



FEUP **FACULDADE DE ENGENHARIA**
UNIVERSIDADE DO PORTO

Inteligência Artificial

3º ano do Mestrado Integrado em Engenharia Informática e Computação (MIEIC)

E3: Redes Neurais para a predição de espécies de anuros

Relatório Intercalar

Grupo E3_2

Elementos do Grupo:

António Cunha Seco Fernandes de Almeida - 201505836- up201505836@fe.up.pt

Diogo Luis Rey Torres - 201506428 - up201506428@fe.up.pt

João Paulo Madureira Damas - 201504088 - up201504088@fe.up.pt

8 de abril de 2018

1.Objetivo

Este trabalho tem como objetivo a aplicação de Redes Neurais para a predição de espécies de anuros, com base nos seus chamamentos.

Para tal será utilizado um data set com atributos acústicos extraídos dos chamamentos efetuados pelos anuros, em função da sua família, género e espécie.

Os anuros estão diretamente relacionados com o ecossistema onde vivem, sendo a sua atividade avaliada por biólogos como indicador de stress ambiental.

As redes neuronais artificiais são modelos simplificados do sistema nervoso central do ser humano. Trata-se de uma estrutura extremamente interconectada de unidades computacionais, frequentemente designadas por neurónios ou nós, com capacidade de aprendizagem. Durante o processo de aprendizagem, dado por um algoritmo de aprendizagem ou de treino, o peso das conexões é ajustada de forma a se atingir um desejado objetivo.

No final do trabalho é esperado que a rede consiga identificar espécies de anuros em função de dados pré-processados, com uma elevada percentagem de eficácia.

2.Descrição

2.1.Especificação

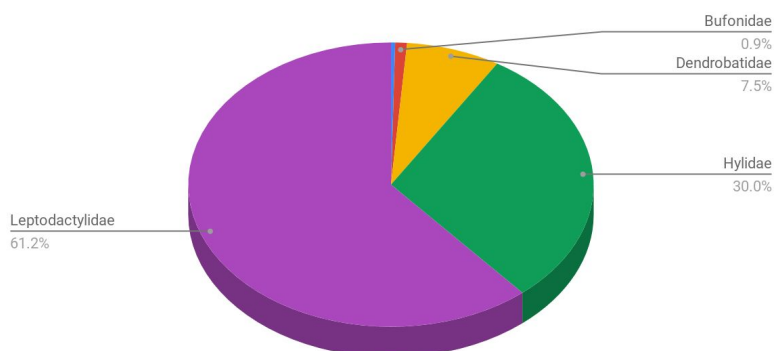
2.1.1. Descrição e análise do dataset.

O *dataset* utilizado intitula-se “Anuran Calls (MFCCs)”, foi retirado do repositório da UCI (Universidade da California, Irvine) [1], e contém registos de áudio de chamamentos de anuros. Conta com 4 famílias, 8 géneros, e 10 espécies, sendo que cada registo corresponde a um anuro.

É um *dataset* multilabel, com 3 colunas: família, género e espécie. Contém ainda uma coluna extra com o identificador da gravação. A distribuição das *labels* é a seguinte:

Famílias	Ocorrências
Bufonidae	68
Dendrobatidae	542
Hylidae	2165
Leptodactylidae	4420

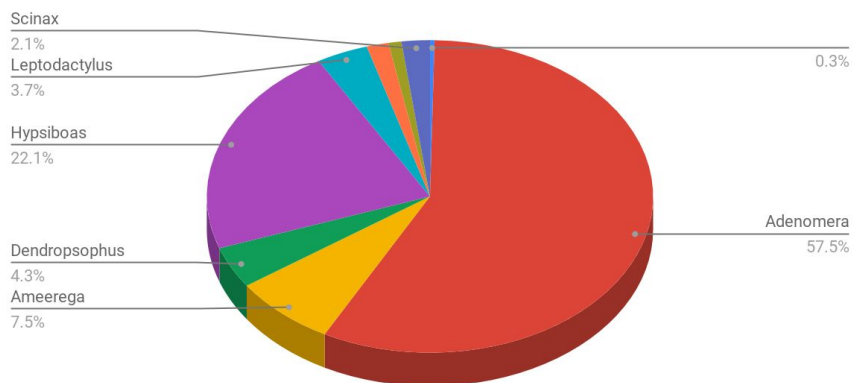
Distribuição Famílias



Legenda 1: Famílias presentes no dataset

Gêneros	Ocorrências
Adenomera	4150
Ameerega	542
Dendropsophus	310
Hypsiboas	1593
Leptodactylus	270
Osteocephalus	114
Rhinella	68
Scinax	148

