

Applications of Machine Learning in Constraining Multi-Scalar Models

Darius Jurčiukonis

Vilnius University, Institute of Theoretical Physics and Astronomy

2025-06-18, EuCAIFCon 2025



ABSTRACT
Machine learning techniques are used to predict theoretical constraints such as unitarity, boundedness from below, and the global minimum in multi-scalar models. This approach has been demonstrated to be effective when applied to various extensions of the Standard Model that incorporate additional scalar multiplets. A high level of predictivity is achieved through appropriate neural network architectures, learning algorithms, and well-prepared training datasets. Machine learning offers a significant computational advantage by enabling faster computations compared to other numerical methods, such as scalar potential minimization. This research investigates the potential of machine learning as an alternative approach for predicting these constraints, potentially improving upon traditional numerical techniques.

DETAILS OF THE COMPUTATIONS

For this research, a desktop computer was used.

System:

- CPU: Intel® Core™ i9-13900K, 24 Cores (8P+16E) 23.5 S.S. GHz
- GPU: NVIDIA GeForce RTX 4090 (24GB GDDR6X)
- RAM: G.Skill 64GB 2x 32GB DDR5 6000MHz
- SOFT: Ubuntu 22.04 LTS, Wolfram Mathematica 13.2

For network training, we utilized a GPU, whereas other computations were performed using a CPU. Given the prevalent use of Wolfram Mathematica in the HEP community, we chose to conduct our machine learning and other computations within Mathematica to ensure seamless integration with other HEP packages.

ML parameters:

- training data: 10^7 examples
- batch size: $2^{14} = 16384$
- training rounds: 500
- neural network: linear nets of 8 layers

The initial network, net-1, is trained using raw data supplemented with an appropriate number of true samples. This network, net-1, is subsequently used to prepare the training data for net-2. Both net-3 and net-4 are trained using the same data as net-2, but they employ larger matrices. These networks, net-3 and net-4, are then utilized to filter the predictions made by net-2.

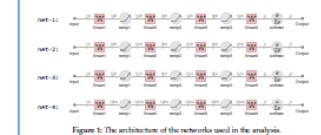


Figure 1: The architecture of the networks used in the analysis.

In the scatter plots, the blue points represent correct computations based on analytical UNI-BFB conditions. The green points depict neural network predictions using the combined networks from net-2 to net-4, while the red points indicate false examples provided by nets.

GENERAL 2HDM

The most general scalar potential for the 2HDM is

$$V = \mu_1^2 \phi_1^\dagger \phi_1 + \mu_2^2 \phi_2^\dagger \phi_2 + (\lambda_1 \phi_1^\dagger \phi_1 + \lambda_2 \phi_2^\dagger \phi_2) + \frac{1}{2} (\phi_1^\dagger \phi_1)^2 + \frac{1}{2} (\phi_2^\dagger \phi_2)^2 + \lambda_3 \phi_1^\dagger \phi_1 \phi_2^\dagger \phi_2 + \lambda_4 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \frac{1}{2} (\phi_1^\dagger \phi_2 + \phi_2^\dagger \phi_1)^2 + \lambda_5 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \lambda_6 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \text{H.c.}$$

where $\mu_{1,2}$ and $\lambda_{1,2,3,4,5,6}$ are real, while the remaining parameters are complex. The quartic part of the scalar potential comprises 10 real parameters [1].

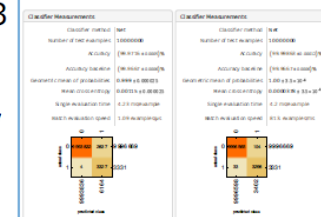


Figure 2: Classifier measurements for the net-1 (left) and net-2 (right).

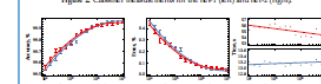


Figure 3: Classifier measurements (accuracy, error and prediction time) as a function from the size of the training data. Blue points represent net-1, while red points correspond to measurements from net-2.

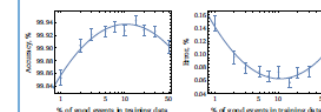


Figure 4: Dependence of net-3 accuracy and error on the percentage of true samples in the new training data.

GENERAL 2HDM

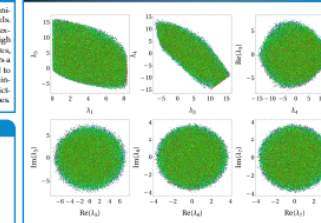


Figure 5: Scatter plots of lambda values for the G2HDM.

