

Veckomöte v6

Deltagare: Eliot, Lene-Kristian

Check-in

- Det e bra med mig! Blev både sjuk i slutet av förra veckan och upptagen med annat som tog tid från studierna (som jag inte hann ta igen under helgen) men nu är man tillbaks sedan måndags. Rutinen att dyka upp här, mötena och alla de nya ansiktena här börjar bli väldigt bekant och detta är fortsatt toppen!

Workshop v6 bullet points

- Wins.
 - Clarity on GNN vs Transformer architecture choices,
 - found relevant papers on ML based reconstruction similar to my project
- Help/Stuck
 - Was interrupted end of last week and got sick as well, did not manage to catch up during weekend
 - Struggling on using ClusterAnalyzer from ldmx-sw but not "stuck" per-say
- Focus this week:
 - Producing nice histogram plots with ClusterAnalyzer,
 - Moving onto using the newly implemented CLUE + ParticleFlow analyzers (?) Lene-Kristian recently implemented

Uppföljning "CLUE-punkten"

- Recap om min kodbas
 - <https://github.com/cat2code/mpetren-msceng-ldmx> (branch "eliot" är working branch och mest uppdaterad)
 - Observera att mitt repo är i detta stadie inte tillräckligt snyggt hanterad för att en person ska kunna ladda ner den och bara köra från sin dator utan problem, det är en del "skitfiler" som bör sitta i en .gitignore (mer än vad det redan är).
 - I.e repot är "tungt" just nu.
 - Men med det sagt, kika gärna på relevanta filer som jag nämner nedan, om du vill.
 - Kod-setup: jag har en trevlig setup på min dator som känns bra.
 - IDE: VS Code + devcontainer.json
 - Öppnar env `ldmx/pro` (nu uppgraderad; samma API som current `ldmx-sw` validation scripts).
 - Pylance interpreter är kopplad mot LDMX Python package -> jag får fin färgkod av koden jag läser.
 - Har en virtual env för användning av python utanför LDMX-SW (analys: skapa plots, skapa CSV filer av output med uproot, GNN: för att leka runt med torch, ...)

- **Mål "bli bekant med CLUE":**

- *Jag har gjort:* Reproducerat `ldmx-sw it_pileup` validation lokalt → skapa inputs + köra overlay/reco/DQM → få `hist.root` med analyzers/histogram.
 - Jag började först försöka snickra ihop egna producers/analyzers, men jag tror jag misstolkade "uppgiften" du gav mig (?) och bytte till att utgå direkt från `ldmx-sw validation examples` (då kom jag igång känns det som).

- **Genomfört (it_pileup workflow):**

- **Inputs genererade:**

- `fire gen_main.py` → `ecal_pn.root`
- `fire gen_pileup.py` → `pileup.root`

- **Huvudkörning:**

- `fire config.py` (OverlayProducer + tracking + ECal/HCal digi/reco + vetoes + trigger + CLUE + DQM)
- Output: `events.root` + `hist.root`
- Körda events: **10** (pga `LDMX_NUM_EVENTS=20` och config gör `//2`)

- **Verifiering i notebook:**

- Histogram i `hist.root` och jag har plottat **9** relevanta distributions.
- Ex: `EcalClusterAnalyzer/*` finns med många nycklar (t.ex. `number_of_clusters`, `mixed_hit_energy`, `energy_percentage`, `unclustered_hits(_percentage)`, residualer mot scoring plane etc.).

- **Status:** Overlay + tracking + calorimeter reco + veto/trigger + CLUE + flera DQM-analyzers **funkar** (`hist.root` skapas, plots går att göra).

- *Jag har inte gjort:* PF/PileupFinder-del failar:

- `PFTTrack` hittar inte `EcalScoringPlaneHitsOverlay_overlay` → `PFlow/PileupFinder` får saknade collections.
- Jag prioriterade att först få CLUE/DQM-histogram på plats; PF debug tar jag som separat nästa steg.

Uppföljning GNN

- Jag har gjort en toy-model i jupyter notebook för att leka runt lite på egen hand (tidig övning inför att applicera på LDMX simulerad data eller att använda en annan arkitektur på liknande sätt.)
 - GNN (GravNetConv) node classification task of gaussian blobs (2D) particle clouds.
 - Synthetic data: 2D gaussian particles generation + background noise. -> symboliserar 1 detector layer och ECal hits i denna.
 - (x, y, energy)

- GravNetConv layers i (PyTorch) dynamiskt konstruerar k-nearest-neighbor graph och gör message-parsing mellan noder.
- Output: Cluster assignment A/B/C + noise assignment
- **Brist:** hittills i bara 2D. Kan utökas till 3D genom att lägga till en z-dimension med fler detektorlager.

Punkter antecknade från förra veckan (v5)

- Lek runt med "toy model" för att känna på GNN lite grand
- Försök att hitta ett så enkelt sätt som möjligt att angripa ECal rechit data med en GNN.
 - "Hur applicerar vi en GNN till ECal"
 - Kika på hur ParticleNet används i LDMX-SW och se ifall det finns saker man kan kopiera därifrån eller ta inspiration i övrigt
 - Annars finns kanske GravNet eller ngt annat...
- Skicka ClusterAnalyzer histogram, ASAP!

Punkter till nästa vecka

- Debug PFlow/PileupFinder crash
 - Gör bug report
- Extrahera data från LDMX simulering (.root filer) som input till en ML arkitektur?
 - Ta inspiration från LDMX ParticleNet
 - Plocka ut 2d projektion kanske 😊
- Läsa mig in i hur track kan utnyttjas i min reconstruction-modell
 - Kolla paper jag hittade förra veckan