COMPSTAT2010 in Paris

Ensembled Multivariate Adaptive Regression Splines with Nonnegative Garrote Estimator

Hiroki Motogaito
Osaka University

Masashi Goto
Biostatistical Research Association, NPO.

JAPAN

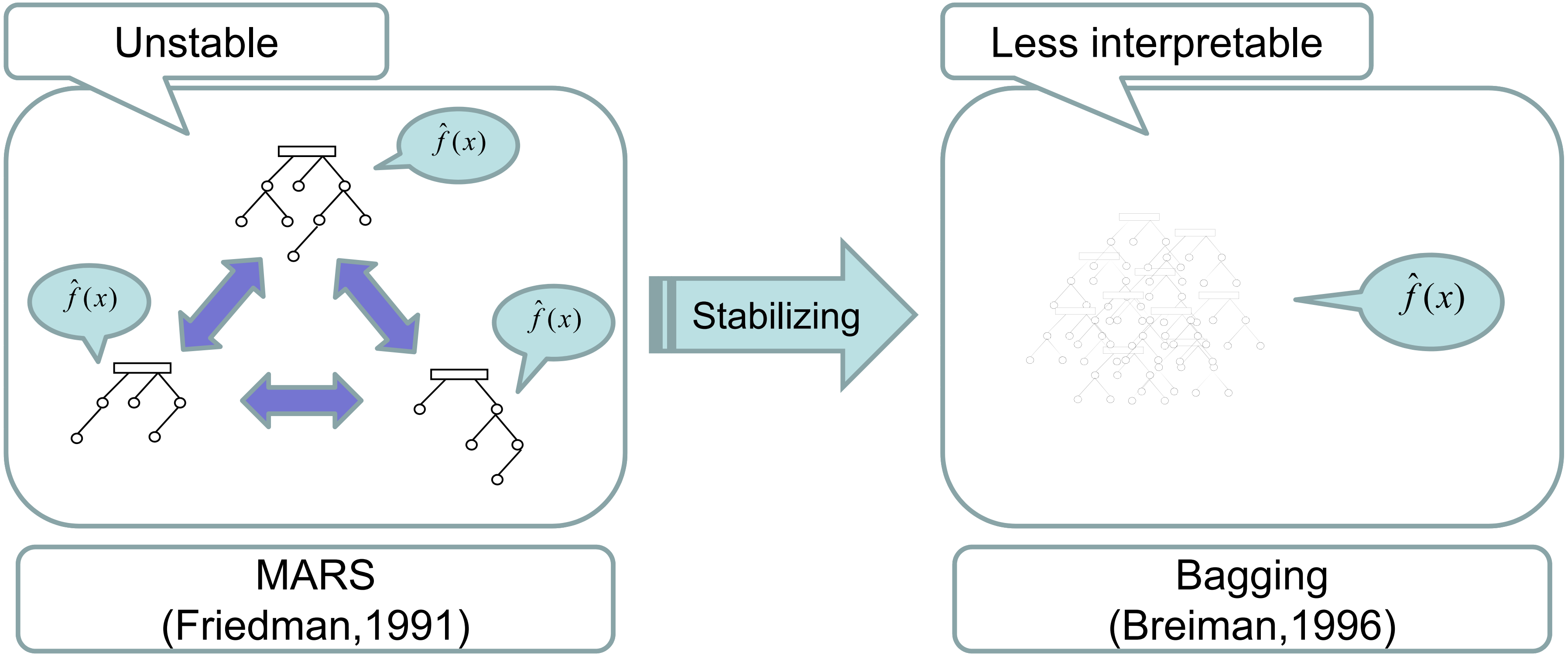
Agenda

- Introduction and motivation
- Tree methods
  - Multivariate Adaptive Regression Splines(MARS)
  - Bagging MARS
- Our method proposed
  - Ensembled MARS with nonnegative garrote
- Example and simulation
- Concluding remarks

Agenda

- Introduction and motivation
- Tree methods
  - ➢Multivariate Adaptive Regression Splines(MARS)
  - ➢Bagging MARS
- Our method proposed
  - ➢Ensembled MARS with nonnegative garrote
- Example and simulation
- Concluding remarks

Introduction and motivation



Motivation

a new version MARS that has both stability and interpretability

Agenda

- Introduction and motivation
- Tree methods
  - Multivariate Adaptive Regression Splines(MARS)
  - Bagging MARS
- Our method proposed
  - Ensembled MARS with nonnegative garrote
- Example and simulation
- Concluding remarks

Multivariate Adaptive Regression Splines(Friedman,1991)

