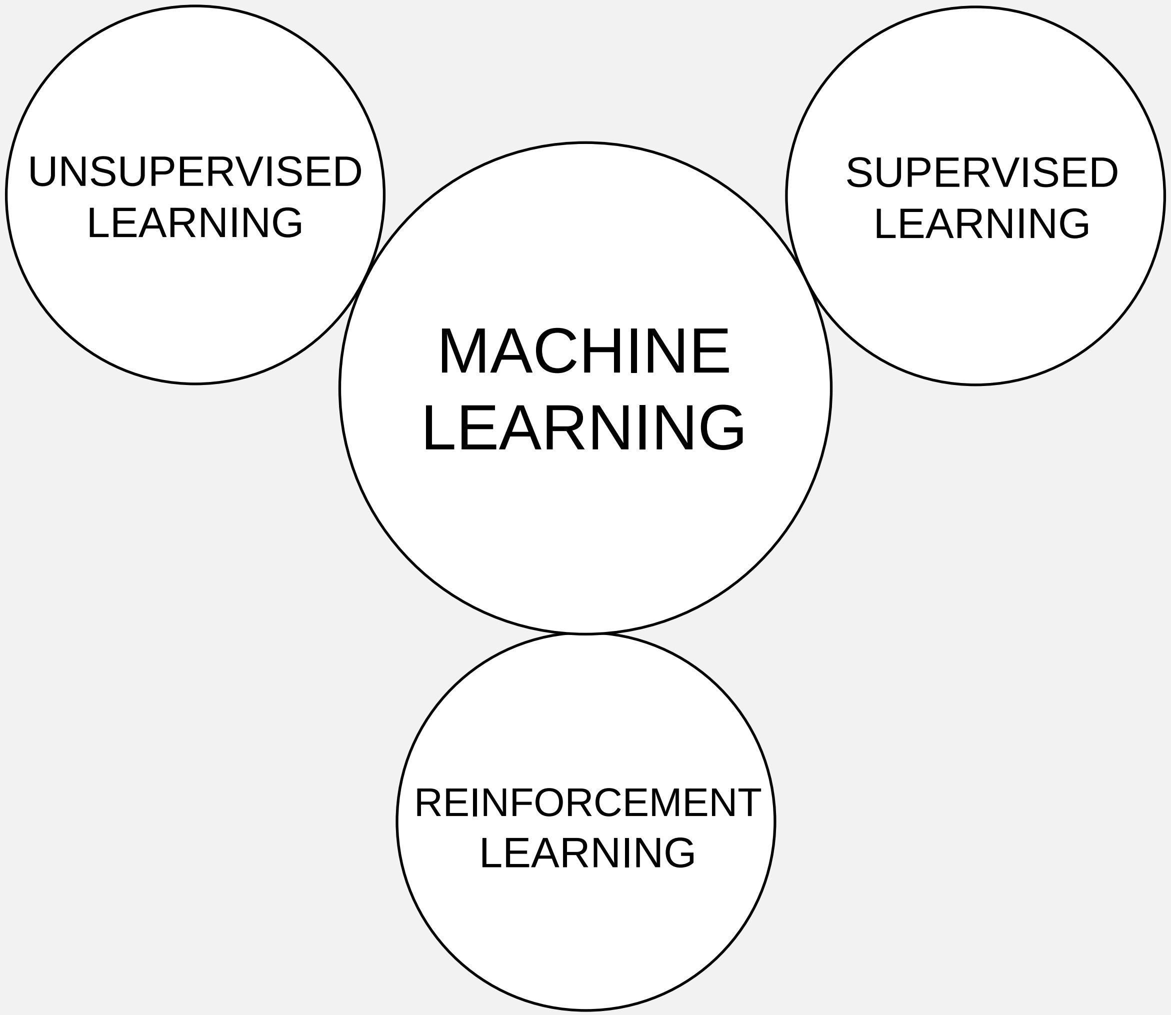
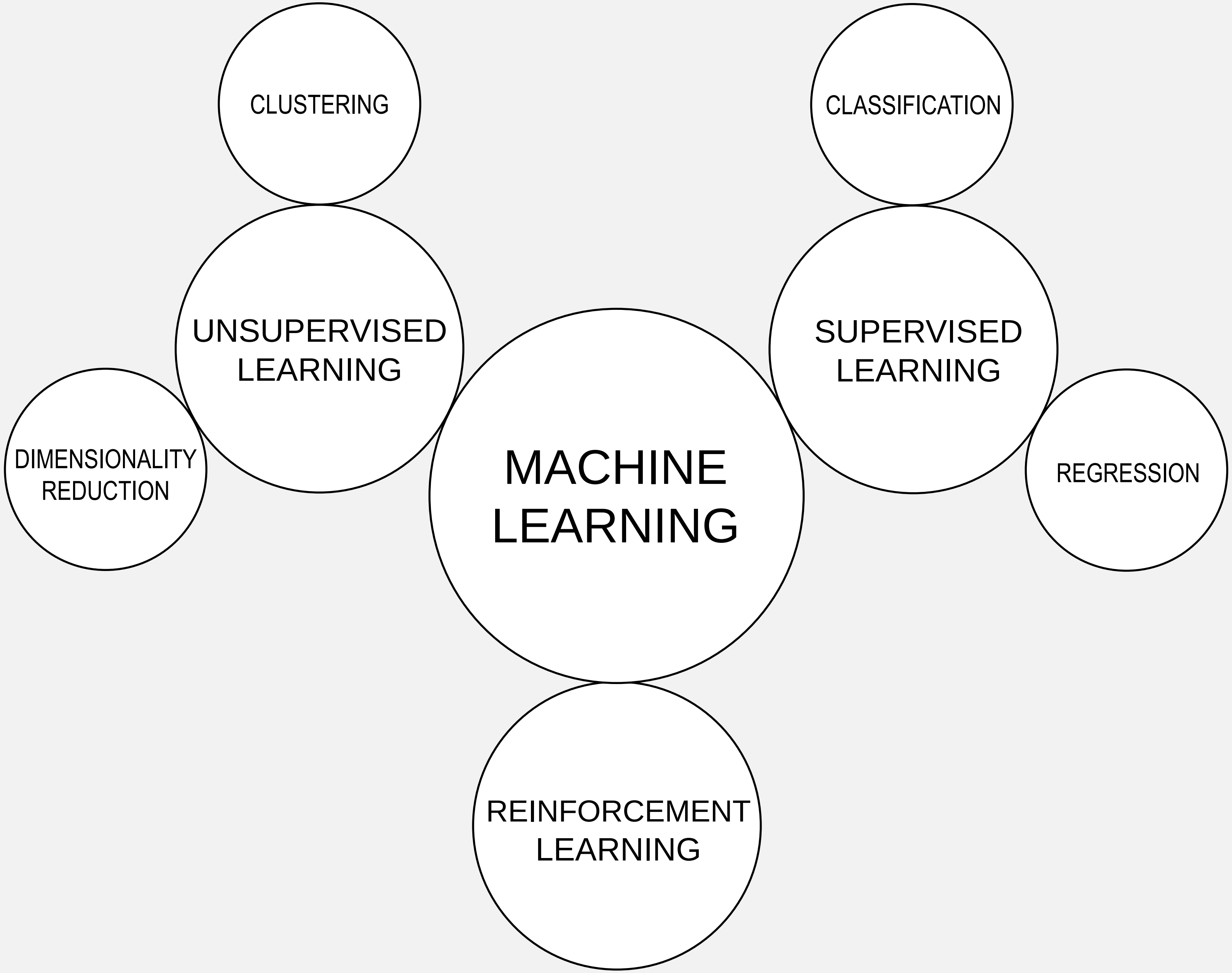
REGRESSION

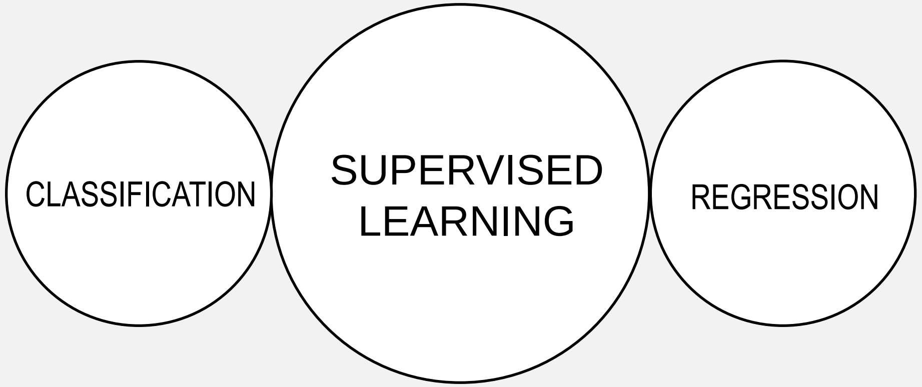
Lecture 8

MAL1, 2024

MACHINE LEARNING


UNSUPERVISED LEARNING
SUPERVISED LEARNING
MACHINE LEARNING
REINFORCEMENT LEARNING


CLUSTERING
CLASSIFICATION
UNSUPERVISED LEARNING
SUPERVISED LEARNING
DIMENSIONALITY REDUCTION
MACHINE LEARNING
REGRESSION
REINFORCEMENT LEARNING

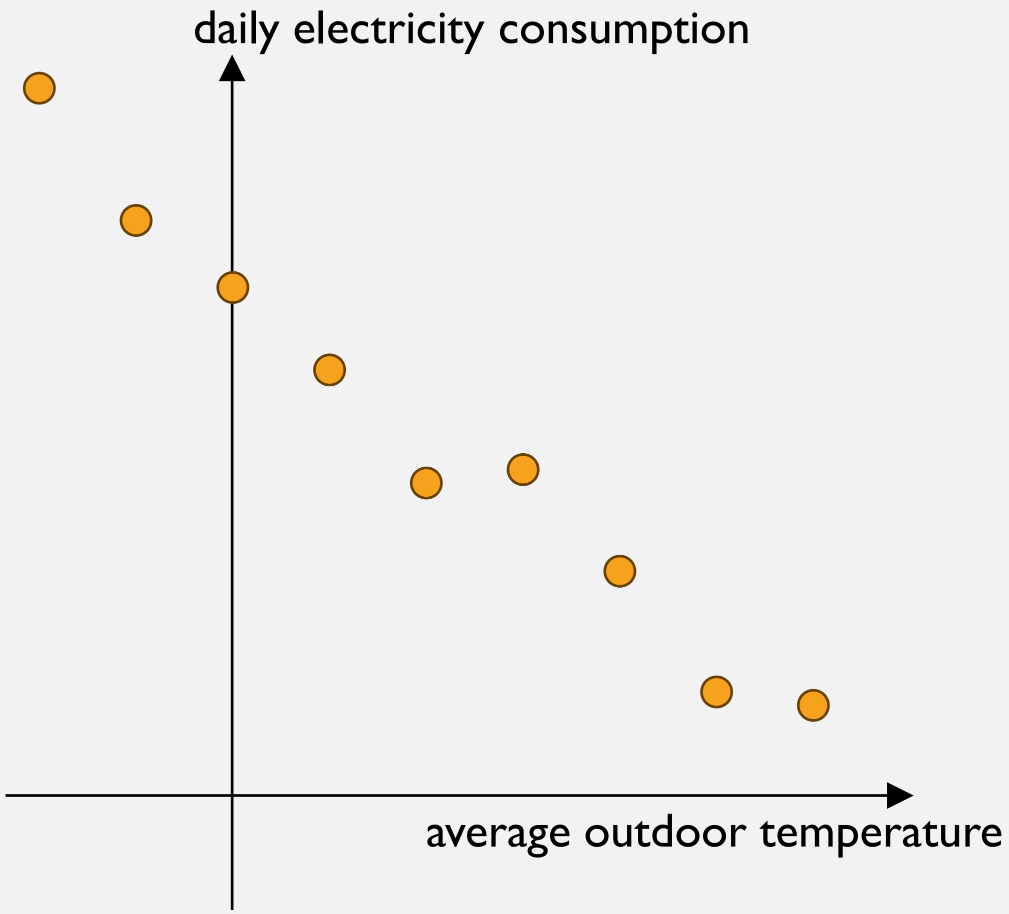

CLASSIFICATION
SUPERVISED LEARNING
REGRESSION

REGRESSION

- Linear regression
- Performance metrics
- Polynomial regression
- Regularization

