Brain Break

10.55

CSE/STAT 416

Course Wrap Up

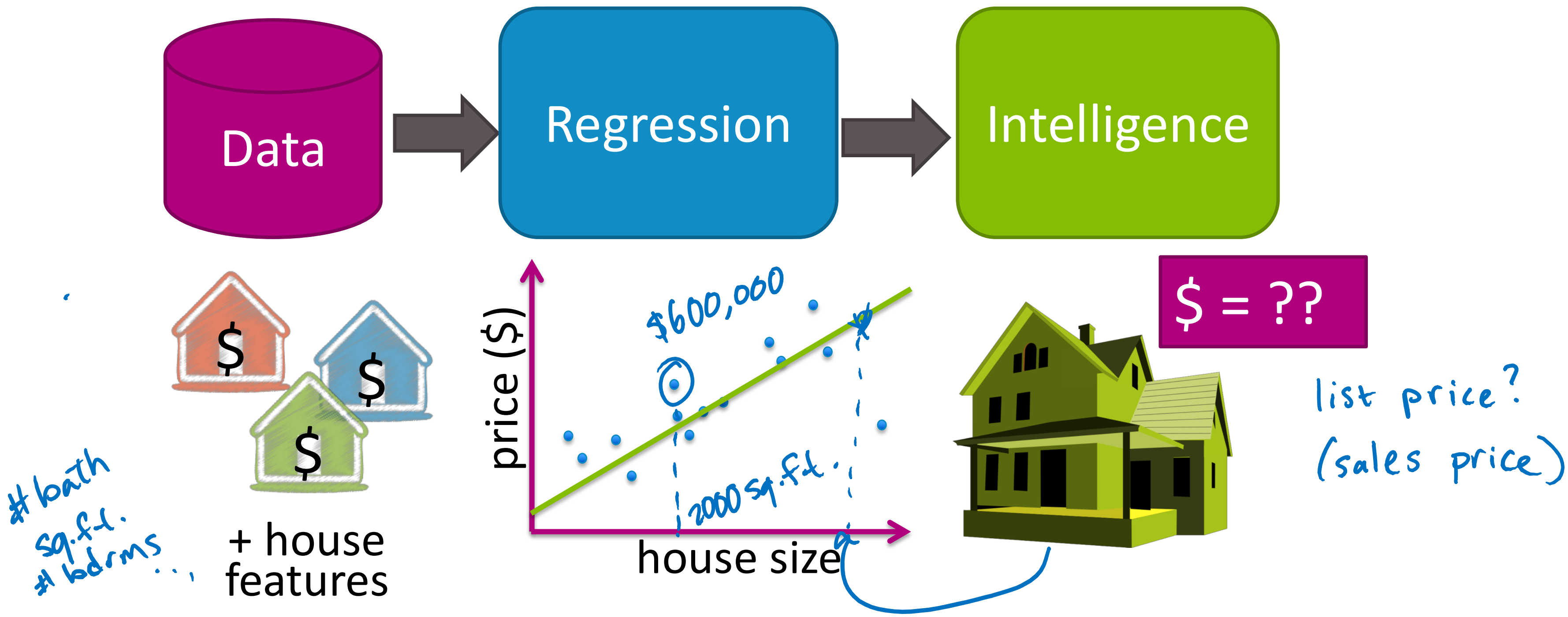
Hunter Schafer
University of Washington
Aug 19, 2019

Slides borrowed from Emily Fox

One Slide

- Regression
- Overfitting
- Training, test, and generalization error
- Bias-Variance tradeoff
- Ridge, LASSO
- Cross validation
- Gradient descent
- Classification
- Logistic regression
- Decision trees
- Boosting
- Precision and recall
- Nearest-neighbor retrieval, regression, and classification
- Kernel regression
- Locality sensitive hashing
- Dimensionality reduction, PCA
- k-means clustering
- Hierarchical clustering
- Unsupervised v. supervised learning
- Recommender systems
- Matrix factorization
- Coordinate descent
- Neural networks
- Convolutional neural networks
- Transfer learning for deep learning

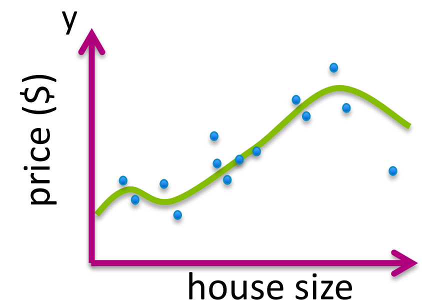
Case Study 1:
Predicting house prices



Regression

$y_i = f(x_i) + \varepsilon_i$ $\hat{y}_i = \hat{f}(x_i)$

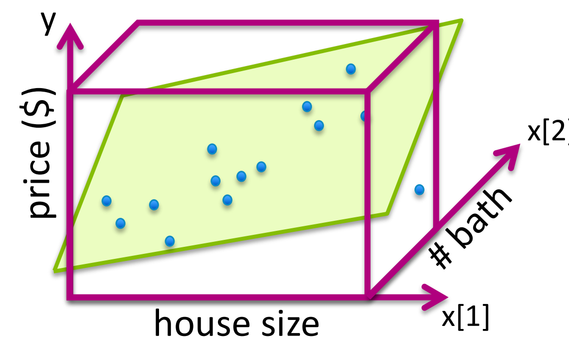
Case study: Predicting house prices



Models

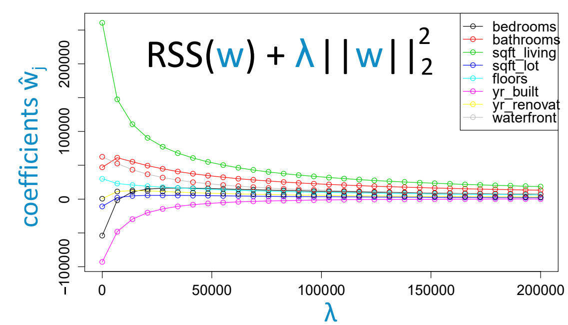
- Linear regression
- Regularization:
  Ridge (L2), Lasso (L1)

LASSO => Sparse



Including many features:

- Square feet
- # bathrooms
- # bedrooms
- Lot size
- Year built
- …



Regression

Case study: Predicting house prices

$RSS(w) = \sum_{i=1}^{n} (w^T h(x_i) - y_i)^2$

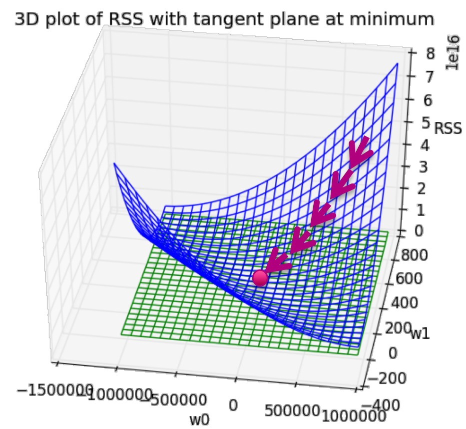
Algorithms

- Gradient descent

$RSS(w_0,w_1) =$
$(\$_{house\ 1}-[w_0+w_1 sq.ft._{house\ 1}])^2$
$+ (\$_{house\ 2}-[w_0+w_1 sq.ft._{house\ 2}])^2 +$
$(\$_{house\ 3}-[w_0+w_1 sq.ft._{house\ 3}])^2 + \ldots$
[include all houses]

