Phân loại tuyến tính

Nhiệm vụ

|  |  |
| --- | --- |
| **X** | **y** |
| -2 | -1 |
| -1 | 1 |
| 0 | -1 |
| 1 | 1 |
| 2 | 1 |

Kiểu phân loại này dựa trên việc xác định dấu của hàm, có dạng:

ở đâu là các đặc điểm mô tả đối tượng, vectơ ( là trọng số của các đặc điểm tương ứng, là ngưỡng quyết định, n là số lượng đặc điểm.

Khi giải bài toán phân loại tuyến tính, chúng ta xác định hai lớp có nhãn lần lượt là 1 và -1. Khi đó lời giải cho bài toán phân loại trong trường hợp này sẽ là biểu thức sau:

Hãy xem một ví dụ. Giả sử có dữ liệu về các đối tượng được mô tả bởi hai đặc tính và một đặc tính kết quả nhận giá trị 1 hoặc -1.

Hãy vẽ dữ liệu này.

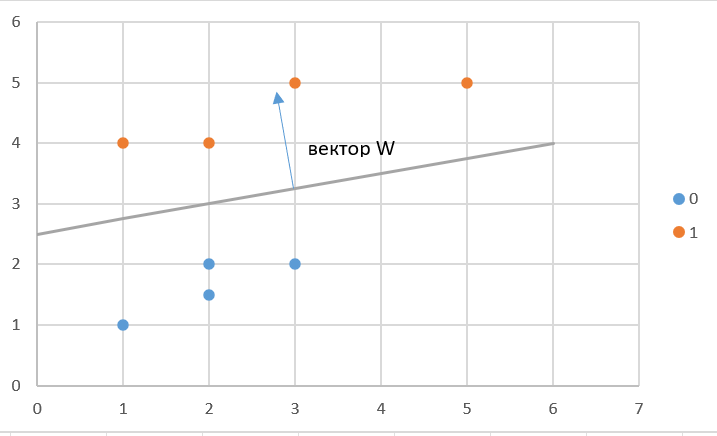
Như bạn có thể thấy trong biểu đồ, bạn có thể thử xây dựng một đường thẳng có thể chia dữ liệu thành các lớp, nghĩa là bên dưới đường thẳng sẽ có các đối tượng có nhãn lớp là -1 và trên +1. Hãy vẽ một đường thẳng lý thuyết.

Như có thể thấy trong hình, đường thẳng đã chia thành công dữ liệu của chúng ta thành 2 lớp. Tất cả các đối tượng nằm dưới đường thẳng sẽ thuộc lớp -1. Khi đó lời giải cho bài toán lọc tuyến tính sẽ như sau:

Nếu , thì đối tượng thuộc lớp +1,

Nếu , thì đối tượng thuộc lớp -1.

Do đó, nếu bạn thay thế các giá trị cụ thể của x và y vào phương trình này, dấu của biểu thức thu được sẽ xác định tư cách thành viên của lớp. Trong thực tế, là tích vô hướng của vectơ X và vectơ W. Đường dựng nên được xác định bởi biểu thức Do đó, vectơ W trực giao với đường dựng nên về mặt này, dấu của tích vô hướng sẽ được xác định bằng cosin của góc giữa chúng. Nếu vật nằm dưới đường thẳng thì góc sẽ tù, tức là giá trị cosin giữa các vectơ đang xét sẽ có dấu -.

**

Làm thế nào để xác định vị trí tối ưu của đường phân chia? Để làm điều này, cần phải xác định hàm mất mát.

Chức năng không thành công trong 2 trường hợp:

1. , nhưng đối tượng thuộc lớp -1;
2. , nhưng đối tượng thuộc lớp +1.

Khi đó hàm mất M được gọi là offset và sẽ có dạng sau:

Hoặc:

Nhưng hàm này không liên tục nên hàm số lớn:

Hàm đa số hóa F là hàm được sử dụng để đánh giá hoặc "đại hóa" một hàm khác bằng cách cung cấp giới hạn trên cho các giá trị của nó , đó là:

Hàm số như vậy có thể là F = (1- M ) 2

Hãy tính toán các vết lõm:

|  |  |
| --- | --- |
| **M** | **Thụt lề** |
| M1 |  |
| M2 |  |
| M3 |  |
| M4 |  |
| M5 |  |

Sau đó, bạn cần giảm thiểu biểu thức:

… =

=

Hãy tối ưu hóa hàm này bằng cách lấy đạo hàm riêng theo và cho chúng bằng 0:

Nghiệm của phương trình sẽ có hai giá trị:

*Do đó, trình phân loại sẽ như sau: nếu 0,4x+0,2>0 thì đối tượng thuộc lớp 1 (nếu không thì thuộc lớp -1).*

Ví dụ được thảo luận trước đó rất đơn giản và việc tìm điểm tối thiểu (giá trị tối ưu của trọng số w i ) không khó. Tuy nhiên, nếu có nhiều biến thì hệ thống sẽ phát triển và rất khó tìm ra giải pháp cho hệ thống như vậy.

Vì vậy, chúng tôi sẽ sử dụng một phương pháp tối ưu hóa khác.

Là một phương pháp tối ưu hóa, chúng tôi sẽ sử dụng phương pháp giảm độ dốc ngẫu nhiên. Phương pháp giảm độ dốc ngẫu nhiên giả định thuật toán cập nhật trọng số sau:

ở đâu , thường là một con số nhỏ, ví dụ 0,05

Bây giờ hãy giải quyết ví dụ của chúng ta bằng cách sử dụng phương pháp giảm độ dốc:

Theo định nghĩa, gradient sẽ là đạo hàm riêng đối với các tham số của chúng tôi và , và chúng tôi đã tìm thấy chúng:

Hãy lấy một điểm tùy ý. Ví dụ: (1,1), đó là:

Chúng tôi tính toán giá trị của gradient tại thời điểm này, thay thế các đơn vị thay vì và . Và sau đó chúng tôi tính toán tọa độ mới của điểm:

Điểm mới sẽ là

Hãy tính thêm một bước gradient nữa

Và do đó, chúng tôi thực hiện các bước giảm độ dốc cho đến khi các hệ số ngừng thay đổi.