Preliminary Project Report

Aditya Varun V
Indian Institute of Technology, Hyderabad
Kandi, Sangareddy
ai22btech11001@iith.ac.in

Arsh Arora
Indian Institute of Technology, Hyderabad
Kandi, Sangareddy
bm22btech11004@iith.ac.in

Surya Saketh Chakka
Indian Institute of Technology, Hyderabad
Kandi, Sangareddy
ai22btech11005@iith.ac.in

Mayank Parasramka
Indian Institute of Technology, Hyderabad
Kandi, Sangareddy
ai22btech11018@iith.ac.in

Saketh Ram Kumar Dondapati
Indian Institute of Technology, Hyderabad
Kandi, Sangareddy
ai22btech11023@iith.ac.in

Abstract

With the increasing need for advanced diagnostic tools, developing technologies capable of real-time health monitoring can significantly reduce the workload on healthcare professionals. Machine Learning (ML) models that can assist in early detection and diagnosis of medical conditions are up and coming. However, training these models requires access to large volumes of diverse medical data, often restricted due to privacy laws like HIPAA. Federated Learning(FL) offers an innovative solution by enabling collaborative model training across institutions without sharing sensitive patient data. This decentralized approach safeguards privacy while utilizing a wide range of datasets. This paper explores FL and techniques to apply to the PT-BXL dataset to advance ECG interpretation.

This paper begins with a short introduction to ECG and the dataset we will be using. Next, we explore suitable model architectures that could be applied to this dataset to classify the time series data. After that, we will briefly overview FL concepts, such as centralized and decentralized federated learning, and the algorithms involved here. These will be used to create the federated learning model for the given dataset.

1. Electrocardiograms (ECG)

An Electrocardiogram (ECG) is a tool that records the heart's electrical activity via electrodes on the skin, which detect electrical changes during each heartbeat. The resulting graph shows time on the horizontal axis and voltage on the vertical axis, allowing healthcare professionals to assess cardiac rhythm and diagnose heart conditions.

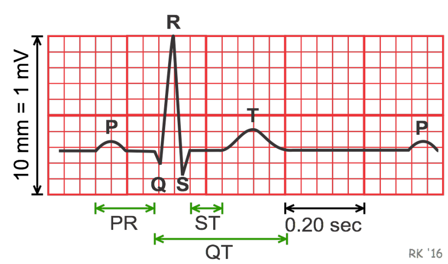

Figure 1. ECG Signal with Grids

Key waveform components include:

1. **P-Wave:** Represents atrial depolarization.
2. **Q-Wave:** The initial negative deflection, indicating early ventricular depolarization.
3. **R-Wave:** The main positive deflection, indicating ventricular depolarization.
4. **S-Wave:** The final negative deflection, indicating late ventricular depolarization.
5. **T-Wave:** Indicates ventricular repolarization, crucial for identifying ischemia or electrolyte imbalances.

1.1. ECG Analysis

ECG analysis involves heart rate calculation, rhythm evaluation, and waveform interpretation. Heart rate is calculated using the RR interval as follows:

$$\text{Heart Rate (bpm)} = \frac{60}{RR \text{ interval (seconds)}}$$

Rhythm evaluation determines if RR intervals are regular or irregular, indicating conditions like normal sinus rhythm or arrhythmias such as atrial fibrillation. P-wave morphology and the PR interval reveal atrial activity; abnormal P-waves may suggest enlargement, while a prolonged PR interval indicates potential AV block.

The QRS complex reflects ventricular depolarization, with width indicating conduction status; a narrow QRS suggests normal conduction, while a wide QRS may indicate a bundle branch block. ST segment and T-wave analysis are crucial for diagnosing ischemia, with ST-segment changes suggesting myocardial infarction and T-wave abnormalities indicating electrolyte imbalances or other cardiac issues. This comprehensive evaluation underscores the ECG's vital role as a diagnostic tool.

1.2. PTBXL and its details

The PTB-XL dataset is a large publicly available dataset designed to advance automatic ECG interpretation algorithms, facilitating research and development in cardiovascular diagnostics. Below is a detailed breakdown of its key features and structure.

1.2.1 Dataset Overview

PTB-XL is the largest open-access 12-lead electrocardiogram (ECG) dataset, comprising 21,837 ECG recordings from 18,885 patients.Each 10-second recording offers high-quality diagnostic information and is annotated with 71 diagnostic labels, organized hierarchically into 5 superclasses and 24 subclasses, which include both normal and pathological conditions, such as myocardial infarction, conduction disturbances, hypertrophy, and ST/T changes.

