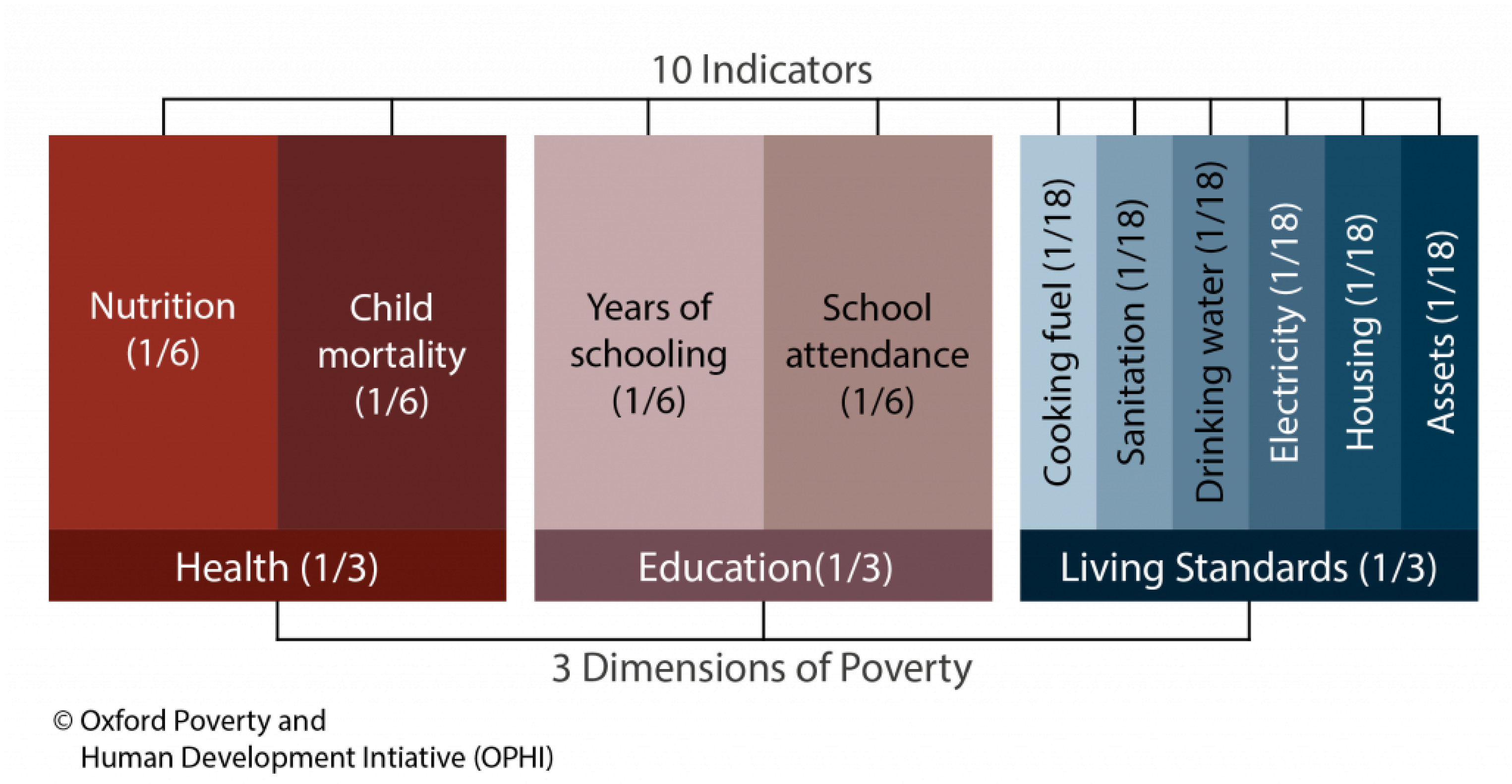
Multidimensional Poverty Prediction

Girela, Ignacio[1,2] & **Koplin, Eric**[3]

[1]CONICET, [2]Universidad Nacional de Córdoba, [3]Vanderbilt University Medical Center

Introduction

- The United Nations Development Program (UNDP) publishes the global Multidimensional Poverty Index (MPI), a weighted composite index based on indicators related to Sustainable Development Goals (SDGs) [2].
- The global MPI complements the World Bank's 2.15$-a-day international poverty measure to monitor progress in key dimensions of human-development.



Motivation

Carrying out multi-topic household surveys is the best way to monitor the socioeconomic welfare status of a society.

Fig 1: Data collection in household surveys

However ...

- Data collection requires high administrative costs.
- Developing countries often depend upon international NGOs or foreign aid.
- Qualified personnel are required for taking anthropometric measures.
- High levels of non-response can undermine the representativeness of a cluster sample.
- Conflicts or dangerous areas block proper data collection.

