Reliable Visual Analytics with Dimensionality Reduction: Quality Evaluation and Interpretation of Projections

Part 2:
Interpretation of Dimensionality Reduction Results

Hyeon Jeon[1], Takanori Fujiwara[2], Rafael M. Martins[3]

1 – Seoul National University
2 – Linköping University
3 – Linnaeus University

EuroVis 2025 Tutorial

EUROVIS 2025
LUXEMBOURG

What you will learn today

- Fundamentals: Interpretation of dimensionality reduction (DR) results
  - Linear DR
    - **Axis level**
  - Nonlinear DR
    - **Observed-pattern level**
      - Univariate focus
      - Composite variable focus
      - Classifier-based
      - Local pattern correlation
    - **Model-mechanism level**
      - Gradient-based
      - Parametric nonlinear DR
- Practices: Interpretations with existing libraries

Tutorial Materials

https://hyeonword.com/dr-tutorial/

Interpretation of dimensionality reduction (DR) results

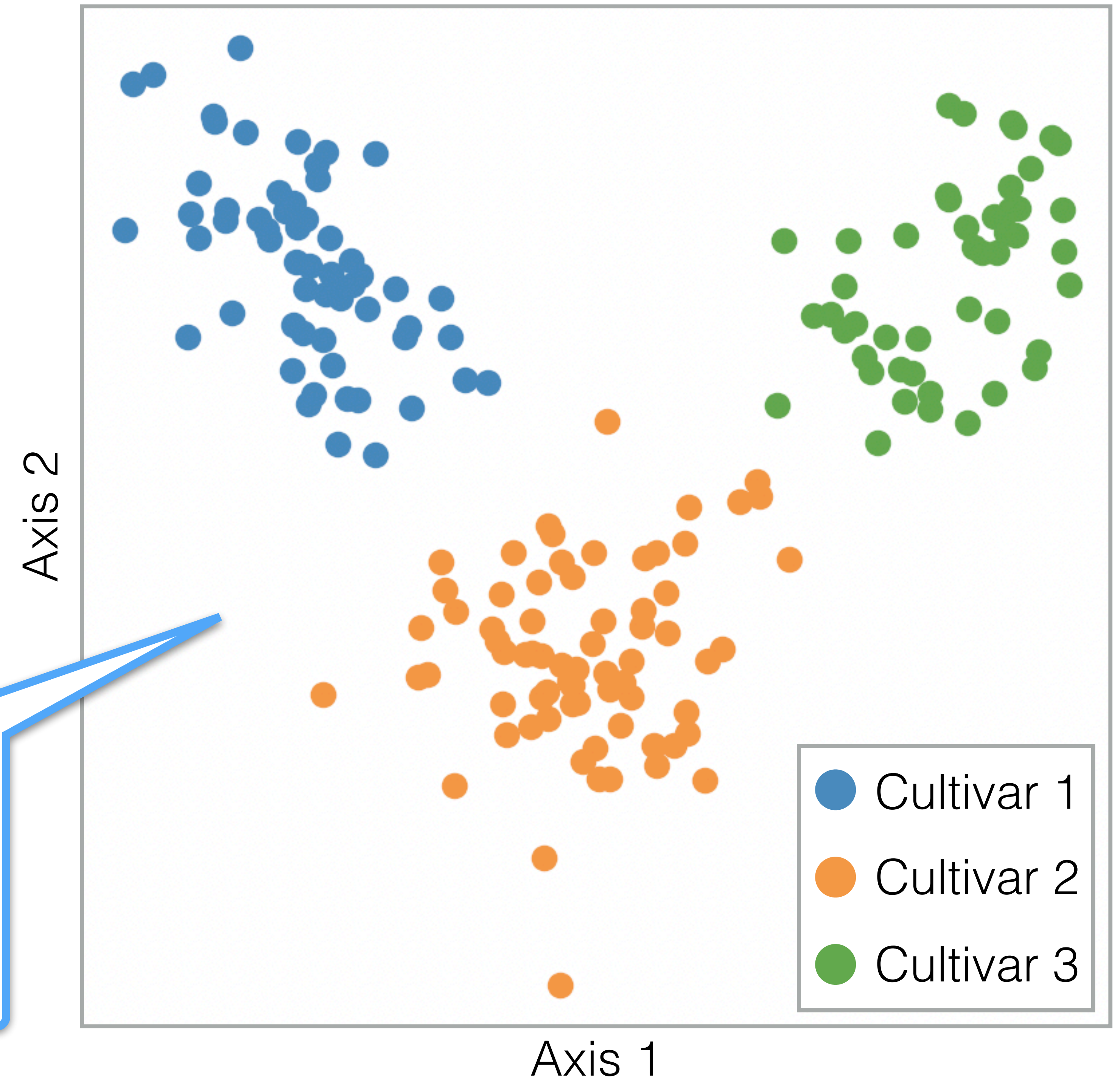
Wine dataset (13D)

	alchol	malic_acid	ash	alcalinity_ash	magnesium	total_phenols	flavanoid
0	1.52	-0.56	0.23	-1.17	1.91	0.81	1
1	0.25	-0.50	-0.83	-2.49	0.02	0.57	0.
2	0.20	0.02	1.11	-0.27	0.09	0.81	1.22
3	1.69	-0.35	0.49	-0.81	0.93	2.49	1.47
4	0.30	0.23	1.84	0.45	1.28	0.81	0.66
...	...	...	...	...	...	...	...

*displayed values are after the standardization

Linear discriminant analysis

LDA



How are they separated?
What are their differences?

- Interpreting a **lower-dimensional space** — **Linear DR**
- Interpreting based on **observed patterns** (e.g., clusters) — **Nonlinear DR**
- Interpreting from a DR **model/mechanism** level — **Nonlinear DR**

No single best way for the interpretation so far (especially, for nonlinear DR)

Interpretation of dimensionality reduction (DR) results

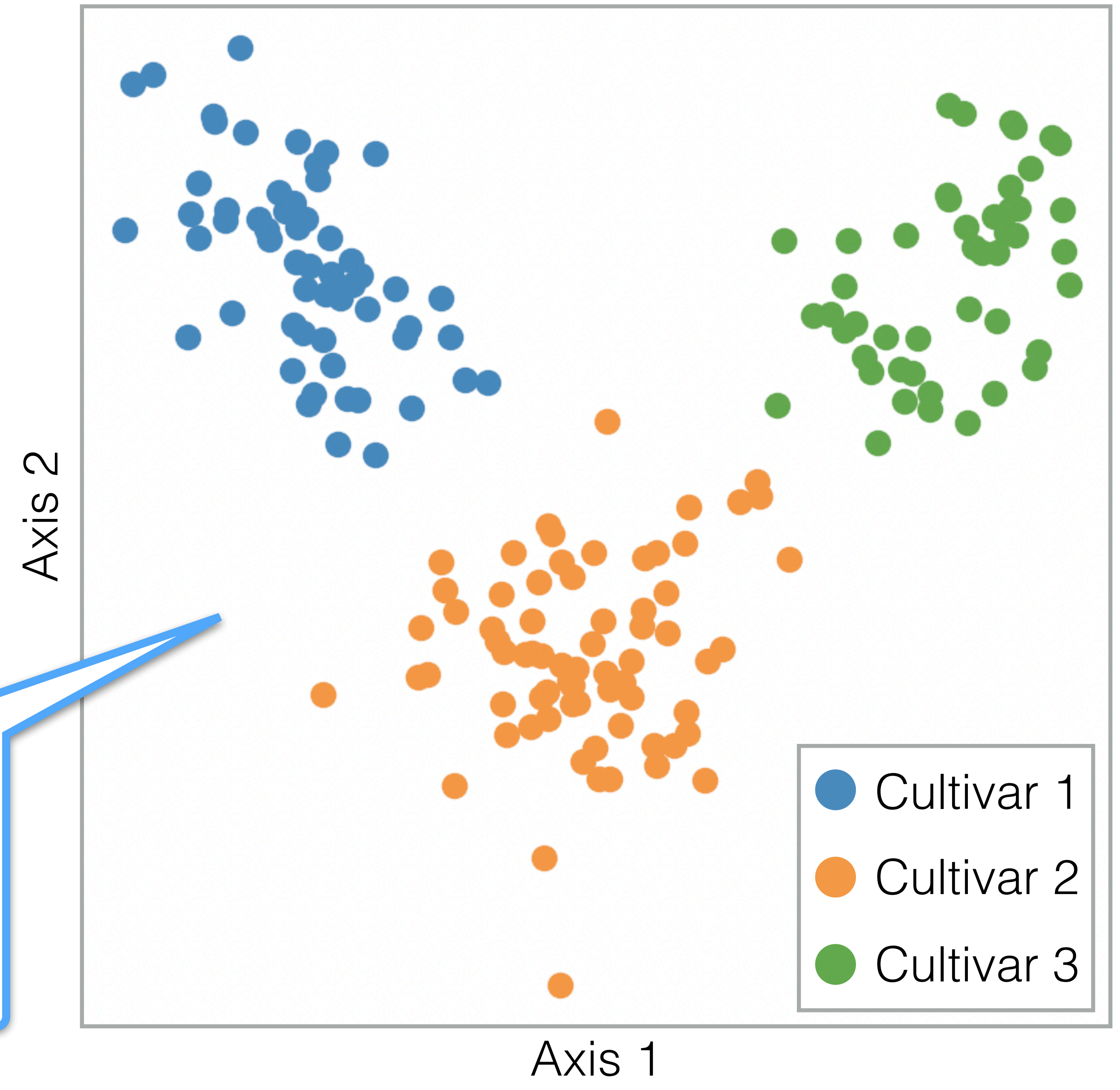
Wine dataset (13D)

	alchol	malic_acid	ash	alcalinity_ash	magnesium	total_phenols	flavanoid
0	1.52	-0.56	0.23	-1.17	1.91	0.81	1
1	0.25	-0.50	-0.83	-2.49	0.02	0.57	0.
2	0.20	0.02	1.11	-0.27	0.09	0.81	1.22
3	1.69	-0.35	0.49	-0.81	0.93	2.49	1.47
4	0.30	0.23	1.84	0.45	1.28	0.81	0.66
...	...	...	...	...	...	...	...

*displayed values are after the standardization

Linear discriminant analysis

LDA



How are they separated?
What are their differences?

- Interpreting a **lower-dimensional space** — **Linear DR**
- Interpreting based on **observed patterns** (e.g., clusters) — **Nonlinear DR**
- Interpreting from a DR **model/mechanism** level — **Nonlinear DR**

No single best way for the interpretation so far (especially, for nonlinear DR)

Linear DR

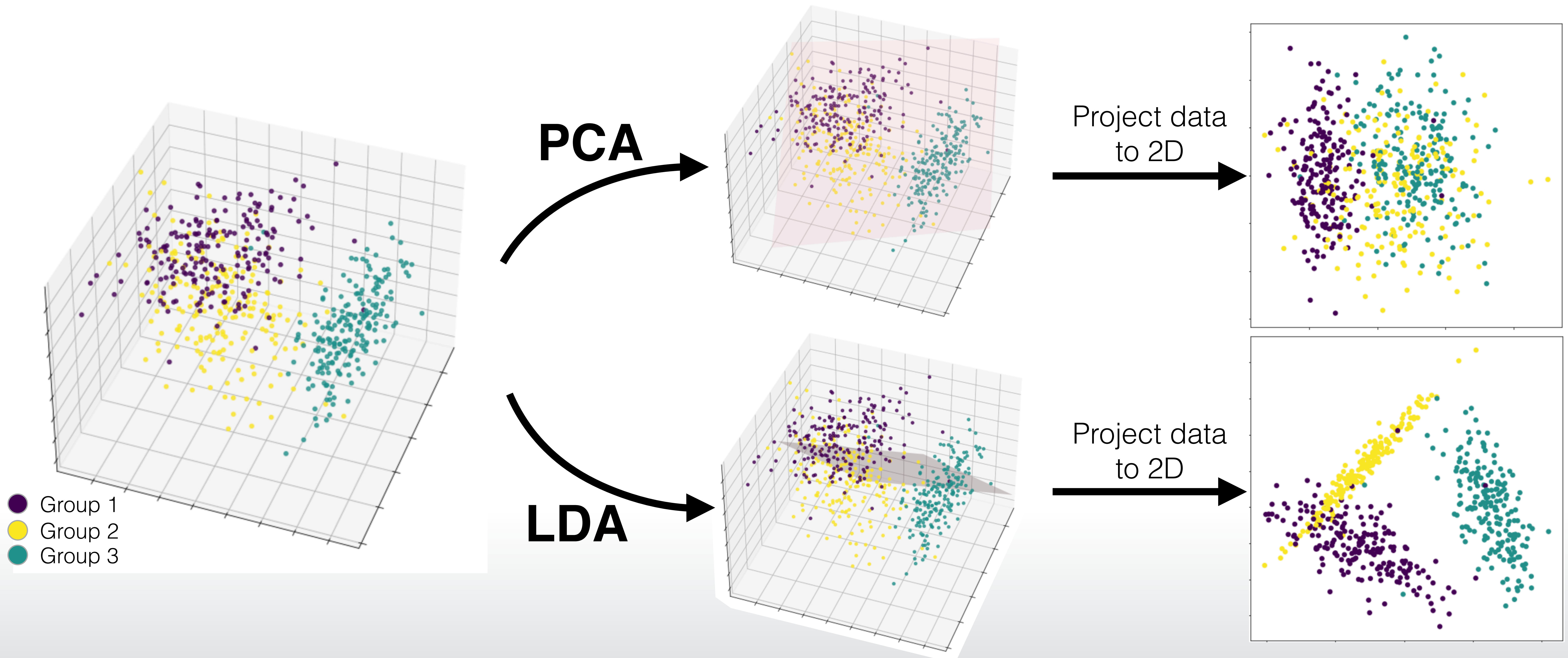
- Finds a linear projection from the original dimensions to low-dimensional axes

$$\mathbf{X} \cdot \mathbf{P} = \mathbf{Y}$$

$\mathbf{X} \in \mathbb{R}^{n \times d}$: original data, $\mathbf{P} \in \mathbb{R}^{d \times d'}$: projection matrix, $\mathbf{Y} \in \mathbb{R}^{n \times d'}$: projected data,

(n: # of instances, d: # of original dimensions, d': # of dimensions after projection)

- Representative methods
  - Principal component analysis (PCA) [Pearson,1901]: Preserves data **variance** as much as possible
  - Linear discriminant analysis (LDA) [Fisher, 1936]: Maximizes the **separation** of predefined groups



Interpreting the axes

Wine dataset (13D)

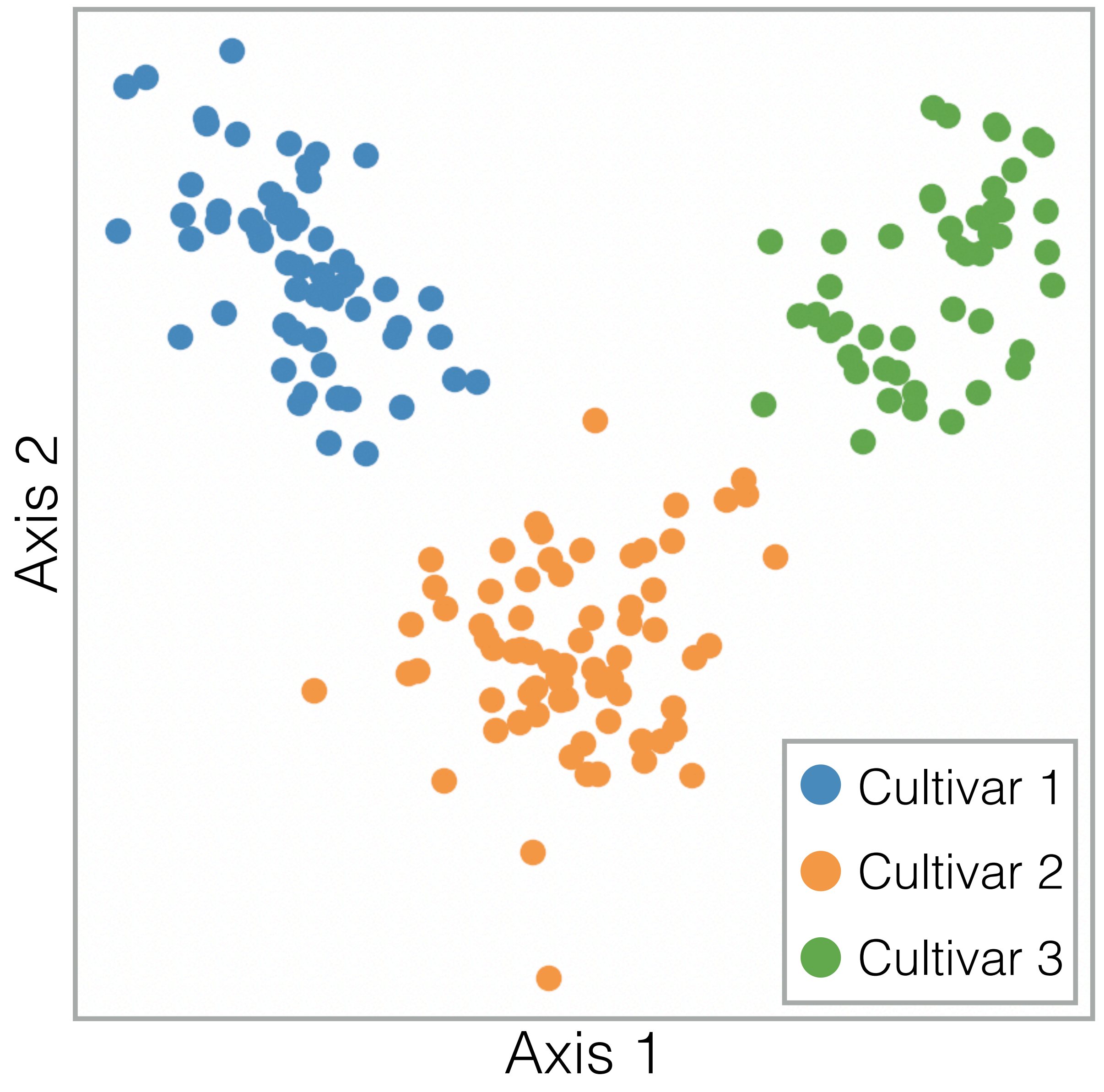
	alchol	malic_acid	ash	alcalinity_ash	magnesium	total_phenols	flavanoids	no
0	1.52	-0.56	0.23	-1.17	1.91	0.81	1.03	
1	0.25	-0.50	-0.83	-2.49	0.02	0.57	0.73	
2	0.20	0.02	1.11	-0.27	0.09	0.81	1.22	
3	1.69	-0.35	0.49	-0.81	0.93	2.49	1.47	
4	0.30	0.23	1.84	0.45	1.28	0.81	0.66	
...	...	...	...	...	...	...	...	

•

	axis 1	axis 2
alchol	-0.20	0.79
malic_acid	0.11	0.38
ash	-0.06	0.72
alcalinity_ash	0.31	-0.54
magnesium	-0.02	-0.01
total_phenols	0.23	-0.02
flavanoids	-1.00	-0.55
nonflav_phenols	-0.11	-0.23
proanthocyanins	0.05	-0.20
color_intensity	0.50	0.65
hue	-0.11	-0.39
od280/od315	-0.50	0.04
proline	-0.51	1.00

=

LDA result



Linear dimensionality reduction

$$\mathbf{X} \cdot \mathbf{P} = \mathbf{Y}$$

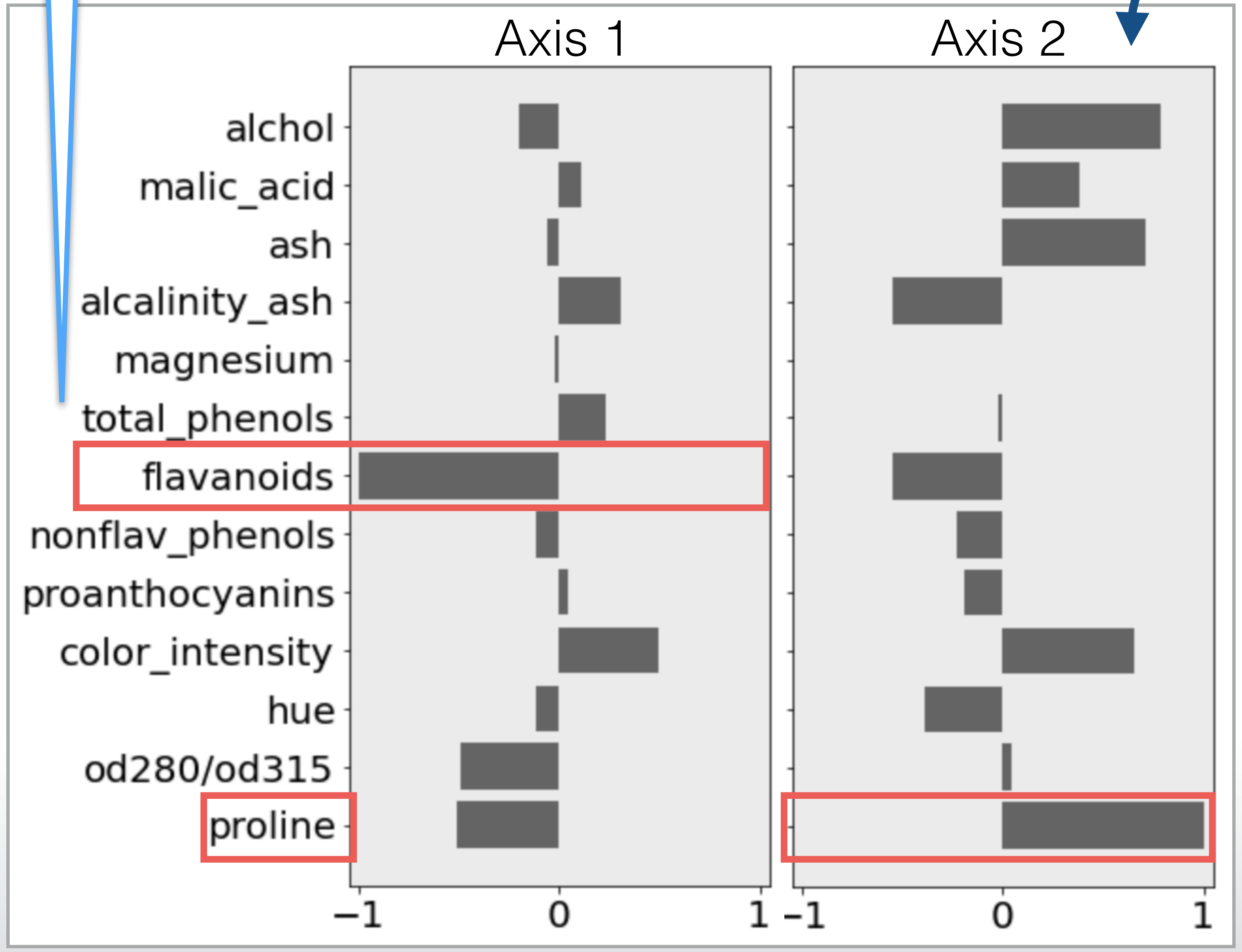
The projection matrix contains the information of the axes

