# Learning Safe Prediction for Semi-Supervised Regression

Yu-Feng Li Han-Wen Zha Zhi-Hua Zhou

Lamda,Nanjing University;AAAI 2017

陈永恒,2017.5.22

# Outline

# Outline

1. Introduction
   - Motivation
   - Related Work
2. The Proposed Method
   - Problem Setting and Formulation
   - Representation to Geometric Projection
   - How the Proposal Works
3. Experiments

# Motivation

- Semi-supervised learning (SSL) concerns the problem on how to improve learning performance via the usage of additional unlabeled data.

# Motivation

- Semi-supervised learning (SSL) concerns the problem on how to improve learning performance via the usage of additional unlabeled data.
- Such a learning framework has received a great deal of attention owing to immense demands in real-world applications

# Motivation

- Semi-supervised learning (SSL) concerns the problem on how to improve learning performance via the usage of additional unlabeled data.
- Such a learning framework has received a great deal of attention owing to immense demands in real-world applications
- Despite the success of SSL, studies reveal that SSL with the exploitation of unlabeled data might even deteriorate learning performance.

# Motivation

- It is highly desirable to study safe SSL scheme that on one side could often improve performance, on the other side will not hurt performance

# Motivation

- It is highly desirable to study safe SSL scheme that on one side could often improve performance, on the other side will not hurt performance
- Recently several proposals have been developed to alleviate such a fundamental challenge for semi-supervised classification (SSC), while the efforts on semi-supervised regression (SSR) remain to be limited.

# Motivation

- It is highly desirable to study safe SSL scheme that on one side could often improve performance, on the other side will not hurt performance
- Recently several proposals have been developed to alleviate such a fundamental challenge for semi-supervised classification (SSC), while the efforts on semi-supervised regression (SSR) remain to be limited.

# Outline

# Related Work

- SSC
  - Safe semi-supervised SVMs(Li and Zhou 2011; 2015)

# Related Work

- SSC
  - Safe semi-supervised SVMs(Li and Zhou 2011; 2015)
  - Safe semi-supervised classifier based on semi-supervised least square classifiers.(Krijthe and Loog 2015)

# Related Work

- SSC
  - Safe semi-supervised SVMs(Li and Zhou 2011; 2015)
  - Safe semi-supervised classifier based on semi-supervised least square classifiers.(Krijthe and Loog 2015)
  - A generic safe SSC framework for variants of performance measures (Li, Kwok, and Zhou 2016)

# Related Work

- SSC
  - Safe semi-supervised SVMs(Li and Zhou 2011; 2015)
  - Safe semi-supervised classifier based on semi-supervised least square classifiers.(Krijthe and Loog 2015)
  - A generic safe SSC framework for variants of performance measures (Li, Kwok, and Zhou 2016)
  - Judging the quality of graph in graph-based SSC(Li, Wang, and Zhou 2016)

# Related Work

- SSC
  - Safe semi-supervised SVMs(Li and Zhou 2011; 2015)
  - Safe semi-supervised classifier based on semi-supervised least square classifiers.(Krijthe and Loog 2015)
  - A generic safe SSC framework for variants of performance measures (Li, Kwok, and Zhou 2016)
  - Judging the quality of graph in graph-based SSC(Li, Wang, and Zhou 2016)

# Related Work

- SSC
  - Safe semi-supervised SVMs(Li and Zhou 2011; 2015)
  - Safe semi-supervised classifier based on semi-supervised least square classifiers.(Krijthe and Loog 2015)
  - A generic safe SSC framework for variants of performance measures (Li, Kwok, and Zhou 2016)
  - Judging the quality of graph in graph-based SSC(Li, Wang, and Zhou 2016)
- SSR
  - Co-training style algoithm for the learning of two semi-supervised regressors (Zhou and Li 2005)

