

## Abstract

Psychiatric disorders (PDs), such as schizophrenia (SZ) and bipolar disorder (BD), affect millions worldwide, yet their diagnosis still relies heavily on subjective clinical assessments. We propose a Time-Series Language Model for PD detection using R-R interval sequences combined with a Contestable Language Model. Our human-centered PD detection system enables users and clinicians to gain more trust in AI-driven diagnostic systems in an explainable and contestable manner.

## From Explainable AI to Contestable AI

**Explainable AI (XAI)** helps users understand AI decisions. **Contestable AI** goes further, allowing users to challenge and potentially overturn those decisions. It emphasizes user agency, not just comprehension.

The requirements for explainability and contestability can be found in the Montr el Declaration of Responsible AI, GDPR, EU AI Act, and Canada's Directive on Automated Decision-Making.

Two core mechanisms for high contestability:

- Built-in safeguards:** automated checks or cross-validating systems to constrain AI behavior.
- Interactive controls:** let users contest, correct, or override decisions using explanations and intervention tools.

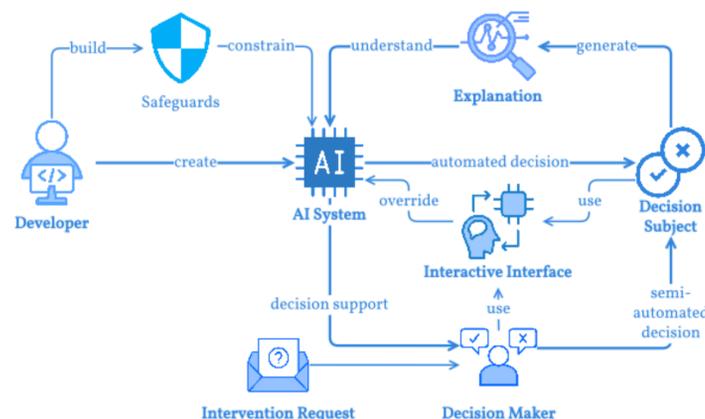


Fig. 1. Human-centered Explainable AI Systems with Contestability

## Human-centered Psychiatric Disorder Detection System

Our human-centered PD detection system consists of:

- Cardiac Monitoring Interface (CMI):** Enables users to record cardiac signals (ECG, heart rate, and RR intervals) using Polar H10/H9 devices.
- Contestable Diagnosis Interface (CDI):** Allows clinicians to review and contest the AI's decisions before finalizing the diagnosis. We utilize *Llama-3.3-70B-Instruct* as the core engine for Contestable AI.

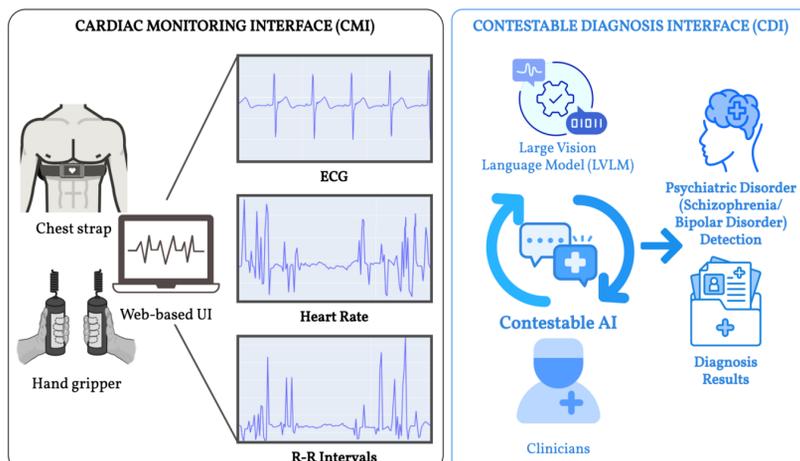


Fig. 2. Human-centered Psychiatric Disorder Detection System with Contestable AI

## Chronos-based Time Series Language Model

We proposed a Chronos-based [4] time series language model called PD-T5 (tiny, mini, small, base variants) for PD detection. This model classifies time series data using a language model architecture (i.e., T5):

- The time series is first transformed into a sequence of tokens through scaling and quantization. A language model is then trained on these tokens using cross-entropy loss.
- After training, probabilistic forecasts are generated by sampling multiple future trajectories based on historical context.
- A classifier head subsequently performs the PD classification.

