Comparison of Feature Selection Methods for ECG Classification(MIT-BIH)

Team Members

- Akshita Chauhan(102203527)
- Yatish Garg(102203171)
- Sidharth Chugh(102303311)
- Tanishq Bansal(102303790)

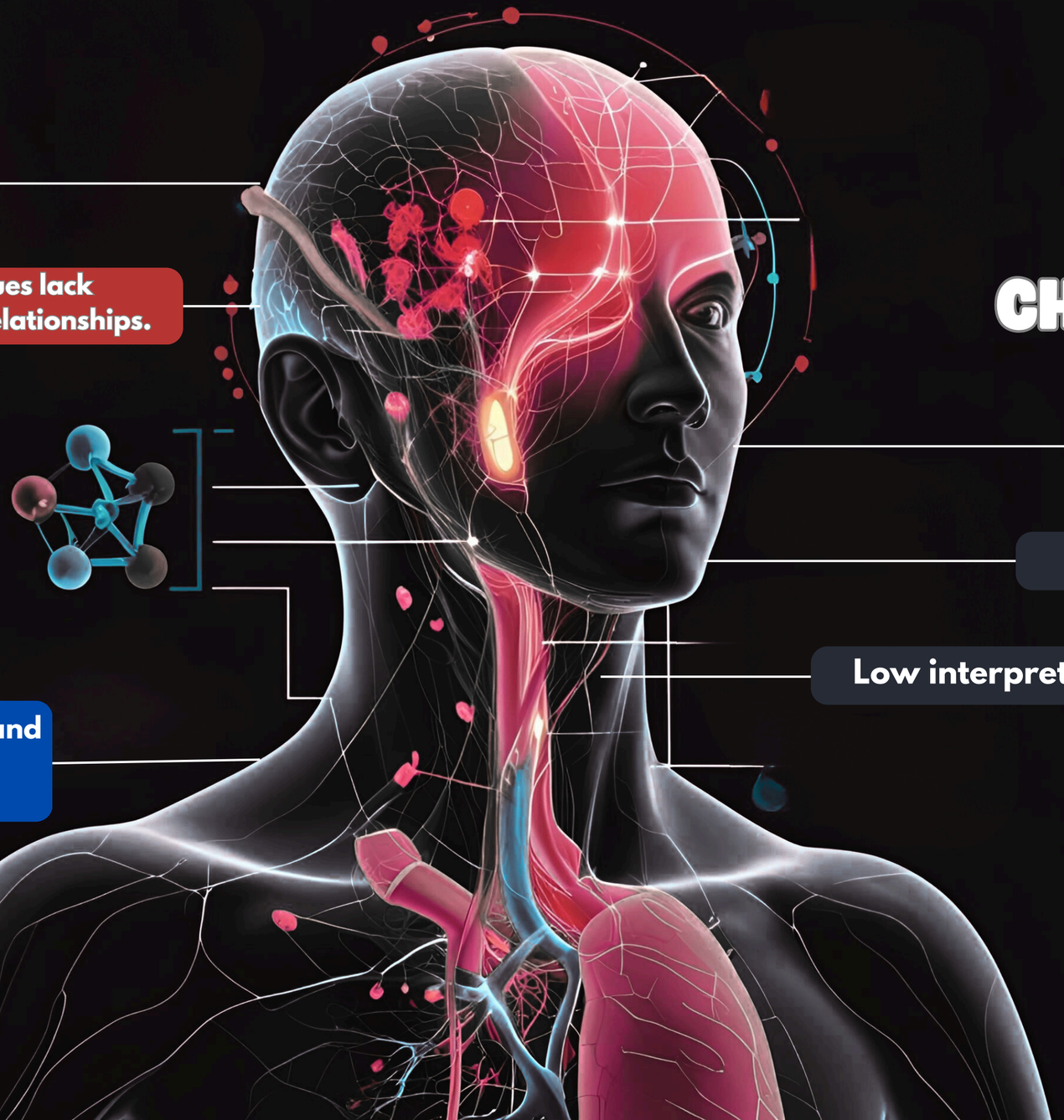
Mentors

- Dr Jyoti Maggu
- Dr Simranjit Kaur

PROBLEM CONTEXT

ECG signals are complex, high-dimensional, and often noisy.

Traditional feature selection techniques lack explainability and ignore inter-feature relationships.



MOTIVATION

xGNNs can model feature interactions and provide attention-based importance scores.

CHALLENGES

Redundancy in features

High computational load

Low interpretability in classical models

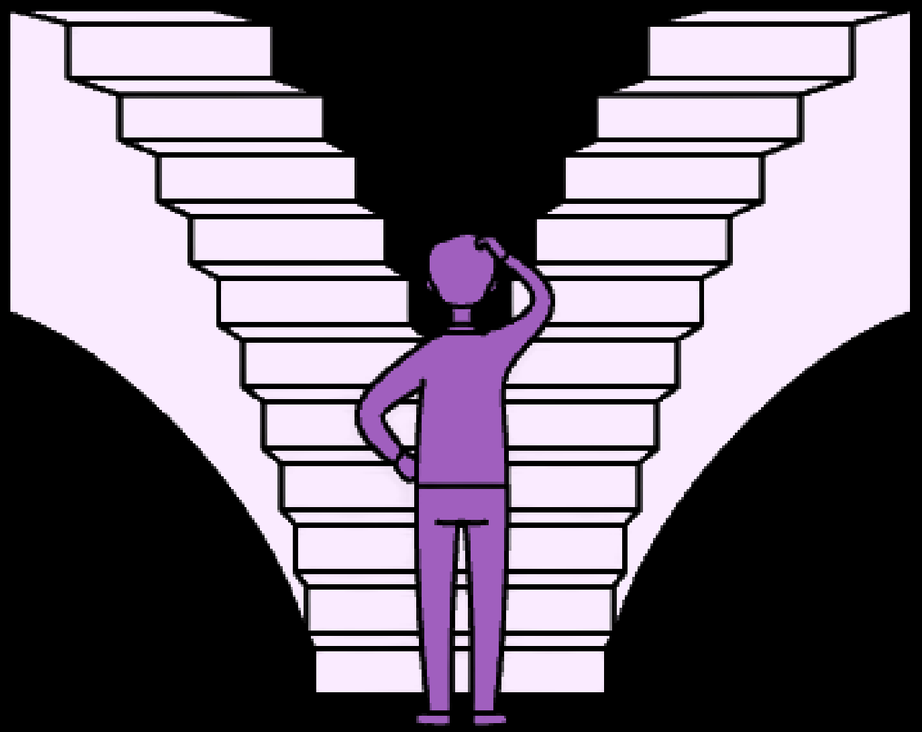
Limitation of Existing Solutions

LASSO Regression

Effective for eliminating features by setting coefficients to zero, but may arbitrarily drop correlated predictors.

