Search for New Physics Using Machine Learning Techniques

Imperial College London

Oskar Högberg and Jakub Pazio
Supervisor: Mark Smith, High Energy Physics Group
Blackett Laboratory, Imperial College London

INTRODUCTION

The Standard Model is the theory of how particles behave and interact. It has produced amazingly accurate predictions but there are still some pieces missing, such as particle candidates for dark matter and gravity[1]. This has sparked a search for New Physics. One promising such search focuses on B-meson decay[3].

$$B \rightarrow K^{*} l^{+} l^{-}$$

This decay is heavily suppressed in the Standard Model and it is therefore sensitive to New Physics[2].

To compare different New Physics models, we use Wilson Coefficients, the most interesting of which are $\mathcal{C}_9$ and $\mathcal{C}_{10}$. These seem to deviate from the standard model. Currently the Wilson coefficients are found through a complicated fit.

Inspired by recent successes[4] using machine learning in high energy physics, we want to improve this fitting method using machine learning. Increased precision in Wilson coefficients could lead to a breakthrough in the search for New Physics.

METHOD

DATA

- The data used are the squared invariant mass, q^2, of the decay along with the angular variables θ_k, θ_l, and ϕ.
- These are generated with Monte-Carlo Simulations.
- For our Neural Network method, the data is then converted into a histogram with 10 bins in each of the four variables.
- Data is split into a training set and a evaluation set to verify performance.

NEURAL NETWORK (NN)

A Neural Network is a directed graph with a series of node layers. Each layer connects to the next through weighted edges which are tuned during training. Each input node is a bin from the histogram and the outputs are the $\mathcal{C}_9$ and $\mathcal{C}_{10}$.

Pros:
- Multiple Outputs
- Can approximate any function
- Performs Regression

Cons:
- Requires additional processing
- Loses information during binning

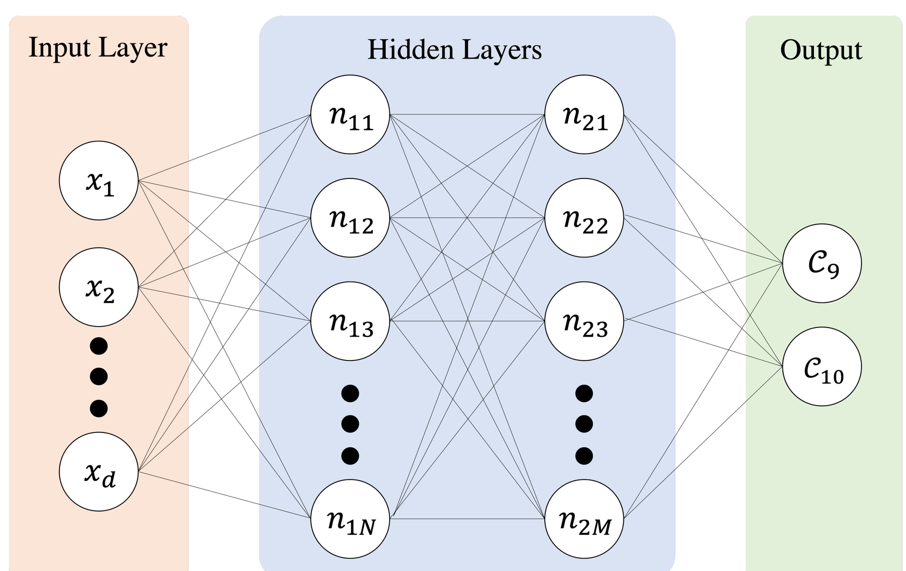
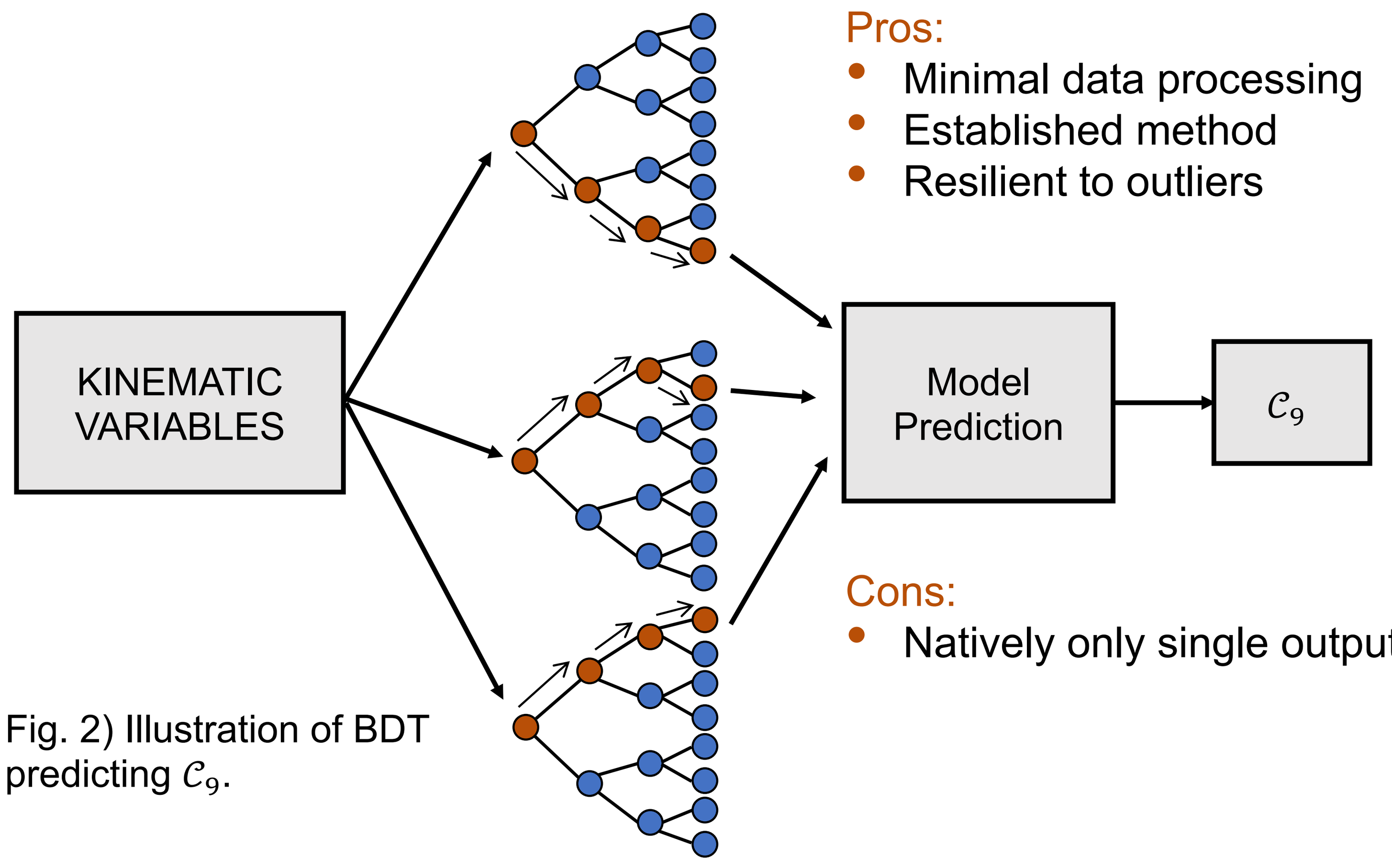


Fig. 1) Representation of our neural network.

BOOSTED DECISION TREES (BDT)

A Boosted Decision Tree Classifier is a collection of shallow decision trees. They are combined to achieve better performance than a single decision tree.



Pros:
- Minimal data processing
- Established method
- Resilient to outliers

Cons:
- Natively only single output

Fig. 2) Illustration of BDT predicting $\mathcal{C}_9$.

The BDT computes the log-probability that an event was generated from a specific $\mathcal{C}_9$. Summing multiple events from a dataset gives the log-likelihood that the dataset was generated from each $\mathcal{C}_9$.

RESULTS

NEURAL NETWORK (NN)

- In Fig. 3 it is shown how predictions are most precise for $\mathcal{C}_9^{NP} < 0$ and $\mathcal{C}_{10}^{NP} > 0$.
- Produces one sigma contours of similar size to current methods, fig. 4.

