



Stacked Ensembles

• Engineer original feature set

- Train L models using strong learners (GLM, RF, GBM, ...)

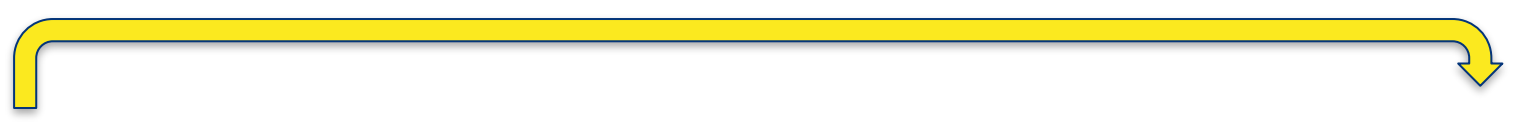
- Control overfitting using k -fold
CV

- Stack these predicted values
to form new feature set

- Train metalearner on stacked predictions

$$\begin{bmatrix} y \end{bmatrix} \approx \begin{bmatrix} \hat{y}_l \end{bmatrix} = f_l \left(\begin{bmatrix} x_1 \dots x_p \end{bmatrix} \right)$$

$$\begin{bmatrix} y \end{bmatrix} \approx g\left(\begin{bmatrix} \hat{y}_1 \dots \hat{y}_L \end{bmatrix}\right)$$

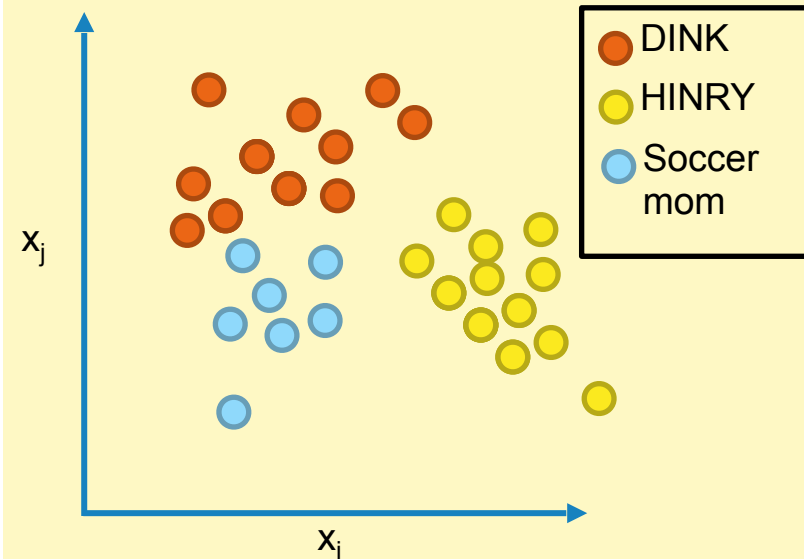




Unsupervised Learning

Clustering:

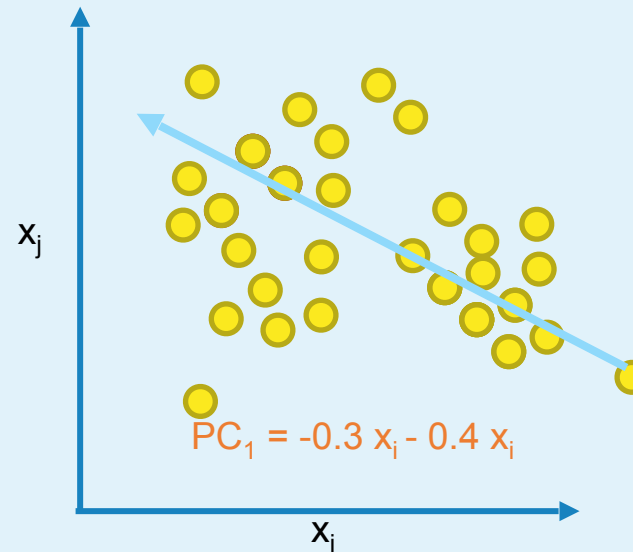
Grouping rows – e.g. creating groups of similar customers



H₂O algos:
k – means

Feature extraction:

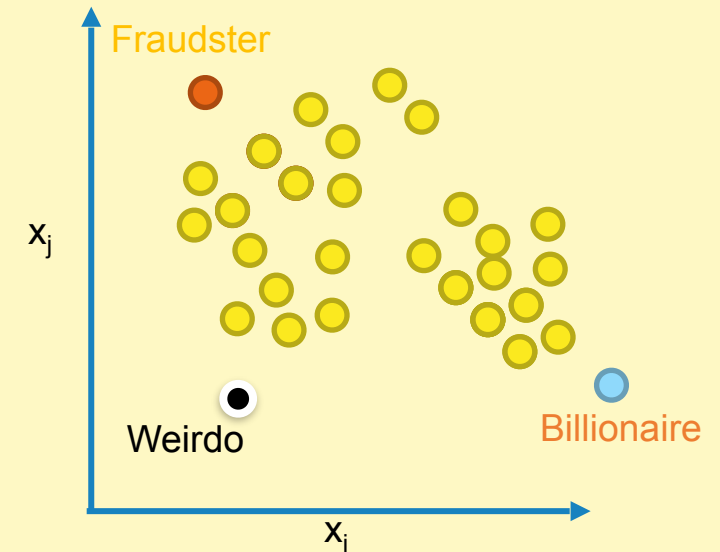
Grouping columns – Create a small number of new representative dimensions



H₂O algos:
Principal components
Generalized low rank models
Autoencoders
Word2Vec

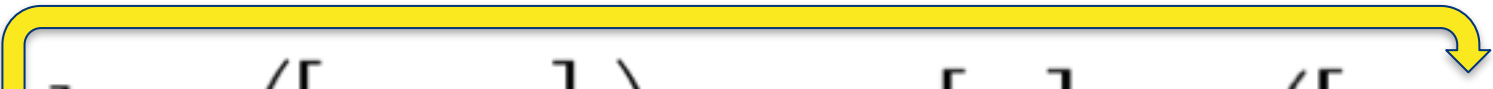
Anomaly detection:

Detecting outlying rows - Finding high-value, fraudulent, or weird customers



H₂O algos:
Principal components
Generalized low rank models
Autoencoders

Stacked Ensembles


$$\begin{bmatrix} y \end{bmatrix} \approx \begin{bmatrix} \hat{y}_l \end{bmatrix} = f_l \left(\begin{bmatrix} x_1 \dots x_p \end{bmatrix} \right) \qquad \begin{bmatrix} y \end{bmatrix} \approx g \left(\begin{bmatrix} \hat{y}_1 \dots \hat{y}_L \end{bmatrix} \right)$$

- Engineer original feature set
- Train L models using strong learners (GLM, RF, GBM, ...)
- Control overfitting using k-fold CV
- Stack these predicted values to form new feature set
- Train metalearner on stacked predictions