Surya Kasturi

Email: kasturisurya@gmail.com
Website: https://www.suryakasturi.com/
Location: Boston, MA

Education

- M.S. Computer Science, University of Delaware, Newark, DE, 2016 – 2018
- B.Tech Electrical Engineering, JNTU Hyderabad, India, 2010 – 2014

Industry Experience

Communal Health, Boston, MA, Co-Founder, 2023–2024

- AI powered scheduling platform for post-discharge elder care
- https://communal.health

Lendbuzz, Boston, MA, Senior Software Engineer, 2021–present

- Lendbuzz is a pre-IPO unicorn startup in auto finance. I work on the Machine Learning team. I help build and maintain the automated loan approval flow and backend infrastructure
- Gen AI: Design and implementation of a chat bot for car dealers to negotiate loan terms; Developed a dialog state tracker using in-context examples and natural language rules in LLM prompt
- Automated Loan Approval: Co-owner of the service; Helped in development and risk analysis of an automated credit profiling algorithm;
- Maintain ML data platform: Identified and fixed a critical bug in the queuing system enabling parallel consumers for the first time.
- Infrastructure modernization: Streamlined legacy code by containerizing services and authoring 20k lines of API spec; created a seed database for enabling a reproducible development and testing environment now adopted by the entire engineering org; Developed a novel visualization tool to aid monolith/microservice transition
- Performance optimizations: Reduced GPRC server latency and timeouts by half by implementing multiprocessing, fast JSON serialization, and db indexing strategies; minimized celery worker maintenance by removing memory leaks

PingAn AI Institute, New York, NY, Machine Learning Engineer, 2018–2021

- Language modeling for chatbots, question answering in documents with Professor Harry Wang from University of Delaware. Founding engineer of the New York office. Lead research, engineering and hiring.
- Led the work of a Conversational AI platform to build chatbots (alternative to Google DialogFlow, IBM Watson and Rasa)
- Research on knowledge-grounded conversational modeling at DSTC7, AAAI 2019 (best in human evaluation)
- Research on domain transfer learning of task oriented chatbots. Schema-guided Dialog State Tracking at DSTC8, AAAI 2020.
- Developed machine reading comprehension models (top results on SQuAD leaderboard from PingAn team)

PwC, Hyderabad, Software Engineer, 2015–2016

- Worked on transition to Google Cloud Platform. Developed novel proof of concepts. Lead a team in building web application to manage, monitor, and audit Google Cloud resources used across the organization. Used C#, VBA, Python, Flask, JavaScript.

Internships

Advanced Digital Sciences Center, UIUC, Research Assistant, Summer 2017

- Neural network distillation of time series forecasting models in Python, PyTorch
- https://iarcs.illinois.edu/news/researchers-optimize-air-conditioning-units-max-efficiency
- https://publish.illinois.edu/chiller-modelling-and-optimization/research/

Indian Institute of Science, Bangalore, Research Assistant, Winter 2013

- Worked on building Face Recognition and Detection models at Computational Intelligence Lab, Department of Aerospace Engineering

Fraunhofer Institute for Communication Systems ESK, Munich, Research Intern, Summer 2013

- Worked in Automotive Connectivity Division. Developed a road traffic simulation tool that allowed predicting optimal advisory speed for vehicles based on stop light timings and traffic

Google's Summer of Code, Summer 2013

- Developed features for SciPy Central website in Python, Django and Postgres
- Became a core contributor and maintained the production website
- https://github.com/scipy/scipycentral

Teaching Experience

TA, CISC 106 Computer Science for Scientist and Engineers, University of Delaware, Fall 2017

- Substitute instructor for three sessions
- Organized bi-weekly lab, graded assignments and exams, office hours

Awards and Honors

- MIT Delta V Accelerator Fellowship, 2023
- Best system design at DSTC7, AAAI 2019
- Google Summer of Code, 2013
- First rank, JNTU Hyderabad, 2013

Related Coursework

- Improving your Product Sense by Shreyas Doshi, Maven, Winter 2025
- 14.740 Foundations of Development Policy at MIT Edx Micromasters program, Fall 2021
- 14.100 Microeconomics at MIT Edx Micromasters program, Fall 2021
- Graduate: Statistical Research Methods, Intro to Machine Learning, Intro to Computer Vision, Intro to Artificial Intelligence, Design and Analysis of Algorithms
- Undergraduate: Digital Image Processing, Digital Signal Processing, Mathematical Methods for Engineers

Activities

- Open source contributions to Cohere for AI and Eleuther AI, 2025
- Reviewer, Manning Publications, 2025
- Reviewer, U.S. NSF, 2025
- Reviewer, ACL, 2025
- Reviewer at Dialog System Technology Conference DSTC7, AAAI 2018