Bias–Variance decomposition of the error in bagged and random aggregated ensemble of SVMs, while varying the cardinality of the data: results and graphics.

Giorgio Valentini
DSI - Dipartimento di Scienze dell' Informazione
Università degli Studi di Milano
e-mail: valentini@dsi.unimi.it

Contents

1 Single, bagged and random aggregated SVMs: gaussian kernel

1.1 Experimental setup

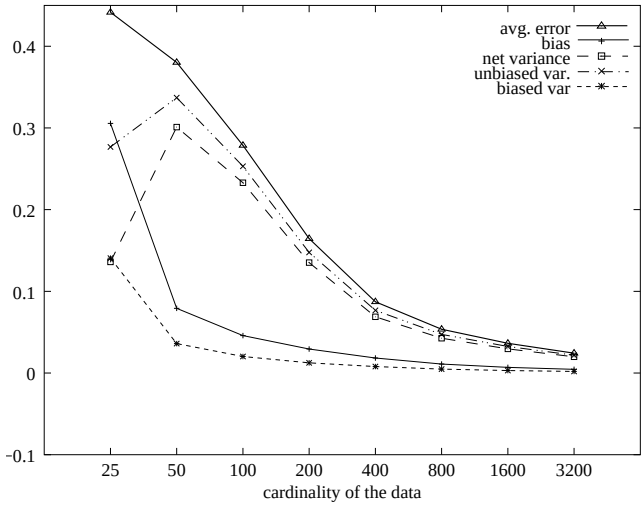
We evaluated the bias–variance decomposition of the error with respect of the cardinality of the training samples using the synthetic P2 data set, considering single, bagged, random aggregated (RA) SVMs.

Characteristics of the experiments:

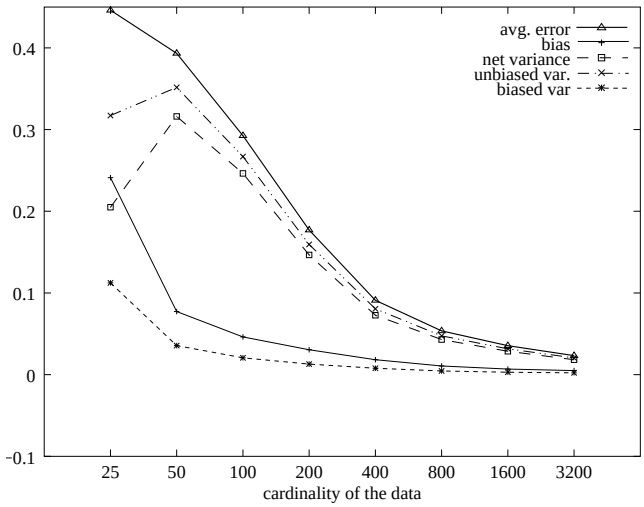
- Bias-variance decomposition of the error evaluated with 200 ensembles (bagging and random aggregating), each one composed by 60 base learners (gaussian SVMs), and with 200 SVMs (single SVMs) for each different sample size of the training sets.
- Single and ensembles of SVMs trained with samples of the following cardinality: $25, 50, 100, 200, 400, 800, 1600, 3200$.
- BV decomposition of the error evaluated on a separated test set composed by 10000 examples.
- We evaluated only the three pairs of values of σ and C:
  - $\sigma = 0.1, C = 100$.
  - $\sigma = 0.5, C = 100$.
  - $\sigma = 5, C = 100$.

1.2 BV decomposition in single, bagged and random aggregated ensembles while varying the cardinality

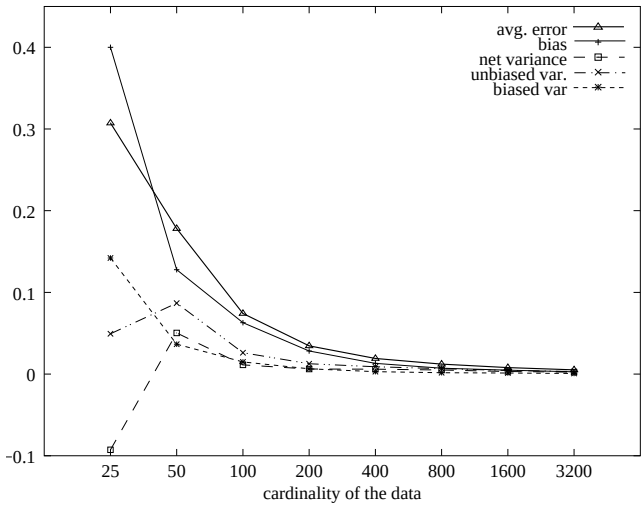
Each figure shows the BV decomposition in single, bagged and random aggregated ensembles while varying the cardinality.



(a)

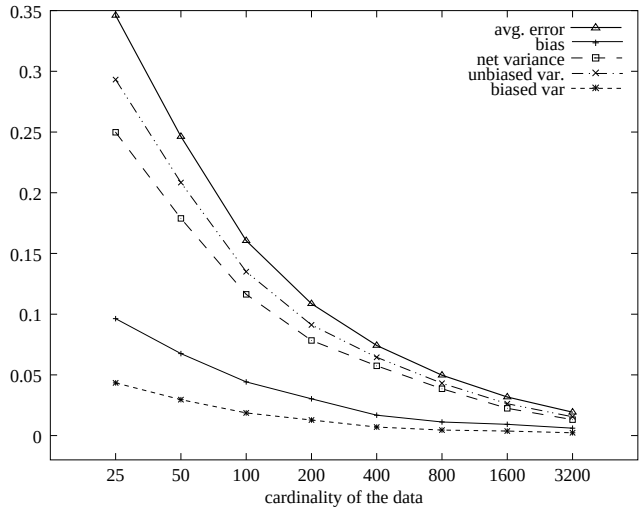


(b)

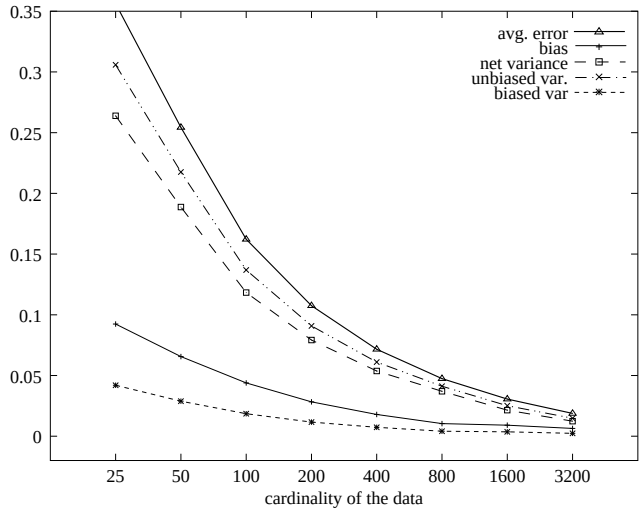


(c)

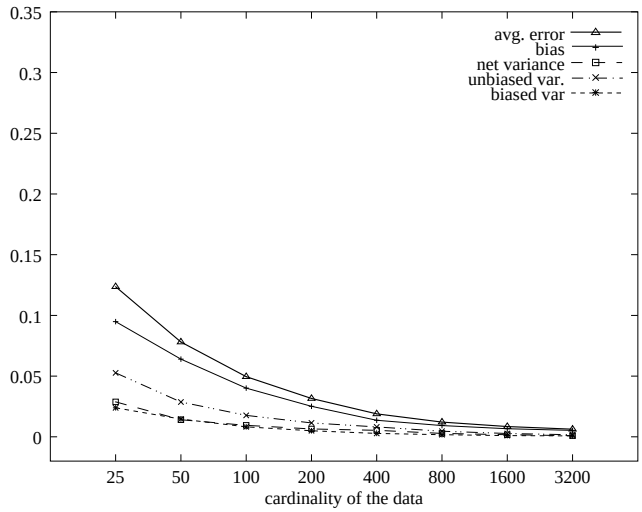
Figure 1: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in single, bagged and random aggreated gaussian SVMs, while varying the cardinality of the data, with $\sigma = 0.1$ and $C = 100$. (a) Single SVM (b) Bagged SVMs (c) Random aggregated SVMs.



(a)

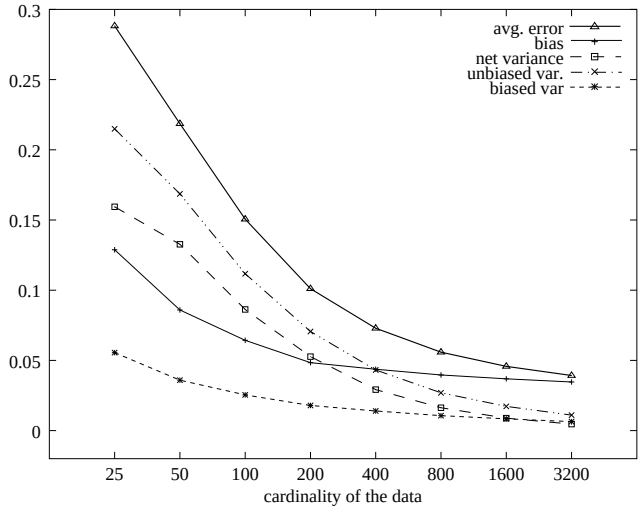


(b)

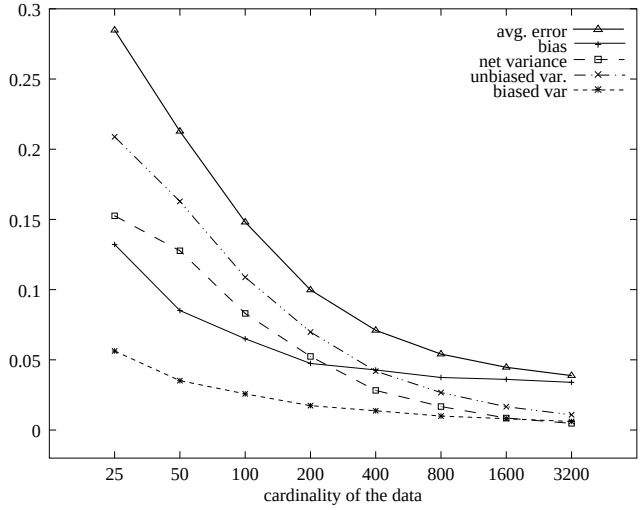


(c)

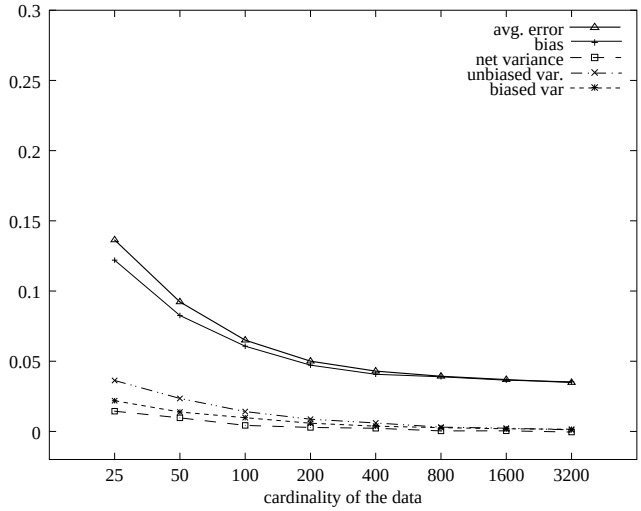
Figure 2: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in single, bagged and random aggregated gaussian SVMs, while varying the cardinality of the data, with $\sigma = 0.5$ and $C = 100$. (a) Single SVM (b) Bagged SVMs (c) Random aggregated SVMs.



