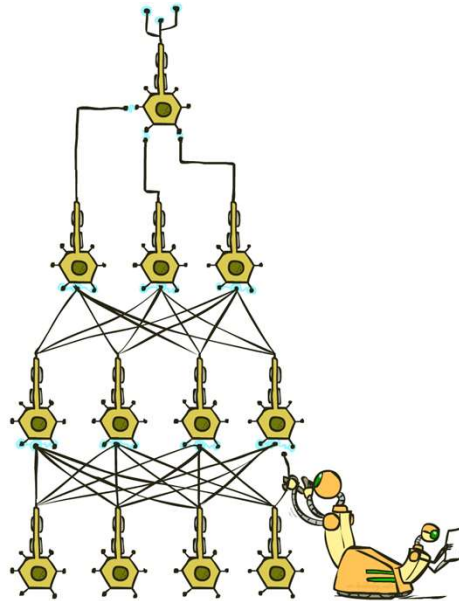


CSE 473: Artificial Intelligence

Multilayer Perceptrons



Steve Tanimoto --- University of Washington

[These slides were created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley. All CS188 materials are available at <http://ai.berkeley.edu>.]

How Many Computers to Identify a Cat?

HOME PAGE TODAY'S PAPER VIDEO MOST POPULAR U.S. Edition

Subscribe: Digital / Home Delivery Log In Register Now Help

The New York Times Business Day **Technology** Search All NYTimes.com

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS

AN ADVISOR WHO KNOWS THE COMPLEXITY OF YOUR FINANCIAL LIFE. AND HOW TO SIMPLIFY IT. Merrill Lynch Wealth Management Bank of America Corporation Connect with your Silicon Valley Advisor

How Many Computers to Identify a Cat? 16,000



Log in to see what your friends are sharing on nytimes.com. Privacy Policy | What's This? Log In With Facebook

What's Popular Now

Restaurant Review: Guy's American Kitchen & Bar in Times Square

Occupancy Offshoot Aims to Erase People's Debts

An image of a cat that a neural network taught itself to recognize. By JOHN MARKOFF Published: June 25, 2012

MOUNTAIN VIEW, Calif. — Inside Google's secretive X laboratory, known for inventing self-driving cars and augmented reality glasses, a small group of researchers began working several years ago on a simulation of the human brain.

There Google scientists created one of the largest neural networks for machine learning by connecting 16,000 computer processors, which they turned loose on the Internet to learn on its own.

Presented with 10 million digital images found in YouTube videos, what did Google's brain do? What millions of humans do with YouTube: looked for cats.

FACEBOOK TWITTER GOOGLE+ E-MAIL SHARE PRINT REPRINTS

Today's Headlines Daily E-Mail Sign up for a roundup of the day's top stories, sent every morning. See Sample | Privacy Policy

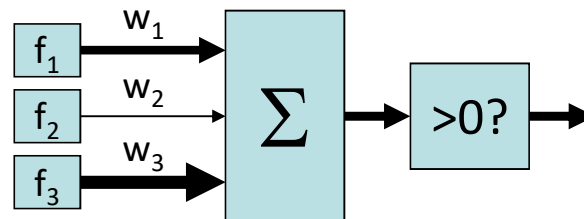
Subscribe to Technology RSS Feeds

- Technology News
- Internet
- Start-Ups
- Business
- Companies
- Computing
- Bits Blog
- Personal Tech
- Pogue's Posts

