Predicting Abnormal Heart Rhythms Using ECG Signals

Nicholas Ali
University of Toronto
Toronto, Canada

Stephanie Mann
University of Toronto
Toronto, Canada

Lynn Tao
University of Toronto
Toronto, Canada

Chloe Yu
University of Toronto
Toronto, Canada

Abstract

This project develops a computationally-efficient model to identify abnormal cardiac rhythms from electrocardiogram (ECG) data by fitting individual heartbeats to a two-variable dyanmical system known as the FitzHugh–Nagumo (FHN) ordinary differential equation [1]. Optimizing the parameters using a least-squares approach and using those as features for a k-Nearest Neighbours (KNN) classifier, this approach leverages parameters of a low-dimensional dynamical system to capture the excitability and recovery dynamics of the cardiac action potentials. By representing a set of 5586 beats from the MIT-BIH Arrhythmia Database after filtering for quality, we found that these compact parameters can capture distinctions between normal and abnormal heart rhythms. Our results show that while the model is able to predict abnormal heart rhythms with 84% accuracy, its performance decreases to 50% when distinguishing between subtypes of abnormal beats, revealing that the heterogeneity of certain arrhythmia subtypes may require higher-dimensional feature representations for detailed detections. These results establish low-complexity dynamical systems, such as FHN, as a interpretable method of capturing the fundamental "excitable" characteristics of heart rhythms under resource-constrained applications, while demonstrating its limitations in more complex cases.

1 Introduction

Accurate classification of cardiac arrhythmias is critical both for clinical diagnosis and for life-threatening emergencies requiring automated external defibrillators (AEDs). The use of AEDs during cardiopulmonary resuscitation (CPR) dramatically improves survival rates, from 10% with CPR alone to 70% when an AED is used [2]. However, current AEDs can identify only four types of arrhythmias: ventricular tachycardia, ventricular fibrillation, asystole, and pulseless electrical activity, while up to 12 additional clinically significant arrhythmias remain undetectable by AEDs [3]. Although recent deep learning approaches have demonstrated success in comprehensive arrhythmia classification from ECG data [4, 5], these methods require substantial computational resources that may be impractical for many real-world use cases. This project explores whether a low-complexity algorithm for fitting heartbeats followed by classification algorithms can achieve meaningful arrhythmia detection with significantly reduced computational demands.

ECGs measure the surface-level electrical potential between two electrodes placed on different body parts [7]. Modified Limb Lead II, or MLII, is the method that produces the strongest ECG signal - this is the method studied in this report. Each heartbeat picked up by an ECG is comprised of three waves that reflect the electrical depolarization and repolarization of the sinoatrial node: P, the QRS complex, and T [6]. Heart rhythms are typically categorized by the heart rate and the QRS width, as shown in Figure 1. Due to constraints in the electrocardiogram data available, our model is trained to categorize normal beats (N) and bundle branch block beats, specifically the left bundle branch block beat (L). This is a blockage in the heart's electrical pathway to the ventricles. As a result, one side of the beat is slower, causing them to beat out of sync. It can be caused by underlying heart diseases, sometimes requiring a pacemaker to resynchronize the heartbeat [8].

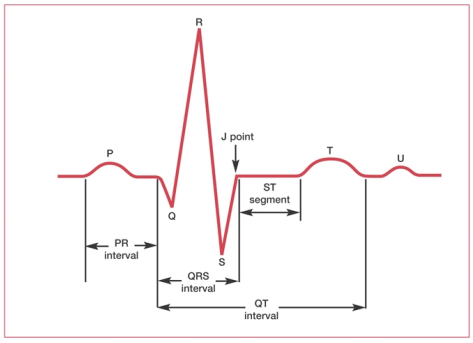


Figure 1: Section of rhythm strip produced by an electrocardiogram, providing a view of electrical activity over one heartbeat. There are 3 main components: the P-, QRS-, and T- waves. [6].

2 Mathematical Model

The FitzHugh-Nagumo (FHN) model is a simplified system of coupled, nonlinear ordinary differential equations introduced in 1961 in the field of neuroscience to model excitable neuron behaviour [1]. It is a simplification of the Hodgkin-Huxley equations [9], which models action potentials in neurons. It retains the combination of a fast activation variable with a slow recovery variable, making it useful for modeling excitability in biological applications, such as heart rhythms.

The equations are given by:

$$\begin{aligned} \frac{dv}{dt} &= v - \frac{v^3}{3} - w + I, \\ \frac{dw}{dt} &= \frac{v + a - bw}{\tau} \end{aligned} \tag{1}$$

where $v(t)$ represents the fast membrane potential that models the ECG waveform. $w(t)$ is a latent recovery variable that controls the timing of repolarization, which is the electrical recovery phase of a heart muscle cell after excitation. While $w(t)$ is not directly observed, it plays a critical role in shaping the dynamics of $v(t)$, both being dependent variables evolving over time.

2.1 Phase Plane and Stability Analysis

The qualitative behaviour of this system is determined by its two nullclines. The intersection of these nullclines represent the fixed

point (v^*, w^*) of the cardiac cycle, representing the resting potential before any stimulus.

The v-nullcline is given by

$$\frac{dv}{dt} = 0 \quad \Rightarrow \quad w^* = v^* - \frac{(v^*)^3}{3} + I, \tag{2}$$

while the w-nullcline is given by

$$\frac{dw}{dt} = 0 \quad \Rightarrow \quad w^* = \frac{v^* + a}{b}. \tag{3}$$

The stability of this fixed point is determined by the Jacobian matrix evaluated at (v^*, w^*):

$$J(v^*, w^*) = \begin{pmatrix} 1 - (v^*)^2 & -1 \\ \frac{1}{\tau} & -\frac{b}{\tau} \end{pmatrix} \tag{4}$$

The eigenvalues λ must satisfy the characteristic equation:

$$\det(J - \lambda I) = \lambda^2 - \mathrm{tr}(J) + \det(J) = 0, \tag{5}$$

where the trace is $\mathrm{tr}(J) = [1 - (v^*)^2] - \frac{b}{\tau}$ and the determinant is $\det(J) = (1 - (v^*)^2)(-\frac{b}{\tau}) + \frac{1}{\tau}$. In order for the fixed point to be stable, as it is physically represents the resting potential of the heart, both eigenvalues must have negative real parts. This occurs when $\mathrm{tr}(J) = \lambda_1 + \lambda_2 < 0$ and $\det(J) = \lambda_1 \lambda_2 > 0$, constraining the fixed points to a physiologically-possible region in the plane.

