Machine Learning from Data

Lecture 10: Spring 2021

Today’s Lecture

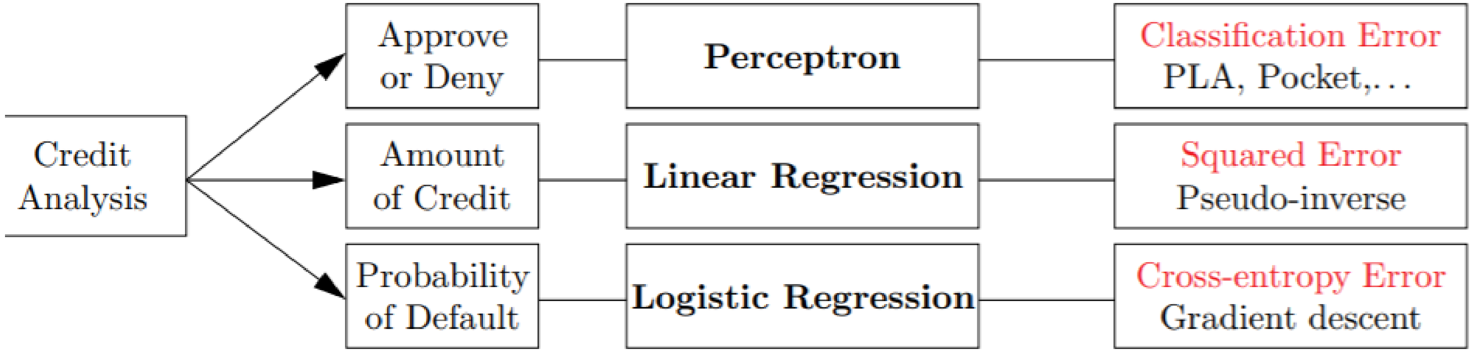
- Non-Linear Transforms
  - Z-Space
  - Polynomial Transforms

Linear Model (Recap)

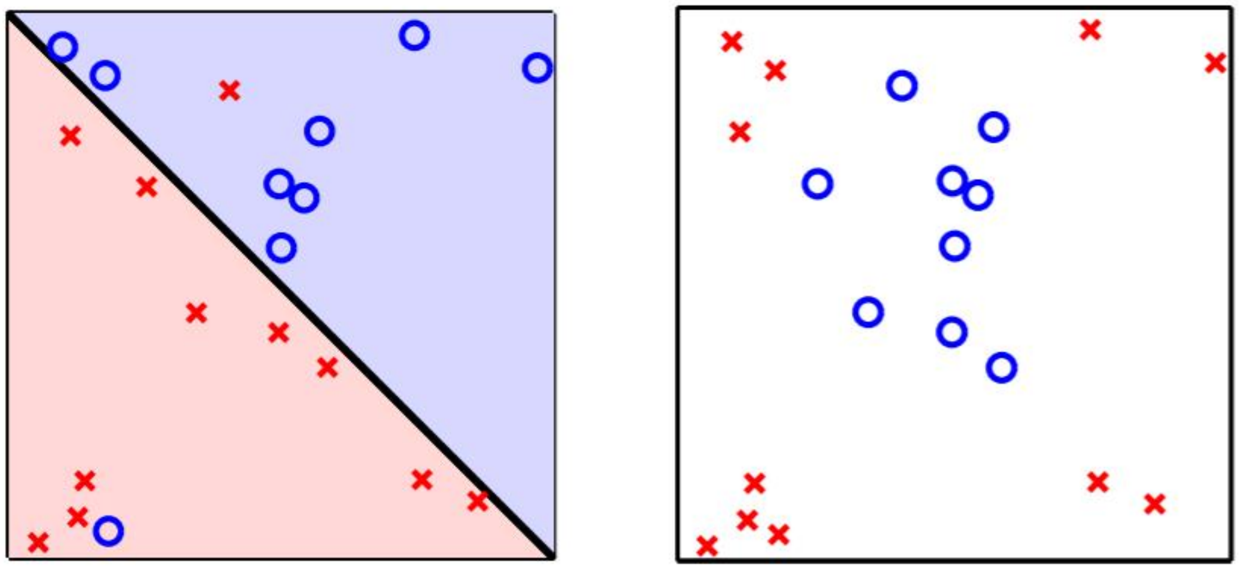
linear in $\mathbf{x}$: gives the line/hyperplane separator