$\downarrow$

$\hat{w}$

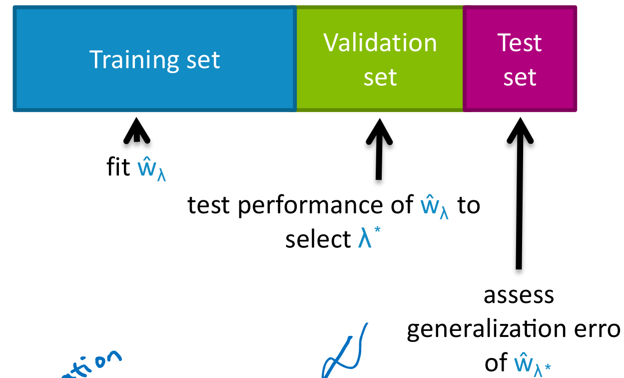


Regression

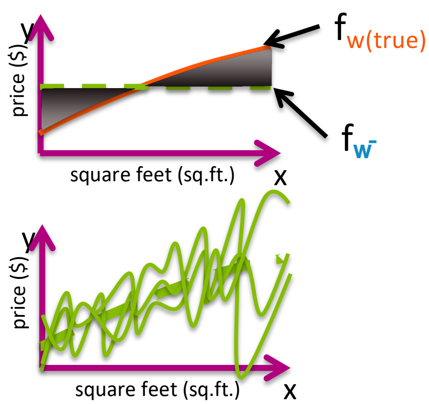
Case study: Predicting house prices

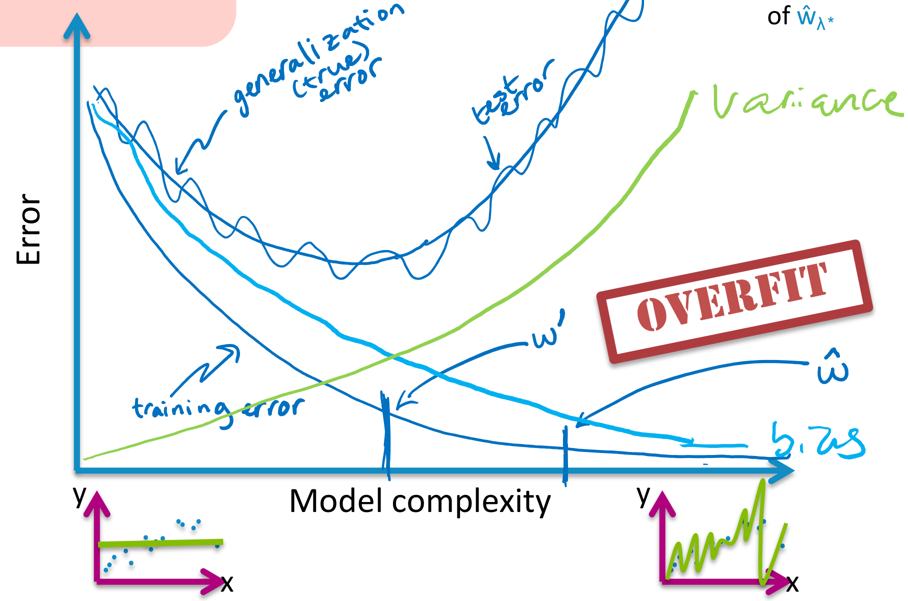
Concepts

- Loss functions, bias-variance tradeoff, cross-validation, sparsity, overfitting, model selection

