

L'apprentissage profond

Par:

KOUAMO Stéphane

Institut National Polytechnique - Houphouet Boigny
Master Data Science

November 7, 2024

Contexte et Définition

Définition

Apprentissage profond (ou deep learning en Anglais) : ensemble de méthodes d'apprentissage automatique tentant de modéliser avec un haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non linéaires.

C'est un domaine de l'Intelligence Artificielle qui a littéralement explosé ces dernières années.

Domaines d'application

- reconnaissance de modèle,
- classification statistique,
- etc.

Contexte et Définition

Contexte

Le traitement de l'image consiste à traiter les vecteurs caractéristiques de l'image : abstraction de l'image.



Contexte et Définition

Contexte

Seulement, la qualité de la reconnaissance dépend de la manière dont on a effectué cette abstraction de l'image. Pour pallier à ça, le deep learning intervient.

L'idée

L'idée est de sauter cette étape d'abstraction de l'image et de passer en paramètre d'entrée du réseau l'ensemble de l'image (réseau profond, trop de données à manipuler, etc.).

- l'algorithme fabrique lui même les **traits caractéristiques** et prédit mieux (avec un taux d'erreur bas),
- on a donc plusieurs images brutes passées en entrée du réseau.

Contexte et Définition

Traits caractéristiques

Image d'une empreinte digitale



Original Image



Binary Image



Thin Image

Contexte et Définition

Contexte

Le *deep learning* fait partie d'une famille de méthodes d'apprentissage automatique fondées sur l'apprentissage de modèles de données.

Par exemple : une image peut être représentée de différentes façons par un vecteur de données en fonction de :

- l'intensité des pixels dont elle est constituée;
- les différentes arêtes;
- les différentes régions de forme particulière;
- etc.

Contexte et Définition

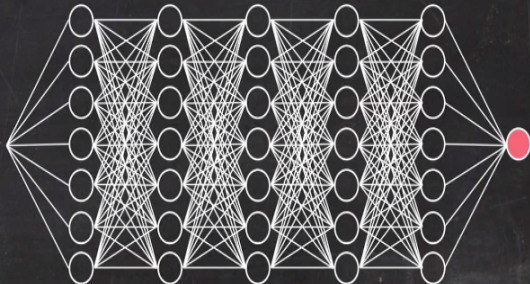
caractéristiques

Les techniques d'apprentissage profond (**deep learning**) constituent une classe d'algorithmes d'apprentissage automatique :

- Utilisant différentes couches d'unité de traitement non-linéaire pour l'extraction et la transformation des caractéristiques. Chaque couche prend en entrée la sortie de la précédente.
- Elles fonctionnent avec un apprentissage à plusieurs niveaux de détails ou de représentations des données. À travers les différentes couches on passe de paramètres de bas niveau à des paramètres de plus haut niveau.
- Ces différents niveaux correspondent à différents niveaux d'abstraction des données.

Contexte et Définition

Ainsi, pour apprendre à reconnaître un véhicule, l'on empilera différentes couches cachées avec chacune une fonction bien spécifique.



Contexte et Définition

caractéristiques

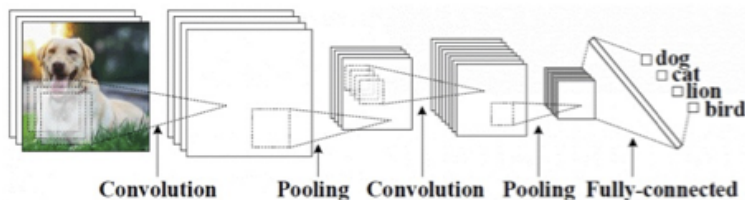
- En somme l'idée, c'est de brancher des perceptrons entre eux, regroupés en couches, de façon hiérarchique. Et d'entraîner toutes les couches en même temps.
- Pooling: Mot anglais signifiant **grouper**, le plus souvent afin d'optimiser. C'est le fait de conserver des ressources (mémoire, sessions, threads, etc.) actives au lieu d'en recréer.

