KAUSTAV BHATTACHARJEE

Data Scientist | AI and Analytics Research

+1 (347) 601 7601 kaustavbhatt94@gmail.com
https://www.linkedin.com/in/kaustav-bhattacharjee/

SUMMARY

Results-oriented Data Scientist with a strong research background in trustworthy AI, specializing in explainable AI (XAI) and privacy-preserving data visualization. Proven ability to develop and deploy innovative, data-driven solutions that enhance transparency, interpretability, and fairness in AI systems. Research published in leading journals and conferences, including Springer Nature, Computer Graphics Forum, VizSec, USEC, HILDA, and ISGT.

EXPERIENCE

PhD Intern, Visual Analytics for Explainable AI

Pacific Northwest National Laboratory (PNNL)

07/2022 - 05/2024 (~2 years)

- Developed an interactive dashboard to visualize and interpret the impact of key weather parameters on a deep learning model, improving stakeholder understanding and facilitating data-driven decision-making.
- Designed a framework for evaluating model performance under varying conditions, ensuring robustness and reliability for real-world grid applications.
- Built a data-driven analytics system for comparing multiple load forecasting models, facilitating model selection and pattern discovery.

Doctoral Researcher

NJIT's Intelligible Information Visualization (NiiV) Lab

08/2019 - Present (~5 years)

- Developed a visual analytics interface to enable incremental recourse planning utilizing ML model explanation methods like SHAP.
- Implemented a privacy-preserving pipeline utilizing a Flask API for efficient triage and disclosure risk analysis of open datasets.
- Created a React and D3.js interface to showcase interpretability of machine learning models such as Random Forest and LambdaMART via sensitivity analysis and feature engineering.
- Utilized Natural Language Processing (NLP) to design a scoring system for assessing dataset utility.
- Developed an interactive dashboard for energy consumption analysis across diverse conditions, leveraging data extraction from relevant datasets.
- Enhanced a visual analytics platform to uncover correlations among climate science metrics, models, and variables.

Associate, Cyber Security

PricewaterhouseCoopers Private Limited (PwC)

08/2016 - 08/2019 (~3 years)

- Applied data science techniques to Managed Security Services (MSS) projects within banking, pharmaceuticals, and mining sectors, focusing on anomaly detection and threat analysis across diverse systems including firewalls, antivirus, and servers.
- Served as a Security Information and Event Management (SIEM) Administrator, establishing incident monitoring lifecycles with RSA Archer SecOps module and configuring a malware analysis lab.
- Managed the proprietary Global Threat Intelligence Platform and implemented the Malware Information Sharing Platform (MISP).
- Developed and deployed visualization dashboards for client management using SQL for data processing and various Web Technologies for front-end implementation.

Intern

Analyse Technology Solutions

06/2015 - 07/2015 (~ 2 months)

- Completed an internship project on analyzing Twitter datasets, including hashtag analysis and sentiment analysis, using Map Reduce (Java).

EDUCATION

Ph.D. in Data Science

New Jersey Institute of Technology

08/2019 - Present *(Graduating soon!)*

GPA: **3.95**/4.0

B.Tech in Information Technology

Institute of Engineering & Management

07/2012 - 08/2016

GPA: **9.13**/10

SKILLS

Data Science & Machine Learning

- *Programming Languages:* Python (libraries: pandas, numpy, scikit-learn, NLTK, TextBlob, Flair), R, C++, Java
- *Machine Learning & AI:* NLP, Text Analysis, Feature Engineering, Classification, Regression, Clustering, Explainable AI, Reinforcement Learning, Recurrent Neural Networks (RNNs), PyTorch, Tensorflow
- *Data Pipelines & Infrastructure:* HuggingFace Pipelines, Docker, Virtual Machines

Data Analysis & Visualization

- *Data Exploration & Wrangling:* Jupyter Notebooks, A/B Testing
- *Data Visualization:* Tableau, D3.js, Plotly, matplotlib, Dash
- *Statistical Analysis:* Descriptive Statistics, Time Series Analysis, Regression Analysis

Web & API Development

- *Front-End Development:* HTML, CSS, Javascript, React.js, Node.js, Material UI (MUI)
- *Back-End Development:* Flask, REST API
- *Development Tools:* Git, Postman, Heroku

