Introduction to Deep Learning

Arijit Mondal
Dept. of Computer Science & Engineering
Indian Institute of Technology Patna
arijit@iitp.ac.in

Course structure

- Introduction to big data problem & representation learning
- Overview of linear algebra and probability
- Basics of feature engineering
- Neural network
- Introduction to open-source tools
- Deep learning network
- Regularization
- Optimization
- Advanced topics
- Practical applications

Evaluation policy

- Mid-sem - 20%
- Project - 40%-60%
- End-sem - 20%-40%
- Paper presentation - 10% (Depending on class size)

Project & Presentation

- Group wise project
- A group can have 2-3 students (Depending on class size)
- Each group will be assigned papers for presentation in the class
- Presentation duration $\sim$30 minutes

Books

- Deep Learning - Ian Goodfellow, Yoshua Bengio, Aaron Courville
- The Elements of Statistical Learning - Jerome H Friedman, Robert Tibshirani, Trevor Hastie
- Reinforcement Learning: An Introduction - Richard S Sutton, Andrew G Barto

Acknowledgement

- Deep Learning Book by Ian Goodfellow, Yoshua Bengio, Aaron Courville
- Presentation by Yann LeCun, Geoff Hinton, Yoshua Bengio
- Various websites for images
- Dr. Jacob Minz (Synopsys)
- IIT KGP Batch of 2001
  - Joydeep Acharya (Hitachi)
  - Sanjeev Kumar (Liv.AI)
  - Mithun Dasgupta (Microsoft)
  - Amit Kumar (Avnera)
  - Mrinmoy Ghosh (Facebook)
  - Animesh Datta (Qualcomm)
  - Bhaskar Saha (PARC)
  - Banit Agrawal (Facebook)

Introduction

Problem Solving Strategies for Big Data

- Need to solve problems efficiently and accurately when the input data is huge ($\sim$ GB, TB order)
- Finding a deterministic algorithm is difficult
  - Need to find out features
  - Requires significant effort for model building
  - Need to have domain knowledge
- Statistical inference is found to be suitable
  - Feature selection is not crucial
  - Model will learn from past data

Applications: Computer vision

- 2d to 3d conversion
- Street view generation
- Image classifications
- Image segmentation

Image source: Internet

Applications: Activity Recognition

- Recognize activities like walking, running, cooking, etc. from still image or video data