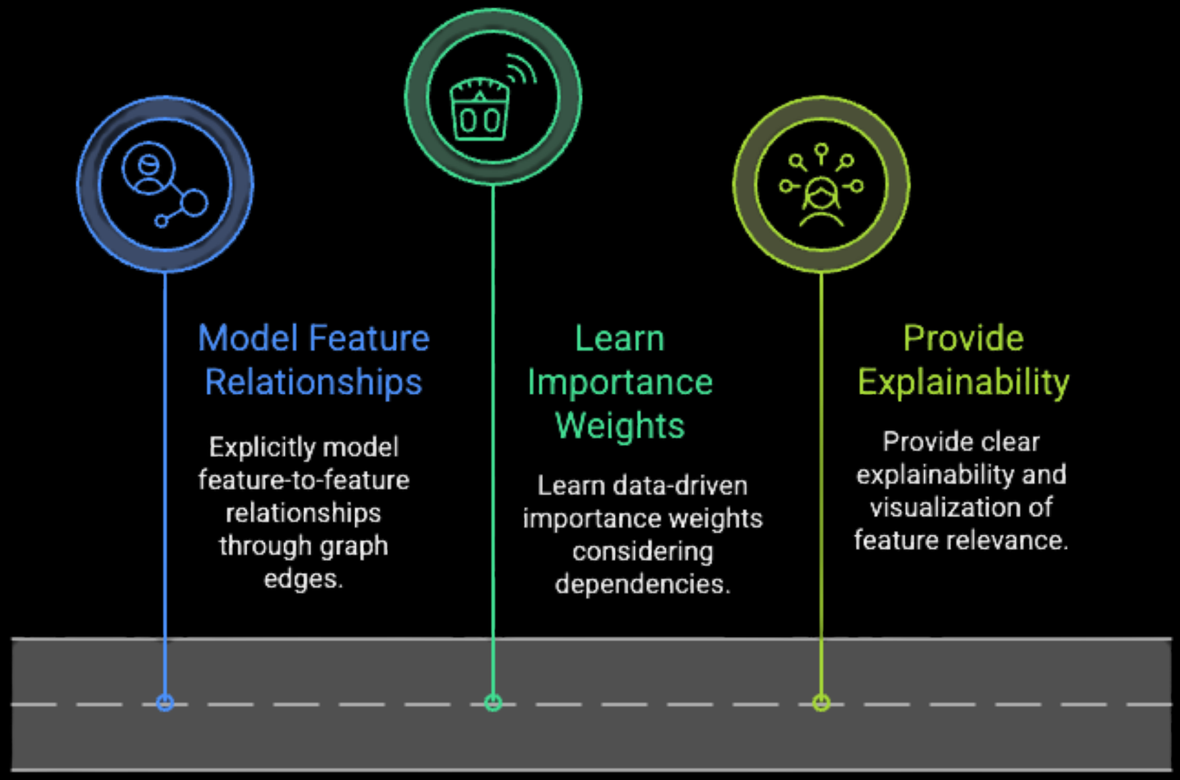
Ridge Regression

Shrinks coefficients without eliminating features, retaining all predictors but potentially including redundant ones.

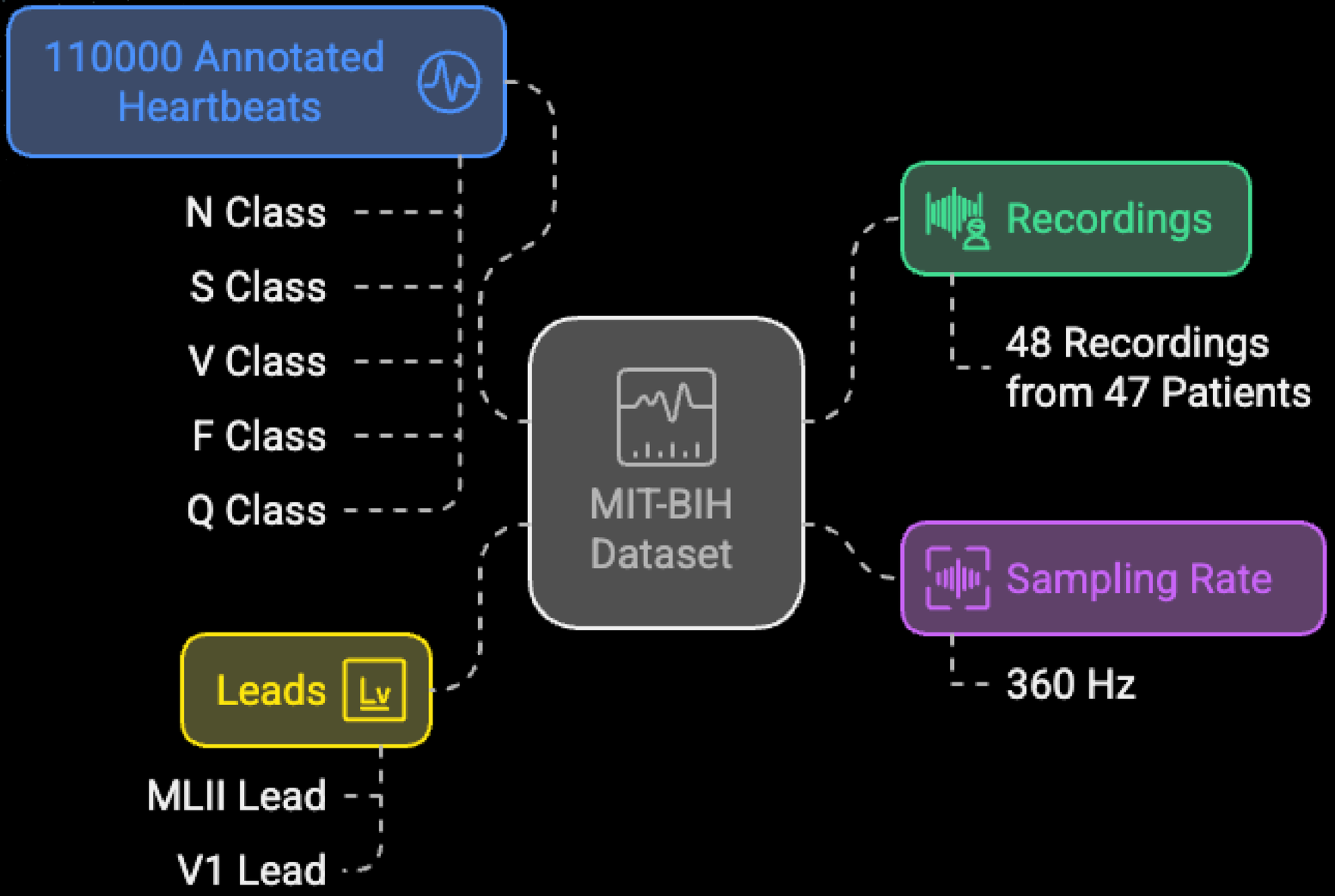


How our solution overcome those limitations?

