Deep Learning

Payam Esfandiari

Spring 15

Outline

- Short History
- Introduction
- Deep Architectures
- Application in Genomics

Short History

- A Class of Machine Learning techniques introduced by Hinton in 2006

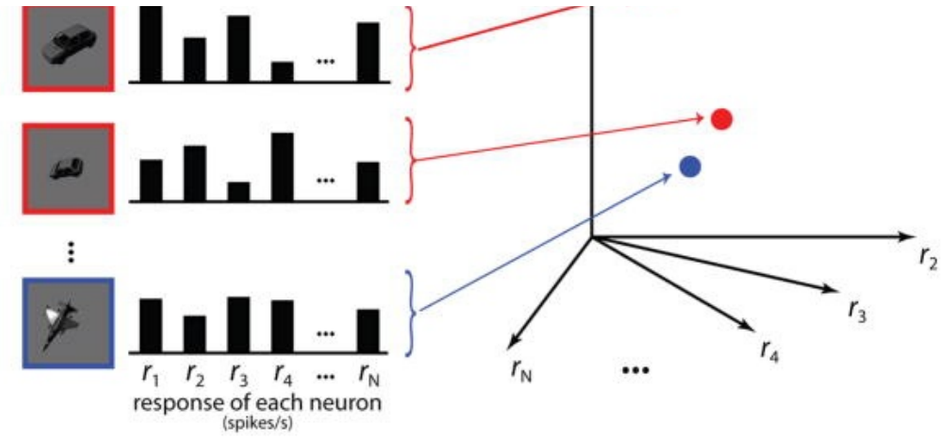
- Andrew Ng popularized it !

Short History

- Deep architecture is designed to mimic Brain Layers :
    - Brain learns layers of features
    - Neurons in the Brain
    - A Great Talk :
        - *How does the brain solve visual object recognition? DiCarlo et al. 2012*

