

Doença Cardiovascular

O presente trabalho visa uma implementação de um sistema que saiba se o paciente teve ou não doença cardiovascular, a partir de três atributos: colesterol, idade e glicemia sérica em jejum.

”

Sumario

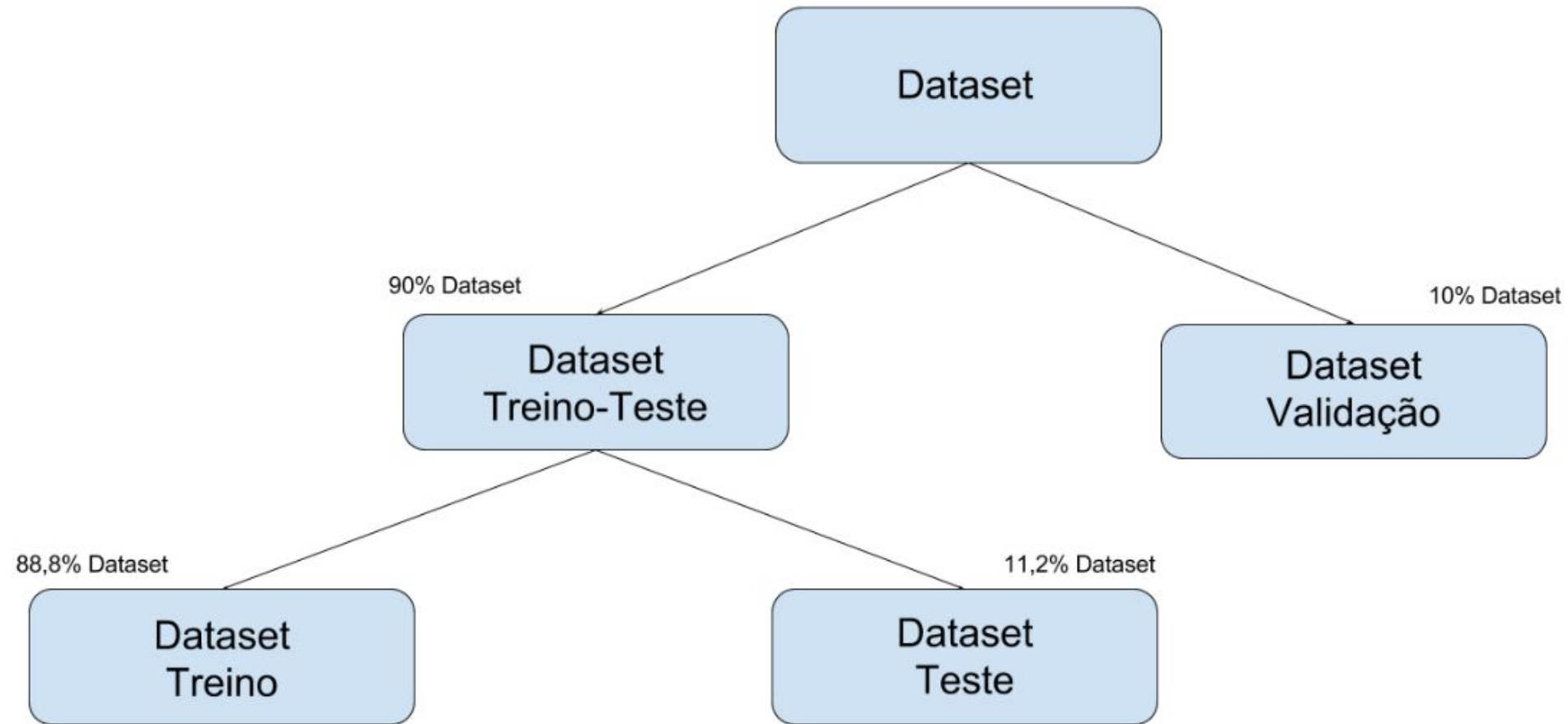
- ▷ Dataset
 - ▶ Estatística sobre o dataset
 - ▶ Matriz de Correlação
- ▷ Metodologia
 - ▶ Métricas
 - ▶ Ferramentas Utilizadas
- ▷ Experimentos e Resultados
 - ▶ Dimensionando a Rede Neural para o problema
 - ▶ Desempenho do modelo com a retirada de uma feature
 - ▶ Definindo os pesos para diferentes erros no treinamento
 - ▶ Validação
- ▷ Conclusão e Perguntas

