

Kraków, 28 grudnia 2025

dr hab. Marcin Wolter

Instytut Fizyki Jądrowej
im. Henryka Niewodniczańskiego
Polskiej Akademii Nauk
ul. Radzikowskiego 152
31-342 Kraków

Recenzja rozprawy doktorskiej:

mgr Sabin Hashmi, „*Real-Time Machine Learning-Enhanced Reconstruction of Long-Lived Tracks in LHCb High-Level Trigger and Calibration Analysis of the UT Detector*”

Promotor: dr hab. inż. Tomasz Szumlak

Recenzja została wykonana na zlecenie Rady Dyscypliny Nauki fizyczne Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie. Podstawa prawna art. 187 Ustawy z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce” (z późn. zm.)

Przedstawiona mi do oceny rozprawa doktorska pana mgr. Sabina Hashmi skupia się na kompleksowym badaniu zastosowania różnorodnych technik uczenia maszynowego do poprawy rekonstrukcji śladów cząstek długożyciowych w eksperymencie LHCb w CERN. Praca obejmuje okres badań prowadzonych w czasie modernizacji detektora przed Run 3. Rozprawa została sporządzona w języku angielskim.

Całość rozprawy jest logicznie podzielona na trzy części:

Część I (Rozdziały 1-2) wprowadza w kontekst fizyczny eksperymentu LHCb, architekturę detektora oraz infrastrukturę obliczeniową. Szczególną uwagę poświęcono detektorom pozwalającym na rekonstrukcję torów cząstek (VELO - Vertex Locator, UT - Upstream Tracker, SciFi - Scintillating Fibre Tracker) oraz systemowi wyzwalania wysokiego poziomu (HLT). Autor prezentuje również szczegółowy opis zastosowanych algorytmów rekonstrukcji torów cząstek, w tym kluczowych dla jego pracy algorytmów Hybrid Seeding i Long-lived Tracking.

Część II (Rozdziały 3-6) stanowi główną część rozprawy, prezentując strategię dwustopniowej selekcji śladów cząstek:

- Etap pierwszy koncentruje się na poprawie selekcji śladów SciFi używanych jako początkowe elementy (tzw. „seeds”) do rekonstrukcji śladów downstream.
- Etap drugi dotyczy właściwej selekcji śladów downstream i eliminacji śladów kombinatorycznych (tzw. „ghost tracks”)

Część III (Rozdziały 7-8) obejmuje kalibrację detektora UT, oraz zawiera podsumowanie wyników.

Do głównych osiągnięć Autora można zaliczyć opracowanie i zintegrowanie dwóch modeli uczenia maszynowego opartych na algorytmie Catboost:

- **Klasyfikator śladów SciFi:** Osiąga redukcję ghost tracks o 87% (implementacja C++) do 96% (implementacja ONNX) w porównaniu z algorytmem bazowym, przy minimalnej stracie efektywności rekonstrukcji.
- **Klasyfikator śladów downstream:** Redukuje ghost tracks o 50-60% przy zachowaniu wysokiej wydajności rekonstrukcji prawdziwych śladów.

Autor porównał modele liniowe (regresja logistyczna) z bardziej zaawansowanymi metodami nieliniowymi, demonstrując przewagę tych ostatnich. Wybór modelu Catboost był uzasadniony przez Autora wyższą wydajnością (AUC ~85% dla downstream, ~97% dla SciFi) oraz możliwością łatwej integracji z frameworkiem Gaudi poprzez eksport do C++ i formatu ONNX.

Istotnym wkładem pana mgr. Sabina Hashmi w prace eksperymentu LHCb jest integracja opracowanych algorytmów z systemem wyzwalań (HLT trigger) eksperymentu. Autor zaimplementował dwie metody włączenia opracowanych algorytmów:

- Bezpośrednią implementację w C++ (niższy narzut obliczeniowy)
- Implementację przez pakiet ONNX Runtime (większa elastyczność)

Szczególnie wartościowa jest analiza wpływu modeli na kluczowe wskaźniki wydajności: efektywność rekonstrukcji, ghost ratio, oraz rozkłady zmiennych kinematycznych. Wyniki pokazują, że nowe modele nie wprowadzają zniekształceń w rozkładach fizycznych (co potwierdzają analizy w frameworku DaVinci dla cząstek K^0_L).

Ostatnia część pracy obejmuje rozdział poświęcony kalibracji detektora UT i prezentuje dwa komplementarne podejścia:

- Wykorzystanie dywergencji Kullbacka-Leiblera (KL) do porównywania rozkładów wartości piedestału między kalibracjami.
- Model prognostyczny oparty na sieciach LSTM do przewidywania przyszłych wartości piedestału.

