

Ibrahim Al Azher

🏠 ibrahimalazhar.com 📧 scholar.google.com/ibrahimalazhar 🐙 github.com/ibrahimalazhar
🌐 www.linkedin.com/in/ibrahim-al-azhar-a68821158 ✉ ibrahimalazhar264@gmail.com 📞 +1 7373735076

AREA OF EXPERTISE

My research area is machine learning, science of science, information extraction, natural language processing, and Large Language Models (LLM). My research goal is to make scientific articles more reliable, credible, robust, reproducible, and transparent to the research community and policymakers. My work also incorporates Large Language Models (LLMs) techniques such as zero-shot, few-shot, chain of thoughts, self-consistency, and fine-tuning alongside Retrieval Augmented Generation (RAG) for text generation and topic modeling. My work also digs deep into robust evaluation metrics for LLM and RAG. My works are on multimodal LLM in images alongside texts in scientific articles, LLM for code vulnerability detection, and image analysis with multimodal LLM. Check out my blog about some of the recent developments on LLM: [ibrahimalazhar/LLM](https://ibrahimalazhar.com/LLM)

EDUCATION

Northern Illinois University

PhD Computer Science; GPA: 3.65

Dekalb, IL

August 2023 - May 2025

Northern Illinois University

M.S Computer Science; GPA: 3.65

Dekalb, IL

August 2023 - May 2025

Shahjalal University of Science and Technology

B.Sc. Computer Science; GPA: 3.69

Sylhet, Bangladesh

March 2016 - March 2021

EXPERIENCE

Research Assistant

Northern Illinois University

August 2023 - Continue

Dekalb, IL

- Extracting texts from scientific articles using API. Incorporating LLM techniques such as zero-shot, few-shot, chain of thoughts, self-consistency, self-improvement, and fine-tuning with Retrieval Augmented Generation (RAG) for text generation.
- LLM for code vulnerability detection, image analysis with large vision language models
- Contributed to the research, publishing papers, and writing of an NSF grant proposal.

Software Engineer

BYSL Global Technology Group

June 2021 - Feb 2022

Dhaka, Bangladesh

- Works on scalable projects from design to coding, testing, and installation using Python, XML, Javascript, PostgreSQL, and jQuery in Linux OS.
- Linux OS. Designed and developed various ERP-based business software and modules like sales, inventory, purchase, website, and e-commerce using Python language in Odoo.

Junior Software Engineer

Unisoft Systems Limited

December 2020 - May 2021

Dhaka, Bangladesh

- Research, assess, and lead the initiation of new technologies to maximize performance. Interact with clients and analyze requirements and develop an application in Python.
- Had written various SQL to generate numerous reports and created various modules using the OOP, Database model.

Published

- **Ibrahim Al Azher**, Venkata Devesh Reddy Seethi, Akhil Pandey Akella, and Hamed Alhoori. “LimTopic: LLM-based Topic Modeling and Text Summarization for Analyzing Scientific Articles’ Limitations.” 2024 ACM/IEEE Joint Conference on Digital Libraries, December 2024, Hong Kong. We introduced LimTopic, a strategy for using topic generation with meaningful titles and descriptions in limitation sections in scientific articles with Topic Modeling and Large Language Models (LLMs). Traditional topic modeling, or zero-shot LLM, fine-tuned LLM cannot extract meaningful titles alongside the relevant text from diverse data sources. [Paper](#). [Arxiv](#). [Code](#).
- **Ibrahim Al Azher**, and Hamed Alhoori. “Generating Suggestive Limitations from Research Articles Using LLM and Graph-Based Approach.” 2024 ACM/IEEE Joint Conference on Digital Libraries, December 2024, Hong Kong. [Paper](#)
- **Ibrahim Al Azher**, and Hamed Alhoori. “Mitigating Visual Limitations of Research Papers.” 2024 IEEE International Conference on Big Data (IEEE BigData 2024), December 2024, Washington D.C. We focus on visual limitations in scientific research papers such as charts, graphs, and diagrams. Our task is to build a multimodal LLM to generate clear, concise descriptions, regenerate visual images, and explainable tabular forms based on images. We integrated QWen, Llava, Llama, and GPT 4o to generate descriptions from pictures as a zero-shot learning. Applying prompt engineering and getting feedback from LLM and LLM as a judge approach is used here. Part of this work was published as a poster. [Paper](#). [PDF](#).
- **Ibrahim Al Azher**, Hamed Alhoori. “Generating Suggestive Limitations from Research Articles Using LLM and Graph-Based Approach” The 11th IEEE International Conference on Data Science and Advanced Analytics (DSAA 2024), October 2024, San Diego, CA . The graph-based method involves constructing a limitation recommendation system using a graph neural network to represent relationships among scientific articles based on their limitations. Node represents a limitation section alongside other sections. Creating a knowledge graph using citation of papers building edges will suggest a new paper’s additional limitations based on the knowledge graph with GraphSage. In addition, we are incorporating the graph to LLM by i) sending the subgraph and ii) sending the textual data to LLM to generate refined limitations.
- **Ibrahim Al Azher**, Sohel Ahmed, and Md. Saiful Islam. “Identifying Author in Bengali Literature by Bi-LSTM with Attention Mechanism.” 24th International Conference on Computer and Information Technology (ICCIT), December 2021. [Paper](#). [PDF](#)
- Harish Verma, Miftahul Jannat Mokarrama, **Ibrahim Al Azher** and Hamed Alhoori. “A Comparative Study of ORKG and LLM identified Research Contributions” 2nd Workshop on Innovation Measurement for Scientific Communication (IMSC) in the Era of Big Data, JCDL, December 2024, Hong Kong.
- Murtuza Shahzad, Joey Wilson, **Ibrahim Al Azher**, and Mona Rahimi. “From Theory to Practice: Code Generation Using LLMs for CAPEC and CWE Frameworks” 2nd International Workshop on Large Language Models for Code (LLM4Code 2025), ICSE 2025
- Christy Muasher-Kerwin, M Courtney Hughes, Michelle Foster, **Ibrahim Al Azher**, Hamed Alhoori. “Exploring Large Language Models for Summarizing and Interpreting an Online Brain Tumor Support Forum.” Sage Journal. [Paper](#)
- MD Saifur Rahman Chowdhury, Nahiyan Habib Khan, Dip Singha, **Ibrahim Al Azher**, Tufael Ahmed, Geerbani Pal Shashi. “Leveraging Self-Sovereign Identity (SSI) with Hyperledger Indy: A Decentralized Identity Ecosystem for Secure Document Management in Bangladesh ” 27th International Conference on Computer and Information Technology (ICCIT) 20-22 December 2024, Cox’s Bazar, Bangladesh.

In Progress

- **Ibrahim Al Azher**, Miftahul Jannat Mokarrama, Zhishuao Guo, Sagnik Ray Choudhury, and Hamed Alhoori. “FutureGen: Trend Analysis and LLM-RAG approach to generate ‘future work’ of the scientific article.’ Generate scientific article future work with LLM and RAG, and LLM feedback to improve the result. Analysis the trend with topic modeling + LLM. [PDF](#)
- **Ibrahim Al Azhar**, Bhanu Kedhaar, Sagnik Ray Choudhury, and Hamed Alhoori. “Recommended limitations from cited papers”. Generating extended limitations from cited papers, the information of the cited documents is stored in the RAG database, and for each paper, RAG will be updated. Propose a new robust framework for retrievers and generators and use RLAIIF for improvement. Also, hierarchical limitations and trend analysis using time series.
- Tamzid Azad, **Ibrahim Al Azher**, and Hamed Alhoori. “Predicting Scholarly Impact with Retrieval-Augmented LLMs”.

- Tamzid Azad, **Ibrahim Al Azher**, and Hamed Alhoori. “How Large Language Models (LLM) affect the Readability Score in Scientific article. ”

Machine Learning works | *Python, Keras, Tensorflow*

Jan 2020 – July 2022

- Predicting mortality, hospital length of stay in health dataset (MIMIC IV). Analysis Amazon reviews datasets with various models and data preprocessing. Apply CNN, LeNet, AlexNet, GoogleNet, ResNet, VGG16, and VGG19 models and transfer learning models on image data.

SKILLS

Programming Languages: C, C++, Python, Javascript, Java, PHP

Framework and Technologies: Django, Laravel, HTML, CSS, Bootstrap, Odoo, Django REST framework

Open Source Library: Keras, Tensorflow, Pytorch, numpy, pandas, scikit learn, matplotlib,

Tools: Oracle Sql, MySQL, Postgresql, Git, Github, CodeBlocks, Pycharm, Visual Studio, Anaconda, Jupyter notebook, Latex, Overleaf

HONORS & AWARDS

- **Google GCP Grant Award**, Google — Feb 2024 to Feb 2025
- **IEEE Travel Grant**, IEEE International Conference on Data Science and Advanced Analytics (DSAA), San Diego, California — October 2024
- **SIGIR Student Travel Grant**, 2024 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Hong Kong — December 2024
- **NIU Student Travel Grant**, October - December 2024
- **3MT Best Paper Presentation Award**, Joint Conference on Digital Libraries (JCDL), Hong Kong — December 2024

OTHER PROFESSIONAL INFORMATION

Reviewer

- [ACL Workshop 2024, 2025], [SDM 2025]

Program Committee

- [Web Science 2025]

Session Chair

- [JCDL 2024]

Volunteer

- [DSAA 2024]