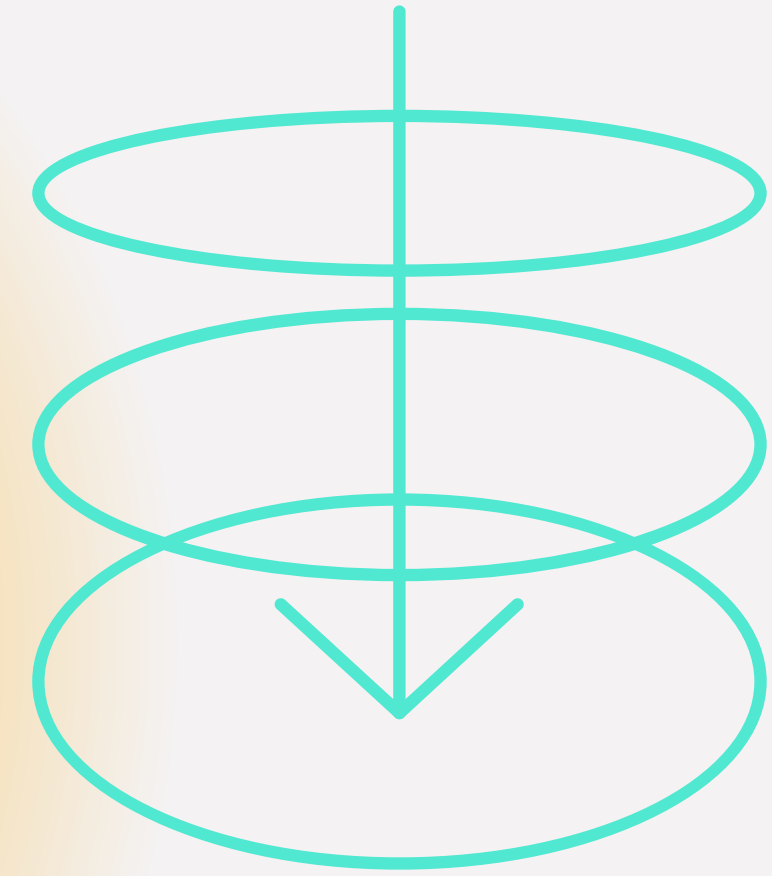


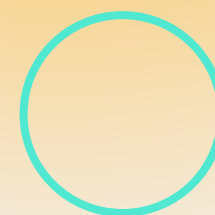
Introduction to

***t-SNE***

non-linear dimensionality  
reduction



Erubiel Tun Moo  
Fernando Rodríguez Zapata  
Braulio Millan Chi  
Michelle Cámara González

