

Machine Learning & Deep Learning

Arturo Devesa, CEO

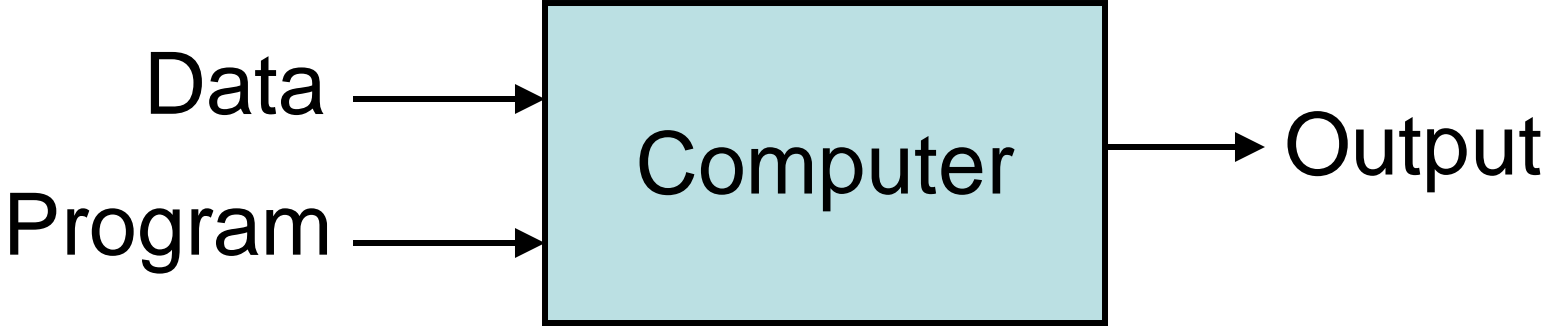
A Few Quotes

- “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)
- “Web rankings today are mostly a matter of machine learning” (Prabhakar Raghavan, Dir. Research, Yahoo)
- “Machine learning is going to result in a real revolution”
(Greg Papadopoulos, CTO, Sun)
- “Machine learning is today’s discontinuity”
(Jerry Yang, CEO, Yahoo)

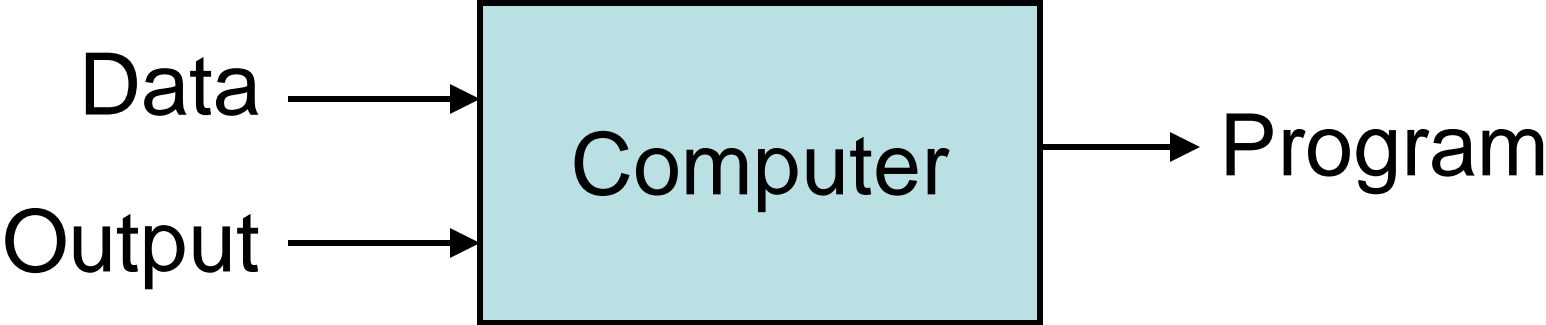
So What Is Machine Learning?

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

Traditional Programming



Machine Learning



Magic?

No, more like gardening

- **Seeds** = Algorithms
- **Nutrients** = Data
- **Gardener** = You
- **Plants** = Programs



Sample Applications

- Web search
- Chatbots, conversational Siri, Alexa
- Computational biology
- Finance
- E-commerce
- Space exploration
- Robotics
- Information extraction
- [Your favorite area]

ML in a Nutshell

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

Representation

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

Evaluation

- Accuracy
- Precision and recall
- Squared error
- Likelihood
- Posterior probability
- Cost / Utility
- Margin
- Entropy
- K-L divergence
- Etc.

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**
 - Rewards from sequence of actions

Inductive Learning

- **Given** examples of a function $(X, F(X))$
- **Predict** function $F(X)$ for new examples X
 - Discrete $F(X)$: Classification
 - Continuous $F(X)$: Regression
 - $F(X) = \text{Probability}(X)$: Probability estimation

What We Cover

- **Supervised learning**
 - Decision tree induction
 - Rule induction
 - Instance-based learning
 - Bayesian learning
 - Neural networks
 - Support vector machines
 - Model ensembles
 - Learning theory
- **Unsupervised learning**
 - Clustering
 - Dimensionality reduction
 - Autoencoders and Deep Belief Nets
 - Generative Adversarial Networks

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 and interactions