# Related Work

- SSC
  - Safe semi-supervised SVMs(Li and Zhou 2011; 2015)
  - Safe semi-supervised classifier based on semi-supervised least square classifiers.(Krijthe and Loog 2015)
  - A generic safe SSC framework for variants of performance measures (Li, Kwok, and Zhou 2016)
  - Judging the quality of graph in graph-based SSC(Li, Wang, and Zhou 2016)
- SSR
  - Co-training style algoithm for the learning of two semi-supervised regressors (Zhou and Li 2005)
  - Coregularization framework using multi-view training data (Brefeld et al.2006)

# Outline

# Problem Setting and Formulation

- Let $\{f_1, f_2, ..., f_b\}$ be multiple SSR predictions and $f_0$ be the prediction of certain direct supervised learner, where $f_i \in \mathbb{R}^u, i = 0, ..., b, b$ and $u$ refer to the number of regressors and unlabeled instances. $f^*$ refers to the groud-truth assignment

# Problem Setting and Formulation

- Let $\{f_1, f_2, ..., f_b\}$ be multiple SSR predictions and $f_0$ be the prediction of certain direct supervised learner, where $f_i \in \mathbb{R}^u, i = 0, ..., b, b$ and $u$ refer to the number of regressors and unlabeled instances. $f^*$ refers to the groud-truth assignment
- Final safe prediction $g(\{f_1, f_2, ..., f_b\}, f_0)$, which often outperforms $f_0$, meanwhile could not be worse than $f_0$.

# Problem Setting and Formulation

We start with a simple scenario to alleviate this challenge, where the weights of SSR regressors are known. Specifically:

- Let $\alpha{=}[\alpha_1; \alpha_2; ...; \alpha_b] \geq 0$ be the weights of individual regressors $f_i$ 's.Larger the weight,closer the regressor is to the ground-truth

# Problem Setting and Formulation

We start with a simple scenario to alleviate this challenge, where the weights of SSR regressors are known. Specifically:

- Let $\alpha = [\alpha_1; \alpha_2; ...; \alpha_b] \geq 0$ be the weights of individual regressors $f_i$ 's.Larger the weight,closer the regressor is to the ground-truth
- We use error to measure the performance gain against $f_0$,i.e,$(\|f_0 - f^*\|^2 - \|f - f^*\|^2)$

# Problem Setting and Formulation

We start with a simple scenario to alleviate this challenge, where the weights of SSR regressors are known. Specifically:

- Let $\alpha = [\alpha_1; \alpha_2; ...; \alpha_b] \geq 0$ be the weights of individual regressors $f_i$ 's.Larger the weight,closer the regressor is to the ground-truth
- We use error to measure the performance gain against $f_0$,i.e,$(\|f_0 - f^*\|^2 - \|f - f^*\|^2)$
- But
  $f^*$ is obviously unknown,so we optimize the following functional instead:

$$\max_{\mathbf{f} \in \mathbb{R}^u} \sum_{i=1}^{b} \alpha_i \left( \|\mathbf{f}_0 - \mathbf{f}_i\|^2 - \|\mathbf{f} - \mathbf{f}_i\|^2 \right) \qquad (1$$

# Problem Setting and Formulation

In reality,the explicit weights of individual regressors is difficult to know.For the sake of simplicity, one assumes that $\alpha$ is is from a convex linear set $\mathcal{M} = \{\alpha | A^T \alpha \leq b; a \geq 0\}$,which is a general form that reflects the relation of individual learners in ensemble learning,where **A** and **b** are task-dependent coefficients.

# Problem Setting and Formulation

In reality,the explicit weights of individual regressors is difficult to know.For the sake of simplicity, one assumes that $\alpha$ is is from a convex linear set $\mathcal{M} = \{\alpha | A^T\alpha \leq b; a \geq 0\}$,which is a general form that reflects the relation of individual learners in ensemble learning,where **A** and **b** are task-dependent coefficients.