For example, Axis 1 is generated by
-0.20 alcohol + 0.11 malic_acid - 0.06 ash …

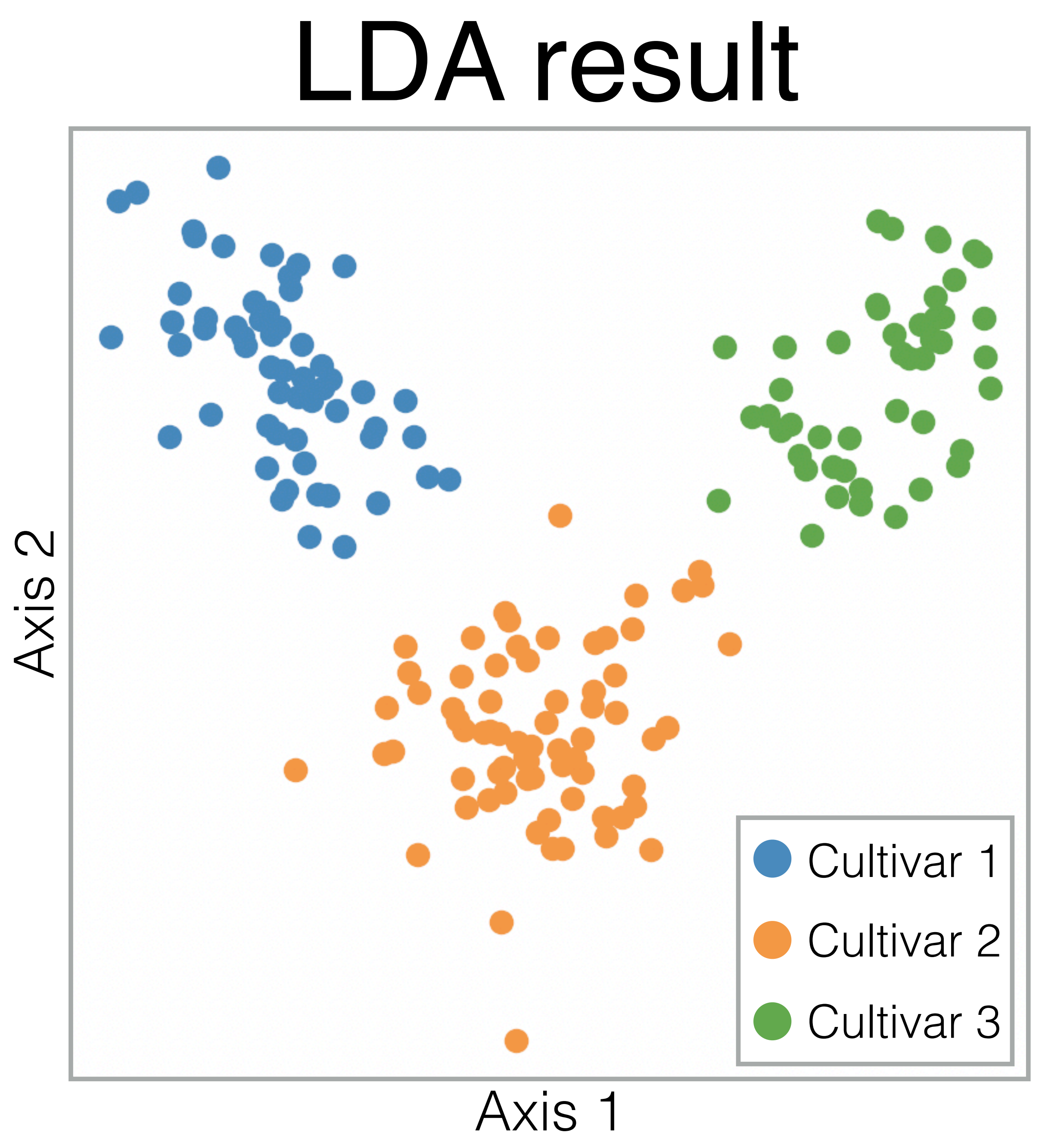
Interpreting the axes

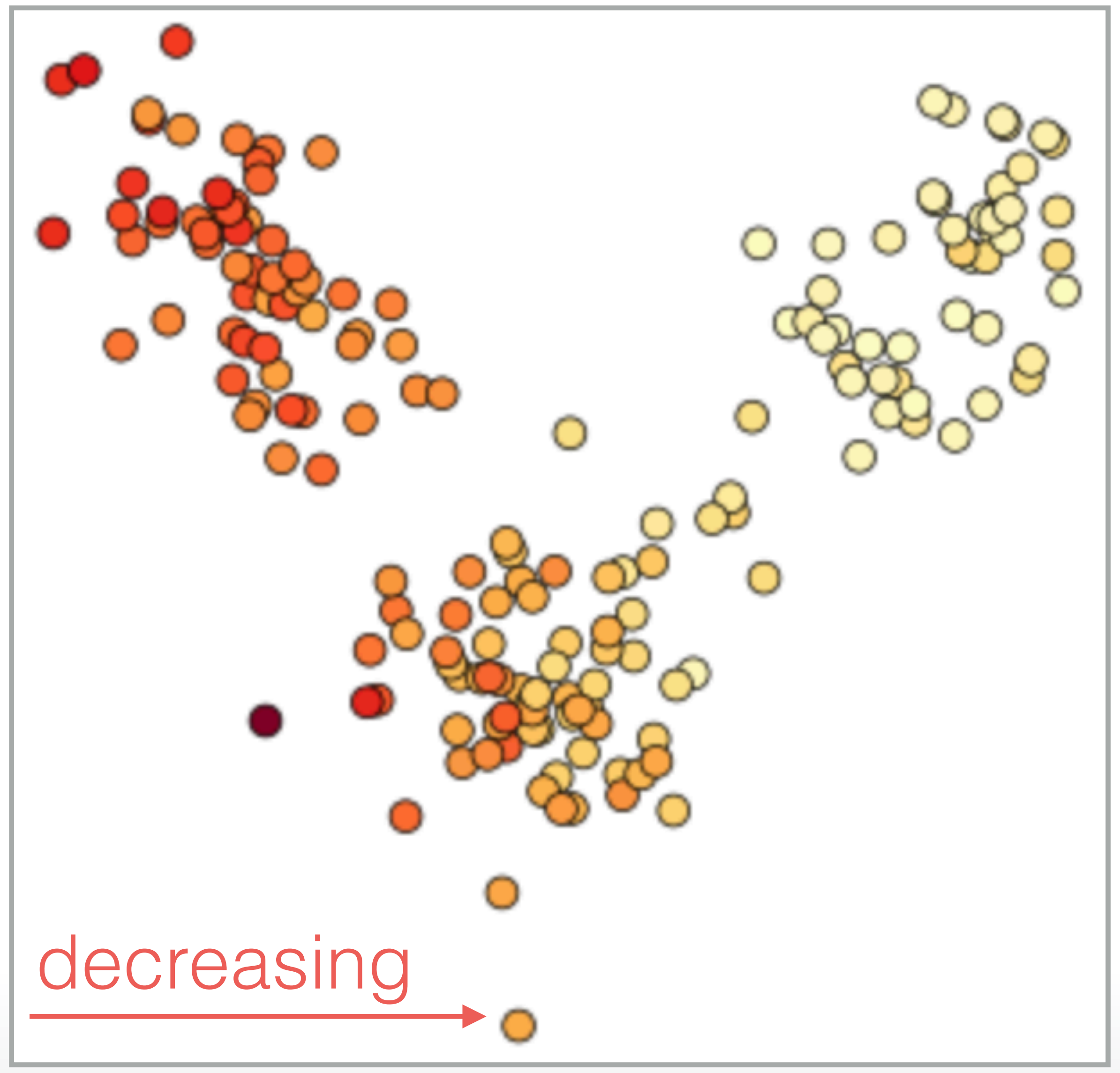
Axes 1 and 2 should be highly influenced by flavanoids and proline, respectively

Visualize

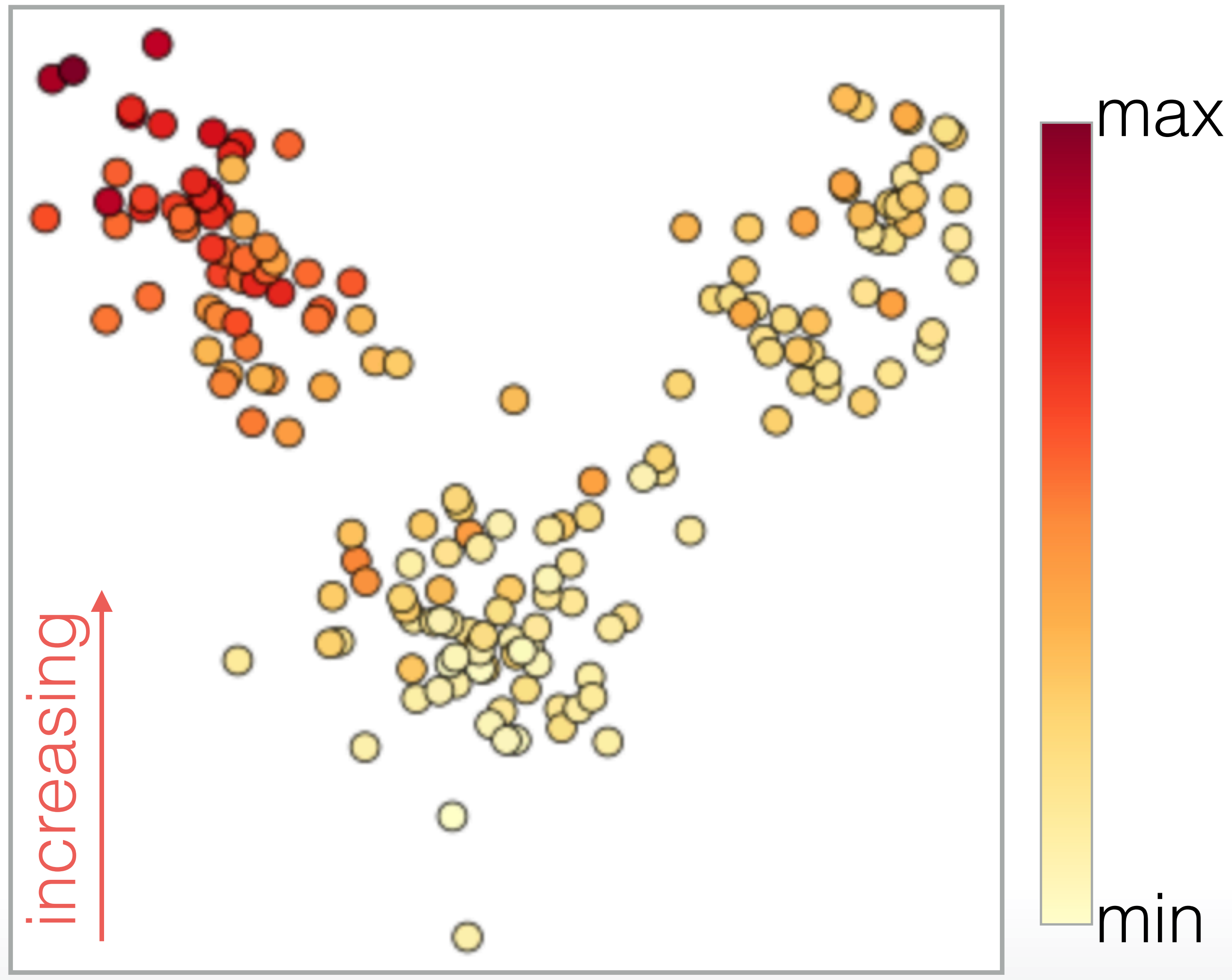


	axis 1	axis 2
alchol	-0.20	0.79
malic_acid	0.11	0.38
ash	-0.06	0.72
alcalinity_ash	0.31	-0.54
magnesium	-0.02	-0.01
total_phenols	0.23	-0.02
flavanoids	-1.00	-0.55
nonflav_phenols	-0.11	-0.23
proanthocyanins	0.05	-0.20
color_intensity	0.50	0.65
hue	-0.11	-0.39



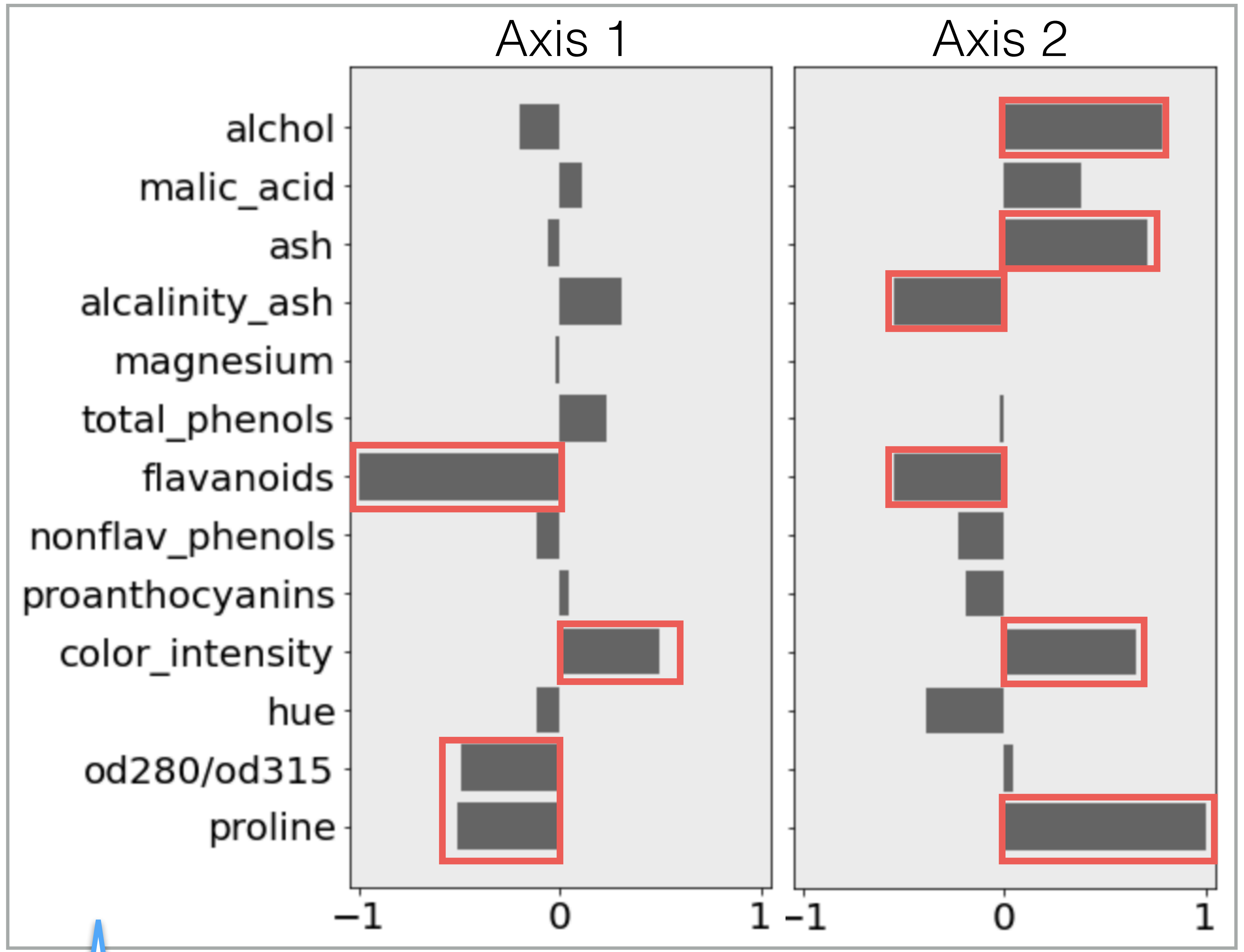


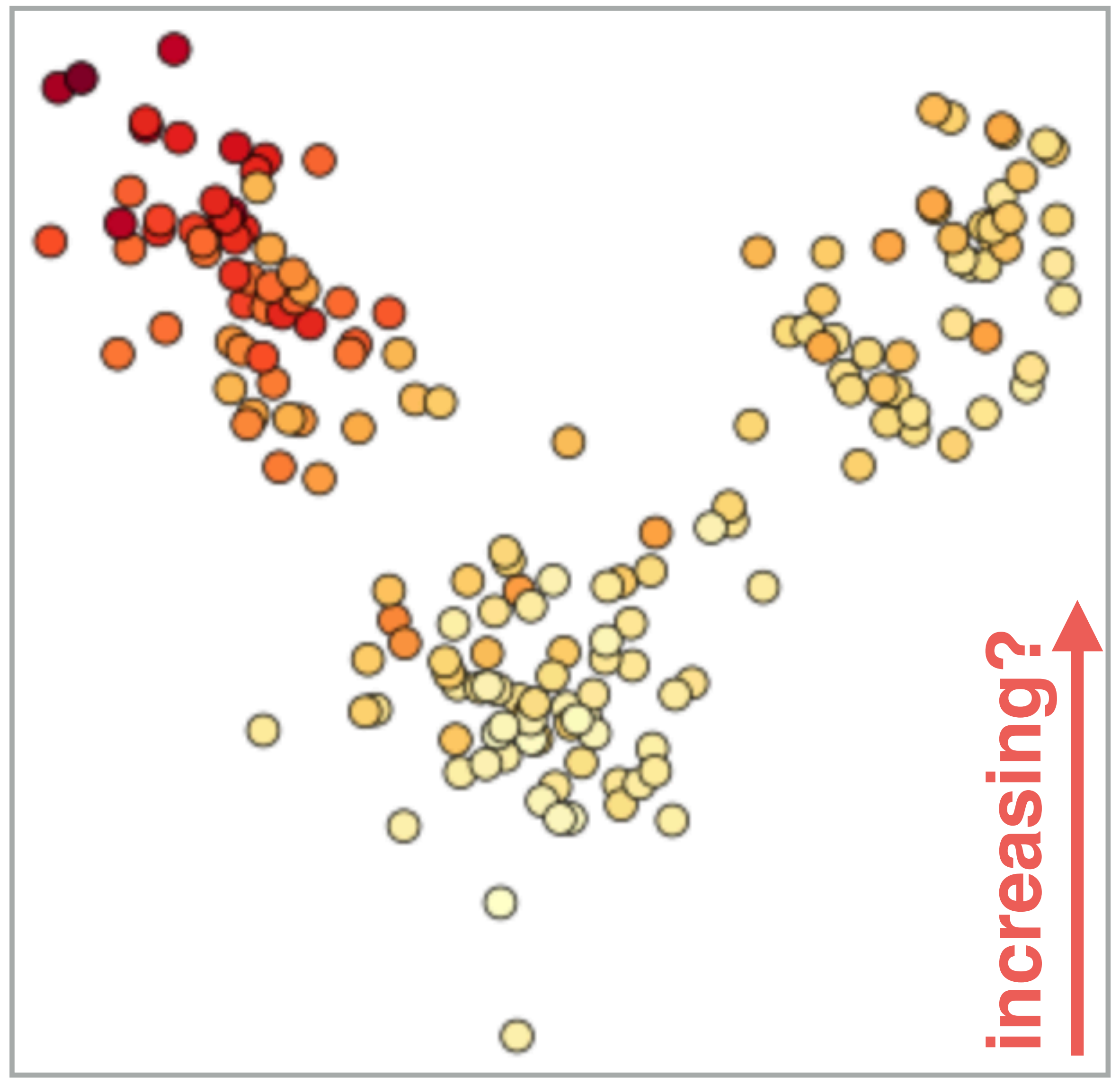
Colored by **flavanoids**



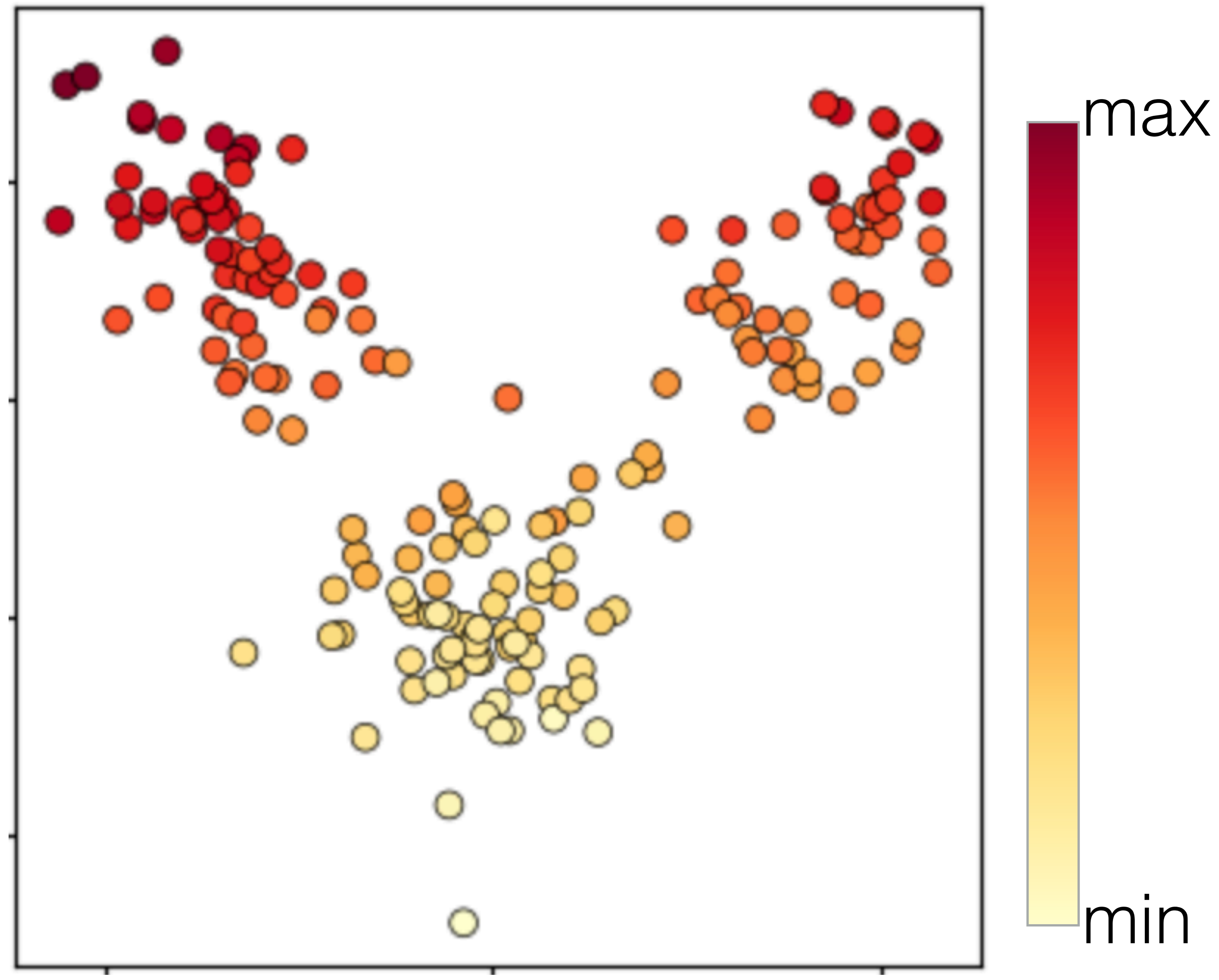
Colored by **proline**

Interpreting the axes





Colored by **proline**
(univariate)



Colored by
0.8 alchol + 0.7 ash - 0.5 alcalinity_ash
- 0.5 flavanoids + 0.7 color_intensity + proline
(composite variable)

To better understand the axes, we should consider relationships of multiple attributes

Do you think linear DR is still explainable if there are many attributes (e.g., 100)?

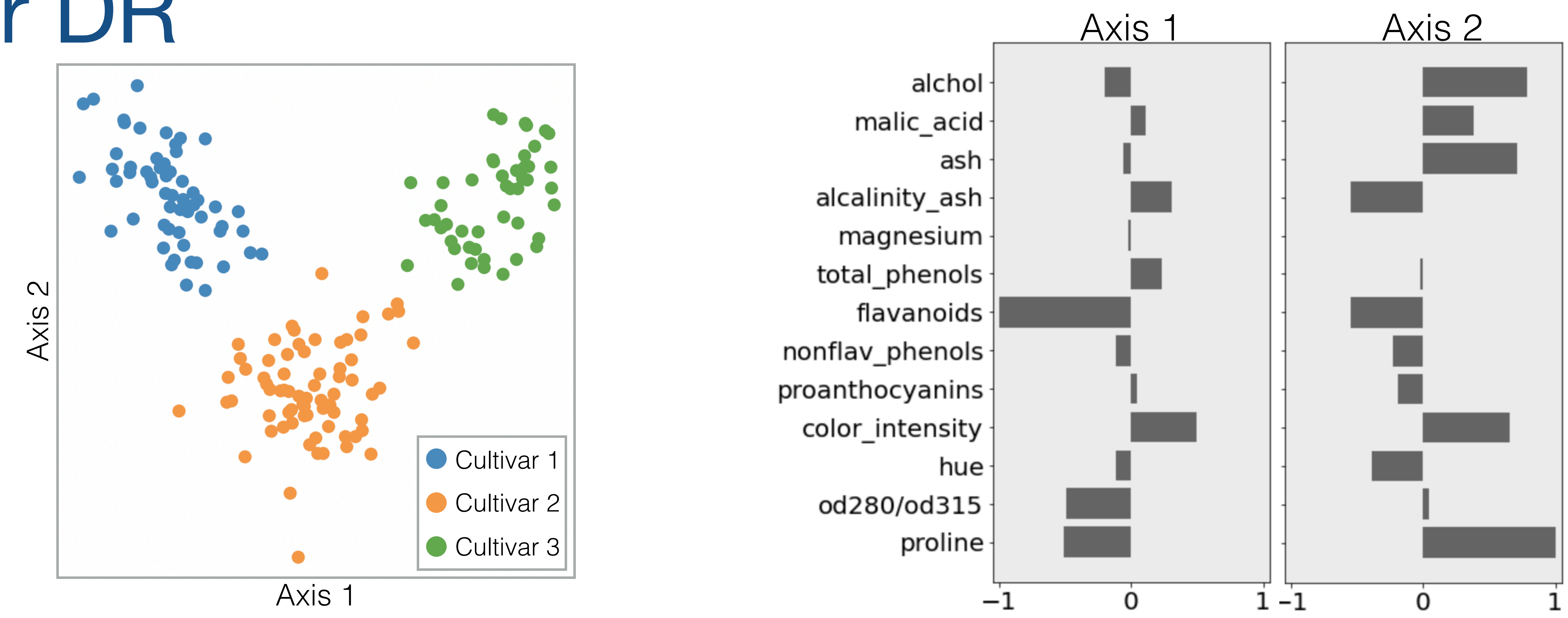
Sparse linear DR

- Finds a linear projection, $\mathbf{P}$, that has a **small number of nonzero** elements
  - $\mathbf{X} \cdot \mathbf{P} = \mathbf{Y}$ $\mathbf{X} \in \mathbb{R}^{n \times d}$: original data, $\mathbf{P} \in \mathbb{R}^{d \times d'}$: projection matrix, $\mathbf{Y} \in \mathbb{R}^{n \times d'}$: projected data,
    (n: # of instances, d: # of original dimensions, d': # of dimensions after projection)

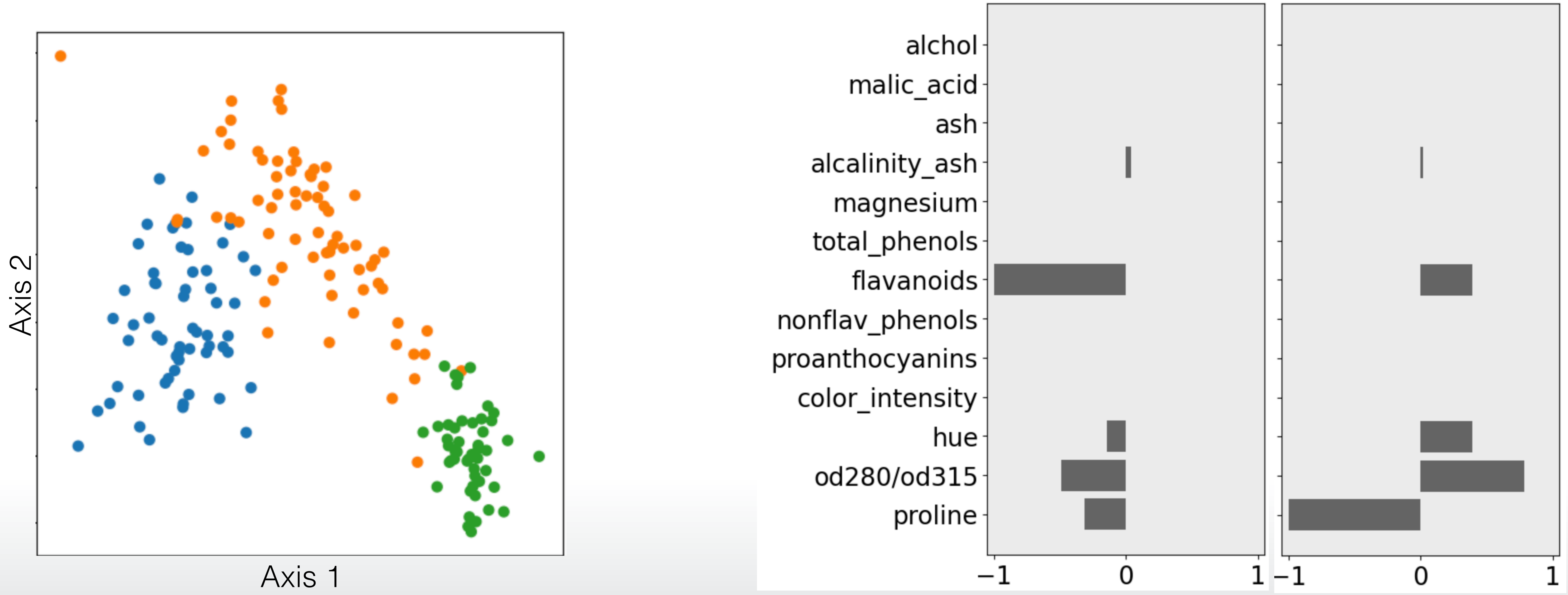
- Representative methods
  - Sparse PCA
    (e.g., Zou et al., “Sparse principal component analysis.” *J. Comput. Graph. Stat.*, 2006.)
  - Sparse LDA
    (e.g., Wen et al., “Robust sparse linear discriminant analysis.” *IEEE Trans. Circuits Syst. Video Technol.*, 2018.)