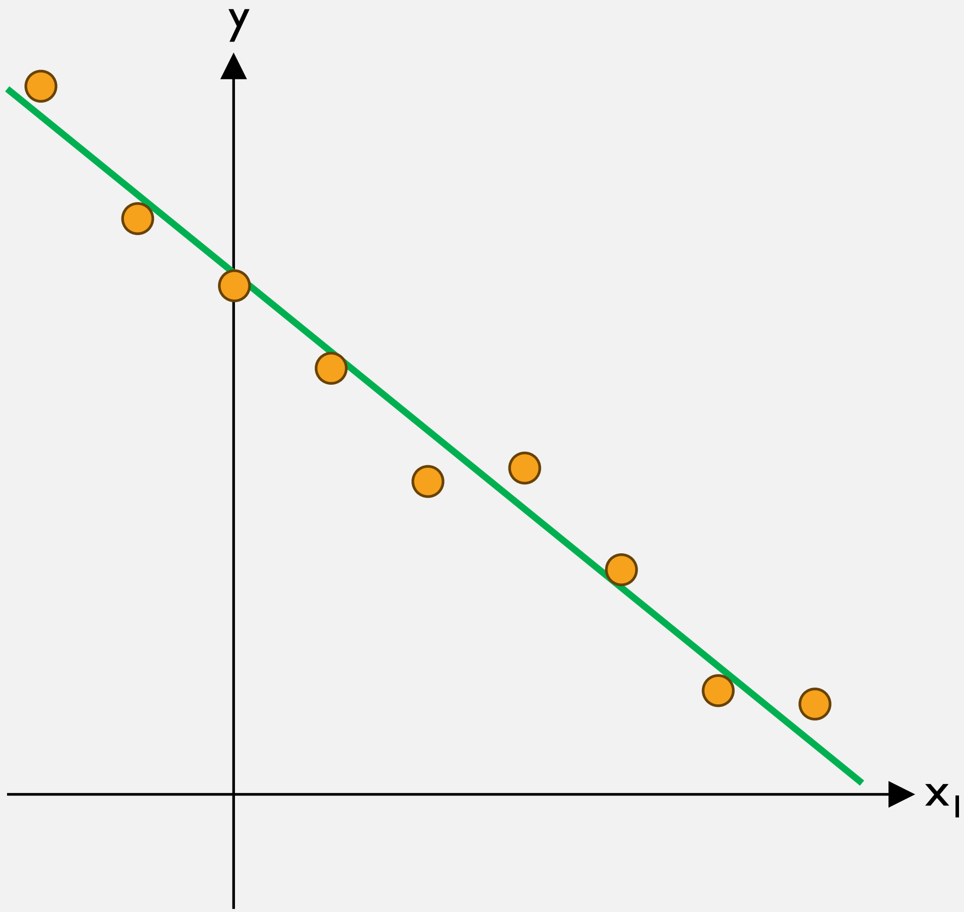
REGRESSION

average outdoor temperature (°C)	daily electricity consumption (kWh)
-10	46.5
-5	37.9
0	33.2
5	27.5
10	20.3
15	21.1
20	14.2
25	6.3
30	5.6



REGRESSION

x_1	y
$\theta_0 + -10 \times \theta_1 \approx$	46.5
$\theta_0 + -5 \times \theta_1 \approx$	37.9
$\theta_0 + 0 \times \theta_1 \approx$	33.2
$\theta_0 + 5 \times \theta_1 \approx$	27.5
$\theta_0 + 10 \times \theta_1 \approx$	20.3
$\theta_0 + 15 \times \theta_1 \approx$	21.1
$\theta_0 + 20 \times \theta_1 \approx$	14.2
$\theta_0 + 25 \times \theta_1 \approx$	6.3
$\theta_0 + 30 \times \theta_1 \approx$	5.6



