Date : / /
Page No.

TSNE (t-Distributed Stochastic Neighbourhood embedding)

→ State of the art / best dim-red technique for visualisation.

→ Where PCA preserves global shape/structure of data

But, TSNE preserves local structure data, means distance b/w two data points.

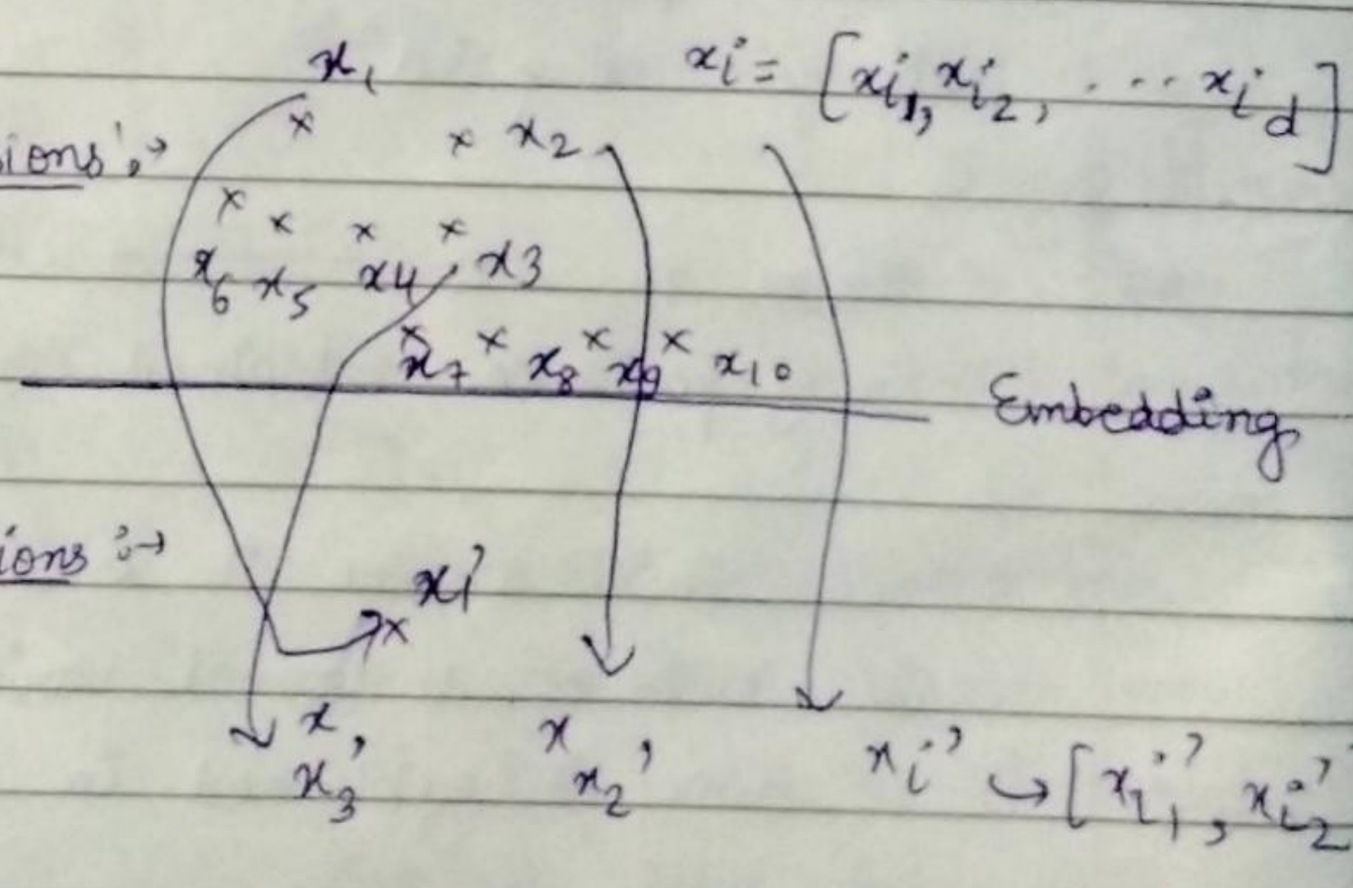
* Neighborhood :→

$N(x_i) = \{ x_j$, such that x_i & x_j are geometrically close $\}$

$|x_i - x_j|^2 = dist^2$

* Embedding :→

d dimensions :→



Embedding is to find corresponding data points for d-dimensions into 2 dimensions.

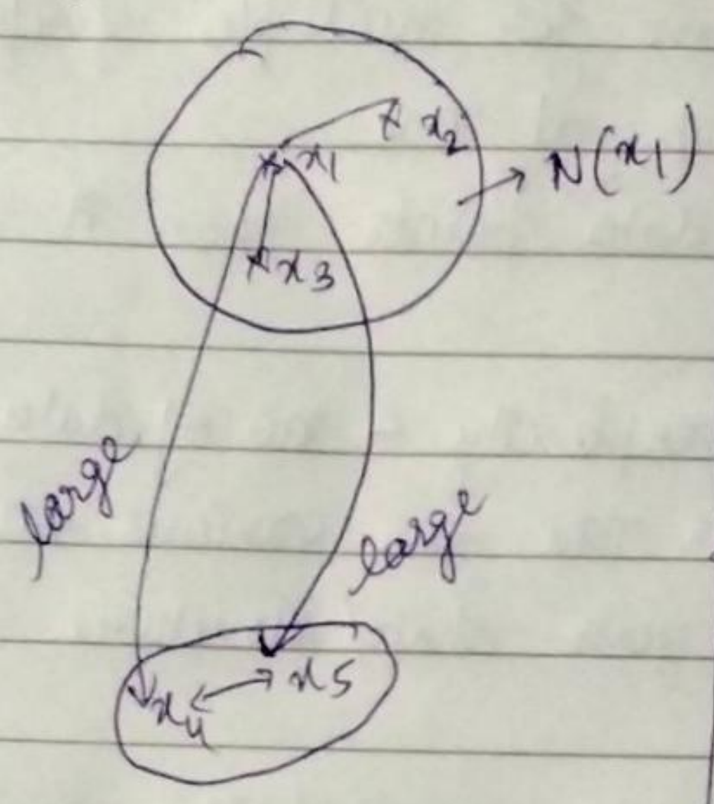
Date : / /
Page No.

* Geometric Intuition of TSNE :→

TSNE is a neighborhood preserving embedding.

d dimensions $x_i \in R^d$

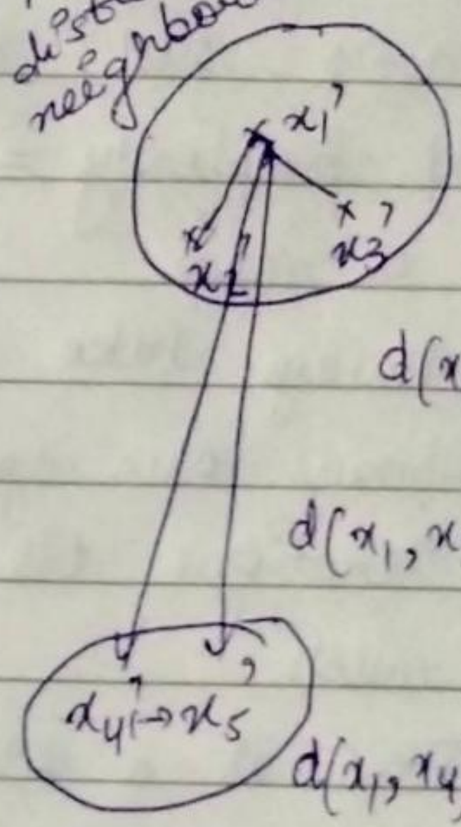
$N(x_1) = \{x_2, x_3\}$



$N(x_4) = \{x_5\}$

2 dimensions $x_i' \in R^2$

preserving in distance neighbourhood



$d(x_1, x_2) \simeq d(x_1', x_2')$

$d(x_1, x_3) \simeq d(x_1', x_3')$

$d(x_1, x_4) \not\simeq d(x_1', x_4')$

$d(x_4, x_5) = d(x_4', x_5')$

* Crowding - problem :→ Sometimes, specially in case of hypercubes, it is impossible to preserve distances b/w data points in 1-D.

* How to apply t-SNE ?

Steps :→ no of iterations → more the no. of iterations, the better is the solution.
Keep iteration till the shape not changes much means upto stable configuration.

Date : / /
Page No.

⇒ Perplexity → no of data points in the neighbourhood whose distance we want to preserve from a point on performing tSNE at a time to convert it into 2 dimensions.

* Points / lessons to keep in mind while running TSNE?→

1. Run the algorithm for multiple perplexity values, don't just stop at one.
2. If perplexity = no of data points then it will create a mess.
   Always take perplexity < no. of data points.
3. Always run your TSNE for various values of steps/ iterations till your shape/ structure doesn't move much.
4. TSNE is a stochastic/ probablistic analysis algorithm means everytime it runs it gives nearly slightly different results but a deterministic algorithm gives exactly same results.

drawback

5. TSNE internally expands dense clusters and shrinks sparse clusters to make them nearly similar dense.
   So, by looking at TSNE we cannot guess about the density of different clusters.

info

6. TSNE doesnot preserve distance bw the clusters.
7. Never conclude anything by just running tSNE for one or two values of perplexities, Specially smaller values of perplexities always run it for multiple values.

Date : / /
Page No.

* Topology :→ study of shapes.

Lesson
⇒ If you are running TSNE for same perplexity value multiple times and getting diff graphs then keep in mind, its not the correct value of perplexity to interpret anything.

But if the shape is roughly the same or just rotated then we can conclude than this perplexity value is stable.

* Epsilon :→ how fast your data will move on running TSNE.

TSNE
⇒ Points about MNIST dataset is→

Imp
1. TSNE group those numbers which visually look similar, for ex → it also group slant ones together and straight ones altogether.

visual similarity plays imp role here

* Random_state :→ It plays important role in ensuring that tSNE shows same output for multiple runs for fixed value of perplexity.

So, put random state = 0.

As, tSNE is a randomized Algorithm.