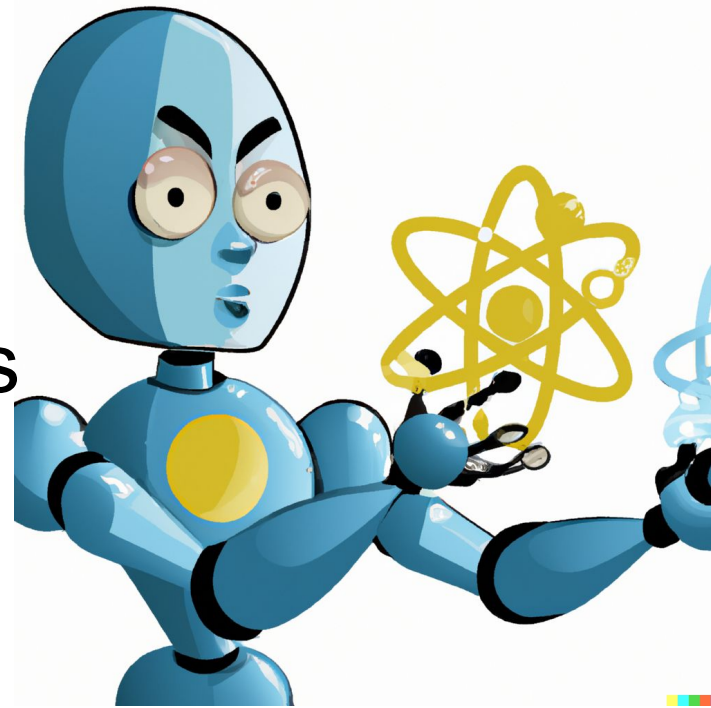


Physics Event Classification Using Large Language Models

2023 AI4EIC Hackathon Tutorial
December 1, 2023



Generated by DALL-E 2 (Prompt: "A robot holding an atom in each hand looking pensive.")

Overview of the Hackathon Challenge

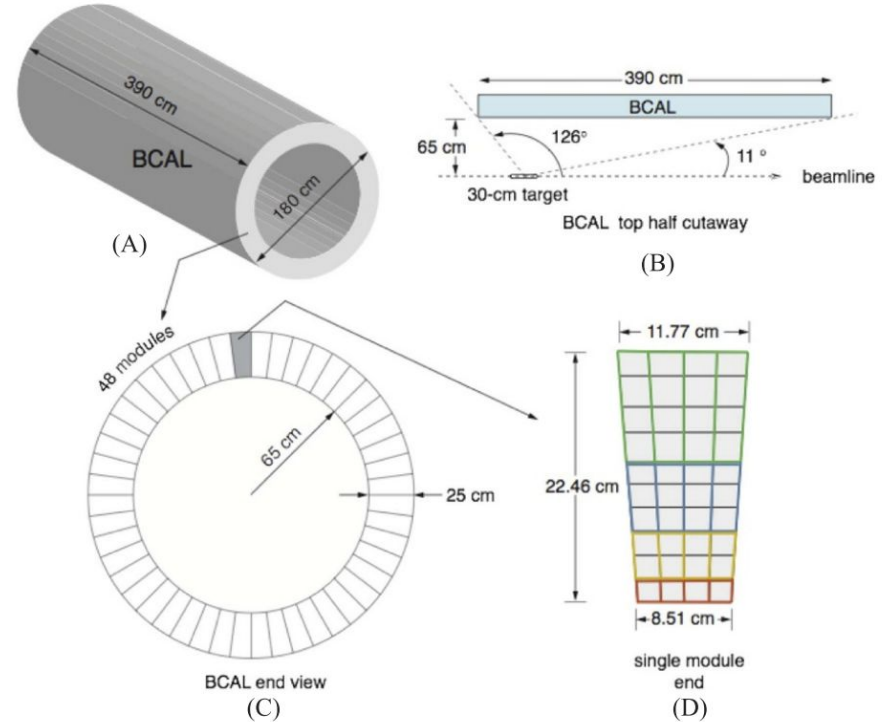
The hackathon will involve using an LLM (ChatGPT-3.5) to write a binary classifier for electromagnetic calorimeter events. This hackathon will involve no programming by the participants. Rather, you will gain experience in prompt engineering with an LLM to refine your physics result.