Choć ta część jest skromniejsza objętościowo od poprzednich, niemniej stanowi solidną podstawę do przyszłych prac nad autonomicznymi systemami kalibracji.

Praca pana mgr. Sabina Hashmi prezentuje pełny cykl rozwoju modeli ML – od wyboru danych treningowych, przez porównanie algorytmów, optymalizację hyperparametrów, aż po wdrożenie w środowisku produkcyjnym. Na podkreślenie zasługuje troska Autora o interpretowalność wyników poprzez konsekwentne stosowanie analizy SHAP. Na podkreślenie zasługuje użyteczność metod rozwijanych przez Autora, a następnie zaimplementowanych przez niego w oprogramowaniu eksperymentu. Zaimplementowane algorytmy znacząco zwiększają wydajność selekcji torów cząstek w eksperymencie LHCb.

Rozprawa wnosi istotny wkład do zastosowań uczenia maszynowego w fizyce wysokich energii, demonstrując możliwość integracji zaawansowanych modeli ML z systemem wyzwalania. Opracowana metodologia dwustopniowej selekcji może być rozszerzona na inne typy śladów w LHCb oraz potencjalnie zaadaptowana w innych eksperymentach. Praca ma bezpośrednie znaczenie praktyczne dla operacji Run 3 eksperymentu LHCb i stanowi solidną podstawę dla dalszych ulepszeń i przyszłych modernizacji (Upgrade II).

Rozprawa doktorska pana mgr. Sabina Hashmi przedstawia kompleksowe badania dotyczące zastosowania metod uczenia maszynowego w rekonstrukcji śladów cząstek długożyciowych w eksperymencie LHCb oraz analizy kalibracji detektora UT (Upstream Tracker). Praca jest dobrze ustrukturyzowana, co świadczy o przemyślanej organizacji materiału. Rozprawa wykazuje głęboką wiedzę Autora w zakresie zarówno fizyki cząstek, jak i rekonstrukcji torów cząstek, ich selekcji oraz zastosowania metod uczenia maszynowego.

Doktorant prezentuje zrozumienie zarówno zagadnień z dziedziny fizyki wysokich energii oraz łączy je z zaawansowanymi metodami obliczeniowymi. Oceniając pracę Doktoranta chciałbym podkreślić wyraźnie widoczną umiejętność samodzielnego prowadzenia badań naukowych, o czym świadczą zarówno systematyczne podejście metodologiczne do opracowania dwustopniowej selekcji śladów, ocena i porównanie różnych modeli uczenia maszynowego (logistyczna regresja, drzewa decyzyjne, gradient boosting, Catboost) oraz implementacja opracowanych metod do środowiska Gaudi (w dwóch wariantach C++ i ONNX Runtime). Należy też zwrócić uwagę na przeprowadzoną przez Doktoranta szczegółową analizę wydajności z użyciem różnych metryk (efficiency, ghost ratio, ROC curves, SHAP analysis).

Pomimo wysokiej jakości merytorycznej, ma do rozprawy szereg uwag. Pozwolę sobie wymienić część z nich:

- W pierwszym rzędzie muszę zwrócić uwagę na brak referencji lub ewentualnie krótkiego wyjaśnienia pojęć, które ułatwiłyby lekturę czytelnikowi. W sekcji 1.1 "Supervised Machine Learning" przydałyby się referencje do poszczególnych metod. Poza tym mnie, jako czytelnikowi pracy, nasuwa się myśl, że może nie jest konieczne opisywanie metod, które nie zostały użyte w pracy (np. metoda Naive Bayes, redukcja wymiarowości za pomocą Principal Component Analysis czy enkoderów). Podobnie, czy metoda Reinforcement Learning było użyte w tej analizie?
- Praca, do której Autor odwołuje się trzy razy opisując metody uczenia maszynowego ([35] Machine Learning for Columnar High Energy Physics Analysis Elliott

Kauffman, Alexander Held, and Oksana Shadura) jest krótkim referatem na CHEP23. Wydaje mi się, że lepiej nadawałby się do tej roli jakiś powszechnie uznany podręcznik.

