Kausik Lakkaraju

Email : kausik0123456789@gmail.com
Linkedin: https://www.linkedin.com/in/kausik-lakkaraju-8a29833a
Mobile : +1-803-238-0960
Github: https://github.com/kausik-l
Webpage: https://kausik-l.github.io/

Education

- **AI Institute, University of South Carolina** — Columbia, SC (USA)
  Ph. D. - Computer Science & Engineering; GPA: 3.81 — *Jan 2021 - Present*
  Courses: *Trusted AI, Neural Networks, Bayesian Networks and Decision Graphs, Theory of Computation, Foundation of Analysis 1 (MATH)*
- **University of South Carolina** — Columbia, SC (USA)
  MS - Computer Science & Engineering; GPA: 3.87 — *Aug 2019 - May 2021*
  Courses: *Artificial Intelligence, Data Mining and Warehousing, Natural Language Processing, Deep Reinforcement Learning, Robotics, Big Data Analytics, Database System Design, Analysis of Algorithms, Computer Architecture*
- **Osmania University** — Hyderabad, Telangana (India)
  B.E. - Electronics and Communication & Engineering; Percentage: 76% — *Aug 2015 - May 2019*
  Major Project: *Multiple color line follower robot with obstacle detection*

Skills Summary

- **Languages**: Python, JavaScript, C++, C, SQL, R.
- **Areas of Research**: Trusted AI, Causal Models, LLMs, Multi-modal Decision Support, Chatbot Development, Rating of AI Systems, AI Fairness

Research Experience

- **University of South Carolina** — Columbia, SC, USA
  Graduate Research Assistant - Dr. Biplav Srivastava — *Jan 2021 - Present*
  - **Rating AI Systems for Bias**:
    - *Rating of AI Systems through a Causal Lens:* This is my Ph.D. thesis topic. As most of the AI systems are known for giving biased or uncertain outputs, we proposed a causality-based method to rate AI systems for bias. We assessed text-based sentiment analysis systems so far and built a tool using the Django framework that aids users in rating their own AI systems. We are currently working on evaluating multi-modal systems.
  - **Chatbot Testing**:
    - *LLM Testing:* We assessed LLM-based chatbots, ChatGPT and Bard, for efficacy and fairness. Currently, we are working on building a principled approach to evaluate LLMs for bias.
    - *SafeChat Framework:* The SafeChat framework we built enables users to build safe and trustworthy chatbots. We performed experiments and showed that chatbots built using SafeChat outperform LLM-based ChatGPT and Bard in the personal finance domain. We are currently adding more trust services to the framework including the rating service.
- **University of South Carolina** — Columbia, SC, USA
  Graduate Research Assistant - Dr. Dezhi Wu — *May 2020 - May 2021*
  - **Full-stack Development**: Involved understanding a clustering algorithm created by a Professor from Physics department at our university. We used Django REST framework to build the back-end. I was also part of front-end development. We used ReactJS for the UI.
  - **Front-end Development**: Involved in building a website for a scientific experiment that Dr. Wu has conducted with some other research scientists. We used ReactJS to build the front-end.

Selected Publications

Research Papers

- Kausik Lakkaraju, Aniket Gupta, Biplav Srivastava, Marco Valtorta and Dezhi Wu. The Effect of Human v/s Synthetic Test Data and Round-tripping on Assessment of Sentiment Analysis Systems for Bias. IEEE TPS 2023.
- Lakkaraju, K., Jones, S. E., Vuruma, S. K. R., Pallagani, V., Muppasani, B., & Srivastava, B. LLMs for Financial Advisement: A Fairness and Efficacy Study in Personal Decision Making. ICAIF 2023.
- Lakkaraju, K., 2022. Why is my System Biased?: Rating of AI Systems through a Causal Lens. In Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22). Association for Computing Machinery, New York, NY, USA, 902. https://doi.org/10.1145/3514094.3539556

