

ELM - Extreme Learning Machines

1 Introdução

Como sabemos, existem redes neurais com retropropagação de informações em sua fase de treinamento. Neste tipo de rede a informação flui nos sentidos diretos, para o cálculo de uma saída, e inverso, para atualização de seus pesos. O maior representante deste tipo de rede neural é o MLP (*Multilayer Perceptron*). O principal algoritmo de treinamento para este tipo de rede é o *Backpropagation* que é baseado no algoritmo do gradiente descendente. Porém, a fase de treinamento deste tipo de rede, em geral, é muito custosa para grandes bases de dados, além de serem extremamente complexas.

As redes ELM ou *Extreme Learning Machines* são baseadas em *Feedforward*. Neste tipo de arquitetura as informações fluem em apenas um sentido - o direto. Isso se dá principalmente pelo fato de que os pesos das camadas escondidas não precisam ser ajustados e os valores da camada de saída são obtidos com OLAM, ou solução por mínimos quadrados. Devido à isto, seu treinamento é extremamente rápido se comparado com redes com retropropagação de informações. Além disso, mostram-se eficazes tanto para tarefas de classificação quanto para tarefas de regressão. A Figura 1 mostra a arquitetura tradicional de uma rede ELM.

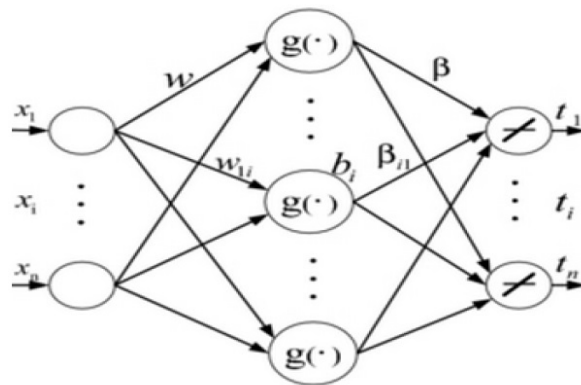


Figura 1: Arquitetura típica de uma rede ELM

Este documento visa relatar a eficiência do método apresentado acima para quatro bases de dados: base de dados "Iris", base de dados "Coluna Vertebral", base de dados

”Dermatologia” é uma base de dados gerada artificialmente. As três primeiras bases serão para testes de classificação enquanto que a base artificial será para testes de regressão. As métricas utilizadas foram a acurácia e o desvio padrão da taxa de acertos para classificação. Para regressão foi utilizado o erro quadrático médio para 20 realizações com uso do método da validação cruzada.

2 Bases de dados

Nesta seção faremos breves esclarecimentos sobre os conjuntos de dados que foram utilizadas para os testes.

2.1 Base de dados da ”Iris”

A base de dados da Iris é, provavelmente, o conjunto de dados mais conhecido quando falamos em reconhecimento de padrões e classificação. Este conjunto de dados se refere à classificação de 3 espécies de flores para plantas do gênero Iris: Iris-setosa, Iris-versicolor e Iris-virginica. Cada padrão é composto por uma tupla com 5 atributos: comprimento da sépala, largura da sépala, comprimento da pétala, largura da pétala e sua classificação. Neste conjunto de dados temos 150 padrões com 50 representantes de cada uma das classes mencionadas acima.

2.2 Base de dados da ”Coluna Vertebral”

A base de dados da coluna vertebral dá conta da classificação de pacientes em relação às duas patologias da coluna vertebral: Hérnia de Disco e Espondilolistese. Cada padrão deste conjunto de dados é composto por uma tupla de 6 atributos: forma e orientação da pélvis, forma e orientação de coluna lombar, incidência pélvica, inclinação pélvica, ângulo de lordose lombar, inclinação sacral, raio pélvico e grau de espondilolistese. Além disso, cada tupla é classificada entre Hérnia de Disco, Espondilolistese e Normal. Neste conjunto de dados temos 310 entradas com 60 entradas relacionadas às pacientes com Hérnia de Disco, 150 entradas relacionadas às pacientes com Espondilolistese e mais 100 entradas relacionadas às pacientes sem patologias.

