Electrocardiogram Signal and Interpretation Through AI: Developing a Diagnostic Tool for Precision Cardiovascular Diagnosis

Ryan M. Levinter

Abstract

When reading electrocardiogram (ECG) scans, diagnostic error is a major source of medical errors, often resulting in postponed or neglected treatment of cardiovascular disease along with the risk of using incorrect medication and amplifying or introducing additional risks. This study addresses these challenges by developing an artificial intelligence (AI) model for automated ECG classification.

A hybrid deep learning architecture that integrates three complementary neural network components was developed and trained on publicly available data from the PTB-XL dataset from PhysioNet. The model culminates into a one-dimensional convoluted neural network (1D-CNN) for feature extraction, a long short-term memory (LSTM) network for temporal sequence modeling, and a temporal attention mechanism to highlight diagnostically significant signal segments. It was trained to classify ECGs into five categories: Normal, Myocardial Infarction, ST/T change, Conduction Disturbance, and Hypertrophy. If an ECG was found to have multiple diagnoses, it was sorted into multiple categories.

The completed model achieved an accuracy of 88.1% and a micro-average area under the receiver operating characteristic curve (AUC-ROC) of 0.936. Performance was strongest for Normal rhythms (Accuracy: 86.3% AI compared to 54% human- F1: 0.867) (Cook et al., 2020), Myocardial Infarction (Accuracy: 87.8%- F1: 0.761), and Conduction Disturbance (Accuracy: 90.1%- F1: 0.767). Performance on Hypertrophy (Accuracy: 88.5%- F1: 0.448) was limited, primarily due to low recall which is a recognized limitation of ECG-based diagnosis for this condition.

To demonstrate a pathway for clinical integration, a functional prototype ECG device was designed and constructed using an AD8232 sensor, Arduino Nano, and 3D printing for fabrication of parts. This custom hardware was successfully integrated with the AI model through a locally hosted web interface, forming a complete end-to-end system for real-time signal acquisition and diagnosis.

These results demonstrate that the model achieves accuracy comparable to performance achieved by humans. It demonstrates potential due to its consistency and scalability, which could reduce diagnostic variability and improve diagnostic accuracy across the board.

1. Introduction

1.1 Clinical Context and Problem Statement

Cardiovascular disease (CVD) remains the leading cause of mortality worldwide, representing an urgent public health crisis that demands improved diagnostic capabilities (Heart Disease Facts, 2024; World Health Organization: WHO, 2025). The electrocardiogram (ECG) serves as a cornerstone of cardiovascular diagnosis, providing non-invasive visualization of cardiac electrical activity. Despite its clinical importance, ECG interpretation suffers from significant human error, with nearly 39% of ECGs being misdiagnosed due to physician fatigue, time pressure, and interpretive variability (Kraik et al., 2025). These diagnostic errors can lead to missed treatment opportunities, unnecessary procedures, and patient harm, highlighting the critical need for tools that provide consistent, objective analysis.

1.2 The Role of Artificial Intelligence in Cardiology

Artificial intelligence (AI) has emerged as a transformative solution in healthcare. In cardiology, AI models have shown remarkable promise in ECG analysis, detecting specific abnormalities like atrial fibrillation with accuracies comparable to expert cardiologists (Zhang et al., 2024). Furthermore, studies have revealed AI's ability to identify subtle patterns imperceptible to human experts, such as detecting asymptomatic left ventricular dysfunction from normal sinus rhythm ECGs (Attia et al., 2019). This trend toward "high-performance medicine" (Topol, 2019) represents a fundamental shift where AI can augment human capabilities and reduce diagnostic variability.

1.3 Limitations of Existing Models and Research Gap

Most existing AI models for ECG analysis focus on binary classification tasks, distinguishing between normal and abnormal rhythms or identifying single diseases. This approach fails to reflect clinical reality, where practitioners must differentiate among multiple cardiac pathologies with overlapping features. A diagnostic tool restricted to detecting only one condition has limited clinical utility, creating a significant gap between prototype AI models and clinically useful diagnostic tools.

1.4 Proposed Approach and Model Architecture

This study addresses these limitations through a hybrid deep learning architecture for multi-class ECG interpretation. The model integrates three complementary components: (1) a 1D-Convolutional Neural Network (CNN) for extracting localized morphological features from raw signal data; (2) a Long

Short-Term Memory (LSTM) network for modeling temporal dependencies and rhythmic patterns; and (3) a temporal attention mechanism that dynamically weights diagnostically relevant signal segments.

1.5 Diagnostic Categories and Dataset

The objective is to classify ECG signals into five clinically relevant categories: Normal Rhythm, Myocardial Infarction, ST/T Change, Conduction Disturbance, and Hypertrophy. Representative examples of these conditions from the PTB-XL dataset are shown in Figures 1 and 2. An electrocardiogram (ECG) is a diagnostic tool which uses a series of leads to capture a waveform of the heart's activity, comprising P waves, QRS complexes, and T waves to assess cardiac rhythm and function.

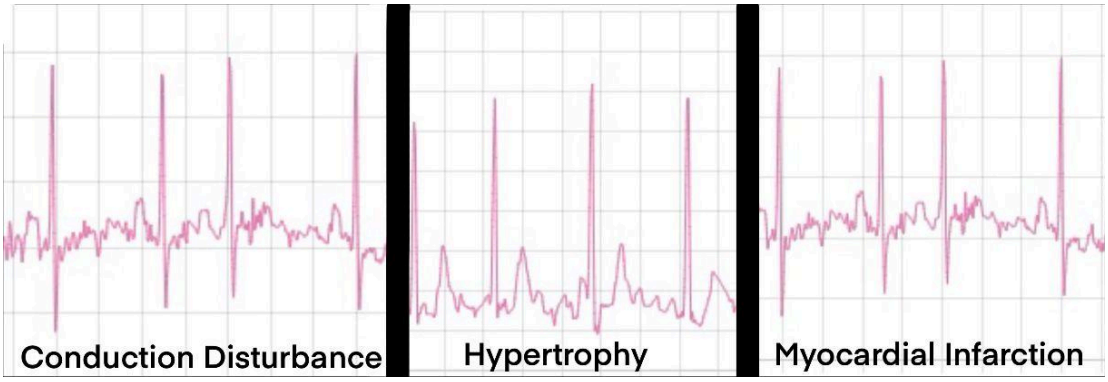


Figure 1: An ECG Scan for Conduction Disturbance (leftmost), Hypertrophy (middle), and Myocardial Infarction (rightmost). Data from PTB-XL was loaded into an interpreter and displayed. Created by student researcher.

