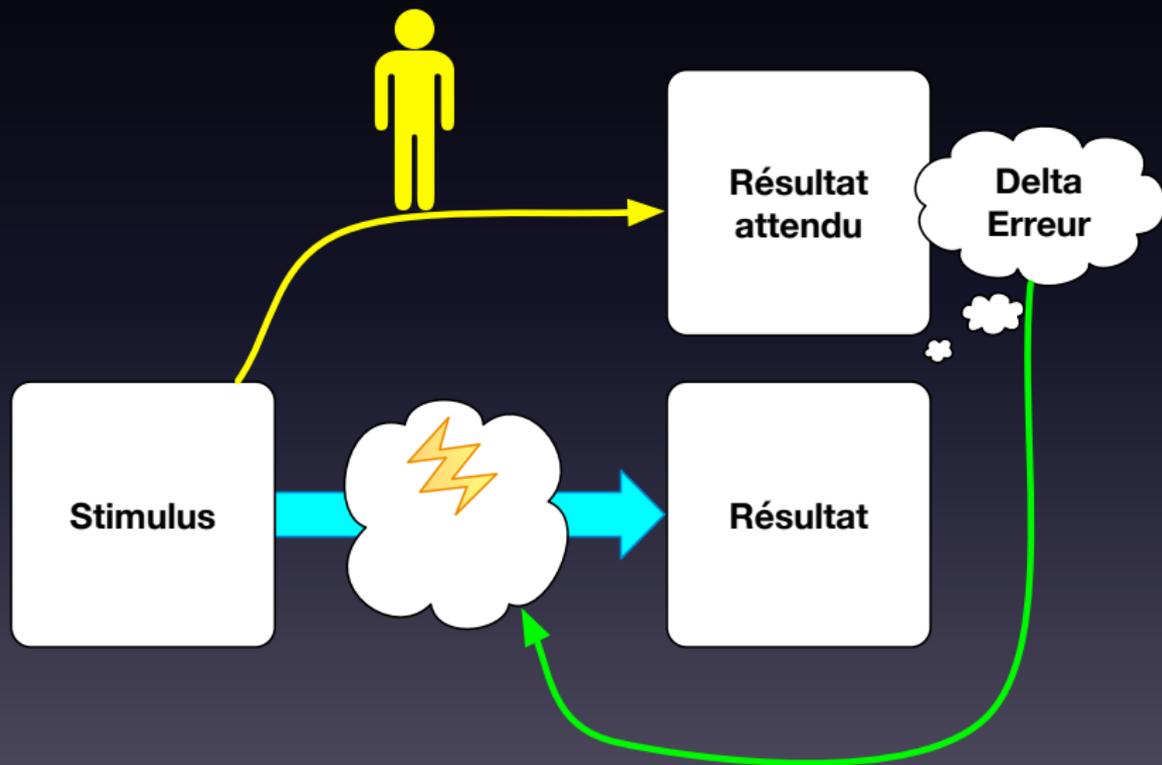


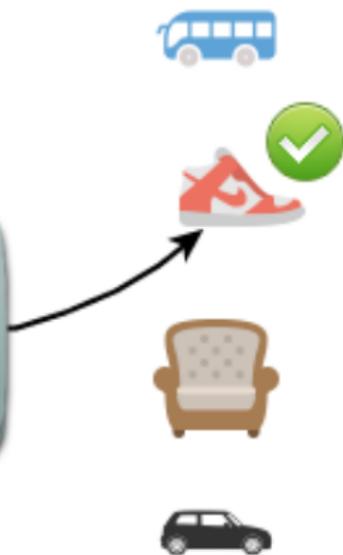
Le langage est-il apprenable à l'aide d'un réseau de neurones?

Lionel Clément – lionel.clement@u-bordeaux.fr
Université Bordeaux

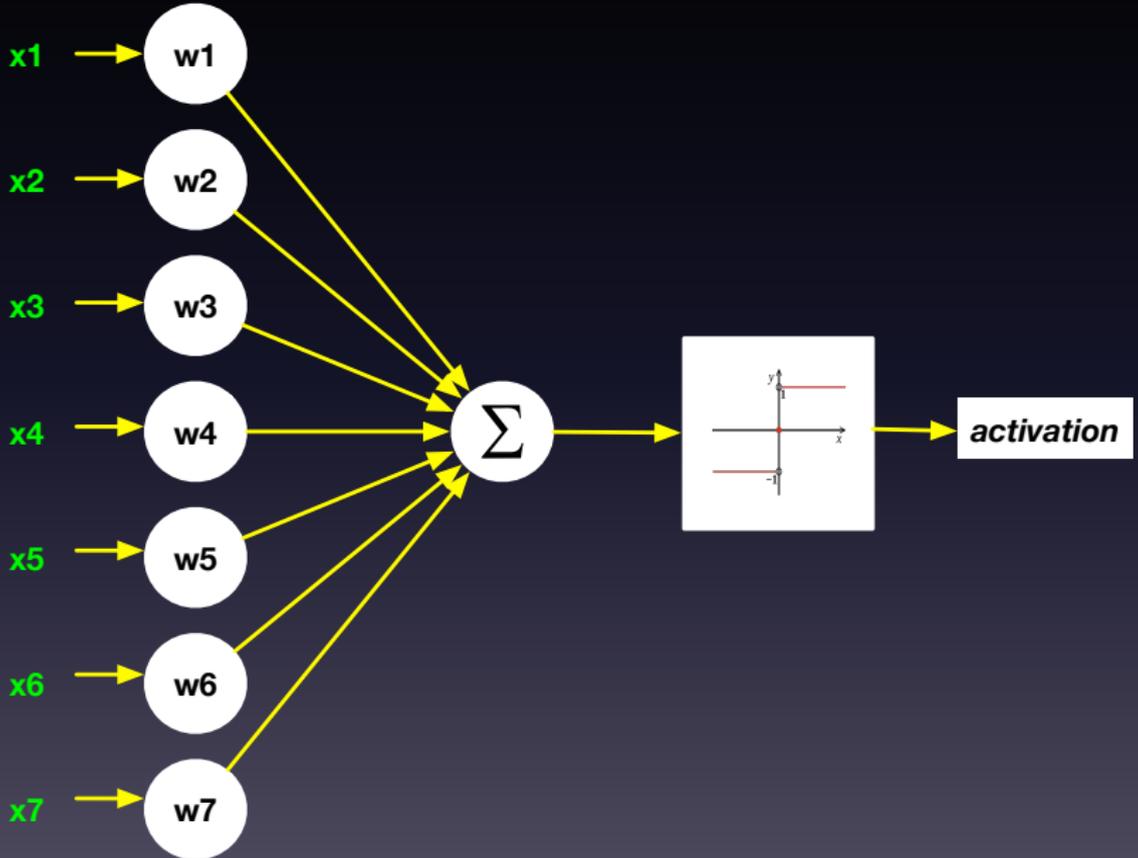
jeudi 25 octobre 2018

Apprentissage comportementaliste (ou béhavioriste)





Perceptron (Cornell 1957)



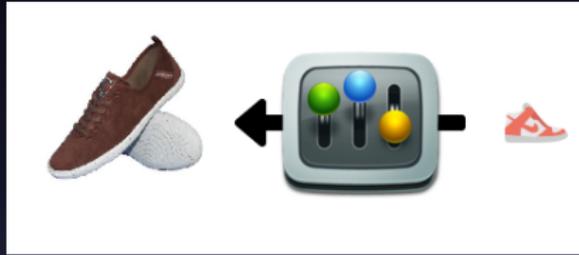
Calcul par réseau de neurones



Deux opérations :

- Somme pondérée des *inputs*
- Fonction non linéaire pour trouver un seuil d'activation