GNN with Attention Approach



Dataset Overview



Steps involved in the approach

Feature Extraction

Extract and classify features from the dataset.

Graph Construction

Construct a graph with features as nodes and relationships as edges.

GNN Training using GAT

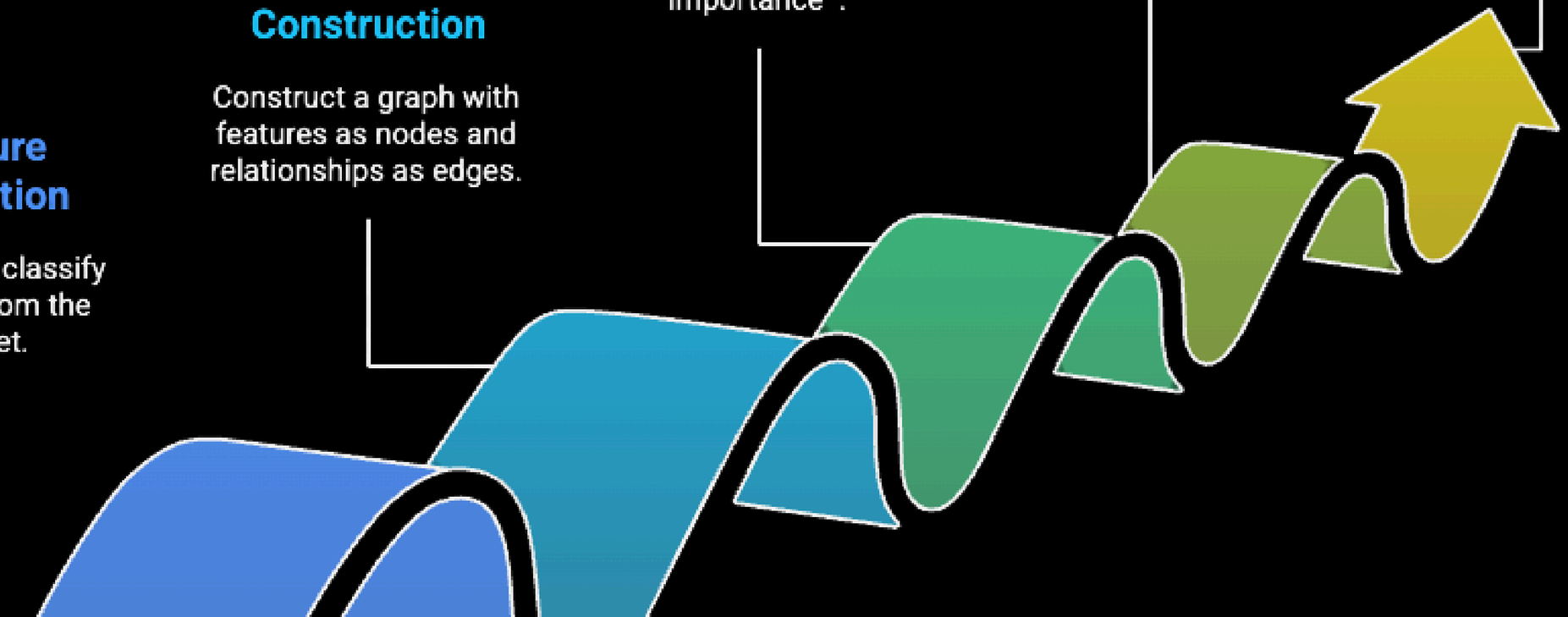
Train a Graph Neural Network to learn feature importance .

Feature Selection

Select the top-k features based on attention weights.

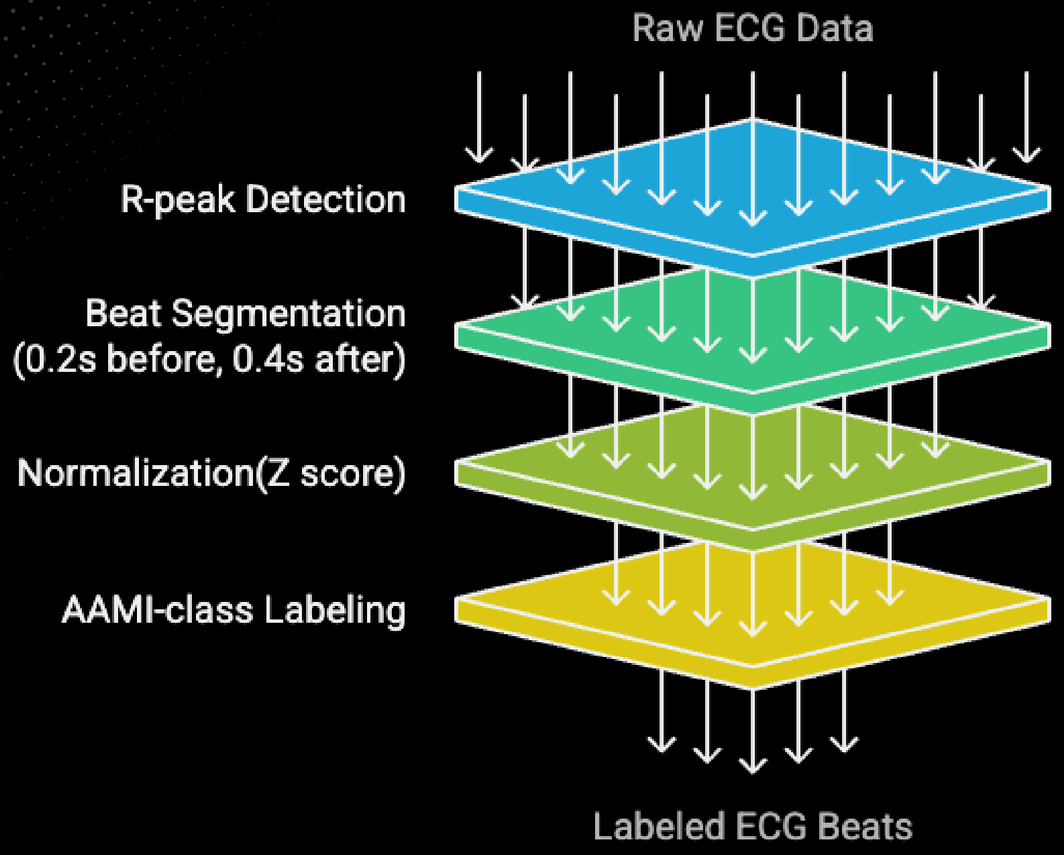
Downstream Classifier

Train a classifier on selected features for final prediction.