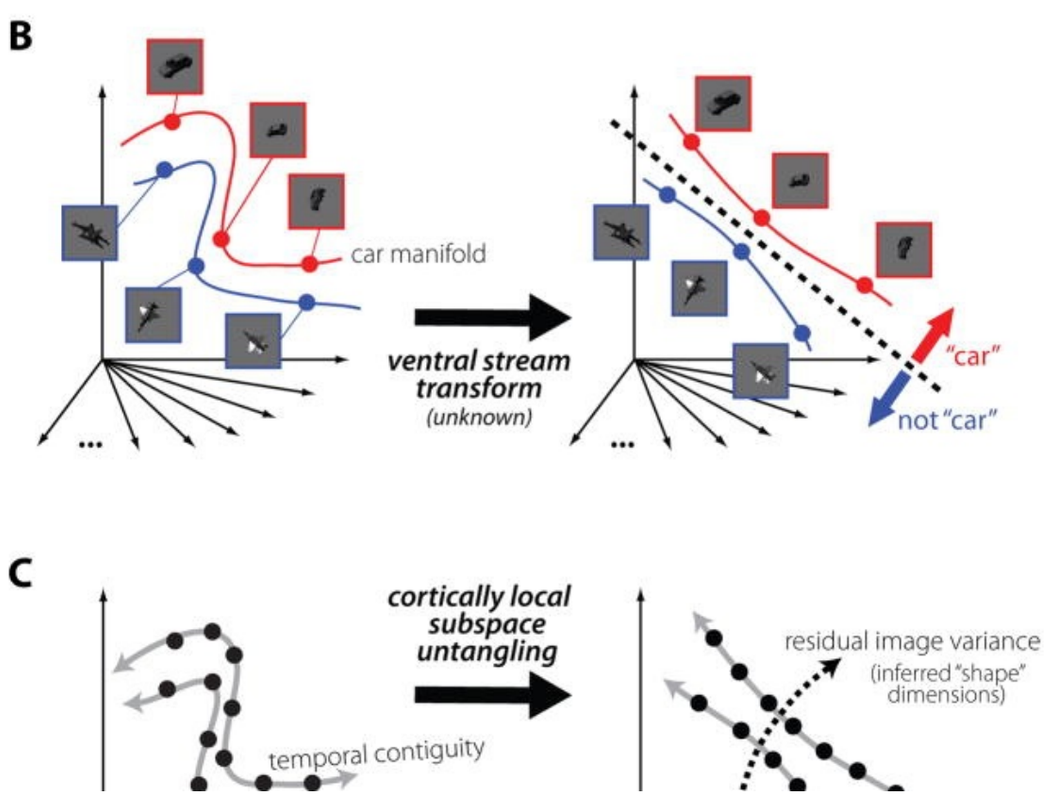
Introduction

- Neurons ***response*** to each object differently



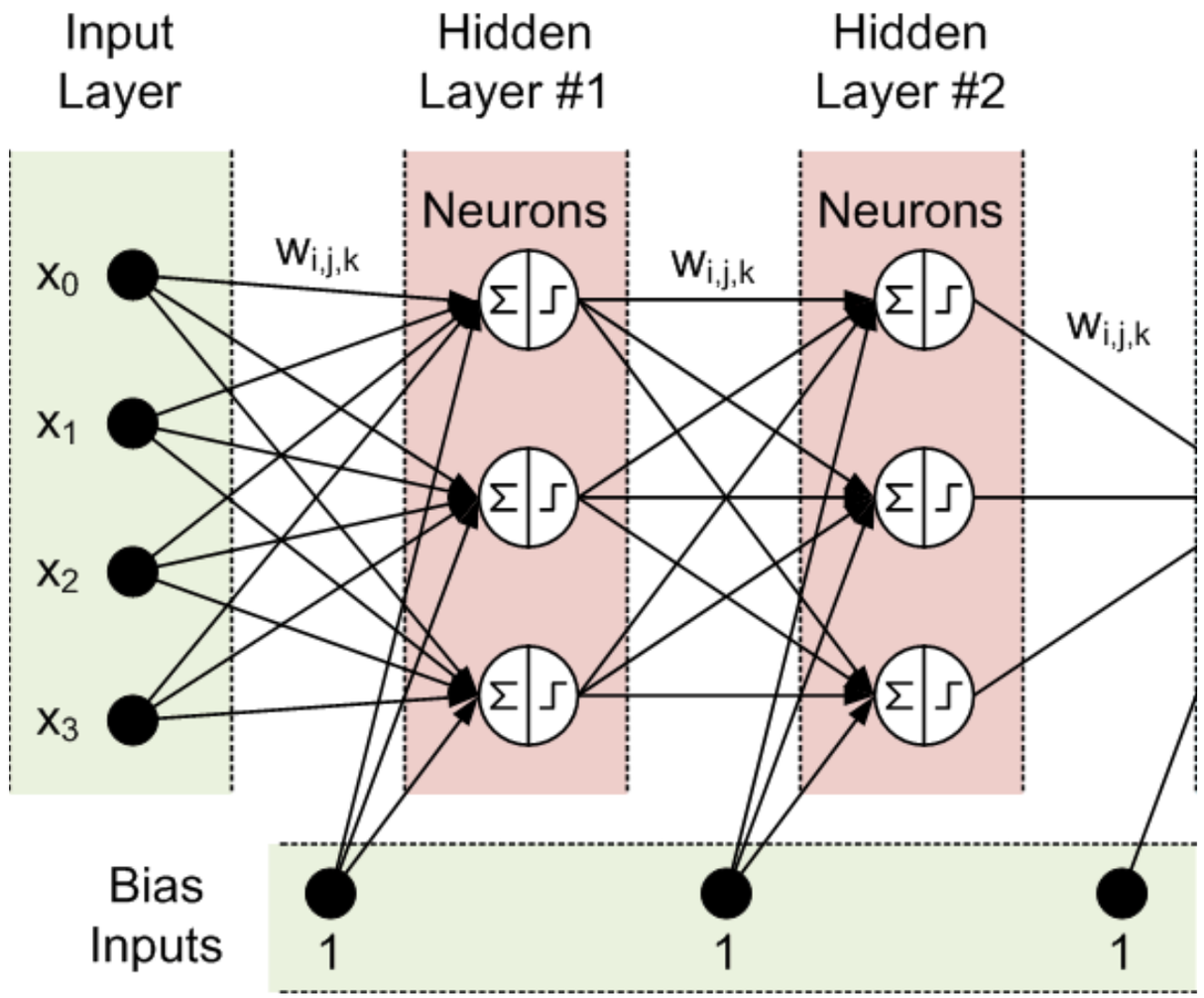
- Each response can be a point in the ***Response Space***

- Our Brain will find separating planes between each class



Introduction

- Basic Concepts :
  - Stack Layers of Neurons
  - Try to find out ***W***
  - Training by ***Back-Propagation***



Introduction

```
initialize network weights (often small random values)
do
   forEach training example ex
      prediction = neural-net-output(network, ex)  // forward pass
      actual = teacher-output(ex)
      compute error (prediction - actual) at the output units
      compute  for all weights from hidden layer to output layer  // backward pass
      compute  for all weights from input layer to hidden layer   // backward pass continued
      update network weights // input layer not modified by error estimate
until all examples classified correctly or another stopping criterion satisfied
return the network
```

Introduction

- So why is it so popular :
    - Scalable
    - Accurate
    - Adaptive
    - Interesting Results

Introduction

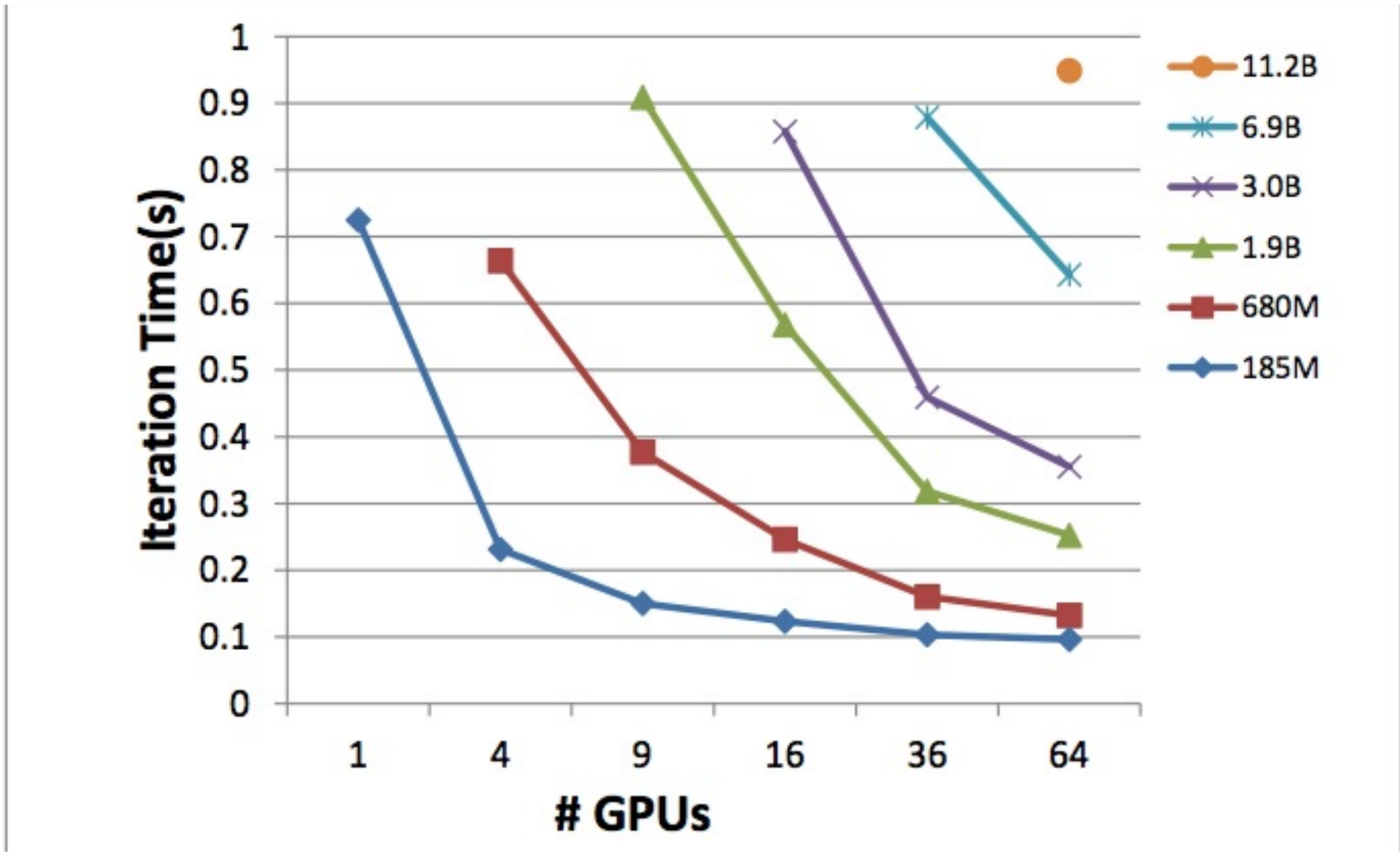
- Several Papers :