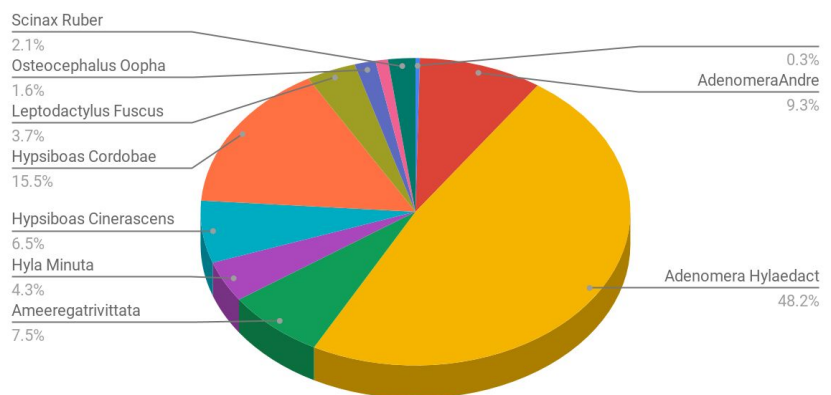
Distribuição Gêneros



Legenda 2: Gêneros presentes no dataset

Espécies	Ocorrências
AdenomeraAndre	672
Adenomera Hylaedact	3478
Ameeregatrivittata	542
Hyla Minuta	310
Hypsiboas Cinerascens	472
Hypsiboas Cordobae	1121
Leptodactylus Fuscus	270
Osteocephalus Oopha	114
Rhinellagranulosa	68
Scinax Ruber	148

Distribuição Espécies



Legenda 3: Espécies presentes no dataset

Analisando a distribuição dos dados, verifica-se uma frequência desequilibrada de dados, sendo 61%, 57%, e 48% dos registos de, respetivamente, Famílias, Gêneros, e

Espécies serem do mesmo tipo. Isto abre a possibilidade de espécies pouco frequentes na no *dataset* na fase de testes não serem corretamente identificadas, para além de uma geral tendência em favor dos *labels* de maior ocorrência.

Para a obtenção deste dataset, foi ainda necessário aos autores pré-processar os dados obtidos pelos aparelhos de medição desde a filtragem até a respetiva segmentação em sílabas. Posteriormente foi realizada a sua extração aplicando Mel-frequency cepstral coefficients (MFCCs) obtendo assim os valores dos atributos presentes em cada coluna na tabela.

2.1.2. Pré-processamento dos dados.

Inicialmente, por não ser identificada utilidade na mesma, foi retirada a coluna do identificador de gravação.

Depois de isoladas colunas correspondentes às *labels*, e dado que são *strings*, foi aplicada uma conversão “One Hot” às mesmas, para poderem ser utilizadas na rede.

Estando planeado o desenvolvimento de várias redes com diferentes tipos de *outputs*, foram ainda isoladas as colunas das *labels* já convertidas, para utilização posterior.

Finalmente, está prevista a divisão dos dados em percentagens ainda a definir, para serem utilizados isoladamente nas fases de treino e teste.

2.1.3. Modelo de aprendizagem a aplicar

No desenvolvimento do trabalho será utilizada uma rede neuronal do tipo Multilayer Perceptron (MLP). Este modelo de rede é do tipo *feedforward*, ou seja, a informação flui apenas num sentido: começa nos neurónios de input, passando pelos neurónios intermédios e acaba nos neurónios de output. Os neurónios intermédios formam as camadas intermédias (*hidden layers*). O número de *hidden layers* e o números de neurónios em cada uma é variável, mas cada neurónio está sempre ligado a todos os neurónios da camada seguinte, sendo que a todas as ligações está associado um peso que determina o nível de ativação do neurónio seguinte.

