



Universidade do Porto

Faculdade de Engenharia

FEUP

Redes Neurais para Reconhecimento de Linguagem Gestual

Relatório Intercalar

Inteligência Artificial

3º ano do Mestrado Integrado em Engenharia
Informática e Computação

Elementos do Grupo:

Diogo Joaquim Araújo Pinto – 201108016 - ei11120@fe.up.pt

Luís Brochado Pinto dos Reis – 201003074 - ei11009@fe.up.pt

Wilson da Silva Oliveira - 201109281 - ei11085@fe.up.pt

22 de abril de 2014

Conteúdo

1	Considerações gerais	3
1.1	Topologia da rede	3
1.2	Tipos de aprendizagem	4
2	Objetivo	4
3	Descrição	4
3.1	Pormenorização do projeto	4
3.2	Concepção de uma rede neuronal multi-camada	4
3.3	Implementação/aplicação do algoritmo "Back-Propagation	5
3.4	Medição detalhada de resultados nos dados de treino e de teste .	5
3.5	Organização do projeto	6
4	Trabalho efetuado	6
5	Resultados esperados e forma de avaliação	6
6	Conclusões	6
7	Recursos	7

1 Considerações gerais

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.**

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

3.1 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".

Para a realização desta rede neuronal, tivemos acesso a uma base de dados que contém várias amostras de palavras da língua gestual Australiana. Para a utilização desta base de dados, foi necessário o pré-processamento das medições contidas na mesma, para poder ser utilizada pela rede neuronal e servir como apoio na identificação de novos casos.

3.2 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.

A camada de saída contém um número de nós igual ao número de gestos reconhecidos pela rede neuronal, estes nós têm como output uma probabilidade indicativa da confiança que a rede neuronal tem na sua decisão.

A camada intermédia é utilizada para auxiliar os cálculos efetuados pela rede neuronal, sendo que ainda precisamos de ajustar e procurar a estrutura interna mais benéfica para a nossa rede neuronal.

3.3 Implementação/aplicação do algoritmo "Back-Propagation"

O algoritmo de "back-propagation" propaga os parâmetros recebidos pela camada de entrada, aplicando-lhes uma função de transferência à medida que transmite os dados às camadas seguintes. Após chegar à camada de saída os dados são propagados para trás até à camada de entrada.

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. Este passo também é conhecido como "Treino" da rede.

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).

3.4 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.

3.5 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 posteriormente 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 fase de otimização da rede neuronal, procurando obter um melhor desempenho, e resolver alguns problemas com a mesma.

5 Resultados esperados e forma de avaliação

À medida que se forem realizando mais testes, o esperado é que o programa a cada chamada necessite de consultar menos pastas da base de dados do que na chamada anterior, pois assim verifica-se que a programa aprendeu o gesto com sucesso e é capaz de o identificar.

Para a utilização da rede neuronal, neste momento é necessário carregar um ficheiro com os valores inseridos manualmente para a rede neuronal poder identifica-lo. Esperamos que até à entrega final nos seja possível implementar uma maneira mais amigável para o utilizador fazer uso desta rede neuronal.

6 Conclusões

Com este projeto foi possível compreender o conceito de redes neurais artificiais, bem como as potencializações da sua utilidade na sociedade. Uma vez que o trabalho ainda não está completo o grupo tem grandes ambições em poder contribuir para a sociedade com este projeto, ajudando a melhorar a percepção de linguagem gestual bem como a identificação de gestos de um modo rápido e eficaz.

7 Recursos

O software utilizado foi o IDE IntelliJ IDEA 13.1.1, para programação em java.

Referências

- [1] Eugénio Oliveira *Slides de IART: http://paginas.fe.up.pt/~eol/IA/1314/APONTAMENTOS/7_RN.pdf* 2013-2014.
- [2] Wikipedia, Redes Neurais Modulares *http://en.wikipedia.org/wiki/Modular_neural_network*.
- [3] Anabela Simões , Ernesto Costa *Inteligência Artificial* 2004: FCA - Editora Informática.
- [4] Peter Norvig, Stuart Russell *Artificial Intelligence A Modern Approach* 2009: Prentice Hall.