A Brief History of Connectionism

By Jonathan Hall

Table of Contents

- Definitions
- Intro to Connectionism
- Old Connectionism
- New Connectionism

Wikipedia Definition

- **Connectionism** is a set of approaches in the fields of artificial intelligence, cognitive psychology, cognitive science, neuroscience and philosophy of mind, that models mental or behavioral phenomena as the emergent processes of *interconnected networks of simple units*. There are many forms of connectionism, but the most common forms use neural network models.

Supervised Learning vs. Unsupervised Learning from Wikipedia

- In machine learning, **unsupervised learning** is a class of problems in which one seeks to determine how the data are organized. It is distinguished from supervised learning (and reinforcement learning) in that the learner is given only unlabeled examples.
- **Supervised learning** is a machine learning technique for deducing a function from training data. The training data consist of pairs of input objects (typically vectors), and desired outputs. The output of the function can be a continuous value (called regression), or can predict a class label of the input object (called classification). The task of the supervised learner is to predict the value of the function for any valid input object after having seen a number of training examples (i.e. pairs of input and target output). To achieve this, the learner has to generalize from the presented data to unseen situations in a "reasonable" way (see inductive bias).
- Just remember: supervised learning is MUCH easier than unsupervised learning.

Linearly Separable vs. Linearly Inseparable

linearly separable

linearly inseparable

Applications of Neural Networks

- Spam Filtering
- Certain robot competitions
- Speech reading

Intro to Connectionism

- Multidisciplinary field
- Human mind = computer
- Connectionist models try to emulate the brain’s structure and processes into a computer

Old Connectionism Contents

- Psychological contributions
- Neuropsychological contributions
- Early models

Psychological Contributions

- Spencer's Connexions:
  - Need to understand how brain works
  - Idea of weighted connections
- James' Associative Memory Model:
  - Recall of one idea can recall related ideas
  - Not all connections are equal
- Thorndike's Connectionism:
  - Law of Exercise or Use or Frequency: repeat the same action over and over → tendency to do that action increases
  - Law of Effect: reward for action increases tendency of action; punishment for action decreases tendency of action

Neuropsychological Contributions

- Lashley's Search for the Engram:
  - Equipotentiality Principle: other parts of the brain can pick up the slack left by one part of the brain
  - Mass Action Principle: reduction in learning capability proportional to amount of brain tissue damaged
- Hebbian Learning:
  - If brain cells A & B interact enough, their compatibility with each other will increase

Early models

- Pandemonium: learning model with 4 layers
  - 4th layer: store and pass data
  - 3rd layer: perform computations on data
  - 2nd layer: sort and weigh results
  - 1st layer: make decisions

- Tested on 2 tasks: distinguish dots from dashes, recognize 10 different hand-drawn characters
- Good work, but tasks were <u>simple</u>

Early Models: Perceptron

- Single layer neural network that learns something
- Learns to classify something with “true” or “false” by studying examples
- Supervised learning algorithm
- Why it’s important: can solve any binary pattern classification problem if a solution exists

Early Models: Perceptron (Cont.)

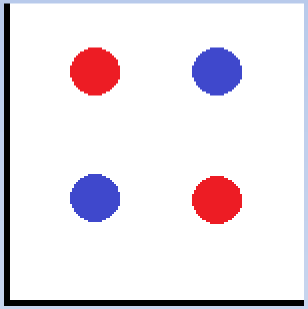
- 3 layers:
  - S-layer: get input
  - A-layer: do computations
  - R-layer: handle output
- A- and R-units only fire when threshold is exceeded

Early Models: Adaline

- Adaline (ADaptive LINear Element)
- Supervised learning algorithm
- Used +1 & -1 for yes/no instead of perceptron’s 1 & 0
- Different method of answering than perceptron

Early Models: Perceptrons Limitations

- "What are neural networks good *for*?"
- Perceptrons couldn't handle linearly inseparable problems (order > 1)
- Linearly inseparable problems require multiple layers in the neural network



XOR Classification

- Summary: neural networks of that time were good at small problems, but bad at larger problems

Importance of Old Connectionism

- Academic: know history
- learn from history, not repeat it
  - one of the guys spent 30yrs. to figure out 2 facts that may not even be completely accurate

New Connectionism Contents

- Interactive Activation and Competition (IAC) Model
- Grossberg’s Instar & Outstar Neurons
- Grossberg’s Adaptive Resonance Theory (ART)
- Multi-Layer Perceptrons (MLP)
- Generalized Delta Rule (GDR)

Interactive Activation and Competition (IAC) Model

- Units organized into “pools”
  - Units in pool compete for strong connection
  - Pool connections: normally bidirectional and excitatory
- 2 types of nodes
  - Instance: connections and communication
  - Property: contain information

IAC in Action

- Goal: obtain data about “Lance”

1) Activate “Lance” property node
2) “Lance” property node does 2 things:
   - Send inhibitory signal to other property nodes in same pool
   - Send excitatory signal to “Lance” instance node
3) “Lance” instance node does 2 things:
   - Send inhibitory signal to all other instance nodes
   - Send excitatory signal to properties of “Lance”
4) Properties of “Lance” nodes do 2 things:
   - Send inhibitory signal to nodes within their respective pools
   - Send excitatory signal back instance nodes that connected to them
5) Eventually everything settles into equilibrium

Grossberg’s Instar & Outstar Neurons

- Instars: learn a specific pattern
  - One pattern per instar
  - Adjusted weight vector means instar gets the right input
- Outstars: transmit a specific pattern
- Summary: Instars receive and recognize the data and outstars transmit data to other neurons

Grossberg’s Adaptive Resonance Theory (ART)

- What is it: Unsupervised Learning Algorithm
- Purpose: Store and classify data, like a vector
- How it works:
  – Give ART network new data
  – ART network checks if new data is similar or identical to existing categories (within a tolerance)
  – If yes, then new data is stored in an existing category and category is modified to include new data
  – If no, then new category is created and stores the new data
- Why ART is good: network can learn new data and remain stable (not crash) while doing so

Kohonen Network

- Self-organized mapping
- How it’s different: one input neuron affects *all* output neurons
- How it works:
  - Get input, like a vector
  - Output neuron with highest value/highest weight does the classification
  - Adjust neurons accordingly
  - Rinse and repeat until training is completed

Multi-Layer Perceptron (MLP)

- Input units
- Hidden units
- Output units
- Units are in layers
- Feed-forward architecture
- Everything is done is parallel
- Why MLP is good: can *theoretically* solve any pattern classification problem
- Why MLP is good in application: ability to learn

Generalized Delta Rule (GDR)

- Very important
- Generalized training procedure for neural networks
- Supervised learning algorithm
- Another way of using the back propagation algorithm

Other Networks

- Recurrent Networks
  - Input→Hidden→Output→State→Back to Hidden
- Value Unit Networks
  - Solve local minima problem by carefully choosing starting point
- Radial Basis Function
  - Little different from standard networks and useful for certain types of problems

Importance of New Connectionism

- Multilayer networks
  - Can (theoretically) train a network to solve problems
- Scientist can choose which network to use to solve a problem

Future of Neural Networks

- Networks that more closely emulate the brain

The End