Kuber Shahi

+1 858-518-4608 | kshahi@ucsd.edu | kubershahi.github.io | linkedin.com/in/kubershahi

EDUCATION

University of California - San Diego San Diego, CA
Master of Science in Computer Science, AI Specialization, GPA: 4.0/4.0 *Sep 2024 – Present*
Coursework: *Stat NLP, ML Systems, ML for Robotics, Medical Imaging, Software Engineering*

Ashoka University Sonipat, India
Bachelor of Science in Computer Science, Minor in Physics, GPA: 3.87/4.00 *Aug 2018 – May 2022*

WORK EXPERIENCE

Data Scientist Jun 2022 – May 2024
Vayana Network

- Optimized financial data processing by developing efficient, scalable ETL pipelines with PySpark on AWS, reducing processing time by 20% and enhancing data reliability for analytics workflows.
- Engineered a graph-based Business Network Analysis platform for the sales team, leveraging clustering algorithms to identify high-value potential customers, resulting in a 15% increase in business growth.
- Collaborated across teams to design and deploy a Management Information System (MIS) with Apache Superset, enhancing reporting, decision-making, and operational efficiency by 15%.
- Spearheaded the development of an NLP-driven PAN Resolution System to resolve duplicate company records across multiple data sources, improving data accuracy and reporting capabilities by 30%.

Data Science Intern Nov 2021 – Mar 2022
Ageless Partners

- Formulated a roadmap for a fitness recommendation system using wearable devices and led 3 co-interns to complete the first milestone, boosting user engagement and retention by 20%.
- Conducted data analysis and visualization using Python and Plotly, supporting the research team's de-aging investigation and achieving a 15% improvement in user recommendations and insights.

Web Developer Jun 2020 – Sep 2020
Techvik

- Redesigned and remodeled the platform's website on WIX, increasing the user traffic by up to 20 %.
- Optimized the website's load time and increased its SEO performance by 30 % through media optimization and extensive keyword tagging.

RESEARCH EXPERIENCE

Graduate Student Researcher Jun 2025 – Present
UC San Diego

- Investigating planning methods for LLM-based travel planning agents, with a focus on evaluating external planner integration for improving multi-step reasoning and constraint handling.

Privacy Preserving Neural Networks Aug 2021 – Dec 2021
Capstone Project, Ashoka University

- Advisors: Prof. Mahavir Jhawar & Prof. Debayan Gupta
- Researched secure multi-party computation (MPC) techniques to enable privacy-preserving neural network training across distributed systems.
- Implemented SecureNN protocols in C++ and analyzed their effectiveness in ensuring data privacy and computational efficiency for real-world, data-sensitive applications. [*Presentation* | *Report* | *GitHub*]

Research Intern May 2021 – Aug 2021
Mphasis Lab, Ashoka

- Advisor:Prof. Mahavir Jhawar
- Researched and implemented Privacy-Preserving Machine Learning (PPML) protocols such as SecureML and BLAZE in C++, evaluating each protocol's efficiency and security for sensitive datasets. [*GitHub*]
- Co-led the development of a secure, optimized PPML protocol for business-specific needs under Prof. Mahavir Jhawar's guidance, improving the protocol's reliability and applicability for real-world solutions.

Secure Machine Learning

Jan 2021 – May 2021

Independent Study Module, Ashoka University

- Advisor: Prof. Debayan Gupta
- Studied the impact of adversarial attacks such as Data Poisoning and Model Evasion on the performance and reliability of machine learning (ML) models.
- Demonstrated the attacks in the *Subpopulation Data Poisoning Attacks* paper to understand the impact of poisoning attacks on real-world machine learning models. [*Presentation* | *GitHub*]

Applied Cryptography

Jan 2021 – May 2021

Independent Study Module, Ashoka University

- Mentor: Prof. Mahavir Jhawar
- Examined and evaluated the security vulnerabilities in email clients that support the two primary forms of end-to-end email encryption: OpenPGP and S/MIME.
- Illustrated the attacks on various email clients outlined in the *Mailto: Me Your Secrets* paper and suggested countermeasures against them. [*Report* | *GitHub*]

