

Introdução ao Aprendizado Não Supervisionado

Prof. Leandro Alvim, D.Sc.

Julho, 2016

Agenda

1. O que é
2. Motivações
3. Aplicações
4. Formalização
5. Exercícios

Aprendizado Não Supervisionado

Definition

Método de aprendizado que consiste de inferir estruturas latentes a partir de observações não rotuladas

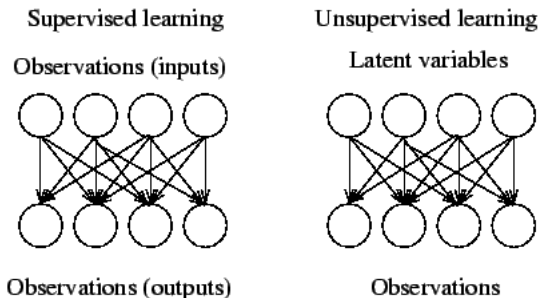


Figura: Diagrama de estrutura de dependência ¹

¹Figura extraída de

http://users.ics.aalto.fi/harri/thesis/valpola_thesis/node34.html

Principais Motivações

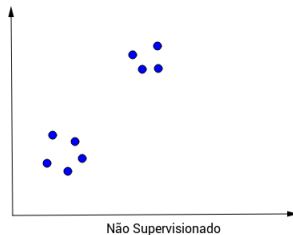
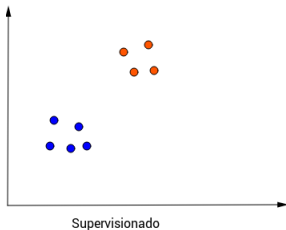
- Existe uma forma informativa de visualizar dados de *alta dimensão*?
- Como podemos descobrir subgrupos entre os exemplos?
- Como reduzir um elevado número de atributos?
- Como podemos aprender a partir de grandes volumes de dados não rotulados de forma a melhorar o aprendizado supervisionado?

Principais Motivações

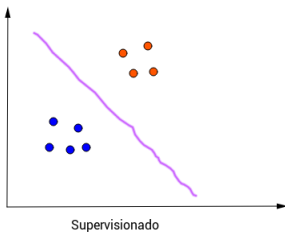
- Existe uma forma informativa de visualizar dados de *alta dimensão*? **Variáveis Latentes**
- Como podemos descobrir subgrupos entre os exemplos? **Agrupamento**
- Como reduzir um elevado número de atributos? **Redução de Dimensionalidade**
- Como podemos aprender a partir de grandes volumes de dados não rotulados de forma a melhorar o aprendizado supervisionado? **Deep Learning**

Exemplo

Exemplos não possuem **gabarito**



Classificação vs Agrupamento



Aplicações - Agrupamento de Palavras

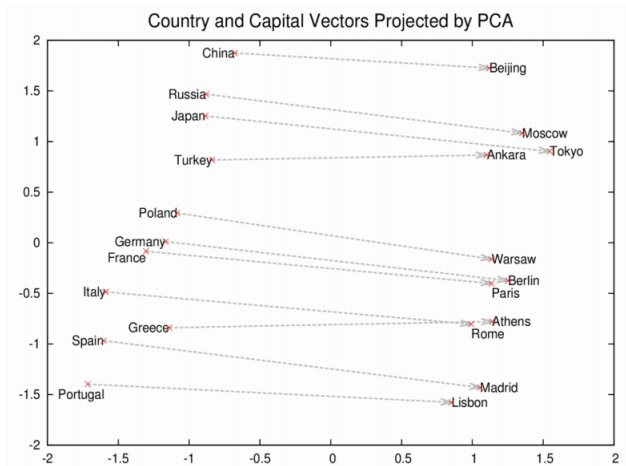


Figura: Método Word2Vec ¹

¹Figura extraída de DeepLearning4J <http://deeplearning4j.org/word2vec>

Aplicações - Geração Características

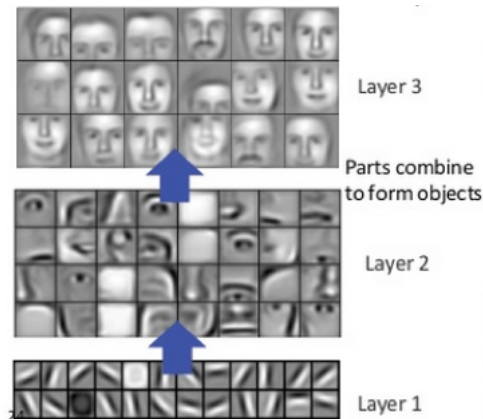


Figura: *Feature Hierarchy*²

²Figura extraída de DeepLearning4J

<http://deeplearning4j.org/neuralnet-overview.html>

Aplicações - Separação de Dados

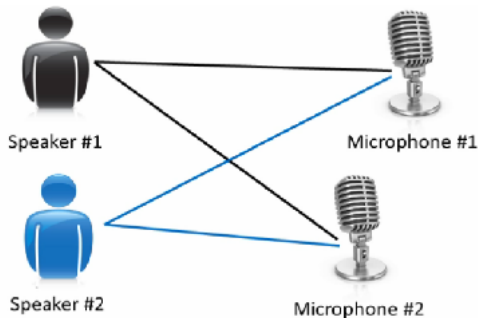


Figura: Cocktail Party Problem [3]

³Figura extraída de *What is unsupervised learning* <https://wingshore.wordpress.com/2014/11/03/what-is-unsupervised-learning-contd/>

Formalizando

Seja um conjunto de treino

$$\{x_1, x_2, \dots, x_n\}$$

$$\text{tal que, } x_i = \begin{bmatrix} 1 \\ a_1 \\ \cdot \\ \cdot \\ a_d \end{bmatrix}$$

em que a_i corresponde ao i -ésimo atributo do exemplo $x_i \in \mathbb{R}^{d+1}$.

Formalizando

Nosso objetivo é:

$$\operatorname{argmin}_z J(z)$$

em que z representa um conjunto de variáveis latentes e,

$$J(z) = \frac{1}{n} \sum_{i=1}^n \text{custo}(h_z(x_i))$$

- Descreva um problema prático que pode ser resolvido por agrupamento