

2 laboratorinis darbas

Šio laboratorinio darbo tikslas yra susipažinti su perceptrono ir vienasluoksniuo perceptrono apmokymu, ieškant hyperplokštumos, tiesiškai atskiriančios dviejų klasių duomenis n -matėje erdvėje. **Atsiskaitymas iki lapkričio 18 d. imtinai.**

Tarkime, turime objektus iš dviejų klasių C_1 ir C_2 . Kiekvienas objektas bus n -matis vektorius (stulpelis), kurio koordinatės vadiname objekto požymiais:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Transponavę stulpelį \mathbf{x} , gauname eilutę $\mathbf{x}' = (x_1, x_2, \dots, x_n)$.

Mokymui naudosime mokymo imtis L_1 ir L_2 , kur $L_1 \subseteq C_1$ ir $L_2 \subseteq C_2$ yra vienodo dydžio N . Tarkime,

$$L_1 = \{\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}, \dots, \mathbf{x}_N^{(1)}\} \quad \text{ir} \quad L_2 = \{\mathbf{x}_1^{(2)}, \mathbf{x}_2^{(2)}, \dots, \mathbf{x}_N^{(2)}\}.$$

Testavimui naudosime testines imtis T_1 ir T_2 , kur $T_1 \subseteq C_1$ ir $T_2 \subseteq C_2$ yra vienodo dydžio M .

Kiekvienas studentas gauna 2-os užduoties varianto numerį k . Žr. lentelę, esančią faile Uzduotys.pdf http://uosis.mif.vu.lt/~valdas/DNT/Laboratoriniai_darbai/

Duomenys, reikalingi 2-ai užduočiai atlikti yra http://uosis.mif.vu.lt/~valdas/DNT/Laboratoriniai_darbai/Laboratorinis_2/Duomenys/

Tai reiškia, kad variantui k turite naudoti duomenis $L_1 = Ak$, $L_2 = Bk$, $T_1 = Ck$, $T_2 = Dk$, $C_1 = Ek$, $C_2 = Fk$ (aibės C_1 ir C_2 reikalingos tik užduotyje 2C).

2A. Perceptronas

Naudodami mokymo imtis L_1 ir L_2 ir perceptrono mokymo taisyklę, raskite mokymo imtis skiriančią hyperplokštumą $\mathbf{w}'\mathbf{x} = 0$ tokią, kad $\mathbf{w}'\mathbf{x} > 0 \quad \forall \mathbf{x} \in L_1$ ir $\mathbf{w}'\mathbf{x} \leq 0 \quad \forall \mathbf{x} \in L_2$. Mokymą pradedame nuo vienetinių svorių, t.y. $\mathbf{w}_{\text{init}} = (1, 1, \dots, 1)'$. Viena po kito perceptronui pateikiame mokymo duomenis, kaitaliodami klases, t.y. mokymo vektorius išdėstome tokia tvarka:

$$\{\mathbf{x}_1^{(1)}, \mathbf{x}_1^{(2)}, \mathbf{x}_2^{(1)}, \mathbf{x}_2^{(2)}, \dots, \mathbf{x}_N^{(1)}, \mathbf{x}_N^{(2)}\}.$$

Viena epocha vadiname visų mokymo vektorių pateikimą perceptronui po 1 kartą, t.y. $2N$ individualių mokymo algoritmo iteracijų.

Perceptrono mokymo taisyklė atrodo taip:

$$\mathbf{w}_{\text{new}} = \begin{cases} \mathbf{w} & \text{jei } \mathbf{w}'\mathbf{x} > 0 \text{ ir } \mathbf{x} \in L_1 \text{ arba } \mathbf{w}'\mathbf{x} \leq 0 \text{ ir } \mathbf{x} \in L_2, \\ \mathbf{w} + \eta\mathbf{x} & \text{jei } \mathbf{w}'\mathbf{x} \leq 0 \text{ ir } \mathbf{x} \in L_1, \\ \mathbf{w} - \eta\mathbf{x} & \text{jei } \mathbf{w}'\mathbf{x} > 0 \text{ ir } \mathbf{x} \in L_2, \end{cases}$$

kur \mathbf{x} yra naujas mokymo vektorius, o η yra mokymo žingsnis. Mokymas vyksta tol, kol per visą epochą neįvyksta nė vienos svorių korekcijos, t.y. kol perceptronas bus apmokytas be klaidų.

Rasti: Naudodami mokymo žingsnius $\eta = 0.1$, $\eta = 1$ ir $\eta = 10$, tris kartus apmokykite perceptroną be klaidų ir apskaičiuokite individualių svorių koreguojančių iteracijų skaičių (t.y. tas iteracijas, kuriose svoriai nesikeičia, skaičiuodami praleidžiamie) bei paskutinės (be klaidų) epochos numerį. Kiekvieną kartą apmokius perceptroną be klaidų raskite testinės klaidos dydį procentais, t.y., kiek procentų vektorių iš imčių T_1 ir T_2 plokštuma $\mathbf{w}'\mathbf{x} = 0$ atskiria klaidingai (čia \mathbf{w} yra be klaidų apmokyto perceptrono svoriai).

