



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PROFESSOR: GEORGE ROSSANY S. DE LIRA  
DISCIPLINA: INTRODUÇÃO À CIÊNCIA DOS DADOS

PERÍODO 2022.1

ALUNO:  
LUAN FÁBIO MARINHO GALINDO (118 110 382)

## **RELATÓRIO**

### **CLASSIFICADOR NEURAL DO ESTADO OPERACIONAL DE PARA-RAIOS**

Campina Grande – PB

Fevereiro de 2023

## **SUMÁRIO**

1. INTRODUÇÃO
2. METODOLOGIAS DE DESENVOLVIMENTO E TESTES
3. ALGORITMO UTILIZADO
4. ANÁLISE DOS RESULTADOS OBTIDOS

## 1. INTRODUÇÃO

O estudo da ciência dos dados é uma área extremamente importante nos dias de hoje, devido a sua capacidade de escalabilidade de processos potencialmente complexos, junto com sua automação. Em particular, neste estudo, será analisado o uso de uma rede *Multilayer Perceptron (MLP)*, ou Perceptron de Múltiplas Camadas, para o desenvolvimento de um algoritmo de treinamento de vários *datasets* com possíveis defeitos na rede elétrica, como o aparecimento de harmônicos que resultam em distorção do sinal. O objetivo principal deste estudo é automatizar o processo de análise da forma de onda, minimizando a intervenção humana e aumentar a eficiência da detecção e tratamento desses possíveis defeitos.

## 2. METODOLOGIAS DE DESENVOLVIMENTO E TESTES

O algoritmo foi desenvolvido a partir de um *script* pré-editado, fornecido pela atividade anterior, juntamente com os *datasets*. A partir disso, inseriu-se os dados no algoritmo e fez-se as análises iniciais desses dados, para posterior tratamento. Destarte, fez-se a eliminação de erros de medição usando a função “*drop*”, da biblioteca *pandas*. Em seguida, foi realizada a subamostragem dos dados, por uso da função “*decimate*”, da biblioteca *scipy.signal*. Com isso, os dados foram concatenados, de modo a reunir os dados classificados como “bons” e “ruins” para efeito de treinamento da rede neural. O tratamento inicial desses dados fez-se por meio da função “*train\_test\_split*”, da biblioteca *scikit-learn*. Essa função é responsável por dividir o conjunto dos dados em subconjuntos para treino e teste da eficiência do algoritmo de forma aleatória e, neste caso, estratificada. Isso se deve ao fato de ser desejado que a divisão dos subconjuntos se dê de forma uniformizada, visando que o algoritmo não fique enviesado e dessa forma, menos eficiente.

## 3. ALGORITMO UTILIZADO

Uma rede MLP (Perceptron Multicamadas) é uma rede neural artificial composta por mais de uma camada de neurônios alimentados em cascata. Elas são usadas para processar dados de forma inspirada pelo cérebro humano e podem ser denominadas redes neurais artificiais (ANN) ou Perceptron de Múltiplas Camadas, conforme já citado. A rede multicamadas possui como característica o fato de haver uma camada de entrada, e outras ocultas (*hidden layers*), sendo que sua quantidade de camadas ocultas é determinada pelo projetista. O treinamento da rede é feito por meio do algoritmo de *backpropagation*, ou propagação reversa, onde os pesos de cada camada são alterados a cada iteração, de modo a minimizar o erro relativo entre a entrada e a saída esperada.

No algoritmo deste estudo, definiu-se alguns fatores da função de treinamento, como o máximo de iterações (500), o número de camadas ocultas (4) e a função de ativação como tangente hiperbólica.

#### 4. ANÁLISE DOS RESULTADOS OBTIDOS

Ao analisar-se os resultados obtidos, obteve-se uma taxa de acerto de 100% para a análise das tensões e uma taxa de acerto de 50% para as correntes. Para tentar melhorar os resultados, variou-se a quantidade de camadas ocultas do modelo, atingindo até 4096 camadas, mas os resultados permaneceram constantes. Com efeito, o algoritmo convergiu com apenas 11 iterações para a corrente e após 60 iterações para as tensões, e obteve as taxas de acertos já citadas em todos os testes realizados, usando os subconjuntos de teste. Os únicos problemas observados foram na plotagem dos gráficos da matriz de confusão, que não foram solucionados, devido a algum problema no ambiente do *Spyder*, que não conseguia realizar a importação da função *plot\_confusion\_matrix*. No entanto, foi possível visualizar a dita matriz com o uso de uma função mais elementar.

Com isso, conclui-se que os resultados foram um pouco aquém do esperado, dado que em uma das análises seus resultados são considerados insuficientes para uma aplicação mais prática. Isso pode ser confirmado pelos resultados obtidos nos testes de RNA presentes no código.