Ananya Kamboj

(857)-415-1503 | anaxkam@gmail.com | linkedin.com/in/ananyakamboj | github.com/ananyakamboj

EDUCATION

University of Massachusetts Amherst — Jan 2022 – Feb 2024
Master of Sciences in Computer Science, Specialization in Data Science — *Amherst, MA*

Guru Gobind Singh Indraprastha University — Aug 2017 – Jun 2021
Bachelor of Science in Computer Science — *New Delhi, India*

TECHNICAL SKILLS

Languages: Python, SQL (Postgres), R, C/C++, JavaScript, HTML/CSS
Developer Tools: Git, Tableau, Power BI, AWS S3, AWS RedShift, Google Cloud Platform, Microsoft Azure
Data Science Techniques: Machine Learning, Deep Learning, NLP, Feature Engineering, Data Pipelines, A/B Testing, End-to-End ML
Frameworks/Libraries: Pandas, NumPy, Matplotlib, Scikit-learn, Keras, Tensorflow, PyTorch, PySpark

EXPERIENCE

Data Science Intern — Jun 2023 – Aug 2023
Panera Bread ⤴ — *St.Louis, MO*

- Analyzed 5TB+ order and survey data for 2000 cafes using BigQuery, resulting in a weekly savings of $25,000 by resolving accuracy issues.
- Contributed to the development of CV-based algorithm for ingredient verification, leading to a notable 30% increase in order accuracy. Won company hackathon among 150 participants.
- Implemented data science and AI/ML techniques to prevent order accuracy issues, optimizing timing and predicting potential issues.

Data Science Student Researcher — Feb 2023 – Jun 2023
Oracle Labs ⤴ — *Amherst, MA*

- Engineered Python tool using ML Privacy and TensorFlow Privacy, reducing researchers' time by 40% in evaluating privacy attacks.
- Researched privacy in machine learning models, assessing attacks' effectiveness across different datasets and tools.
- Analyzed data to recognize overfitting's impact on privacy, prompting timely mitigation measures.

Decision Analytics Associate — May 2021 – Jan 2022
ZS Associates ⤴ — *Pune, India*

- Built customer segments by analyzing real-world oncology data (RWD) extracted using SQL from AWS S3 and Redshift for improved cancer treatment.
- Achieved 5% increase in product uptake through quantitative analysis of patient journeys using statistical models in Python including regression.
- Communicated analytical findings to stakeholders and clients to drive informed decisions and commercial success.

Data Science Intern — Aug 2020 – Nov 2020
Aarna Analytics ⤴ — *Kerela, India*

- Conducted exploratory data analysis on diverse datasets, extracting insights for informed decision-making.
- Implemented linear regression for stock price prediction across multiple client datasets, enhancing financial forecasting accuracy.
- Experimented with time series forecasting and XGBoost algorithms resulting in increased prediction accuracy by 20% compared to baseline models.

PROJECTS

Prompting Paradox in LLMs | *Python, TensorFlow, NLP*

- Explored prompt impact on in-context learning, achieving 6.95-point F1 score improvement with MetaICL over GPT-2.
- Identified label space significance with a 5% performance decrease using random labels.

PRIMIA: Privacy through MIA | *Python, TensorFlow, Machine Learning, Git*

- Investigated factors impacting ML model vulnerability to membership inference attacks.
- Developed Python tool with ML Privacy/TensorFlow Privacy for attack evaluation over multiple datasets, boosting efficiency by 40%.