Abstract

***Aims**: Automatic abnormal ECG detection algorithms are crucial for assisting in treatment of heart health problems and saving lives. This study's objective is to verify whether it is possible to detect outliers in ECG data using unsupervised learning methods or more specifically probabilistic and deep learning models. An outlier for our case is an abnormal ECG signal, a one belonging mostly to a sick person while the normal case is that of a normal healthy person.*

***Methods**: The methods used are Autoencoders, Variational Autoencoders, Diffusion Models, Partial Shift Variational Autoencoder (our implementation of half VAE half prediction model), Normalizing flows and Gaussian Mixture Models (GMMS). Where the first four models try to reconstruct the original signal from the input. The models are trained on normative data and are fed by mixed data for classification, the error in reconstruction and the loss according to the model are the classification metrics, with higher scores meaning abnormality. The models are trained and tested on the PTB-XL and CPSC 2018 datasets, which contain 12-Lead signals of different categories.*

***Results**: The results have shown that the models distinguish between normal and abnormal data to a specific degree, more specifically the AUC scores on the PTB-XL and CPSC datasets for the best two models are: 0.88 and 0.83 for the VAE, 0.80 and 0.74 for the PSVAE. Moreover, the VAE achieved an AUC of 0.89, 0.80 and 0.81 for Conduction Disturbance, Myocardial Infarction, and ST/T Change respectively.*

***Conclusion**: This indicates that a VAE when optimized itself and fed with more proper data may be able to detect outliers in ECG data. Additionally, the first four models could generate samples from noise, hence they could be used to produce new samples without human measurements.*