

INTRODUCTION TO ML

Narges Norouzi

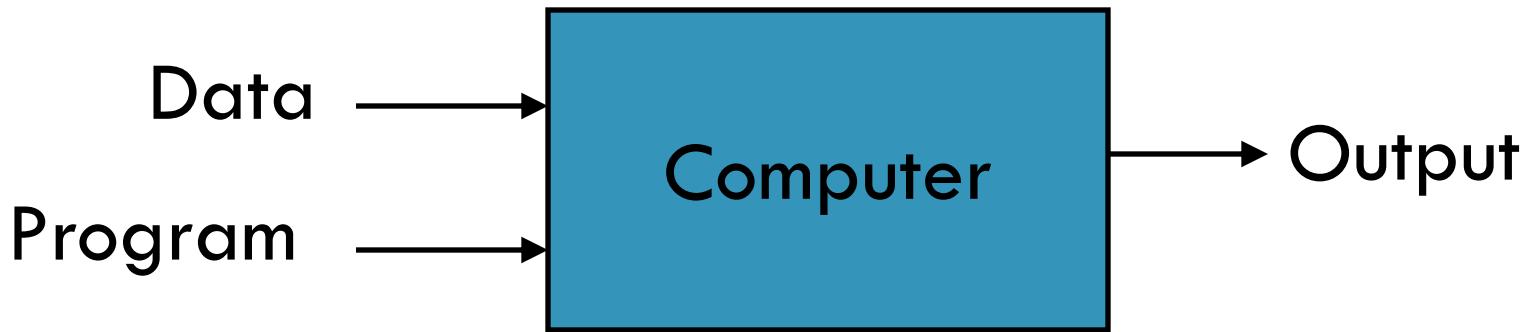
A FEW QUOTES

- “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates, Chairman, Microsoft)
- “Machine learning is the next Internet” (Tony Tether, Director, DARPA)
- “Machine learning is the hot new thing” (John Hennessy, President, Stanford)
- “Machine learning is going to result in a real revolution” (Greg Papadopoulos, CTO, Sun)

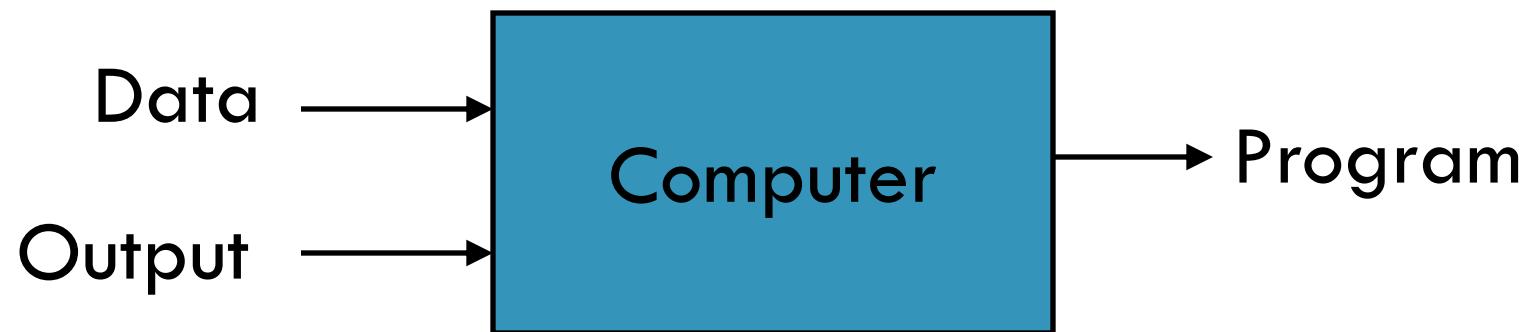
WHAT IS ML?

- Automating automation
- Getting computers to program themselves
- Writing software is the bottleneck
- Let the data do the work instead!

Traditional Programming



Machine Learning



TYPES OF DATA

- **Time series**
 - Finances, clickstream, demand/sale forecast, natural phenomena, biological data, ...
- **Spatial data**
 - Advertising, cell phone and network data, VR gaming, space exploration, ...
- **Textual data**
 - Web, social media, legal documents, articles, debugging, ...
- **Image and video data**
 - Photographs, artworks, ...