1. 0 Dataset

Features

- ▷ Colesterol
- ▷ Glicemia
- ▷ Idade
- ▷ Desfecho

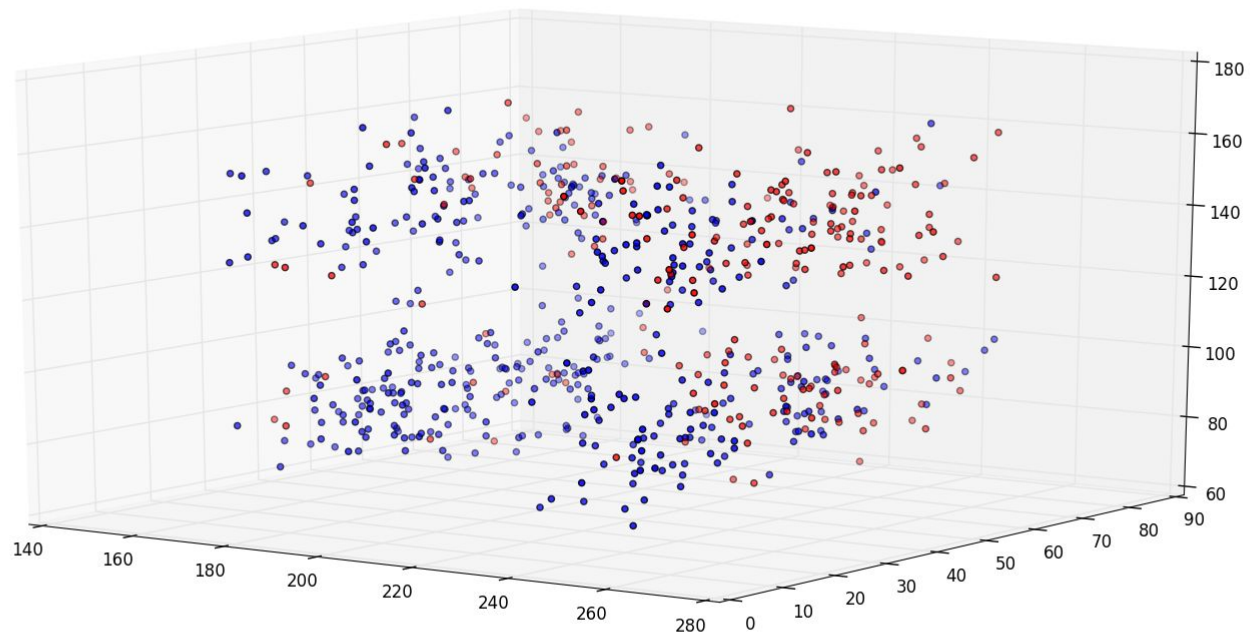
Features



Distribuição do Desfecho

Desfecho	Quantidade Individuos	Porcentagem
Com a Doença	281	30,78%
Sem a Doença	593	69,22%
Total	874	100%

Plot do Dataset



Matriz de Correlação

	Colesterol	Idade	Glicemia	Saída
Colesterol	1	0.0315	0.0083	0.3270
Idade	0.0315	1	0.069	0.3198
Glicemia	0.0083	0.069	1	0.233
Saída	0.3270	0.3198	0.233	1

2. Metodologia

Métricas

Para definir qual modelo escolher, utilizaremos duas métricas de comparação:

- ▷ Especificidade
- ▷ Sensibilidade

Sensibilidade

$$\frac{VP}{VP + FN}$$

VP = Verdadeiro Positivo

FN = Falso Negativo

Especificidade

$$\frac{VN}{VN + FP}$$

VN = Verdadeiro Negativo

FN = Falso Positivo

Ferramentas Utilizadas

- ▷ Linguagem Julia, versão 0.5
 - ▶ Pycall
 - ▶ IJulia
- ▷ Python, versão 3.5.2
 - ▶ IPython
 - ▶ Scikit-Learn
 - ▶ Theanets
 - ▶ Numpy
 - ▶ Matplotlib

