

RNA - Rede ELM

Problemas: Íris e Coluna Vertebral

Átila Camurça Alves

06 de Outubro de 2018

1 Introdução

ELM, do Inglês *Extreme Learning Machine*, é um novo algoritmo de aprendizagem para redes neurais de camada escondida única. Em comparação com o algoritmo de aprendizado de redes neurais convencionais, ele supera a velocidade lenta de treinamento e os problemas de ajuste excessivo. O ELM é baseado na teoria de minimização de riscos empíricos e seu processo de aprendizado precisa de apenas uma única iteração. O algoritmo evita várias iterações e minimização local. Ele tem sido usado em vários campos e aplicativos por causa da melhor capacidade de generalização, robustez e controlabilidade e rápida taxa de aprendizado.

2 Problemas

2.1 Íris

O problema da Íris é a classificação de uma espécie de flor. Essa base de dados é formada por 3 categorias: Setosa, Versicolor e Virgínica, onde:

- Setosa é classificada como classe $[1 \ 0 \ 0]$, com 50 itens na base;
- Versicolor é classificada como classe $[0 \ 1 \ 0]$, com 50 itens na base;
- Virgínica é classificada como classe $[0 \ 0 \ 1]$, com 50 itens na base;

2.2 Coluna Vertebral

O problema da Coluna Vertebral é a classificação para identificar se um paciente sofre ou não de algum problema na coluna. Essa base de dados possui as seguintes categorias: Hernia, Spondylolisthesis e Normal, onde:

- Hernia é classificada como classe $[1\ 0\ 0]$, com 60 itens na base;
- Spondylolisthesis é classificada como classe $[0\ 1\ 0]$, com 150 itens na base;
- Normal é classificada como classe $[0\ 0\ 1]$, com 100 itens na base;

3 Resultados

Os resultados de ambos problemas foram dentro do esperado, usando 50 neurônios na camada escondida. Seguem os resultados apresentando a matriz de confusão para as realizações com maior taxa de acerto.

3.1 Íris

No problema da Íris os resultados encontrados encontram-se acima de 86,00% na taxa de acerto.

Pior resultado:

```
==== Sumário =====
Num. Pred corretas: 26 de 30
      Realização: 12
Matriz de Confusão: [10 0 0;0 11 1;0 3 5]
      Taxa de acerto: 86.6667
```

Melhor resultado:

```
==== Sumário =====
Num. Pred corretas: 30 de 30
      Realização: 11
Matriz de Confusão: [8 0 0;0 13 0;0 0 9]
      Taxa de acerto: 100
```

Sumário:

```
==== Sumário Geral ====
      Acurácia: 95.3333
      Desvio Padrão: 3.9589
```

3.2 Coluna Vertebral

No problema da Coluna Vertebral os resultados encontrados encontram-se acima de 79,00% na taxa de acerto.

Pior resultado:

```
==== Sumário ====
Num. Pred corretas: 49 de 62
      Realização: 15
Matriz de Confusão: [6 0 6;0 27 4;3 0 16]
      Taxa de acerto: 79.0323
```

Melhor resultado:

```
==== Sumário ====
Num. Pred corretas: 58 de 62
      Realização: 17
Matriz de Confusão: [11 0 2;0 37 1;1 0 10]
      Taxa de acerto: 93.5484
```

Sumário:

```
==== Sumário Geral ====
      Acurácia: 86.8548
      Desvio Padrão: 3.598
```

4 Conclusão

Devido a falta de um algoritmo de seleção de parâmetros, por exemplo busca em grade com validação cruzada de k-partes, os resultados obtidos foram dentro do esperado mas poderiam ser melhores e mais eficientes.

Repositório com código-fonte: <https://github.com/atilacamura/rna-2018-1>

Link para download: <https://github.com/atilacamura/rna-2018-1/archive/master.zip>