FINDING THETA'S

FINDING THETA'S

$$\mathrm{SSE} = \sum_i \left(y^{(i)} - \hat{y}^{(i)}\right)^2 = \sum_i \left(y^{(i)} - \theta^T x^{(i)}\right)^2$$

minimize → take the derivative wrt. all θ's and set equal to zero:

$$\frac{\partial}{\partial \theta_j} \mathrm{SSE} = \frac{\partial}{\partial \theta_j} \sum_i \left(y^{(i)} - \theta^T x^{(i)}\right)^2 = 2 \sum_i \left(y^{(i)} - \theta^T x^{(i)}\right) x_j^{(i)} = 0$$

summarize in matrix form:

$$2\boldsymbol{X}^T(\boldsymbol{X\theta} - \boldsymbol{y}) = 0$$

FINDING THETA'S

$$\boldsymbol{\theta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & -10 \\ 1 & -5 \\ 1 & 0 \\ 1 & 5 \\ 1 & 10 \\ 1 & 15 \\ 1 & 20 \\ 1 & 25 \\ 1 & 30 \end{bmatrix} \quad \boldsymbol{y} = \begin{bmatrix} 46.5 \\ 37.9 \\ 33.2 \\ 27.5 \\ 20.3 \\ 21.1 \\ 14.2 \\ 6.3 \\ 5.6 \end{bmatrix}$$

$$\boldsymbol{X}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -10 & -5 & 0 & 5 & 10 & 15 & 20 & 25 & 30 \end{bmatrix}$$

$$\boldsymbol{X}^T\boldsymbol{X} = \begin{bmatrix} 9 & 90 \\ 90 & 2400 \end{bmatrix}$$

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \begin{bmatrix} \frac{8}{45} & -\frac{1}{150} \\ -\frac{1}{150} & \frac{1}{1500} \end{bmatrix}$$

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T = \begin{bmatrix} \frac{11}{45} & \frac{19}{90} & \frac{8}{45} & \frac{13}{90} & \frac{1}{9} & \frac{7}{90} & \frac{2}{45} & \frac{1}{90} & -\frac{1}{45} \\ -\frac{1}{75} & -\frac{1}{100} & -\frac{1}{150} & -\frac{1}{300} & 0 & \frac{1}{300} & \frac{1}{150} & \frac{1}{100} & \frac{1}{75} \end{bmatrix}$$

$$\boldsymbol{\theta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \begin{bmatrix} 33.72 \\ -1.009 \end{bmatrix}$$