Sparse linear DR

LDA



Sparse LDA



Interpretation of dimensionality reduction results

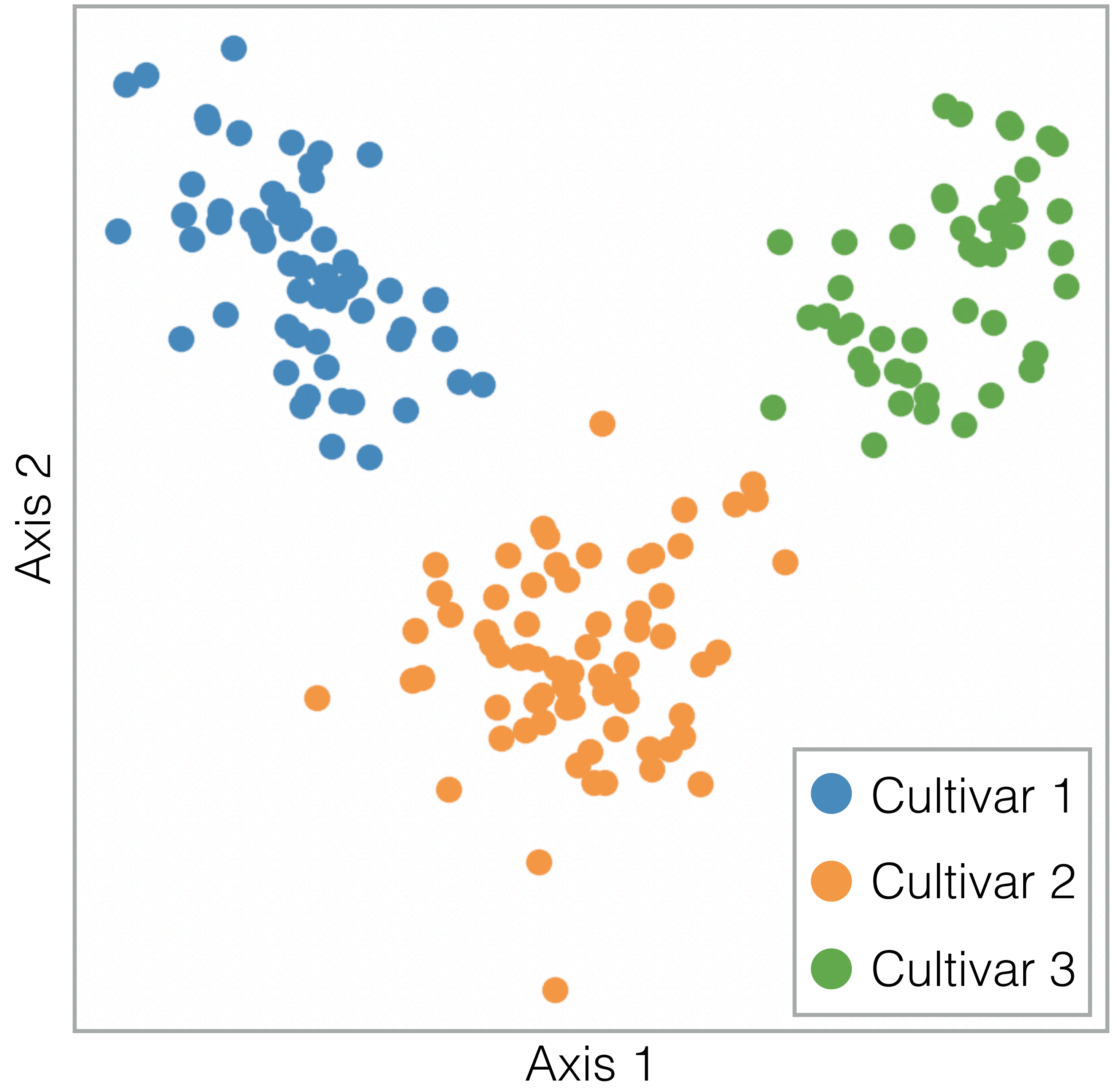
Wine dataset (13D)

	alchol	malic_acid	ash	alcalinity_ash	magnesium	total_phenols	flavanoid
0	1.52	-0.56	0.23	-1.17	1.91	0.81	1
1	0.25	-0.50	-0.83	-2.49	0.02	0.57	0.
2	0.20	0.02	1.11	-0.27	0.09	0.81	1.22
3	1.69	-0.35	0.49	-0.81	0.93	2.49	1.47
4	0.30	0.23	1.84	0.45	1.28	0.81	0.66
...	...	...	...	...	...	...	...

*displayed values are after the standardization

Linear discriminant analysis

LDA



- Interpreting a **lower-dimensional space** — **Linear DR**
- Interpreting based on **observed patterns** (e.g., clusters) — **Nonlinear DR**
- Interpreting from a DR **model/mechanism** level — **Nonlinear DR**

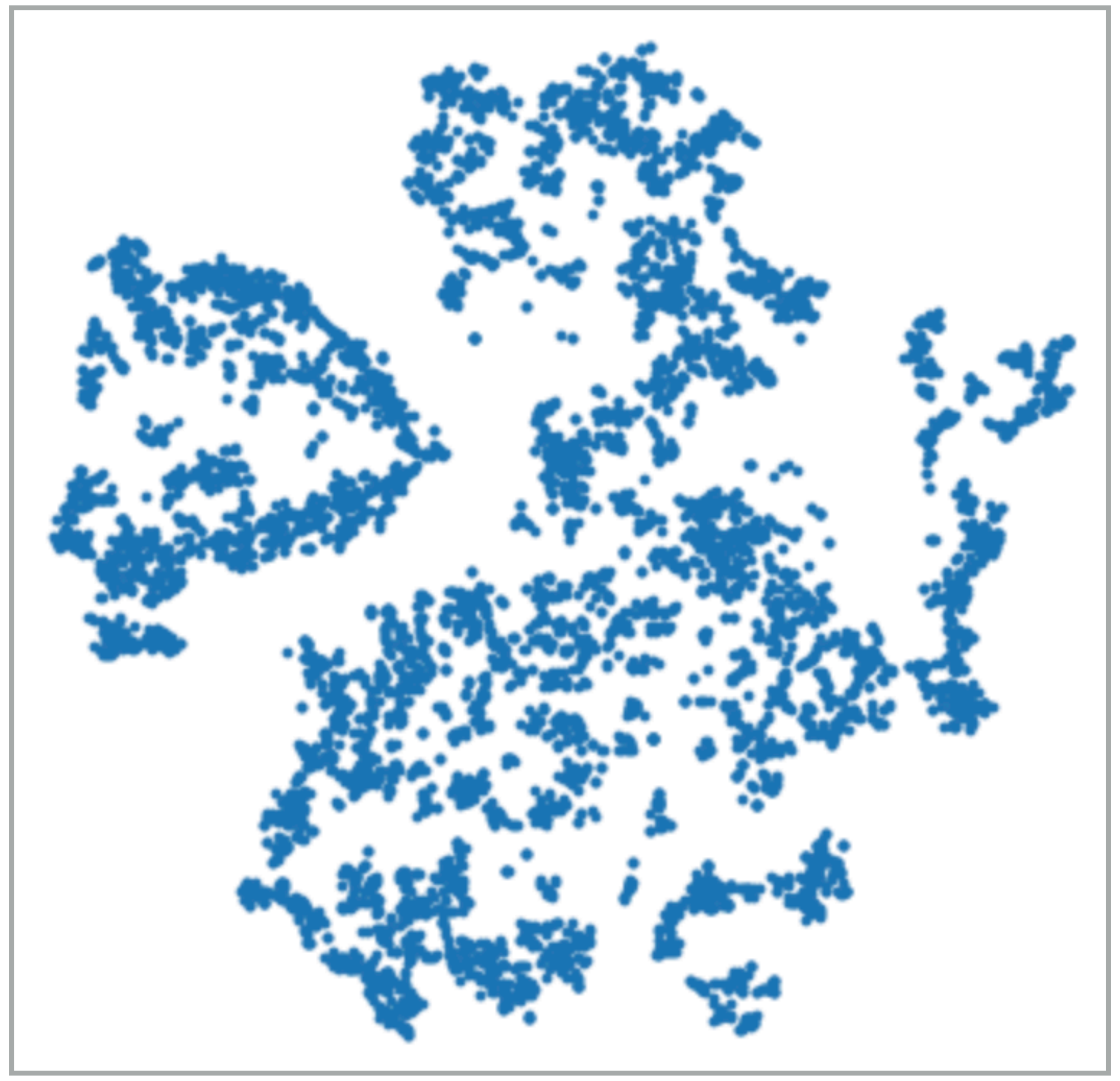
Nonlinear DR

- Does not produce a linear projection from the original to the low-dimensional space
- Many commonly used nonlinear methods such as UMAP, t-SNE, Metric MDS do not provide a **parametric mapping** from the original to the low-dimensional space
  - Parametric mapping $f_\theta : \mathbf{X} \rightarrow \mathbf{Y}$ where θ are parameters
  - Without the parametric mapping, we do not have the information how the original data is projected onto the low-dimensional space.

Instead of understanding the low-dimensional axes, we can try to perform the interpretation based on visual patterns appeared in a DR result

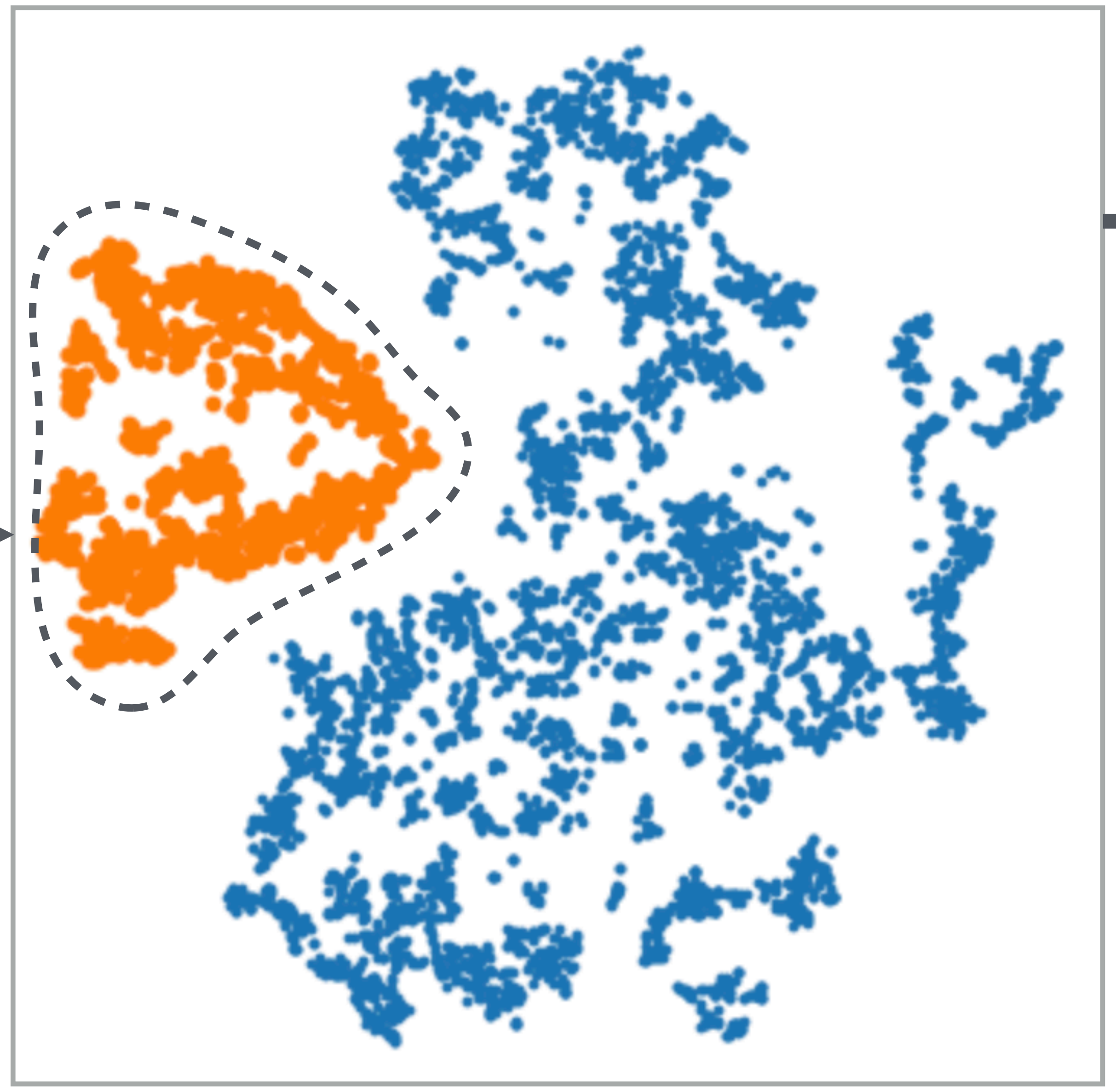
Interpreting by comparing groups/clusters

Food nutrition dataset (12D)



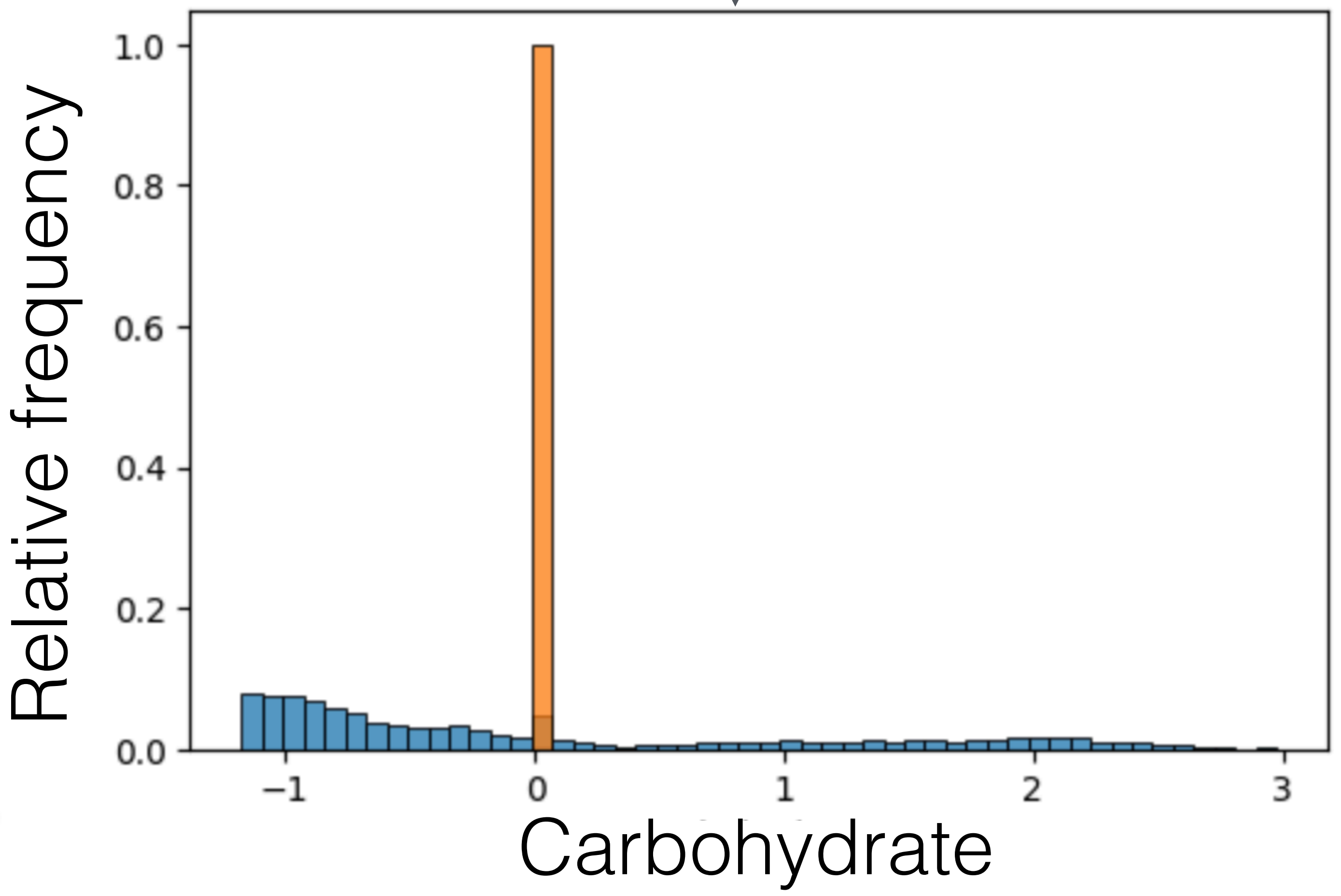
t-SNE result

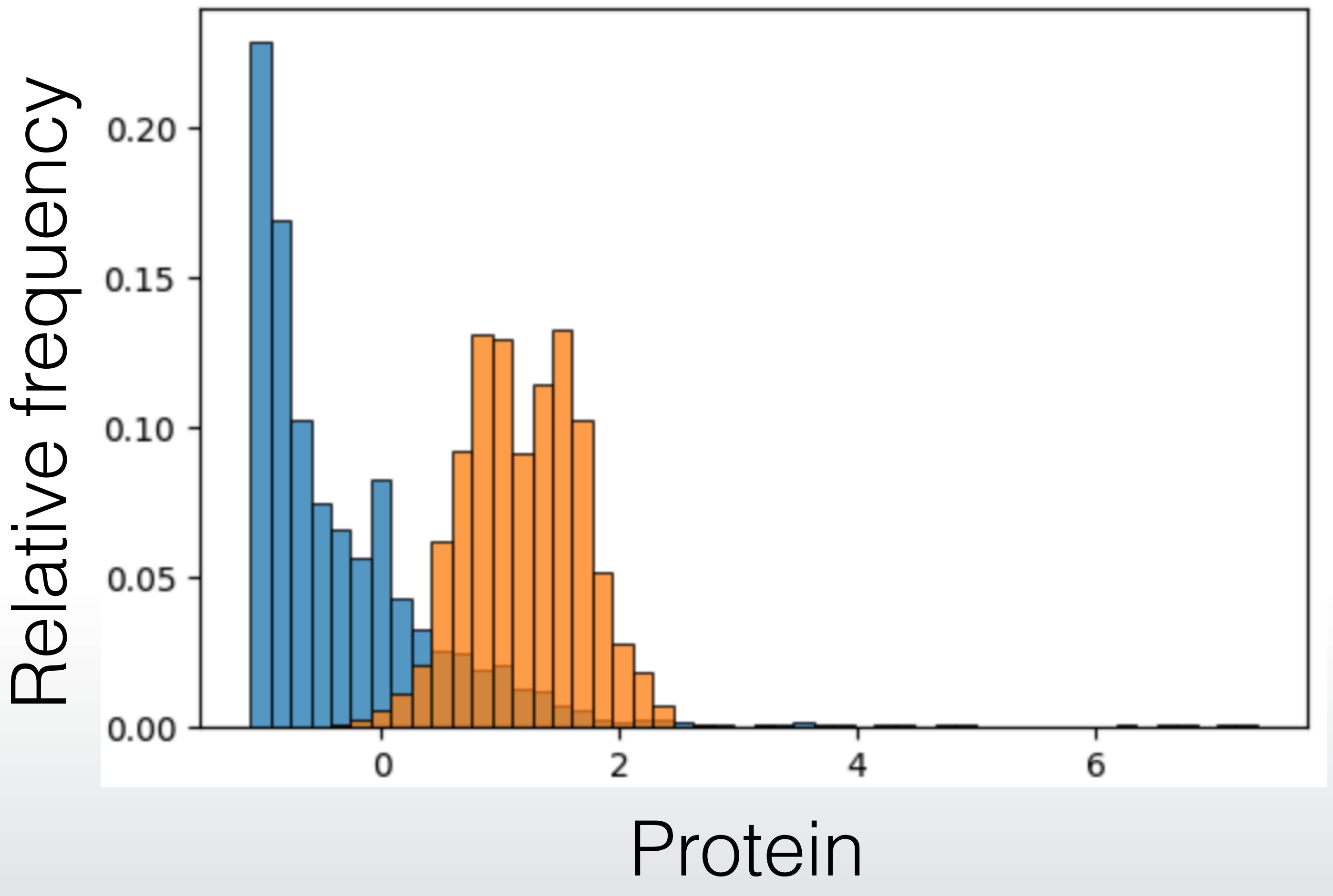
Lasso selection



One cluster is selected

Select attributes with statistical measures (e.g., t-test, histogram intersection, LDA); then, visualize them with histograms, etc.





Selected foods have a medium level of carbohydrate but have a lot of protein

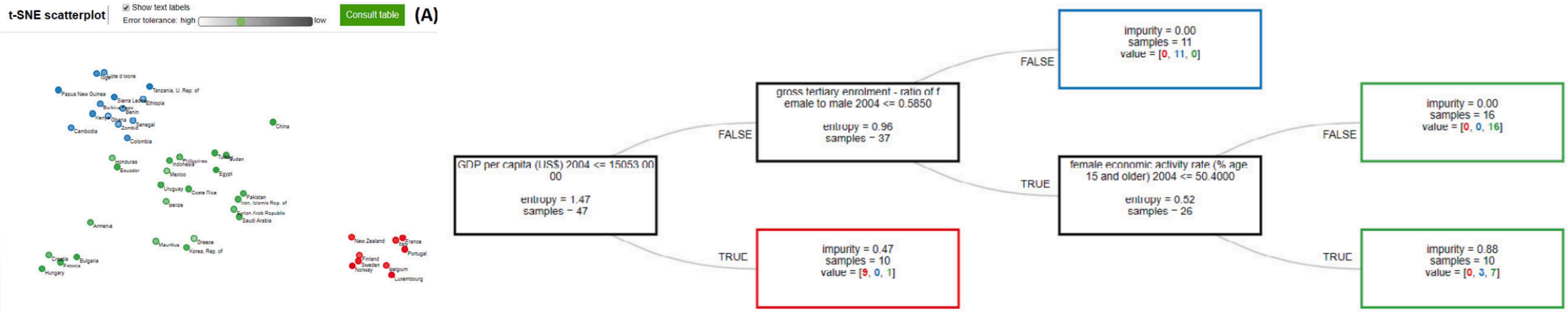
Existing approaches for comparing groups/clusters

- Univariate focus
  - e.g., t-student test-based attribute selection

    Marcílio-Jr et al., “Contrastive analysis for scatterplot-based representations of dimensionality reduction.” *C&G*, 2021.
- Composite variable focus
  - e.g., comparative analysis using linear DR

    Fujiwara et al., “Supporting analysis of dimensionality reduction results with contrastive learning.” *IEEE TVCG*, 2020.
- Classifier-based
  - e.g., building a simple model that classifies clusters in a DR result

    Bibal et al., “IXVC: An interactive pipeline for explaining visual clusters in dimensionality reduction visualizations with decision trees.” *Array*, 2021.

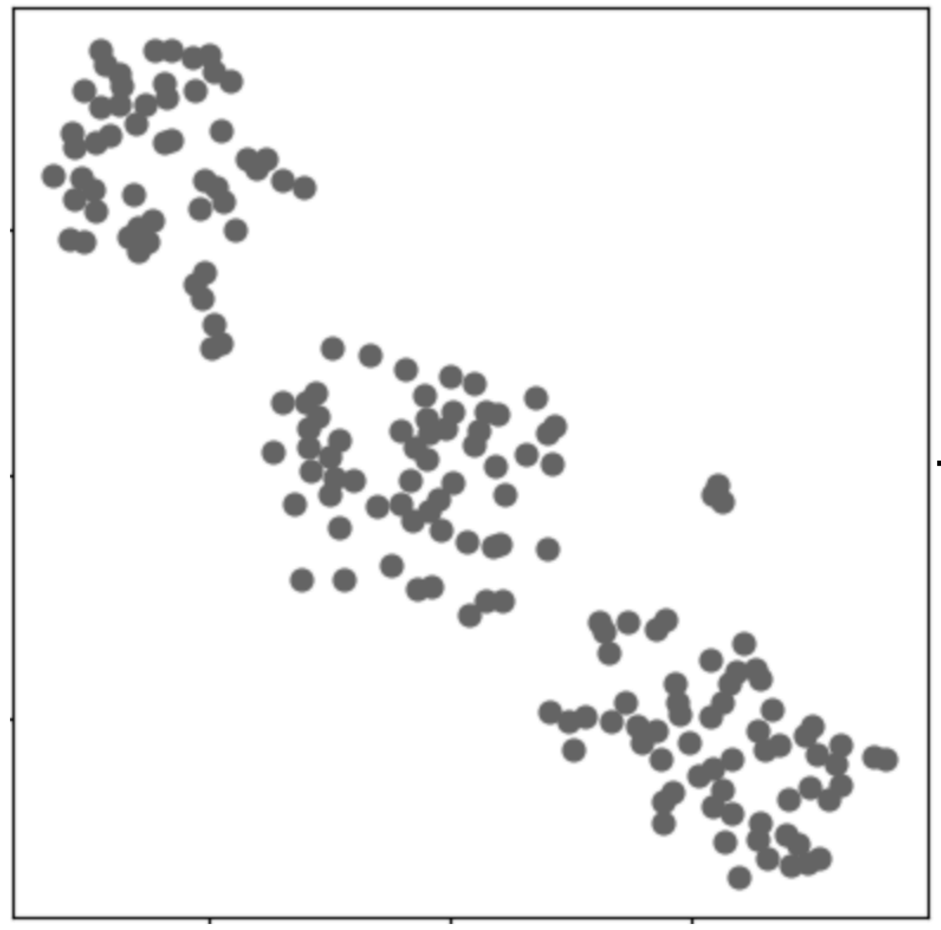


Existing approaches for comparing groups/clusters

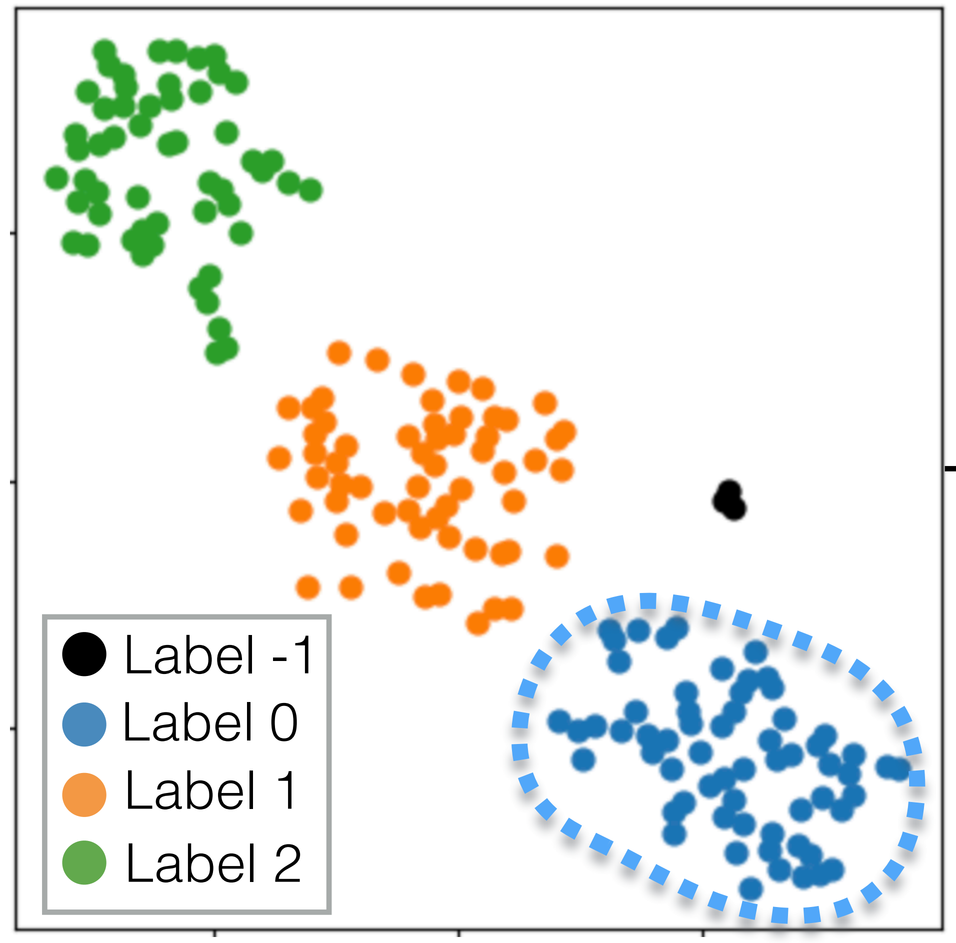
- Univariate focus
  - e.g., t-student test-based attribute selection
    Marcílio-Jr et al., "Contrastive analysis for scatterplot-based representations of dimensionality reduction." *C&G*, 2021.
- Composite variable focus
  - e.g., comparative analysis using linear DR
    Fujiwara et al., "Supporting analysis of dimensionality reduction results with contrastive learning." *IEEE TVCG*, 2020.
- Classifier-based
  - e.g., building a simple model that classifies clusters in a DR result
    Bibal et al., "IXVC: An interactive pipeline for explaining visual clusters in dimensionality reduction visualizations with decision trees." *Array*, 2021.

Comparing clusters/groups using linear DR: PCA-based

- Applying **PCA** for each cluster
  (e.g., `PCA(n_components=1).fit(data[label==0])`)
  - Extract variance-related information of each cluster—**variety factors**

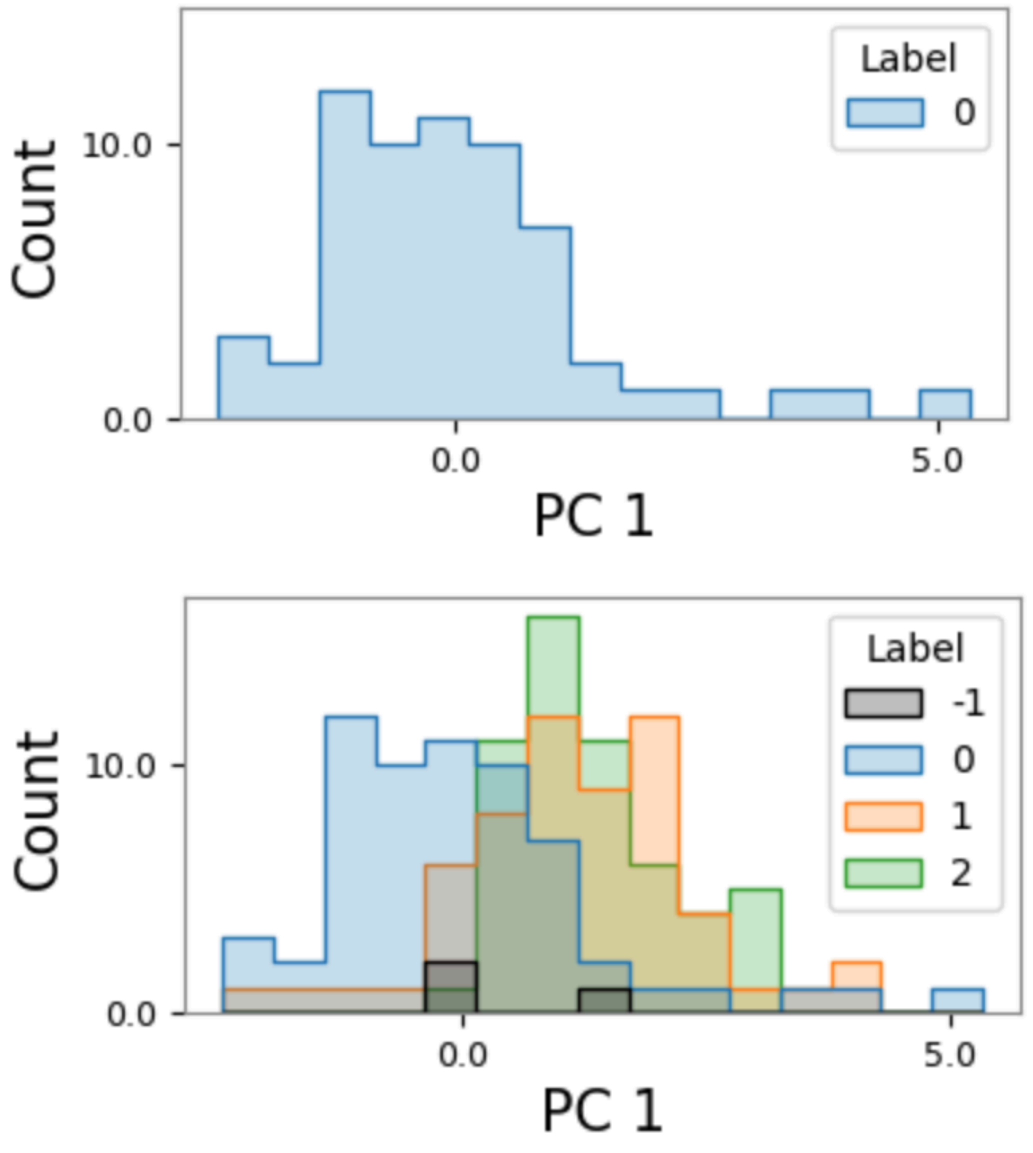


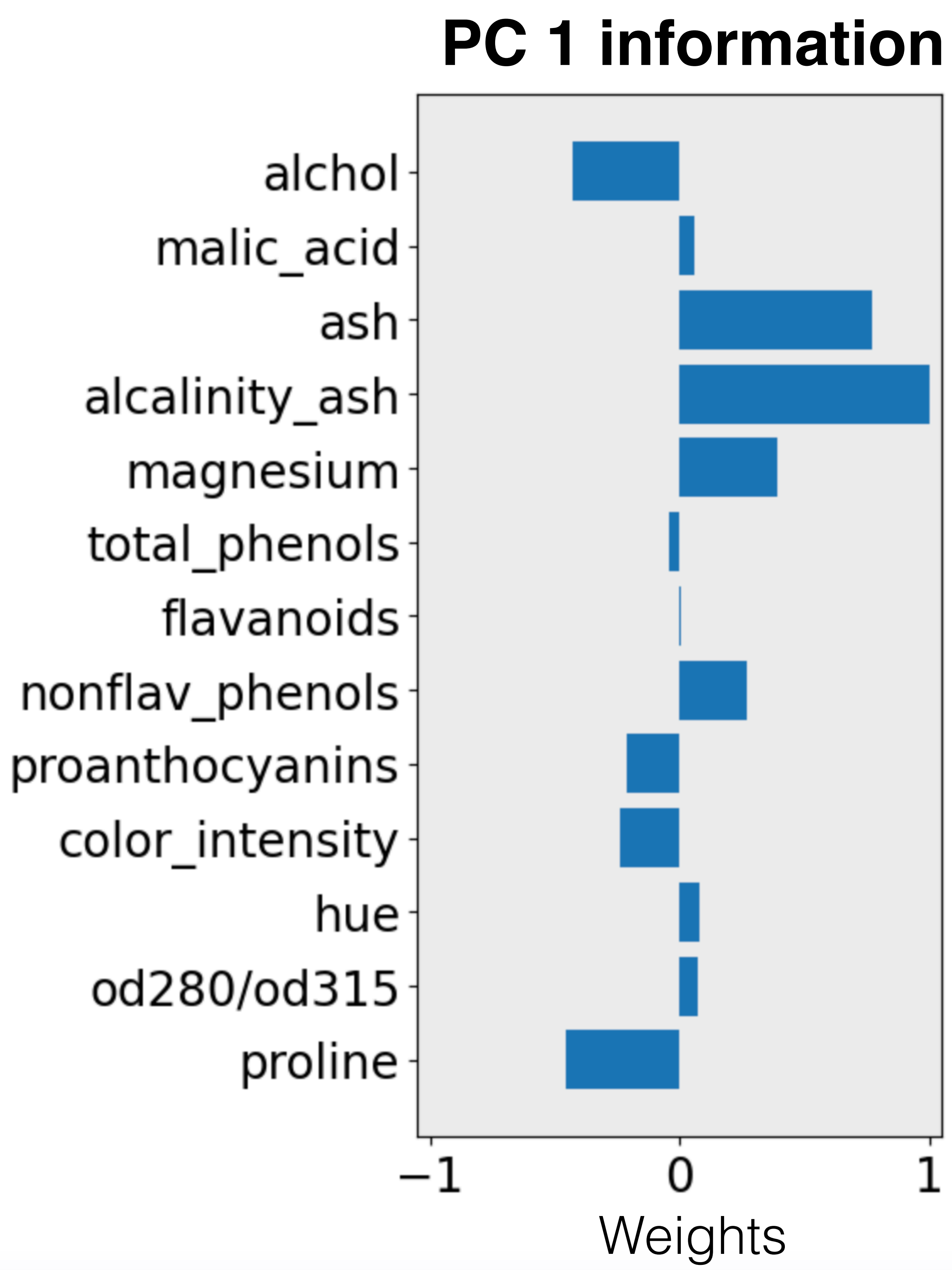
t-SNE result
(Wine dataset)



After clustering
(e.g., with DBSCAN)

PCA to
extract PC 1

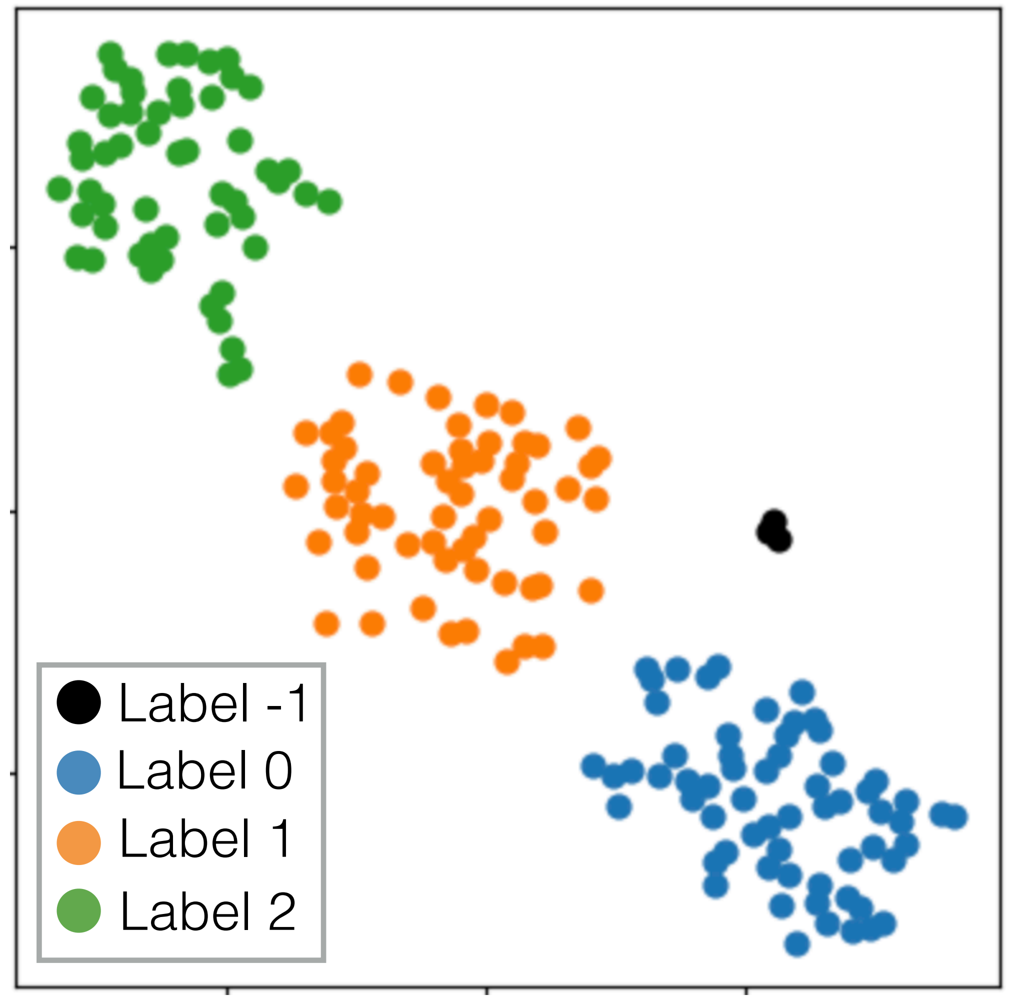




demo script

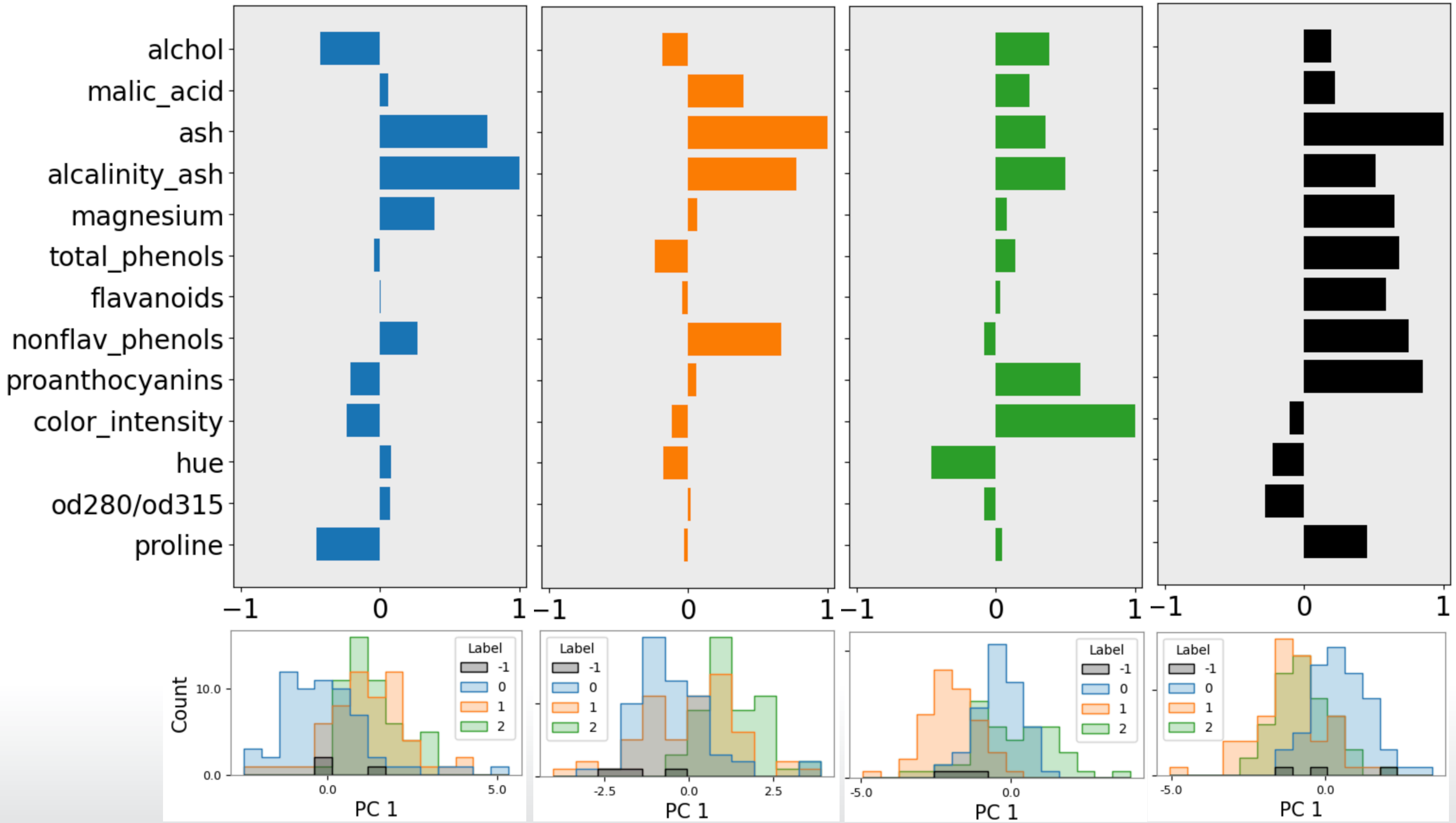
Comparing clusters/groups using linear DR: PCA-based

- Applying **PCA** for each cluster
  (e.g., `PCA(n_components=1).fit(data[label==0])`)
  - Extract variance-related information of each cluster—**variety factors**
  - But, there is no consideration of differences between one cluster and others



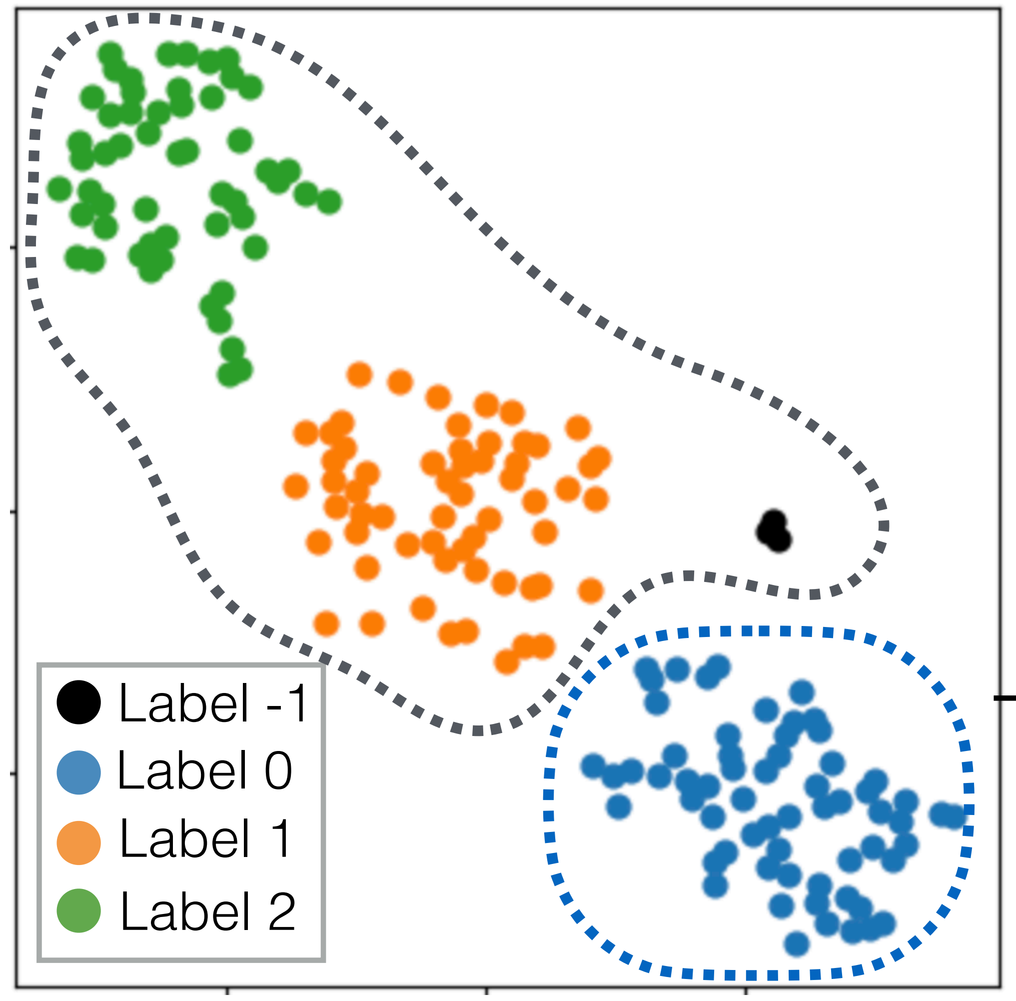
t-SNE result after clustering

PCA to extract PC 1



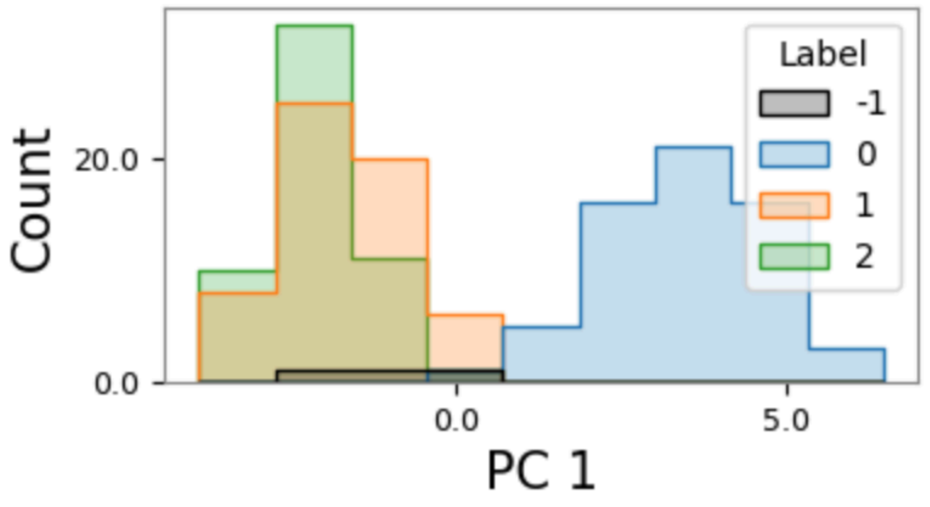
Comparing clusters/groups using linear DR: LDA-based