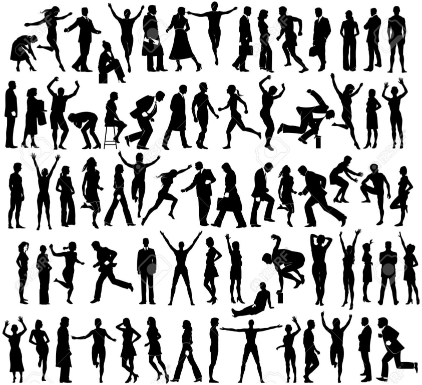


Image source: Internet

Applications: Image Captioning

- Automated caption generation for a given image

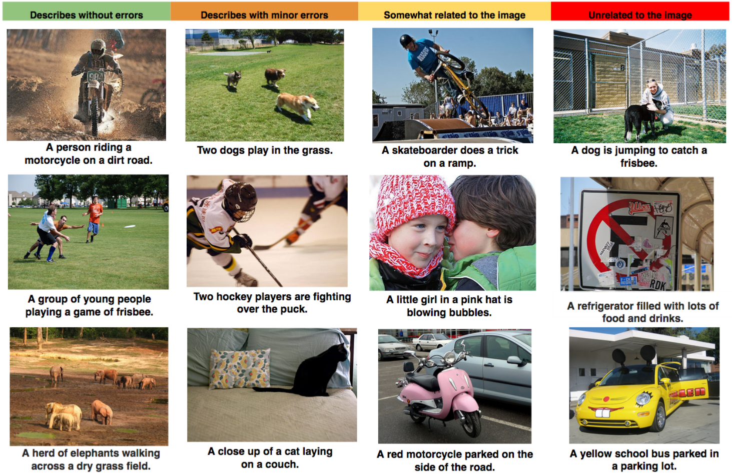


Image source: Internet

Applications: Object Identification

- Identify objects in still image or in video stream

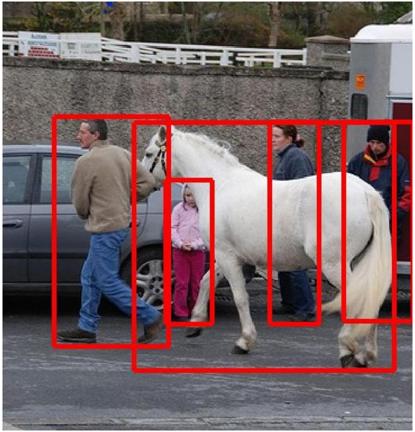


Image source: Internet

Applications: Automated Car

- Self driving car

Image source: Internet

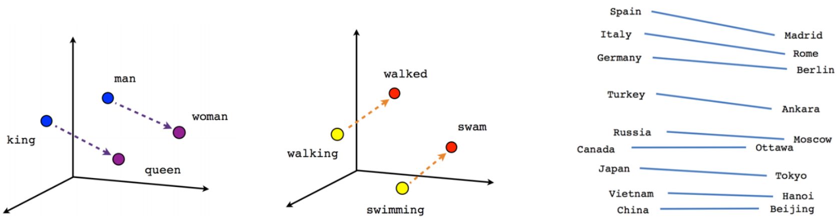
Applications: Drones & Robots

- Managing movement of robot or drones

Image source: Internet

Applications: Natural Language Processing

- Recommender system
- Sentiment analysis
- Question answering
- Information extraction from website
- Automated email reply



Male-Female Verb tense Country-Capital

Image source: Internet

Applications: Speech processing

- Conversion of speech into text
- Generation of particular voice for the given text

Image source: Internet

Other possible applications

- Write a story/text and generate a video/image of it
- Conversion of speech from one language to another language in real time
- Weather prediction
- Genomics
- Drug discovery
- Particle physics

Issue of Representation

- Representation of data in an efficient/structured manner is crucial for solving problems more effectively
  - Searching of a set of elements in a given list (sorted/unsorted)
  - Arithmetic operations on Arabic and Roman numerals
  - Primality test of n when n is represented as $11111\ldots111$ (n-number of one)
- Structured representation can help in predicting future values

Learning representation/feature

- Traditional approaches
  - Pattern recognition
    - Input, output of the problem
- End to end learning
  - System automatically learns internal representation

AI-ML Tasks

- Heavily depends on features
- Requires good domain knowledge
- Feature extraction is not easy job
    - Identify a car
        - How to describe wheel
        - Shadow/brightness
        - Obscuring element

Representation Learning

- Learned representation often result in better performance compared to hand design
- Allows the system to rapidly adapt to new task
- Need to discover a good set of features
- Manual design of features is nearly impossible

Design of Features

- Goal is to separate out variation factors
- These factors are separate sources of influence
- It may exist as unobserved object or unobserved forces that affect observable quantity
    - Speech - Factors are age, sex, accent, etc
    - Image - Position, color, brightness, etc.

Deep Learning

- Try to address the problem of representation learning
- Representation are expressed in terms of other simpler representation
- Develop complex concept using simpler concept

