

Optical Character Recognition

Jorge Melo & Xavier Pinho

1. Introdução

O objetivo deste trabalho é construir vários classificadores com capacidade de prever algarismos escritos por um dado utilizador através de uma aplicação de fácil utilização. Com esse intuito, desenvolveram-se seis classificadores, usando três diferentes funções de ativação, uma com memória associativa e outra sem.

Cada classificador foi avaliado através da sua accuracy e da curva ROC.

O utilizador é convidado a escrever um dado conjunto de algarismos que serão avaliados pelo modelo que este desejar.

2. Métodos

- **Dataset**

Os datasets de treino e de teste foram desenhados por ambos elementos do grupo e um membro exterior ao grupo através da função `mpaper`, tendo sido obtida uma matriz 256×500 e outra 256×50 , respetivamente. Os vetores target correspondents apresentam dimensões 10×500 e 10×50 também respetivamente.

A matriz treino foi dividida em treino e validação na proporção 85 para 15, respetivamente, com o objetivo de evitar um possível overfitting.

- **Arquitetura da Rede Neuronal**

Para construir os diferentes classificadores usaram-se três funções de ativação: Hardlim, Purelin e Logsig. Cada uma destas deu origem a duas Redes Neurais, uma com memória associativa (AM) e outra sem memória associativa (WAM), através da poderosa ferramenta matemática, MatLab.

Para cada classificador, o learning rate foi de 0.5, o número de épocas de 1000 e o goal de $1e-6$.

3. Resultados

Para avaliar a accuracy dos modelos foi usada a função `accuracyTeste.m`, que compara o output da simulação de cada uma das redes neuronais com os dados de test. Para os classificadores com memória associativa foi criada uma matriz de números perfeitos (256×500) com a ajuda do ficheiro `Perfect_numbers.mat`.

	Purelin	Hardlim	Logsig
AM + Classificador	10%	8%	12%
Classificador	94%	92%	98%

Tabela 1 – Resultados para a accuracy dos vários modelos.

Observando accuracy de cada classificador nos dados de teste, conclui-se que o melhor classificador é o Logsig sem memória associativa.

Para representar graficamente a especificidade e a sensibilidade do modelo, decidimos fazer o plot da curva ROC para os dados de test. Apenas se fez o plot dos dois classificadores com melhor accuracy. Estes plots podem ser vistos nas figuras 1 e 2.

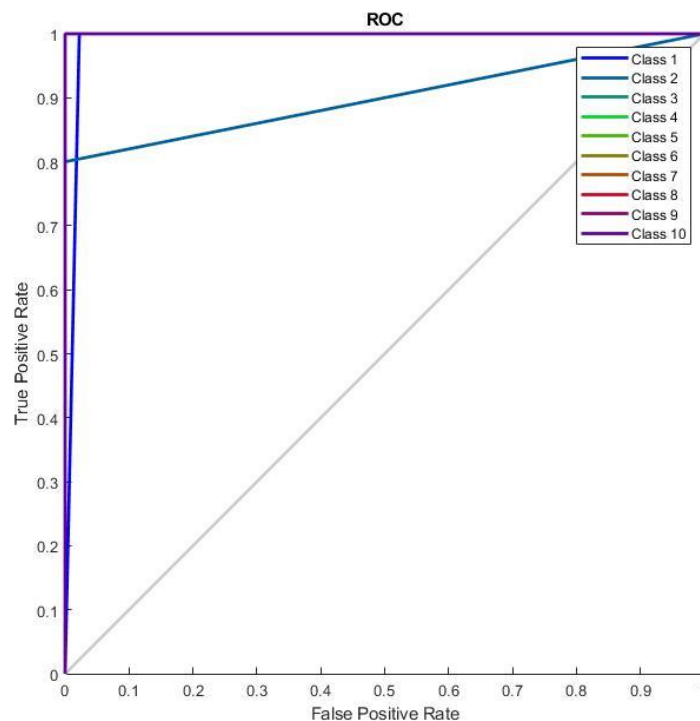


Figura 1 – Curva ROC para o classificador Logsig sem memória associativa.