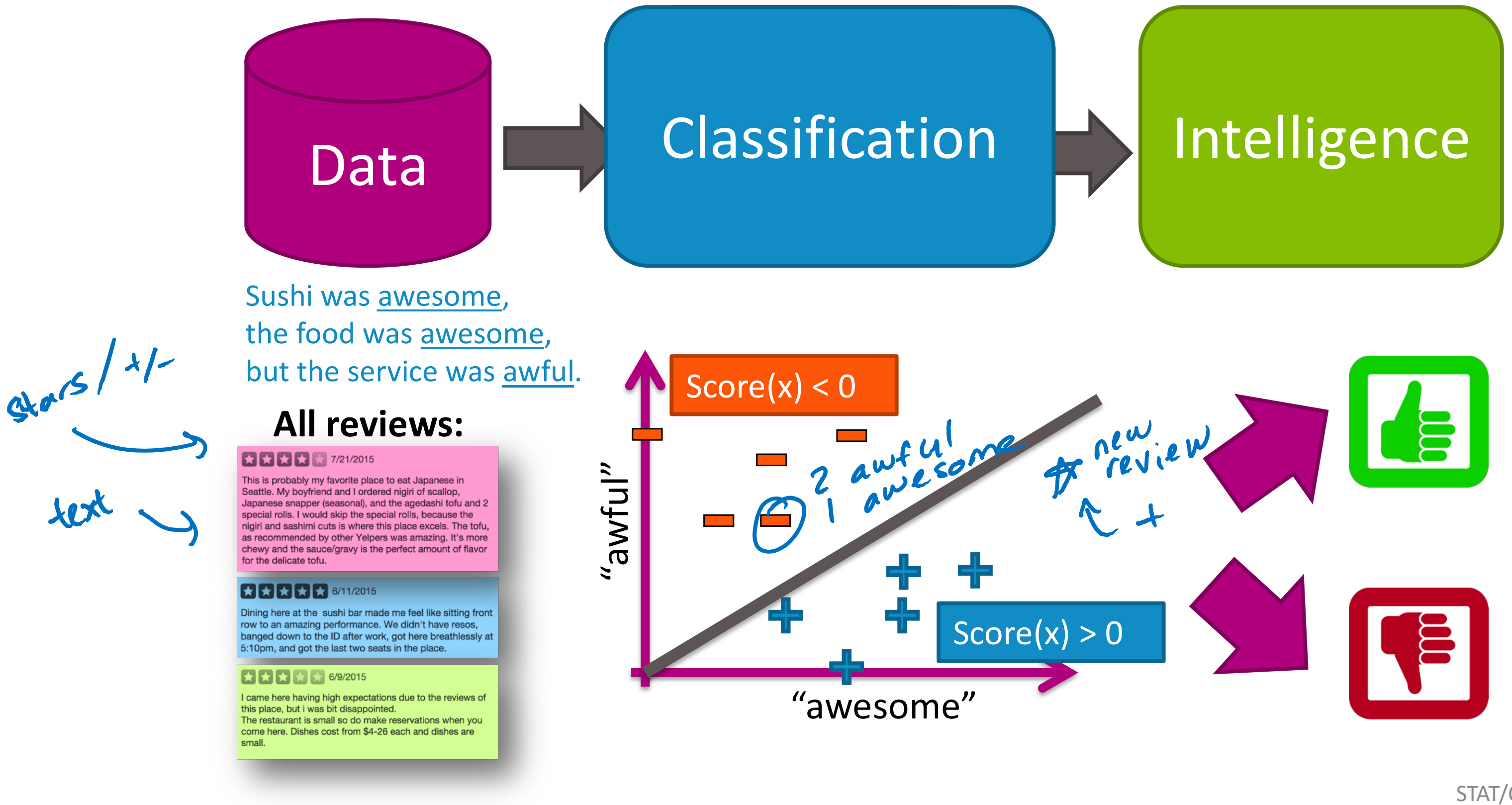


1. Noise
2. Bias
3. Variance



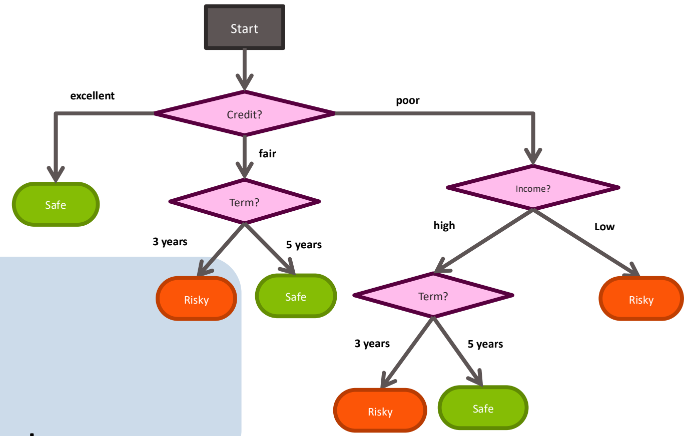


Case Study 2:
Sentiment analysis



Classification

Case study: Analyzing sentiment



Models

- Linear classifiers (logistic regression)
- Multiclass classifiers
- Decision trees
- Boosted decision trees and random forests