1.2.2 Demographics and Data Balance

The dataset is balanced by gender, with 52% male and 48% female patients. Patient ages range from 0 to 95 years, with a median age of 62. This diversity ensures the dataset is representative of both healthy and pathological ECG patterns, which is crucial for developing generalizable ML models.

1.2.3 Recording Details and Preprocessing

The ECG data was recorded between 1989 and 1996 using Schiller AG devices at a sampling rate of 400 Hz and subsequently upsampled to 500 Hz for higher temporal resolution. A downsampled version at 100 Hz is also available for ease of use. Preprocessing steps included removing spikes (common at the recording's start and end) and converting to a standardized 16-bit binary format. Noise, baseline drift, and electrode issues were annotated to enhance the dataset's utility for real-world applications.

1.2.4 Annotations and Hierarchical Structure

Each ECG recording is annotated with diagnostic labels following the SCP-ECG standard. These labels include multiple diagnostic statements and are grouped into a hierarchical structure with 5 superclasses:

- **Normal (NORM)**
- **Conduction Disturbances (CD)**
- **Myocardial Infarction (MI)**
- **Hypertrophy (HYP)**
- **ST/T Changes (STTC)**

Each superclass is further divided into 24 subclasses, representing more specific diagnoses. This organization supports multi-label classification tasks and allows ML algorithms to be trained on both coarse-grained and fine-grained diagnostic categories.

1.2.5 Signal Quality and Validation

Approximately 73.7% of the ECGs in PTB-XL were validated by human cardiologists, ensuring high-quality annotations, while 26.3% were automatically annotated by ECG devices, with some undergoing secondary validation. Signal quality metadata, such as noise presence, baseline drift, and electrode issues, was manually annotated by technical experts, providing essential information for evaluating the robustness of ML models to real-world ECG data imperfections.

1.2.6 Cross-Validation and Benchmarking

PTB-XL features predefined cross-validation folds for standardized evaluation of ML models, divided into 10 stratified folds. The 9th and 10th folds are designated for validation and testing, with the 10th fold containing only ECG recordings validated by human cardiologists, ensuring reliability for final model evaluation. This structure prevents data leakage by placing all ECGs from a single patient in the same fold.

2. Model Architectures

This section provides an overview of the deep learning model architectures we will utilise for the classification task with the PTB-XL database.

2.1. Transformers

Transformers have shown good results in modeling long-range dependencies for sequential data, and hence are a good choice for the model. This section will cover changes made to the traditional transformer architecture to account for the time series nature of the data.

2.1.1 Time Series Modifications

Positional Encoding

- **Absolute Encoding:** Encoding the position in the sequence absolutely, does not exploit all the information of the time series data.
- **Learnable Encoding:** This has shown better results than the above since the encoding can adapt to specific tasks. References utilizing this: Zerveas *et al*. [15], Lim *et al*. [4].
- **Timestamp Encoding:** These may provide additional information for the classification since it gives information about the time of the data-point occurrence. These may be used as additional positional encoding information and have been used in the following: Zhou *et al*., 2021 [16], Wu *et al*., 2021 [13], Zhou *et al*., 202 [17]

Attention Module To reduce the time and memory complexities, several alternatives to the original transformer have been proposed. This reduction of complexity has been based on introducing sparsity into the attention mechanism or using the low rank property of the self-attention matrix to speed up the computation. Using these algorithms may be helpful under constrained circumstances.

Methods	Time Complexity	Memory Complexity	Steps
Transformer [Vaswani et al., 2017[12]]	$O(N^2)$	$O(N^2)$	N
LogTrans [Li et al., 2019[3]]	$O(N \log N)$	$O(N \log N)$	1
Informer [Zhou et al., 2021[16]]	$O(N \log N)$	$O(N \log N)$	1
Autoformer [Wu et al., 2021[13]]	$O(N \log N)$	$O(N \log N)$	1
Pyraformer [Liu et al., 2022a[5]]	$O(N)$	$O(N)$	1
Quatformer [Chen et al., 2022[2]]	$O(2cN)$	$O(2cN)$	1
FEDformer [Zhou et al., 2022[17]]	$O(N)$	$O(N)$	1

Table 1. Transformer Complexities with Different Attention Modules

2.1.2 Transformer Model Architectures

The following are previously used transformer models that may help with implementation of the transformer model architecture for the PTBXL data and other time series data:

- **Zerveas *et al*. [15]**: Introduces an embedding layer in Transformer that learns embedding vectors for each position index jointly with other model parameters.
- **Lim *et al*. [4]**: Utilizes an LSTM network to encode positional embeddings, which can better exploit sequential ordering information in time series.
- **Masked Transformer for Electrocardiogram Classification**. [18] Provides an alternative transformer training strategy using an encoder-decoder transformer on the Vision transformer architecture.