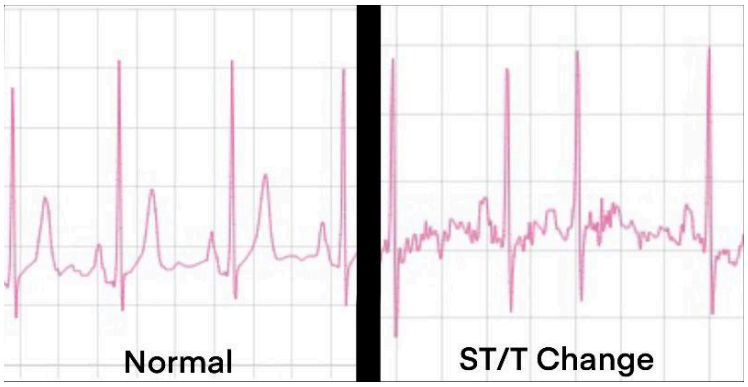


Figure 2: An ECG Scan for a normal scan (leftmost) and ST/T Change (rightmost). Data from PTB-XL was loaded into an interpreter and displayed. Created by student researcher.

1.6 Study Rationale, Objectives, and Hypothesis

The development of a custom ECG device is motivated by the need to demonstrate a practical, end-to-end pathway from physical signal acquisition to AI-powered diagnosis, moving beyond a purely software-based prototype. Furthermore, existing AI models often lack clinical utility by focusing on binary classification, creating a gap that this study addresses through a multi-class approach.

The primary objectives of this study are to: (1) develop and train a hybrid AI model for multi-class ECG interpretation; (2) design and build a functional prototype ECG device for signal acquisition; and (3) integrate the model and hardware into a unified system via a web interface for real-time diagnosis.

The researcher hypothesized that a hybrid deep learning model incorporating convolutional, recurrent, and attention mechanisms will achieve diagnostic accuracy comparable to human performance on a multi-class ECG classification task.

2. Methods

2.1. Data Acquisition and Preprocessing

ECG scans were obtained from the publicly available PTB-XL database hosted on PhysioNet (Wagner et al., 2020). The PTB-XL dataset contains 21,799 records from 18,869 patients, meaning some patients contributed multiple ECG recordings which allow the model to be trained to understand multiple-conditions using 12-lead ECGs. Each recording is 10 seconds in duration and sampled at both 500 Hz and 100 Hz, which defines the temporal resolution required when acquiring new scans.

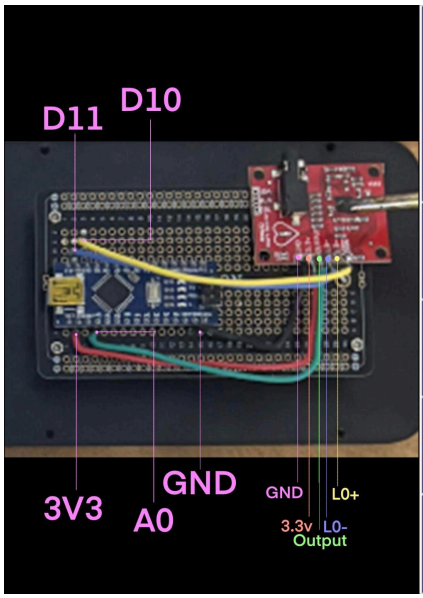
For this study, the researcher restricted analysis to the five most prevalent diagnostic super classes within the dataset's structured hierarchy: Normal ECG, Myocardial Infarction, ST/T Change, Conduction Disturbance, and Hypertrophy.

A two-step normalization procedure was applied: DC offset removal by centering each lead to zero mean, followed by Z-score normalization to unit variance. This standardization accelerates model convergence and ensures feature detectors respond to relative morphological shapes rather than absolute voltage values.

2.2. Electrocardiogram Hardware Design

To enable real-time data acquisition and demonstrate a pathway for clinical integration, a prototype ECG device was built using off-the-shelf components. It is critical to contextualize this prototype by comparing it to a standard 12-lead clinical ECG. A diagnostic-grade 12-lead ECG uses 10 electrodes (placed on the limbs and chest) to derive 12 different electrical "views" of the heart. This comprehensive perspective is essential for precise anatomical localization of abnormalities, such as determining the specific region of a myocardial infarction.

The prototype developed for this study utilizes an AD8232 ECG sensor module, which is not recommended for clinical use, but usable for a prototype (Dias et al., 2022). The AD8232 was chosen for its accessibility and robust documentation for prototyping. To read the signal, an Arduino Nano microcontroller was selected due to its dedicated analog ports, small footprint, and low power consumption, which minimizes noise compared to a system requiring an external analog-to-digital converter (ADC). The researcher wrote out the wiring diagram (Figure 3).



MC Port	AD8232 Port
3V3	3.3V
GND	GND
D10	LO+
D11	LO-
A0	OUTPUT

Figure 3: Wiring Table for Arduino Nano and AD8232 and a diagram of the ECG wiring. Illustrates the complete circuit schematic for the prototype ECG device, showing how the AD8232 sensor, Arduino Nano, and patient electrodes are interconnected to form a functional unit. The electrical connections required to interface the AD8232 ECG sensor with the Arduino Nano microcontroller, ensuring correct voltage, ground, and signal routing for data acquisition. Created by student researcher.

To minimize interference, the prototype was enclosed in a custom 3D-printed case designed in multiple components (Figures 4 through 6). The base was printed in polyethylene terephthalate glycol (PETG) for flexibility and includes built-in standoffs to secure the breadboard. The walls were printed using polyethylene terephthalate glycol plus (PLA+), chosen for its rigidity and ability to withstand high-speed printing during the prototyping phase. The lid used PETG for chemical resistance in case something were to spill. Finally, a thermoplastic polyurethane (TPU) skid-plate was printed around the base to increase friction and reduce slipping during use. The skid-plate, base, and walls feature screw holes, allowing four screws to hold them together, while the top of the walls features inlaid neodymium magnets which allow the lid inlaid with opposing magnets to stay in place.

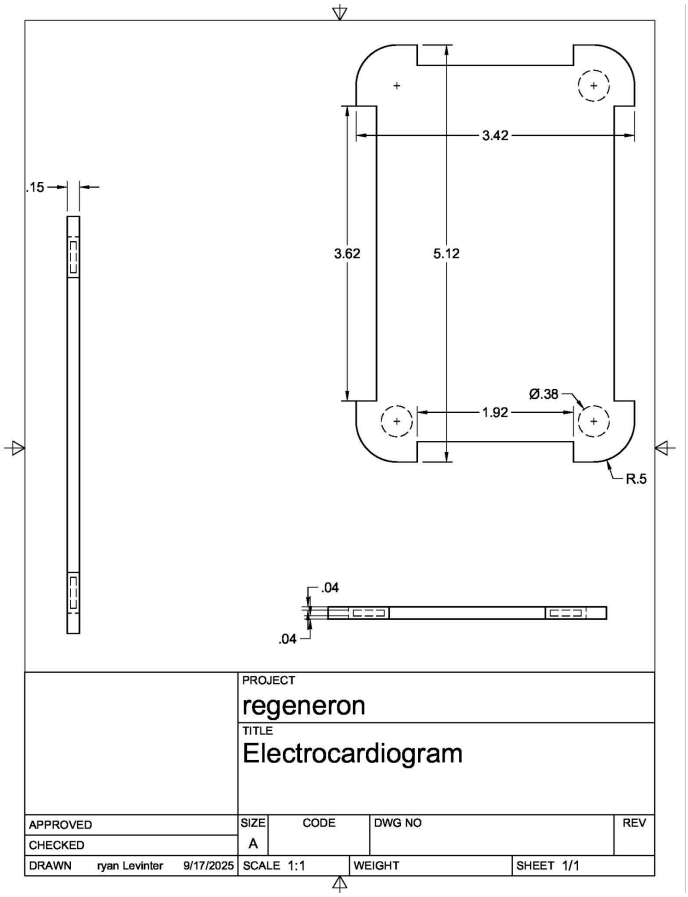


Figure 4: The schematics for the top of the ECG encasement. Shows the design of the enclosure lid, which uses inlaid neodymium magnets to secure it to the walls, allowing for easy access to the internal components. Created by student researcher.

