Weight Learning from Cost Matrix in Weighted Least Squares Model Based on Genetic Algorithm

Zhu Hong

October 15, 2018

Background

$$\mathbf{C} = \begin{bmatrix} 0 & \cdots & c_{1m} \\ \vdots & \vdots & \vdots \\ c_{m1} & \cdots & 0 \end{bmatrix}_{m \times m}$$

Motivation

- The existing cost-sensitive learning methods cannot guarantee the overall misclassification cost attains its minimum value.
- The weights in Weighted Least Squares model are generally unknown and finding these weights is usually difficult.

Contribution

- We propose a weight learning method for constructing a cost-sensitive classifier.
- We complete the transition from the misclassification cost matrix (MCM) to the weights in the weighted least square model.

Model Construction

- Objective Function:

$$\text{Min: } E = w_1 \sum_{i=1}^{N_1} \left(\sum_{j=1}^{\tilde{N}} \boldsymbol{\beta}_j g\left(\boldsymbol{\alpha}_j \cdot \boldsymbol{x}_i + b_j \right) - \boldsymbol{t}_i \right)^2$$

$$+ w_2 \sum_{i=N_1+1}^{N_2} \left(\sum_{j=1}^{\tilde{N}} \boldsymbol{\beta}_j g\left(\boldsymbol{\alpha}_j \cdot \boldsymbol{x}_i + b_j \right) - \boldsymbol{t}_i \right)^2$$

$$+ \cdots$$

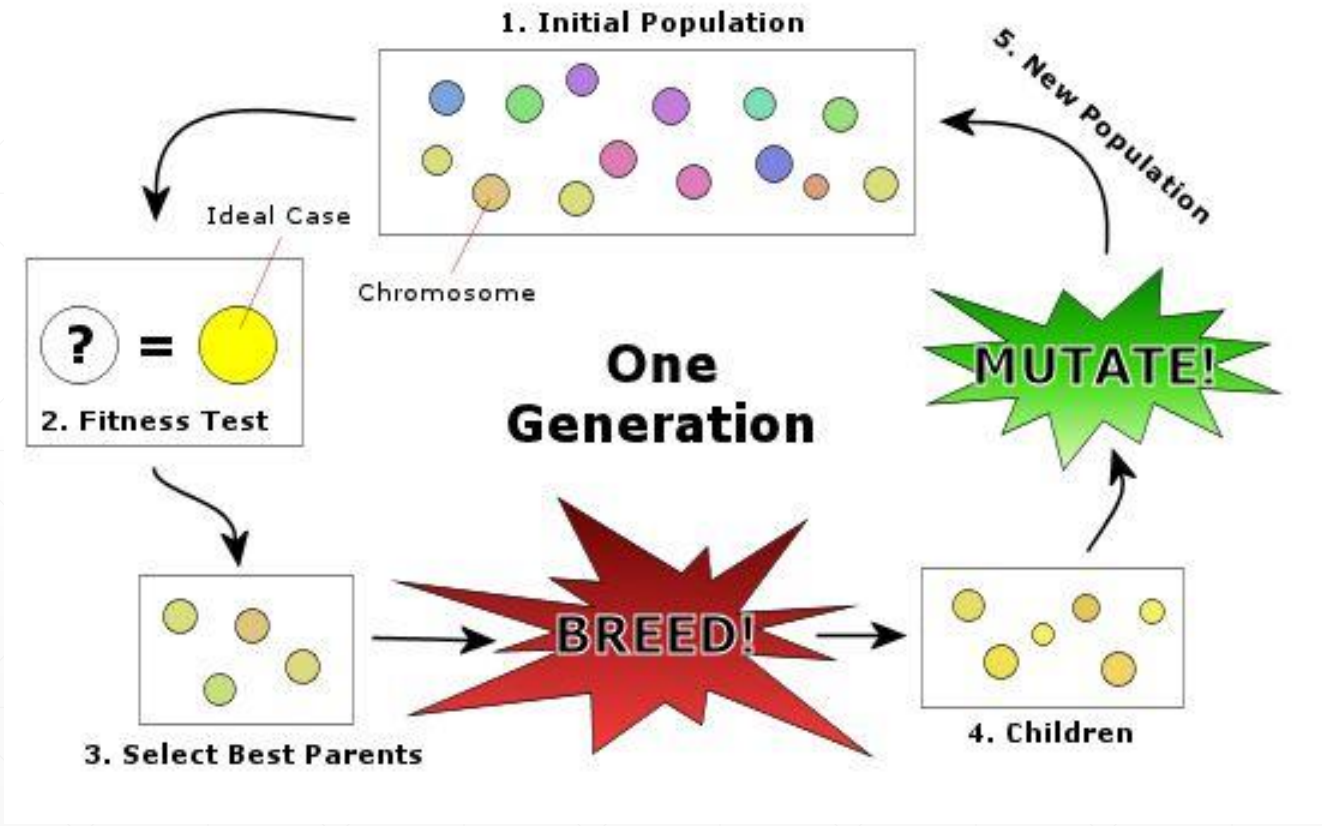
$$+ w_m \sum_{i=N_{m-1}+1}^{N_m} \left(\sum_{j=1}^{\tilde{N}} \boldsymbol{\beta}_j g\left(\boldsymbol{\alpha}_j \cdot \boldsymbol{x}_i + b_j \right) - \boldsymbol{t}_i \right)^2$$