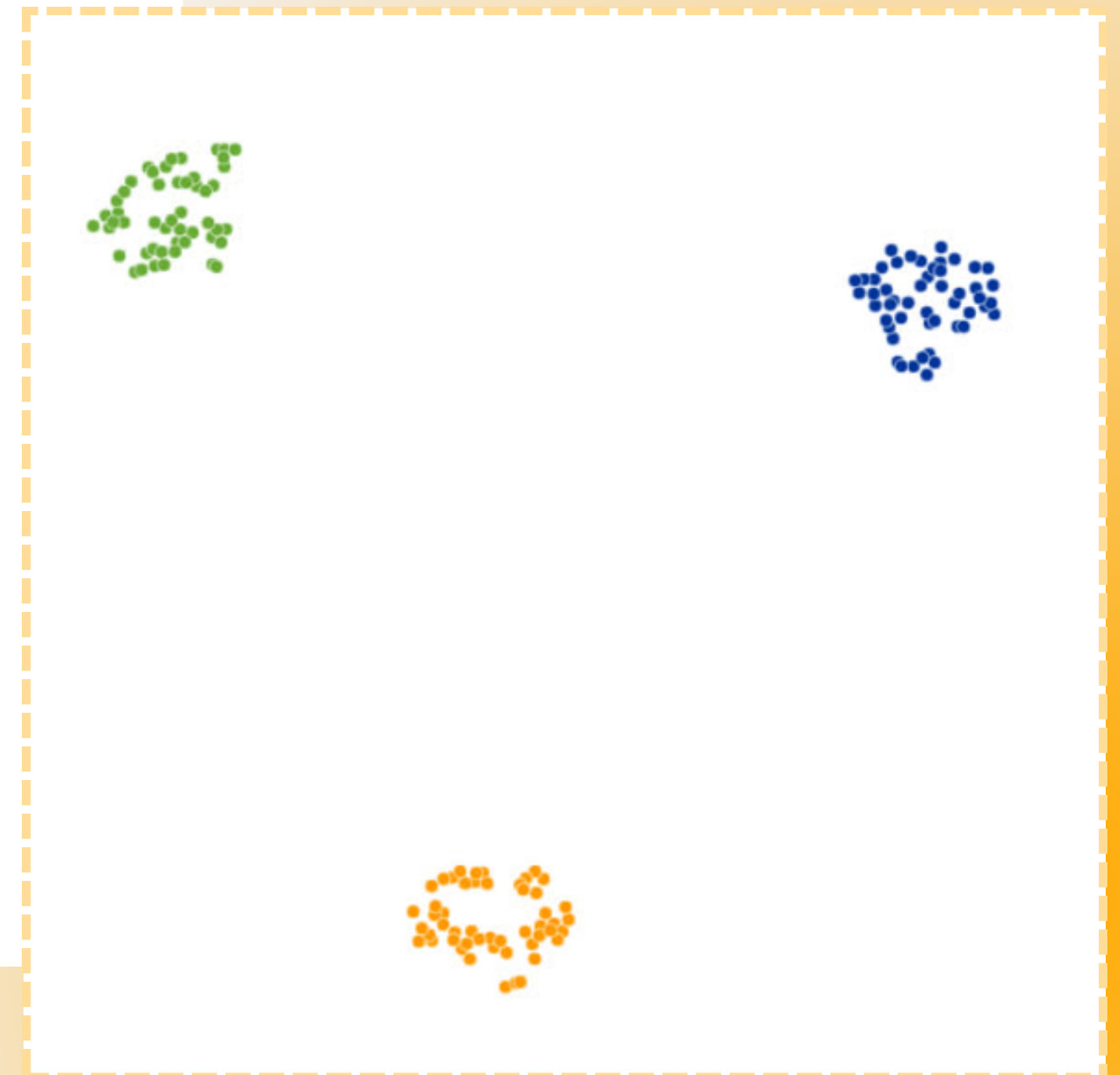


9°A

Team 5

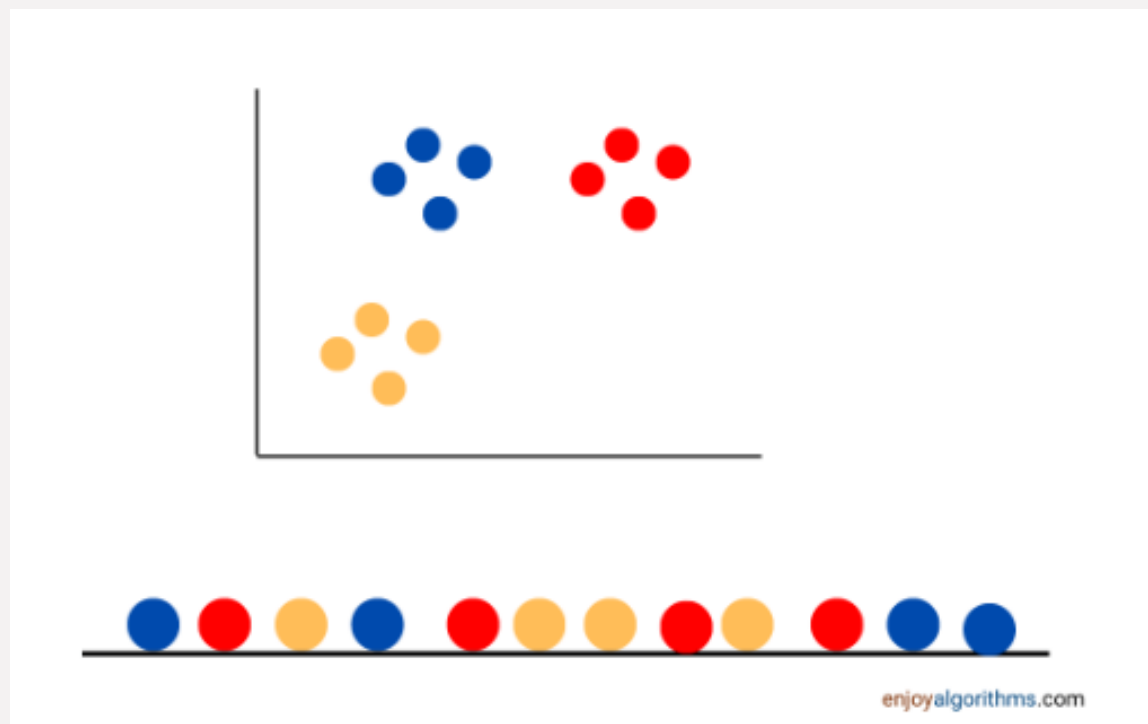
## 01 - Definition

- Unsupervised dimensionality reduction technique.
- For data exploration and visualizing high-dimensional data.
- It separate data that cannot be separate by a line.
- Easiest to use when features are numeric.



