

Elements of Machine Learning 2023

Lista 03

02.marzo.2023

En esta tarea se requiere realizar visualizaciones de *manifold learning* sobre dos conjunto de datos: (1) un conjunto de datos componentes en vinos; y (2) el *Social Happy Index* de países.

1. El conjunto de datos **wines.csv**, contiene información sobre varios componentes químicos asociados a tres diferentes diferentes cepas de vino.

Elija 3 métodos de *Manifold Learning* (t-SNE debe ser una de ellas), y realice visualizaciones en 2D o 3D para este conjunto. A partir de las visualizaciones, tratar de:

- Identificar alguna explicación para los ejes.
- Identificar posibles grupos o clústers.
- Verificar si dichos clúster separan bien a las tres diferentes cepas de vino.

A partir de sus visualizaciones, explicar las principales diferencias entre las tres cepas.

Más información sobre el origen de estos datos puede encontrarse en <http://archive.ics.uci.edu/ml/machine-learning-databases/wine>.

2. El índice de felicidad *Happy Planet Index* es un intento de medir el bienestar sostenible para todos, el cual puede encontrarse en el sitio <http://happyplanetindex.org/>.

Puedes descargar los datos del índice de felicidad 2016 en <http://happyplanetindex.org/resources> o también están disponibles en el archivo `hpi-data-2016.xlsx`.

Usar el método SOM para encontrar visualizaciones útiles para el conjunto de datos del índice de felicidad y de sus variables. Complementar estas visualizaciones con otros métodos, y hacer un análisis completo a partir de los datos.

Redactar su informe con sus análisis de datos, destacando sus conclusiones e *insights* más importantes. Preparar una presentación (de sólo 2 diapositivas), con un resumen informativo.