

ENHANCED AI LLM-BASED PAPER REVIEWER FOR THE CMS EXPERIMENT

Annunziata Álvarez-Cascos

Supervisor: Florian Rehm

BE-CSS DSB



Automating Peer Review in CMS Experiment Publications with LLMs

We aim to automate and enhance the peer review of CMS experiment papers using a fine-tuned Llama 3.1 model.

This should result in improved paper quality and streamline the review process.

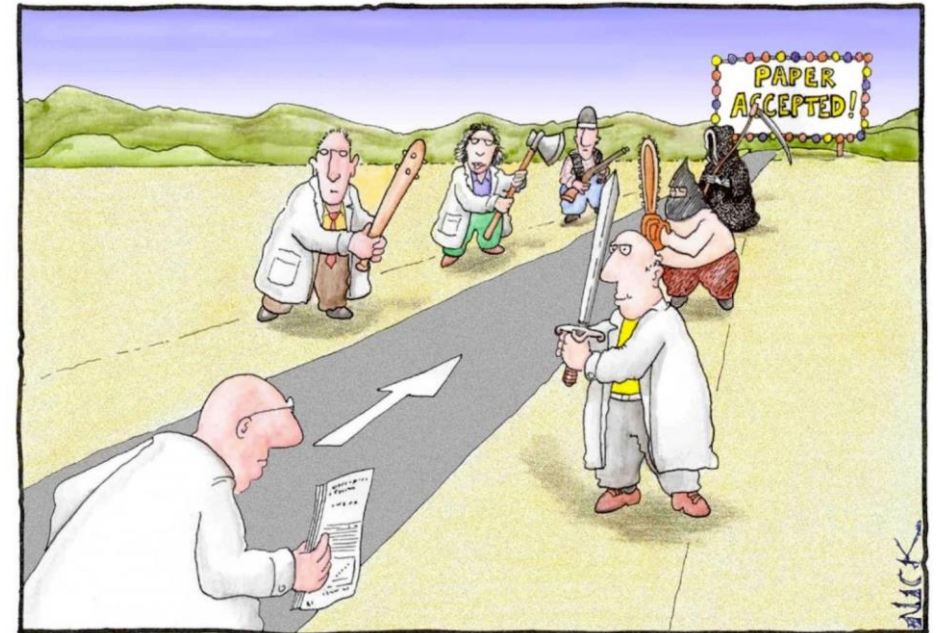
Table of Contents

- 01 Problem Statement
- 02 Proposed Solution
- 03 Process Overview
- 04 Training Data
- 05 Input and Output Examples
- 06 Status Overview
- 07 Next Steps

Problem Statement

Peer reviewing is **time-consuming**:

- Diverse writing styles.
- **Inconsistencies in writing styles** cause misunderstandings and slow down the review process.
- **Not all authors follow recommended guidelines**, leading to readability issues.



<https://scribblygumblog.wordpress.com/2015/09/30/a-first-timers-guide-to-peer-review/>

Proposed Solution

Utilize the advantages of AI to improve the peer review process.

Develop an LLM model to streamline peer reviews in the CMS experiment.

Fine-tune the model with over 1300 LaTeX documents from CMS.

How?

Use Llama 3.1 8B

Implement PEFT
(Parameter Efficient
Fine-Tuning)

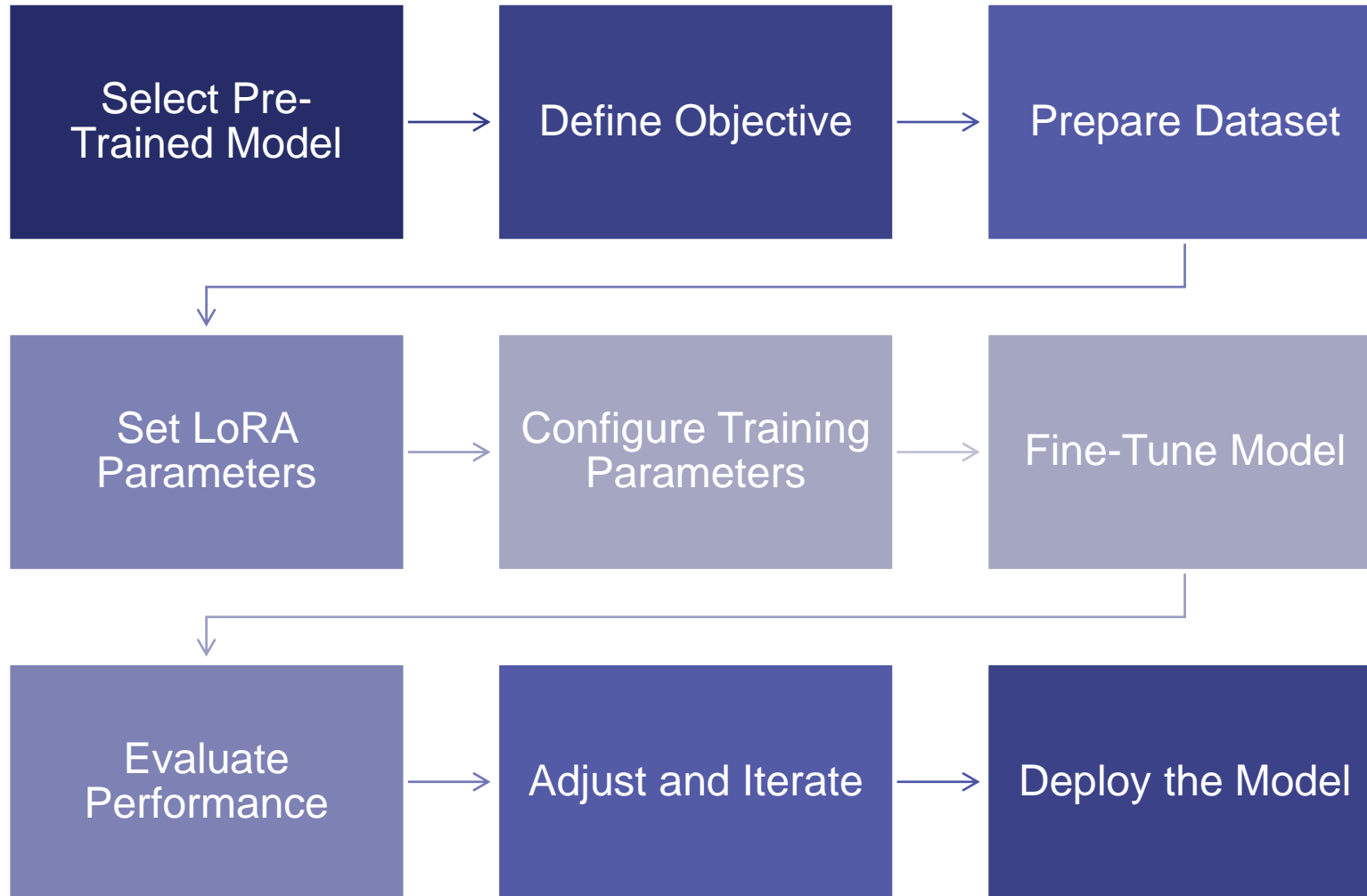
Inference

- **8B Parameters:** Trained on vast text data to generate human-like responses.
- **Instruction-Tuned:** Optimized to follow human instructions effectively.
- **Versatile Applications:** Suitable for a wide range of tasks

- Allows models to be fine-tuned with fewer parameters, making training more efficient.
- **LoRA (Low-Rank Adaptation):** Reduces the number of trainable parameters by using low-rank matrices, saving computational resources.

- Using the fine-tuned model to get the desired outputs, allowing it to apply learned knowledge to provide instant results.

