# A super-awesome jedi power title

Magnus Halvorsen

September 17, 2014

**Abstract**

# Chapter 1

# Introduction

## 1.1   Artificial Neural Networks

An Artificial Neural Network (ANN) is a computational model that is used for machine learning and pattern recognition. The name and basic concept is inspired by how the animal brain uses a network of neurons to recognize and classify various objects.

A neuron can intuitively be viewed as a probabilistic classifier. Depending on the input data it will return the probability that the data belongs to a certain *class*.

A neuron takes as input a set of data, which is associated with a respective weight. The data and the weight are then used to calculate an activation function *f(...)* which outputs the probability that the input belongs to a certain class. If it is a binary classifier *f(...)* will output a value between 0 and 1. More formally a neuron is defined as follows:

$$Input : \{x_1, x_2, \ldots, x_n\} = \mathbf{x}$$

$$Output : h_{\mathbf{w},\mathbf{b}}(\mathbf{x}) = f(\mathbf{w^T}\mathbf{x}) = f(\sum_{i=1}^{n} w_i x_i + b)$$

$\mathbf{w}$ is the connection weights, b is the neuron bias and *f(...)* is the activation function. *f(...)* tends to be either:

$$Sigmoid : f(z) = \frac{1}{1 - e^{-z}}, \in [0, 1]$$

or

$$Hyperbolic tagent : f(z) = tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}, \in [-1, 1]$$

An ANN consist of $n_l$ layers, each containing a set of neurons. The first layer is *the input layer*, and the last layer is *the output layer*. The layers in

between are called *the hidden layers*, because their values are not observed in the training set. The input is provided to the input layer which calculates its activation functions (one for each neuron). The result is propagated to the first hidden layer, and continues up until the output layer - which provides the final result. This is known as a *feedforward neural network.*

The network takes in two parameters, $(\mathbf{W}, \mathbf{b}) = (\mathbf{w^{(1)}}, \mathbf{b^{(1)}}, \mathbf{w^{(2)}}, \mathbf{b^{(2)}}, \ldots, \mathbf{w^{(n_l)}}, \mathbf{b^{(n_l)}})$. Then $w_{ij}^l$ denotes the weight between unit j in layer l, and unit i in layer l+1. $b_i^l$ denotes the bias associated with unit i in layer l+1.

# Chapter 2

# Related Work

Related work stuff

# Chapter 3

# Method

My special method.

# Chapter 4

# Discussion

Discuss dis

# Chapter 5

# Conclusion

I HAVE CONCLUDED!