Microsoft Kinect RGBD

Random (Forest) Improves generalization

Ensemble of trees

Reduces overfitting by taking a majority vote over the predictions of multiple trees.

The trees in the ensemble need to vary from one another. They are all constructed from the same algorithm, but using a varied training set and a random vector Θ sampled from a distribution.

$S: \{a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6 \ a_7 \ a_8\}$

$S:$ original training set size m

done with replacement

$\{S'\}:$ random subsamples of S of size m'

$S': \{a_6 \ a_3 \ a_1 \ a_1\}$

$S': \{a_8 \ a_3 \ a_2 \ a_4\}$

$I: \{attributes\}$ $I_1, I_2 \ldots$

Generate a sequence of feature subsets $I_1, I_2 \ldots$

Each subset contains randomly selected k features. $k < d$.

$\Theta: \{S', I_t\}$

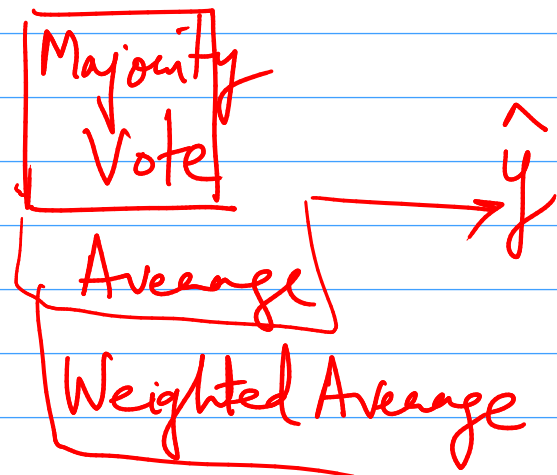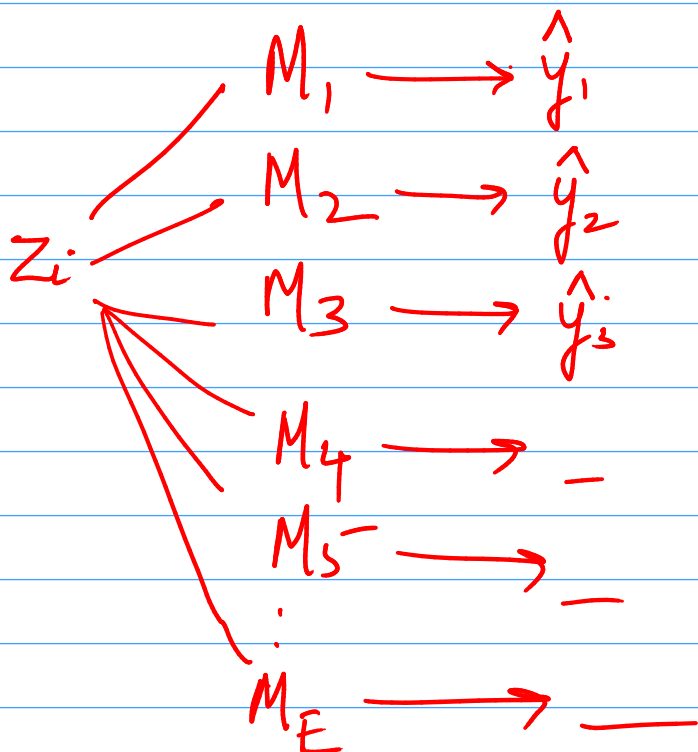On each sample S' we apply the decision tree growing algorithm (A.)
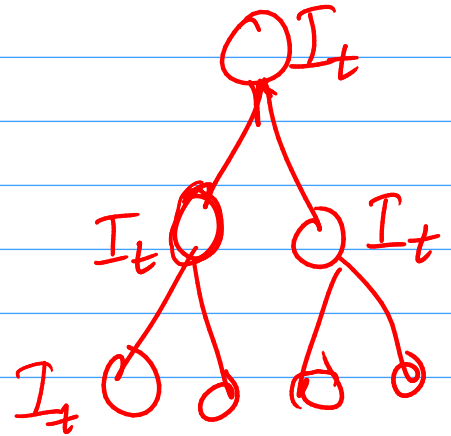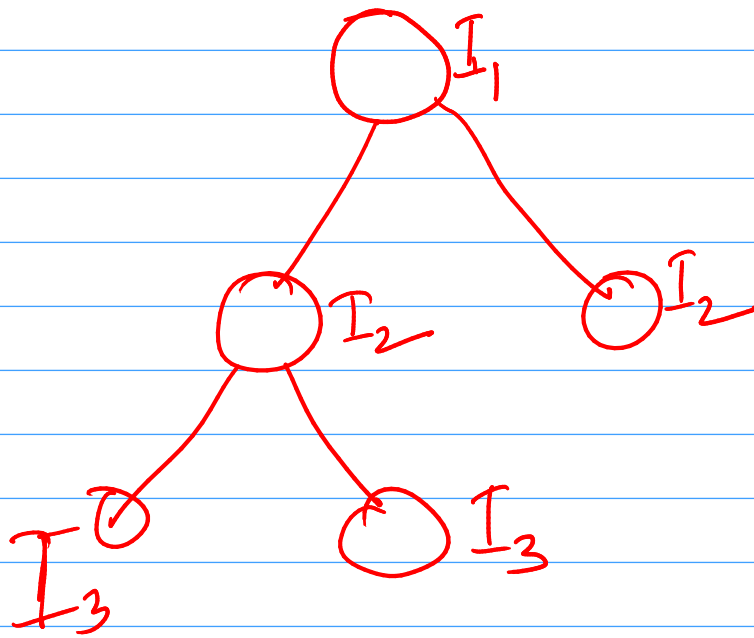
At each splitting stage, the split feature is selected from the feature set I_t.

Training set $S' \rceil I_t \rceil$

$\begin{matrix} S' \\ I \end{matrix} \rceil$ $\begin{matrix} S' \\ I: I_1 \ I_2 \ I_3 \ldots \end{matrix}$

Majority Vote

I_1

I_2 I_2

I_3 I_3

I_t

I_t I_t

I_t

$M_1 \longrightarrow \hat{y}_1$

$M_2 \longrightarrow \hat{y}_2$

$M_3 \longrightarrow \hat{y}_3$

$M_4 \longrightarrow \underline{\quad}$

$M_5 \longrightarrow \underline{\quad}$

$M_E \longrightarrow \underline{\quad}$

z_i

| Majority Vote |
| Average |
| Weighted Average |

$\longrightarrow \hat{y}$

test

$z_i \longrightarrow$ | Ensemble | $\longrightarrow (\hat{y}) \longrightarrow$ error
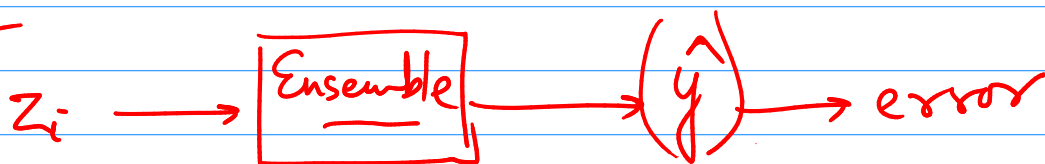
Adaboost $\longleftarrow$ Ensemble

l Loss: Error incurred on a test example

L Risk: Expected (Averaged) loss. — Expectation E

L_D True Risk: Expected loss over the unseen test examples — All possible examples ~ D

Empirical Risk: Expected loss over the training set.
L_S — sample S.

Error rate, prob. of error

Powerful models do not always perform well if you train them on a finite dataset.

When the training data S is (limited), it can depict accidental relationships between x_i and y_i

Assume that the training set of m examples $S = \{x_1, x_2 \ldots x_m\}$ is sampled again and again and a model is learned for every sample S. S' m'

$S \begin{cases} S' \, M_1 \\ S' \, M_2 \\ S' \, M_3 \\ S' \, M_4 \end{cases}$

Let such learned models be $M_1, M_2 \, M_3 \ldots$
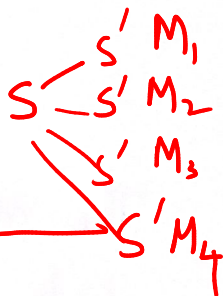
Let the test examples $z_1 \, z_2 \, z_3 \cdots z_t$ be fixed.

Bias in the error would imply that some particular test examples will be consistently misclassified even when you try different models $M_1 \, M_2 \cdots$ etc.

For example, k-nearest neighbour classifier can consistently misclassify some (-1) labelled examples if $(+1)$ labelled examples

5-nearest

Adopt the label of the nearest neighbour

K-nearest neighbour