Large Language Models (LLMs)

- Large Language Models could potentially transform the way that science is done
- Have already resulted in novel methods and discoveries in the medical field
- Applications for nuclear and particle physics are in their nascent stages
- A number of LLM talks at yesterday morning's session (Karthik Suresh, Daniel Murnane, and others)

GlueX Barrel Calorimeter (BCAL)

- Cylindrical electromagnetic sampling calorimeter
- Scintillating fibers interleaved with lead
- Consists of 48 modules with 4 readout layers



Data

- Neutron and photon showers in the BCAL
- Consists of 14 variables (see the doc on Indico for complete information)
 - Radial depth
 - Energy deposited in each layer
 - Energy fraction deposited in each layer
 - Shower width in z , r , θ , φ and t
- Split into training and testing data
 - Training data contains a column for the PID: 0 for neutrons, 1 for photons
 - Testing data contains a column for the Event ID

Amazon Web Service Instance

- The code from ChatGPT and the calorimetry data will be located on your own AWS instance
- *ssh user@address_we_provide_you*
- You will be responsible for setting up your own conda environment and installing necessary Python packages (details in the PDF on Indico)

Login at <http://34.224.57.224:5500/login>

✕

⋮

main app

chatGPT

login


logout


submit


Welcome to AI4EIC Hackathon 2023


Login to get started

Enter your credentials

Team Name 

Username 

Password 



Creating a Session and Setting the Context

main app

chatGPT

login

logout

submit

×

Welcome to AI4EIC Hackathon 2023

Hi FirstName1 LastName1 🗨️ from Awesome Team 👤

Start a Chat Session

Session Name

Session-0

Session Context

You are a good chat bot.

Start Chat Session

Pushing Code to AWS



How can I help you?



Write a Hello World program in Python.



Certainly! Here's a simple "Hello, World!" program in Python:

```
print("Hello, World!")
```

Please slow down...

Push to AWS instance

Ask your question



Executing Code and Submitting

- Run your code from within the workspace directory of AWS
- Submit a CSV file with two columns: eventID (column 1) and PID (column 2)
- All team members may submit as many times as they would like
- Only the top score for each part will be counted



main app

chatGPT

login

logout

submit



Welcome to AI4EIC Hackathon 2023

Submit new Solution

Question Number:

None



Select



main app

chatGPT

login

logout

submit

Welcome to AI4EIC Hackathon 2023

Submit new Solution

Question Number:

Question 1



Select





main app

chatGPT

login

logout

submit

Submit new Solution

Question Number:

Question 1



Select

You selected: Question 1

Please ensure your data is in this format:

eventID	PID
8	0
9	1

File Path

Enter the path relative to \$WORKSPACE Only `.csv` files`

Submit for grading

Evaluation

- Your score for each part will be the fraction of correctly classified events, times 100
- The total score will be the sum of the scores from the two parts
- Uncertainty in the score will be calculated
- The tiebreaker for teams within uncertainty of each other will go to the team with the least amount of prompts

Select

Please ensure your data is in this format:

eventID	PID
8	0
9	1

File Path

test_predictions.csv

Submit for grading

Success!

