

Classification algorithms for the recognition of partial discharge pulses by different sources

S. Licciardi¹, A. Di Fatta¹, *Student member, IEEE*, G. Ala¹, *Senior member, IEEE*
P. Romano¹, *Senior member, IEEE* and A. Imburgia¹, *Member, IEEE*

¹Engineering Department, University of Palermo, Viale delle Scienze, 90128, Palermo, Italy,
silvia.licciardi@unipa.it 0000-0003-4564-8866, alessio.difatta@unipa.it, guido.ala@unipa.it,
pietro.romano@unipa.it, antonino.imburgia@unipa.it

Partial discharges pulses are phenomena of localized ionization which occurs in the presence of high electric fields in limited portions of an insulating system, without leading to a complete electrical failure. These events can manifest themselves in different forms, depending on the location and configuration of the defect within the insulating material. The techniques traditionally used for interpreting the data collected include the use of time and frequency domain analysis, and are distinguished according to the nature of the voltage applied (AC or DC). Recognition and classification of discharges in DC still remains an open problem. In the present paper, Artificial learning is used to classify three different types of experimental partial discharges. Machine learning algorithms, including Decision Trees and Discriminant Analysis, and a neural network of the LSTM type have been implemented. The optimization of the hyperparameters of the various algorithms provided high levels of accuracy while maintaining low computational costs.

Index Terms—Machine learning, deep learning, LSTM neural networks, partial discharges pulses.

I. INTRODUCTION

In the field of Artificial intelligence (AI), an increasingly preponderant space has been conquered by Artificial learning (AL). Machine learning (ML) [1] and Deep Learning (DL) [2] are the most significant examples. Both methodologies are composed of algorithms that, through an iterative analysis of the input data (inputs/datasets), acquire the ability to “learn” the patterns in which these inputs are part of, the relationships or the significant structures that characterize them. Such algorithms are able to learn relationships between input and output, identify latent structures in data or choose optimal strategic criteria. The difference between ML and DL lies in the complexity of the patterns that they are able to recreate and identify, and consequently in the more complex structure of the ‘deep’, compared to the more simple machine learning. Both share targets, such as classification or regression, but the so-called artificial neural networks (ANN), of which the DL is composed, are computational models inspired by the workings of the human neural network and capable of extracting more complex information. While ML uses simpler algorithms, such as Decision Tree (DT), Discriminant Analysis (Linear and Quadratic) (DA), Logistic Regression (LR) and so on, DL makes use of many hidden layers (hidden layers), each of them can extract a different feature from the datasets according to the type of mathematical function used. ML requires preprocessing of the data by the user, not so the DL, since it automatically learns from raw data. Finally, ML is suitable for problems with moderate datasets while DL requires large amounts of data and computing power, being particularly effective for complex problems such as image, voice or natural

language recognition, beyond the analysis of time series (also feasible with ML but in a less complex form). Given these characteristics, the fields of application are innumerable [3]-[4], as well as the possibilities to combine NN architectures even in hybrid form [4]. This study will use a Long Short Term Memory (LSTM) neural network [5]-[6] for the recognition of partial discharges in the presence of high electric fields and will show a comparison with selected classification algorithms belonging to ML, derived from Numerical Analysis. Partial discharges (PD) are phenomena of localized ionization which occurs in the presence of high electric fields in limited portions of an insulating system. They can occur in different forms, depending on the location and configuration of the defect within the insulating material. The main types include internal (I), surface (S) and corona (C) discharges. The signals analysed in this study were obtained by measurements on test specimens made to simulate the three different discharge phenomena, as shown in Fig. 1. For the study and separation of these

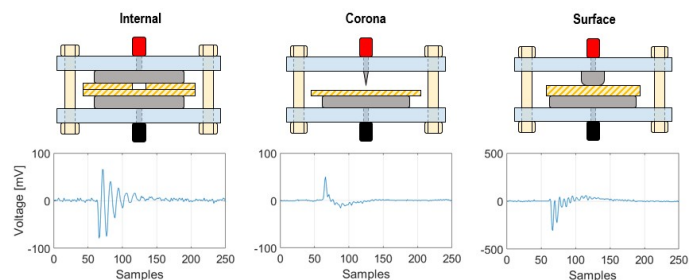


Fig. 1. Types of acquired partial discharge signals: internal, surface, corona.

types of discharges, an experimental setup was developed in compliance with the IEC 60270 standard, which serves as the international reference for the detection and characterization of PDs. The acquisition of discharges continued until a

dataset of approximately $2 \cdot 10^5$ signals was obtained, whose waveforms are also shown in Fig. 1. Recently, research has focused on the application of statistical methods to identify the key characteristics of partial discharges, as amplitude, duration of pulse shape and frequency of occurrence [7]-[8]. However, the recognition and classification of discharges in DC still remains a major problem. An AL-based approach can offer significant benefits. The neural networks of DL can for example automatically learn complex characteristics from data, making them particularly suitable for identifying nonlinear models and correlations that might escape traditional methods. This allows for a more accurate classification of defect types and conditions. In addition, NN can be used to develop predictive monitoring systems, improving reliability and reducing downtime by detecting anomalies or degradation trends early on. The methodologies used in this paper were implemented in Matlab software using a simple CPU and did not require parallelization. The simulations aim to minimize the computational costs associated with each chosen algorithm, always ensuring high levels of accuracy. This was achieved through a thorough analysis of the mathematical structure of the considered architectures and the optimization of the hyperparameters of selected ML algorithms as well as of the built DL network.

