ±0.039	0.012 ±0.037	0.597 ±0.067	0.044 ±0.036	0.027 ±0.027	0.753 ±0.006	0.305 ±0.14	0.281 ±0.134	0.717 ±0.026	0.643 ±0.087	0.633 ±0.024
svmguide1	0.794 ±0.002	0.715 ±0.002	0.447 ±0.015	0.789 ±0.002	0.786 ±0.003	0.653 ±0.005	0.743 ±0.011	0.780 ±0.007	0.784 ±0.002	0.785 ±0.003	0.791 ±0.002	0.779 ±0.003	0.763 ±0.003	0.793 ±0.002	0.79 ±0.003	0.791 ±0.003
svmguide3	0.631 ±0.015	0.266 ±0.024	0.305 ±0.039	0.282 ±0.021	0.617 ±0.014	0.44 ±0.035	0.445 ±0.245	0.603 ±0.005	0.059 ±0.093	0.464 ±0.116	0.161 ±0.195	0.62 ±0.013	0.518 ±0.066	0.519 ±0.043	0.587 ±0.022	0.572 ±0.049
vehicle0	0.720 ±0.023	0.090 ±0.057	0.413 ±0.028	0.094 ±0.027	0.666 ±0.012	0.547 ±0.055	0.175 ±0.221	0.599 ±0.029	0.086 ±0.112	0.212 ±0.148	0.033 ±0.049	0.653 ±0.026	0.328 ±0.074	0.278 ±0.094	0.337 ±0.204	–
vowel0	0.870 ±0.012	0.458 ±0.046	0.829 ±0.011	0.461 ±0.020	0.845 ±0.017	0.840 ±0.011	0.817 ±0.023	0.839 ±0.009	0.523 ±0.253	0.805 ±0.026	0.740 ±0.099	0.842 ±0.014	0.707 ±0.069	0.839 ±0.019	0.813 ±0.018	–
w8a	0.903 ±0.002	0.798 ±0.004	0.897 ±0.002	0.797 ±0.003	0.892 ±0.002	0.895 ±0.001	0.877 ±0.002	0.902 ±0.001	0.584 ±0.274	0.897 ±0.002	0.887 ±0.013	0.884 ±0.004	0.865 ±0.008	0.879 ±0.003	0.891 ±0.003	0.879 ±0.009
yeast1	0.663 ±0.014	0.342 ±0.027	0.311 ±0.028	0.332 ±0.03	0.637 ±0.011	0.381 ±0.048	0.084 ±0.079	0.603 ±0.03	0.031 ±0.051	0.174 ±0.158	0.008 ±0.024	0.574 ±0.016	0.224 ±0.049	0.461 ±0.04	0.388 ±0.154	0.531 ±0.102
yeast3	0.747 ±0.019	0.027 ±0.044	0.591 ±0.019	–	0.681 ±0.013	0.612 ±0.062	0.292 ±0.229	0.607 ±0.029	0.034 ±0.072	0.475 ±0.212	–	0.327 ±0.068	0.527 ±0.048	0.090 ±0.042	0.283 ±0.191	–
yeast5	0.717 ±0.029	0.014 ±0.044	0.642 ±0.022	–	0.626 ±0.021	0.480 ±0.072	0.503 ±0.035	0.651 ±0.029	0.004 ±0.012	0.38 ±0.107	–	–	0.014 ±0.044	–	0.412 ±0.189	–
yeast6	0.872 ±0.031	–	0.76 ±0.029	–	0.761 ±0.025	0.658 ±0.043	0.547 ±0.223	0.789 ±0.024	–	0.391 ±0.202	–	0.122 ±0.092	–	–	0.452 ±0.219	–
yest4	0.778 ±0.029	–	0.645 ±0.044	–	0.664 ±0.039	0.485 ±0.103	0.502 ±0.036	0.675 ±0.014	–	0.461 ±0.082	–	–	–	–	0.281 ±0.223	–

Table C3 The results of different methods in terms of F1 score.

F1	OHGD	CSOGD				AdaC2	CSB2	UOB	RUSB			OOB	OUB	WOOB	WOUB	OEB
		$\mathbf{Cost}_I$	$\mathbf{Cost}_{II}$	$\mathbf{Sum}_I$	$\mathbf{Sum}_{II}$				**1**	**2**	**3**					
a8a	0.668 ±0.002	0.654 ±0.001	0.547 ±0.003	0.669 ±0.002	0.656 ±0.003	0.625 ±0.002	0.667 ±0.003	0.656 ±0.003	0.615 ±0.028	0.622 ±0.003	0.667 ±0.007	0.657 ±0.002	0.619 ±0.004	0.675 ±0.001	0.641 ±0.004	0.648 ±0.004
a9a	0.668 ±0.001	0.656 ±0.001	0.545 ±0.001	0.669 ±0.002	0.653 ±0.002	0.624 ±0.002	0.668 ±0.001	0.657 ±0.002	0.632 ±0.022	0.625 ±0.002	0.664 ±0.005	0.658 ±0.002	0.621 ±0.004	0.675 ±0.002	0.643 ±0.003	0.650 ±0.003
ecoli4	0.329 ±0.056	–	0.151 ±0.018	–	0.174 ±0.027	0.178 ±0.063	0.167 ±0.049	0.191 ±0.018	0.112 ±0.00	0.15 ±0.069	0.112 ±0.00	0.160 ±0.050	–	–	0.131 ±0.030	–
ijcnn1	0.548 ±0.003	0.582 ±0.003	0.446 ±0.002	0.587 ±0.002	0.512 ±0.003	0.562 ±0.002	0.588 ±0.002	0.528 ±0.003	0.329 ±0.087	0.388 ±0.016	0.505 ±0.031	0.568 ±0.002	0.443 ±0.01	0.579 ±0.002	0.452 ±0.011	0.468 ±0.009
led7digit	0.492 ±0.063	0.292 ±0.076	0.420 ±0.026	0.237 ±0.096	0.504 ±0.052	0.530 ±0.047	0.505 ±0.066	0.515 ±0.03	0.157 ±0.003	0.519 ±0.099	0.219 ±0.054	0.605 ±0.059	0.520 ±0.057	0.647 ±0.062	0.470 ±0.092	0.460 ±0.041
magic	0.688 ±0.001	0.651 ±0.002	0.549 ±0.003	0.676 ±0.002	0.676 ±0.003	0.590 ±0.003	0.651 ±0.025	0.676 ±0.002	0.579 ±0.066	0.661 ±0.003	0.682 ±0.003	0.687 ±0.001	0.654 ±0.008	0.680 ±0.002	0.677 ±0.001	0.675 ±0.002
mushroom	0.930 ±0.002	0.891 ±0.001	0.891 ±0.002	0.938 ±0.002	0.939 ±0.001	0.924 ±0.004	0.920 ±0.003	0.897 ±0.008	0.930 ±0.002	0.930 ±0.002	0.931 ±0.002	0.930 ±0.004	0.921 ±0.009	0.923 ±0.006	0.907 ±0.004	0.915 ±0.006
musk2	0.482 ±0.007	0.487 ±0.004	0.478 ±0.004	0.504 ±0.005	0.498 ±0.003	0.514 ±0.006	0.474 ±0.006	0.469 ±0.006	0.449 ±0.019	0.459 ±0.012	0.468 ±0.017	0.479 ±0.005	0.404 ±0.014	0.479 ±0.003	0.412 ±0.011	0.415 ±0.0012
pageblocks0	0.641 ±0.017	0.616 ±0.009	0.474 ±0.009	0.625 ±0.010	0.601 ±0.011	0.633 ±0.014	0.668 ±0.01	0.647 ±0.016	0.254 ±0.146	0.632 ±0.013	0.185 ±0.000	0.688 ±0.009	0.617 ±0.023	0.702 ±0.008	0.602 ±0.031	0.586 ±0.0024
rna	0.874 ±0.002	0.742 ±0.001	0.761 ±0.001	0.879 ±0.001	0.884 ±0.001	0.810 ±0.005	0.501 ±0.001	0.765 ±0.003	0.500 ±0.000	0.787 ±0.007	0.694 ±0.016	0.828 ±0.002	0.579 ±0.007	0.820 ±0.002	0.734 ±0.005	0.743 ±0.003
segment0	0.875 ±0.044	0.805 ±0.009	0.667 ±0.018	0.809 ±0.011	0.858 ±0.016	0.795 ±0.024	0.808 ±0.029	0.684 ±0.032	0.268 ±0.027	0.539 ±0.149	0.437 ±0.024	0.806 ±0.03	0.494 ±0.02	0.862 ±0.037	0.49 ±0.03	0.535 ±0.026