2.2. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) have emerged as effective models for analyzing ECG data, offering robust performance in diagnosing cardiac abnormalities on the PTB-XL dataset. CNNs can process raw ECG data directly without extensive preprocessing, making them suitable for real-time analysis in clinical settings.

2.2.1 CNN Architectures and Performance

- **ResNet and Inception-Based Models by Strodthoff *et al*. [11]**: The ResNet and Inception-based CNN architectures have demonstrated superior performance in classifying ECG data from the PTB-XL dataset. These models exploit skip connections, making the network deeper without suffering from vanishing gradients. In particular, the xResNet1d101 model has been highlighted for its very high accuracy across various ECG classification tasks.
- **Lightweight CNN Model by Saglietto *et al*. [8]**: A lightweight CNN model trained on single-lead ECG data (Lead D1) from the PTB-XL dataset achieved promising results despite a performance drop compared to 12-lead setups. The average AUC difference was around 8.7%. However, this gap was significantly reduced when a second lead (D2) was incorporated, demonstrating that simpler CNN models can still deliver strong performance even with reduced lead inputs.

3. Federated Learning (FL)

FL is a method in which a central deep model is trained in a dispersed manner, avoiding issues of data privacy. This is done by training a deep model locally at a node and then uploading the results. In healthcare, for example, the model can be trained by a hospital. This is significant because Artificial Intelligence (AI) can be used in many healthcare applications, such as diagnosis, text identification in reports, etc. In this process, the node would only return its results, not the intermediate data used for local training. Therefore, the patient's data itself is never actually collected by the central system and, as a result, remains confidential. Furthermore, this way, many medical institutions can collaborate to build a good model without the fear of data leakage. FL exchanges a minimal amount of data to train deep networks.

4. Literature Review

4.1. Centralised Federated Learning(CFL)

McMahan *et al*. [7] proposed the first well-known FL algorithm, the ***FederatedAveraging*** algorithm. In the proposed model, there is a central aggregator and a fixed set of K nodes. A synchronous update scheme is assumed, which proceeds in rounds of communication. At the start of each round, a random fraction of nodes C is chosen, after which the current model parameters are sent to each of the chosen nodes. Then, the selected nodes locally compute results and send them to the server.

Server executes:
initialize w_0
for each round $t = 1, 2, \ldots$ **do**
  $m \leftarrow \max(C \cdot K, 1)$
  $S_t \leftarrow$ (random set of m clients)
  for each client $k \in S_t$ **in parallel do**
    $w^k_{t+1} \leftarrow$ ClientUpdate(k, w_t)
  $m_t \leftarrow \sum_{k \in S_t} n_k$
  $w_{t+1} \leftarrow \sum_{k \in S_t} \frac{n_k}{m_t} w^k_{t+1}$ *// Erratum*[4]

ClientUpdate(k, w)**:** *// Run on client k*
$\mathcal{B} \leftarrow$ (split $\mathcal{P}_k$ into batches of size B)
for each local epoch i from 1 to E **do**
  for batch $b \in \mathcal{B}$ **do**
    $w \leftarrow w - \eta \nabla \ell(w; b)$
return w to server

Figure 2. The ***FederatedAveraging*** algoithm

Each client k calculates $g_k = \nabla F_k(w_t)$, the average gradient on the local data using the current model parameters w_t. Then, the aggregator applies the update given by $w_{t+1} \leftarrow w_t - \eta \sum_{k=1}^{CK} \frac{n_k}{n} g_k$. In other words, the aggregator takes the weighted sum of the gradients formulated at the chosen nodes. The gradient finding at each node can be characterized by two more parameters, E and B. These parameters are the number of epochs and the local minibatch size used at each node, respectively.

4.2. Decentralised Federated Learning (DFL)

DFL is an advanced extension of the traditional FL paradigm, which emerged in 2018. Unlike CFL approaches, it distributes the aggregation of model parameters between the neighbouring participants [6]. In DFL, each client k computes the local gradient $g_k = \nabla F_k(w_t)$ based on its own dataset and the current model parameters w_t. Instead of sending the gradients to a central server, clients exchange these locally computed results with a subset of their neighbours.