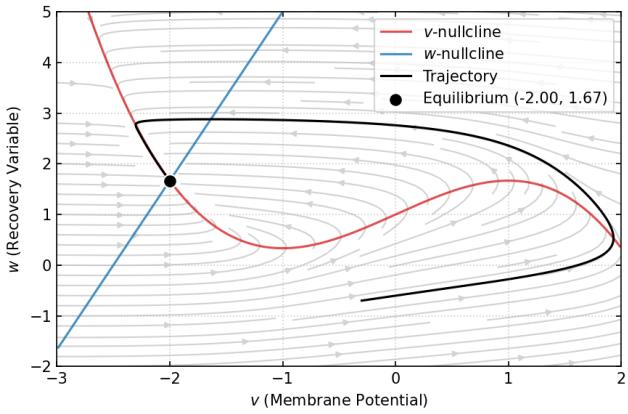


Figure 2: Phase portrait of the FHN equation using the parameters obtained from one heartbeat. The black trajectory depicts a single cardiac cycle after excitation as it returns to equilibrium.

Depolarization to the QRS region occurs when a small stimulus triggers a large change due to the cubic term $v - \frac{v^3}{3}$. For small perturbations lower than the threshold, the $-w$ term acts as a drag force that returns v to rest. However, for large perturbations, $v - \frac{v^3}{3}$ dominates, causing v to accelerate rapidly to depolarization. Following this, the slow variable w increases and eventually overpowers the cubic term to drive $v(t)$ back towards the baseline, and initiating the repolarization phase. As shown by the sample trajectory in Figure 2, the potential $v(t)$ is disturbed (depolarization), forming a half-loop following the vector field, before following the v-nullcline in back to equilibrium as w slowly decays (repolarization). These stability conditions ensure that the heart returns to rest after each beat and that the system is excitable. A sample calculation under these conditions is provided in Appendix A.1.

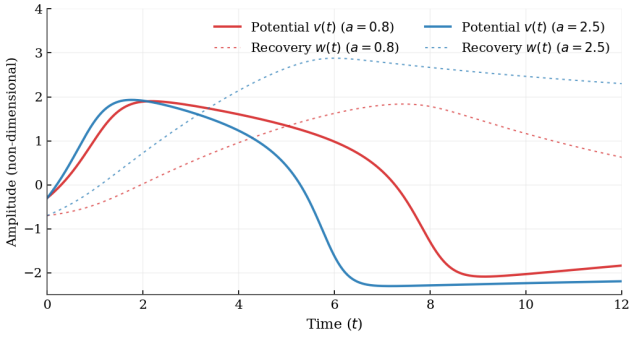


Figure 3: Comparison of membrane potential $v(t)$ (solid) and recovery variable $w(t)$ (dotted) for varying excitability thresholds. Higher values of a (blue) exhibit quicker repolarization.

2.2 Intepretation of FHN Parameters

The parameters of the model are $\{a, b, \tau, I, v_0, w_0\}$, which include the system constants (a, b, τ, I) and the initial conditions (v_0, w_0):

- a, b **(Excitability Threshold):** These two parameters control the position of the linear w-nullcline ($w = \frac{v+a}{b}$). a controls the vertical shift while b determines the slope. Together, they determine how quickly the recovery variable w responds to changes in the potential. Smaller values of b and larger values of a create a steeper w-nullcline, meaning it is less easily excitable and returns to equilibrium faster. As shown in Figure 4, $a = 2.5$ leads to quicker repolarization compared to $a = 0.8$ under the same starting conditions.
- τ **(Recovery Time Constant):** Unlike a and b which affect the position of the nullclines, τ affects the time scale separation between fast and slow dynamics within v. When τ is large, w evolves much more slowly than v, creating a rapid QRS complex upstroke followed by a slower recovery T. The transition to the relaxation section T is significantly influenced by the degree of this time scale separation [10], thus influencing the QT duration.
- I **(Applied Current):** This represents an external driving current, shifting the v-nullcline vertically. For individual beat modelling, it has a small influence on the model, as it mostly affects the timing of the beats rather than the shape of individual beats.

These FHN parameters implicitly encode the stability and excitability properties of the FHN system due to their influence on the Jacobian matrix and its eigenvalues through the trace and determinant. By using these parameters as features for classification, we test the hypothesis that normal and abnormal heart rhythms differ in their fundamental dynamical properties, including stability properties and trajectories. Normal rhythms may exhibit consistent stability with eigenvalues clustered in similar positions of the complex plane, while abnormal beats may exhibit different eigenvalue distributions based on the subtype. This provides a theoretical foundation as to why FHN parameters may be able to distinguish between the key features for arrhythmia classification.

3 Methodology

3.1 Data Preprocessing and Wave Extraction

The ECG data used in this study was obtained from the MIT-BIH Arrhythmia Database [11], a widely used benchmark dataset for cardiac signal analysis. Each record in the database consists of a raw ECG waveform file and a corresponding annotation file. The waveform files contain voltage measurements over time for each heartbeat, while the annotation files provide beat-level labels, indicating whether a beat is normal or belongs to a specific type of arrhythmia. As shown in Figure 4, the dataset is comprised of majority Normal (N) beats, guiding the mitigation of class imbalances in later processes.

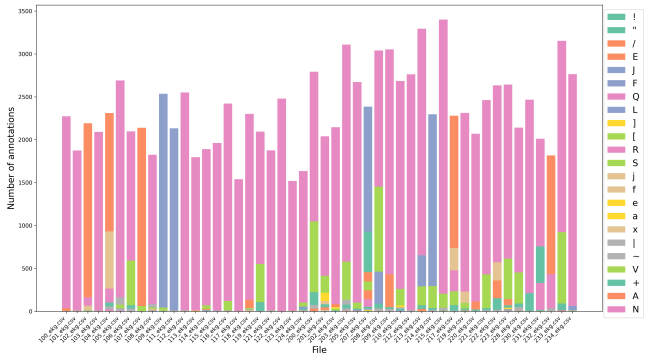


Figure 4: Number of annotations for each category in each raw ECG file.