RN Convolutionnel

- l'apprentissage profond consistera alors introduire des couches de convolution;
- et des couches de Pooling.

Contexte et Définition

RN Convolutionnel



Raisons du succès du deep learning

- amélioration des processeurs de cartes graphiques (GPU) ;
 - amélioration des algorithmes d'apprentissage ;
 - disponibilité des données d'apprentissage (plusieurs base de données existantes : imagenet, mnist, CCO, ...).
 - etc.
-
- Des résultats mathématiques montrent que certaines classes de fonctions, représentables de manière compacte avec un réseau de profondeur d , nécessiteraient un nombre exponentiel de paramètres avec un réseau de profondeur $d + 1$.

- L'architecture des réseaux profonds permet la réutilisation de paramètres ou de caractéristiques extraites, propriété désirable pour la modélisation de fonctions complexes.
- Le cerveau humain (plus particulièrement le cortex visuel), est construit comme une architecture profonde, avec plusieurs sous-régions fonctionnelles (V1,V2,MT) arrangées en niveaux d'abstraction de plus en plus élevés.
- En plus d'extraire des caractéristiques de plus en plus abstraites, une autre propriété désirable serait de les démêler, afin que les différents facteurs qui expliquent les variations des données soient séparés. Ceci serait très intéressant pour l'apprentissage multitâches, ainsi que pour l'adaptation de domaine.

Exemples d'algorithmes

Réseau de neurones convolutionnel

Réseau de neurones convolutionnel

$$z_{x,y}^j(A) = f\left(\sum_{r,s} w_{x,y,r,s}^j I_{r,s}(A) + t^j\right)$$

L'objectif de cette partie du réseau est d'estimer la probabilité :

$$p = p[M(A, B)/\Delta z(A, B)] = p(M/\Delta z);$$

d'une correspondance entre A et B, étant donnée la preuve Δz provenant du filtre de convolution.

Exemples d'algorithmes

Réseau de neurones localement connectés

Mathématiquement l'activation d'un neurone dans le réseau est définie par :

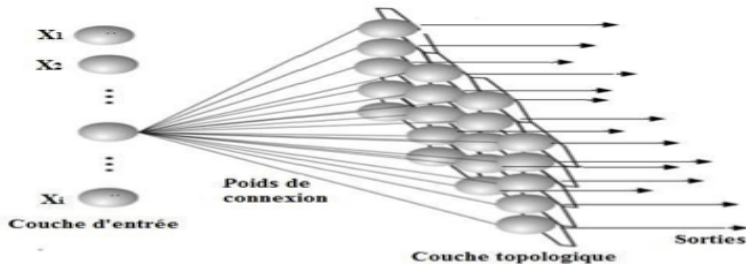
$$\alpha_{ij}(n+1) = \frac{I_{ij} + \sum_{kl \in N_r(ij)} [w_{ij;kl} \alpha_{kl}(n)]}{\sum_{ij} (I_{ij} + \sum_{kl \in N_r(ij)} [w_{ij;kl} \alpha_{kl}(n)])} \cdot E$$

Exemple : Entrée=vecteur 1D, Sortie=vecteur 1D.

Problème : classer homme/femme en fonction d'un vecteur.

Exemples d'algorithmes

Réseau de Kohonen



Exemples d'algorithmes

Réseau de neurones récurrents

un RNN est un RNA présentant des connexions récurrentes. Il est généralement constitué de neurones interconnectées interagissant non-linéairement et pour lequel il existe au moins un cycle dans la structure. Les unités sont reliées par des synapses qui possèdent un poids. La sortie d'un neurone est une combinaison non linéaire de ses entrées.

Ce type de réseau de neurones est adapté pour le traitement des series temporelles (reconnaissance du son, de l'écriture manuscrite, etc.)

Exemples d'algorithmes

Rappel : le but du principe d'apprentissage est de minimiser l'erreur entre une prédiction et la vérité.