(a)



(b)

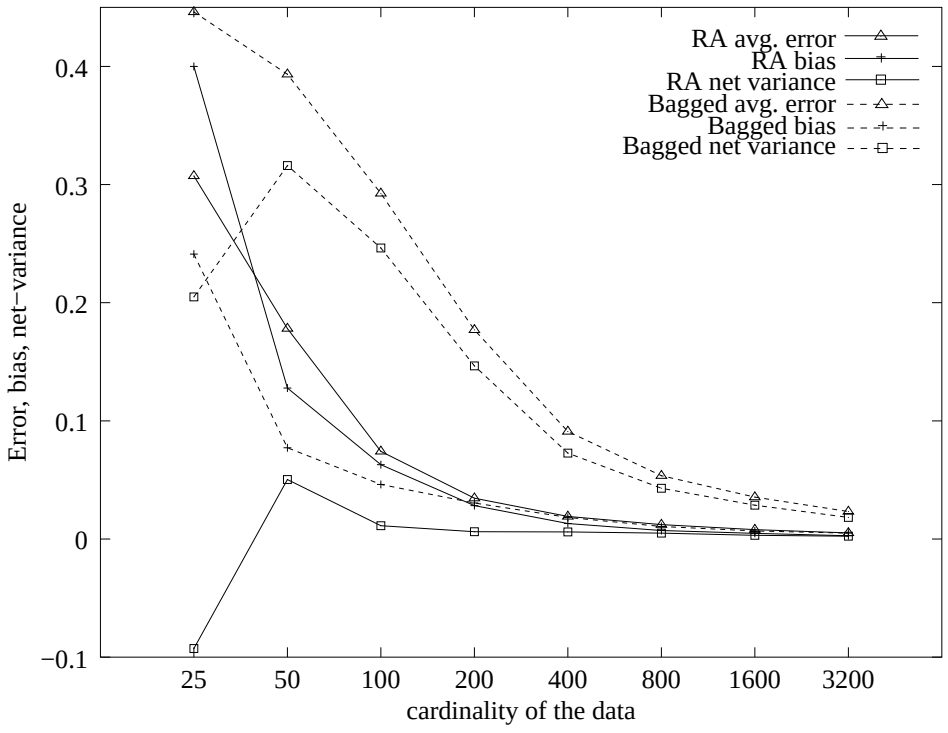


(c)

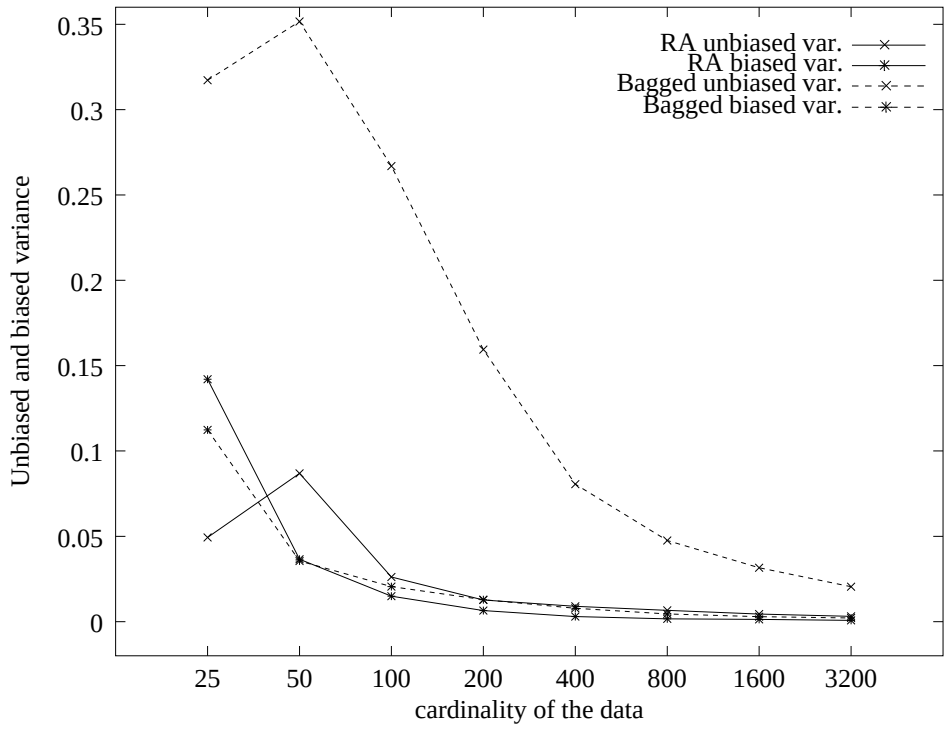
Figure 3: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in single, bagged and random aggreated gaussian SVMs, while varying the cardinality of the data, with $\sigma = 5$ and $C = 100$. (a) Single SVM (b) Bagged SVMs (c) Random aggregated SVMs.

1.3 Comparison of BV decomposition in single, bagged and random aggregated ensembles while varying the cardinality

Each figure shows in the same graph the BV decomposition in random aggregated (RA) and bagged ensembles of SVMs. Continuous lines: RA; dashed lines:bagged ensembles.

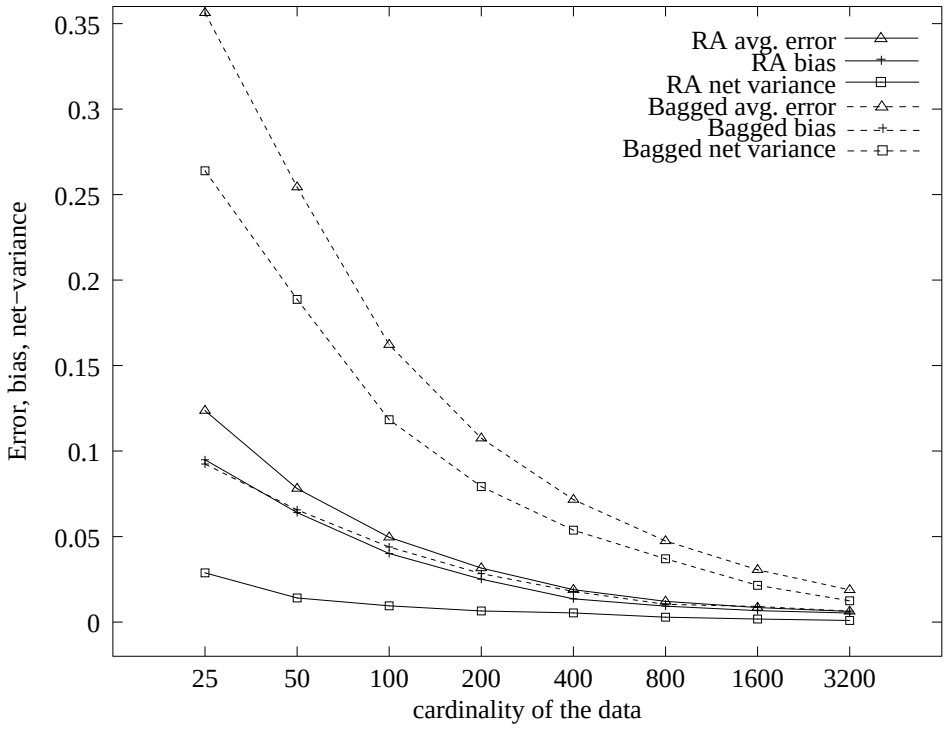


(a)

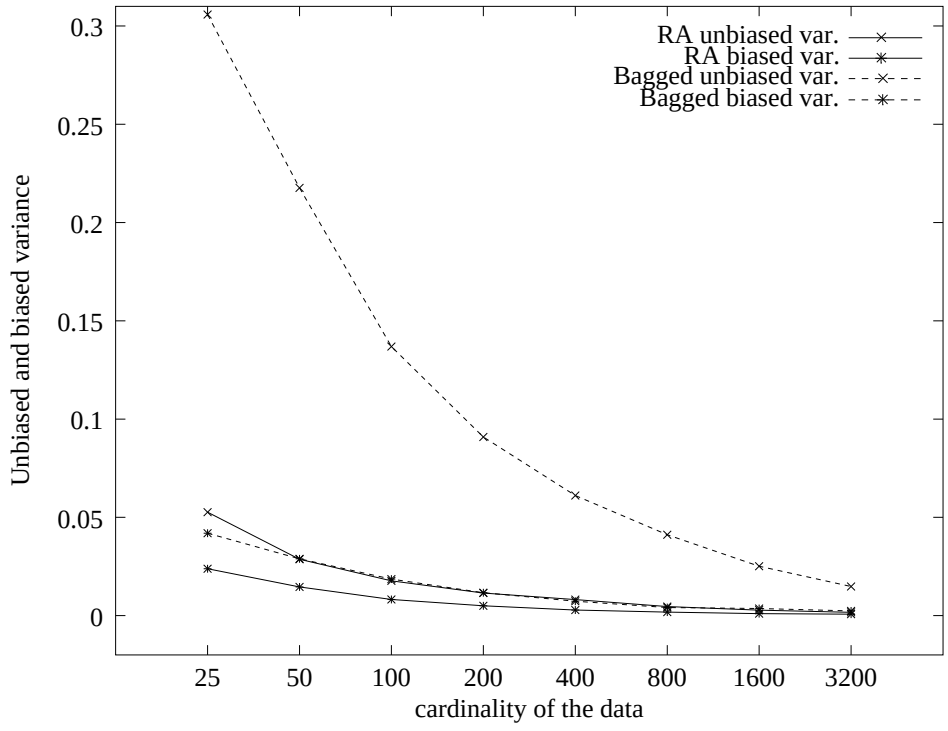


(b)

Figure 4: Comparing bias-variance decomposition of the error in bagged and random aggreated SVMs, while varying the cardinality of the data, with $\sigma = 0.1$ and $C = 100$. (a) Comparing average error, bias and net variance (b) Comparing unbiased and biased variance.

