Practical work 1: The study of ECG Heartbeat Categorization

I. Introduction

In this work, we study the problem of ECG heartbeat categorization using a one-dimensional Convolutional Neural Network (1D CNN). The main objectives are to analyze the dataset, pre-process the data, and implement a classification model to distinguish between different classes of heartbeats.

II. Dataset and Exploratory Data Analysis

A. Dataset Description

The dataset used in this study is the MIT-BIH Arrhythmia Database, a benchmark resource for the classification of ECG heartbeats. The five classes of heartbeats in the dataset are:

- **N**: Normal beat
- **S**: Supraventricular ectopic beat
- **V**: Ventricular ectopic beat
- **F**: Fusion beat
- **Q**: Unknown beat

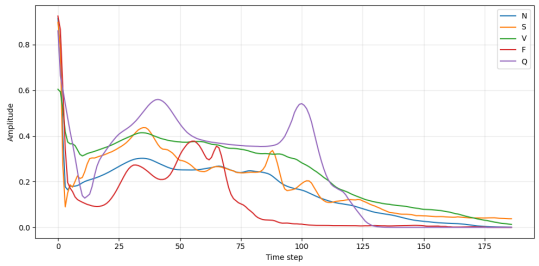

Fig. 1. Average ECG waveform by class

B. Class Distribution

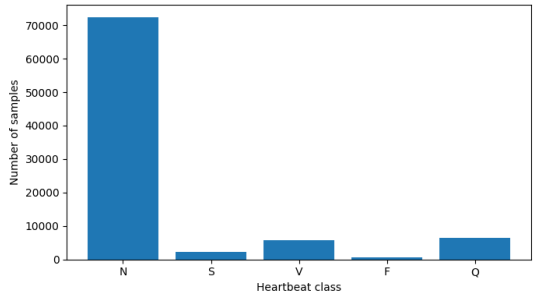

Fig. 2. Class distribution of heartbeat types

A first insight into the dataset reveals a highly imbalanced class distribution, as illustrated in Figure 2. The majority of samples belong to the **N** (Normal) class, significantly outnumbering the remaining arrhythmic classes (**S**, **V**, **F**, and **Q**).

- **N**: 72,471 samples
- **S**: 2,223 samples
- **V**: 5,788 samples
- **F**: 641 samples
- **Q**: 6,431 samples

III. Methodology

The dataset was first split into training and validation sets (90%/10%), and features were standardized. We addressed class imbalance by providing computed class weights during training. For classification, we implemented a 1D Convolutional Neural Network (CNN) with two convolutional and max pooling layers for temporal feature extraction, followed by a dense layer with dropout for regularization, and a softmax output layer to predict the 5 heartbeat classes.

IV. Evaluation and Results

The performance of the 1D CNN model was evaluated on the test set using a confusion matrix (Figure 3) and key classification metrics (Table I). The model's overall weighted F1-score of 0.967 and accuracy of 96.4% indicate effective heartbeat classification, despite class imbalance. The confusion matrix shows that the model performs excellently across all heartbeat classes.

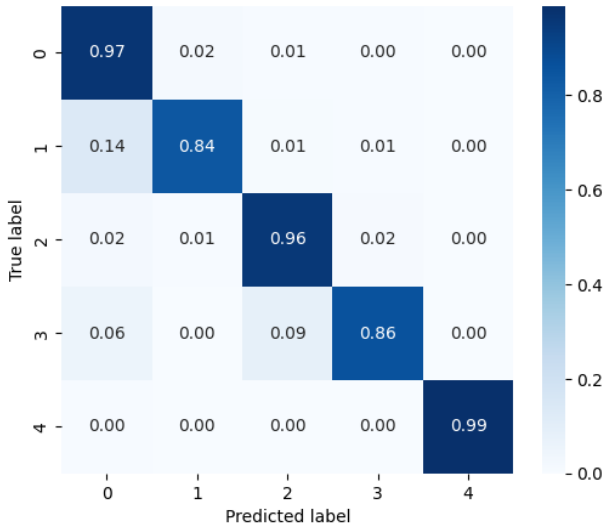

Fig. 3. Confusion matrix of the 1D CNN model

TABLE I
CNN CLASSIFICATION REPORT

Class	Precision	Recall	F1-score	Support
N	0.9930	0.9674	0.9800	18118
S	0.5742	0.8417	0.6827	556
V	0.9016	0.9558	0.9279	1448
F	0.5472	0.8580	0.6683	162
Q	0.9713	0.9888	0.9800	1608
Accuracy			**0.9642**	21892
Macro avg	0.7975	0.9223	0.8478	21892
Weighted avg	0.9714	0.9642	0.9667	21892

V. CONCLUSION

In this study, we explored the ECG Heartbeat Categorization dataset and implemented a 1D Convolutional Neural Network for heartbeat classification. Our model achieved an accuracy of **96.4%** on the test set, demonstrating its effectiveness in distinguishing between five heartbeat classes. Further improvements may involve more advanced architectures or signal augmentation to enhance predictive performance.