II. CLASSIFICATION ALGORITHMS

The neural network used for the classification of the experimental data described above is a unidirectional Long short term memory (LSTM), particularly suitable for time signals recognition due to its intrinsic structure. It is a recurrent neural network (RNN) that avoids the problem of the reversal or explosion of the gradient (see [6] for the details and a complete explanation of the used and optimized hyperparameters meaning). In the following table we report, after several simulations, the most efficient results, in terms of accuracy and training time, obtained by comparing the use of LSTM and selected ML methods (Decision Tree (Fine and Optimizable), Discriminant Analysis (Quadratic and Optimizable)), chosen for the lower computational costs at the same performance. The setting of the hyperparameters is a determining factor for performance; at the end of the simulations, the following levels have been set for the LSTM: number of layers (5), state activation function (tanh), gate activation function (sigmoid), initial learning rate (10^{-3}), optimization methods (Adam), mini batch size (512), number of epochs (1 or 2), loss function (crossentropy), shuffle applied to every epoch, initialization of weights via Gloriot method and other model parameters such as the initial number of hidden states (200) or the initial bias (0). Dataset consists of 107.400 units of evenly distributed partial discharges, of which 90% used for training, 4% for validation and 6% for testing. In Fig. 2 confusion matrices of a) DT (fine) and b) LSTM are given. They were obtained from experimental signals (only I and S) on which the algorithms had not been trained. In the main diagonal the correct assignments.

III. RESULTS AND CONCLUSION

The simulations carried out have been optimal in terms of classification of the signals used. All the algorithms uti-

TABLE I
CLASSIFICATION ALGORITHMS - RESULTS

Parameters	DT (F)	DT (O)	DA (Q)	DA (O)	LSTM
Accuracy	99,9%	100%	99,8%	99,8%	100%
Training Time (min)	0,11	1,46	0,23	15,09	9,53

Comparison of the performance of Decision Tree (Fine and Optimizable), Discriminant Analysis (Quadratic and Optimizable) and LSTM.

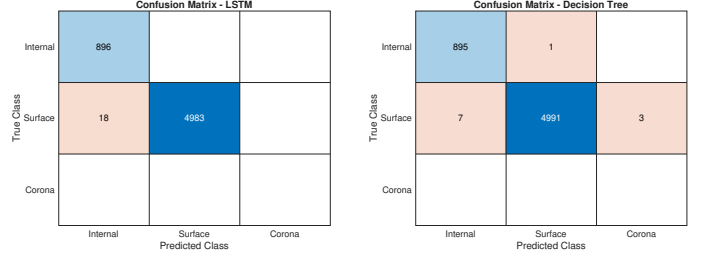


Fig. 2. Confusion matrices of 1° LSTM neural network and 2° Decision tree (fine), obtained of signals not belonging to the dataset.

lized were able to classify, with an accuracy between 99.8% and 100%, the partial discharges of the trainset and testset. The computational times of some selected ML algorithms are considerably shorter but the LSTM did not require a preprocessing of the data. Furthermore, the provided time indicates the training time of the NN but, once trained, the prediction is instantaneous. The studies in progress are aimed at ascertaining the generalization capacity of the algorithms under consideration. The aim is to investigate to what extent the neural network and machine learning algorithms are able to generalize the results obtained, namely correctly classifying a type of signal which, although comparable to the categories considered, differs considerably both in absolute terms and in terms of trend.

REFERENCES

- [1] E. Alpaydin, *Introduction to Machine Learning*, 3rd ed., Adaptive Computation and Machine Learning; MIT Press: Cambridge, MA, USA, 2014.
- [2] I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*, MIT press, 2016, <http://www.deeplearningbook.org>.
- [3] A. Entezari, A. Aslani, R. Zahedi, Y. Noorollahi, Artificial intelligence and machine learning in energy systems: A bibliographic perspective, *Energy Strategy Reviews*, vol. 45, 2023, 101017.
- [4] S. Licciardi, G. Ala, E. Francomano et al., "Neural Network Architectures and Magnetic Hysteresis, overview and comparisons", *Mathematics*, MDPI, 12, 3363, 2024, Special Issue Mathematical Applications in Electrical Engineering.
- [5] Y. Bengio, N. Boulanger-Lewandowski, R. Pascanu, "Advances in optimizing recurrent networks", *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, 2013, pp. 8624–8628.
- [6] G. Ala, P. Catrini, M. Ippolito, M. La Villetta, S. Licciardi et al., "Deep Learning for Smart Grid and Energy Context", In *Proceedings of Asia Meeting on Environment and Electrical Engineering (EEE-AM)*, Hanoi, Vietnam, (November 13-15 2023), pp. 1–6.
- [7] A. Imburgia, A. Di Fatta, P. Romano, G. Rizzo, V.L. Vigni, G. Ala, "A Study on Partial Discharges Pattern Recognition Under DC Voltage Through Clustering Algorithms and Cross Correlation Filter," in *IEEE Trans. Diel. Electr. Insulat.*, vol. 30, no. 6, pp. 2543–2550, Dec. 2023.
- [8] A. Di Fatta et al., "Partial Discharge Separation by Using Pulses Cross-Correlation," in *2023 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP)*, East Rutherford, NJ, USA, 2023, pp. 1–4.