

Sistema de Identificação de Espécies de Animais Utilizando Redes Neurais

João Gabriel Fazio Pauli¹, Fernando Seiji Onoda Inomata²

¹Universidade Estadual do Oeste do Paraná (Unioeste) - Cascavel - PR - Brazil

joao.gabriel.fazio@gmail.com, seijinomata0@gmail.com

Abstract. *This paper presents a system that identifies different animal species using a neural network trained on their physical and behavioral characteristics. The system allows users to input information about an animal and predicts the likely species based on this data. The paper describes the problem addressed, the system architecture, algorithms used, training parameters, results, and conclusions.*

Resumo. *Este artigo apresenta um sistema que realiza a identificação de diferentes espécies de animais usando uma rede neural treinada em suas características físicas e comportamentais. O sistema permite que os usuários insiram informações sobre um animal e preveja as espécies prováveis com base nesses dados. O artigo descreve o problema abordado, a arquitetura do sistema, algoritmos usados, parâmetros de treinamento, resultados e conclusões.*

1. Introdução

A identificação de animais é uma tarefa importante para a biologia, ecologia e outros campos científicos. Este projeto propõe uma abordagem baseada em aprendizado de máquina para classificar as espécies desses animais com base em características observáveis. O sistema utiliza uma rede neural treinada com dados previamente coletados, permitindo identificar espécies de forma eficiente e precisa.

2. Problema Abordado

O problema abordado é de **classificação**, onde o objetivo é identificar as espécies de um animal com base em entradas fornecidas pelo usuário. O sistema solicita respostas a perguntas sobre características físicas e comportamentais, e então utiliza um modelo de rede neural modelo MLP para classificar a espécie do animal descrito.

3. Arquitetura Utilizada

A aplicação web é construída utilizando o framework Flask, que funciona como uma interface web para coletar as informações desejadas pelo usuário, e o backend processa essas entradas e utiliza um modelo treinado para fazer a predição da espécie. A estrutura do sistema inclui:

- Carregamento e processamento de dados de arquivos CSV.
- Treinamento do modelo de rede neural usando 'scikit-learn'.
- Interface interativa para entrada e exibição dos resultados.

4. Algoritmo de Treinamento Utilizado

Para o treinamento da Rede Neural, foi utilizado o algoritmo MLPClassifier da biblioteca scikit-learn, treinado utilizando o método de backpropagation para realizar os ajustes dos pesos das conexões na rede neural.

5. Parâmetros Utilizados

- **Variáveis de entrada:** Características como: hair, feathers, eggs, milk, airborne, aquatic, predator, toothed, backbone, breathes, venomous, fins, has four or more legs, tail, domestic e catsize.
- **Variáveis de saída:** Classe do animal e seu nome.
- **Camadas Ocultas:** 3
- **Quantos neurônios em cada camada:** 64, 32, 16
- **Codificação dos Dados:** Os dados são transformados em valores binários (0 ou 1), indicando a presença ou ausência de uma característica. A variável de saída é uma classe numérica representando diferentes tipos de animais.
- **Função de Ativação:** ReLU.
- **Taxa de Aprendizado:** 0.001
- **Otimizador:** Adam.
- **Demais Parâmetros:** O modelo pode ser ajustado via GridSearchCV para encontrar a melhor combinação de hiperparâmetros, incluindo o número de camadas, neurônios, e funções de ativação.

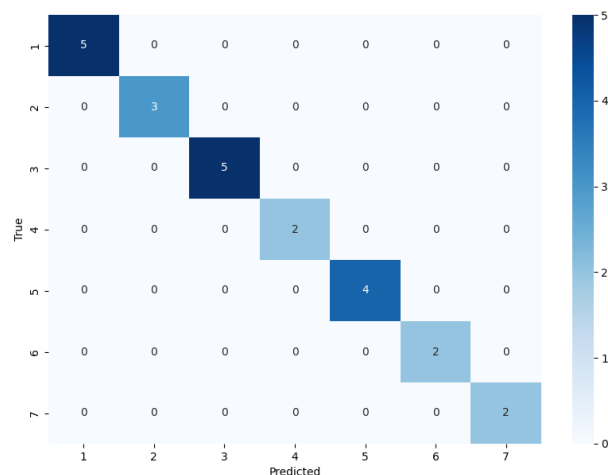


Figura 1. Matriz de confusão do modelo.

6. Resultados e Conclusões

Após o treinamento, o modelo foi testado com um conjunto de dados separado, e os resultados mostraram um percentual de acerto de aproximadamente 94%. A matriz de confusão é apresentada na Figura 1, indicando a eficácia da classificação. Além do mais, é válido ressaltar que durante os processos de treinamento da rede neural, ocorreu uma certa dificuldade na questão de normalização da variável "legs", onde a normalização não

estava sendo realizada corretamente, ocasionando em uma perda significativa de precisão no porcentual final.

Os resultados indicam que a abordagem trabalhada é eficaz para identificar a maioria das espécies de animais presentes no banco de dados. No entanto, há vários espaços para melhorias, como por exemplo, a adição de novas características poderia permitir uma distinção mais clara entre as espécies, visando evitar que 2 mesmas espécies possuam os mesmos atributos. Além disso, aumentar significativamente o número de animais no banco de dados utilizado fortaleceria ainda mais o modelo, proporcionando dados mais diversificados para o treinamento. Com essas melhorias, seria possível obter resultados mais precisos e um desempenho superior ao atual.

7. Referências

Referências

- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
- Grinberg, M. (2018). *Flask*.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. *Springer Series in Statistics*, 2.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). *Scikit-learn: Machine Learning in Python*.
- [Goodfellow et al. 2016] [Hastie et al. 2009] [Pedregosa et al. 2011] [Grinberg 2018]