## Project Title: Domain Specialized Large Language Models for Factual Knowledge

**Project Description:** This project focuses on developing domain-specialized large language models (LLMs) with enhanced factual knowledge in specific fields such as news domain. The intern will explore methods to fine-tune, adapt, and evaluate LLMs for improved accuracy and domain relevance. Tasks may involve curating domain-specific datasets, injecting factual updates, and assessing model performance against expert benchmarks. The goal is to create more reliable and knowledgeable LLMs that can support domain-specific applications with high factual integrity.

**Project Type:** Research and development

**Duties/Activities:** The students will be asked to work on one or more of the following:

- Specialized LLMs model design and development
- Deploy the model and create an API

**Required Skills:**

- **Fundamental Knowledge** of AI/Machine learning/NLP
- **Programming Proficiency:** Strong programming skills in Python
- **Problem-Solving and Research Ability:** The ability to analyze and implement AI/ML solutions, work with large-scale datasets

**Preferred Intern Academic Level:** We accept all levels: PhD, MSc, Senior undergrad students enrolled in CS, CSE.

**Learning Opportunities:** You will gain hands-on experience with the technical foundations of large language models (LLMs), including their training, fine-tuning, and evaluation. With strong dedication and contributions, there is potential for a research publication based on your work.

**Expected Team Size:** 2-3 people

**Mentors**

Firoj Alam (fialam@hbku.edu.qa, https://firojalam.one/)