5. The Perceptron Algorithm

The Perceptron Algorithm

0:00 / 0:00 1.50x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Perceptron Concept Questions 1

1/1 point (graded)
Remember that the Perceptron Algorithm (without offset) is stated as the following:

Perceptron $\left(\left\{\left(x^{(i)}, y^{(i)}\right), i=1, \ldots, n\right\}, T\right)$:

initialize $\theta = 0$(vector);
for $t = 1, \ldots, T$ do
for $i = 1, \ldots, n$ do

if $y^{(i)}\left(\theta \cdot x^{(i)}\right) \leq 0$ then
update $\theta = \theta + y^{(i)} x^{(i)}$

What does the Perceptron algorithm take as inputs among the following? Choose all those apply.

- [x] Training set
- [x] T - the number of times the algorithm iterates through the whole training set
- [ ] Test set
- [ ] θ
- [ ] θ_0

✔

Solution:

The perceptron algorithm takes T and the training set as input, and aims to learn the optimal "θ","θ_0"

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Perceptron Update 1

1/1 point (graded)
Now consider the Perceptron algorithm with Offset. Whenever there is a "mistake" (or equivalently, whenever $y^{(i)}\left(\theta \cdot x^{(i)} + \theta_0\right) \leq 0$ i.e. when the label y^i and $h\left(x\right)$ do not match), perceptron updates

$$\theta \text{ with } \theta + y^{(i)} x^{(i)}$$

and

$$\theta_0 \text{ with } \theta_0 + y^{(i)}.$$

More formally, the Perceptron Algorithm with Offset is defined as follows:

Perceptron $\left(\left\{\left(x^{(i)}, y^{(i)}\right), i = 1, \ldots, n\right\}, T\right)$:
initialize $\theta = 0$(vector); $\theta_0 = 0$(scalar)
for $t = 1, \ldots, T$ do
for $i = 1, \ldots, n$ do
if $y^{(i)}\left(\theta \cdot x^{(i)} + \theta_0\right) \leq 0$ then
update $\theta = \theta + y^{(i)} x^{(i)}$

update $\theta_0 = \theta_0 + y^{(i)}$

In the next set of problems, we will try to understand why such an update is a reasonable one.

When a mistake is spotted, do the updated values of θ and θ_0 provide a better prediction? In other words, is

$$y^{(i)}\left((\theta + y^{(i)}x^{(i)}) \cdot x^{(i)} + \theta_0 + y^{(i)}\right)$$

always greater than or equal to

$$y^{(i)}\left(\theta \cdot x^{(i)} + \theta_0\right)$$

- Yes, because $\theta + y^{(i)}x^{(i)}$ is always larger than θ
- Yes, because $\left(y^{(i)}\right)^2 \|x^{(i)}\|^2 + \left(y^{(i)}\right)^2 \geq 0$
- No, because $\left(y^{(i)}\right)^2 \|x^{(i)}\|^2 - \left(y^{i}\right)^2 \leq 0$
- No, because $\theta + y^{(i)}x^{(i)}$ is always larger than θ

Solution:

Comparing the two terms,

$$y^{(i)}\left((\theta + y^{(i)}x^{(i)}) \cdot x^{(i)} + \theta_0 + y^{(i)}\right) - y^{(i)}\left(\theta \cdot x^{(i)} + \theta_0\right) = \left(y^{(i)}\right)^2 \|x^{(i)}\|^2 + \left(y^{(i)}\right)^2 = \left(y^{(i)}\right)^2 \left(\|x^{(i)}\|^2 + 1\right) > 0$$

the first is always greater than the latter. Considering that our goal is to minimize the training error, the update always makes the training error decrease, which is desirable.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Perceptron Update 2

0 points possible (ungraded)

For a given example i, we defined the training error as 1 if $y^{(i)}\left(\theta \cdot x^{(i)} + \theta_0\right) \leq 0$, and 0 otherwise:

$$\varepsilon_i\left(\theta, \theta_0\right) = \left[\left[y^{(i)}\left(\theta \cdot x^{(i)} + \theta_0\right) \leq 0\right]\right]$$

Say we have a linear classifier given by θ, θ_0. After the perceptron update using example i, the training error $\varepsilon_i\left(\theta, \theta_0\right)$ for that example can (select all those apply):

- [ ] Increase
- [x] Stay the same
- [x] Decrease

✔

Solution:

From the previous problem, we saw that $y^i\left(\theta \cdot x+\theta_0\right)$ increases after the perceptron update. Thus $\left[\left[y^i\left(\theta \cdot x+\theta_0\right) \leq 0\right]\right]$ becomes zero or stays 1.

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 1 Linear Classifiers and Generalizations (2 weeks):Lecture 2. Linear Classifier and Perceptron / 5. The Perceptron Algorithm

Add a Post

Show all posts by recent activity

Recitation 2
Many of the questions below will be addressed in the recitation video coming out next week! 2
Pinned Staff

Lecture Question
Dear Staff, Hello, can you please explain to me the following. how did the professor get the update ? I don't understand how did he move from yi(θ... 3

Deadline missed, can I still earn a certificate?
If I missed the deadline for completing some homework/exercise, can I still earn the course certificate by completing them before May 12, 2020? Th... 2

error in video
I believe that there is an error in the video. It claims that after the first execution of the loop, if the new value for theta is used in the check for a mat... 3

How to choose an appropriate value? 1

Error on i-th example after the update 3

Perceptron algorithm with non-binary features
Hello. I have a question. Is it a requirement for the perceptron algorithm that every element of the feature vector be binary —either 0 or 1—? Woul... 2

Success of algorithm for linearly separable sets
Nice proof for those that are interested http://www.cs.columbia.edu/~mcollins/courses/6998-2012/notes/perc.converge.pdf Uses a definition of x t... 1

What about noise?
I'm new to machine learning, but it it true that Perceptron Algorithm might get stuck in a loop if any of the training data was wrong? If, for example, ... 2

? Perceptron Update 2
I am confused about this question. In the answer choice, I see something related to "change. However, what I have understood from the question i... 8

Question regarding the step for weights update 2

? Need clarification/help on when it is considered an error when the label and classifier meet the condition (video @ 2:28).
In the video **(2:28)** , it has ***[[y_i (theta*x_i + theta_0) <=0]]***. The lecture says that if the result is either ***0 or Negative, then it is conside... 2

glitch in course progress. Where is Question 18?

Learn About Verified Certificates