Helixco Cavity

Jaewoong Lee

Ulsan National Institute of Science and Technology

jwlee230@unist.ac.kr

2020-08-24

Overview

t-SNE

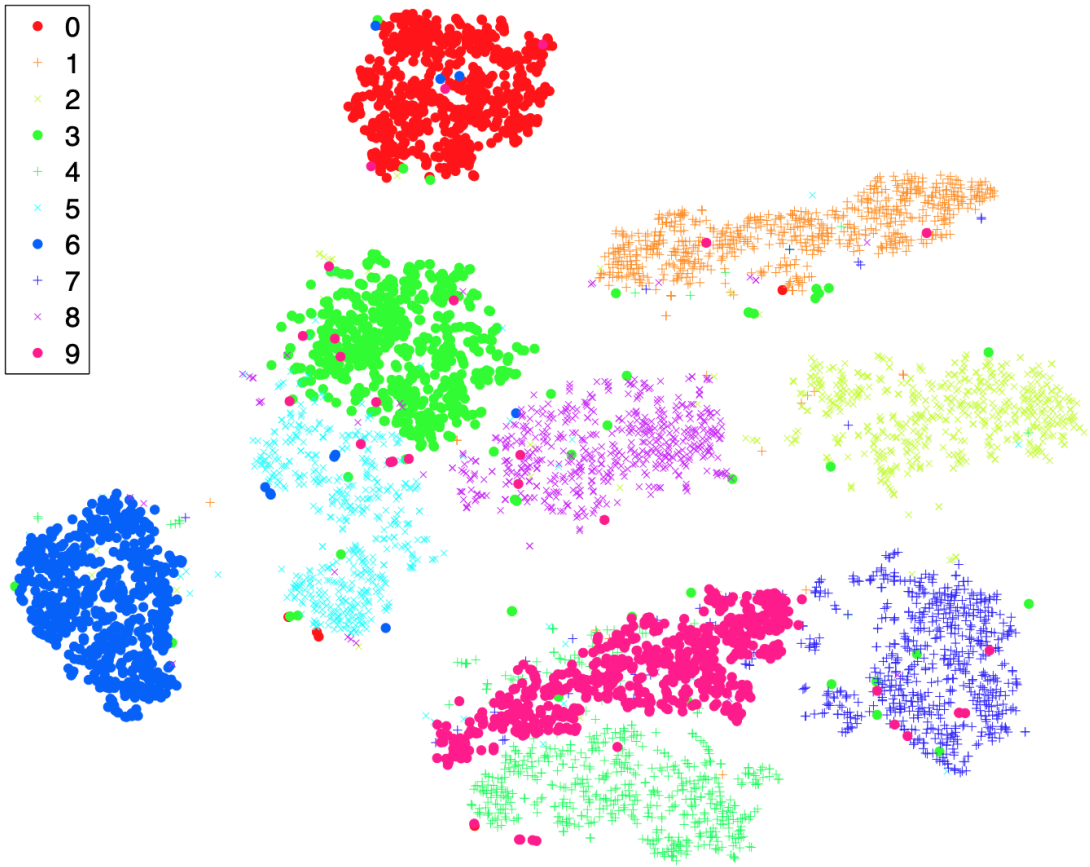


Figure: Visualizations of handwritten digits from the MNIST data set (Maaten & Hinton, 2008)

Programming Methods

- Docker (Merkel, 2014)
- QIIME 2
- Scikit-learn (Pedregosa et al., 2011; Buitinck et al., 2013)