For signal processing and waveform extraction, we used the NeuroKit2 library, which provides functions for detecting R-peaks, P-waves, T-waves, and other ECG features [12]. NeuroKit2 was chosen for its accuracy and convenience in handling large datasets. We focused specifically on the Modified Limb Lead II, or MLII, as it is commonly used in arrhythmia detection and contains clear representations of the P-QRS-T complexes. Individual beats were extracted from each recording based on detected P-onsets and T-offsets, ensuring that each beat contained a complete cardiac cycle.

To align the original dataset annotations with the detected beats, each beat was assigned the label of the R-peak falling within its P-wave onset to T-wave offset interval. Formally, let $[P_i, T_i]$ denote the interval of the i-th detected beat, and R_j denote the j-th annotated R-peak, with annotation S_j. The beat i is assigned the annotation S_j if $R_j \in [P_i, T_i]$. This approach ensures correct labeling even when peak detection differs between the NeuroKit2 and original dataset slightly for abnormal beats. Beats with missing channels, corrupted signals, or alignment errors were skipped and excluded from further analysis. After alignment, each beat in the dataset contains a *Symbol* column corresponding to its annotation, which is then concatenated into a single dataset used in subsequent FHN fitting and classification steps.

3.2 FitzHugh-Nagumo Model Fitting

To capture the dynamics of individual ECG beats, we fit the FitzHugh-Nagumo (FHN) model to each segmented beat. The model consists of a coupled system of ordinary differential equations, as defined in Section 2.

For a given parameter vector $\theta = \{a, b, \tau, I, v_0, w_0\}$, the FHN system is numerically integrated over the duration of the beat using LSODA (Livermore Solver for Ordinary Differential Equations) via the `solve_ivp` function from the `scipy` library [13]. LSODA is an adaptive solver that automatically detects stiffness, making it well suited for systems exhibiting both fast and slow dynamics.

We simulate the system using the initial conditions v_0 and w_0 over the time interval of the beat, producing a predicted waveform $v_{\text{sim}}(t)$. To fit the model, we use a *least-squares* method to minimize the discrepancy between the simulated waveform and the actual ECG segment $v_{\text{ECG}}(t)$. The six parameters $\theta = \{a, b, \tau, I, v_0, w_0\}$ are optimized to minimize the loss function $L(\theta)$, which is defined as the *mean-squared error*:

$$L(\theta) = \sum_{i=1}^{N} \left(v_{\text{sim}}(t_i, \theta) - v_{\text{ECG}}(t_i)\right)^2 \tag{6}$$

Optimization is performed using the Nelder-Mead method, which is a numerical method that iteratively explores the parameter space to locate a minimum of $L(\theta)$ [14]. For each candidate parameter set proposed by the optimizer, the FHN system is re-integrated using LSODA to evaluate the corresponding loss value. Minimizing $L(\theta)$ ensures that the simulated waveform closely matches the observed ECG, yielding a compact parametric representation of each beat suitable for downstream classification.

On top of the loss, another metric used to evaluate the fit is the *coefficient of determination*, R^2. This is a statistical measure of how well a variable is replicated by a model, based on the proportion of the total variance of the model. It is defined with the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (v_{\text{ECG}}(t_i) - v_{\text{sim}}(t_i))^2}{\sum_{i=1}^{N} (v_{\text{ECG}}(t_i) - \bar{v}_{\text{ECG}})^2} \tag{7}$$

where $\bar{v}_{\text{ECG}}$ is the mean of the observed ECG signal. Values of R^2 close to 1 indicate that the FHN model captures most of the variance in the observed data.

3.3 Filtering, Grouping, and Class Balancing

To assess the separability of the extracted FHN parameters, we applied a *t-distributed Stochastic Neighbor Embedding (t-SNE)* [15] to the full set of beats. t-SNE is a nonlinear dimensionality reduction technique that projects high-dimensional data into a 2-D space while preserving local similarities. This allows us to visualize clusters of beats and see whether different beats in the same class are close together, and whether beats from different classes have features distinct enough to be separated. This analysis guided subsequent processing steps in this section, namely filtering poorly fitted beats, selecting which classes to predict, and balancing class sizes to improve classification performance.

To remove poorly fitting beats caused by incorrect initial guesses or non-converging solutions, we filtered the fitted results using thresholds on the loss L and R^2. Various threshold combinations were tested, and the values yielding the highest overall classification

accuracy (discussed in Section 4.1) were selected. Experimentally, the thresholds were found to be approximately $L < 0.1$ and $R^2 > 0.6$.

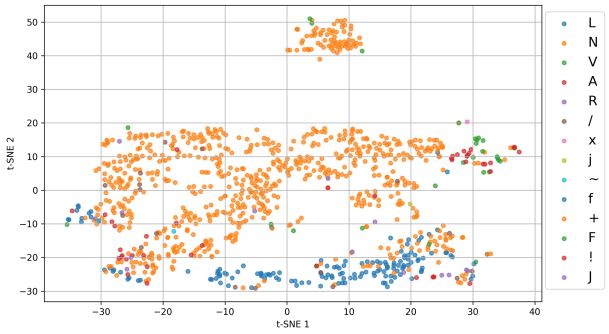


Figure 5: t-SNE visualization of FHN parameters for a subsample of filtered ECG beats.

There are significantly more categories compared to features, many categories with small sample sizes, and an imbalance in the amount of data points in each category; as such, similar categories were grouped together for classification. Several grouping strategies were explored, including:

- **Binary:** Normal (N) vs. Abnormal (all other categories)
- **N/L/Other:** Normal (N) vs. Left bundle branch block (L) vs. Other

As shown in Figure 5, the N-beats dominate the graph at the top middle and center with slight L-beat clustering at the bottom middle. These categories contain the largest number of data points as shown in Figure 4 and thus were chosen as one set of groupings. The binary classification was chosen for its simplicity and applicability. Differentiating between a normal and abnormal heartbeat is highly relevant to real-world applications. This approach also considers all of the data points using a simple binary split. The N-beats and L-beats have noticeable clustering and we hypothesize that our model may be able to classify these beat types accurately, although there may be inaccuracies with classifying the remaining beat types since they appear less clearly defined. Other classifications were considered, but we did not move forward with them, mainly due to a lack of data for alternate beat types.

