Logical Neural Networks: Opening the black box
Logical Normal Form Network

Daniel Braithwaite

May 7, 2017

Given a simple feed foward network using neurons resembleing NAND gates is unable to learn any given boolean formula we switch our interest to something else. We know that any boolean expression can be represented in Conjunctive Normal Form (CNF) or Disjunctive Normal Form (DNF). So it seems prudent to ask wether its possible to construct a feedfoward netural netowrk which can learn these underlying representations. Findings [1] seem to sugest that it is possible however there is limited justification for some claims and because of this is diffcult to reproduce. We will take this general concept and reproduce the research in an attempt to develop a better understanding.

1 Noisy Neurons

During a previous investigation of Logical Neural Networks the concept of Noisy-OR and Noisy-AND neurons where derived [2]. Here we will simply state them as.

Definition 1.1. A **Noisy-OR** neuron is a perceptron with activation $a = 1 - e^{-z}$ where W is our weights, X is our inputs, b is our byas term and $z = WX + b$. We contrain each $w_i \in W$ and b to be in the interval $[0, \inf]$

Definition 1.2. A **Noisy-OR** neuron is a perceptron with activation $a = e^{-z}$ where W is our weights, X is our inputs, b is our byas term and $z = WX + b$. We contrain each $w_i \in W$ and b to be in the interval $[0, \inf]$

2 Logical Normal Form Networks

A Logical Normal Form Network is a neural net which stasfys one of the folowing definitions

Definition 2.1. CNF-Network A **CNF-Network** is a three layer network where neurons in the hidden layer consist soley of Noisy-OR's and the output layer is a single Noisy-AND.

Definition 2.2. DNF-Network A **DNF-Network** is a three layer network where neurons in the hidden layer consist soley of Noisy-AND's and the output layer is a single Noisy-OR.

It is worth noteing that CNF and DNF form can have the nots of atoms in there clauses, a simple way to account for this is to double the number of inputs where one represents the atom and the next represents the negation of that atom.

3 Learnability Of Boolean Gates

Theoretiaclly it makes sense for these networks to be able to learn the CNF and DNF representations of various boolean expressions, however before we attempt arbatary boolean functions we would like to start with something simple, namely expressions such as NOT, AND, NOR, NAND, XOR and IMPLIES. Results of which where promising, we where able to achieve a low error and from inspecting the weights we could see that the networks where infact learning the correct CNF and DNF representations.

4 Learnability Of Interesting Expressions

We now wish to see if we can use these networks to learn more interesting boolean formula, starting with expressions of 3 variables. However we encounter our first issue with this approach, despite being able to learn the given expressions and achieve a low error the weights are now sometimes not interpretable. This sugests that we need to use a different paramaterisation for our neurons. Specifically something which limits our neurons functionality so that we can maintian a well defined operation.

References

[1] HERRMANN, C., AND THIER, A. Backpropagation for neural dnf-and cnf-networks. *Knowledge Representation in Neural Networks, S* (1996), 63–72.

[2] LAURENSON, J. Learning logical activations in neural networks, 2016.