Autres algorithmes

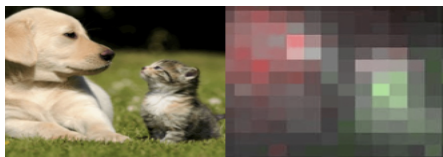
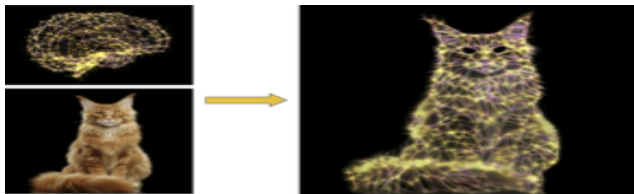
AlphaGo, Autopilot Tesla, Google Translate, Prisma, DeepDream, etc.

Exemple : Analyse de l'image

Entrée= $image(matrice2D)$, Sortie= $carte(matrice2D)$.

Problème : Détecter les objets dans un flux vidéo.

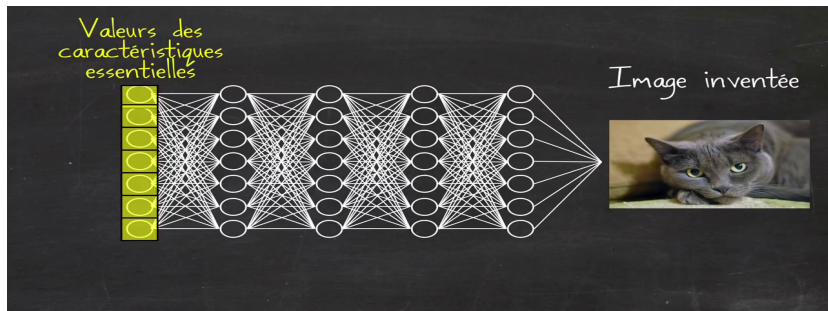
Exemples d'algorithmes



Modèles génératifs

Prendre un ensemble d'éléments en entrée et produire en sortie une image (pourrait être très utile pour la voiture qui roule toute seule pour les non-voyants);

Un modèle génératif consiste à effectuer le chemin inverse d'un réseau de neurones profonds.



Modèles génératifs

Auto-encodeur

Un auto-encodeur est un réseau de neurones artificiels utilisé pour l'apprentissage non supervisé de caractéristiques discriminantes. L'objectif étant d'apprendre une représentation (encodage) d'un ensemble de données, généralement dans le but de réduire la dimension de cet ensemble.

Etude de cas: prédiction des prix du maïs

- Effectuer un apprentissage profond avec 3 couches cachées.
- Comblent les vides des valeurs cibles.
- Prédire le prix journalier du maïs.

Echantillonnage

Utilisation de la bases de données 'Prix du maïs'

- Utilisation de la bases de données 'Prix du maïs' ,
- Utilisation de la cross validation pour diviser les bases de données et plusieurs échantillons (80 – 20, 60 – 40, et 50 – 50).

Apprentissage : algorithme proposé

Principe

- Initialiser les neurones de la première couche de classification par les prototypes de la première classe de données.
- Initialiser les neurones de la seconde couche de classification par les prototypes de la deuxième classe de données.
- Initialiser les neurones de la troisième couche cachée par les données de la base en fonction des classes prédéfinies.
- Appliquer un vecteur en entrée.
- Appliquer la mesure de similarité à la première couche de classification.

$$h_j = H(P(M(x_i - w_{ij}))) \quad 1 \leq i \leq N, \quad 1 \leq j \leq K$$

Apprentissage: algorithme proposé

- Sélectionner le neurone vainqueur.
- Appliquer la mesure de similarité à la seconde couche de classification.

$$g_k = \sqrt{\sum_{j=1}^{K1} (h_j - z_{jk})^2} \quad 1 \leq k \leq K1,$$

- Selectionner le neurone vaniqueueur.
- Appliquer la mesure de similarité à la couche de correspondance (si vainqueur classe correcte).

$$h'_l = \sqrt{\sum_{k=1}^{K2} (g_k - z'_{kl})^2} \quad 1 \leq l \leq K2,$$

Apprentissage: algorithme proposé

- Sélectionner le neurone vainqueur.
- Calculer l'erreur entre la sortie obtenue et celle attendue.

