Lista 12

Celso Henrique de Souza Lopes

Exercício 1)

Para que uma rede MLP pudesse ser treinada corretamente, em 1986 D. E. Rumelhart e seus colegas, fizeram uma alteração fundamental na arquitetura do perceptron: substituíram a função step pela função logística (ou sigmoide). Isso foi essencial porque a função step contém apenas segmentos planos, portanto, não há gradiente com o qual se trabalhar (i.e., o algoritmo do gradiente descendente não pode se mover em uma superfície plana, ou seja, com gradiente igual a zero), enquanto isso, a função logística possui derivada diferente de zero e bem definida em todos os pontos, permitindo que o algoritmo do gradiente descendente faça progresso a cada passo.

Exercício 3)

Utilizando o exemplo activation\_functions.ipynb como base, cite três funções de ativação diferentes das que vimos. Plote a função e sua derivada.

Exercício 3)

a) X tem dimensão 10x1.

b) Wh tem dimensão 10x50 e bh tem dimensão 50x1.

c) Wo tem dimensão 50x3 e bo tem dimensão 3x1.

d) Y tem dimensão 3x1

e) f(.) -> função de ativação ReLu. y = f( Wo\*f(Wh\*X + bh) + bo )

Exercício 4)

Para classificar emails em spam ou ham é necessário apenas 1 neurônio, pois a saída é binária. A função de ativação mais apropriada é a função logística. Para classificar a base de dados MNIST são necessários 10 neurônios, pois cada saída de cada neurônio corresponde à probabilidade estimada da classe correspondente. A camada de saída é tipicamente modificada substituindo as funções de ativação individuais por uma função softmax, que retorna um valor probabilístico entre 0 e 1.

Exercício 5)

Passo de aprendizagem e momento. Se o modelo estiver sobreajustando, pode-se variar esses parâmetros, de forma que eles sejam ótimos.