

Article

Not peer-reviewed version

Machine Learning Approach to Shield Optimization at Muon Collider

[Luca Castelli](#) *

Posted Date: 28 January 2025

doi: 10.20944/preprints202501.2125.v1

Keywords: Muon; Collider; MDI; Machine-Learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Machine Learning Approach to Shield Optimization at Muon Collider

Luca Castelli ^{1,2} ¹ Dipartimento di Fisica, Sapienza Università di Roma; luca.castelli@uniroma1.it² INFN Sezione di Roma

Abstract: Muon collisions are considered a promising mean for exploring the energy frontier, leading to a detailed study of the possible feasibility issues. Beam intensities of the order of 10^{12} muons per bunch are needed to achieve the necessary luminosity, generating a high flux of secondary and tertiary particles from muons decay that reach both the machine elements and the detector region. To limit the impact of this background on the physics performance tungsten shieldings have been studied. A machine learning-based approach to the geometry optimization of these shieldings will be discussed.

Keywords: muon; collider; MDI; machine-learning

1. Introduction

The Muon Collider is a proposed next-generation facility in high-energy physics [1]. Muons, as fundamental particles in the lepton family, are particularly suited for collider applications due to their relatively large mass, about 200 times heavier than electrons, combined with their point-like nature. This makes muons less prone to radiative energy losses compared to electrons when accelerated to high energies, allowing to build a circular facility to explore the multi-TeV center-of-mass energy region. However, one of the primary challenges associated with the Muon Collider arises from the intrinsic instability of muons. They decay in the beam pipe, and the resulting decay products interact with the collider components, generating a flux of photons, neutrons, and charged particles that reach the detector, known as Beam-Induced Background (BIB). Effective mitigation strategies are essential to reduce BIB and ensure the desired detector performance and measurement accuracy. To this purpose a tungsten cone-shaped shield (Nozzle) has been designed inside the detector to absorb the incoming particles and reduce their energy. The original geometry of this object was proposed by MAP [2] for a $\sqrt{s} = 1.5$ TeV machine, therefore optimizations for the $\sqrt{s} = 3, 10$ TeV Muon Collider are required. This contribution aims to describe the activities undertaken and ongoing concerning the Nozzle geometry optimization for the 3 TeV Muon Collider exploiting machine learning (ML) algorithms.

2. Beam-Induced Background

At $\sqrt{s} = 3$ TeV, the muon decay rate is expected to reach 2.34×10^5 decays per meter per beam for each bunch crossing. This results in a significant flux of particles, estimated to be on the order of 10^8 per bunch crossing, widely distributed in time relative to the bunch crossing. These particles, known as Beam-Induced Background (BIB), pose a considerable challenge to the detector, particularly the vertex detector near the Interaction Point (IP).

To mitigate the impact of BIB, a nozzle design based on the MAP configuration has been developed. The nozzle features a cone-shaped tungsten core extending from a few centimeters from the IP to the edge of the detector, forming a 6-meter-long structure. Near the IP, the nozzle's angular aperture is 10° to address the critical impact of BIB on the vertex detector, narrowing to 5° at a distance of 1 meter. To further reduce neutron production from interactions within the tungsten core, the nozzle is coated with borated polyethylene. The detector layout and nozzle geometry are illustrated in Figure 1.

The time of arrival of BIB particles at the detector surface is highly dispersed relative to the bunch crossing. A time-based rejection strategy is employed, using a readout window from -1 ns to $+15$ ns around the bunch crossing to significantly reduce off-time background events. The composition of the BIB before and after time discrimination is detailed in Table 1, where the first column shows the initial particle flux and the second column the flux remaining after time rejection. Additionally, Figure 2 displays the full time distribution and energy spectra for each particle species, based on the MAP configuration.

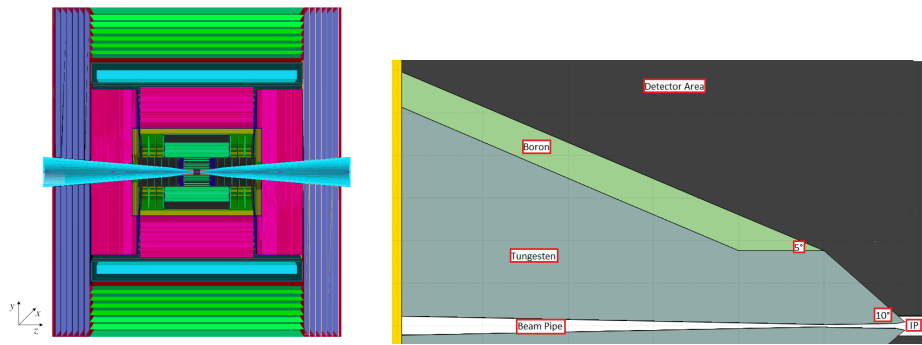


Figure 1. Left: The detector used for $\sqrt{s} = 3$ TeV Muon Collider studies. The nozzles are shown in cyan. Right: Section of left nozzle.