$$s = \mathbf{w}^{\mathrm{T}}\mathbf{x}$$

linear in $\mathbf{w}$: makes the algorithms work



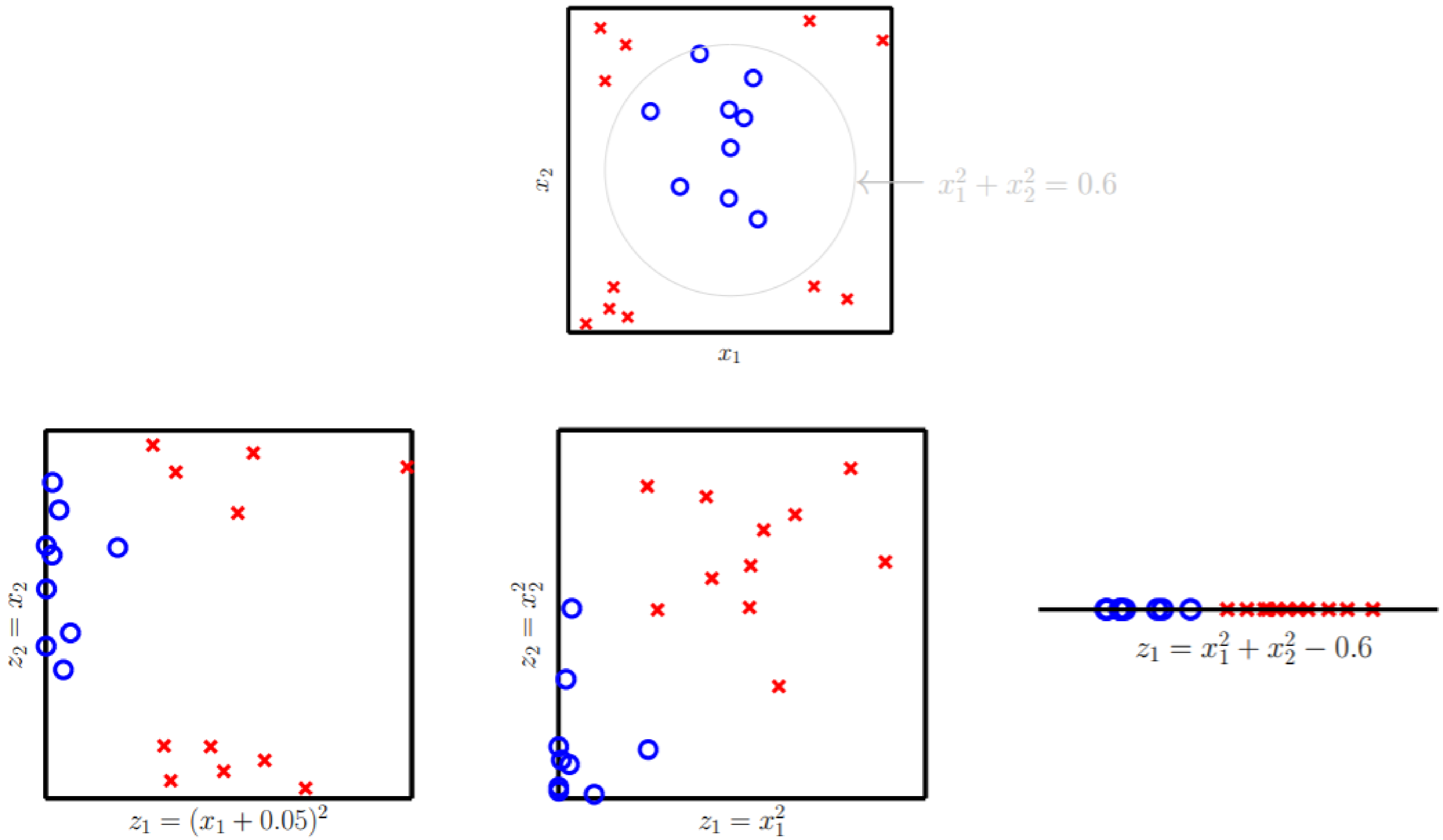
Limits of the Linear Model



(a) Linear with outliers   (b) Essentially nonlinear

To address (b) we need something more than linear.

