



Universidade do Porto
Faculdade de Engenharia

FEUP

Redes Neurais para a Predição de Espécies de Anuros

Relatório Intercalar

Inteligência Artificial

3º ano do Mestrado Integrado em Engenharia Informática e Computação

Elementos do Grupo:

João Almeida – up201504874 – j.almeida@fe.up.pt

João Mendes – up201505439 – up201505439@fe.up.pt

Ricardo Santos – up201503716 – up201503716@fe.up.pt

8 de abril de 2018

Conteúdo

1	Objetivo	3
2	Descrição	3
2.1	Especificação.....	3
2.1.1	Descrição e análise do dataset	3
2.1.2	Pré-processamento dos dados.....	4
2.1.3	Modelos de aprendizagem a aplicar	4
2.1.4	Arquitetura da Rede Neuronal	5
2.1.5	Configuração Prevista da Rede Neuronal.....	5
2.2	Trabalho Efetuado	5
2.3	Resultados esperados e forma de avaliação	6
3	Conclusões	6
4	Recursos	7

1 Objetivo

A saúde ambiental é um problema cada vez mais presente no planeta. De modo a identificar problemas ecológicos em fases precoces, a comunidade científica tem rastreado as variações populacionais de anuros em programas de monitorização bioacústica. Estes programas utilizam redes de pequenos sensores que coletam, processam e transmitem a informação áudio relacionada com diversas variáveis ambientais.

O objetivo deste trabalho é a implementação de uma rede neuronal artificial multi-camada (com utilização do algoritmo *Back-Propagation*) capaz de prever espécies de anuros a partir de qualquer *dataset* obtido nestes programas de monitorização.

Neste relatório intercalar encontra-se especificado a descrição e análise do *dataset* (secção 2.1.1), o pré-processamento dos dados (secção 2.1.2), os modelos de aprendizagem a aplicar (secção 2.1.3), a arquitetura da rede neuronal (secção 2.1.4), a configuração prevista para a rede (secção 2.1.5), bem como o trabalho efetuado até à data (secção 2.2) e os resultados esperados e forma de avaliação (secção 2.3). No final, encontram-se as conclusões (secção 3) e os recursos utilizados para a realização do trabalho (secção 4).

2 Descrição

2.1 Especificação

2.1.1 Descrição e análise do *dataset*

Os programas de monitorização bioacústica deste *dataset* foram realizados no *campus* da Universidade Federal do Amazonas e na Mata Atlântica, ambos no Brasil, e em Córdoba, na Argentina. No total, foram recolhidos 60 registos áudio correspondentes a 60 espécimes (sapos individuais) que pertencem a 1 ordem (Anuro), 4 famílias, 8 géneros e 10 espécies diferentes, de acordo com a taxonomia de Lineu. Cada registo áudio equivale a um sinal bioacústico $X = \{x_1, x_2, \dots, x_N\}$, uma sequência em ordem ao tempo de tamanho N , na qual cada valor representa o nível de pressão acústica (ou amplitude). Este sinal correspondente ao chamamento de cada espécime de anuro.

Na **fase de segmentação**, cada sinal é dividido em **sílabas** $x_k = \{x_t, x_{t+1}, \dots, x_{t+n}\}$ – a unidade elementar utilizada em classificação com base em dados bioacústicos. O pré-processamento destes programas envolveu determinar o início e o fim das sílabas. Cada linha do ficheiro do *dataset* corresponde a uma sílaba, totalizando **7195 sílabas**.

Na **fase de extração de características**, cada sílaba é representada por um conjunto de características, denominados **Descritores de Baixo Nível** (*Low Level Descriptors*, **LLDs**, em inglês). Os LLDs utilizados nesta pesquisa são os **Coefficientes**

Cepstrais da Frequência-Mel (*Mel-Frequency Cepstral Coefficients*, **MFCCs**, em inglês). É feita uma análise espectral baseado num banco de 44 filtros triangulares logaritmicamente espaçados no domínio da frequência. Cada sílaba é representada por um conjunto de coeficientes (normalizados entre -1 e 1, dado as sílabas terem durações diferentes) $MFCC(x_k) \rightarrow c_k$, isto é, $X \rightarrow \{(c_1, s_1), (c_2, s_1), \dots, (c_k, s_l)\}$, em que cada $c_k = [c_1, c_2, \dots, c_l]$ é um vetor de características com l coeficientes, e s_l é o nome da espécie. A **utilização dos coeficientes** na análise dos dados é mais **robusta**, mais **compacta** e mais **fácil de reconhecer**, comparativamente a usar diretamente o sinal de entrada, constituindo assim os **22 atributos de identificação dos dados do nosso modelo**.

