From black box to glass box: Making UMAP interpretable with exact feature contributions

We transform UMAP from a black box into a glass box. By learning the embedding function with a certain type of deep network, we can compute equivalent linear mappings of the input features that exactly reconstruct each embedding, revealing the heretofore hidden logic of UMAP.

Purpose

UMAP is a ubiquitous tool for low-dimensional visualization of high-dimensional datasets. UMAP learns a low-dimensional mapping from the nearest-neighbor graph structure of a dataset, often producing visually distinct clusters of data that align with known labels (e.g., cell types in a gene expression dataset). While the learned relationship between the input features and the embedding positions can be useful, the nonlinear UMAP embedding function also makes it difficult to directly interpret the mapping in terms of the input features.

Here, we show how to enable interpretation of the nonlinear mapping through a modification of the parametric UMAP approach, which learns the embedding with a deep network that is locally linear (but still globally nonlinear) with respect to the input features. This allows for the computation of a set of exact feature contributions as linear weights that determine the embedding of each data point. By computing the exact feature contribution for each point in a dataset, we directly quantify which features are most responsible for forming each cluster in the embedding space. We explore the feature contributions for a gene expression

dataset from this "glass-box" augmentation of UMAP and compare them with features found by differential expression.

View the notebook

The **full pub** is available [here](#).

The **source code** to generate it is available in [this GitHub repo](#) (DOI: [10.5281/zenodo.17478720](#)).

In the future, we hope to host [notebook pubs](#) directly on our publishing platform. Until that's possible, we'll create stubs like this with key metadata like the DOI, author roles, citation information, and an external link to the pub itself.

Contributors (A–Z)

- **Audrey Bell**: Critical Feedback
- **James R. Golden**: Conceptualization, Formal Analysis, Investigation, Software, Visualization, Writing
- **Evan Kiefl**: Validation
- **George Sandler**: Critical Feedback, Visualization
- **Ryan York**: Supervision