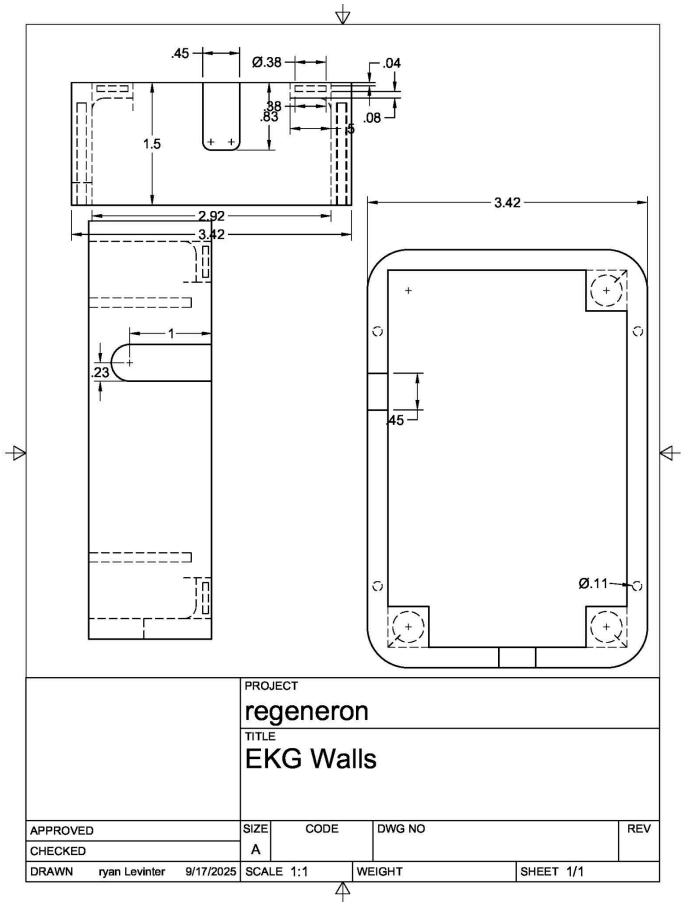


Figure 5: The schematics for the walls of the ECG encasement. Displays the walls of the enclosure, which includes integrated magnets to connect to the lid as well as screw holes to connect to the base and skidplate. Created by student researcher.

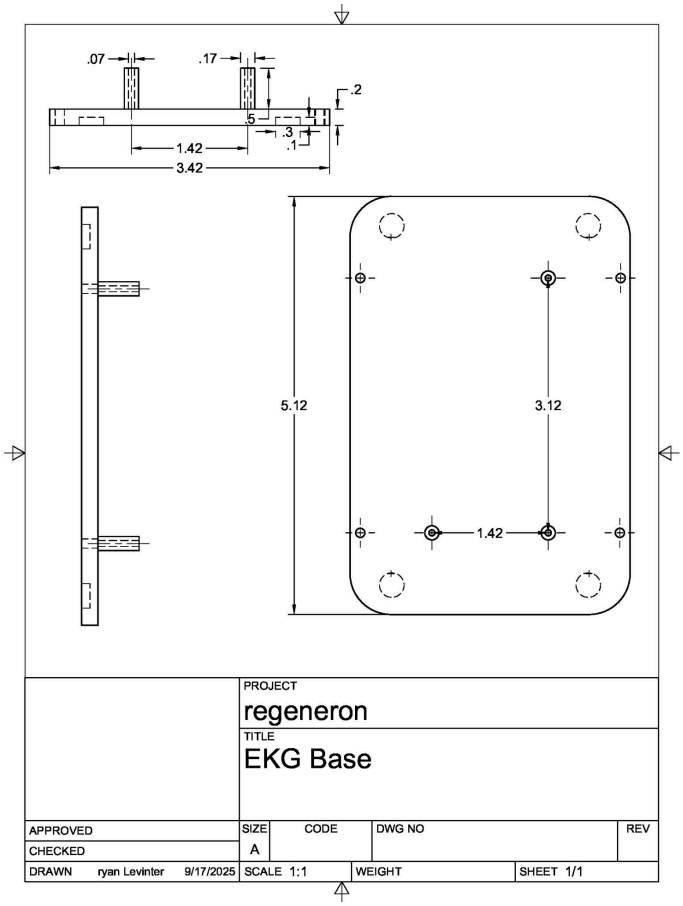


Figure 6: The schematics for the bottom of the ECG encasement. Displays the base of the enclosure, which includes integrated standoffs to firmly secure the internal printed circuit board and prevent movement during use. Created by student researcher.

2.3. Device Enclosure and Fabrication

To house the electronic components, a custom enclosure was designed (Figures 4 through 6). All components were then printed using the Elegoo Centauri Carbon (Figure 7). This printer was chosen due to its ability to print at high speeds, making prototyping easier, and its ability to use more difficult materials such as TPU. After the components were fabricated, the researcher assembled the components (Figure 8).

Figure 7: The walls being printed using Polylactic Acid Plus in the Elegoo Centauri Carbon. Depicts the physical fabrication process of the enclosure walls via 3D printing, using PLA+ filament for a strong and precise prototype structure. Created by student researcher.

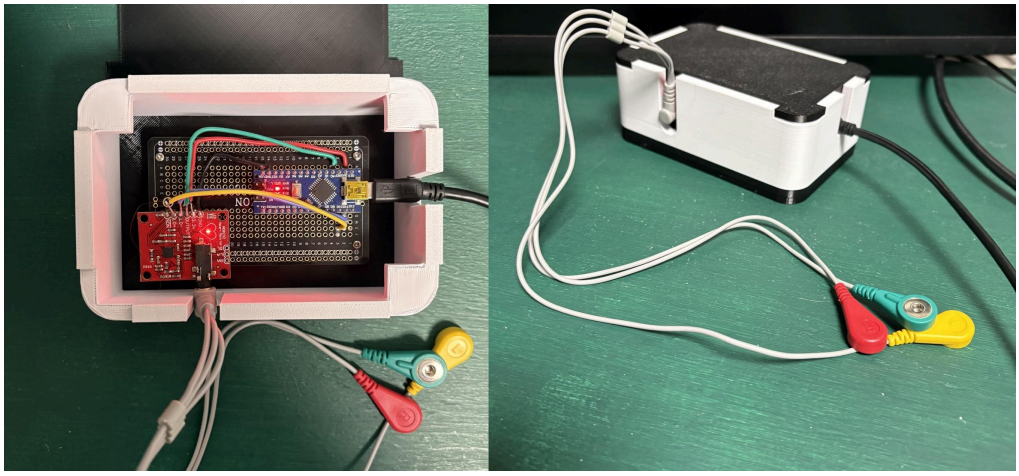


Figure 8: The assembled ECG. After the 3D printing components were completed and the breadboard was soldered, the researcher connected each piece. Created by student researcher.