Without further knowledge to determine the weights of individual regressors, one aim to optimize the worst-case performance gain as follow:

$$\max_{\mathbf{f} \in \mathbb{R}^u} \min_{\boldsymbol{\alpha} \in \mathcal{M}} \sum_{i=1}^{b} \alpha_i \left( \|\mathbf{f}_0 - \mathbf{f}_i\|^2 - \|\mathbf{f} - \mathbf{f}_i\|^2 \right) \qquad (2)$$

# Outline

1. Introduction
   - Motivation
   - Related Work
2. The Proposed Method
   - Problem Setting and Formulation
   - Representation to Geometric Projection
   - How the Proposal Works
3. Experiments

# Representation to Geometric Projection

- Note that Eq.(2) is concave to $f$ and convex to $\alpha$, and thus it is recognized as saddle-point convex-concave optimization.

# Representation to Geometric Projection

- Note that Eq.(2) is concave to $f$ and convex to $\alpha$, and thus it is recognized as saddle-point convex-concave optimization.
- However, it is non-trivial to be solved efficiently because of poor convergence rate.

# Representation to Geometric Projection

- Note that Eq.(2) is concave to $f$ and convex to $\alpha$, and thus it is recognized as saddle-point convex-concave optimization.
- However, it is non-trivial to be solved efficiently because of poor convergence rate.
- In order to alleviate the computational overload and understand how Eq.(2) works, we in the following show that Eq.(2) can be formulated as a geometric projection issue that help address the above concerns.

- By setting the derivative of Eq.(2) w.r.t. f to zero, Eq.(2) has a closed-form solution w.r.t. f as

$$\mathbf{f} = \sum_{i=1}^{b} \alpha_i \mathbf{f}_i, \quad (3$$

- By setting the derivative of Eq.(2) w.r.t. f to zero, Eq.(2) has a closed-form solution w.r.t. f as

$$\mathbf{f} = \sum_{i=1}^{b} \alpha_i \mathbf{f}_i, \quad (3$$

- By substituting Eq.(3) into Eq.(2),we have

$$\min_{\boldsymbol{\alpha} \in \mathcal{M}} \| \sum_{i=1}^{b} \alpha_i \mathbf{f}_i - \mathbf{f}_0 \|^2 \quad (4$$

- By setting the derivative of Eq.(2) w.r.t. f to zero, Eq.(2) has a closed-form solution w.r.t. f as

$$\mathbf{f} = \sum_{i=1}^{b} \alpha_i \mathbf{f}_i, \quad (3$$

- By substituting Eq.(3) into Eq.(2),we have

$$\min_{\boldsymbol{\alpha} \in \mathcal{M}} \| \sum_{i=1}^{b} \alpha_i \mathbf{f}_i - \mathbf{f}_0 \|^2 \quad (4$$

- By expanding the quadratic form in Eq.(4), it can be rewritten as:

$$\min_{\boldsymbol{\alpha} \in \mathcal{M}} \quad \boldsymbol{\alpha}^\top \mathbf{F} \boldsymbol{\alpha} - \mathbf{v}^\top \boldsymbol{\alpha} \quad (5$$

- By setting the derivative of Eq.(2) w.r.t. f to zero, Eq.(2) has a closed-form solution w.r.t. f as

$$\mathbf{f} = \sum_{i=1}^{b} \alpha_i \mathbf{f}_i, \quad (3$$

- By substituting Eq.(3) into Eq.(2),we have

$$\min_{\boldsymbol{\alpha} \in \mathcal{M}} \| \sum_{i=1}^{b} \alpha_i \mathbf{f}_i - \mathbf{f}_0 \|^2 \quad (4$$

- By expanding the quadratic form in Eq.(4), it can be rewritten as:

$$\min_{\boldsymbol{\alpha} \in \mathcal{M}} \quad \boldsymbol{\alpha}^\top \mathbf{F} \boldsymbol{\alpha} - \mathbf{v}^\top \boldsymbol{\alpha} \quad (5$$

Where $F \in \mathbb{R}^{b \times b}$ is a linear kernel matrix of $f_i$'s, $F_{ij} = f_i^T f_j, \forall 1 \leq i, j \leq b$ and $v = [2f_1^T f_0; \ldots; 2f_b^T f_0]$