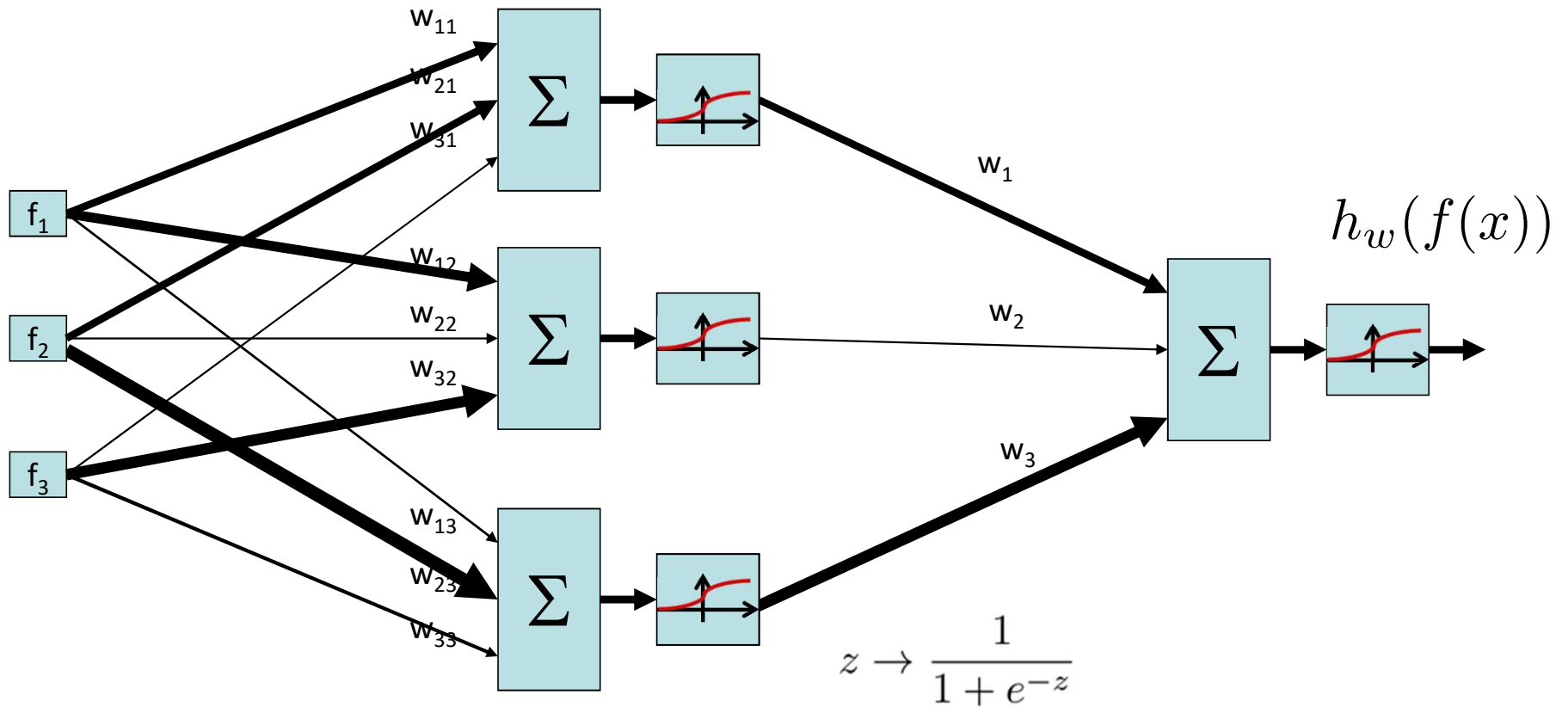
MOST E-MAILED MOST VIEWED

“Google Brain”
[Le, Ng, Dean, et al, 2012]