Simple to Complex Features

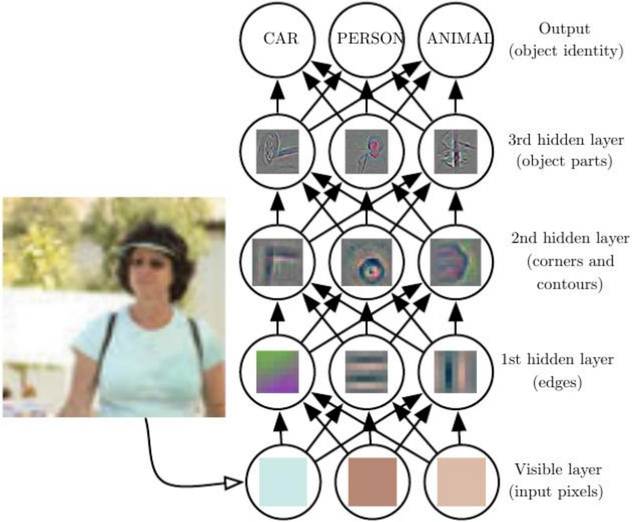


Image source: Deep Learning Book

Simple to Complex Features

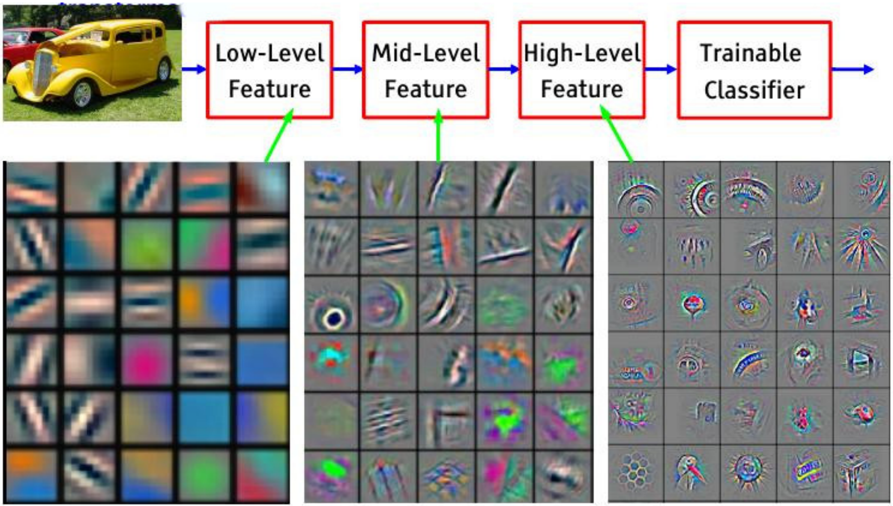


Image source: Deep Learning Tutorial by Yann LeCun Marc'Aurelio Ranzato, ICML, 2013

Conventional Machine Learning

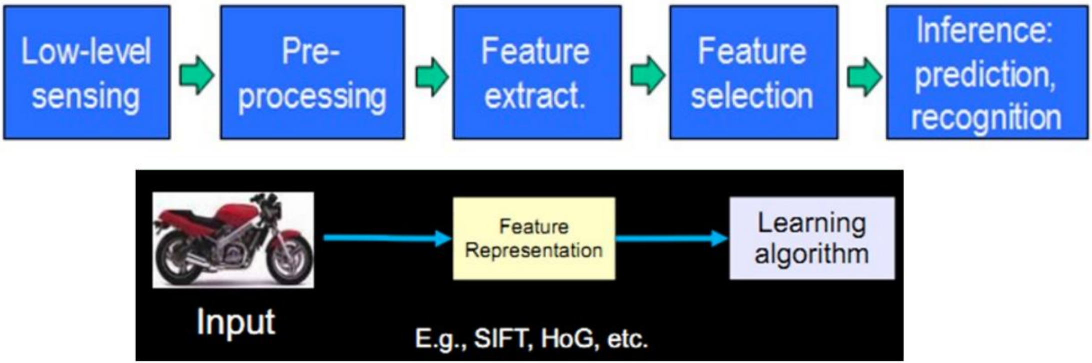
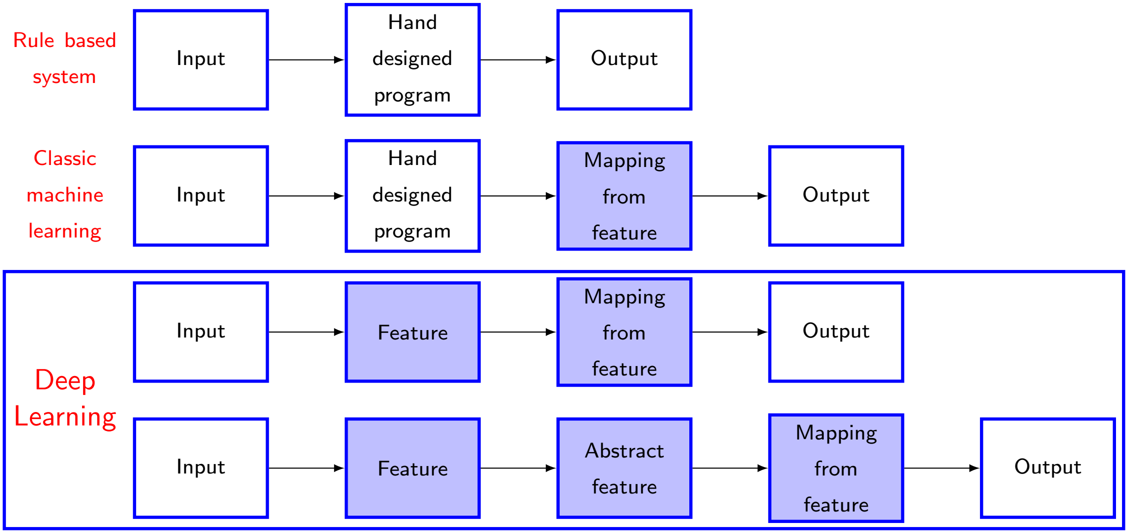


