DEIS - Departamento de Engenharia Informática e de Sistemas
ISEC - Instituto Superior de Engenharia de Coimbra

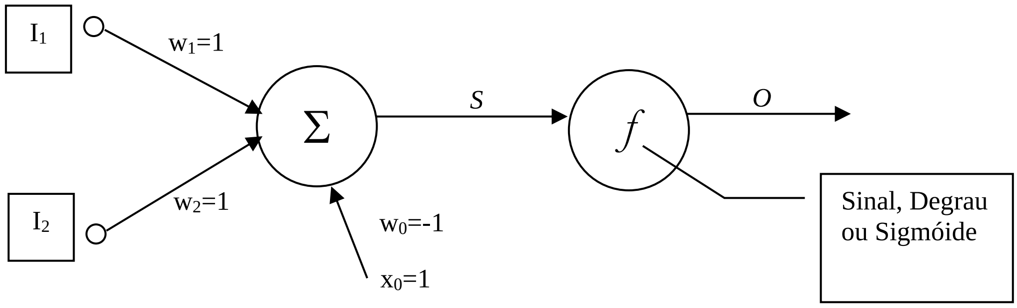
Conhecimento e Raciocínio 2017/2018

# Prática 1 - Redes Neuronais

---

**1.** A figura seguinte representa uma unidade de 2 entradas cujos valores podem ser, apenas, do tipo lógico, isto é, 0 ou 1. Construa uma "tabela de verdade" para a saída *O* quando se trata de:

**a)** Um perceptrão de função sinal
**b)** Um perceptrão de função degrau (de que função lógica se trata ?)
**c)** Uma unidade sigmóide




**1. a)** $I_1=0, I_2=0$ $S = w_1.I_1 + w_2.I_2 + w_0.x_0 = 1\times0+1\times0+1\times(-1) = -1$

$-1<0 \Rightarrow O= -1$

$I_1=0, I_2=1$ $S = 1\times0+1\times1+1\times(-1) = 0$

$0=0 \Rightarrow O= -1$

$I_1=1, I_2=0$ $S = 1\times1+1\times0+1\times(-1) = 0$

$0=0 \Rightarrow O= -1$

$I_1=1, I_2=1$ $S = 1\times1+1\times1+1\times(-1) = 1$

$1>0 \Rightarrow O= +1$

| $I_1$ | $I_2$ | O |
|---|---|---|
| 0 | 0 | -1 |
| 0 | 1 | -1 |
| 1 | 0 | -1 |
| 1 | 1 | +1 |

**b)** O somatório é realizado como em a). Apenas a função *f* muda, produzindo agora uma saída de *O* de valor 0 se $S\leq -w_0$, ou 1 se $S>-w_0$. Assim:

| $I_1$ | $I_2$ | O |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

Como se deduz da tabela, o perceptrão implementa a função lógica AND.

**c)**

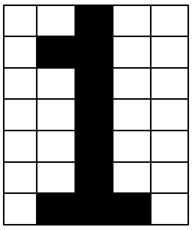
O somatório é realizado como em a). A função $f$ muda, produzindo agora uma saída $O$ de valor

$$O = \sigma(S) = \frac{1}{1+e^{-S}}$$

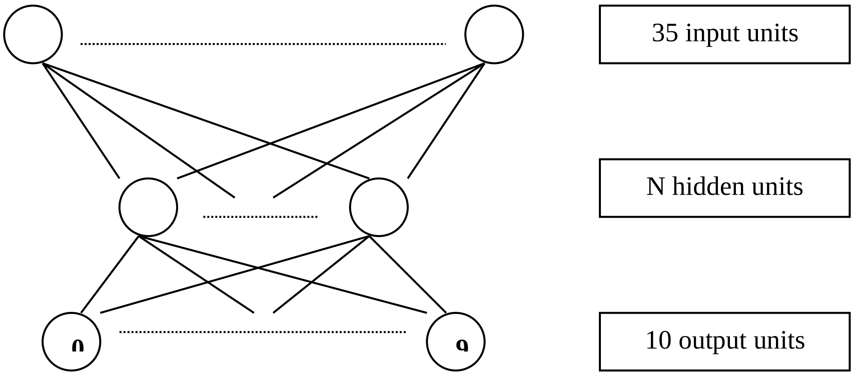
| $I_1$ | $I_2$ | S | Output |
|---|---|---|---|
| 0 | 0 | | $O = \sigma(S) = \frac{1}{1+e^{-(-1)}} = \frac{1}{1+e}$ |
| 0 | 1 | | 1/2 |
| 1 | 0 | | 1/2 |
| 1 | 1 | | $O = \frac{1}{1+e^{-}1}$ |

**2.** Pretende implementar-se uma rede neuronal para a identificação dos algarismos "0" a "9". Suponha que cada algarismo é representados numa matriz de 7 linhas por 5 colunas.

**a)** Desenhe uma rede cuja topologia lhe pareça adequada ao fim em vista e enumere as suas principais características, tais como: Tipo de unidades utilizadas, número de camadas, etc.
**b)** De acordo com a alínea a), para o algarismo representado em seguida quais seriam as entradas aplicadas à rede e as saídas pretendidas ?



**2. a)** Uma rede possível utilizaria p.e. 35 unidades de entrada "ligadas" a cada pixel da matriz 7*5, algumas unidades escondidas (p.e. 5) e 10 unidades de saída, cada uma para identificar cada algarismo. As unidades deveriam ser do tipo sigmoide, porque permitem a aprendizagem de funções mais complexas. Nesta topologia, cada algarismo poderia ser representado por um array de 35 números: "0" para os pixeis brancos e "1" para os pixeis negros que desenham o algarismo.




**b)** Entrada: [00100 01100 00100 00100 00100 00100 01110]
Alvo: [0100000000]

**3.** Um perceptrão de 3 entradas, *threshold=0* e saída dada pela função sinal, é treinado com os exemplos 110 e 111, cujas saídas alvo são respectivamente -1 e 1.

**a)** Terminado o treino, a saída do perceptrão será -1 quando e só quando a sequência apresentada para identificação for:

(a) 000 or 110 or 011 or 101
(b) 010 or 100 or 110 or 101
(c) 000 or 010 or 110 or 100

Justifique o seu raciocínio

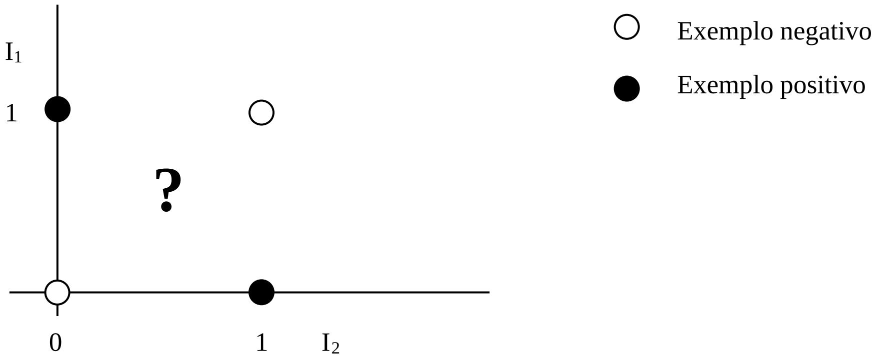
**b)** Qual o tipo de números que o perceptrão está habilitado a reconhecer como exemplos positivos ?

**3. a)** Designemos por $x_2x_1x_0$ os bits que compõem os exemplos apresentados ao perceptrão. Como no treino são utilizados apenas dois exemplos que diferem entre si apenas no bit menos significativo sendo os restantes bits 1, o perceptrão terá de ajustar os coeficientes de ponderação $w_2$ e $w_1$ para valores próximos de 0, dando a $w_0$ um valor alto de modo a poder distinguir o exemplo 110 do exemplo 111.
Nestas condições, a única sequência que produzirá **sempre** a saída -1 é a (c) porque para ela os bits $x_0$ de todos os números têm valor 0 (o valor dos restantes será "anulado" pelos coeficientes $w_2$ e $w_1$).
Além disso, os números da sequência (c) são todos os que podem produzir saída 0 (não há mais nenhum de 3 bits capaz de o fazer porque terá um bit $x_0$ de valor 1 e portanto produzirá a saída +1). Logo, **só** os exemplos da sequência (c) poderão produzir uma saída -1.

**b)** De acordo com o exposto em a) este perceptrão será capaz de reconhecer (como exemplos positivos) os **números ímpares** de 0 a 7.