1.Building High-level Features Using Large Scale Unsupervised Learning - Andrew Ng - 2012

Traditional ML model :
Feature Extraction —> Machine Learning
Instead : Learn better features —> Cluster them

2.Deep learning with COTS HPC systems - Andrew Ng - ICML 2013

Training a huge system is overwhelming
Build a huge system on GPU clusters



Time taken to perform a mini-batch update for all weights in large neural networks of sizes ranging from 180 million parameters up to 11.2 billion parameters

Introduction

3. Unsupervised feature learning for audio classification using convolutional deep belief networks - Andrew Ng - NIPS 2009

4. Learning features from music audio with deep belief networks - ISMIR 2010

5. Audio-based music classification with a pre trained convolutional network - ISMIR 2012

6. Analyzing drum patterns using conditional deep belief networks - ISMIR 2012

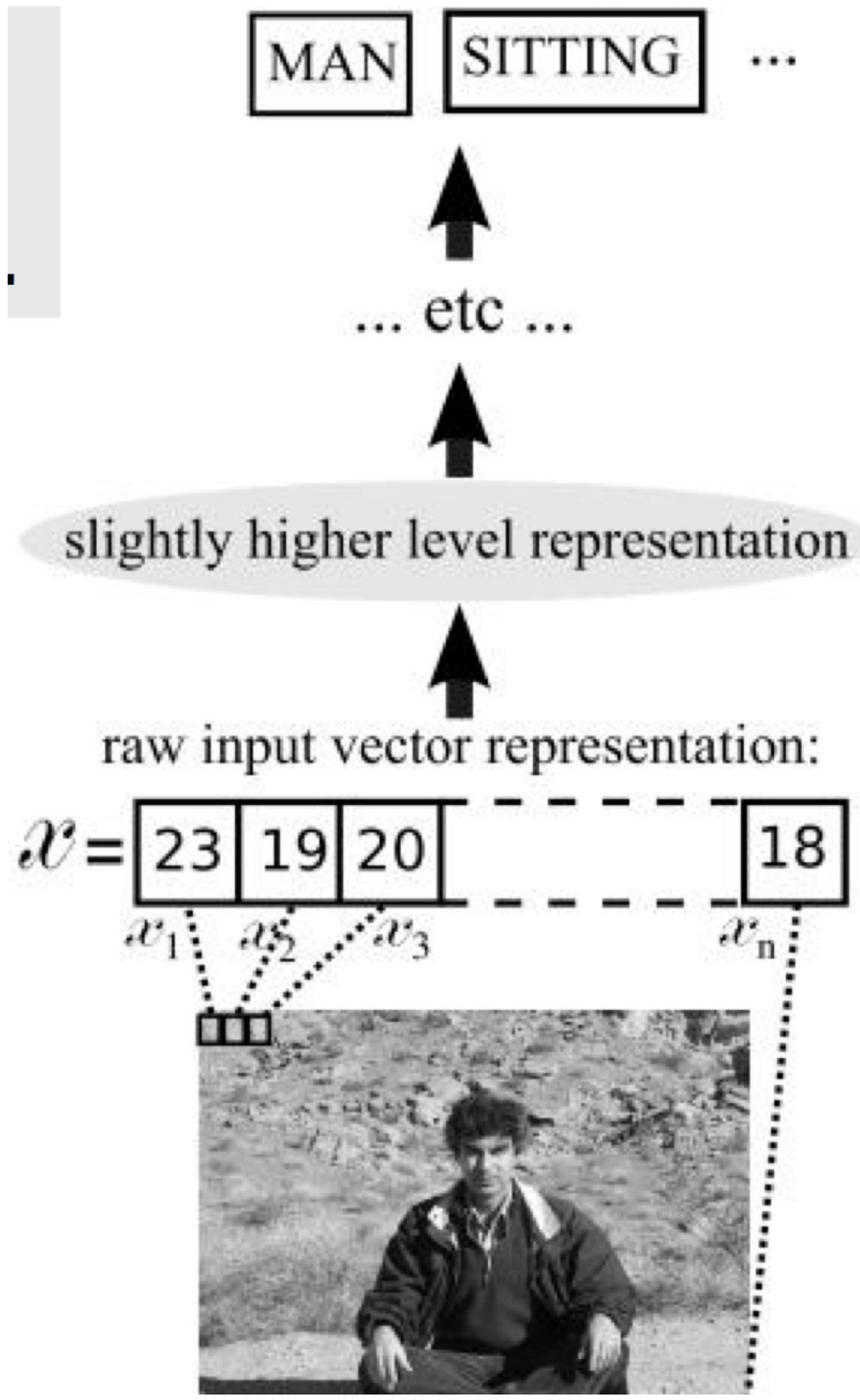
Deep Architecture

- So Deep learning aim at :
  - learning feature hierarchies :
    representation learning
  - learning representations of data that make it easier to extract useful information
    Representation Learning: A Review and New Perspectives

Deep Architecture

How does it work ?

We learn layers of features



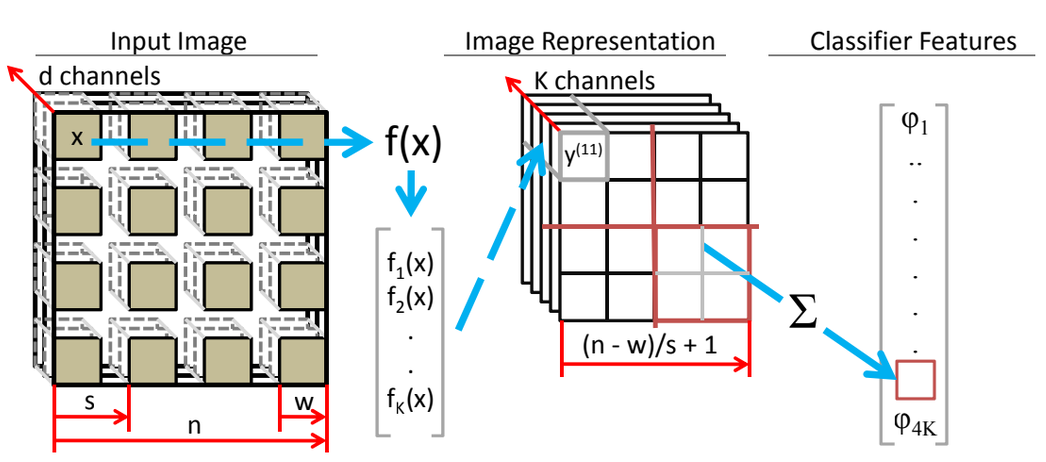
Deep Architecture

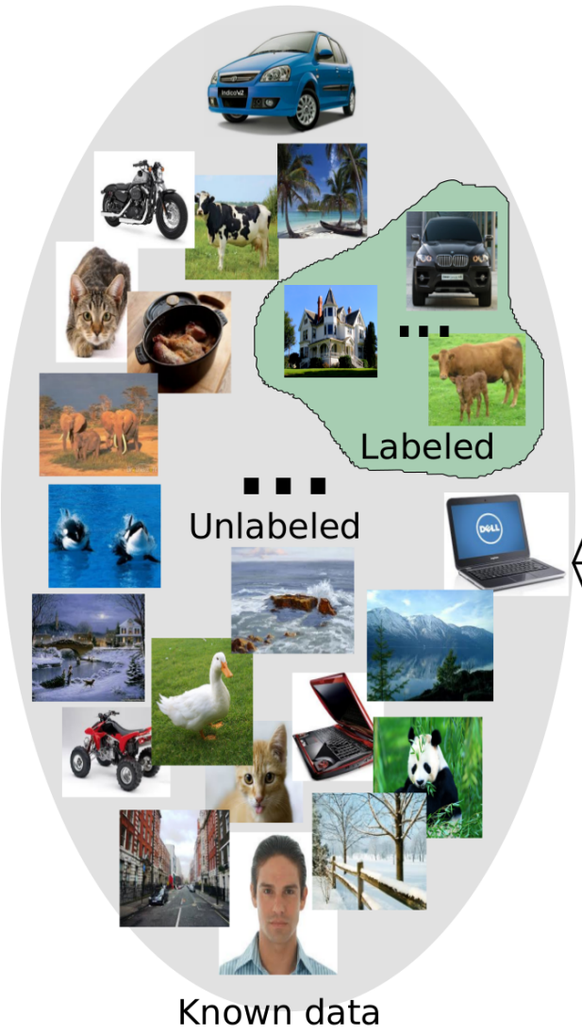
- An Analysis of Single-Layer Networks in Unsupervised Feature Learning - Andrew Ng - NIPS 2010