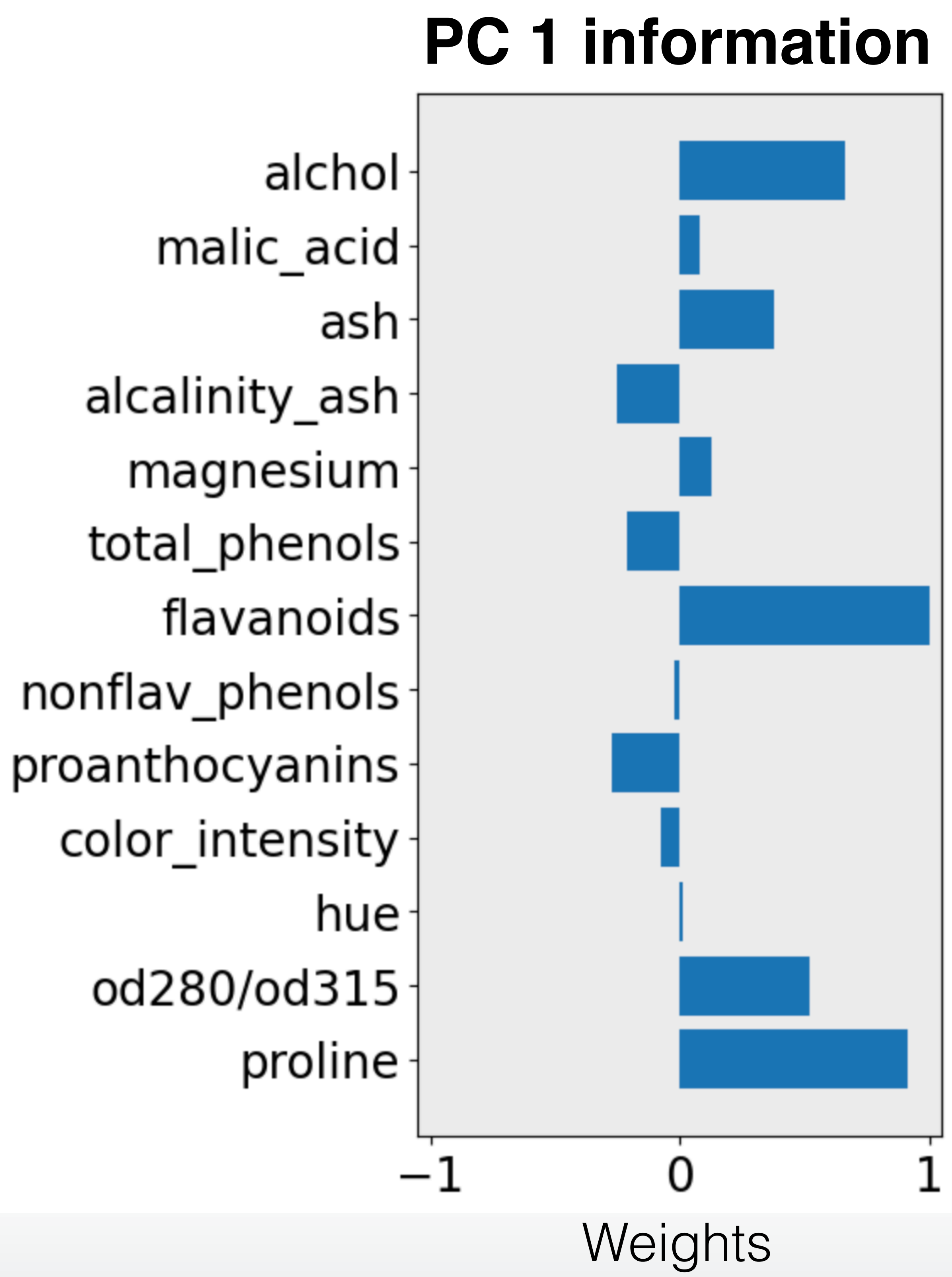
- Applying linear discriminant analysis (**LDA**) to distinguish one cluster and others (e.g., `LDA(n_components=1).fit(data, label==0)`)
  - Find differences between one cluster and others—**differentiating factors**



t-SNE result after clustering

Classification
Label 0 and **others**
with LDA and then
extraction of PC 1

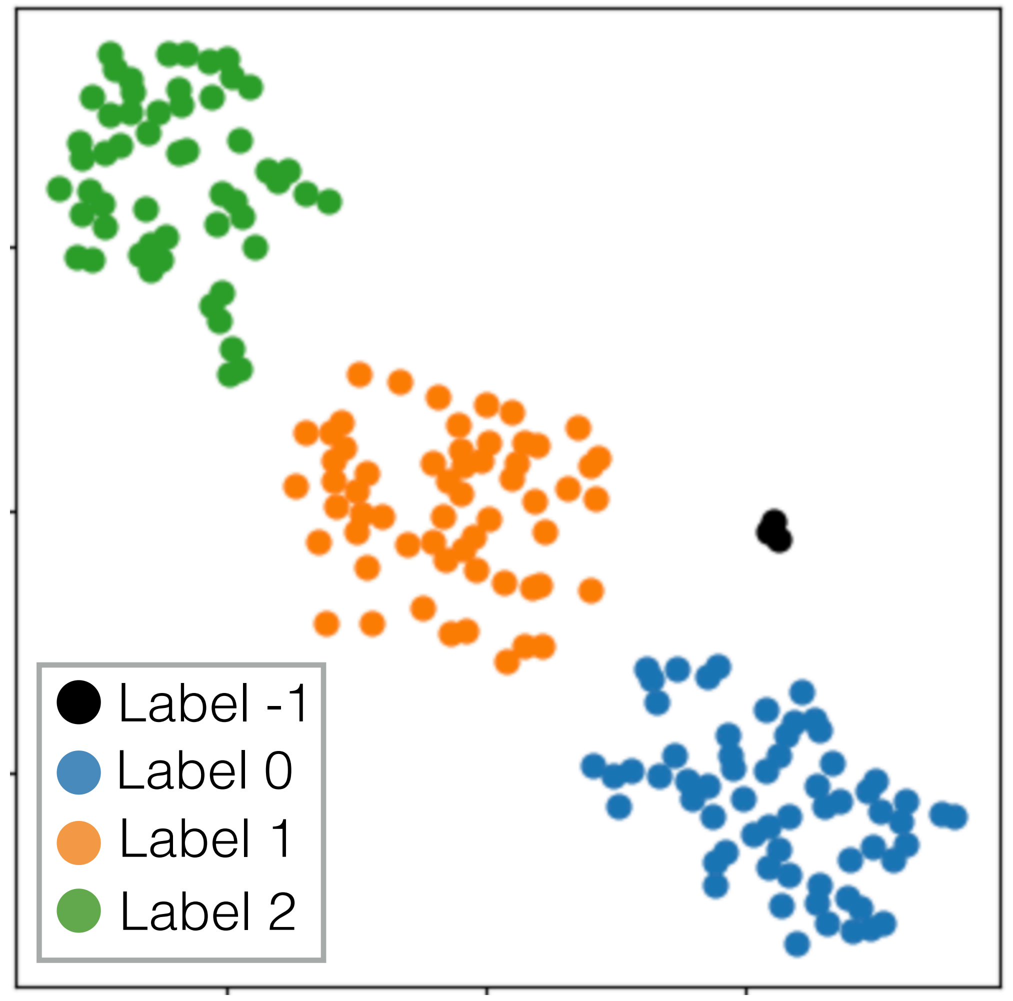




demo script

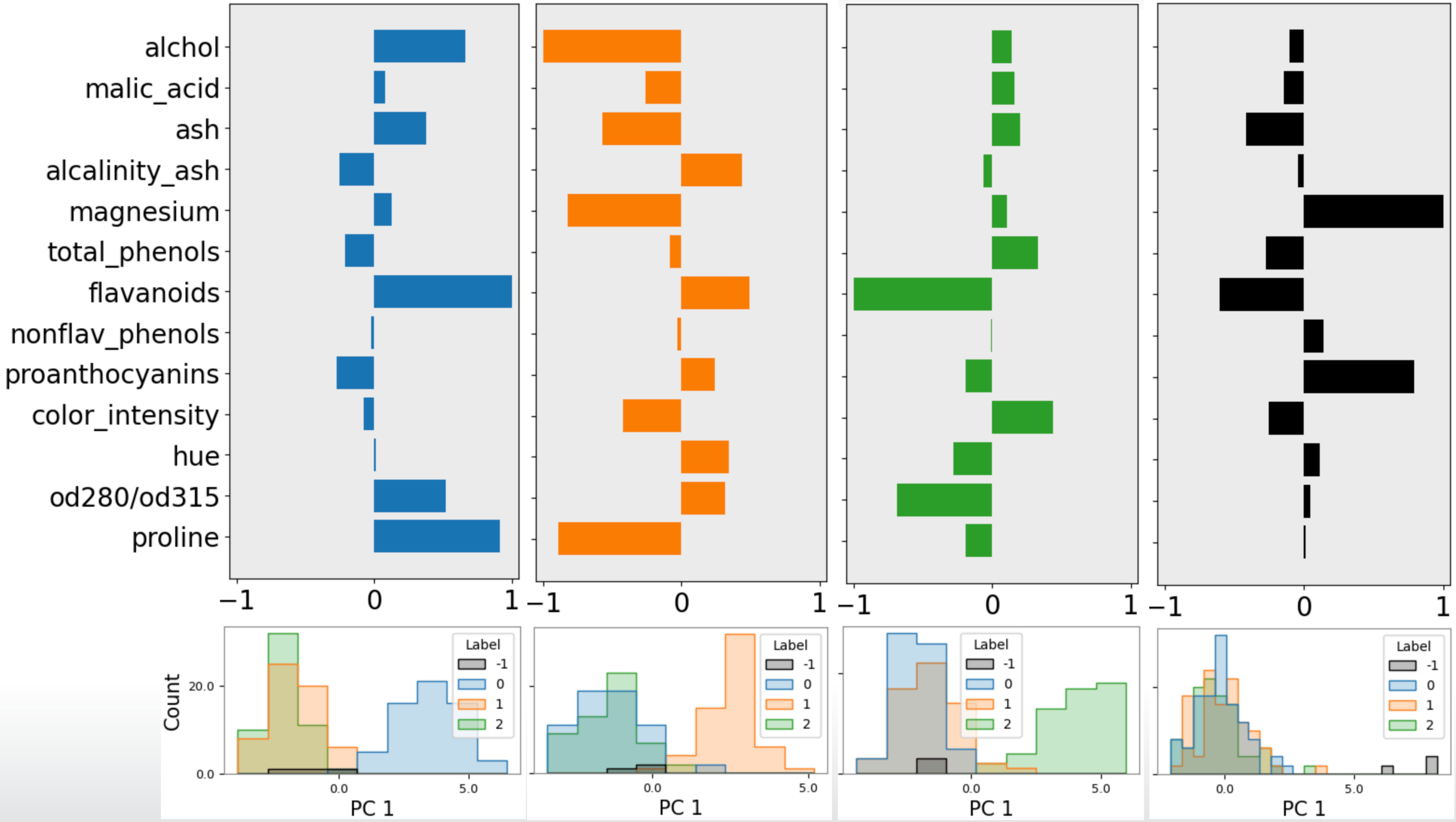
Comparing clusters/groups using linear DR: PCA-based

- Applying linear discriminant analysis (**LDA**) to distinguish one cluster and others (e.g., `LDA(n_components=1).fit(data, label==0))`
  - Find differences between one cluster and others—**differentiating factors**
  - But, may lose overall/variance information related each cluster



t-SNE result after clustering

LDA to extract PC 1



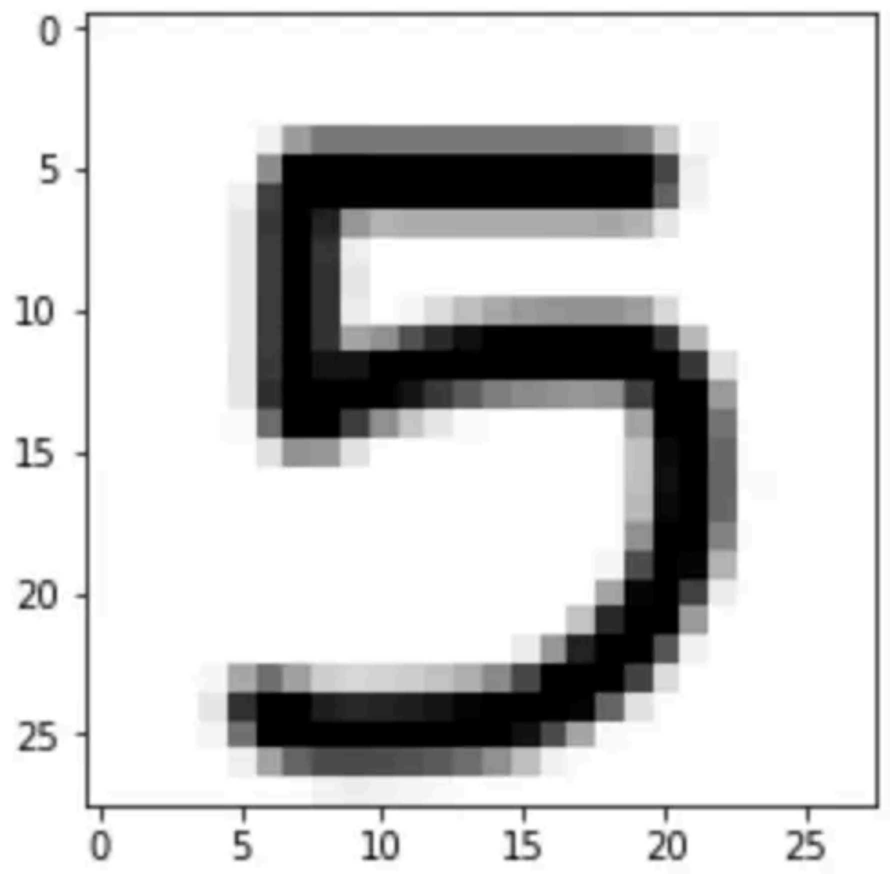
ccPCA: Contrasting clusters in PCA

Fujiwara et al., “Supporting analysis of dimensionality reduction results with contrastive learning.” *IEEE TVCG*, 2020.

- Extract salient factors in one group relative to others while considering their distinction—**characterizing factors**

MNIST Dataset

70,000 hand-written digits



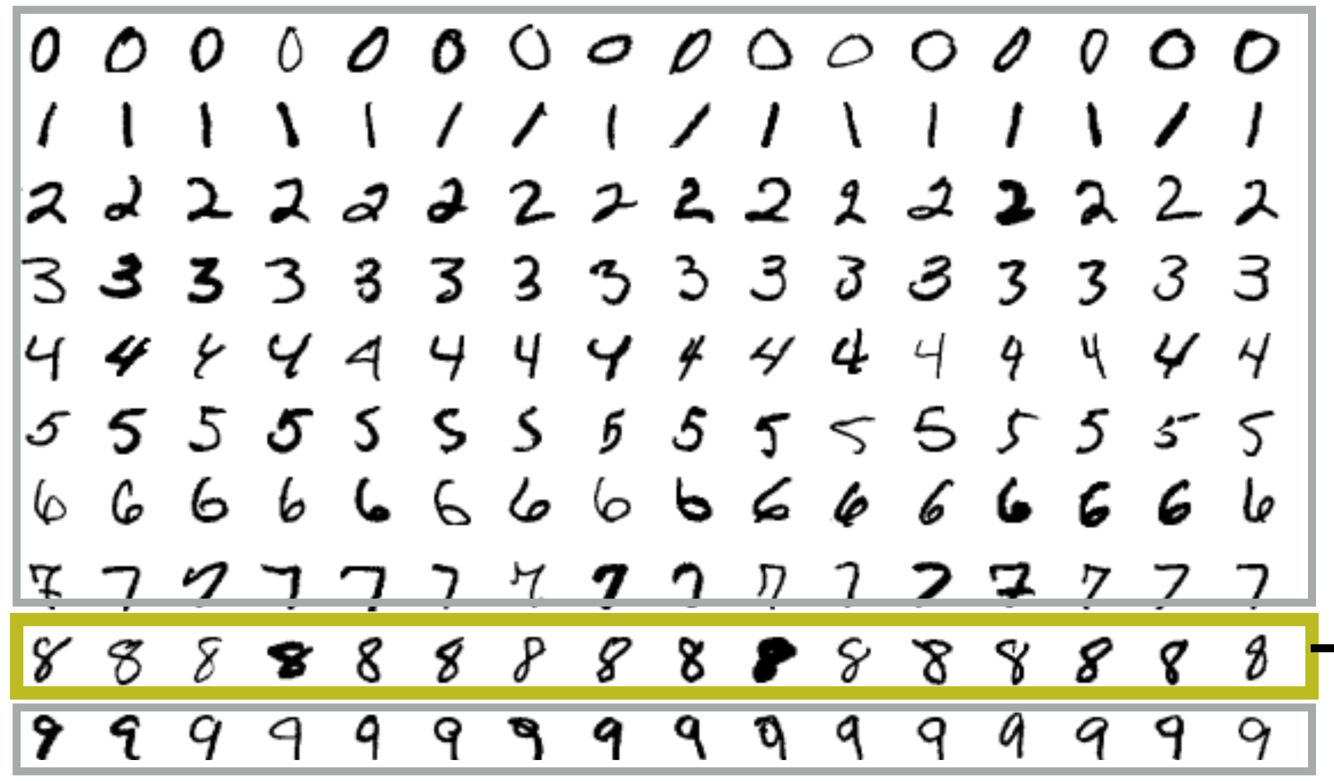
28 pixels x 28 pixels = 784 dimensions

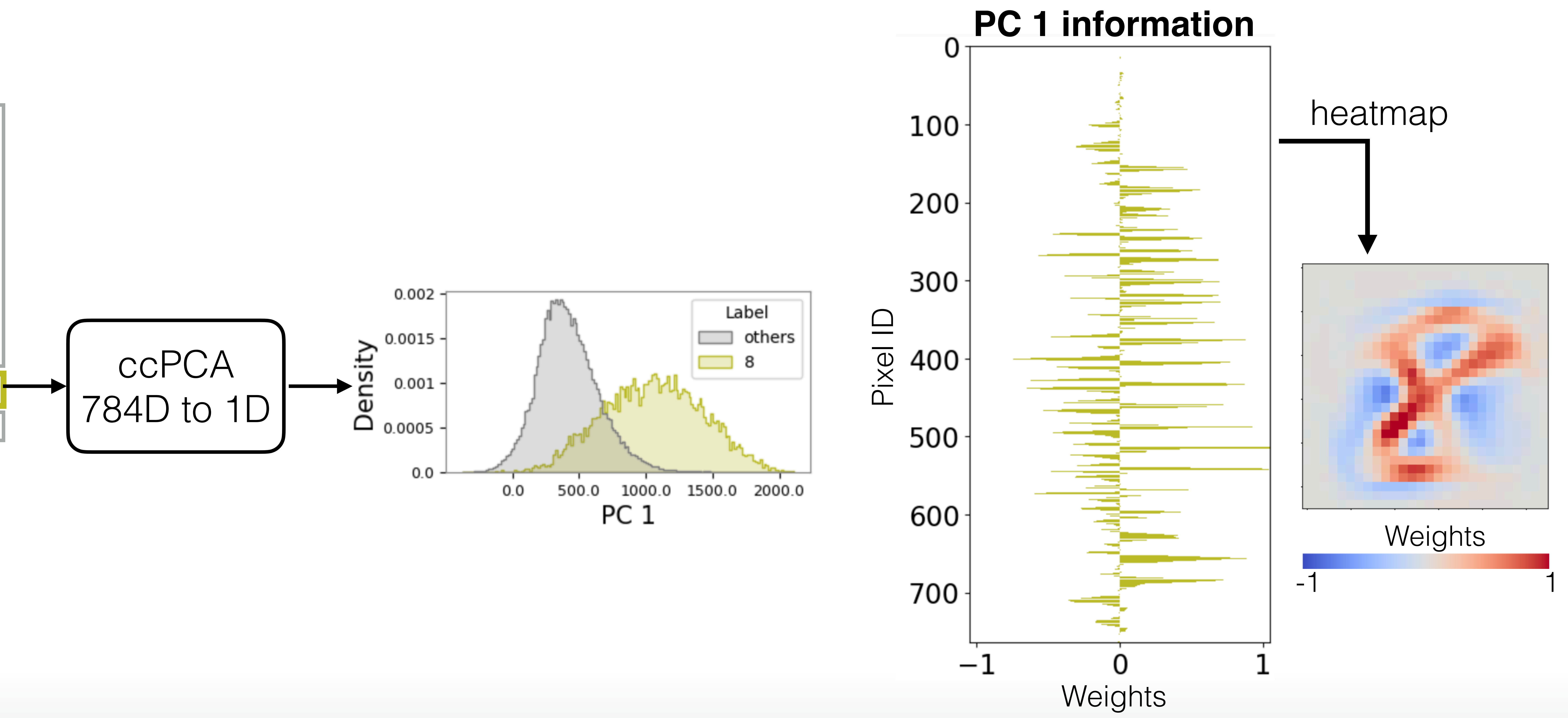
ccPCA: Contrasting clusters in PCA

Fujiwara et al., “Supporting analysis of dimensionality reduction results with contrastive learning.” *IEEE TVCG*, 2020.

- Extract salient factors in one group relative to others while considering their distinction—**characterizing factors**

MNIST Dataset





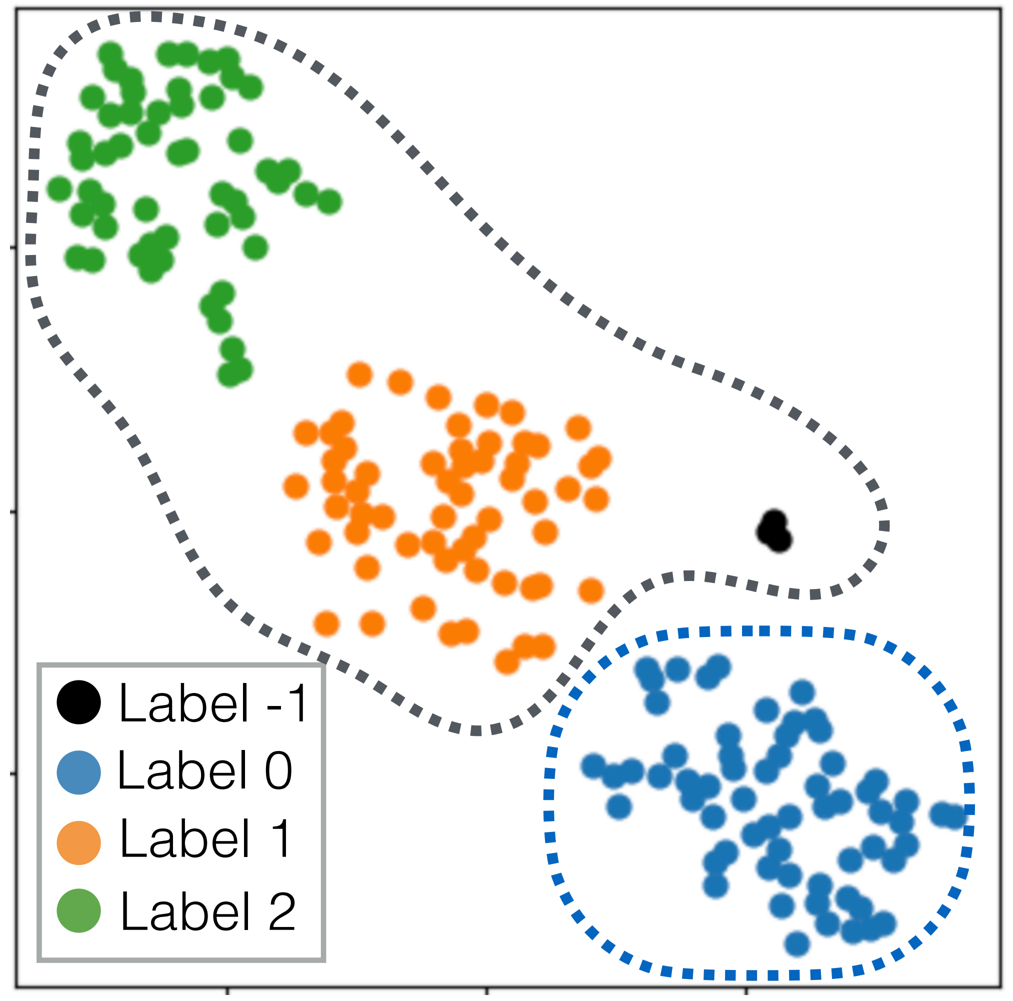
ccPCA: Contrasting clusters in PCA

Fujiwara et al., “Supporting analysis of dimensionality reduction results with contrastive learning.” *IEEE TVCG*, 2020.