While this is a novel and promising approach, further investigation is needed to determine whether it is worth pursuing, especially compared to our previous model's parameter efficiency [1].

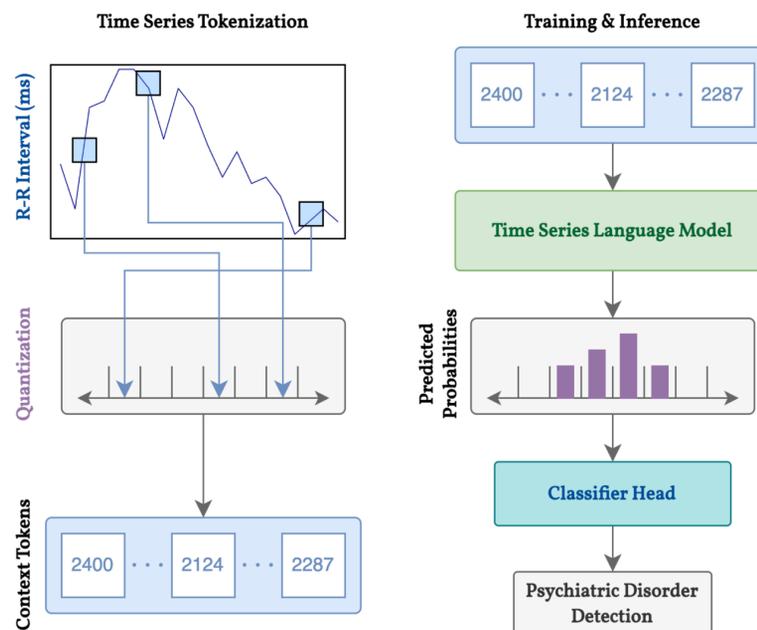


Fig. 3. Chronos-based Time Series Language Model for PD Detection

## Preliminary Results

	#Params	Acc	Pre	Rec	AUC
Misgar et al. [2]	7M	<u>0.833</u>	0.884	0.767	0.906
Buza et al. [3]	-	<u>0.833</u>	-	-	<u>0.910</u>
Nguyen et al. [1] (Ours)	2M	<b>0.900</b>	<b>0.962</b>	<u>0.833</u>	<b>0.933</b>
PD-T5-tiny (Ours)	8M	0.581	0.533	0.506	0.526
PD-T5-mini (Ours)	20M	0.65	0.627	0.700	0.699
PD-T5-small (Ours)	46M	<u>0.833</u>	0.857	0.800	0.890
PD-T5-base (Ours)	200M	<b>0.900</b>	<u>0.927</u>	<b>0.867</b>	<b>0.933</b>

Table 1. The leave-one-out cross-validation performance results are presented with the highest values indicated in **bold**, and the second-highest values underlined

