  
PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA  
DOUTORADO EM MÉTODOS DE APOIO À DECISÃO

**TRABALHO 1: ANÁLISE DE CRÉDITO BANCÁRIO**

**Gabriel Baruque**

**2012325**

Rio de Janeiro

2020

GABRIEL BARUQUE

**TRABALHO 1: ANÁLISE DE CRÉDITO BANCÁRIO**

Trabalho apresentado à disciplina ELE2394 – Redes Neurais I, do programa de Pós Graduação em Métodos de Apoio à Decisão da Coordenação de Engenharia Elétrica da PUC-Rio como requisito parcial de avaliação.

Professora: Marley Vellasco

Rio de Janeiro

2020

# Introdução

O objetivo deste trabalho é, a partir de uma base de dados, explorar diferentes configurações de redes neurais e técnicas de pré-processamento dos dados, a fim de se entender quais são os reflexos que essas mudanças trazem no resultado da classificação.

Para demonstrar um resumo dos resultados, serão construídas tabelas, que representarão as diferentes configurações utilizadas. Os elementos modificados durante o desenvolvimento do trabalho são:

* Normalização dos atributos de entrada
* Número de camadas
* Quantidade de neurônios em cada camada
* Número de épocas de treinamento
* Utilização de um conjunto de validação
* Codificação binária dos atributos de entrada
* Agrupamento de categorias

Uma discussão sobre os resultados será feita, trazendo as possíveis conclusões que podem ser feitas sobre o tema.

# Resultados

As tabelas a seguir mostram os resultados obtidos de acordo com cada configuração da rede, seguindo a ordem estipulada no enunciado do trabalho.

Tabela do Grupo I

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | GRUPO I | | | | | | | | |
|  | Configuração | | | | | | | | |
|  | a) | b) (hidden layer) | | c) (épocas) | | | d) | e) | f) |
|  | - | a) | b) | 5 | 500 | 5000 | valid.20% | valid.20% | valid.20% |
| Hidden Layer 1 | 3 | 3 | 5 | 2 | 2 | 2 | 9 | 3 | 3 |
| Hidden Layer 2 | - | - | - | 2 | 2 | 2 | - | - | 3 |
| Hidden Layer 3 | - | - | - | 2 | 2 | 2 | - | - | 3 |
| Class. Correta em % | 53.03 | 89.77 | 89.95 | 53.03 | 90.29 | 91.16 | 90.12 | 90.12 | 90.29 |
| Class Incorreta em % | 46.97 | 10.23 | 10.05 | 46.97 | 9.71 | 8.84 | 9.88 | 9.88 | 9.71 |
| Mean Abs. Error | 0.50 | 0.16 | 0.15 | 0.50 | 0.16 | 0.15 | 0.16 | 0.16 | 0.16 |
| Root Mean Sq. Error | 0.50 | 0.29 | 0.29 | 0.50 | 0.29 | 0.28 | 0.29 | 0.28 | 0.29 |
| Relative Abs. Error (%) | 99.46 | 31.40 | 29.80 | 99.58 | 31.88 | 30.71 | 31.49 | 32.41 | 32.72 |
| Root Relative Sq. Error (%) | 100.27 | 57.83 | 57.15 | 100.14 | 57.41 | 55.12 | 57.29 | 56.95 | 58.78 |

Tabela Grupo II

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | GRUPO I | | | | | | | | |
|  | Configuração | | | | | | | | |
|  | a) | b) (hidden layer) | | c) (épocas) | | | d) | e) | f) |
|  | - | a) | b) | 2 | 100 | 10000 | valid. 30% | - | - |
| Hidden Layer 1 | 10 | 10 | 20 | 10 | 10 | 10 | 10 | 10 | 5 |
| Hidden Layer 2 | - | - | - | - | - | - | - |  | 5 |
| Hidden Layer 3 | - | - | - | - | - | - | - |  | - |
| Class. Correta em % | 50.61 | 89.08 | 89.08 | 87.87 | 89.08 | 88.04 | 88.91 | 88.73 | 89.77 |
| Class Incorreta em % | 49.39 | 10.92 | 10.92 | 12.13 | 10.92 | 11.96 | 11.09 | 11.27 | 10.23 |
| Mean Abs. Error | 0.50 | 0.14 | 0.14 | 0.17 | 0.14 | 0.13 | 0.14 | 0.14 | 0.14 |
| Root Mean Sq. Error | 0.55 | 0.29 | 0.29 | 0.32 | 0.29 | 0.32 | 0.29 | 0.30 | 0.29 |
| Relative Abs. Error (%) | 99.52 | 27.61 | 27.77 | 33.60 | 27.61 | 25.14 | 28.75 | 28.98 | 28.20 |
| Root Relative Sq. Error (%) | 109.60 | 58.18 | 58.48 | 63.02 | 58.18 | 63.36 | 58.76 | 59.66 | 58.73 |