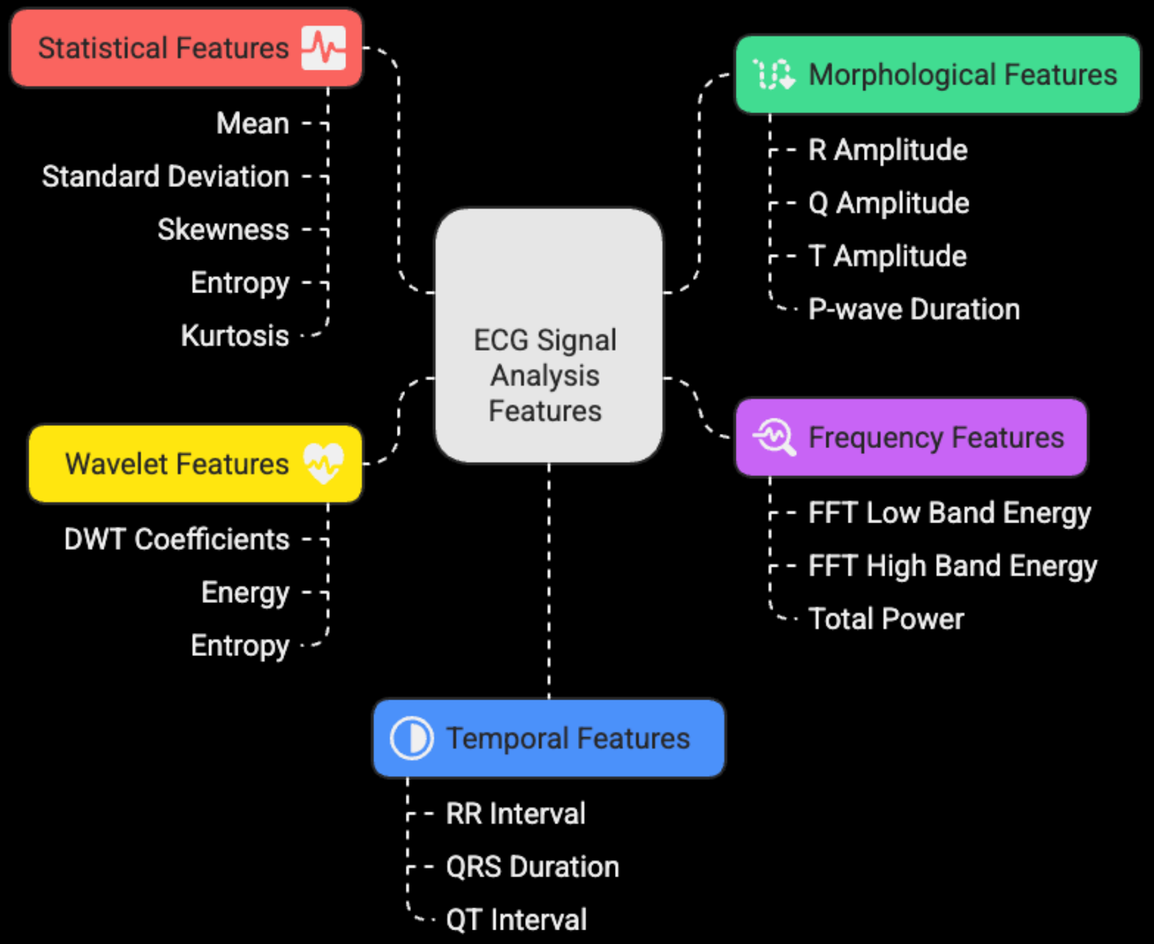


Feature Extraction

ECG Data Processing Funnel



ECG Signal Analysis Features



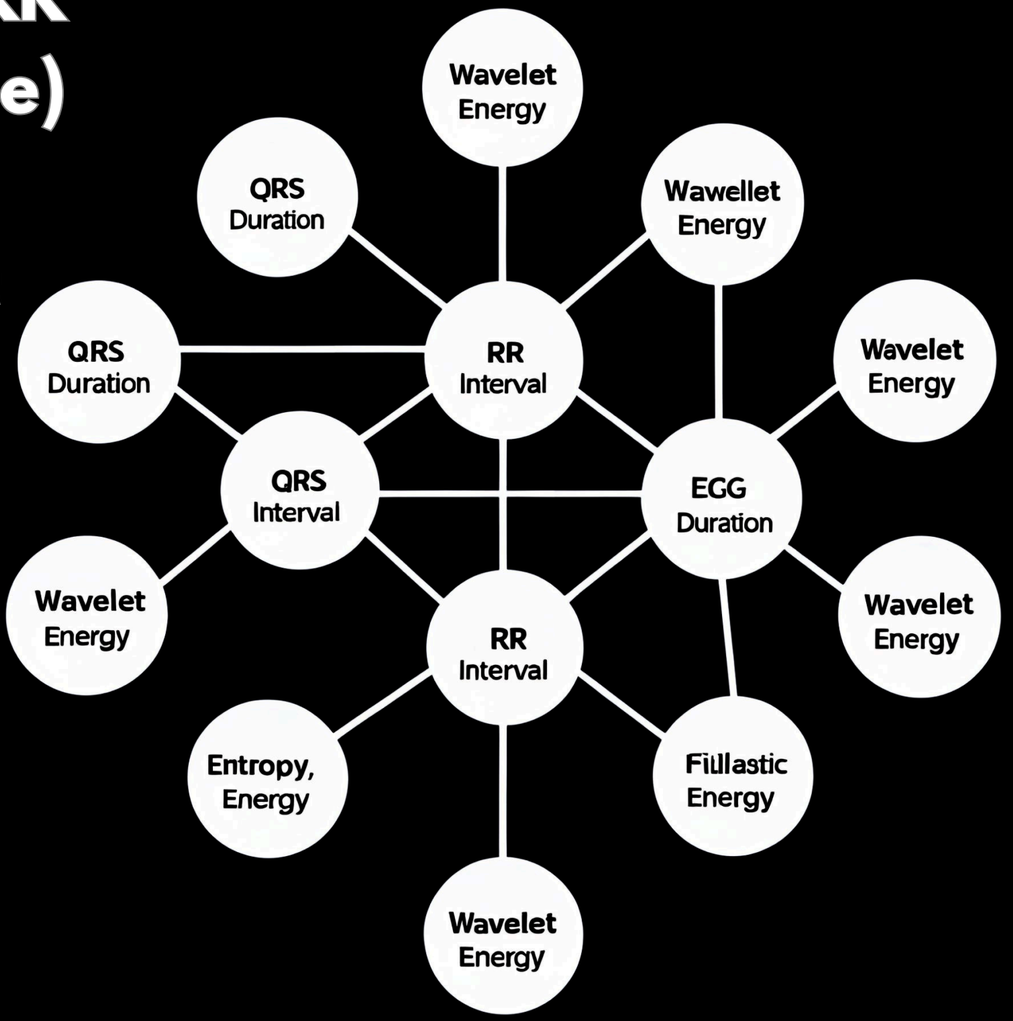
Graph Construction

Each ECG feature (e.g., QRS Duration, RR Interval, Entropy, T Wave Amplitude) becomes a node in a graph.

Edges are created between nodes to represent relationships or dependencies, such as:

- Statistical Correlation (e.g., Pearson correlation > threshold)
- Mutual Information (shared predictive value)
- Domain Knowledge (clinical association between intervals)

This feature graph captures how features influence each other, which traditional models ignore.



GNN Training using GAT

Process Feature Graph

The GAT begins by processing the feature graph.

Learn Node Attention

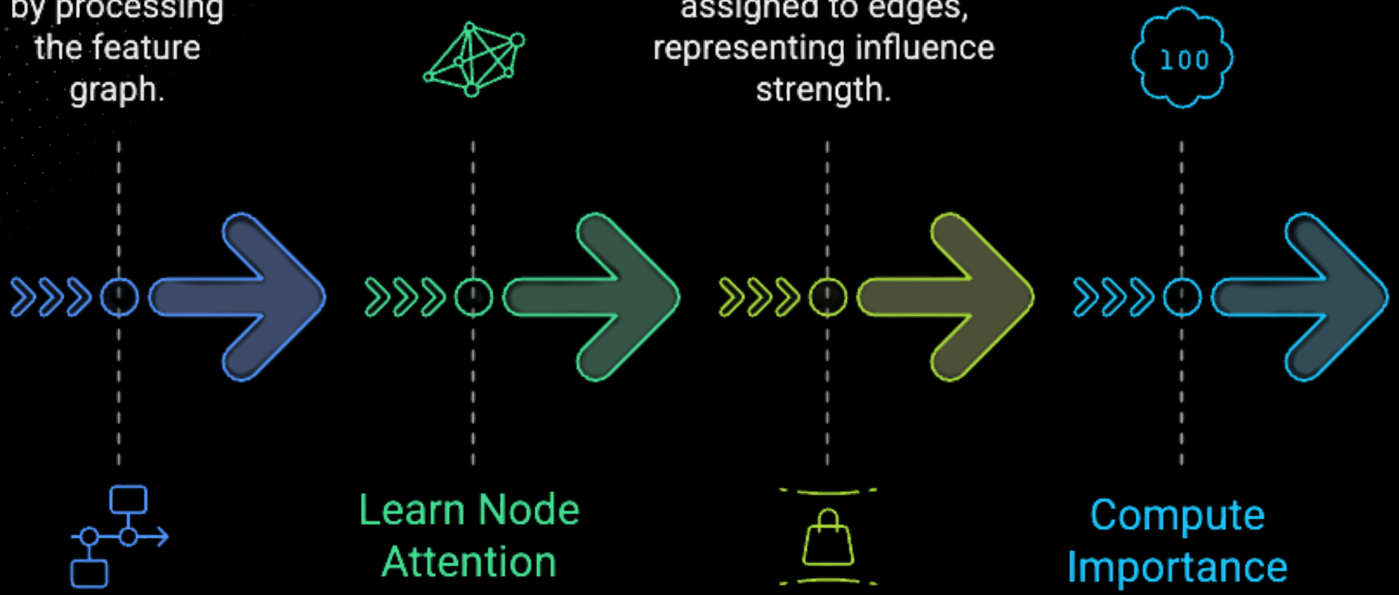
The model learns how much each node should attend to its neighbors.

