INTRODUCTION

Deep Neural Networks

Session 01

Pramod Sharma
pramod.sharma@prasami.com

2

Agenda

- Introduction
- AI vs ML vs Deep Learning
- MP Neuron
- Coding Exercise

3

References

- Deep Learning, Ian Goodfellow, Yoshua Bengio, Aaron Courville
- Neural Networks and Learning Machines, Simon Haykin
- Pattern Recognition and Machine Learning, Christopher M. Bishop
- Deep Learning with Python - François Chollet
- Hands-On Machine Learning with Scikit-Learn and TensorFlow
- TensorFlow Deep Learning Cookbook
- Reinforcement Learning with TensorFlow: A Beginner's Guide to Designing Self-learning Systems with TensorFlow and OpenAI Gym Sayon Dutta
- Hands-On Reinforcement Learning with Python: Master Reinforcement and Deep Reinforcement Learning Using OpenAI Gym and TensorFlow Sudharsan Ravichandiran
- Deep Reinforcement Learning Hands-On: Apply Modern RL Methods, with Deep Q-networks, Value Iteration, Policy Gradients, TRPO, AlphaGo Zero and More Maxim Lapan

pra-sâmi

4

"Theory exam– 40% weightage

Lab exam – 40% weightage

Internal exam – 20% weightage"

Evaluation method

pra-sâmi

5 Logistics

pra-sâmi

6 Logistics

- We encourage 'discussion groups':
  - Study groups
  - Whatsapp groups
- Expect you to complete your assignment individually!
- No group assignments unless stated otherwise
- Code is small part of it
- Pay special attention to **inline comments**
  - Comments should focus on what you were trying to implement

pra-sâmi

7

Journey So far

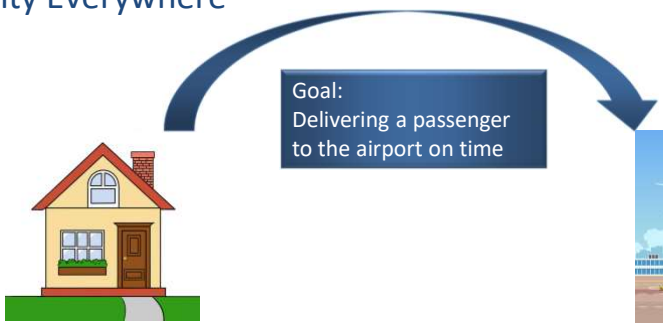
Search → Search and Inference → Uncertainty → Bayesian Network

Listen → Reflect → Implement → Share

pra-sâmi

8

Uncertainty Everywhere



- The agent forms a plan, lets say… A90,
  - Leave home 90 minutes before the flight departs
  - Driving at a reasonable speed
- Are you certain "*Plan A90 will get us to the airport in time*."?
  - Not in absolute sense but with some riders

pra-sâmi

9

Uncertainty Everywhere

- How about other plans, such as A180,
  - Might increase the agent's belief that it will get to the airport on time,
  - But also increase the likelihood of a long wait
- Probability is an agent's measure of belief in some proposition — subjective probability.
- An agent's belief depends on its prior belief and what it observes.

10

Agent In Uncertain Environment

- Agents don't have complete knowledge about the world.
- Agents need to make (informed) decisions given their uncertainty.
- It isn't enough to assume what the world is like.
  - Example: wearing a seat belt.
- An agent needs to reason about its uncertainty.
- When an agent takes an action under uncertainty, it is gambling ⇒ probability

11

Overview

- Nature is a continuum where as math is discreet values
  - Old film based images were continuous painting of colors where as digital images are pixels
- Brain works differently than our mathematical computations
- Brain is highly complex, nonlinear and parallel computer
- Neural networks are supposed to be inspired fro
- Highly generalized form, a Neural N nulate manner in which brain performs a task

