Parameterising ECG Waveforms to Uncover Pathology: A Phase-Folded Approach

Isha Agrawal

University of Aberdeen

MSc Data Science, 2024/2025

Abstract

The growing impact of cardiovascular disease (CVD) highlights the urgent need for accessible and automated diagnostic tools. Traditionally, 12-lead ECGs require clinical expertise to deploy and interpret, which is impractical for out-of-clinic examination. This project aims to address this issue by developing an automated ECG classification system using parameterised waveform features, including PQRST amplitudes and durations. Phase-folding was applied to standardize signals from lead V6 into a single cardiac cycle, allowing efficient comparison and reducing computational complexity. The extracted parameters were used to train a LightGBM model for both binary (Normal vs Abnormal) and multiclass classification. The binary model achieved high performance with an accuracy of 83% and an AUC of 0.90, while the multiclass model achieved an accuracy of 68%. These findings demonstrate that even with a reduced, parameterised input, high classification accuracy can be achieved. This approach shows potential for scalable, out-of-clinic ECG analysis and could contribute to early detection. To improve model performance, limitations such as class imbalance and potential signal distortion from phase-folding must be addressed.

1. Introduction

Cardiovascular diseases (CVDs) remain one of the leading causes of mortality and a major contributor to impaired quality of life. According to the World Health Organisation (2019), approximately 17.9 million deaths occur each year due to CVDs. Early diagnosis is critical for effective treatment and preventing serious cardiac events. Electrocardiography (ECG) is a widely used, non-invasive medical examination that detects irregularities in the heart's rhythm and conduction by measuring its electrical activity (Institute for Quality and Efficiency in Health Care, 2023). However, due to its 12-lead placement and the need for professional interpretation, it can be challenging to perform in out-of-clinic settings. The objective of this project is to overcome these challenges by creating an automated classification of the ECG signal through parameterisation of key waveform features to uncover pathology. These features are used to train models capable of both binary (Normal vs Abnormal) and multiclass classification across multiple cardiac conditions. This project differs from previous studies by solely using lead V6 and applying phase-folding to standardize the ECG signal, making our analysis more suitable for out-of-clinic diagnosis.

1.1. Electrocardiography

The standard 12-lead ECG consists of 10 electrodes placed on different areas of the body to provide 12 perspectives of the heart's electrical signals. The 4 limb electrodes are placed on the right arm, left arm, right leg, and left leg; the 6 chest electrodes are placed on specific chest areas. (Institute for Quality and Efficiency in Health Care, 2023)

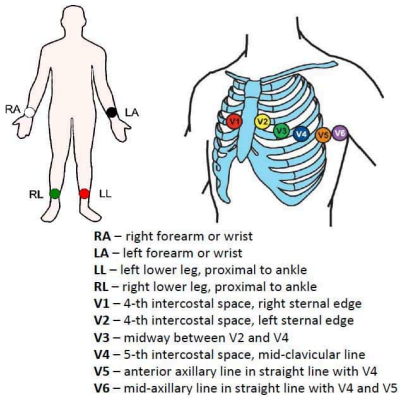

Figure 1. Standard 12-ECG Electrode placement. (AIMCARDIO, 2020)

1.2. PQRST Components

The ECG trace produced looks like a waveform that repeats with each heartbeat; its components include the PQRST waves. The P wave is the first upward signal and represents atrial depolarisation; the atria contract, pushing blood into the ventricles. The QRS complex represents ventricular depolarisation; electrical signals cause the ventricles to contract. The T wave is the repolarisation of the ventricles after they contract. These components provide insights into heart rate, rhythm, abnormalities within the heart, and evaluate atrial and ventricular function. (Institute for Quality and Efficiency in Health Care, 2023)

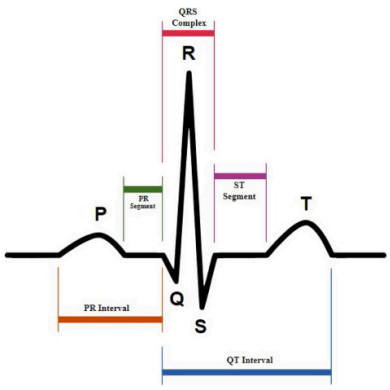

Figure 2. A typical ECG waveform illustrating the PQRST features (Madona, Basti, and Zain, 2021)

1.3. Disease Superclasses

The appearance of the ECG waveform varies depending on the cardiac condition. This project focuses on five categories: Normal ECG (NORM), Cardiac Hypertrophy (HYP), Myocardial Infarction (MI), ST/T segment changes (STTC), and Conduction Disturbances.
The Normal ECG reflects the standard ECG seen in patients without heart abnormalities.
Cardiac Hypertrophy is when the heart muscle thickens or enlarges due to increased workload. It can affect different parts of the heart, e.g., Left Ventricular Hypertrophy, where the wall of the left ventricle thickens. (ECGpedia, 2024)
Myocardial Infarction, commonly known as a heart attack, occurs when blood flow to the heart is blocked, usually due to plaque buildup or blood clots in the coronary arteries. The blockage causes a lack of oxygen in the affected area, which leads to tissue death. (Cleveland Clinic, 2024)
Conduction Disturbances are abnormal heart rhythms caused by disruptions in the heart's normal electrical conduction; this can result in the heart beating too slowly, too quickly, or irregularly. (National Heart, Lung, and Blood Institute, 2024)

ST/T changes are alterations in the ST segment or T wave of the ECG waveform, typically occurring when the heart is unable to repolarize properly due to injury or ischaemia. The normal ST segment is flat and isoelectrical; however, elevation or depression can occur due to conditions such as MI and Ischaemia. (ECG Waves, 2018)

1.4. Objective

The primary objective of this project is to parameterise ECG waveforms to uncover the pathology associated with specific cardiac conditions and develop an automated diagnostic model. Instead of using the whole ECG waveform, we extracted key features such as amplitudes of the P wave, QRS complex, and T wave, along with corresponding intervals like the PR interval and QRS duration. The goal is to find the appropriate parameters that help to differentiate between normal and abnormal cardiac health. ECG signals are time-series data consisting of thousands of data points per lead, which makes analysis computationally demanding and time-consuming. Parameterisation reduces the computational complexity and memory usage as only key points are used for analysis, making the whole process more efficient.

