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

- **Artificial Intelligence (AI) Institute, University of South Carolina** — Columbia, SC (USA)
  Doctor of Philosophy (Ph.D.) in Computer Science & Engineering; GPA: 3.81 — *December 2025 (Expected)*
  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)
  Master of Science in Computer Science & Engineering; GPA: 3.87 — *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)
  Bachelor of Engineering in Electronics and Communication & Engineering; Percentage: 76% — *May 2019*
  Major Project: *Multiple color line follower robot with obstacle detection*

Skills Summary

- **Languages**: Python, JavaScript, C++, C, MySQL, R.
- **Frameworks**: Tensorflow, PyTorch, Pandas, Matplotlib, Numpy, Keras, Tableau, Power BI, Hadoop.
- **Areas of Research**: Causal Inference, Trusted AI, Foundation Models, AI Fairness, Time-Series Forecasting, Analytics, Statistics, Generative AI, Multi-modal Decision Support, Chatbot Development, Rating of AI Systems, Explainable AI (XAI), AI in Finance, Machine Learning, Deep Learning, Image Segmentation

Relevant Research Experience

- **Mayo Clinic** — Rochester, MN, USA
  Data Science Intern - Dr. Hamid R. Tizhoosh — *Sep 2024 - Jan 2025*
  - **Segmentation of CT scans**: I worked on segmenting CT scans using foundation model based segmentation models like SAM2.
  - **Evaluating patching algorithms**: Patching algorithms are designed for the efficient search and retrieval of digital histopathology data. I evaluated these algorithms using foundation models as the backbone, testing their performance across various settings.

- **University of South Carolina** — Columbia, SC, USA
  Graduate Research Assistant - Dr. Biplav Srivastava — *Jan 2021 - Present*
  - **Rating AI Systems for Robustness**:
    - *Rating of AI Systems through a Causal Lens (Ph.D. thesis topic.):* As most AI systems are known for giving biased or uncertain outputs, we proposed a causality-based method to rate AI systems for robustness under different settings. I worked on assessing text-based sentiment analysis systems, multi-modal and foundation model based time-series forecasting models so far and built a tool called ARC that aids users in rating AI systems. I am currently exploring how existing XAI methods can complement and improve our rating approach.
  - **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 chatbots in the personal finance domain. I am currently adding more trust services to the framework including the rating service.

Selected Publications

Research Papers

- Lakkaraju, K., Valluru, S. L., & Srivastava, B. (2025). Holistic Explainable AI (H-XAI): Extending Transparency Beyond Developers in AI-Driven Decision Making. arXiv preprint arXiv:2508.05792.
- Lakkaraju, Kausik, et al. "On Creating a Causally Grounded Usable Rating Method for Assessing the Robustness of Foundation Models Supporting Time Series." arXiv preprint arXiv:2502.12226 (2025).
- Lakkaraju, Kausik, et al. "Rating Multi-Modal Time-Series Forecasting Models (MM-TSFM) for Robustness Through a Causal Lens." arXiv preprint arXiv:2406.12908 (2024).

- K. Lakkaraju, B. Srivastava and M. Valtorta, "Rating Sentiment Analysis Systems for Bias Through a Causal Lens," in IEEE Transactions on Technology and Society, doi: 10.1109/TTS.2024.3375519.
- K. Lakkaraju, A. Gupta, B. Srivastava, M. Valtorta and D. Wu, "The Effect of Human v/s Synthetic Test Data and Round-Tripping on Assessment of Sentiment Analysis Systems for Bias," in 2023 5th IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), Atlanta, GA, USA, 2023 pp. 380-389. doi: 10.1109/TPS-ISA58951.2023.00053
- Kausik Lakkaraju, Sara E Jones, Sai Krishna Revanth Vuruma, Vishal Pallagani, Bharath C Muppasani, and Biplav Srivastava. 2023. LLMs for Financial Advisement: A Fairness and Efficacy Study in Personal Decision Making. In Proceedings of the Fourth ACM International Conference on AI in Finance (ICAIF '23). Association for Computing Machinery, New York, NY, USA, 100–107. https://doi.org/10.1145/3604237.3626867
- 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. $https://doi.org/10.1007/978-3-031-00129-1_49$

Patents

- Srivastava, Biplav, Vishal Pallagani, Revathy Chandrasekaran Venka, Vedant Khandelwal, and Kausik Lakkaraju. "Multimodal retrieval and execution monitoring using rich recipe representation." U.S. Patent 12,332,873, issued June 17, 2025.
- 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 12,067,983, issued August 20, 2024.
- Srivastava, Biplav, Kausik Lakkaraju, and Marco Valtorta. "Assigning trust rating to ai services using causal impact analysis." U.S. Patent Application 18/448,369, filed February 22, 2024.

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.
- **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. LIME was used to provide an explanation for the predictions made by the models. **Keywords**: Trusted AI, XAI, Deep Learning.
- **Deep Reinforcement Learning Based Chatbot (Dec '20)**: 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.

Certifications & Prizes

- Invited to present my research at ACM AIES 2022, ACM FAccT 2025, and IJCAI 2025 Doctoral Consortiums with full travel support.
- Conducted a tutorial on the topic 'Evaluating and Rating AI Systems for Trust and its Applications to Finance' at the ACM ICAIF 2024 conference.
- Awarded $ 1,000 NSF student travel grant to attend the IEEE TPS 2023 conference.
- Recipient of the best graduate student poster presentation award at a university-wide event that was held at the University of South Carolina in April 2023 for a poster presented on the topic, 'Rating of AI Systems through a Causal Lens'.
- Presented a student track paper (on 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) conducted by *Siemens Healthineers* certification program for two consecutive years.