x_1	y
-10	46.5
-5	37.9
0	33.2
5	27.5
10	20.3
15	21.1
20	14.2
25	6.3
30	5.6

in practice, the pseudoinverse is computed using SVD

REGRESSION

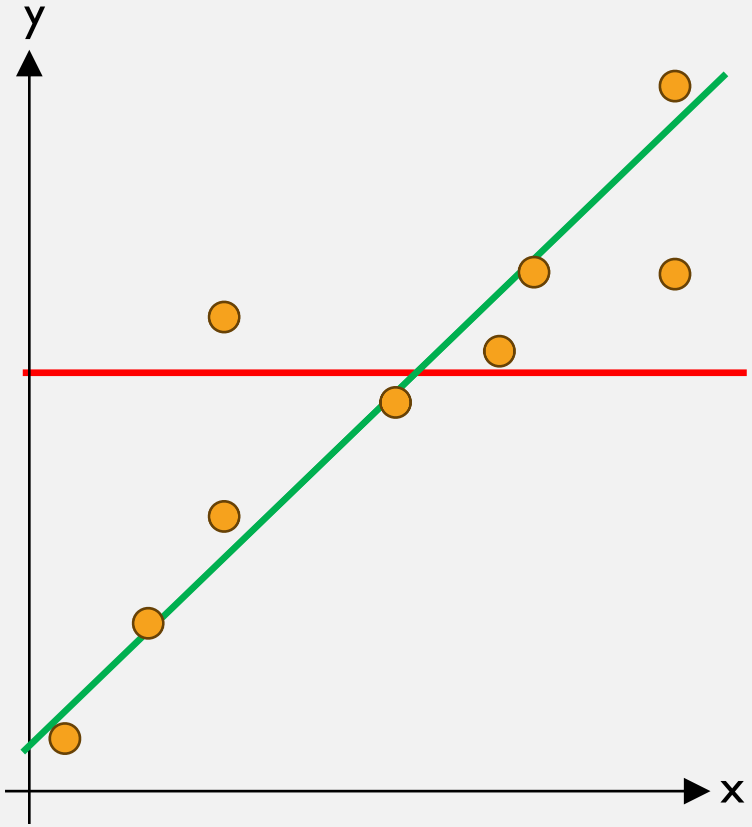
- Linear regression
- Performance metrics
- Polynomial regression
- Regularization

SSE AND FRIENDS

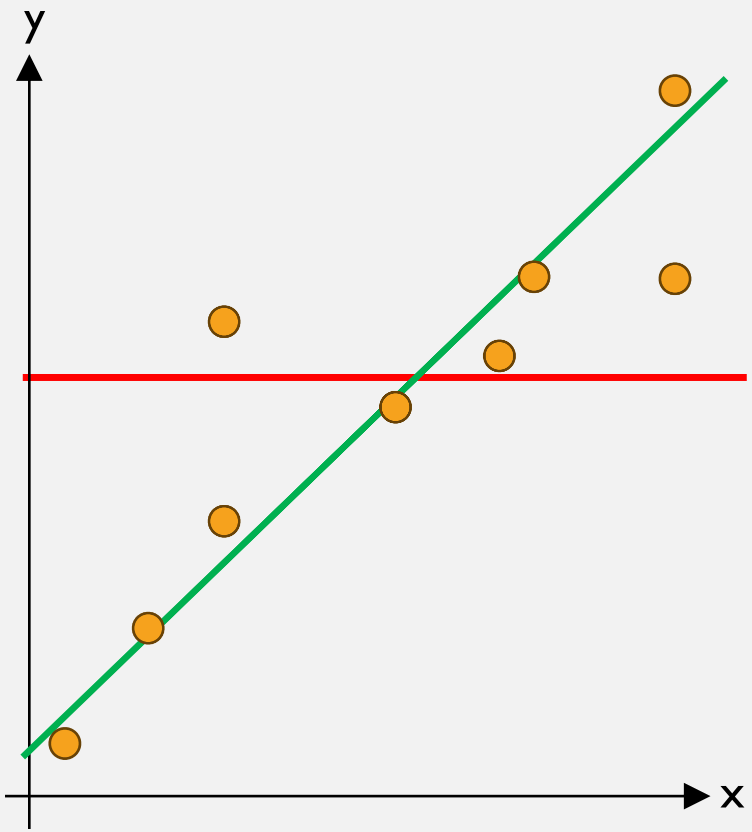
- The smaller the SSE, the better the model …

… so what should we use as our performance metric?