1.5. Process

This is the process the raw ECG data will go through:

1. Data Preprocessing: Highpass and Lowpass Filtering
2. Phase Folding
3. Lead V6 Extraction
4. PQRST Detection
5. Computing relevant intervals and durations
6. Machine Learning using the parameters as features.

2. Data

This study uses the PTB-XL dataset, a large publicly available ECG database. It contains 21,799 records from 18,869 patients, each record consisting of 10-second ECG segments, sampled at either 100 Hz or 500 Hz. The dataset covers ages ranging from 0 to 95 years, and it is balanced concerning sex, with 52% male and 48% female. Up to two cardiologists annotated the ECG records using a set of 71 standardized statements based on the SCP-ECG format. These statements describe the form, rhythm, and diagnosis in a consistent and machine-readable format; these labels were further combined into five superclasses (Wagner et al., 2020).
The PTB-XL dataset is ideal for this project due to its size and detailed labelling of the disease classes.

3. Method

To begin the analysis, a user-defined function named load_ecg was implemented to load individual ECG recordings along with their corresponding metadata. Our analysis used a sampling frequency of 500 Hz to ensure high signal quality, which is essential for precise feature extraction. Lower sampling frequencies, such as 100 Hz, may lead to loss of important details and result in reduced accuracy (Kwon et al., 2018).

3.1. Preprocessing

After loading the raw ECG data, preprocessing was performed using high-pass and low-pass filtering. A high-pass filter was applied at 0.5 Hz to remove the baseline drift that may have been caused by patient movement or respiration. A low-pass filter was applied at 40 Hz to remove any high-frequency noise caused by electrical interferences or muscle artifacts. These steps ensured the data was prepared for analysis.

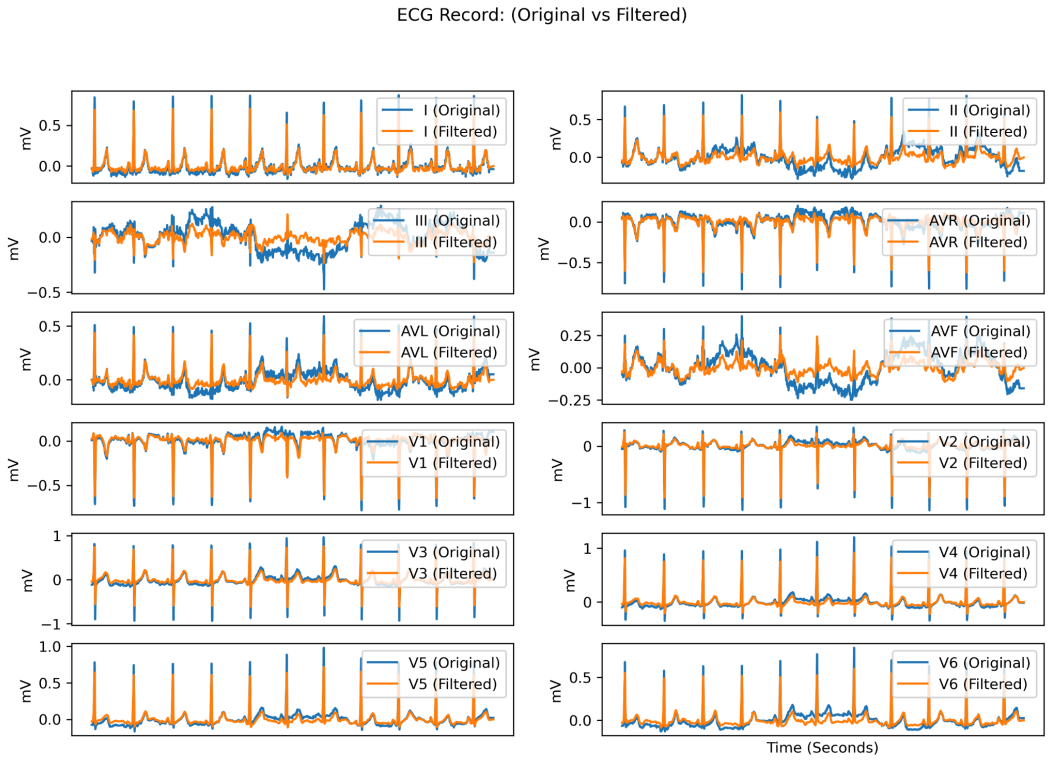

Figure 3. Comparison of raw and filtered 12-lead ECG signal. High-pass and low-pass filters were applied.

3.2. Phase Folding

Phase folding is a method used to identify repeating patterns in periodic signals. In time-series analysis, it involves dividing the signal according to a specific position or known duration. These cycles are then aligned, making it easier to observe any differences (Wang et al., 2024).

In this project, phase folding was applied to the ECG signal to align cardiac cycles for analysis. R peaks were chosen as reference points due to their high amplitude and easy identification. R peak detection was performed using the find_peaks function from the SciPy library. This function finds the local maxima within a 1-D array by comparing neighbouring values (SciPy, 2024). The threshold for R peak detection was set to 30% of the maximum value of the summed signal, and the minimum distance between consecutive R peaks was set to 40% of the sampling frequency. Initially, the minimum distance was set to 60%; however, it was observed that at higher heart rates, some of the R peaks were not detected due to the proximity. The minimum distance was reduced to 40% to ensure robust peak detection at different heart rates.
The average RR interval (distance between consecutive R peaks) was calculated to fold the signal; 80% of the average RR interval was used for the segment length to focus more on relevant features. The ECG signal was divided into these segments, one segment corresponded to one cardiac cycle. An average ECG signal was generated from the segments to represent the typical characteristics of the ECG waveform. Phase-folding and averaging the ECG signal allows the data to be consistent across all patients. Instead of analyzing each cardiac cycle individually, we now have one representative cycle; this allows faster data processing, especially when dealing with large datasets.

