

Project Title

Mitigating and Evaluating Hallucinations in Large Language Models

Project Description:

A hallucination occurs when a model generates false, misleading, or nonsensical information that is not grounded in real-world data or facts. This project focuses on mitigating hallucinations in FANAR models. Students will interact with the model at various stages to evaluate and refine its outputs by curating new datasets, designing evaluation metrics, and developing methods to detect hallucinations. They will collaborate closely with the team responsible for the post-training of FANAR models, gaining hands-on experience in improving model robustness and accuracy.

Project Type: Research or Engineering

Basic AI/ML engineering, including reading, understanding, and implementing academic papers.

Internship Batch: Open to both batches, priority will be on the first batch.

- **Batch 1:** May 11 to July 10, suitable for Education City students, i.e., CMUQ, TAMUQ and HBKU students
- **Batch 2:** May 25 to July 24, suitable for QU university students

Duties/Activities:

The role may include at least two of the following tasks:

- Reading and analyzing academic papers.
- Large-scale data curation.
- Assisting in the development of training routines.

Required Skills:

- Machine Learning & AI – Understanding of NLP models and deep learning concepts.
- Python Programming – Experience with libraries like Hugging Face and PyTorch, and familiarity with relevant open-source projects on GitHub is a big plus.
- Data Processing – Ability to clean, curate, and analyze large datasets.
- Research & Critical Thinking – Reading academic papers and applying findings.
- Collaboration & Communication – Working with researchers, documenting findings, and presenting insights.

Preferred Intern Academic Level:

Open to all levels, with preference for senior-level students.

Learning Opportunities:

Gain hands-on experience in real-world AI/ML engineering.

Expected Team Size: *it is preferable to have team projects*

Multiple positions are available, involving both individual and team-based tasks.

Mentors

Name: H. T. Sencar, E. Altinisik, M. Fatehkia

email: hsencar@hbku.edu.qa, ealtinisik@hbku.edu.qa, mfatehkia@hbku.edu.qa