- Model form

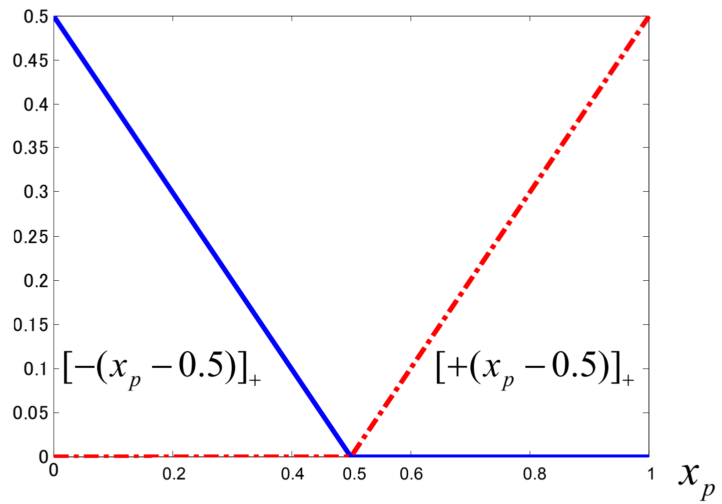
Regression model

$$\hat{f}_{\mathrm{MARS}} = \hat{\beta}_0 + \sum_{m=1}^{M} \hat{\beta}_m B_m(\mathbf{x})$$

Basis function

$$B_m(\mathbf{x}) = \prod_{k=1}^{K_m} [i_{(k,m)}(x_{p(k,m)} - t_{(k,m)})]_+^q$$

- Algorithms
  - Forward stepwise
    - Increase basis functions
  - Backward stepwise
    - Prune off
    - Select the best tree



q=1 and knot t=0.5

Bagging (Breiman,1996)

- Model form(Bagging MARS)

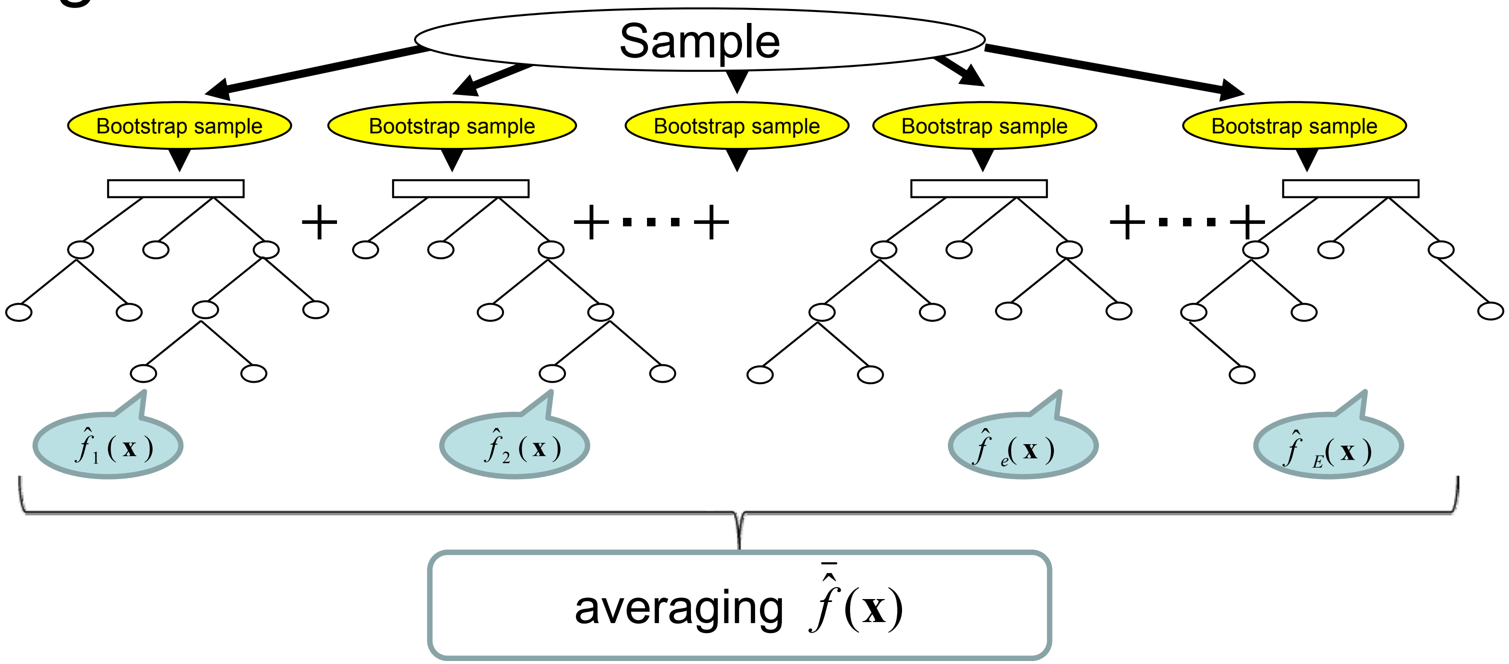
Regression model

$$\hat{f}_{\text{Bagging MARS}} = \frac{1}{E}\sum_{e=1}^{E}\hat{f}_e(\mathbf{x})$$