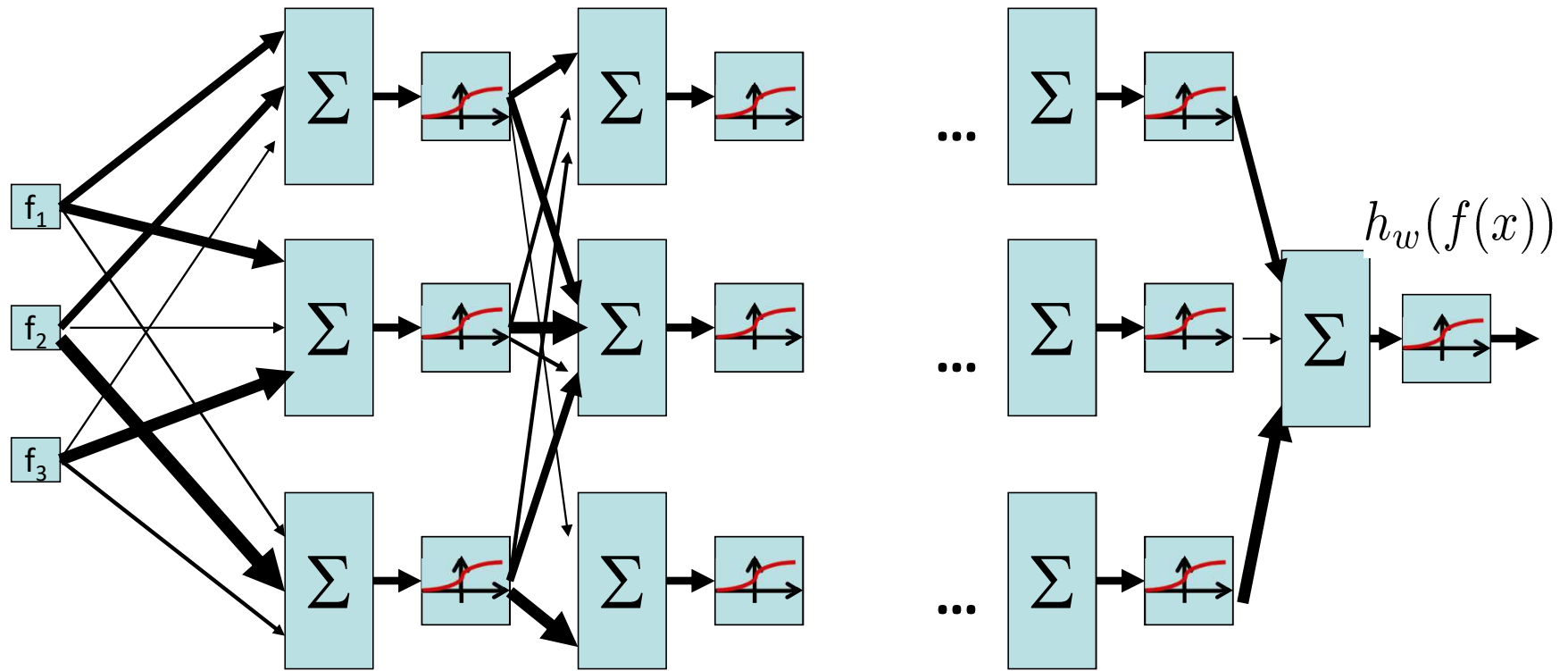
Perceptron



Two-Layer Neural Network

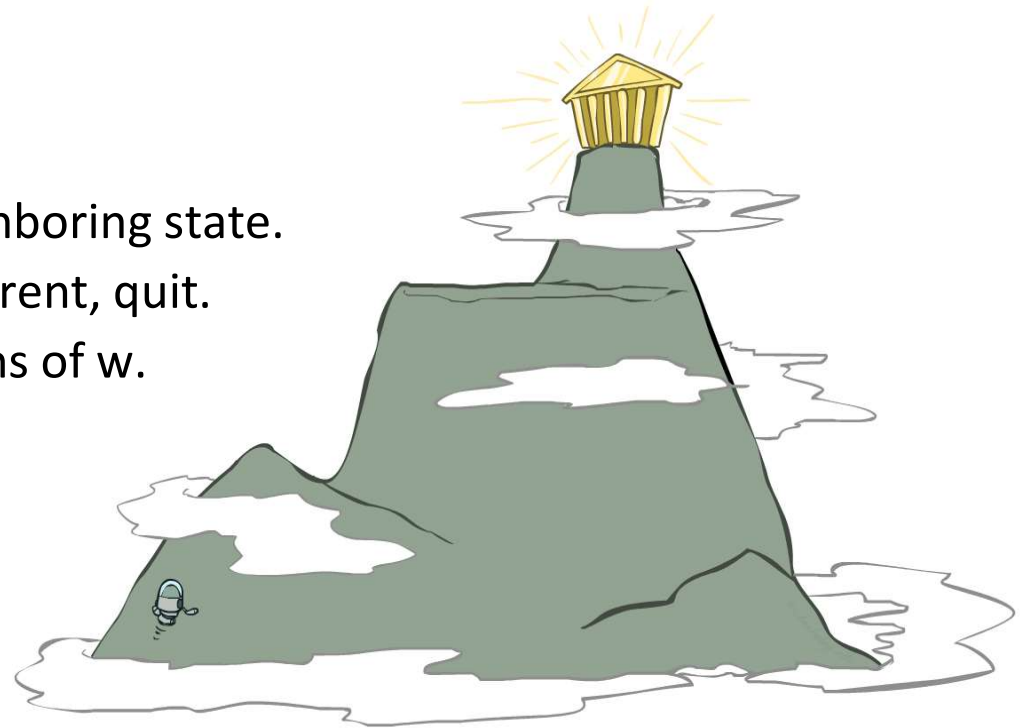


N-Layer Neural Network



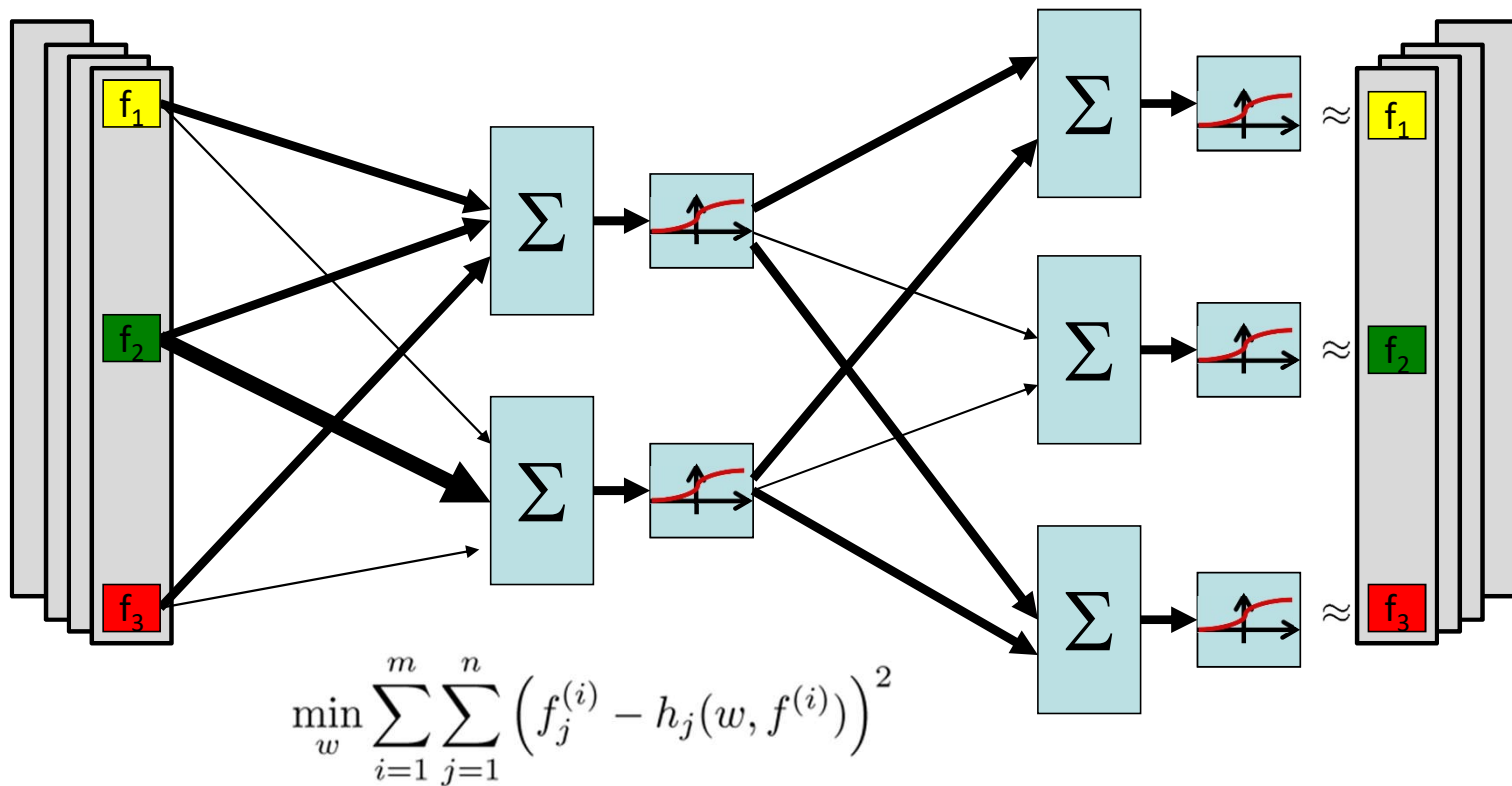
Hill Climbing

- Simple, general idea:
 - Start wherever
 - Repeat: move to the best neighboring state.
 - If no neighbors better than current, quit.
 - Neighbors = small perturbations of w .
- Property
 - Many local optima



--> How to find a good local optimum?

