

CERN Summer Student Programme 2025 参加報告

京都大学 理学研究科

田中 智也

tanaka.tomoya.84n@st.kyoto-u.ac.jp

2025 年 (令和 7 年) 8 月 22 日

1 はじめに

私は 6 月 30 日から 8 月 22 日までの 8 週間、CERN Summer Student Programme 2025 に参加した。本プログラムは世界中の学生が CERN に 8~13 週間滞在し、国際交流を行いながら、講義の聴講や workshop の参加を行いつつ、supervisor のもとで研究を行うというものである。私は supervisor の Davide Di Croce 氏と Johannes Junggeburth 氏の指導のもと、New Generation Trigger (NGT) チームの一員として ATLAS 実験に関するプロジェクトに取り組んだ。

2 活動内容

2.1 講義

プログラム期間前半の 5 週間は、午前中に 3 コマ 3 時間の講義が行われた。講義の内容は標準模型や天文学などの物理に関するものから、超伝導磁石といった加速器工学、さらには高エネルギー物理を用いた放射線治療についてなど、非常に多岐にわたる内容を学ぶことができた。受講生の熱量も高く、講義後は質問が相次ぐ光景が何度も見られ、国際的な環境での学びの違いを感じることができた。

2.2 Workshop

プログラム中には、ROOT の使い方や霧箱、シリコンセンサーといった様々な種類の workshop が開かれた。私は MadGraph の workshop に参加し、ファインマンダイアグラムの自動描画や、断面積の計算などを学ぶことができた。

2.3 研究内容

2.3.1 概要

ATLAS 実験は、大型ハドロン衝突型加速器 (LHC) に設置された大型汎用検出器 (ATLAS 検出器) を用いた実験である。LHC で起こる陽子衝突によって生じる粒子を ATLAS 検出器で捉えることで、標準模型の検証や新物理の探索を行っている。LHC は 2030 年頃に High Luminosity LHC (HL-LHC) にアップグレードされ、1 bunch あたりの衝突数である pile-up が 60 から 200 に増加する予定である。この高いイベントレートに対応するべく、検出器側のアップグレードも併せて行われている。

アップグレード後の Phase-II におけるソフトウェアトリガーである Event Filter (EF) では、オフラインに近い方法でトラックの構成が行われる予定である。本研究は、EF でのミューオントラックのステップの一つである、ミューオンのセグメント探索に関するものである。

ミューオンのセグメントとは直線に近似されたミューオンの軌道のことであり、セグメントを組み合わせることで最終的なトラックが決定される (図 1 参照)。新しい手法におけるセグメントの再構成はいくつかのステップに分かれており、第一段階として bucket と呼ばれるミューオンのヒットの集合が作成される。ここでのヒットは、Monitored Drift Chamber (MDT) などの精密測定をする検出器から、Thin Gap Chamber (TGC) などのトリガー発行に用いられる検出器まで、すべてのミューオンスペクトロメータでのヒットを含む。Bucket は、bucket filter という機械学習を用いたモデルによって、中にあるセグメントの数に応じて分類される。その後、segment regressor によって bucket 内部のセグメントのパラメータ (位置および方向) が推定される。Segment regressor には現在 one-segment regressor と two-segment regressor があり、それぞれ bucket にあるセグメントの本数に応じて使い分けられる予定である。

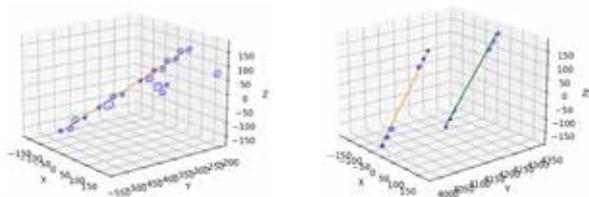


図 1: ミューオンセグメントの例。オレンジの線がセグメントであり、青い点はヒット、リングはMDTのドリフト半径を示している。左はセグメントが1本、右は2本の場合。

セグメントの特定には、バックグラウンドの存在や条件を満たすセグメントが複数再構成されることによる曖昧さなどの困難がある。これらの問題の解決のため、segment regressor は Graph Neural Network (GNN) を用いて開発されている。GNN を用いることで、より高い精度でセグメントを再構成し、かつセグメントの候補数を従来よりも減らすことを目標としている。プログラム期間中、私は one-segment regressor ・ two-segment regressor の開発と性能評価、およびモデルの最適化に取り組んだ。