Image source: Deep Learning by Yann LeCun, Yoshua Bengio & Geoffrey Hinton

Deep Learning Model

- Feed forward deep network or multilayer perceptron
- Mathematical functions that map input to output
- Composed of simpler functions
- Each layer provides a new representation
- Learning right representation

Representation learning



Depth of network

- Number of sequential instruction must be executed to evaluate the architecture
  - Length of the longest path
- Depth of the model

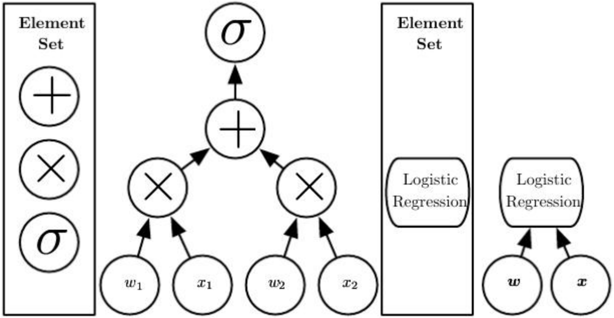


Image source: Deep Learning Book

History

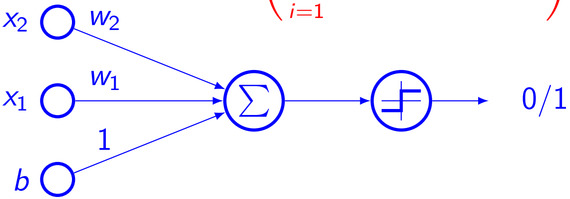
- Has many names and view point
    - Cybernetics (1940-1960)
    - Connectionism (1980-1990) (neural net)
    - Deep learning (2006+)
- More useful as the amount of data is increased
- Models have grown in size as increase in computing resources
- Solving complex problem with increasing accuracy

Learning Algorithm

- Early learning algorithm
  - How learning happen in brain?
  - Computational model of biological learning
- Neural perspective of DL
  - Brains provide a proof by example
  - Reverse engineer the computational principle behind the brain and duplicate its functionality

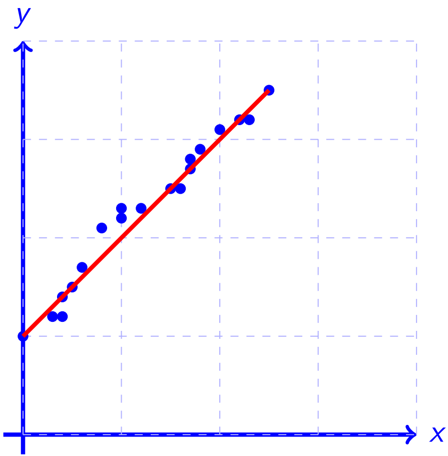
History of basic model

- The first learning machine: the Perceptron
  - Built at Cornell, 1960
- The perceptron was simple linear classifier on top of simple feature extractor
- Most of the practical applications of ML today use glorified linear classifiers or glorified template matching.
- Significant effort is required from the expert for identifying relevant features
- Typically it will solve $y = sign\left(\sum_{i=1}^{N}(w_i \times f_i(X) + b)\right)$

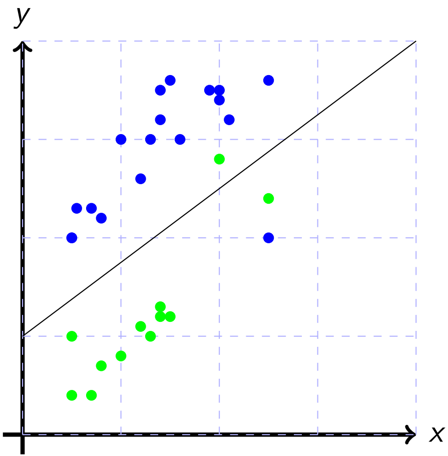


Broad Categories of Problem

- Regression

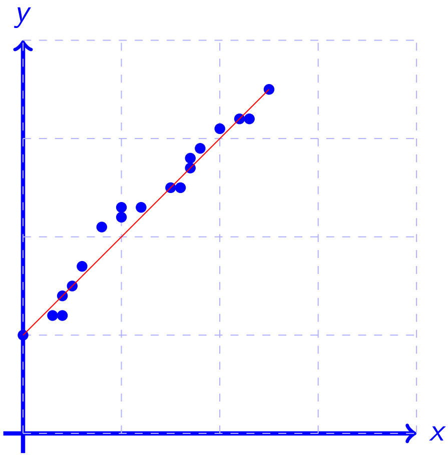


- Classification

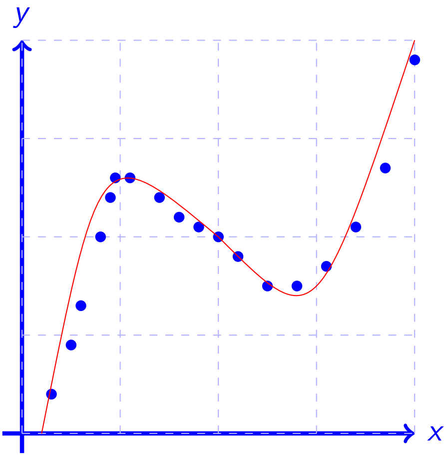


Regression

- Regression (linear)

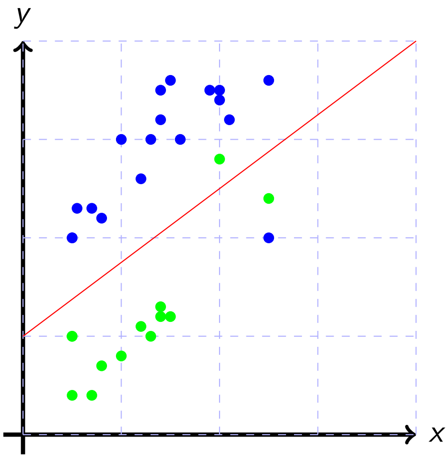


- Regression (Non-linear)

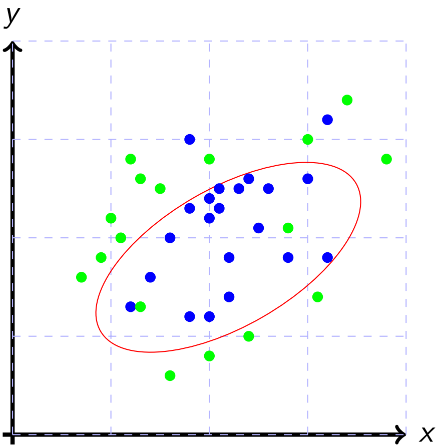


Classification

- Linear

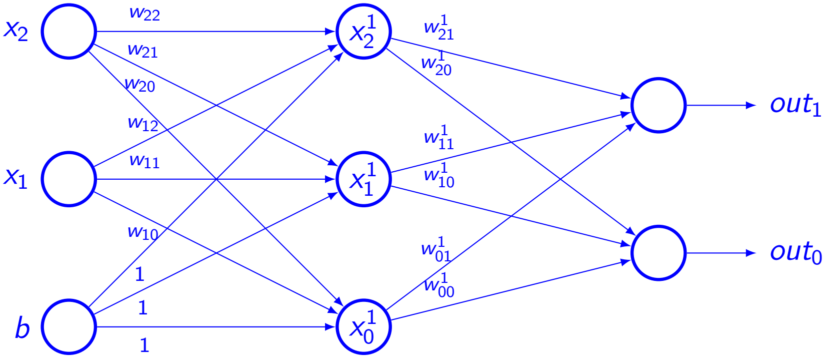


- Non-linear

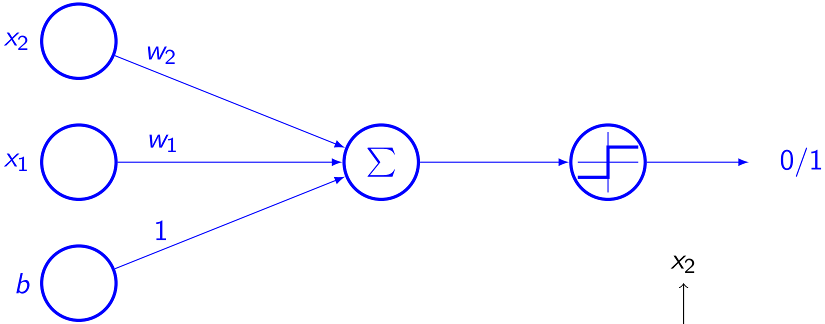


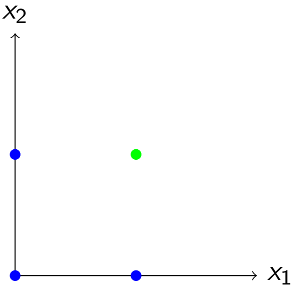
Artificial Neural Network

- A simple model

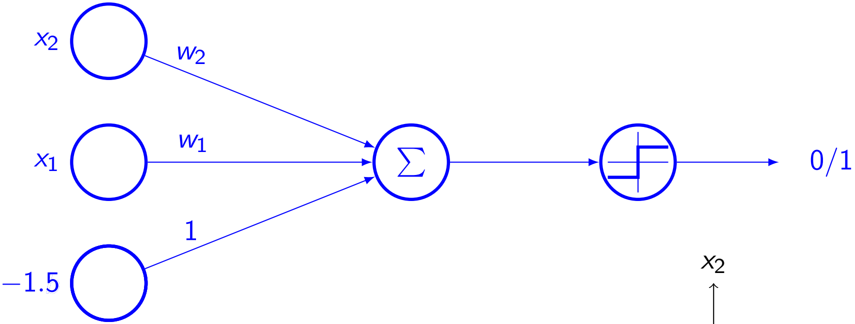


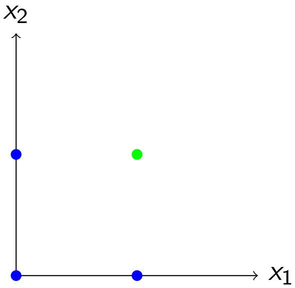
Example NN: AND gate



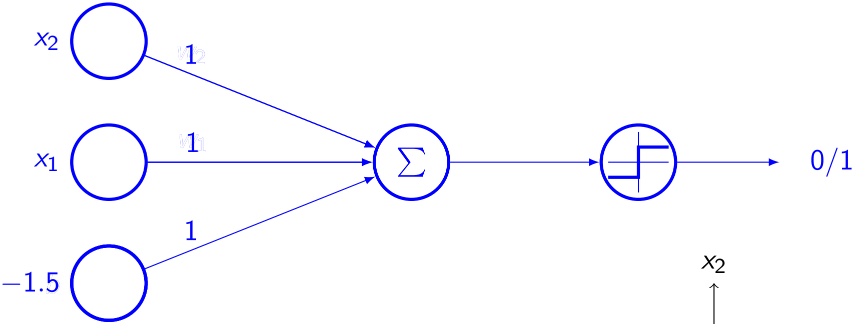


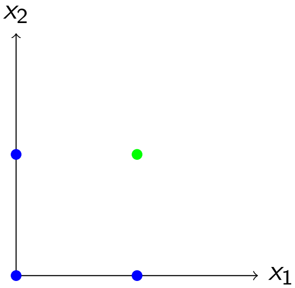
Example NN: AND gate



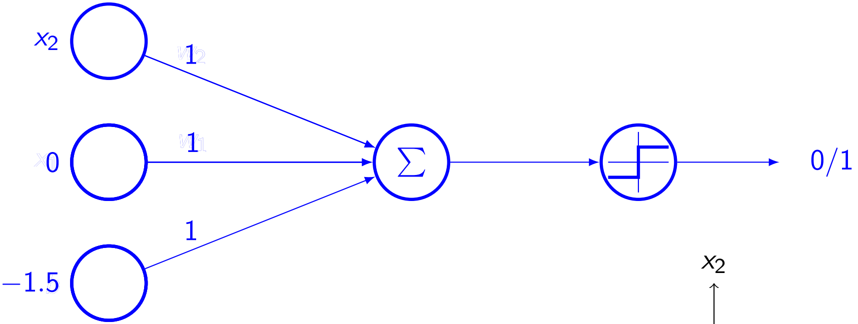


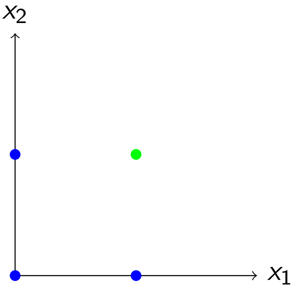
Example NN: AND gate



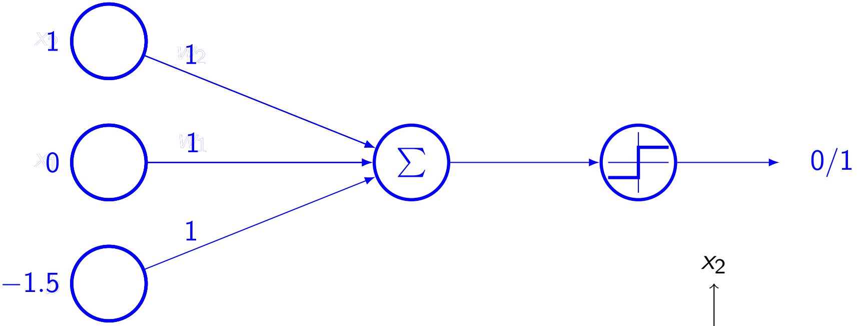


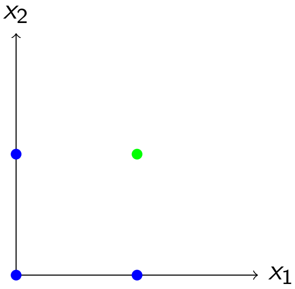
Example NN: AND gate



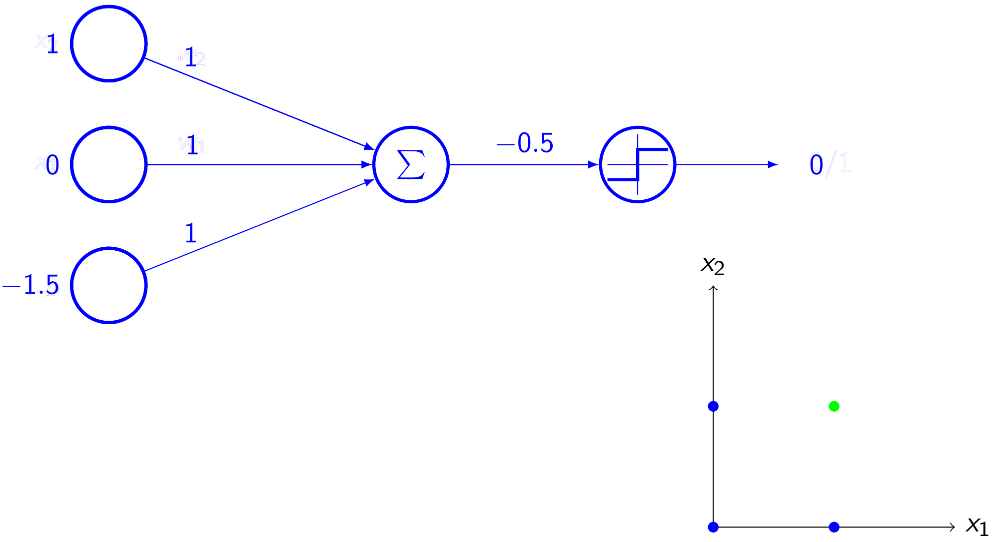


Example NN: AND gate

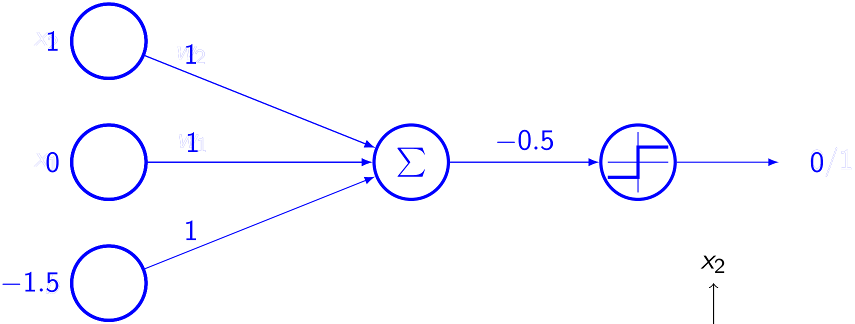


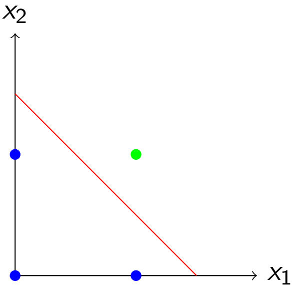


Example NN: AND gate

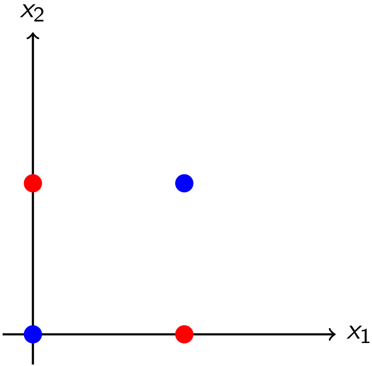


Example NN: AND gate

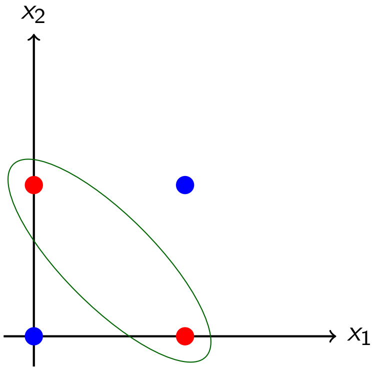




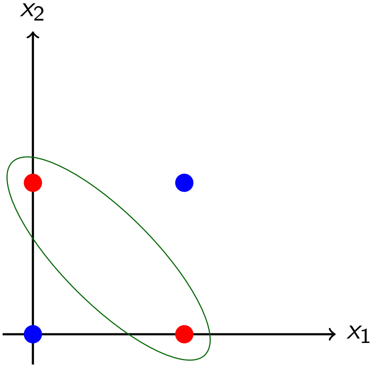
Example NN: XOR gate

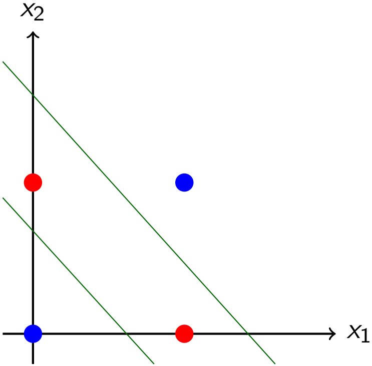


Example NN: XOR gate



Example NN: XOR gate





Distributed representation

- Each input should be represented by many features
- Each feature should be involved in the representation of many possible inputs
- Example: car, flower, birds — red, green, blue
  - 9 neurons
  - For each combination of color and object
- Distributed neurons
  - 3 Neurons for color
  - 3 Neurons for object
  - Total 6 neurons

Popularization of Neural Network

- Most of the theory of neural network was developed in the 1980s
- Started gaining popularity around 4-5 years ago
    - Geoffrey Hinton and Alex Krizhevsky winning the ImageNet competition where they beat the nearest competitor by a huge margin (2012)

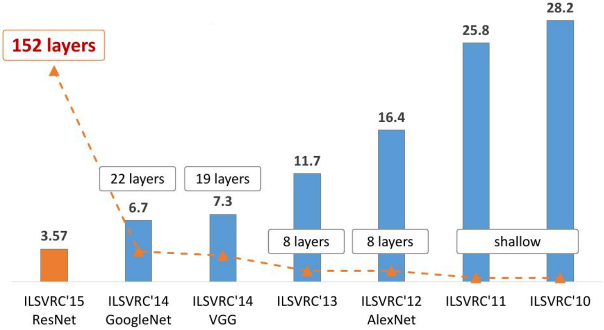


Image source: Deep Residual Learning by Kaiming He, et.al.

Popularity

- Increase data size
  - Computing resources are available
  - Accepting performance 5000 labeled example per category
  - 10 million for human performance
- Increasing model size
- Increasing accuracy, complexity, real world impact
- Used by many companies
  - Google, Microsoft, Facebook, IBM, Baidu, Apple, Adobe, Nvidia, NEC, etc.
- Availability of good commercial & open-source tools
  - Theano, Torch, DistBelief, Caffe, TensorFlow, Keras, etc.