KEY PROJECTS

Efficient English-Maithili Translation | *Machine Translation, Efficient Fine-Tuning*   Nov 2024 – Dec 2024

- Developed a lightweight translation model for the Maithili language, optimizing for better accuracy and easy integration into different web applications.
- Applied Low-Rank Adaptation (LoRA) and whole model fine-tuning to develop a preliminary model, achieving comparable accuracy while addressing training data and computational limitations. [*GitHub*]

Gamified Kanban Board | *Team Project, Agile Development, GitHub Actions*   Nov 2024 – Dec 2024

- Collaboratively developed a gamified Kanban board for task management with reward-based incentives, leveraging CI/CD pipelines for efficient development, automated testing, and seamless deployment.
- Designed and implemented task representation as interactive task cards with timers to track progress and set up linting and formatting workflows to ensure project-wide code quality and consistency. [*GitHub*]

Headline Generation | *Text Summarization, LLM Benchmarking, Hugging Face*   Mar 2022 – May 2022

- Engineered a headline generation model by fine-tuning Google's Pegasus LLM (568M parameters) on a cloud-based data pipeline using PyTorch and CUDA, generating concise and accurate headlines from news articles. [*GitHub*]

aCERT: Academic Credential Verification | *Web Apps, MERN Stack, Web3*   Apr 2022 – May 2022

- Led the development of a React-based web application for securely uploading and verifying academic credentials on the Ropsten Testnet, utilizing Express.js for backend functionality and Truffle for smart contract deployment. [*GitHub*]

Synthesizing DFAs using RNNs | *Sequence Learning, Language Representation*   Apr 2021 – May 2021

- Implemented an RNN-based model in Python to automate the synthesis of Deterministic Finite Automata (DFAs), demonstrating the model's ability to learn and represent formal languages. [*GitHub*]

Age and Gender Detection | *Image Classification, Deep Learning, Model Optimization*   Nov 2020 – Dec 2020

- Designed and optimized CNN models using Keras and TensorFlow for age and gender classification in facial images, comparing traditional and deep learning approaches to achieve high accuracy. [*GitHub*]

TEACHING EXPERIENCE

Undergraduate Teaching Assistant

Aug 2020 – May 2021

CS Department, Ashoka University

- **Discrete Mathematics** (CS - 1104), Spring 2021, Prof. Subash Bhalla, Class Size: 100. *Tasks:* holding weekly office hours, setting and grading assignments and test papers, and facilitating online lectures.
- **Probability & Statistics** (CS - 1208), Monsoon 2020, Prof. Mahavir Jhawar, Class Size: 80. *Tasks:* holding weekly office hours, grading assignments and test papers, and facilitating online lectures.

STEM Facilitator

Jun 2019 – Jul 2019

Maa Anandmayee Memorial School

- Partnered with MakersBox Delhi to establish a STEM lab at the school by leading equipment selection, procurement, and setup while teaching students hands-on projects in cutting-edge technologies like Robotics, IoT, VR, and Lego.

Technical Skills

Languages	Python, C/C++, JavaScript, SQL, NoSQL, Shell Scripting (Bash), R
Technologies	PyTorch, TensorFlow, Triton, Pandas, NumPy, Scikit-learn, React, Node.js, Django, Apache Spark, Hive/Hadoop, Selenium
Dev Tools	AWS, Git, Docker, Linux, HuggingFace, LangChain, Apache Superset, Dash, Power BI

Leadership & Activities

Resident Assistant, Office of Student Affairs (OSA)	2021 – 2022
Peer Mentor, Office of Learning Support (OLS)	2020- 2021
IT Team, Ashoka University Election Commission (AUEC)	Feb 2020
Event Organizer, Ashoka University International Student Association (AUISA)	2018- 2020
Ashoka Volleyball Team, Ashoka University	2018- 2020

Awards & Scholarships

Bronze Medal, CS Awards, Ashoka University	Jun 2022
Silver Medal, CS Awards, Ashoka University	Jun 2021
Dean's List, All 8 Semesters, Ashoka University	All 8 Semesters
Full Scholarship on Tuition and Residence, Ashoka University	2018- 2022