Intuitive Explanations Cluster Analysis Data Mining Algorithms +1

What is an intuitive explanation of DBSCAN?

1 Answer

Andy Wilson, Origami, visualization and large data.
Answered Jul 31, 2015

Take a large group of people and have them all stand in a field. You're going to use DBSCAN to identify separate crowds within the group.

You have to choose two numbers beforehand. First, how close do two people have to be to be "close"? Let's take something intuitive: if I can reach out and put my hand on top of someone's head, he's close by. About 3 feet. Second, how many people have to be close to you in order to form a crowd? Let's say three. Two is company, but three's a crowd.

Now here's how you run DBSCAN.

Have everyone count the number of people close to them. All the people with at least 3 nearby neighbors hold up green flags. They are the *core nodes*. They are in the middle of crowds. Remember, it's not enough to be able to touch each other's hands: to be close by, we have to be able to put hands on one another's heads.

Then we will have some people who are not yet holding up flags. They have fewer than 3 neighbors and one of those neighbors holds a green flag. Such people hold up yellow flags. They are the *edge* of a crowd.

Finally, we've got people who have fewer than 3 neighbors. None of those neighbors holds a green flag. These people hold up red flags. They are the *outliers*, not part of any crowd. (They're the radical individualists.)

Suppose that I am holding a green flag. I can identify all the other people in my crowd as follows: if I can make a chain of green-flag people (core points) with hands on one another's heads from myself to the target person, we are part of the same crowd. This chain cannot involve any yellow-flag people (edge points). To keep the different crowds straight, put numbers on the different green flags. Everyone with green flag #1 can reach everyone else with green flag #1 through a chain of hands on heads. Everyone with green flag #2 can reach everyone else holding green flag #2 but not anyone holding green flags #1 or #3. You get the idea.

There's more on Quora...

Pick new people and topics to follow and see the best answers on Quora.

Update Your Interests

Related Questions

What is an intuitive explanation of MapReduce?

How do I choose value of epsilon in DBSCAN?

What is an intuitive explanation of counterfactual regret minimization?

What is an intuitive explanation for PCA?

What is an intuitive explanation for the expectation maximization (EM) algorithm?

What is an intuitive explanation of the Master Theorem?

How do I select parameter values for "dbscan()" in r?

Is DBSCAN useful for visualization?

Is DBSCAN the best clustering algorithm?

What is an intuitive explanation of the Viterbi algorithm?

+ Ask New Question

More Related Questions

Question Stats

13 Public Followers

3,189 Views

Last Asked Oct 24, 2014

Edits

Or suppose that I hold a yellow flag. I get to choose which crowd to be part of. If someone with green flag C (#1, #2, #3 or whatever) can put her hand on my head, I can say I'm part of crowd C. If more than one person has his hand on my head then the guy running the game gets to choose whether I'm part of just one cluster or whether I hold onto my multiple memberships is up to the guy running the game.

You're done. Each distinct number that appears on a green flag (#1, #2, #3, ...) identifies a separate cluster. People holding red flags -- the *outlier nodes* -- are not part of any cluster. The *edge nodes* (people with yellow flags) can be considered part of whichever cluster is most convenient for you.

Does that help?