Софийски университет "Св. Климент Охридски"

Факултет по математика и информатика

МАШИННО САМООБУЧЕНИЕ

спец. Изкуствен интелект, I курс, зимен семестър
учебна година 2024/2025

Изготвил:
Кристиян Симов
фак. номер 4MI3400288

Дата:
16. 11. 2024 г.
София

Домашна работа №3

Съдържание

1 Решение на задача №1

Единствената нужна промяна на алгоритъм **НАУЧИ-ЕДНО-ПРАВИЛО** от Лекция 5 Табл.5-2, която ще позволи той да работи коректно ако някой от атрибутите е с непрекъснати стойности, е да прибавим генериране (вътре в цикъла "за всяка хипотеза h от *Кандидат-хипотези*"от първа точка) на ограничения c от вида $A_i > v_{i,j}$ и $A_i <= v_{i,j}$, $j \in \{1, n'\}$ за всеки непрекъснат атрибут A_i, където $v_{i,1}, ..., v_{i,n'}$ са намерените възможни гранични стойности за атрибута A_i, чрез сортиране на всички n' на брой примери, които се покриват от текущата хипотеза h в цикъла (в началото това коректно са всички примери, тъй като h е най-общата хипотеза, покриваща всичките n примера, т.е $n' = n$) и намирането на средно аритметичното на всеки две съседни стойности по A_i в сортираната последователност $a_{i1}, .., a_{in'}$. Така получаваме $v_{i,1} = \frac{a_{i1}+a_{i2}}{2}, ..., v_{i,n'} = \frac{a_{i(n'-1)}+a_{in'}}{2}$ и след това допълнителните възможни ограничения $(A_i > v_{i,j})$ и $(A_i <= v_{i,j})$, $j \in \{1, n'\}$, които искахме да генерираме, тъй като сред тях е възможно да изберем (ако някое от тях има най-добра евреистика) новата конюнкция към правилото което строим. Това правим за всеки непрекъснат атрибут A_i. Тоест вече избираме c от обединението на множеството *Всички-ограничения* (предварително генерираните ограничения за дискретните атрибути) и всички така генерирани допълнителни правила за непрекъснатите атрибути. Това правим за всяка хипотеза от цикъла, защото ако тя покрива различни примери, това променя вида на сортираната последователност и съответно изчисленията при средните стойности. Всяка спецификация чрез кое да е ограничение би могла потенциално да промени множеството от покриваните примери.

2 Решение на задача №2

Нека има два персептрона (A и B), чиято повърхност на решение се определя с формулата:

$$w_0 + w_1x_1 + w_2x_2 > 0$$

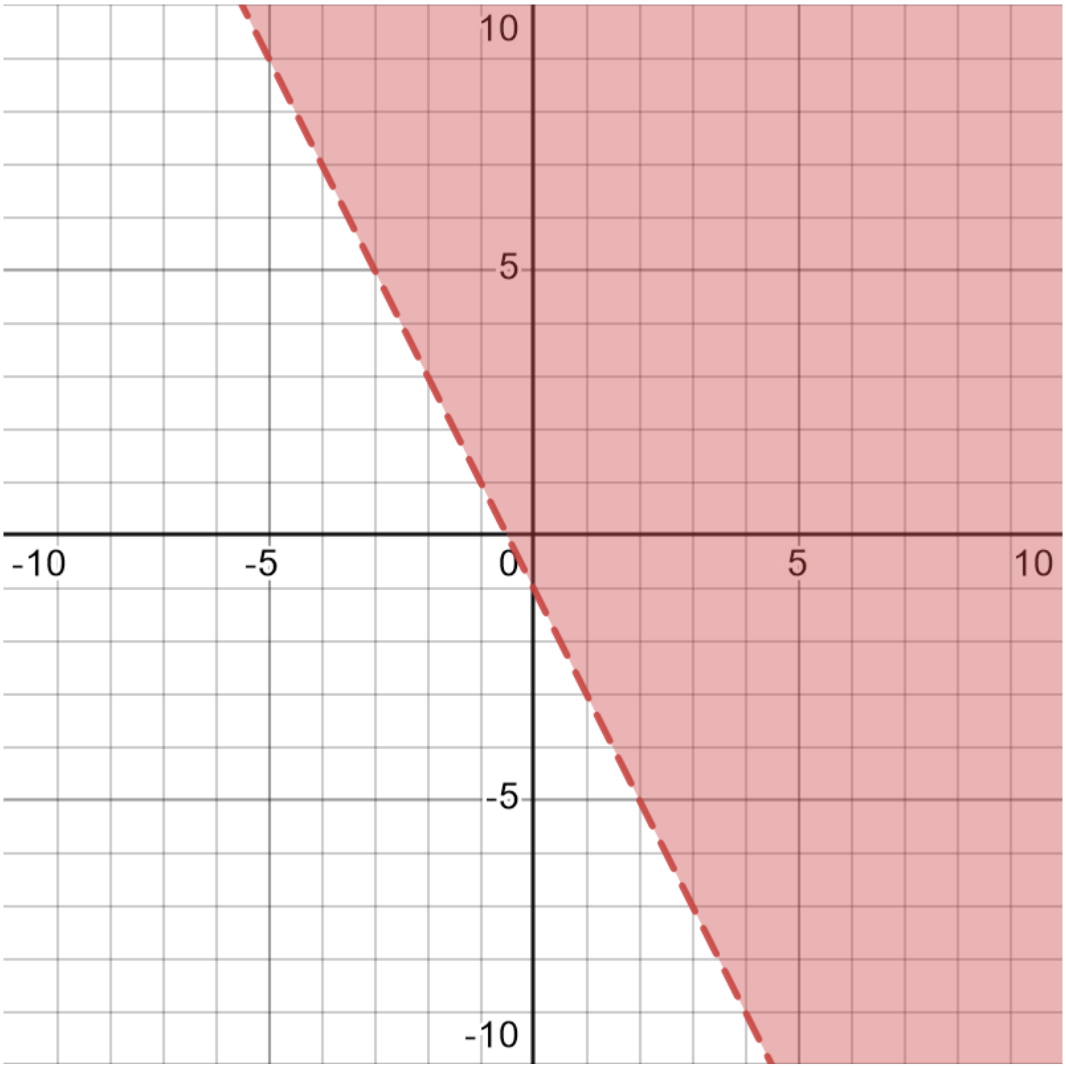
Нека персептрон A има следните стойности на теглата:

$$w_0 = 1, w_1 = 2, w_2 = 1$$

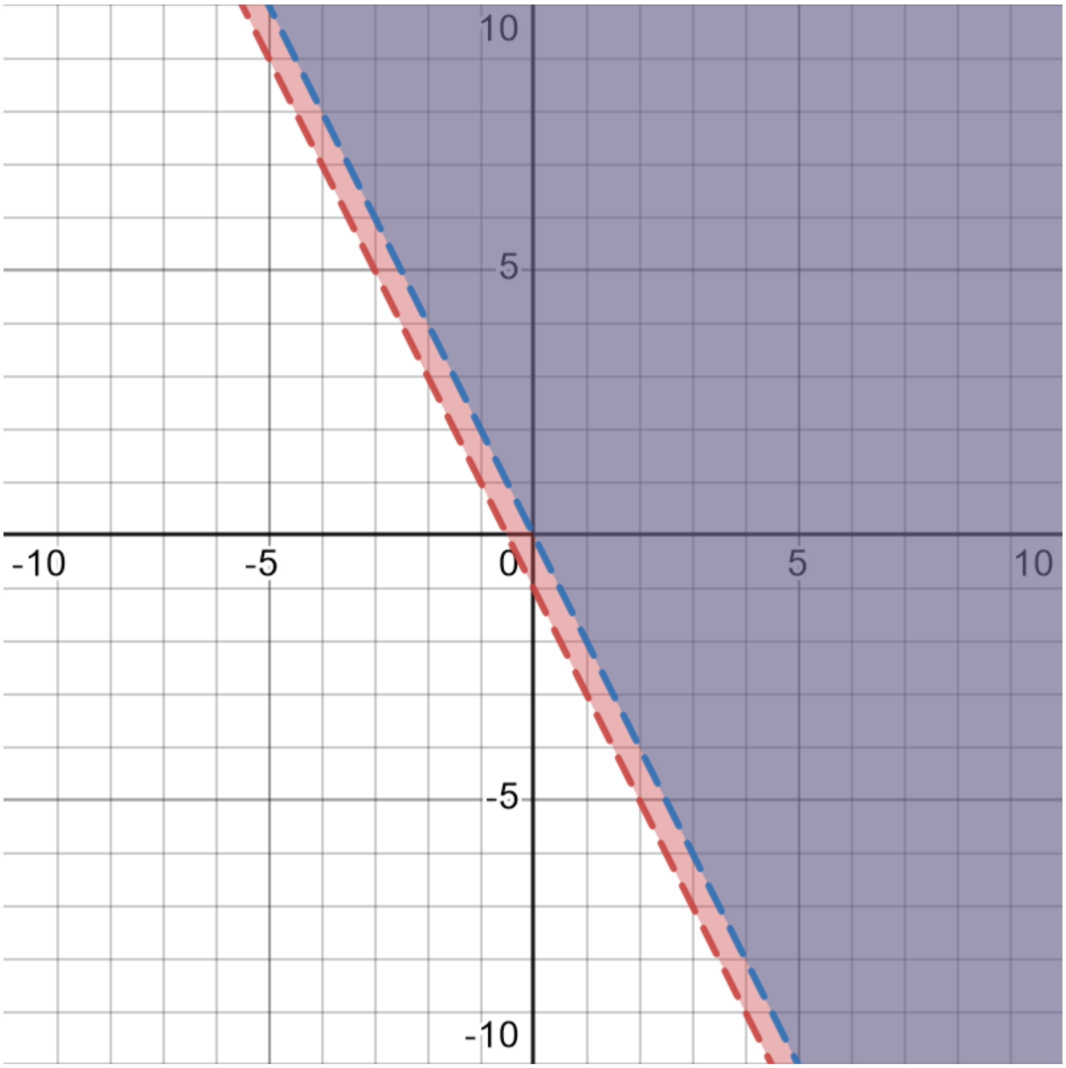
Нека персептрон B има следните стойности на теглата:

$$w_0 = 0, w_1 = 2, w_2 = 1$$

В случая имаме x_1 и x_2 (два входа) следователно за персептрони A и B n-мерните хиперплоскости, които отделят n-мерното пространство на примерите, представляват прави в двумерно пространство:

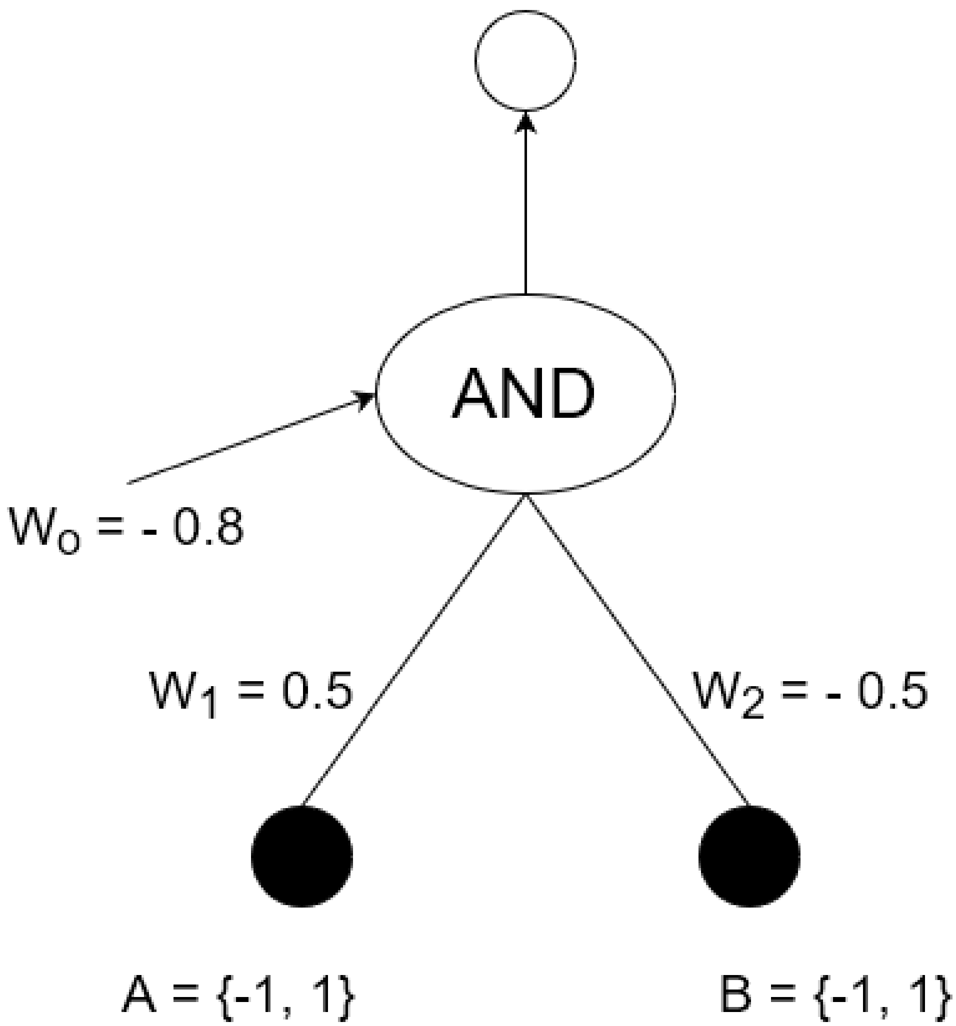


Фигура 1: Персептрон A. В червен пунктир виждаме графиката на правата $1+2x_1+x_2=0$, която отделя пространството, а със светло червено е запълнено множеството от стойности които са решение на неравенството $1+2x_1+x_2>0$