2.3 Base de dados da ”Dermatologia”

Esta base de dados está relacionada à classificação de pacientes portadores de patologias dermatológicas. Mais especificamente, cada padrão está relacionado a uma entre as 6 patologias mencionadas a seguir: Psoríase, Dermatite Seborreica, Líquen Plano, Pitiríase Rósea, Dermatite Crônica e Pitiríase Rubra Pilar. Cada padrão deste conjunto de dados é composto por uma tupla com 34 atributos. Dentre estes atributos podemos citar: Eosinófilos de Infiltrado, Infiltrado de LPN, Fibrose da Derma Papilar, entre outros. Este conjunto de dados possui 366 entradas com 112 instâncias de Psoríase,

61 instancias de Dermatite Seborreica, 72 instancias de Líquen Plano, 49 instancias de Pitiríase Rósea, 52 instancias de Dermatite Crônica e 20 instancias de Pitiríase Rubra Pilar.

2.4 Base de dados artificial

Para testar a eficiência do método para a aproximação de funções, foi gerada uma base de dados com 1000 entradas. Este conjunto de dados é baseada em uma função senoidal com um certo ruído adicionado.

3 Testes e Resultados

Nesta seção serão apresentados os resultados obtidos após a aplicação do Perceptron multi-camadas para os três conjuntos de dados apresentados na seção anterior.

3.1 Resultados para base de dados da "Iris"

Neste conjunto foram 10 realizações. Além disso, a saída foi obtida com uso do OLAM. A taxa de acerto para cada realização é mostrada na Figura 2.

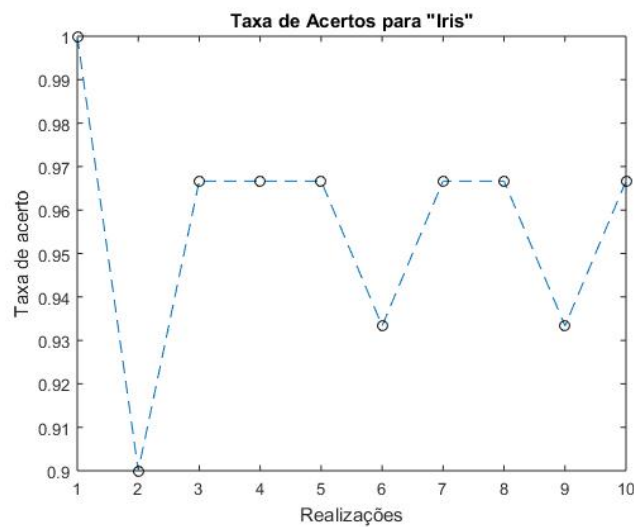


Figura 2: Taxa de acerto para a base de dados da "Iris"

3.2 Resultados para base de dados da "Coluna Vertebral"

Neste conjunto foram 10 realizações. Além disso, a saída foi obtida com uso do OLAM. A taxa de acerto é mostrada abaixo, na Figura 3.

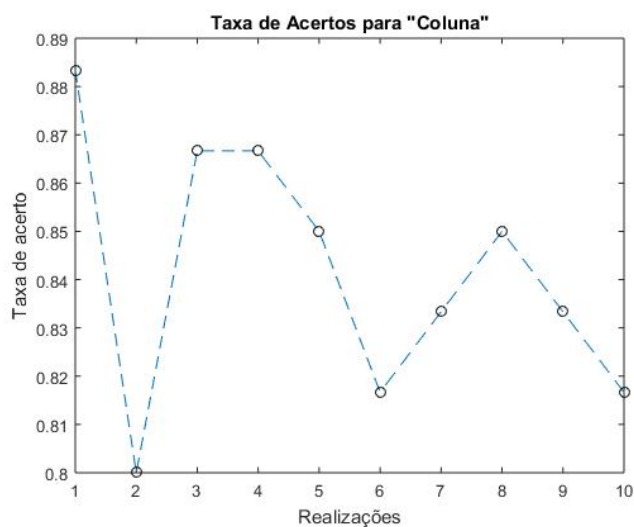


Figura 3: Taxa de acerto para a base de dados da "Coluna Vertebral"

3.3 Resultados para base de dados "Dermatologia"

Neste conjunto de dados foram 10 realizações. Além disso, a saída foi obtida com uso do OLAM. A taxa de acerto é mostrada abaixo, na Figura 4.