ML IN A NUTSHELL

- Tens of thousands of machine learning algorithms
- Hundreds new every year
- Every machine learning algorithm has three components:
 - **Representation**
 - **Evaluation**
 - **Optimization**

REPRESENTATION

- Decision trees
- Sets of rules / Logic programs
- Graphical models (Bayes/Markov nets)
- Neural networks
- Support vector machines
- Model ensembles

EVALUATION

- Accuracy
- Precision and recall
- Squared error
- Likelihood
- Posterior probability
- Cost / Utility
- Error margin
- Entropy

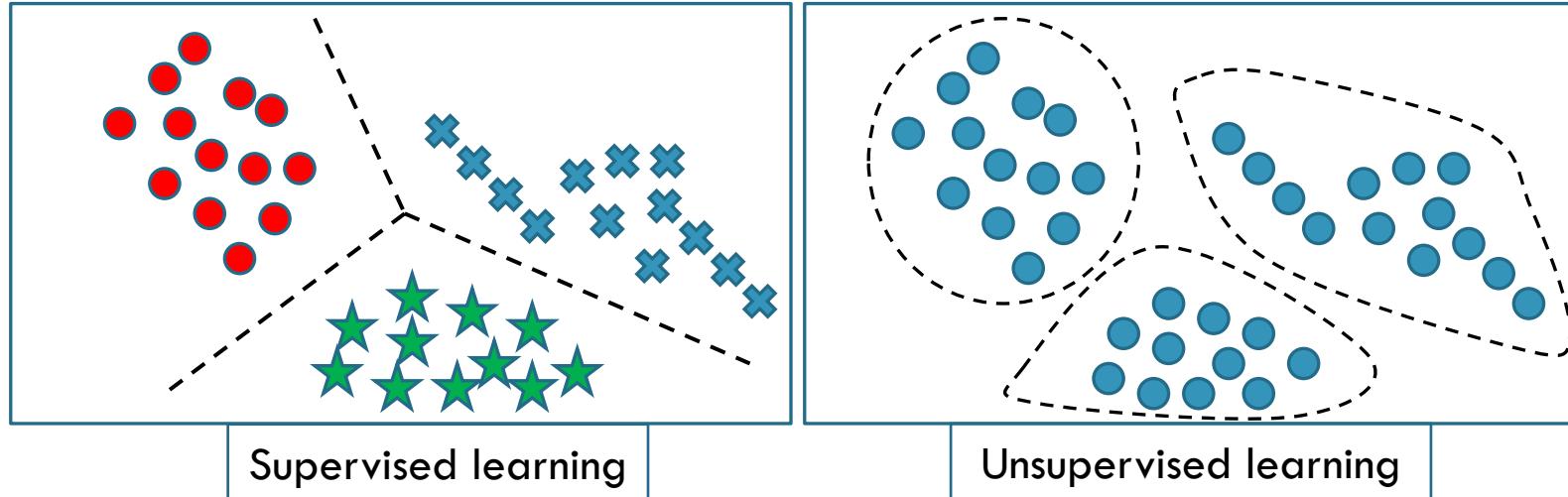
OPTIMIZATION

- Combinatorial optimization
 - E.g.: Greedy search
- Convex optimization
 - E.g.: Gradient descent
- Constrained optimization
 - E.g.: Linear programming

