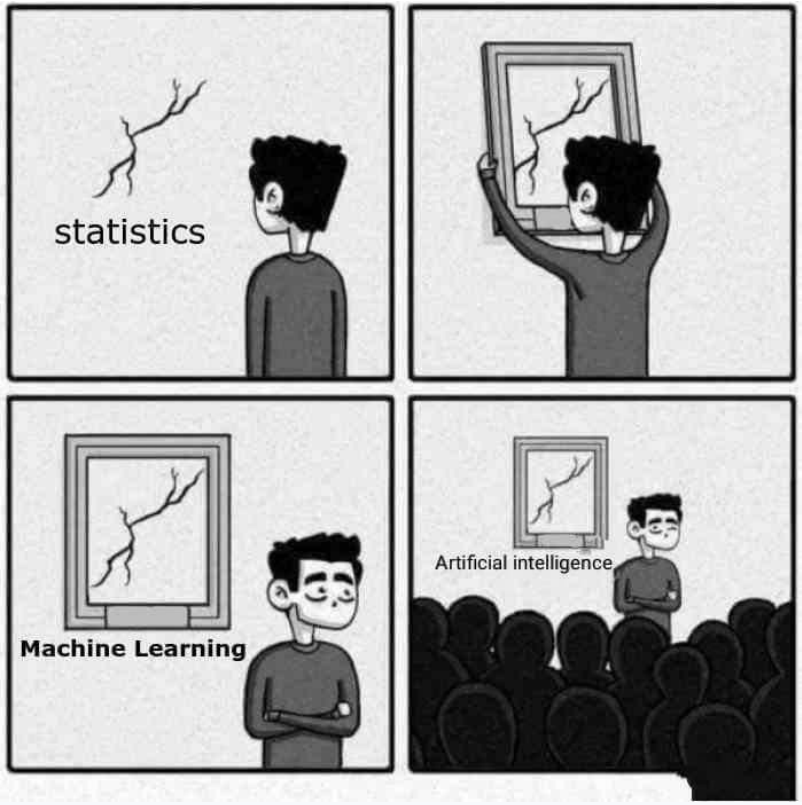
EGM702 – Photogrammetry and Advanced Image Analysis

Week 5, Part 3: Machine Learning

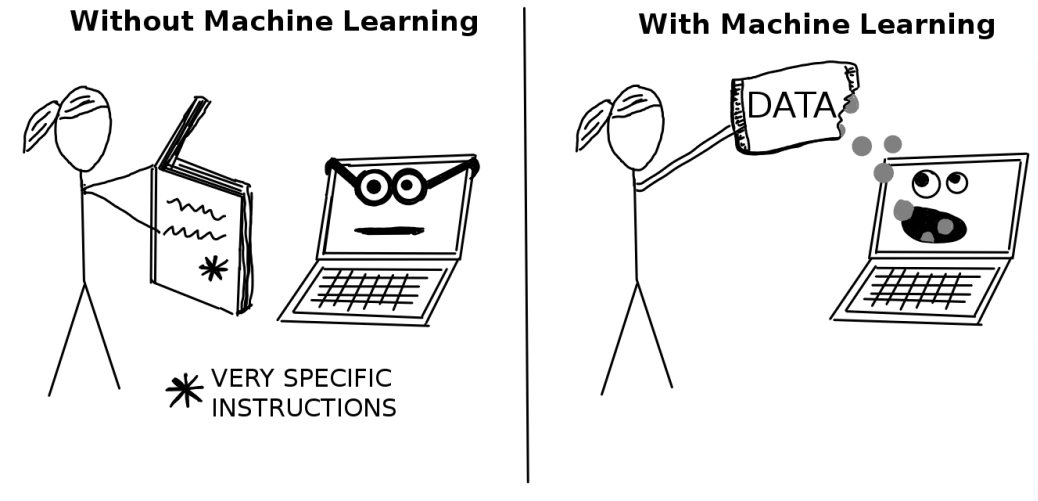
What is Machine Learning?



xkcd.com

What is Machine Learning?

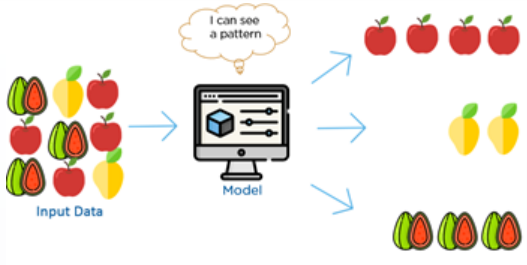
- Machine learning:
  - Computer modelling of learning processes
  - Using statistics to find patterns in data
- Train models without explicitly programming them, using an algorithm
- Deep learning: ML algorithms inspired by human brain
- Algorithm: a process or set of rules to follow in calculations or problem-solving