To balance the training set and prevent bias toward more frequent classes, two methods were attempted: (1) **oversampling** smaller classes, also known as "bootstrapping", by duplicating entries, and (2) **undersampling** larger classes by randomly eliminating some entries to reduce size. This ensured that all categories under comparison (e.g. Normal vs. Abnormal) had equal representation, reducing the risk of overfitting and class imbalance that can skew the prediction results.

3.4 K-Nearest Neighbour Classification

For each beat in the balanced dataset, we extracted the FHN parameters $\{a, b, \tau, I, v_0, w_0\}$ as feature vectors. In addition, we included some standard features of the ECG waveform, specifically the QRS width and P–T segment width, which capture temporal characteristics of the beats that are relevant for distinguishing different arrhythmia types. We trained a K-nearest neighbors (KNN) classifier using these features to assign each beat to its corresponding class. For each beat, the k nearest neighbours were determined using the Euclidean distance of the feature vectors in a multidimensional space:

$$d(\mathbf{x}, \mathbf{x}_i) = \sqrt{\sum_{j=1}^{d} (x_j - x_{i,j})^2}, \tag{8}$$

where $\mathbf{x}$ is the feature vector of the beat to classify, $\mathbf{x}_i$ are training vectors, and d is the number of features. The predicted class was assigned based on the majority label among the $k = 5$ nearest neighbours, as this k-value had the highest accuracy. Various combinations of the FHN and width features were tested to assess the significance of each variable by comparing the accuracy of subsequent results.

To account for multiple beats coming from a single patient, we employed a *Leave-One-Group-Out* (LOGO) cross-validation strategy [16], where each group corresponds to all beats from a single recording. This ensures that the classifier is tested on entirely unseen recordings, providing a realistic estimate of generalized performance. This reduces the risk of overfitting and bias from individual recordings or class imbalances, teaching the model to generalize to all patients rather than "memorizing" patterns within specific individuals.

3.5 Performance Metrics

To evaluate the classification performance, we calculate the precision, recall, and F1-score.

Precision measures the proportion of predicted positives that are true positives:

$$Precision = \frac{\Sigma TP}{\Sigma(TP + FP)} \tag{9}$$

Recall measures the proportion of true positives that are predicted as positive:

$$Recall = \frac{\Sigma TP}{\Sigma(TP + FN)} \tag{10}$$

The F1-score is the harmonic mean of these two metrics, which can serve as a measure of the overall model performance:

$$F1 = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \tag{11}$$

4 Results

4.1 Fitting Quality

Figure 6 shows three representative examples of fitted beats from recording `100_ekg.csv`. The grey curve represents the original ECG segment, the blue points show the subsampled data points used for fitting, and the red dotted curve displays the FHN model fit after optimization. They have a loss of <0.1 and R^2 of >0.9. These

examples illustrate the model's ability to capture the overall morphology of both normal and abnormal beats, reproducing the rapid QRS upstroke, ST segment plateau, and gradual T wave recovery.

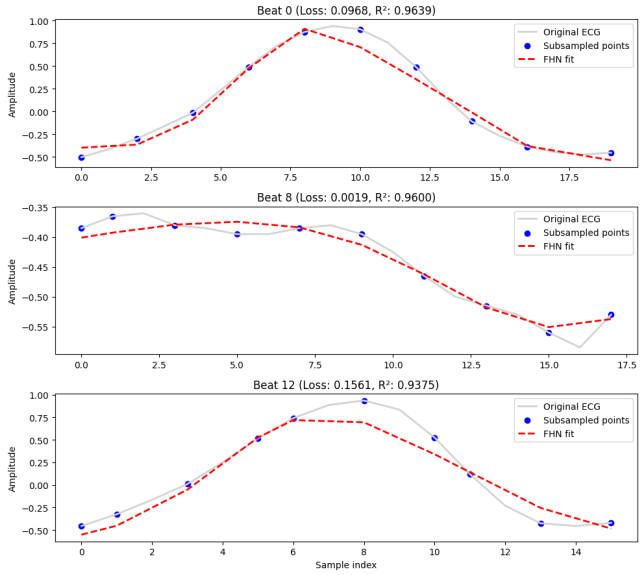


Figure 6: Original ECG segments, subsampled points, and FHN model fit for three sample beats from `100_ekg.csv`

Across the entire MIT-BIH dataset, we attempted to fit 63982 heartbeats. Of these, 719 beats failed to converge or produced unrealistic parameter values. Beats that failed to converge often corresponded to parameter values that violated the stability conditions discussed in Section 2.1. When the optimization proposed parameters producing $tr(J) > 0$ or $\det(J) < 0$, the resulting system exhibited unstable oscillations rather than the excitable pulse observed in real ECG data. To ensure data quality, we applied filtering thresholds of $L < 0.1$ (mean squared error) and $R^2 > 0.8$ (coefficient of determination), which removed poorly fitting beats. After this filtering step, 5586 beats remained for classification.

The quality metrics for the filtered dataset show strong agreement between the FHN model and observed ECG waveforms. The mean loss across all fitted beats is $L_{\text{mean}} = 0.0568$, with median absolute error MAE = 0.0582 and mean $R^2 = 0.97$. These metrics indicate that the FHN model provides a compact yet accurate parametric representation of individual heartbeat dynamics suitable for classification models.

4.2 Feature Separability Analysis

Table 1 presents the median and interquartile range (IQR) for all FHN parameters and width features, grouped by beat classification. This reveals both distinguishing features and overlapping characteristics across beat types. Parameters such as a (excitability threshold) and τ (recovery time constant) show moderate separation between normal and abnormal beats, with beats labelled N having median $a \approx 0.45$ and beats labelled L having $a \approx 0.57$. The recovery time constant τ displays similar trends, with $\tau \approx 1.29$ for N and $\tau \approx 1.00$ for L. This is consistent with the hypothesis that abnormal rhythms affect the excitation-recovery dynamics represented by these parameters.

