**Inteligência Artificial**

Lista de Exercícios – Aprendizado de Máquina - Profa. Heloisa

| Recomendações |
| --- |
| - Para resolver exercícios de construção de uma árvore de decisão, seja completa ou com apenas alguns passos, toda a informação deve ser indicada: quais exemplos de cada classe “desceram” por cada ramo quando um nó é definido, como foi calculada entropia e ganho de informação;  - O cálculo de entropia envolve a função logaritmo na base 2. No final da lista está disponível uma tabela com os valores (p, log2p e p.log2p) já calculados de forma aproximada, com 2 casas decimais. Esses valores podem ser usados nos exercícios;  - Alternativamente, exercícios do tipo citado no item anterior podem já fornecer alguns valores de ganho de informação de atributos específicos ou dicas de valores obtidos dependendo da proporção de exemplos em cada classe;  - Quando todos os exemplos de um conjunto são da mesma classe, a entropia é zero. Quando isso ocorre, não é necessário fazer os cálculos detalhados.  - Para calcular o ganho de informação de um atributo que vai ser escolhido para a raiz da árvore, lembre-se que a entropia do conjunto original nem sempre é igual a 1; se a classe não for binária ou se for binária mas o número de exemplos em cada classe for diferente, é preciso calcular essa entropia;  -Se ao calcular a entropia de um conjunto notar que a proporção de exemplos nas classes é a mesma de um conjunto para o qual já calculou a entropia anteriormente no mesmo exercício, não é necessário refazer os cálculos (ver exercício 2);  - As informações solicitadas referentes aos passos da indução de árvore de decisão podem ser indicadas textualmente, não sendo obrigatório fazer o desenho da árvore. Nesse caso, todas as informações pedidas devem estar presentes, inclusive quais são os exemplos de cada classe usados no cálculo da entropia de cada conjunto (ver exercício 5); |

1) Uma árvore de decisão deve ser construída para determinar se uma pessoa deve ou não esperar por uma mesa em um restaurante. Os dados disponíveis registram casos de clientes que frequentam restaurantes e cada dado é descrito pelos atributos: *Cliente*, que indica quantas pessoas estão no restaurante (com valores nenhum, algum e cheio); *Faminto*, que indica se o cliente estava ou não faminto quando chegou (com valores sim ou não); *Tipo*, que indica o tipo do restaurante (com valores francês, italiano, tailandês ou só serve hambúrguer) e *Sex/sab*, que indica se é sexta ou sábado (com valores sim ou não). Calcule o ganho de informação para o atributo *Client*e. Utilize a tabela de logaritmos no final da lista.

| Exemplo | Cliente | Faminto | Tipo | Sex/sab | Esperar |
| --- | --- | --- | --- | --- | --- |
| 1 | alguns | sim | frances | não | sim |
| 2 | cheio | sim | tailandês | não | não |
| 3 | alguns | não | hambúrguer | não | sim |
| 4 | cheio | sim | tailandês | sim | sim |
| 5 | cheio | não | frances | sim | não |
| 6 | alguns | sim | italiano | não | sim |
| 7 | nenhum | não | hambúrguer | não | não |
| 8 | cheio | sim | hambúrguer | sim | sim |
| 9 | cheio | não | hambúrguer | sim | não |
| 10 | nenhum | não | tailandês | não | não |