3. Experimentos e Resultados

Experimento de Dimensionalidade da Rede

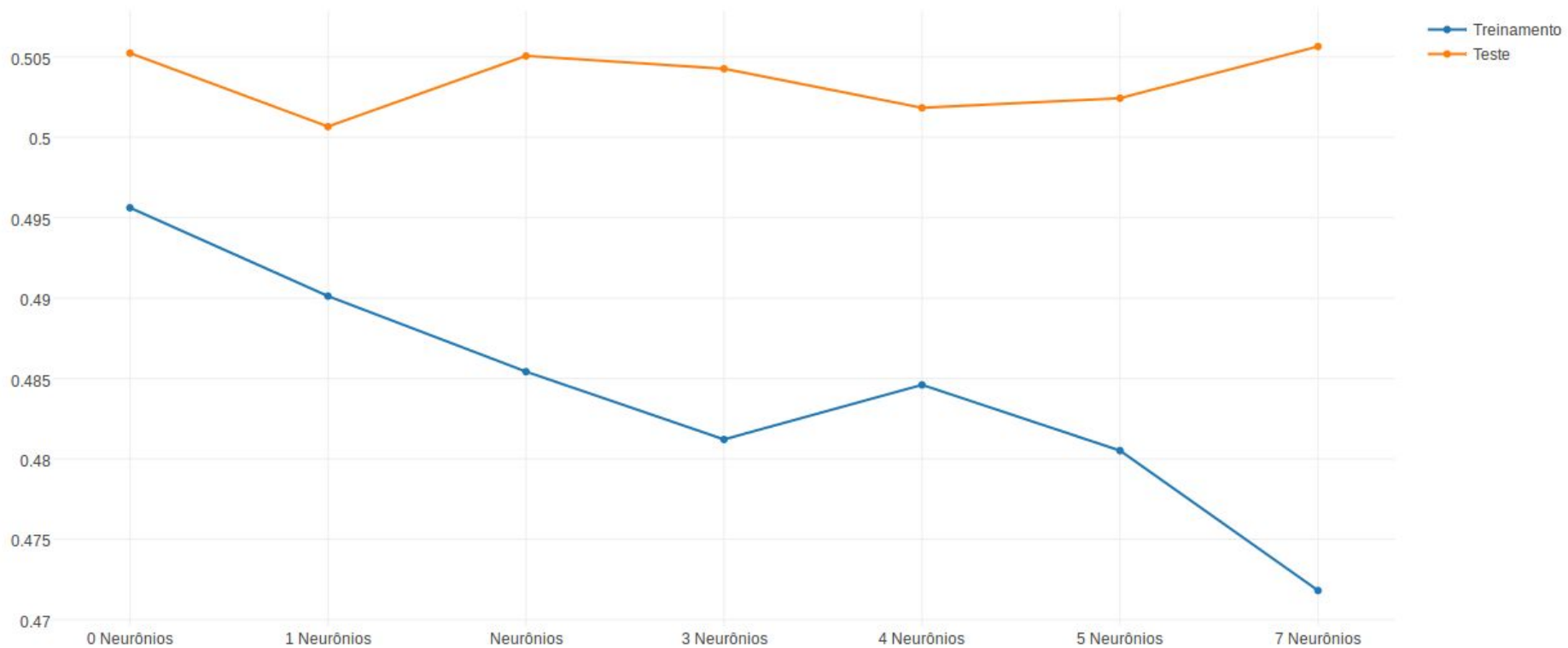
Foram Escolhidos 7 modelos de Rede Neural para este experimento:

- ▷ Sem camada Escondida
- ▷ 1 Camada de 1 Neurônios
- ▷ 1 Camada de 2 Neurônios
- ▷ 1 Camada de 3 Neurônios
- ▷ 1 Camada de 4 Neurônios
- ▷ 1 Camada de 5 Neurônios
- ▷ 1 Camada de 7 Neurônios

Experimento de Dimensionalidade da Rede

- ▶ Na camada escondida foi utilizado neurônios com Tangente Hiperbólica e na camada de entrada e saída neurônios lineares.
- ▶ Taxa de Aprendizagem: 0.01
- ▶ Gradiente descendente estocástico
- ▶ Função de Erro: Entropia Cruzada
- ▶ Validação Cruzada: K-Fold, com $K = 8$

Experimento de Dimensionalidade da Rede



Experimento de Dimensionalidade da Rede

Modelo	Sensibilidade	Especificidade
Sem camada Escondida	5.62e-01	8.74e-01

Experimento de Retirada de uma Feature

- ▷ Rede Neural com 2 neuronios na camada de entrada e saída, rede escolhida no experimento de dimensionalidade.
- ▷ Taxa de Aprendizagem: 0.01
- ▷ Gradiente descendente estocástico
- ▷ Função de Erro: Entropia Cruzada
- ▷ Validação Cruzada: K-Fold, com $K = 8$

Experimento de Retirada de uma Feature

Features	Erro	Sensibilidade	Especificidade
Colesterol e Idade	5.33e-01	4.78e-01	8.71e-01
Colesterol e Glicemia	5.60e-01	4.31e-01	8.58e-01
Glicemia e Idade	5.65e-01	4.28e-01	8.52e-01

Experimento de alteração de pesos

- ▷ Rede Neural com 3 neuronios na camada de entrada e 2 neurônios na saída, rede escolhida no experimento de dimensionalidade.
- ▷ Pesos Escolhidos para compor o experimento: 1,2,3,4,8
- ▷ Validação Cruzada: K-Fold, com $K = 8$

Experimento de alteração de pesos

Peso	Sensibilidade	Especificidade
1	5.64e-01	8.73e-01
2	7.68e-01	7.04e-01
3	8.60e-01	6.17e-01
4	9.00e-01	5.22e-01
8	9.40e-01	2.80e-01

Validação

- ▶ Rede Neural com 3 neuronios na camada de entrada e 2 neurônios na saída, rede escolhida no experimento de dimensionalidade.
- ▶ Taxa de Aprendizagem: 0.01
- ▶ Gradiente descendente estocástico
- ▶ Função de Erro: Entropia Cruzada

Modelo	Sensibilidade	Especificidade
Sem camada Escondida	9,09e-01	6.06e-01

3. Conclusão.

Conclusão.

Como podemos perceber através da implementação e dos experimentos, o algoritmo se comportou da maneira muito boa, tendo uma Sensibilidade e Especificidade, sem demonstrar uma grande variação na Validação.

Obrigado!

Perguntas?

Codigo e Relatorio disponivel em:
<https://github.com/hugdiniz/Infarto>