

Lista de Exercícios 09 - AM2

Assunto: Redes Neurais

Professor: Murilo Naldi

1 - Qual é o objetivo do parâmetro *bias* em uma rede neural?

- a) Ajustar o limite de ativação e transladar a combinação das entradas
- b) Reduzir o número de iterações de treinamento necessárias
- c) Ajudar a reduzir o *overfitting* do modelo
- d) Aumentar a precisão do modelo
- e) Permitir o uso de múltiplas camadas

2 - O que é o algoritmo de retropropagação em redes neurais?

- a) Uma técnica usada para determinar o número ideal de camadas ocultas em uma rede neural
- b) Uma técnica usada para otimizar os pesos do modelo
- c) Uma técnica usada para determinar a função de ativação ideal para usar
- d) Uma técnica usada para determinar o número ideal de neurônios em uma camada

3 – Explique como funciona o algoritmo de retropropagação e porque ele é tão importante para as redes neurais?

4- O que é a Regra Delta? Por que a regra de treinamento utilizada na retro-propagação recebe o nome de Regra Delta Generalizada?

5- O que é uma função de ativação? Cite alguns exemplos de funções deste tipo e suas características.

6 – Porque a função de ativação Relu é preferida entre as funções de ativação

6 - Qual tipo de rede neural é largamente usada para tarefas de classificação de imagem?

- a) Rede neural recorrente
- b) Rede neural autoencoder
- c) Rede neural convolucional
- d) Rede neural perceptron multicamada

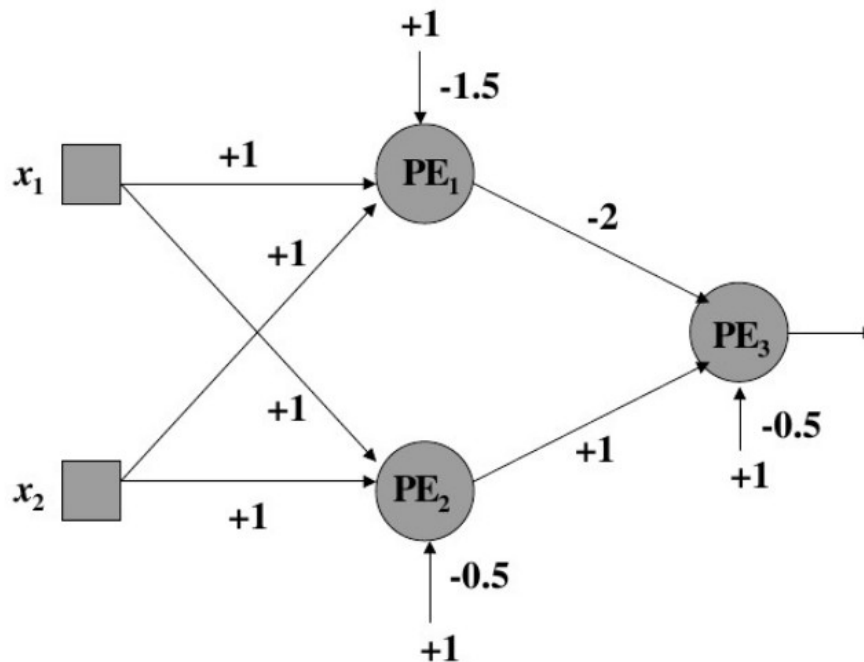
7 - Qual é o objetivo do *dropout* em uma rede neural?

- a) Aumentar o número de iterações de treinamento necessárias
- b) Reduzir o número de neurônios no modelo
- c) Aumentar o número de camadas ocultas no modelo
- d) Evitar o *overfitting* do modelo

8 – Explique usando fórmulas quais as diferenças dos cálculos de ajuste dos pesos da camada de saída e das camadas intermediárias de uma MLP.

9 – Suponha que os padrões de entrada de uma rede MLP são divididos em 2 classes e que alguns padrões de uma classe são muito semelhantes a alguns padrões da outra classe. Neste caso você preferiria utilizar uma taxa de aprendizado grande ou pequena no treinamento? Por que? E se os padrões de ambas as classes forem bastante diferentes?

10 - Mostre que a rede MLP abaixo com 3 neurônios resolve o problema do XOR (ou exclusivo) construindo (a) regiões de decisão (interpretação gráfica) e (b) uma tabela verdade para a rede (propagação dos estímulos de entrada).



11. Como é definido o gradiente local e porque ele é importante?
12. Explique o Teorema da Aproximação Universal e como ele é importante para aplicações de aprendizado de máquina?
13. O que é aprendizado profundo e porque ele pode reduzir o número de neurônios necessários para aproximar uma função?
14. Durante o aprendizado de uma rede profunda, o que causa o viés indutivo e o que causa variância? Existe uma relação entre os dois? Como regularizar?
15. O que é a fase de pré-treino em redes profundas? Qual é o seu objetivo?
16. Como funciona a técnica *Dropout* e qual seu objetivo?
17. Uma taxa de aprendizado muito baixa pode fazer com que uma rede profunda demore muitas iterações para aprender e, conseqüentemente, pode causar sobreajuste. Uma taxa muito grande aumenta a variância e pode tornar a rede instável. Como o algoritmo *Adagrad* e seus derivados tentam solucionar o problema de selecionar um bom valor para a taxa de aprendizado?
18. O que torna as redes convolutivas tão vantajosas para reconhecimento de imagens, em relação a outras redes profundas? Podemos obter tal vantagem para outros tipos de dados? Quais? E porque?
19. Convoluções são feitas por filtros aplicados em um conjunto de dados ao mesmo tempo. Isso pode ser vantajoso se comparado a uma rede completamente conectada. Porque?
20. O que é a técnica de *padding* e como ela ajuda no processo de convolução? E o *stride*? Como essas duas técnicas interferem no tamanho da saída?

21. Dados que possuem múltiplos canais de entrada precisam de um *kernel* capaz de concatenar os resultados. Como isso pode ser feito e qual o tamanho final do *kernel*?
22. Uma vez feito o processo de convolução, é preciso agregar os resultados. Qual técnica utilizamos para fazer tal agregação e como ela é feita?
23. Existem diferentes tipos e variedades de redes convolucionais. Entretanto, a maioria possui uma estrutura topológica (tipos e organização das camadas) comum. Quais são as principais características desta topologia e quais são as funcionalidades de cada camada.