However, substantial overlap exists in several parameters where there is negligible differences between classes. For example, $I \approx 0.45$ for N, while $I \approx 0.48$ for L, both with large interquartile ranges spanning approximately 0.2. This overlap suggests that I, which primarily affects baseline offset rather than beat morphology, provides limited discriminative power. Similarly, the initial condition w_0 does not have much distinction across beat types, as it represents the initial condition of the latent recovery variable rather than a beat-specific characteristic.

Table 1: Median and Interquartile Range (IQR) for FHN and Width Features

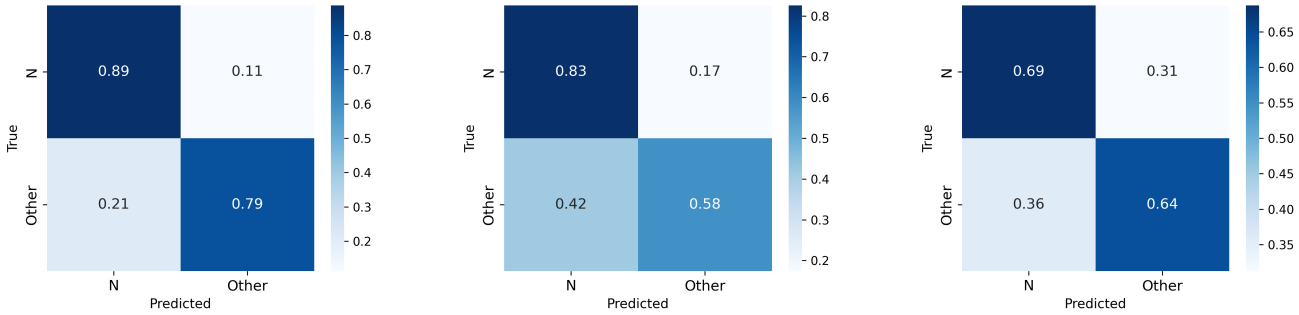
Symbol	A	L	N
a	0.57 (0.37)	0.52 (0.26)	0.45 (0.13)
b	0.98 (0.19)	1.04 (0.04)	0.99 (0.07)
tau	1.12 (0.57)	1.00 (0.15)	1.29 (0.32)
I	0.53 (0.33)	0.48 (0.22)	0.45 (0.12)
QRS Width	34 (20)	23 (2)	18 (2)
PT Width	143 (72)	206 (15)	146 (49)
v0	-0.58 (0.58)	-0.26 (0.19)	-0.48 (0.18)
w0	0.001 (0.003)	0.001 (0.002)	0.001 (0.002)

To visualize this separability in a 2-D space, Figure 5 shows a 2-D t-SNE plot of the FHN features for all beats. Normal (N) beats form a relatively compact cluster, as labelled in orange. This indicates that normal cardiac rhythms share consistent FHN parameter profiles, with similar excitability thresholds, recovery dynamics, and width characteristics. In contrast, abnormal beats are more dispersed throughout the feature space and partially overlap with both the normal cluster and each other. For example, the L category, labelled in blue, is more clustered than other abnormal beats, which show little discernible structure. This shows that an FHN fitting may not be generalizable for all types of abnormal heart rhythms.

4.3 Binary Classification

The filtered dataset of 5586 beats exhibits substantial class imbalance, with normal beats comprising a larger proportion than abnormal beats. By applying the undersampling method discussed in Section 3.3, this resulted in a balanced dataset of 2892 total beats with 50% normal (N) and 50% abnormal (Other). We also tested the oversampling method, which yielded similar classification performance. Given the comparable results, undersampling was chosen as the primary balancing method for simplicity and computational efficiency, as it reduces the total training set size while maintaining equal class representation.

A confusion matrix visualizes classification performance by showing the counts of true versus predicted labels. For our binary task, the rows represent the actual beat type (N or Other), while columns represent the predicted type. Diagonal entries (top-left and bottom-right) indicate correct predictions: normal beats correctly identified as normal (true negatives) and abnormal beats correctly identified as abnormal (true positives). Off-diagonal entries represent errors: normal beats misclassified as abnormal (false



(a) FHN and width features combined **(b) FHN-parameter-only features** **(c) Width-only features**

Figure 7: Comparison of binary classification performance under different feature sets.

positives, top-right) and abnormal beats misclassified as normal (false negatives, bottom-left).

Comparing the three confusion matrices in Figure 7, we observe that the combined feature set in (a) produces the strongest diagonal concentration, indicating the highest proportion of correct classifications. The FHN-only features in (b) show moderate performance for true positives of N, but there are increased misclassifications in both directions, showing weakened diagonal concentration. The width-only features in (c) exhibit the weakest diagonal concentration and the most diffuse pattern, suggesting substantial confusion between normal and abnormal beats.

For the binary classification, the model is able to distinguish between normal (N) and abnormal beats using the full feature set with an accuracy of 84%, as shown in Table 2. This performance exceeds that obtained using either the waveform width features alone (67%) or the FHN parameters alone (71%). The F1-score for the normal and abnormal classes are 0.84 and 0.83, respectively, meaning that the model maintains a strong balance between precision and recall.

The complementary nature of these feature groups explains the performance gain. Width-based features (QRS and PT widths) capture temporal characteristics that many arrhythmias, such as bundle branch blocks, follow. However, as seen by (c), width alone cannot capture differences in the underlying excitation-recovery dynamics, as beats with similar durations may have different trajectories of $v(t)$. By combining both feature types, the classifier gains sensitivity to both temporal (widths corresponding to duration) and excitation–recovery dynamics in heart rhythms. The resulting joint feature representation enhances the separability between normal and abnormal beats in feature space, leading to improved classification performance.

Table 2: Classification report for the binary classifier (N vs. Other).