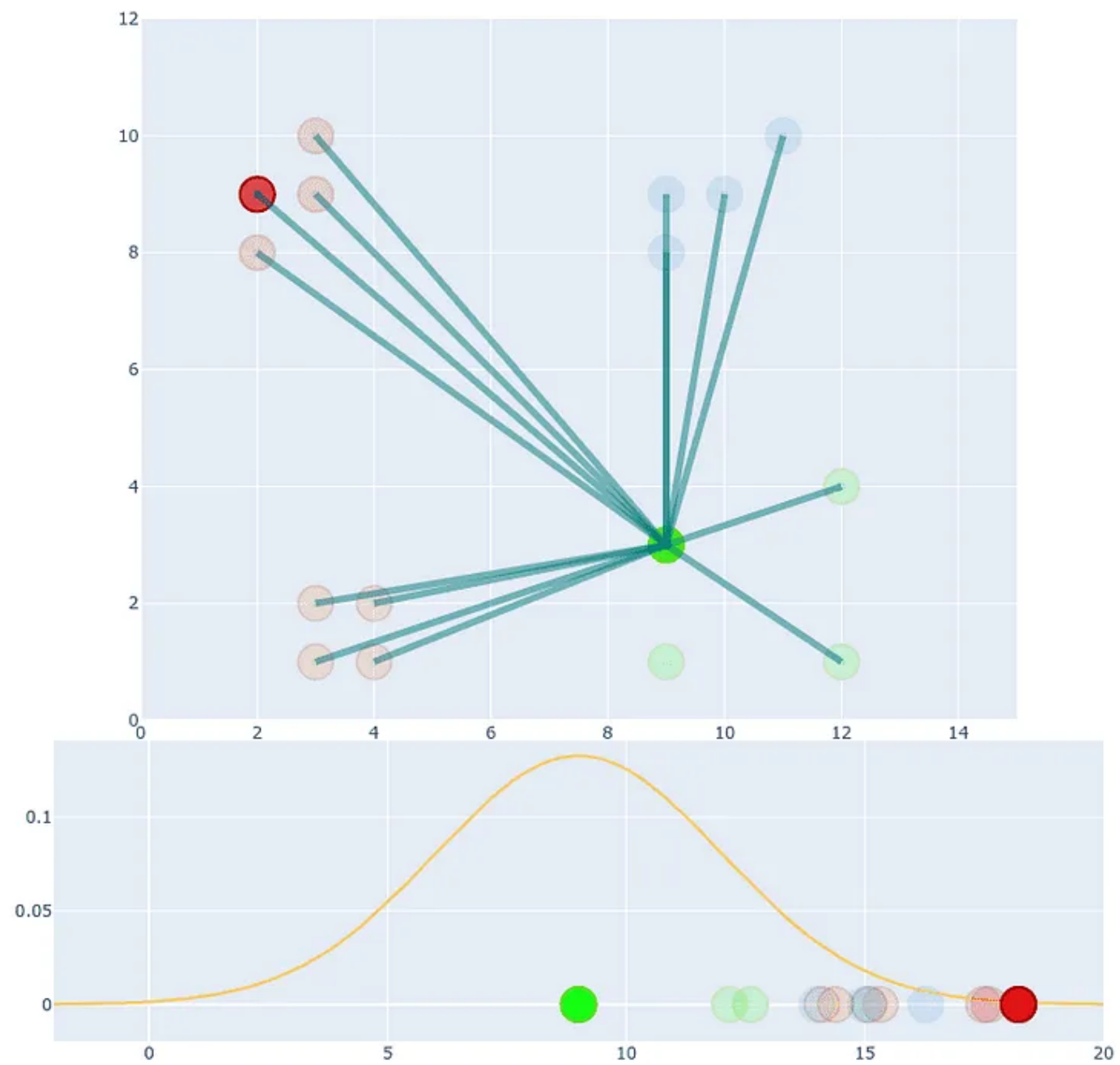
## 02 - How it works

1. Similarity Measure
2. Mapping to Lower Dimension
3. Divergence Minimization
4. Cluster Formation



## 03 - Steps

1. Calculate joint probabilities (Gaussian).
2. Low dimension space.
3. Calculate joint probabilities (t-distribution).
4. Calculate divergence.
5. Optimize.

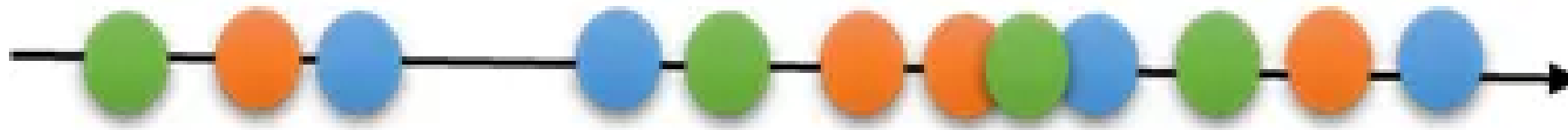


$$p_{j|i} = \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / 2\sigma_i^2)}$$

