

# Sachin K Rao

✉ [sachinkrao7@gmail.com](mailto:sachinkrao7@gmail.com) | ☎ +91-9632613073 | 🌐 [github.com/sachinrao24](https://github.com/sachinrao24)  
🌐 [linkedin.com/in/sachinkrao7](https://www.linkedin.com/in/sachinkrao7) | Blog: [sachinrao.hashnode.dev](https://sachinrao.hashnode.dev) | 🌐 [sachinrao24.github.io](https://sachinrao24.github.io)

## EDUCATION

- **RNS Institute of Technology** Bangalore, India  
*Bachelor of Engineering - Information Science and Engineering* 2019-2023  
**Courses:** Artificial Intelligence and Machine Learning, Data Structures and Algorithms, Automata Theory and Computability, Big Data Analytics, Database Management Systems, Computer Networks and Security, Operating Systems, File Structures

## EXPERIENCE

- **Catalyst Management Services** Bangalore, India  
*Data Scientist* July 2023 - Present  
Catalyst Management Services is a consulting firm that has been active in the social sector for over 30 years, with a commitment to transform social enterprises and work towards achieving Sustainable Development Goals (SDGs).
  - Developed a hybrid search system combining semantic search using **Weaviate**'s text2vec-contextionary model and full-text search using **MongoDB** indexing.
  - Deployed Streamlit apps on **AWS EC2** and used **Docker** for containerization, including a **REST API** built with **Flask**, for effective integration with other organizational dashboards.
  - Served as project manager for a lead-sourcing initiative for the Catalyst Group, setting project milestones, conducting stakeholder reviews, hiring and mentoring two interns, and achieving a **100-hour/month** reduction in manual effort through automation.
  - Implemented a **RAG** pipeline using **LlamaIndex** and **OpenAI embeddings** with keyword, BM25, and vector retrievers to classify and contextualize text data.
  - Built **logistic regression**, **random forest**, and **gradient boosting** models like **XGBoost**, **CatBoost** to predict dropout rates and identify at-risk students for an education NGO, achieving **88.2%** accuracy.
  - Built data pipelines to extract, transform, and load data from a digital education platform into a **PostgreSQL** database on **ClickHouse** and wrote comprehensive unit tests for each stage.
  - Conducted field research with the state government of Andhra Pradesh, collaborating with healthcare professionals to build models predicting anemia risk in pregnant women in rural areas.

### *Data Science Intern*

*April 2023 - July 2023*

- Developed web-scraping scripts using **Requests**, **BeautifulSoup**, and **Selenium** to scrape over **100,000 articles** from static and dynamic websites, reducing manual workload by over **50 hours/month**.
- Fine-tuned an NER model using manually annotated data through **Doccano** and fine-tuned a DeBERTa-v3 zero-shot classification model to classify articles with **91%** accuracy, outperforming baseline models by **12%**.
- Built a multi-page **Streamlit** application with caching and a customized UI to analyze epidemiological patterns, with visualizations using **Plotly**, **Matplotlib**, and **PyDeck**.

## PUBLICATIONS

- S. K. Rao, "N-ary relation extraction using web-scraped data," 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), Bangalore, India, 2023, pp. 611-615, doi: 10.1109/ICAECIS58353.2023.10170239.
- <sup>1</sup>S.K.Rao et al., "Identifying Hyperpartisanship in News Media Using Bi-LSTM and Decoding-Enhanced BERT," 2024 Asian Conference on Intelligent Technologies (ACOIT), Bangalore, India, 2024.
- <sup>2</sup>S.K.Rao et al., "Graph-Based Representation and Detection of Cyberbullying on Social Media Using GCNs," 2024 Asian Conference on Intelligent Technologies (ACOIT), Bangalore, India, 2024.
- <sup>3</sup>S.K.Rao et al., "Impact of Misinformation on Reddit User Behavior: A Multimodal Transformer-Based Detection Model," 2024 Asian Conference on Intelligent Technologies (ACOIT), Bangalore, India, 2024.
- S.K.Rao et al., "FormFit: A physical exercise form correction system using BlazePose," Proceedings of the 1st International Conference on Computing and Intelligent Information Systems (ICCIIS 2024), Bangalore, India, pp. 490-497, Volume 1 (1st ed.). CRC Press. <https://doi.org/10.1201/9781032711157>

---

<sup>1</sup>, <sup>2</sup> and <sup>3</sup> were accepted and submitted for inclusion in IEEE Xplore.