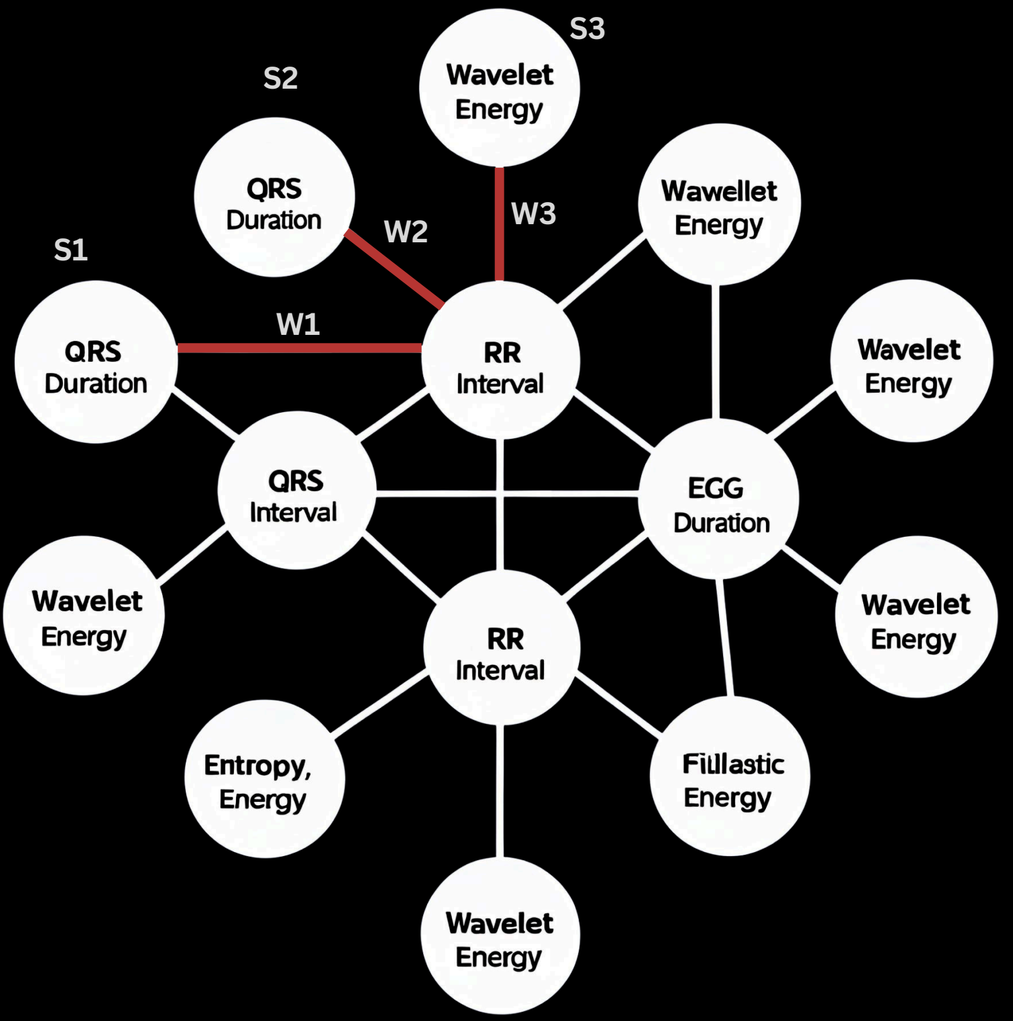
Assign Attention Weights

Attention weights are assigned to edges, representing influence strength.

Compute Importance Scores

Importance scores are computed for each node, reflecting feature criticality.





Feature Selection

Train GAT Model

The GAT model is trained on the feature graph.

Assign Importance Scores

Each feature is assigned an importance score based on attention.

Rank Features

Features are ranked according to their importance scores.

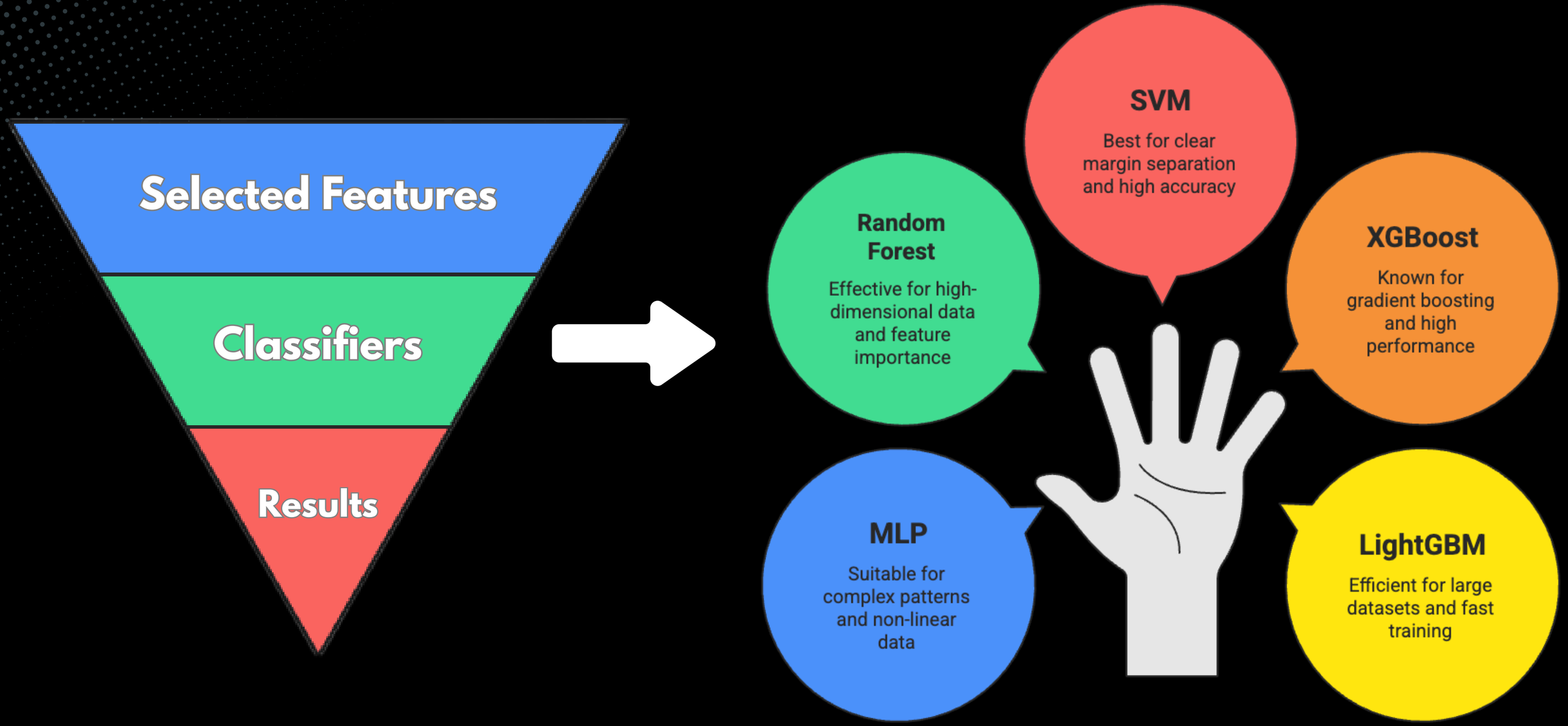
Apply Selection Strategy

A threshold or top-K strategy is used to select features.

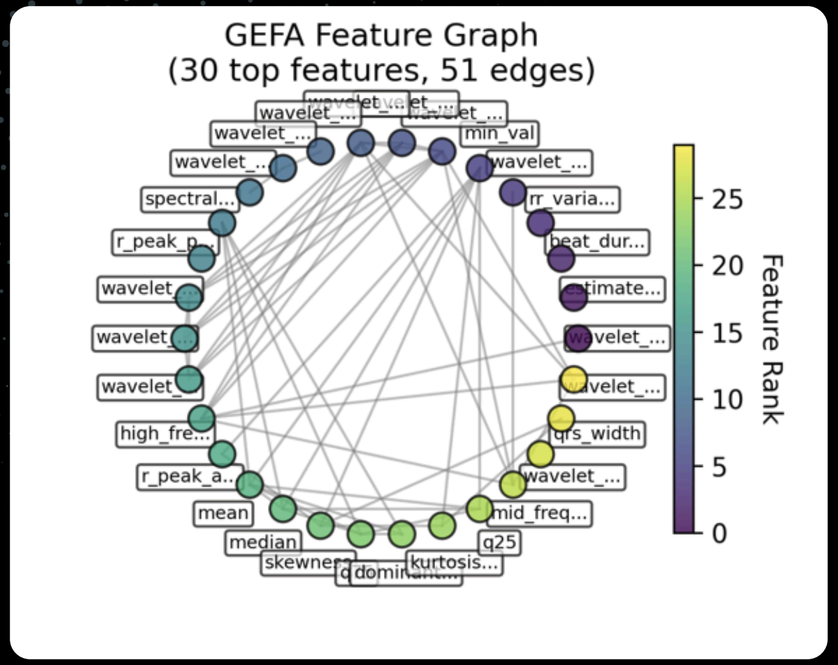
Drop Less Important Features

Features with low attention scores are removed.

