

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 4МІ3400288

Дата:

23. 10. 2024 г.
София

Домашна работа №1



Съдържание

1	Решение на задача №1	2
2	Решение на задача №2	4
3	Решение на задача №3	6
3.1	11
3.2	11
3.3	11
3.4	12

1 Решение на задача №1

Нека в контекста на изграждане на хипотези означим с $?$ - всички стойности на атрибут са възможни, а с \emptyset - нито една не е възможна.

Нека означим множествата с допустимите стойности (без \emptyset и $?$) за всеки от n -те атрибута с фамилията $A = \{A_1, A_2, \dots, A_n\}$.

Нека $m_1 = |A_1|, m_2 = |A_2|, \dots, m_n = |A_n|$

Тогава за задачата от Лекция 1 имаме $m_1 = 3, m_2 = 2, \dots, m_6 = 2$

По вероятностни съображения, тъй като можем да избираме първо по 3 начина, после по 2, ... , накрая отново по 2, получаваме, че броя на всички различни възможни примери е точно числото:

$$E_6 = \prod_{i=1}^6 m_i = 3 * 2^5 = 3 * 32 = 96$$

За броя на възможните хипотези, които в тази задача имат вида $\langle a_1, \dots, a_6 \rangle$, където $a_i \in A_i$ или $a_i = ?$, или $a_i = \emptyset$, правим аналогично съображение на предходното, но към всяка мощност m_i трябва да прибавим 2 (броим $?$ и \emptyset , като валидни стойности за a_i). Така бихме получили:

$$\prod_{i=1}^6 (m_i + 2) = 5 * 4^5 = 5 * 1024 = 5120,$$

но ще преброим всеки вектор $\langle a_1, \dots, a_6 \rangle$, където за някое i имаме $a_i = \emptyset$. Тези вектори за нашите цели ефективно са като вектора $\langle \emptyset, \emptyset, \dots, \emptyset \rangle$ и затова нека преброим само него и към броя възможности прибавяме само 1-ца (за $?$). Така броя на различните възможни хипотези спада на едва:

$$H_6 = 1 + \prod_{i=1}^6 (m_i + 1) = 1 + 4 * 3^5 = 1 + 4 * 243 = 1 + 972 = 973$$

Нека добавим множество A_7 с мощност $m_7 = 3$ (по условие) към фамилията A . Тогава вече ще имаме брой възможни примери равен на:

$$E_6 * m_7 = \left(\prod_{i=1}^6 m_i \right) * m_7 = \prod_{i=1}^7 m_i = 96 * 3 = 288$$

От своя страна, броят хипотези ще се измени по следния начин:

$$1 + \left(\prod_{i=1}^6 (m_i + 1) \right) * (m_7 + 1) = 1 + 972 * 4 = 1 + 3888 = 3889$$

Или изразено чрез H_6 :

$$1 + (H_6 - 1) * (m_7 + 1) = H_6 * (m_7 + 1) - m_7 = 973 * 4 - 3 = 3892 - 3 = 3889$$

Нека по-общо прибавим $(n+1)$ -во множество A_{n+1} ($\emptyset \notin A_{n+1}$ и $? \notin A_{n+1}$) с мощност k , т.е. $m_{n+1} = |A_{n+1}| = k$, към фамилията A .

Тогава ако функцията $E_n = E(n)$ описва нарастването на броя възможни примери, а $H_n = H(n)$ - на броя възможни хипотези то:

$$\begin{aligned} E(n+1) &= E(n) * m_{n+1} = E(n) * k \\ H(n+1) &= H(n) * (m_{n+1} + 1) - m_{n+1} = H(n) * (k + 1) - k \end{aligned}$$

2 Решение на задача №2

Нека в множество D имаме обучаващи примери от вида $\langle x, c(x) \rangle$, които в зависимост от реда на постъпване при обучението са индексирани по следния начин (в обратен ред спрямо таблицата от Лекция 1):

$$x_1 = \langle \text{Слънце}, \text{Топъл}, \text{Висока}, \text{Силен}, \text{Студена}, \text{Промяна} \rangle +$$

$$x_2 = \langle \text{Дъжд}, \text{Студен}, \text{Висока}, \text{Силен}, \text{Топла}, \text{Промяна} \rangle -$$

$$x_3 = \langle \text{Слънце}, \text{Топъл}, \text{Висока}, \text{Силен}, \text{Топла}, \text{Същото} \rangle +$$

$$x_4 = \langle \text{Слънце}, \text{Топъл}, \text{Нормална}, \text{Силен}, \text{Топла}, \text{Същото} \rangle +$$