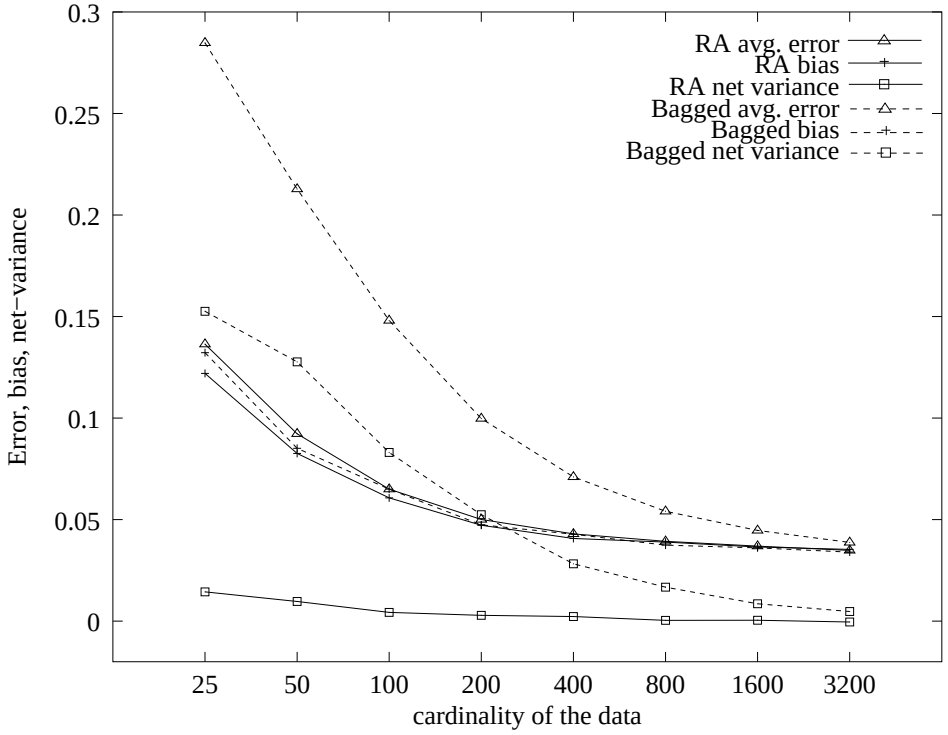


(a)

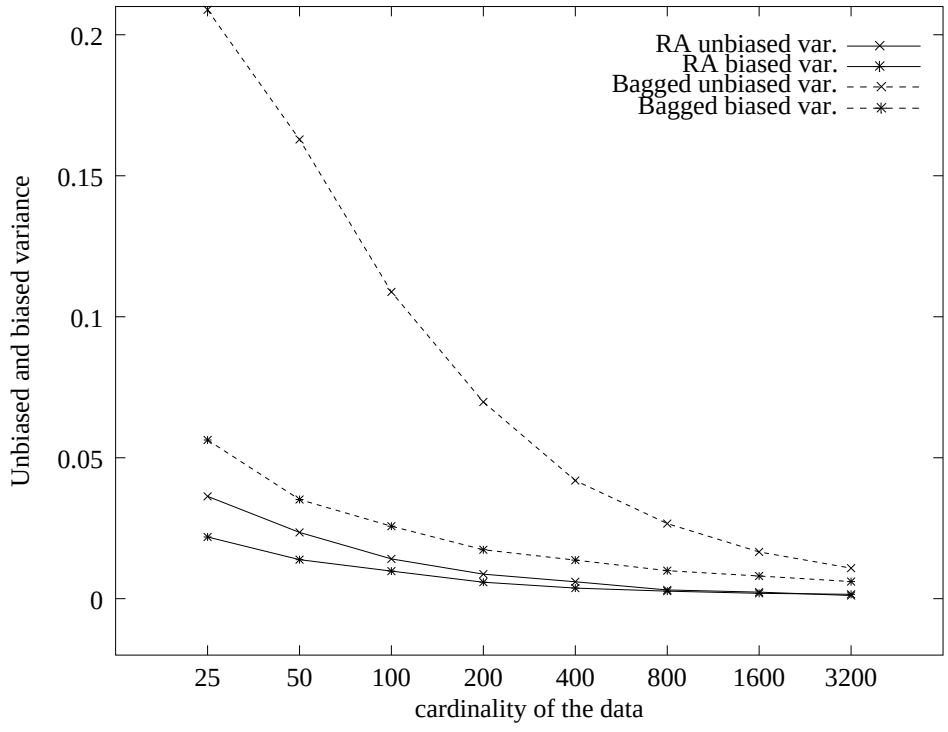


(b)

Figure 5: Comparing bias-variance decomposition of the error in bagged and random aggreated SVMs, while varying the cardinality of the data, with $\sigma = 0.5$ and $C = 100$. (a) Comparing average error, bias and net variance (b) Comparing unbiased and biased variance.



(a)



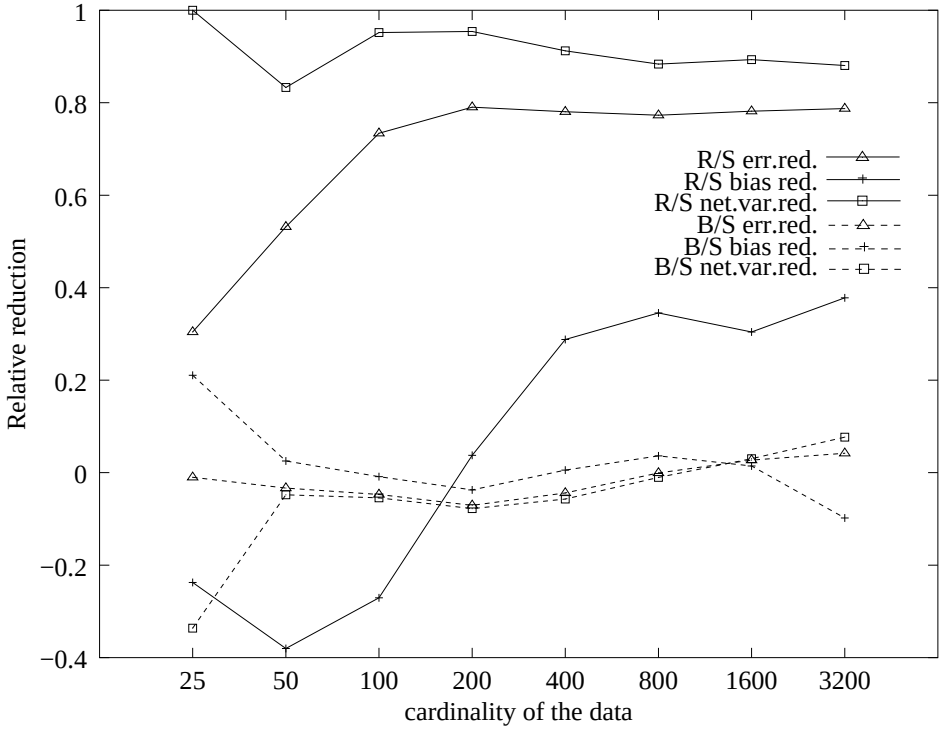
(b)

Figure 6: Comparing bias-variance decomposition of the error in bagged and random aggreated SVMs, while varying the cardinality of the data, with $\sigma = 5$ and $C = 100$. (a) Comparing average error, bias and net variance (b) Comparing unbiased and biased variance.

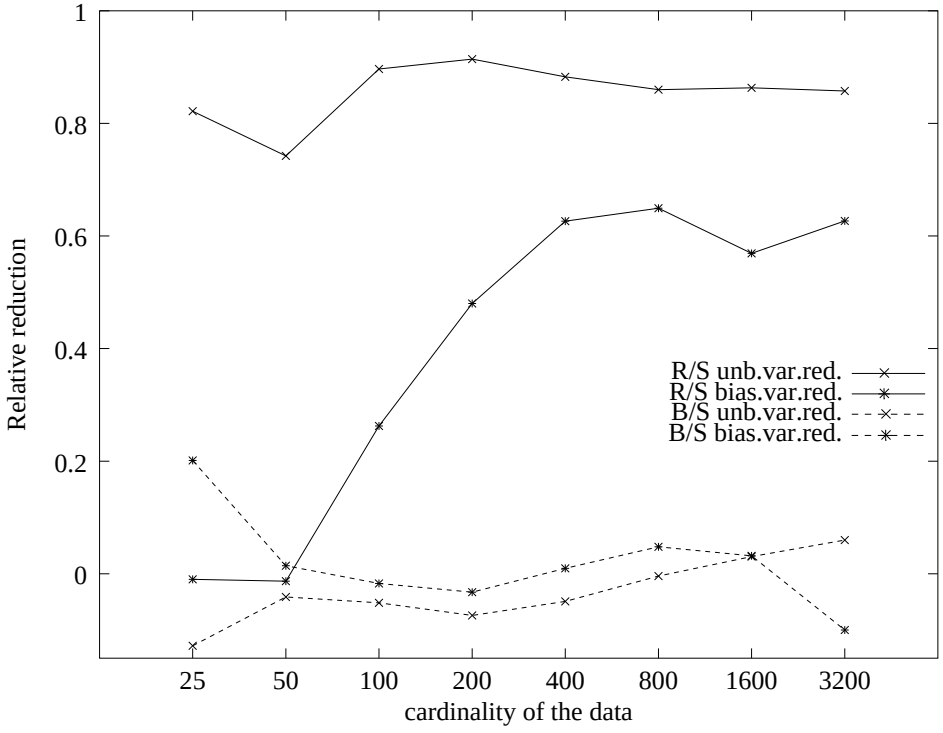
1.4 Comparison of the relative error, bias and variance reduction in RA and bagged ensembles of SVMs with respect to single SVMs