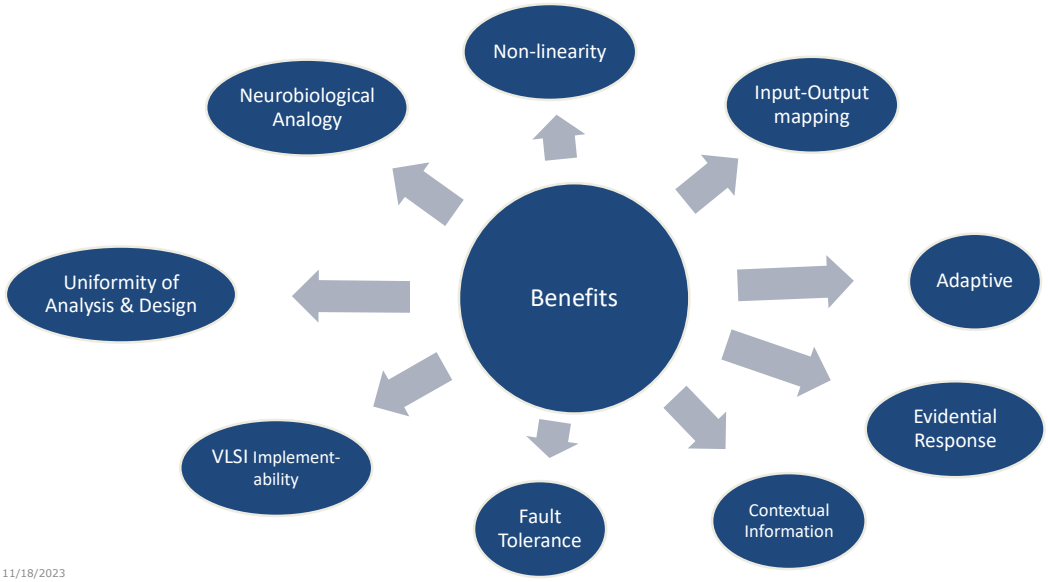
All models are wrong... some models are useful!

Is this how our brain works? Really!!

pra-sâmi

12

Benefits of Neural Networks



pra-sâmi

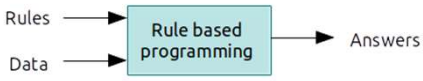
13

What's Being Played...

Can a Rule based system answer based on the data provided?

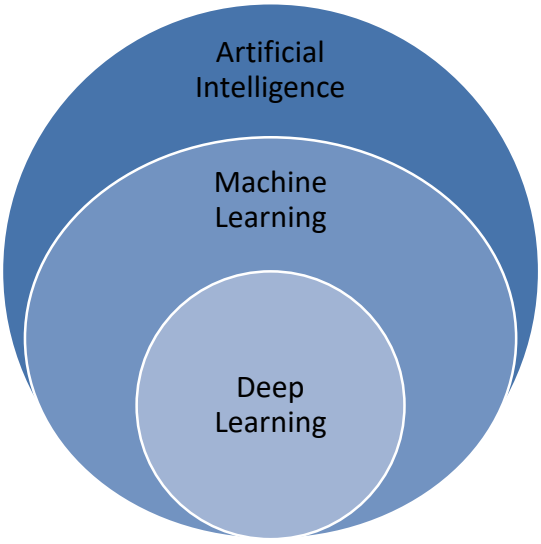
We as humans can easily guess!

Time	Day	Type
9:00 AM	Weekday	News
11:00 AM	Weekday	K-Soaps
3:00 PM	Weekday	Soaps
5:00 PM	Weekday	Soaps
6:00 PM	Weekday	Cartoons
9:00 PM	Weekday	Sports, Movies
4:00 PM	Weekday	???

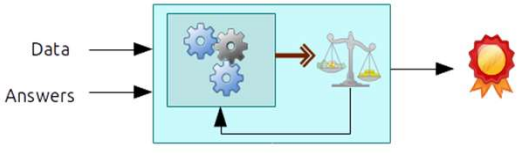


14

AI vs ML vs Deep Learning



- Used interchangeably
- AI is a broader concept, it includes basic AI to Deep learning.
- Machine learning: enabling Machines to Learn from the past incidents (available data).
- Deep Learning: One can say that it tries to copy information processing patterns found in the human brain



pra-sâmi

15

Artificial Intelligence vs. Machine Learning

Artificial Intelligence

- Create intelligent machines that can simulate Human thinking capability and behavior
- A man-made thinking power
- No pre-programing needed
- Algorithms which can work with their own "intelligence"
- Algorithms such as Reinforcement learning algorithm and deep learning neural networks being used in multiple places such as Siri, Google's AlphaGo, AI in Chess playing, etc.
- Based on capabilities, AI can be classified into three types:
  - Weak AI
  - General AI
  - Strong AI
- Currently, we are working with weak AI and general AI. The future of AI is Strong AI for which it is said that it will be more intelligent than humans (???)