Тогава алгоритъмът CANDIDATE-ELIMINATION ще премине през следните стъпки:

0) Инициализация

$$G_0 \leftarrow \langle ?, ?, ?, ?, ?, ? \rangle$$

$$S_0 \leftarrow \langle \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset \rangle$$

$$1) \quad x_1 = \langle \text{Слънце}, \text{Топъл}, \text{Висока}, \text{Силен}, \text{Студена}, \text{Промяна} \rangle +$$

Примерът е положителен. В G_0 няма несъвместими хипотези, остава както е. S_0 е несъвместима, тъй като е твърде специфична, премахваме я. В съответствие с x_1 добавяме хипотеза s_1 към S_1 .

$$S_1 \leftarrow \langle \text{Слънце}, \text{Топъл}, \text{Висока}, \text{Силен}, \text{Студена}, \text{Промяна} \rangle$$

$$G_1 \leftarrow G_0$$

2) $x_2 = \langle \text{Дъжд}, \text{Студен}, \text{Висока}, \text{Силен}, \text{Топла}, \text{Промяна} \rangle -$

Примерът е отрицателен. В S_1 няма несъвместими хипотези, остава както е. G_1 е несъвместима, тъй като е твърде обща, премахваме я. В съответствие с x_2 и S_2 добавяме само хипотези $g_{2.1}$, $g_{2.2}$ и $g_{2.5}$ към G_2 :

$$S_2 \leftarrow S_1$$

$$G_2 \leftarrow \langle \text{Слънце}, ?, ?, ?, ?, ? \rangle$$

$$G_2 \leftarrow \langle ?, \text{Топъл}, ?, ?, ?, ? \rangle$$

$$G_2 \leftarrow \langle ?, ?, ?, ?, \text{Студена}, ? \rangle$$

3) $x_3 = \langle \text{Слънце}, \text{Топъл}, \text{Висока}, \text{Силен}, \text{Топла}, \text{Същото} \rangle +$

Примерът е положителен. S_2 е несъвместима, тъй като е твърде специфична. Обобщаваме единствената хипотеза в нея до s_3 и получаваме S_3 . Вече хипотезата $g_{2.5}$ от G_2 е несъвместима с резюмето s_3 на всички положителни примери до този момент, премахваме я от G_2 и получаваме G_3 .

$$S_3 \leftarrow \langle \text{Слънце}, \text{Топъл}, \text{Висока}, \text{Силен}, ?, ? \rangle$$

$$G_3 \leftarrow G_2 \setminus \langle ?, ?, ?, ?, \text{Студена}, ? \rangle$$

4) $x_4 = \langle \text{Слънце}, \text{Топъл}, \text{Нормална}, \text{Силен}, \text{Топла}, \text{Същото} \rangle +$

Примерът е положителен. S_3 е несъвместима, тъй като е твърде специфична. Обобщаваме единствената хипотеза в нея до s_4 и получаваме S_4 . Новополученото резюме s_4 е съвместимо с хипотезите от G_3 . Следователно за G_4 получаваме G_3 .

$$S_4 \leftarrow \langle \text{Слънце}, \text{Топъл}, ?, \text{Силен}, ?, ? \rangle$$

$$G_4 \leftarrow G_3 = \{ \langle \text{Слънце}, ?, ?, ?, ?, ? \rangle, \langle ?, \text{Топъл}, ?, ?, ?, ? \rangle \}$$

Очаквано, получихме същия резултат като при $x_4 \rightarrow x_3 \rightarrow x_2 \rightarrow x_1$. \square

3 Решение на задача №3

Нека бележим хипотезите (правоъгълниците) по следния съкратен начин:

$$h_{\square} = \langle a, b, c, d \rangle$$

Да приемем, че всичките обучаващи примери (както и бъдещи примери за класификация) се побират в двумерно пространство на характеристиките (атрибутите) с размерност 10 на 10, център $(0, 0)$ и дискретни стойности - единствено положителни цели числа.