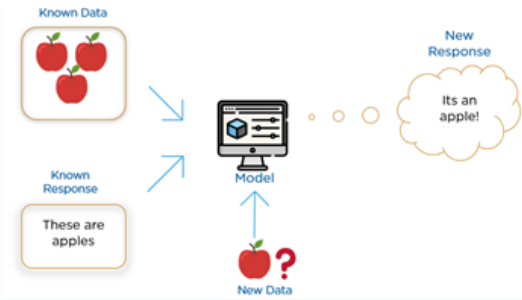


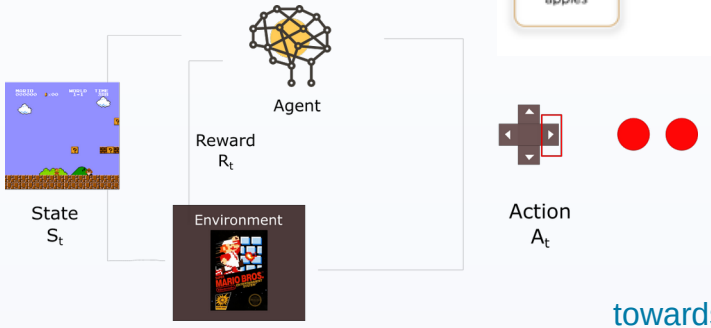
C. Molnar

Machine Learning approaches

- Unsupervised Learning:
  - Algorithm is trained with unlabeled data
  - Finding structure (patterns) in the data
- Supervised Learning:
  - Algorithm trained using labeled data: example input + output value
  - Inferring a function to map (classify) new, unseen data
- Reinforced Learning:
  - Algorithm learns by trial and error to achieve objective
  - ‘Reward’ or ‘punish’ based on behavior



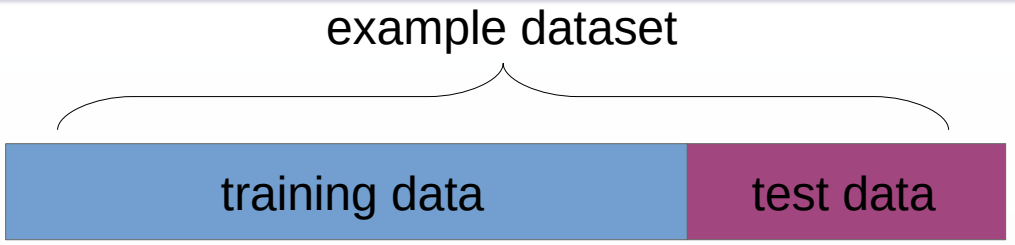




towardsdatascience.com

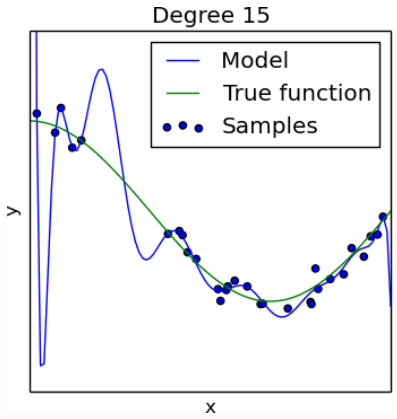
Training data and test data

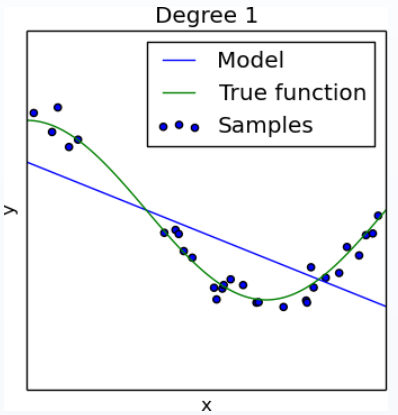
- Start with labeled example dataset (i.e., training samples)
- Use portion of these to train the model
- Use either rest of the example dataset or external dataset to test model
  - Test data should have similar statistical properties to training data

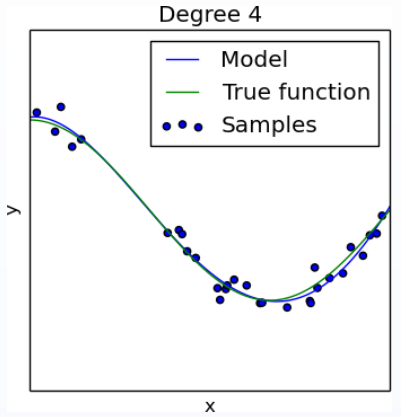


The perils of overfitting

- Overfitting (over-training): model too closely matches the dataset
  - i.e., has more parameters than are justified by dataset
  - Model is unable to handle new (unseen) data
- Underfitting (under-training): model doesn't match the structure of the data
  - i.e., using a linear model for nonlinear data
  - Model will not adequately predict new data
- We want our models to be generalizable (works well with new, unseen, data)







scikit-learn.org

Summary

- Machine learning: computer modelling of learning processes
- ML comes in three main approaches: unsupervised, supervised, and reinforced learning
- Use input data to train a ML model, split into training and test datasets
- Want models to be predictive without overfitting training data

Additional resources

- What is Machine Learning? [IBM]
- Interpretable Machine Learning [C. Molnar]
- Machine Learning Methods [computerphile]
- Deep Learning [computerphile]
- Machine Learning Crash Course [Google]