Class	Precision	Recall	F1-score	Support
Normal	0.81	0.89	0.85	1430
Abnormal	0.88	0.79	0.83	1430
Accuracy		0.84		2860

4.4 Multi-Class Classification

Extending beyond binary classification, we evaluated the model's ability to distinguish between Normal (N) beats, left bundle branch block (L) beats, and all other types combined (Other). Figure 8 presents the confusion matrix for this three-class classification.

The overall accuracy for the three-class problem is 50%, a notable decrease from the 84% binary accuracy. Table 3 breaks down the performance by class, with highly asymmetric results: Normal (F1 = 0.01), L (F1 = 0.80), and Other (F1 = 0.50).

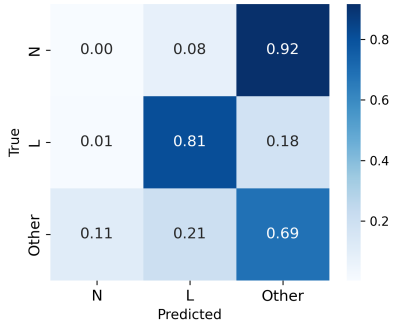


Figure 8: Classification of N vs. L vs. Other on a confusion matrix.

While the model is able to detect L with a moderate F1-score of 0.79, the severe degradation in N beats reveals that N and L beats may be more similar to each other than either is to the heterogeneous Other category in certain feature dimensions. N beats appear well-separated in the t-SNE embedding (Figure 5), however its ubiquity may prevent the model from accurately gauging its feature boundaries for classification, especially when more specific classifications are employed.

To investigate this, we visualized the t-SNE embeddings with approximated linear decision boundaries for both classification tasks in Figures 9 and 10. These visualizations reveal that in the binary classification setting, the decision boundary clearly separates the N (orange) from the Other (blue) category. The boundary passes through a region of relatively low data density and there are few

misclassified points, indicating good class separation that explains the strong binary classification performance (84% accuracy).

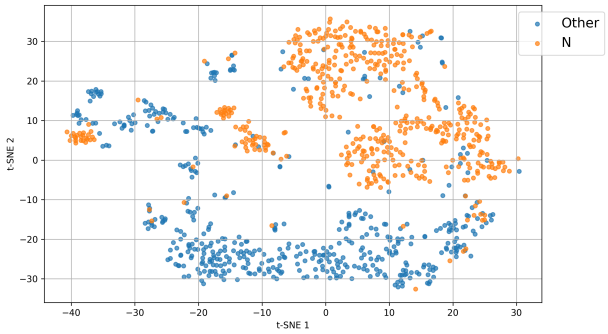


Figure 9: t-SNE embedding of binary classification data (N vs. Other)

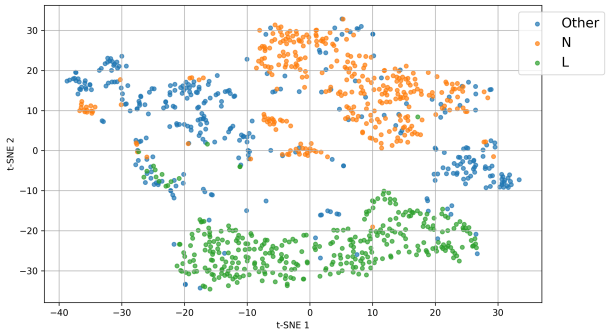


Figure 10: t-SNE embedding of three-class classification data (N vs. L vs. Other)

Figure 10 shows that extracting L beats from the original Other group causes the Other category (blue) to become significantly more diffuse and exhibits greater overlap with the N category (orange). In the binary classification, L beats, which form a compact, well-defined cluster, contributed to creating a clear abnormal class. Their distinctive FHN parameter profile help "anchor" the Other category and provided a strong contrast against N beats. When L is removed and placed into its own category, the remaining Other beats are predominantly scattered arrhythmias (R, B, V, etc.) that lack cohesive clustering properties. Without the anchor provided by L beats, the distinction between the Normal and Other categories becomes less defined.

Table 3: Classification report for multi-class classifier (N vs. L vs. Other).

Class	Precision	Recall	F1-score	Support
N	0.03	0.00	0.01	636
L	0.74	0.81	0.77	636
Other	0.39	0.69	0.49	636
Accuracy		0.50		1908

As seen in Figure 10, a higher proportion of the Other (blue) points are now mixed with the Normal (orange) points of the feature space. This spatial overlap was less apparent in the binary setting because the concentrated L cluster pulled the overall Other average away from Normal. By removing this category, the decision boundary must now separate Normal from a more diffuse Other category, leading to increased confusion between the two classes.

The strong performance of the binary classification shows that abnormal cardiac rhythms share distinct features in their excitation-recovery dynamics compared to normal heart rhythms. Specifically, parameters a and τ showed the clearest separation between normal and abnormal beats, as shown in Table 1. This aligns with the understanding that arrhythmias are often due to changed excitability thresholds or disrupted depolarization and repolarization patterns. For example, the left bundle branch block (L) delays the electrical signal to the heart, which is also evident in the distinct recovery time constant $\tau = 1.00$ compared to normal ($\tau = 1.29$) [8].

5 Limitations and Next Steps

The performance of the model is affected by the selected dataset, resampling bias, beat segmentation method, and the FHN model limitations.

Dataset: While the MIT-BIH Arrhythmia Database is a good benchmark dataset, only 25 of its 48 records are complete [11]. As well, there is a significant lack of abnormal beat types which prevented us from assessing specific types of abnormal beats beyond L-beats. The majority of entries represent "Normal" beats. Several abnormal heart conditions are only represented by a few patients, or in some cases, a single patient [11]. This results in a restricted data set, and impacts the diversity of the data the model can categorize.

Resampling Risks: Much of the data reflects normal heart rates, rendering the model biased. This explains the model's poor performance when categorizing N-beats compared to abnormal beats. To compensate for this limitation the model currently under-samples normal conditions by randomly removing some normal data and over-samples abnormal data through duplication. However, under-sampling can risk losing critical data, and oversampling can cause the model to become specialized and biased towards the limited available heart conditions.

Beat Segmentation Method: The model focuses on singular beats rather than a sequence of beats. While this is sufficient for arrhythmia types that solely depend on dynamics in single beats or widths, this would be insufficient for extending the model for applications in heart conditions that exhibit features over multiple beats [17].