Each figure shows in the same graph the relative error, bias and variance reduction in RA and bagged ensembles of SVMs with respect to single SVM. For instance, the relative error reduction for bagging is computed in the following way:

$$\text{Relative error reduction} = \frac{\text{Single SVM error} - \text{Bagging error}}{\text{Single SVM error}}$$

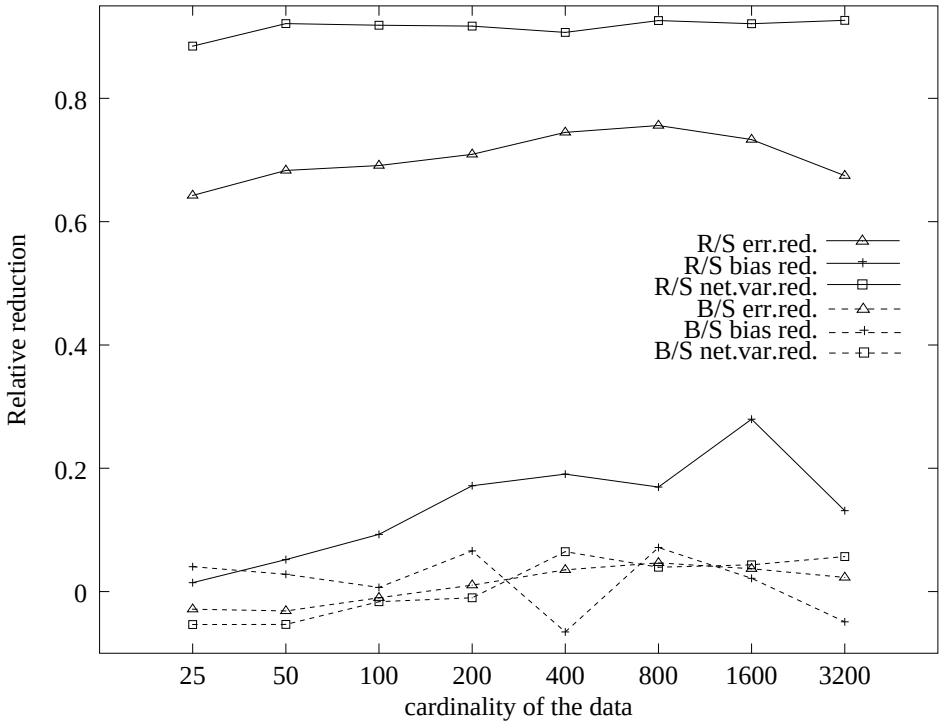


(a)

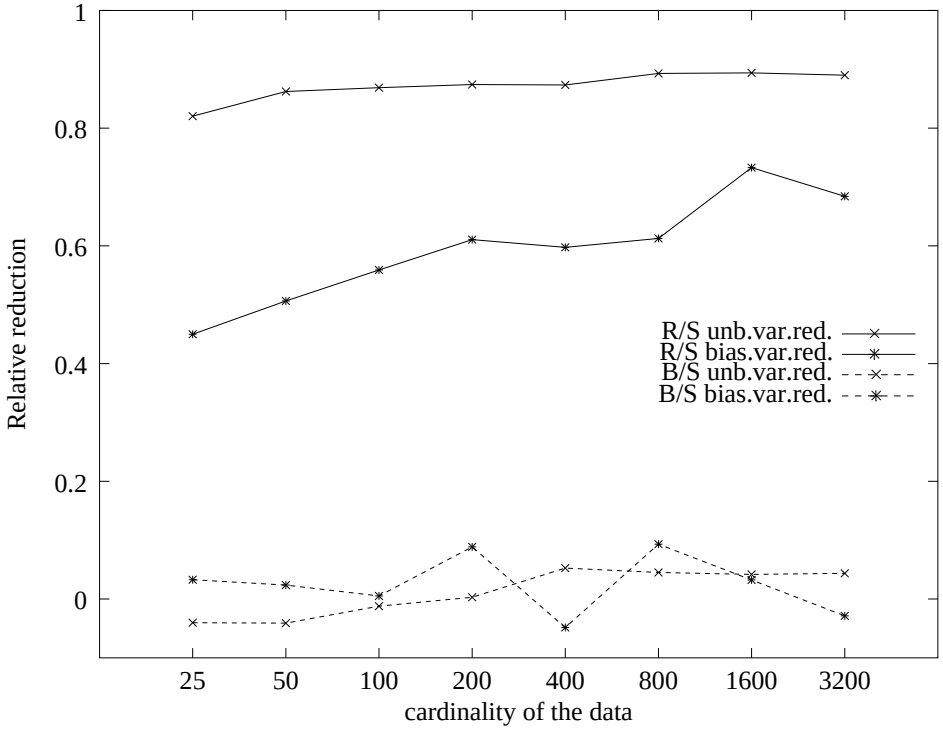


(b)

Figure 7: Comparing relative error and bias-variance reduction in bagged and random aggreated SVMs with respect to single SVMs, while varying the cardinality of the data, with $\sigma = 0.1$ and $C = 100$. Continuous lines refer to random aggregated SVMs, dashed lines to bagged SVMs. R/S stands for Random aggregated vs. single SVMs and B/S bagged vs. single. Negative values indicate better results of single SVMs. (a) Comparing relative reduction of the error, bias and net variance (b) Comparing relative unbiased and biased variance reduction.

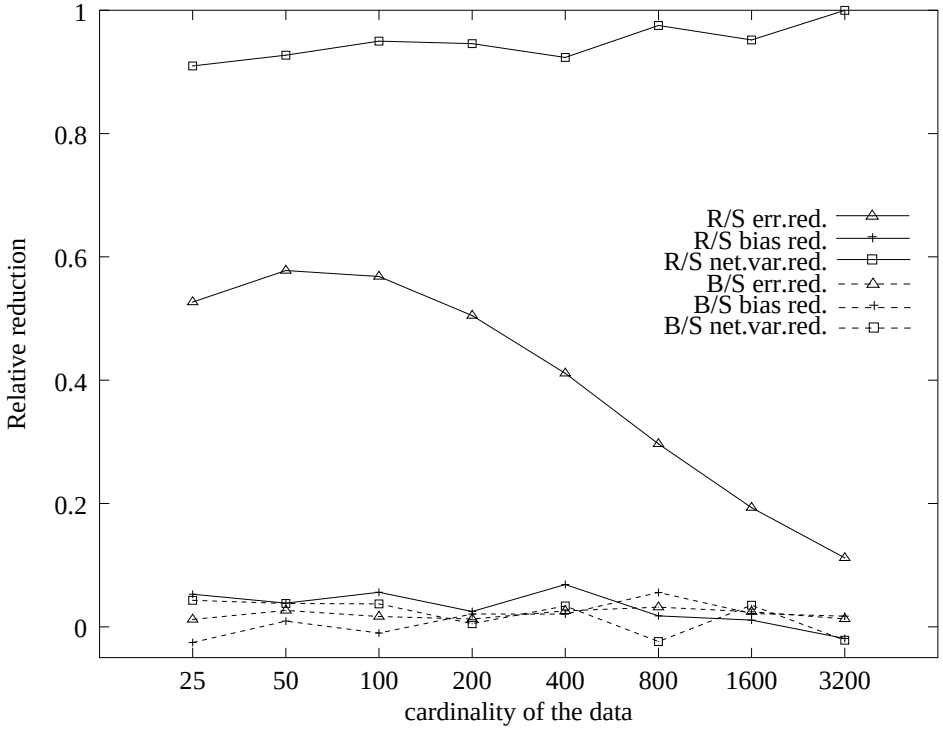


(a)

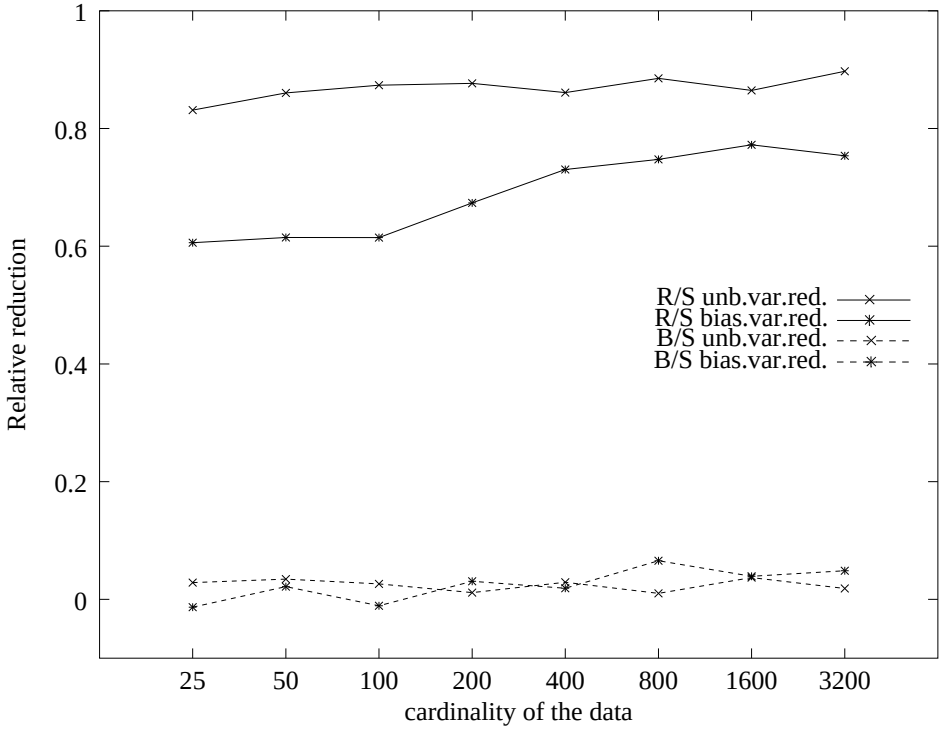


(b)

Figure 8: Comparing relative error and bias-variance reduction in bagged and random aggreated SVMs with respect to single SVMs, while varying the cardinality of the data, with $\sigma = 0.5$ and $C = 100$. Continuous lines refer to random aggregated SVMs, dashed lines to bagged SVMs. R/S stands for Random aggregated vs. single SVMs and B/S bagged vs. single. Negative values indicate better results of single SVMs. (a) Comparing relative reduction of the error, bias and net variance (b) Comparing relative unbiased and biased variance reduction.



(a)



(b)

Figure 9: Comparing relative error and bias-variance reduction in bagged and random aggreated SVMs with respect to single SVMs, while varying the cardinality of the data, with $\sigma = 5$ and $C = 100$. Continuous lines refer to random aggregated SVMs, dashed lines to bagged SVMs. R/S stands for Random aggregated vs. single SVMs and B/S bagged vs. single. Negative values indicate better results of single SVMs. (a) Comparing relative reduction of the error, bias and net variance (b) Comparing relative unbiased and biased variance reduction.