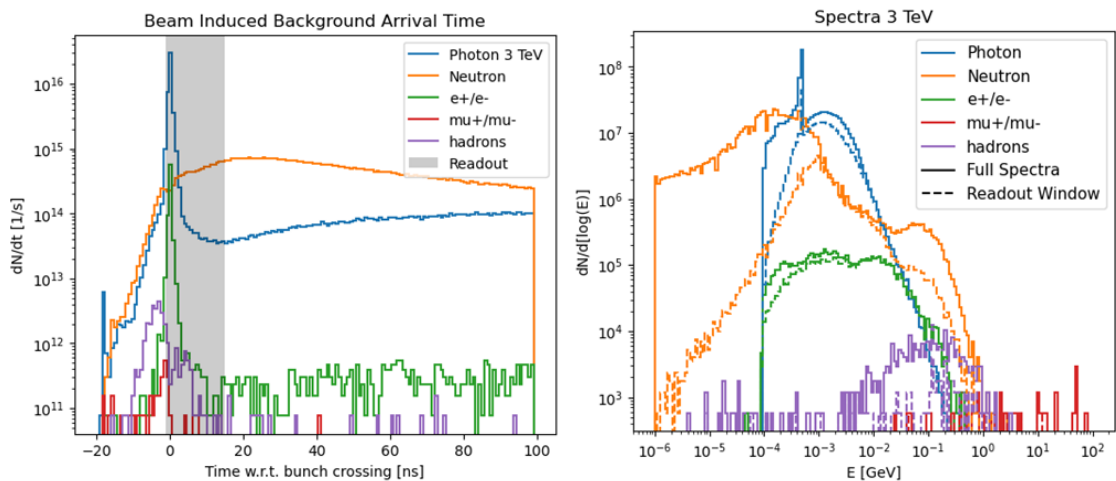


Figure 2. Left: Arrival times of the main BIB components at the detector relative to the bunch crossing. The light gray band between -1 and 15 ns indicates the assumed readout window for the detector. Right: Energy spectra of the BIB particles that reach the detector within the readout window. The solid and dashed lines refer to the center-of-mass energy of 3 TeV and 1.5 TeV, respectively.

Table 1. Particle species per Bunch Crossing contributing to the Beam-Induced Background at $\sqrt{s} = 3$ TeV with the MAP. First column consider the whole flux entering in the detector. Second column considers the application on a Readout Window of $[-1, 15]$ ns with respect to bunch crossing.

Particle Species	Flux	Flux - R.W.
Photons	8.48×10^7	4.22×10^7
Neutrons	7.35×10^7	7.18×10^6
Electron/Positron	7.00×10^5	5.75×10^5
Hadrons	2.43×10^4	4.91×10^3
Muons	2.42×10^3	6.76×10^2

The simulation of the BIB which led to the plots shown in Figure 2 has been achieved by implementing the design in Figure 1 (right) together with the beam pipe and all the machine elements in FLUKA [3,4], exploiting the two step simulation:

- **Step 1:** Muon decays along the beam pipe are simulated and the interaction with the environment are propagated up to the detector area. There the incoming particles are killed in the simulation and their kinematic information are scored.
- **Step 2:** Particles arriving on the detector surface are propagated inside the nozzles. The flux surviving the interactions inside the Nozzles and reaching the detector is scored as the Beam-Induced Background.

This approach allows saving a significant amount of time since it is not necessary to repeat the first step each time the nozzles geometry changes, considering that only the second step of the simulation takes on average 4 days for a 1.6% of the expected decays for a bunch crossing, considering only one beam.