Ganho de informação do atributo Cliente

Cliente

| 2 | não |
| --- | --- |
| 4 | sim |
| 5 | não |
| 8 | sim |
| 9 | não |

cheioalgunsnenhum

| 7 | não |
| --- | --- |
| 10 | não |

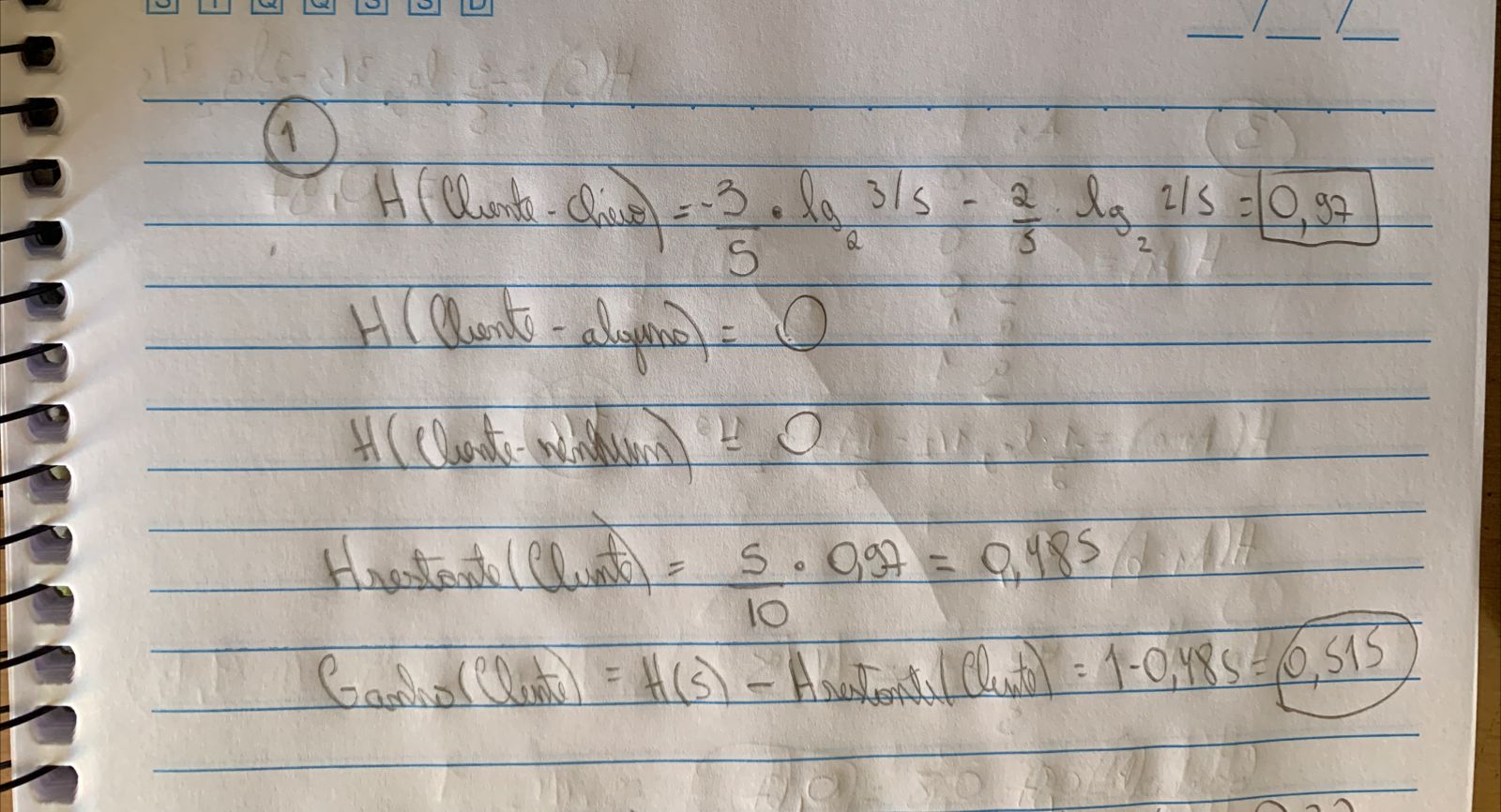
| 1 | sim |
| --- | --- |
| 3 | sim |
| 6 | sim |

H(Cliente-cheio) = -(2/5.log2 2/5) – (3/5.log2 3/5) = -(0,4.log2 0,4) – (0,6.log2 0,6) = -(0,4. -1,32) – (0,6.-0,74) = 0,53 + 0,44 = 0,97

H(Cliente=alguns) = 0 (todos os exemplos da mesma classe)

H(Cliente=nenhum) = 0 (todos os exemplos da mesma classe)

HRestante(Cliente) =5 /10\*H(Cliente=cheio) + 3/10\*H(Cliente=alguns) + 2/10 \* H(Cliente=nenhum) = 0,5 \* 0,97 + 0,3 \* 0 + 0,2 \* 0 = 0,485



Ganho(Cliente) = H(S) – Hrestante(Cliente) = 1 – 0,485 = 0,515

2) Calcule o ganho de informação do atributo “Faminto” para os dados do exercício anterior. Utilize a tabela de logaritmos no final da lista.