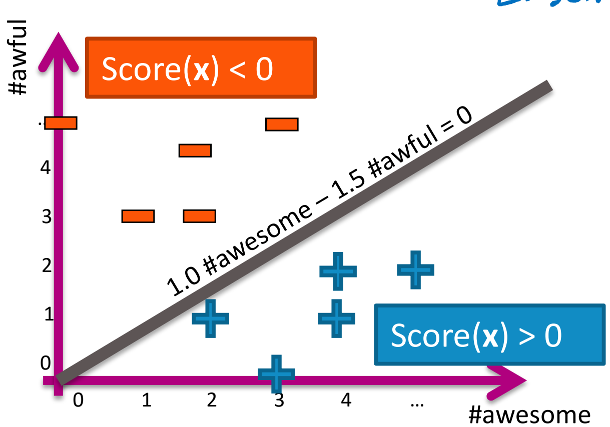
Ensemble Method

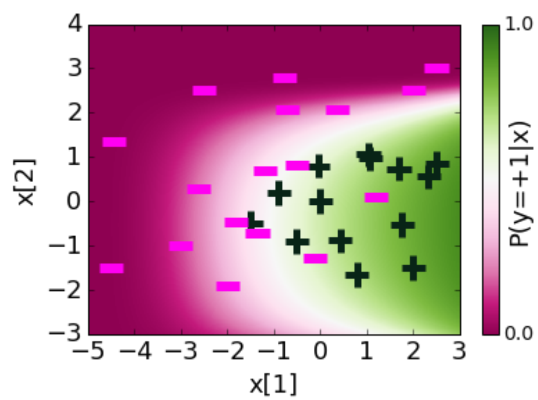
AdaBoost

$$\hat{f}(x) = \text{sign}\left(\sum_{t=1}^{T} \hat{w}_t \hat{f}_t(x)\right)$$

model weight

$\alpha_i \leftarrow$ dataset weight



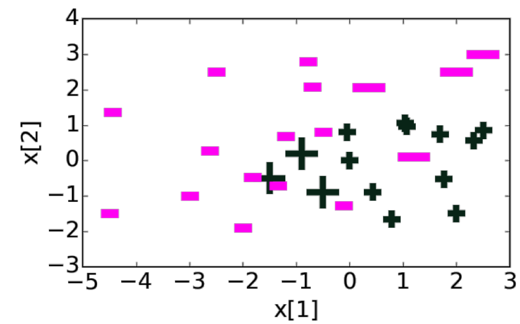


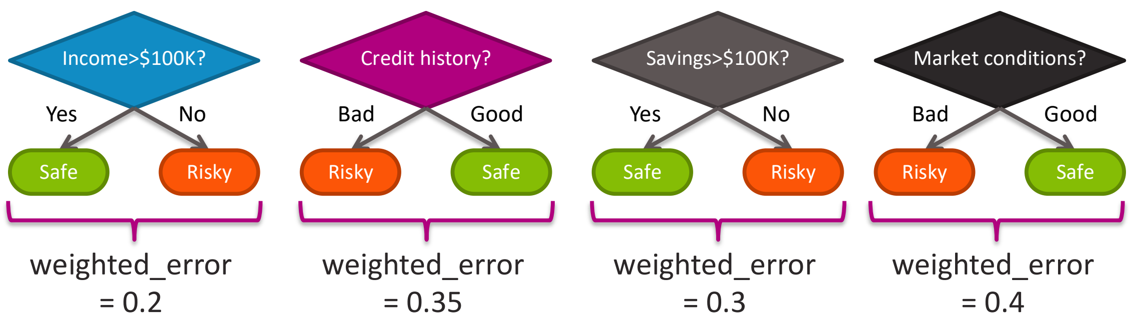
Classification

Case study: Analyzing sentiment

Algorithms

- Boosting
- Learning from weighted data





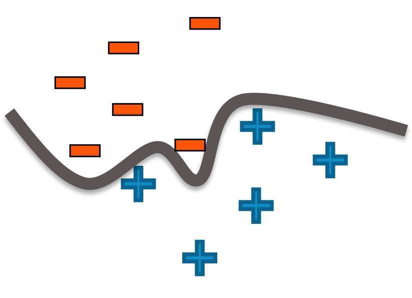
Classification

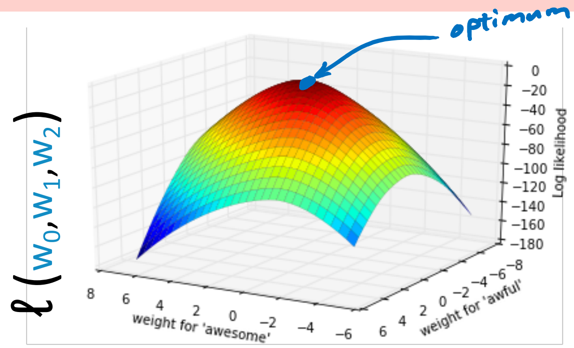
Case study: Analyzing sentiment

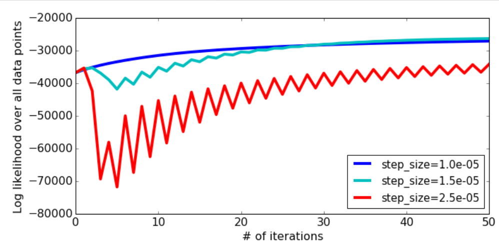
Accuracy
vs class imbalance

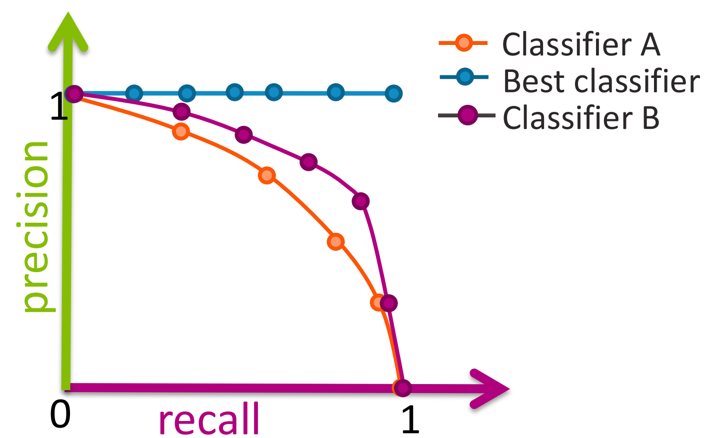
Concepts

- Decision boundaries, maximum likelihood estimation, ensemble methods, random forests
- Precision and recall









Case Study 3:
Document retrieval

Case Study 3+:
Document structuring for retrieval

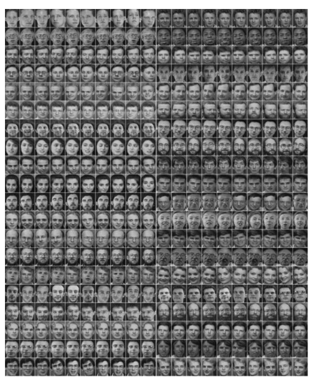
Bag of Words
TF-IDF

Euclidean
Manhattan
Cosine

Case Study 3++:
Dimensionality reduction

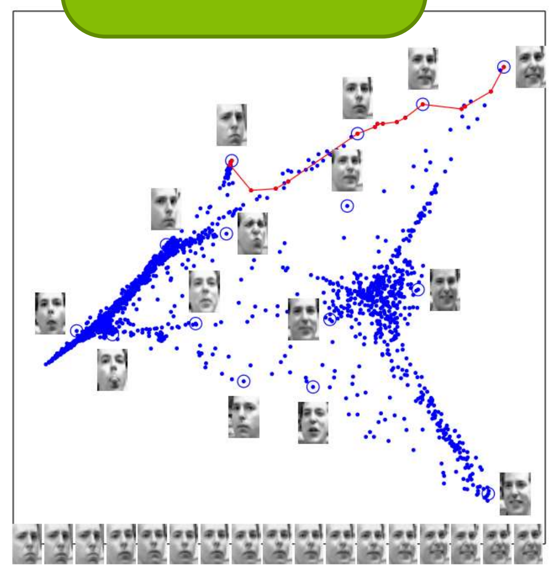
Data → PCA → Intelligence

simple



Images with thousands or millions of pixels

Can we give each image a coordinate, such that similar images are near each other?



[Saul & Roweis '03]