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}$$

# Gaussian distribution

# t distribution

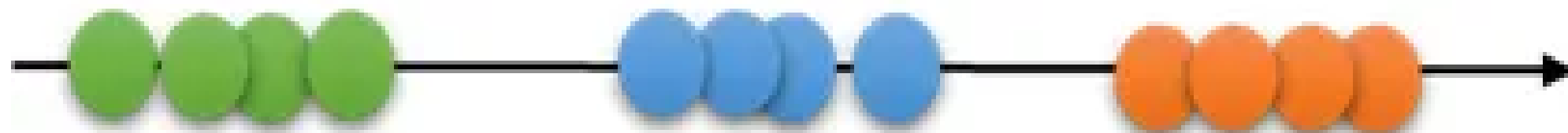


$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}}.$$

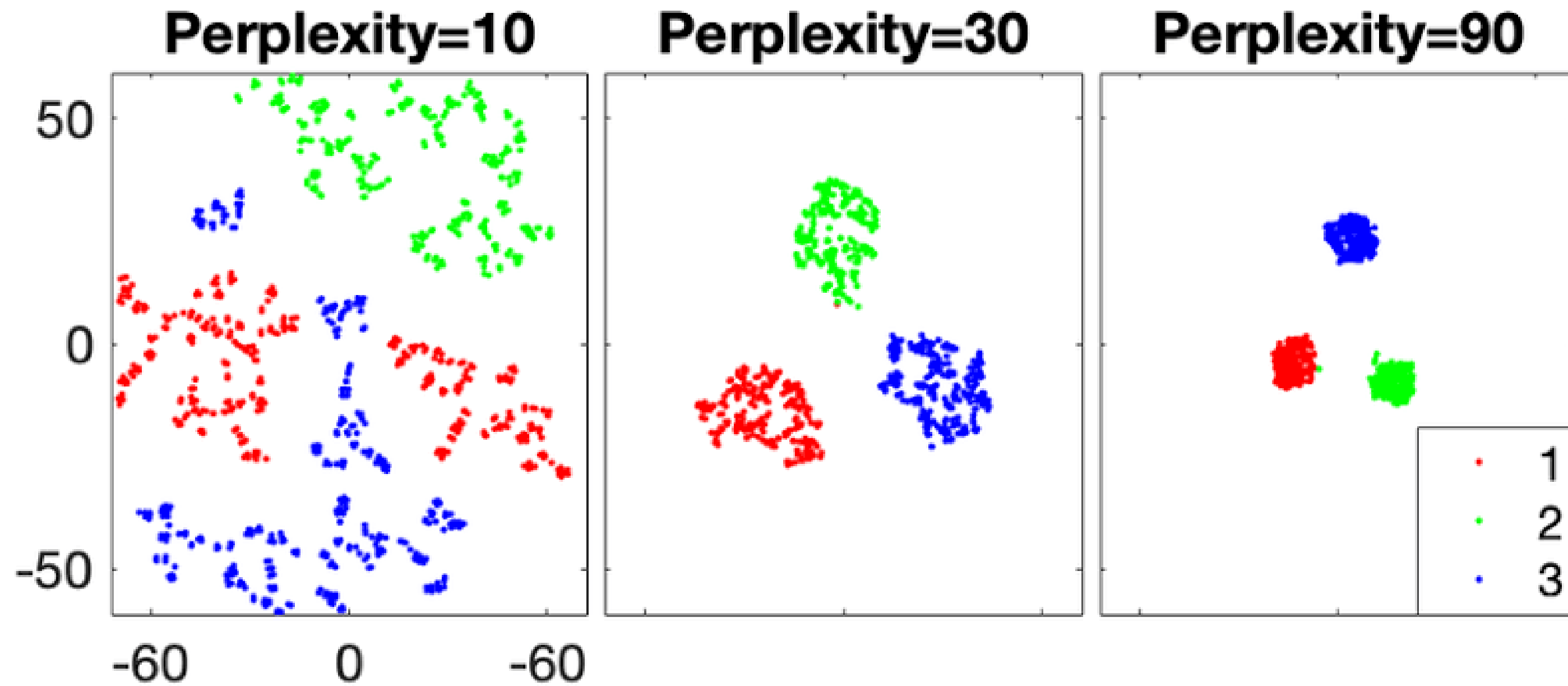
# Kullback-Leiber divergence

$$C = D_{\text{KL}}(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log \left( \frac{P(x)}{Q(x)} \right)$$

