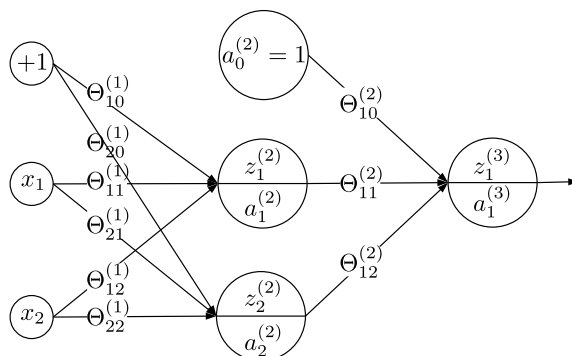


1 Redes Neurais I

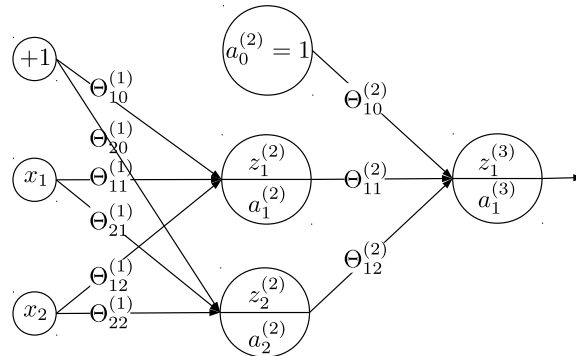
1. Comente sobre a motivação computacional e a biológica que levaram ao conceito de rede neural. Quais as principais características do “algoritmo” do cérebro que inspiraram essas redes?
2. Comente sobre a hipótese do algoritmo de aprendizado único no cérebro e algumas aplicações do conceito na medicina.
3. Faça o esquema, como no slide de aula, dos neurônios artificiais de McCulloch-Pitts, Perceptron e Logístico.
4. Defina os conceitos de “unidade de ativação”, “peso” e “bias” no neurônio logístico.
5. Construa uma pequena rede de neurônios logísticos como vistos em aula (uma camada escondida com duas unidades) que resolva o problema x_1 XOR x_2 . Escolha pesos e biases diferentes do slide da aula e mostre também o valor na saída de cada unidade de ativação.
6. Descreva o processo para se chegar a uma rede neural partindo de uma regressão logística.
7. Considere o diagrama a seguir. Defina nele o que seriam as camadas de



entrada, saída e escondida, $a_i^{(j)}$, $z_i^{(j)}$, $\Theta_{ab}^{(j)}$ e quais seriam os biases.

8. Considere a rede neural a seguir com as entradas e pesos/biases fornecidos na tabela. Calcule a saída da rede, mostrando os cálculos intermediários. NOTA: Esta é exatamente a rede que está no slide. Procure resolver sem consultar e depois use a resposta no slide para checar.

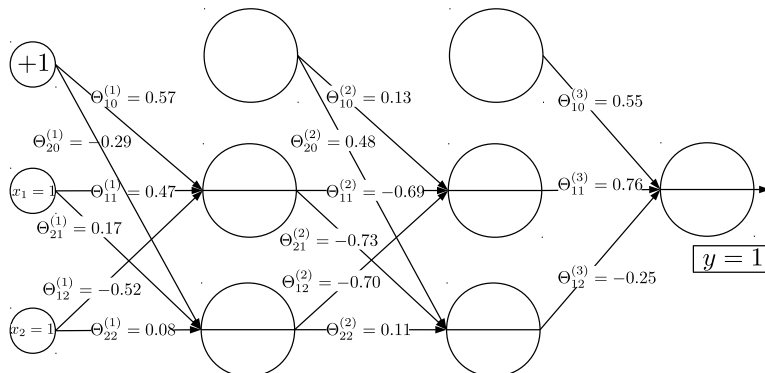
$x_1 = 0.5$	$x_2 = 1.0$				
$\Theta_{10}^{(1)} = -0.1$	$\Theta_{20}^{(1)} = 0.2$	$\Theta_{11}^{(1)} = 1.3$	$\Theta_{21}^{(1)} = -0.5$	$\Theta_{12}^{(1)} = 0.6$	$\Theta_{22}^{(1)} = 0$
$\Theta_{10}^{(2)} = 2.1$	$\Theta_{11}^{(2)} = -1.5$	$\Theta_{12}^{(2)} = -0.3$			



9. Mostre o pseudo-código do algoritmo vetorizado que calcula a saída da rede (*forward*) com uma camada escondida. Qual adaptação seria feita para generalizar para mais camadas?
10. Comente sobre a intuição segundo a qual uma rede neural vai aprendendo novos atributos em cada camada. Qual a característica principal que distingue atributos aprendidos em diferentes camadas?
11. Descreva como a abordagem “um-*vs*-todos” é usado nas redes neurais multi-classes.

2 Redes Neurais II

1. Defina matematicamente a função de custo de uma rede neural multi-classes. Defina os termos que aparecem na expressão.
2. Considere a rede neural a seguir. Faça uma atualização dos pesos/biases usando *backpropagation* e mostrando os cálculos intermediários. NOTA: Este é exatamente o exemplo do slide. Procure resolver sem consultar e depois use a resposta no slide para checar.



3. Descreva, com suas palavras e usando a notação e as equações da aula, o passo-a-passo do algoritmo de *backpropagation* para calcular o gradiente sobre um exemplo de treinamento.
4. Mostre o pseudo-código do *backpropagation* para m exemplos, vetorizando o que for possível, conforme o que vimos em aula.
5. Comente sobre as implicações de se usarem pesos e biases com o mesmo valor na inicialização. Como isto pode ser resolvido?
6. Qual a importância da checagem do gradiente no *backpropagation*? Por que esta etapa deve ser removida da versão final do algoritmo?
7. Descreva matematicamente o cálculo do gradiente numérico.
8. Descreva com suas palavras o passo-a-passo da definição e treinamento de uma rede neural, desde a definição da arquitetura até o gradiente descendente.
9. Comente sobre as implicações do fato de a função de custo de uma rede neural ser altamente não-convexa.
10. Descreva as adaptações mais usuais para o uso de uma rede neural em problemas de regressão.