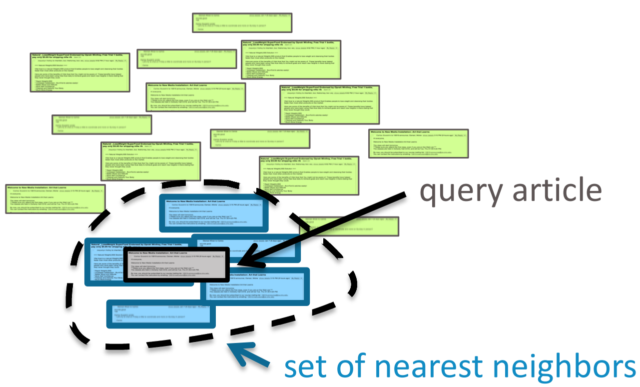
Clustering & Retrieval

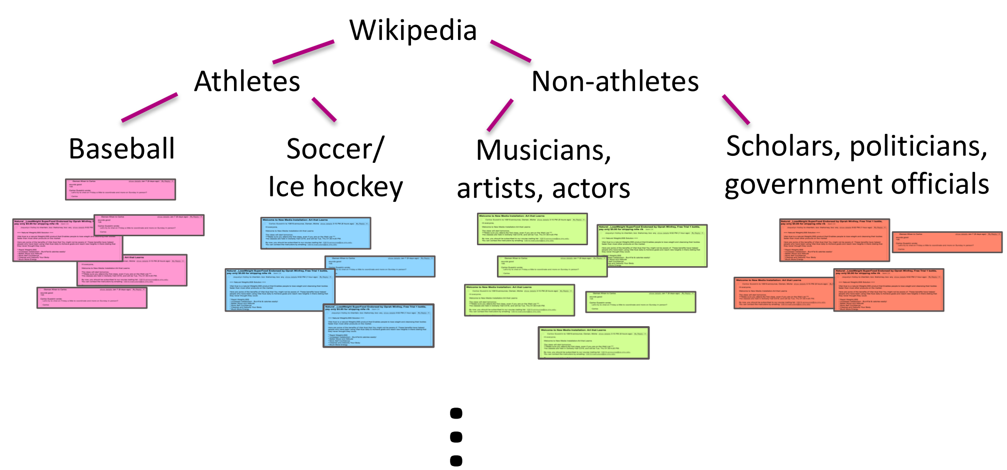
Case study: Finding documents

Models

- Nearest neighbors / Kernel methods
- Clustering
- Hierarchical clustering

Agglomerative



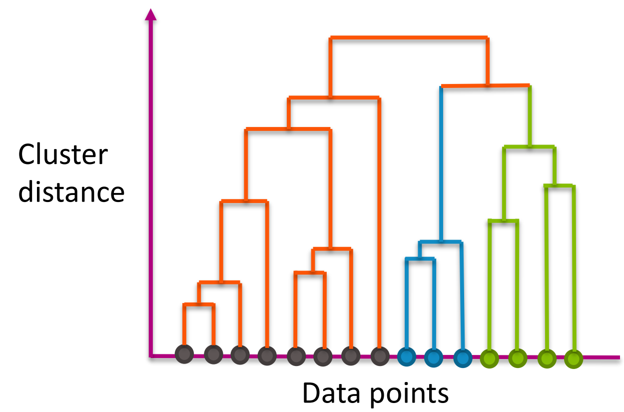


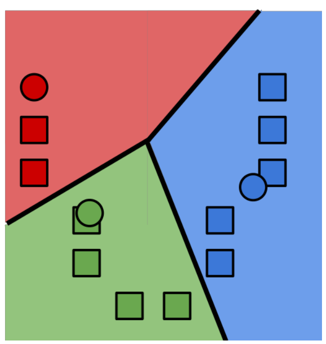
Clustering & Retrieval

Case study: Finding documents

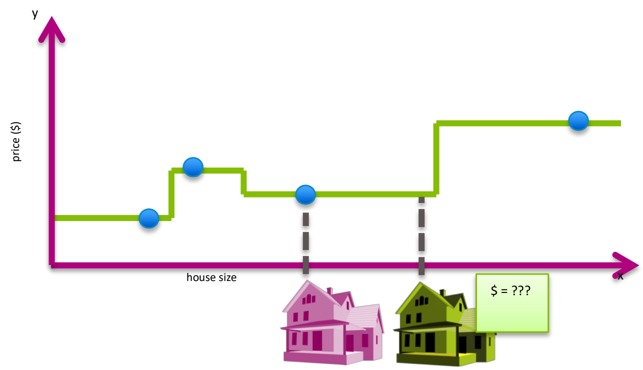
Algorithms

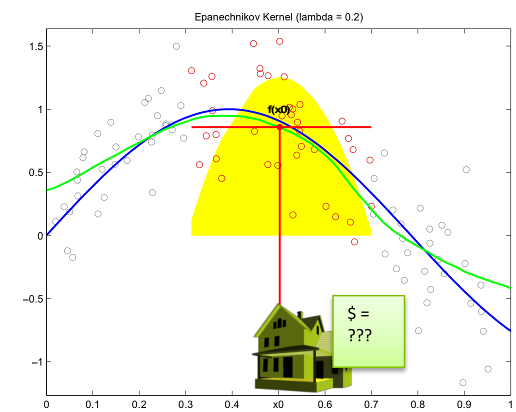
- k-means , k-means++
- Locality-sensitive hashing (LSH)
- NN regression and classification
- Kernel regression
- Agglomerative and divisive clustering
- PCA





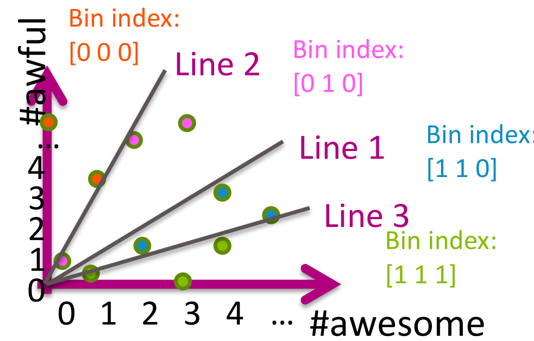
Lloyds Algorithm





Clustering & Retrieval

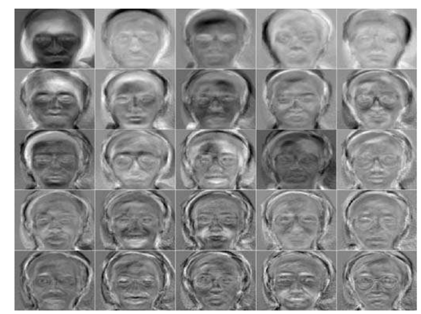
Case study: Finding documents

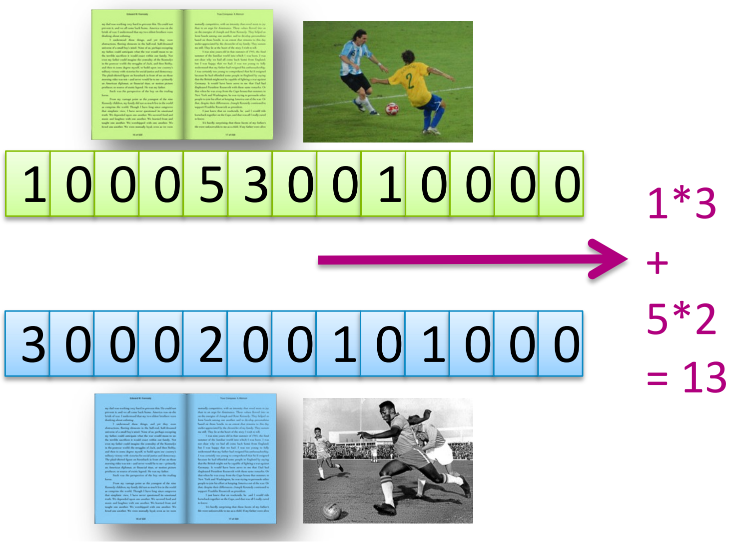


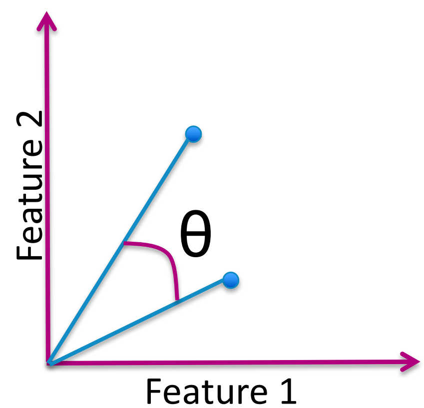
Concepts

- Distance metrics, kernels, approximation algorithms, dimensionality reduction

Principal components:

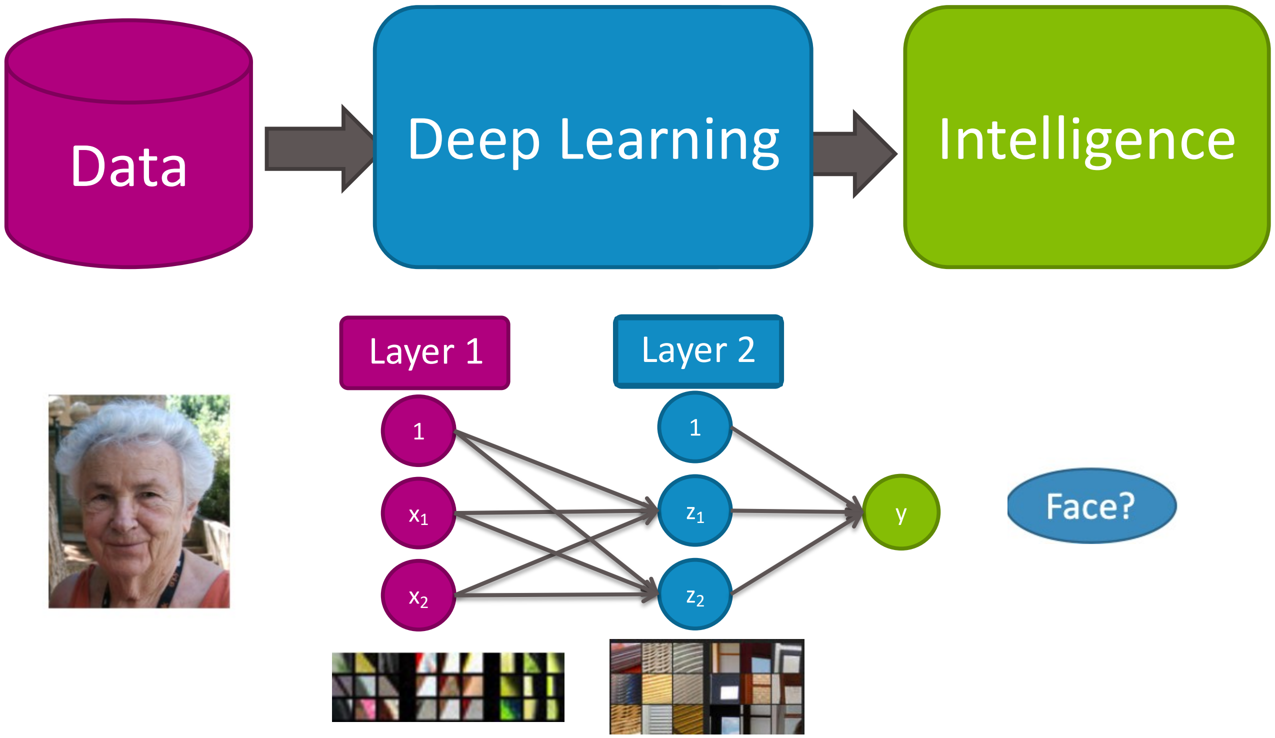






Reconstructing:

Case Study 4:
Image classification



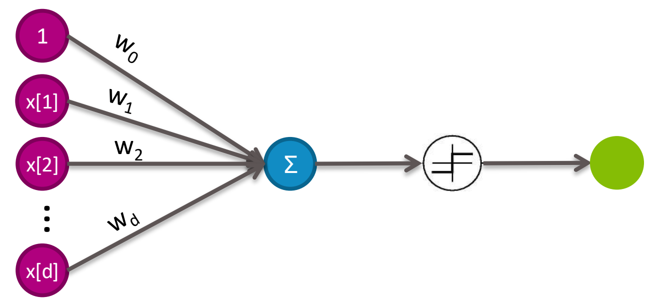
Deep Learning

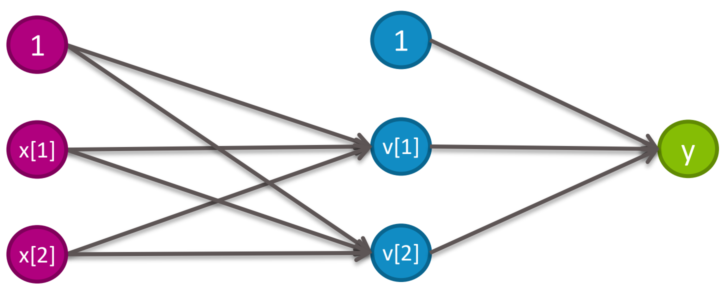
Case study: Image classification

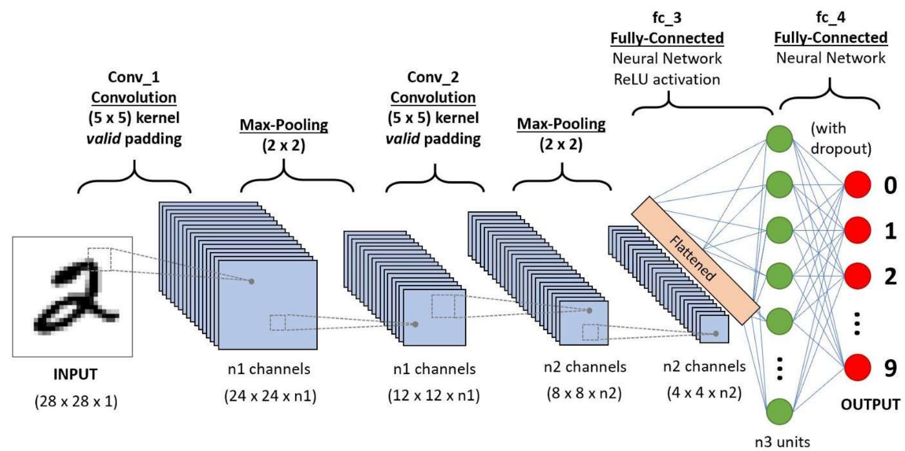
Fully Connected

Models

- Perceptron
- General neural network
- Convolutional neural network





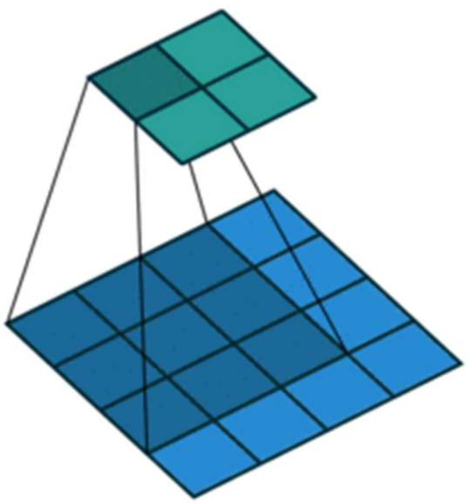


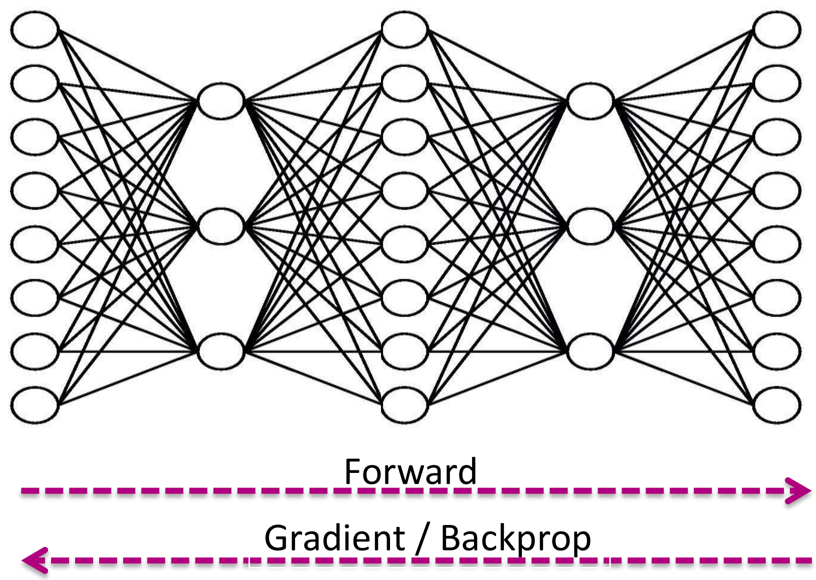
Deep Learning

Case study: Image classification

Algorithms

- Convolutions
- Backpropagation (high level only)



