CSE 473: Artificial Intelligence
Multilayer Perceptrons

How Many Computers to Identify a Cat?

Perceptron

\[ f_1 + f_2 + f_3 > 0? \]

Two-Layer Neural Network

\[ \sum \sum \sum > 0? \]

Hill Climbing

- Simple, general idea:
  - Start wherever
  - Repeat: move to the best neighbor
  - If no neighbors better than current, quit

- 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