The model updates are then aggregated in a decentralized manner, where each node updates its parameters using the model parameters from neighbouring nodes with $w^k_{t+1} \leftarrow w^k_t - \eta \sum_{j \in N(k)} w^j_t$ where $N(k)$ represents the set of neighboring nodes of client k. This decentralised aggregation allows the system to operate asynchronously, where nodes may have varying communication delays or participate in different update schedules. Compared to CFL, DFL improves the scalability and addresses the limitations of having a single point of failure, trust dependencies, and bottlenecks at the central server node. Like CFL, DFL can also be characterised by two parameters: E, the number of Epochs and B, the local minibatch size used at each node.

Healthcare Institutions are inclined towards DFL frameworks over CFL due to their abundant patient privacy data, computational resources, and storage capabilities. Healthcare institutions widely employ DFL frameworks in various studies [1, 9, 10, 14].

References

[1] Saleh Baghersalimi, Tomas Teijeiro, Amir Aminifar, and David Atienza. Decentralized federated learning for epileptic seizures detection in low-power wearable systems. *IEEE Transactions on Mobile Computing*, 2023. 4

[2] Weiqi Chen, Wenwei Wang, Bingqing Peng, Qingsong Wen, Tian Zhou, and Liang Sun. Learning to rotate: Quaternion transformer for complicated periodical time series forecasting. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, page 146–156, New York, NY, USA, 2022. Association for Computing Machinery. 3

[3] Shiyang Li, Xiaoyong Jin, Yao Xuan, Xiyou Zhou, Wenhu Chen, Yu-Xiang Wang, and Xifeng Yan. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting, 2020. 3

[4] Bryan Lim, Sercan Ö. Arık, Nicolas Loeff, and Tomas Pfister. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4):1748–1764, 2021. 3

[5] Shizhan Liu, Hang Yu, Cong Liao, Jianguo Li, Weiyao Lin, Alex X. Liu, and Schahram Dustdar. Pyraformer: Low-complexity pyramidal attention for long-range time series modeling and forecasting. In *International Conference on Learning Representations*, 2022. 3

[6] Enrique Tomás Martínez Beltrán, Mario Quiles Pérez, Pedro Miguel Sánchez Sánchez, Sergio López Bernal, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez, and Alberto Huertas Celdrán. Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges, 2023. 4

[7] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data, 2023. 3

[8] Andrea Saglietto, Daniele Baccega, Roberto Esposito, Matteo Anselmino, Veronica Dusi, Attilio Fiandrotti, and Gaetano Maria De Ferrari. Convolutional neural network (cnn)-enabled electrocardiogram (ecg) analysis: a comparison between standard twelve-lead and single-lead setups. *Frontiers in Cardiovascular Medicine*, 11:1327179, 2024. Erratum in: Front Cardiovasc Med. 2024 Mar 14;11:1396396. 3

[9] Micah J Sheller, G Anthony Reina, Brandon Edwards, Jason Martin, and Spyridon Bakas. Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part I 4*, pages 92–104. Springer, 2019. 4

[10] Micah J Sheller, Brandon Edwards, G Anthony Reina, Jason Martin, Sarthak Pati, Aikaterini Kotrotsou, Mikhail

Milchenko, Weilin Xu, Daniel Marcus, Rivka R Colen, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific reports*, 10(1):12598, 2020. 4

[11] Nils Strodthoff, Patrick Wagner, Tobias Schaeffter, and Wojciech Samek. Deep learning for ecg analysis: Benchmarks and insights from ptb-xl. *IEEE Journal of Biomedical and Health Informatics*, 25(5):1519–1528, 2021. 3

[12] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2023. 3

[13] Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting, 2022. 3

[14] Jie Xu, Benjamin S Glicksberg, Chang Su, Peter Walker, Jiang Bian, and Fei Wang. Federated learning for healthcare informatics. *Journal of healthcare informatics research*, 5: 1–19, 2021. 4

[15] George Zerveas, Srideepika Jayaraman, Dhaval Patel, Anuradha Bhamidipaty, and Carsten Eickhoff. A transformer-based framework for multivariate time series representation learning, 2020. 3

[16] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(12):11106–11115, 2021. 3

[17] Tian Zhou, Ziqing Ma, Qingsong Wen, Xue Wang, Liang Sun, and Rong Jin. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting, 2022. 3

[18] Ya Zhou, Xiaolin Diao, Yanni Huo, Yang Liu, Xiaohan Fan, and Wei Zhao. Masked transformer for electrocardiogram classification, 2024. 3