Machine Learning

- An application or subset of AI
- Allows machines to learn from data without being programmed explicitly
- Uses a massive amount of structured and semi-structured data
- It can work only on data it has seen
- For unknown cases it becomes unresponsive or unreliable
- Being used for online recommender system, for Google search algorithms, Email spam filter, Facebook Auto friend tagging suggestion, etc.
- It can be divided into three types:
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning

pra-sâmi

16

Deep Learning

- Large Neural Networks

"Using brain simulations, hope to:
Make learning algorithms much better and easier to use,
Make revolutionary advances in machine learning and AI,
I believe this is our best shot at progress towards real AI."

- Andrew Ng

- Learning successive layers of increasingly meaningful representations
- Modern network contain hundreds of successive layers
- Successive layers are learned via "neurons" connected via neural network

Some concepts were inspired by how our brain works
It is NOT a replica of human brain!!!

pra-sâmi

17

Deep Learning

- Why Deep Learning is more practical today?
  - Availability of large computing power
  - Availability of large datasets
- Most flavors of the old generations of learning algorithms, performance will plateau
- Deep learning that is scalable
  - Performance just keeps getting better as more and more data is fed
- Most value today is coming from supervised learning
- Eventually, we will see benefits of unsupervised learning

18

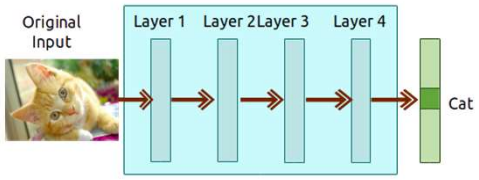
Deep Learning

- Usually a neural network contains
  - Input Layer
  - Hidden layers [1 … n]
  - Output layer
- We may call network with 1 to 2 hidden layer as shallow
- Network with 10 or more layers as deep
  - No set demarcation!
- I guess, scientists just got excited when someone labeled them as deep network
- Intelligent software to automate routine tasks, understand speech or images, make diagnosis in medicine and support basic scientific research

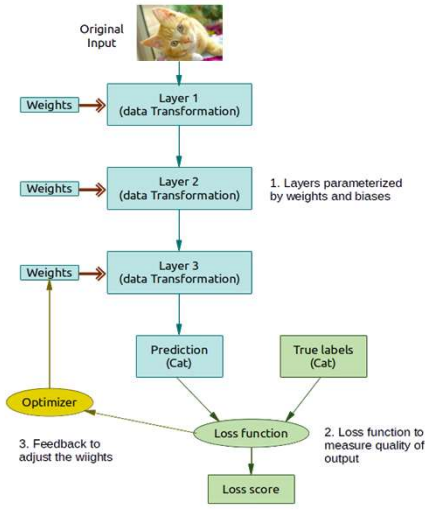
19

How deep learning works…

- As the images are processed through the layers



- The representations are increasingly filtered, purified and distilled to make them more meaningful



pra-sâmi

20

What has been achieved so far

- Learn to see and hear… so natural to humans but elusive to machines earlier
- Image classification
- Speech recognition
- Handwriting reco
- Writing style recognition (who was the author)
- Improved machine translation
- Text-to-speech conversion
- Digital assistants such as Google Now and Amazon Alexa
- Little autonomous driving
- Improved ad targeting, as used by Google, Baidu, and Bing
- esults on the web
- Ability to answer natural-language questions
- Superhuman games playing: chess, go…

Still long way to go…
Human-level general intelligence too far away…

pra-sâmi

21

Neurons

pra-sâmi

22

To play or not to play!

id	Dry Weather	Low Temp	Homework Done	Team Members	Equipment	Ground	Played
1	1	1	1	1	0	1	1
2	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1
4	0	1	0	1	1	1	0
5	0	0	1	1	1	0	0
6	0	0	0	0	0	1	0

pra-sâmi

23

Neurons

- Features:
  - Is it raining?
  - Is it too hot?
  - Have I completed my homework?
  - Are sufficient players ready?
  - Is cricket equipment ready?
  - Is ground available?
- Depending on the feature values, you may get to play or not
- Features like homework and availability of ground can be considered as 'inhibitory'.

id	Dry Weather	Low Temp	Homework Done	Team Members	Equipment	Ground	Played
1	1	1	1	1	0	1	1
2	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1
4	0	1	0	1	1	1	0
5	0	0	1	1	1	0	0
6	0	0	0	0	0	1	0

- Notes :
  - Aggregator function is sum and threshold can be 3.
  - Assign 0 or 1 if a parameter is in favor or not

Given sufficient data point, we can train an algorithm to make such simple decisions for us.

