

3) Principal Component analysis (PCA)

Function PCA (Data, D, Integer K):

Standardised \leftarrow STANDARDISE-DATA (D)

CovMatrix \leftarrow COMPUTE-COVARIANCE-MATRIX
(Standardised)

Eigenvalues, Eigenvectors \leftarrow Eigen-Decomposition
(CovMatrix)

Sorted \leftarrow Sort-Eigen Vectors - By - Eigen values (Eigenvalues
Eigenvectors)

Top-K-Vectors \leftarrow select-Top-K-Eigen vectors (Sorted, K)

Reduced Data \leftarrow Project-data (Standardised,
Top-K-vectors)

Return Reduced Data

function STANDARDISE-DATA (Data D):

for Each feature in D:

Subtract mean and divide by Standard variation

Return Standardised D

function Project-data (Data, Eigenvectors):

Return - Matrix-Multiplication (Data, eigenvectors)

4) Ada Boost (Adaptive Boosting) algorithm

Function ADABOOST (Dataset D , Integer T):

Initialize weights $w_i = 1/n$ for each training sample (x_i, y_i)

Classifiers $\leftarrow \{\}$

Alphas $\leftarrow \{\}$

for t from 1 to T do:

Classifier $h_t \leftarrow \text{TRAIN-WEAK-LEARNER}$
(D , weights w)

Error $\epsilon_t \leftarrow \sum [w_i * I(h_t(x_i) \neq y_i)]$

Alpha $\alpha_t \leftarrow 0.5 * \log((1 - \epsilon_t) / \epsilon_t)$

for each i from 1 to n do:

$w_i \leftarrow w_i * \exp(-\alpha_t * y_i * h_t(x_i))$

normalize weights: $w_i \leftarrow w_i / \sum w_i$

Append h_t to Classifiers

Append α_t to Alphas

Return Classifiers, Alphas

Function PREDICT-ADABOOST (Classifiers,

Alphas, Instance x):

Total $\leftarrow 0$

For t from 1 to T do:

Total \leftarrow Total + $\alpha_t * h_t(x)$

Return SIGN (Total)


09.04