shuttle04	0.912 ±0.06	0.854 ±0.016	0.892 ±0.009	0.862 ±0.011	0.932 ±0.024	0.909 ±0.022	0.807 ±0.07	0.887 ±0.076	0.381 ±0.34	0.808 ±0.092	0.309 ±0.124	0.943 ±0.036	0.803 ±0.048	0.959 ±0.009	0.758 ±0.052	0.748 ±0.047
spambase	0.889 ±0.004	0.871 ±0.002	0.790 ±0.004	0.894 ±0.002	0.893 ±0.002	0.843 ±0.004	0.864 ±0.005	0.884 ±0.003	0.886 ±0.003	0.881 ±0.003	0.889 ±0.003	0.882 ±0.01	0.877 ±0.004	0.889 ±0.004	0.882 ±0.004	0.854 ±0.005
splice	0.768 ±0.009	0.671 ±0.008	0.666 ±0.002	0.738 ±0.009	0.750 ±0.009	0.675 ±0.006	0.649 ±0.001	0.609 ±0.07	0.649 ±0.001	0.650 ±0.000	0.748 ±0.007	0.636 ±0.024	0.627 ±0.022	0.707 ±0.047	0.695 ±0.039	0.597 ±0.026
svmguide1	0.765 ±0.003	0.704 ±0.003	0.635 ±0.004	0.759 ±0.002	0.757 ±0.003	0.706 ±0.002	0.739 ±0.006	0.749 ±0.008	0.760 ±0.002	0.760 ±0.002	0.763 ±0.002	0.758 ±0.003	0.746 ±0.003	0.765 ±0.002	0.760 ±0.003	0.762 ±0.003
svmguide3	0.450 ±0.016	0.131 ±0.022	0.392 ±0.007	0.146 ±0.020	0.433 ±0.015	0.405 ±0.01	0.403 ±0.021	0.419 ±0.006	0.386 ±0.004	0.400 ±0.019	0.388 ±0.010	0.440 ±0.014	0.409 ±0.013	0.365 ±0.039	0.413 ±0.02	0.391 ±0.039
vehicle0	0.545 ±0.025	–	0.416 ±0.007	0.019 ±0.011	0.485 ±0.014	0.447 ±0.022	0.377 ±0.009	0.423 ±0.013	0.38 ±0.002	0.379 ±0.003	0.381 ±0.001	0.492 ±0.017	0.375 ±0.007	0.143 ±0.081	0.402 ±0.019	–
vowel0	0.545 ±0.019	0.339 ±0.053	0.431 ±0.021	0.342 ±0.023	0.503 ±0.020	0.529 ±0.022	0.555 ±0.033	0.479 ±0.020	0.257 ±0.098	0.471 ±0.042	0.337 ±0.073	0.546 ±0.032	0.473 ±0.026	0.595 ±0.024	0.412 ±0.048	–
w8a	0.302 ±0.003	0.303 ±0.002	0.307 ±0.002	0.302 ±0.002	0.279 ±0.003	0.34 ±0.001	0.338 ±0.002	0.303 ±0.004	0.153 ±0.111	0.291 ±0.012	0.268 ±0.034	0.333 ±0.002	0.315 ±0.002	0.332 ±0.001	0.315 ±0.004	0.315 ±0.005
yeast1	0.533 ±0.016	0.202 ±0.028	0.462 ±0.005	0.193 ±0.031	0.503 ±0.013	0.471 ±0.008	0.449 ±0.002	0.486 ±0.014	0.449 ±0.001	0.448 ±0.006	0.448 ±0.000	0.514 ±0.008	0.445 ±0.004	0.33 ±0.043	0.474 ±0.017	0.407 ±0.105
yeast3	0.388 ±0.021	–	0.258 ±0.007	–	0.318 ±0.010	0.27 ±0.041	0.24 ±0.05	0.263 ±0.017	0.198 ±0.001	0.264 ±0.057	0.198 ±0.00	0.162 ±0.053	0.211 ±0.019	–	0.207 ±0.014	–
yeast5	0.152 ±0.015	–	0.124 ±0.007	–	0.106 ±0.007	0.161 ±0.040	0.135 ±0.029	0.119 ±0.010	0.066 ±0.000	0.095 ±0.023	0.066 ±0.000	–	–	–	0.077 ±0.029	–
yeast6	0.279 ±0.030	–	0.188 ±0.013	–	0.152 ±0.011	0.283 ±0.039	0.185 ±0.072	0.172 ±0.014	0.058 ±0.000	0.111 ±0.041	0.058 ±0.000	–	–	–	0.08 ±0.024	–
yest4	0.141 ±0.021	–	0.113 ±0.016	–	0.086 ±0.011	0.145 ±0.041	0.107 ±0.023	0.095 ±0.012	0.046 ±0.000	0.081 ±0.035	0.046 ±0.000	–	–	–	0.069 ±0.040	–

Table C4 The comparison with chunk based methods in terms of AUC.