Neste tipo de rede neuronal, a aprendizagem é supervisionada. Por outras palavras, através do treino por pares input-output conhecidos, a rede procurar determinar um mapeamento o mais genérico possível que possa determinar o output para qualquer input que lhe seja posteriormente dado.

Para isso, é utilizado o algoritmo Backpropagation. Este algoritmo consiste em duas etapas de propagação em sentidos inversos: inicialmente, dado um input à rede, o estado é propagado no sentido direto (entrada, intermédia(s), saída) com a ajuda de uma função de transferência, tipicamente a função sigmóide, por se assemelhar à função degrau (*heaviside step*), conseguindo simultaneamente ter uma variação suave. O resultado da propagação produz o output determinado pela rede, sendo este comparado com o valor real. Em caso de igualdade, a rede está corretamente adaptada e não serão precisas alterações. Caso contrário, é efetuado o cálculo do erro utilizando a derivada da função de transferência em cada neurónio de saída, sendo depois os valores propagados no sentido inverso (saída, intermédia(s), entrada). Esta propagação é utilizada para ajustar os pesos das ligações de forma a minimizar a função de perda. De forma a que as variações nos pesos não oscilem demasiado e prejudiquem o processo de aprendizagem em caso de saltos incorretos, no

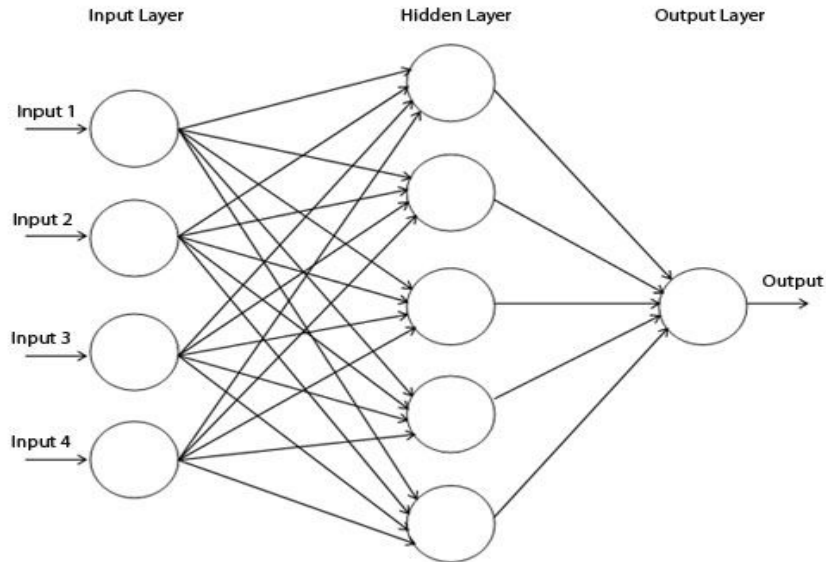
ajuste dos pesos é utilizado um fator de aprendizagem, um valor entre 0 e 1, tipicamente baixo (mas não muito baixo) de forma a atenuar estas variações.

Este procedimento é repetido até a rede convergir para um estado onde os erros começam a diminuir de frequência e, por conseguinte, os ajustes nos pesos das ligações começam a estabilizar.

2.1.4. Redes neuronais: arquitectura e configuração prevista da rede

Numa rede *feedforward*, como a que será utilizada, os neurónios estão organizados em camadas. Cada neurónio está ligado exatamente a todos os neurónios da camada seguinte (exceto os de saída). Assim sendo, pode-se dizer que a informação flui num sentido, correspondendo a informação de saída de uma camada à informação de entrada na próxima.