  - Introduced a single layer feature representation
  - Feature extraction process is unsupervised
  - Easy to understand system

Deep Architecture

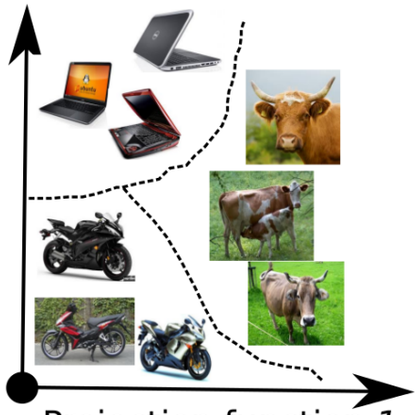
- Extract patches of data
- Learn feature-mapping using unsupervised learning algorithm
- Extract features
- Pool features together over regions of input data
- Train a linear classifier





Sampling

Training

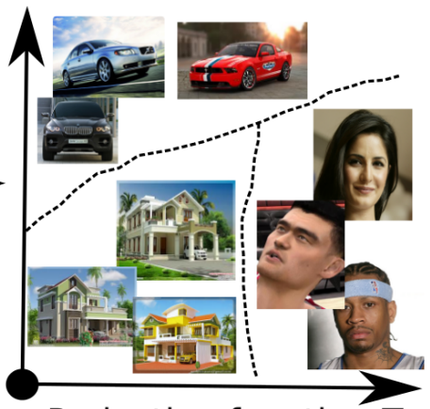



An image

Projecting


Vector *1*

Training



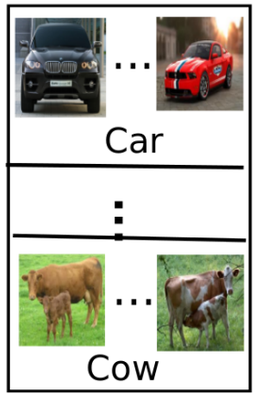
Projecting


vector *T*

Concat-enating


Projection vector

Feature learning



Labeled images

Training

Classifiers


An unlabeled image

Car

Classification with the learned feature

Application in Genomics

- Now imagine our data comes from Genomics
    - We start by sampling from raw data
    - Give those samples to an unsupervised algorithm e.g K-means
    - From the result of K-means we will form a new set of features
    - We use those features to represent our raw data in the supervised learning algorithm e.g SVM

Application in Genomics

- So why we do this :
    - We don’t need any labels to get Features
    - We reduce the size of dataset by a large factor (From 3.5 GB to 45 MB)
    - It’s scalable
    - We get better accuracy

Any Question?

A list of papers

- Representation Learning: A Review and New Perspectives, Yoshua Bengio, Aaron Courville, Pascal Vincent, Arxiv, 2012.
- The monograph or review paper Learning Deep Architectures for AI (Foundations & Trends in Machine Learning, 2009).
- Deep Machine Learning – A New Frontier in Artificial Intelligence Research – a survey paper by Itamar Arel, Derek C. Rose, and Thomas P. Karnowski.
- Graves, A. (2012). Supervised sequence labelling with recurrent neural networks(Vol. 385). Springer.
- Schmidhuber, J. (2014). Deep Learning in Neural Networks: An Overview. 75 pages, 850+ references, http://arxiv.org/abs/1404.7828, PDF & LATEX source & complete public BIBTEX file under http://www.idsia.ch/~juergen/deep-learning-overview.html.
- Building High-level Features Using Large Scale Unsupervised Learning Quoc V. Le, Marc'Aurelio Ranzato, Rajat Monga, Matthieu Devin, Kai Chen, Greg S. Corrado, Jeffrey Dean, and Andrew Y. Ng, ICML 2012.