- Extract salient factors in one group relative to others while considering their distinction — **characterizing factors**

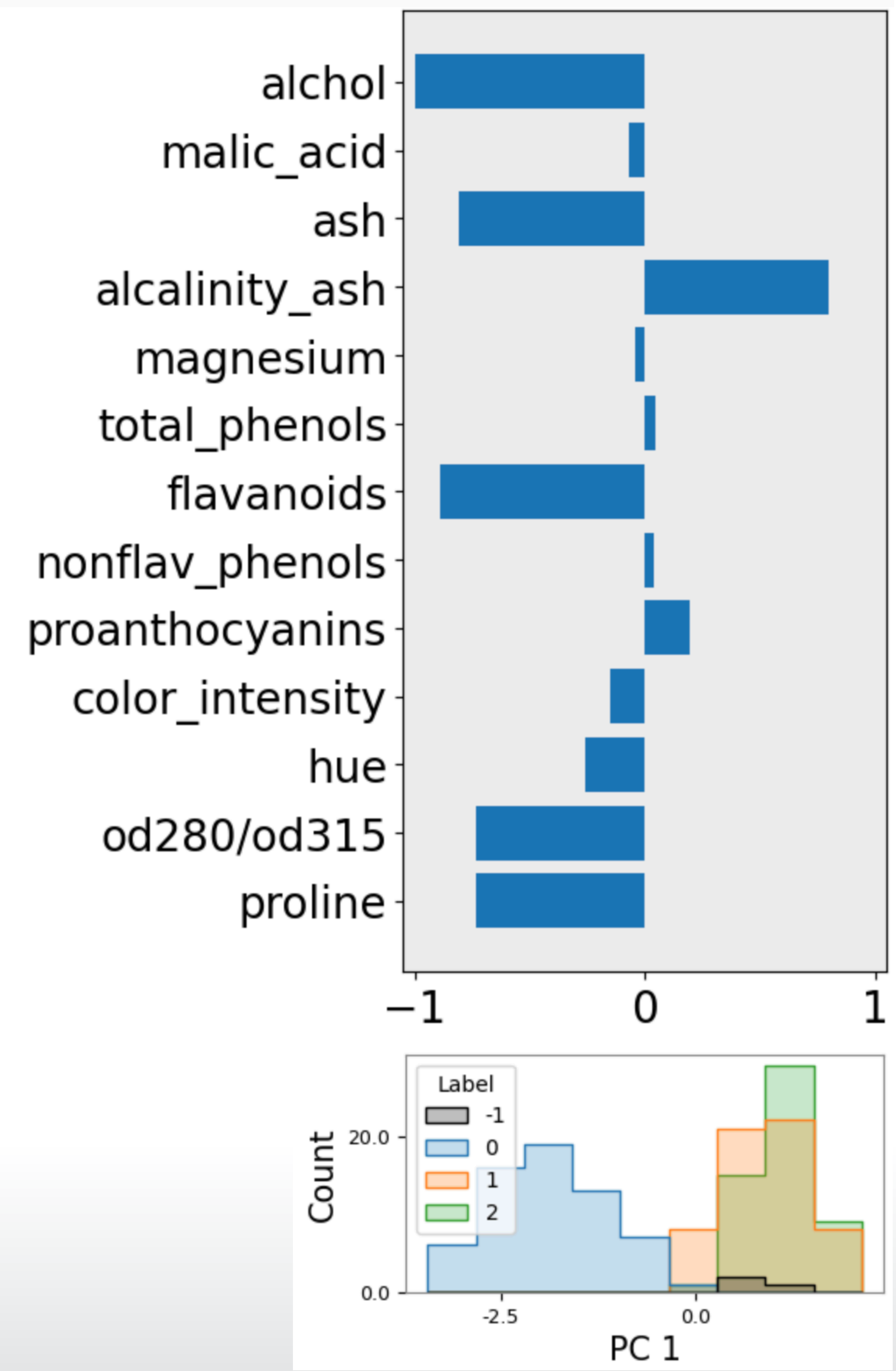
Comparing clusters/groups using linear DR: ccPCA-based

- Applying **ccPCA** to contrast one cluster with others
  (e.g., `CCPCA(n_components=1).fit(data[label==0], data[label!=0])`)
  - Balance the preservation of variance within a cluster and the distinction from others
    - better match with t-SNE and UMAP's algorithms



t-SNE result after clustering

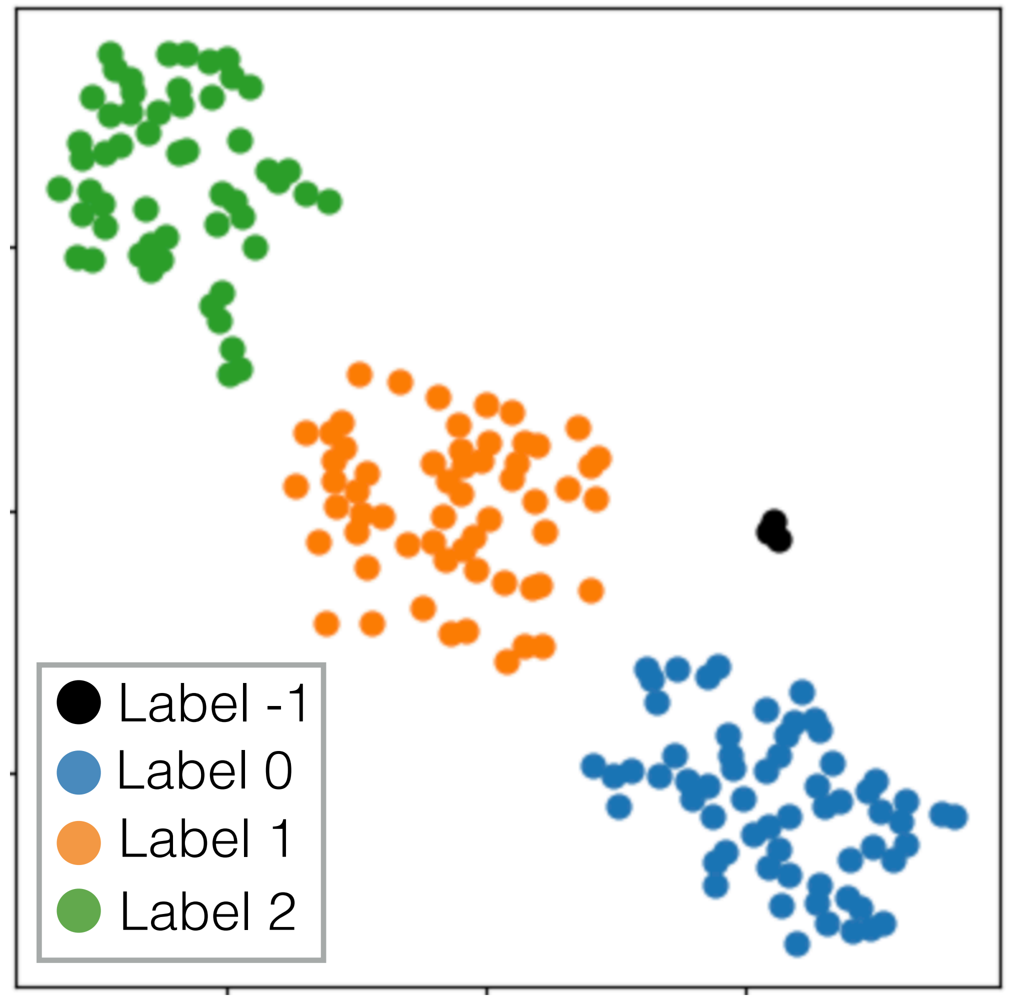
ccPCA to extract PC



demo script

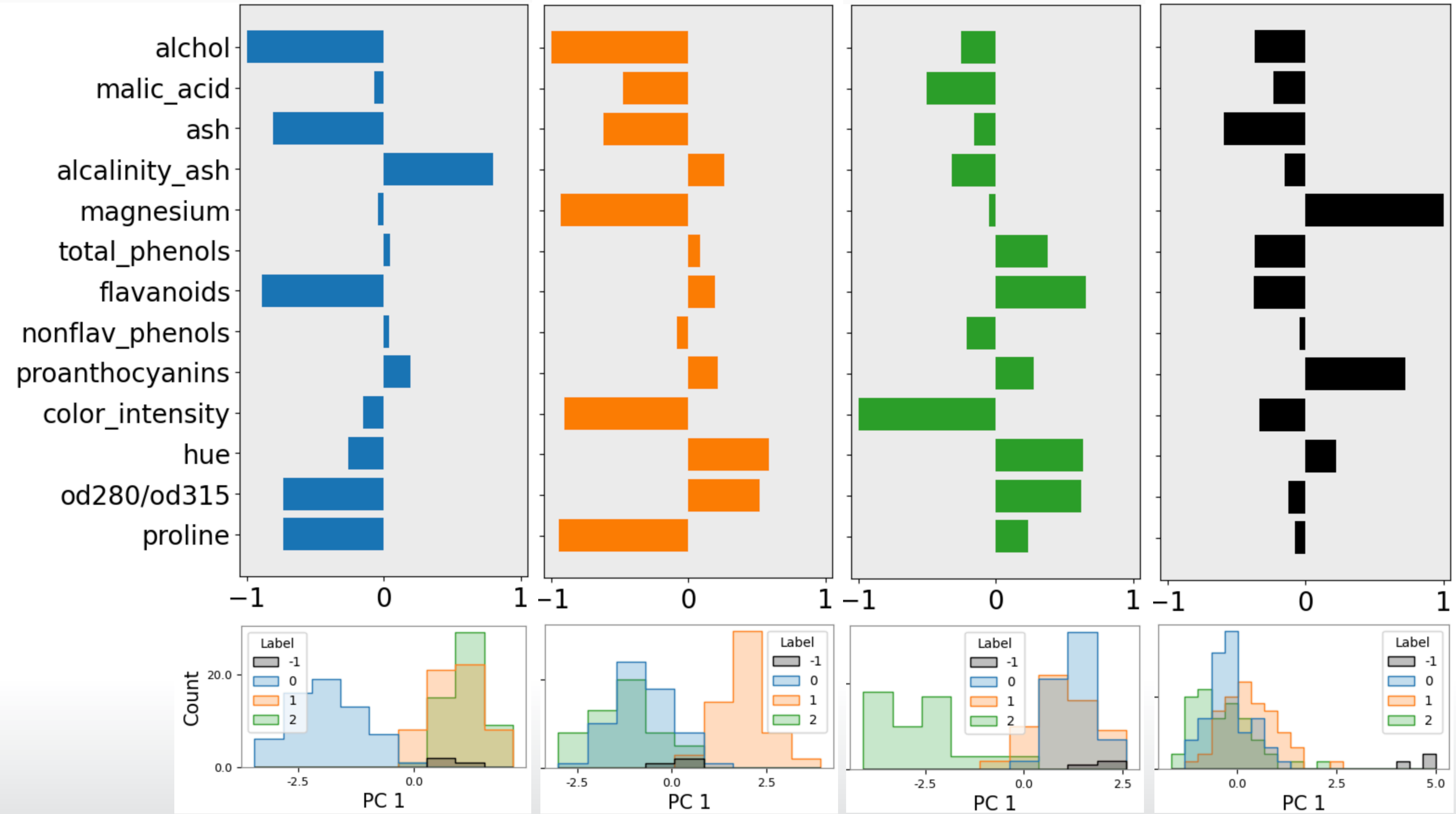
Comparing clusters/groups using linear DR: ccPCA-based

- Applying **ccPCA** to contrast one cluster with others
  (e.g., `CCPCA(n_components=1).fit(data[label==0], data[label!=0])`)
  - Balance the preservation of variance within a cluster and the distinction from others
    - better match with t-SNE and UMAP's algorithms



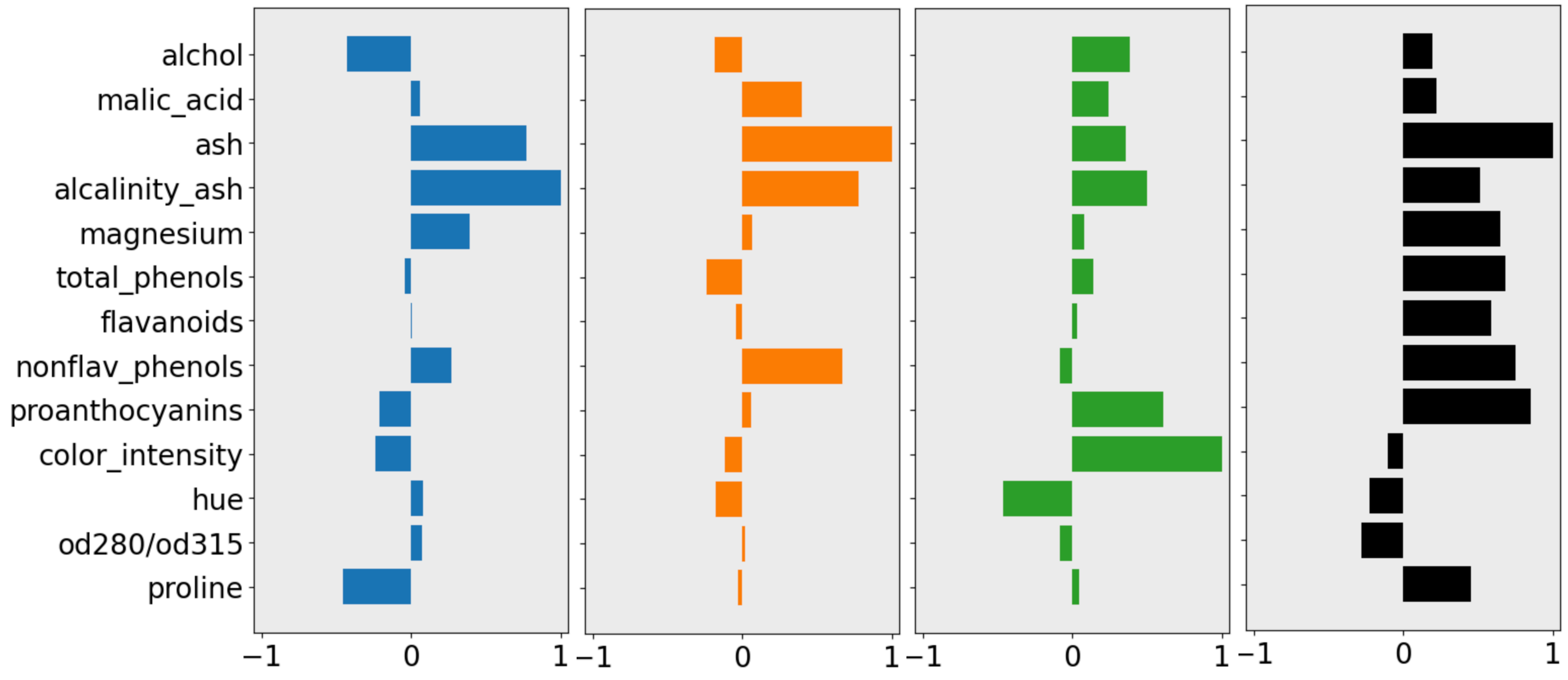
t-SNE result after clustering

ccPCA to extract PC

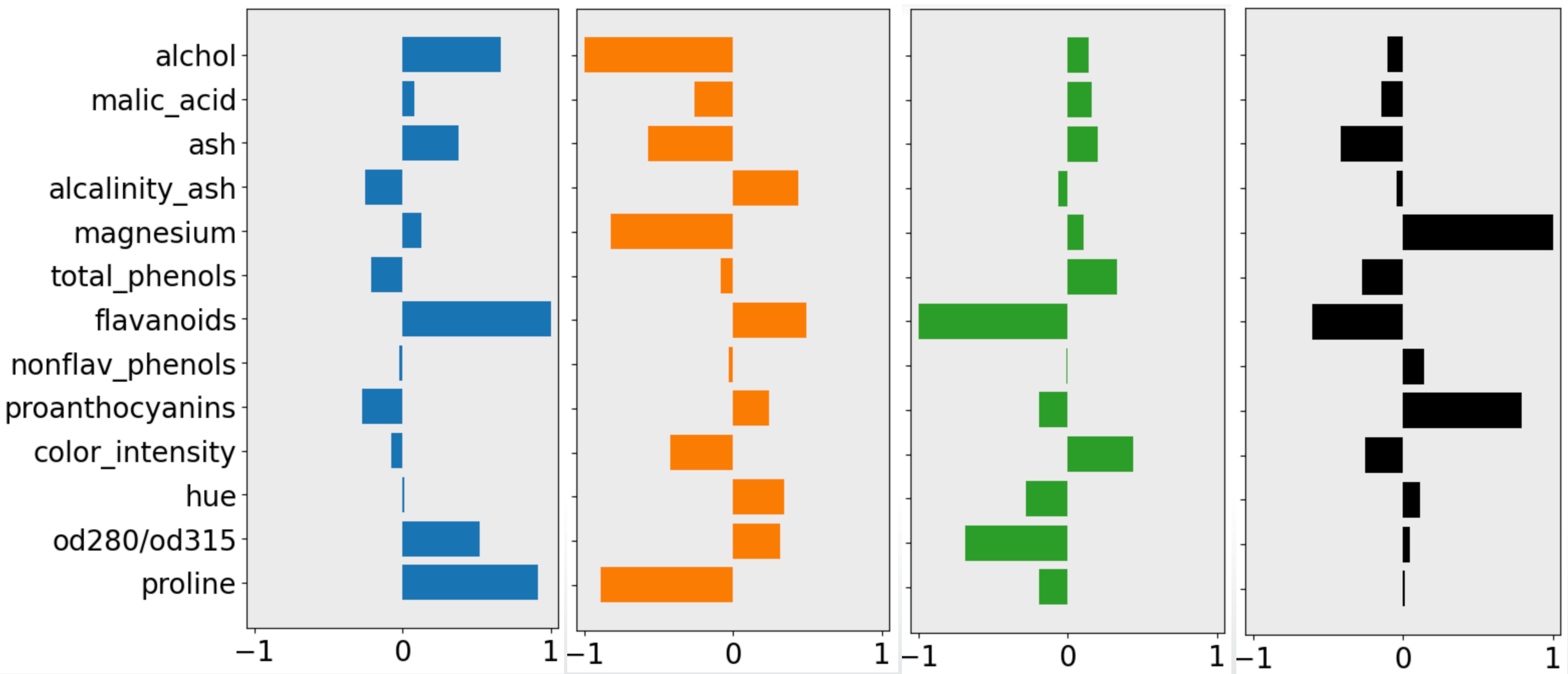


Comparing clusters/groups using linear DR

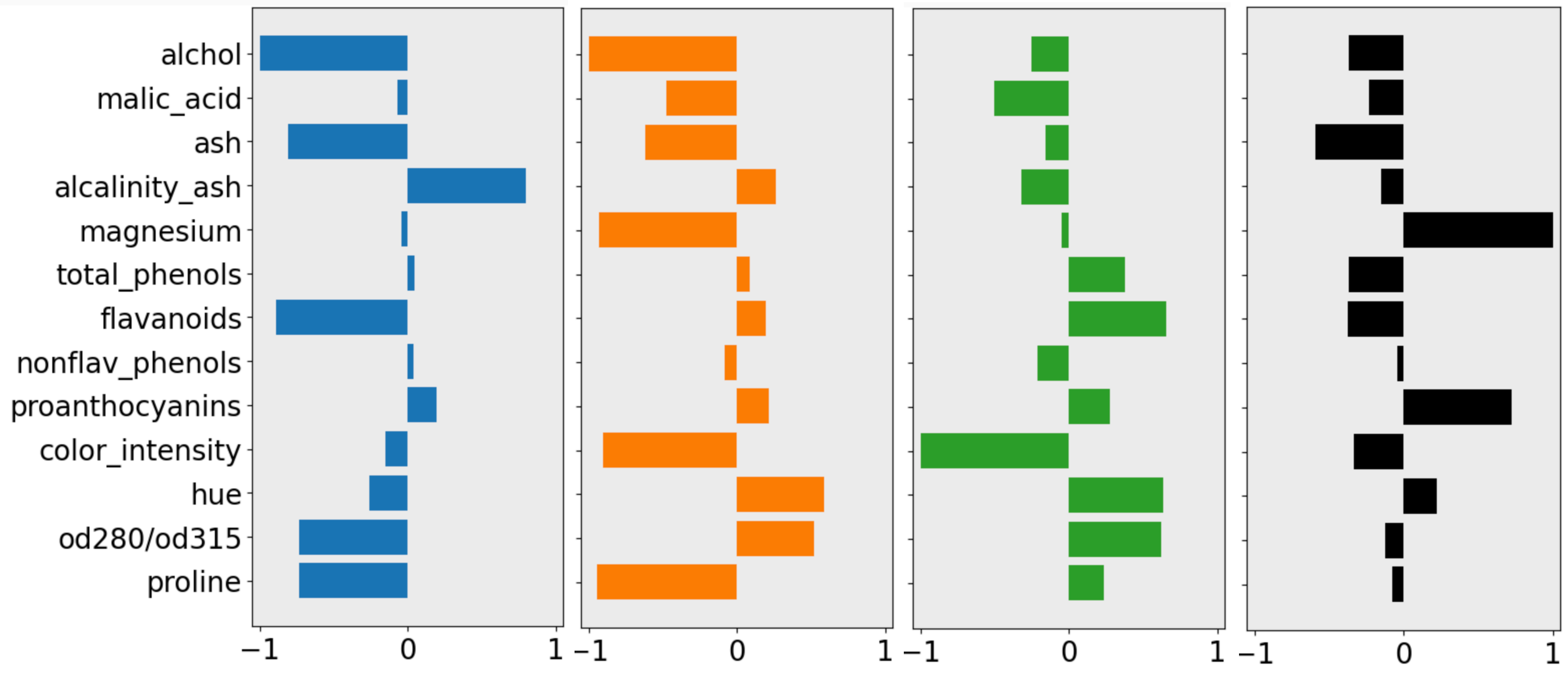
PCA



LDA



ccPCA



Coding exercise (20 minutes)

- Colab notebook link: http://bit.ly/3ZFOoqv

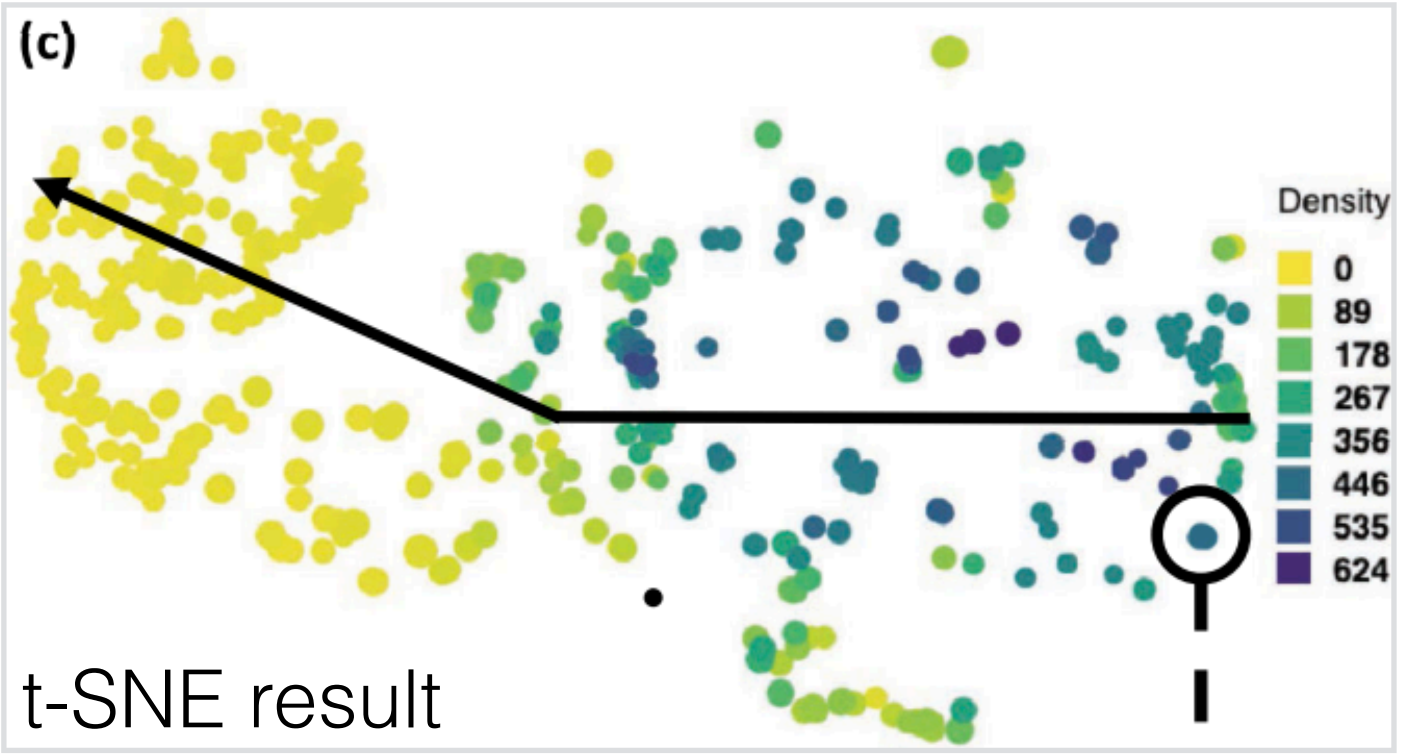
- Select one dataset
- Apply t-SNE or UMAP
- Label instances shown in the DR result (manually or apply clustering)
- Try multiple different interpretation approaches
  1. Univariate statistics-based attribute selection/ranking
  2. PCA
  3. LDA
  4. ccPCA
- Compare outcomes from the above approaches

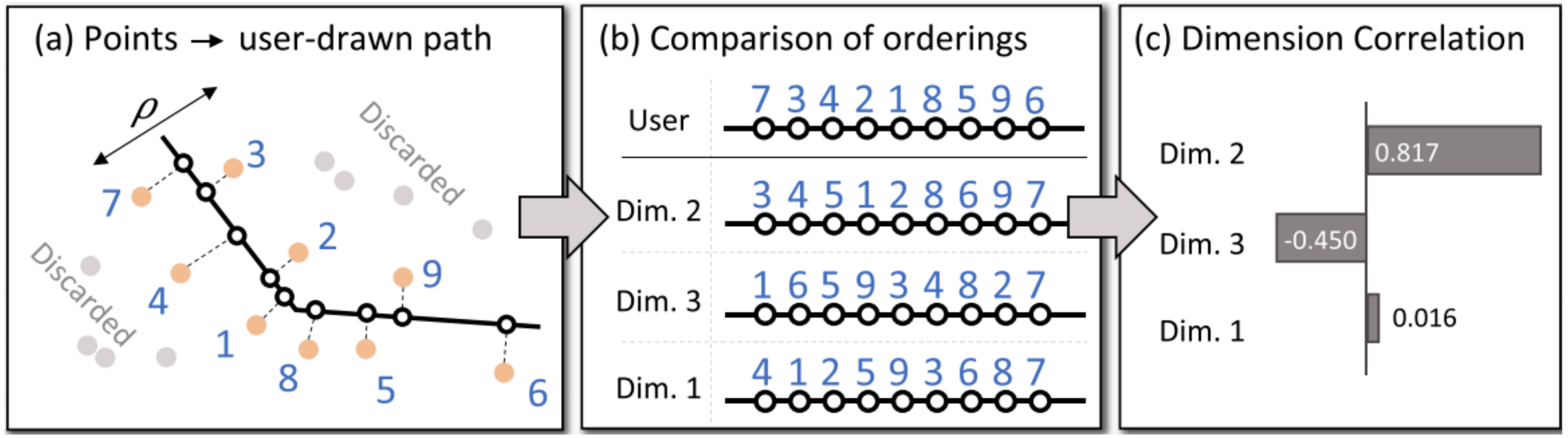
More flexible comparison

- Unified linear comparative analysis (ULCA)
  Fujiwara et al., “Interactive dimensionality reduction for comparative analysis.” *IEEE TVCG*, 2022.
  https://github.com/takanori-fujiwara/ulca
  - More explicitly control of how strongly linear DR **separates** groups and **preserves/eliminates variance** of each group

Other interpretation approach based on observed patterns

- Local direction/path
  - e.g., checking a correlation between a user-drawn path and each attribute
    Chatzimparmpas et al., "t-viSNE: Interactive assessment and interpretation of t-SNE projections." *IEEE TVCG*, 2020.