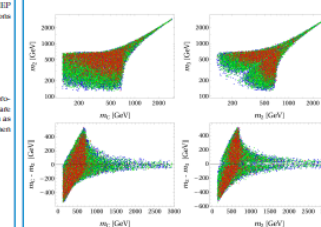


Figure 6: Scatter plots of Higgs masses for the G2HDM.

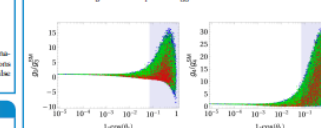


Figure 7: Scatter plots of Higgs trilinear and trilinear couplings for the G2HDM.

SM + QUADRUPLT

The scalar potential in the SM + 4plet [2] is of the form $V = V_{SM} + V_{4plet}$, where

$$V_{SM} = \mu_1^2 \phi_1^\dagger \phi_1 + \frac{1}{2} (\phi_1^\dagger \phi_1)^2 + \mu_2^2 \phi_2^\dagger \phi_2 + \frac{1}{2} (\phi_2^\dagger \phi_2)^2 + \lambda_1 \phi_1^\dagger \phi_1 \phi_2^\dagger \phi_2 + \lambda_2 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \lambda_3 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \text{H.c.}$$

where ϕ_1 are SU(2)-invariant polynomials of the SM fields Φ and of the fields from the additional multiplet Δ ; $\mu_{1,2}$ and $\lambda_{1,2,3}$ are real parameters. The quartic part of the scalar potential comprises 3 real parameters.



Figure 8: Classifier measurements for the net-1 (left) and net-2 (right).

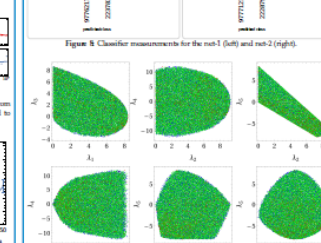


Figure 9: Scatter plots of lambda values for the SM + quadruplet.

SM + SIXPLET

The scalar potential in the SM + 6plet [2] is of the form $V = V_{SM} + V_{6plet}$, where

$$V_{SM} = \mu_1^2 \phi_1^\dagger \phi_1 + \frac{1}{2} (\phi_1^\dagger \phi_1)^2 + \mu_2^2 \phi_2^\dagger \phi_2 + \frac{1}{2} (\phi_2^\dagger \phi_2)^2 + \lambda_1 \phi_1^\dagger \phi_1 \phi_2^\dagger \phi_2 + \lambda_2 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \lambda_3 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \text{H.c.}$$

where $\mu_{1,2}$ and $\lambda_{1,2,3}$ are real parameters. The quartic part of the scalar potential comprises 6 real parameters.

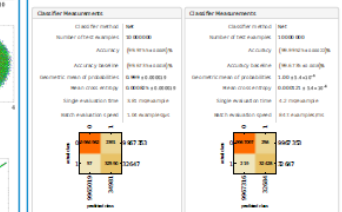


Figure 10: Classifier measurements for the net-1 (left) and net-2 (right).

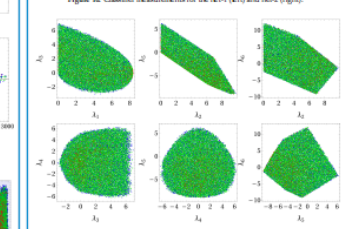


Figure 11: Scatter plots of lambda values for the SM + sixplet.

RESULTS

Table 1: Percentage of true samples for which the neural networks are generally randomly (new data). The "BFB" column displays the percentage of true samples from new data, while the "BFB-B" column indicates the percentage from datasets refined by UNI conditions. The aligned three Higgs boson model (A3HDM) and the CP-conserving left-right model (L3RM) were previously studied in Ref. [3].

Model	UNI	BFB-A	BFB-B	UNI-BFB
G2HDM (10)	1.02	23.3	5.25	0.03
SM + 4plet (5)	12.8	30.0	17.3	2.2
SM + 6plet (5)	5.4	22.2	6.07	0.33
L3RM (10)	0.41	18.4	0.7	0.0028
A3HDM (10)	0.18	1.73	0.05	0.00009

Table 2: Computation times (in seconds) needed to identify 1000 true samples. The second column presents the computation time employing both UNI and global minimization methods. The third column depicts the time taken using only neural networks. The fourth column outlines the time required when utilizing neural networks, followed by a check using global minimization.

Model	UNI+min	neural nets	neural nets+min	ratio-1	ratio-2
G2HDM	131	1.7	35	77	3.7
SM + 4plet	53	0.12	19	441	2.7
SM + 6plet	496	0.76	45	653	11
L3RM	2224	18	78	125.5	105.5
A3HDM	61800	1090	1170	56.7	52.6

Table 3: Percentage of true samples within predicted bounds, worked using precise analytical UNI-BFB conditions.

Model	net-1	net-2	net-3	net-4
G2HDM	52.58	96.97	97.98	> 99.0
SM + 4plet	98.99	98.8.99.1	99.2	> 99.5
SM + 6plet	92.94	98.8.99.2	99.0	> 99.5
L3RM	42.47	97.762	97.93	> 98
A3HDM	13.15	85.88	89.92	> 97

Table 3: Percentage of true samples within predicted bounds, worked using precise analytical UNI-BFB conditions.

CONCLUSIONS

- This analysis demonstrates that machine learning techniques can effectively predict UNI and BFB constraints in multi-scalar models.
- Simple linear networks can achieve high prediction accuracy, though they require appropriately prepared and sizeable training data samples.
- Machine learning techniques can significantly reduce computing time in comparison to the global minimization technique.

ACKNOWLEDGEMENTS

This work has received funding from the Research Council of Lithuania (LMTLT) under Contract No. S-CERN-242.

REFERENCES

- [1] D. Jurčiukonis and I. Lavoura, JHEP 12 (2018) 044, [arXiv:1807.04244].
- [2] A. Mileev, D. Jurčiukonis and I. Lavoura, arXiv:2105.01272.
- [3] D. Jurčiukonis, PUS EPS-HEP2023 (2024) 404, [arXiv:2401.03130].
- [4] D. Jurčiukonis and I. Lavoura, JHEP 07 (2023) 166, [arXiv:2303.16633].
- [5] D. Rucka, D. Jurčiukonis and I. Lavoura, Nucl. Phys. B 966 (2023) 116207, [arXiv:2310.12073].

Introduction

There are two important theoretical constraints that one must impose on the scalar potential.

- **UNI**: all the (tree-level) scalar–scalar scattering amplitudes must respect **unitarity**.
- **BFB**: the potential **must have a minimum**, viz. they prevent the existence of directions in field space along which the potential is unbounded from below.

The simple scalar potential in the SM + multiplet is of the form

$V = V_{\text{SM}} + V_{\text{multiplet}}$, where

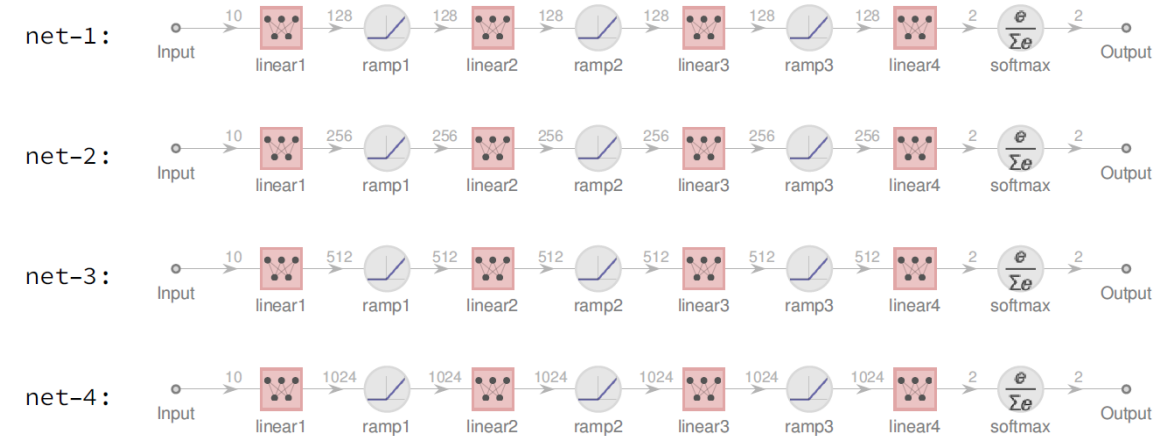
$$V_{\text{SM}} = \mu_1^2 F_1 + \frac{\lambda_1}{2} F_1^2,$$

$$V_{\text{multiplet}} = \mu_2^2 F_2 + \frac{\lambda_2}{2} F_2^2 + \lambda_3 F_1 F_2 + \lambda_4 F_4 + \sum_{i=5}^{t+3} \lambda_i F_i,$$

where F_i are $SU(2)$ -invariant polynomials of the SM fields Φ and of the fields from the additional multiplet Δ_n .