Нека в множество D имаме обучаващи примери от вида $\langle x, c(x) \rangle$, които в зависимост от реда на постъпване при обучението са индексирани по следния начин (напълно произволно, с изключение на това че първо са индексирани положителните примери, с цел стегнатост на решението - ако имаме резюмето на положителните по начало, ще можем с лекота да отхвърлим множество от хипотези за G):

$$x_1 = \langle 6, 5 \rangle +$$

$$x_2 = \langle 5, 3 \rangle +$$

$$x_3 = \langle 4, 4 \rangle +$$

$$x_4 = \langle 5, 1 \rangle -$$

$$x_5 = \langle 1, 3 \rangle -$$

$$x_6 = \langle 2, 6 \rangle -$$

$$x_7 = \langle 5, 8 \rangle -$$

$$x_8 = \langle 9, 4 \rangle -$$

Тогава алгоритъмът CANDIDATE-ELIMINATION ще премине през следните стъпки:

0) Инициализация

$$S_0 \leftarrow \langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$$

$$G_0 \leftarrow \langle 0, 10, 0, 10 \rangle$$

1) $x_1 = \langle 6, 5 \rangle +$

Примерът е положителен. G_0 е съвместима. S_0 е твърде специфична, тъй като е несъвместима. Обобщаваме я и получаваме S_1 - правоъгълник, който покрива точно една точка - самата x_1 .

$$S_1 \leftarrow \langle 6, 6, 5, 5 \rangle$$

$$G_1 \leftarrow G_0$$

2) $x_2 = \langle 5, 3 \rangle +$

Примерът е положителен. G_1 е съвместима. Остава както е. S_1 е твърде специфична и значи несъвместима. Обобщаваме я и получаваме S_2 - правоъгълник, който покрива точно двете точки x_1 и x_2 .

$$S_2 \leftarrow \langle 5, 6, 3, 5 \rangle$$

$$G_2 \leftarrow G_1$$

3) $x_3 = \langle 4, 4 \rangle +$

Примерът е положителен. G_2 е съвместима. Остава както е. S_2 е твърде специфична и значи несъвместима. Обобщаваме я и получаваме S_3 - правоъгълник, който покрива точно трите точки x_1 , x_2 и x_3 .

$$S_3 \leftarrow \langle 4, 6, 3, 5 \rangle$$

$$G_3 \leftarrow G_2$$

$$4) \quad x_4 = \langle 5, 1 \rangle -$$

Примерът е отрицателен. S_3 е съвместима. Остава както е. G_3 е твърде обща и значи несъвместима. Специфицираме я и получаваме g_4 - единствен правоъгълник, който е съвместим едновременно с S_3 и с новата точка x_4 .

$$S_4 \leftarrow S_3$$

$$G_4 \leftarrow \langle 0, 10, 2, 10 \rangle$$

$$5) \quad x_5 = \langle 1, 3 \rangle -$$

Примерът е отрицателен. S_4 е съвместима. Остава както е. G_4 е твърде обща и значи несъвместима. Специфицираме я и получаваме g_5 - единствен правоъгълник, който е съвместим едновременно с S_4 и с новата точка x_5 .

$$S_5 \leftarrow S_4$$

$$G_5 \leftarrow \langle 2, 10, 2, 10 \rangle$$

$$6) \quad x_6 = \langle 2, 6 \rangle -$$

Примерът е отрицателен. S_5 е съвместима. Остава както е. g_5 от G_5 е твърде обща и значи несъвместима. Специфицираме я и получаваме $g_{6.1}$ и $g_{6.2}$ - два възможни правоъгълника, които са съвместими едновременно с S_5 и с новата точка x_6 .

$$S_6 \leftarrow S_5$$

$$G_6 \leftarrow \{ \langle 3, 10, 2, 10 \rangle, \langle 2, 10, 2, 5 \rangle \}$$

7) $x_7 = \langle 5, 8 \rangle -$

Примерът е отрицателен. S_6 е съвместима. Остава както е. $g_{6.1}$ от G_6 е твърде обща и значи несъвместима. Специфицираме я и получаваме $g_{7.2}$, а $g_{7.1}$ е съвместимата $g_{6.2}$ от G_6 . Накарая, в G_7 имаме единствено два възможни правоъгълника, които са съвместими едновременно с S_6 и с новата точка x_7 .

$$S_7 \leftarrow S_6$$

$$G_7 \leftarrow (G_6 \setminus \{3, 10, 2, 10\}) \cup \{\langle 3, 10, 2, 7 \rangle\}$$

$$G_7 = \{\langle 2, 10, 2, 5 \rangle, \langle 3, 10, 2, 7 \rangle\}$$

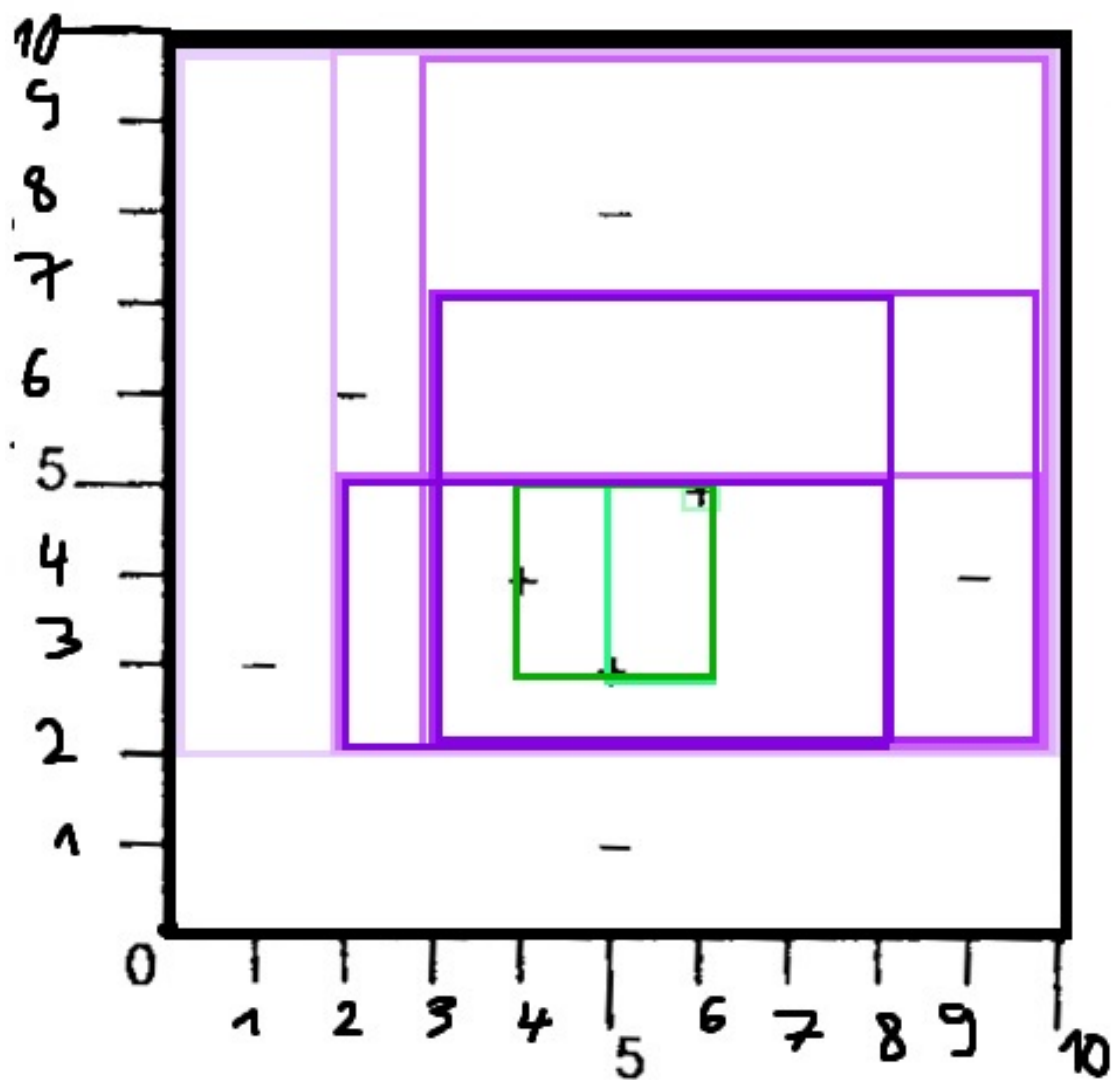
8) $x_8 = \langle 9, 4 \rangle -$

Примерът е отрицателен. S_7 е съвместима. Остава както е. $g_{7.1}$ и $g_{7.2}$ от G_7 са твърде общи и значи несъвместими. Специфицираме ги и получаваме $g_{8.1}$ и $g_{8.2}$ - два възможни правоъгълника, които са съвместим едновременно с S_7 и с новата точка x_8 .

$$S_8 \leftarrow S_7 (= S_6 = \dots = S_3 = \langle 4, 6, 3, 5 \rangle)$$

$$G_8 \leftarrow \{\langle 2, 8, 2, 5 \rangle, \langle 3, 8, 2, 7 \rangle\}$$

Край на обучението.