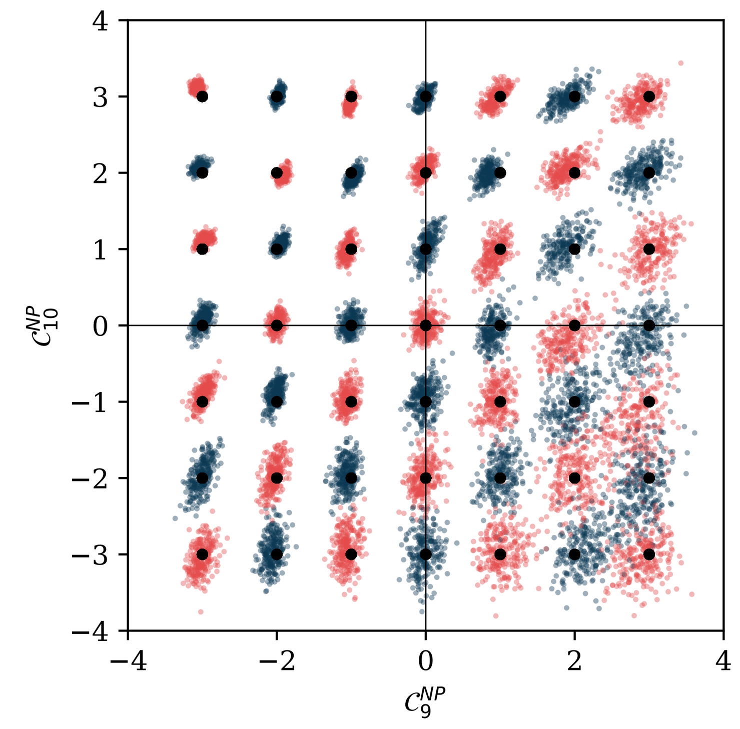


Fig. 3) Grid with true values (black) with predictions (red and blue).

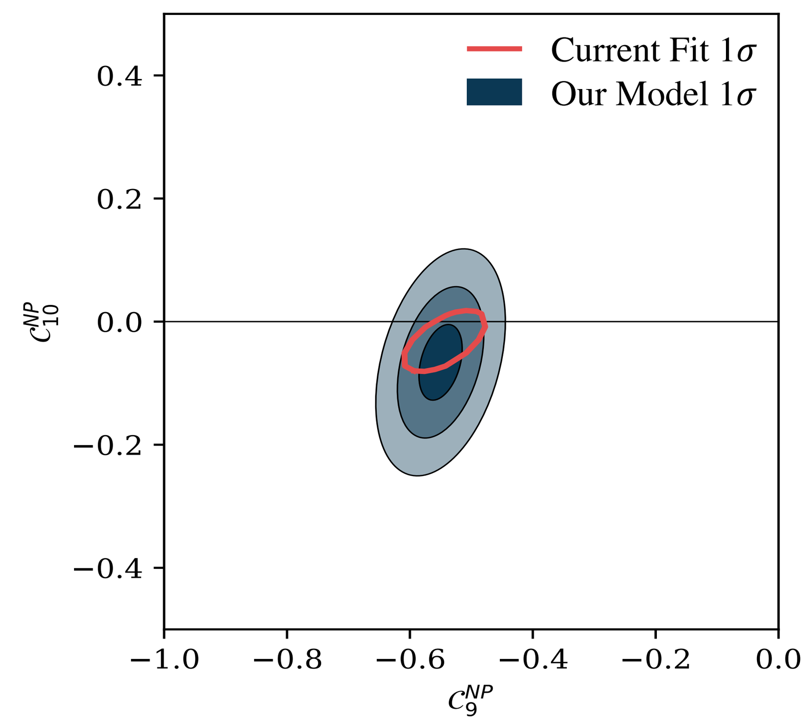


Fig. 4) One to three sigma contours of predictions and current best fit.

BOOSTED DECISION TREES

- Can distinguish data generated with New Physics from data generated with Standard Model physics within $\Delta\mathcal{C}_9^{NP} = 0.5$, see Fig. 5.
- Can correctly identify most likely $\mathcal{C}_9^{NP}$ to precision of 0.5 over whole range -3 to 3, see Fig. 6.

