Machine Learning Models for Seizure Detection: Deployment Insights for e-Health IoT Platform

Gabriel Puerta, Frédéric Le Mouël, Oscar Carrillo

Machine Learning Models for Seizure Detection: Deployment Insights for e-Health IoT Platform

Gabriel Puerta[1], Frédéric Le Mouël[1] and Oscar Carrillo[1]

***Abstract*— This work focused on evaluating some machine learning models for their application in e-Health IoT platforms used to detect epileptic seizures. The evaluation is based on two groups of metrics: statistical validation and computational complexity. These metrics determine relevant factors for the models' selection based on the intrinsic efficiency of ML models to detect seizures and their IoT appropriateness to reduce computation and, therefore, energy use. The evaluation scenario defines an EFC architecture with Edge, Fog, and Cloud layers, where the models are deployed initially in the cloud layer for training and validation, to later be deployed in the fog and edge layers for use. Results highlight that GBC and XGBC models present better performance when executed from the cloud; LR, NB and SNN models can be trained from fog nodes, and finally, SLR and MLP can be deployed and used from edge nodes. MLP especially presents a good balance between a low computational cost and a high accuracy in seizure detection.**

I. INTRODUCTION

Multiples authors state that Internet of things (IoT)-based systems constitute the beginning of the fourth industrial revolution (4.0), promising several changes in all aspects of human life and, especially, new paradigms in the healthcare industry [1]. Artificial Intelligence (AI) proposes solutions to some e-Health challenges, but - as requiring high computational capabilities, has to adapt to IT infrastructure constraints to be geographically close to the patient [2]. For the above reasons, this work focuses on analysing some AI models adopted in Machine Learning (ML) on top of an EFC architecture for e-Health (Edge/Fog/Cloud) [3]. We illustrate these models relevancy with epilepsy seizure detection and prediction [4], either by evaluating the accuracy of seizure detection and by the computational training cost on each EFC part.

II. MATERIALS AND METHODS

The evaluation scenario was tested with a system classifying and predicting seizures in patients with epilepsy. Several ML techniques were used and compared: Simple Linear Regression (SLR), Logistic Regression (LR), K-Nearest Neighbors (KNN), Stochastic Gradient Descent (SGD), Naïve Bayes (NB), Decision Tree Classifier (DT), Random Forest (RF), Gradient Boosting Classifier (GBC), Extremely Random Trees (ETC), XG Boost Classifier (XGBC), Sequential model Neural Network (SNN) and Multi-Layer Perceptron neural network (MLP). Two data sets were used: one containing 4097 electroencephalography (EEG) readings from 500 patients [5] and a second with EEG records of 23 pediatric subjects with intractable seizures [6]. A 70/30 training/testing strategy was applied while respecting the prevalence between the two subsets - percentage of patients having the epileptic characteristic to be detected. Metrics used for model evaluations are in Table I.

III. RESULTS

Results are presented in Table II. Three groups can be identified: (1) highly-performant models - with GBC and XBGC, but having an important computational cost; (2) low-computational cost models - with SLR and MLP, where only MLP is performing well; (3) reasonably costly models - with LR, NB and SNN, where only NB and SNN are performing well. Other models belong to the first group, are costly but sub-optimal compared to GBC and XBGC.

[1]Univ Lyon, INSA Lyon, Inria, CITI, EA3720, 69621 Villeurbanne, France

Models	Ccost	Metrics	Formula
SLR	$O(p^2n+n^3)$	Accuracy	$\frac{(tp+tn)}{(tp+tn)}$
LR	$O(np)$		
KNN	$O(npk_{neighbors})$	Recall	$\frac{tp}{(tp+fn)}$
SGD	$O(n^2p+p^3)$		
NB	$O(np)$	Specificity	$\frac{tn}{(tp+tn)}$
DT	$O(n^2p)$		
RF	$O(n^2pk_{trees})$	Precision	$\frac{tp}{(tp+fp)}$
GBC	$O(npk_{trees})$		
ETC	$O(npk_{trees})$	F1 - Score	$2* \frac{Precision*Recall}{Precision+Recall}$
XBGC	$O(npk_{trees})$		
SNN	$O(n^3)$	AUC	$\frac{1}{2}*\frac{tp}{(tp+fn})+\frac{tn}{(tn+fp)}$
MLP	$O(n^2)$		

TABLE I
MODEL'S COMPUTATIONAL METRIC & PERFORMANCE METRICS

Models	AUC	Accu	Recall	Preci	Speci	F1	Ccost
SLR	0.6400	0.6640	0.5309	0.7234	0.7970	0.6124	**0.31**
LR	0.6278	0.6571	0.5303	0.7105	0.7839	0.6073	*0.46*
KNN	0.9939	0.6206	0.2417	0.9974	0.9994	0.3891	0.93
SGD	0.5558	0.5746	0.5059	0.5865	0.6433	0.5433	0.83
NB	*0.9858*	*0.9391*	*0.9019*	*0.9744*	*0.9763*	*0.9368*	*0.6*
DT	0.9837	0.9781	0.9575	0.9987	0.9969	0.9777	0.99
RF	0.9983	0.9700	0.9488	0.9909	0.9913	0.9694	0.78
GBC	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	1.05
ETC	1.0000	0.9972	0.9994	0.9950	0.9950	0.9972	0.86
XBGC	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	0.92
SNN	*0.9950*	*0.9872*	*0.9919*	*0.9827*	*0.9825*	*0.9873*	*0.61*
MLP	*0.9981*	*0.9916*	*0.9900*	*0.9931*	*0.9931*	*0.9916*	**0.04**

TABLE II
PERFORMANCE COMPARISON & COMPUTATIONAL COST

IV. CONCLUSIONS

From the different metrics, we can conclude that GBC and XGBC models have to be trained and used from the Cloud. NB and SNN are relevant for fog nodes, and MLP perfectly fits for edge nodes due to its good balance between a low computational cost and a high accuracy in seizure detection.

REFERENCES

[1] P. Sethi and S. R. Sarangi, "Internet of Things: Architectures, Protocols, and Applications", Int. J. Electr. Comput. Eng. 2017.
[2] P. H. Vilela, J. J. P. C. Rodrigues, P. Solic, K. Saleem, and V. Furtado, "Performance evaluation of a Fog-assisted IoT solution for e-Health applications", Future Gener. Comput. Syst., vol. 97, pp. 379–386, 2019.
[3] M. McClellan, C. Cervelló-Pastor, and S. Sallent, "Deep learning at the mobile edge: Opportunities for 5G networks", Appl. Sci., vol. 10, no. 14, 2020.
[4] F. Forooghifar, A. Aminifar, L. Cammoun, et al., "A Self-Aware Epilepsy Monitoring System for Real-Time Epileptic Seizure Detection", Mobile Netw., 2019.
[5] R. G. Andrzejak, K. Lehnertz, F. Mormann, et al., "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state", Phys. Rev. E, 64, 061907, 2001.
[6] A. Shoeb, "Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment", PhD Thesis, MIT, 2009.