Raspoznavanje uzoraka
Algoritmi za rješavanje zadataka

Fisherova analiza

1. izračunaj srednje vrijednosti

$$\mathbf{m}_i = \frac{1}{n_i}\sum_{i=1}^{n}\mathbf{x}_i$$

2. izračunaj matrice raspršenja

$$S_j = \sum_{x\in\omega_j}(\mathbf{x}_i - \mathbf{m}_j)(\mathbf{x}_i - \mathbf{m}_j)^T$$

3. izračunaj

$$S_w = \sum_{i=1}^{K} S_i$$

4. za slučaj dva razreda

$$\mathbf{w} = S_w^{-1}(\mathbf{m}_1 - \mathbf{m}_2)$$

5. za slučaj više razreda

   - izračunaj srednju vrijednost svih uzoraka $\mathbf{m}$
   - izračunaj

   $$S_B = \sum_{i=1}^{K} n_i(\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T$$

   - izračunaj

   $$(S_w^{-1}S_B - \lambda\mathbf{I})\mathbf{w} = 0$$

6. normaliziraj rješenje

Ho-Kashyapov algoritam

1. proširi i invertiraj uzorke
2. izračunaj $\mathbf{X}^{\#}$

$$\mathbf{X}^{\#} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$$

3. ponavljaj
   - $\mathbf{w}(k) = \mathbf{X}^{\#}\mathbf{b}(k)$
   - $\mathbf{e}(k) = \mathbf{X}\mathbf{w}(k) - \mathbf{b}(k)$
   - $\mathbf{b}(k+1) = \mathbf{b}(k) + C[\mathbf{e}(k) + |\mathbf{e}(k)|]$
   - ako je $\mathbf{e}(k) < 0$, primjeri nisu linearni separabilni
4. dok je $\mathbf{e}_j \neq 0$

Perceptron za dva razreda

1. proširi i invertiraj uzorke
2. ponvaljaj
   - izračunaj $\mathbf{w}(k)^T\mathbf{x}$
   - ako je ≤ 0 korigiraj

$$\mathbf{w}(k+1) = \mathbf{w}(k) + c \cdot \mathbf{x}_i$$

3. dok nisu svi uzorci ispravno klasificirani

Bayesov naivni klasifikator

1. ako je potrebno izračunaj $\mathbf{m}_i$ za sve klase
2. ako je potrebno izračunaj kovarijacijske matrice C_i

$$C_j = \frac{1}{N_i}\sum_{i=1}^{N}(\mathbf{x}_i - \mathbf{m}_j)(\mathbf{x}_i - \mathbf{m}_j)^T$$

3. izračunaj $|C_1|$, $|C_2|$, C_1^{-1}, C_2^{-1}
4. decizijska funkcija zadana je sa

$$d_i = \mathrm{lnp}(\omega_i) - \frac{1}{2}\mathrm{ln}|C_i| - \frac{1}{2}\left(\mathbf{x} - \mathbf{m}_i\right)^T C_i^{-1}\left(\mathbf{x} - \mathbf{m}_i\right)$$

5. izračunaj $d_1 - d_2 = 0$
6. (*izmnoži sve matrice da dobiješ zapis u obliku polinoma*)

SVM
Postavljamo kriterijsku funkciju $J(\mathbf{w}, b, \boldsymbol{\lambda})$:

$$J(\mathbf{w}, b, \boldsymbol{\lambda}) = \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} - \sum_{i=1}^{n} \lambda_i \cdot (d_i(\mathbf{w}^{\mathrm{T}}\mathbf{x} + b) - 1)$$

gdje suma ima onoliko članova koliko imamo primjera za učenje.
Jednadžbe koje opisuju zavisnost $\mathbf{w}$ i $\boldsymbol{\lambda}$ dobijemo slijedećim vrijednostima:

- $\frac{\partial J(\mathbf{w}, b, \boldsymbol{\lambda})}{\partial \mathbf{w}} = 0$
- $\frac{\partial J(\mathbf{w}, b, \boldsymbol{\lambda})}{\partial b} = 0$

pri čemu mora vrijediti

- $\lambda_i \cdot [d_i(\mathbf{w}^{\mathrm{T}}\mathbf{x} + b) - 1] = 0, \quad i = 1, \ldots, n$
- $\lambda_i \geq 0$