SSE AND FRIENDS



SSE AND FRIENDS



$$SST = \sum_i \left(y^{(i)} - \bar{y}\right)^2$$ total deviation from mean

$$SSE = \sum_i \left(y^{(i)} - \hat{y}^{(i)}\right)^2$$ explained part

$$SSR = \sum_i \left(\hat{y}^{(i)} - \bar{y}\right)^2$$ unexplained part

CODE EXAMPLE

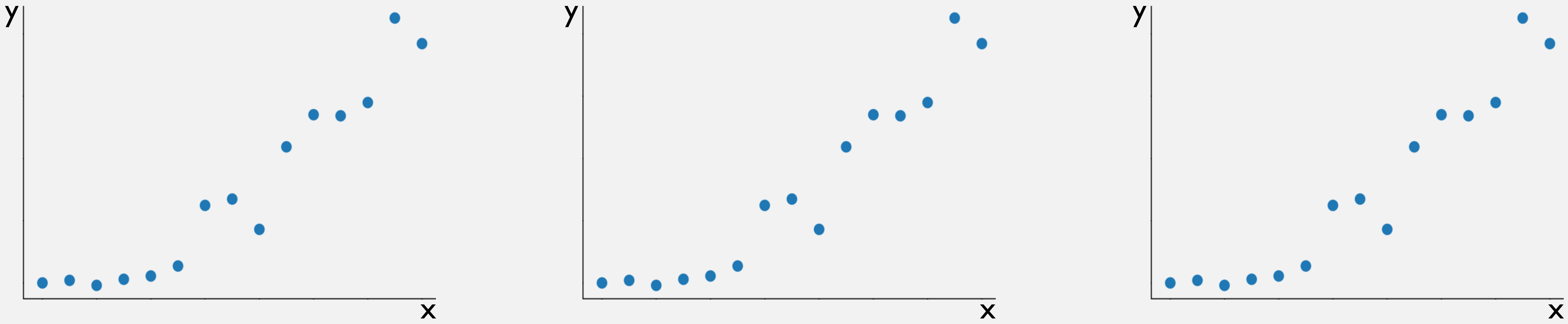
Jupyter Notebook **Regression - Hitters**

REGRESSION

- Linear regression
- Performance metrics
- Polynomial regression
- Regularization

POLYNOMIAL REGRESSION

UNDERFITTING AND OVERFITTING



REGRESSION

- Linear regression
- Performance metrics
- Polynomial regression
- Regularization

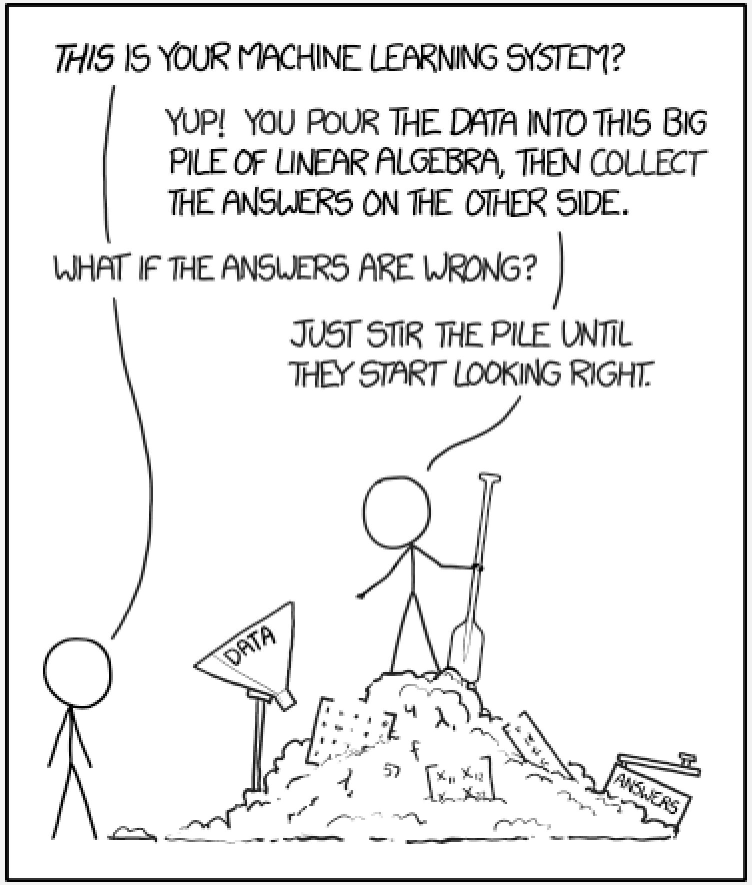
REGULARIZATION

THE OPTIMAL REGULARIZATION PARAMETER

RIDGE VS LASSO REGRESSION

CODE EXAMPLE

Jupyter Notebook **Regression - Hitters**


THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS