

# Redes Neurais para Reconhecimento de Linguagem Gestual

Diogo Pinto, Luís Brochado, Wilson Oliveira

24 de Abril de 2014

## 1 Introdução

No Séc. XIX dois psicólogos (*Alexander Bain e William James*) propuseram que o cérebro e a memória estavam organizados em redes de neurónios interligados. William James concluiu ainda que as ligações entre neurónios são ligações elétricas e que a memória é apenas um conjunto de *circuitos* aprendidos.

Um século depois, a ideia do duo de psicólogos foi utilizada para tentar criar máquinas que conseguissem aprender e "pensar" sem intervenção humana. Assim nasceu o conceito de **Rede Neuronal Artificial**, que é nada mais nada menos do que um programa criado pelo Homem para tentar simular o comportamento do sistema nervoso de forma a levar a máquina a raciocinar e tirar conclusões autonomamente.

Após anos de trabalho chegou-se a um consenso relativamente à caracterização de uma rede neuronal. Este fica completamente caracterizada quando são conhecidas:

- A unidade básica (conhecida como neurónio artificial);
- A topologia da rede (estrutura das ligações desta);
- A regra de aprendizagem.

Mencionando o neurónio artificial, estamos preparados para falar das camadas das redes neuronais e dos diferentes tipos de ligações entre neurónios. Para abordar a topologia de uma rede é necessário compreender o conceito de camada. Uma camada é um conjunto de neurónios que desempenham a mesma função.

### 1.1 Topologia da rede

Existem três tipos de topologia para uma rede neuronal. São eles

- **Redes totalmente conectadas**, em que cada neurónio da rede está ligado a todos os outros;
- **Redes de camada única**, que contém apenas camada de entrada e camada de saída;

- **Redes de múltiplas camadas**, que além das camadas de entrada e saída, também possuem uma ou mais camadas intermédias (também denominadas de camadas escondidas).

Uma rede neuronal é sempre constituída por duas camadas no mínimo, uma camada de entrada e uma camada de saída. Tal como os nomes indicam, a camada de entrada tem como objetivo receber informação e passá-la à camada de saída ou à(s) camada(s) intermédia(s). A camada de saída tem como função fornecer o resultado final. Já a camada intermédia, se existente, é responsável pela correção do resultado final em função do input da camada de entrada.

Assim, podemos considerar as diferentes ligações entre neurónios:

- **Ligação direta**, são ligações um-a-um entre um neurónio de uma camada e da camada que se lhe procede, ou seja, neste caso o número de neurónios nas várias camadas tem de ser igual.
- **Ligação inter-direta**, as ligações entre camadas são múltiplas, ou seja, não existe relação um-a-um e o mesmo neurónio numa camada pode estar ligado a mais de um neurónio da camada que lhe precede.
- **Ligação intra-direta**, são ligações completas ou aleatórias entre nós da mesma camada.

Falta então mencionar os diferentes métodos de aprendizagem.

## 1.2 Tipos de aprendizagem

- **Aprendizagem por Reforço;**
- **Aprendizagem supervisionada;**
- **Aprendizagem não supervisionada.**

O método de aprendizagem escolhido para este projeto foi o de aprendizagem supervisionada. Com este método de aprendizagem a rede produz a uma resposta ao input, após o que um "supervisor" apresenta como resposta correta. Caso as respostas sejam diferentes, a diferença entre as duas respostas é utilizada para modificar o peso das ligações existentes.

## 2 Objetivo

Este projeto visa a criação de um programa que seja capaz de aprender e reconhecer linguagem gestual através da implementação de redes neuronais.

## 3 Descrição

### 3.0.1 Pormenorização do projeto

Este projeto foca-se na criação de um programa que consiga treinar apropriadamente uma Rede Neuronal Artificial, a reconhecer linguagem gestual, usando o algoritmo "Back-Propagation".

O conjunto de dados deve ser cuidadosamente analisado de forma a verificar a eventual necessidade de pré-processamento. O modelo obtido deve depois ser utilizado na classificação de novos casos.

### 3.0.2 Concepção de uma rede neuronal multi-camada

A camada de entrada contém os atributos de identificação dos dados, 11 parâmetros para o reconhecimento de cada mão, perfazendo um total de 22 parâmetros sendo eles:

- Posição X expressa em metros, em relação a um ponto ligeiramente abaixo do queixo;
- Posição Y expressa em metros, em relação a um ponto ligeiramente abaixo do queixo;
- Posição Z expressa em metros, em relação a um ponto ligeiramente abaixo do queixo;
- Rotação no eixo do X, medida num valor entre -0.5 e 0.5, sendo 0 a posição da palma da mão plana horizontalmente. Se o valor for positivo significa que a palma da mão está virada para cima na perspectiva do signatário. Para obter a medida em graus, multiplicar por 180;
- Rotação no eixo dos Y, medida num valor entre -1.0 e 1.0, sendo 0 a posição da palma para a frente na perspectiva do signatário;
- Curvatura do dedo polegar medida entre 0 e 1, sendo que 0 significa o dedo esticado e 1 o dedo totalmente dobrado;
- Curvatura do dedo indicador medida entre 0 e 1, sendo que 0 significa o dedo esticado e 1 o dedo totalmente dobrado;
- Curvatura do dedo médio medida entre 0 e 1, sendo que 0 significa o dedo esticado e 1 o dedo totalmente dobrado;
- Curvatura do dedo anelar medida entre 0 e 1, sendo que 0 significa o dedo esticado e 1 o dedo totalmente dobrado;
- Curvatura do dedo mindinho medida entre 0 e 1, sendo que 0 significa o dedo esticado e 1 o dedo totalmente dobrado.

**No entanto, as medidas de dobra dos dedos não são totalmente exatas.**

Já a camada de saída contém a classificação obtida e a(s) camada(s) intermédia(s) auxilia(m) no funcionamento da rede neuronal.

### 3.0.3 Implementação/aplicação do algoritmo "Back-Propagation"

### 3.0.4 Medição detalhada de resultados nos dados de treino e de teste

Com a arquitetura atual, para 8 palavras a rede demora cerca de 160 iterações pela base de dados para ser propriamente ensinada.

## 3.1 Organização do projeto

O trabalho foi dividido em três partes. A primeira consiste na criação da rede neuronal e aplicação do algoritmo de retropropagação ("Back-Propagation") e apresentar os resultados bem como as estatísticas. A segunda parte baseia-se na otimização da primeira parte, de modo a obter respostas mais rapidamente.

A terceira parte trata-se da implementação de funcionalidades adicionais, como por exemplo, uma GUI ou o reconhecimento de gestos através de imagens e/ou vídeo. Esta última parte será previamente discutida com os docentes da disciplina.

O grupo está a considerar construir uma rede neuronal modular caso esta se demonstre uma opção útil na otimização do programa. Uma rede neuronal modular atribui determinada tarefa que inicialmente seria desempenhada por um ou vários neurónios, a uma outra "sub"rede neuronal contida na rede principal.

## 4 Trabalho efetuado

A primeira parte do projeto, enunciada na alínea anterior, já se encontra completa, sendo que o grupo está a iniciar a segunda parte sendo para tal necessária a ajuda de um docente devido a algumas dúvidas que surgiram.

## 5 Algoritmo

### 5.1 Alimentação para a frente ("Feedforward")

Uma rede neuronal feedforward é uma rede em que os neurónios (unidade básica de uma rede neuronal) apenas estão ligados a neurónios das camadas seguintes.

Deste modo, aproveitamos o facto da nossa rede neuronal ser uma rede "feedforward" e criámos um método com o mesmo nome, cujo objetivo é propagar a **informação** pela rede neuronal. Assim, dividimos o problema em duas partes. A primeira parte consiste em propagar os valores por toda a rede (do início para o fim), parte essa que está implementada no método "feedforward".

A segunda parte consiste na aplicação do algoritmo de retropropagação, a ser explicada mais à frente.

**Devemos explicar os diferentes tipos de redes que há expôr as fórmulas que usamos**

### 5.2 Retropropagação

Ao contrário do feedforward, a retropropagação não é um tipo de rede neuronal mas um algoritmo que como o nome indica, faz a retropropagação da **informação**, sendo que o sentido em que a rede propaga a dita **informação** é da camada de saída para as anteriores.

Isto é, após o cálculo dos erros na camada de saída, os erros devem ser propagados para as camadas intermédias e à medida que o erro vai sendo propagado para trás, o peso das ligações entre camadas é alterado de modo a diminuir o erro.

**Falta apresentar as fórmulas visto que já temos uma breve descrição de como funciona**

### 5.3 Atualização dos pesos das ligações

Atualizar o peso das ligações (é equivalente a minimizar o erro) é o modo de se reduzir para zero, ou aproximadamente zero, o erro na camada de saída, tornando a rede neuronal fidedigna.

Este processo é feito através da função de custo quadrático e da sua derivada, sendo que a primeira é a função a minimizar no algoritmo de retropropagação (apresentando  $\mathbf{x}$  exemplos a um rede com  $\mathbf{y}$  saídas).

**Falta explicar como funciona e apresentar as fórmulas**

## 6 Conclusões

A escrever

## 7 Recursos