FARJANA SULTANA SAMIA

AI/NLP Researcher — Ph.D. Student, Iowa State University

+1 515-357-3579 — fssamia@iastate.edu — linkedin.com/in/farjanasultanasamia — farjana-02.github.io/Samia

PROFESSIONAL SUMMARY

Ph.D. researcher specializing in NLP, Large Language Models, and Retrieval-Augmented Generation (RAG) with a year of industry experience deploying NLP systems at production scale. Passionate about applying foundation models and LLM techniques to complex, real-world systems.

EDUCATION

Ph.D. in Computer Science — Iowa State University, Ames, IA *Jan 2024 – Present*

- GPA: 3.92 / 4.00

B.Sc. in Computer Science and Engineering — *Feb 2017 – Dec 2021*
Manarat International University, Bangladesh

- GPA: 3.90 / 4.00 — Graduated with Distinction; 1st rank among 48 graduates

RESEARCH EXPERIENCE

Graduate Researcher Assistant — Data and Knowledge Mining Lab — *Jul 2024 – Present*
Iowa State University, Ames, IA

- Conducting advanced research on NLP and LLMs including transformer/foundation models, knowledge graphs, and RAG pipelines for complex information retrieval tasks
- Building Entity Linking (EL) and Relation Extraction (RE) systems that integrate structured knowledge with language models
- Applying prompt engineering and fine-tuning of foundation models to solve domain-specific language understanding problems

PUBLICATION

Bridging Lexical Divergence: LLM-Assisted, Cost-Efficient, Zero-Shot Scientific Entity Linking *Findings of EMNLP 2026*
Contribution: Developed an annotation-free pseudo-labeling and filtering pipeline for model fine-tuning.

PROFESSIONAL EXPERIENCE

Graduate Teaching Assistant — (COM S 327, COM S 227) — *Jan 2024 – Present*
Iowa State University, Ames, IA

- Advanced Programming Techniques: mentoring students in algorithm design, debugging, memory management, pointers, templates, and STL (C++)
- Object Oriented Programming(OOP): led lab sessions, supported students with Java and OOP concepts, graded assignments, and collaborated with instructors

Software Engineer (NLP) — Technometrics Limited, Dhaka, Bangladesh *Dec 2022 – Dec 2023*

- Built Bengali multi-label topic classification model using fine-tuned BERT (Flair, PyTorch, FastText) achieving 85% accuracy across 17 topic classes

- Developed Bengali aspect-based sentiment analysis system (PyABSA) to predict per-aspect polarity scores within a single sentence
- Engineered social media and news scraping bot (Selenium, MongoDB, Multiprocessing) with automated scheduling and REST API delivery; includes Word Cloud, trending topics, and viral post analytics
- Built scalable scraping pipeline covering 25 Bengali blogs and 15 news portals using Scrapy and MongoDB

KEY PROJECTS

Advanced RAG System with Multi-Query & Contextual Compression

- Designed RAG pipeline using GPT and LLaMA with chunking, re-ranking, and contextual compression to improve LLM response accuracy
- Deployed as an interactive web app using Funix, enabling non-technical stakeholders to query the system via browser

Sentiment Analysis with Aggregated Token Attention Network (ATAN)

- Designed multi-model sentiment classifier on Amazon Fine Food Reviews using BERT variants (BERT-Base, DistilBERT, multilingual); achieved 77% accuracy via token-level attention aggregation
- Applied lemmatization, class balancing (60K samples, 3 classes), and feature engineering for optimized performance

Object Recognition – CIFAR-10 & Red Blood Cell Counting via Transfer Learning

- CNN-based image classification on CIFAR-10 (60K images, 10 classes) using Keras; cell detection using Faster RCNN, SSD, and YOLO — demonstrates versatility beyond NLP

TECHNICAL SKILLS

Languages: Python (expert), C/C++, C#

AI/ML: PyTorch, TensorFlow, Keras, HuggingFace Transformers, LangChain, LlamaIndex, FAISS

NLP/LLM: BERT, DistilBERT, GPT, LLaMA, RAG, Entity Linking, Knowledge Graphs,Prompt Engineering, Model Fine-Tuning

Web & APIs: Flask, Django, FastAPI, HTML/CSS, JavaScript, Elasticsearch, SQL

Databases: MySQL, MongoDB, Neo4j (Graph DB), Pinecone (Vector DB), Semantic Search

Scraping: Scrapy, Selenium, BeautifulSoup, Requests

Tools & OS: Git, Jupyter, VS Code, PyCharm, Linux (Ubuntu)