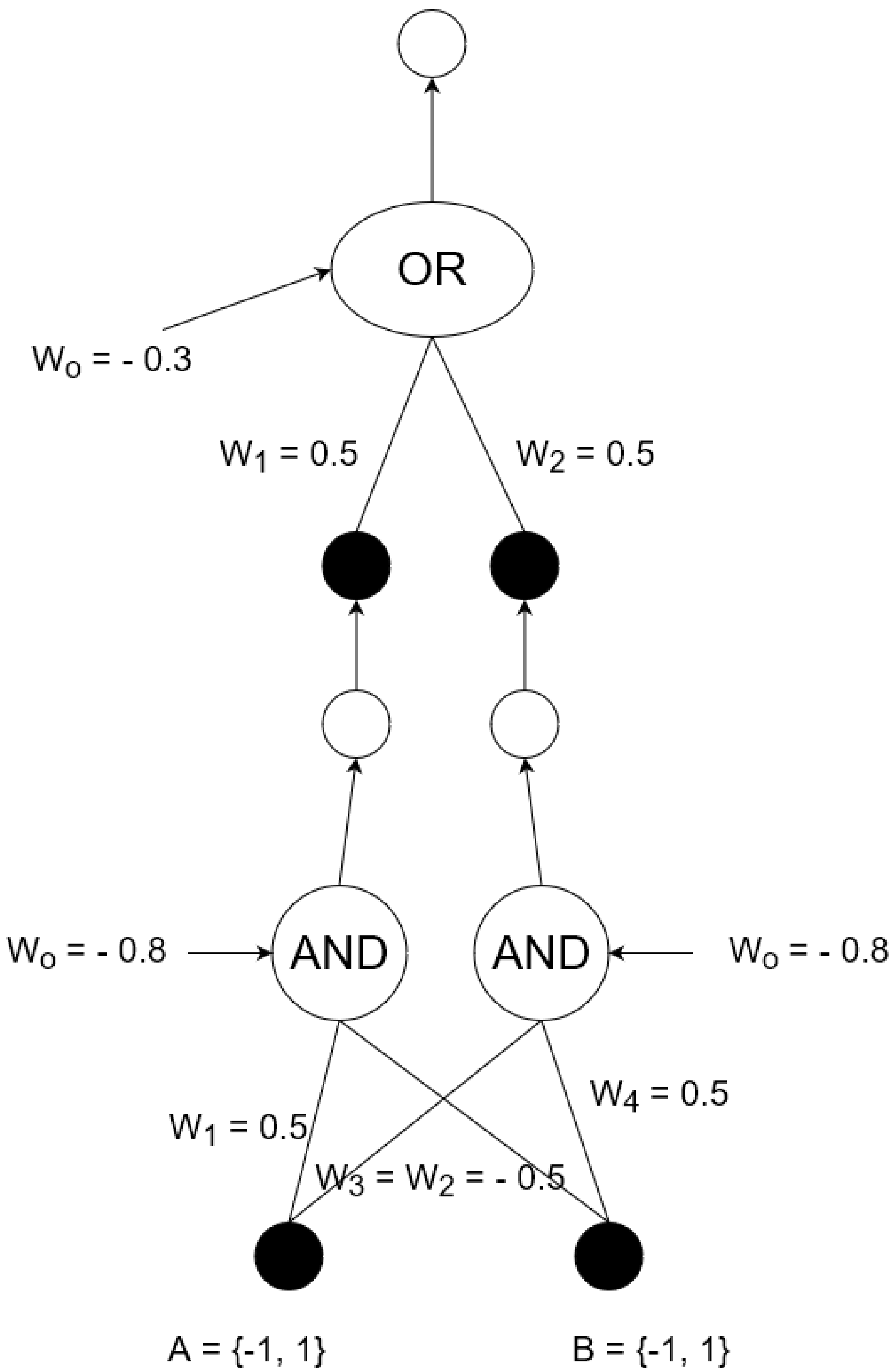


Фигура 2: Персептрони A и B. Върху предходната Фигура 1 е нанесена със син пунктир графиката на правата $2x_1 + x_2 = 0$, която отделя пространството, а със светло син е запълнено множеството от стойности които са решение на неравенството $2x_1 + x_2 > 0$. Виждаме, че то е подмножество на множеството от решения на оцветената в светло червено област от решения на неравенството $1 + 2x_1 + x_2 > 0$. Това означава, че всяка точка от синята област е решение за червената, но обратното не е в сила. Следователно $A >_g$ B.

3 Решение на задача №3



Фигура 3: а) $A \wedge \neg B$



Фигура 4: b) $(A \wedge \neg B) \vee (\neg A \wedge B)$

4 Решение на задача №4

Ще изведем правилото за обучение чрез градиентното спускане на един линеен възел, чийто изход o се задава от формулата:

$$o = w_0 + w_1 x_1 + w_1 x_1^2 + ... + w_n x_n + w_n x_n^2$$

Правилото за обучение чрез градиентното спускане в покомпонентен вид е следното:

$$\Delta w_i = -\eta \tfrac{\delta E}{\delta w_i}$$

Пресмятаме стойността на вектора от производни $\frac{\delta E}{\delta w_i}$, който формира градиента:

$$\tfrac{\delta E}{\delta w_i} = \tfrac{\delta}{\delta w_i} \tfrac{1}{2} \sum_{d \in D} (t_d - o_d)^2 = \frac{1}{2} \sum_{d \in D} 2(t_d - o_d) \frac{\delta}{\delta w_i} (t_d - o_d) =$$

$$= \sum_{d \in D} (t_d - o_d) \frac{\delta}{\delta w_i} (t_d - \vec{w}\vec{x}) = \sum_{d \in D} (t_d - o_d)(-x_{id} - x_{id}^2)$$

Така в крайна сметка за Δw_i получаваме:

$$\Delta w_i = -\eta \tfrac{\delta E}{\delta w_i} = -\eta \sum_{d \in D} (t_d - o_d)(-x_{id} - x_{id}^2) = \eta \sum_{d \in D} (t_d - o_d)(x_{id} + x_{id}^2)$$