2018-09-17

Nankai-Baidu
Joint Laboratory

Parallel and Distributed
Software Technology Lab

- ML Condential: Machine Learning on Encrypted Data
  Thore Graepel, Kristin Lauter, and Michael Naehrig

- ML Confidential Protocol
- Linear Means (LM) Classier
- Fisher's Linear Discriminant(FLD) Classier
  (on a publicly available data set)

□ML Confidential Protocol

✓Key Generation: The Data Owner executes the HE.Keygen algorithm publishes the public key, securely stores the private key locally

✓Encryption and Upload of Training Data: ··· encrypt preprocessed version of the training set data, i.e. sufficient statistics

Training:

Classification:

Verification of the Learned Model:

□Linear Means (LM) Classier

determines w and c such that

$f(x;\ w, c) = 0$ defines a hyper-plane midway on and orthogonal to the line through the two class conditional means.

It can be derived as the Bayes optimal decision boundary in the case that the two class-conditional distributions have identical isotropic Gaussian distributions.

□Linear Means (LM) Classier

$x \in R^n, y = \{+1, -1\}$

a linear classifier of the form $A(x;\ w, c) = sign\big(f(x;\ w, c)\big)$

$f(x;\ w, c) = w^T x - c$

Step 0. $I_y = \{i \in \{1, \cdots, m\} | y_i = y\}$

the index set of training examples with label y

$y = +1$: N_{+1}

$y = -1$: N_{-1}

□Linear Means (LM) Classier

$x \in R^n, y = \{+1, -1\}$

a linear classifier of the form $A(x;\ w, c) = sign\big(f(x;\ w, c)\big)$

$f(x;\ w, c) = w^T x - c$

Step 2. $\boldsymbol{m_y} = \frac{1}{N_y}\boldsymbol{s_y}$ $\boldsymbol{s_y} = \sum_{i \in I_y} x_i$

Step 3. $\boldsymbol{w}^* = \boldsymbol{m_{+1}} - \boldsymbol{m_{-1}}$ $\boldsymbol{w}^{*T} x_0 - \boldsymbol{c} = \boldsymbol{0}$

$x_0 = (\boldsymbol{m_{+1}} + \boldsymbol{m_{-1}})/2$

$\boldsymbol{c}^* = \boldsymbol{w}^{*T} x_0$

- Linear Means (LM) Classier

Step 3. $\boldsymbol{w}^* = \boldsymbol{m}_{+1} - \boldsymbol{m}_{-1} \quad \boldsymbol{c}^* = (\boldsymbol{m}_{+1} - \boldsymbol{m}_{-1})^T(\boldsymbol{m}_{+1} + \boldsymbol{m}_{-1})/2$

$$\boldsymbol{m}_y = \frac{1}{N_y}\boldsymbol{s}_y$$

$$A(x;\ w,c) = sign\big(f(x;\ w,c)\big)$$

$$f(x;\ w,c) = \boldsymbol{w}^{*T}x - \boldsymbol{c}^*$$

$$\tilde{\boldsymbol{f}}^*(\boldsymbol{x};\tilde{\boldsymbol{w}}^*,\tilde{\boldsymbol{c}}^*) = 2N_{+1}^2 N_{-1}^2 f(x;\ w,c)$$

Nankai-Baidu
Joint Laboratory

Parallel and Distributed
Software Technology Lab

- Linear Means (LM) Classier

compute $N_{-1}\boldsymbol{s}_{+1}$ and $N_{+1}\boldsymbol{s}_{-1}$ instead of $\boldsymbol{m}_{+1}$ and $\boldsymbol{m}_{-1}$

$$\widetilde{\boldsymbol{w}}^{*} = N_{-1}\boldsymbol{s}_{+1} - N_{+1}\boldsymbol{s}_{-1} = N_{+1}N_{-1}(\boldsymbol{m}_{+1} - \boldsymbol{m}_{-1}) = N_{+1}N_{-1}\boldsymbol{w}^{*}$$
$$\tilde{\boldsymbol{c}}^{*} = 2N_{+1}^{2}N_{-1}^{2}\boldsymbol{c}^{*}$$

$$\begin{aligned}\tilde{\boldsymbol{f}}^{*}(\boldsymbol{x};\widetilde{\boldsymbol{w}}^{*},\tilde{\boldsymbol{c}}^{*}) &= 2N_{+1}^{2}N_{-1}^{2}{\boldsymbol{w}^{*}}^{\boldsymbol{T}}\boldsymbol{x} - 2N_{+1}^{2}N_{-1}^{2}\boldsymbol{c}^{*} \\ &= 2N_{+1}^{2}N_{-1}^{2}f(x;\, w,c)\end{aligned}$$
$$sign\left(\tilde{\boldsymbol{f}}^{*}(\boldsymbol{x};\widetilde{\boldsymbol{w}}^{*},\tilde{\boldsymbol{c}}^{*})\right) = sign\big(f(x;\, w,c)\big)$$

