
Algorithm 4: Supervised learning cut level algorithm

input : Training set $\mathbf{T} = \mathbf{Z}_{S_1}^{train}$ and test set $\mathbf{U} = \mathbf{Z}_{S_1}^{test}$
output: Matrix $\tilde{\mathbf{X}}^{(s^*)}$ of aggregated-SNP at best cut level s^*

hierarchy \leftarrow Constrained-HAC on \mathbf{T}
cutlevel \leftarrow Initialize levels where to cut hierarchy
for $s \leftarrow$ Sequence(cutlevel) do
 $\tilde{\mathbf{T}}^s \leftarrow \text{Aggregating}(\mathbf{T}, \text{hierarchy}, \text{cutlevel}[s]);$
 $\tilde{\mathbf{U}}^s \leftarrow \text{Aggregating}(\mathbf{U}, \text{hierarchy}, \text{cutlevel}[s]);$
 ridgecoef $\leftarrow \text{RidgeRegression}(\mathbf{y}_{S_1}^{train} \sim \tilde{\mathbf{T}}^s);$
 $\mathbf{y}_{S_1}^{pred} \leftarrow \text{Predict}(\tilde{\mathbf{U}}^s, \text{ridgecoef});$
 $\text{AUC}[s] \leftarrow \text{ROC}(\mathbf{y}_{S_1}^{test}, \mathbf{y}_{S_1}^{pred});$
end
 $s^* \leftarrow \text{Which}(\text{cutlevel}, \text{Max}(\text{AUC}));$
 $\tilde{\mathbf{X}}^{(s^*)} \leftarrow \text{Aggregating}(\mathbf{Z}, \text{hierarchy}, \text{bestlevel});$