2.4. Artificial Intelligence & Machine Learning Model

The diagnostic model was developed in PyTorch (Paszke et al., n.d.) as a hybrid deep learning architecture that integrates convolutional, recurrent, and attention-based components for comprehensive ECG analysis.

The PTB-XL dataset was partitioned into training (80%), validation (10%), and test (10%) subsets. This stratified partition was applied to maintain consistent class distribution across all subsets, preventing data leakage and providing an unbiased estimate of the model's generalization performance on unseen data. The model was trained exclusively on the training set, with the validation set used for hyperparameter tuning and early stopping, and the final reported results computed on the held-out test set.

The input to the model was a 2D tensor with dimensions of 1,000 × 12. This structure was directly determined by the PTB-XL dataset, where each recording is a 10-second segment sampled at 100 Hz (10 sec * 100 samples/sec = 1,000 time steps per lead). The dataset provides these 10-second segments as standardized, clinically accepted snapshots. While the initial moments of a physical ECG recording can be noisy as electrodes stabilize, the PTB-XL dataset comprises high-quality, diagnostic-grade excerpts where such artifacts have been minimized, allowing the model to learn from the full, stable rhythm.

The model architecture integrated three components in sequence. First, a 1D-CNN with two layers (32 and 64 filters, kernel size=5, padding=2, ReLU activation) extracted spatial features from the raw ECG signals. Notably, pooling layers were omitted to maintain the full 1,000-time-step resolution for subsequent temporal analysis.

The CNN output was reshaped via tensor permutation from (batch_size, 64 features, 1000 timesteps) to (batch_size, 1000 timesteps, 64 features) to create a sequence for the Bidirectional LSTM network. The LSTM (128 hidden units, bidirectional=True) processed this sequence to capture temporal patterns and rhythm dependencies in both forward and backward directions.

A Temporal Attention mechanism then computed importance weights for each of the 1,000 time steps across the LSTM's 256-dimensional outputs (128 × 2 for bidirectional). These weights were used to create a context vector as a weighted sum, focusing the model on clinically relevant segments. Finally, this context vector was passed to a fully connected layer with sigmoid activation for multi-label classification across the five diagnostic categories

The model was trained for multi-label classification using Binary Cross-Entropy Loss with Logits (BCEWithLogitsLoss) and the Adam optimizer with a learning rate of 1e-3. Training proceeded for 15 epochs with a batch size of 32. Model performance was evaluated on a held-out validation set after each epoch to monitor for overfitting.

The output of this training process was a fully trained, serialized model file containing the final learned parameters for the entire hybrid architecture. This saved model is the core diagnostic engine that can be loaded by the inference script to make predictions on new, unseen ECG data.

Training was conducted on an Intel Core i5-1035G7 laptop with 16 GB DDR4 RAM. The PyTorch framework was configured to automatically use GPU acceleration if available; however,

training proceeded on the CPU as no compatible GPU was detected. This demonstrates the model's feasibility on standard hardware without specialized accelerators.

2.5. Statistical Analysis

Model performance was evaluated using standard classification metrics: Accuracy (overall correctness), Precision (reliability of positive predictions), Recall (sensitivity to true positives), F1-Score (harmonic mean of precision and recall), and AUC-ROC (discrimination ability across thresholds).

Additionally, a 95% confidence interval for the overall accuracy was calculated using the normal approximation method for a binomial proportion to quantify the uncertainty of the performance estimate.

2.6. Testing and Classification of Custom ECG Signals

The completed model was deployed through a locally hosted web interface. ECG signals were transmitted from the Arduino Nano to an external computer via mini-USB. The interface was built using HTML, CSS, and JavaScript, with Python and Flask used to load and run the pretrained AI model. Flask was chosen over Django for its flexibility in lightweight, modular design (GeeksforGeeks, 2025). This process is shown in Figure 9.

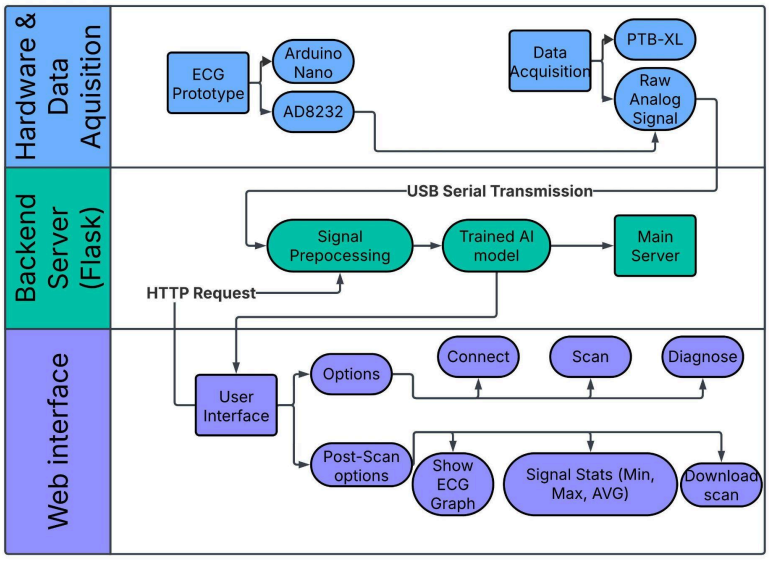


Figure 9: System architecture for the ECG diagnostic tool. Shows the three main components: the ECG hardware, the Flask server with the AI model, and the web interface, along with how data flows between them. Created by student researcher.

Figure 10 shows the user interface (ui). The ui features three buttons near the top: “Connect to Arduino” which connects to the ECG, “Begin Scan (5-second read)” which begins a scan for five seconds, and “Diagnose” which runs the scan data through the AI model and prints a diagnosis. Next to this is an indicator which changes color depending on the connection status of the ECG which automatically disconnects after scanning to avoid accidentally beginning another scan. Below the buttons and indicator is the display for the numerical values of the scan and below this displays the minimum, maximum, and average signal values to help doctors analyze the data. Below this is the display from the AI interpretation.

Figure 10: The web interface for the ECG. Shows the user-friendly web application that allows operators to connect the hardware, initiate a 5-second ECG scan, and view the AI-generated diagnosis in real-time. Created by student researcher.

3. Results

3.1. Training Progression

Training history further supports the robustness of the model. As shown in Figures 11 through 15, all recorded statistical elements improved across epochs, confirming that the training process successfully prevented overfitting, as evidenced by consistent improvement on validation metrics. The moving average of batch loss also showed smooth convergence, indicating a stable optimization process.