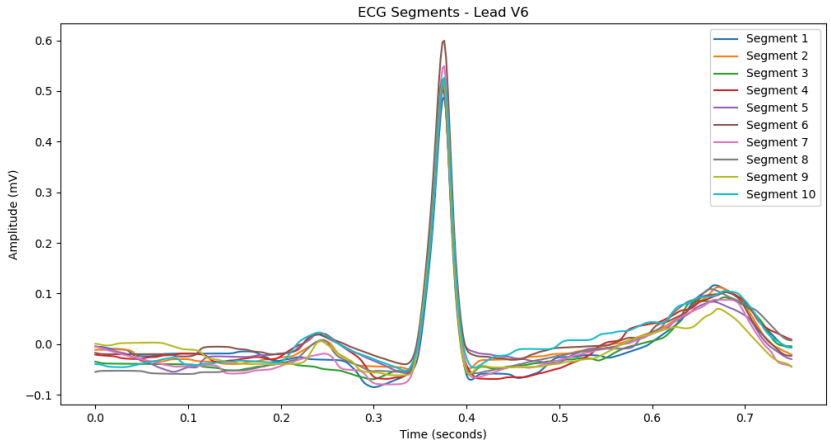

Figure 4. Phase-folded and aligned ECG beat in lead V6. The phase-folding standardises the ECG signal into a single cardiac cycle.

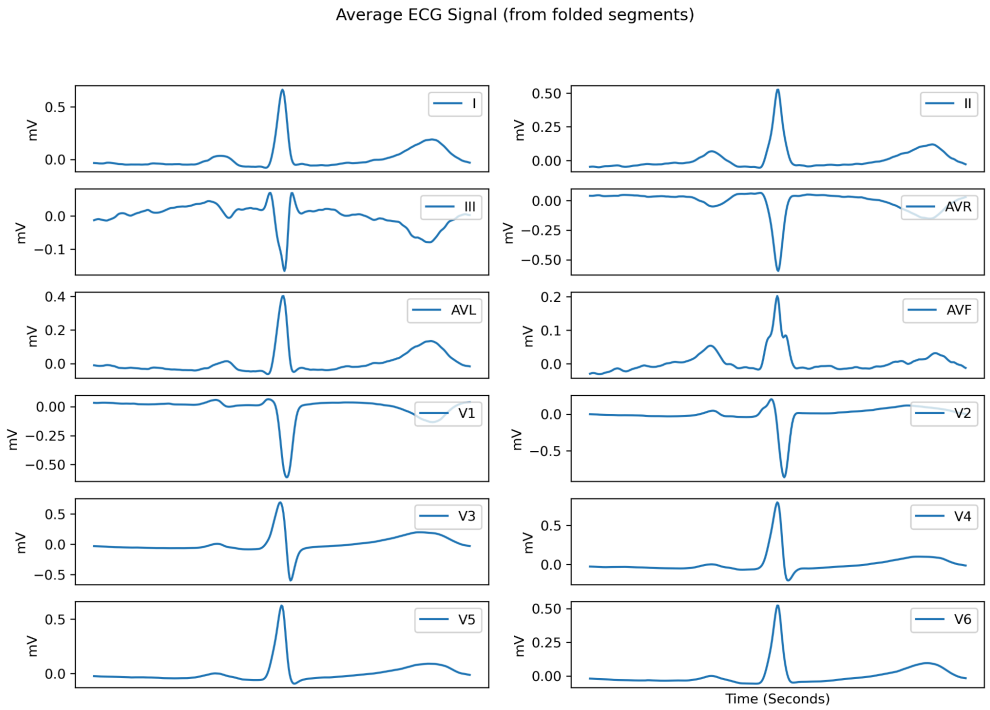

Figure 5. Averaged Phase-Folded 12 Lead ECG, providing a consistent basis for comparison of waveform features.

3.3. Lead V6

After the preprocessing and phase folding, lead V6 was selected to do feature extraction. According to a study conducted by Ramirez et al. (2024), lead V6 contains unique diagnostic features that enhance disease classification. It is positioned to capture the electrical activity of the lateral wall of the left ventricle, which makes it useful for diagnosing conditions arising from this region of the heart, like left ventricular hypertrophy and MI.

3.4. PQRST Detection

After lead V6 extraction, we identified the key features on the ECG waveform that would help with classification. These features included: P wave onset, P wave peak, Q wave, R peak, S wave, T wave onset, T wave peak, and T wave end. A user-defined function, detect_pqrst, was implemented to locate these features; the find_peaks function from the SciPy library was used. R peaks were detected using the technique previously applied during phase folding. Q and S waves were identified as the local minima immediately before and after the R wave, respectively. The P wave peak was detected by looking for the most prominent peak before the Q wave. P wave onset was identified as the first consistent upward slope before the P peak. T wave features were located in the window after the S wave. T wave onset was located as the first point in this region, where the slope starts to rise. The T wave peak was selected as the highest local maximum after the S wave. T-wave end was identified as the point where the slope flattens, and the gradient becomes zero. After correctly identifying these features, intervals

were calculated in milliseconds, such as QRS duration (ms), PR interval (ms), and ST interval (ms), along with the heart rate.

