

Rahasya Barkur (SHE/HER)

4134660747 | BARKURRAHASYA@GMAIL.COM | LINKEDIN.COM/IN/RAHASYA/ | GITHUB.COM/RAHASYABARKUR

EDUCATION

University of Massachusetts Amherst

MA, US

Master of Science in Computer Science, GPA: 3.83/4

Sep'24 - May'26

Courses: Neural Networks, Computer Vision, Reinforcement Learning, Advanced NLP, Optimization, Applied Information

Retrieval, Software Engineering Theory and Principles, Data Science Fundamentals

National Institute of Technology Karnataka (NITK)

Surathkal, India

Bachelor of Technology in Electronics and Communication, GPA: 8.25/10

Aug'17 - May'21

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, Scala, R, C++, C

ML Libraries: Keras, Tensorflow, Pytorch, OpenCV, Scikit-Learn

Big Data & Cloud Technologies: PySpark, Kafka, Hadoop, Hive, Delta Lake, Parquet, AWS, Airflow, AbInitio, Talend

DB & CI/CD Tools: Oracle SQL Developer, MongoDB, Git, LightSpeed, Docker, Jenkins, OpenShift, UDeploy

PROFESSIONAL EXPERIENCE

Data Science Intern, Dell Technologies Austin, TX

May'25 - Aug'25

- Automated microservice creation workflows using a Crew AI Agent framework, accelerating feature development.
- Optimized PySpark queries and engineered an automated data integrity check pipeline, feeding into a Power BI dashboard for real-time reliability monitoring in a big data ecosystem.

Data Engineer, Citi India

Jul'21 - Aug'24

- Developed and deployed a Streaming PySpark Scala application enabling clients to retrieve on-demand billing statements
- Designed and developed a snowflake data warehousing schema to streamline cash transaction processing across multiple sources and downstream applications, optimizing data flow and improving analytics query performance
- Coordinated database development with optimized indexing based on backend service requirements, reducing response time for the recap reconciliation web application from 4-5 minutes to under 1 minute.
- Streamlined system upgrades, cloud integration, and Kafka migration, Oracle DB 11G to 19C, Hadoop, facilitating a full transition away from legacy tools (e.g., TeamCity, Confluent, UDeploy).

Natural Language Processing Intern, Citi India

May'20 - Jun'20

- Automated insights extraction from Earnings Call Recordings using advanced NLP algorithms, saving 120+ analyst hours monthly and expediting report generation

Machine Learning Intern, NITK India

May'19 - Jul'19

- Applied U-Net model for segmentation of malignant and benign cells from hospital-provided samples, integrating a user-friendly GUI for streamlined labeling by medical professionals.

PROJECTS & PUBLICATIONS

Verifying Dynamic RAG | Supervisor: Prof. Haw-Shiuan Chang

Jan'25 - Apr'25

- Developed a validation framework by decomposing questions and validating claims in the answers using dynamic RAG.

Benchmarking Object Counting in Multimodal Video-LLMs | Supervisor: Prof. Grant Van Horn

Sep'24 - Dec'24

- Created **VisCount**, a benchmark dataset to evaluate object counting abilities of MultiModal-LLMs, on images and videos

GeoLocator | Supervisor: Prof. Brian Levine, Rescue Lab

Sep'24 - Dec'24

- Enhanced the open-source GeoCLIP model with script-country detection using CLIP embeddings and web data, and deployed application to extract image locations without metadata, aiding law enforcement with 20% improved accuracy.

Land Cover Change Segmentation from Satellite Images | Supervisor: Prof. Shyam Lal, NITK

Dec'19 - Aug'22

- Published in IEEE Transactions on Emerging Topics in Computational Intelligence: [RSCDNet: A robust Deep learning architecture for change detection from bi-temporal high-resolution remote sensing images](#).
- Assembled an encoder-decoder neural network architecture with novel attention modules to generate binary change maps from bi-temporal images, achieving a 12% improvement in F1 score over state-of-the-art models at the time.
- Further extension to the work was carried on bi-temporal **multispectral** images having 13 bands utilised and **hyperspectral** images having 198 bands. Abundance mapping estimation using hyperspectral unmixing reduced computation costs.

Speech Emotion Recognition | Supervisor: A V Narasimhadhan, NITK

May'20 - May'21

- Published [EnsembleWave An ensemble approach for Automatic Speech Emotion Recognition](#), as part of IEEE CONECCCT 2022 Conference on Electronics, Computing and Communication Technologies
- Designed a domain generalised bagging ensemble model that increased the F1 score by 12% while reduced FLOPS by 70%

EXTRA CURRICULAR ACTIVITIES

- Working as a Career Advisor in Manning College of Information & Computer Sciences at University of Massachusetts Amherst
- Participated in HackUmass 2024, a 30-hour hackathon. Built a VS Code extension, Spaghetti Sniffer, that humorously flags bad code using an LLM (Ollama) backed by graph-based Abstract Syntax Tree analysis.