□Fisher's Linear Discriminant(FLD) Classier

This algorithm is similar to the Linear Means classier,

but does take into account the class-conditional covariances.

maximizes the separation $S = \frac{\sigma_{inter}^2}{\sigma_{intra}^2} = \frac{w^T D w}{w^T C w}$

D the *between-class* covariance matrix

C the total *within-class* covariance matrix

- Fisher's Linear Discriminant(

want $|m_1 - m_2|$ to be large

and $s_1^2 + s_2^2$ to be small

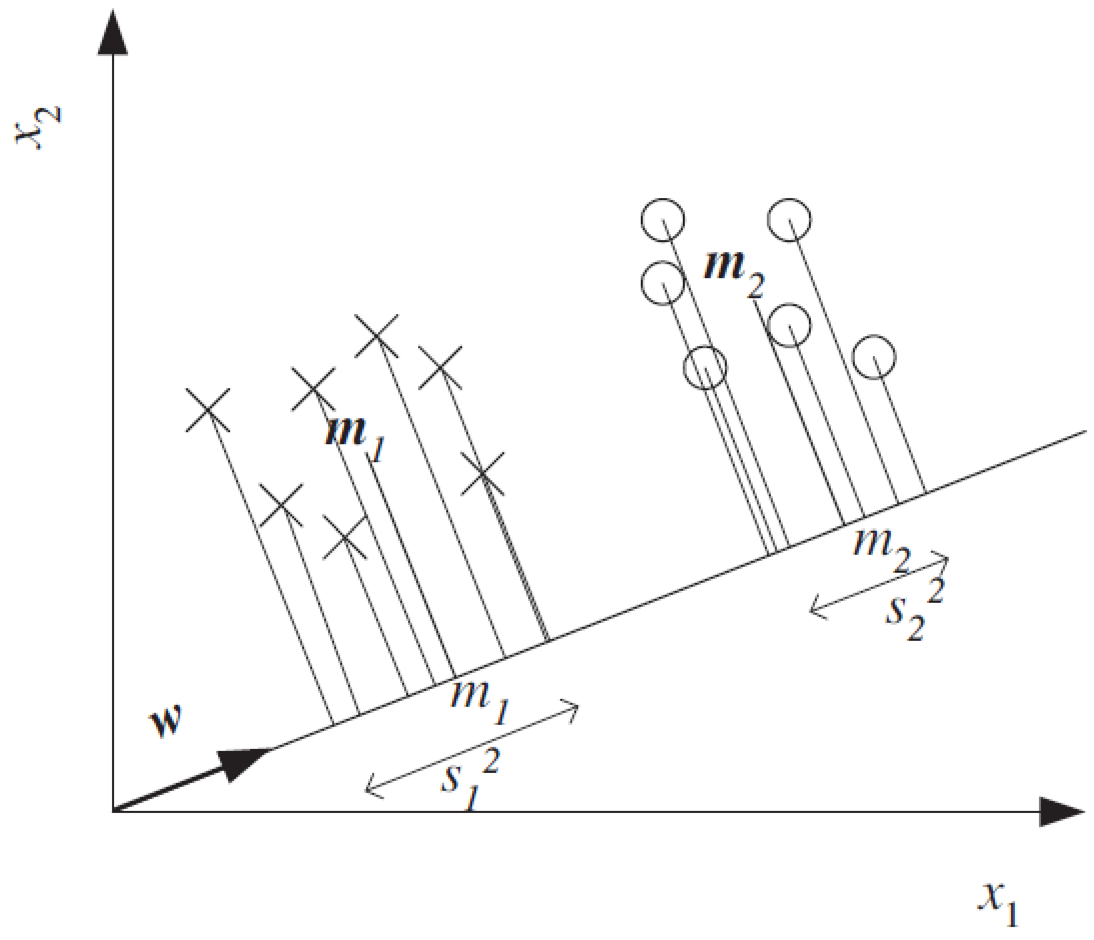
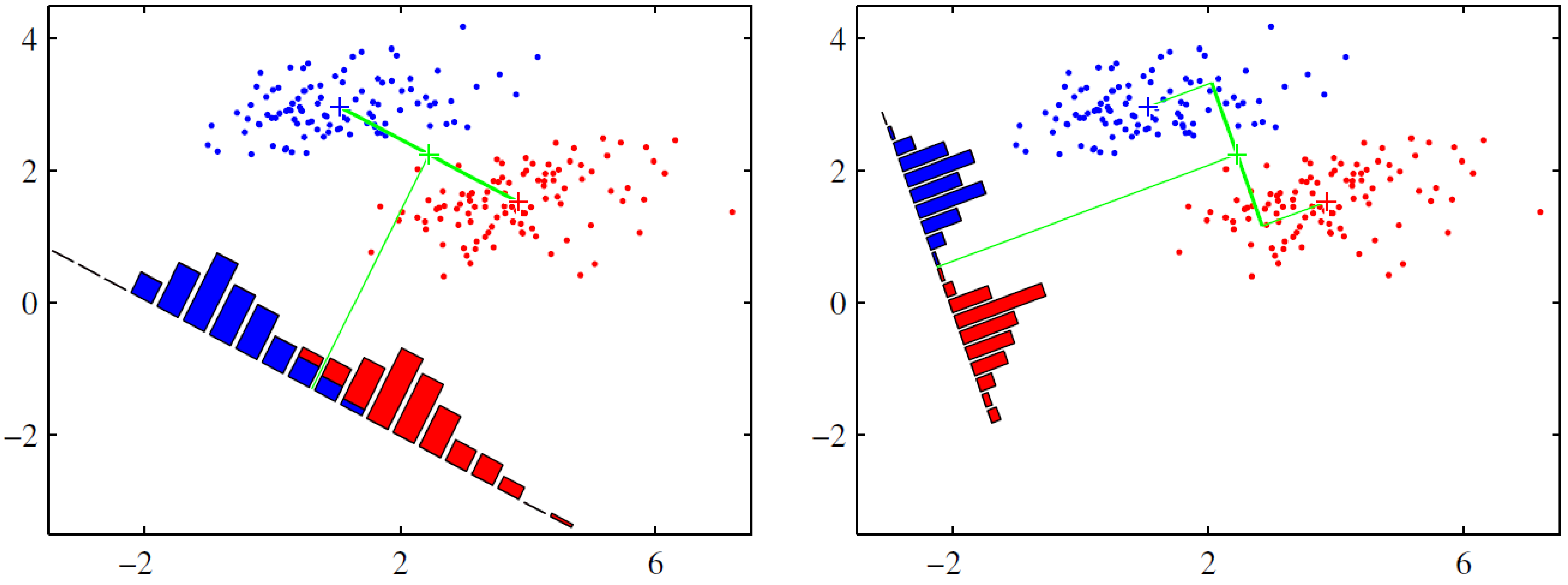