$$C = KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$$



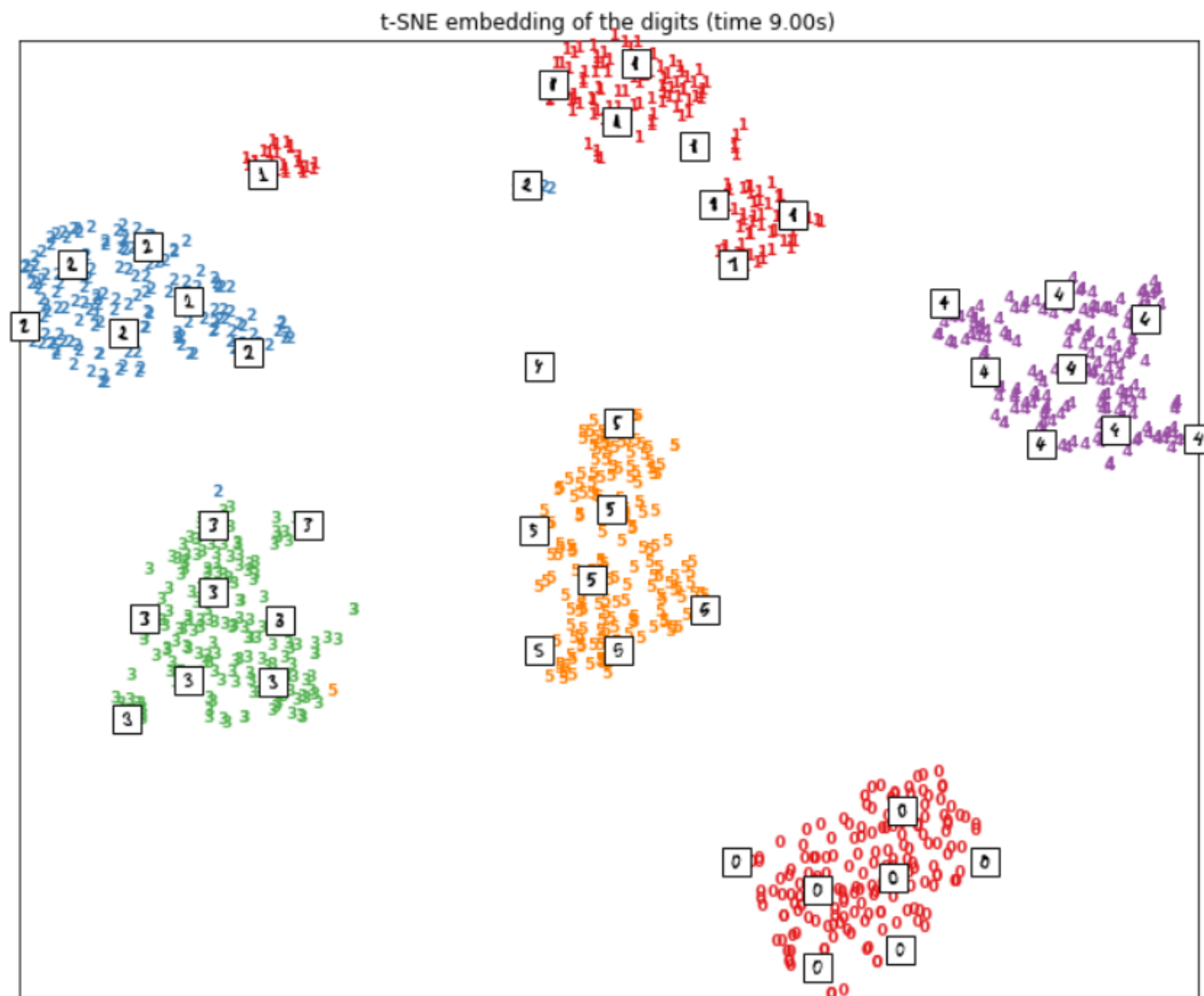
## 04 - t-SNE perplexity





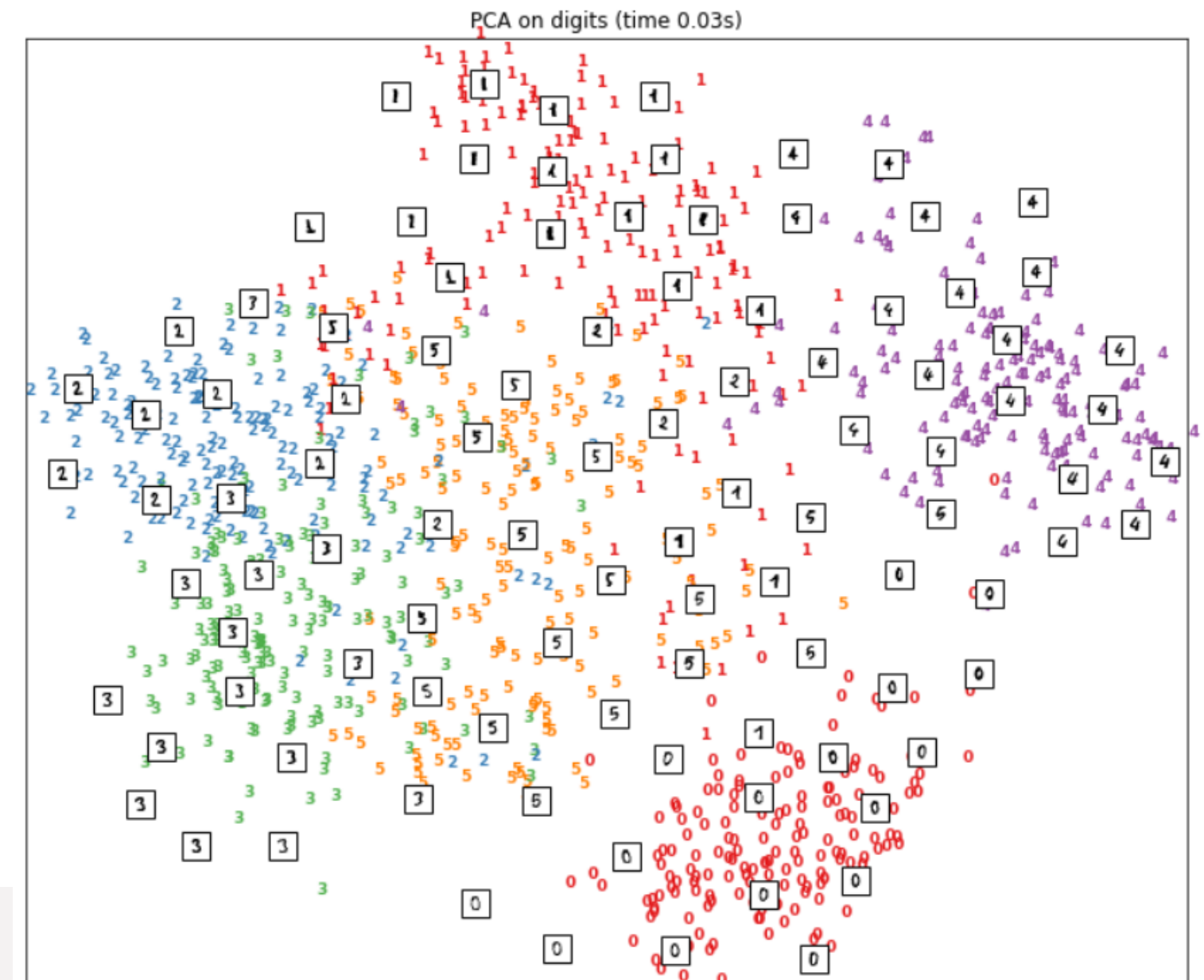
## 05 - TSN-E vs PCA

- Nonlinear technique.
- Lower dimensional space.



*Dimension reduction with the t-SNE method*

- Preserving large pairwise distances.
- Maximize variance.



*Dimension reduction with ACP*



## 06 - Advantages and Disadvantages

### Advantages

1. Handles Non-linear Data
2. Preserves Local Structure

### Disadvantages

1. Computational Complexity
2. Non-Deterministic



## 07-References

- Harshit. (n.d.). T-SNE algorithm in machine learning. Enjoyalgorithms.com. Retrieved 19 September 2023, from <https://www.enjoyalgorithms.com/blog/tsne-algorithm-in-ml>
- (N.d.). Datacamp.com. Retrieved 19 September 2023, from <https://www.datacamp.com/tutorial/introduction-t-sne>
- (burnpiro), K. E. (2020, April 13). t-SNE clearly explained. Retrieved September 19, 2023, from Towards Data Science website: <https://towardsdatascience.com/t-sne-clearly-explained>
- Soroker, A. (2020, August 4). T-SNE explained — math and intuition - the Startup - medium. Retrieved September 19, 2023, from The Startup website: <https://medium.com/swlh/t-sne-explained-math-and-intuition>

**THANKS!**