QIIME 2 Workflow

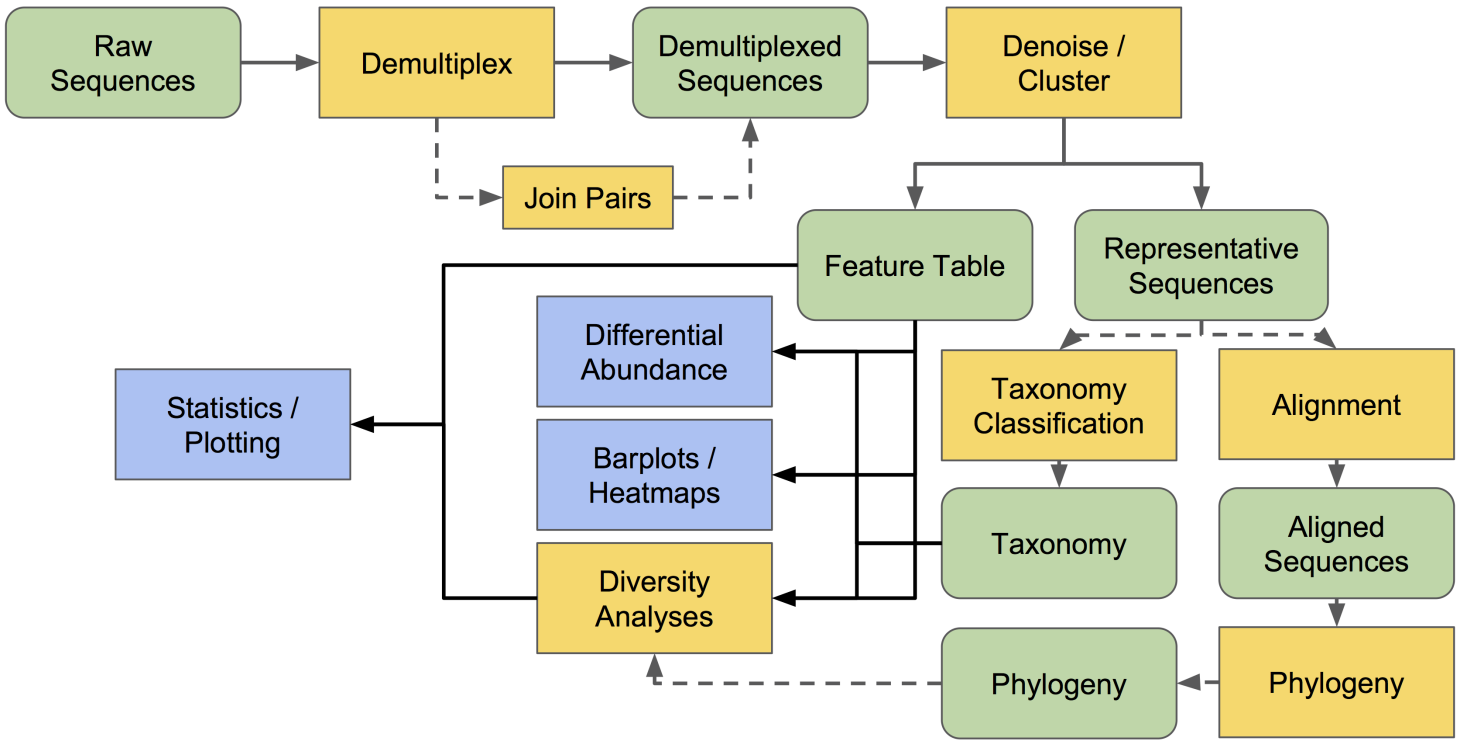
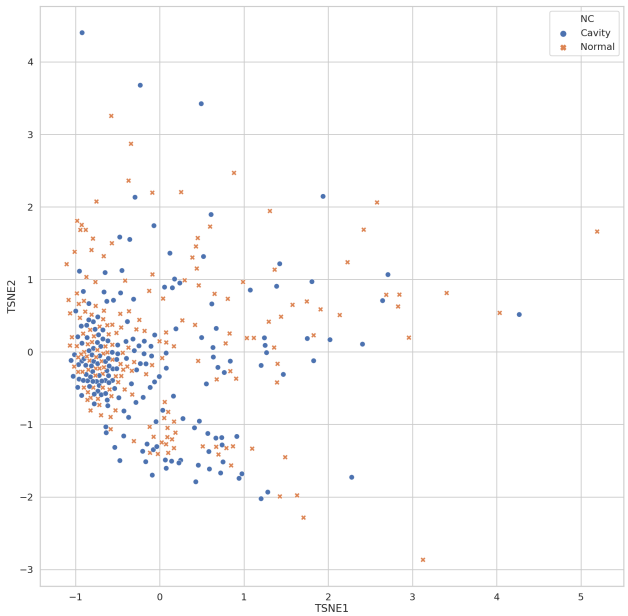


