Sheaves as a Framework for Understanding and Interpreting Model Fit

Henry Kvinge, Brett Jefferson, Cliff Joslyn, Emilie Purvine

Motivation:

- In machine learning (ML) we usually think of datasets as sets without additional structure between individual datapoints.
- On the other hand, current benchmark datasets are often large and complex with extensive metadata that defines subpopulations with distinct dynamics. Understanding performance on subpopulations is important.
- In such a setting, it makes more sense to identify a dataset as a *dataspace*, with a topology defined by metadata.
- **We show how (pre)sheaves on this dataspace can be used to better understand and interpret ML models.**

The process

(1) Put a domain-informed topology on a dataset

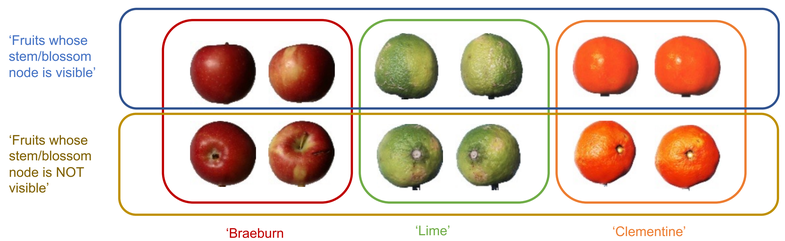
We think of a dataset $D = \{x_i\}_{i \in I} \subseteq Y$ as a function $f: I \to Y$ on an underlying set I.
The set I could be:

- A set of sensors from which we have collected readings,
- A set of genes for which we are measuring expression levels,
- A set of ID's for images we will be encoding with a deep network.

A dataset often comes with metadata attached to it. Elements of an image dataset may be split by:

- Labels on objects in the image,
- The camera that was used to take images,
- The location or time the image was taken.

Example: *Fruits 360* (Muresan, 2018)

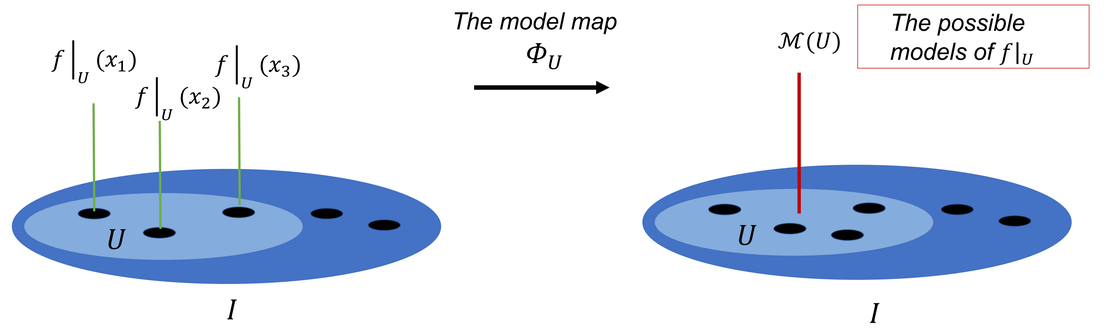


We assume that dataset D has metadata that allows us to identify significant subpopulations $U_1, U_2, \ldots, U_k$.

We use $U_1, U_2, \ldots, U_k$ as the subbase of a new database topology $\mathcal{T}$ on D.

(2) Create a *data sheaf* and *model presheaf*

For a presheaf $\mathcal{F}$ on topology $\mathcal{T}$, an *assignment* is a choice of sections, $\{a_U\}_{U \in \mathcal{T}}$.



$\mathcal{D}(U)$ consists of all functions $f: U \to Y$

Each element $f \in \mathcal{D}(U)$ *corresponds to a measurement of data on* $U \in \mathcal{T}$

$\mathcal{M}(U)$ consists of a user chosen family of models (e.g., linear models).

A section of $\mathcal{M}(U)$ is a model of data values from U

Data sheaf $\mathcal{D}$ — Model presheaf $\mathcal{M}$

Take-away:

- *The data that we collect is an assignment for a sheaf, the set of functions we use to model subsets of the dataset is an assignment for a presheaf.*
- *The process of modeling data can be understood as a map Φ from the data sheaf to the model presheaf.*

(3) Measure of model fit by closeness to being a presheaf morphism

Question: How can we use this framework to understand statistics on various subpopulations of a dataset?
One answer: Look at the extent to which the model map fails to be a presheaf morphism.

$$\begin{array}{ccc} \mathcal{D}(U) & \xrightarrow{\Phi_U} & \mathcal{M}(U) \\ \downarrow Res_V^U & & \downarrow Res_V^U \\ \mathcal{D}(V) & \xrightarrow{\Phi_V} & \mathcal{M}(V) \end{array}$$

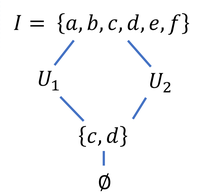
The dataset + metadata

Researcher	# of papers
a	5
b	6
c	8
d	7
e	4
f	5

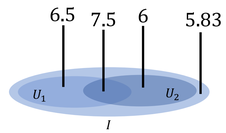
Known collaborations

- $U_1 = \{a, b, c, d\}$
- $U_2 = \{c, d, e, f\}$

The topology



The model sheaf: data mean



Analysis

- Biggest difference from global mean arises from $\{c, d\}$, researchers who are in two collaborations.
- Collaboration U_1, deviates more from the global mean than U_2.

Example (bibliometrics)

U.S. DEPARTMENT OF ENERGY BATTELLE
PNNL is operated by Battelle for the U.S. Department of Energy

For additional information, contact:
Henry Kvinge | henry.kvinge@pnnl.gov

Pacific Northwest NATIONAL LABORATORY
www.pnnl.gov