Figure 6.7 Two-dimensional, two-class data projected on w.



4
2
0
−2
−2
2
6
4
2
0
−2
−2
2
6

- Fisher's Linear Discriminant(FLD) Classier

maximizes the separation $S = \frac{\sigma_{inter}^2}{\sigma_{intra}^2} = \frac{w^T D w}{w^T C w}$

Taking the gradient w.r.t. w and setting it to zero.

$Cw^* \propto \boldsymbol{d}$ $\quad\quad \boldsymbol{d} = \boldsymbol{m}_{+1} - \boldsymbol{m}_{-1}$

$>> C^{-1} \; >> w^*$

cost function : $E(w) = \frac{1}{2} ||Cw - d||^2$

$\nabla_w E(w) = Cw - d$ $\quad\quad w_{j+1} = w_j - \eta \cdot \nabla_w E(w_j)$

use gradient descent to find the solution w^*

□Fisher's Linear Discriminant(FLD) Classier

use gradient descent to find the solution w^*

$$w_{j+1} = w_j - \eta \cdot \nabla_w E(w_j)$$

defining $w_0 = \mathbf{0}$

$$w_r = \eta \left(\sum_{j=0}^{r-1} (I - \eta C)^j \right) \boldsymbol{d}$$

This series converges if ……,

which can be ensured by choosing η sufficiently small.

- Fisher's Linear Discriminant(FLD) Classier

$$\boldsymbol{w}_r = \eta \left(\sum_{j=0}^{r-1} (I - \eta C)^j \right) \boldsymbol{d}$$

When $\eta < 1$

$$\widetilde{\boldsymbol{w}}_{\boldsymbol{r}} = \left(N_{+\mathbf{1}}^{\mathbf{3}} N_{-\mathbf{1}}^{\mathbf{3}} \eta^{-\mathbf{1}} \right)^{\boldsymbol{r}} \cdot \boldsymbol{w}_r$$

$$\eta^{-\mathbf{1}} \in \boldsymbol{Z}$$

resulting in the score function being a multiple of the original score function.

- HElib
- Logistic Regression