

# t-SNE

## THEORY

- 1) t-SNE stands for t-distributed stochastic neighbor embedding. It is a statistical method for visualizing high-dimensional data by giving each datapoint a location in a two or three-dimensional map.
- 2) Pairwise similarities between data points for high dimensional object is given by the following equation:

$$P_{j|i} = \frac{\exp(-||X_i - X_j||^2/2\sigma_i^2)}{\sum_{i \neq j'} \exp(-||X_i - X_{j'}||^2/2\sigma_i^2)}$$

- 3) For low dimensional object:

$$Q_{i,j} = \frac{(1 + ||Y_i - Y_j||^1)^{-1}}{\sum_{l,k} (1 + ||Y_l - Y_k||^2)^{-1}}$$

- 4) Conditional probabilities can be made symmetrical using

$$P_{i,j} = \frac{P_{i|j} + P_{j|i}}{2N}$$

- 5) Kullback-Leibler Divergence is given by:

$$D_{KL}(P, Q) = \sum_{i,j} P_{i,j} \log \left( \frac{P_{i,j}}{Q_{i,j}} \right)$$

## QUIZ

- 1) How the minimization of the Kullback–Leibler divergence with respect to the points is performed?
- 2) The bandwidth of the gaussian kernel is decided by which method?
- 3) There are how many stages of t-SNE?

## ANSWER

- 1) Using Gradient Descent
- 2) Bisection Method
- 3) 2