A. Milagre, DJ, L. Lavoura, 2505.05272

Technique



- The initial network, net-1, is trained using raw data.
- This network, net-1, is subsequently used to prepare the training data for net-2.
- Both net-3 and net-4 are trained using the same data as net-2.
- These networks, net-3 and net-4, are then utilized to filter the predictions made by net-2.

Results

Computation times (in seconds) needed to identify 1000 true samples.

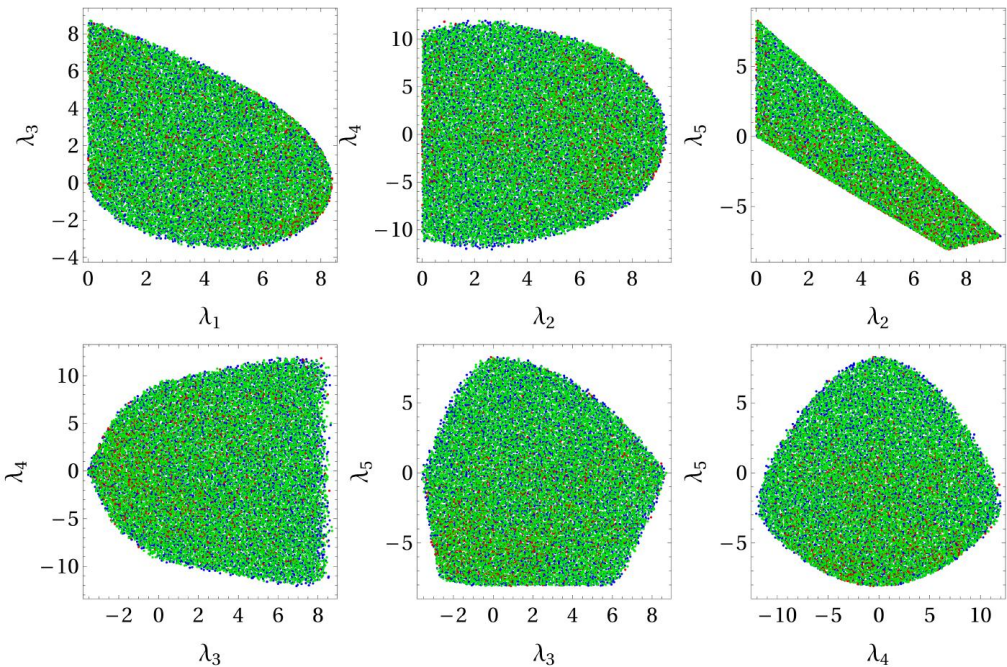
Model	UNI+min.	neural nets	neural nets+min.	ratio-I	ratio-II
G2HDM	131	1.7	35	77	3.7
SM + 4plet	53	0.12	19	441	2.7
SM + 6plet	496	0.76	45	653	11
LRM	2224	18	78	123.5	28.5
A3HDM	61800	1090	1170	56.7	52.8

Percentage of true samples within predicted results (from raw data), verified using analytical UNI+BFB conditions.

Model	net-1	net-2	net-3,4	net-2 – net-4
G2HDM	52-55	96-97	97-98	> 99.0
SM + 4plet	98-99	98.8-99.1	~ 99.2	> 99.5
SM + 6plet	92-94	98.8-99.2	~ 99.0	> 99.5
LRM	42-47	91-92	91-93	> 98
A3HDM	13-15	85-88	89-92	> 97

This work has received funding from the Research Council of Lithuania (LMTLT) under Contract No. S-CERN-24-2.

Conclusions



- Machine learning techniques can effectively predict UNI and BFB constraints in multi-scalar models.
- Simple linear networks can achieve high prediction accuracy, though they require appropriately prepared training data samples.
- Machine learning techniques can significantly reduce computing time in comparison to the global minimization technique.