The Relationship Between PCA plots, LDA plots, and actual Class Separability

Bruce R. Linnell, PhD

2005

Once at an e-nose conference, I heard several presentations which stated that since the different classes they were trying to identify overlapped in their two-dimensional PCA plots, there was no separation between the classes. This is an erroneous conclusion!

The best way to accurately determine whether the classes are well separated is to use some kind of *classifier* (neural network, Linear, K-nearest-neighbor, etc.) with some kind of *error estimator* (Leave-One-Out, N-fold-cross-validation, Bootstrap, etc.) in the original feature space (using no transforms such as PCA or LDA), to calculate the estimated probability of success (Ps). If the sample ratio (or "SR", the ratio of the number of examples per class to the number of features per example) is small, use *feature selection* to remove unneeded features.

Figure 1 shows the relationship between the actual separation in feature space (as measured by the Linear classifier and the Holdout estimator) versus the apparent separation in a 2-D PCA projection (measured in the same manner), for real data, using several full-rank, high-SR University of California at Irvine (UCI) datasets, as well as all of the e-nose datasets that I have access to. The farther a dataset is from the diagonal line, the worse the correlation between the apparent separation on a PCA plot and the true separation. This shows that **even when data appears highly overlapped in a 2-D PCA plot (near the bottom), there is still a good chance that it might be much more separable (on the right side) in the original feature space**.

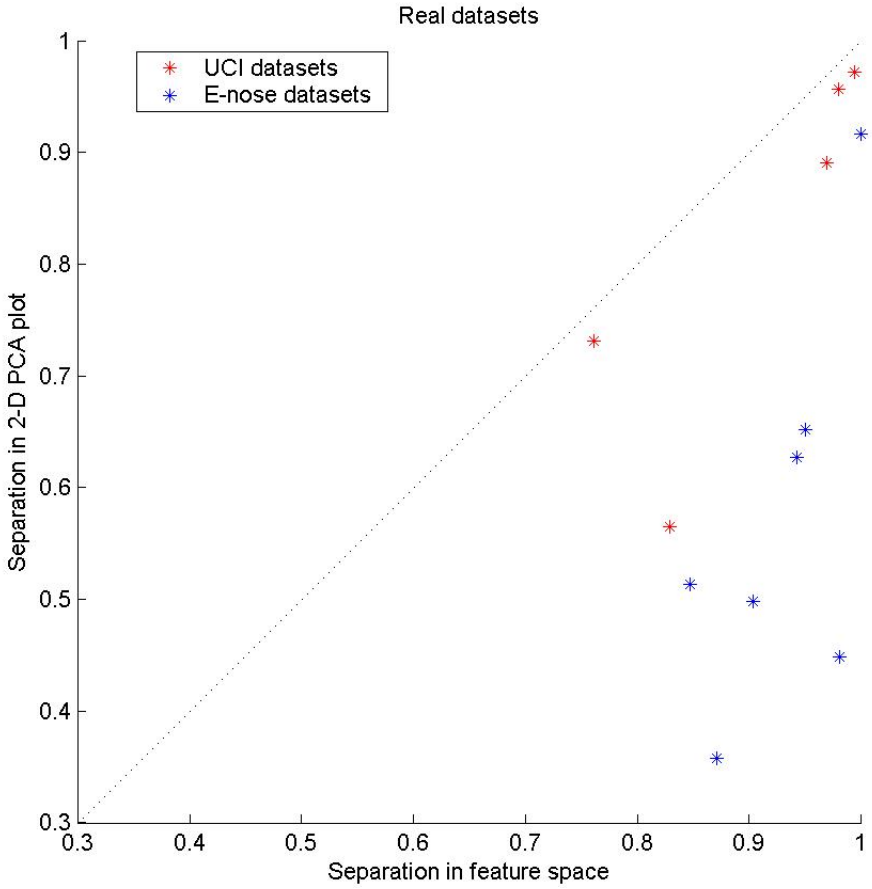

Figure 1

For example, the PCA plot of one of the e-nose datasets is shown in Figure 2, where the data appears highly overlapped, but the linear separation in the original feature space is actually 98%.

