

Exercício 3 - Perceptron Simples

Matheus Bitarães de Novaes

08 de Dezembro de 2021

1 Introdução

Esta atividade abordará a construção de uma rede neural que soma elementos das redes RBF e ELM.

As bases de dados utilizadas são a *Breast Cancer* e a *Statlog(heart)*.

Primeiramente o modelo ELM será implementado e avaliado nos dois *datasets*. O modelo ELM consiste em adicionar uma camada intermediária de neurônios, com pesos aleatórios, com o objetivo de aumentar a dimensão do conjunto de entradas e obter-se um conjunto linearmente separável após treinamento dos pesos da camada de saída.

Após implementação do modelo, os *datasets* foram divididos em treinamento e teste (numa proporção de 30%), normalizados (para ficarem entre 0 e 1) e executados 50 vezes. o número de neurônios na camada escondida foi de 15 (valor selecionado após experimentação)

As acurácias encontradas foram as abaixo:

- *Breast Cancer Dataset*: 0.8303 ± 0.0314
- *Statlog(heart) Dataset*: 0.8289 ± 0.0375

Os boxplots com as distribuições das acurácias podem ser vistos nas figuras 1 e 2.

Após isto, o modelo RBF com *k-means* foi implementado. O modelo RBF é similar ao ELM no sentido de ter camadas intermediárias na rede buscando linearizar a fronteira entre as classes. A diferença é que o RBF realiza este aumento de dimensão através de funções de bases radiais.

Os dados foram divididos em treinamento e teste (numa proporção de 30%), normalizados (para ficarem entre 0 e 1) e executados 50 vezes. O número de *clusters* para as redes dos dois modelos foi definido como 2, após experimentação.

As acurácias encontradas foram as abaixo:

- *Breast Cancer Dataset*: 0.7888 ± 0.03827
- *Statlog(heart) Dataset*: 0.7869 ± 0.03941

Os boxplots com as distribuições das acurácias podem ser vistos nas figuras 3 e 4.

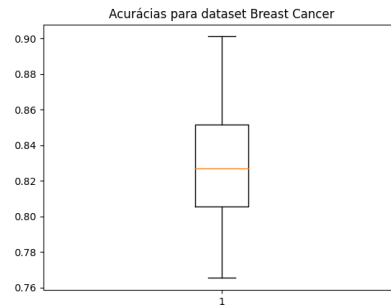


Figure 1: Boxplot das acurácias para o Breast Cancer Dataset utilizando ELM

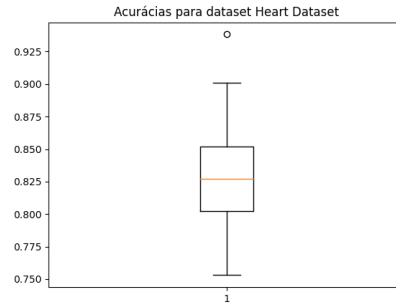


Figure 2: Boxplot das acurácias para o Statlog(heart) Dataset utilizando ELM

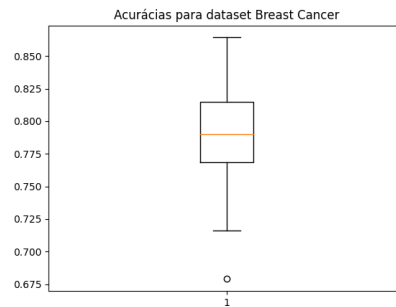


Figure 3: Boxplot das acurácias para o Breast Cancer Dataset utilizando RBF

2 Conclusões

Dados os experimentos realizados com as duas implementações, podemos notar que o ELM teve um desempenho superior ao RBF. Porém, notou-se du-

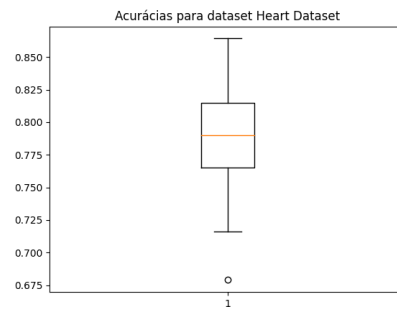


Figure 4: Boxplot das acurácias para o Statlog(heart) Dataset utilizando RBF

rante a implementação que, ao aumentar o numero de *clusters* utilizados, a acurácia diminuí, o que pode indicar um erro na implementação (que foi feita manualmente, em python, com base no material disponibilizado).