



Graph Convolutional Networks for Temporal Vulnerability Classification in Obstetric Hospitalization Flows from SUS

Tiago Mota de Oliveira - UFPR | Marcel Antunes Raposo - UFPR | Claudemir Casa - UFPR

Felipe Lourenço - USP | Luciano Silva - UFPR

UFPR: Universidade Federal do Paraná | USP: Universidade de São Paulo

tiago.mota@governor.ltd · maraposo.research@gmail.com · claudemir.casa@governor.ltd · felipe.lourenco.ca@gmail.com · luciano@ufpr.br

Introduction

Municipalities of a regional health network are not independent observations but nodes connected by recurrent patient referral flows, i.e. network-structured data addressed by graph neural networks for time series [2]. We formulate intermunicipal obstetric hospitalization flows from SUS as monthly hub-referral graphs and evaluate whether a Graph Convolutional Network (GCN) [1] can classify municipalities into three vulnerability levels defined from a composite dependency index.

Contributions: the hub-referral graph formulation, a controlled diagnosis of a negative graph-learning result, evidence separating unstable labels from a stable model ranking, and a visual analytics prototype for spatial and temporal interpretation.

Data and Problem Formulation

SIH/SUS [4] records of 2025 (ICD-10 chapter O) yield 35,022 obstetric hospitalizations for the 29 municipalities of the 2nd Health Region of Paraná. Each monthly graph G_t links municipality i to the hub c (Curitiba) whenever $F_{t[i, c]} > 0$, with the dependency ratio $D_{i,t}$ as edge attribute, so every G_t is a star centred on Curitiba. Nodes carry seven features (volume, outflow and accessibility), and the vulnerability index is:

$$V_{i,t} = 0.4 \tilde{D}_{i,t} + 0.3 \tilde{C}_{i,t} + 0.2 \tilde{T}_i + 0.1 \tilde{R}_i$$

weighting normalized dependency on Curitiba, flow to Curitiba, travel time and road distance. It is discretized into low (< 0.33), moderate ($0.33 \leq < 0.66$) and high (≥ 0.66): index-derived analytical classes, not clinically validated risk categories.

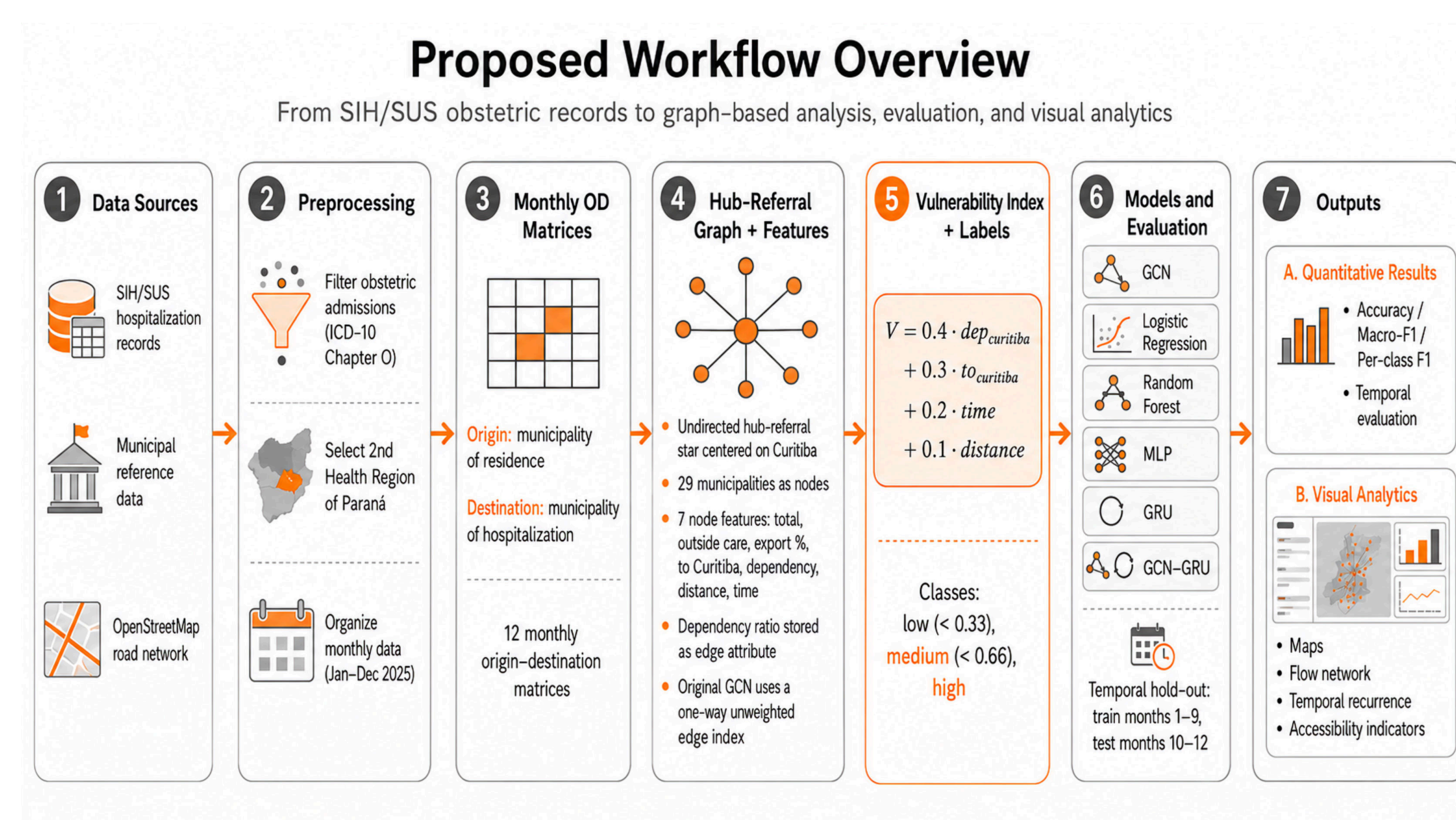


Figure 1. Overview of the proposed workflow.

Models and Experimental Setup

A two-layer GCN ($7 \rightarrow 16 \rightarrow 3$) is compared with Logistic Regression, Random Forest, an architecture-matched MLP (GCNConv replaced by Linear), a GRU and a GCN-GRU [3]. Temporal hold-out: training on months 1–9, test on months 10–12 ($29 \times 3 = 87$ node instances). Every stochastic model is repeated over 30 seeds and the graph is ablated (random-edge, reversed, symmetrized, dependency-weighted).

Code and materials: github.com/motaoliveiraufpr/susflow

Table 1. Graph ablation and baseline performance.

Model	Acc	Macro-F1	F1-C2
Unweighted graph ablation			
GCN (one-way edge index)	0.920 ± 0.011	0.787 ± 0.029	0.483 ± 0.078
GCN (random-edge)	0.738 ± 0.038	0.662 ± 0.088	0.497 ± 0.241
GCN (dependency-weighted)	0.879 ± 0.026	0.705 ± 0.028	0.312 ± 0.056
GCN (reverse-edge)	0.960 ± 0.007	0.896 ± 0.005	0.750 ± 0.000
GCN (symmetrized)	0.974 ± 0.005	0.938 ± 0.016	0.853 ± 0.044
Logistic Reg.	0.954	0.934	0.889
Random Forest	0.954	0.934	0.889
MLP (no graph)	0.977 ± 0.000	0.951 ± 0.000	0.889 ± 0.000
Class-balanced & recurrent (G=graph, T=time)			
GCN + bal. (G)	0.913 ± 0.013	0.819 ± 0.017	0.592 ± 0.031
GRU + bal. (T)	0.968 ± 0.009	0.936 ± 0.030	0.863 ± 0.084
GCN-GRU + bal. (G+T)	0.956 ± 0.025	0.900 ± 0.039	0.768 ± 0.085
MLP + bal. (-)	0.987 ± 0.004	0.988 ± 0.013	0.989 ± 0.034

Results

- **The graph does not help:** the MLP, i.e. the GCN without message passing, is best (0.977 vs. 0.920 accuracy; 0.951 vs. 0.787 macro-F1).
- **Graph construction matters:** reversing every edge raises accuracy to 0.960 and symmetrizing to 0.974 ($p < 10^{-4}$), while random edges (0.738) and $D_{i,t}$ -weighted edges (0.879) hurt.
- **Balancing and recurrence:** class balancing improves the macro-F1 of MLP, GCN and GRU; GCN-GRU (0.956) beats the GCN but stays below the balanced MLP (0.987).
- **Sensitivity:** alternative index constants change up to 43% of labels (k as low as 0.16), yet the MLP beats the GCN in all 40 evaluations (39 configurations).

