# Saad Sher Alam

Urbana, IL | +1 (447) 902 6776 | saadsa2@illinois.edu | LinkedIn | GitHub

## Education

**University of Illinois Urbana-Champaign (UIUC)**　　　　　　　　　　　　　　May 2026
*Master of Science in Computer Science - MS CS*
**Teaching Assistant:** Data Structures (Fall'24)

**Lahore University of Management Sciences (LUMS)**　　　　　　　　　　　　　May 2024
*Bachelor of Science in Computer Science - BS CS*　　　　　　　　　　　　　**CGPA:** 3.85
**Teaching Assistant:** Data Structures (Spring'23), Data Science (Fall'23), Network Security (Spring'24)

## Experience

**SysNet Group, UIUC**　　　　　　　　　　　　　　　　　　　　　　May 2023 - May 2024
*Research Assistant*
- Devised a framework to analyze cost-savings for discrepant cloud-based applications by collecting HTTP traffic for each unit test. In the best case, **20.3%** savings can be realized for GET requests, and **48.4%** for POST requests.
- Led application testing to reveal discrepancies between cloud services (Azure Storage, DynamoDB, S3) and their emulators (Azurite, LocalStack). Out of the **12** open-source applications that I tested, **7** exhibit discrepancies.
- Submitted findings (under-review) at the IEEE/ACM International Conference on Software Engineering **(ICSE'25)**.
- Worked on a wrapper for SDK APIs of Azure Blob, Azure Queue, Azure Table, AWS DynamoDB, and AWS S3, to facilitate grammar-based and mutation-based fuzzing using fuzingbook in Python.

**Networks and Systems Group, LUMS**　　　　　　　　　　　　　　　　May 2022 - May 2024
*Research Data Scientist*
- Discovered latent buffer losses in YouTube streaming that account for **7%** of a 2GB data plan; proposed a solution using double-ads which minimized unnecessary buffer losses by **1.2x.**
- Collected data for **17600** YouTube videos and **46600** YouTube ads across **8** countries. This is the first large corpus of YouTube data, with **15+** features, that includes buffer data for each video. The dataset is publicly available.
- Published findings in the Proceedings of the ACM Web Conference 2024 **(WWW'24)**.
- Devised a systematic methodology to scrape data from YouTube using Selenium, Chromedriver, and PyTube API.

## Projects

**Guftaar** | *React, JavaScript, Node.js, MongoDB*
- Developed a first-of-its-kind m-health platform in Pakistan to provide virtual therapy to PWS (people who stutter).
- Successfully deployed the website and signed up our first certified speech instructor within 24 hours of deployment.

**HospitEase** | *React, JavaScript, Node.js, MongoDB*
- Developed a full-stack hospital management system by gathering requirements from Mughal Labs.
- Automated the process of appointment scheduling, generating medical records, and rating/review system.

**SherLang: A LISP-like Programming Language** | *C*
- Built my own LISP-like programming language with the following features: polish arithmetic, s-expressions, functions, conditionals, recursive functions, strings, and the standard library.
- Implemented a handwritten parser from scratch for SherLang.

**RAFT: A Distributed Consensus Protocol** | *Golang*
- Implemented RAFT as described in: In Search of an Understandable Consensus Algorithm (Ongaro et al.)
- Created a fault-tolerant key-value server on top of RAFT.

**Authorship Attribution** | *Python, Sickit-learn, PyTorch, NumPy, Matplotlib*
- Collected 5000 tweets across 5 trending Twitter handles using public Twitter APIs.
- Evaluated performance of kNN, Neural Networks, and Random Forests for authorship attribution.
- Generalized findings to 2 feature representations: TF-IDF and Word Embeddings.

## Technical Skills

**Languages**: Python, C/C++, JavaScript, Golang, SQL
**Tools and Frameworks:** React, Node.js, Selenium, Pandas, NumPy, TensorFlow, PyTorch, OpenCV, Keras, Docker, Git