

# Atividade 5 de Lab de IC sobre Redes Neurais

---

## Parte 1

Na parte 1 havia o objetivo de visualizarmos a convergência utilizando o algoritmo **Back Propagation**. Atráves deste algoritmo obtivemos valores consideráveis tanto para a função de perda quanto para a acurácia. Foi possível visualizar que não seriam necessárias mais que 500 épocas para atingir resultados consideráveis. Embora o algoritmo seja simples é possível perceber sua robustez na otimização do problema. Foi possível perceber, também, a capacidade do algoritmo em fazer um ajuste não linear para a obtenção das respostas desejadas.

## Parte 2

Nessa parte do projeto foi possível realizar uma comparação entre os modelos **RandomForest** do **ScikitLearn** e duas redes neurais montadas utilizando o **Keras**. Com a **RandomForest** foi possível obter uma acurácia de **0.771**, resultado relevante porém pouco maior que a média estimada primeiramente. O primeiro modelo de rede neural utilizando o **Keras** com uma camada escondida obteve uma acurácia de **0.766**, valor esse pouco menor que o anterior, exibindo uma certa similaridade entre os modelos. Porém as curvas **roc-auc** mostraram que o **RandomForest** possui um ajuste melhor ao problema utilizando estes parâmetros iniciais.

Na segunda parte do projeto houve a proposta da criação de um modelo de 2 camadas escondidas e 6 nós em cada camada, juntamente com uma variação entre os parâmetros. Algo que causou certa surpresa foi o fato do modelo ter apresentado o melhor ajuste justamente com os primeiros parâmetros descritos na atividade. Demais **solvers** e **learning\_rates** não apresentaram resultados satisfatórios estando todos abaixo da primeira configuração.

É importante pontuar que após a geração 400 os resultados não demonstraram ganhos tão relevantes, de modo que poderíamos ter parado o algoritmo na geração 500.

## Parte 3

A parte três consiste na comparação e na tentativa de melhoria do modelo de classificação de números escritos. Mesmo o **modelo\_1** e o **modelo\_2** sendo diferentes podemos compará-los utilizando as métricas disponíveis como **perda** e **acurácia**.

Dentre os dois modelos, o **modelo\_1** possui uma perda maior no treino, porém é composto por uma estrutura mais barata que gera uma acurácia bem próxima do **modelo\_2**. Com base neste parâmetro, o **modelo\_1** parece mais atrativo.

O comportamento da curva de treino e da curva de validação nos dois modelos apresenta um comportamento bem próximo. A partir de um determinado momento os dois modelos começam a apresentar um certo **overfitting** onde a curva de validação, após o seu mínimo, apresenta um acréscimo no decorrer das épocas. Embora o comportamento próximo, o valor da perda mínimo e o tempo de aquisição do mesmo ainda irão determinar a qualidade de cada modelo.

Em termos da acurácia, os gráficos possuem um comportamento próximo ao dos gráficos de perda, sendo a diferença o fato dos gráficos agora serem de curvas crescentes. Da mesma forma, através do gráfico de

acurácia, podemos perceber que a partir de um determinado ponto não há mais ganho na validação, apresentando o mesmo **overfitting** citado no gráfico de perda.

Dados os gráficos acreditamos que as métricas de acurácia acabam sendo mais significativas uma vez que demonstram, de maneira limitada, o ajuste de um modelo ao problema. Assim, essa métrica acaba sendo a mais significativa, na maioria das vezes, na seleção do modelo.

Com os testes realizados, não conseguimos obter resultados relevadamente melhores.