$$o_i = \frac{1}{N} \sqrt{\sum_{i=1}^N (h'_l - w'_{li})^2} \quad 1 \leq l \leq K2,$$

- Vérifier les seuils.
- Modifier le vainqueur par.

$$w_{ij}(n+1) = \frac{x - i + \sum_{ij} [w_{ij} \cdot \alpha_{kl}(n)]}{\sum_{jl} (x_i + \sum_{li} [w'_{li} \cdot \alpha_{kl}(n)])} \cdot E$$

- Mettre à jour les prototypes des neurones vainqueurs.

$$m'_{n+1} = \frac{m'_n + x'_{n+1}}{n+1}$$

- Retour à la ligne 5 jusqu' à l'ensemble de donnée terminé.

Conclusion

- Le deep learning qui est une façon particulière de faire du machine learning a trouvé le chemin du succès grâce à l'amélioration des GPU, la disponibilité des données et l'augmentation des performances des ordinateurs en terme de puissance de calcul.
- Le deep learning est indiqué pour résoudre des problèmes complexes faisant intervenir plusieurs paramètres.
- Cependant le nombre de couche cachée reste un mystère : il n'existe pas de fondement théorique sur la façon dont ce dernier doit être choisi.
- Ici c'est le résultat qui justifie que la méthode employée pour résoudre son problème est la bonne ou pas.
- Le développement des modèles génératifs pourrait être une porte ouverte pour les débuts d'explication sur comment doit être fait le choix du nombre de couche cachée.

-  R. Cappelli, A. Lumini, D. Maio and D. Maltoni, "Synthetic fingerprint-Database Generation", proceeding of the 16th International Conference on Pattern Recognition (ICPR 2002), Québec City, Vol.3, pp 744 – 747, August 2002.
-  Y. Le Cun, Y. Bengio and G. Hinton. "Deep learning". Macmillan Publishers Limited. Nature, Vol.521, pp. 436 – 444. May 2015. DOI : 10.1038/nature14539.
-  Hinton G. and al. "Deep Neural Networks for Acoustic Modelling in Speech Recognition". IEEE Signal Processing Magazine. Digital Object Identifier 10 : 1109/MSP : 2012 : 2205597. November 2012.
-  J.J. Hopfield. "Neural networks and physical systems with emergent collective computational abilities". Proceedings of the National Academy of Sciences. (1982). pp. 2554 – 2558

-  J. Jiang. "Image compression with neural networks, A survey". Signal Processing : Image Communication, (1998).
-  Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. " *Deep Residual Learning for Image Recognition*". Computer Vision and Pattern Recognition. *arXiv* : 1512 : 03385[cs : CV], 2016.
-  Y.Minami, S.Tamura, H.Sawai, K.Shikano, " *Output Smoothing for TDNN Using Information on Input Vectors Neighborhood in Input Layer, Hidden Layer* ", Proc. ASJ, 1 : 3 : 18, pp.35 – 36 (Mar. 1990) (in Japanese).
-  M. Minsky and S. Papert. *Perceptrons*. MIT Press, Cambridge MA, 1969.
-  O. Moreira-Tamayo et J. Pineda de Gyvez. " *Preprocessing Operators for Image Compression Using Cellular Neural Networks*". Dans IEEE International Conference on Neural Networks, juin 1996. pp.1500 – 1505.



Lettvin, J. Y., McCulloch, Maturana, H.R., W. Sturgis. et W. H. Pitts. " *What the frog's eye tells the frog's brain*".
Proceedings of the I.R.E. (1959). Vol.47, N° : 11, pp.
1940 – 1951.



F. Rosenblatt. " *The Perceptron: A Probabilistic Model for
Information Storage and Organization in the Brain*".
Psychological Review. Vol. 65, N°. 6, pp. 386 – 408. 1958.



D.E. Rumelhart, G.E. Hinton and R.J. Williams. " *Learning
Internal representations by error propagation*". In Parallel
Distributed Processing, Cambridge, MA : MIT Press,
1 : 318 – 362, 1986.