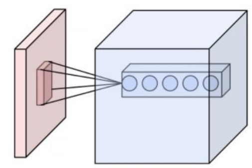


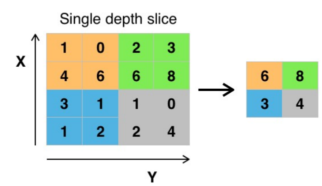
Deep Learning

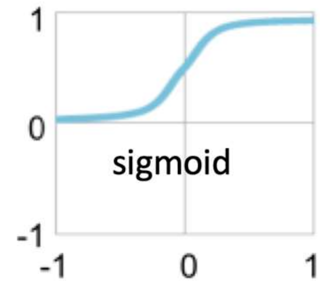
Case study: Image classification

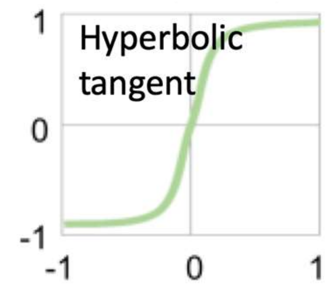
Concepts

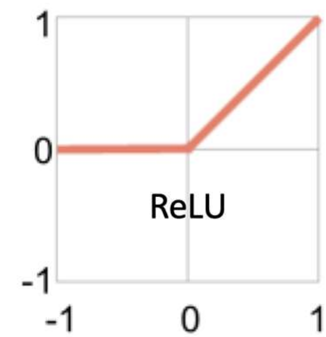
- Activation functions, hidden layers, architecture choices

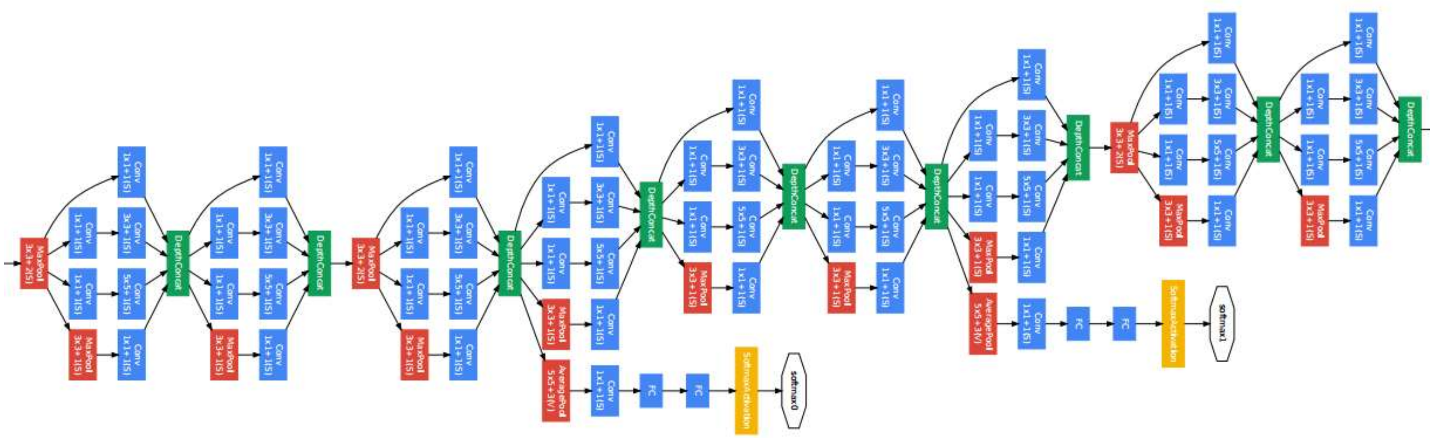








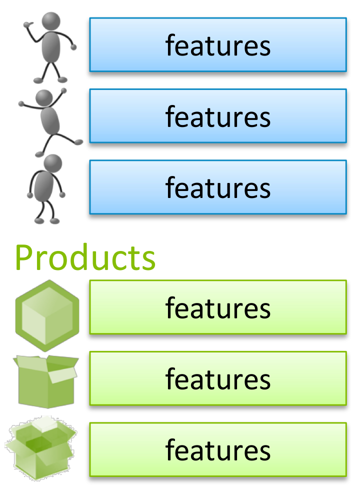




Case Study 5:
Product recommendation

Your past purchases:

+ purchase histories of all customers



Recommended items:

Recommender Systems & Matrix Factorization

Case study: Recommending Products

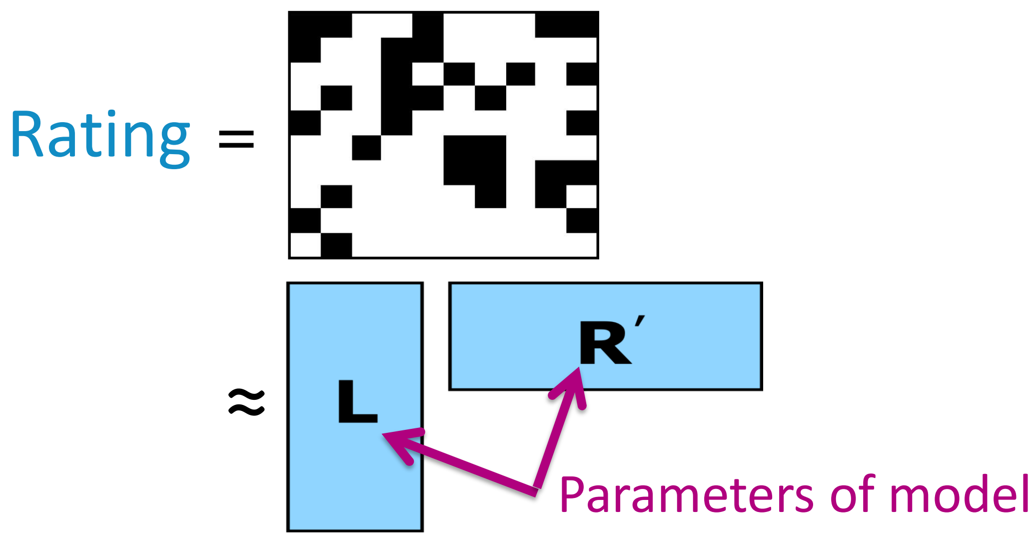
Models

- Collaborative filtering
- Matrix factorization

Popularity

Co-occurence matrix

Featurized MF

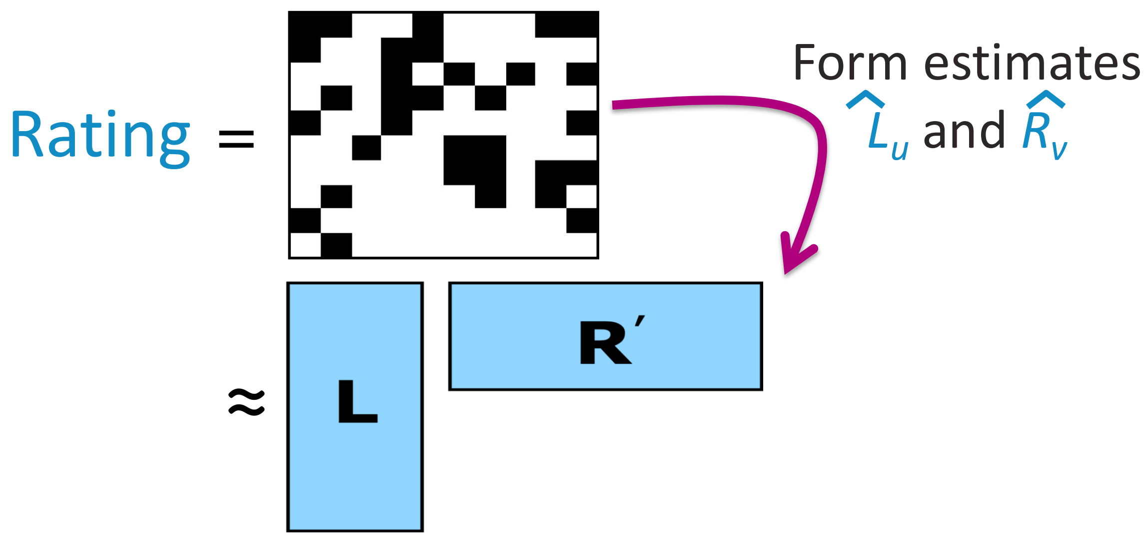


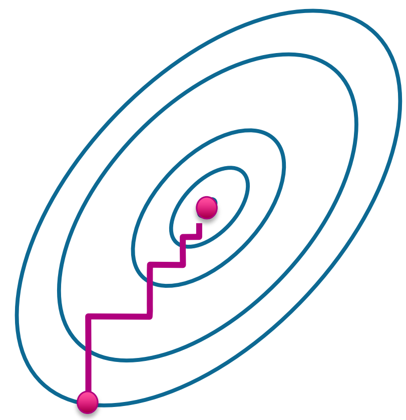
Recommender Systems & Matrix Factorization

Case study: Recommending Products

Algorithms

- Coordinate descent



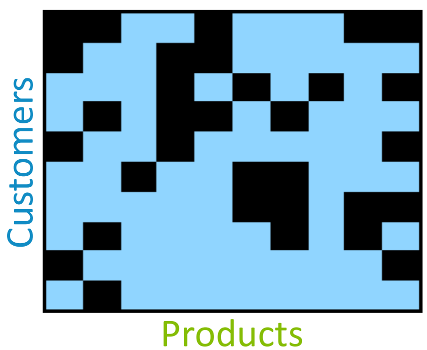


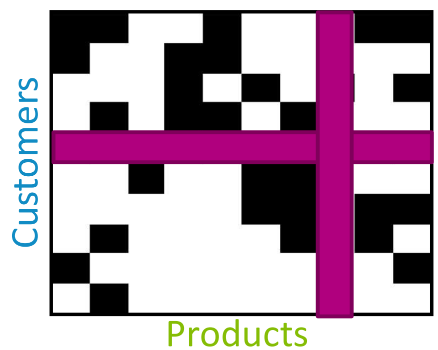
Recommender Systems & Matrix Factorization

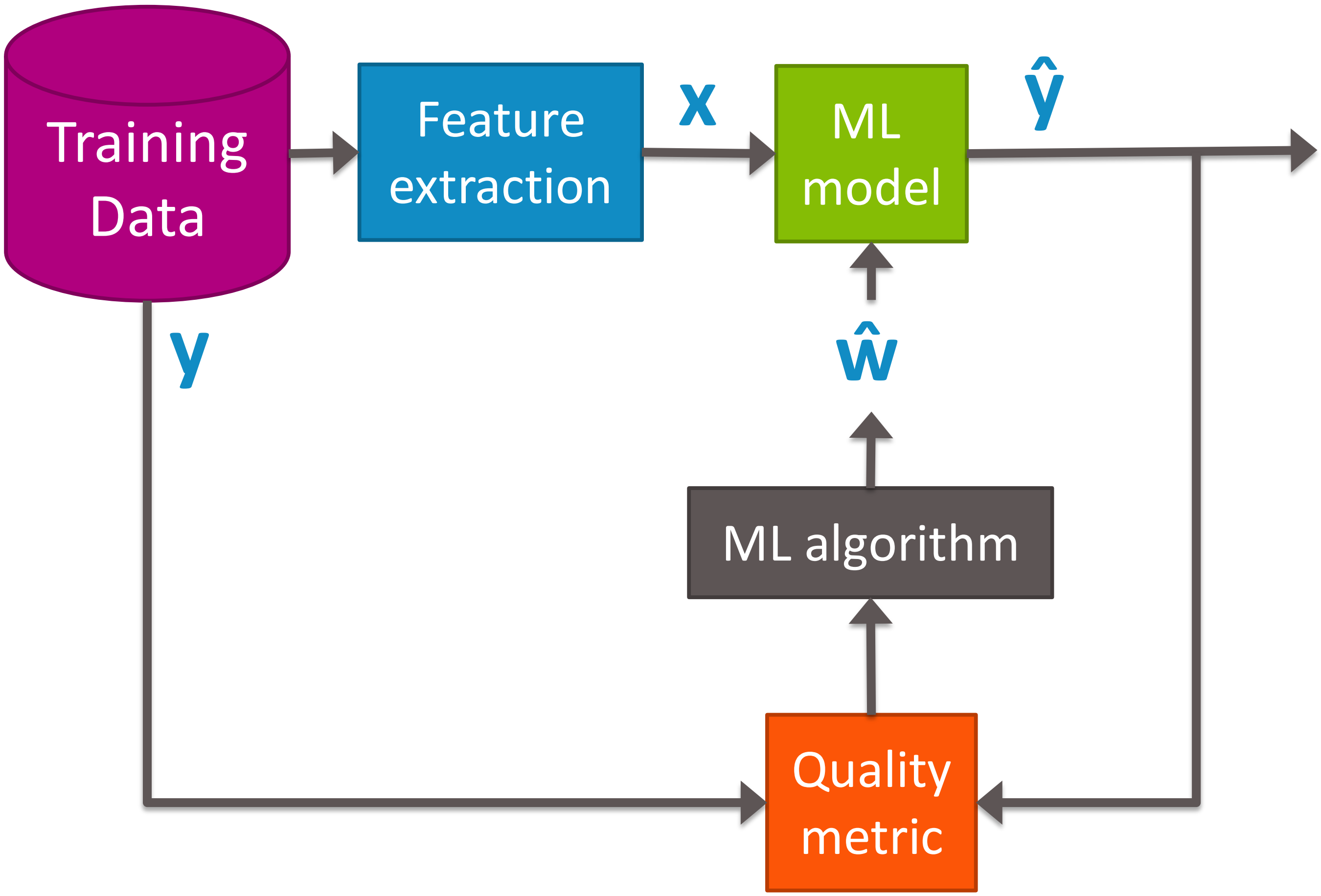
Case study: Recommending Products

Concepts

- Matrix completion, cold-start problem






Training Data
Feature extraction
x
ML model
ŷ
y
ŵ
ML algorithm
Quality metric

Big Picture

Improving the performance at some task through experience!

- Before you start any learning task, remember fundamental questions that will impact how you go about solving it

What is the learning problem?

What model?

With what optimization algorithm?

How will you evaluate the model?

From what experience?

What loss function are you optimizing?

Are there any guarantees?

Congrats on finishing CSE/STAT 416!
Thanks for the hard work!