Фигура 1: На изображението виждаме в нюансите на зеления цвят последователните хипотези за S , а в нюансите на лилавото - последователните хипотези за G , генерирани в хода на изпълнението на алгоритъма. Финалните хипотези са нанесени в съответните най-тъмни нюанси.

3.1

Финалният вид на S след обучението е следният:

$$S = \langle 4, 6, 3, 5 \rangle$$

Видът на S е нанесен в най-тъмния нюанс на зеленото на чертежа в Фиг.1 на стр.10

3.2

Финалният вид на G след обучението е следният:

$$G = \{\langle 2, 8, 2, 5 \rangle, \langle 3, 8, 2, 7 \rangle\}$$

Видът на G е нанесен в най-тъмния нюанс лилавото на чертежа в Фиг.1 на стр.10

3.3

Нека хипотезите от $G = \{\langle 2, 8, 2, 5 \rangle, \langle 3, 8, 2, 7 \rangle\}$ означим така:

$$g_1 = \langle 2, 8, 2, 5 \rangle$$

$$g_2 = \langle 3, 8, 2, 7 \rangle$$

Тогава пример, който ще намали пространството на версиите задължително когато постъпи, се намира в участъка $\langle 3, 8, 5, 7 \rangle$ (в g_2 , но извън g_1), например нека това е:

$$x_{decr} = \langle 4, 6 \rangle$$

$$1.1) \quad x_{decr} = \langle 4, 6 \rangle +$$

Примерът е положителен. Тогава хипотезата g_1 вече няма да е съвместима с него и ще трябва да я премахнем. В допълнение ще трябва да обобщим единствената хипотеза s_1 от S , така че тя да покрие новия

положителен пример, което обаче няма да промени броят хипотези. В крайна сметка броят хипотези намалява (съответно и пространството на версиите).

$$1.2) \quad x_{decr} = \langle 4, 6 \rangle -$$

Примерът е отрицателен. Тогава хипотезата g_2 вече няма да е съвместима с него и ще трябва да я премахнем, тъй като тя ще се специфицира до нова хипотеза g_3 , която попада изцяло вътре в g_1 . Така в крайна сметка в G ще остане само g_1 . Следователно отново правим същия извод като в края на 1.1).

Пример, който няма да намали пространството на версиите задължително когато постъпи, се намира в участъка извън пространството заключено от G , например нека това е:

$$x_{decr} = \langle 9, 4 \rangle$$

$$2.1) \quad x_{decr} = \langle 9, 4 \rangle -$$

Примерът задължително ще е отрицателен, според модела нямаме друг случай за него. Той е съвместим и с S , и с G . Следователно няма нужда да правим промени по хипотезите и така пространството на версиите не се променя (тоест също така не намалява).

3.4

По дефиниция, за да се научи едно понятие абсолютно точно, е необходимо границите S и G да се "стегнат" до една и съща граница съдържаща единствена хипотеза.

В случая искаме това да е точно хипотезата:

$$h_{\square} = \langle 3, 5, 2, 9 \rangle$$

1) Първо да конструираме S , която има същия вид като h_{\square} .

За целта можем да използваме минимално 2 обучаващи положителни примера, които ако свържем биха образували диагонала на хипотезата (правоъгълника $\langle 3, 5, 2, 9 \rangle$), например:

$$x_1 = \langle 3, 2 \rangle + \text{ и } x_2 = \langle 5, 9 \rangle +$$

Така ще получим $S_2 = \langle 3, 5, 2, 9 \rangle$ за две стъпки от $S_0 = \langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$

2) Сега остана да "свиваме" G докато $G \equiv S$.