Databases & Cloud Platforms

- *Databases:* MySQL, Postgres, MongoDB, SQLAlchemy
- *Cloud Platforms:* Amazon Web Services (AWS), Google Cloud

Other Skills

- *Web Mining & Automation:* Selenium
- *Big Data Technologies:* Hadoop, MapReduce, Hive
- *Malware Detection:* x64Dbg, Regshot, CEF Explorer
- *Security Monitoring & Incident Management:* ArcSight, Splunk, RSA Archer

PUBLICATIONS

Forte: An Interactive Visual Analytic Tool for Trust-Augmented Net Load Forecasting

IEEE Innovative Smart Grid Technologies North America (ISGT NA 2024)

Kaustav Bhattacharjee, Soumya Kundu, Indrasis Chakraborty, Aritra Dasgupta

https://doi.org/10.1109/ISGT59692.2024.10454191

Developed an interactive visual analytics tool (Forte) to enhance trust in AI-powered net load forecasting. Forte empowers human analysts to explore the relationships between net load predictions and various input factors, like weather data. It also enables evaluation of model performance under diverse conditions, fostering informed decision-making in the energy sector.

TRIVEA: Transparent Ranking Interpretation using Visual Explanation of Black-Box Algorithmic Rankers

The Visual Computer, Springer Nature Journal

Jun Yuan, Kaustav Bhattacharjee, Akm Islam, Aritra Dasgupta

https://doi.org/10.1007/s00371-023-03055-x

Designed TRIVEA, an interactive visualization system leveraging Explainable AI (XAI) techniques like LIME and SHAP to interpret black-box ranking models (e.g., LambdaMART, RankBoost, SVM, RankNet) used in machine learning. TRIVEA empowers users to understand how different factors influence model rankings, assess agreement between models, and thereby facilitate transparent and trustworthy decision-making.

VALUE: Visual Analytics driven Linked data Utility Evaluation

Workshop on Human-In-the-Loop Data Analytics (HILDA)

Kaustav Bhattacharjee, Aritra Dasgupta

https://dl.acm.org/doi/10.1145/3597465.3605225

Developed VALUE, a human-in-the-loop visual analytics workflow for evaluating the utility of joining open datasets. VALUE empowers researchers to identify and prioritize high-utility data linkages, fostering data-driven insights through a transparent and interactive exploration process.

Power to the Data Defenders: Human-Centered Disclosure Risk Calibration of Open Data

2023 Symposium on Usable Security and Privacy (USEC)

Kaustav Bhattacharjee, Aritra Dasgupta

https://www.ndss-symposium.org/ndss-paper/auto-draft-352/

Investigated human-centered approaches to mitigate disclosure risks and potential leakage of sensitive information in open data.

PRIVEE: A Visual Analytic Workflow for Proactive Privacy Risk Inspection of Open Data

2022 IEEE Symposium on Visualization for Cyber Security (VizSec)

Kaustav Bhattacharjee, Akm Islam, Jaideep Vaidya, Aritra Dasgupta

https://doi.org/10.1109/VizSec56996.2022.9941431

Designed PRIVEE, a visual analytics workflow empowering data defenders to proactively identify and mitigate privacy risks in open data through interactive visualizations.

Privacy-Preserving Data Visualization: Reflections on the State of the Art and Research Opportunities

Computer Graphics Forum Journal

Kaustav Bhattacharjee, Min Chen, Aritra Dasgupta

https://doi.org/10.1111/cgf.14032

Focused on the systemic analysis of the approaches, methods and techniques used for handling data privacy in visualization. Identified the gaps and the future research opportunities in this area.

ACHIEVEMENTS

Won the Best Poster prize at the NYC Privacy Day Spring 2024

Awarded NSF Travel Grant to attend Symposium on Usable Privacy and Security (SOUPS 2023)

Won the Best Presentation Award at NJIT GSA Research Day 2023

Delivered a lightning talk at the Symposium of Usable Privacy and Security (SOUPS) 2023

Awarded Gold medal from Tata Consultancy Services (TCS) for overall performance in both student and co-curricular activity in the year 2016

Received A.P.J Abdul Kalam Techno-Wiz award for excellent contribution to the annual college technical fest Innovación 2015