

# Neural Network for Machine Learning - Geoffrey Hinton

Shubhra Aich

## Contents

<b>1</b>	<b>Ways to make neural networks generalize better</b>	<b>2</b>
<b>2</b>	<b>Combining multiple neural networks to improve generalization</b>	<b>2</b>

## **1 Ways to make neural networks generalize better**

Different regularization methods have different effects on the learning process. For example L2 regularization penalizes high weight values. L1 regularization penalizes weight values that do not equal zero. Adding noise to the weights during learning ensures that the learned hidden representations take extreme values. Sampling the hidden representations regularizes the network by pushing the hidden representation to be binary during the forward pass which limits the modeling capacity of the network.

## **2 Combining multiple neural networks to improve generalization**

So we should aim to make the individual predictors disagree without making them being poor predictors. The art is to have individual predictors that make errors very different from one another but each one is fairly accurate.

In case of mixtures of experts, the aim of data clustering is not to find clusters of input vectors which are similar. Rather, what we are interested in is the similarity in the input output mappings. In other words, we want each cluster to have a relationship between input and output that can be well-modeled by one local model.