2 Single, bagged and random aggregated SVMs: polynomial kernel

2.1 Experimental setup

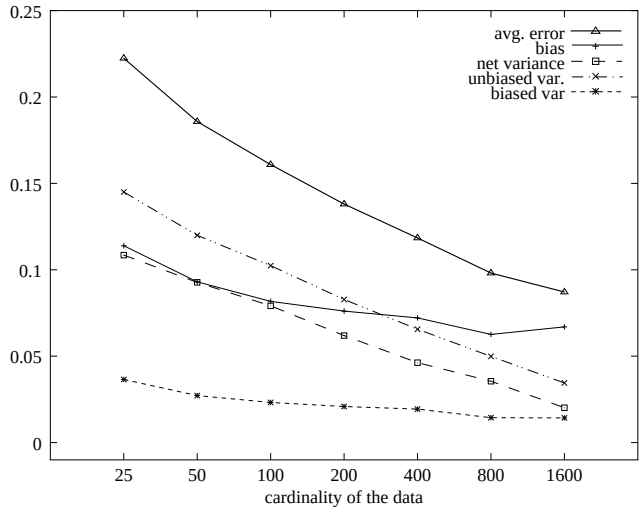
We evaluated the bias–variance decomposition of the error with respect of the cardinality of the training samples using the Spam data set (UCI repository), considering single, bagged, random aggregated (RA) SVMs.

Characteristics of the experiments:

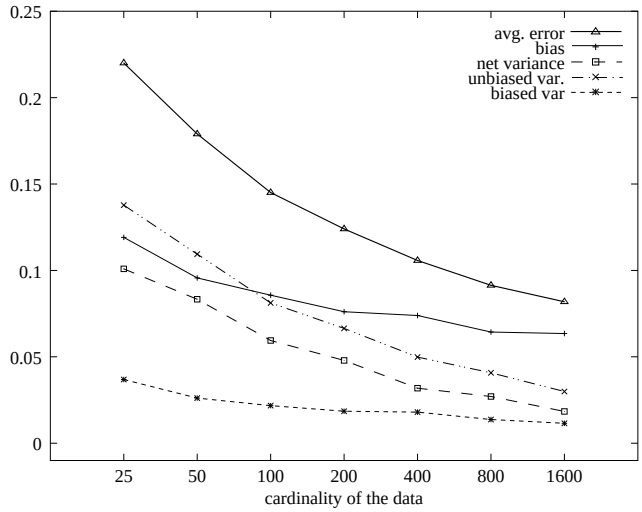
- Bias-variance decomposition of the error evaluated with 200 ensembles (bagging and random aggregating), each one composed by 60 base learners (gaussian SVMs), and with 200 SVMs (single SVMs) for each different sample size of the training sets.
- Single and ensembles of SVMs trained with samples of the following cardinality: $25, 50, 100, 200, 400, 800, 1600$.
- BV decomposition of the error evaluated on a separated test set composed by 2300 examples.
- We evaluated only the three pairs of values of the polynomial *degree* and C:
  - $degree = 2, C = 100$.
  - $degree = 3, C = 100$.
  - $degree = 4, C = 100$.

2.2 BV decomposition in single, bagged and random aggregated ensembles while varying the cardinality

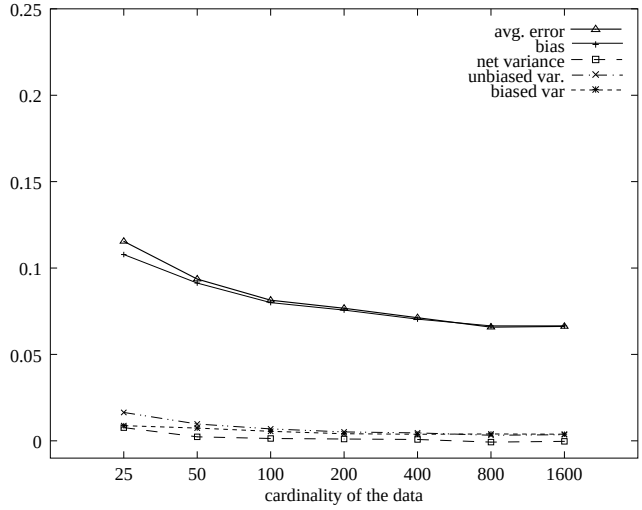
Each figure shows the BV decomposition in single, bagged and random aggregated ensembles while varying the cardinality.



(a)

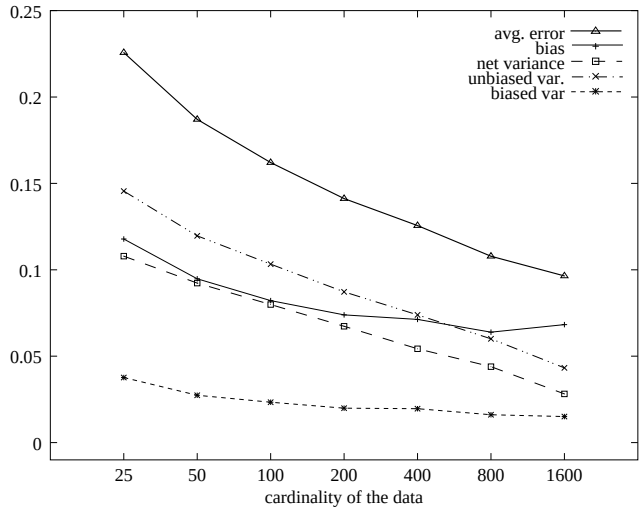


(b)

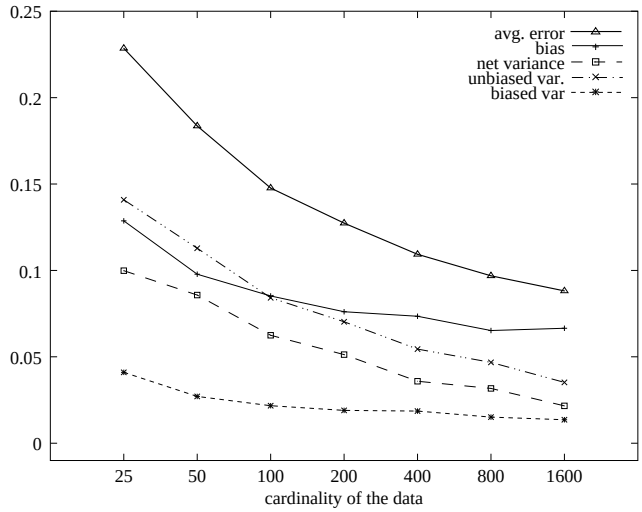


(c)

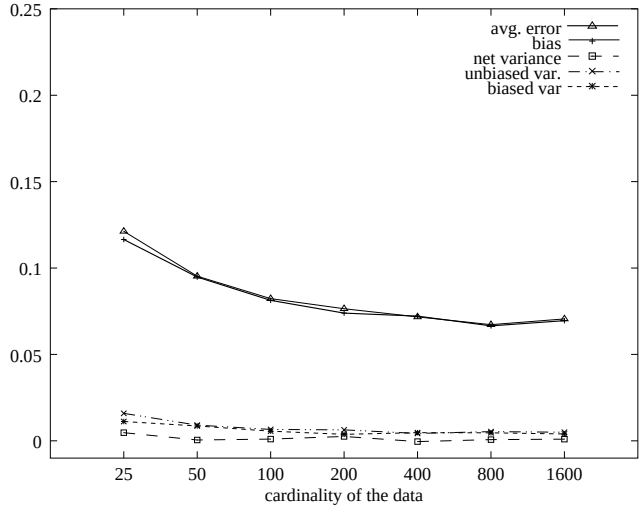
Figure 10: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in single, bagged and random aggregated polynomial SVMs, while varying the cardinality of the data, with degree=2 and $C = 100$. (a) Single SVM (b) Bagged SVMs (c) Random aggregated SVMs.



(a)

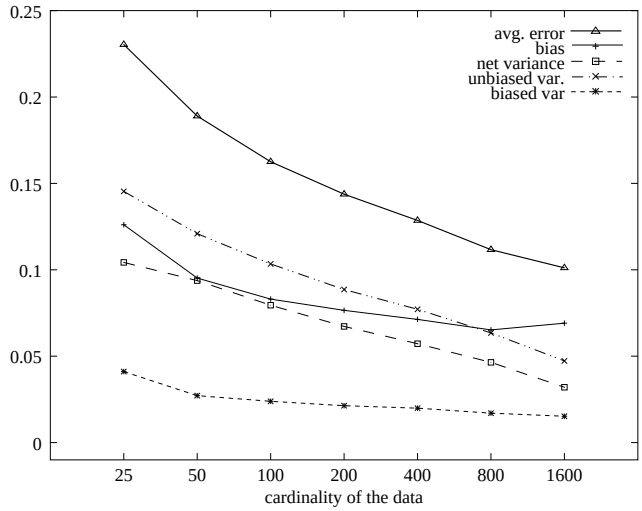


(b)

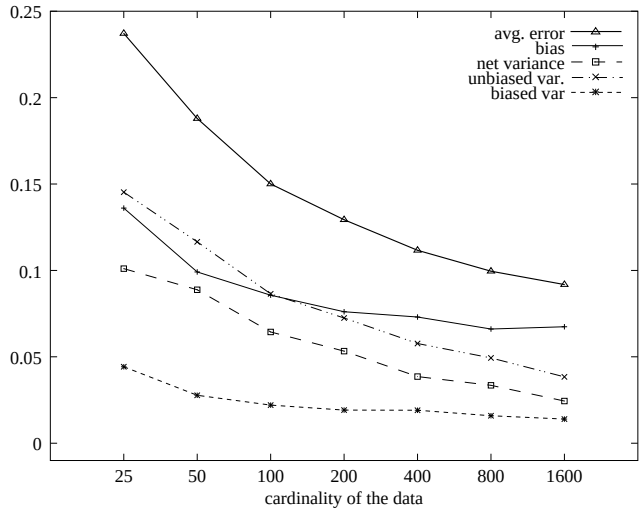


(c)

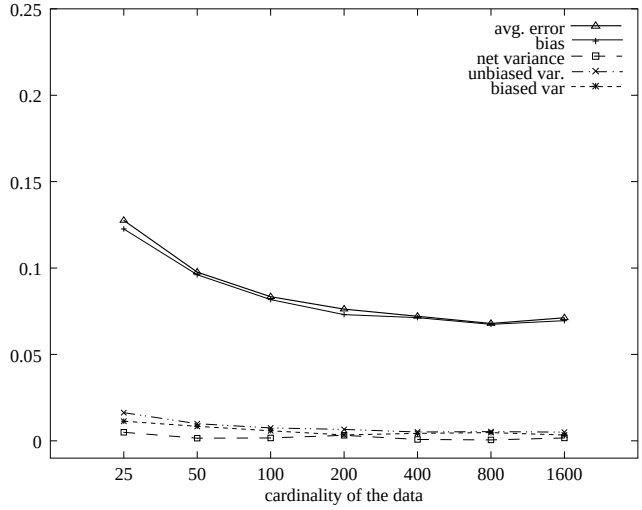
Figure 11: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in single, bagged and random aggregated polynomial SVMs, while varying the cardinality of the data, with degree=3 and $C = 100$. (a) Single SVM (b) Bagged SVMs (c) Random aggregated SVMs.