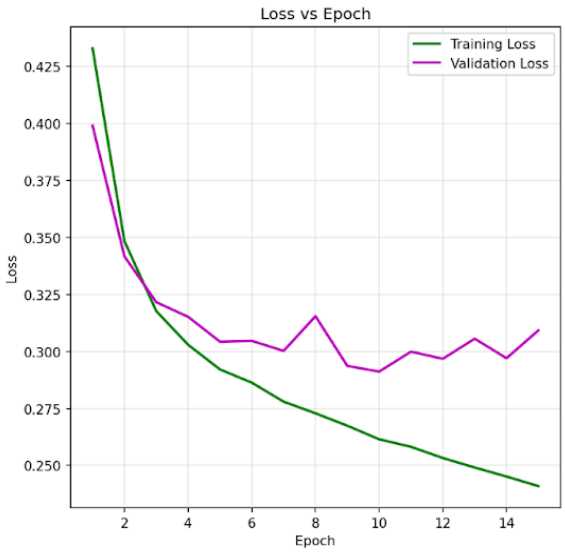


Figure 11: The progression AI model loss while training. Plots the steady decrease in training and validation loss over epochs, indicating that the model successfully learned from the data without overfitting. Created by student researcher.

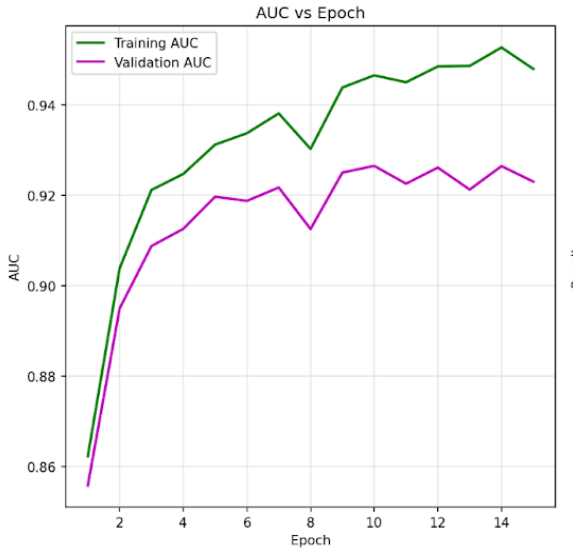


Figure 12: The progression AI model AUC while training. Shows the consistent improvement in the Area Under the Curve (AUC) on the validation set across training epochs, demonstrating the model's increasing ability to discriminate between classes. Created by student researcher.

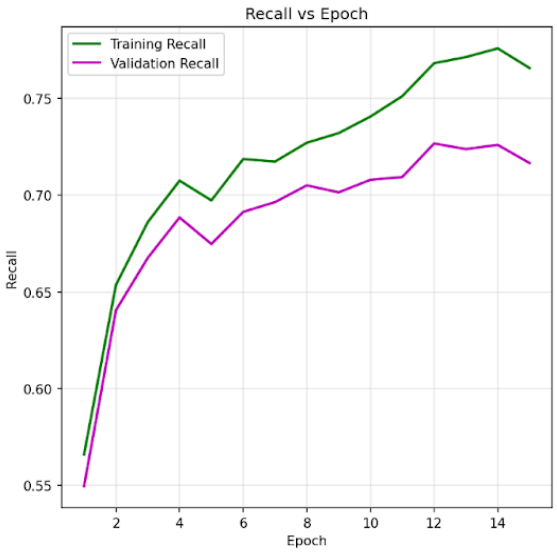


Figure 13: The progression AI model recall while training. Illustrates the model's improving sensitivity over time, showing its increasing capability to correctly identify true positive cases for each cardiac condition during training. Created by student researcher.

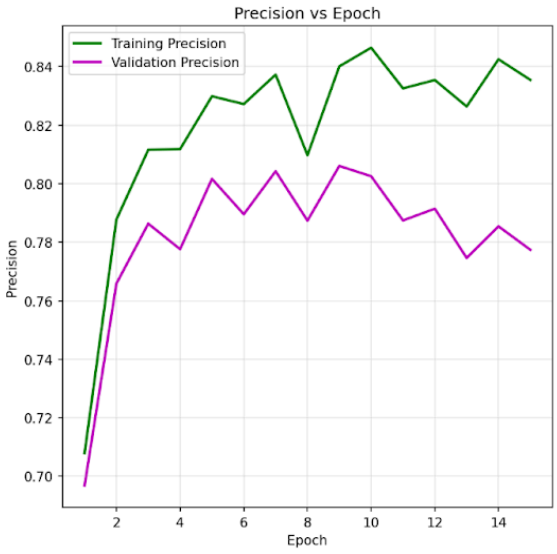


Figure 14: The progression AI model precision while training. Charts the increase in precision over epochs, reflecting the model's improving accuracy whenever it predicts a positive case for a specific cardiac condition. Created by student researcher.

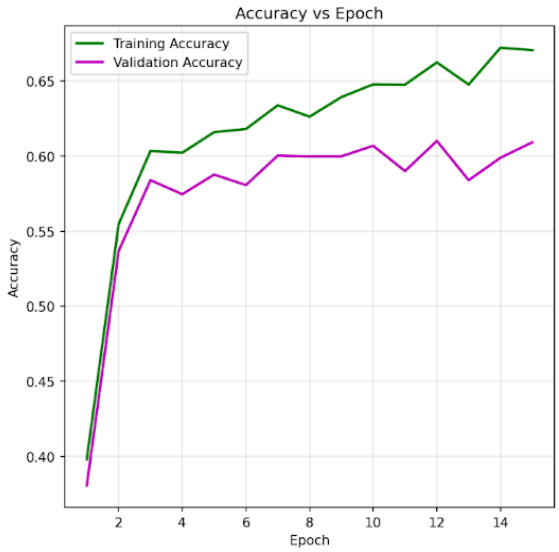


Figure 15: The progression AI model accuracy while training. Displays the rising training and validation accuracy over successive epochs, confirming the model's overall learning progression and stabilization. Created by student researcher.

This progression provides confidence in the model's ability to generalize beyond the training data, reinforcing the reliability of its performance on the held-out test set.

3.2. Model Performance

The model's overall accuracy of 88.1% was accompanied by a 95% confidence interval of [0.867, 0.895], indicating high precision in the performance estimate. Furthermore, the model's accuracy was statistically significant compared to the baseline no-information rate of 48.7% (one-sample proportion z-test, $z = 59.4$, $p < 0.001$), confirming it learned diagnostic patterns beyond simple class frequency.

Given the class imbalance in the dataset, where the 'NORM' class was most frequent, the F1-score was prioritized for per-class evaluation to manage the compromise between precision and recall, which is particularly important for underrepresented classes like Hypertrophy (HYP).

The proposed hybrid CNN–LSTM–Attention model demonstrated strong performance in multi-class ECG classification. On the held-out test set of 2,180 ECG scans, the model achieved an overall accuracy of 88.1% and a micro-averaged AUC of 0.936, indicating reliable discrimination across the five diagnostic categories (Figure 16).