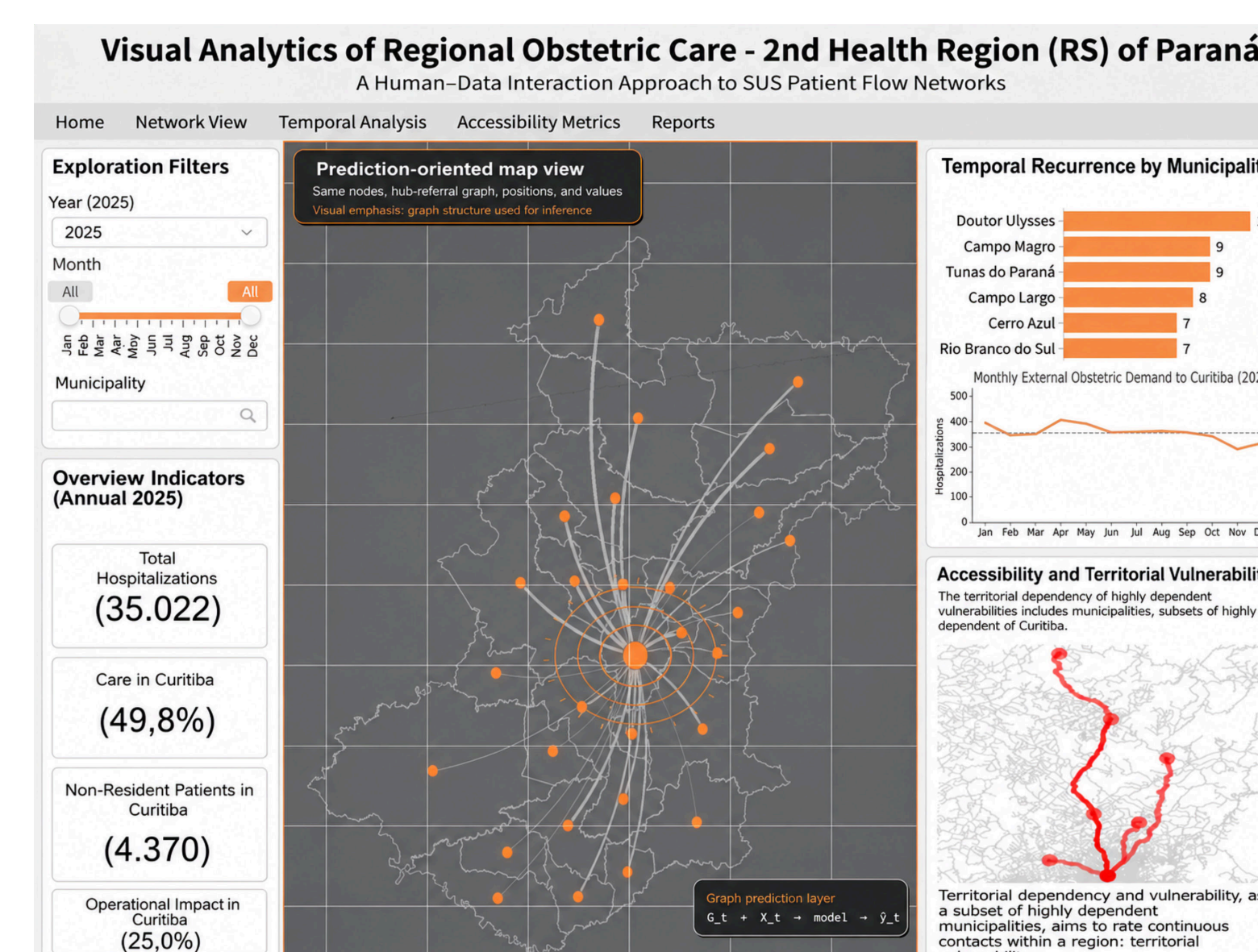


Figure 2. Visual analytics support for interpreting predicted vulnerability patterns in obstetric hospitalization flows.

Visual analytics prototype. A prediction-oriented map of the hub-referral graph G_t and node features X_t , with temporal recurrence, demand and accessibility views. The map shows the contemporaneous classification $\hat{y}_t$, not next-month forecasting; it is an illustrative aid, not a deployed decision-support system.

Conclusion

Each monthly graph is a star centred on a zero-filled Curitiba, so message passing attenuates the signals that define the label — a property of the graph construction, not of graph learning as such. The labelling is fragile, but the model ranking is not.

Future work: full origin-destination graphs, alternative supervision signals and clinical validation of vulnerability classes.

REFERENCES

- [1] Kipf and Welling (2017). Semi-supervised classification with graph convolutional networks.
- [2] Jin et al. (2024). A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection.
- [3] Zhao et al. (2020). T-GCN: A temporal graph convolutional network for traffic prediction.
- [4] DATASUS. Morbidade Hospitalar do SUS (SIH/SUS).

Acknowledgment: This study was financed in part by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES): Finance Code 001.