- Lakkaraju, K., Hassan, T., Khandelwal, V., Singh, P., Bradley, C., Shah, R., Agostinelli, F., Srivastava, B., & Wu, D. (2022). *ALLURE: A Multi-Modal Guided Environment for Helping Children Learn to Solve a Rubik's Cube with Automatic Solving and Interactive Explanations.* Proceedings of the AAAI Conference on Artificial Intelligence, 36(11), 13185-13187. https://ojs.aaai.org/index.php/AAAI/article/view/21722

- Lakkaraju, K., Palaiya, V., Paladi, S.T., Appajigowda, C., Srivastava, B., Johri, L. (2022). Data-Based Insights for the Masses: Scaling Natural Language Querying to Middleware Data. In: , et al. Database Systems for Advanced Applications. DASFAA 2022. Lecture Notes in Computer Science, vol 13247. Springer, Cham. $https://doi.org/10.1007/978-3-031-00129-1_49$

- Biplav Srivastava, Kausik Lakkaraju, Mariana Bernagozzi, Marco Valtorta, Advances in Automatically Rating the Trustworthiness of Text Processing Services, AAAI Spring Symposium, San Francisco, 2023

Patents

- Srivastava, Biplav, Kausik Lakkaraju, Revathy Venkataramanan, Vishal Pallagani, Vedant Khandelwal, and Hong Yung Yip. "Robust useful and general task-oriented virtual assistants." U.S. Patent Application 17/714,508, filed November 10, 2022.

Academic & Personal Projects

- **YOLOR Vs. YOLOv6 Face-off: A Comparison of SOTA Object Recognition Models (Dec '22)**: We deep-dived into the working of YOLOR and YOLOv6 object recognition models, which were released in 2022, and evaluated them on specific tasks to compare their performance. We tested the original models and also tested them after fine-tuning them on a different blood cells dataset which classifies different blood cells. **Keywords**: Computer Vision, Object Recognition.
- **Plant Disease Explanation (Dec '21)**: I trained neural network models like CNN and InceptionV3 (using transfer learning) on two well-known datasets, PlantDoc and PlantVillage to predict the plant name and plant disease (if it is diseased). In addition to this, I used LIME explainer which would provide the explanation for why the system had to take a particular decision for a certain sample. **Keywords**: Trusted AI, XAI, Deep Learning.
- **Explainable Pet Class Predicting System (Dec '21)**: Used the 'Pawpularity' dataset from Kaggle to train neural network models to predict a 'Pawpularity score' based on how cute a particular image of the pet is. I added the explainability module to the project so that one would know what features contributed to the outcome given by the black-box neural network model. **Keywords**: Trusted AI, XAI, Deep Learning.
- **Indian Sign Language Recognition using OpenCV (Dec '20)**: My project allows one to create an Indian Sign Language (ISL) dataset on their own in a few minutes. I trained a CNN model on one such dataset created by me. It was able to achieve a test accuracy of more than 98%. **Keywords**: Computer Vision, Deep Learning, Image Processing.
- **Deep Reinforcement Learning Based Chatbot (Dec '20)**: I trained a Seq2Seq model using the 'Self-Critical Sequence Training (SCST)' method on movie dialogs dataset to create a chatbot (for chitchat). **Keywords**: Deep Reinforcement Learning, Chatbot.
- **Covid-19 Tracker (July '22)**: Built a UI using ReactJS for tracking worldwide and country-wise covid cases and deployed it using firebase. (click here). **Keywords**: Front-end, ReactJS.

Certifications & Prizes

- I was granted a student travel award of $ 1000 to attend the IEEE TPS 2023 conference.

- I was granted the best graduate student poster presentation award at a university-wide event that was held at the University of South Carolina for a poster I presented on my Ph.D. dissertation topic, 'Rating of AI Systems through a Causal Lens'.

- Presented a student track paper (on my Ph.D. dissertation topic) at the AIES 2022 conference which was held at the University of Oxford (UK).

- Secured first prize in ITT (Innovation Think Tank) certification program for two consecutive years.

- **Certifications:** Building Basic GANs (Coursera), Python for Data science (IBM), Deep Learning A-Z (Udemy), Machine Learning (A-Z)(Udemy), Django Full Stack Web Development (Udemy).