Many Non-Linear Transforms May Work

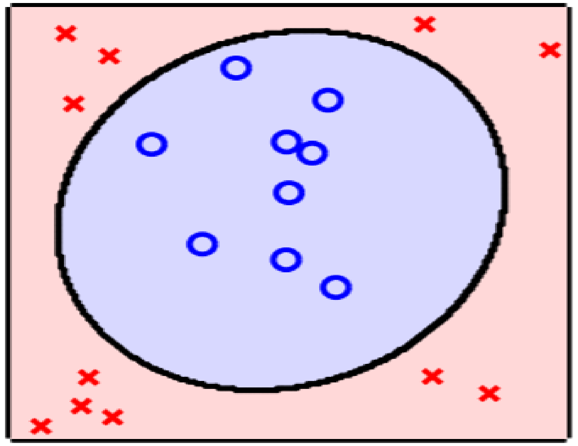


Choose Transform before Looking at the data

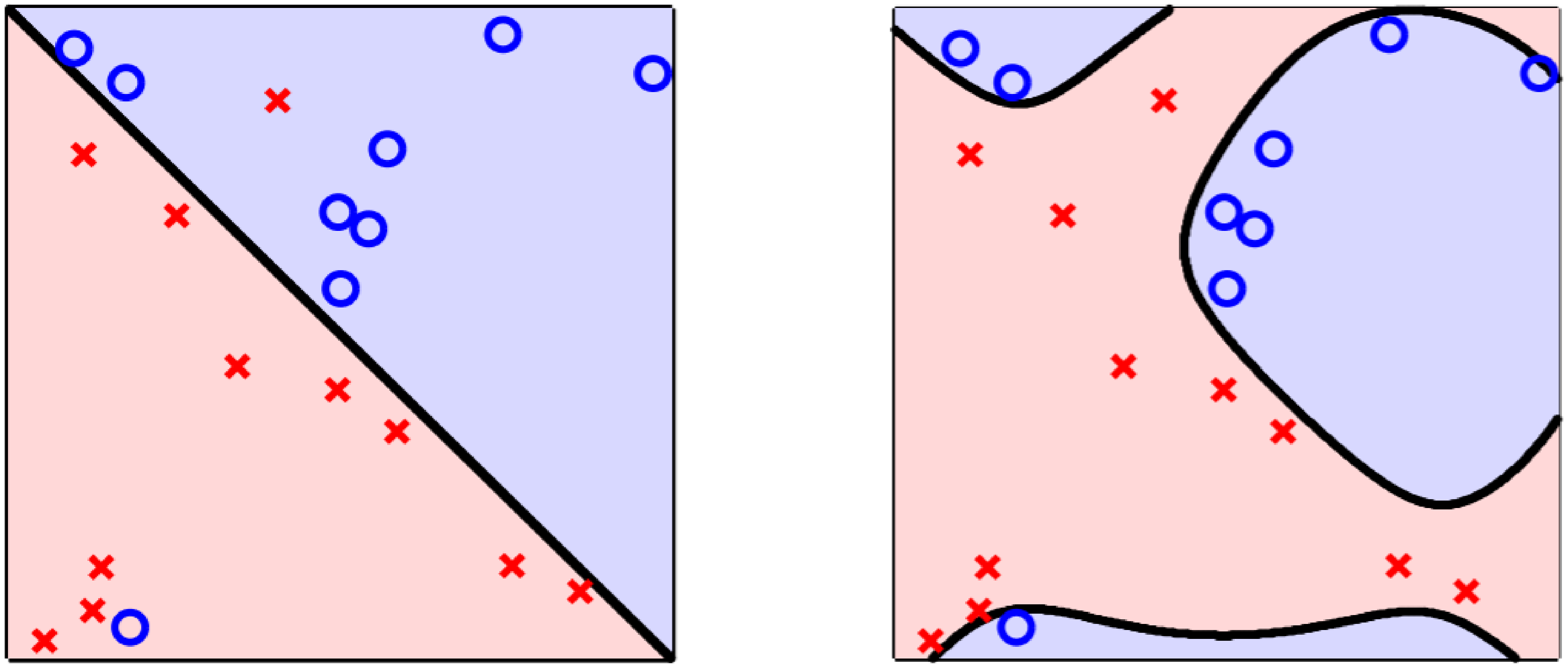
After constructing features carefully, **before** seeing the data . . .