Faminto

sim não

| 3 | sim |
| --- | --- |
| 5 | não |
| 7 | não |
| 9 | não |
| 10 | não |

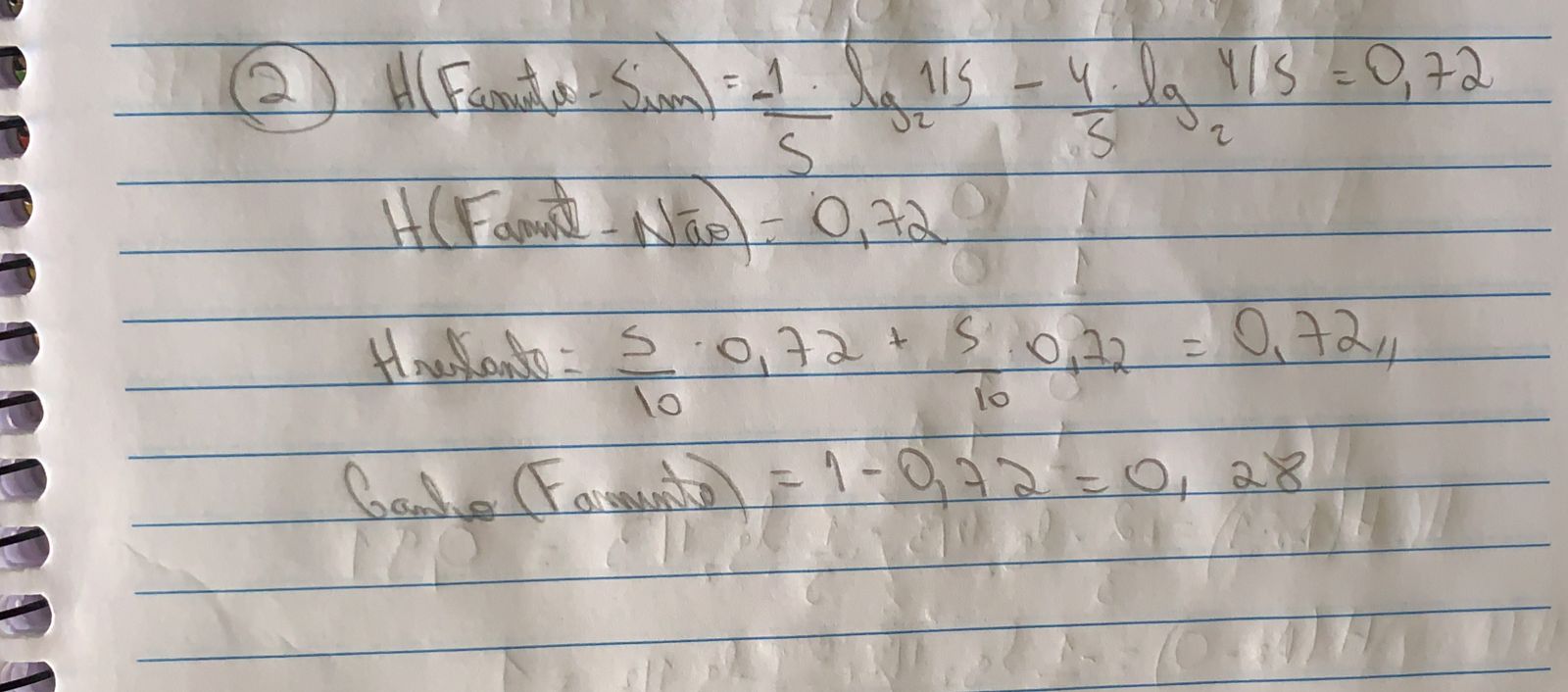
| 1 | sim |
| --- | --- |
| 2 | não |
| 4 | sim |
| 6 | sim |
| 8 | sim |

H(Faminto=sim) = -(1/5.log2 1/5) – (4/5.log2 4/5) = -(0,2.log2 0,2) – (0,8.log2 0,8) = -(0,2. -2,32) – (0,8.-0,32) = 0,46 + 0,26 = 0,72

H(Faminto=não) = 0,72 (mesma proporção de classes do caso anterior)