- Relational Expression between Output Weights and Weights of the Least Squares Model:

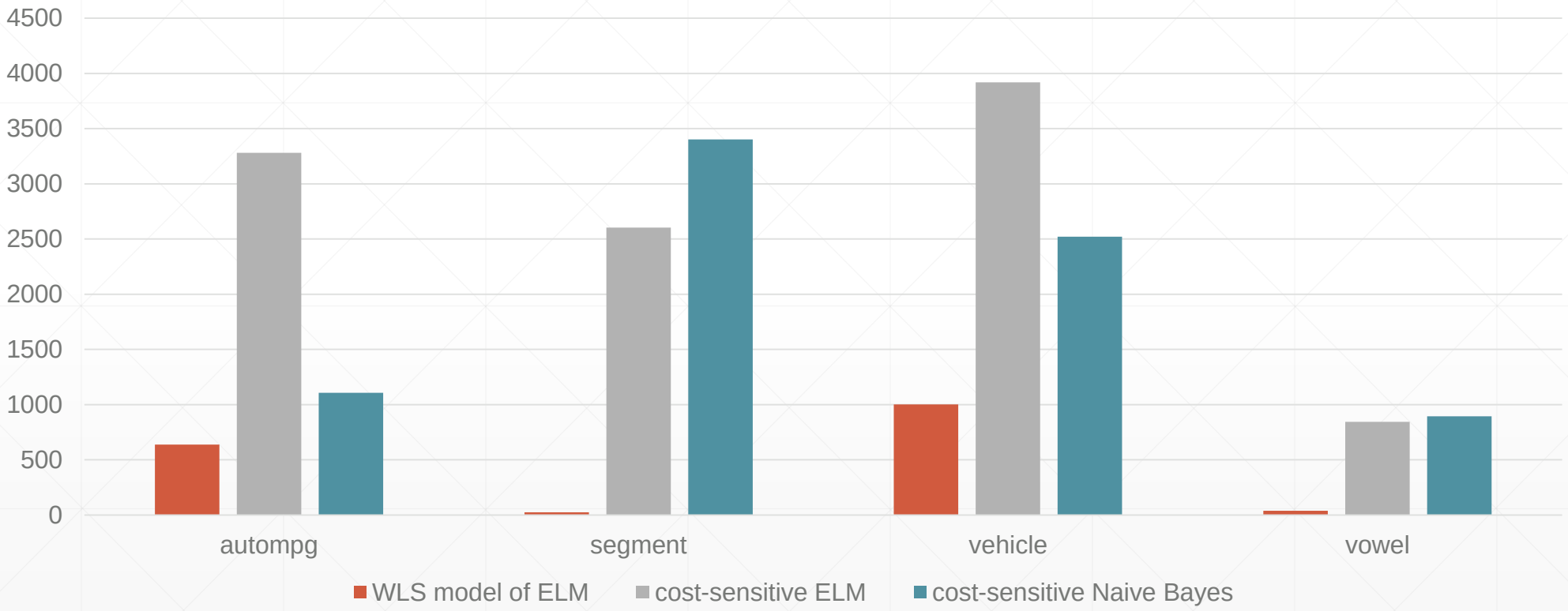
$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} w_1\mathbf{H}_1 \\ \vdots \\ w_m\mathbf{H}_m \end{pmatrix}^{+} \begin{pmatrix} w_1\mathbf{T}_1 \\ \vdots \\ w_m\mathbf{T}_m \end{pmatrix}$$

where $\mathbf{H} = \begin{bmatrix} g(\boldsymbol{\alpha}_1 \cdot \boldsymbol{x}_1 + b_1) & \cdots & g(\boldsymbol{\alpha}_{\widetilde{N}} \cdot \boldsymbol{x}_1 + b_{\widetilde{N}}) \\ \vdots & \vdots & \vdots \\ g(\boldsymbol{\alpha}_1 \cdot \boldsymbol{x}_N + b_1) & \cdots & g(\boldsymbol{\alpha}_{\widetilde{N}} \cdot \boldsymbol{x}_N + b_{\widetilde{N}}) \end{bmatrix}_{N \times \widetilde{N}}$ is the hidden layer output matrix, “+” is the Moore-Penrose generalized inverse operator.

Genetic Algorithm



Experimental Validation



Conclusion

- We propose a weight learning method for constructing a WLS model of ELM.
- We complete the transition from the misclassification cost matrix (MCM) to the weights in the weighted least square model.
- We conduct comparative experiments to verify the validity of the proposed method.