# Representation to Geometric Projection

**Algorithm 1** The SAFER Method

**Input**: multiple SSR predictions $\{\mathbf{f}_i\}_{i=1}^b$ and certain direct supervised regression prediction $\mathbf{f}_0$

**Output**: the learned prediction $\bar{\mathbf{f}}$

1: Construct a linear kernel matrix $\mathbf{F}$ where $F_{ij} = \mathbf{f}_i^\top \mathbf{f}_j$, $\forall 1 \leq i, j \leq b$

2: Derive a vector $\mathbf{v} = [2\mathbf{f}_1^\top \mathbf{f}_0; \ldots; 2\mathbf{f}_b^\top \mathbf{f}_0]$

3: Solve the convex quadratic optimization Eq.(5) and obtain the optimal solution $\boldsymbol{\alpha}^* = [\alpha_1^*, \ldots, \alpha_b^*]$

4: Return $\bar{\mathbf{f}} = \sum_{i=1}^b \alpha_i^* \mathbf{f}_i$

It is not hard to find that Eq.(4) is a geometric projection problem.Specifically, let $\Omega = \{f| \sum_{i=1}^{b} \alpha_i f_i, \alpha \in \mathcal{M}\}$,Eq.(4) can be rewritten as,

$$\bar{\mathbf{f}} = \arg\min_{\mathbf{f}\in\Omega} ||\mathbf{f} - \mathbf{f}_0||^2, \tag{6}$$

# Representation to Geometric Projection

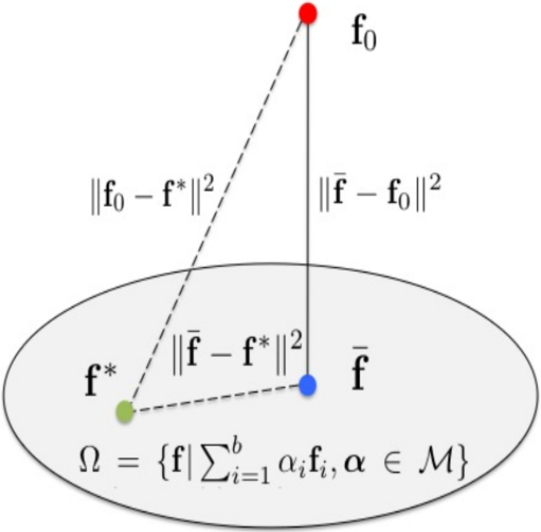


Figure 1: Illustration of the intuition of our proposal via the

# Outline

# How the Proposal Works

Before going into the detail analysis,we notice from Figure 1 that $\|\bar{f} - f^*\|$ should be smaller than $\|f_0 - f^*\|$ if $f^* \in \Omega$.Such an observation motivates us to derive the following results.

**Theorem 1.** $\|\bar{\mathbf{f}} - \mathbf{f}^*\|^2 \leq \|\mathbf{f}_0 - \mathbf{f}^*\|^2$, *if the ground truth label assignment* $\mathbf{f}^* \in \Omega = \{\mathbf{f} | \sum_{i=1}^{b} \alpha_i \mathbf{f}_i, \boldsymbol{\alpha} \in \mathcal{M}\}$.

Before going into the detail analysis,we notice from Figure 1 that $\|\bar{f} - f^*\|$ should be smaller than $\|f_0 - f^*\|$ if $f^* \in \Omega$.Such an observation motivates us to derive the following results.

**Theorem 1.** $\|\bar{\mathbf{f}} - \mathbf{f}^*\|^2 \leq \|\mathbf{f}_0 - \mathbf{f}^*\|^2$, *if the ground truth label assignment* $\mathbf{f}^* \in \Omega = \{\mathbf{f} | \sum_{i=1}^{b} \alpha_i \mathbf{f}_i, \boldsymbol{\alpha} \in \mathcal{M}\}$.

