

Machine Learning

- Last lab!
- Computer program: a set of instructions telling a computer (hardware via a computer) what to do.
- **Supervised learning:** use training data (examples) to “teach” the algorithm what input maps to what output, based on the training set. Here it matters what your structure/algorithm is, but also decisions made in creating your training data
 - Example, looking for criminals, looked at pictures of people smiling vs. not and the people not smiling were deemed criminals (on the computer)
 - Uses preset categories
 - Regression
 - Population growth prediction
 - Estimating life expectancy
 - Market forecasting
 - Weather forecasting
 - Advertising popularity prediction
 - Classification
 - Diagnostics
 - Identity fraud detection
 - Image classification
- **Unsupervised learning:** no inputs or preset categories or labels to match. This can both be useful in producing unbiased information about a data set- but also is not constrained by your question (the computer doesn’t know physics, so might create a model based on superficial feature or error)
 - Computer is kinda guessing
 - Doesn’t know anything about the question but will try to detect things- example: will try to detect people with short hair
 - Clustering
 - Customer segmentation
 - Targeted marketing
 - Recommender systems
 - Dimensionality reduction
 - Feature elicitation
 - Big data visualization
 - Meaningful compression
 - Structure discovery
- **Reinforcement Learning:** seen more with decision learning, task learning, robots trying to get something right and then trying it again
 - Game AI
 - Skill acquisition
 - Learning tasks
 - Robot navigation

- Real time decisions

Models

- Naive Bayes
 - Good starting point, especially for well separated classes
 - Fast, so easy to test
 - Assumes everything is drawn from a gaussian distribution
 - Training set trains the model into the naive bayes and then comes out with new classified data

<https://scikit-learn.org/stable/tutorial/index.html>