

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones

Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

Apprentissage profond, réseau de neurones, etc.

Antoine BIARD & Vincent BODIN

May 20, 2014

Introduction

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

- Souvent, l'extraction d'une représentation repose sur une connaissance *a priori* des données ;
- quantité de données différentes pose la question de la généralité de cette méthode ;
- volonté d'extraire - de manière non-supervisée - une représentation qui agrège l'information ;
- méthode profonde semblent produire ce résultat avec des représentations de plus en plus abstraites.

Table of Contents

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)
Restricted
Boltzmann
Machine
(RBM)
Deep Belief
Networks
(DBN)
Deep
Boltzmann
Machine
(DBM)

1 Revue élémentaire des méthodes de deep learning

- Réseau de neurones
- Multilayer Perceptron (MLP)
- Restricted Boltzmann Machine (RBM)
- Deep Belief Networks (DBN)
- Deep Boltzmann Machine (DBM)

2 Implémentation et résultats

Réseau de neurones

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones

Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

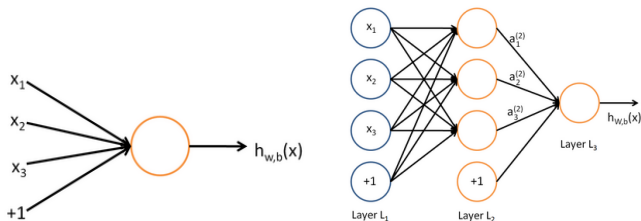


Figure: (gauche) Un neurone avec trois entrées (x_1, x_2, x_3) et un *offset* ; (droite) un réseau de neurones de taille (3,3,1) avec des *offset* aux deux premières couches.

Sortie d'un neurone simple

$$h_{w,b}(x) = \sigma(w^T x) = \sigma \left(\sum_{i=1}^p w_i x_i + b \right) \quad (1)$$

Multilayer Perceptron (MLP)

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones

**Multilayer
Perceptron
(MLP)**

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

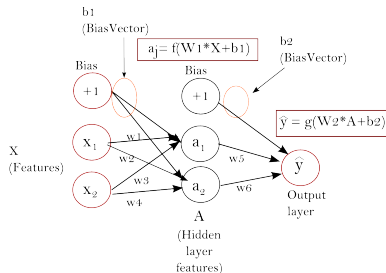
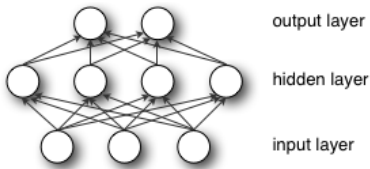


Figure: (gauche) Un MLP avec une seule couche de variables cachées ; (droite) structure de MLP avec l'ajout de biais et les poids.

Multilayer Perceptron (MLP)

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones

Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

Sortie d'un nœud

$$s(x) = \sigma(Wx + b) \quad (2)$$

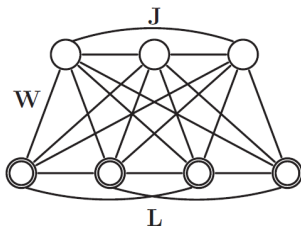
Le MLP diffère par la manière d'entraîner :

Forward pass. On part des variables visibles v et on remonte le graphe avec les poids de la structure.

Backpropagation. Chemin inverse en descendant le graphe et en comptabilisant les erreurs commises, mise à jour les poids.

Machine de Boltzmann

General Boltzmann Machine



Restricted Boltzmann Machine

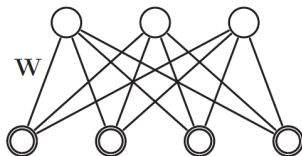


Figure: (gauche) Machine de Boltzmann générale ; (droite) RBM.

Énergie

$$E(v, h; \theta) = -v^T W h - \frac{1}{2} v^T L v - \frac{1}{2} h^T J h \quad (3)$$

(pas de terme de biais ici)

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

Restricted Boltzmann Machine

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

**Restricted
Boltzmann
Machine
(RBM)**

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

- La factorisation dans le graphe donne :

$$\begin{aligned} p(h|v) &= \prod_i p(h_i|x) \\ p(v|h) &= \prod_j p(x_j|h) \end{aligned} \quad (4)$$

- Les probabilités conditionnelles valent :

$$\begin{aligned} p(h_i = 1|v) &= \sigma \left(\sum_j W_{ji} x_j + d_i \right) \\ p(x_j = 1|h) &= \sigma \left(\sum_i W_{ji} h_i + b_j \right) \end{aligned} \quad (5)$$

Deep Belief Networks (DBN)

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones

Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

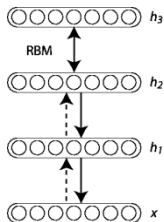


Figure: Un *deep belief network* (DBN), la dernière couche est non orientée, tandis que toutes les autres le sont.

$$p(x, h^1, \dots, h^l) = \left(\prod_{k=0}^{l-2} p(h^k | h^{k+1}) \right) p(h^{l-1}, h^l) \quad (6)$$

Entraînement par couches des RBM (*pre-training*) ;
échantillonnage pour entrée dans MLP (*fine-tuning*).