FHN Model Limitations: As shown in Figure 5, the clusters are not well-formed for abnormal heart rates, indicating that data is not significantly differentiated by the FHN parameters. The FHN model can approximate the features, however it does not well-represent noisy abnormal beats. Additionally, the FHN model is a generic model, simplified from the Hodgkin-Huxley model. It cannot fully approximate the key spiking behavior of bursting, and it lacks realistic action potential duration (APD) restitution [18], [19]. Key spiking behavior of bursting refers to rapid action potentials followed by dormant periods [20]. APD refers to how long

a cardiac cell stays depolarized, and APD restitution is the rate-dependent change over this duration [21]. The FHN model cannot reproduce the rate-dependent APD restitution as it cannot model multi-variable dynamics and history-dependent recovery [22]. To model this behavior, a set of ODEs that can switch between multiple timescales is required. This may require higher-dimensional features or more sophisticated dyanmical models. This would result in more specialized feature vectors and aid the kNN model in differentiating between different beat types, aiding the model to identify specific heart conditions.

Further work could involve training and testing the model on more patients with more abnormal beat types, reducing the reliance on undersampling and oversampling techniques caused by large class imbalances between arrhythmia subtypes. Additionally, merging the FHN model with more detailed models, such as the Hindmarsh–Rose and/or Monodomain model, may improve the representation of complex arrhythmia dynamics [23]. The Hindmarsh–Rose model consists of three coupled differential equations that capture bursting and spiking behaviors not represented by the FHN model [24]. The Monodomain model simulates how electrical signals spread through cardiac tissue [23]. As a computationally efficient simplification of the Bidomain model, it combines systems of ordinary and partial differential equations to capture propagation of electrical signals in networks of intracellular and extracellular space. Incorporating such spatial dynamics into the current classification model could enable the modelling of continuous sequences of beats, allowing for a temporal analysis of the beats over multiple cycles [25]. Moreover, the model can easily be coupled with the FHN model through its ionic current term [26].

6 Conclusion

This project demonstrates that the FitzHugh–Nagumo ordinary differential equation model can accurately capture essential dynamics of cardiac rhythms and serve as a foundation for arrhythmia classification. Our approach achieves 84% accuracy for binary classification in distinguishing normal beats from abnormal beats using only eight features: six FHN parameters and two width features. This is a drastic reduction in dimensionality and computational expenses compared to deep learning methods that typically have hundreds to thousands of features [27]. However, abnormal classes displayed significant overlap in feature space, achieving only 50% accuracy when distinguishing between normal beats, left bundle branch block beats, and other arrhythmia types. This limitation indicates that the low-dimensional parameter space of the FHN model alone is not enough to reliably separate the various arrhythmia subtypes. Further work can involve training the model on a larger dataset to assess its effectiveness on other beat types, integrating additional features such as key spiking behaviour of bursting and APD restitution, as well as the integration of higher-dimensional models to achieve an optimal balance between performance and computational efficiency.

References

[1] Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, jul 1961.

[2] Martin Andrews. AED Statistics in Canada — aed.ca. https://aed.ca/blogs/news/aed-statistics-in-canada?srsltid=AfmBOooQQbwB34rJRhPz8bSVeaM_SCqX4OP0H-8LztlM09-bkanbyOEF.

[3] Mount Sinai. Types of Arrhythmias - Mount Sinai Heart | Mount Sinai - New York — mountsinai.org. https://www.mountsinai.org/care/heart/services/arrhythmia/types. [Accessed 04-10-2025].

[4] PMcardio. PMcardio: AI that reads ECGs in seconds | Powerful Medical — powerfulmedical.com. https://www.powerfulmedical.com/pmcardio-individuals/.

[5] George Wright. AI stethoscope could detect heart conditions in seconds — bbc.com. https://www.bbc.com/news/articles/c2l748k0y77o, 2025.

[6] Euan A. Ashley and Josef Niebauer. *Cardiology Explained.* Remedica, London, 2004. Conquering the ECG. Available from: https://www.ncbi.nlm.nih.gov/books/NBK2214/.

[7] Danbi Gwon, Hakyung Cho, and Hangsik Shin. Feasibility of a waistband-type wireless wearable electrocardiogram monitoring system based on a textile electrode: Development and usability study. *JMIR mHealth and uHealth*, 9(5):e26469, 2021. Preprint available at ResearchGate (Dec 2020).

[8] Cedars-Sinai Health Library. Left bundle branch block. https://www.cedars-sinai.org/health-library/diseases-and-conditions/l/left-bundle-branch-block.html, 2025. Accessed: 2025-12-19.

[9] Alan L. Hodgkin and Andrew F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544, aug 1952.

[10] Daniel Cebrián-Lacasa, Pedro Parra-Rivas, Daniel Ruiz-Reynés, and Lendert Gelens. Six decades of the fitzhugh–nagumo model: A guide through its spatio-temporal dynamics and influence across disciplines. *Physics Reports*, 1096:1–39, December 2024.

[11] George B. Moody and Roger G. Mark. MIT-BIH Arrhythmia Database (Version 1.0.0). PhysioNet, 2005. Dataset.

[12] Dominique Makowski, Tam Pham, Zen J. Lau, Jan C. Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and S. H. Annabel Chen. Neurokit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4):1689–1696, 2021.

[13] A. C. Hindmarsh. Odepack, a systematized collection of ode solvers. *IMACS Transactions on Scientific Computation*, 1:55–64, 1983.

[14] Saša Singer and John A. Nelder. Nelder–mead algorithm. *Scholarpedia* (Online), 2009. Available at: http://www.scholarpedia.org/article/Nelder-Mead_algorithm.

[15] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008.

[16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[17] Mayo Clinic. Heart Arrhythmia. https://www.mayoclinic.org/diseases-conditions/heart-arrhythmia/symptoms-causes/syc-20350668. [Accessed 04-10-2025].

[18] Rose T. Faghih, Ketan Savla, Munther A. Dahleh, and Emery N. Brown. Broad range of neural dynamics from a time-varying fitzhugh–nagumo model and its spiking threshold estimation. *IEEE Transactions on Biomedical Engineering*, 59(3):816–823, 2012.