**Theorem 2.** $\bar{\mathbf{f}}$ *has already achieved the maximal worst-case performance gain against* $\mathbf{f}_0$, *if the ground truth* $\mathbf{f}^* \in \Omega$. *Specifically,* $\bar{\mathbf{f}}$ *is the optimal solution of the following functional,*

$$\bar{\mathbf{f}} = \underset{\mathbf{f} \in \mathbb{R}^u}{\operatorname{argmax}} \min_{\mathbf{f}^* \in \Omega} \left( \|\mathbf{f}_0 - \mathbf{f}^*\|^2 - \|\mathbf{f} - \mathbf{f}^*\|^2 \right)$$

# How the Proposal Works

- To understand our proposal more comprehensively, in the following we investigate how the performance of our proposal is affected when the condition previously mentioned is violated.

- To understand our proposal more comprehensively, in the following we investigate how the performance of our proposal is affected when the condition previously mentioned is violated.
- Specifically,let $\lambda^* = [\lambda_1^*, ..., \lambda_b^*] \in \mathcal{M}$ be the optimal solution of the following functional:

$$\boldsymbol{\lambda}^* = \underset{\boldsymbol{\lambda} \in \mathcal{M}}{\operatorname{argmin}} \| \sum_{i=1}^{b} \lambda_i \mathbf{f}_i - \mathbf{f}^* \|^2$$

# How the Proposal Works

- To understand our proposal more comprehensively, in the following we investigate how the performance of our proposal is affected when the condition previously mentioned is violated.
- Specifically,let $\lambda^* = [\lambda_1^*, ..., \lambda_b^*] \in \mathcal{M}$ be the optimal solution of the following functional:

$$\boldsymbol{\lambda}^* = \operatorname*{argmin}_{\boldsymbol{\lambda} \in \mathcal{M}} \| \sum_{i=1}^{b} \lambda_i \mathbf{f}_i - \mathbf{f}^* \|^2$$

- and $\epsilon$ be the residual, i.e., $\epsilon = f^* - \sum_{i=1}^{b} \lambda_i^* f_i$, reflecting the degree of violation.

# How the Proposal Works

- Suppose $f_i$' s are normalized into $[0:1]$,we then have the following result for the proposed method.

# How the Proposal Works

- Suppose $f_i$' s are normalized into $[0:1]$,we then have the following result for the proposed method.
- **Theorem 3.** *The increased loss of the proposed method against* $\mathbf{f}_0$*, i.e.,* $\left(\|\bar{\mathbf{f}} - \mathbf{f}^*\|^2 - \|\mathbf{f}_0 - \mathbf{f}^*\|^2\right)$*, is at most* $\min\{2\|\epsilon\|_1/u, 2\|\epsilon\|_2/\sqrt{u}\}$.

- Suppose $f_i$' s are normalized into $[0:1]$,we then have the following result for the proposed method.
- **Theorem 3.** *The increased loss of the proposed method against* $\mathbf{f}_0$*, i.e.,* $\left(\|\bar{\mathbf{f}}-\mathbf{f}^*\|^2-\|\mathbf{f}_0-\mathbf{f}^*\|^2\right)$*, is at most* $\min\{2\|\epsilon\|_1/u, 2\|\epsilon\|_2/\sqrt{u}\}$.
- Theorem 3 discloses that when the required safeness condition is violated, the worst-case increased loss of our proposed method is only related to the norm of the residual (in other words, the quality of regressors),and has nothing to do with other factors,e.g.,the quantity of regressors.

Table 1: Mean Square Error (mean±std) for the compared methods and SAFER using 5 and 10 labeled instances. For all methods, if the performance is significantly better/worse than the baseline 1NN method, the corresponding entries are bolded/boxed (paired t-tests at 95% significance level). The average mean square error on all the experimental data sets is listed for comparison. The win/tie/loss counts are summarized and the method with the smallest number of losses against 1NN is bolded.

| 5 labeled instances | | | | | | |
|---|---|---|---|---|---|---|
| Dataset | 1NN | Self-$k$NN | Self-LS | COREG | Voting | SAFER |
| abalone | .017 ± .007 | **.014 ± .003** | **.013 ± .004** | **.013 ± .003** | **.012 ± .003** | **.013 ± .003** |
| bodyfat | .024 ± .008 | .025 ± .009 | [.054 ± .016] | .026 ± .008 | [.031 ± .011] | .025 ± .009 |
| cadata | .090 ± .031 | **.073 ± .023** | **.067 ± .022** | **.069 ± .028** | **.069 ± .022** | **.070 ± .023** |
| cpusmall | .027 ± .012 | [.031 ± .008] | [.050 ± .021] | [.031 ± .009] | .024 ± .006 | .028 ± .009 |
| eunite2001 | .052 ± .017 | **.037 ± .015** | **.024 ± .012** | **.037 ± .011** | **.031 ± .013** | **.032 ± .010** |
| housing | .042 ± .007 | .043 ± .009 | [.048 ± .012] | .041 ± .008 | .042 ± .009 | .041 ± .009 |
| mg | .071 ± .035 | **.057 ± .015** | **.053 ± .011** | **.054 ± .019** | **.054 ± .013** | **.053 ± .013** |
| mpg | .029 ± .012 | .030 ± .012 | [.040 ± .014] | .031 ± .012 | .031 ± .012 | .030 ± .012 |
| pyrim | .032 ± .009 | **.027 ± .005** | [.063 ± .012] | .029 ± .011 | **.025 ± .007** | **.025 ± .005** |
| space_ga | .005 ± .002 | .005 ± .003 | [.030 ± .005] | **.004 ± .002** | [.008 ± .002] | **.004 ± .002** |
| Ave. Mse. | .039 | .034 | .044 | .033 | .033 | .032 |
| win/tie/loss against 1NN | | 5/4/1 | 4/0/6 | 5/4/1 | 5/3/2 | **6/4/0** |

| 10 labeled instances | | | | | | |
|---|---|---|---|---|---|---|
| Dataset | 1NN | Self-$k$NN | Self-LS | COREG | Voting | SAFER |
| abalone | .020 ± .010 | **.014 ± .005** | **.013 ± .004** | **.012 ± .003** | **.012 ± .003** | **.013 ± .005** |
| bodyfat | .019 ± .005 | .019 ± .007 | .041 ± .013 | .020 ± .006 | .023 ± .009 | .018 ± .007 |
| cadata | .083 ± .029 | **.063 ± .012** | **.056 ± .007** | **.054 ± .010** | **.057 ± .009** | **.060 ± .013** |
| cpusmall | .024 ± .012 | .027 ± .008 | .042 ± .004 | .028 ± .008 | **.020 ± .005** | .025 ± .008 |
| eunite2001 | .044 ± .014 | **.037 ± .013** | **.020 ± .006** | **.031 ± .009** | **.029 ± .009** | **.029 ± .007** |
| housing | .039 ± .010 | .036 ± .009 | .036 ± .009 | **.035 ± .005** | **.034 ± .008** | **.035 ± .009** |
| mg | .062 ± .019 | **.046 ± .015** | **.048 ± .011** | **.045 ± .015** | **.043 ± .014** | **.045 ± .014** |
| mpg | .022 ± .007 | .020 ± .006 | .030 ± .014 | .021 ± .007 | .021 ± .008 | .020 ± .006 |
| pyrim | .023 ± .006 | **.021 ± .005** | .052 ± .014 | .022 ± .006 | **.020 ± .007** | **.020 ± .006** |
| space_ga | .004 ± .001 | **.003 ± .001** | .028 ± .002 | **.003 ± .001** | .006 ± .001 | **.003 ± .001** |
| Ave. Mse. | .034 | .029 | .037 | .027 | .026 | .027 |
| win/tie/loss against 1NN | | 6/3/1 | 4/1/5 | 6/3/1 | 7/1/2 | **7/3/0** |

Table 2: Mean Absolute Error (mean±std) and Mean $\epsilon$-insensitive Error (mean±std, $\epsilon = 0.05$) for the compared methods and SAFER using 10 labeled instances.

| Mean Absolute Error | | | | | | |
|---|---|---|---|---|---|---|
| Dataset | 1NN | Self-$k$NN | Self-LS | COREG | Voting | SAFER |
| abalone | .100 ± .025 | **.089 ± .020** | **.086 ± .018** | **.083 ± .015** | **.081 ± .018** | **.086 ± .019** |
| bodyfat | .108 ± .013 | .107 ± .018 | .164 ± .026 | .114 ± .015 | .119 ± .023 | .105 ± .018 |
| cadata | .216 ± .037 | **.195 ± .022** | **.189 ± .016** | **.182 ± .023** | **.189 ± .019** | **.192 ± .023** |
| cpusmall | .073 ± .014 | .078 ± .007 | .168 ± .010 | .081 ± .008 | .092 ± .008 | .076 ± .007 |
| eunite2001 | .162 ± .023 | **.152 ± .027** | **.108 ± .016** | **.138 ± .018** | **.132 ± .021** | **.133 ± .017** |
| housing | .137 ± .018 | .135 ± .023 | .140 ± .023 | .135 ± .016 | .131 ± .022 | .132 ± .022 |
| mg | .188 ± .029 | **.166 ± .025** | **.176 ± .017** | **.168 ± .026** | **.163 ± .023** | **.164 ± .025** |
| mpg | .110 ± .014 | .107 ± .018 | .138 ± .029 | .112 ± .020 | .109 ± .022 | .105 ± .018 |
| pyrim | .105 ± .014 | .107 ± .011 | .174 ± .021 | .111 ± .012 | **.095 ± .016** | **.099 ± .014** |
| space_ga | .050 ± .005 | **.043 ± .005** | .131 ± .004 | **.041 ± .005** | .060 ± .004 | **.042 ± .004** |
| Ave. Mae. | .125 | .118 | .147 | .116 | .117 | .114 |
| win/tie/loss against 1NN | | 5/4/1 | 4/1/5 | 5/2/3 | 5/2/3 | **6/4/0** |

| Mean $\epsilon$-insensitive Error | | | | | | |
|---|---|---|---|---|---|---|
| Dataset | 1NN | Self-$k$NN | Self-LS | COREG | Voting | SAFER |
| abalone | .062 ± .023 | **.049 ± .017** | **.046 ± .015** | **.044 ± .012** | **.042 ± .014** | **.046 ± .016** |
| bodyfat | .065 ± .013 | .065 ± .017 | .118 ± .025 | .070 ± .014 | .076 ± .021 | .063 ± .017 |
| cadata | .170 ± .037 | **.149 ± .021** | **.143 ± .015** | **.136 ± .022** | **.143 ± .018** | **.146 ± .023** |
| cpusmall | .039 ± .013 | .043 ± .007 | .122 ± .009 | .046 ± .007 | .053 ± .007 | .041 ± .007 |
| eunite2001 | .117 ± .023 | **.108 ± .026** | **.066 ± .015** | **.094 ± .017** | **.089 ± .020** | **.089 ± .016** |
| housing | .095 ± .017 | .092 ± .022 | .096 ± .021 | .092 ± .014 | .088 ± .020 | .089 ± .021 |
| mg | .143 ± .029 | **.121 ± .024** | **.130 ± .017** | **.123 ± .025** | **.118 ± .023** | **.119 ± .024** |
| mpg | .069 ± .013 | .066 ± .016 | .094 ± .028 | .069 ± .019 | .067 ± .020 | .064 ± .016 |
| pyrim | .066 ± .012 | .065 ± .010 | .129 ± .020 | .068 ± .012 | **.055 ± .015** | **.059 ± .013** |
| space_ga | .016 ± .004 | **.011 ± .003** | .087 ± .004 | **.010 ± .003** | .023 ± .003 | **.010 ± .002** |
| Ave. Mee. | .084 | .077 | .103 | .075 | .075 | .073 |
| win/tie/loss against 1NN | | 5/4/1 | 4/1/5 | 5/3/2 | 5/2/3 | **6/4/0** |