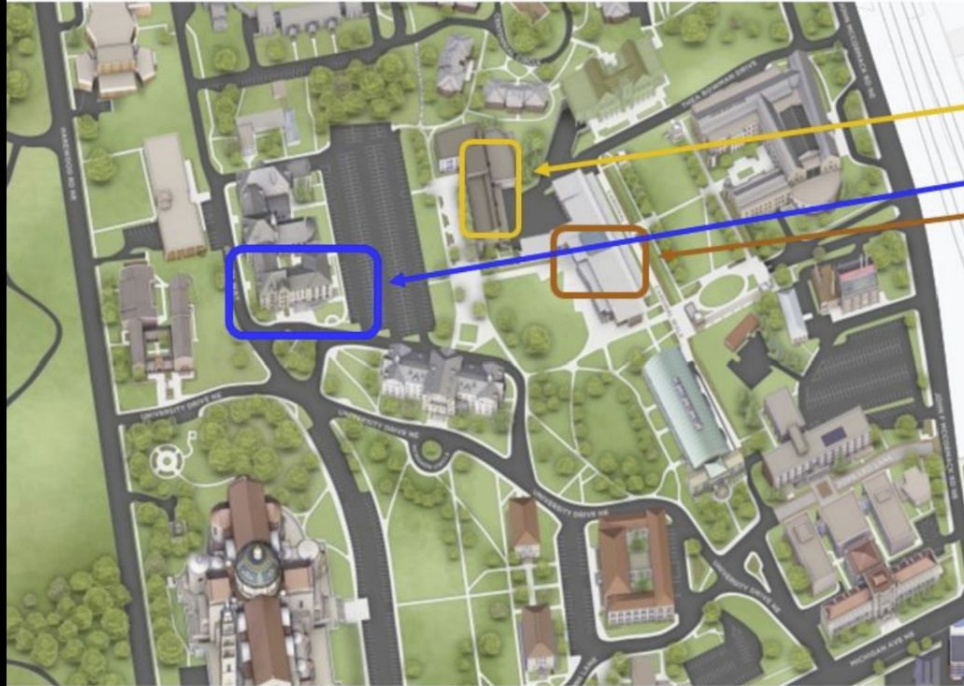
Wow, what a score!

96.2445 \pm 0.043333012819327485

Hackathon Logistics and Support

- Runs from 10:00-17:00 ET
- Teams in-person at CUA have four rooms to choose from:
 - Caldwell Auditorium
 - Hannan Hall Room 203
 - Hannan Hall Room 231
 - Pryzbyla Center Room 321
- Karthik Suresh, James Giroux and Patrick Moran will be available to help answer questions
- You can contact us on the AI4EIC Slack or in the Zoom room if you are not in Caldwell Auditorium
- [Zoom link](#)

Rooms for the Hackathon



For the hackathon on Friday, the following rooms are available:

- Hannan Hall (Physics building)
Rooms 203 and 231
- Caldwell auditorium
- Pryzbyla Room 321



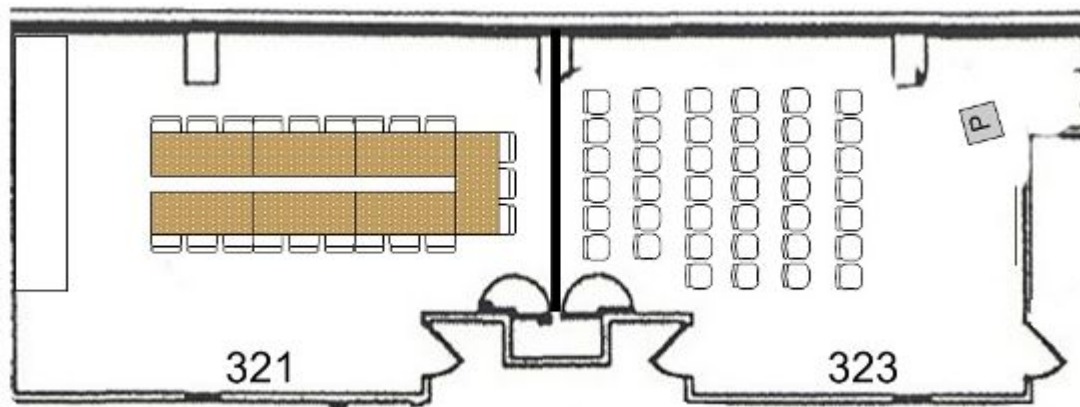
Caldwell Hall



Pryzbyla University Center



Hannan Hall



Pryzbyla Room 321