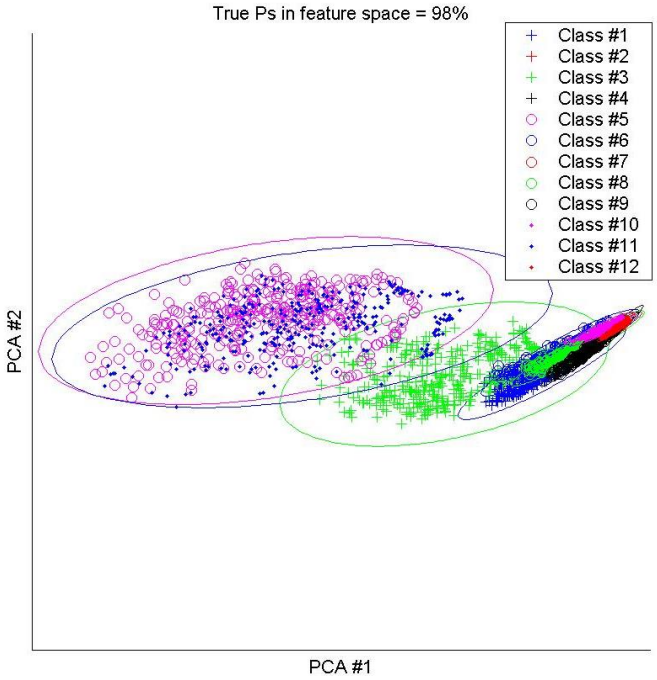

Figure 2

These results imply that the dimensions which contain the largest variance (which is what PCA finds) may not contain the best separability information.

The apparent separation between classes in a Linear Discriminant Analysis (LDA) plot also has no correlation to the classification success rate of the original data. LDA plots can both over- and under-exaggerate the true class separation, and over-exaggeration is very common with small-SR data, as shown in Figure 3. The four classes in the left LDA plot appear to be highly overlapped, yet the actual classification success rate is 94%. Conversely, the classes in the right plot appear to be well separated, while the actual success rate is only 65%.

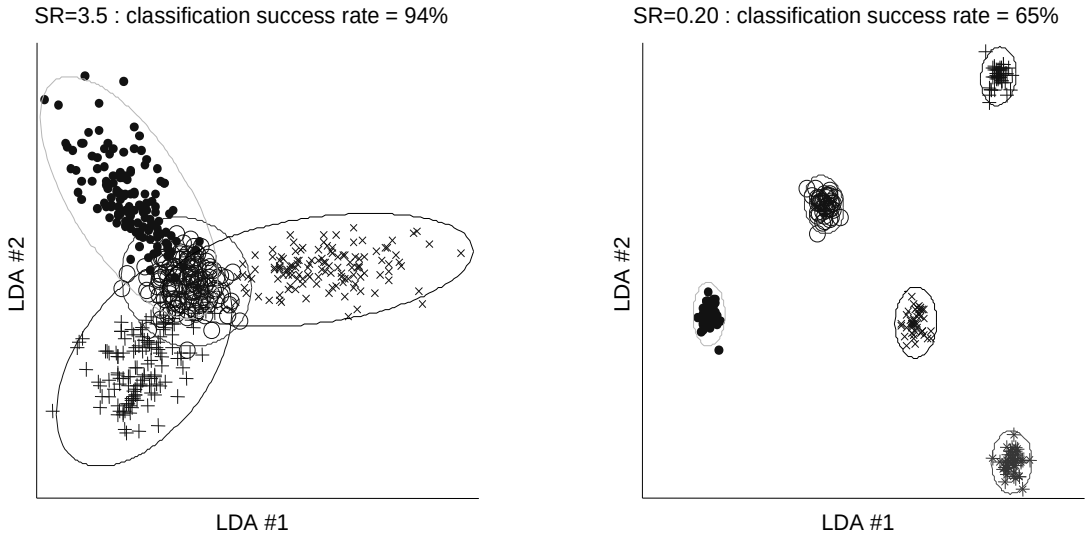

Figure 3

The basic reason for these results is that **both PCA and LDA projections lose information**, and when reduced to only 2 dimensions to be graphed, lots of the information that provides separability may be lost.

I feel that this is an important point because in the presentations mentioned, when the e-nose data did not appear separable in the PCA plots, the researchers chose to use different technologies (such as mass spectroscopy) to identify their classes. Considering that the e-nose is still a relatively new and struggling technology, reports of poor performance due to using the wrong tools for the job or misunderstanding the tools that are being used can only harm the potential for future growth of this field.