Zadatak je pronaći optimalne vrijendosti λ_i. Potrebno je ispitati sve kombinacije λ_i koje zadovoljavaju uvjete:

$$\lambda_i \cdot [d_i(\mathbf{w}^{\mathrm{T}}\mathbf{x} + b) - 1] = 0 \;\; \text{za} \quad i = 1, \ldots, n.$$

Potrebno je samo paziti da je primjer $\mathbf{x}_i$ potporni vektor ako je $\lambda_i > 0$.
Vrijednosti hiperravnine dobiju se na sljedeći način:

- $\mathbf{w} = \sum_{i=1}^{n} \lambda_i d_i \mathbf{x}_i$
- $b = d_i - \mathbf{w}^T \mathbf{x}_i$

SVM - Dualni problem

1. definiraj vektor **E**; dimenzije su jednake broju primjera, a elemeti odgovaraju 1/-1 ovisno o klasi
2. definiraj vektor **b**; dimenzije su jednake broju primjera, svi elementi su jednaki 0
3. definiraj $\mathbf{A}$ = - $\mathbf{I}$; dimenzije broj primjera $\times$ broj primjera
4. definiraj vektor **c**; svi elementi jednaki -1, broj elemenata jednak broju primjera
5. definiraj $\mathbf{Q}_{ij} = d_i \cdot d_j \mathbf{x}_i^T \mathbf{x}_j$; za sve kombinacije primjera

6. za dobivene $\boldsymbol{\lambda}$ izračunaj

$$\mathbf{w} = \sum_{i=1}^{N} \lambda_i d_i \mathbf{x}_i$$

$$b = \pm 1 - \mathbf{w}^T \mathbf{x}_i, \text{za bilo koji } \mathbf{x}_i$$

PCA

1. izračunaj srednju vrijednost uzoraka
2. izračunaj matricu $\mathbf{K}$ (kovarijacijsku) ili matricu $\mathbf{R}$ (korelacijsku)
   - $\mathbf{K} = \frac{1}{N-1} \sum\limits_{i=1}^{N} (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^T$
   - $\mathbf{R} = \sum\limits_{j=1}^{K} \frac{1}{N_j} \mathrm{p}(\omega_j) \sum\limits_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^T$
3. pronađi svojstvene vrijednosti rješavajući $|\mathbf{K} - \lambda \mathbf{I}| = 0$
4. uzmi n najvećih svojstvenih vrijednosti i riječi sustav $[\mathbf{K} - \lambda \mathbf{I}]\mathbf{e} = 0$
5. napravi matricu $\mathbf{A}$ od n svojstvenih vrijednosti

$$A = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_n^T \end{bmatrix}$$

6. transformiraj ulazne podatke tako da množiš

$$y = \mathbf{A}(\mathbf{x} - \mathbf{m})$$

Potencijalne funkcije

1. $d_0(\mathbf{x}) = 0$
2. ponavljaj
   - izračunaj $d_k(\mathbf{x}_i)$
   - ako je $\mathbf{x}_i \in \omega_1$ i $d_k(\mathbf{x}_i) < 0 \quad d_{k+1}(\mathbf{x}) = d_k(\mathbf{x}) + K(\mathbf{x}_k, \mathbf{x})$
   - ako je $\mathbf{x}_i \in \omega_2$ i $d_k(\mathbf{x}_i) > 0 \quad d_{k+1}(\mathbf{x}) = d_k(\mathbf{x}) - K(\mathbf{x}_k, \mathbf{x})$
   - inače $d_{k+1}(\mathbf{x}) = d_k(\mathbf{x})$

3. dok ne razvrstaš pravilo sve primjere

Poopćene linearne funkcije

1. transformiraj podatke tako da koristiš rekurziju

$$g^r(\mathbf{x}) = \left(\sum_{p_1=1}^{l} \sum_{p_2=p_1}^{l} \cdots \sum_{p_r=p_{r-1}}^{l} w_{p_1,p_2,\cdots,p_r} x_{p_1} \cdots x_{p_r} \right) + g^{r-1}(\mathbf{x})$$

$$g^0(\mathbf{x}) = w_0$$

2. nad transformiranim podacima koristi bilo koji algoritam naveden prije

Euklidska udaljenost:

$$D = ||\mathbf{x}_k - \mathbf{x}_l|| = \sqrt{\sum_{i=1}^{n}(x_{k_i} - x_{l_i})^2}$$

n - dimenzija vektora
Euklidska udaljenost invarijantna je na translaciju i rotaciju.
Udaljenost Minkovskog:

$$D(\mathbf{x}_k, \mathbf{x}_l) = \left(\sum_{j=1}^{n} |\mathbf{x}_k - \mathbf{x}_l|^s)\right)^{\frac{1}{s}}$$