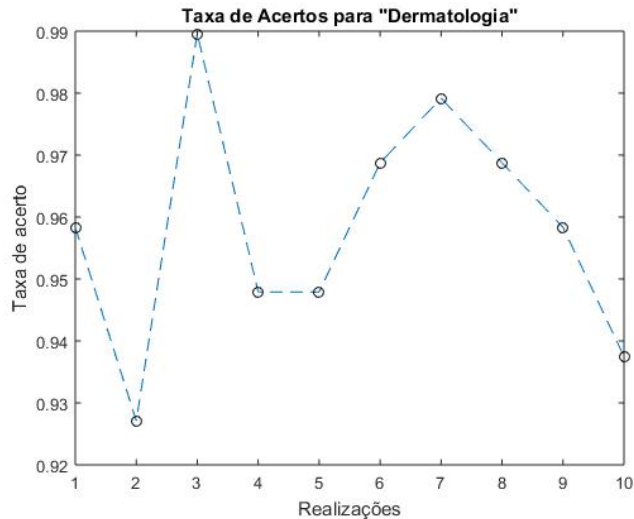


Figura 4: Taxa de acerto para a base de dados "Dermatologia"

3.4 Resultados para base de dados Artificial

Para este conjunto foram feitas 10 realizações. A função de ativação é linear. Os resultados visuais são mostrados na Figura 5.

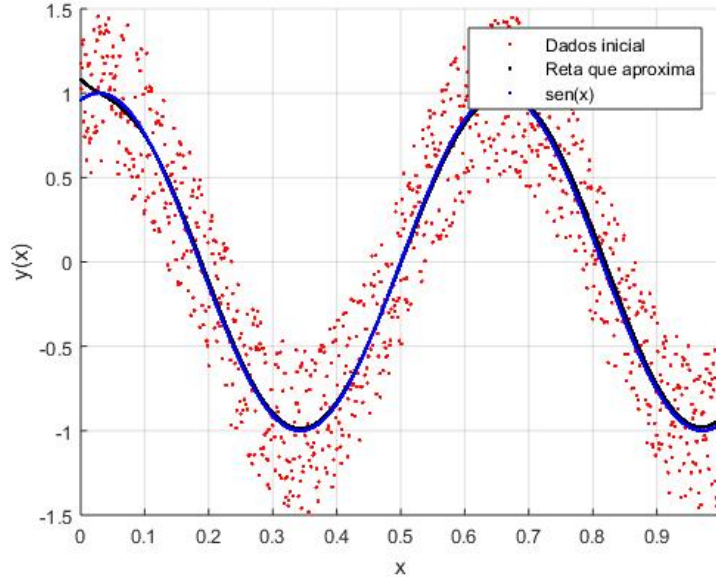


Figura 5: Resultados para regressão

3.5 Resultados gerais para todas as bases de dados utilizadas

A tabela abaixo resume os resultados obtidos para a utilização de todos os conjuntos de dados apresentados na seção 2 para classificação.

Base de dados	Acurácia	Desvio padrão
Iris	0.9633	0.02460
Coluna Vertebral	0.8533	0.0526
Dermatologia	0.9604	0.0128

Para o conjunto de dados artificial, temos um RMSE médio de 0.2926 com desvio padrão de 0.0070 para 10 realizações.

4 Conclusão

A partir dos resultados mostrados acima podemos comprovar que as redes ELM são bastante eficazes tanto para classificação quanto para regressão. Além disso, o seu treinamento é, de fato, bastante veloz devido ao OLAM.