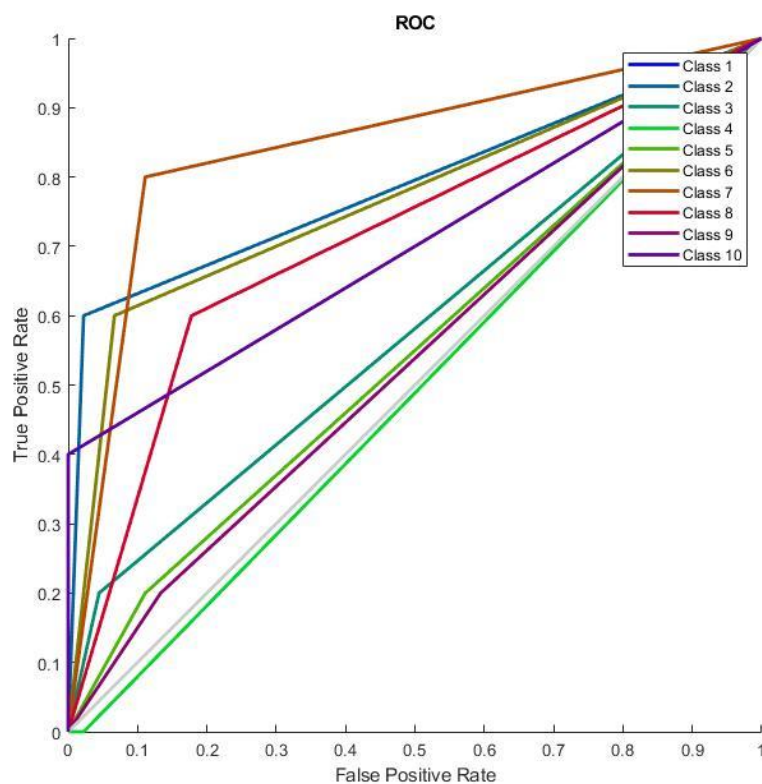


Figura 2 – Curva ROC para o classificador Logsig com memória associativa.

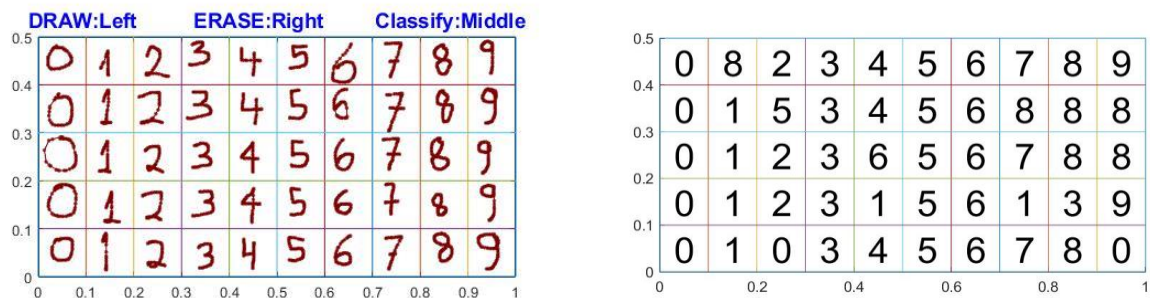


Figura 3 – Representação do input de um utilizador e os resultados da classificação.

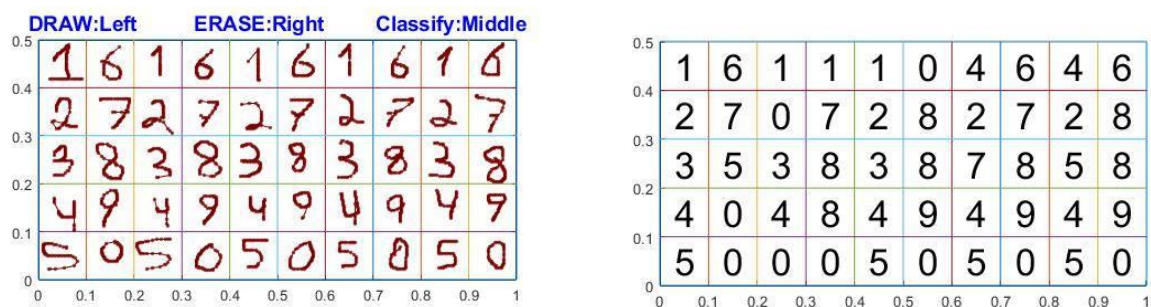


Figura 4 – Representação do input de um utilizador externo ao grupo e os resultados da classificação.

4. Conclusão

O objetivo de construir uma rede neuronal capaz de distinguir algarismos foi atingido.

Como já referido nos resultados, o melhor classificador é o Logsig sem Memória associativa, e o melhor classificador com Memória Associativa é o Logsig; desta forma conclui-se que a função transferência que apresenta melhores resultados é precisamente o Logsig (sigmoide).

Para confirmar os resultados obtidos através dos dados de teste, desenvolvemos uma aplicação que permite ao utilizador desenhar algarismos e obter a previsão do modelo escolhido. Os resultados para o membro do grupo (11 errados em 50) foram superiores aos do membro exterior ao grupo (16 errados em 50), isto pode ser explicado pelo facto do dataset de treino com que as redes neuronais foram treinadas foi construído pelos membros do grupo. Isto mostra que os modelos não estão idealmente generalizados, uma vez que o dataset de treino tem pouca variabilidade para além do facto de o considerarmos muito escasso. Ainda assim podemos considerar os resultados bastante razoáveis.

Também se observou que os classificadores com Memória Associativa apresentam resultados muito inferiores aos classificadores sem Memória Associativa, esta diferença poderá explicada pela limitação dos dados utilizados na Memória Associativa e pela sua unidirecionalidade.