- s=2 $\rightarrow$ Euklidska udaljenost
- s=1 $\rightarrow$ Manhattan ili "city block" udaljenost

$$D(\mathbf{x}_k, \mathbf{x}_l) = \sum_{j=1}^{n} |\mathbf{x}_k - \mathbf{x}_l|$$

- s $\rightarrow \infty \rightarrow$ Čebišavljeva udaljenost

$$D(\mathbf{x}_k, \mathbf{x}_l) = max\{\sum_{j=1}^{n} |\mathbf{x}_k - \mathbf{x}_l|\}$$

Težinska udaljenost Minkovskog

$$D(\mathbf{x}_k, \mathbf{x}_l) = \left(\sum_{j=1}^{n} \mathrm{w}_j |\mathbf{x}_k - \mathbf{x}_l|^s)\right)^{\frac{1}{s}}$$

w_j – težina pojedine značajke
$\mathrm{w}_j \geq 1$
Mahalanobisova udaljenost

$$D(\mathbf{x}_k, \mathbf{x}_l) = (\mathbf{x}_k - \mathbf{x}_l)^T \mathbf{C}^{-1} (\mathbf{x}_k - \mathbf{x}_l)$$

C – kovarijacijska matrica dobivena iz skupa uzoraka za učenje (S_N)
Ako je C = I onda je Mahalanobisova udaljenost jednaka kvadratu Euklidske udaljenosti.
Normirani korelacijski koeficijenti

$$C(\mathbf{x}_k, \mathbf{x}_l) = \frac{\sum_{j=1}^{n}(x_{k_j} - m)(x_{l_j - m})}{\sqrt{\sum_{j=1}^{n}(x_{k_j} - m)^2 \sum_{j=1}^{n}(x_{l_j - m})^2}}$$

m - srednja vrijednost značajke

"Hi kvadrat" udaljenost

$$D(\mathbf{x}_k, \mathbf{x}_l) = \sum_{j=1}^{n} \frac{1}{x_{o_j}} \left(\frac{x_{k_j}}{x_{k_0}} - \frac{x_{l_j}}{x_{l_0}} \right)^2$$

$$x_{o_j} = \sum_{k=1}^{N} x_{k_j}$$

$$x_{k_0} = \sum_{j=1}^{N} x_{k_j}$$

$$x_{l_0} = \sum_{j=1}^{n} x_{l_j}$$

N - broj uzoraka u skupu za učenje S_N

Mjere sličnosti za binarne uzorke

$\mathbf{x}_l$ i $\mathbf{x}_k$ opisani su binarnim značajkama.
Hammingova udaljenost

$$D(\mathbf{x}_k, \mathbf{x}_l) = \sum_{j=1}^{n} d_j$$

$$d_j = \begin{cases} 1 & \text{ako je } x_{k_j} \neq x_{l_j} \\ 0 & \text{ako je } x_{k_j} = x_{l_j} \end{cases}$$

Težinska Hammingova udaljenost

$$D(\mathbf{x}_k, \mathbf{x}_l) = \sum_{j=1}^{n} \mathrm{w}_j d_j$$

$\mathrm{w}_j \geq 1$ – težina pojedinih začajki

	1	0
1	a	b
0	c	d

Russelova i Raova sličnost

$$S(\mathbf{x}_k, \mathbf{x}_l) = \frac{a}{a+b+c+d}$$

Jaccardova i Needhamova sličnost

$$S(\mathbf{x}_k, \mathbf{x}_l) = \frac{a}{a+b+c}$$

Sokalova i Sneathova sličnost

$$S(\mathbf{x}_k, \mathbf{x}_l) = \frac{a}{a + 2(b + c)}$$

Tanimotova sličnost

$$S(\mathbf{x}_k, \mathbf{x}_l) = \frac{a + d}{a + d + 2(b + c)}$$