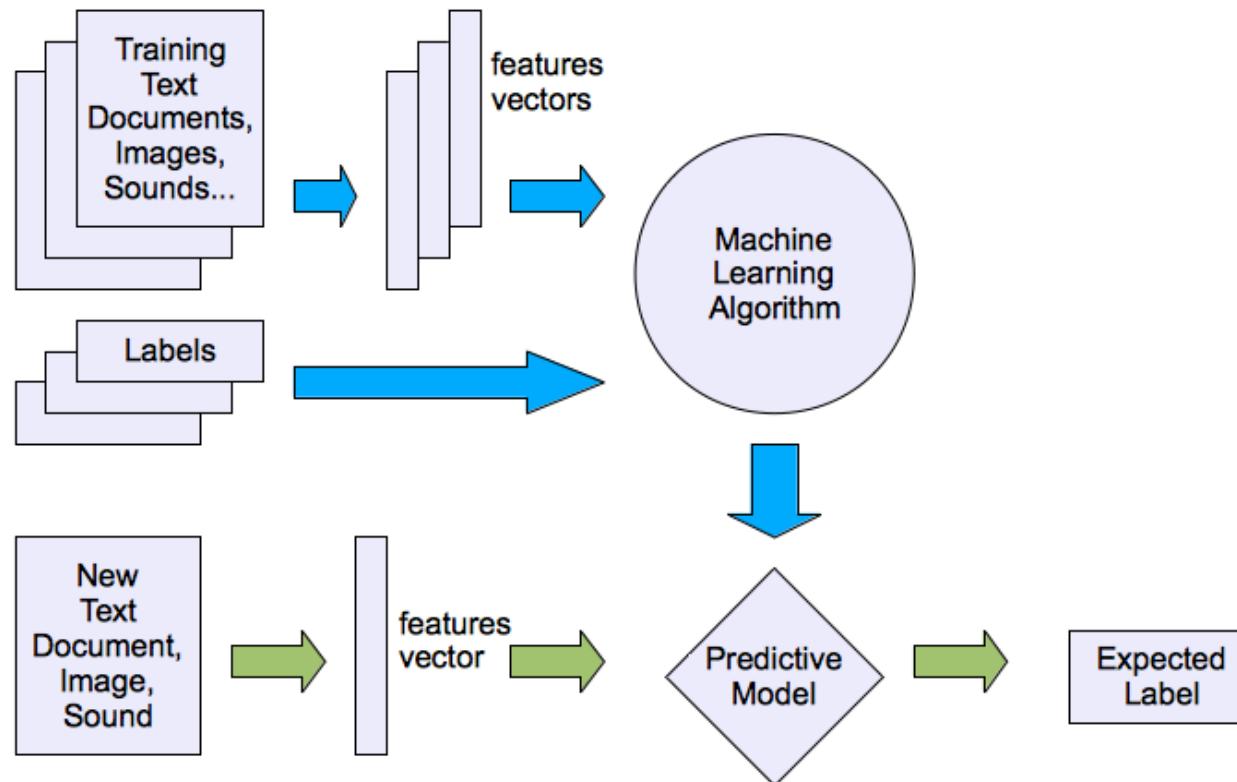
TYPES OF LEARNING

- **Supervised (inductive) learning**
 - Training data includes desired outputs
- **Unsupervised learning**
 - Training data does not include desired outputs
- **Semi-supervised learning**
 - Training data includes a few desired outputs
- **Reinforcement learning**
 - Learn action to maximize payoff
- Fairly new category: **Self-supervised learning**

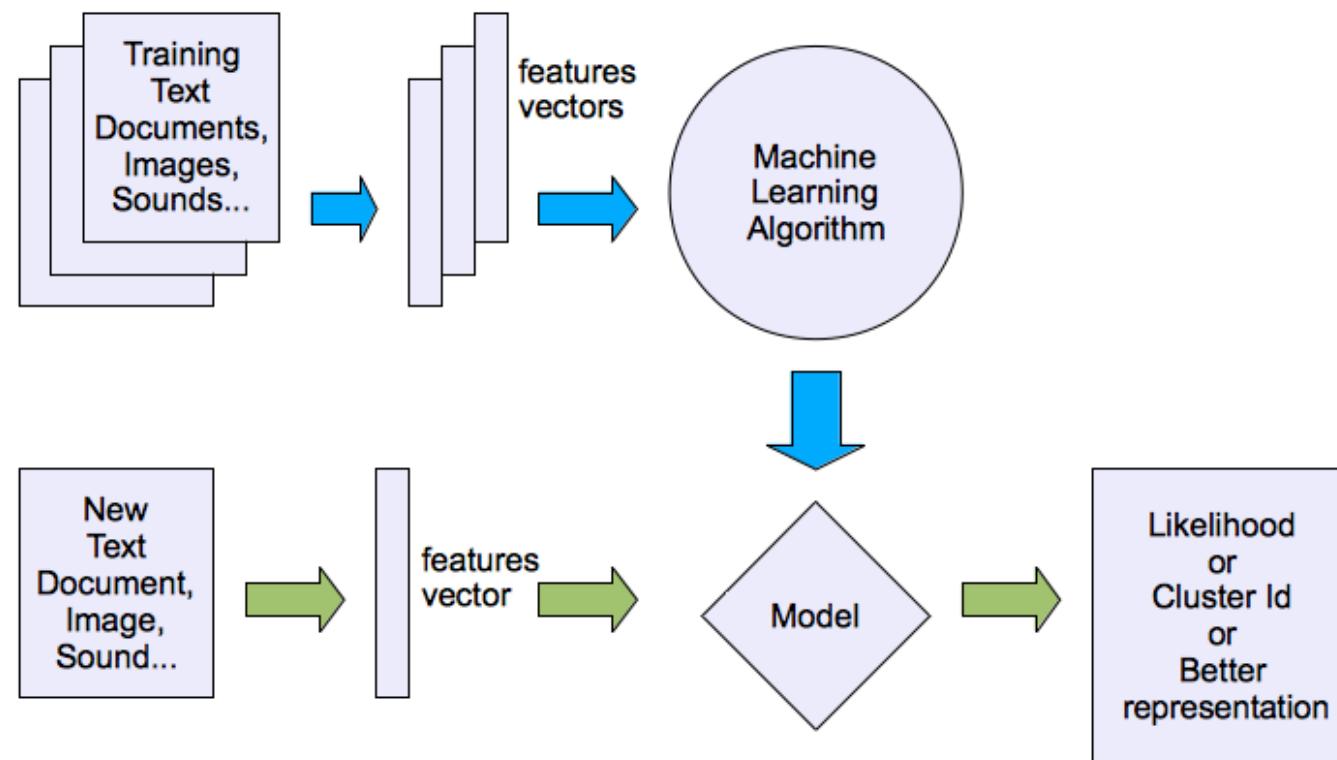
ALGORITHMS



SUPERVISED LEARNING



UNSUPERVISED LEARNING



ML IN PRACTICE

- Understanding domain, prior knowledge, and goals
- Data integration, selection, cleaning, pre-processing, etc.
- Learning models
- Interpreting results
- Consolidating and deploying discovered knowledge
- Loop

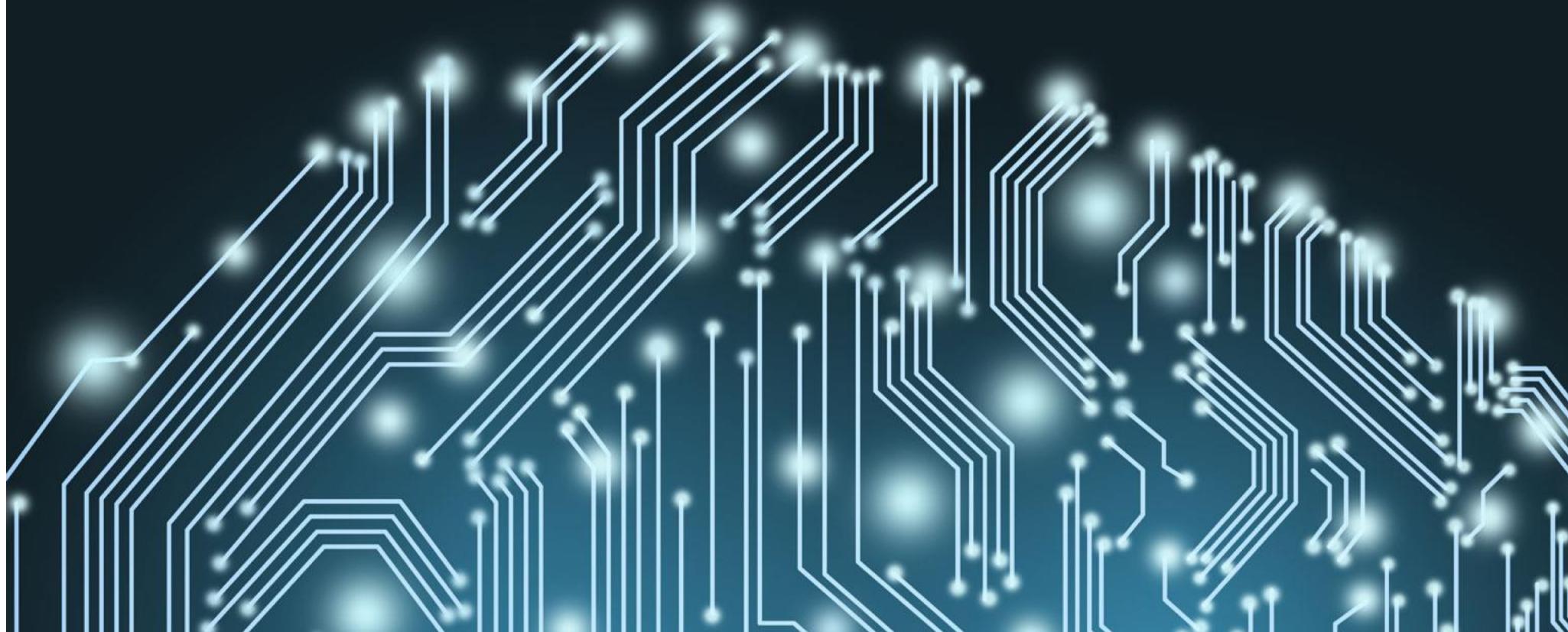
CSE 144

- We will focus on core concepts of machine learning
 - Feature engineering
 - Regression analysis
 - Neural networks
 - Deep neural networks and its applications
 - CNN
 - RNN
 - Autoencoder

PYTHON TUTORIAL

bit.ly/pythonoverview3

NARGES NOROUZI



QUESTIONS?