2.3.2 One-segment regressor

まず私が行ったことは one-segment regressor の性能評価である。本モデルは Edge Convolutional モデルをもとに作られた GNN であり、1 つの bucket から 1 つのセグメントのパラメータ (位置: x, y , 方向: θ, ϕ) を予測する。入力データについて、bucket 内の各ヒットについて位置 (3 変数)、ステーションのインデックス、MDT のドリフト半径、レイヤー番号の 6 変数が特徴量となり、bucket はヒットをノードとする全結合グラフに変換されて渡される。また、入力データ、ラベルともに規格化されている。モデルに関しては、PyTorch から機械学習の汎用フレームワークである ONNX へと変換されている。ONNX の runtime が ATLAS のソフトウェアである ATHENA に組み込まれていることから、ONNX 変換後の性能が重要となる。

結果は図 2 のグラフに示した。座標は bucket に固有なローカル座標であり、 x 方向を検出器の感度が悪い方向 (MDT の場合であればワイヤ方向) に取っているため x 方向および ϕ のエラーが大きくなっている。なお、bucket 内部のヒットが x の情報を持たない場合はセグメントの x 座標を 0 としているため、 $x = 0$ に多くのイベントが存在する。一方で、 y と θ に関しては強い線形関係が見られ、regressor が精度良く予測できていることが示された。エラー率を Mean Squared Error (MSE) を予測値の範囲で割った値として定義すると、 y につい

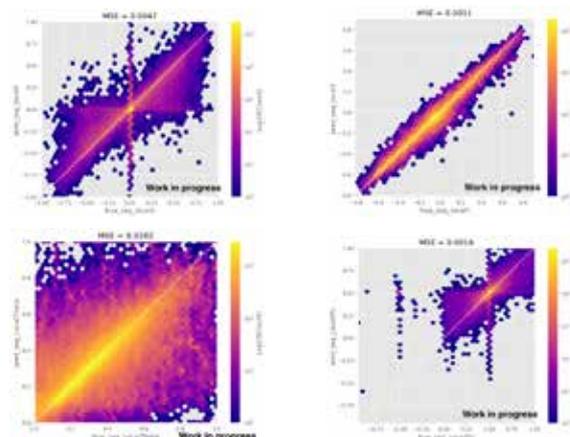


図 2: One-segment regressor の truth segment (横軸) と予測された segment (縦軸) のパラメータのエラー。左上から x, y, θ, ϕ 。

て 0.06%, θ についておよそ 2% であった。

また、速度に関しては、ONNX に変換したモデルのスタンドアロンでの推論速度を NVIDIA H100 NVL を用いて測定した。Batch サイズは 128 であり、推論は並列に行われた。結果は 319 ms/batch, 2.5 ms/bucket であった。なお、速度の評価は bucket の生成からトラックの特定までの一貫したプロセスにかかる時間で行われるものであるため、この値は参考値である。

2.3.3 Two-segment regressor

次に、two-segment regressor の開発を行った。このモデルは 2 本のセグメントを 1 つの bucket 内部から特定し、パラメータを推定する。One-segment regressor と同様に Edge Convolutional モデルを用い、出力する変数の数はセグメントが 2 本になるため 8 つに変更されている。また、2 本のセグメントは y 座標の絶対値が小さいほうが 1 本目になるよう順序付けられている。

結果は図 3 の通りである。なお、この結果は ONNX に変換せずに得られたものである。概ね one-segment regressor と同様の結果であり、 y, θ のグラフは強い線形相関を、 x, ϕ のグラフはやや広い分布を示した。また、1 本目と 2 本目のセグメントで結果に大きな差は見られなかった。エラー率に関しては、いずれの変数においても 2% 以下であった。

2.3.4 モデルの最適化

また、モデルのアーキテクチャを変更することで、さらなる性能の改善にも取り組んだ。これまでのモデルで使用されていた Edge Convolutional モデルに加えて、GAT (Graph Attention Network) や SAGE (Sample and Aggregate) といった他の GNN モデルのブロックを組み合