pra-sâmi

24

MP Neuron

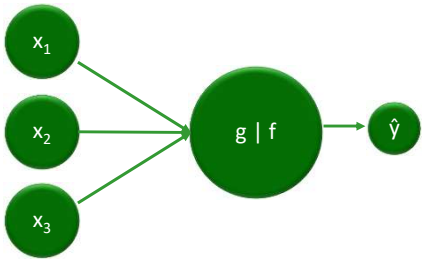
- In 1943 Warren S. McCulloch, a neuroscientist, and Walter Pitts, a logician, published "A logical calculus of the ideas immanent in nervous activity" in the Bulletin of Mathematical Biophysics
- In this paper McCulloch and Pitts tried to understand how the brain could produce highly complex patterns by using many basic cells that are connected together
- These basic brain cells are called neurons, and McCulloch and Pitts gave a highly simplified model of a neuron in their paper
- The McCulloch and Pitts model of a neuron, which we will call an MCP neuron for short, has made an important contribution to the development of artificial neural networks -- which model key features of biological neurons

pra-sâmi

25

MP Neuron

- Neurons receive signals and produce a response
- In this model:
  - All inputs are binary i.e. [0,1]
  - Inputs are “inhibitory” or “excitatory”.
  - Inhibitory have maximum influence on the model
  - It has an aggregator ‘g’ and a function ‘f’
  - There is a threshold
  - If g is more than threshold, $\hat{y}$ = 1 else 0

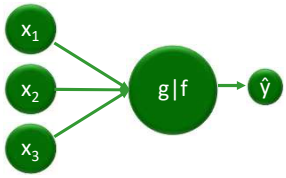


- $\hat{y}$ = 0 if any x_i is inhibitory, else g(x) = $\sum x_i$
- $\hat{y}$ = 1 if g(x) $\geq$ threshold else $\hat{y}$ = 0

pra-sâmi

26

MP Neuron

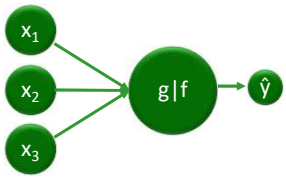


id	Dry Weather	Low Temp	Homework Done	Team Members	Equipment	Ground	Sum	Played
1	1	1	1	1	0	1	5	1
2	1	1	1	1	1	1	6	1
3	1	1	1	1	1	1	6	1
4	0	1	0	1	1	1	4	0
5	0	0	1	1	1	0	3	0
6	0	0	0	0	0	1	1	0

pra-sâmi

27

MP Neuron



id	Dry Weather	Low Temp	Homework Done	Team Members	Equipment	Ground	Sum	Played
1	1	1	1	1	0	1	5	1
2	1	1	1	1	1	1	6	1
3	1	1	1	1	1	1	6	1
4	0	1	0	1	1	1	4	0
5	0	0	1	1	1	0	3	0
6	0	0	0	0	0	1	1	0

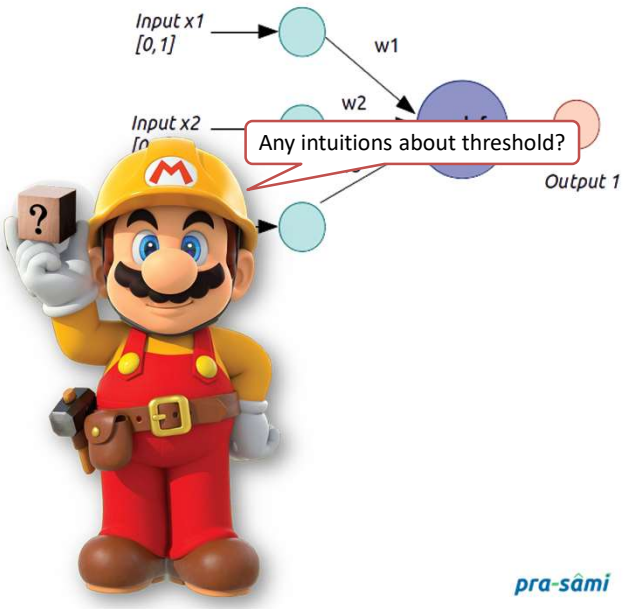
The logic is straight forward. Let's implement this model on a dataset.