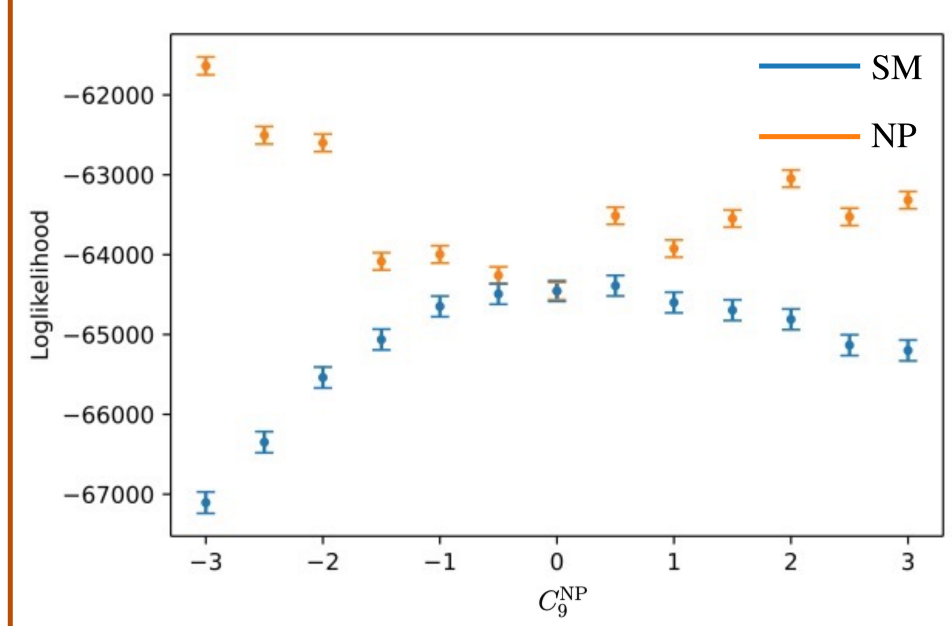


Fig. 5) Log-likelihood prediction of a dataset being generated from New Physics.

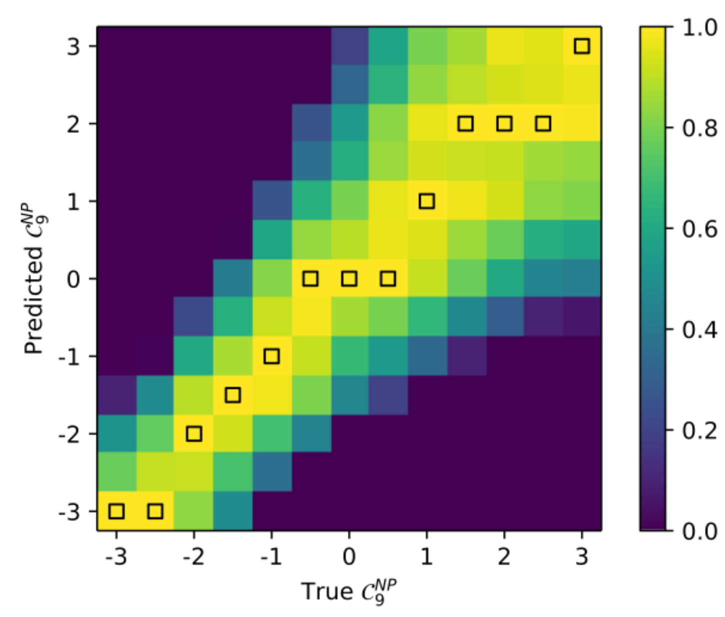


Fig. 6) Log-likelihood for a set of $\mathcal{C}_9^{NP}$ spaced 0.5 apart. Black square indicates the most likely coefficient.

Conclusion

- To aid the search for New Physics, we have developed two machine learning methods for determining the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$.
- Using a BDT, a dataset can be correctly determined to originate from New Physics or the Standard Model. The correct value for $\mathcal{C}_9$ can also be determined in most cases.
- Using a NN, both $\mathcal{C}_9$ and $\mathcal{C}_{10}$ are determined with similar confidence to current methods.
- Both models are more accurate in $\mathcal{C}_9$ than $\mathcal{C}_{10}$.
- Both models become more accurate for smaller $\mathcal{C}_9$ and larger $\mathcal{C}_{10}$

References

[1] L. collaboration, R. Aaij, C. A. Beteta, T. Ackernley, B. Adeva, M. Adinolfi, and H. Afsharnia, “Test of lepton universality in beauty-quark decays,”
[2] M. Alguer´o, B. Capdevila, S. Descotes-Genon, J. Matias, and M. Novoa-Brunet, “$b \to sl^+l^-$ global fits after moriond 2021 results,”
[3] U. Egede, T. Hurth, J. Matias, M. Ramon, and W. Reece, “New physics reach of the decay mode $\overline{B_d} \to K^{+-*}l^+l^-$,” vol. 2010.
[4] J. Collins, K. Howe, and B. Nachman, “Anomaly detection for resonant new physics with machine learning,” vol. 121, no. 24, p. 241803.
[5] C. F. M. D. S. Anders Andreassen, Ilya Feige, “Junipr: a framework for unsupervised machine learning in particle physics,”

jakub.pazio18@imperial.ac.uk, oskar.hogberg18@imperial.ac.uk
Physics Department, Imperial College London