Tabela Grupo III

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | GRUPO I | | | | | | | | |
|  | Configuração | | | | | | | | |
|  | a) | b) (hidden layer) | | c) (épocas) | | | d) | e) | f) |
|  | - | I | II | 5000 | 3000 | 10000 | valid. 20% | valid. 20% | valid. 20% |
| Hidden Layer 1 | 4 | 10 | 50 | 4 | 50 | 20 | 20 | 20 | 20 |
| Hidden Layer 2 | 3 | - | - | 3 | - | 10 | 10 | 10 | 10 |
| Hidden Layer 3 | 2 | - | - | 2 | - | 5 | 5 | 5 | 5 |
| Class. Correta em % | 43.50 | 89.08 | 89.43 | 87.35 | 87.87 | 88.39 | 90.47 | 89.60 | 89.60 |
| Class Incorreta em % | 56.50 | 10.92 | 10.57 | 12.65 | 12.13 | 11.61 | 9.53 | 10.40 | 10.40 |
| Mean Abs. Error | 0.50 | 0.15 | 0.15 | 0.18 | 0.15 | 0.17 | 0.15 | 0.15 | 0.16 |
| Root Mean Sq. Error | 0.50 | 0.29 | 0.29 | 0.31 | 0.33 | 0.32 | 0.30 | 0.29 | 0.30 |
| Relative Abs. Error (%) | 100.98 | 29.34 | 29.95 | 36.67 | 30.33 | 33.20 | 30.82 | 30.59 | 31.95 |
| Root Relative Sq. Error (%) | 101.11 | 58.57 | 58.96 | 62.24 | 65.37 | 65.08 | 59.15 | 58.33 | 60.27 |

# discussão

Em geral, os resultados de todos os testes realizados em cada grupo foram muito semelhantes. Porém alguns pontos podem ser destacados:

* A normalização dos atributos de entrada, em todos os grupos, fez com que o desempenho da rede aumentasse muito, chegando próximo aos 90%. Portanto pode-se concluir que a normalização dos atributos é fator extremamente importante para a utilização de redes neurais, principalmente por fazer com que a ordem de grandeza das entradas seja semelhante, evitando uma maior influência de alguma entrada específica sobre as outras.
* Nos testes realizados, não houve diferença considerável entre a utilização de uma única camada escondida, ou várias camadas. Na verdade, em alguns casos, a rede com mais camadas escondidas obteve um desempenho um pouco menor, provavelmente por ser complexa demais para o problema em questão.
* No Grupo I, configuração c), ao se utilizar uma rede com 3 camadas escondidas e 2 neurônios em cada uma, apenas 5 épocas foram insuficientes para que os pesos dos neurônios fossem aprendidos. O desempenho da rede foi de aproximadamente 53% nesse caso.
* Também no Grupo II, configuração c), foi utilizada uma rede com 10 neurônios e uma camada escondida, treinando por apenas 2 épocas. Nesse caso, o acerto na classificação foi de aproximadamente 87,9%, menor do que os testes feitos com 100 e 10000 épocas.
* Com várias tentativas de observar um *overfitting*, em todas as configurações, esse problema não aconteceu, mesmo com uma quantidade muito grande de épocas para treinamento. No Grupo I, foram utilizadas 5000 épocas. No Grupo II, 10000. No grupo III 5000, 3000 e 10000 épocas. Percebe-se que em alguns casos, a rede perde um pouco da sua capacidade de classificação correta com tantas épocas de treinamento. Já em outros casos, há uma pequena melhora.
* A utilização do conjunto de validação, apresentou em geral uma melhora muito suave na classificação, nos Grupos II e III. Porém, o mesmo é importante uma vez que em problemas onde ocorra o *overfitting*, esse conjunto é responsável por auxiliar o momento certo de parar o treinamento através do *early stopping*, exatamente para evitar esse problema.
* A codificação binária dos atributos de entrada melhorou ligeiramente os resultados no Grupo III. Essa codificação é feita de modo que os atributos categóricos possam ser tratados como valores (requerido por alguns métodos de classificação), de modo que não haja ordem entre esses valores. Em geral é um método importante para tratar dados categóricos.
* O agrupamento de algumas categorias foi responsável por minimizar a quantidade de entradas que não continham informações, ou que possuíam pouca representatividade de informação para a rede. Porém, como em outros parâmetros, a melhora na classificação por conta desse pré-processamento, foi mínima.