## PROJECTS

---

- **Transmission Control Protocol implementation using Rust:**
  - Designed a multi-threaded packet processing engine for a custom TCP/IP stack in accordance with **RFC 793 standards**, enabling real-time handling of network events.
  - Designed custom **TCP listener** and stream abstractions, integrating read/write operations, flow control, and graceful shutdown mechanisms for high-performance data communication.
- **Sequence Tagging using RNN and Bi-LSTM:**
  - Implemented an **RNN**-based acceptor to distinguish between two languages using synthetic datasets of 2,500 samples.
  - Built a sequence tagging model using **Bi-LSTM** achieving **93.2%** accuracy for Part-of-Speech tagging and **91.6%** accuracy for Named-Entity Recognition.
- **Tweet Trooper - Offensive Tweet Detection:** Winning project at IEEE CodXpress 2023
  - Built a Firefox web extension to identify and blur offensive tweets dynamically by modifying Twitter's DOM.
  - Trained a **Bi-LSTM** model on a cyberbullying tweets dataset, achieving an **accuracy of 89.2%**.
  - Hosted the model using a **Flask** server exposed via **Ngrok**.
- **Neural Network with Backpropagation for Image Classification in C:**
  - Implemented a neural network with backpropagation in **C** and designed custom data structures to manage layers, weights, biases and gradients.
  - Trained and evaluated models on the MNIST dataset for handwritten digit recognition, achieving **87.6% accuracy**.
  - Developed functions for data preprocessing, including scaling, normalization, and one-hot encoding.

## SKILLS

---

- **Languages:** Python, C, C++, Java, SQL, Rust
- **Data Analysis & Machine Learning:** Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn, Power BI, PyTorch, TensorFlow, Keras, HuggingFace, NLTK, LlamaIndex
- **Databases & Tools:** PostgreSQL, MySQL, MongoDB, ClickHouse, SQLAlchemy, dbt, Weaviate
- **Web Frameworks & APIs:** Flask, FastAPI, Streamlit, REST
- **DevOps, Web Scraping & Automation:** AWS EC2, Docker, Git, GitHub, GitHub Actions, BeautifulSoup, Selenium, Requests, Bash/Shell scripting

## ACHIEVEMENTS

---

- **Winner** of the IEEE CodXpress national hackathon, 2023, among over 80 participating teams.
- **Best Paper Award** at the National Conference on Computing Technologies, 2023, organized by the RNS Institute of Technology.
- **Best Project Award** at the Intercollege Project Competition, 2023, organized by the Bangalore Institute of Technology, with over 30 participating teams.
- **Cricket:** Represented Bangalore Cricketers in the U-19 Division 1 league and City Cricketers in the Division 4 Karnataka State League.
- **Speedcubing:** Best All India Rank of 19 for solving a Rubik's Cube blindfolded.

## LEADERSHIP AND VOLUNTEERING

---

- **re.search(club, RNSIT)** Bangalore, India  
*Founder and Chair* *April 2023 - April 2024*
  - Led and mentored a team of **80+ students**, guiding them in research projects, technical topics in machine learning, and writing academic papers.
  - Instituted a weekly blog titled "The Cognos," and increased social media engagement across multiple platforms, growing LinkedIn following to **500+ members**.
  - Recruited and built a core team of eight students, including a head of content, head of outreach, and a vice-chair, to help manage and expand the club's activities.
- **St. Broseph Foundation** Bangalore, India  
*Chapter Member, West Bangalore* *Nov 2023 - Present*
  - Managed finances, logistics, and outreach for a rabies vaccination event in collaboration with the local government veterinary hospital and an animal welfare group, *People For Animals*.
  - Lead the chapter's winter initiative to collect and distribute clothing to homeless individuals.
- **DataKind** Bangalore, India  
*Chapter Member, Bangalore* *Dec 2022 - Present*
  - Collaborated with an NGO focused on implementing the concept of Circular Economy to drive sustainability and optimize waste monitoring systems.
  - Designed and streamlined data pipelines, enabling automated updates to a Looker dashboard for real-time insights and reporting on waste usage.