Objective

Assuming it is not possible to collect data of a multi-topic household survey in certain areas, we aim to **predict the multidimensional poverty levels of Primary Sample Units (PSU) using only the information of the closest clusters** $\longrightarrow$ *Inductive Learning task*

Data

Household surveys $\longrightarrow$ Structured data

We use the Nigeria 2018 Demographic Health Survey (DHS). Household survey data have strict privacy policies and permissions. Then, geographical information only includes the (slightly disturbed) centroids of the PSU.

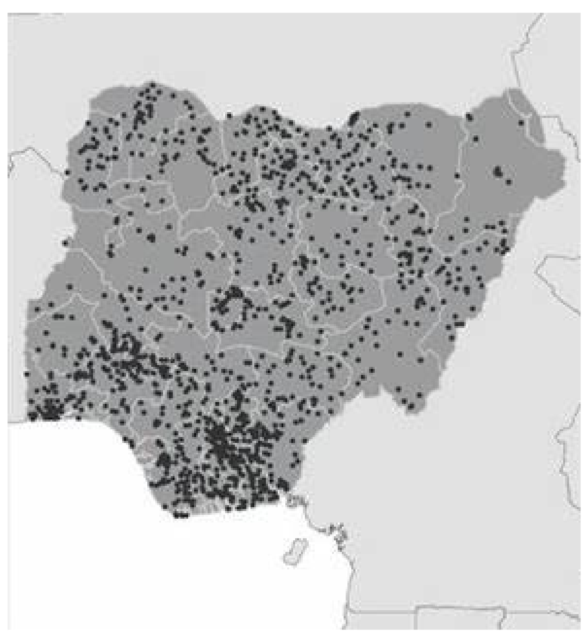


Fig 2: Clusters distribution Nigeria DHS

Methodology & Experiments

Given that we need to learn the importance of each cluster's neighbor information, we apply a **Graph Attention Network** (GAT) for regression [3, 1].

Procedure:

1. Compute euclidean distance matrix based on PSU centroids to build the adjacency matrix.
2. Select sociodemographic features
   (a) Without including MPI indicators (simplistic model using only basic demographic features)
   (b) Including MPI indicators
3. Split graph data: train/validation/test sets.
4. Build GAT architecture

```
print(model)

GAT(
  (gat1): GATv2Conv(53, 32, heads=16)
  (norm1): LayerNorm((512,), eps=1e-05, elementwise_affine=True)
  (gat2): GATv2Conv(512, 1, heads=1)
  (norm2): LayerNorm((256,), eps=1e-05, elementwise_affine=True)
  (gat3): GATv2Conv(256, 16, heads=1)
)
```

5. Train the model and predict the MPI level in test clusters.
6. Compare results with a spatial RBF kernel-based model and a benchmark (a kernel that includes MPI indicators yielding the lowest error).

Results

Nodes: 1389
Edges: 8956
Features: (a) 33 and (b) 43

	(a) Without MPI indicators		(b) With MPI indicators		
	GAT	RidgeKernel	GAT	RidgeKernel	Benchmark
MSE	0.0186	0.0125	**0.00641**	0.0126	0.00102
r^2	0.473	0.549	**0.819**	0.543	0.963

Table 1: Comparison of results

Conclusions

Predicting multidimensional poverty levels in areas where data is not available is a critical challenge in the context of poverty reduction strategies.

Widely used spatial statistical learning methods rely on measurement features of target clusters (features in test sets). **In cases where features are unavailable due to administrative costs associated with survey data collection, Graph Attention Networks demonstrate significant potential in predicting cluster-level multidimensional poverty.**

Discussion & Analysis

Limitations:

- GIS data derived from household surveys is seldom openly accessible.
- Living areas (rural and urban) should be take into account when building the adjacency matrix.

Further research:

- A more detailed exploration of results: in which 'kind of' clusters the model perform poorly? (e.g., isolated rural areas? rural areas far from water sources?).
- Explore how partially observed (easy-to-access) data of target nodes can be included in the model (e.g., satellite imagery).
- A similar approach can be applied with numerous countries where each at a province level (nodes). Then, it is possible to predict MPI with neighbor countries.

References

[1] Shaked Brody, Uri Alon, and Eran Yahav. "How Attentive are Graph Attention Networks?" In: *International Conference on Learning Representations.* 2022.

[2] United Nations Development Programme (UNDP) and Oxford Poverty and Human Development Initiative (OPHI). *Global Multidimensional Poverty Index 2024: Unpacking deprivation bundles to reduce multidimensional poverty.* 2024.

[3] Petar Veličković et al. "Graph attention networks". In: *International conference on learning representations.* 2018.

Code available here: