The Abdus Salam International Centre for Theoretical Physics

Introduction to machine learning and neural networks

Professor Clement Onime (onime@ictp.it)

1. *International Centre for Theoretical Physics (ICTP), Trieste, Italy*

Objectives

Objectives

- **Demystify Machine Learning**
  - **Relationship to Artificial Intelligence (A.I)**
  - **Learning algorithms and methodology**
  - **Specialised systems**
    - **Recommender systems**
    - **Neural networks**
- **Examples and Applications**

DEMYSTIFY

Introduction

- AI at work
- New capability for computers

Examples:

- Database mining
  Large datasets from growth of automation/web.
  E.g., Web click data, medical records, biology, engineering
- Applications can't program by hand.
  E.g., Autonomous helicopter, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision.
- Self-customizing programs
  E.g., Amazon, Netflix product recommendations
- Understanding human learning (brain, real AI).

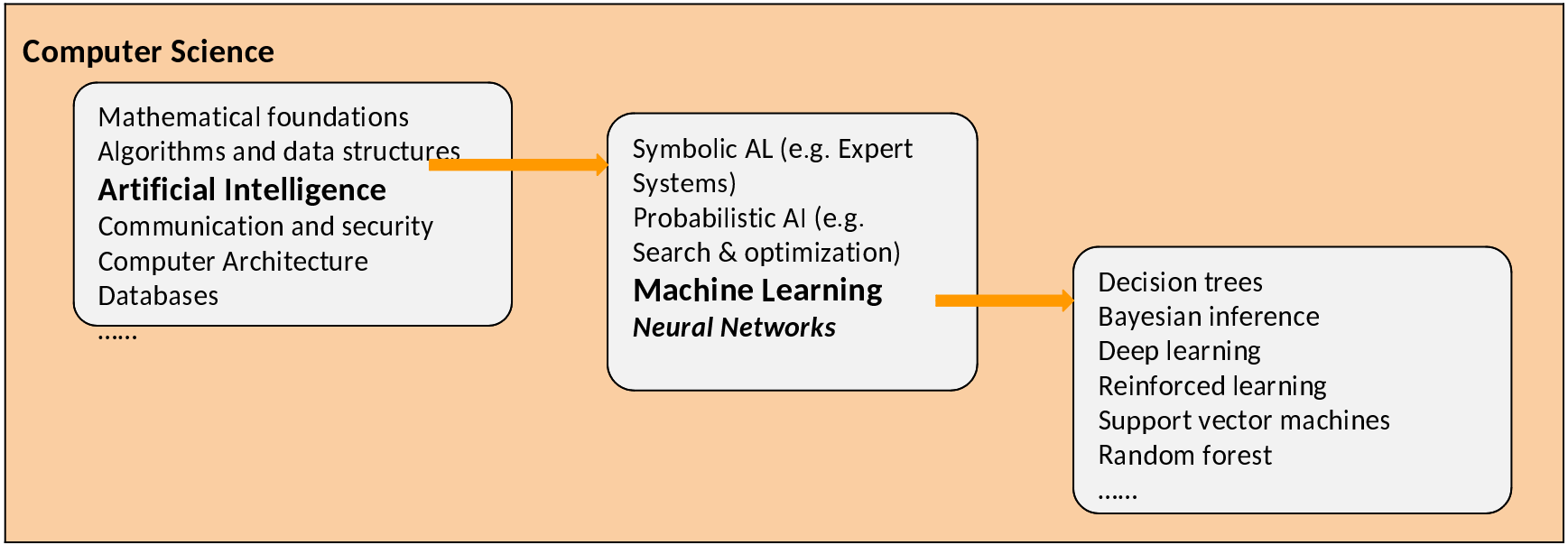
Artificial Intelligence: approaches

- Example - Excelling at playing the game of chess

Symbolic AI	Mathematical/Statistical AI	Machine Learning Approach
"Let us sit down with the world's best chess player and put his/her knowledge into a computer program"	*"Let us simulate all the different possible moves and the associated outcomes at each single step and go with the most likely to win"*	**"Let us show millions of examples or real life and simulated games (won and lost) to the program, and let it learn from experience"**

From AI to ML

- Artificial Intelligence (AI) is a branch or Computer Science that uses algorithms and techniques to mimic human intelligence
- Machine Learning (ML) is one of several AI techniques for sophisticated cognitive tasks



Machine Learning paradigm

- Machine Learning is a particularly interesting technique because it represents a paradigm shift within AI

Traditional AI techniques

- **Static** – hard-coded set of steps and scenarios
- **Rule Based** – expert knowledge
- **No generalization** – handling special cases is difficult

---

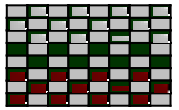
Machine Learning

- **Dynamic** – evolves with data, finds new patterns
- **Data driven** – discovers knowledge
- **Generalization** – adapts to new situations and special cases

Machine Learning definition

Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.

Tom Mitchell (1998) Well-posed Learning Problem: A computer program is said to *learn* from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.



Quick Example: One man’s ham is another man’s spam

"A computer program is said to *learn* from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T in this setting?

- Classifying emails as spam or not spam. T
- Watching you label emails as spam or not spam. E
- The number (or fraction) of emails correctly classified as spam/not spam. P
- None of the above—this is not a machine learning problem.

ML Basics

HOW MACHINES LEARN

Machine learning algorithms:

- Supervised learning
- Unsupervised learning
- *Others:*
  - *Reinforcement learning*

Supervised and Unsupervised Learning

Supervised Learning:

Predicting values. **Known** targets.

User inputs correct answers to learn from. Machine uses the information to guess new answers.

REGRESSION:
Estimate continuous values
(Real-valued output)

CLASSIFICATION:
Identify a unique class
(Discrete values, Boolean, Categories)

Unsupervised Learning:

Search for structure in data. Unknown targets.

User inputs data with undefined answers. Machine finds useful information hidden in data

CLUSTER ANALYSIS:
Group into sets

DENSITY ESTIMATION:
Approximate distribution

DENSITY REDUCTION:
Select relevant variables

Supervised and Unsupervised Learning

Supervised Learning:

Regression

- Linear Regression
- Ordinary Least Squares Regression
- LOESS (Local Regression)

Classification

- Decision Trees
- K-Nearest Neighbors
- Support Vector Machine
- Logistic Regression
- Naïve Bayes
- Random Forests9

Unsupervised Learning:

Cluster Analysis

- K-Means Clustering
- Hierarchical Clustering

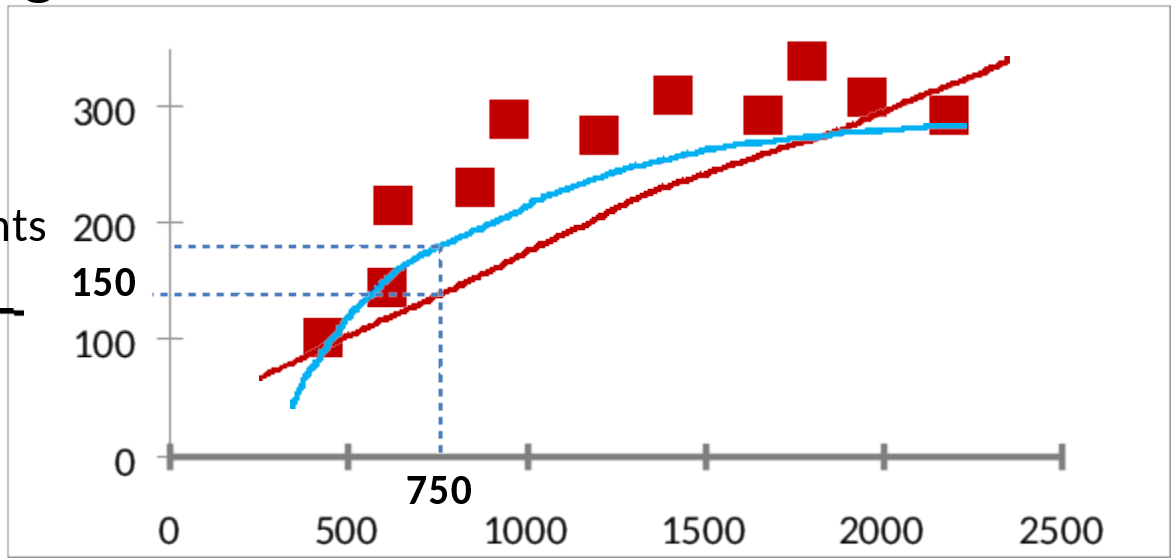
Dimension Reduction

- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)

After training

SAMPLE APPLICATIONS

Linear regression

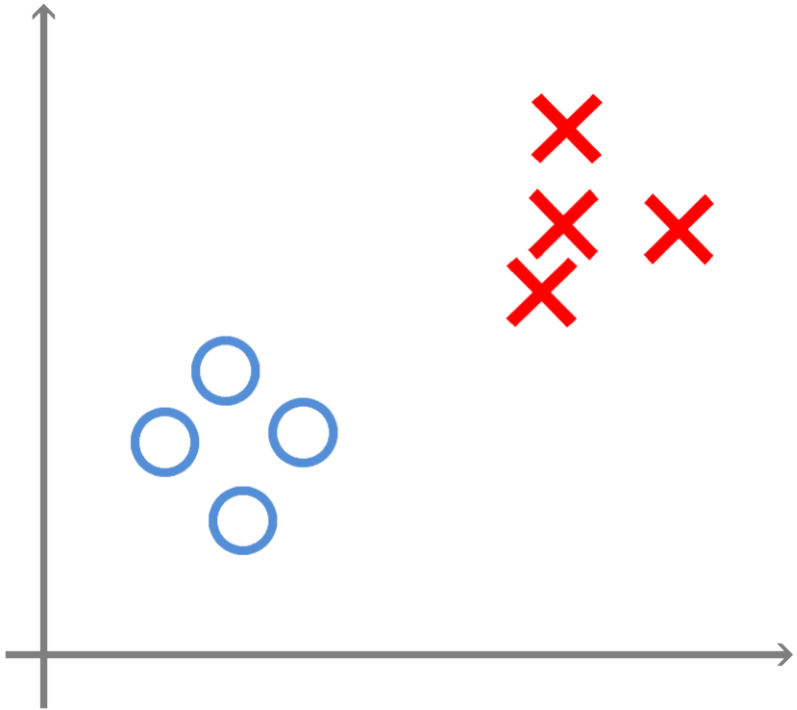


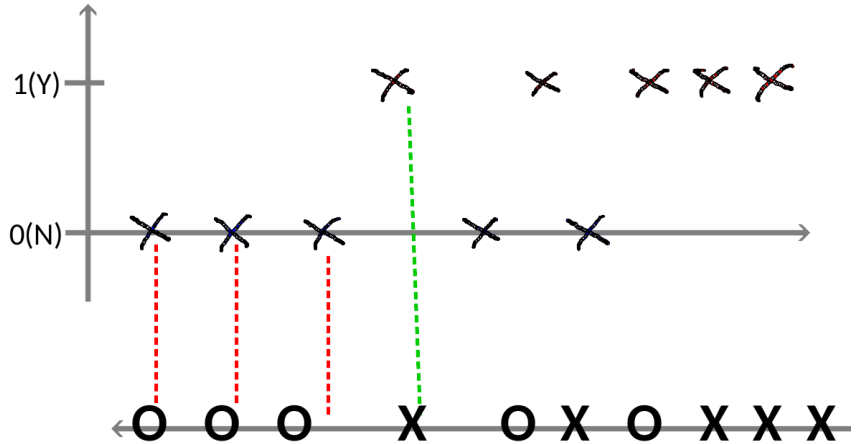
Supervised Learning

"right answers" given for training

Regression: Predict continuous valued output

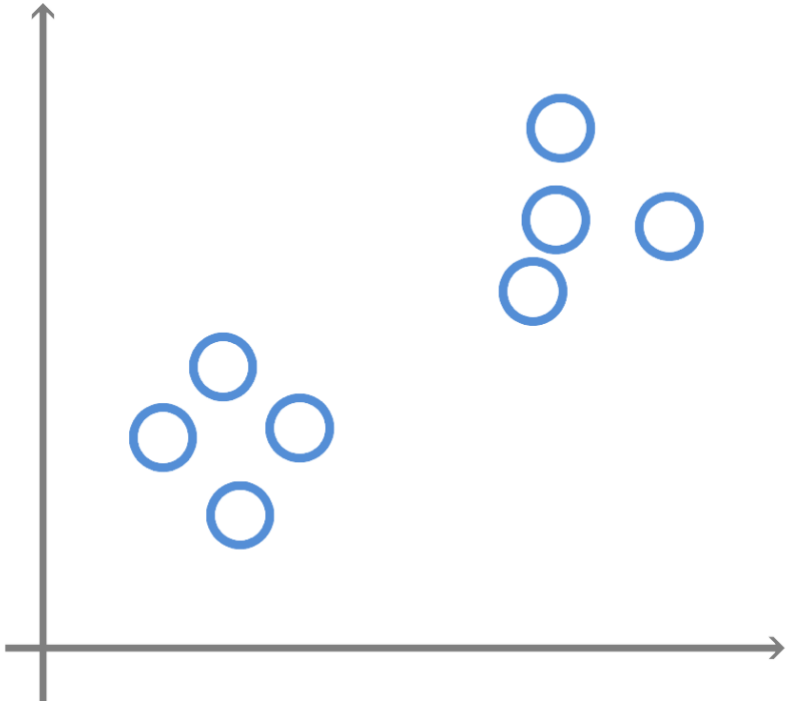
Cassification

