

# Deep learning pour la Data Visualisation



Étudiants :  
D'Harréville Renaud et Jaunet Théo

Encadrant :  
Vuillemot Romain

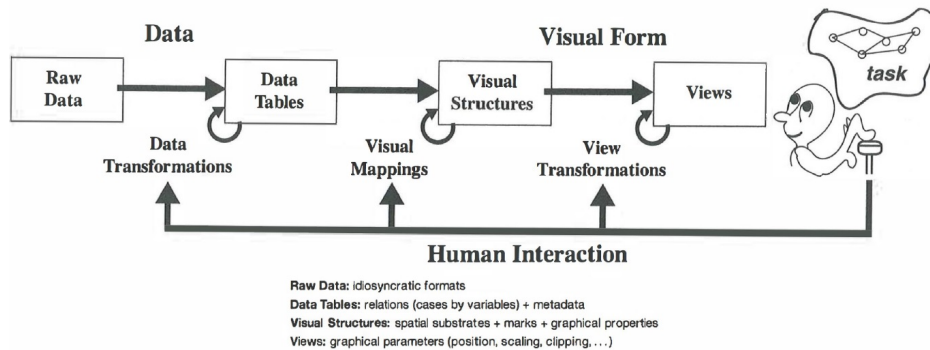
## Table des matières

<b>1</b>	<b>Contexte</b>	<b>2</b>
1.1	La visualisation des données . . . . .	2
1.2	Le Deep Learning . . . . .	2
<b>2</b>	<b>Problèmes ouverts en visualisation de données</b>	<b>3</b>
<b>3</b>	<b>Travaux attendus</b>	<b>3</b>
3.1	Études des méthodes de deep Learning . . . . .	3
3.2	Créations automatiques de visualisations à partir de données . . . . .	4
3.3	Transition entre les visualisations . . . . .	4
3.4	Transfert de style entre plusieurs visualisation . . . . .	4

# 1 Contexte

## 1.1 La visualisation des données

Dans un monde, où les données numériques sont de plus en plus importantes, la visualisation graphique de celles-ci devient un enjeu majeur pour la bonne compréhension, et la diffusion des informations qu'elles contiennent. Le problème étant complexe et en partie subjectif, il est difficile de créer une seule solution efficace pour toutes les situations rencontrées. Cependant, il est possible de retirer un certain nombre d'étapes permettant, à chacune de celles-ci, de bien construire un graphique. En voici un exemple illustré :



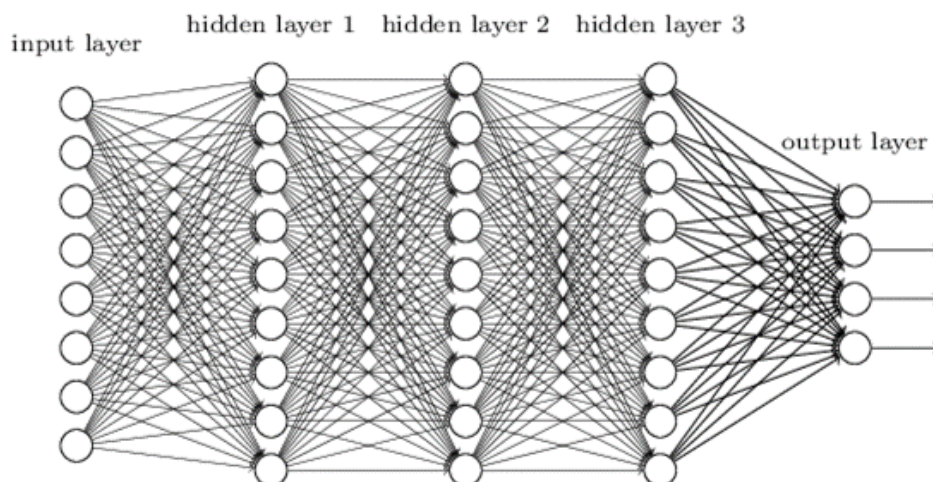
[Card, Mackinlay, Shneiderman, Readings in Information Visualization: Using Vision to Think, 1999]

- 1) En entrée du processus, il y a les données brutes (Raw Data), que l'on souhaite visualiser. Mais ces données sont en général trop brutes, et trop nombreuses pour que l'on puisse toutes les considérer. On doit donc faire une sélection de ce qui est utile.
- 2) Cette sélection nous envoie sur la deuxième étape (Data Tables).
- 3) De ces données, on définit 'à la main' la structure souhaitée de la visualisation, ce que l'on souhaite voir afficher, ainsi que la forme caractéristique de notre graphique.
- 4) Et c'est seulement à partir de là que la visualisation peut être générée.