[19] R.H. Clayton, O. Bernus, E.M. Cherry, H. Dierckx, F.H. Fenton, L. Mirabella, A.V. Panfilov, F.B. Sachse, G. Seemann, and H. Zhang. Models of cardiac tissue electrophysiology: Progress, challenges and open questions. *Progress in Biophysics and Molecular Biology*, 104(1):22–48, 2011. Cardiac Physiome project: Mathematical and Modelling Foundations.

[20] Patrick A. Morelo, Mauricio Girardi-Schappo, Bianca L. Paulino, et al. Recovering from cardiac action potential pathologies: a dynamic view. *Research Square*, 2024. Preprint.

[21] M. E. L. Hardy, E. Pervolaraki, O. Bernus, and E. White. Dynamic action potential restitution contributes to mechanical restitution in right ventricular myocytes from pulmonary hypertensive rats. *Frontiers in Physiology*, 9:205, 2018.

[22] Xinwei Han, Yao Chen, Weihua Gao, Juel Xue, Xiaodong Han, Zuxiang Fang, Cuiwei Yang, and Xiaomei Wu. Study of the restitution of action potential duration using the artificial neural network. *Mathematical Biosciences*, 207(1):78–88, 2007.

[23] Y. Wang, L. Cai, X. Luo, et al. Simulation of action potential propagation based on the ghost structure method. *Scientific Reports*, 9(1):10927, 2019.

[24] J. L. Hindmarsh and R. M. Rose. A model of neuronal bursting using three coupled first order differential equations. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 221(1222):87–102, 1984.

[25] Henrik Finsberg. fenicsx-beat – an open source simulation framework for cardiac electrophysiology. *Journal of Open Source Software*, 10(114):8416, 2025.

[26] Thomas Roy, Yves Bourgault, and Charles Pierre. Analysis of time-stepping methods for the monodomain model. Technical Report hal-01609274v2, HAL Archives, 2017.

[27] Yasir Ansari, Omar Mourad, Khalid Qaraqe, and Erchin Serpedin. Deep learning for ecg arrhythmia detection and classification: an overview of progress for period 2017–2023. *Frontiers in Physiology*, 14:1246746, 2023.

A Example of Fixed Point Analysis

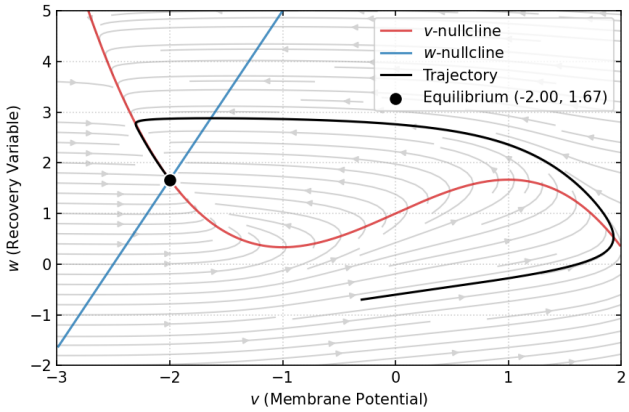


Figure A.1: Phase portrait of a sample FHN solution under the conditions $a = 2.5$, $b = 0.3$, $\tau = 5.2$, and $I = 1$. The fixed point is located at $(-2.00, -1.67)$.

To verify that the conditions in Figure A.1 represents a valid excitable cardiac rhythm, we check the stability conditions at the fixed point $(v^*, w^*) = (-2.00, 1.67)$. The Jacobian matrix evaluated at this point is:

$$J(v^*, w^*) = \begin{pmatrix} 1 - (v^*)^2 & -1 \\ \frac{1}{\tau} & -\frac{b}{\tau} \end{pmatrix} = \begin{pmatrix} 1 - (-2.00)^2 & -1 \\ \frac{1}{5.2} & -\frac{0.3}{5.2} \end{pmatrix} \tag{12}$$

$$J(v^*, w^*) = \begin{pmatrix} -3.00 & -1 \\ 0.192 & -0.058 \end{pmatrix} \tag{13}$$

The eigenvalues λ satisfy the characteristic equation:

$$\det(J - \lambda I) = \lambda^2 - \text{tr}(J)\lambda + \det(J) = 0 \tag{14}$$

Computing the trace and determinant:

$$\begin{aligned} \text{tr}(J) &= [1 - (v^*)^2] - \frac{b}{\tau} \\ &= [1 - (-2.00)^2] - \frac{0.3}{5.2} \\ &= [1 - 4.00] - 0.058 \\ &= -3.00 - 0.058 \\ &= -3.058 \end{aligned} \tag{15}$$

$$\begin{aligned} \det(J) &= \left(1 - (v^*)^2\right)\left(-\frac{b}{\tau}\right) + \frac{1}{\tau} \\ &= (-3.00)(-0.058) + 0.192 \\ &= 0.174 + 0.192 \\ &= 0.366 \end{aligned} \tag{16}$$

The characteristic equation becomes:

$$\lambda^2 + 3.058\lambda + 0.366 = 0 \tag{17}$$

Using the quadratic formula:

$$\begin{aligned} \lambda_{1,2} &= \frac{-3.058 \pm \sqrt{(3.058)^2 - 4(0.366)}}{2} \\ &= \frac{-3.058 \pm \sqrt{9.351 - 1.464}}{2} \\ &= \frac{-3.058 \pm \sqrt{7.887}}{2} \\ &= \frac{-3.058 \pm 2.808}{2} \end{aligned} \tag{18}$$

This gives two real eigenvalues:

$$\lambda_1 = \frac{-3.058 + 2.808}{2} = \frac{-0.250}{2} = -0.125 \tag{19}$$

$$\lambda_2 = \frac{-3.058 - 2.808}{2} = \frac{-5.866}{2} = -2.933 \tag{20}$$

We verify the stability conditions: $\text{tr}(J) = -3.058 < 0$ and $\det(J) = 0.366 > 0$. Both eigenvalues are real and negative, confirming that the fixed point is a stable node. This parameter set produces valid excitable behaviour: the system rests at a stable equilibrium and returns to rest after a large disruption.