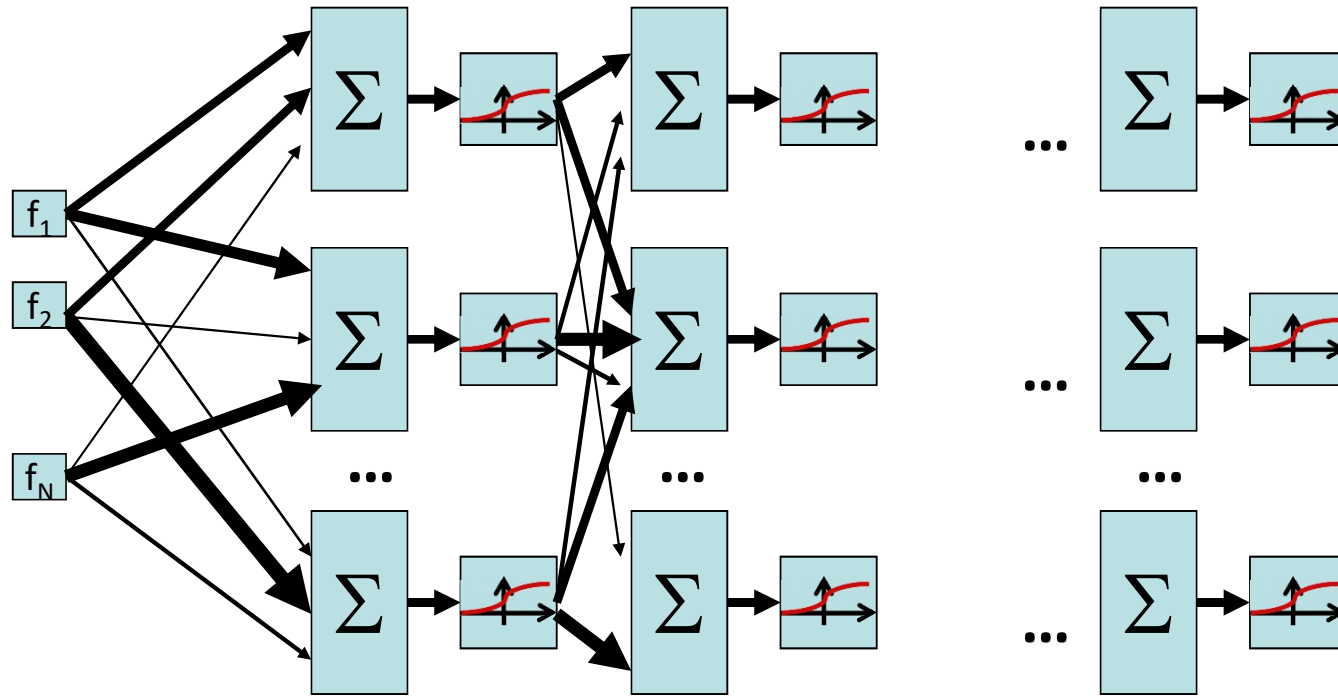
Auto-Encoder (Crude Idea Sketch)



Training Procedure: Stacked Auto-Encoder

- Auto-encoder
 - Layer 1 = “compressed” version of input layer
- Stacked Auto-encoder
 - For every image, make a compressed image (= layer 1 response to image)
 - Learn Layer 2 by using compressed images as input, and as output to be predicted
 - Repeat similarly for Layer 3, 4, etc.
- Some details left out
 - Typically in between layers responses get agglomerated from several neurons (“pooling” / “complex cells”)

Final Result: Trained Neural Network



Final Result: Trained Neural Network

