Geometric Transformers for Physiological Time Series Prediction: A Phase-Space Embedding Approach

Kevin R. Haylett
Manchester,UK

June 2026
Pensée

Abstract

This paper presents a novel framework for predicting physiological events, such as heart attacks, using a transformer-based architecture inspired by nonlinear dynamical systems theory. Building on the geometric attention mechanism previously proposed for language models, we reinterpret physiological signals (e.g., ECG, heart rate, respiration) as time series suitable for phase-space embedding via Takens' theorem. We introduce a model that replaces standard dot-product attention with geometric delay embeddings and distance-based weighting, enabling interpretable trajectory modeling in a high-dimensional manifold. We demonstrate the applicability of this architecture for anticipating critical state transitions in medical signals and outline a training pipeline suitable for clinical use.

1 Introduction

Modern physiological monitoring generates rich time series data: ECGs, heart rate, SpO2, blood pressure, and more. Accurately forecasting transitions such as arrhythmia, ischemia, or cardiac arrest remains a central challenge in healthcare. Existing methods—ranging from statistical models to deep neural networks—struggle with interpretability and often fail to capture the nonlinear dynamics inherent in biological systems.

This paper proposes a reimagined transformer architecture grounded in the principles of nonlinear dynamical systems. We treat physiological signals as dynamical trajectories, using Takens' embedding theorem to project these signals into a higher-dimensional phase space. This phase space trajectory is then modeled using a geometric attention mechanism, which replaces the traditional dot-product attention in transformers.

Our approach generalizes prior work on geometric attention in language models by applying it to real-valued, multivariate medical signals. The model is designed to detect early signs of system bifurcation or failure, offering an interpretable and extensible framework for high-risk time series prediction.

2 Background

2.1 Takens' Embedding and Nonlinear Time Series

Takens' theorem shows that the dynamics of a smooth deterministic system can be reconstructed from time-delayed versions of a single observed variable. For a time series $x(t)$, one constructs vectors of the form:

$$v_i = [x(t_i), x(t_i - \tau), x(t_i - 2\tau), \dots, x(t_i - (m-1)\tau)] \in \mathbb{R}^m \tag{1}$$

This transformation enables recovery of the system's attractor geometry.

In physiological contexts, signals such as ECG or HRV exhibit nonlinear behavior that is poorly captured by linear models. Delay embeddings allow for reconstruction of the system's state space, revealing hidden structure in the signal dynamics.

2.2 Transformers and Attention

Transformers have revolutionized sequence modeling, but their dot-product attention is agnostic to system dynamics. It computes similarity based on learned projections, without explicit representation of trajectory or curvature. Our proposal replaces this with geometric attention that reflects the actual distances between embedded states in phase space.

3 Geometric Transformer for Physiological Signals

3.1 Input Preprocessing

- **Signal normalization**: Raw signals (e.g., ECG in mV) are normalized per patient/session.
- **Multivariate input**: Inputs $x(t) \in \mathbb{R}^d$ can include multiple signals (e.g., HR + SpO2 + BP).
- **Embedding layer**: Identity or linear projection to $\mathbb{R}^k$. No tokenization required.

3.2 Delay Embedding Layer

We construct the phase space trajectory via delay embeddings:

$$v_i = [x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau}] \in \mathbb{R}^{m \cdot k} \tag{2}$$

where τ is a delay (often 1 or 2), and m is the embedding dimension.

3.3 Geometric Attention Mechanism

Given delay embeddings v_i, we compute pairwise distances:

$$w_{ij} = \exp\left(-\frac{\|v_i - v_j\|^2}{\sigma^2}\right) \tag{3}$$

$$z_i = \sum_j w_{ij} v_j \tag{4}$$

This yields a smooth, data-adaptive manifold trajectory $z_i \in \mathbb{R}^{m \cdot k}$, projected onto a hypersphere to ensure curvature stability.

3.4 Transformer Layer and Decoder

Each layer consists of:

- GeometricAttention
- Feedforward network (nonlinear projection)

- Layer normalization
- Optional smoothness loss:

$$L_{\text{smooth}} = \sum_i \|z_i - z_{i-1}\|^2 \tag{5}$$

The decoder predicts either:

- **Next signal state** (e.g., next ECG point)
- **Risk score** (e.g., heart attack probability)
- **Categorical label** (e.g., normal vs arrhythmic)

4 Training Strategy

4.1 Dataset

- MIT-BIH Arrhythmia Database (for ECG)
- MIMIC-III/IV (for ICU multivariate time series)

4.2 Loss Function

$$L = L_{\text{task}} + \lambda_1 L_{\text{smooth}} \tag{6}$$

Where L_{task} is MSE for regression or cross-entropy for classification.

4.3 Optimization

- AdamW optimizer
- Learning rate scheduler with warm-up
- Batch size: 32–64 depending on GPU

5 Benefits Over Traditional Models

Feature	**RNN/LSTM**	**CNN**	**Geometric Transformer**
Captures nonlinearity	✓	✓	✓✓
Interpretability	×	×	✓
Early warning	Limited	Static	✓
Visualizable attractors	No	No	Yes

6 Use Case: Predicting Cardiac Events

- **Input**: 10s sliding ECG window sampled at 250 Hz
- **Delay embedding**: $m = 3$, $\tau = 5$ (captures $\sim$120 ms cycles)
- **Output**: Risk probability for ventricular fibrillation within next 30s
- **Deployment**: Real-time monitoring with 1s refresh; Integrated into ICU dashboards or wearables

7 Reflections and Future Work

This paper reframes physiological signal prediction through the lens of geometric dynamics. The proposed transformer model, grounded in Takens' embedding and geometric attention, offers a new paradigm for forecasting critical events by modeling smooth, interpretable trajectories in phase space.

Future work will:

- Explore learnable delays τ and embedding orders m
- Extend to multimodal signals (EEG, respiration, blood pressure)
- Apply dynamical systems metrics (e.g., Lyapunov exponents)
- Benchmark against clinical-grade AI models

8 Conclusion

By unifying concepts from nonlinear dynamics and transformer architectures, we offer a principled, interpretable, and scalable framework for real-time prediction of health-critical physiological events. The geometric transformer turns medical signals into navigable trajectories on manifolds of meaning, opening new frontiers in both AI safety and clinical diagnostics.