## 1.2 Le Deep Learning

Depuis les derniers succès des recherches sur le Deep learning, ce domaine ne fait que prendre de l'ampleur. Il permet en effet de faire un certain nombre de tâches jusqu'alors très complexes dans le domaine de l'intelligence artificielle, de manière relativement simple. Sa vocation principale est à la classification de données mises en entrée. Les exemples classiques sont ceux de la classification des objets, ou la détection de visages. Le fonctionnement du deep learning est simple : Il reprend l'idée des neurones interconnectés dans le cerveau. Le réseau neuronal réel étant trop complexe, l'idée est d'en simuler un grâce à un système de couches successives de neurones virtuelles. (voir images ci-dessous)

# Deep neural network



<http://www.jtoy.net/2016/02/14/opening-up-deep-learning-for-everyone.html>

## 2 Problèmes ouverts en visualisation de données

En combinant donc les deux domaines il peut être intéressant de se pencher sur des problèmes de visualisation des données difficiles au moyen du deep learning. Voici donc certains problèmes difficiles qu'il est envisageable d'attaquer au moyen du deep learning :

- Mapping automatique de données et marques graphiques : [3]
- Transitions entre visualisation : [2]
- Compléter les données manquantes sur un graphique [1]
- Ajouter un style à une visualisation :  
<http://vis.berkeley.edu/papers/d3decon/>

## 3 Travaux attendus

Parmi les problèmes cités ci-dessus, Il a été choisi de se concentrer sur trois problèmes jugés principaux lors de ce projet. Les sous-chapitres ci-dessous présentent donc un peu plus en détail ces problèmes, ainsi qu'une première étape de prise en main et de compréhension de l'outil de deep learning.

### 3.1 Études des méthodes de deep Learning

Nous allons étudier le comportement d'un réseau neuronal sous contraintes d'une limite de noeuds en sortie. Pour cela, nous allons premièrement entraîner un réseau neuronal

a l'aide d'images de diagrammes en barres afin, d'en premier temps d'étudier le fonctionnement du deep learning, mais aussi pour en extraire les caractéristiques réduites. En effet, le réseau neuronal cherche à réduire le nombre de noeuds vers un nombre de plus en plus petit. Ce nombre peut être imposé. Cela permet d'étudier les caractéristiques, ressorties par le deep learning, qui différencient les styles de graphes.

### 3.2 Créations automatiques de visualisations à partir de données

Les possibilités offertes par le deep learning n'étant pas encore bien définies pour ce domaine, cette section laissera grandement place à de la recherche. Au fur et à mesure, l'objectif est de se rapprocher d'une génération bout à bout de la visualisation. Et cela va d'abord passer par la compréhension et l'analyse des possibilités offertes par du deep learning sur notre problème. Il faudra donc, plus précisément déterminer ce qui peut être fait lors des différentes étapes de la génération. Et ainsi, pouvoir en même temps manipuler les paramètres à chaque niveau, afin de se réajuster à nos préférences.

### 3.3 Transition entre les visualisations

Cette nouvelle visualisation est créée et paramétrée soit manuellement, soit par l'algorithme vu au dessus, avec d'autres paramètres. L'objectif de cette partie est de voir comment les transitions peuvent être faites de façon générique, sans passer par du code brut. Pour cette étape, il faut donc étudier les possibilités du deep learning notamment dans le domaine vidéo. En effet, il est envisageable que le deep learning puisse permettre de générer une transition d'une visualisation A vers une visualisation B et l'inverse.

### 3.4 Transfert de style entre plusieurs visualisation

Le transfert de style est un outil de deep learning, notamment en matière d'art. Plus particulièrement, il permet d'appliquer sur une peinture le style d'un autre tableau/image. Nous allons donc étudier l'application de cet outil dans le cas de visualisation de données afin de pouvoir découvrir les opportunités qu'il offre. Notamment pour respecter les codes couleurs, mises en formes, échelles mais également pour que deux graphes, à l'aide des caractéristiques primitives issues du deep learning, puissent être modifiés pour former un nouveau graphe avec le contenu du premier mais le style du second.

## Références

- [1] Cyntrica Eaton, Catherine Plaisant, and Terence Drisd. Visualizing missing data : Graph interpretation user study. In *IFIP Conference on Human-Computer Interaction*, pages 861–872. Springer, 2005.
- [2] Jeffrey Heer and George Robertson. Animated transitions in statistical data graphics. *IEEE transactions on visualization and computer graphics*, 13(6) :1240–1247, 2007.
- [3] Jock Mackinlay. Automating the design of graphical presentations of relational information. *Acm Transactions On Graphics (Tog)*, 5(2) :110–141, 1986.