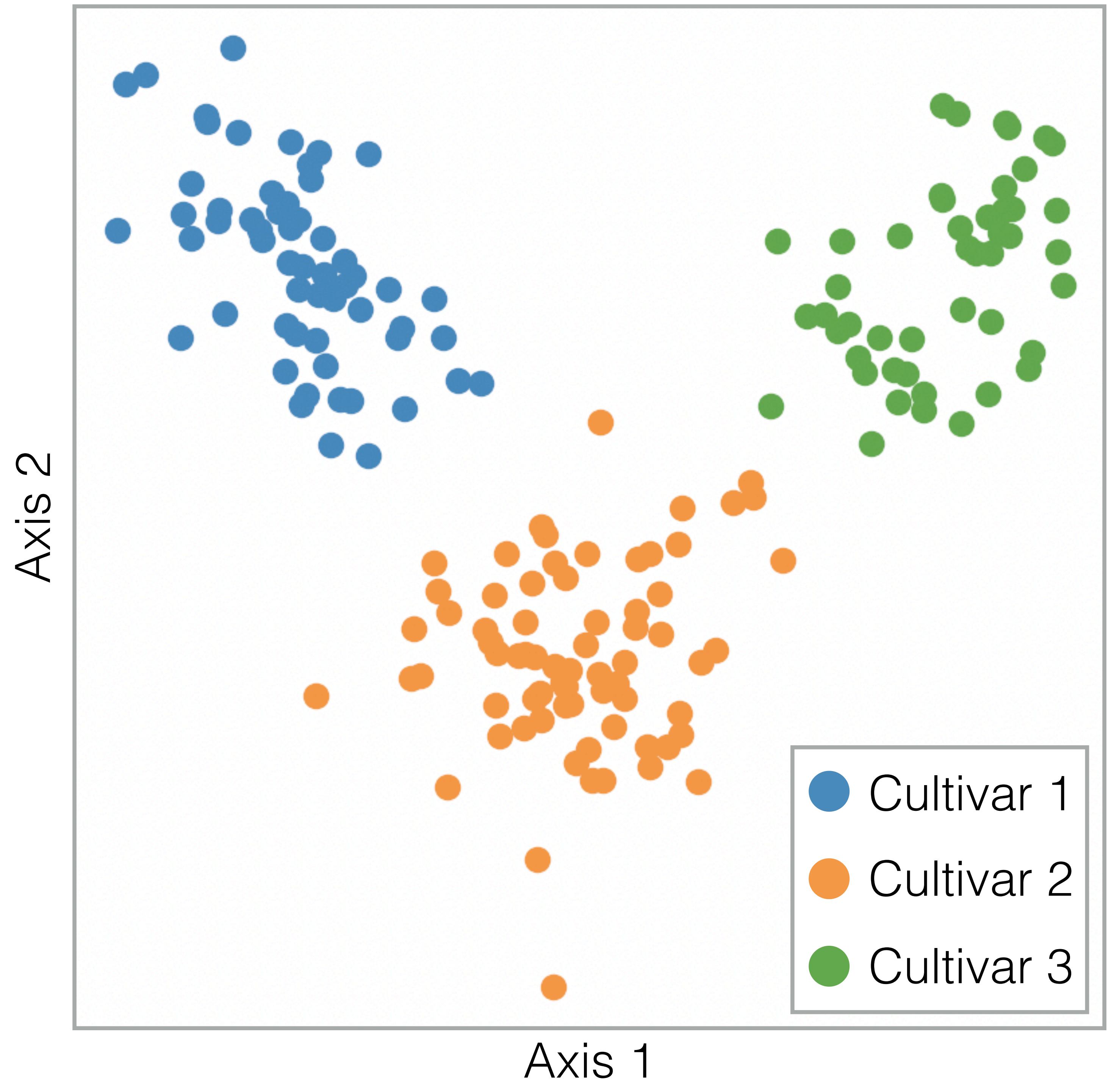
Interpretation of dimensionality reduction results

Wine dataset (13D)

	alchol	malic_acid	ash	alcalinity_ash	magnesium	total_phenols	flavanoid
0	1.52	-0.56	0.23	-1.17	1.91	0.81	1.
1	0.25	-0.50	-0.83	-2.49	0.02	0.57	0.
2	0.20	0.02	1.11	-0.27	0.09	0.81	1.22
3	1.69	-0.35	0.49	-0.81	0.93	2.49	1.47
4	0.30	0.23	1.84	0.45	1.28	0.81	0.66
...	...	...	...	...	...	...	...

*displayed values are
after the standardization

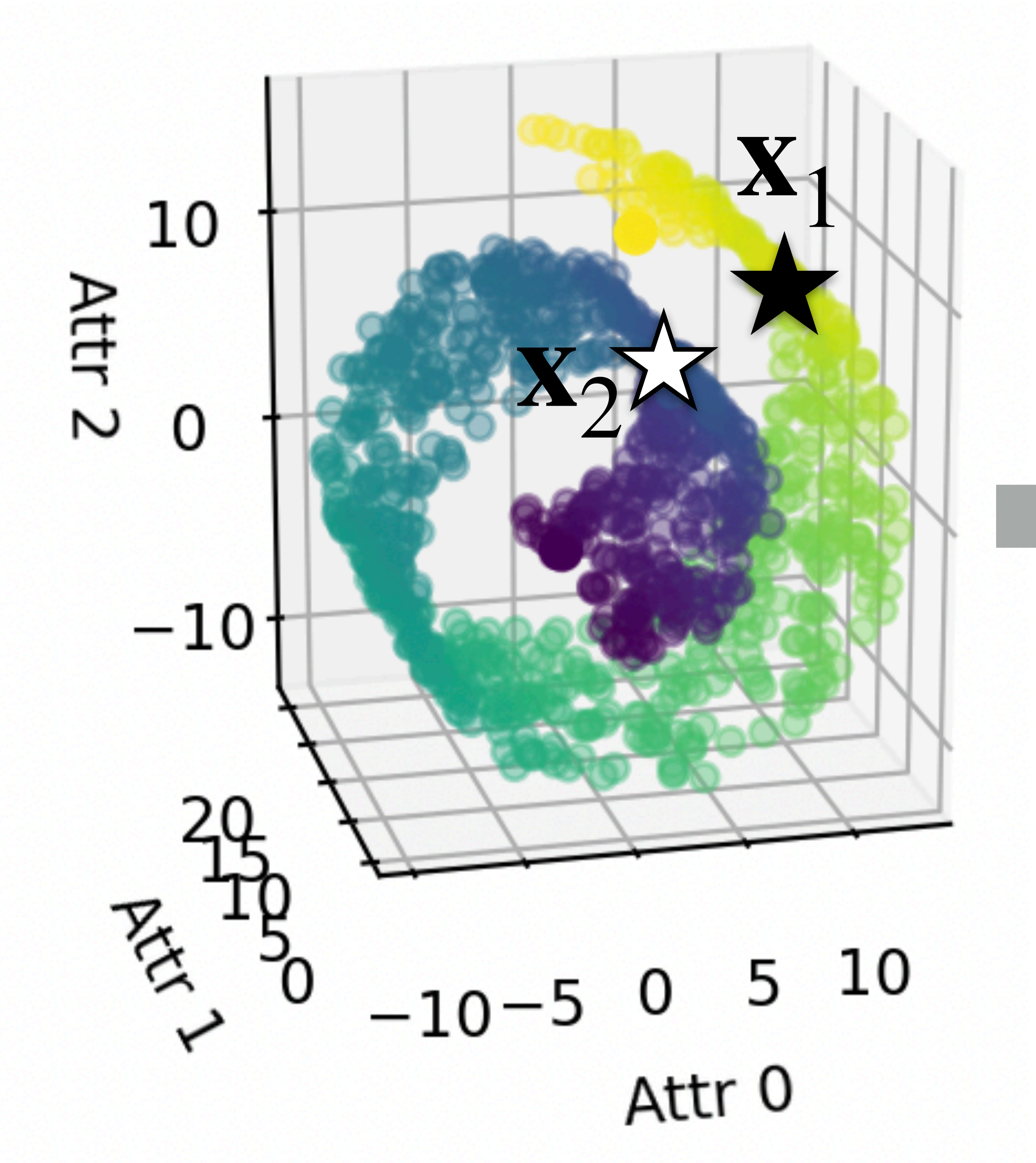
Linear discriminant analysis
LDA



- Interpreting a **lower-dimensional space** — **Linear DR**
- Interpreting based on **observed patterns** (e.g., clusters) — **Nonlinear DR**
- Interpreting from a DR **model/mechanism** level — **Nonlinear DR**

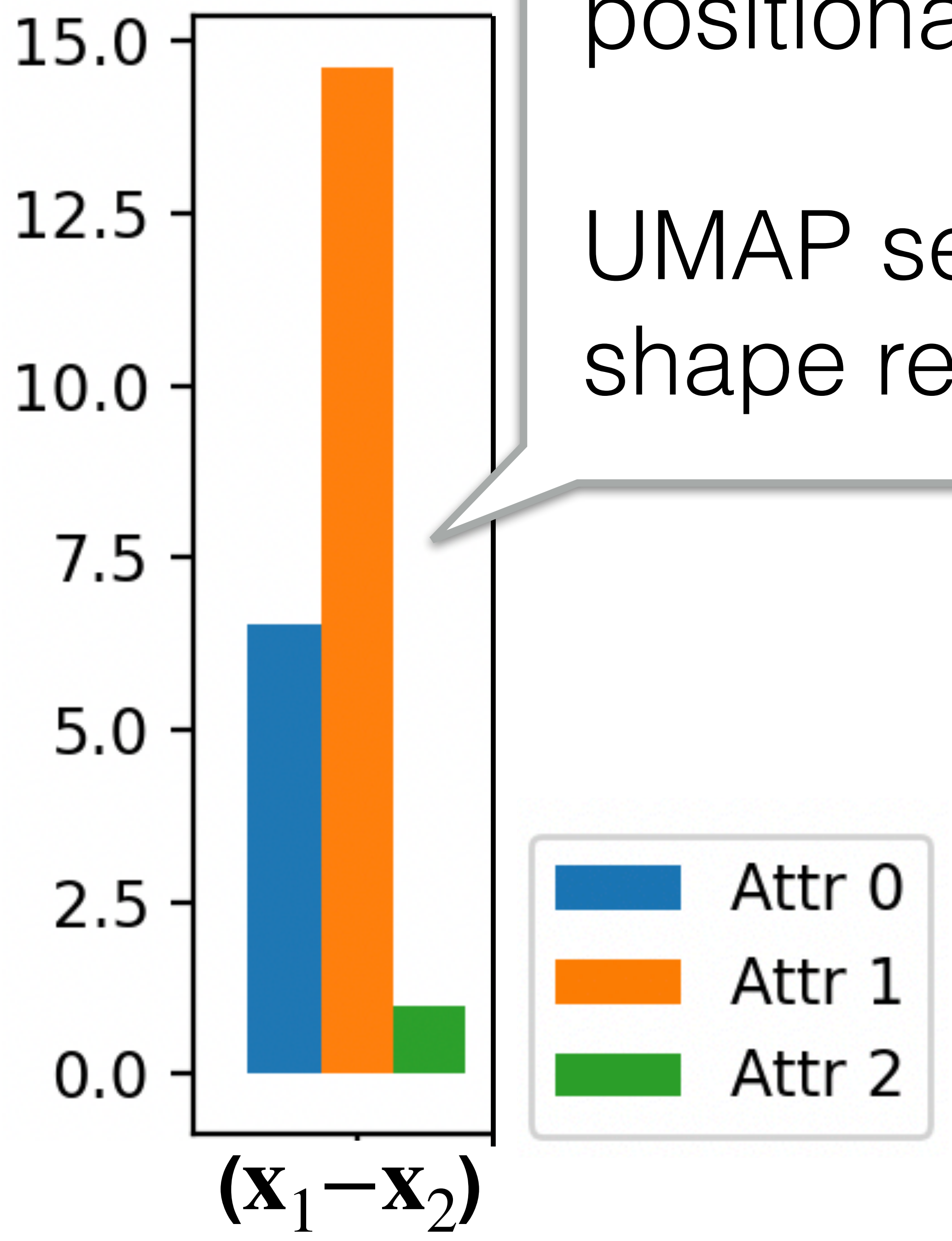
Critical problem of observed-level interpretations

- Comparing (groups of) instances in a DR result does not consider how the DR method projected the instances



Swiss-roll dataset (3D)

UMAP result



Attr1 strongly contributes to the positional difference in the DR result?

UMAP seems to capture the spiral shape related to Attr 0 and Attr 2.

Existing approaches for DR model/mechanism-level interpretations

- Gradient-based
  - Approximately derive **gradient** of each low-dimensional coordinate with respect to their high-dimensional coordinate and use the gradient to understand the DR result

    e.g., Corbugy et al., "Gradient-based explanation for non-linear non-parametric dimensionality reduction." *Data Mining and Knowledge Discovery*, 2024.
    Faust et al., "DimReader: Axis lines that explain non-linear projections." *IEEE TVCG*, 2019.

- Parametric nonlinear DR
  - Use **neural networks-based DR** optimization to produce a parametric mapping ($f_\theta : \mathbf{X} \rightarrow \mathbf{Y}$) and then interpret results based on the mapping

    e.g., Zang et al., "DMT-EV: An explainable deep network for dimension reduction." *IEEE TVCG*, 2024.

  - Build a **substitute parametric model** that mimics a mapping from the original to low-dimensional space or from the low-dimensional to original space

    e.g., Espadoto et al., "UnProjection: Leveraging inverse-projections for visual analytics of high-dimensional data." *IEEE TVCG*, 2023.

Existing approaches for DR model/mechanism-level interpretations

- Gradient-based
  - Approximately derive **gradient** of each low-dimensional coordinate with respect to their high-dimensional coordinate and use the gradient to understand the DR result
    e.g., Corbugy et al., “Gradient-based explanation for non-linear non-parametric dimensionality reduction.” *Data Mining and Knowledge Discovery*, 2024.
    Faust et al., “DimReader: Axis lines that explain non-linear projections.” *IEEE TVCG*, 2019.

- Parametric nonlinear DR
  - Use **neural networks-based DR** optimization to produce a parametric mapping ($f_\theta : \mathbf{X} \rightarrow \mathbf{Y}$) and then interpret results based on the mapping
    e.g., Zang et al., “DMT-EV: An explainable deep network for dimension reduction.” *IEEE TVCG*, 2024.
  - Build a **substitute parametric model** that mimics a mapping from the original to low-dimensional space or from the low-dimensional to original space
    e.g., Espadoto et al., “UnProjection: Leveraging inverse-projections for visual analytics of high-dimensional data.” *IEEE TVCG*, 2023.

Gradient-based interpretation

Corbugy et al., “Gradient-based explanation for non-linear non-parametric dimensionality reduction.” *Data Mining and Knowledge Discovery*, 2020.

- t-SNE’s objective function

$$\arg\min_{\mathbf{Y}} \mathrm{KL\!-\!Divergence}(\mathbf{P}, \mathbf{Q})$$

$\mathbf{P}$: Similarities in the original space

$$p_{ij} = \frac{1}{2n}(p_{i|j} + p_{j|i})$$

$$p_{j|i} = \frac{e^{-\frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_j\|_2^2 \sigma_i^{-2}}}{\sum_{\ell \neq i} e^{-\frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_\ell\|_2^2 \sigma_i^{-2}}}$$

$\mathbf{Q}$: Similarities in the low-dimensional space

$$q_{ij} = \frac{(1 + \|\mathbf{y}_i - \mathbf{y}_j\|_2^2)^{-1}}{\sum_{r=1}^{n} \sum_{\ell \neq r}(1 + \|\mathbf{y}_r - \mathbf{y}_\ell\|_2^2)^{-1}}$$

- Apply Gould et al.’s approximation formula to derive the gradient of the low-dimensional coordinate, $\mathbf{y}_i$, with respect to the original data, $\mathbf{x}_i$

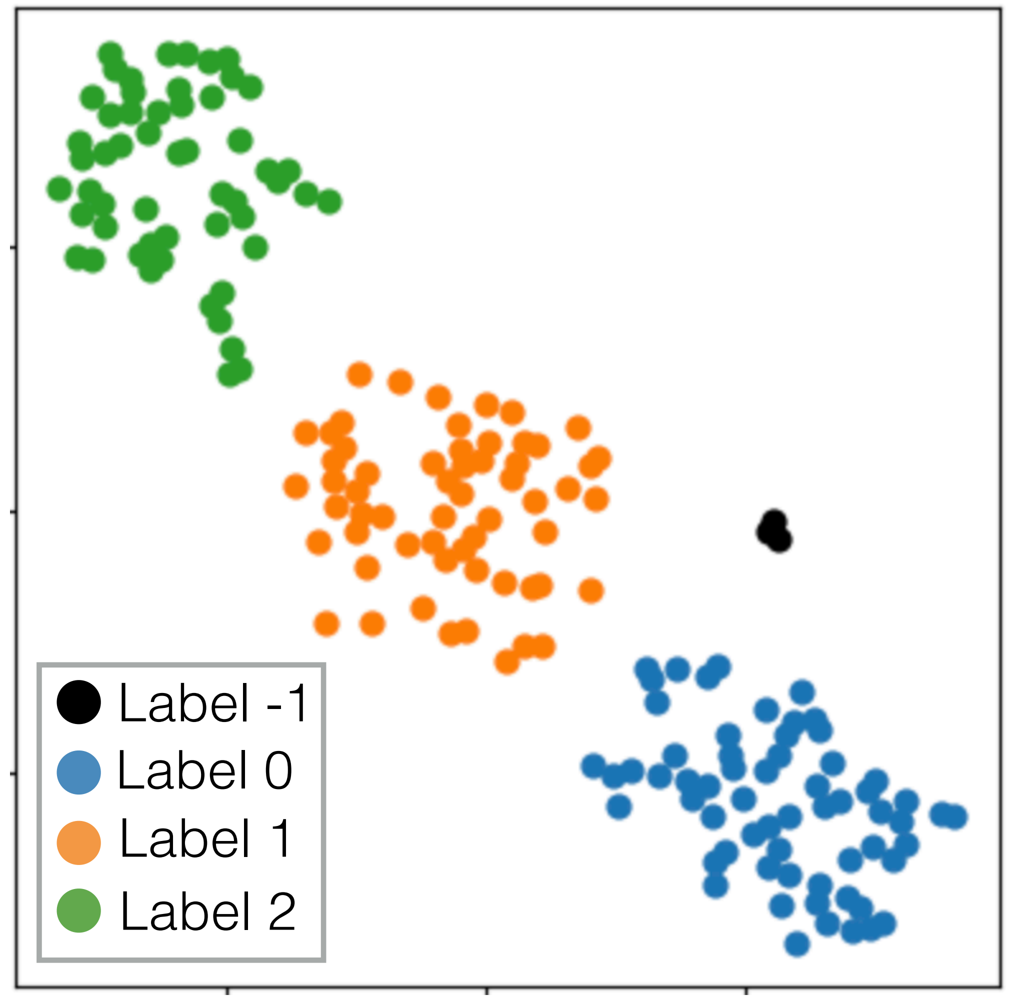
Gould et al., “On differentiating parameterized argmin and argmax problems with application to bi-level optimization.” *arXiv*, 2016.

$$\nabla_{\mathbf{y}_i} f(\mathbf{x}_i, \mathbf{y}_i) = 4 \sum_{j \neq i} (p_{ij} - q_{ij})(1 + \|\mathbf{y}_i - \mathbf{y}_j\|_2^2)^{-1}(\mathbf{y}_i - \mathbf{y}_j)$$

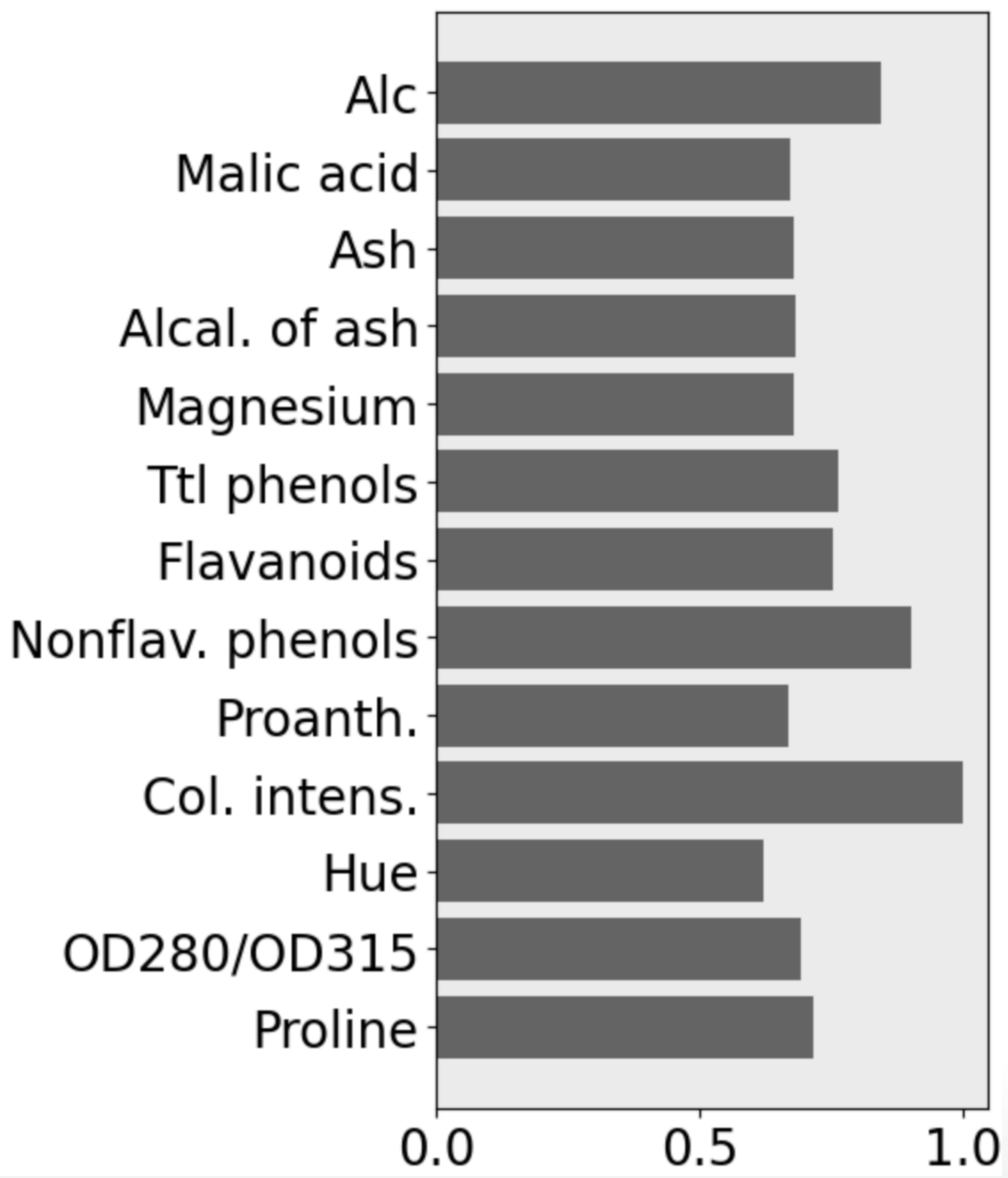
Gradient-based interpretation

Corbugy et al., “Gradient-based explanation for non-linear non-parametric dimensionality reduction.” *Data Mining and Knowledge Discovery*, 2020.

