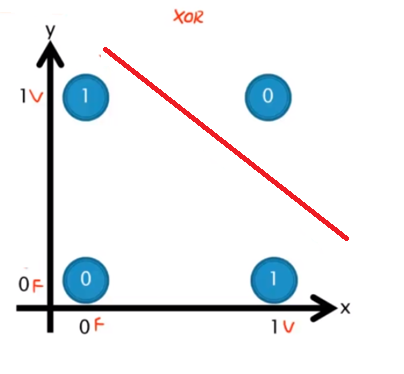
1) Tendo em vista um plano cartesiano com duas classes, quando passamos uma reta sobre essas duas classes não conseguimos separar as mesmas, isso é um problema não linearmente separável. No exemplo da porta XOR, temos:



0 0 = 0 (falso)

0 1 = 1 (verdadeiro)

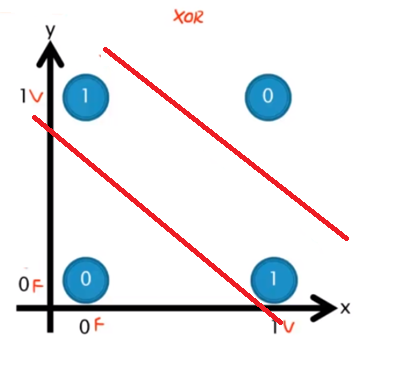
1 0 = 1 (verdadeiro)

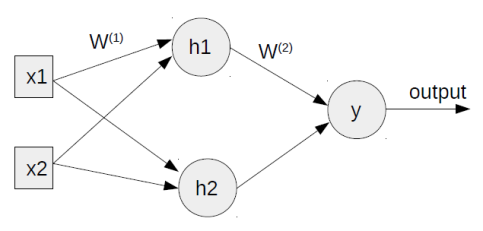
1 1 = 0 (falso)

Ou seja, quando passamos uma reta entre as classes 0 e 1 tendo em vista a porta lógica XOR, vemos que não é possível separar as mesmas.

Um Perceptron de Rosenblatt (1958) consegue resolver somente funções linearmente separáveis, uma vez que o mesmo não consegue gerar um hiperplano para separar os dados.

Sendo assim, para resolver problemas dos que não são linearmente separáveis temos que usar redes mais complexas, como a multilayer perceptron, inserindo camadas ocuttas, as quais os neurônios possuem funções de ativação não lineares. Portanto, se eu passar duas retas (uma reta que resolve o problema da porta lógica AND e uma reta da porta lógica OR) no gráfico acima da porta XOR, que representam, cada um, um neurônio da camada o culta, aliada a uma função de ativação, podemos resolver o este problema não linear, como representado pelo gráfico abaixo.





2) O gradiente de descida é usado na manipulação de ajustes de pesos, com intuito de encontrar a combinação de pesos que o erro seja o menor possível. A ideia é que o gradiente seja o mínimo possível, para isso usamos o cálculo do gradiente. Temos erros iniciais, vai aplicando o cálculo do gradiente até encontrar o menor ponto de gradiente possível, o chamado mínimo global. Portanto, à medida que os pesos são ajustados, o gradiente vai descendo.

O delta da camada escondida é dada pela multiplicação da derivada da função de ativação do neurônio da camada oculta, peso posterior e o delta de saída(o delta de saída deve ser calculado antes do delta da camada de saída.).

Para calcular o Delta de saúda, calculo a derivada da função de ativação, logo após eu calculo do parâmetro delta de saída, sendo este delta de saída = erro(resultado esperado - resultado previsto)\*derivadaDaFunçãodeAtivação.

Por fim, faz-se necessário encontrar a direção até o mínimo global.