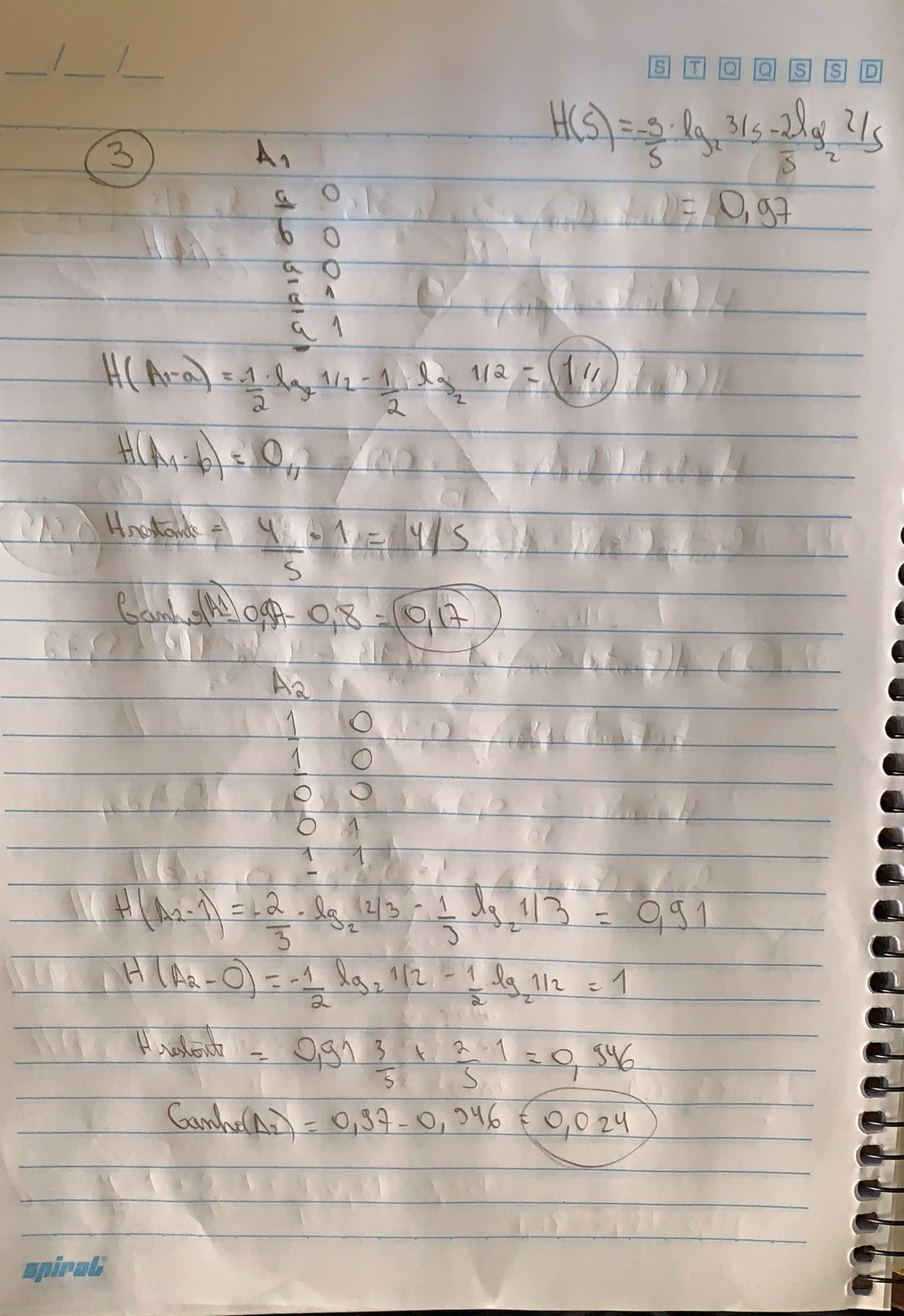
HRestante(Faminto) =5 /10\*H(Faminto=sim) + 5/10\*H(Faminto=não) = = 0,5 \* 0,72 + 0,5 \* 0,72 = 0,72

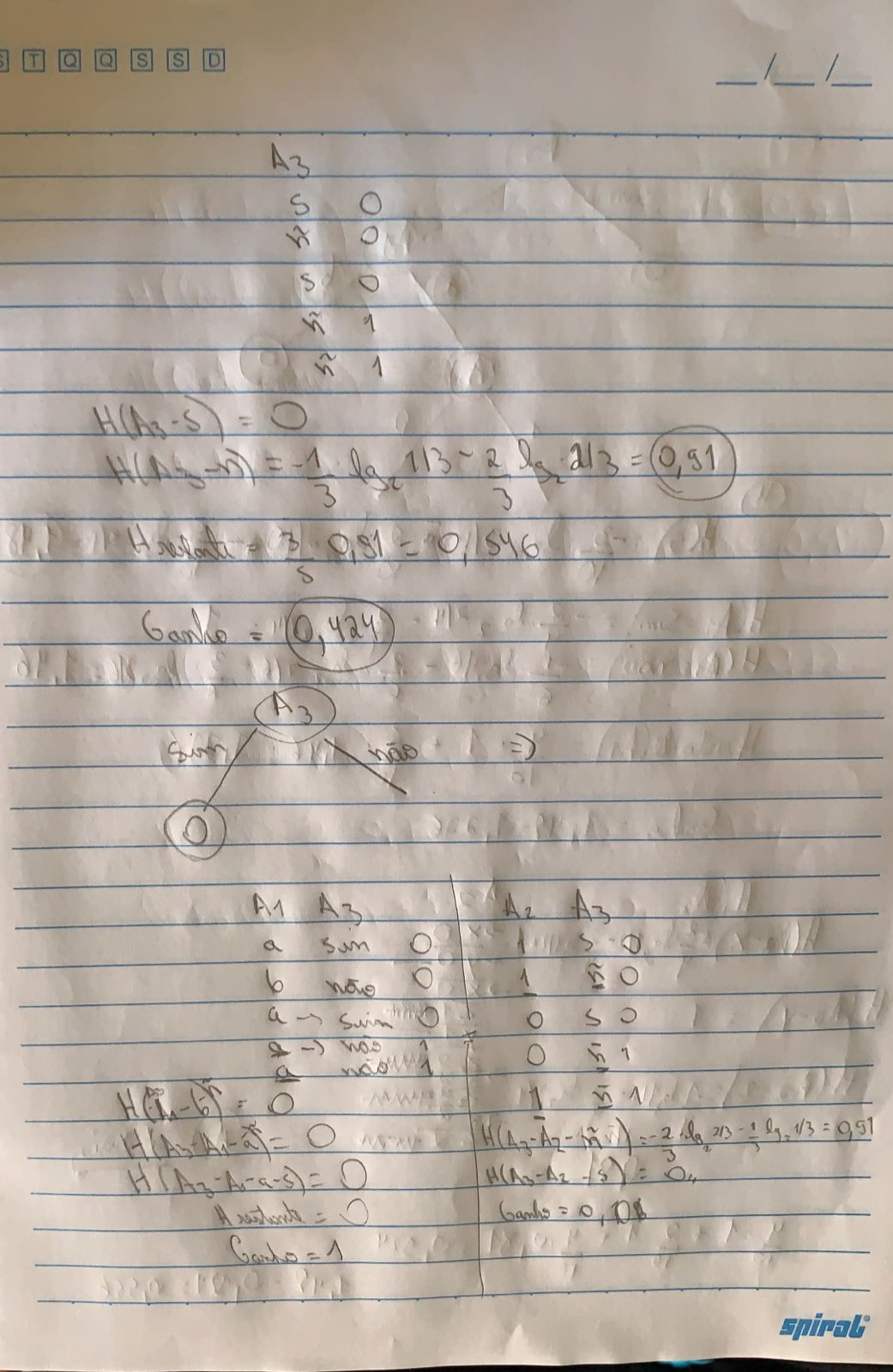
Ganho(Faminto) = H(S) – Hrestante(Faminto) = 1 – 0,72 = 0,28

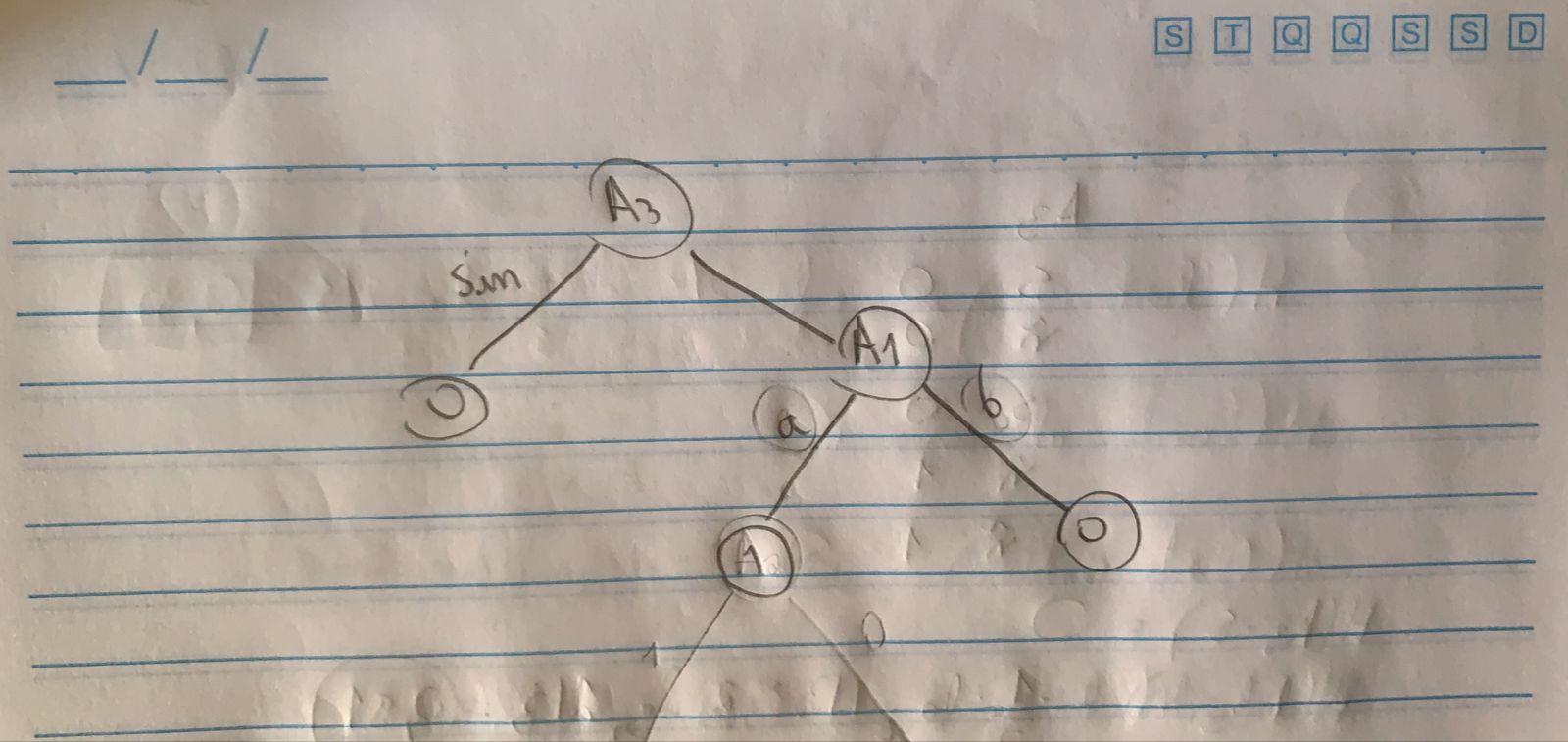


3) Considere o conjunto de dados com três atributos (A1, A2 e A3) e uma classe binária (Classe). Construa uma árvore de decisão com esses dados, usando o algoritmo ID3. Indique, a cada passo, o atributo escolhido em cada nó, os exemplos de cada classe cobertos pelo ramo sendo construído e os cálculos de entropia e ganho de informação que foram necessários durante o processo. Usar a tabela de logaritmos no final da lista.

| Exemplos | A1 | A2 | A3 | Classe |
| --- | --- | --- | --- | --- |
| x1 | a | 1 | sim | 0 |
| x2 | b | 1 | não | 0 |
| x3 | a | 0 | sim | 0 |
| x4 | a | 0 | não | 1 |
| x5 | a | 1 | não | 1 |







4) Considere o conjunto de dados com dois atributos (A1, A2) e uma classe com três valores possíveis (Classe). Supondo que uma árvore de decisão vai ser construída com esses dados, calcule o ganho de informação para cada um dos atributos A1 e A2. Qual atributo seria escolhido para ser atribuído à raiz da árvore? Usar a tabela de logaritmos no final da lista.

| Exemplos | A1 | A2 | Classe |
| --- | --- | --- | --- |
| x1 | sim | x | 0 |
| x2 | não | y | 0 |
| x3 | sim | x | 0 |
| x4 | não | z | 1 |
| x5 | não | y | 1 |
| x6 | sim | z | 2 |
| x7 | não | x | 2 |
| x8 | não | z | 2 |
| x9 | sim | x | 2 |
| x10 | não | z | 2 |

Entropia do conjunto original

H(S) = - (3/10 log2 3/10) – (2/10 log2 2/10) - (5/10 log2 5/10) = - (0,3 log2 0,3) – (0,2 log2 0,2) - (0,5 log2 0,5) = 0,52 + 0,46 + 0,5 = 1,48

H(A1 = sim) = -(0,5. log2 0,5) - 0,5. log2 0,5) = 0,5 + 0,5 = 1

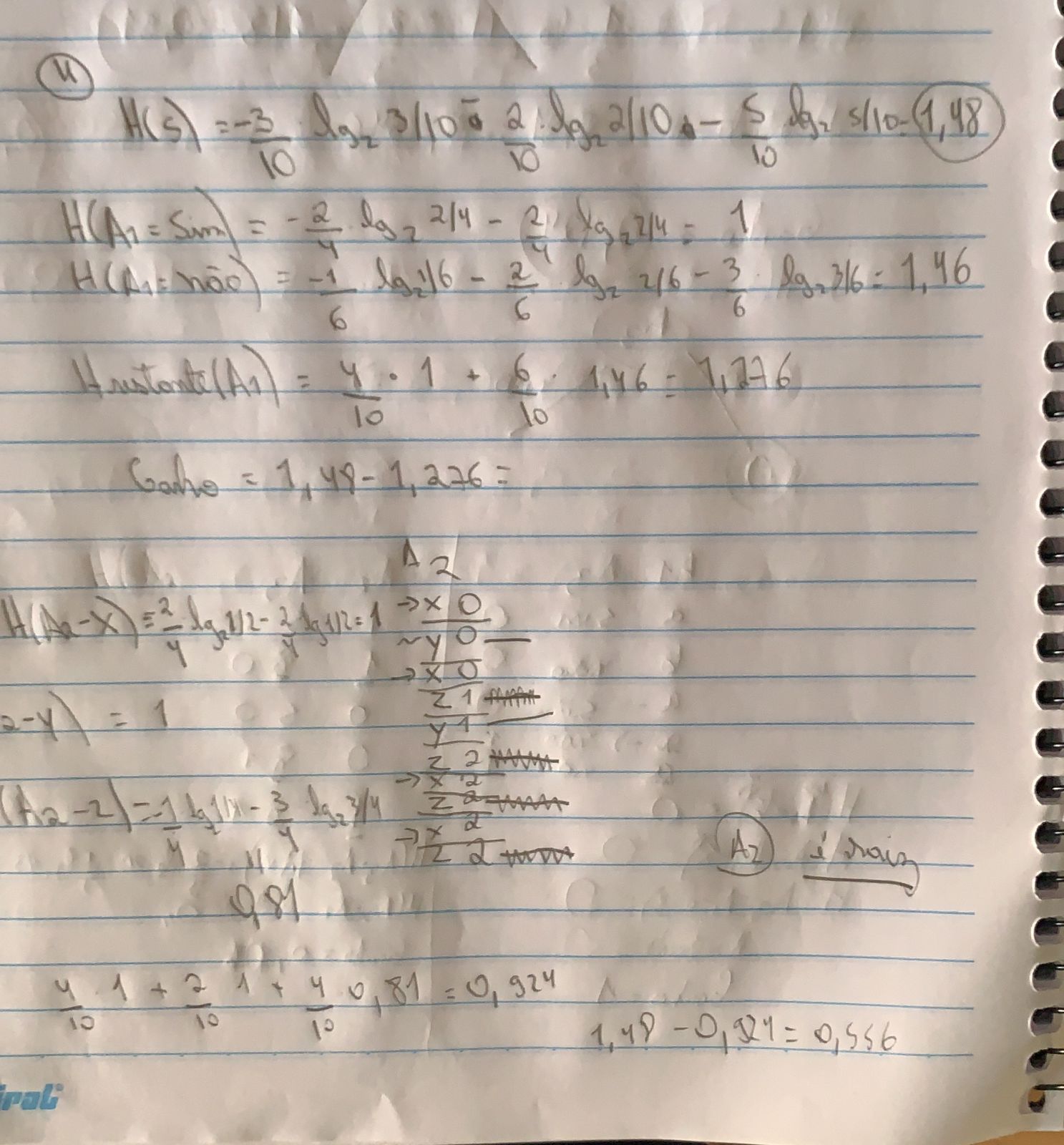
H(A1 = não) = - (1/6. log2 1/6) - (2/6. log2 2/6) - (3/6. log2 3/6) =

= - (0,17. log2 0,17) - (0,33. log2 0,33) - (0,5. log2 0,5) =

= 0,43 + 0,53 + 0,5 = 1,46

HRestante(A1) = 4/10 \* H(A1 = sim) + 6/10 \* H(A1 = não) = 0,4 \* 1 + 0,6 + 1,46 = 0,4 + 0,876 = 1,276 Ganho(A1) = 1,48 – 1,276 = 0,204

(Falta calcular o ganho de A2 e dizer qual dos 2 seria escolhido)



5) Uma árvore de decisão deve ser construída a partir dos dados relacionados na tabela a seguir, usando o algoritmo C4.5. Calcule o ganho de informação dos atributos A1 e A2 e diga qual seria escolhido para ficar na raiz da árvore. Utilize a tabela de logaritmos no final da lista.

| Exemplo | A1 | A2 | Classe |
| --- | --- | --- | --- |
| 1 | a | 1 | 0 |
| 2 | b | 1 | 0 |
| 3 | a | 0 | 1 |
| 4 | b | 0 | 1 |
| 5 | a | 1 | 1 |
| 6 | a | 0 | 1 |

Calcular a entropia do conjunto de dados original S:

E(S) = -1/3\*log21/3 – 2/3\*log22/3 = 0,53 + 0,40 = 0,93

Calcular o ganho de informação de cada atributo:

Atributo A1

Valor a: 3, 5, 6 (classe 1) 1 (classe 0)

Valor b: 4 (classe 1) 2 (classe 0)

H(A1=a) = -0,75\* log20,75 – 0,25\*log20,25 = -0,75 \* -0,31 – (0,25 \* -2) = 0,31 + 0,5 = 0,81 H(A1=b) = -0,5\* log20,5 – 0,5\*log20,5 = 1

E\_Restante(A1) = 4/6 \* 0,81 + 2/6 \* 1 = 0,66 \* 0,81 + 0,33 \* 1 = 0,5346 + 0,33 = 0,8646 Ganho de informação (A1) = E(S) – E\_Restante(A1) = 0,93 – 0,8646 = 0,0654

Atributo A2

Valor 0: 3, 4, 6 (classe 1)