	AUC	Precision	Recall	F1	Accuracy
NORM	0.952	0.858	0.945	0.867	0.863
MI	0.931	0.849	0.739	0.761	0.878
STTC	0.944	0.859	0.705	0.751	0.879
CD	0.935	0.899	0.747	0.767	0.901
HYP	0.845	0.718	0.307	0.448	0.885
Overall score	0.921	0.837	0.689	0.719	0.881

Figure 16: The scores (individual and overall) of the model after using the test script. Summarizes the model's performance, showing high AUC and F1-scores for Normal, Myocardial Infarction, and Conduction Disturbance classes, but revealing a low recall for Hypertrophy, which limits its F1-score. Created by student researcher.

Finally, a confusion matrix was generated to visualize the model's classification behavior across the five diagnostic categories (Figure 17). This revealed specific patterns of misclassification—for instance, the model frequently confused true Hypertrophy cases with other conditions, reflected in its relatively low recall for this class.

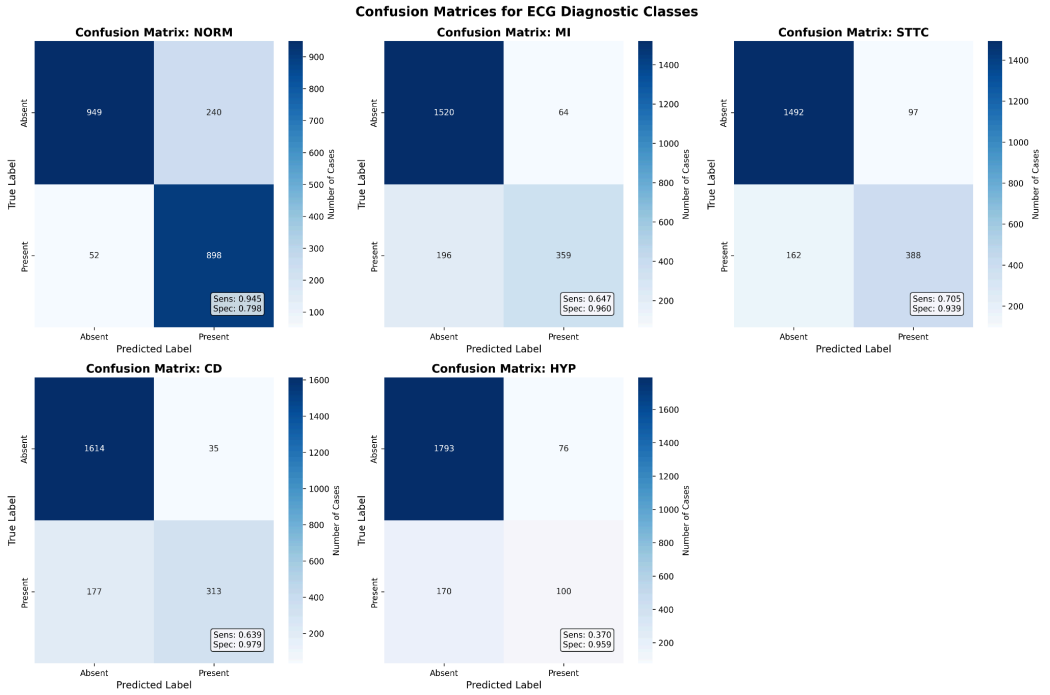


Figure 17: The confusion matrix generated after the model was trained. Reveals specific patterns of misclassification, such as the model frequently confusing true Hypertrophy cases for other conditions, which explains the low recall for that class. Created by student researcher.

3.3. Per-Class Performance

Analysis of per-class results revealed clear strengths and weaknesses. The model performed best at identifying Normal rhythms (NORM), achieving consistently high scores across precision, recall, and F1. Performance was also strong for Myocardial Infarction (MI) and Conduction Disturbance (CD), with high AUC values (0.931 and 0.935) and balanced F1-scores (0.761 and 0.767), demonstrating reliable detection of clinically significant abnormalities (Figure 16).

Performance for ST/T Change (STTC) revealed a tradeoff: while precision was strong (0.859), recall was more limited (0.667), indicating that the model identified STTC cases accurately when predicted but failed to capture a subset of true instances.

The most significant weakness was observed in Hypertrophy (HYP). Despite achieving a reasonable precision score (0.718), recall was very low (0.307). This disparity suggests that the model correctly classified Hypertrophy when predicted but misclassified the majority of true HYP cases into other categories (Figure 17).

3.4. Web Interface

The successful deployment of the AI model was demonstrated through a functional, locally hosted web application. This integrated system provides an end-to-end pipeline from physical ECG data acquisition to AI-powered diagnostic interpretation, validating the model's practicality. The interface, depicted in Figure 18, offers a user-friendly portal with three primary control buttons: "Connect to Arduino" to establish communication with the hardware, "Begin Scan (5-second read)" to initiate data collection, and "Diagnose" to execute the AI model on the captured signal.

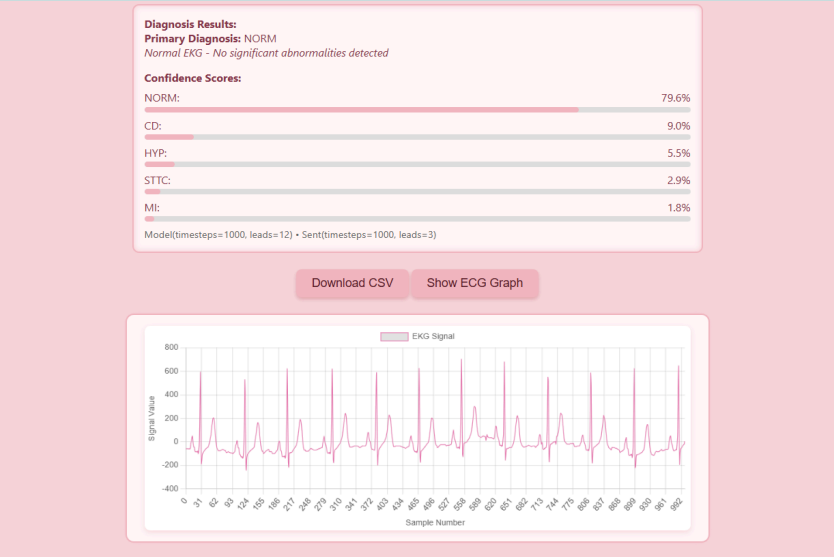


Figure 18: The web interface. Display of the results of the web interface when given a scan from PTB-XL of a normal scan which it did not use to train. Created by student researcher.

The interface provides rapid visual feedback, displaying the live ECG waveform and key signal statistics (minimum, maximum, and average voltage) to aid in data quality assessment. Most critically, the "Diagnosis Results" panel displays the model's multi-label diagnosis, listing all detected conditions with their corresponding probabilities. The connection status is indicated by a color-changing indicator that automatically disconnects after scanning to prevent accidental data overwriting. This integrated prototype successfully proves that the trained model can function outside a static testing environment and can be deployed as a cohesive tool for real-time signal analysis.

3.5. Hardware Integration

The operational latency of the integrated system was benchmarked in a testing environment. The end-to-end process, from scan initiation to the display of the AI diagnosis, was consistently completed in

under 10 seconds, demonstrating feasibility for rapid analysis. Figure 19 illustrates the proposed clinical lead placement.