2.1.2 Pré-processamento dos dados

O *dataset* será dividido em dois *sets*, os quais designaremos por **training set**, com cerca de 75% dos elementos, e **test set**, com os restantes 25%. Como indicam os nomes, o primeiro *set* será utilizado para o treino da rede neuronal, e o segundo será utilizado para testá-la após esse treino.

A divisão será feita de modo relativamente arbitrário, no entanto teremos alguns cuidados a tomar: o **primeiro será o de garantir que há pelo menos um espécime de cada espécie nos dois sets**; o segundo será o de garantir que, **sendo selecionada uma sílaba de um dado espécime para fazer parte de um desses sets, todas as sílabas desse espécime deverão ser adicionadas ao mesmo set**.

Adotamos estas regras por forma a **melhorar as capacidades de generalização** da rede neuronal, de acordo com estudos previamente realizados por outros autores que determinaram que estas são afetadas negativamente caso haja uma mistura das sílabas de um dado espécime entre os *sets* de treino e teste (mais concretamente, é **criado um bias na precisão de predição da rede**), que se **deve ao facto de todas as sílabas de um dado espécime terem fortes semelhanças entre si**. Este tópico é explorado mais aprofundadamente no artigo “*How to Correctly Evaluate an Automatic Bioacoustics Classification Method*” de Juan G. Colonna, João Gama, e Eduardo F. Nakamura^[1], ao qual também **recorreremos para o método de avaliação dos resultados dos testes que efetuaremos**.

2.1.3 Modelos de aprendizagem a aplicar

O desafio proposto é o de atribuir o nome de uma espécie a uma nova sílaba utilizando os coeficientes **MFCC** obtidos dessa sílaba. Para isto, implementaremos uma **rede neuronal** treinada sob **um método de classificação supervisionada**, através de um algoritmo de *backpropagation* (ou *backwards propagation of error*), procurando **balancear**

os pesos nas camadas internas, na tentativa de criar um modelo com a capacidade de prever a classificação de novas amostras dentro das espécies consideradas. A **avaliação** do *output* será realizada através de uma certa ***cost function***, que será minimizada através do algoritmo ***gradient descent***.

2.1.4 Arquitetura da Rede Neuronal

Em particular, a rede neuronal a ser utilizada terá uma **arquitetura *feedforward* multi-camada**. Essas camadas serão ***fully connected***, e serão testadas várias configurações de camadas interiores. Terá como camada de entrada os valores **MFCC** de cada sílaba, logo, **22 atributos de entrada**, e como camada de saída **10 valores entre 0 e 1**, que serão arredondados para estes valores e interpretados como pertencendo a uma dada espécie (daí, **um *output* correto terá sempre a forma de um vetor unitário**). Procuramos com este *design* que a rede tenha maior facilidade em conseguir estabelecer relações entre espécies e os **MFCCs** correspondentes, através da conexão de *inputs* de diferentes espécies a *outputs* inequivocamente diferentes.

2.1.5 Configuração Prevista da Rede Neuronal

Testaremos configurações **com 3 a 7 camadas interiores**, com **número de células variando entre 3 e 25**. Uma tentativa de configuração em particular consistirá em fazer as últimas 3 camadas, camada de *output* inclusive, ter um número de células correspondente ao número de famílias, género, e espécies do *dataset* considerado, nessa ordem, na expectativa de que as potenciais semelhanças entre **MFCCs** de espécies diferentes nos mesmos super-grupos taxonómicos sejam “agrupadas” nessas camadas interiores. Treinaremos também a rede sob vários ***learning rates***, não podendo fazer de momento nenhuma previsão sobre qual o valor mais adequado.

2.2 Trabalho Efetuado

De momento, temos uma implementação incompleta da rede neuronal a ser utilizada, criada em **C++**, **sem recorrer a bibliotecas externas**. Resta terminar a implementação, realizar os **testes**, e fazer potenciais ajustes tanto à implementação como à rede obtida após cada sessão de treino-teste.

Pretendemos também comparar os resultados da nossa implementação com os resultados obtidos de treinos e testes com este *dataset* realizados em redes neuronais provenientes de bibliotecas já existentes, como a ***TensorFlow***, caso a implementação desta não constitua um obstáculo na conclusão da nossa implementação.