Each tree

$\hat{f}_e(\mathbf{x})$: MARS model

- Algorithms



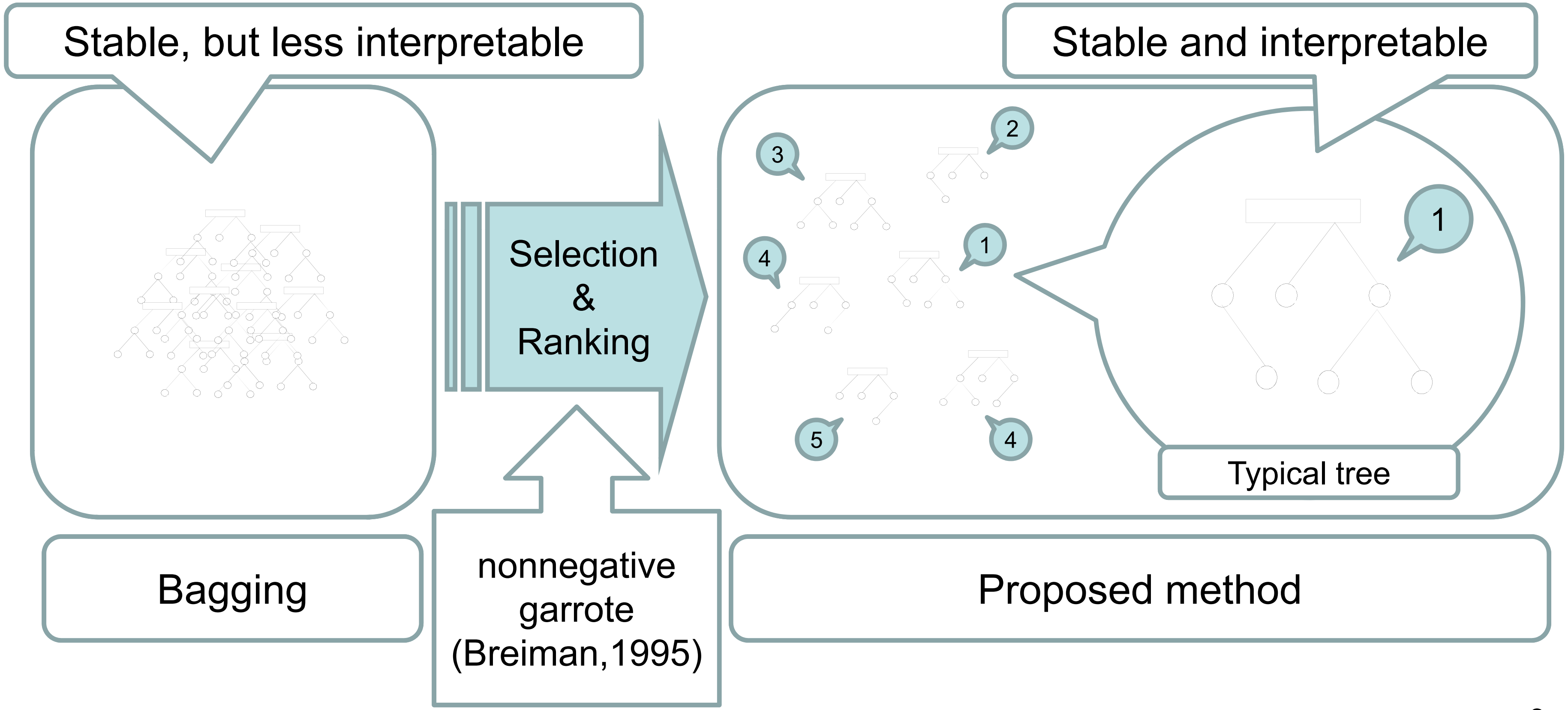
Agenda

- Introduction and motivation
- Previous research
  - Multivariate Adaptive Regression Splines(MARS)
  - Bagging MARS
- Our method proposed
  - Ensembled MARS with nonnegative garrote
- Example and simulation
- Concluding remarks

Proposed method

Motivation

a new version MARS that has both stability and interpretability



Ensembled MARS with non-negative garrote（1/2）

- Model form

Regression model

$$\hat{f} = \sum_{e=1}^{E} \hat{c}_e \hat{f}_e(\mathbf{x})$$

Each tree

$\hat{f}_e(\mathbf{x})$: MARS model , $\hat{c}_e$: non-negative garrote estimator

