0

LDA vs. SVM for Dimensionality Reduction

svm dimensionality-reduction

Whats the difference between LDA and Linear SVM for dimensionality reduction. I am little confuse as LDA also looks for projection that separates the classes of data and SVM we also look for hyperplane ...

share improve this question

RiZ
1 • 1 • 1

asked Feb 1 '15 at 2:17

Can you spell out your acronyms? Are you referring to *linear discriminant analysis* & *support vector machines*? – **gung - Reinstate Monica** ♦ Feb 1 '15 at 2:30

Both are classification algorithms, not designed for dimensionality reduction. PCA is probably closer to what you are looking for. – **Marc Claesen** Feb 1 '15 at 10:13

This discussion contains some information on using LDA and CCA for dimensionality reduction (also see references in my answer's UPDATE). – **Aleksandr Blekh** Feb 4 '15 at 8:02

add a comment

2 Answers

order by votes

4

None of the algorithms are for dimensionality reduction.(but see last comment)

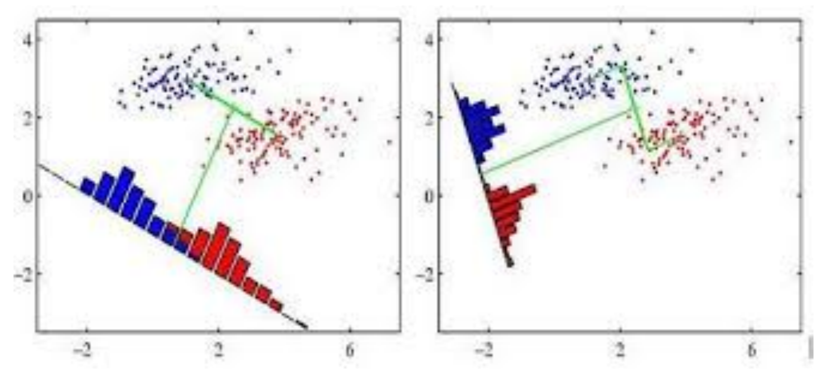
In terms of classification, LDA finds the hyperplane that "best separates" **all** the data points while linear SVM looks for the hyperplane that "best separates" **only** the points in the frontier between the two classes.

Thus LDA has to do with the mean of each of the two classes (since it deals with all the points and the mean is a "good summary" of all the points). If the two classes are basically spherical then the LDA hyperplane is perpendicular to the line linking the two means of the points.

What is confusing is that most of the explanations of LDA talks about "**projection** of the data that best separates the classes". What is confusing is that the projection is onto a line and not a hyperplane - the LDA projection line (or LDA direction) is the line orthogonal to the separating hyperplane.

Take the image below (I did not create the image - found through Google but the original site no longer exists. I will remove it if it violates any copyright).

The LDA is the pane in the right. The green line is the LDA **direction** . The LDA hyperplane is the line perpendicular to the green that separates the two regions (further confusing is that in a 2D data the separating hyperplane is also a line!!)



The figure in sklearn example of LDA shows the separating hyperplane of the LDA (also notice that since the data are not spherically distributed in each class the LDA separating hyperplane is **not** perpendicular to the line joining the two centers/means)

Finally, LDA can be seen as an **extreme** dimensionality reduction (contrary to what I wrote first). It reduces the data to **one** dimension - if you have to remember only one dimension of the data and you still want to classify it, the best thing you can do is to project the data to the LDA direction!!!

share improve this answer

Jacques Wainer
3,925 • 1 • 16 • 29

answered Feb 4 '15 at 7:11

0

Although mostly used for classification, LDA can be used for dimensionality reduction too: "Data reduction entails a sequence of unit-variance, linear discriminant variables $\beta_k^T x$, chosen to successively maximize $\beta_k^T \Sigma_{\mathrm{Bet}} \beta_k$, with Σ_{Bet} the between-class covariance matrix. These discriminant variables represent a subspace for which the class centroids are spread out as much as possible." (from Hastie T, Buja A, Tibshirani R. Penalized discriminant analysis. The Annals of Statistics. 1995 Feb 1:73-102).

Regarding SVMs, a corresponding dimensionality-reduction technique would be Kernel-PCA.

share improve this answer

Luca Citi
799 • 2 • 6

answered May 31 '17 at 9:07

Your Answer

Body

Add picture

Log in

OR

Name

Email