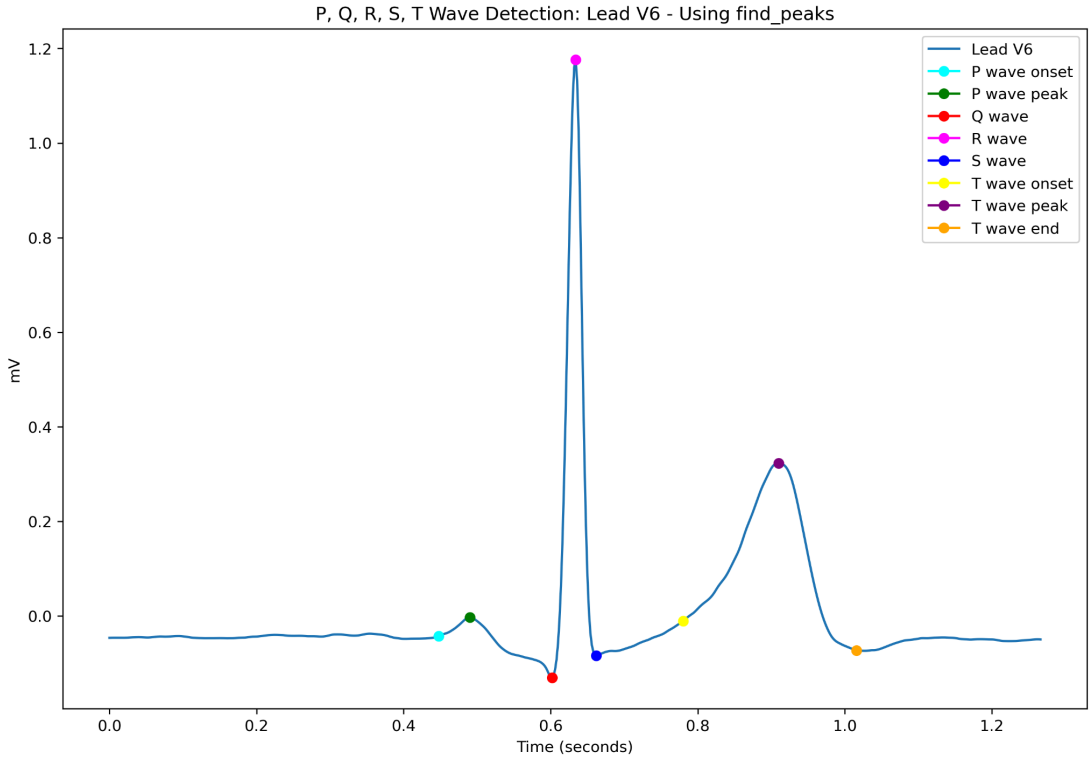

Figure 6. Highlighted PQRST features in lead V6, essential for the parameterization of the ECG signal.

3.5. Machine Learning

After PQRST feature extraction, the next step was to train a machine learning model to classify ECGs based on the diagnostic superclass. To train the model, each ECG ID needed a corresponding diagnostic superclass. The scp_codes were extracted from the PTB-XL database and were converted into Python dictionaries using ast.literal_eval. Each scp_code contained multiple diagnosis codes for each ECG; therefore, they needed to be mapped onto a broader diagnostic superclass. This was achieved by implementing a function called aggregate_diagnostic, which iterated through the scp_codes and matched them with the definitions provided in the scp_statements.csv and returned a superclass. The PQRST features were merged with the diagnostic superclass in preparation for machine learning. The PTB-XL dataset provides a strat_fold column that assigns each ECG ID to a fold to ensure even distribution between each fold. There are 10 folds in total; I started by defining the test_fold as 10 and the train fold as the remaining 1 to 9. However, K-fold cross-validation was performed on the dataset to improve model performance. This method divides the dataset into K (K = 10) equally sized folds. The model is trained K times, using K-1 folds each time for training and the remaining fold for validation. The average of the values is taken as an overall performance

estimate (scikit learn, 2009). X_train was defined as all the PQRST features, and y_train was defined as the diagnostic superclass.

The next step was data cleaning, which involved checking for missing values. Rows with missing diagnostic labels were excluded from the model. For ECGs associated with multiple diagnostic classes, only the first label was used for classification; this made it compatible with LightGBM's framework, which only expects a single label. Some features in the dataset contained missing values, particularly related to the T wave, such as T onset, T peak, T end, T Amplitude, and the ST segment. T-end was removed entirely from the classification as it contained a very high proportion of missing values, which would have degraded the model's performance. K-Nearest Neighbors (KNN) imputation was used to estimate missing values for the remaining features. This method identifies the k=5 most similar data points and averages them to fill in missing values (scikit-learn.org, n.d). KNN imputation is more effective than discarding the missing values or replacing them with the mean or median. ST interval was not imputed using KNN; instead, it was calculated with the known S and T values.

Additional features were engineered from the original ECG measurements to improve model performance. Amplitude ratios such as the R-to-S ratio and T-to-R ratio were computed to capture the relationship between amplitudes.

3.6. Random Forest vs LightGBM

At first, a random forest classifier was chosen for classification. This algorithm uses multiple decision trees during training to improve accuracy and avoid overfitting. To improve accuracy, LightGBM (Light Gradient Boosting Machine) was used, which is a gradient boosting framework that uses tree-based learning algorithms. It builds decision trees to learn patterns in the data, and each new tree corrects the errors of the previous ones. It is well-suited for handling large-scale data, has a faster training speed and higher efficiency compared to other frameworks (lightgbm.readthedocs.io, n.d.).

4. Results

4.1. LightGBM for Multiclass Classification

Label	Precision	Recall	F1 – Score	Support
CD	0.54	0.42	0.47	2456
HYP	0.41	0.12	0.18	1054
MI	0.47	0.19	0.27	2409
NORM	0.75	0.9	0.82	9426
STTC	0.64	0.75	0.69	5060
Accuracy			0.68	20405
MacroAvg	0.56	0.48	0.49	20405
WeightedAvg	0.65	0.68	0.65	20405

Table 1. Classification Report of LightGBM. This table presents the precision, recall, and F1-score for each class along with the model accuracy for multiclass classification.

The LightGBM model achieved an accuracy of 68%, which correctly predicts the class for most of the ECG samples. The Normal class had the best performance, with 75% precision, 90% recall, and an F1 score of 82%. STTC was the next best performing class with 64% precision, 75% recall, and an F1-score of 69%. Conduction Disturbance had a moderate precision (54%), recall (42%), and F1-Score (47%), indicating the model's performance at identifying this class was average. MI and Hypertrophy had poor performance with F1-Scores of 27% and 18% respectively.

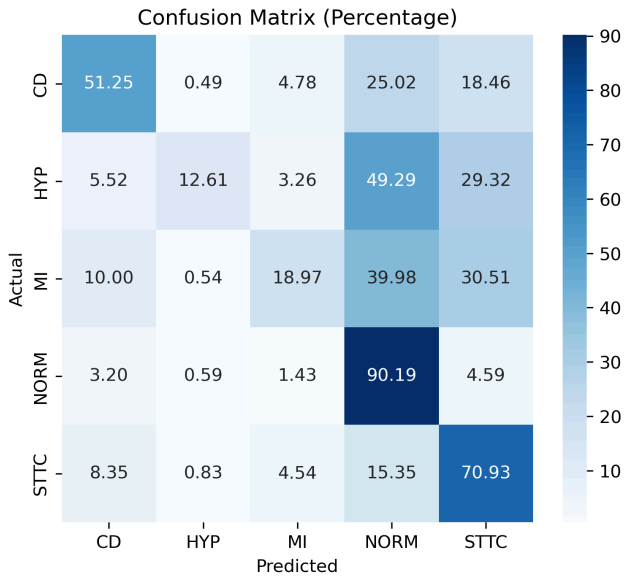

Figure 7. Confusion Matrix of Multiclass Classification showing actual vs predicted values for each class.