- Complementary tools
  - Global feature importance
    - Compute attribute importances by summing the lengths of the gradient vectors at the coordinates of all instances in the DR result



t-SNE result

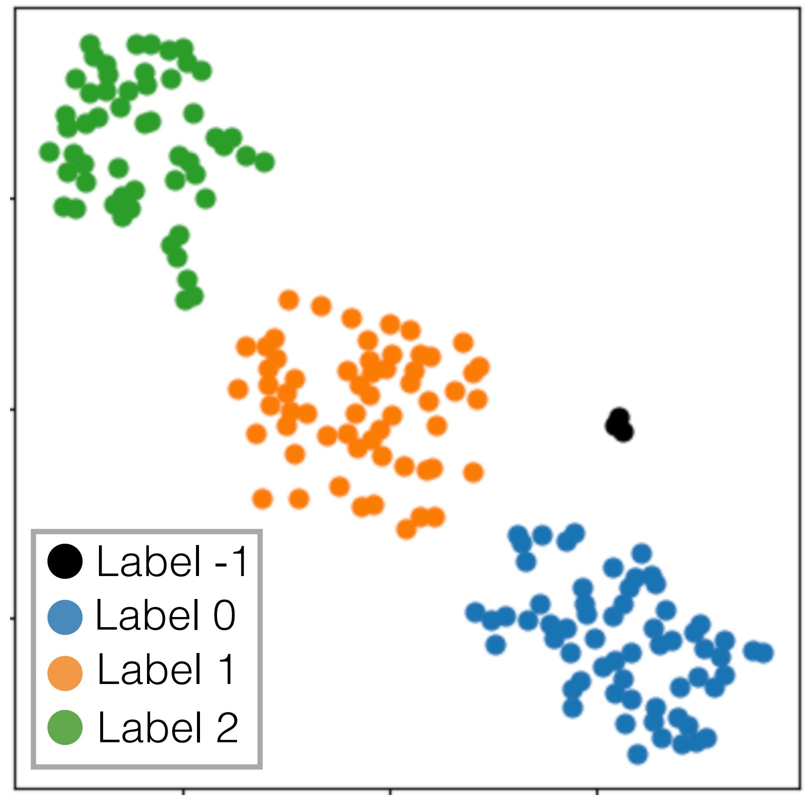


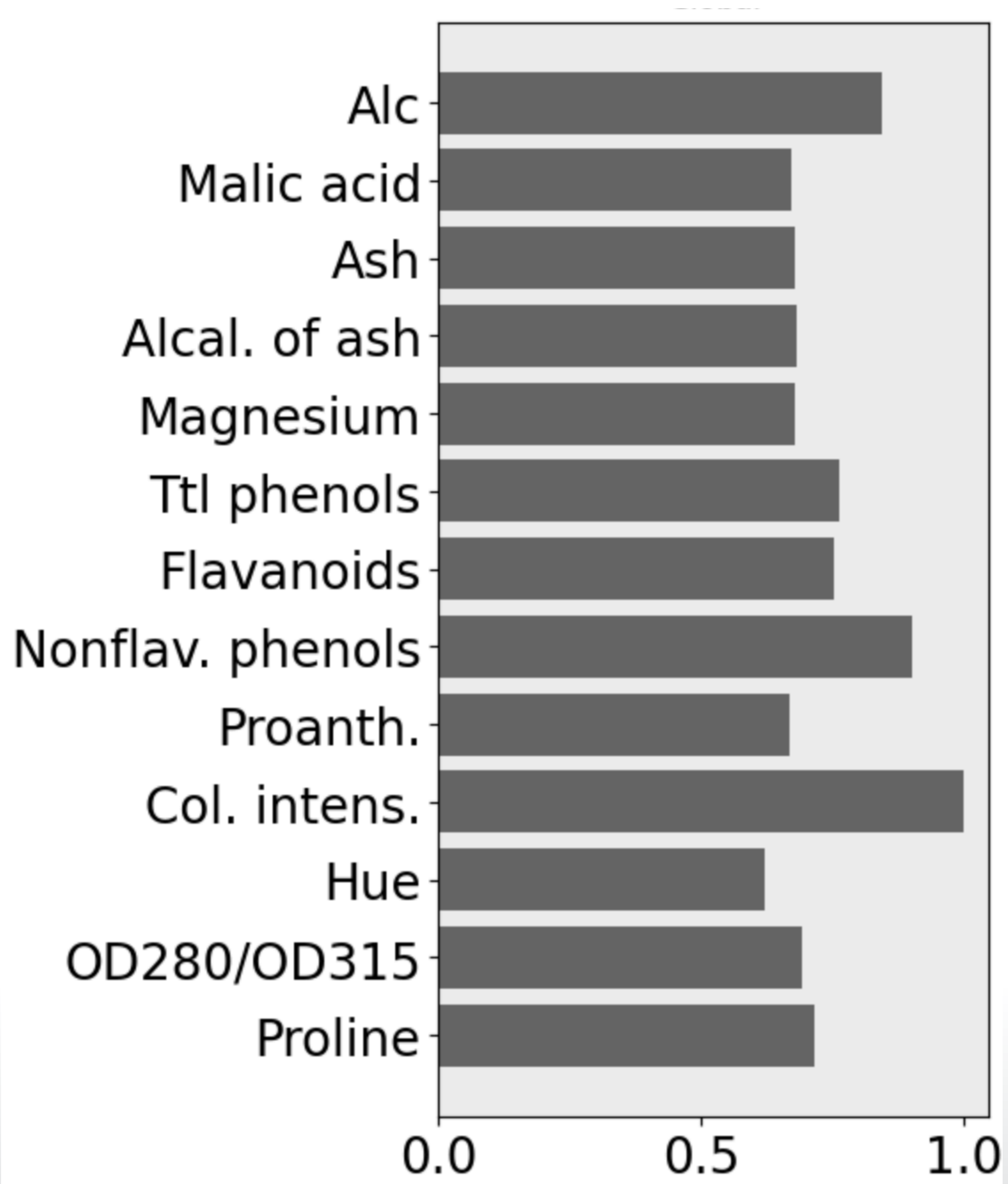
Global feature importance

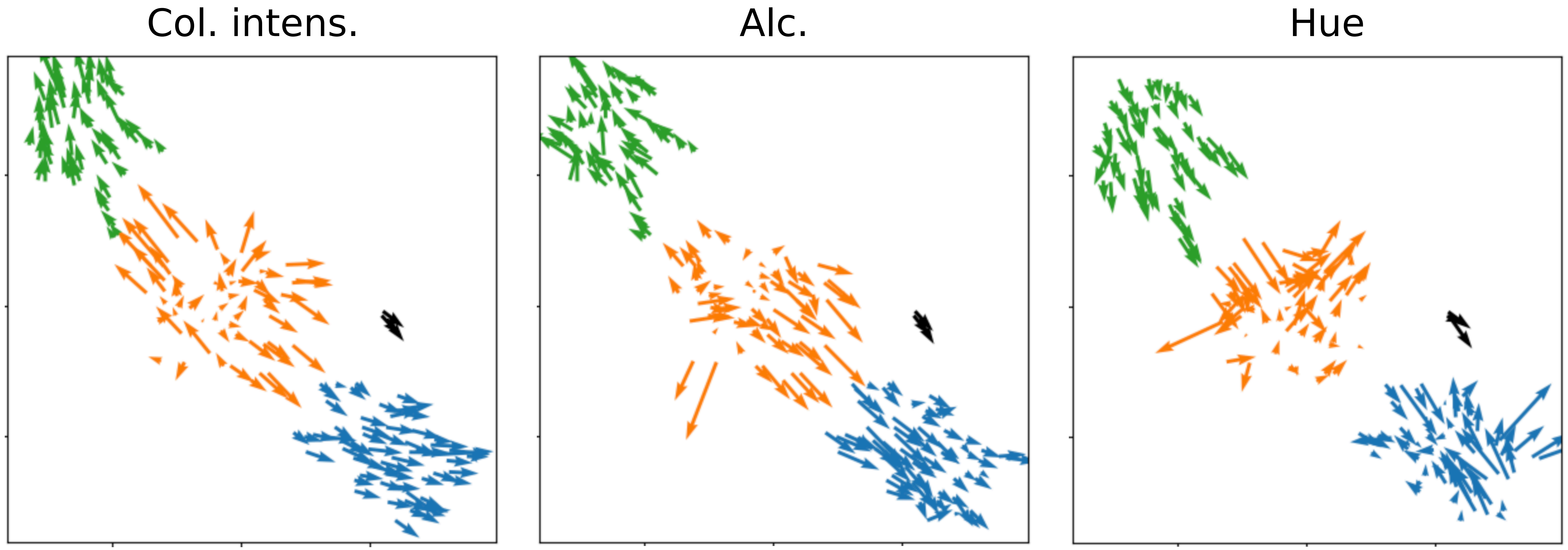
Gradient-based interpretation

Corbugy et al., “Gradient-based explanation for non-linear non-parametric dimensionality reduction.” *Data Mining and Knowledge Discovery*, 2020.

- Complementary tools
  - Global feature importance
  - Vector field visualization for a selected attribute



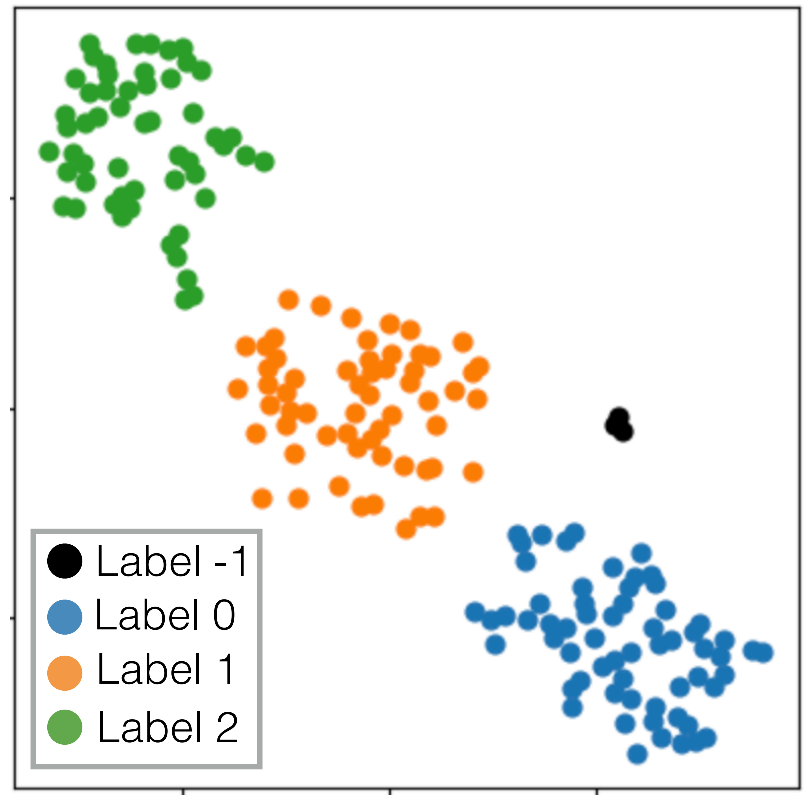




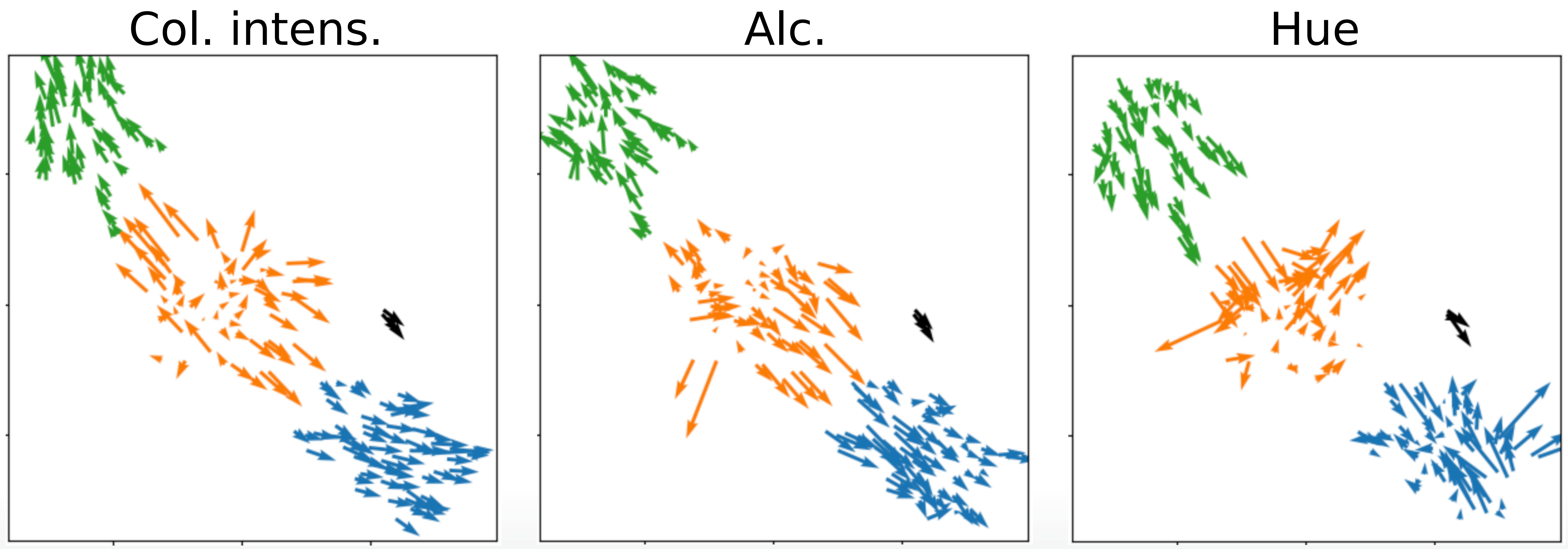
Vector field visualizations

Gradient-based interpretation

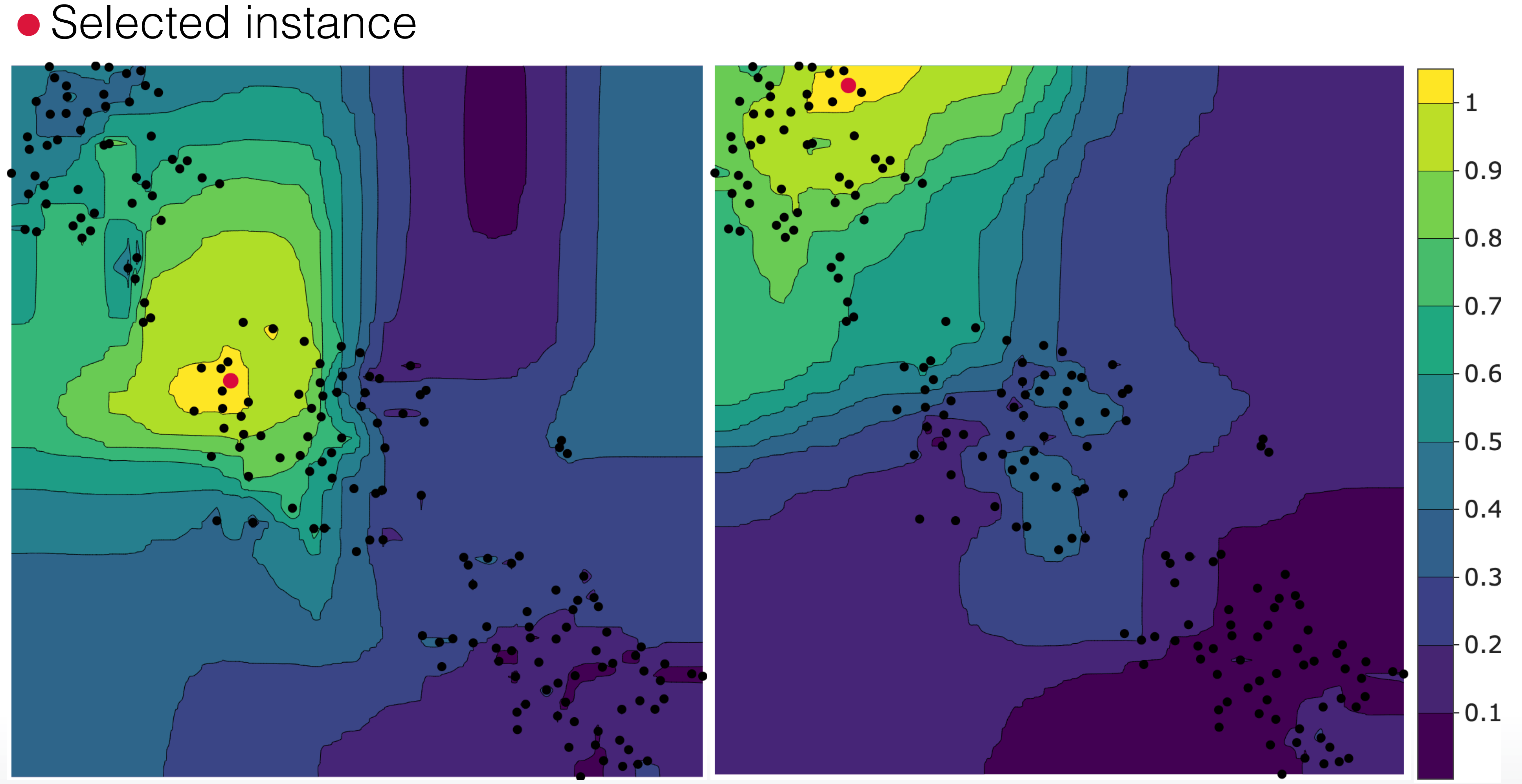
Corbugy et al., “Gradient-based explanation for non-linear non-parametric dimensionality reduction.” *Data Mining and Knowledge Discovery*, 2020.



- Complementary tools
  - Global feature importance
  - Vector field visualization for a selected attribute
  - Explanation scope from a selected instance
    - Based on similarities of the gradient vectors



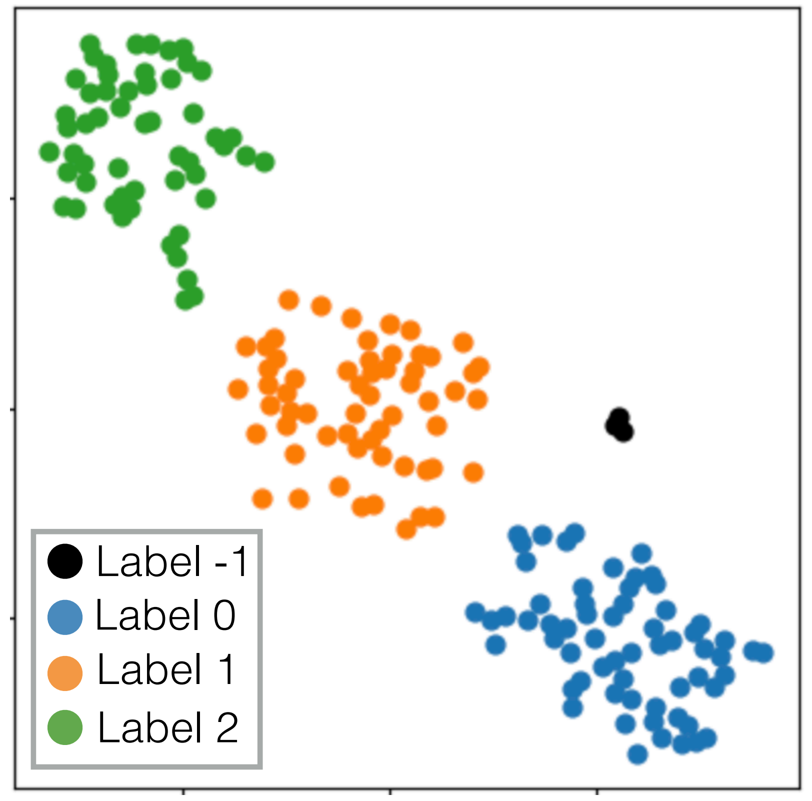
Vector field visualizations



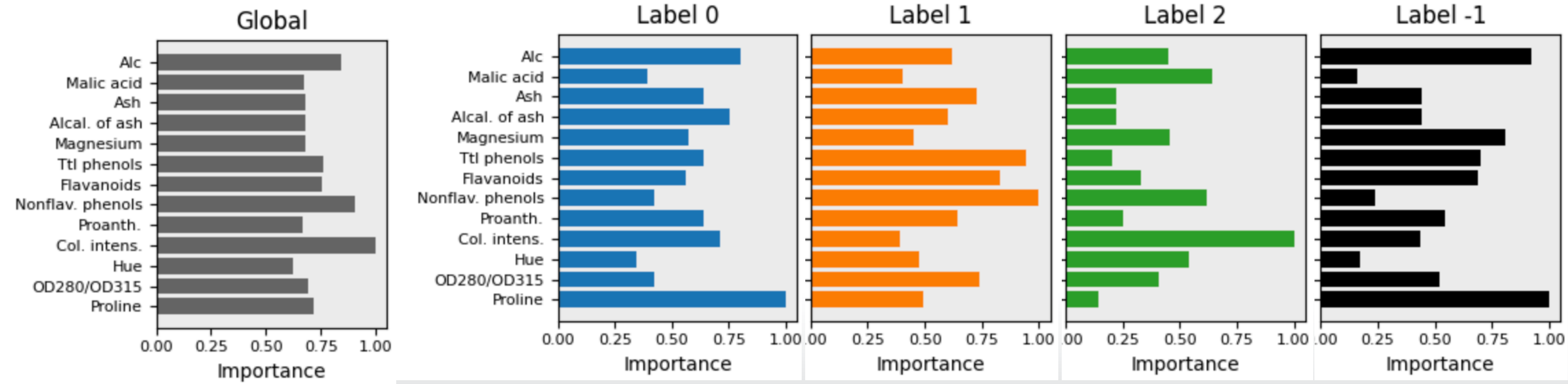
Explanation scope visualizations

Gradient-based interpretation

Corbugy et al., “Gradient-based explanation for non-linear non-parametric dimensionality reduction.” *Data Mining and Knowledge Discovery*, 2020.



- Complementary tools
  - Global feature importance
  - Vector field visualization for a selected attribute
  - Explanation scope from a selected instance
  - Cluster-level feature importance (newly made for this tutorial)
    - The sum of each cluster’s gradient vector lengths



Coding exercise (20 minutes)

- Colab notebook link: https://bit.ly/4kCNPWD

- Select one dataset
- Apply t-SNE
- Apply the gradient-based interpretation method
- Use the complementary visualizations to interpret the t-SNE result
- Compare the interpretation result with the previous coding exercise results (e.g., PCA)

Existing approaches for DR model/mechanism-level interpretations

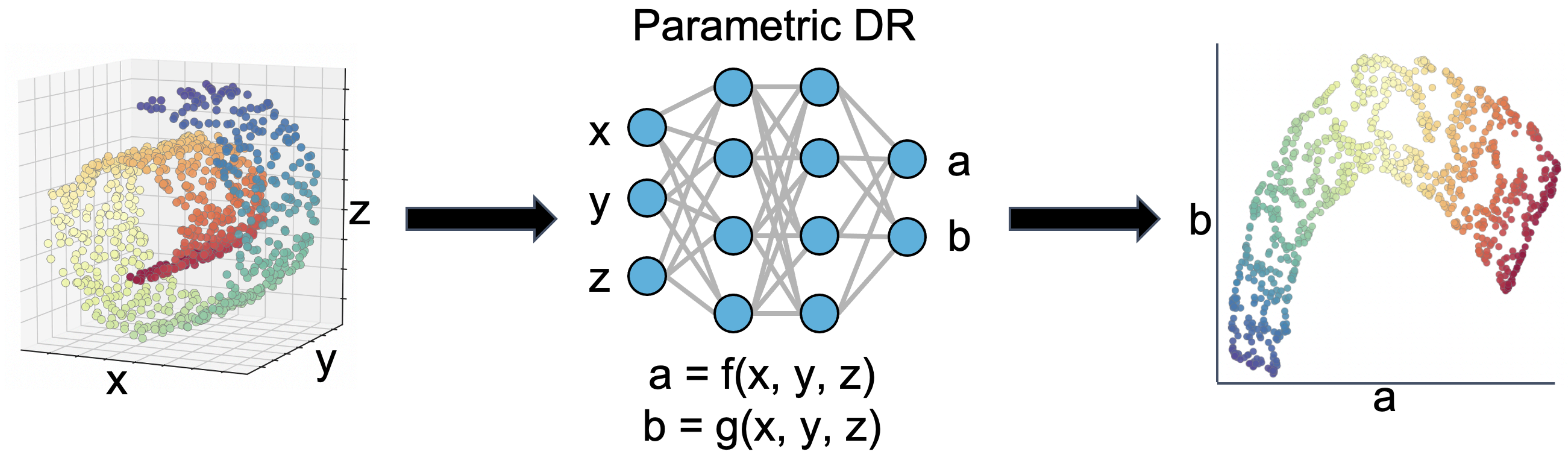
- Gradient-based
  - Approximately derive **gradient** of each low-dimensional coordinate with respect to their high-dimensional coordinate and use the gradient to understand the DR result
    e.g., Corbugy et al., “Gradient-based explanation for non-linear non-parametric dimensionality reduction.” *Data Mining and Knowledge Discovery*, 2024.
    Faust et al., “DimReader: Axis lines that explain non-linear projections.” *IEEE TVCG*, 2019.

- Parametric nonlinear DR
  - Use **neural networks-based DR** optimization to produce a parametric mapping ($f_\theta : \mathbf{X} \rightarrow \mathbf{Y}$) and then interpret results based on the mapping
    e.g., Zang et al., “DMT-EV: An explainable deep network for dimension reduction.” *IEEE TVCG*, 2024.
  - Build a **substitute parametric model** that mimics a mapping from the original to low-dimensional space or from the low-dimensional to original space
    e.g., Espadoto et al., “UnProjection: Leveraging inverse-projections for visual analytics of high-dimensional data.” *IEEE TVCG*, 2023.

Parametric nonlinear DR

- (typically) uses neural networks to have a parametric mapping from the original space to the low-dimensional space
  - Parametric UMAP, parametric t-SNE, etc.

    Sainburg et al., “Parametric UMAP embeddings for representation and semisupervised learning.” *Neural Computation*, 2021.
    Van der Maaten, “Learning a parametric embedding by preserving local structure.” *PMLR*, 2009.
    Hinterreiter et al., “ParaDime: A framework for parametric dimensionality reduction.” *CGF*, 2023.



- Due to the parametric mapping, we can apply various existing interpretation methods designed for deep learning
  - e.g., integrated gradients Sundararajan et al., “Axiomatic Attribution for Deep Networks.” *arXiv*, 2017

Remaining challenges

- Evaluation
  - How can we say which interpretation methods are better than others?
  - How can we ensure that we interpret a DR result well?
- Limited existing work on the DR model/mechanism-level interpretations
- Limited availability of source code
  (no source code, too old to install, no documentation/examples, etc.)
- Many more…

What you learned today

- Fundamentals: Interpretation of dimensionality reduction (DR) results
  - Linear DR
    - **Axis level**
  - Nonlinear DR
    - **Observed-pattern level**
      - Univariate focus
      - Composite variable focus
      - Classifier-based
      - Local pattern correlation
    - **Model-mechanism level**
      - Gradient-based
      - Parametric nonlinear DR
- Practices: Interpretations with existing libraries

Tutorial Materials

https://hyeonword.com/dr-tutorial/

References

Python libraries

- PCA, LDA, t-SNE: Scikit-learn https://scikit-learn.org/
- ccPCA: https://github.com/takanori-fujiwara/ccpca
- ULCA: https://github.com/takanori-fujiwara/ulca
- DimVis: https://github.com/parisa-salmanian/DimVis
- t-SNE gradients explanation: https://github.com/sady410/tsne_gradients_explanation

Referred papers in today's slides

- Bibal et al., "IXVC: An interactive pipeline for explaining visual clusters in dimensionality reduction visualizations with decision trees." *Array*, 2021.
- Chatzimparmpas et al., "t-viSNE: Interactive Assessment and Interpretation of t-SNE Projections." *IEEE TVCG*, 2020.
- Corbugy et al., "Gradient-based explanation for non-linear non-parametric dimensionality reduction." *Data Mining and Knowledge Discovery*, 2020.
- Espadoto et al., "UnProjection: Leveraging inverse-projections for visual analytics of high-dimensional data." *IEEE TVCG*, 2023.
- Faust et al., "DimReader: Axis lines that explain non-linear projections." *IEEE TVCG*, 2019.
- Fujiwara et al., "Supporting analysis of dimensionality reduction results with contrastive learning." *IEEE TVCG*, 2020.
- Fujiwara et al., "Interactive dimensionality reduction for comparative analysis." *IEEE TVCG*, 2022.
- Gould et al., "On differentiating parameterized argmin and argmax problems with application to bi-level optimization." *arXiv*, 2016.
- Hinterreiter et al., "ParaDime: A framework for parametric dimensionality reduction." *CGF*, 2023.
- Marcílio-Jr et al., "Contrastive analysis for scatterplot-based representations of dimensionality reduction." *C&G*, 2021.
- Sainburg et al., "Parametric UMAP embeddings for representation and semisupervised learning." *Neural Computation*, 2021.
- Salmanian et al., "DimVis: Interpreting visual clusters in dimensionality reduction with explainable boosting machine." *MLVis*, 2024.
- Van der Maaten, "Learning a parametric embedding by preserving local structure." *PMLR*, 2009.
- Sundararajan et al., "Axiomatic Attribution for Deep Networks." *arXiv*, 2017.
- Wen et al., "Robust sparse linear discriminant analysis." *IEEE Trans. Circuits Syst. Video Technol.*, 2018.
- Zang et al., "DMT-EV: An explainable deep network for dimension reduction." *IEEE TVCG*, 2024.
- Zou et al., "Sparse principal component analysis." *J. Comput. Graph. Stat.*, 2006