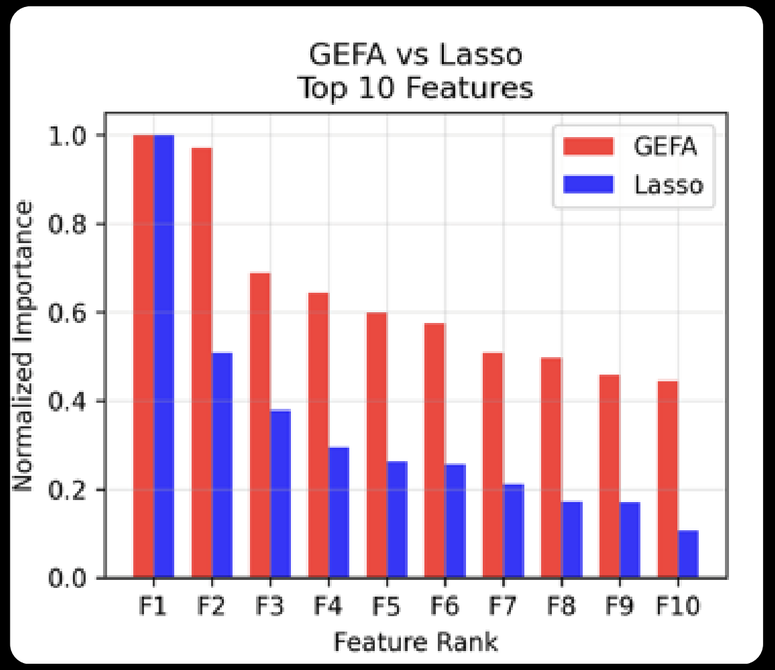
Downstream Classifiers for Final Prediction