- Metoda SHapley Additive exPlanations (SHAP) jest używana przez Doktoranta bardzo często. Jednakże jej omówienie zawiera tylko pięć linii tekstu i odnośnik do pracy [47]. Moim zdaniem praca wiele by zyskała na bardziej szczegółowym wytłumaczeniu metody SHAP, znanej między innymi z teorii gier.
- Opisane pakiety oprogramowania powinny mieć podane odnośniki bibliograficzne, np. nie mają ich pakiety Optuna czy DaVinci.
- Rysunki 22 i 41 zawierają szereg zmiennych przedstawionych jako "boxplot". Rozkłady przedstawione na rysunku są jednowymiarowe, czy rysunki nie byłyby bardziej czytelne, gdyby przedstawić je w formie histogramów?
- W rysunki 25 i 44 przedstawiają korelacje między zmiennymi. Widać na nim silnie skorelowane pary zmiennych. Czy nie można by usunąć jednej z tak skorelowanych zmiennych?
- Równania 18 i 19 są identyczne, zamiast je powtarzać można w tekście odwołać się do pierwszego równania.
- W rozdziale 4 sekcja 6 "Model Comparison" nie jest zaznaczone, jakiego pakietu z zaimplementowanymi metodami uczenia maszynowego używał Autor. Czy był to pakiet Catboost czy też np. Scikit-learn dla innych metod? Referencja "ProkhorenkovaCatboost2017" powinna znaleźć się w spisie literatury, nie bezpośrednio w tekście pracy.
- Z tabeli 6 nie wynika, że modele oparte na metodzie "boosting" dają najlepsze wyniki, MLP daje podobne rezultaty. Również Catboost nie jest lepszy od innych metod typu boosting lub MLP.
- Tabele 7 i 13 zawierają listę parametrów metody Catboost wybranej do optymalizacji za pomocą pakietu Optuna. Które z tych hyperparametrów były faktycznie optymalizowane?
-
- W tabelach 5 i 11 prawdopodobieństwa (p) dla "True Track" i "Ghost Track" nie sumują się do jeden. Czy wynika to z tego, że mamy jeszcze jakąś inną klasę torów?
- Rysunek 50 a: dlaczego na rysunku został użyty tak duży zakres SHAP value, rysunek robi się nieczytelny.
- Rysunek 65: dopasowanie krzywej Gaussa do przedstawionego rozkładu nie uwzględnia stałego tła, odnoszę wrażenie, że lepsze wyniki dałoby dopasowanie funkcji Gauss+constants.
- Brakuje wyjaśnienia dla "Kullback-Leibler divergence" ani odnośnika do literatury (str. 113). Dla LSTM (str. 115) też brakuje odnośnika do literatury.
- Dlaczego klasyfikator śladów SciFi zaimplementowany w C++ osiąga redukcję „ghost tracks” o 87% a w ONNX aż o 96%?

Po lekturze rozprawy doktorskiej nasuwa mi się kilka dodatkowych pytań do Autora:

- O ile model LSTM daje lepsze przewidywania niż np. średnia krocząca z przeszłych kalibracji? Ile zyskujemy stosując model LSTM? Średnia krocząca czy jakaś prosta wielkość mogłaby posłużyć jako benchmark.
- Czy rozważano zastosowanie innych architektur sieci neuronowych (np. Graph Neural Networks) do rekonstrukcji śladów?
- Dlaczego zdecydowano się na Catboost zamiast np. XGBoost, który jest szerokostosowany w fizyce cząstek? Czy łatwiej go było zaimplementować we frameworku eksperymentu LHCb?
- Rysunki 28 i 36 przedstawiają znaczenie zmiennych w modelu Logit i Catboost. Widzimy zdecydowane różnice pomiędzy modelami, np. w Catboost X jest bardzo znaczącą zmienną, w modelu liniowym już nie. Czy jest znane jakieś wytłumaczenie tego faktu?
- Jakie są dalsze plany rozwoju przedstawionych algorytmów?

Biorąc pod uwagę całokształt pracy i jej znaczący wkład w analizę eksperymentu LHCb oraz w rozwój i zastosowanie metod uczenia maszynowego w rekonstrukcji torów cząstek w eksperymentach fizyki wysokich energii oceniam pracę mgr. Sabina Hashmi zdecydowanie pozytywnie.

Podsumowując stwierdzam, że recenzowana rozprawa spełnia wymagania ustawy o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki określonych w art. 187 ustawy z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce” (z późn. zm.) i wnioskuję o jej dopuszczenie do dalszych etapów postępowania o nadanie stopnia doktora w dziedzinie nauk ścisłych i przyrodniczych w dyscyplinie nauki fizyczne.



dr hab. Marcin Wolter