3. Nozzle Geometry Optimization

As mentioned in the Introduction, the nozzle geometry was originally designed for a $\sqrt{s} = 1.5$ TeV machine. Therefore, it must be adapted to the energy stages currently envisioned for the Muon Collider, namely 3 and 10 TeV in the center of mass. The primary goal of this adaptation is to minimize the Beam-Induced Background (BIB) in the detector, thereby reducing its impact on event reconstruction. Additionally, reducing the nozzle size offers several advantages, particularly cost-wise. Given that tungsten has a density of approximately 19 g/cm^3 , depending on the alloy, even a slight reduction in nozzle volume can lead to a substantial decrease in material requirements. This reduction also lowers the overall weight that needs to be supported, simplifying the design of the nozzle's supporting structure. Finally, decreasing the nozzle volume creates more space for detectors, thereby enhancing the acceptance of the entire experiment. This work discusses the optimization achieved and the studies ongoing for the $\sqrt{s} = 3 \text{ TeV}$ stage.

3.1. FLUKA Simulation

The first step in the optimization process focused on identifying how the geometry influences the resulting Beam-Induced Background (BIB). To this end, several FLUKA simulations were conducted, varying three key geometrical parameters: the angle of the nozzle tip, the position along the beam axis where the aperture changes, and the base radius. Figure 3 illustrates these parameters.

The simulations used a sample size corresponding to 1.6% of the expected decays, approximately 10^4 muons. This sample size ensures a negligible statistical error in the flux of particles entering the detector surface. Figure 4 presents the composition of the BIB for the original nozzle design compared to a tested configuration. For photons, neutrons, and e^+ / e^- , the statistical error is negligible.

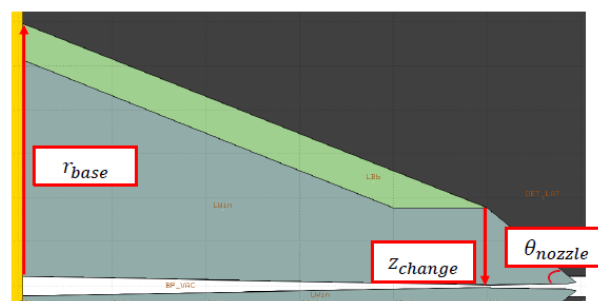


Figure 3. Geometrical parameters changed during the first studies on the Beam-Induced Background.

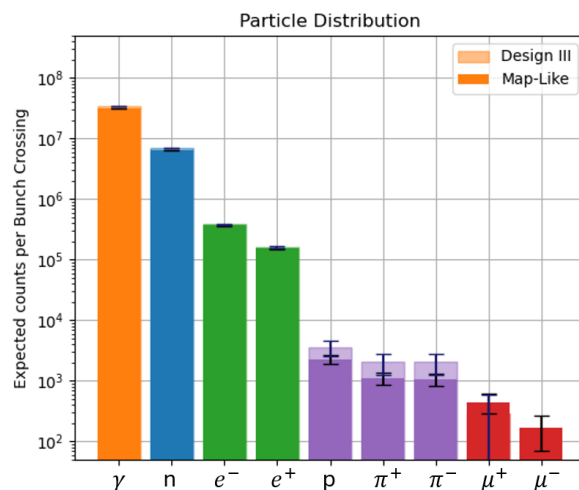


Figure 4. Components of the Beam-Induced Background with the statistical error obtained with the original nozzle design (solid) and a test configuration (faded).

The results indicated that careful optimization of the last meter of the nozzle is critical, whereas material savings can be achieved further from the Interaction Point (IP). However, due to the sensitivity of the geometry—where even small variations result in significant increases in Beam-Induced Background (BIB)—no definitive guidelines for modifications could be established.

Additionally, an assessment of the borated polyethylene coating was made after analyzing the neutron fluence distribution shown in Figure 5. Observations revealed that the highest concentration of particles does not occur on the external surface. Consequently, it was proposed to relocate the coating internally, placing a layer of tungsten on the outside. This configuration would help block photons emitted during neutron capture in the boron.

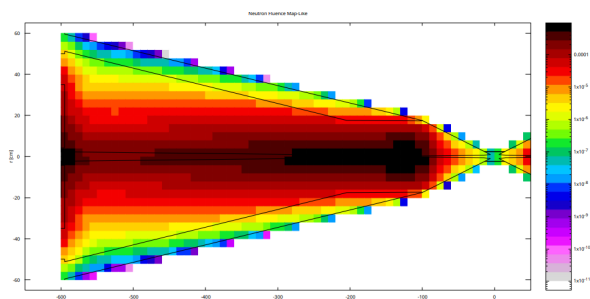


Figure 5. Neutron fluence in the nozzle. The highest concentration of particle is concentrated in the internal part.

3.2. Machine Learning Studies

Manual optimization achieved by iterating FLUKA simulations is not the most effective strategy for finalizing the nozzle design due to the extensive computational time required for a single simulation. To address this challenge, an alternative approach leveraging ML algorithms has been adopted.

