

# ILC Experiment

## Preliminary Results

Franck Dubard

Applied Statistics and Machine Learning Research Group  
Linear Accelerator Laboratory, Paris Sud

dubard@lal.in2p3.fr

March 18, 2013

# Table of Contents

- 1 Prerequisite
- 2 Development

# Table of Contents

- 1 Prerequisite
  - Questions over Deep Learning
  - Family of algorithms
  - The Denoising Auto-associators/Auto-encoders
- 2 Development

# Question over Deep Learning

# Question over Deep Learning

What ?

Sub-field of machine learning.

# Question over Deep Learning

What ?

Sub-field of machine learning.

When ?

Since the 80's,

# Question over Deep Learning

What ?

Sub-field of machine learning.

When ?

Since the 80's, but it is actually only booming until a few years ago.

# Question over Deep Learning

What ?

Sub-field of machine learning.

When ?

Since the 80's, but it is actually only booming until a few years ago.

Why ?

Learning several levels of representations, corresponding to a hierarchy of features or factors or concepts.


Most of the algorithms are framed as unsupervised learning.



# Family of algorithms

- 2 different paradigms, but the connections are more tenuous when we consider deeper models<sup>1</sup>

---

<sup>1</sup>For more details, please refer to the paper named *Representation Learning: A Review and New Perspectives* by Y. Bengio, A. Courville and P. Vincent 

# Family of algorithms

- 2 different paradigms, but the connections are more tenuous when we consider deeper models<sup>1</sup>

## Probabilistic Graphical Model

RBM<sup>1</sup> and its derivatives.

---

<sup>1</sup>Restricted Boltzmann Machines

## Neural Network Model


The AAs<sup>1</sup>/AEs<sup>2</sup> variants.

---

<sup>1</sup>Auto-Associators

<sup>2</sup>Auto-Encoders

---

<sup>1</sup>For more details, please refer to the paper named *Representation Learning: A Review and New Perspectives* by Y. Bengio, A. Courville and P. Vincent 

# Family of algorithms

- 2 different paradigms, but the connections are more tenuous when we consider deeper models<sup>1</sup>

## Probabilistic Graphical Model

RBM<sup>1</sup> and its derivatives.

---

<sup>1</sup>Restricted Boltzmann Machines

## Neural Network Model

The AAs<sup>1</sup>/AEs<sup>2</sup> variants.

---

<sup>1</sup>Auto-Associators

<sup>2</sup>Auto-Encoders

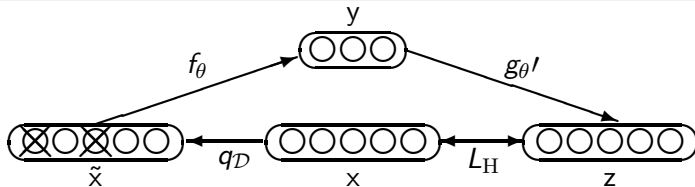
## Choice

We decided to use AAs/AEs and more precisely, DAAs/DAEs.

---

<sup>1</sup>For more details, please refer to the paper named *Representation Learning: A Review and New Perspectives* by Y. Bengio, A. Courville and P. Vincent

# The DAAs/DAEs Principle



- Clean input  $\mathbf{x} \in [0, 1]^d$  is **partially destroyed**, yielding **corrupted input**:  $\tilde{\mathbf{x}} \sim q_D(\tilde{\mathbf{x}}|\mathbf{x})$ .
- $\tilde{\mathbf{x}}$  is mapped to **hidden representation**  $\mathbf{y} = f_\theta(\tilde{\mathbf{x}})$ .
- From  $\mathbf{y}$ , we **reconstruct**  $\mathbf{z} = g_{\theta'}(\mathbf{y})$ .
- Train parameters to minimize the **cross-entropy "reconstruction error"**  $L_H(\mathbf{x}, \mathbf{z}) = H(\mathcal{B}_x || \mathcal{B}_z)$ , where  $\mathcal{B}_x$  denotes multivariate Bernoulli distribution with parameter  $\mathbf{x}$ .

*Extracting and Composing Robust Features with Denoising Autoencoders, ICML 2008*

# Table of Contents

## 1 Prerequisite

## 2 Development

- Toy Problem
- Real Problem
- Preliminary Results

# Toy Problem

## Description

We use our method on one of this year ICDAR<sup>1</sup> competitions dataset, the Handwritten Digit and Digit String Recognition Competition<sup>2</sup>.

---

<sup>1</sup>International Conference on Document Analysis and Recognition

<sup>2</sup><http://caa.tuwien.ac.at/cv1/research/icdar2013-hdrc>

# Toy Problem Results

The hyperparameters are :

- Learning Rate : 25%
- Corruption Level : 90 %

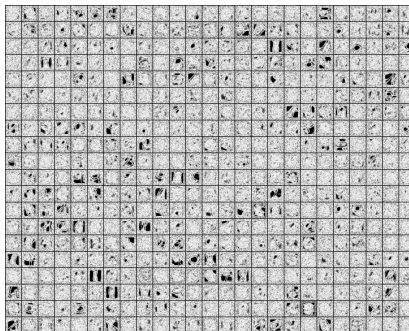


Figure : 500 filters of size  $28 \times 28$

# Toy Problem Results

The hyperparameters are :

- Learning Rate : 25%
- Corruption Level : 90 %

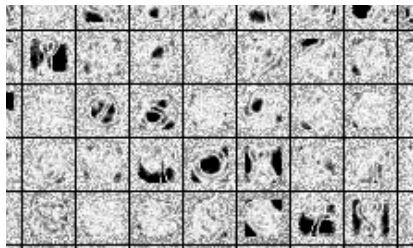


Figure : some filters began to show the shape of recognizable number



# Real Problem

## Data Set Characteristics

We have 6150 examples.

Each one are a vector of 9720 ( $18 \times 18 \times 30$ ) numbers.

Each number are part of  $\mathbb{R}_{\geq 0}$ , which represents the amount of deposit energy.

# Preliminary Results

Simulation...