The performance of the LightGBM classification model was further evaluated using a confusion matrix, which gives us an insight into the true vs predicted class labels for all superclasses.

- Normal ECGs were the most accurately classified, with 90.19% of instances correctly predicted.
- STTC was correctly predicted in 70.93% of cases.
- Conduction Disturbances were accurately classified in 51.25% of cases.
- Myocardial Infarction had only 18.97% correct classifications.
- Hypertrophy had the lowest true positives, with only 12.61% of cases correctly predicted.

Misclassifications were also observed:

- Normal ECGs were misclassified, particularly as STTC (4.59%) and Conduction Disturbance (3.2%).
- STTC cases were misclassified as Normal (15.35%) and Conduction Disturbance (8.35%).
- Conduction Disturbances were incorrectly labelled as Normal (25.02%) and STTC (18.46%).
- Myocardial Infarction cases were often misclassified as Normal (39.98%) and STTC (30.51%).
- Hypertrophy was predominantly misclassified as Normal (49.29%) and STTC (29.32%).

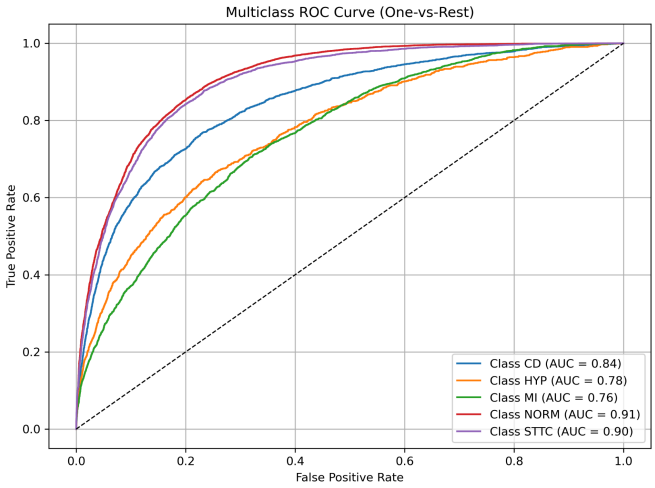

Figure 8. ROC curves for the multiclass classification model. This figure illustrates the performance of the model across multiple cardiac conditions, showing the true positive rate vs the false positive rate for each class.

Area Under the Curve (AUC) is the area under the Receiver Operating Characteristic (ROC) Curve. It measures how well the model can distinguish between classes on a scale from 0 to 1. The AUC values were relatively high for all classes: 0.91 for Normal, 0.9 for STTC, 0.84 for Conduction Disturbance, 0.78 for Hypertrophy, and 0.76 for MI. These values indicate that, despite the low precision and recall for some classes (Hypertrophy and MI), the model still has a strong ability to differentiate between classes.

4.2. LightGBM for Normal vs Abnormal Classification

Label	Precision	Recall	F1 – Score	Support
Abnormal	0.8	0.84	0.82	8473
Normal	0.86	0.82	0.84	9884
Accuracy			0.83	18357
MacroAvg	0.83	0.83	0.83	18357
WeightedAvg	0.83	0.83	0.83	18357

Table 2. Classification Report of LightGBM. This table presents the precision, recall, and F1-score for each class along with the model accuracy for binary classification.

In addition to the multiclass classification, the model was evaluated to perform binary classification into Normal and Abnormal categories. The LightGBM model achieved an overall accuracy of 83%, reflecting strong performance in distinguishing between healthy and diseased ECG patterns. Normal ECGs were identified with a precision of 86%, a recall of 82%, and an F1-score of 84%, meaning that the majority of predictions labeled as Normal were accurate. Abnormal ECGs had a precision of 80%, a recall of 84%, and an F1-score of 82%. The precision was slightly lower compared to the Normal class, suggesting a higher false-positive rate. The macro and weighted averages converge at 0.83, demonstrating consistent performance in both classes.

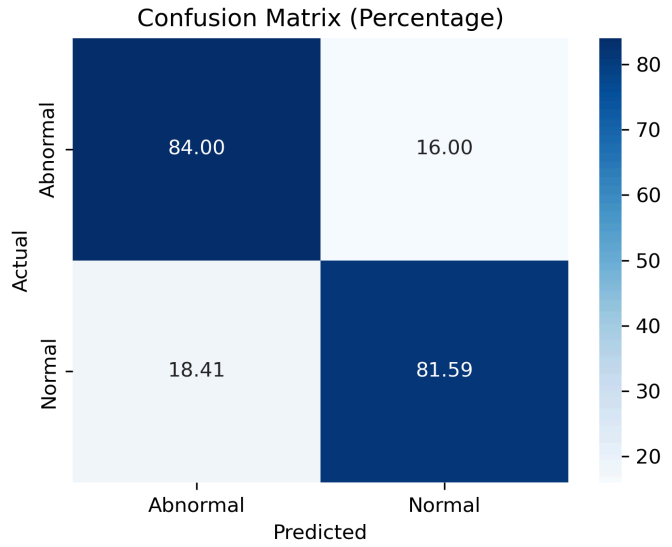

Figure 9. Confusion Matrix of Binary Classification showing actual vs predicted values for each class.

The confusion matrix for binary classification gave us deeper insights into the model's behaviour. The model correctly classified 81.59% of Normal ECGs and 84% of Abnormal ECGs. It produced 16% of false negatives, where Abnormal ECGs were misclassified as Normal, and 18.41% of false positives, where Normal ECGs were mislabelled as Abnormal.

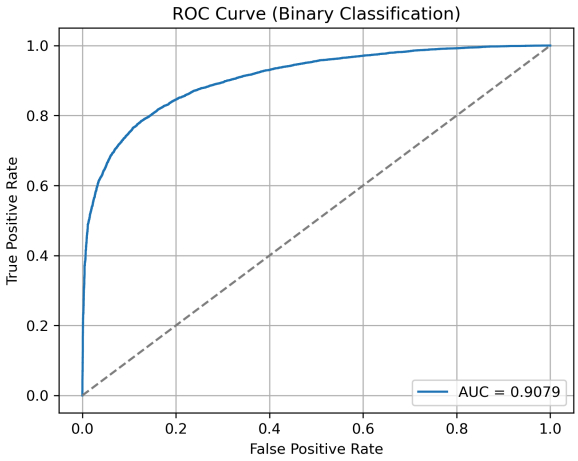

Figure 10. ROC curves for the binary classification model. This figure illustrates the performance of the model across normal vs abnormal ECGs, showing the true positive rate vs the false positive rate.

The binary classification model achieved an AUC of 0.90, indicating excellent overall performance. This high AUC value demonstrates the model's strong ability to differentiate between the two classes across various threshold settings.

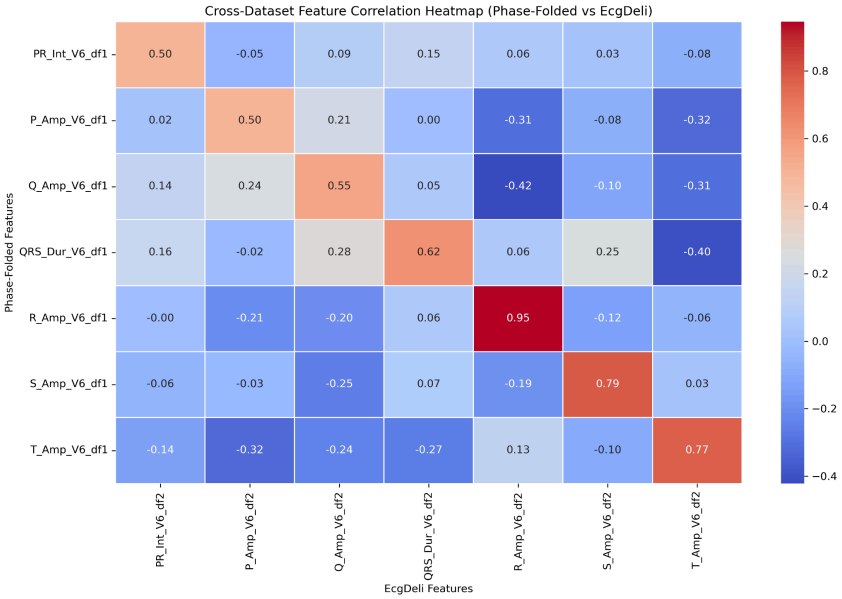

Figure 11. Cross-correlation heatmap comparing lead V6 features extracted using the Phase-Folded method and EcgDeli (from the PTB-XL dataset). High correlations for R (0.95), S (0.79), and T (0.77) amplitudes demonstrate strong agreement between methods.

EcgDeli is a set of standardized ECG features introduced in the PTB-XL paper by Wagner et al. (2020), designed to support structured waveform analysis for machine learning. A cross-correlation analysis was conducted to assess the alignment between Phase-Folded and EcgDeli features. The highest correlation was observed for the R Amplitude (0.95), followed by the S (0.79) and T amplitudes (0.77). These results confirmed consistency between the two methods of feature extraction.

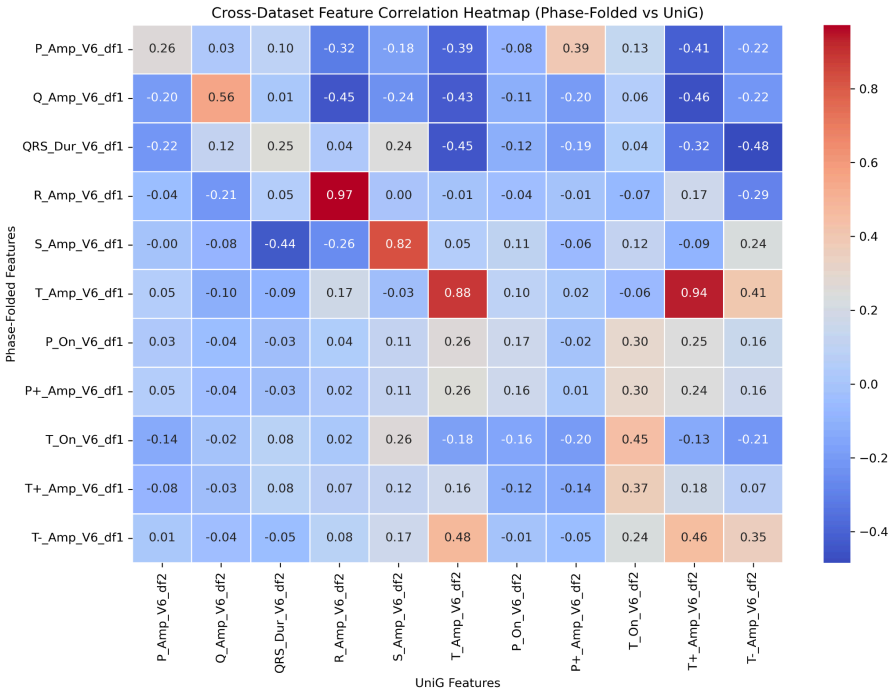

Figure 12. Cross-correlation heatmap comparing lead V6 features extracted using the Phase-Folded method and UniG (from the PTB-XL dataset). High correlations for R (0.97), S (0.82), and T (0.88) amplitudes demonstrate strong agreement between methods.

Another cross-correlation analysis was performed to assess the alignment between features extracted using the Phase-Folded method and those from the UniG feature set. UniG is an independently derived set of ECG features, distinct from EcgDeli, and includes clinically relevant waveform parameters such as amplitudes and timings across different ECG segments. High positive correlations were observed between corresponding features, such as R amplitude (0.97), T amplitude (0.88), and S amplitude (0.82). These results demonstrate a strong degree of agreement between the two methods, reinforcing the validity of the Phase-Folded approach and its ability to extract meaningful cardiac features comparable to those obtained through established pipelines like UniG.

5. Discussion

This study evaluates the effectiveness of phase-folded ECG waveform parameterization in identifying diagnostic patterns. In multiclass classification, the model showed the strongest performance in detecting Normal ECGs and ST-T Changes, which could be due to the clearer morphological features these classes present, like consistent PQRST waves or T wave deviations. STTC has distinct patterns in the T wave, which is why the model was able to distinguish these better than the other disease classes. The highest F1-score was presented by the Normal class (82%), followed by STTC (69%), highlighting their greater separability in the feature space. The model presented moderate difficulty in classifying Conduction Disturbances, resulting in an F1-score of 47%. Conversely, the classification performance for Myocardial Infarction and

Hypertrophy was considerably lower, with F1-scores of 27% and 18%, respectively. These results emphasize that the model struggles with separating these conditions from others. A major contributing factor is class imbalance; hypertrophy had the fewest samples in the dataset, which possibly limited the model's ability to generalize its ECG patterns. Another contributing factor is the overlapping waveform features between classes, making it challenging to distinguish between certain cardiac conditions in both automated and manual analysis. For instance, MI and STTC both present with ST-segment deviations and abnormal T waves, while MI and Conduction Disturbances may share features such as a widened QRS and an abnormal ST-T morphology (Deshpande, 2014). The AUC values for multiclass classification were quite high, ranging from 0.76 to 0.91, indicating strong model performance in distinguishing between different classes. This means adjusting the classification thresholds could help balance the precision and recall for individual classes, potentially enhancing the accuracy and F1-scores of the model.

The binary classification demonstrated better performance compared to the multiclass classification, with an overall accuracy of 83%. By reducing the number of categories, the model becomes less complex and reduces the risk of misclassifications due to shared waveform features. The precision and recall scores were balanced across both classes, meaning the slight class imbalance did not affect model performance. The binary model achieved an AUC value of 0.9, which means that the model has a 90% chance of correctly identifying Normal or Abnormal ECGs, even in the presence of overlapping features. The high AUC value is substantial for clinical settings as it indicates a reduced likelihood of a missed diagnosis.

A similar study conducted by Smigiel et al. (2021) portrayed similar findings with an accuracy of 88.2% for binary classification and 72% for 5-class classification. These results were achieved by using 12-lead ECGs with convolutional neural networks. Another study conducted by Palczynski et al. (2022) conveyed parallel findings with an accuracy of 88.9% for binary classification and 75.2% for multiclass classification using FSL and XGBoost. These findings highlight that our approach of phase-folding and feature extraction provides similar results to studies employing different methods.

Both classification models have their clinical advantages and disadvantages. The multiclass classification model provides a comprehensive diagnosis by detecting multiple cardiac conditions. However, the increased number of classes introduces a higher risk of misclassification. On the contrary, binary classification offers a faster, more accurate diagnosis by distinguishing between Normal and Abnormal ECGs. This model is less prone to misclassification and achieves higher accuracy, making it a practical screening tool for early diagnosis.

The limitations should be considered when interpreting the results of this study. By averaging the ECG waveform through the phase-folding process, some details may be lost. The technique may potentially suppress subtle abnormalities such as ectopic beats (extra or premature heartbeats). This could result in the model misinterpreting features that are useful for diagnosing certain cardiac conditions. Since phase-folding relies on periodicity, cases of irregular

heart rhythms or arrhythmias can result in incorrect alignment of the ECG signal. The technique relies on consistent RR intervals for alignment; however, arrhythmias present variable RR intervals, which will disrupt this process (Soos and McComb, 2020). This can mask clinically significant characteristics and reduce the model's ability to detect certain cardiac conditions.

Despite its limitations, phase-folding offers several advantages for ECG signal analysis. This technique involves averaging the signal, which eliminates any random noise, resulting in a clearer waveform. Phase-folding reduces the data size vastly by producing a single representative ECG, rather than multiple segments. This saves computational time and complexity during feature extraction and model training, especially for large datasets. Since each ECG is standardized to one cardiac cycle, it becomes easier to compare patterns across patients or disease groups. Cross-correlation was performed with other studies, such as ECGdeli and UniG, and our phase-folded results show similar findings (Nils Strodthoff et al., 2023).

Lead V6 was used for our analysis, as it offers several diagnostic advantages that make it suitable for ECG interpretation. This lead is placed on the left side of the chest, which makes it less susceptible to noise and movement-related artefacts compared to the limb leads. The QRS complex and T wave are well defined in V6, making it ideal for assessing duration and amplitudes. V6 has also been proven to capture unique characteristics that may be missed by other leads (Ramirez et al., 2024). Using a single lead simplifies the model as it reduces the number of input features, which in turn lowers computational complexity. By using the same lead across all ECG IDs, there is less variability in the signal morphology caused by lead placement, resulting in a more effective training process. V6 is well-suited for out-of-clinic diagnosis as it's simpler to deploy than a 12-lead ECG. It can be used in wearable ECG devices for monitoring the electrical activity of the heart. Despite its advantages, using only lead V6 may reduce the diagnostic accuracy, as some conditions are better detected in other leads. For instance, an inferior myocardial infarction is typically identified in leads II, III, and aVF, while changes in V6 are not too prominent (Lome, n.d).

Several improvements can be made to enhance the model’s performance. To address class imbalance in the dataset, techniques such as undersampling the majority class (Normal) or oversampling the minority classes (HYP and MI) may help to improve the model’s classification. Fine-tuning the classification thresholds, as suggested by the ROC curves, may increase F1-scores for certain superclasses, promoting consistent performance across classes. Incorporating an extra lead that provides a different view of the heart may help in identifying conditions that are missed by lead V6. Additionally, more parameters, such as the P offset, durations of the P and T waves, and the area under the waveform segments, could be calculated to expand the feature set. Although feature engineering was performed, further assessing feature importance and removing less informative ones may improve model accuracy. More patient features, such as BMI, smoking status, and history of cardiac events, would strengthen the model's ability to make accurate predictions. If we were to create a wearable device or mobile application linked to the model, the patient could self-report symptoms such as chest pain, shortness of breath, or heavy breathing, supporting more personalised out-of-clinic diagnosis.

6. Conclusion

The findings of this project demonstrate that combining ECG waveform parameterisation with machine learning results in a clinically efficient method for automated ECG classification. By applying phase-folding to standardize ECG signals from lead V6, the method simplifies ECG analysis without compromising diagnostic accuracy. The binary classification model outperformed the multiclass model with higher accuracy, making it suitable as a quick screening tool, whereas the multiclass model provides more detail on the specific cardiac condition. This approach is clinically relevant, especially for out-of-clinic settings where quick, understandable results are essential. Future improvements, such as handling class imbalance and including additional patient features, may further enhance the diagnostic performance of the model.

References

AIMCARDIO (2020). ▷ *12 Lead Placement guide with diagram [VIDEO]*. [online] Used Medical Equipment We specialize in stress Systems. Available at: https://aimcardio.com/blog/12-lead-placement-guide-with-diagram/ [Accessed 28 April 2025].