pra-sâmi

28

Code Example1 – MP Neurons

- Need a dataset with plenty of features and binary output
- Breast Cancer dataset from scikit-learn
  - data = sklearn.datasets.load_breast_cancer()
- Its features are a continuous and we need binary
  - Use pandas pd.cut to bin the columns
  - X_bin = X . apply (pd.cut, bins=2, labels=[1,0])
  - For b in range [0, num_features+1]
    - Sum it by row and compare with b
- Converted file is in the shared folder



pra-sâmi

29

pra-sâmi

ADDITIONAL MATERIAL

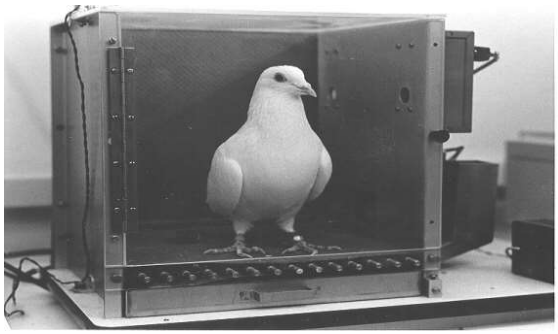
pra-sâmi

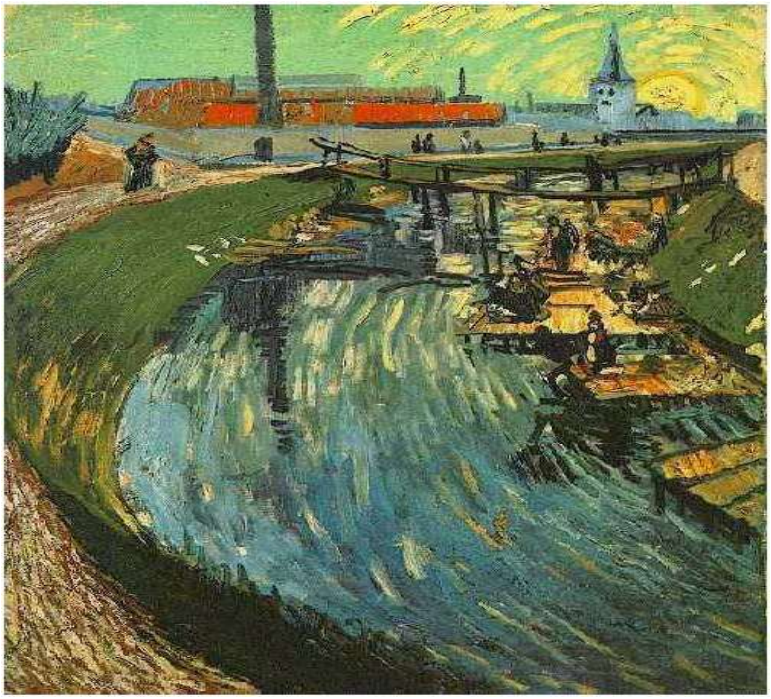
31

Biological Neural Nets

- Pigeons as art experts (Watanabe et al. 1995)

- Experiment:
  - Pigeon in Skinner box
  - Present paintings of two different artists (e.g. Chagall / Van Gogh)
  - Reward for pecking when presented a particular artist (e.g. Van Gogh)





Biological Neural Nets

- Pigeons were able to discriminate between Van Gogh and Chagall
  - With 95% accuracy on train set (when presented with pictures they had been trained on)
  - Discrimination, still 85% successful for previously unseen paintings of the artists

- Pigeons do not simply memorise the pictures
- They can extract and recognise patterns (the 'style')
- They generalise from the already seen to make predictions

- This is what neural networks (biological and artificial) are good at (unlike conventional computer)

35

Brain and Machine

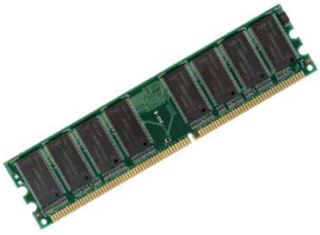
- The Brain
  - Pattern Recognition
  - Association
  - Complexity
  - Noise Tolerance

- The Machine
  - Calculation
  - Precision
  - Logic

pra-sâmi

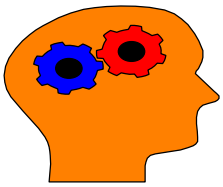
36

The contrast in architecture



- The Von Neumann architecture uses a single processing unit;
  - Tens of millions of operations per second
  - Absolute arithmetic precision

- The brain uses many slow unreliable processors acting in parallel



pra-sâmi

37

The biological inspiration

- Features of the Brain
  - Ten billion (10^{10}) neurons
  - On average, several thousand connections
  - Hundreds of operations per second
  - Die off frequently (never replaced)
  - Compensates for problems by massive parallelism

- The brain has been extensively studied by scientists
- Vast complexity prevents all but rudimentary understanding
- Even the behavior of an individual neuron is extremely complex
- Single "percepts" distributed among many neurons
- Localized parts of the brain are responsible for certain well-defined functions (e.g. vision, motion).