Figure: QIIME 2 Workflow

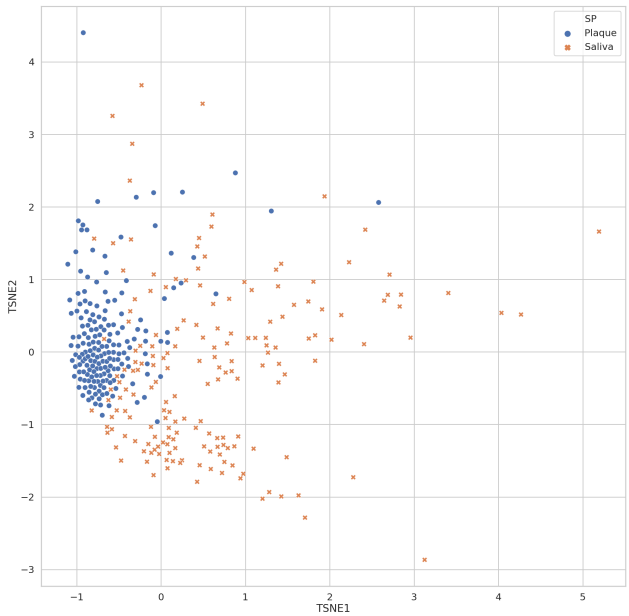
Yields I

- t-SNE with every bacterium

t-SNE with every bacterium



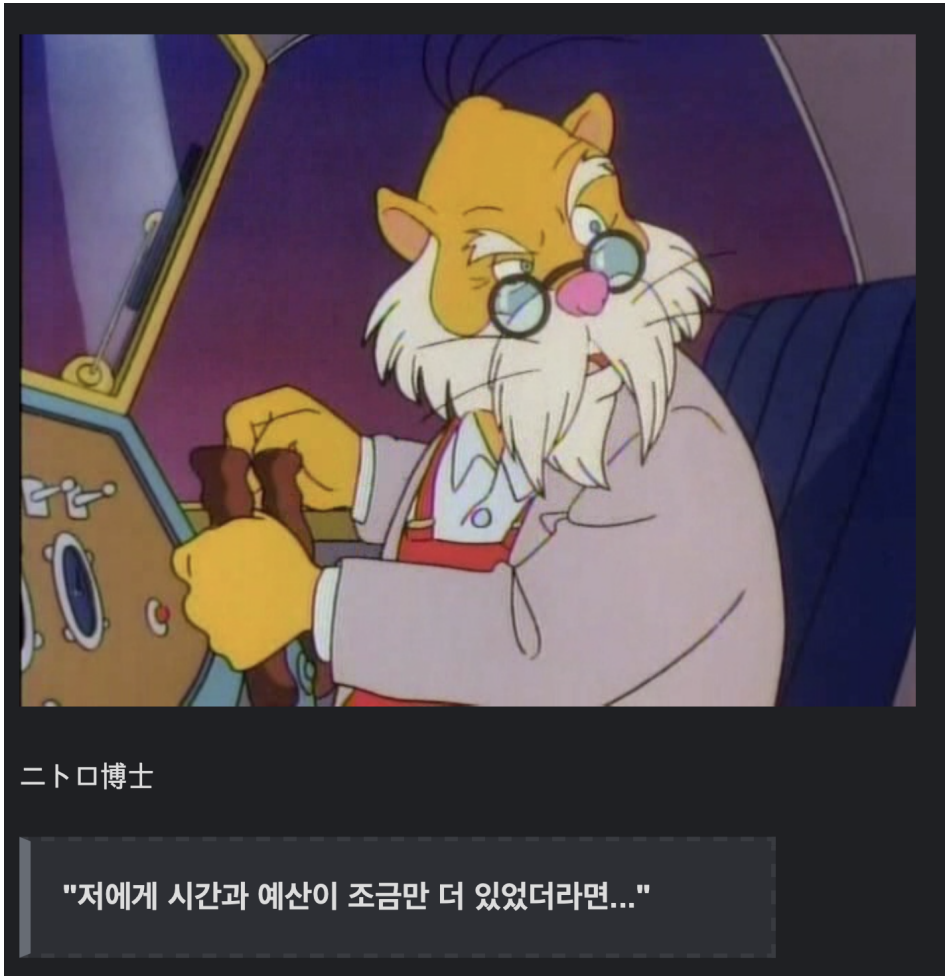
(a) Normal vs. Cavity



(b) Saliva vs. Plaque

$\therefore$ We need to select bacteria with feature importance.

Requirements I



Expectations I

- Improved classification

References I

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., ... Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *Ecml pkdd workshop: Languages for data mining and machine learning* (pp. 108–122).

Maaten, L. v. d., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, *9*(Nov), 2579–2605.

Merkel, D. (2014). Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, *2014*(239), 2.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.