AUC	OHGD	HDWE	KNORAU2	
			NON	ROS
a8a	0.816 ±0.003	0.829 ±0.001	0.601 ±0.016	0.551 ±0.017
a9a	0.815 ±0.002	0.831 ±0.001	0.614 ±0.011	0.564 ±0.015
ecoli4	0.791 ±0.094	0.94 ±0.007	0.941 ±0.006	0.92 ±0.022
ijcnn1	0.871 ±0.003	0.903 ±0.000	0.903 ±0.000	0.703 ±0.004
led7digit	0.834 ±0.038	0.926 ±0.01	0.926 ±0.007	0.847 ±0.018
magic	0.757 ±0.002	0.457 ±0.033	0.675 ±0.004	0.423 ±0.02
mushroom	0.931 ±0.002	0.917 ±0.004	0.923 ±0.003	0.926 ±0.002
musk2	0.798 ±0.007	0.6 ±0.015	0.761 ±0.005	0.699 ±0.007
pageblocks0	0.883 ±0.009	0.871 ±0.045	0.905 ±0.003	0.888 ±0.014
rna	0.91 ±0.003	0.73 ±0.02	0.66 ±0.002	0.659 ±0.002
segment0	0.97 ±0.011	0.848 ±0.007	0.856 ±0.002	0.543 ±0.014
shuttle04	0.982 ±0.007	0.978 ±0.004	0.974 ±0.006	0.995 ±0.002
spambase	0.911 ±0.004	0.881 ±0.006	0.867 ±0.004	0.864 ±0.006
splice	0.769 ±0.01	0.546 ±0.027	0.524 ±0.002	0.786 ±0.002
svmguide1	0.614 ±0.022	0.794 ±0.005	0.813 ±0.007	0.883 ±0.008
svmguide3	0.632 ±0.017	0.762 ±0.003	0.403 ±0.037	0.399 ±0.016
vehicle0	0.718 ±0.026	0.782 ±0.024	0.731 ±0.029	0.667 ±0.014
vowel0	0.863 ±0.02	0.915 ±0.008	0.91 ±0.004	0.842 ±0.026
w8a	0.901 ±0.003	0.885 ±0.000	0.95 ±0.002	0.95 ±0.002
yeast1	0.662 ±0.018	0.711 ±0.003	0.413 ±0.013	0.409 ±0.007
yeast3	0.743 ±0.024	0.89 ±0.002	0.891 ±0.006	0.731 ±0.012
yeast5	0.715 ±0.031	0.97 ±0.001	0.97 ±0.001	0.94 ±0.004
yeast6	0.856 ±0.048	0.977 ±0.001	0.977 ±0.001	0.923 ±0.009
yeast4	0.779 ±0.028	0.965 ±0.001	0.965 ±0.001	0.869 ±0.009

Table C5 The comparison with chunk based methods in terms of GMEANS.

GMEANS	OHGD	HDWE	KNORAU2	
			NON	ROS
a8a	0.814 ±0.002	0.703 ±0.009	0.683 ±0.014	0.635 ±0.017
a9a	0.815 ±0.001	0.699 ±0.006	0.695 ±0.01	0.649 ±0.014
ecoli4	0.797 ±0.067	–	0.029 ±0.059	0.81 ±0.176
ijcnn1	0.870 ±0.002	–	–	0.617 ±0.008
led7digit	0.837 ±0.034	0.168 ±0.118	0.209 ±0.136	0.827 ±0.093
magic	0.758 ±0.001	0.349 ±0.096	0.262 ±0.03	0.326 ±0.043
mushroom	0.933 ±0.002	0.914 ±0.004	0.921 ±0.003	0.924 ±0.002
musk2	0.778 ±0.007	0.696 ±0.009	0.788 ±0.004	0.771 ±0.007
pageblocks0	0.886 ±0.007	0.085 ±0.061	0.396 ±0.044	0.791 ±0.027
rna	0.914 ±0.002	0.352 ±0.09	0.696 ±0.002	0.695 ±0.002
segment0	0.964 ±0.011	0.033 ±0.032	0.069 ±0.023	0.66 ±0.012
shuttle04	0.983 ±0.007	0.778 ±0.059	0.711 ±0.101	0.95 ±0.025
spambase	0.911 ±0.003	0.88 ±0.004	0.873 ±0.004	0.87 ±0.005
splice	0.768 ±0.009	0.255 ±0.09	0.056 ±0.01	0.484 ±0.061
svmguide1	0.794 ±0.002	0.767 ±0.005	0.758 ±0.011	0.883 ±0.008
svmguide3	0.631 ±0.015	0.024 ±0.018	0.412 ±0.028	0.417 ±0.017
vehicle0	0.720 ±0.023	0.457 ±0.158	0.749 ±0.028	0.743 ±0.015
vowel0	0.870 ±0.012	0.182 ±0.118	0.059 ±0.069	0.793 ±0.047
w8a	0.903 ±0.002	0.424 ±0.007	0.144 ±0.015	0.27 ±0.015
yeast1	0.663 ±0.014	–	0.39 ±0.02	0.383 ±0.016
yeast3	0.747 ±0.019	–	0.094 ±0.043	0.773 ±0.016
yeast5	0.717 ±0.029	–	–	0.553 ±0.094
yeast6	0.872 ±0.031	–	–	0.355 ±0.121
yeast4	0.778 ±0.029	–	–	0.455 ±0.064

Table C6 The comparison with chunk based methods in terms of F1 score.

F1	OHGD	HDWE	KNORAU2	