Building on the insights from the studies summarized in the previous section, a set of geometrical parameters has been identified to redefine the nozzle geometry. Specifically, the borated polyethylene has been incorporated inside the tungsten, and a step has been introduced to reduce material usage in areas where it is not required. The defined parameters are listed below, and the new nozzle design is illustrated in Figure 6:

- r_{base} : Radius of the nozzle at the end of the final focusing stage.
- r_{boron} : Outer radius of the borated polyethylene section of the nozzle, expressed as a fraction of r_{base} .
- z_{step} : Position along the beam axis where the nozzle size is reduced.
- r_{step} : Magnitude of the step, defined as a fraction of the nozzle's radial size at position z_{step} .

- z_{change} : Position along the beam axis where the angular aperture of the nozzle changes.
- θ_{tip} : Angular aperture at the tip of the nozzle.
- z_{tip} : Position along the beam axis of the nozzle tip.
- r_{tip} : Radial coordinate of the nozzle tip.

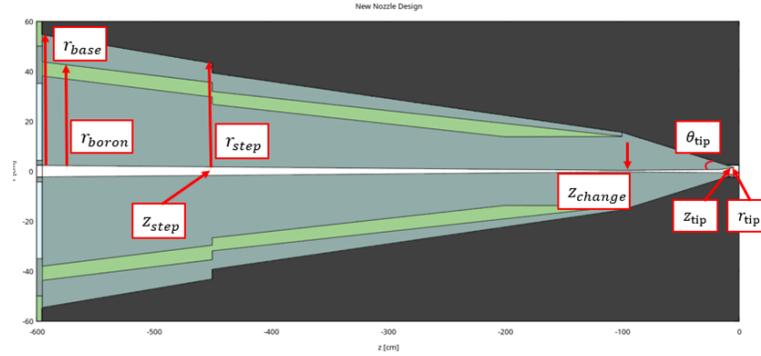


Figure 6. Geometry layout of the nozzle considered for the Machine-Learning studies. In figure are reported the parameters described in text section.

The first step in this study was to construct a dataset for training the ML algorithms. Given the objective of significantly reducing computational time, the simulated decays were limited to only 0.02% of a bunch crossing. This approach enabled the creation of a dataset comprising $1.3 \cdot 10^4$ elements within a reasonable timeframe while maintaining adequate statistical accuracy.

Figure 7 compares the photon energy spectrum of the Beam-Induced Background obtained using the original nozzle design with 1.6% decays simulated (blue) to that of a design studied with ML algorithms and 0.02% decays simulated (black). The difference in the error bar size is noticeable primarily in the higher-energy region of the spectrum, whereas the relative error in the most densely populated energy range remains around 10^{-4} .

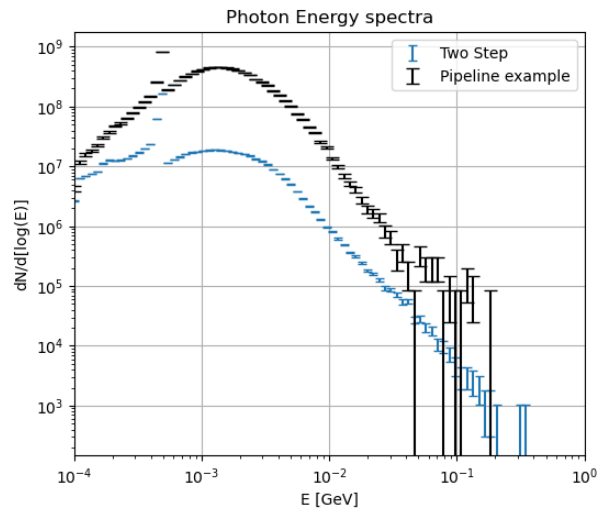


Figure 7. Photon energy spectrum obtained with original nozzle design and 1.6% of expected decay simulated (blue) compared to one of the $1.3 \cdot 10^4$ low statistics simulation used for the ML studies (black). The absolute number of photon differs because different configurations are considered.

The second step of this study aimed to develop a model capable of parametrizing the particle flux entering the detector surface as a function of the geometrical parameters listed above. The integrated flux alone is not an ideal metric, as different particle types impact various detector components differently, and their flux does not vary coherently with changes in nozzle geometry. However, evaluating the effect of each geometry on individual sub-detectors would significantly increase the simulation time, counteracting the primary goal of this study.

To address this challenge, Deep Neural Networks [5], Random Forest [6], and XGBoost [7] regressors were evaluated. Data transformation techniques were also explored, given that the parameters span different phase spaces and small variations in some parameters can lead to large changes in flux. This behavior can negatively affect the performance of ML algorithms.

The performance of the models was assessed by measuring the relative deviation between the predicted flux and the simulated flux on a test dataset, defined as:

$$\Delta[\%] = \frac{Flux_{sim} - Flux_{pred}}{Flux_{sim}} \cdot 100 \quad (1)$$

where $Flux_{sim}$ is the flux obtained from low-statistics FLUKA simulations, and $Flux_{pred}$ is the flux predicted by the trained model.

Among the tested models, XGBoost paired with the Standard Scaler [8] applied to the dataset features (i.e., the parameters) produced the best results. The performance of the model was quantified by fitting the Δ distribution on the test dataset with a Gaussian function, yielding an expected value of $\bar{\Delta} = -0.12\%$ with a standard deviation of $\sigma = 5.24\%$. This demonstrated that the XGBoost model could reliably serve as a surrogate for the FLUKA simulation. Using the trained model, a pseudo-dataset was generated to identify parameter sets that simultaneously minimized the Beam-Induced Background (BIB) and the total volume of the nozzle. The optimized geometry, compared to the original design (Figure 8), was subsequently tested using high-statistics FLUKA simulations.

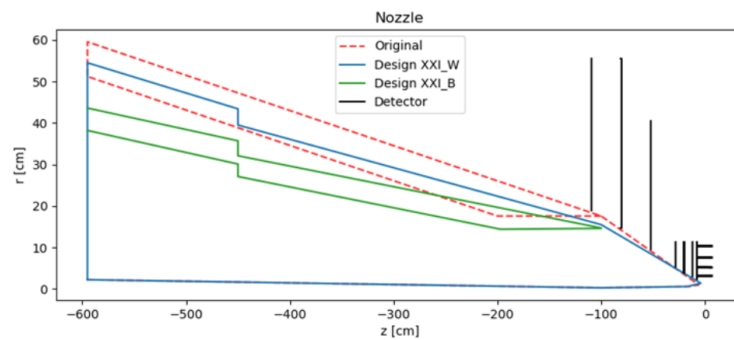


Figure 8. Comparison of the optimized nozzle geometry (blue) featuring the repositioned borated polyethylene (green) with the original design (dashed red). The vertex and tracker detector layers are also depicted (black).

The new design reduced the photon flux by 34% and the e^+/e^- flux by 27% at the cost of an 86% increase in the neutron flux. This trade-off is considered favorable because the primary challenge for event reconstruction is the occupancy in the tracking system, which is primarily affected by the e^+/e^- flux and is improved with the optimized nozzle. Furthermore, the hadronic calorimeter, which is impacted by the neutron flux, is located farther from the interaction point, and the neutron flux is distributed over a larger surface area, mitigating its overall impact. Figure 9 reports the energy spectrum of the Beam-Induced Background comparing the new design to the original one.

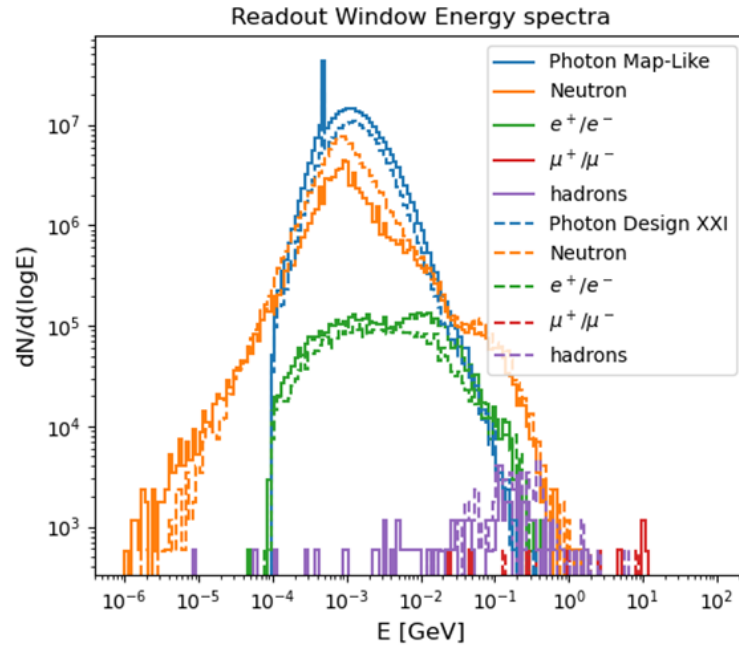


Figure 9. Energy spectrum of the Beam-Induced Background obtained with the high-statics FLUKA simulation and the optimized design (dashed) compared to the original geometry (solid). The results are obtained after the application of a Readout window of $[-1, +15]$ ns with respect to the bunch crossing.