(a)



(b)

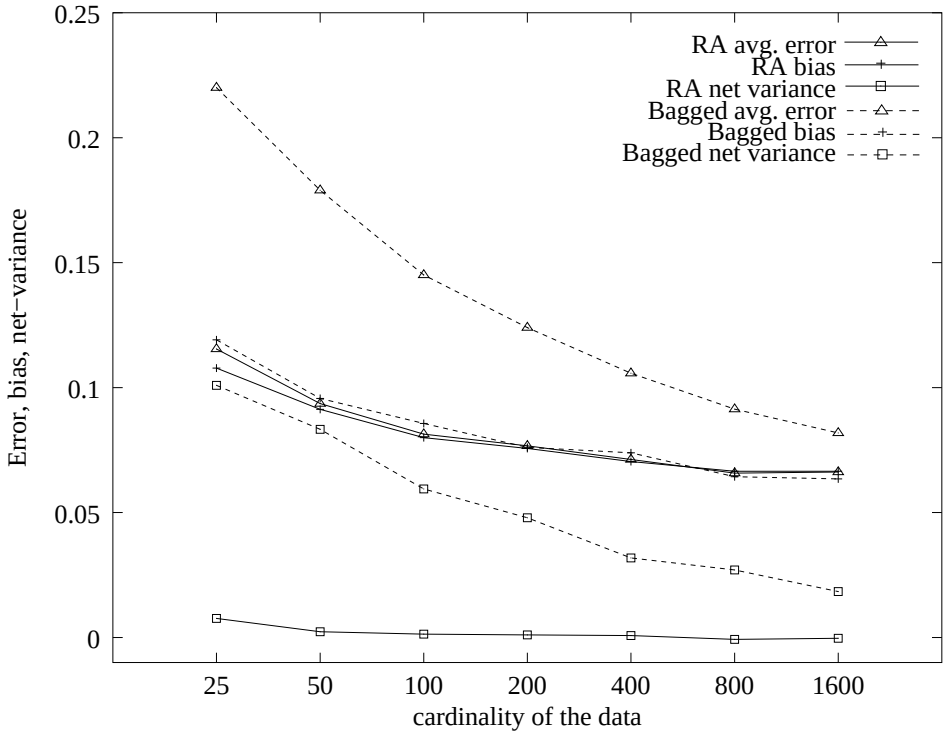


(c)

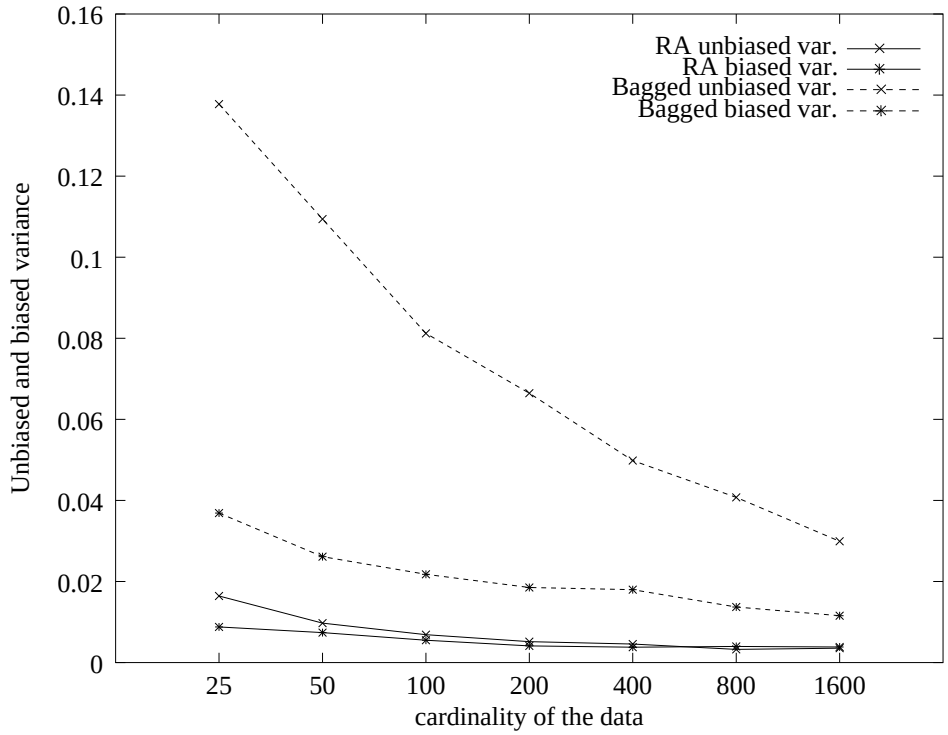
Figure 12: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in single, bagged and random aggregated polynomial SVMs, while varying the cardinality of the data, with degree=4 and $C = 100$. (a) Single SVM (b) Bagged SVMs (c) Random aggregated SVMs.

2.3 Comparison of BV decomposition in single, bagged and random aggregated ensembles while varying the cardinality

Each figure shows in the same graph the BV decomposition in random aggregated (RA) and bagged ensembles of SVMs. Continuous lines: RA; dashed lines:bagged ensembles.

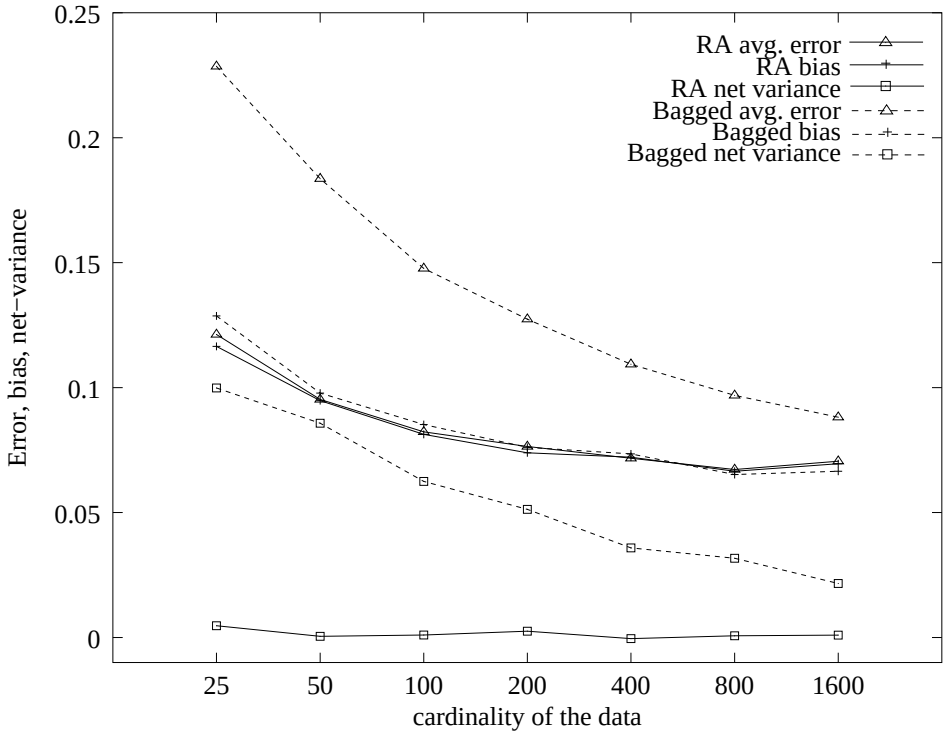


(a)

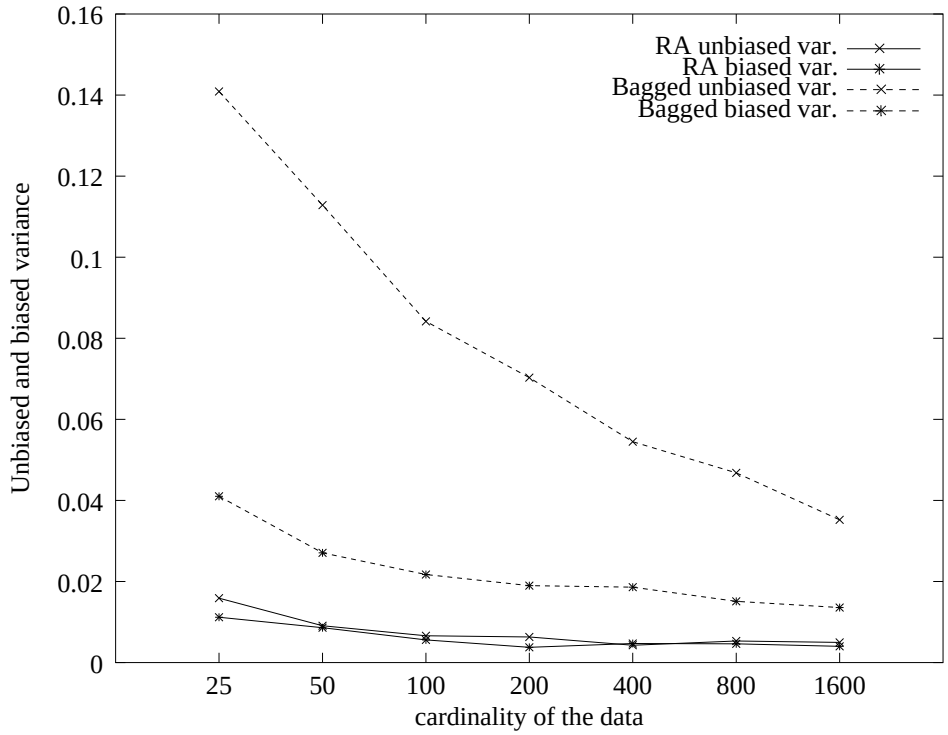


(b)

Figure 13: Comparing bias-variance decomposition of the error in polynomial bagged and random aggreated SVMs, while varying the cardinality of the data, with degree = 2 and $C = 100$. (a) Comparing average error, bias and net variance (b) Comparing unbiased and biased variance.

