

**Redes Neuronais 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

11 de abril de 2018

Conteúdo

[1 Objetivo 3](#_Toc510015496)

[2 Descrição 3](#_Toc510015497)

[2.1 Especificação 3](#_Toc510015498)

[2.1.1 Descrição e análise do dataset 3](#_Toc510015499)

[2.1.2 Pré-processamento dos dados 4](#_Toc510015500)

[2.1.3 Modelos de aprendizagem a aplicar 4](#_Toc510015501)

[2.1.4 Arquitetura da Rede Neuronal 5](#_Toc510015502)

[2.1.5 Configuração Prevista da Rede Neuronal 5](#_Toc510015503)

[2.2 Trabalho Efetuado 6](#_Toc510015504)

[2.3 Resultados esperados e forma de avaliação 6](#_Toc510015505)

[3 Conclusões 7](#_Toc510015506)

[4 Recursos 8](#_Toc510015507)

# 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 áudiorelacionada 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).

# Descrição

* 1. Especificação
     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 Córdova, 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 , uma sequência em ordem ao tempo de tamanho , 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**  – 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 **Coeficientes Cepstrais da Frequência-Mel** (*Mel-Frequency Spectral Coefficients*, ***MFCCs***, em inglês). É feita uma análise espectral baseado num banco de 44 filtros triangulares logaritimicamente 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) , isto é, , em que cada é um vetor de caraterísticas com coeficientes, e 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**.

* + 1. **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. Futuramente tencionamos que estas percentagens sejam variáveis para proceder melhor analisar o comportamento da rede em diferentes circunstâncias.

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****,* **todos as sílabas desse espécime deverão ser adicionadas ao mesmo *set***.

Adotamos estas regras de modo a **melhorar** as **capacidades de generalização** da rede neuronal e ainda 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 chamamento 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 chamamento 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.

* + 1. **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 **balançar 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 de uma função ***gradient descent****,* calculada através do algoritmo de *backpropagation.*

* + 1. **Arquitetura da Rede Neuronal**

Em particular, a rede neuronal a ser utilizada terá uma **arquitetura *feedforward* multi-camada,** onde a informação se move apenas numa direção*.* 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 que serão 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.

* + 1. **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 sejam agrupadas nessas camadas interiores. Como **hiper-parâmetros** testaremos vários ***learning rates***, não podendo fazer de momento nenhuma previsão sobre qual o valor mais adequado.

* 1. Trabalho Efetuado

De momento, temos uma implementação incompleta da rede neuronal a ser utilizada, realizada em **C++, sem recorrer a bibliotecas exteriores**. Resta terminar a implementação e realizar os **testes**, e fazer potenciais ajustes tanto à implementação como à rede obtida após cada sessão de treino-teste. 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. Esta rede seria implementada recorrendo à biblioteca *tensorflow* (<https://www.tensorflow.org/>) - utilizando a linguagem *Python* – já reputada pela sua eficiência, e comparamos os resultados obtidos com os da rede por nós implementada.

* 1. Resultados esperados e forma de avaliação

A fase final do projeto será dedicada à implementação de uma medição dos resultados obtidos, que poderá ou não validar os nossos resultados. Para além disso vamos tentar avaliar a ***accuracy*** da rede neuronal e realizar algum tipo de estatística relativo ao comportamento da rede ao longo do tempo.

Iremos submeter o nosso *dataset* à rede neuronal, com diferentes configurações de camadas/células (neurónios), e será de esperar que independentemente destas ou do tamanho dos training/test sets a rede devolverá o correto espécime de anuro, avaliaremos o comportamento do mesmo input para tirarmos conclusões quanto à implementação.

A comparação com uma rede neuronal construída em *tensorflow* irá permitir-nos avaliar o desempenho da nossa implementação, e, ao longo da fase de desenvolvimento, levar ao seu melhoramento constante. Esta comparação poderá ajudar-nos a confirmar os resultados obtidos. Já no campo dos resultados esperados, estimamos construir uma rede completamente funcional e o mais generalista possível para futuro uso.

# Conclusões

Podemos concluir até agora que redes neuronais são um tema bastante interessante e que fizemos uma boa aposta na escolha do tema. A pesquisa e aprendizagem tem sido deveras consumidora de tempo mas o grupo sente que não tem sido em vão, pois este é um assunto que é cada vez mais falado hoje em dia e que tem aplicações infinitas no mundo real, como é exemplo a predição de espécies de anuros.

Quanto ao nosso projeto em específico, a implementação da rede até agora tem sido um processo de tentativa e erro, estes últimos têm vindo a ser cada vez menos recorrentes com o afunilamento da nossa aprendizagem. Ao nível de programação, temos usado uma linguagem familiar ao grupo – *C++* - que tem sido uma ferramenta útil e não nos tem faltado nada até ao momento.

Em termos teóricos tem sido um processo algo exigente devido a termos de passar por um processo de auto aprendizagem sem o apoio das aulas teóricas e com todo o trabalho a que estamos a ser sujeitos, contudo, o grupo permanece otimista na elaboração deste aliciante projeto.

# 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.