			NON	ROS
a8a	0.668 ±0.002	0.6 ±0.008	0.538 ±0.009	0.511 ±0.009
a9a	0.668 ±0.001	0.597 ±0.005	0.543 ±0.006	0.516 ±0.008
ecoli4	0.329 ±0.056	–	0.027 ±0.057	0.539 ±0.108
ijcnn1	0.548 ±0.003	–	–	0.261 ±0.006
led7digit	0.492 ±0.063	0.14 ±0.106	0.17 ±0.105	0.469 ±0.048
magic	0.688 ±0.001	0.283 ±0.09	0.798 ±0.002	0.214 ±0.05
mushroom	0.930 ±0.002	0.91 ±0.004	0.917 ±0.003	0.92 ±0.002
musk2	0.482 ±0.007	0.406 ±0.008	0.515 ±0.006	0.478 ±0.008
pageblocks0	0.641 ±0.017	0.067 ±0.047	0.291 ±0.032	0.566 ±0.035
rna	0.874 ±0.002	0.287 ±0.088	0.649 ±0.001	0.648 ±0.001
segment0	0.875 ±0.044	0.018 ±0.018	0.047 ±0.018	0.367 ±0.008
shuttle04	0.912 ±0.06	0.751 ±0.058	0.686 ±0.102	0.936 ±0.027
spambase	0.889 ±0.004	0.852 ±0.006	0.84 ±0.005	0.838 ±0.006
splice	0.768 ±0.009	0.233 ±0.093	0.682 ±0.001	0.481 ±0.064
svmguide1	0.765 ±0.003	0.828 ±0.01	0.856 ±0.004	0.892 ±0.008
svmguide3	0.450 ±0.016	0.015 ±0.012	0.392 ±0.01	0.392 ±0.009
vehicle0	0.545 ±0.025	0.362 ±0.134	0.588 ±0.021	0.576 ±0.016
vowel0	0.545 ±0.019	0.146 ±0.099	0.048 ±0.056	0.476 ±0.034
w8a	0.302 ±0.003	0.117 ±0.001	0.06 ±0.007	0.143 ±0.007
yeast1	0.533 ±0.016	–	0.477 ±0.006	0.476 ±0.006
yeast3	0.388 ±0.021	–	0.075 ±0.035	0.425 ±0.011
yeast5	0.152 ±0.015	–	–	0.307 ±0.047
yeast6	0.279 ±0.030	–	–	0.173 ±0.049
yeast4	0.141 ±0.021	–	–	0.184 ±0.024

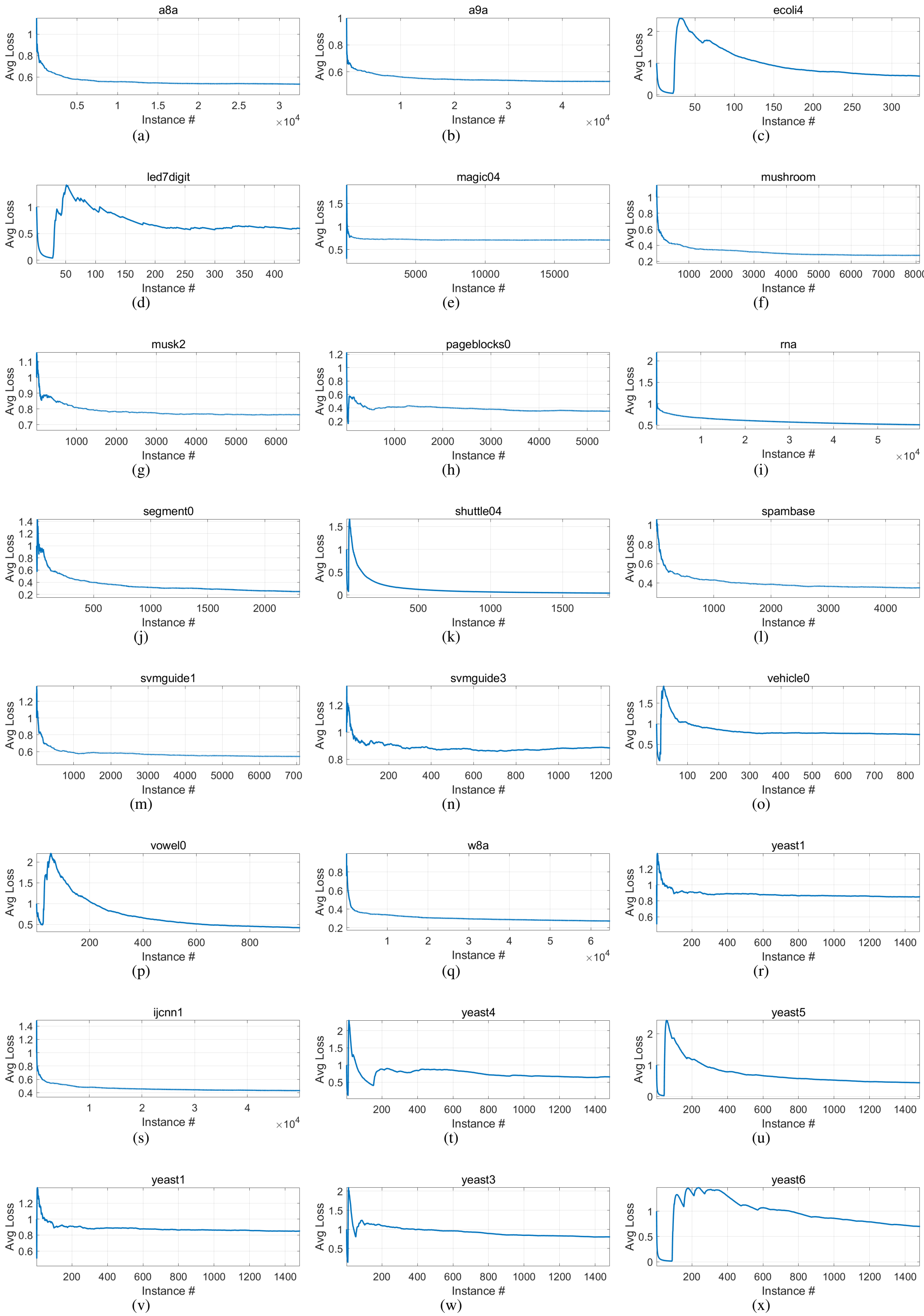


Figure C25 The average accumulative loss curves of the proposed method.