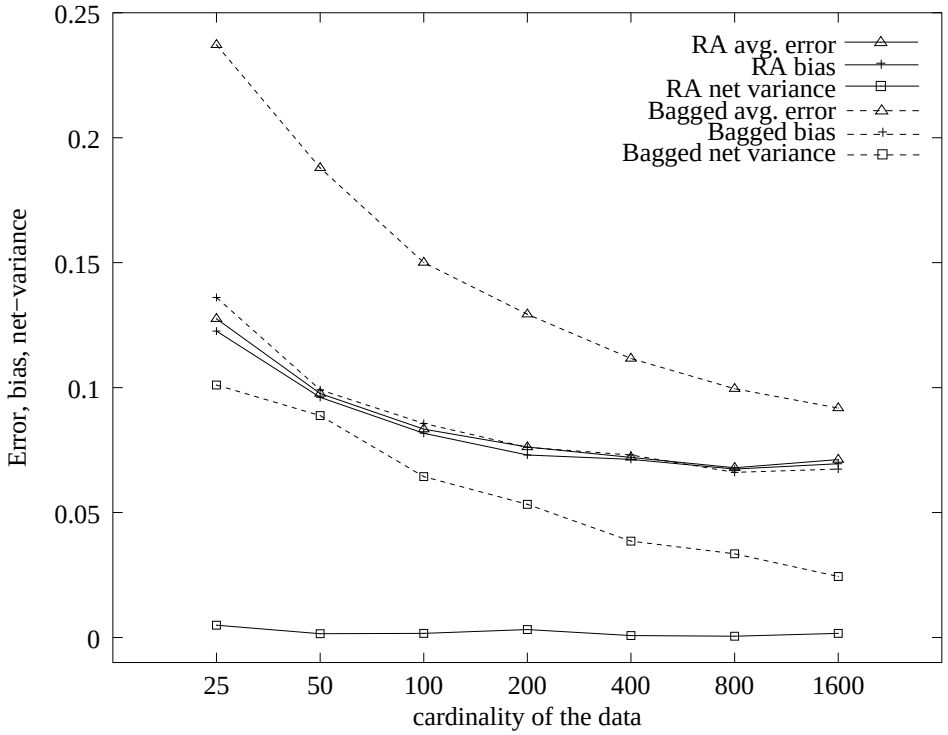


(a)

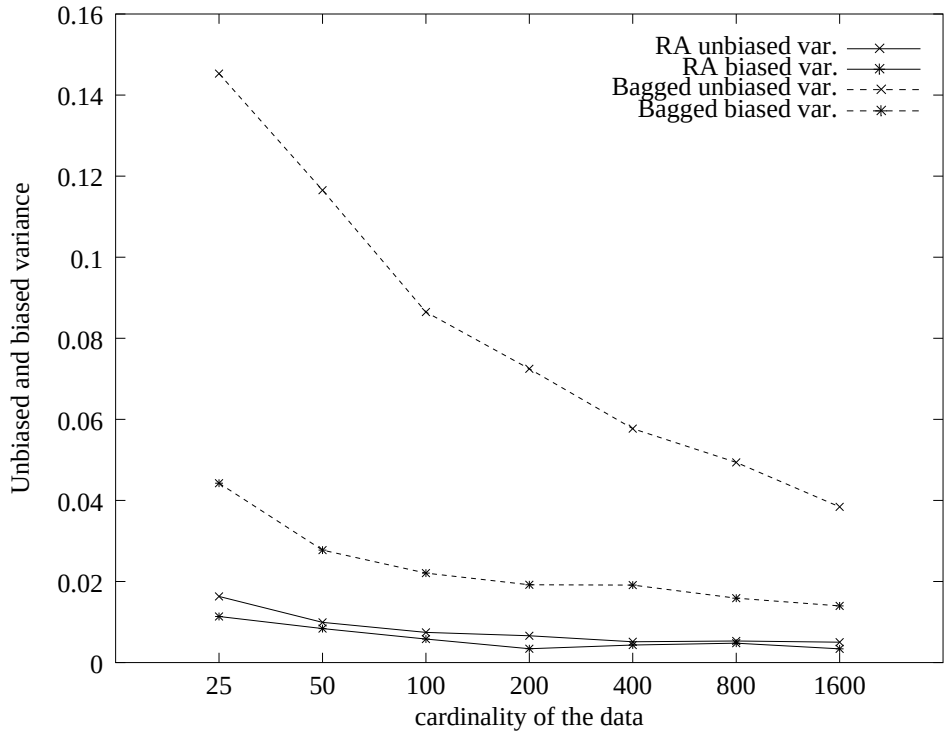


(b)

Figure 14: Comparing bias-variance decomposition of the error in polynomial bagged and random aggreated SVMs, while varying the cardinality of the data, with degree = 3 and $C = 100$. (a) Comparing average error, bias and net variance (b) Comparing unbiased and biased variance.



(a)



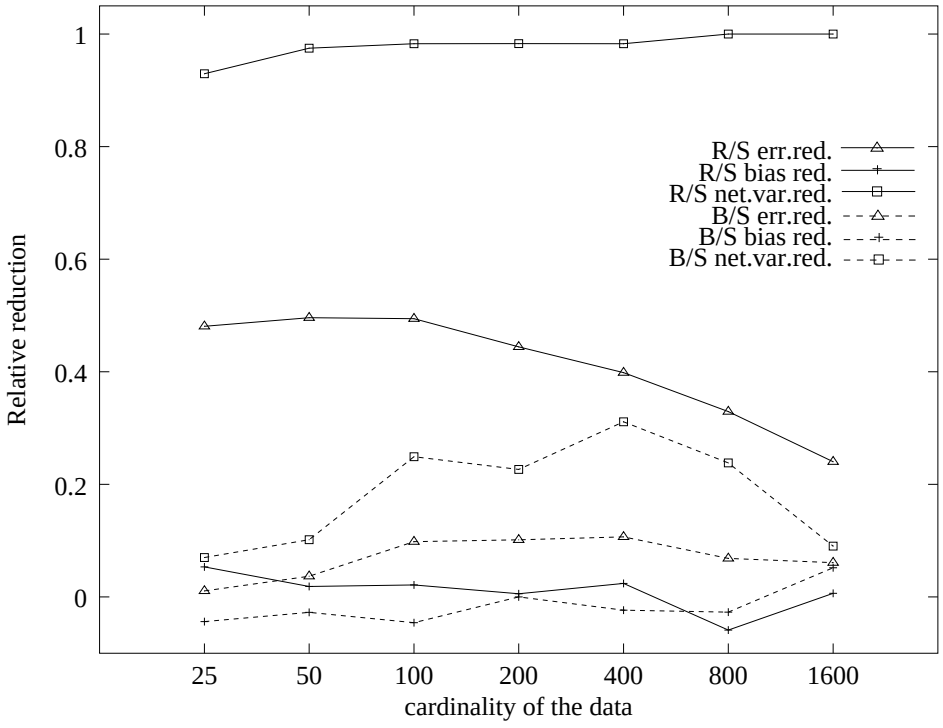
(b)

Figure 15: Comparing bias-variance decomposition of the error in polynomial bagged and random aggreated SVMs, while varying the cardinality of the data, with degree = 4 and $C = 100$. (a) Comparing average error, bias and net variance (b) Comparing unbiased and biased variance.

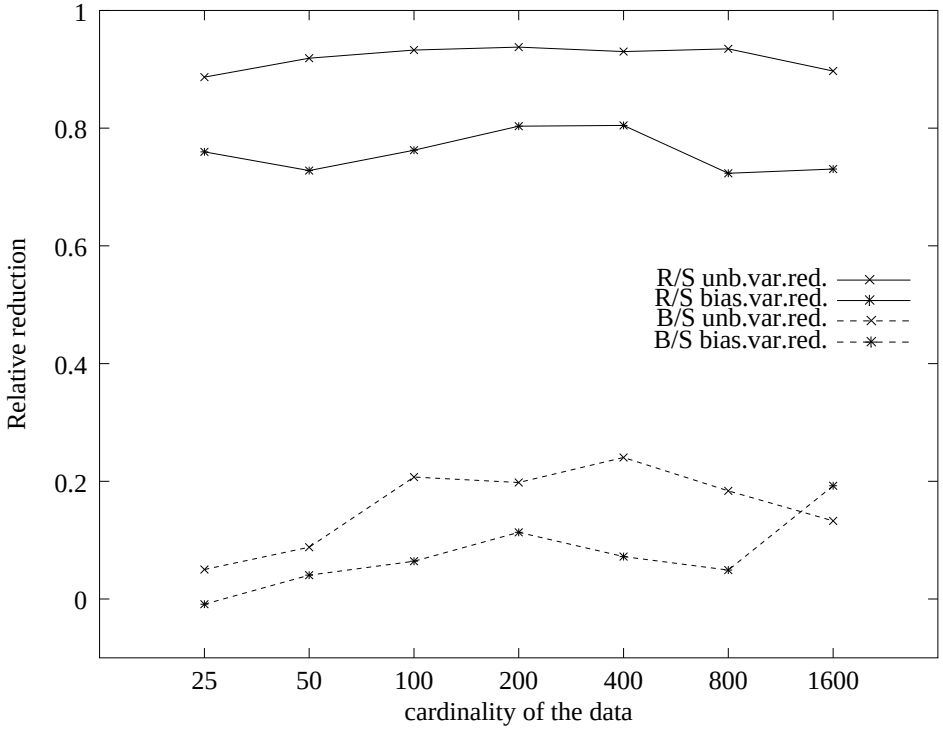
2.4 Comparison of the relative error, bias and variance reduction in RA and bagged ensembles of SVMs with respect to single SVMs

Each figure shows in the same graph the relative error, bias and variance reduction in RA and bagged ensembles of SVMs with respect to single SVM. For instance, the relative error reduction for bagging is computed in the following way:

$$\text{Relative error reduction} = \frac{\text{Single SVM error} - \text{Bagging error}}{\text{Single SVM error}}$$