. . . if you think linear is not enough, try **the 2nd order polynomial transform**.

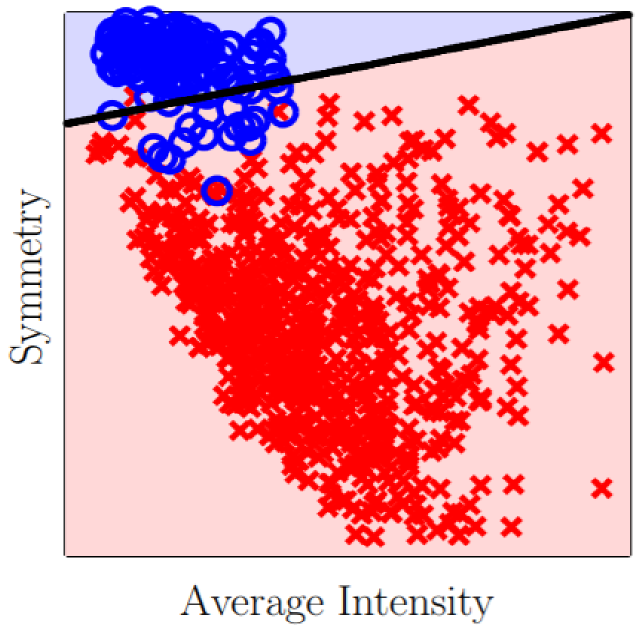
$$\begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} = \mathbf{x} \longrightarrow \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} 1 \\ \Phi_1(\mathbf{x}) \\ \Phi_2(\mathbf{x}) \\ \Phi_3(\mathbf{x}) \\ \Phi_4(\mathbf{x}) \\ \Phi_5(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$



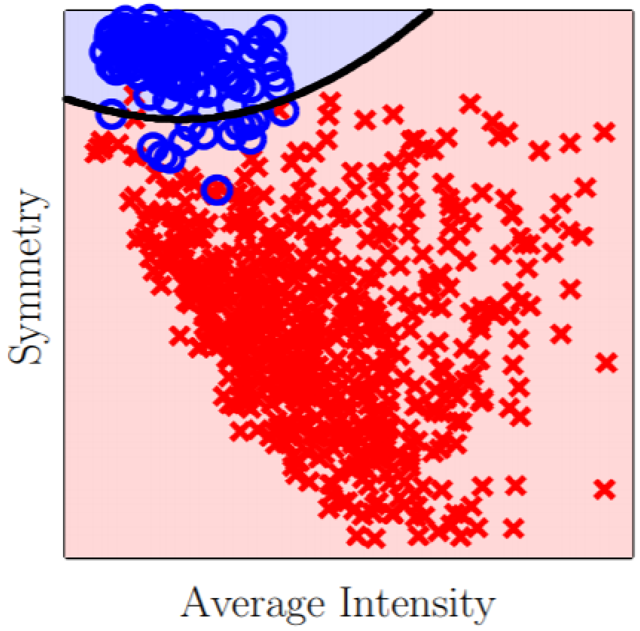
Be Careful with Feature Transforms



Digits Data Again



Linear model
$E_{\text{in}} = 2.13\%$
$E_{\text{out}} = 2.38\%$



3rd order polynomial model
$E_{\text{in}} = 1.75\%$
$E_{\text{out}} = 1.87\%$

Use the Linear Model!

- First try a linear model – simple, robust and works.

- Algorithms can tolerate error plus you have nonlinear feature transforms.

- Choose a feature transform *before* seeing the data. Stay simple.

  Data snooping is hazardous to your E_{out}.

- Linear models are fundamental in their own right; they are also the building blocks of many more complex models like support vector machines.

- Nonlinear transforms also apply to regression and logistic regression.

Thanks!