Process Overview



Training Data

Collection Process

Web Scrapping

More than 1300 peer-reviewed LaTeX documents

```
DatasetDict({
  train: Dataset({
    features: ['text'],
    num_rows: 115253
  })
  test: Dataset({
    features: ['text'],
    num_rows: 6066
  })
})
```

Each row contains a paragraph from the documents.

```
' \\caption{Invariant mass\n      distributions for the  $\\mu\\mu$  system used to reconstruct the\n       $\\mu$  normalization sample. The plot on the left show\n\\begin{table}[hbt]\n  \\begin{center}\n    \\caption{Summary of the systematic uncertainties for the\ninclusive fiducial cross section measurements.\\label{tab:syst}\n  \\section{Determination of the strong coupling constant}\\label{sec:alphas}}\n  Similar to the  $X$  to  $Z$  to  $4\\ell$  study above, ten spin-two hypotheses, listed in Table~\\ref{table-scenarios},\nand three spin-one hypotheses, including\n  A detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant\nkinematic variables, can be found in Ref.~\n  The differential cross sections are measured for two rapidity ranges:  $\\abs{y} \\leq 0.6$  and  $0.6 < \\abs{y} < 1.2$ , as well as for the entire range  $\\abs{y} < 1.2$ .
```

Input and Output Examples

Original paragraph (Sections):

where angular polarization coefficients A_i (pt^{PZ} , Y^{PZ} , M^{PZ}) represent the ratio of cross section with a certain type of PZ boson polarization to unpolarized one. The coefficients reflect the processes occurring at the parton level, so all of them are expected to be negligible at $\text{pt}^{\text{Z}} \sim 0$ except for A_4 , which connected with the electroweak nature of PZ boson decay to couple of leptons, occurring with parity violation. This phenomena is also known as forward-backward asymmetry (A_{FB}) . Coefficients A_0 and A_2 describe longitudinal (L) and transverse (T) polarizations of PZ boson. A_1 - interference of L- and T- polarization states. An equation $A_0 = A_2$, famous as Lam-Tung relation [\cite{LumTung_1978}](#), have a sense of system rotation invariance. However, this ratio does not satisfied due to the non-coplanarity of the interaction planes of involved partons and colliding protons, and other higher-order QCD effects. A_3 , like A_4 based on axial and vector electroweak couplings of quarks and leptons and could be used for precise measurements of sqs . A_5 , A_6 and A_7 are expected to be small and may appear non-zero at NNLO QCD.

Corrected paragraph (Sections):

Corrected paragraph:

The angular polarization coefficients A_i (pt^{PZ} , Y^{PZ} , M^{PZ}) represent the ratio of the cross section with a certain type of PZ boson polarization to the unpolarized one. The coefficients reflect the processes occurring at the parton level, so all of them are expected to be negligible at $\text{pt}^{\text{Z}} \sim 0$ except for A_4 , which is connected with the electroweak nature of PZ boson decay to a couple of leptons, occurring with parity violation. This phenomenon is also known as forward-backward asymmetry (A_{FB}) . The coefficients A_0 and A_2 describe longitudinal (L) and transverse (T) polarizations of the PZ boson, respectively. A_1 describes the interference of L- and T- polarization states. The equation $A_0 = A_2$, known as the Lam-Tung relation [\cite{LumTung_1978}](#), is a consequence of system rotation invariance. However, this ratio does not satisfy due to the non-coplanarity of the interaction planes of the involved partons and colliding protons, and other higher-order QCD effects. A_3 , like A_4 , is based on axial and vector electroweak couplings of quarks and leptons and could be used for precise measurements of sqs . A_5 , A_6 , and A_7 are expected to be small and may appear non-zero at NNLO QCD.

Changes made (Sections):

- Added commas to improve sentence structure and readability.
- Added articles ("a", "the") to improve sentence structure and clarity.
- Changed "have a sense of" to "is a consequence of" to improve clarity and precision.
- Added "respectively" after listing the polarizations described by A_0 and A_2 .
- Changed "could be used for precise measurements of sqs " to "could be used for precise measurements of sqs " (added dollar signs for LaTeX formatting).
- Changed "does not satisfied" to "does not satisfy" (correct verb form).
- Changed "like A_4 based on" to "like A_4 , which is based on" to improve sentence structure.
- Changed "a couple of leptons" to "a couple of leptons" (added quotation marks for clarity).
- Changed "phenomena" to "phenomenon" (singular form) for consistency.

Status Overview

Training & Validation Accuracies: The model shows signs of learning, though there is still room for improvement.

Data & Accuracy: We are continuing to explore different approaches and parameters to achieve appropriate outputs.

Model Efficiency: Fine-tuning can take multiple days to weeks on a small 4-GPU cluster.