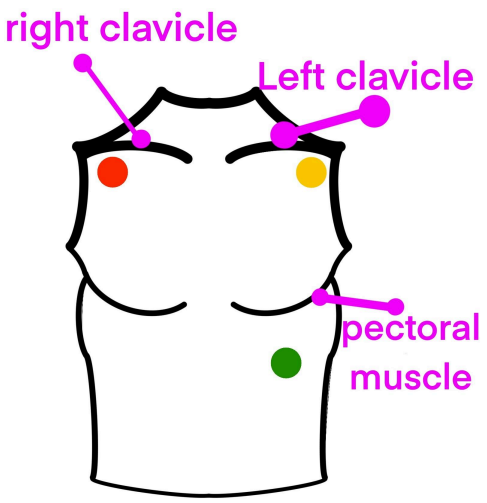


Figure 19: The placement of leads. The RA lead (red) is placed directly under the right clavicle. The LA lead (yellow) is placed directly below the left clavicle. The LL lead (green) is placed below the left pectoral muscle. Created by student researcher.

4. Interpretation and Discussion

4.1. Interpretation of Results

The hybrid CNN–LSTM model with temporal attention demonstrated strong potential as a clinical decision-support tool. Its primary strength lies in its multi-class diagnostic capability, which moves beyond binary classification to better reflect clinical reality. The model's consistency and scalability offer clear advantages over human readers, who are subject to fatigue and inter-observer variability.

4.2. Diagnostic Performance

The model's performance profile aligns with the inherent diagnostic clarity of different cardiac conditions. Its high accuracy for Normal rhythms, Myocardial Infarction, and Conduction Disturbance is likely because these conditions manifest through distinct, well-defined electrophysiological signatures that the model's feature extractors can learn reliably.

In contrast, the poor recall for Hypertrophy is not a fundamental flaw of the model but rather a reflection of a known, well-documented limitation of ECG-based diagnosis for this condition. Standard ECG criteria for hypertrophy are highly specific but notoriously insensitive compared to imaging modalities like echocardiography (Pewsner et al., 2007). The model's struggle with ST/T Change, characterized by good precision but lower recall, mirrors the clinical challenge of these often subtle, transient, and heterogeneous findings.

4.3. Clinical Implications

These findings support the viability of AI-assisted ECG interpretation as a means to reduce diagnostic error and variability. Potential applications include making the Triage systems more efficient by allowing for improved prioritization of urgent cases, such as MI, for rapid physician review (Yancey & O'Rourke, 2023), second-reader support to provide clinicians with an automated "second opinion," and screening tools to identify at-risk individuals in community health settings or via wearable devices where an at-risk patient could be monitored over an extended time for diagnostic purposes. It would also allow doctors to be notified of immediate complications.

The results align with growing evidence that AI can extract clinically valuable information from ECGs beyond traditional interpretation. Similar to how Attia et al. (2018) demonstrated that AI could identify ventricular dysfunction from normal-appearing ECGs, the model's ability to detect multiple conditions simultaneously suggests that AI-assisted ECG analysis could serve as a powerful screening tool for comprehensive cardiovascular assessment.

4.4. Comparison with Previous Studies

This study advances prior work by addressing a clinically relevant multi-class classification task, moving beyond the binary distinctions that dominate the literature. The model's performance compares favorably with earlier reports of binary detection systems (Martínez-Sellés & Marina-Breysse, 2023). The integration of a temporal attention mechanism is a key innovation that enhances not only performance but also interpretability, offering insight into which ECG segments drive predictions.

4.5. Limitations

Key limitations include the lack of prospective clinical validation, potential dataset bias due to the larger amount of normal records compared to other results from the PTB-XL database, and the use of non-clinical-grade hardware (AD8232) for the prototype (Dias et al., 2022). This is due to the 3-lead system compared to the 12-leads used in clinical settings (Figure 19). Future research should focus on

external validation, expanding diagnostic classes to include arrhythmias, and employing explainable AI (XAI) techniques to increase clinician trust and facilitate real-world integration (Kelly et al., 2019).

4.6. Future Work

Future research should expand diagnostic classes to include arrhythmias and channelopathies, integrate real-time analysis for continuous monitoring, and employ explainable AI to increase clinician trust (Violet Turri, 2022). Most importantly, external validation on additional independent clinical datasets is essential to establish generalizability. The dataset should also be expanded and taken from multiple sources to reduce the risk of statistical bias.

5. Conclusion

Cardiovascular disease remains a leading global health burden, and the electrocardiogram is a critical diagnostic tool whose effectiveness is limited by human error, fatigue, and variability in the clinician's training. This study addressed these challenges by developing an AI model capable of multi-class ECG interpretation. The model achieved an overall accuracy of 88.1% across five key readings—normal rhythm, myocardial infarction, ST/T change, conduction disturbance, and hypertrophy—demonstrating that AI can perform complex differential diagnostics with high proficiency.

The potential impact of this technology is substantial as a consistent and scalable AI could reduce diagnostic errors, improve early detection of critical conditions, and standardize ECG interpretation across diverse healthcare settings. While this remains a prototype requiring clinical validation, the results highlight the promise of AI in augmenting human expertise, marking a step toward increased precision of cardiovascular medicine where timely and reliable diagnosis benefits every patient.

This aligns with the broader vision articulated by Topol (2018) of AI-driven "high-performance medicine" that can democratize expertise, reduce diagnostic errors, and ultimately transform patient care through more consistent and accessible diagnostic capabilities.

6. Acknowledgments

I would like to express my gratitude to my mentor, Dr. Steven Lindo, for his support and feedback throughout the project. I would also like to thank my teacher, Ms. Frank, for her support. Lastly I would like to thank the contributors of the PTB-XL dataset, which was essential for this work.

Reference

AI's Ascendance in Medicine: A Timeline. (n.d.). Cedars-Sinai. https://www.cedars-sinai.org/discoveries/ai-ascendance-in-medicine.html

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Saleh, K. B., Badreldin, H. A., Yami, M. S. A., Harbi, S. A., & Albekairy, A. M. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04698-z

Attia, Z. I., Kapa, S., Lopez-Jimenez, F., McKie, P. M., Ladewig, D. J., Satam, G., Pellikka, P. A., Enriquez-Sarano, M., Noseworthy, P. A., Munger, T. M., Asirvatham, S. J., Scott, C. G., Carter, R. E., & Friedman, P. A. (2018). Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. *Nature Medicine*, *25*(1), 70–74. https://doi.org/10.1038/s41591-018-0240-2

Buttner, E. B. (2024, October 8). *The ST segment*. Life in the Fast Lane • LITFL. https://litfl.com/st-segment-ecg-library/?utm_source=chatgpt.com

Choi, E., mp2893, Bahadori, M. T., College of Computing, Georgia Institute of Technology, Schuetz, A., Stewart, W. F., Research Development & Dissemination, Sutter Health, Sun, J., & College of Computing, Georgia Institute of Technology. (n.d.). Doctor AI: Predicting clinical events via recurrent neural networks. In *Proceedings of Machine Learning for Healthcare 2016 JMLR W&C Track Volume 56* (Vol. 56). https://proceedings.mlr.press/v56/Choi16.pdf

Conduction Disorders | NHLBI, NIH. (2022, March 24). NHLBI, NIH. https://www.nhlbi.nih.gov/health/conduction-disorders

Cook, D. A., Oh, S., & Pusic, M. V. (2020). Accuracy of physicians' electrocardiogram interpretations. *JAMA Internal Medicine*, *180*(11), 1461. https://doi.org/10.1001/jamainternmed.2020.3989

Dias, W. G. A., Silva, M. C. C. N., & De Oliveira, E. M. (2022). Assessing the AD8232 sensor's effectiveness on telemedicine kits: checking the AD8232 sensor. *Research Society and Development*, *11*(11), e431111133778. https://doi.org/10.33448/rsd-v11i11.33778

Do you know the symptoms of a heart attack? (n.d.). Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/16818-heart-attack-myocardial-infarction

ECG Learning Center - An introduction to clinical electrocardiography. (n.d.). https://ecg.utah.edu/lesson/2

Electrocardiogram. (2025, September 15). Johns Hopkins Medicine. https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electrocardiogram

Electrocardiogram (ECG or EKG) - Mayo Clinic. (n.d.). https://www.mayoclinic.org/tests-procedures/ekg/about/pac-20384983

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2018). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

GeeksforGeeks. (2025, July 12). *Differences Between Django vs Flask*. GeeksforGeeks. https://www.geeksforgeeks.org/python/differences-between-django-vs-flask/

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal Fundus photographs. *JAMA*, *316*(22), 2402. https://doi.org/10.1001/jama.2016.17216

Hancock, E. W., Deal, B. J., Mirvis, D. M., Okin, P., Kligfield, P., & Gettes, L. S. (2009). AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram. *Circulation*, *119*(10). https://doi.org/10.1161/circulationaha.108.191097

Heart disease - Symptoms and causes. (n.d.). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-causes/syc-20353118

Heart Disease Facts. (2024, October 24). Heart Disease. https://www.cdc.gov/heart-disease/data-research/facts-stats/index.html

Hhackford. (2025, March 25). *CHM releases AlexNet source code*. CHM. https://computerhistory.org/blog/chm-releases-alexnet-source-code/

Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-09954-8

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*(1). https://doi.org/10.1186/s12916-019-1426-2

Kingma, D. P., & Ba, J. (2014, December 22). *Adam: A method for stochastic optimization*. arXiv.org. https://arxiv.org/abs/1412.6980

Kligfield, P., Gettes, L. S., Bailey, J. J., Childers, R., Deal, B. J., Hancock, E. W., Van Herpen, G., Kors, J. A., Macfarlane, P., Mirvis, D. M., Pahlm, O., Rautaharju, P., & Wagner, G. S. (2007). Recommendations for the standardization and interpretation of the electrocardiogram. *Journal of the American College of Cardiology*, *49*(10), 1109–1127. https://doi.org/10.1016/j.jacc.2007.01.024

Kraik, K., Dykiert, I. A., Niewiadomska, J., Ziemer-Szymańska, M., Mikołajczak, K., Kreń, M., Kukiełka, P., Martuszewski, A., Harych, T., Poręba, R., Gać, P., & Poręba, M. (2025). The most common errors in automatic ECG interpretation. *Frontiers in Physiology*, *16*. https://doi.org/10.3389/fphys.2025.1590170

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Left ventricular hypertrophy - Symptoms and causes. (n.d.). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/left-ventricular-hypertrophy/symptoms-causes/syc-20374314

Martínez-Sellés, M., & Marina-Breysse, M. (2023). Current and future use of artificial intelligence in electrocardiography. *Journal of Cardiovascular Development and Disease*, *10*(4), 175. https://doi.org/10.3390/jcdd10040175

Ojha, N., & Dhamoon, A. S. (2023, August 8). *Myocardial infarction*. StatPearls - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK537076/

Paszke, A., University of Warsaw, Facebook AI Research, Facebook AI Research, Facebook AI Research, Google, Facebook AI Research, Self Employed, Facebook AI Research, NVIDIA, Orobix, Oxford University, Xamla, Facebook AI Research, Facebook AI Research, Nabla, Twitter, Qure.ai, Facebook AI Research, . . . Facebook AI Research. (n.d.). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada*. https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf

Pewsner, D., Jüni, P., Egger, M., Battaglia, M., Sundström, J., & Bachmann, L. M. (2007). Accuracy of electrocardiography in diagnosis of left ventricular hypertrophy in arterial hypertension: systematic review. *BMJ*, *335*(7622), 711. https://doi.org/10.1136/bmj.39276.636354.ae

Professional, C. C. M. (2025, March 19). *Right atrial enlargement*. Cleveland Clinic. https://my.clevelandclinic.org/health/symptoms/23576-right-atrial-enlargement

PTB-XL, a large publicly available electrocardiography dataset v1.0.3. (2022, November 9). https://physionet.org/content/ptb-xl/1.0.3/

Thygesen, K., Alpert, J. S., Jaffe, A. S., Chaitman, B. R., Bax, J. J., Morrow, D. A., & White, H. D. (2018). Fourth Universal Definition of Myocardial Infarction (2018). *Journal of the American College of Cardiology*, *72*(18), 2231–2264. https://doi.org/10.1016/j.jacc.2018.08.1038

Topol, E. J. (2018). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Karnofsky, H. (2025, January 6). *AI Has Been Surprising for Years*. Cargenie Endowment for International Peace. https://carnegieendowment.org/research/2025/01/ai-has-been-surprising-for-years?lang=en

Wagner, P., Strodthoff, N., Bousseljot, R., Kreiseler, D., Lunze, F. I., Samek, W., & Schaeffter, T. (2020). PTB-XL, a large publicly available electrocardiography dataset. *Scientific Data*, *7*(1). https://doi.org/10.1038/s41597-020-0495-6

What is Explainable AI? (2022, January 17). SEI Blog. https://www.sei.cmu.edu/blog/what-is-explainable-ai/

World Health Organization: WHO. (2025, July 31). *Cardiovascular diseases (CVDs)*. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Yancey, C. C., & O'Rourke, M. C. (2023, August 28). *Emergency department triage*. StatPearls - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK557583/

Zhang, Y., Xu, S., Xing, W., Chen, Q., Liu, X., Pu, Y., Xin, F., Jiang, H., Yin, Z., Tao, D., Zhou, D., Zhu, Y., Yuan, B., Jin, Y., He, Y., Wu, Y., Po, S. S., Wang, H., & Benditt, D. G. (2024). Robust Artificial Intelligence tool for Atrial fibrillation diagnosis: Novel development approach incorporating both atrial electrograms and surface ECG and evaluation by Head‐to‐Head comparison with Hospital‐Based Physician ECG readers. *Journal of the American Heart Association*, *13*(3). https://doi.org/10.1161/jaha.123.032100

Zheng, S., Cui, X., & Ye, Z. (2025). Integrating artificial intelligence into radiological cancer imaging: from diagnosis and treatment response to prognosis. *Cancer Biology and Medicine*, *22*(1), 6–13. https://doi.org/10.20892/j.issn.2095-3941.2024.0422