

10 grudnia 2022

dr hab. Marcin Wolter

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

**Ocena rozprawy doktorskiej mgr. inż. Macieja Witolda Majewskiego „Application of machine learning methods for the analysis of the calibration of the LHCb VELO detector, studies of irradiated silicon pixel sensors and reconstruction of the neutrino interaction in LARTPC detector”**

Przedstawioną mi do recenzji pracę przeczytałem z dużym zainteresowaniem. Dysertacja doktorska mgr. inż. Macieja Witolda Majewskiego została przygotowana pod kierunkiem promotora prof. dr. hab. inż. Tomasza Szumlaka oraz promotora pomocniczego dr. inż. Bartłomieja Rachwała. Praca ta stanowi znaczący wkład w program badawczy eksperymentu LHCb, w który zaangażowane są krakowskie grupy z Akademii Górniczo-Hutniczej oraz Instytutu Fizyki Jądrowej PAN.

Praca skupia się na zagadnieniach zastosowania metod sztucznej inteligencji do kalibracji detektorów, analizy ich pracy oraz rekonstrukcji oddziaływań cząstek. Jej wiodącym tematem jest użycie narzędzi uczenia maszynowego i sztucznej inteligencji do optymalizacji pomiarów. Doktorant zaproponował i wdrożył szereg nowatorskich metod kalibracji detektorów oraz rekonstrukcji parametrów mierzonych cząstek.

Rozprawa składa się z ośmiu rozdziałów oraz wstępu i wniosków. Rozdział pierwszy stanowi krótki, kilkustronicowy wstęp teoretyczny do fizyki wysokich energii ze szczególnym uwzględnieniem łamania symetrii CP oraz uniwersalności leptonowej - konsekwencji Modelu Standardowego i głównych tematów badawczych eksperymentu LHCb w CERNie. Jeśli łamanie uniwersalności byłoby odkryte przez ten eksperyment, byłoby to równoznaczne odkryciu Nowej Fizyki. W następnym rozdziale opisany jest akcelerator LHC oraz eksperyment LHCb, a szczególny nacisk jest położony na detektor Velo (ang. *Vertex Locator*) służący do precyzyjnego pomiaru wierzchołka oddziaływania. W swojej pracy badawczej Autor skupia się na pomiarach z użyciem tego detektora oraz jego kalibracji z użyciem nowatorskich metod, stąd dokładny opis tego elementu aparatury badawczej.

Rozdział trzeci stanowi ogólne wprowadzenie do metod uczenia maszynowego, ze szczególnym uwzględnieniem metod użytych przez Autora w jego pracy. Dotyczy to w szczególności redukcji wymiarów, rekurencyjnych sieci neuronowych służących do analizy ciągów czasowych oraz metod uczenia wzmocnionego (ang. *Reinforcement learning*). Osobny podrozdział poświęcony jest zagadnieniu klasteryzacji.

Następne rozdziały rozprawy doktorskiej opisują algorytmy i metody, w których skonstruowaniu i zastosowaniu Autor miał znaczący udział. Dotyczą one w pierwszym rzędzie detektora wierzchołka Velo, przy czym rozdział czwarty opisuje prace obejmujące bardzo szeroko rozumianą kalibrację pierwszej wersji detektora, tzw. *Strip Velo*, a rozdział piąty jego następcy, pikselowego detektora *Velo Pix*. Oprócz wyznaczania piedestałów, poziomów wyzwalań, maskowania klastrów pikseli zastosowano techniki redukcji wymiarowości (PCA, autoencoder) w celu lepszego zobrazowania na dwuwymiarowej mapie wielokanałowego detektora. Bardzo ciekawy jest też model, oparty na rekurencyjnych sieciach neuronowych, służący do prognozowania czasu następnej kalibracji, a także propozycja pomiaru naświetlenia detektora strumieniem cząstek na podstawie zmiany zastępczej funkcji aktywacji (ang. *surrogate activation function*) w detektorze *Velo Pix*.

W rozdziale szóstym Autor omawia *Titanię*, *framework* służący do tworzenia narzędzi monitorujących detektor oraz webowy serwis obsługujący bazę danych, nazwany *Storck*, których rozwojem Doktorant kierował. Są to ogólnodostępne serwisy „open source”, które mogą zostać zastosowane w dowolnym eksperymencie.

Krótki, pięciostronicowy rozdział 5 stanowi wprowadzenie do eksperymentów neutrinowych oraz detektora LARTPC (*Liquid Argon Time Projection Chamber*). Sam Autor zaznaczył, że nie przeprowadził analizy fizycznej danych z eksperymentów neutrinowych, ale za to przygotował ciekawy algorytm rekonstrukcji śladów cząstek w detektorze LARTPC. Metoda oparta jest na algorytmach wzmocnionego uczenia (*reinforcement learning*), a testowana na publicznie dostępnych danych.

Bibliografię stanowią 92 pozycje, z których część stanowią strony internetowe zawierające opisy pakietów oprogramowania lub repozytoria kodu. Jest to w pełni zrozumiałe w pracy o charakterze komputerowym.

Uważam, że lektura rozprawy doktorskiej byłaby łatwiejsza, gdyby w rozdziale szóstym opisującym programy *Storck* i *Titania* zreferowane były zasady działania obu serwisów, zamieszczone schematy blokowe, ale szczegóły zostały przeniesione do dodatku na końcu rozprawy. W szczególności dotyczy to fragmentów kodu w języku python umieszczonych w tekście.

Bardzo ciekawa jest praca niezwiązana z eksperymentem LHCb, polegająca na oznaczaniu pochodzenia śladów w detektorze LARTPC. Zastosowana metoda „reinforcement learning” pozwala na implementację „agenta” poruszającego się po obszarze detektora i klasyfikującego ślady. Jedyne niedosyt, jaki mi pozostawiła lektura tego rozdziału stanowi brak porównania z innymi metodami rekonstrukcji, jak choćby



z użyciem sieci GNN (Graph Neural Network) czy też innych, klasycznych metod. Takie, choćby krótkie porównanie, przedstawiłoby motywację stojącą za użyciem zastosowanych algorytmów.

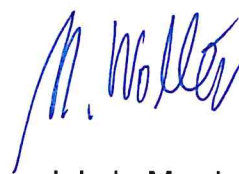
Poza sugestiami przedstawionymi powyżej nie mam innych znaczących uwag krytycznych do przedstawionej mi do recenzji pracy. Rozprawa potwierdza, że Autor zdobył umiejętności pozwalające na stosowanie zaawansowanych metody analizy danych z użyciem metod uczenia maszynowego, znajomość zarówno detektorów używanych w fizyce wysokich energii jak i metod tworzenia narzędzi programistycznych umożliwiających optymalne ich wykorzystanie. Jestem przekonany, że praca włożona przez pana mgr. inż. Macieja Witolda Majewskiego w przygotowanie od strony programistycznej eksperymentu LHCb zaowocuje w przyszłości szeregiem ciekawych analiz fizycznych przeprowadzonych na zebranych danych.

Redagując swoją dysertację pan mgr inż. Maciej Majewski nie ustrzegł się drobnych potknięć edytorskich. Lekturę dysertacji utrudnia fakt, że nie wszystkie ilustracje są zamieszczone zgodnie z kolejnością numeracji (np. Fig. 4.3.5 występuje przed Fig. 4.3.1), choć można to tłumaczyć chęcią wykorzystania pustego miejsca przed rysunkami zajmującymi całą stronę. Także strona 19 pozostawiona została w większości pusta. W pracy występuje także szereg drobnych błędów literowych.

Niezależnie od przedstawionych powyżej uwag uważam, że przedstawiony mi do recenzji elaborat stanowi bardzo wartościowe i oryginalne opracowanie, które wzbogaca wiedzę dotyczącą stosowania zaawansowanych metod statystycznych i uczenia maszynowego w optymalizacji danych zbieranych z detektorów oraz pracy samych detektorów. Po zakresie prac przedstawionych w dysertacji widać, że Autor miał bardzo znaczący wkład w przygotowanie i operacje detektora LHCb tworząc liczne narzędzia softwarowe. Na podkreślenie zasługuje, że znalazł także czas na ćwiczenie w zastosowaniu uczenia maszynowego w detektorze LARTPC, wykraczające poza zakres prac wykonywanych dla eksperymentu LHCb.

Rozprawę doktorską mgr inż. Macieja Witolda Majewskiego uważam za bardzo dobrą. Autor czuje, jakiemu celowi będzie mają służyć opracowywane i wdrożone systemy wykorzystujące uczenie maszynowe oraz dobrze rozumie ich znaczenie dla eksperymentu LHCb.

Uważam, że rozprawa doktorska mgr. inż. Macieja Witolda Majewskiego spełnia wymagania stawiane rozprawom doktorskim i wnoszę o dopuszczenie Autora do dalszych etapów przewodu doktorskiego.



dr hab. Marcin Wolter