- Algorithms

> - Generate Bagging trees.
> - Attach c_e on each tree and estimate $\hat{c}_e$ using nonnegative garrote(Breiman,1995).
>   — Select candidate trees(If $\hat{c}_e = 0$, the tree is removed).
> - Get $\hat{f} = \sum_{e=1}^{E} \hat{c}_e \hat{f}_e(\mathbf{x})$
>   - Interpretable structure through typical tree(max $\hat{c}_e$)

Ensembled MARS with non-negative garrote (2/2)

non-negative garrote (Breiman,1995)

$$\{\hat{c}_p\}_1^P = \arg\min_{\{c_p\}_1^P} \sum_{n=1}^{N} (Y_n - \sum_{p=1}^{P} c_p \hat{\beta}_p x_n^{(p)})^2, \quad \text{subject to} \quad c_p \geq 0, \quad \sum_{p=1}^{P} c_p = s \ ,$$

where $\hat{\beta}_p$ is the least square estimator and $1 \leq s \leq P$.

Ensembled MARS with non-negative garrote

$$\{\hat{c}_e\}_1^E = \arg\min_{\{c_e\}_1^E} \sum_{n=1}^{N} (Y_n - \sum_{e=1}^{E} c_e \hat{f}_e(\mathbf{x}_n))^2, \quad \text{subject to} \quad c_e \geq 0, \quad \sum_{e=1}^{E} c_e = 1$$

where $\hat{f}_e(\mathbf{x}_n)$ is MARS model.

characteristics

- All $c_e = 1/E$ indicates Bagging.
- Selection of optimal s is unnecessary($s = 1$).

Agenda

- Introduction and motivation
- Previous research
  - Multivariate Adaptive Regression Splines(MARS)
  - Bagging MARS
- Our method proposed
  - Ensembled MARS with non-negative garrote
- Example and simulation
- Concluding remarks

Literature example

Prostate cancer data (Stamney *et al*.,1989: Tibshirani,1996)

- y : Level of prostate-specific antigen
- $\mathbf{x} = (x_1, \ldots, x_8)^{\mathrm{T}}$: Clinical measures

  x_1 : Log of tumor size

  x_2 : Weight of prostate

  x_3 : Patient’s age

  x_4 : Log of benign prostatic hyperplasia amount

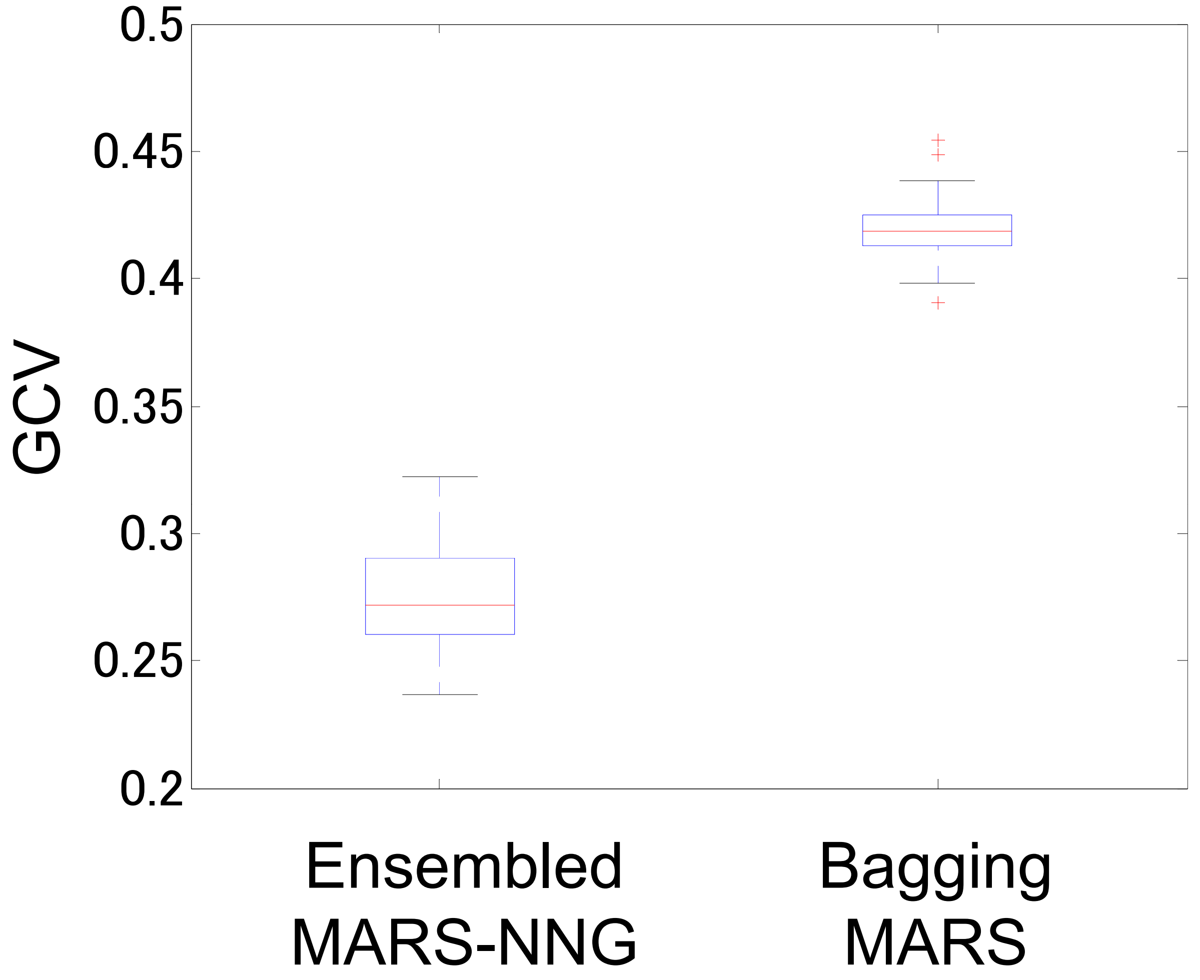
  x_5 : Dummy variables of whether it is metastasizing to seminal vesicle