Deep Boltzmann Machine (DBM)

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)
Restricted
Boltzmann
Machine
(RBM)
Deep Belief
Networks
(DBN)
Deep
Boltzmann
Machine
(DBM)

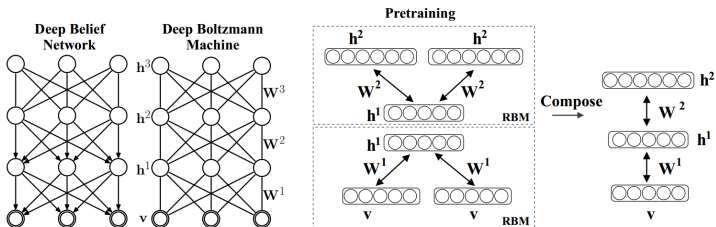


Figure: (extrême gauche) Un DBN avec sa dernière couche non orientée ; (gauche) graphe de DBM, toutes les couches sont non orientées ; (droite) *pre-training* des RBM couche par couche ; (extrême droite) composition des couches pour former le DBM.

Table of Contents

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

1 Revue élémentaire des méthodes de deep learning

2 Implémentation et résultats

- Implémentation
- Résultats

Données utilisées

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

Travail sur la base de données MNIST

- MNIST réduite : 8×8 pixels, 1800 digits
- MNIST originale : 28×28 pixels, 50000 digits (figure 6)

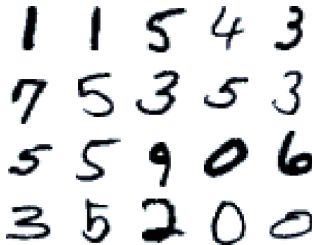


Figure: Exemple de digits de la base MNIST.

Implémentation du RBM

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

Gradient analytique

$$\frac{\partial \log(P(x; \theta))}{\partial \theta} = - \sum_h P(h|x) \frac{\partial E(x, h; \theta)}{\partial \theta} + - \sum_{x, h} P(x, h) \frac{\partial E(x, h; \theta)}{\partial \theta} \quad (7)$$

Plusieurs techniques de calcul :

- Contrastive divergence (CD)
- Persistent Contrastive Divergence

Utilisation de la bibliothèque scikit-learn qui implément CD

Implémentation du MLP

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

	precision	recall	f1-score	support
0	0.90	0.96	0.93	174
1	0.67	0.58	0.62	184
2	0.86	0.87	0.86	166
3	0.75	0.74	0.75	194
4	0.86	0.85	0.86	186
5	0.75	0.80	0.77	181
6	0.90	0.94	0.92	207
7	0.84	0.90	0.87	154
8	0.80	0.65	0.72	182
9	0.69	0.78	0.73	169
avg / total	0.80	0.80	0.80	1797

Figure: Résultats avec le MLP seul et la petite base MNIST - méthode de classification SVM

Processus complet

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones

Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

Algorithm 1 Entrainement d'un DBN

```
H ← digits
for currentRBM in RBMs do
    Entraîner currentRBM à partir de H
    if currentRBM n'est pas le dernier then
        Echantillonner une base H
    end if
    previousRBM ← currentRBM
end for
MLP.weights ← RBMs.weights
Entraînement du MLP à partir de digits
Entraînement de la regression logistique
Test à partir d'une base de test
```

Résultats du RBM

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

	precision	recall	f1-score	support
0	0.99	0.99	0.99	174
1	0.92	0.95	0.93	184
2	0.95	0.98	0.97	166
3	0.97	0.91	0.94	194
4	0.97	0.95	0.96	186
5	0.93	0.93	0.93	181
6	0.98	0.97	0.97	207
7	0.95	1.00	0.97	154
8	0.90	0.88	0.89	182
9	0.91	0.93	0.92	169
avg / total	0.95	0.95	0.95	1797

Figure: Résultats avec le RBM seul et la petite base MNIST -
méthode de classification régression logistique.

Résultats du DBN

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

	precision	recall	f1-score	support
0	0.99	0.99	0.99	174
1	0.94	0.92	0.93	184
2	0.93	0.98	0.95	166
3	0.94	0.91	0.92	194
4	0.97	0.91	0.94	186
5	0.94	0.91	0.92	181
6	0.98	0.95	0.96	207
7	0.93	0.98	0.96	154
8	0.86	0.92	0.89	182
9	0.89	0.93	0.91	169
avg / total	0.94	0.94	0.94	1797

Figure: Résultats avec le DBN et la petite base MNIST - méthode de classification régression logistique.

Résultats du DBN

Apprentissage
profond,
réseau de
neurones,
etc.

Antoine
BIARD &
Vincent
BODIN

Introduction

Revue
élémentaire
des
méthodes de
deep learning

Réseau de
neurones
Multilayer
Perceptron
(MLP)

Restricted
Boltzmann
Machine
(RBM)

Deep Belief
Networks
(DBN)

Deep
Boltzmann
Machine
(DBM)

	precision	recall	f1-score	support
0.0	0.97	0.98	0.98	1312
1.0	0.99	0.98	0.98	1604
2.0	0.95	0.96	0.95	1348
3.0	0.94	0.95	0.95	1427
4.0	0.97	0.97	0.97	1362
5.0	0.96	0.93	0.94	1280
6.0	0.97	0.98	0.97	1397
7.0	0.96	0.95	0.96	1461
8.0	0.94	0.96	0.95	1390
9.0	0.94	0.94	0.94	1419
avg / total	0.96	0.96	0.96	14000

Figure: Résultats avec le DBN et la grande base MNIST - méthode de classification régression logistique.