Deshpande, A. (2014). ST-segment elevation: Distinguishing ST elevation myocardial infarction from ST elevation secondary to nonischemic etiologies. *World Journal of Cardiology*, [online] 6(10), p.1067. doi:https://doi.org/10.4330/wjc.v6.i10.1067 [Accesed 3 May 2025].

Institute for Quality and Efficiency in Health Care (2023). *What is an electrocardiogram (ECG)?* [online] *www.ncbi.nlm.nih.gov*. Institute for Quality and Efficiency in Health Care (IQWiG). Available at: https://www.ncbi.nlm.nih.gov/books/NBK536878/. [Accessed 28 April 2025].

Cleveland Clinic, 2024. *Heart Attack: Symptoms & Treatment*. [online] Available at: https://my.clevelandclinic.org/health/diseases/16818-heart-attack-myocardial-infarction [Accessed 28 April 2025].

ECGpedia, 2024. *Chamber Hypertrophy and Enlargement*. [online] Available at: https://en.ecgpedia.org/index.php?title=Chamber_Hypertrophy_and_Enlargment [Accessed 28 April 2025].

ECG Waves, 2018. *The ST segment: physiology, normal appearance, ST depression & ST elevation*. [online] Available at: https://ecgwaves.com/st-segment-normal-abnormal-depression-elevation-causes/ [Accessed 28 April 2025]

Kwon, O., Jeong, J., Kim, H.B., Kwon, I.H., Park, S.Y., Kim, J.E. and Choi, Y.,(2018). *Electrocardiogram Sampling Frequency Range Acceptable for Heart Rate Variability Analysis*. *Healthcare Informatics Research*, [online] 24(3), p.198. Available at: https://doi.org/10.4258/hir.2018.24.3.198 [Accessed 30 April 2025].

lightgbm.readthedocs.io. (n.d.). *Welcome to LightGBM's documentation! — LightGBM 3.3.5 documentation*. [online] Available at: https://lightgbm.readthedocs.io/en/stable/ [Accessed 2 May 2025]

Lome, S. (n.d.). *Top 5 MI ECG Patterns You Must Know*. [online] www.healio.com. Available at: https://www.healio.com/cardiology/learn-the-heart/ecg-review/ecg-interpretation-tutorial/stemi-mi-ecg-pattern. [Accessed 3 May 2025]

Madona, P., Basti, R.I. and Zain, M.M. (2021). PQRST wave detection on ECG signals. *Gaceta Sanitaria*, [online] 35(2), pp.S364–S369. doi:https://doi.org/10.1016/j.gaceta.2021.10.052 [Accessed 28 April 2025]

National Heart, Lung, and Blood Institute (NHLBI), 2024. *Arrhythmias - Conduction Disorders*. [online] Available at: https://www.nhlbi.nih.gov/health/conduction-disorders [Accessed 28 April 2025].

Nils Strodthoff, Mehari, T., Nagel, C., Aston, P., Sundar, A., Graff, C., Jørgen Kanters, Haverkamp, W., Olaf Doessel, Loewe, A., Bär, M. and Schaeffter, T. (2023). *PTB-XL+, a comprehensive electrocardiographic feature dataset*. [online] Physionet.org. Available at: https://physionet.org/content/ptb-xl-plus/1.0.1/features/#files-panel [Accessed 27 April 2025].

Pałczyński, K., Śmigiel, S., Ledziński, D. and Bujnowski, S. (2022). Study of the Few-Shot Learning for ECG Classification Based on the PTB-XL Dataset. *Sensors*, 22(3), p.904. doi:https://doi.org/10.3390/s22030904. [Accessed 4 May 2025]

Ramirez, E., Ruiperez-Campillo, S., Casado-Arroyo, R., Merino, J.L., Vogt, J.E., Castells, F. and Millet, J. (2024). The art of selecting the ECG input in neural networks to classify heart diseases: a dual focus on maximizing information and reducing redundancy. *Frontiers in Physiology*, 15. doi:https://doi.org/10.3389/fphys.2024.1452829. [Accessed 1 May 2025]

scikit learn (2009). *3.1. Cross-validation: Evaluating Estimator Performance — scikit-learn 0.21.3 Documentation*. [online] Scikit-learn.org. Available at: https://scikit-learn.org/stable/modules/cross_validation.html [Accessed 3 May, 2025]

scikit-learn.org. (n.d.). *sklearn.impute.KNNImputer — scikit-learn 0.23.1 documentation*. [online] Available at: https://scikit-learn.org/stable/modules/generated/sklearn.impute.KNNImputer.html [Accessed 5 May 2025]

SciPy, 2024. *scipy.signal.find_peaks*. SciPy v1.11.4 Reference Guide. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html [Accessed 30 April 2025].

Śmigiel, S., Pałczyński, K. and Ledziński, D. (2021). ECG Signal Classification Using Deep Learning Techniques Based on the PTB-XL Dataset. *Entropy*, 23(9), p.1121. doi:https://doi.org/10.3390/e23091121. [Accessed 5 May 2025]

Soos, M.P. and McComb, D. (2020). *Sinus Arrhythmia*. [online] PubMed. Available at: https://www.ncbi.nlm.nih.gov/books/NBK537011/ [Accessed 3 May 2025]

Wagner, P., Strodthoff, N., Bousseljot, R.D., Kreiseler, D., Lunze, F.I., Samek, W. and Schaeffter, T., 2020. *PTB-XL, a large publicly available electrocardiography dataset*. Scientific Data, 7(1), p.154. Available at: https://www.nature.com/articles/s41597-020-0495-6 [Accessed 29 April 2025].

Wang, K., Ge, J., Willis, K., Wang, K. and Zhao, Y. (2024). The GPU phase folding and deep learning method for detecting exoplanet transits. *Monthly Notices of the Royal Astronomical*

Society, [online] 528(3), pp.4053–4067. doi:https://doi.org/10.1093/mnras/stae245. [Accessed 29 April 2025].

World Health Organisation (2019). *Cardiovascular diseases*. [online] Who.int. Available at: https://www.who.int/health-topics/hypertension/cardiovascular-diseases?utm_source=chatgpt.com [Accessed 27 April 2025]