4. Further Studies

While an initial optimization of the nozzle geometry has been achieved, more advanced ML methods may enhance the design further. The approach described in the previous section has its limitations, as the pseudo-dataset must be manually analyzed to identify an optimal configuration. To address this constraint, two parallel approaches are currently under investigation.

The first approach involves implementing a Bayesian Optimization Loop [9], which iteratively performs FLUKA simulations based on informed predictions of optimal parameters. The second approach leverages a Deep Neural Network (DNN) to approximate the XGBoost model and, by extension, the FLUKA simulations. The goal of this approach is to obtain a differentiable algorithm (the DNN), enabling the application of the Gradient Descent method [10] to analytically determine the configuration that minimizes the chosen metric.

While the Bayesian optimization method is still under study, the DNN-based approach has already been tested. However, it did not yield an improved algorithm because the metric used in the study was the integrated Beam-Induced Background (BIB) flux in the detector. The optimization process led to the selection of the largest possible geometry within the given parameter space, which was suboptimal.

To effectively exploit this method, a more nuanced metric M is required, defined as:

$$M = a \cdot \gamma + b \cdot n + c \cdot e + d \cdot V \quad (2)$$

Here:

- γ, n, e : the total fluxes of photons, neutrons, and electrons, respectively, in the detector;
- V : the total volume of the nozzle;
- a, b, c : parameters reflecting the respective impacts of photons, neutrons, and electrons on the sub-detectors;
- d : a parameter related to the nozzle's total volume.

Ongoing studies aim to evaluate the parameters a, b, c , and d to accurately reflect the trade-offs between particle fluxes and nozzle volume.

References

1. Accettura, C.; et al. Towards a Muon Collider. *Eur. Phys. J. C* **2023**, *83*, [arXiv:2303.08533]. <https://doi.org/https://doi.org/10.1140/epjc/s10052-023-11889-x>.
2. Alexahin, Y.; Gianfelice-Wendt, E.; Kapin, V. Muon Collider lattice concepts. *Journal of Instrumentation* **2018**, *13*, P11002. <https://doi.org/10.1088/1748-0221/13/11/P11002>.
3. Battistoni, G.; Boehlen, T.; Cerutti, F.; Chin, P.W.; Esposito, L.S.; Fassò, A.; Ferrari, A.; Lechner, A.; Empl, A.; Mairani, A.; et al. Overview of the FLUKA code. *Annals of Nuclear Energy* **2015**, *82*, 10–18. Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2013, SNA + MC 2013. Pluri- and Trans-disciplinarity, Towards New Modeling and Numerical Simulation Paradigms, <https://doi.org/https://doi.org/10.1016/j.anucene.2014.11.007>.
4. Ahdida, C.; Bozzato, D.; Calzolari, D.; Cerutti, F.; Charitonidis, N.; Cimmino, A.; Coronetti, A.; D'Alessandro, G.L.; Donadon Servelle, A.; Esposito, L.S.; et al. New Capabilities of the FLUKA Multi-Purpose Code. *Frontiers in Physics* **2022**, *9*. <https://doi.org/10.3389/fphy.2021.788253>.
5. Yi, H.; Shiyu, S.; Xiusheng, D.; Zhigang, C. A study on Deep Neural Networks framework. In Proceedings of the 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016, pp. 1519–1522. <https://doi.org/10.1109/IMCEC.2016.7867471>.
6. Breiman, L. Random Forests. *Machine Learning* **2001**, *45*, 5–32.
7. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016, KDD '16, p. 785–794. <https://doi.org/10.1145/2939672.2939785>.
8. de Amorim, L.B.; Cavalcanti, G.D.; Cruz, R.M. The choice of scaling technique matters for classification performance. *Applied Soft Computing* **2023**, *133*, 109924. <https://doi.org/10.1016/j.asoc.2022.109924>.
9. Frazier, P.I. A Tutorial on Bayesian Optimization, 2018, [arXiv:stat.ML/1807.02811].
10. Ruder, S. An overview of gradient descent optimization algorithms, 2017, [arXiv:cs.LG/1609.04747].

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.