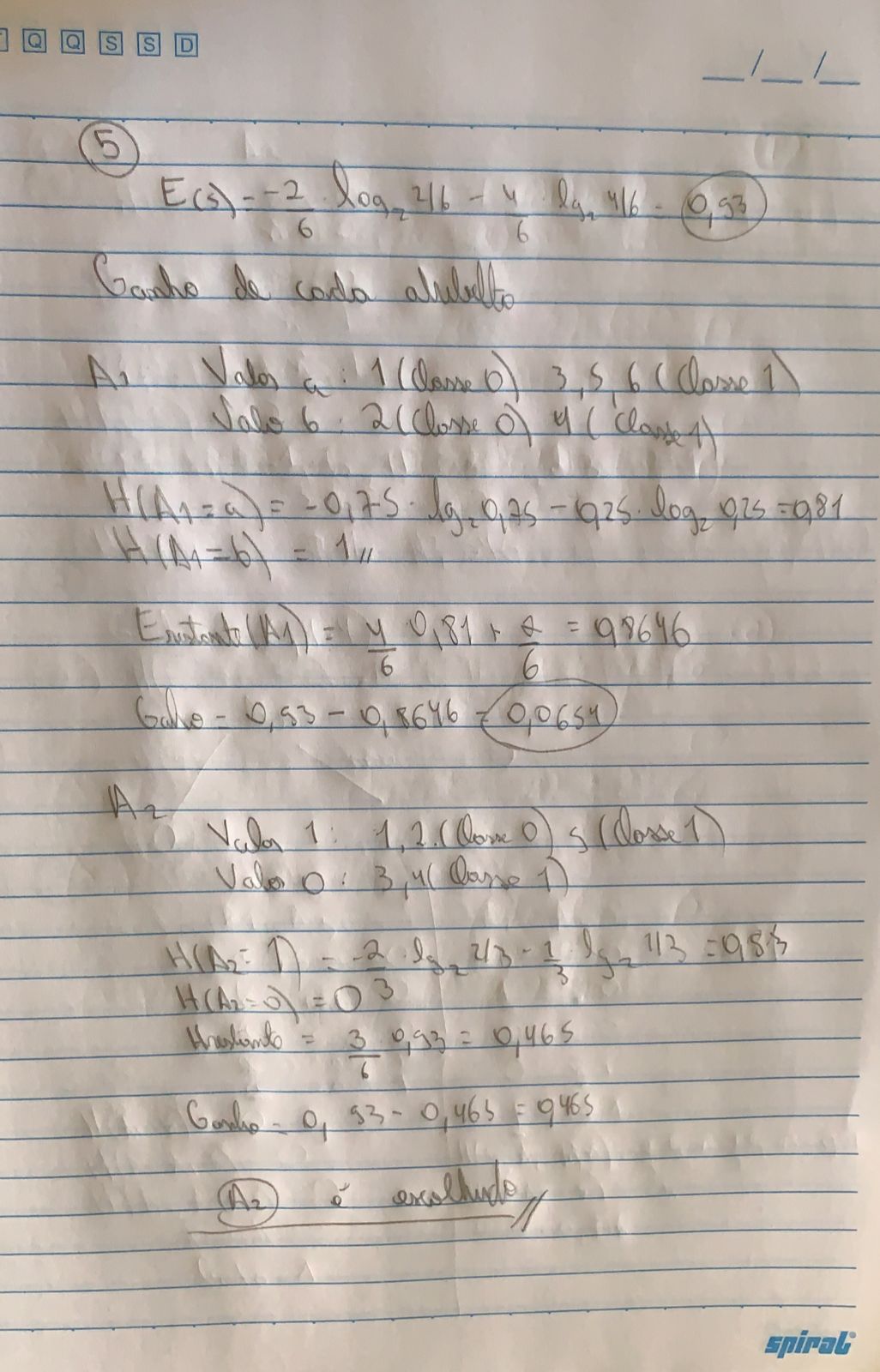
Valor 1: 5 (classe 1) 1,2 (classe 0)

E(0) = -1\* log21 – 0\*log20 = 0 (Todos os exemplos da mesma classe)

E(1) = -0,33\* log20,33 – 0,66\*log20,66 = 0,53 + 0,40 = 0,93

E\_Restante(A2) = 1/2 \* 0 + 1/2 \* 0,93 = 0,465

Ganho de informação (A2) = E(S) – E\_Restante(A2) = 0,93 – 0,465 = 0,465 O atributo escolhido é o atributo A2 que tem o maior ganho de informação.



Informações úteis:

| n | log2n | nlog2n |
| --- | --- | --- |
| 0,17 | -2,56 | -0,43 |
| 0,2 | -2,32 | -0,46 |
| 0,25 | -2,00 | -0,50 |
| 0,3 | -1,74 | -0,52 |
| 0,33 | -1,60 | -0,53 |
| 0,4 | -1,32 | -0,53 |
| 0,5 | -1,00 | -0,50 |
| 0,6 | -0,74 | -0,44 |
| 0,66 | -0,60 | -0,40 |
| 0,75 | -0,42 | -0,31 |
| 0,8 | -0,32 | -0,26 |
| 1 | 0,00 | 0,00 |