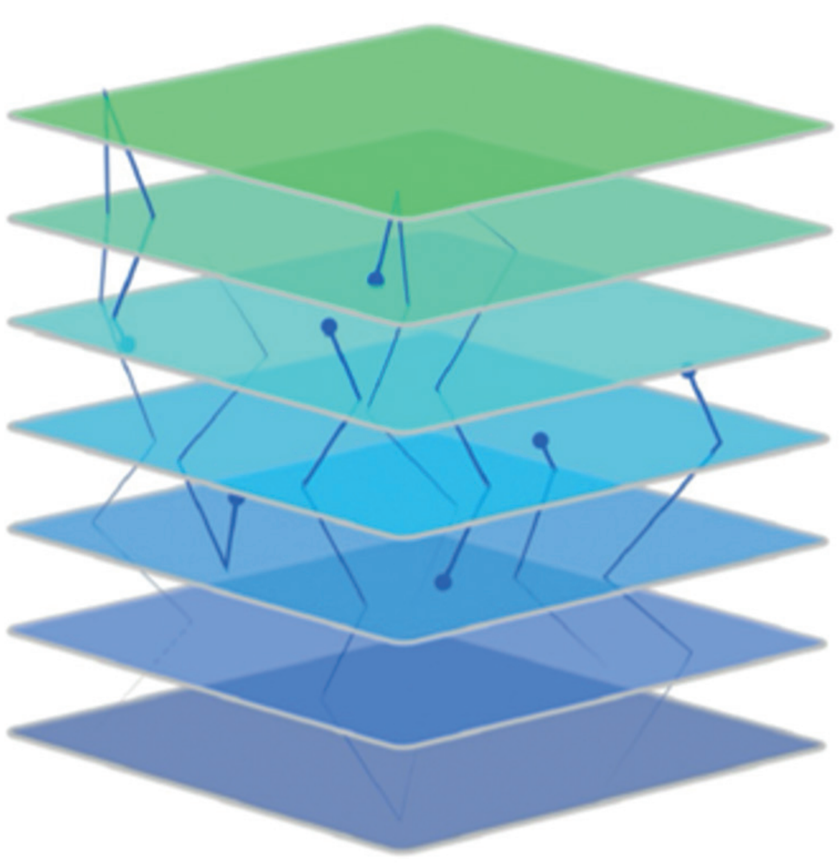
Unimarconi
LA PRIMA UNIVERSITÀ DIGITALE ITALIANA

Harnessing HPC and GNN for Explainable Biomarker Discovery from Wearable Data

Scientific Motivation

Wearable-derived biomarkers (e.g., from **VitalPatch**® from MediBiosense Ltd) show great promise for precision medicine, but clinical adoption is limited by fragmented data, low interpretability, and lack of temporal context. ST-GNNs offer potential to integrate multimodal signals, yet most models are not scalable or explainable. We identify the need for a transparent, HPC-ready platform to turn wearable data into dynamic, clinically meaningful insights.



Explainability & Visualization

- Feature attributions are computed using SHAP and LIME, including support for temporal attention weights
- An interactive dashboard visualizes the impact of physiological signals and contextual events over time on model predictions
- Language models (LLMs) generate temporally-aware summaries to support clinician interpretation and hypothesis generation

Next Steps: Towards Clinical Validation

- Launch prospective, multi-center trials using data from VitalPatch and other wearable devices
- Leverage existing biosignals (e.g., respiration, posture, temperature) to refine biomarker discovery
- Expand compatibility with additional sensor platforms for broader clinical applicability
- Develop LLM-assisted annotation pipelines with clinician-in-the-loop feedback
- Scale validation on EuroHPC and edge deployments with privacy-by-design architectures

What We Propose

- We present a **scalable and explainable AI platform** that transforms multimodal wearable data into **dynamic spatio-temporal graphs.** Our system:
- Learns **adaptive temporal attention weights** across past observations, rather than using fixed windows
- Aggregates multimodal signals across layers (physiological, behavioural, etc.) using Graph Neural Networks
- Produces explainable outputs via SHAP/LIME and **natural language summaries via LLMs**
- Is fully deployable on HPC and cloud infrastructures

The platform is currently being tested on real-world physiological and contextual data acquired via VitalPatch, a Class II medical-grade wearable device.

Mathematical Model: Attention-based Node Update Rule

$$h_i^{(t+1)} = \sigma\left(\sum_{j\in N_i} \alpha_{ij}^{(t)} W h_j^{(t)}\right), \quad \alpha_{ij}^{(t)} = \text{softmax}\left(\frac{(Qh_i^{(t)})^\top (Kh_j^{(t)})}{\sqrt{d}}\right)$$

Legend:

All variables are clearly defined and aligned with standard attention-based GNN literature:

- $h_i^{(t)}$: node embedding at time step t
- N_i: neighborhood of node i
- W: learnable weight matrix
- Q, K: projection matrices for query and key
- $\alpha_{ij}^{(t)}$: attention weight from node j to i
- σ: non-linear activation function
- d: dimensionality of the embedding space

An optional recurrence is also included via a GRU:

$$H^{(t+1)} = \text{GRU}(H^{(t)}, \text{GNNLayer}(H^{(t)}))$$

This extension captures longer-term temporal dependencies within the dynamic graph framework.

Scalability on HPC

- The architecture supports multi-graph batching, enabling parallel training of temporal attention-based GNNs across time windows and cohorts
- Optimized for GPU/TPU acceleration, with sparse matrix ops and graph sampling techniques (e.g., neighbor sampling, subgraph training)
- Deployable on exascale systems (e.g., EuroHPC, Frontier) via PyTorch-Geometric or DGL backends
- Compatible with federated learning and edge-cloud frameworks, ensuring privacy preservation and regulatory compliance in real-world clinical environments

Contact
Silvano Coletti, Francesca Fallucchi, Department of Engineering,
Università degli Studi Guglielmo Marconi, Rome, Italy
s.coletti@studenti.unimarconi.it

References and technical documentation available upon request