  x_6 : Log of capsular penetration

  x_7 : Gleason score

  x_8 : Gleason score’s ratio of 4 or 5
- Sample size : $N = 97$

Literature example

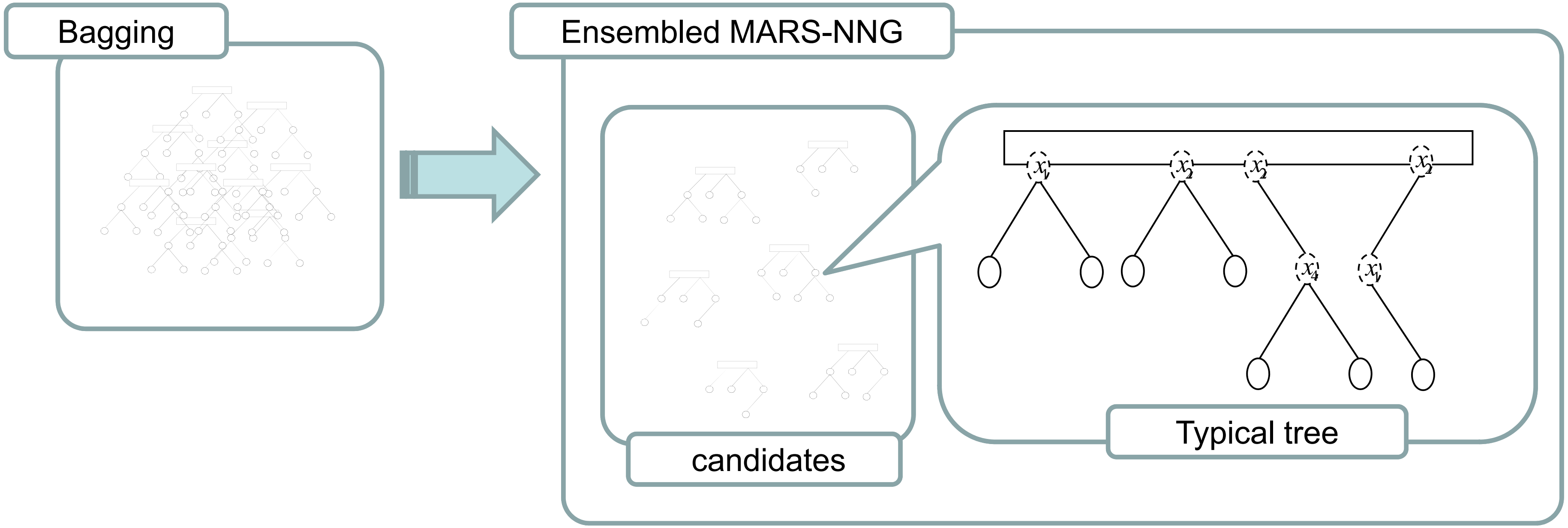


Literature example

- Number of trees

Bagging: 97 → Ensembled MARS-NNG: 9

- Structure



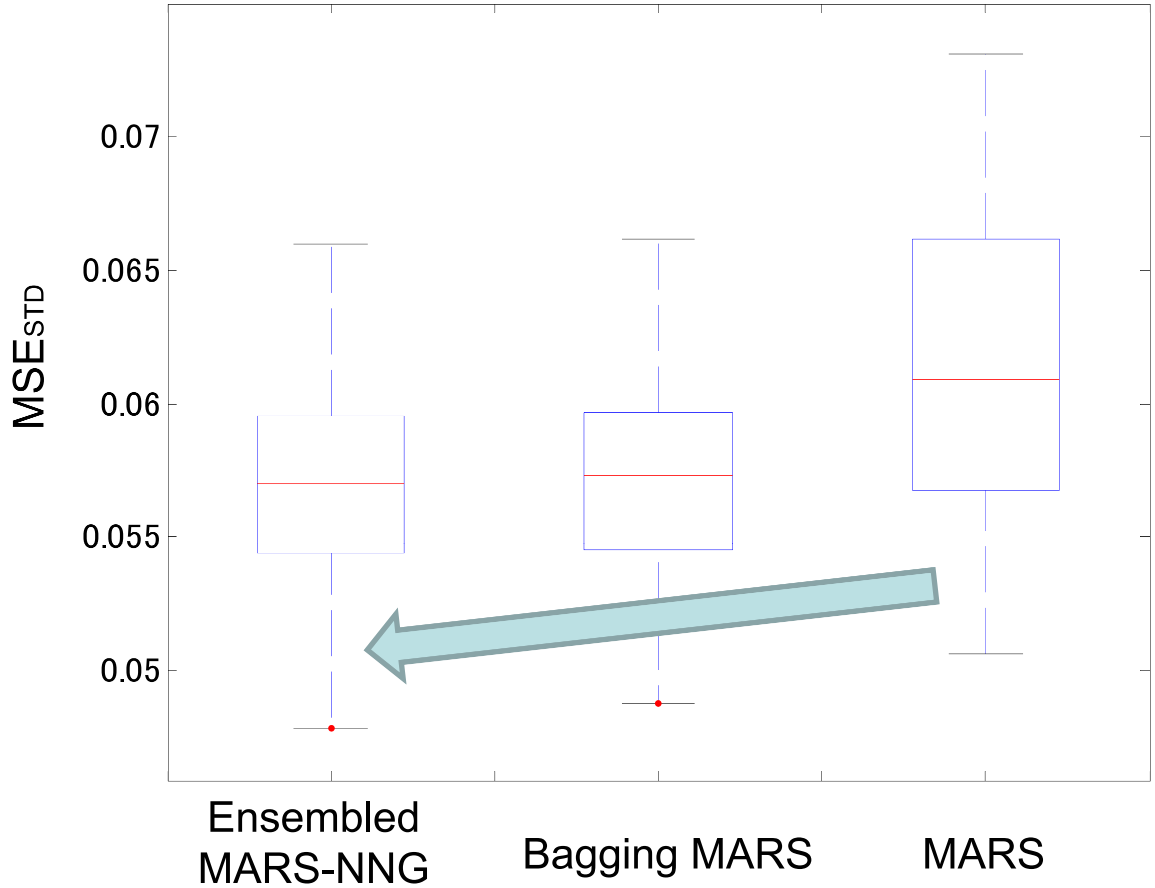
Small simulation

- Design
  - Model(Friedman,1991)

$$y = 10\sin(\pi x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5 + \varepsilon, \quad \text{where } \varepsilon \text{ is } N(0,1).$$

  - Training sample size: 100
  - Testing sample size: 1,000
  - Number of simulation: 100
- Method
  - MARS, Bagging MARS, Ensembled MARS-NNG
- Evaluation
  - MSE_{STD}(Standardized mean square error)

Small simulation



Number of trees	11.6 (averaged)	100	1

Agenda

- Introduction and motivation
- Previous research
  - Multivariate Adaptive Regression Splines(MARS)
  - Bagging MARS
- Our method proposed
  - Ensembled MARS with non-negative garrote
- Example and simulation
- Concluding remarks

Concluding remarks

- We proposed a new ensembled method of MARS.
  - Our method proposed is stable and interpretable.

- Ensembled MARS-NNG provided superior or comparable results to MARS and Bagging MARS.

References

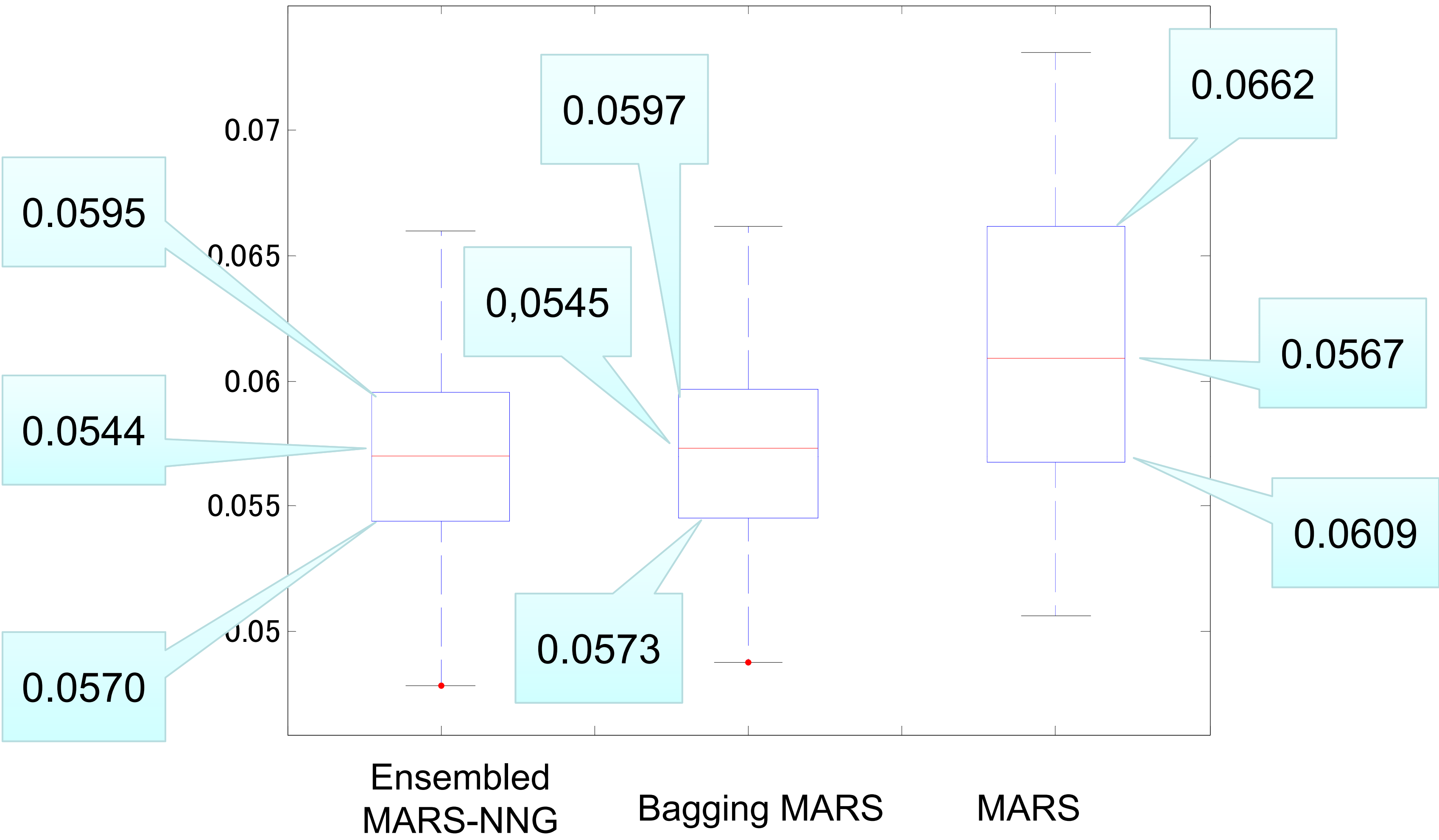
- Breiman, L. (1995) Better subset regression using the nonnegative garrote. *Technometrics*, **37**, 373-384.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, **24**, 123-140.
- Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1984). *Classification And Regression Trees*. Wadsworth.
- Friedman, J. H. (1991) Multivariate Adaptive Regression Splines (with discussion). *Annals of Statistics,***19**, 1-141.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Ann. Statist*., **29(5),** 1189-1232.
- Meinshausen, N. (2009): Node Harvest: simple and interpretable regression and classification. *Arxiv preprint arXiv*:0910.2145.
- Motogaito, H., Sugimoto, T. & Goto, M. (2007): Multivariate Adaptive Regression Splines with Non-negative Garrote Estimator. *Japanese J. Appl. Statist.*, **36**, 99-118 (in Japanese).
- Yuan, M. & Lin, Y. (2007) On the non-negative garrote estimator. *J. R. Statist. Soc., B* **69(2)**, 143-161.

Thank you very much for your attention

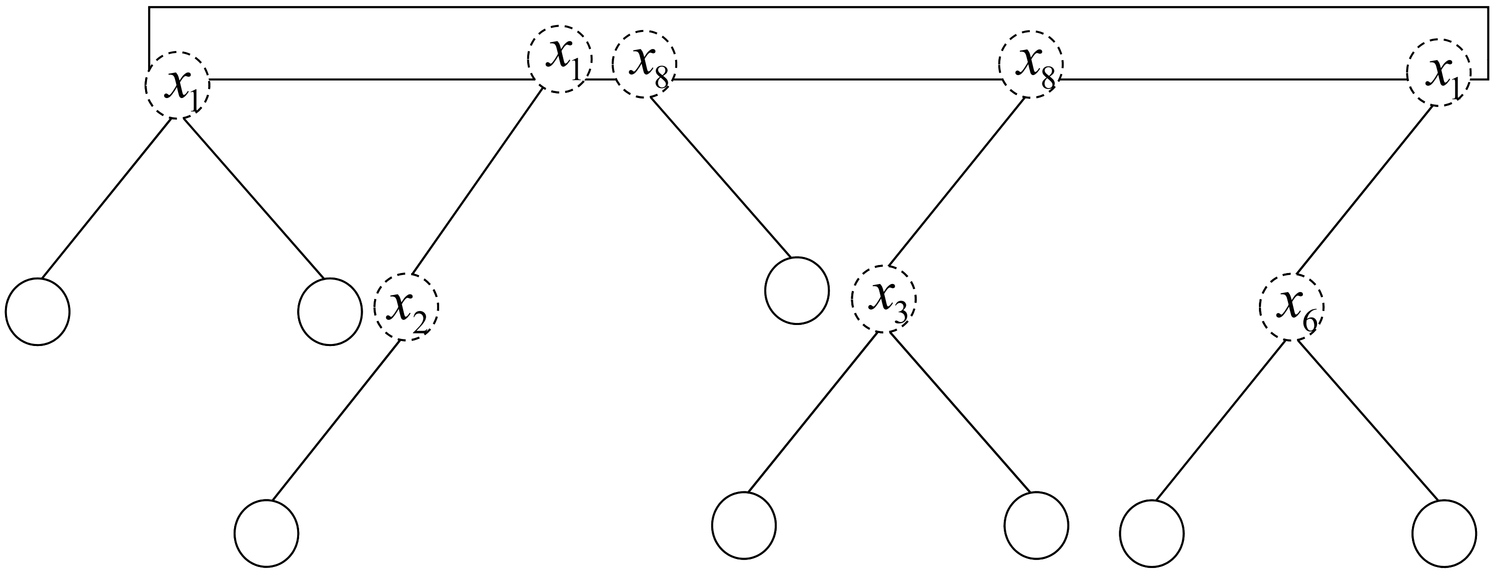
h-motogt@sigmath.es.osaka-u.ac.jp

Back up

Small simulation

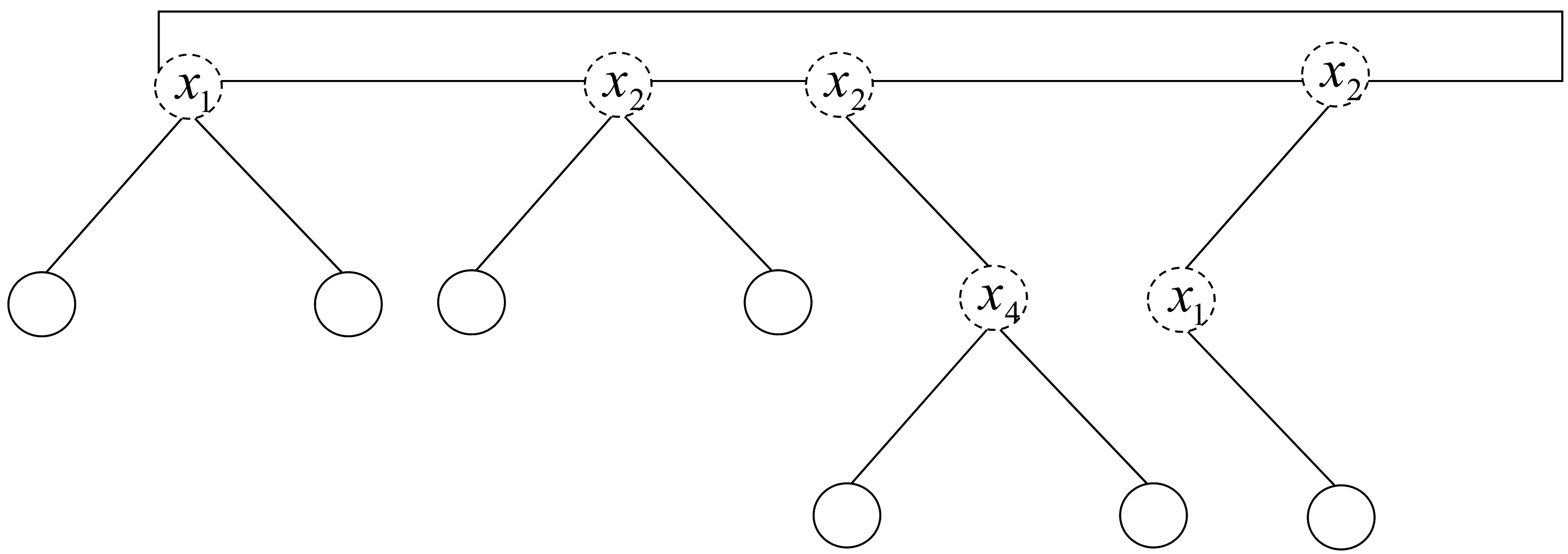


Literature example



MARS

Literature example



Ensembled
MARS-NNG