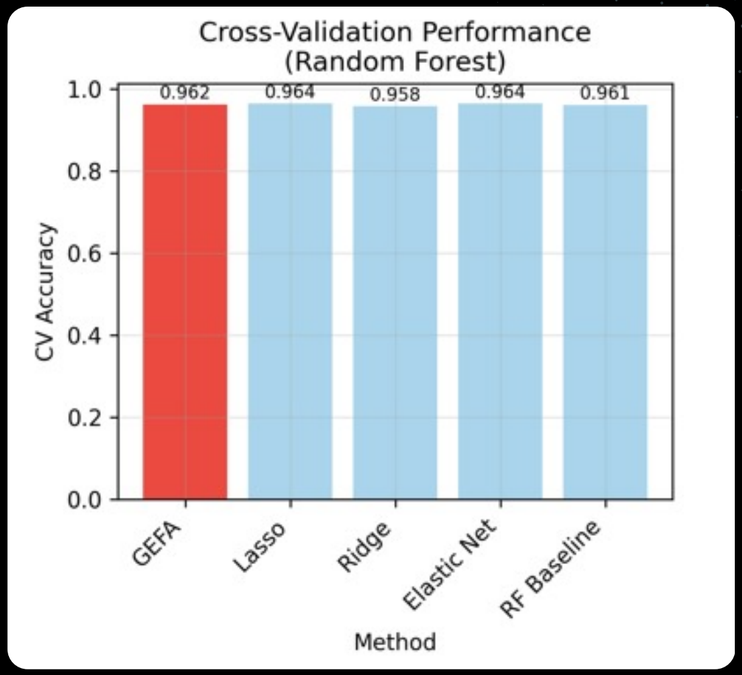
Feature Graph

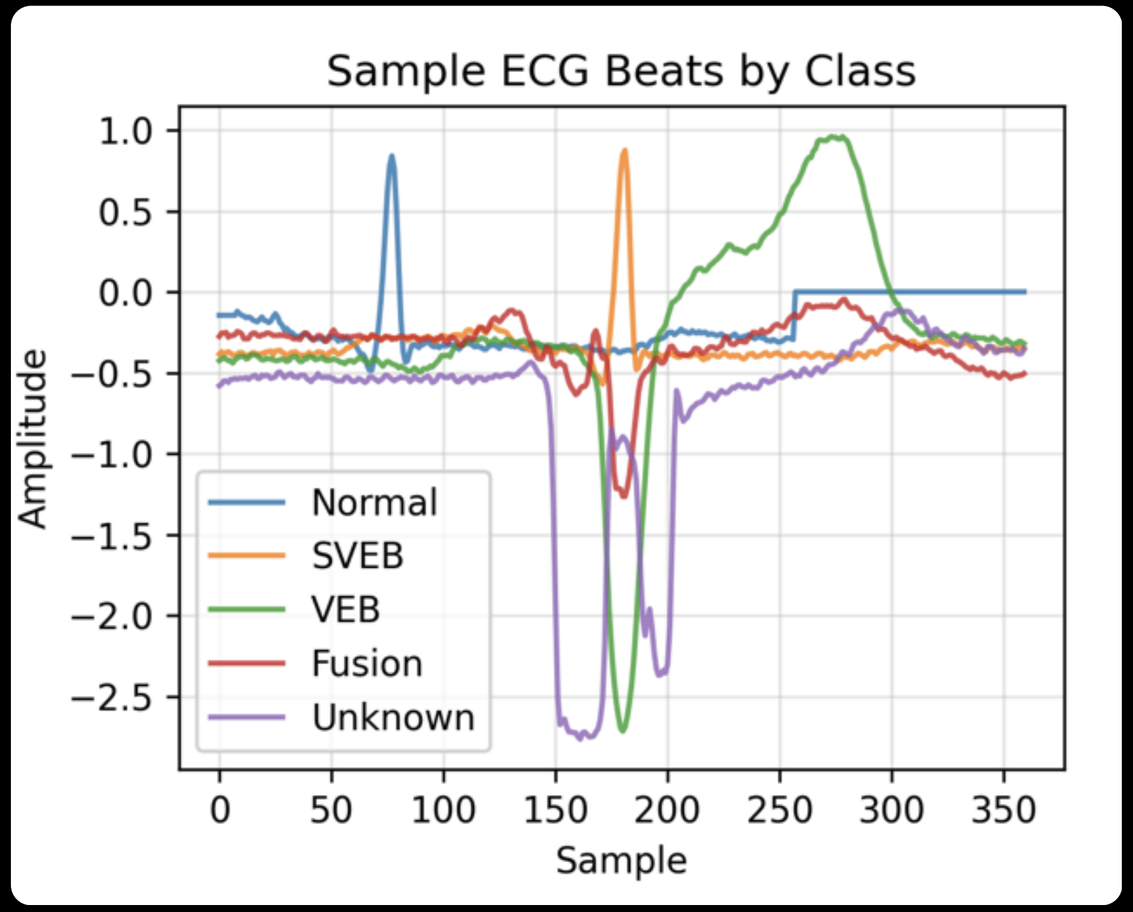


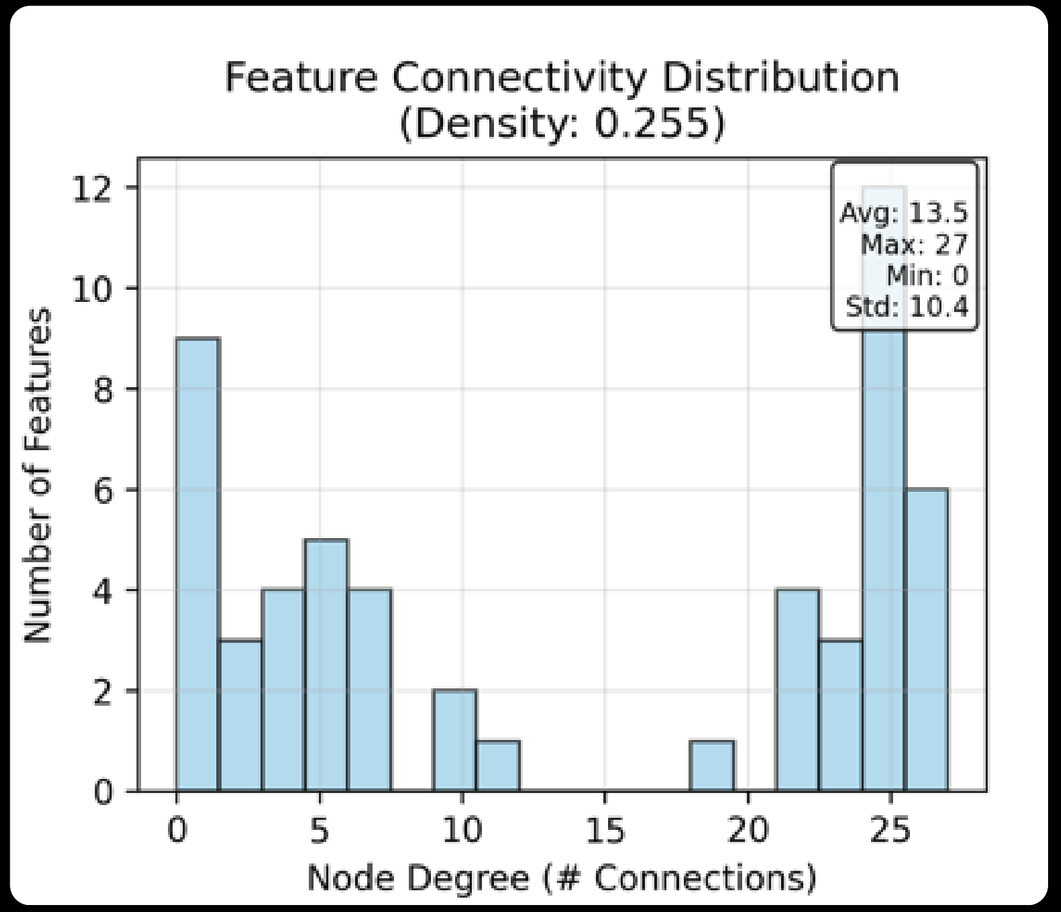
Feature Ranks



CV Comparison




Sample ECG Beats by Class
Amplitude
Sample
Normal
SVEB
VEB
Fusion
Unknown
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0
50
100
150
200
250
300
350


Feature Connectivity Distribution
(Density: 0.255)
Number of Features
Node Degree (# Connections)
Avg: 13.5
Max: 27
Min: 0
Std: 10.4
0
2
4
6
8
10
12
0
5
10
15
20
25

Outputs and Final Results

Top-K features

📌 Top-k Features Selected by GEFA:
wavelet_mean_2, estimated_hr, beat_duration, rr_variability, wavelet_mean_3, min_val, wavelet_std_2, wavelet_std_1, wavelet_energy_2, wavelet_mean_4, wavelet_mean_5, wavelet_mean_6, spectral_centroid, r_peak_position, wavelet_max_2, wavelet_max_1, wavelet_energy_1, high_freq_power, r_peak_amplitude, mean, median, skewness, q75, dominant_freq, kurtosis_stat, q25, mid_freq_power, wavelet_mean_1, qrs_width, wavelet_energy_3

Classification Report

```
================================================================================================================================================
🏆 FEATURE SELECTION COMPARISON RESULTS
================================================================================================================================================
✅ Successfully completed 5 methods
Method          RF Test    RF CV      SVM Test   SVM CV     XGB Test   XGB CV     LGBM Test  LGBM CV    MLP Test   MLP CV     Avg Score
------------------------------------------------------------------------------------------------------------------------------------------------
   GEFA         0.9675     0.9605     0.9575     0.8261     0.9694     0.9615     0.9677     0.9651     0.9575     0.9164     0.9449     ⭐
   Lasso        0.9686     0.9643     0.9575     0.8730     0.9711     0.9656     0.9720     0.9699     0.9635     0.9304     0.9536
   Ridge        0.9618     0.9577     0.9677     0.8399     0.9728     0.9643     0.9669     0.9646     0.9643     0.8967     0.9457
   Elastic Net  0.9660     0.9643     0.9584     0.8733     0.9626     0.9669     0.9686     0.9666     0.9550     0.9291     0.9511
   RF Baseline  0.9626     0.9607     0.9643     0.7157     0.9703     0.9669     0.9669     0.9691     0.9686     0.6976     0.9143
```

THANK YOU!