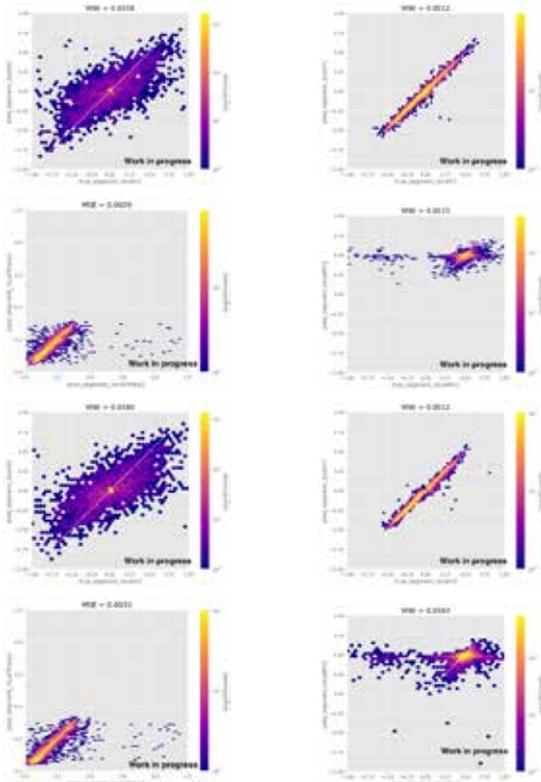


図 3: Two-segment regressor の truth segment (横軸) と予測された segment (縦軸) のパラメータのエラー。上 2 行は 1 本目, 下 2 行は 2 本目のセグメントのグラフ。

合わせたカスタムモデルや, Transformer モデルを作成し, テストを行った。これらのアーキテクチャの変更に加えて, ハイパーパラメータの最適化も試みた。最適化ツールとして, Katib experiment を用いた。Katib experiment は最適なハイパーパラメータを Bayesian や Tree of Parzen Estimators (TPE) といったアルゴリズムを用いて探索する。今回は Multivariate-TPE を用いた最適化を試みた。私がプログラム期間中所属していた NGT チームにおいては, 本ツールを使うのは私が初めてであったということもあり, バグや細かな仕様に阻まれ実行には多くの時間を費やした。また, 最適化には実行を開始してから 2 週間と長時間かかることもあって, 結果はまだ得られておらず, 本記事の執筆現在も取り組んでいる最中である。

3 CERN での生活

私は期間中 CERN 内のホステルに滞在していたため, 研究を行っていた Building 40 やレストランに非常に近く, 快適な生活を送ることができた。休日はほかの summer student と毎週のようにスイスやフランスの観光地へと足を伸ばし, 充実した日々を過ごした。

4 今後の抱負

本プログラムでは, 今まで触れてこなかった機械学習に関して, 実装レベルで深く学ぶことができ, また高エネルギー物理学への応用を知ることができた。機械学習は解析等でも必須の技術といえるため, 様々なモデルに触れながら学べたことは非常に良い経験であった。今後の研究にも活かしていきたい。

また, 日々の生活においては, 英語でのコミュニケーションの難しさに幾度となく直面した。様々な国籍の人たちと会話や議論を行うためには, 英語でスムーズに話せるようになる必要があると痛感した。スピーキングは慣れや練習回数に大きく依存する能力であると思うので, 今後は自分から会話を行うことを心がけていきたい。

5 今後このプログラムへ望むこと

やや難しい要求かもしれないが, 現在 8 週間であるプログラム期間について, より長い期間を選べるようにすることを希望したい。CERN の Member State から来ている学生と同様に 3ヶ月程度滞在できると, 研究をさらに深めることができ, 国際交流を行える期間が増えることから, より有意義なプログラムになると考える。

6 謝辞

今回このプログラムに参加するにあたって, 多くの方々にお世話になりました。まず初めに, このプログラムを支援している KEK の加速器科学国際育成事業 (IINAS-NX) に感謝申し上げます。応募の際には, 京都大学の中家剛教授に推薦状を書いていただき, 木河達也助教には応募書類の添削をしていただくなど, 大変お世話になりました。また, KEK 国際企画課の職員の皆様に渡航の準備を支援していただき, 現地においては小作様に現地での生活に関して色々教えていただきました。現地での supervisor であった Davide Di Croce 氏と Johannes Junggeburth 氏には, ATLAS 実験全般のことから機械学習の細かな仕様に関することまで, 多くのことを教えていただきました。同僚であった Bård Raddum Wikmark さんや, 日本人としてともにプログラムに参加した杉本さん, 関山さん, 田中さんをはじめとする summer student の方々にも大変お世話になりました。皆様のご支援のおかげで, 本プログラムを非常に有意義なものとすることができました。本当にありがとうございました。