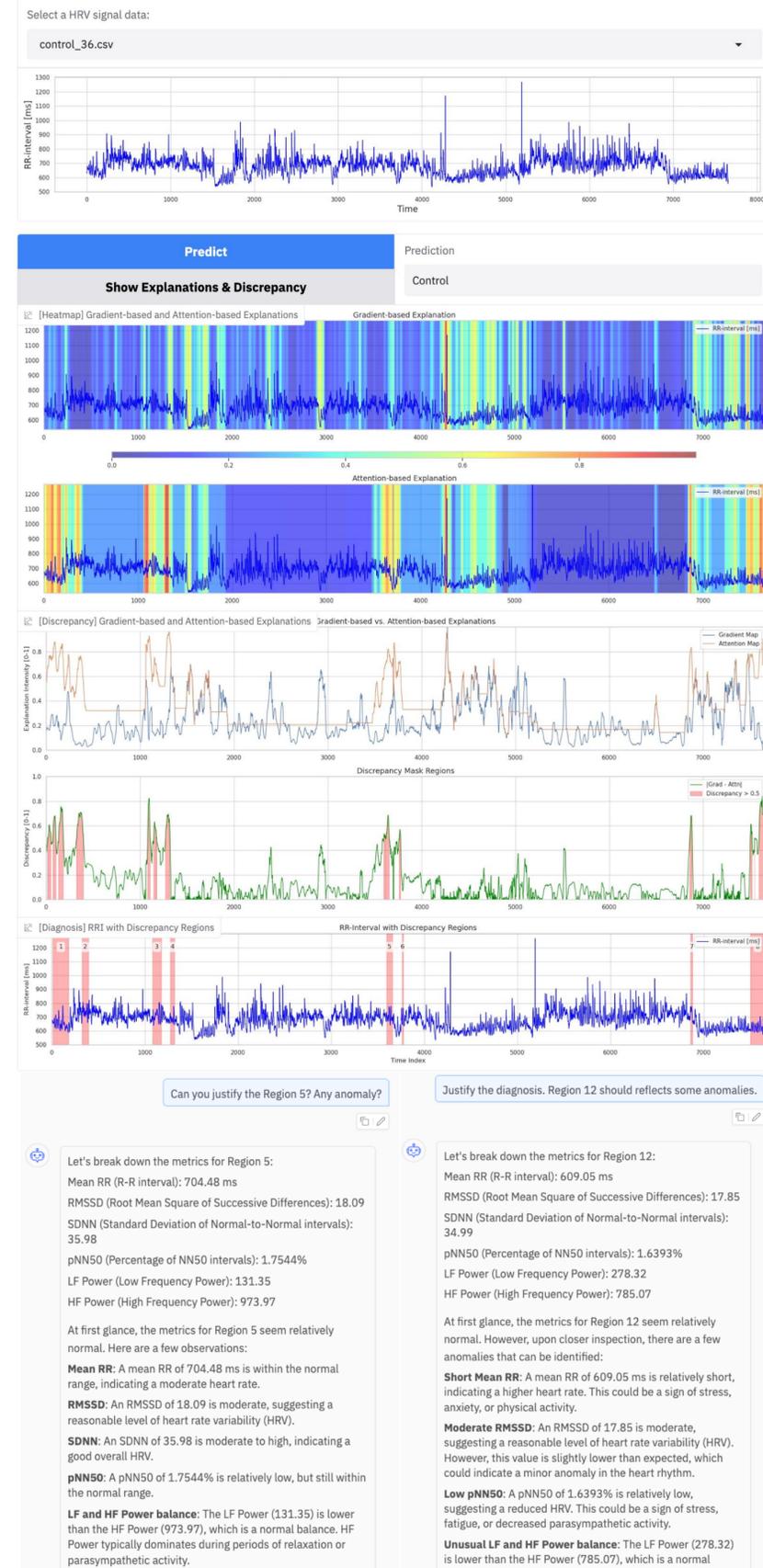


Fig. 4. Contestable Diagnosis Interface (CDI) including PD detection result, Self-Adversarial Explanation (SAE), and LVLM as the Contestable AI

## References

- Hung Nguyen, Alireza Rahimi, Veronica Whitford, H el ene Fournier, Irina Kondratova, Ren e Richard, and Hung Cao. 2025. Human-centered Explainable Psychiatric Disorder Diagnosis System using Wearable ECG Monitors. In *The 29th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2025)*.
- Muzafar Mehraj Misgar and MPS Bhatia. 2024. Unveiling psychotic disorder patterns: A deep learning model analysing motor activity time-series data with explainable AI. *Biomedical Signal Processing and Control* 91 (2024), 106000.
- Krisztian Buza et al., 2023. A Simple and Effective Classifier for the Detection of Psychotic Disorders based on Heart Rate Variability Time Series. In *Workshop on Bioinformatics and Computational Biology WBCB 2023*.
- Ansari, A.F et al., 2024. Chronos: Learning the language of time series.