

Classificação de Flores com Perceptron

Inicialização

- Os pesos w_1, w_2, \dots, w_n e o bias b são iniciados com zeros.
- O usuário define a proporção dos dados que será usada no treinamento, e a taxa de aprendizado que deseja utilizar.
- No programa são criados 3 perceptrons, um para cada flor.
- Cada flor é tratada como classificação binária (vs todas as outras). Um perceptron é treinado para cada uma das três espécies com os dados de treinamento.

Processo

1. Com a entrada $x = (x_1, x_2, \dots, x_n)$ e uma saída esperada y (sendo 0 ou 1)
2. Aplica a função de ativação:

$$saída = \begin{cases} 1 & \text{se } z \geq 0 \\ 0 & \text{se } z < 0 \end{cases}$$

3. Compara com a saída esperada:
Se a saída gerada for diferente da esperada, o perceptron atualiza os pesos.

Atualização dos pesos

$$w_i = w_i + \alpha * (y - \hat{y}) * x_i$$

$$b = b + \alpha * (y - \hat{y})$$

- y = saída esperada
- \hat{y} = saída do perceptron

- α = taxa de aprendizado, evita que a correção seja brusca demais
- x_i = valor de entrada, indica quais dimensões do espaço precisam de mais correção

Teste

- Cada perceptron passa pelo teste com seus respectivos dados.
- Devido a aleatoriedade com que os dados são embaralhados, a taxa de precisão se mostrou um tanto inconsistente, com ênfase na versicolor, porém os demais tiveram resultados satisfatórios, geralmente acima de 80%.

Análise

Os dados das tabelas a seguir foram construídos com taxa de aprendizado igual a 0.1 e porcentual de treinamento igual a 90%.

Setosa

comprimento da sépala	largura da sépala	comprimento da pétala	largura da pétala	Precisão (%)
0.15	0.33	-0.60	-0.27	100
0.14	0.47	-0.73	-0.27	100
0.20	0.51	-0.75	-0.31	100
0.09	0.44	-0.79	-0.36	100
0.08	0.41	-0.74	-0.35	100
0.18	0.49	-0.73	-0.31	100
0.32	0.89	-1.45	-0.71	100
0.04	0.58	-0.77	-0.35	100
0.12	0.32	-0.61	-0.30	100
0.22	0.64	-0.98	-0.49	100

Versicolor

comprimento da sépala	largura da sépala	comprimento da pétala	largura da pétala	Precisão (%)
0.25	-3.21	0.67	-2.37	60

0.97	-3.13	0.6	-3.35	60
0.65	-2.57	0.48	-3.11	73.33
1.85	-3.83	0.36	-2.67	80
-0.58	-1.64	1.18	-2.04	53.33
1.64	-3.65	1.14	-3.08	86.67
1.35	-3.96	0.87	-2.62	53.33
1.23	-4.31	-0.36	-3.09	53.33
1.30	-3.18	0.92	-1.52	73.33
1.46	-3.41	0.36	-2.06	73.33

Virginica

comprimento da sépala	largura da sépala	comprimento da pétala	largura da pétala	Precisão (%)
-2.94	-2.68	4.11	4.33	100
-2.37	-1.99	3.32	3.12	93.33
-3.35	3.40	5.01	4.95	100
-3.02	-2.51	4.56	3.58	100
-1.99	-2.15	3.48	3.53	66.67
-3.83	-2.68	4.17	3.59	73.33
-3.07	-2.41	4.47	3.85	100
-3.03	-1.98	4.21	3.49	93.33
-3.22	-2.91	4.25	3.16	86.67
-3.21	-3.39	5.16	3.84	100

Conclusões

Após alguns procedimentos empíricos foi constatado:

- Taxa de aprendizado ideal é entre 0.1 e 0.5
- A porcentagem ideal para o treinamento é entre 70% e 90%.
- A precisão da classe **Setosa** foi consistentemente alta.
- A classe **Versicolor** apresentou maior dificuldade de classificação.

- A classe **Virginica** obteve resultados intermediários, mas satisfatórios.

Setosa:

- Comprimento da sépala - Baixa influência.
- Largura da sépala - Indicativo de que é setosa.
- Comprimento da pétala - Indicativo de que não é setosa.
- Largura da sépala - Indicativo de que não é setosa.

Versicolor:

- Comprimento da sépala - Alto indicativo de que é versicolor.
- Largura da sépala - Alto indicativo de que não é versicolor.
- Comprimento da pétala - Indicativo de que é versicolor.
- Largura da sépala - Alto indicativo de que não é versicolor.

Virginica

- Comprimento da sépala - Alto indicativo de que não é virginica.
- Largura da sépala - Alto indicativo de que não é virginica.
- Comprimento da pétala - Alto indicativo de que é virginica.
- Largura da sépala - Alto indicativo de que é virginica.