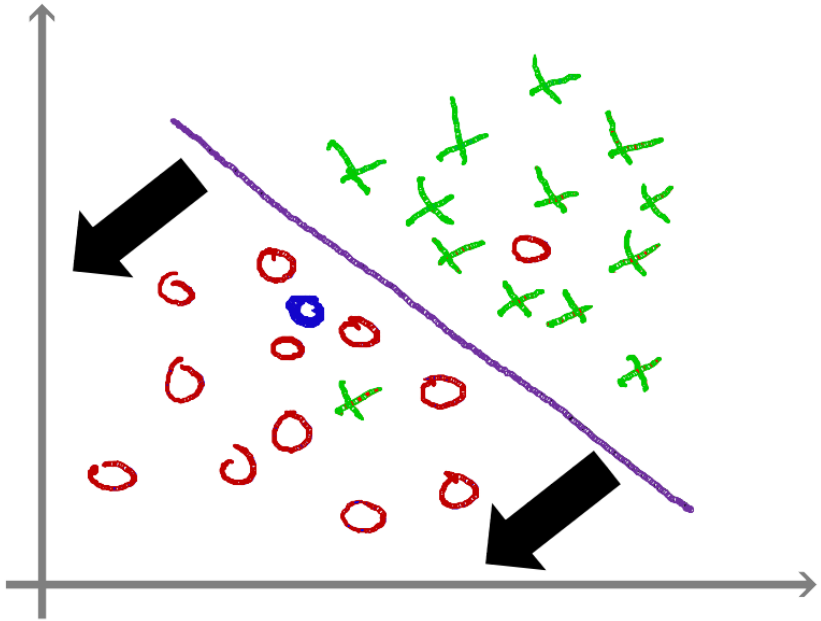




Diffrentiate betwen circle(s) and x(s); dogs & cats; cancer & non-cancer, etc...

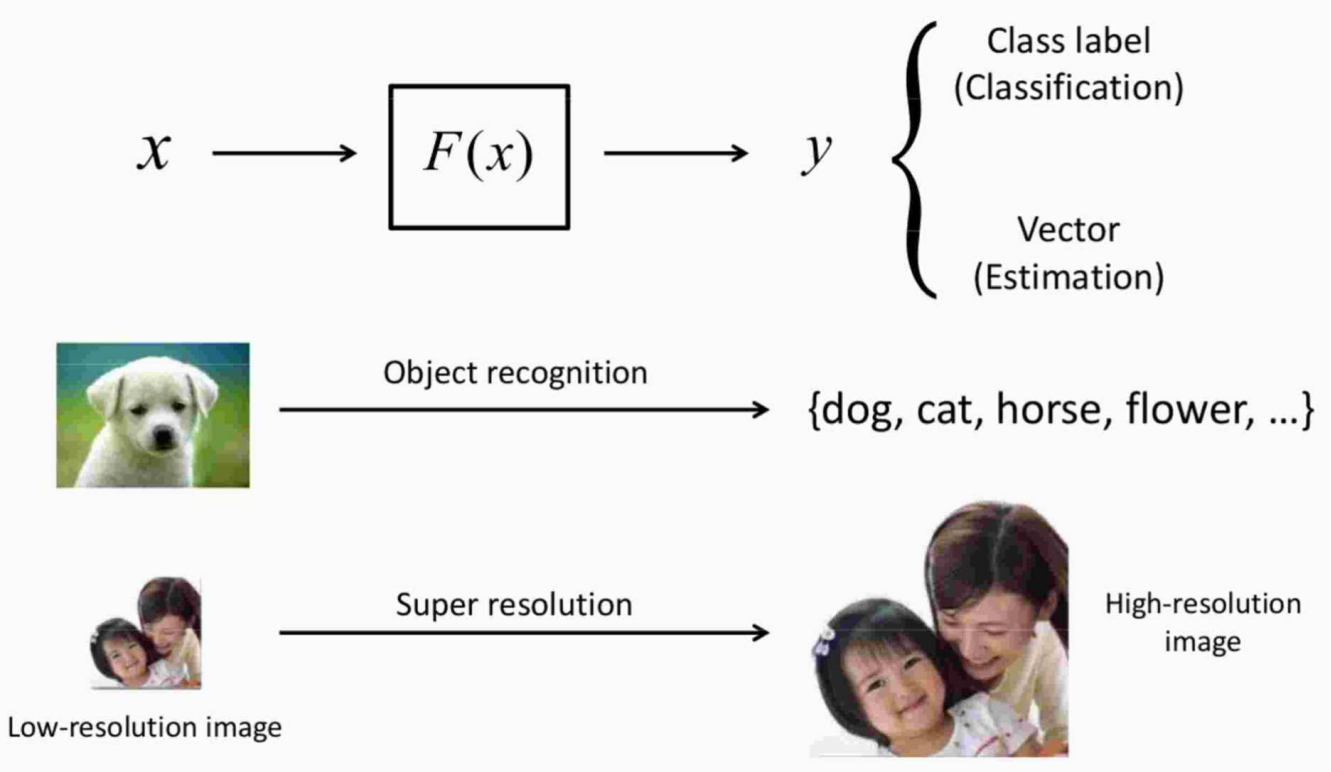
Unsupervised learning





ML can quickly identify and corelate using properties such as Clump size, uniformity of sizes, shapes, etc..

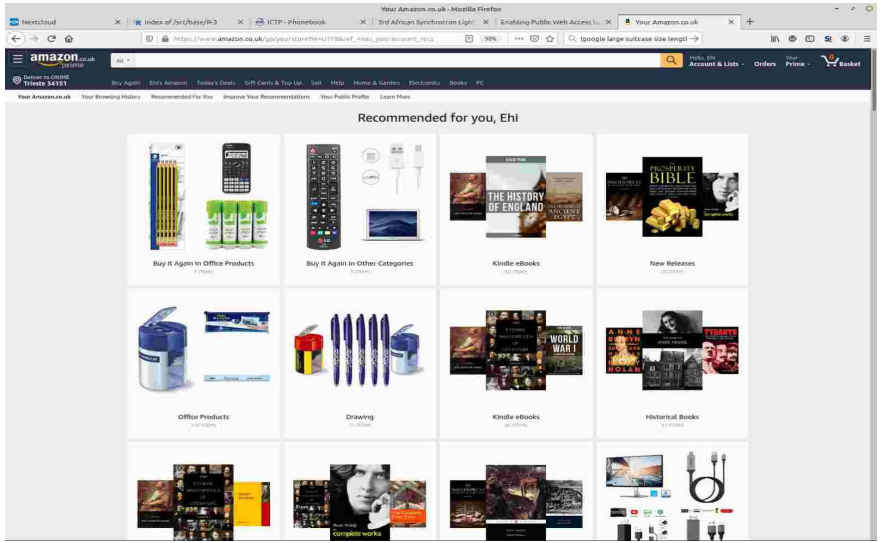
Advanced applications



After training

SPECIALIZED ML SYSTEMS

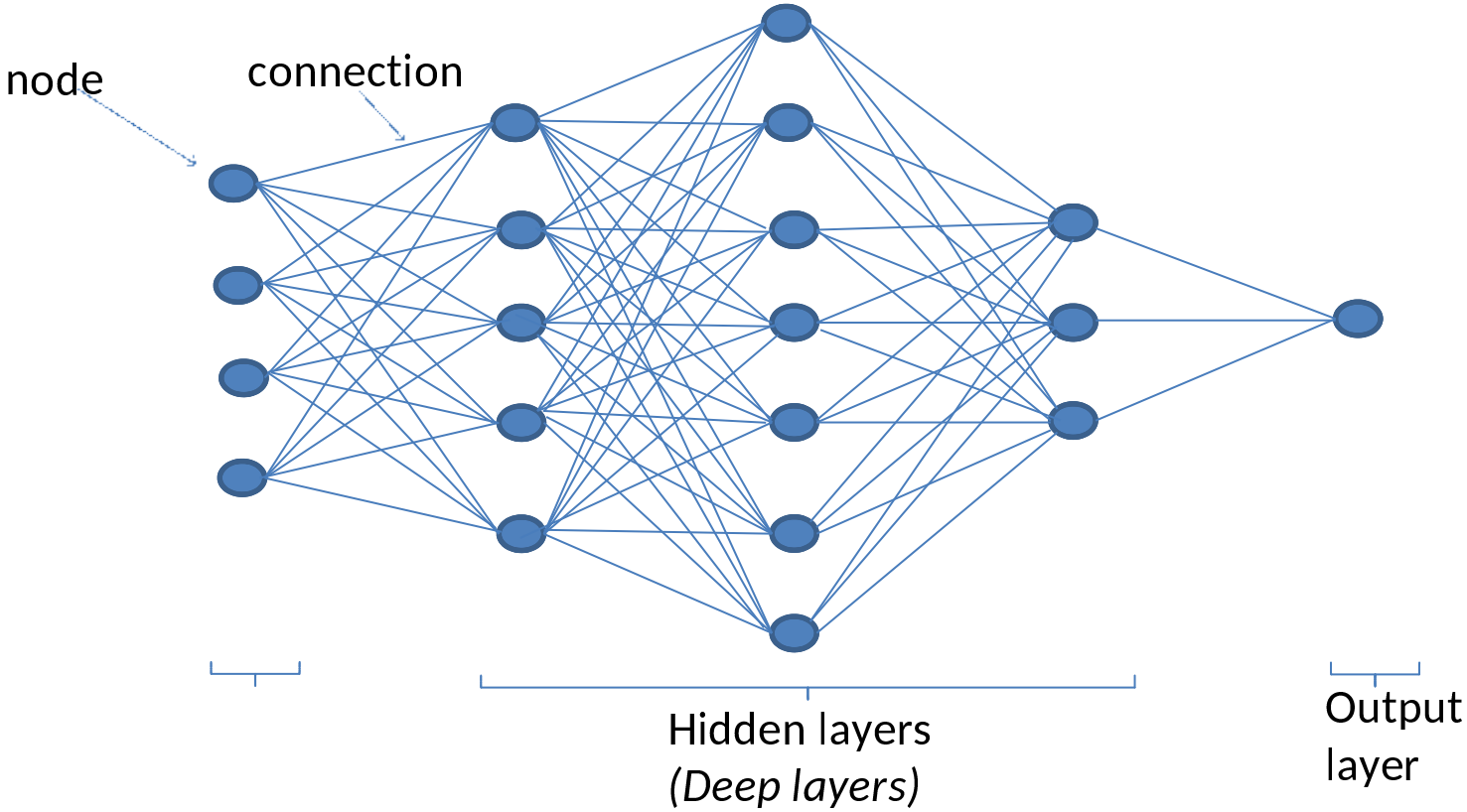
Recommender systems



- Special class of ML systems
  - Does not require lots of data to start & gets better with more data
  - Content/item or user based filtering
- Similarities
- Euclidean distances

Used by shopping and entertainment platforms

Artificial Neural Network (ANN)



Importance of Deep Learning

(Neural Networks)

Techniques	Accuracy	Detection Rate (Precision)	False Alarm Rate
Support Vector Machine (SVM)	94.65%	85.45%	5.2%
Artificial Neural Networks (ANN)	99.71%	99.68%	0.12%
Bayesian Network	97.52%	97.04%	2.50%
K- Nearest Neighbour (KNN)	97.15%	96.84%	2.88%
Fuzzy Logic Based System	95.2%	86.84%	1.15%
Decision Trees	97.93%	98.52%	2.19%
Logistic Regression	94.7%	77.8%	2.9%

"A Comparative Analysis of Various Credit Card Fraud Detection Techniques" by Yashvi Jain, NamrataTiwari, ShripriyaDubey,Sarika Jain International Journal of Recent Technology and Engineering (IJRTE)ISSN: 2277-3878, Volume-7 Issue-5S2, January 2019

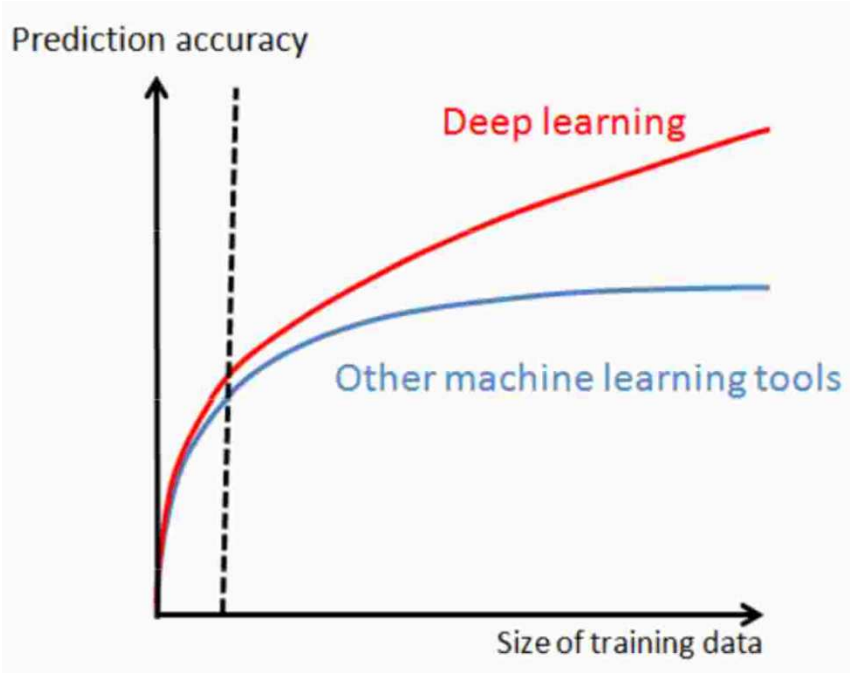
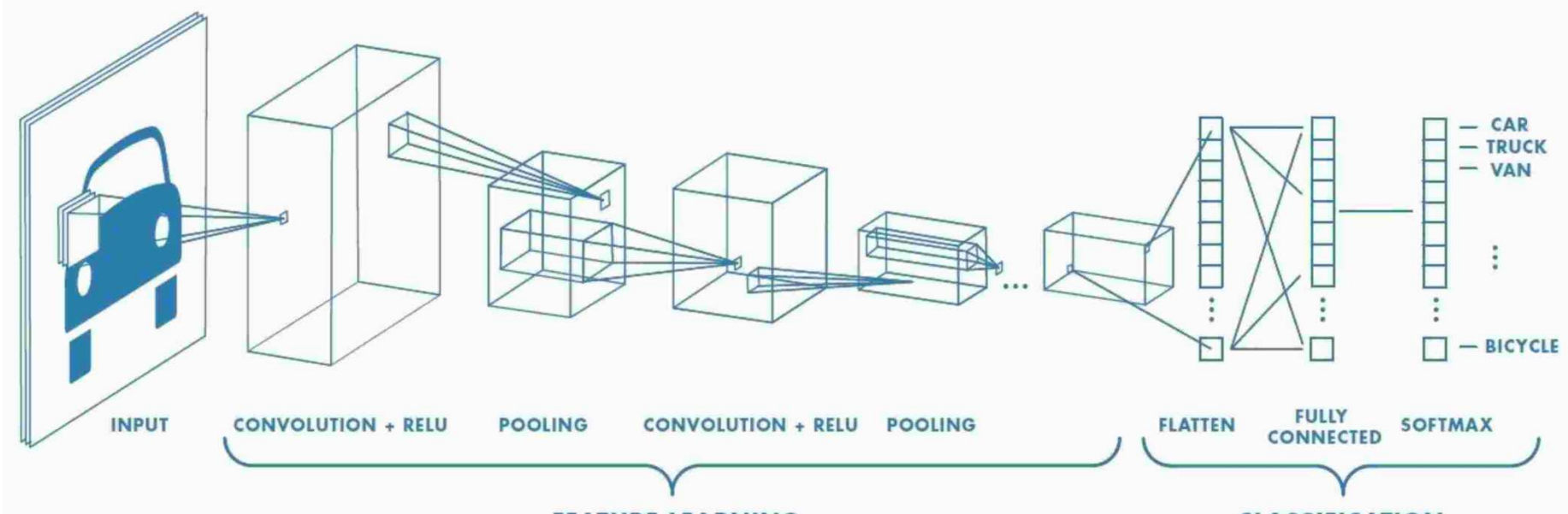
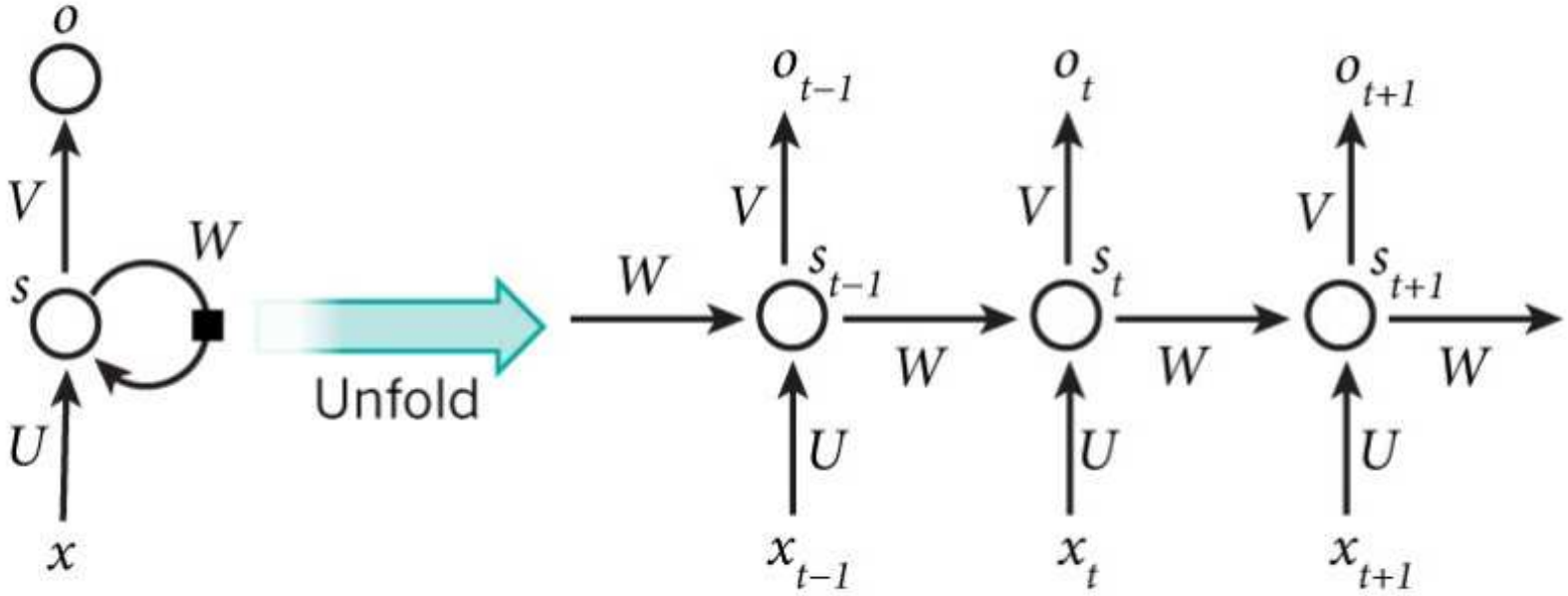


Image from "Deep Learning in Computer Vision" by Xiaogang Wang

Convolutional Neural Network (CNN)

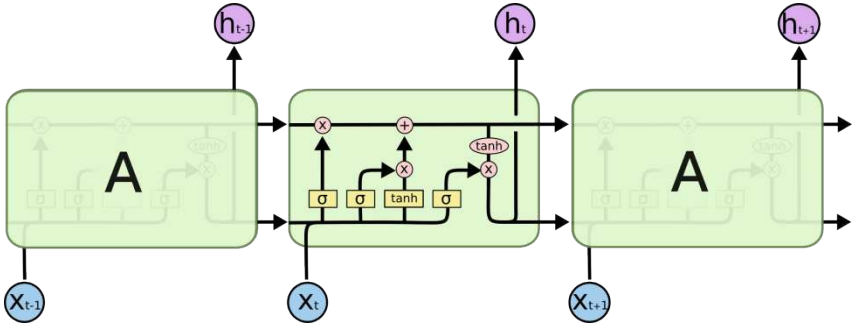


Reccurent Neural Network (RNN)



Long Short Term Memory (LSTM)

- RNNs are composed of a single neural networks
- LSTMs specialised RNNs have four neural networks iinteracting in a specialized way.
  - 3 sigmoids and a tanh
- LSTM work by protecting stored states.



NOTES

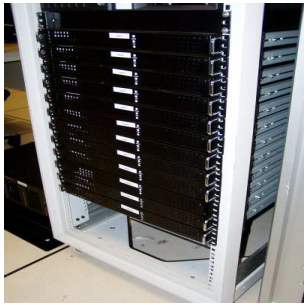
Features in ML

A Feature:

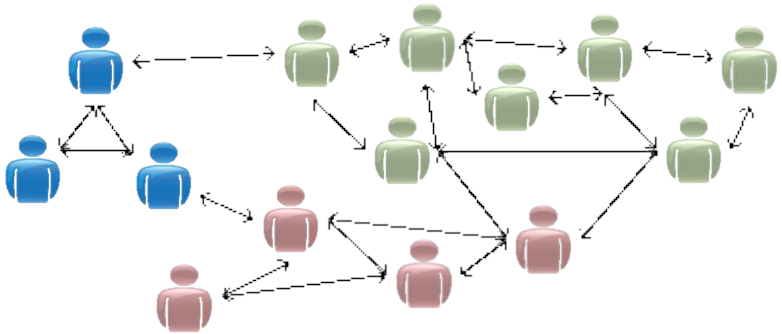
- a data attribute used to make a comparison
  - Quantify attributes of an object (size, weight, color, shape, density) in a way a computer can understand
- Quality is important
  - A good feature helps discriminates between classes

The correct identification and engineering of Features goes a long way in a sucessful ML model.

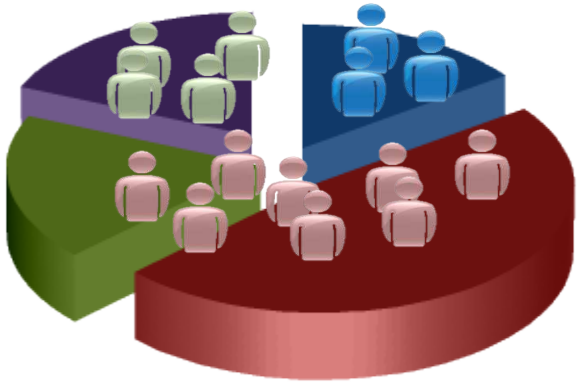
Choosing the right ML algorithm depends on the give task/data



Organize computing clusters



Social network analysis



Market segmentation

Analysis of Astronomy data

Hints for choosing algorithms

- Regression task
  - Linear, Logistic, Lasso, Ridge, Bayesian
- Clustering task
  - K-Means, DBSCAN, GMM
- Classifications task
  - Naïve Bayes, K-Nearest Neighbors, Support Vector Machines
- Images/video data
  - Convolutional Neural Networks, Deep Learning
- Time series data
  - Recurrent Neural Networks, Long Short Term Memory (LSTM)
- Expert Knowledge base
  - Recommender Systems

References

1. Fundamentals of Machine Learning: *Lecture(s) by Asst. Prof. Solomon Gizaw, Director Computational Data Science Graduate studies Programme, Addis Ababa University, Ethiopia and Dr. Ekpe Okorafor, African University of Science and Technology (Abuja, Nigeria) and Accenture Big Data Centre, Texas, U.S.A*

2. https://developers.google.com/machine-learning/crash-course/ml-intro

3. https://see.stanford.edu/Course/CS229/47

4. A Comparative Analysis of Various Credit Card Fraud Detection Techniques, Yashvi Jain, NamrataTiwari, ShripriyaDubey,Sarika Jain, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-7 Issue-5S2, January 2019