Entraînement par rétropropagation



À partir de la sortie désirée pour un certain *input*

- Dérivation des fonctions composées



caractéristiques
permettant
d'identifier
une image



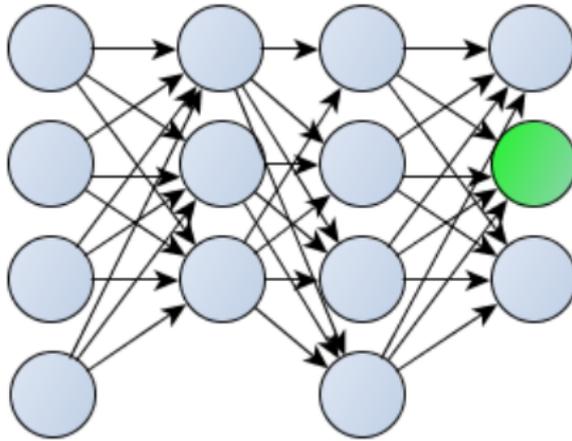
Qu'est-ce qui caractérise l'élément à analyser ?

i.e. Séquence sonore en reconnaissance de la parole :

- Hauteur ? ●
- Timbre ? ●
- Durée ? ●
- Intensité ? ●

Français ● – Espagnol ●, ● – Chinois ●, ●, ● – Anglais ●, ●

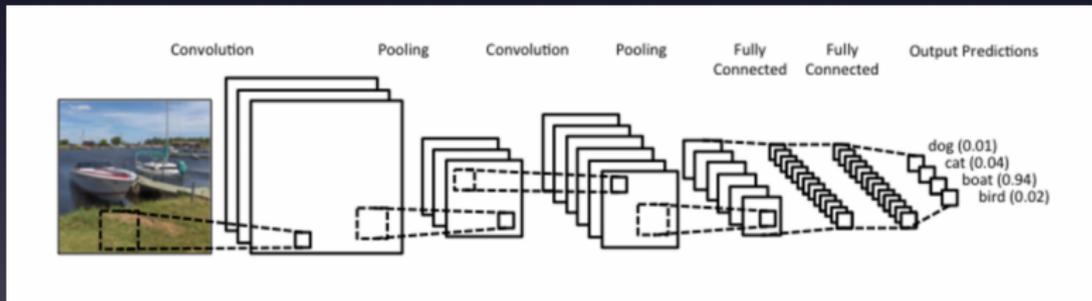
Deep Learning



Réseau de neurones à convolution (Fukushima 82, Le Cun 89)

Être robuste aux petites variations (*i.e* image : translation, rotations, déformations).

- Les “neurones” ne correspondent plus à des pixels, mais à des “régions” de localité.
- Un ensemble de “régions” est aggloméré en un seul élément, etc.



in Denny Britz 2015

⇒ Identifier *in fine* les **caractéristiques** de l'image.

Propriétés

- Avantage : pas besoin d'experts
- Inconvénient : besoin d'une très grosse quantité de données

Applications

- Reconnaissance de la parole (commande vocale)
- Catégorisation de l'image
- Reconnaissance du visage
- Classification de texte
- Reconnaissance d'obstacles
- *etc*

Ça marche pour la langue ?

- Oui assez bien pour :
 - Reconnaissance automatique de la parole
 - Synthèse vocale
 - Reconnaissance d'entités nommées
 - Analyse de sentiment
 - Mémoire de traduction
 - Production d'extraits
 - Indexation de contenu
 - Recherche d'information
 - Classification de documents
- Les résultats se font attendre pour :
 - Traduction automatique
 - Extraction d'information
 - Génération automatique de texte
 - Résumé automatique de texte

Ça marche pour la langue ?

- Assez bien pour :
 - Traitement de la parole
 - Traitement des grammaires locales

- Moins bien (sinon peu ou prou) pour :
 - Syntaxe
 - Sémantique
 - Pragmatique
 - Discours

Le langage

Arrêtons-nous à trois niveaux du langage :

- **Phonologie**

Reconnaissance de quelques dizaines de phonèmes, de leurs propriétés distributives.

- **Morphologie**

Reconnaissance de quelques dizaines de milliers de morphèmes (unités minimales qui font sens), de leurs propriétés distributives et combinatoires.

Les propriétés combinatoires s'articulent avec le lexique (quelques centaines de milliers d'entrées).

- **Syntaxe**

Organisation de l'ordre et de la structure des mots en relation avec le sens, le discours et les actes illocutoires visés. La syntaxe s'articule avec le lexique.

Ça marche pour la langue ?