2B. Vienasluoksnis perceptronas

Naudodami mokymo imtis L_1 ir L_2 ir atgalinio skleidimo algoritmą, apmokykite vienasluoksnį perceptroną be klaidų ir raskite mokymo imtis skiriančią hyperplokštumą $\mathbf{w}'\mathbf{x} = 0$ tokią, kad $\mathbf{w}'\mathbf{x} > 0 \forall \mathbf{x} \in L_1$ ir $\mathbf{w}'\mathbf{x} \leq 0 \forall \mathbf{x} \in L_2$. Mokymą pradėkite iš to paties taško, kaip ir tiesinį perceptroną, t.y. $\mathbf{w}_{\text{init}} = (1, 1, \dots, 1)'$. Vieną po kito vienasluoksniui perceptronui pateikiame mokymo duomenis, išdėstytus bet kuria tvarka (kadangi naudojame paketinį mokymą, t.y. svorių korekciją vyksta po visos epochos, praleidus kiekvieną mokymo vektorių iš mokymo imčių L_1 ir L_2 lygiai 1 kartą).

Perceptrono nuostolių funkcija yra vidutinė kvadratinė klaida

$$E = \frac{1}{4N} \sum_{i=1}^2 \sum_{j=1}^N (d_j^i - o_j^i)^2,$$

kur o_j^i yra perceptrono išėjimas, gaunamas įėjimo vektoriui $\mathbf{x}_j^{(i)}$ (i -osios klasės j -ajam vektoriui), o d_j^i yra norimas perceptrono išėjimas šiam vektoriui, t.y.

$$d_j^i = \begin{cases} 1, & \text{jei } \mathbf{x}_j^{(i)} \in L_1 \\ 0, & \text{jei } \mathbf{x}_j^{(i)} \in L_2. \end{cases}$$

Naudojant logistinę aktyvacijos funkciją vienasluoksnių perceptrono svorių koregavimo taisyklė atrodo taip:

$$\mathbf{w}_{\text{new}} = \mathbf{w} + \frac{\eta}{2N} \sum_{i=1}^2 \sum_{j=1}^N (d_j^i - o_j^i) o_j^i (1 - o_j^i) \mathbf{x}_j^{(i)},$$

kur η yra mokymo žingsnis. Mokymas vyksta tol, kol mokymo klaida (kurią skaičiuojame po kiekvienos epochos) virsta nuliu.

Rasti: Naudodami mokymo žingsnius $\eta = 0.1$, $\eta = 1$ ir $\eta = 10$, tris kartus apmokykite vienasluoksnį perceptroną be klaidų ir apskaičiuokite reikalingų epochų skaičių. Kiekvieną kartą apmokius vienasluoksnį perceptroną be klaidų raskite testinės klaidos dydį procentais (t.y. kiek procentų vektorių iš imčių T_1 ir T_2 plokštuma $\mathbf{w}'\mathbf{x} = 0$ atskiria klaidingai (čia \mathbf{w} yra be klaidų apmokyto vienasluoksnių perceptrono svoriai).

2C. Vienasluksnio perceptrono mokymas tiesiškai neatskiriamų duomenų atveju

Duotos klasės C_1 ir C_2 (duomenų aibės Ek ir Fk , kur k yra varianto numeris). Iš kiekvienos klasės išrenkame po 5 atsitiktinius mokymo vektorius, apmokome netiesinį perceptroną (kaip aprašyta skyrelyje 2B) ir apskaičiuojame mokymo bei testinę klaidą procentais (testinės imtys sutampa su klasėmis C_1 ir C_2). Mokymą pradedame nuo atsitiktinių pradinųjų svorių \mathbf{w}_{init} , tolygiai pasiskirčiusių intervale $(-0.1, 0.1)$. Šį kartą mokome ne iki nulinės mokymo klaidos, o tol, kol atsitiks vienas iš šių įvykių: (1) epochų skaičius pasieks 1000 arba (2) skirtumas tarp dviejų paskutinių nuostolių funkcijos reikšmių pasidarys labai mažas, t.y. $|E_{\text{new}} - E_{\text{old}}| < 0.001$. Čia E_{old} yra vidutinė kvadratinė klaida prieš svorių perskaičiavimą (t.y. po priešpaskutinės epochos), o E_{new} yra vidutinė kvadratinė klaida po svorių perskaičiavimo (t.y. po paskutinės epochos).

Mokymo procesą kartojame 10 kartų (kiekvieną kartą su $5 + 5$ atsitiktinių vektorių) ir apskaičiuojame vidutinę mokymo ir vidutinę testinę klaidas. Po to tą patį eksperimentą atliekame 10 kartų su atsitiktinėmis mokymo imtimis dydžio $10 + 10$ bei $20 + 20$. Mokymo žingsnis $\eta = 1$. Jei $\eta = 1$ duoda blogus rezultatus, papildomai atlikite tuos pačius eksperimentus su jūsų pasirinktu mokymo žingsniu.

Rasti: Naudodami atsitiktines mokymo imtis dydžio 10 ($5 + 5$), 20 ($10 + 10$) ir 40 ($20 + 20$) raskite vidutinę mokymo bei vidutinę testinę klaidas. Taip pat apskaičiuokite vidutinį epochų skaičių iki sustojimo.