

Inteligência Computacional Aplicada

Avaliação Parcial II

Ronnasayd de Sousa Machado - 20161008636

6 de novembro de 2016

Sumário

1	MLP-Backpropagation Annealing Data-set	5
1.1	Considerações	5
1.2	Topologia da rede neural	6
2	MLP-Backpropagation Cardiotocography Data-Set	7
2.1	Considerações	7
2.2	Topologia da rede neural	7

Capítulo 1

MLP-Backpropagation Annealing Data-set

1.1 Considerações

Foram utilizados os dados disponíveis em "<https://archive.ics.uci.edu/ml/datasets/Annealing>" para classificação de cinco classes de processos além de uma classe indefinida. O *Data-set* apresenta 798 amostras com 38 atributos. A classe (1) apresenta 8 amostras, a classe (2) 88 amostras, a classe (3) 608 amostras, a classe (4) 0 amostras, a classe (5) 60 amostras e a classe (U) 34 amostras. Como a classe (4) não apresenta nenhuma amostras para treino decidiu-se, que a classe indefinida (U) seria a nova classe (4). Desta forma, foram classificadas 5 classes de processos.

Do total de 38 atributos apenas 13 foram utilizados, sendo estes:

- 1. family: GB,GK,GS,TN,ZA,ZF,ZH,ZM,ZS
- 3. steel: R,A,U,K,M,S,W,V
- 4. carbon: continuous
- 5. hardness: continuous
- 7. condition: S,A,X
- 8. formability: 1,2,3,4,5
- 9. strength: continuous
- 12. surface-quality: D,E,F,G

6 CAPÍTULO 1. MLP-BACKPROPAGATION ANNEALING DATA-SET

- 32. shape: COIL, SHEET
- 33. thick: continuous
- 34. width: continuous
- 35. len: continuous
- 37. bore: 0000,0500,0600,0760

Já que a rede neural trabalha apenas com valores numéricos as seguintes medidas foram propostas. Para os valores perdidos, representados pelo sinal de (?) no *Data-set*, estes foram substituídos pelo numero zero. Para os atributos selecionados que não eram contínuos, estes tiveram suas classes enumeradas em ordem crescente a partir do numero 1 em sua ordem de apresentação. Por fim todos os atributos foram normalizados em uma faixa de $[0,1]$.

Para lidar com o fato do desbalanceamento de classes. utilizou-se a técnica de *Oversampling*, que consiste em repetir as amostras das classes minoritárias. A classe (1) foi repetida 8 vezes totalizando 64 amostras e a classe (4) foi repetida 2 vezes totalizando 68 amostras.

As tabelas com os valores obtidos se encontram no diretório *"/TabelasAnnealing"*

1.2 Topologia da rede neural

Para a rede neural (MLP-Backpropagation) utilizou-se a seguinte topologia:

- Uma camada oculta com 3 neurônios
- Quantidade máxima de épocas igual 30000
- Função logística para todas as camadas
- Critério de parada de 10^{-5} para $|EQM_K - EQM_{K-1}|$
- Taxa de Aprendizagem de 0.01 e Taxa de Momentum de 0.9

Capítulo 2

MLP-Backpropagation Cardiotocography Data-Set

2.1 Considerações

Foram utilizados os dados disponíveis em "<https://archive.ics.uci.edu/ml/datasets/Cardiotocography>" para classificação de dez classes. O *Data-set* apresenta 2126 amostras com 21 atributos. Todos os atributos foram utilizados, no entanto estes foram normalizados na faixa de $[0,1]$.

As tabelas com os valores obtidos se encontram no diretório *"TabelasCtg"*

2.2 Topologia da rede neural

Para a rede neural (MLP-Backpropagation) utilizou-se a seguinte topologia:

- Uma camada oculta com 20 neurônios
- Quantidade máxima de épocas igual 30000
- Função logística para todas as camadas
- Critério de parada de 10^{-5} para $|EQM_K - EQM_{K-1}|$
- Taxa de Aprendizagem de 0.001 e Taxa de Momentum de 0.95