Futuramente, caso haja oportunidade, contemplaremos a criação de uma arquitetura diferente sobre a implementação atual (nomeadamente, uma **rede convolucional**), e subsequente **treino** e **teste** dessa rede, comparando os resultados obtidos com os da rede atualmente considerada.

2.3 Resultados esperados e forma de avaliação

Como mencionado anteriormente, utilizaremos uma variante da metodologia ***k*-Cross-Validation** (**k-CV**) para avaliação dos resultados obtidos conforme descrita no artigo “*How to Correctly Evaluate an Automatic Bioacoustics Classification Method*”, que se divide em duas estratégias: **One-against-All (1AA)** e **One-against-One (1A1)**.

Resumidamente, a **primeira** estratégia consiste em separar todas as sílabas de um espécime das restantes, sendo que as primeiras ficam no *set* de teste e as últimas no *set* de treino. A **rede é treinada e testada contra as sílabas desse espécime**, sendo que após isto repetimos o processo **iterando pelos espécimes** até que todos tenham feito parte do *set* de teste uma vez (este método é portanto semelhante ao método **Leave-One-Out-CV**, mas **utilizando o conjunto de sílabas de um espécime em vez de sílabas singulares**).

Já a **segunda** estratégia é **semelhante à primeira**, com a **diferença de que o *set* de treino é adicionalmente subdividido por espécie**. O método prossegue com o treino da rede com um desses *subsets* de treino, e **testada com o *set* de teste** (ainda com as sílabas de apenas um espécime), seguido pelo **treino com outro *subset* de treino** (diferente do primeiro), e **testada novamente com o *set* de teste** (igual), até que **todos os *subsets* de treino tenham sido utilizados**. Após isto, o processo todo é **repetido** iterando pelos espécimes da mesma forma que para a primeira estratégia.

Com estas estratégias, procuramos obter **resultados que demonstrem a capacidade de generalização** da rede neuronal sem que esta sofra de um ***bias* de precisão**. Compararemos os nossos resultados com os resultados detalhados no mesmo artigo, procurando obter valores semelhantes aos indicados nos testes realizados noutras redes neuronais (**SVMs**).

3 Conclusões

Em suma, temos a expectativa de que os resultados que vamos obter com o algoritmo de aprendizagem automática desenvolvido por nós sejam fiáveis, pois não só temos as ferramentas necessárias, como temos os resultados de trabalhos precedentes como termo de comparação.

4 Recursos

1) Dataset

COLONNA, J. G.; CRISTO, M.; NAKAMURA, E. F.; GORDO, M. (2018).

UCI Machine Learning Repository

[<http://archive.ics.uci.edu/ml/datasets/Anuran+Calls+%28MFCCs%29>].

Manaus, Brasil: Universidade Federal do Amazonas

2) COLONNA, J. G.; CRISTO, M.; SALVATIERRA, M.; NAKAMURA, E. F.

An Incremental Technique for Real-Time Bioacoustic Signal Segmentation.

Expert Systems with Applications, v. 42, p. 7367-7374, 2015.

3) COLONNA, J. G.; GAMA, J.; NAKAMURA, E. F.

How to Correctly Evaluate an Automatic Bioacoustics Classification Method.

In: 17th Conference of the Spanish Association for Artificial Intelligence (CAEPIA).

Lecture Notes in Computer Science. 986ed.: Springer International Publishing, 2016, v. , p. 37-47.

4) COLONNA, J. G.; PEET, T.; FERREIRA, C. A.; JORGE, A. M.; GOMES, E. F.; GAMA, J. (2016, July).

Automatic Classification of Anuran Sounds Using Convolutional Neural Networks.

In Proceedings of the Ninth International C* Conference on Computer Science & Software Engineering (No. C3S2E '16, pp. 73-78). ACM.

5) COLONNA, J. G.; CRISTO, M.; NAKAMURA, E. F. (2014, August).

A Distributed Approach for Classifying Anuran Species Based on Their Calls.

In Pattern Recognition (ICPR), 2014 22nd International Conference on (pp. 1242-1247). IEEE.

6) NIELSEN, M.

Neural Networks and Deep Learning, Chapter 1 (Last accessed on 8th April 2018).

<http://neuralnetworksanddeeplearning.com/chap1.html>