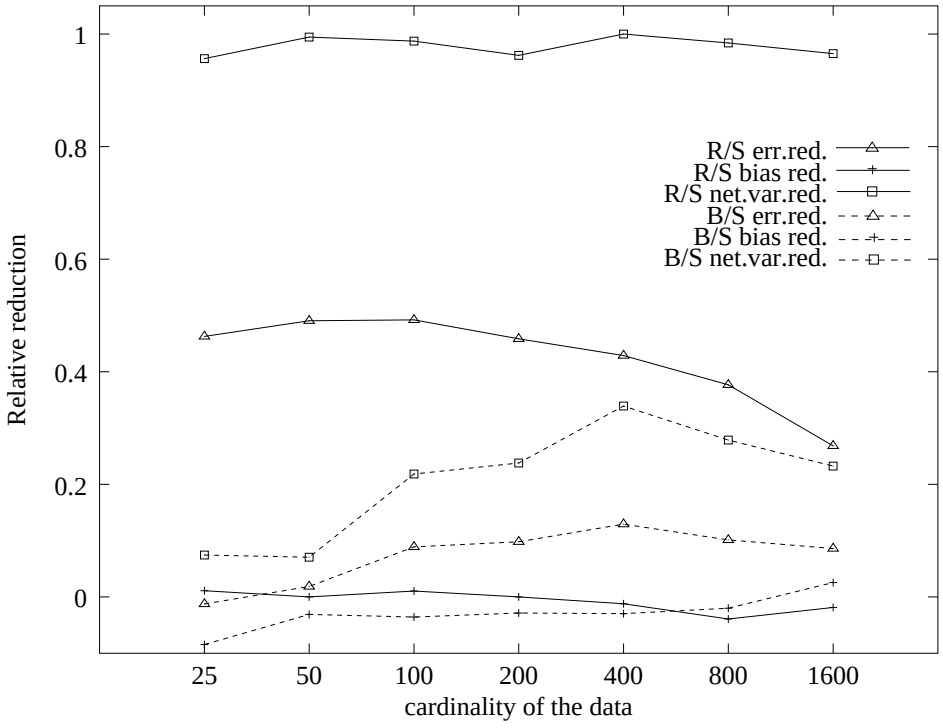


(a)

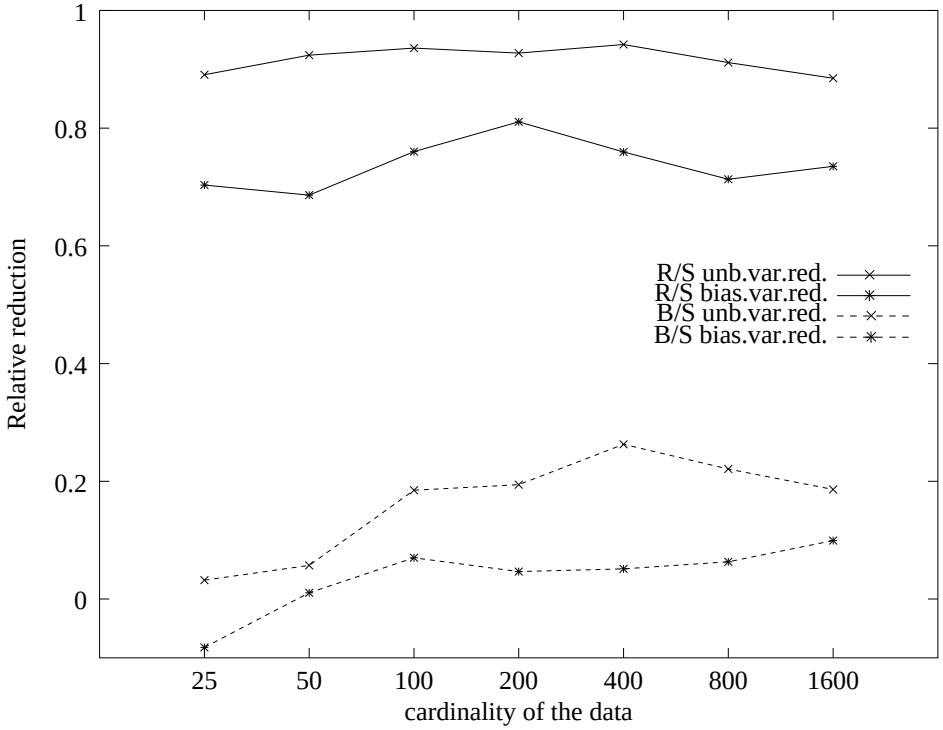


(b)

Figure 16: Comparing relative error and bias-variance reduction in polynomial bagged and random aggreated SVMs with respect to single SVMs, while varying the cardinality of the data, with degree $= 2$ and $C = 100$ with the Spam data set. Continuous lines refer to random aggregated SVMs, dashed lines to bagged SVMs. R/S stands for Random aggregated vs. single SVMs and B/S bagged vs. single. Negative values indicate better results of single SVMs. (a) Comparing relative reduction of the error, bias and net variance (b) Comparing relative unbiased and biased variance reduction.

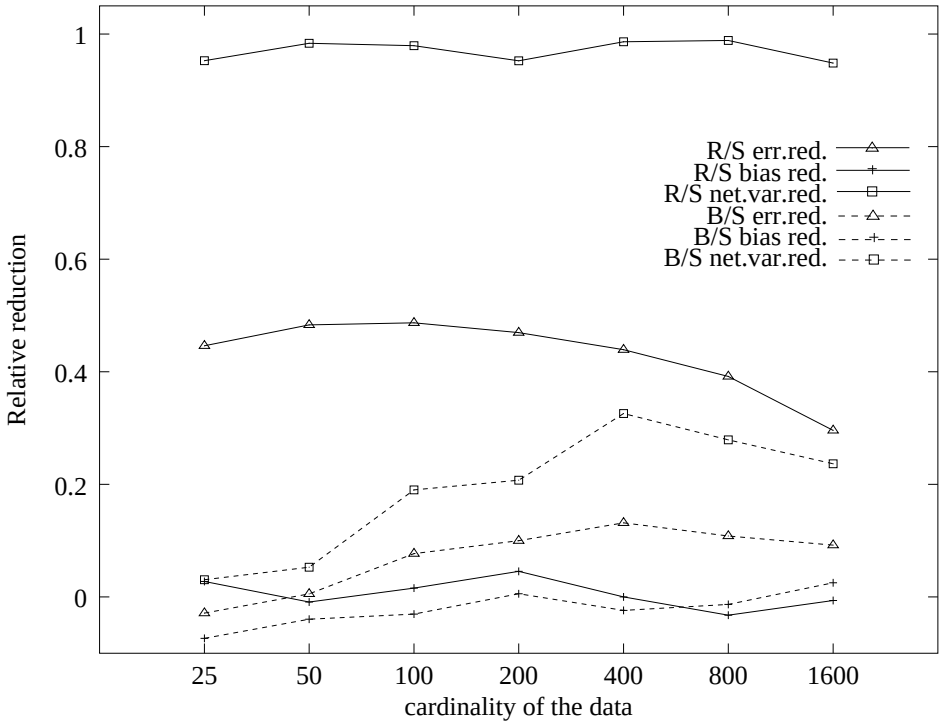


(a)

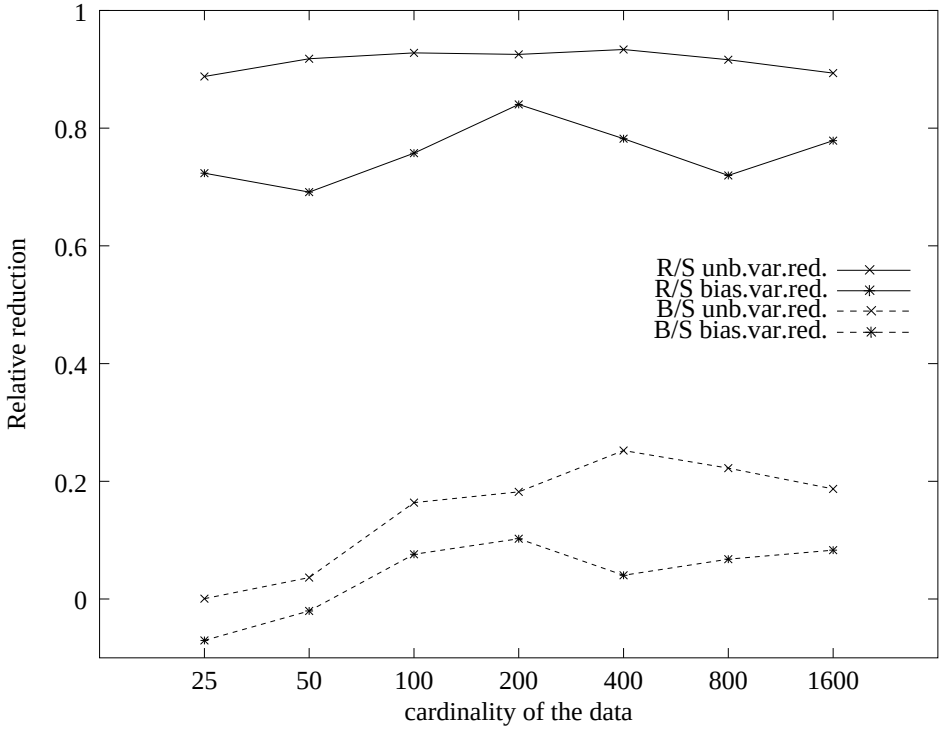


(b)

Figure 17: Comparing relative error and bias-variance reduction in polynomial bagged and random aggreated SVMs with respect to single SVMs, while varying the cardinality of the data, with degree = 3 and $C = 100$ with the Spam data set. Continuous lines refer to random aggregated SVMs, dashed lines to bagged SVMs. R/S stands for Random aggregated vs. single SVMs and B/S bagged vs. single. Negative values indicate better results of single SVMs. (a) Comparing relative reduction of the error, bias and net variance (b) Comparing relative unbiased and biased variance reduction.



(a)



(b)

Figure 18: Comparing relative error and bias-variance reduction in polynomial bagged and random aggreated SVMs with respect to single SVMs, while varying the cardinality of the data, with degree $= 4$ and $C = 100$ with the Spam data set. Continuous lines refer to random aggregated SVMs, dashed lines to bagged SVMs. R/S stands for Random aggregated vs. single SVMs and B/S bagged vs. single. Negative values indicate better results of single SVMs. (a) Comparing relative reduction of the error, bias and net variance (b) Comparing relative unbiased and biased variance reduction.