Machine Learning Task – Idoven

Explanation of the work done and lessons learned

In this notebook, the primary focus was on a Machine Learning task aimed at classifying ECG signals as part of the Data Scientist position at Idoven.

The task began by importing the necessary libraries, including key tools for data manipulation, visualization, and signal processing.

ECG signal data was loaded from Physionet with a 100 Hz sampling frequency due to computational constraints, though higher frequencies would likely improve the results.

The metadata containing labels for classification was stored in a DataFrame. As discussed in the notebook, several features in the metadata were removed due to their critical information, which had been elaborated by experts and could lead to data leakage. Outside of the notebook, the use of these features was studied, resulting in a performance increase up to 0.85 in the ROC AUC metric. However, I ultimately considered this to be data leakage, as these features would only be available in a setting where an expert cardiologist is present, thus indicating an environment where the expert is directly providing diagnoses through these features.

Additionally, a notable imbalance between the 'NORM' and 'HYP' labels was observed. Although this issue was not directly addressed in the project, managing such imbalances is essential for future work, potentially using techniques like undersampling or applying weights.A custom function was developed to visualize the 12-lead ECGs in a format commonly used by cardiologists, aiding in the interpretation of the signals.

The signals exhibited noise, requiring a filtering process. A standard filtering pipeline from the NeuroKit library was employed, including a 0.5 Hz high-pass Butterworth filter and a powerline filter (50 Hz).

The decision to apply Fourier transform analysis helped identify noise in low frequencies and confirm the effectiveness of the filtering techniques.

The notebook also explored methods for extracting relevant features from ECG signals, such as QRS complex analysis, which can be critical for accurate disease classification. Median beats extraction was explored as the median beats correspond to a very efficient way to sum up redundant information in the signal while reducing the computational cost. Additionally, a hand-made ECG featurizer was implemented to enhance the quality of features derived from the ECG signals.

Exploration of Models

Several machine learning models were evaluated on the classification of ECG signals into categories such as 'NORM', 'MI', 'STTC', 'CD', and 'HYP'. These models were trained using both metadata features from the patients as well as information extracted from the ECG signals. This dual approach helped achieve a better understanding of the data's predictive capacity.

- Convolutional Neural Networks (CNNs): CNNs were applied directly to classify the raw ECG signals. These networks excel at capturing spatial relationships between leads and signal segments, automatically learning relevant features for accurate classification. CNNs demonstrated strong performance by extracting informative patterns from the signals.
- CNNs with Residual Connections: To enhance feature learning and avoid issues like vanishing gradients in deeper networks, residual connections were introduced into the CNN architecture. These connections help networks learn more complex features by allowing certain layers to bypass others, leading to better performance when dealing with subtle ECG variations across classes.
- Long Short-Term Memory (LSTM) Networks: Given that ECG signals are time-series data, LSTM models were leveraged to capture temporal dependencies in the signals. LSTMs are highly effective for retaining information from previous time steps, helping the model learn sequential patterns critical for distinguishing between cardiovascular conditions.
- Variational Autoencoders (VAEs): The use of VAEs presents a promising path forward, especially when dealing with limited or poorly labeled data, a very recurrent challenge on health data. VAEs allow for the unsupervised learning of a latent space representation from large amounts of unlabeled data. Once a latent space is learned, a smaller amount of reliable labeled data can be used to train a classifier. Additionally, VAEs offer high explainability, as the latent space dimensions can be manipulated and studied. By encoding and reconstructing ECGs with slight variations in specific latent dimensions, the relevance of particular features can be directly analyzed.

These analyses achieved relatively good results, with an average ROC AUC of 0.74 on the best scenario. The primary objective of this work was to explore various techniques and models instead of concentrating solely on hyperparameter tuning, which only requires of significant time to validate changes in hyperparameters. This exploratory approach emphasizes potential and promising avenues for future research, particularly with access to greater computational resources and extended development time.

Lessons Learned:

- Computational Limitations: Working on a local machine required the use of lower-frequency ECG signals, highlighting the need for scalable resources, such as cloud-based environments or high computational power computers, in future iterations to handle higher-resolution data and more complex computations.

- Data Imbalance: The observed class imbalance suggests that future work should incorporate strategies like resampling or class weighting to prevent skewed results toward the majority class.

- Signal Preprocessing: Effective preprocessing, particularly for noisy ECG signals, is crucial for achieving clean and interpretable features. Though NeuroKit provided reliable standard filters, exploring more advanced denoising techniques could further improve the signal quality.

- Potential Use of VAE (Variational Autoencoder):
    - Unsupervised Learning: The lack of highly trustworthy labeled data is a recurrent issue in health data, and VAEs offer a promising solution. By first training in an unsupervised manner with abundant unlabelled or poorly labeled data, VAEs can help retain the scarce, high-quality labeled data for a second supervised training phase, improving overall model performance.
    - High Explainability: VAEs also offer enhanced explainability, as they can easily encode and decode data from a latent space. This allows researchers to isolate specific latent dimensions deemed important by the classifier and investigate their role in ECG signal interpretation. By adjusting specific latent dimensions while keeping others constant, it's possible to understand the impact of certain features on the model's decisions.

Exploring Future Directions:

A wide range of potential directions for future work was analyzed. The use of median beats and a hand-crafted ECG featurizer demonstrated promise for enhancing feature quality. Additionally, integrating CNNs with residual connections alongside patient metadata showed potential for improving classification performance. The incorporation of Variational Autoencoders (VAEs) is particularly promising, as it addresses the challenges associated with limited labeled data and provides better insight into how key features contribute to classification.