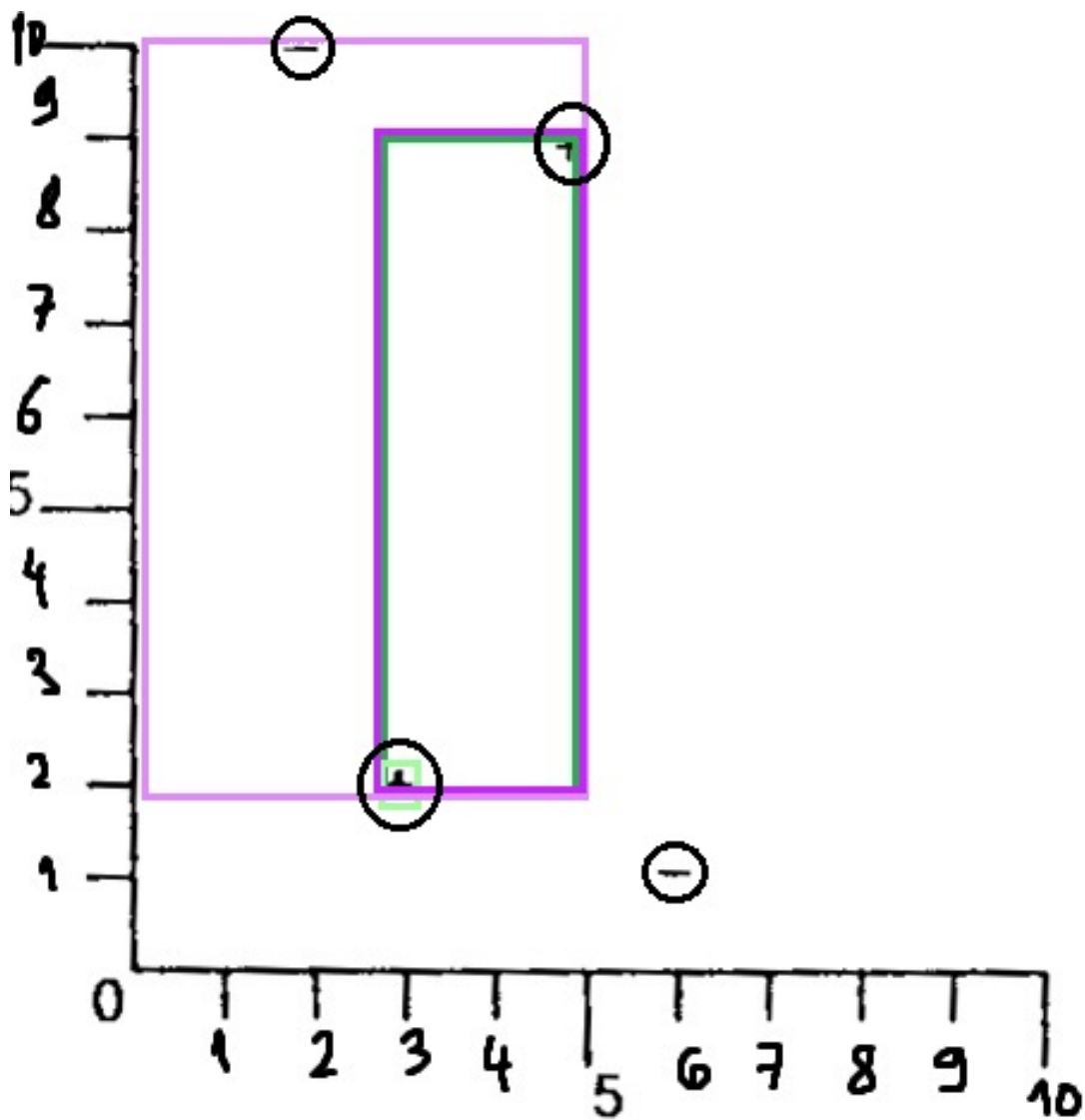
За целта можем да използваме минимално 2 обучаващи отрицателни примера, които заради стъпка 1) е нужно да са непосредствено извън S_2 и дигонално разположени по посока обратния диагонал на този образуван от x_1 и x_2 , например така:

$$x_3 = \langle 6, 1 \rangle - \text{ и } x_4 = \langle 2, 10 \rangle -$$

Така ще получим $G_2 = \langle 3, 5, 2, 9 \rangle$ за две стъпки от $G_0 = \langle 0, 10, 0, 10 \rangle$:

$$\langle 0, 10, 0, 10 \rangle \rightarrow \langle 0, 5, 2, 10 \rangle \rightarrow \langle 3, 5, 2, 9 \rangle$$

Визуално пояснение на гореизложеното в 1) и 2) може да бъде видяно във Фиг.2 на стр.13 (следващата и последна страница).



Фигура 2: На изображението виждаме в нюансите на зеления цвят последователните хипотези за S , а в нюансите на лилавото - последователните хипотези за G , генерирани в хода на изпълнението на построението от 1) и 2). Финалните хипотези за S и G са нанесени в съответните най-тъмни нюанси. Както се вижда те ще съвпаднаат, както искахме, и то с минималния брой от 4 необходими точки - 2 положителни за S и 2 отрицателни за G .