- Ce qui *marche*
 - **Phonologie**
Phones ↔ Phonèmes
 - Reconnaissance automatique de la parole
 - Synthèse vocale
 - **Morphologie**
Grammaires locales (Automates finis)
 - Reconnaissance d'entités nommées
 - Analyse de sentiment
 - Mémoire de traduction
 - Production d'extraits
 - Indexation de contenu
 - Recherche d'information
 - Classification de documents
- Ce qui ne *marche* pas
 - **Syntaxe** pour le moins
 - Traduction automatique
 - Extraction d'information
 - Génération automatique de texte
 - Résumé automatique de texte

Apprentissage des langues

Pourquoi cette difficulté ?

- Caractéristiques (*Features*) introuvables par réseaux convolutifs ?
- Taille des données trop petite ?
- **Mais est-ce que la langue s'apprend ?**
Qu'est-ce qui est "appris" par un réseau de neurones ?
 - Ce qui s'apprend effectivement ?
 - Des universaux de langue ?

Et finalement : est-ce qu'un réseau de neurone est adapté à cette tâche ?

Apprentissage des langues

Paradoxe contre le béhaviorisme

Argument de la pauvreté du stimulus (*Noam Chomsky 1980*)

- Le langage auquel est exposé l'enfant est compatible avec de nombreuses grammaires.
- Les données auxquelles sont exposés les enfants ne sont pas celles qui permettent de distinguer ce qui est grammatical de ce qui ne l'est pas
- Par conséquent la grammaire ne découle pas des données auxquelles sont exposés les enfants.

Apprentissage des langues

Histoire racontée par Hérodote (V^e sc av jc)

- Le langage n'est pas inné.
- Paradoxe avec constat de Chomsky \Rightarrow comment apprend-on la langue alors ?
- La capacité à apprendre la langue est-elle innée ?

Apprentissage des langues

Piaget et Chomsky

- Piaget et Chomsky s'accordent sur ceci :
 - Le langage est un produit de l'intelligence ou de la raison et non pas d'un apprentissage behavioriste.
 - L'origine rationnelle du langage suppose l'existence d'un **noyau fixe**.

Donc passons sur la question de l'innéité et voyons ce **noyau fixe**.

Apprentissage des langues

Noyau fixe

Noam Chomsky 1957

- *The man is here. Is the man here ?*
- *The man will leave. Will the man leave ?*

H_1 Lire de gauche à droite et quand on rencontre *is*, *will*, *etc.*, transposer cette occurrence au début.

H_2 Lire de gauche à droite et quand on rencontre *is*, *will*, *etc.* qui suit un **groupe nominal**, transposer cette occurrence au début.

Apprentissage des langues

Paradoxe

Noam Chomsky 1957

- *The man who is here is tall. Is the man who is here tall ?*
- *The man who is tall will leave. Will the man who is tall leave ?*
- *Is the man who here is tall ?*
- *Is the man who tall will leave ?*

⇒ L'hypothèse H_2 prédit ces exemples et H_1 est réfutée par ces mêmes exemples.

Aucun enfant anglophone ne commet ces erreurs.

Alors comment un enfant sait H_2 ? Comment l'apprend-il ?

Apprentissage des langues

Noam Chomsky 1981

- *Each of the men likes the others.*

→

- *The men like each other.*
- *Each of the men expects (John to like the others.*
- *Each of the men was surprised at (John's hatred of the others.*
- *Each of the men liked (John's stories about the others.*

→

- *The men expect (John to like each other.*
- *The men wer surprised at (John's hatred of each other.*
- *The men liked (John's stories about each other.*

Apprentissage des langues

SSC :

...X... (... Y ...) ...

The men expect (John to like **each other**).

X (SUJET **X** **Y**)

- *The candidates expect (each other to win).*
- *The men heard (stories about each other.*

Apprentissage des langues

Un réseau de neurones est assurément fait pour :

- Reproduire un mécanisme d'apprentissage behavioriste
Reconnaissance de la parole, analyse de *sentiment*, etc.
- Développer une heuristique pour un algorithme complexe
Jeux, problème du sac-à-dos, etc.

Il est légitime de l'employer pour des apprentissages de *Features*

- Reconnaissance de visage
- Classification d'images
- Détection d'obstacles

Est-il légitime pour trouver les propriétés syntaxiques de la langue ?