

Homework #3: Classifiers and Decision Boundaries

Capabilities of models are determined by the shape of their classification boundaries, allowed by the type of the model and its parameters, and by the learning algorithm's ability to fit it. The purpose of this task is to explore decision boundaries of some models that we discussed.

Part 1

Play with the Paint Data widget to draw some datasets with a binary class variable, and use the Polynomial classification to see the shape of decision boundary that you can get with different degrees of polynomial expansion. (Enable "Show contours" for a more informative picture.)

Provide snapshots of Polynomial classification that demonstrate some nice examples of data that:

- is linearly separable,
- is not separable linearly, but can be separated by 2nd degree polynomials,
- requires 3rd degree polynomials,
- requires 4th degree (be creative!)

Use a separate instance of the Paint Data widget for each of those, because you will need this same data sets in the second part.

Part 2

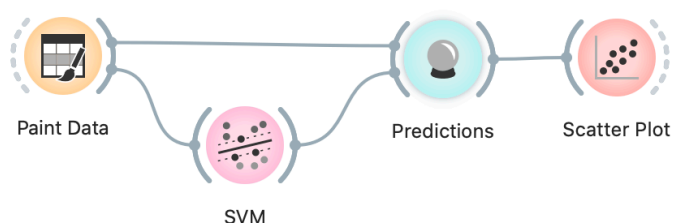
Check whether support vector machines can indeed do the same without explicit polynomial expansion, and add some trees for a good measure. So, replace the Polynomial Classification with SVM, and try

1. Linear SVM
2. SVM with 2nd degree polynomial kernels (make sure you set c to 1)
3. SVM with 3rd degree polynomial kernels (set c to 1)
4. SVM with RBF kernels

See if they match the performance of polynomial classification with the corresponding polynomial degree. RBF kernels are more general and should do handle all data sets.

5. Try classification trees of different depths (in the Tree widget, change the parameter "Limit the maximal tree depth to") and see what they can and can't handle.

SVMs and trees cannot show decision boundaries in the same way as Polynomial classification, so to check whether the model works or not, use Predictions and Scatter Plot.



To produce the report, press Ctrl-C (or Cmd-C) in the widget to copy the image (and not the entire widget!) to the clipboard.

Submit your homework as a short report in PDF (include a title of the homework, your name and email) and submit to it bzupan@gmail.com with subject "DM-HW3" (no spaces in the subject title!). The deadline for this homework is **January 22nd**.