Ao nível das camadas, este tipo de redes são constituídas por **uma input layer**, primeira camada da rede, para onde é passada a informação inicial (assim sendo, o número de neurónios corresponde ao tamanho de dados de input), **uma ou mais hidden layers**, cada uma com um número de neurónios variável, auxiliaadoras da deteção de padrões na informação e **uma output layer**, onde as conclusões finais (resultado do reconhecimento) são guardados. O número de neurónios varia consoante o tipo de resultado que se pretende obter.



Legenda 4: Exemplo de rede feedforward com uma camada de cada tipo.

No caso deste trabalho, o número de hidden layers e a quantidade de neurónios das mesmas não está ainda definida. A partir de uma configuração base, serão aplicados e testados ajustes sucessivos, até a eficácia da rede ser satisfatória. Contudo, já está prevista a camada de input a utilizar: 22 neurónios correspondentes às colunas de cada registo do dataset.

Serão desenvolvidas duas redes: com um único output, correspondente à label de Espécie, por ser mais específica que as restantes; com 3 outputs, correspondentes às labels presentes, numa tentativa de testar se a informação adicional aumenta a eficácia dos resultados. Na prática o número de neurónios de output será maior em ambas as redes, dada a necessidade da aplicação de “One Hot Encoding”, mas conceptualmente, será esta a abordagem a seguir.

2.2.Trabalho Efectuado

Até à data, o grupo desenvolveu as seguintes fases do trabalho:

- Pesquisa acerca das tecnologias utilizadas na área para problemas relacionados com Redes Neurais Artificiais. Após testes com diversas ferramentas, o grupo optou por utilizar Keras [2], com Tensorflow [3] como *backend*.
- Pré-processamento dos dados, como descrito na secção 2.1.2;
- Treino de rede neuronal com arquitetura base, para testar as tecnologias escolhidas e verificar a processamento dos dados;
- Verificação da eficácia da rede obtida, com base nos dados de teste.

2.3.Resultados esperados e forma de avaliação

As redes serão treinadas e testadas usando informação proveniente de um único dataset. Desta forma, é necessária a divisão do mesmo em conjuntos independentes para testar a eficácia do modelo, sem risco do mesmo ser tendencioso em relação aos seus dados de treino.

Desta forma, o dataset será dividido em dois conjuntos, numa proporção inicial 70/30. O primeiro conjunto servirá como training set, ou seja, será utilizado pela rede durante o seu treino. O segundo conjunto servirá de teste ao treino efetuado. Consoante os resultados de eficácia obtidos, as percentagens poderão ser progressivamente ajustadas de forma a encontrar a quantia mínima ideal para a rede aprender a um nível que lhe permita obter predições posteriores a um nível satisfatório.

Poderá ainda ser feita uma divisão 60%-20%-20%, introduzindo um conjunto de validação, utilizado para aperfeiçoar as camadas intermédias das redes. Contudo, essa abordagem será utilizada caso a primeira não obtenha resultados satisfatórios.

3.Conclusões

A monitoração de populações biológicas é bem conhecida por ser uma tarefa complexa que envolve altos custos operacionais. Portanto, o desenvolvimento de novas metodologias capazes de medir quantidades de indivíduos em populações biológicas específicas sem contacto direto é desejado. O reconhecimento de espécies e indivíduos, com base na análise acústica de suas chamadas (Bioacústica), é possível para muitos animais e tem se mostrado uma ferramenta útil no estudo e monitoramento de espécies animais.

Para a realização deste trabalho é proposta uma metodologia supervisionada para identificação automática de anuros baseando-se na utilização das redes neuronais do tipo Multilayer Perceptron(MLP), que na fase de aprendizagem utiliza o algoritmo backpropagation. Esta metodologia já foi utilizada em testes preliminares para a escolha da framework a utilizar, tendo sido obtidos resultados bastante satisfatórios.

Assim, analisando o esforço até aqui realizado, é possível concluir que a metodologia adoptada será viável para cumprir o objetivo de obter uma elevada percentagem de eficiência na classificação dos resultados.

4. Recursos

[1] UCI Machine Learning Repository - <http://archive.ics.uci.edu/ml/datasets/Anuran+Calls+%28MFCCs%29>

[2] Keras - <https://keras.io/>

[3] Tensorflow - <https://www.tensorflow.org/>