---

**4.** A figura seguinte mostra 4 exemplos de treino para uma rede neuronal baseada em perceptrões de função degrau, à qual são aplicadas as entradas $I_1$ e $I_2$.

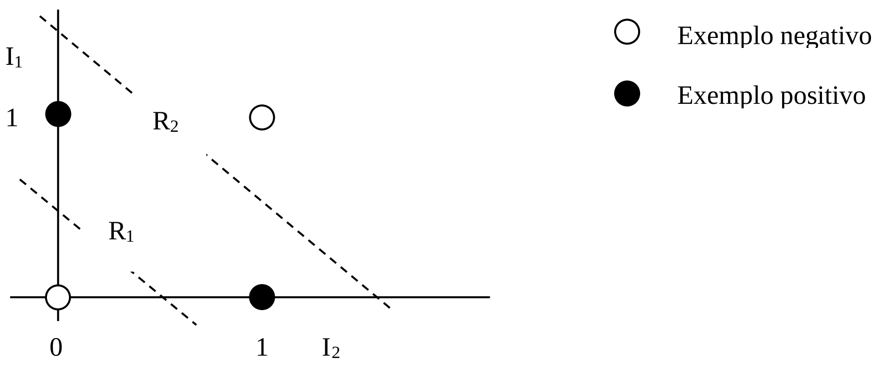



**a)** Qual a função lógica que se pretende que a rede aprenda ? Justifique.
**b)** Desenhe as possíveis superfícies de decisão.
**c)** Projecte uma rede neuronal que possa aprender esta função, assinalando os valores de todos os coeficientes de ponderação e limites (*thresholds*) intervenientes.

**4. a)** A função lógica é o XOR porque os exemplos positivos são os correspondentes a entradas $I_1$ e $I_2$ de diferente valor lógico, isto é:

| I1 | I2 | Output | |
|---|---|---|---|
| 0 | 0 | 0 | |
| 0 | 1 | 1 | (positivo) |
| 1 | 0 | 1 | (positivo) |
| 1 | 1 | 0 | |

**b)** Dado que a rede se baseia em perceptrões, as superfícies de decisão são lineares. Além disso, estas superfícies separam os exemplos positivos dos negativos. Sendo assim são necessárias duas rectas localizadas conforme representado em seguida ou aproximadamente nessas zonas:




**c)** Cada unidade consegue apenas implementar uma superfície de decisão linear (exemplos linearmente separáveis). Como no desenho figuram 2 retas, isto sugere que se utilizem 2 unidades. Contudo, 2 unidades dão origem a 2 saídas, e pretende-se apenas uma, cujo valor será 0 ou 1 consoante as entradas aplicadas. Uma solução consistirá em combinar as saídas das 2 unidades através de uma terceira que dispare quando ambos as anteriores derem uma saída positiva. Ou seja, uma unidade a funcionar como um "AND".

A equação da superfície de decisão é:

$$x_2 = -\frac{w_1}{w_2} . x_1 - \frac{w_0}{w_2}$$

Que é da forma *y=mx+b* (ou seja, uma reta)

Reta R1: Declive m=-1 => $-\frac{w_1}{w_2} = -1$ Por exemplo, $w_1$=1 e $w_2$=1

Ordenada na origem b=+0.5 $-\frac{w_0}{w_2} = +0.5$ Como $w_2$=1, então $w_0$=-0.5

(de facto esta unidade funciona como um "OR" e este são os coeficientes para esta função)

Reta R2: Declive m=-1

Mas nesta reta pretendemos que os exemplos positivos fiquem abaixo dela (e não acima, como habitualmente). Que fazer?

Declive m=-1 => $-\frac{w_1}{w_2} = -1$

Ora, note-se que além da solução acima, existe outra: $w_1$=-1 e $w_2$=-1

Donde ordenada na origem b=+0.5 $-\frac{w_0}{w_2} = +1.5$ Como $w_2$=-1, então $w_0$=+1.5

De facto, se e testar esta solução verifica-se que cumpre o pretendido:

$s = w_1 x_1 + w_2 x_2 + w_0$ Se s>0 então out=+, senão out=-

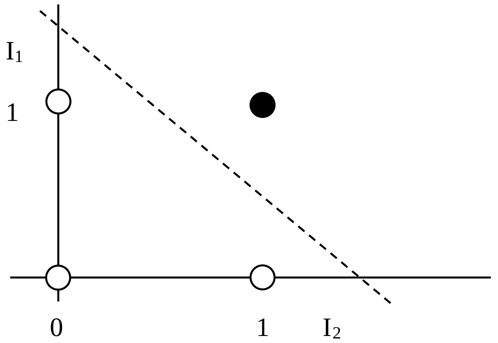
| $x_1$ | $x_2$ | s | o |
|---|---|---|---|
| 0 | 0 | 1.5 | 1 |
| 0 | 1 | 0.5 | 1 |
| 1 | 0 | 0.5 | 1 |
| 1 | 1 | -0.5 | 0 |

E portanto esta configuração da unidade cumpre o pretendido (funciona como um "NAND")

NOTA - Regra Prática: se se pretender trocar o "lado" positivo de uma superfície de decisão com o lado negativo, basta multiplicar todos os coeficiente sinápticos por -1

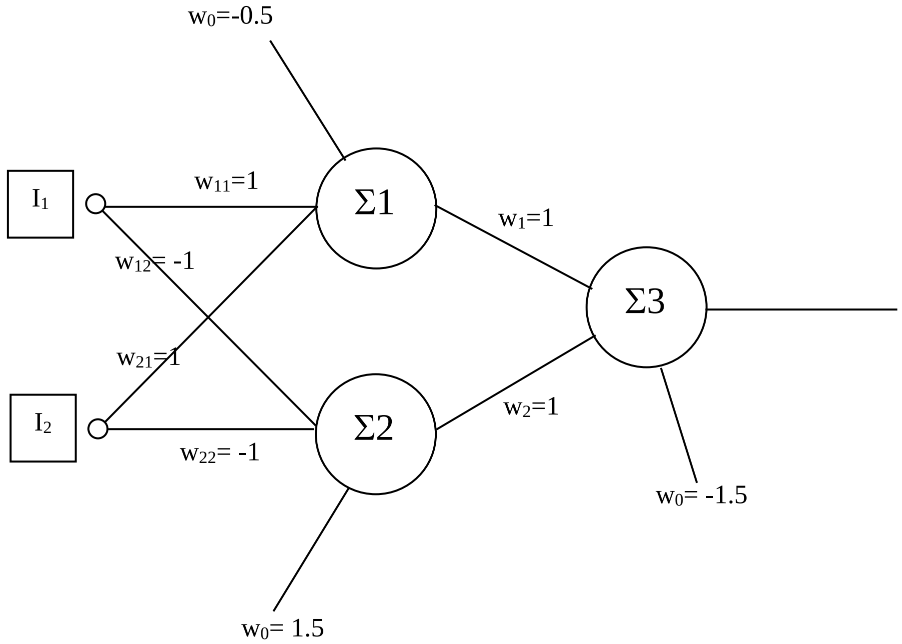
A 3ª unidade, que junta as saídas das anteriores, é, pelo acima exposto, um "AND", cuja configuração é conhecida (ver figura seguinte)

$w_1$=1 e $w_2$=1 e $w_0$=-1.5




Para um AND, representado na figura ao lado, a condição é I1+I2>1.5, o que pode ser conseguido com um perceptrão de coeficientes 1 e *threshold*=-1.5.

Portanto, a configuração da rede seria:




Façamos o teste final: Sejam O1, O2 e O3 as saídas dos perceptrões P1, P2 e P3, respectivamente. Como são de função degrau, a sua saída é 0 ou 1. Então:

Com **I1=0** e **I2=0** resulta:

U1: Como a soma de I1 e I2 é 0, é inferior a $-w_0$=-(-0.5), e portanto O1=0
U2: Como a soma de I1 e I2 é 0, é superior a $-w_0$=-(+1.5), e portanto O2=1
U3: As entradas de P3 serão 0*1 + 1*1=1. Como 1<-(-1.5), **O3=0**

Com **I1=0** e **I2=1** resulta:

U1: 0*1 + 1*1 = 1. Como 1>0.5, O1 = 1
U2: 0*-1 + 1*-1 = -1. Como –1>-1.5, O2=1
U3: 1*1 + 1*1 = 2. Como 2>-(-1.5), **O3=1**

Com **I1=1** e **I2=0** resulta:

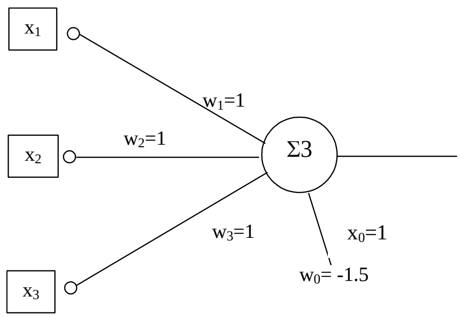
Análogo ao anterior (simétrico), logo **O3=1**

Com **I1=1** e **I2=1** resulta:

U1: 1*1 +1*1 = 2. Como 2>0.5, O1 = 1
U2: 1*-1 + 1*-1 = -2. Como –2<-1.5, O2=0
U3: 1*1 + 0*1 =1. Como 1<-(-1.5), **O3=0**

E portanto o comportamento da rede é exatamente a de um XOR: ela segue a tabela de verdade apresentada em a).

**5.** Considere a unidade representada na figura seguinte:




**a)** Desenhe a sua superfície de decisão

$w_1x_1 + w_2x_2 + w_3x_3 \geq -w_0x_0$

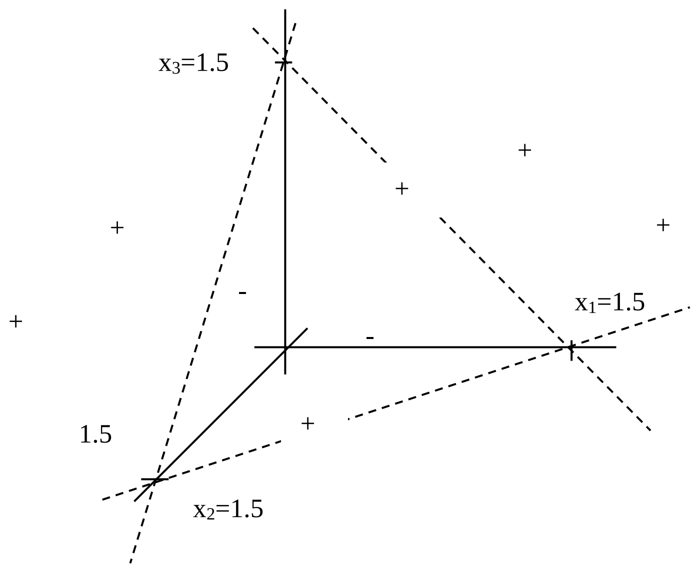
Como, pela figura, $w_{1..3}=1$ e $w_0.x_0=-1.5$, tem-se:

$x_1 + x_2 + x_3 \geq 1.5$

A superfície de decisão é dada por

$x_1 + x_2 + x_3 = 1.5$

Trata-se da equação de um plano (fórmula geral $ax+by+cz=d$) que intersecta os eixos x, y, z no ponto 1.5




**b)** De entre os valores representados na figura ($x_{1...3}$, $w_{0...3}$) quais os que influenciam a superfície de decisão? De que modo?

$w_1$, $w_2$ e $w_3$ determinam a inclinação do plano
O grupo $w_0.x_0$ determina a distância à origem dos eixos

**c)** Considerando $x_0=1$ haverá algum limite quanto ao número máximo de entradas negativas que um exemplo pode conter para ser classificado como positivo? Porquê?

Para um exemplo ser positivo tem de cumprir a condição $x_1 + x_2 + x_3 \geq 1.5$
Logo, 3 entradas negativas nunca poderão constituir um exemplo positivo.
Contudo, 2 entradas negativas e uma positiva poderão constituir um exemplo positivo sse

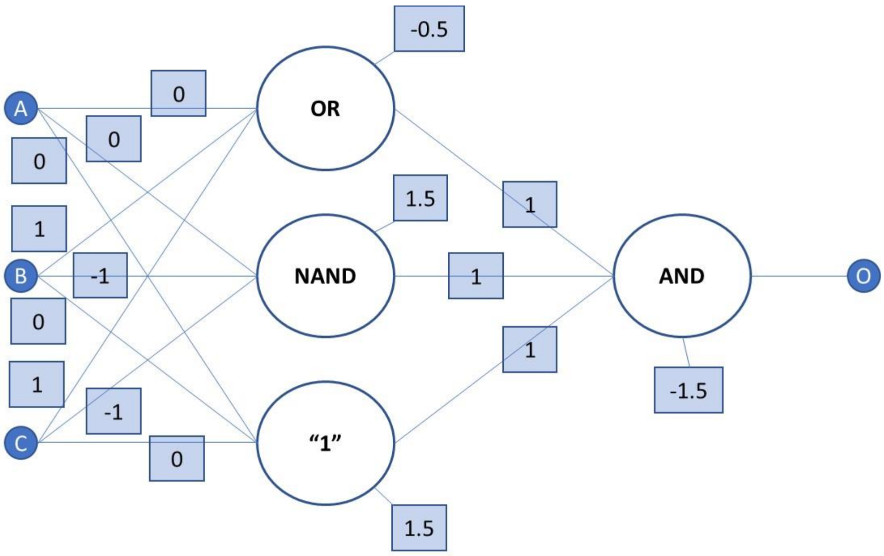
$x_1 \geq 1.5 - x_2 - x_3$ ou $x_2 \geq 1.5 - x_1 - x_3$ ou $x_3 \geq 1.5 - x_1 - x_2$

Logo, o número máximo de entradas negativas para um exemplo poder ser considerado como positivo, é de duas.

---

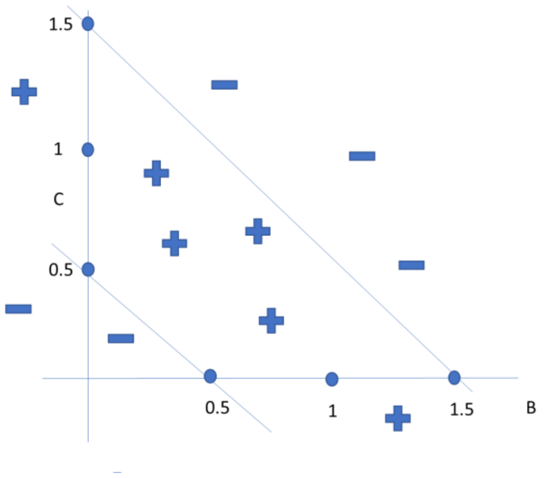
**6.** **a)** Dado o seguinte dataset, complete o diagrama seguinte indicando todos os valores de *w* e as funções lógicas que cada unidade deve realizar:

| A | B | C | O |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 |
| 0 | 1 | 0 | 1 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 |
| 1 | 1 | 1 | 0 |




NOTA1: o input A é irrelevante. O bias da unidade "1" poderia ter outro valor qq desde que >0. O bias do AND poderia ser inferior (-1.5 é o mínimo aceitável)

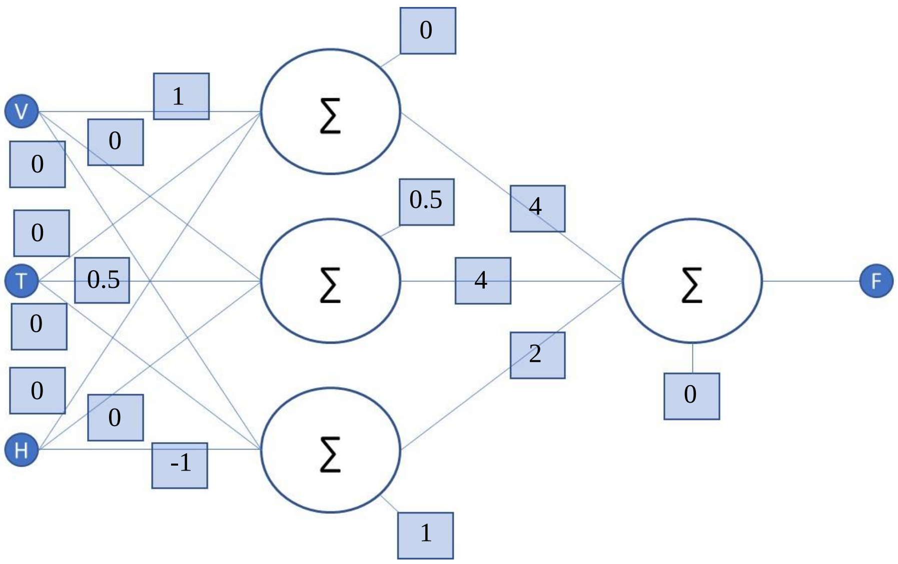
**b)** Desenhe a superfície de decisão desta rede neuronal



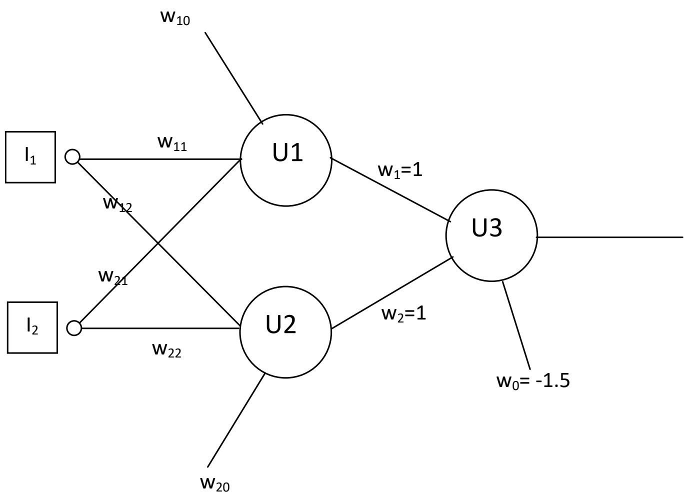

NOTA: A entrada A não foi representada uma vez que não interessa. Se fosse representada, então teríamos outro eixo A e as retas seriam planos (paralelos a esse eixo). Neste caso o desenho teria de ser feito em perspetiva

---

**7.** Para a previssão da intensidade de fogos florestais em função do vento, temperatura e huidade, implementou-se uma rede neuronal de 3 entradas: V, T e H, respetivamente vento, temperatura e humidade. O vento varia no intervalo [0,1], a temperatura em [-1,1] e a humidade em [0,1]. Pretende obter-se uma saída F que represente a intensidade do fogo no intervalo [0,10]. Neste modelo o vento deve ter um peso de 2, a temperatura de 2 e a humidade de 1. Complete o diagrama seguinte indicando os valores de todos os coeficientes sinápticos e considerando que todas as unidades são lineares.



**8.** Considere a seguinte rede neuronal:




**a)** A que condição devem obedecer os coeficientes sinápticos de U1 e U2 para que as suas superfícies de decisão sejam paralelas? (apresente apenas as equações de funcionamento de U1 e U2, a condição que garante o resultado pretendido e a justificação do seu raciocínio)

EQUAÇÕES: $\begin{cases} w_{11}I_1 + w_{21}I_2 + w_{10} = 0 \\ w_{12}I_1 + w_{22}I_2 + w_{20} = 0 \end{cases}$

Basta saber que o declive da superficie de decisão das unidades é dado pelo quociente entre os coef. sinápticos. Posto isto, basta escrever as eq. básicas e obter o valor desse declive para cada unidade, igualando-os conforme pedido

CONDIÇÃO: $\frac{w_{11}}{w_{21}} = \frac{w_{12}}{w_{22}}$

Justificação: Representando I2 em função de I1, o declive das superfícies de decisão é dado pelos quocientes apresentados na CONDIÇÂO acima, pelo que para as rectas sejam paralelas, basta garantir que sejam iguais.

**b)** E, para além disso, cruzarem os eixos a uma distância de 2 unidades entre si?

CONDIÇÃO: $\left| \frac{w_{10}}{w_{21}} - \frac{w_{20}}{w_{22}} \right| = 2$

De forma semelhante, basta saber que para a ordenada na origem é $w_0$ que importa. Apenas haveria que ter em linha de conta a utilização do módulo da diferença, dado que nada é dito acerca de qual das retas deve aparecer "por cima" da outra

Justificação: Nas condições da justificação dada em a), as ordenadas na origem das superfícies de decisão são dadas pelos quocientes da CONDIÇÃO acima, pelo que para garantir que as retas distam de 2 unidades basta garantir que o módulo da diferença entre estas ordenadas é igual a 2 (se são paralelas, o cruzamento em X será também a 2 unidades de distância entre elas)