

SHREEYA GAD

gadshreeya@gmail.com | [Personal Website](#) | [LinkedIn](#)

Senior ML + Systems Engineer experienced in designing & scaling safety-critical machine learning systems, high-volume data infrastructure, and production ML pipelines.

EDUCATION

Cornell University (Aug 2017 – Dec 2021)

Master of Engineering, Computer Science

Bachelor of Arts, Computer Science | Business Minor for Engineers

WORK EXPERIENCE

Meta, Senior Software Engineer, Machine Learning (Mar 2025 - current)

- Led a cross-functional team of 10+ engineers to design and deploy next-generation content integrity classifiers in Threads' recommendation and ranking stack, improving platform safety for 400M+ users.
- Built CPMM & KNN safety classification models and integrated them into real-time inference systems, reducing harmful-content propagation and shrinking the blast radius of integrity attacks.
- Developed recurrent training proposal for Threads integrity ecosystem to reduce model age and improve recall.

Meta, Software Engineer, Systems / Infrastructure (Apr 2022 - Mar 2025)

- Scaled Privacy-Aware A/B Test platform for Ads Ranking models from 10 to 3000+ experiments by redesigning pipelines, automating CI/CD, and optimizing SQL/DB clusters to improve reliability, availability and adherence to privacy compliance.
- Led cross-org effort to reduce e2e latency of the data platform by >50% by overhauling metric computation, removing infra bottlenecks, and modernizing scheduling logic—unlocking significantly faster model iteration & saving \$6.5M in compute.

Meta, Software Engineer Intern, Machine Learning (June 2021 - Aug 2021)

- Identified best feature selection method for ads ranking model to mitigate data loss from Apple privacy updates, yielding 11% improvement in offline gains over baseline, through model exploration and experimentation.
- Automated multiple feature selection methods into model building automation pipeline for platform-wide usage.

Gusto, Software Engineer Intern, Full Stack (June 2020 - Aug 2020)

- Shipped Expense Reporting feature to 1,400+ customers in production in 12 weeks w/ Faster Payments Experience team.
- Used Ruby On Rails / React to build web application that intakes and processes reimbursement requests for small businesses.

Publicis Sapient, Software Developer Intern, Full Stack (June 2019 - Aug 2019)

- Shipped social listening React application to increase product visibility for Consumer Direct Marketing team in 12 weeks.
- Scraped hundreds of news articles to gather product sentiment (w/ BeautifulSoup, Scikit-learn) & displayed analytics.

PROJECTS

Mood Tracker, Master of Engineering, Cornell University

- Developed sentiment ML classifier that analyzes user statuses and tracks their mood over time, hosted via a React website.
- Trained RNN to 90% accuracy w/ PyTorch, incorporated online learning, and persisted model updates using AWS.

Anonymization Station, Master of Engineering, Cornell University

- Developed tool under Professor Kilian Weinberger that anonymizes faces in an image via pixel encryption.
- De-anonymizes faces when provided w/ the correct security key to protect the identity of minors on the internet.

Listen & Learn, Cornell University

- Built ML recommendation engine for podcasts using TF-IDF similarity between user input & episode descriptions.
- Scraped episode metadata (w/ BeautifulSoup, Spotify API) and uploaded TF-IDF vectors of episode descriptions to S3.

Additional Projects: [OOPLang](#) (custom OO interpreter in OCaml), [NATE-OA](#) (Wikipedia chatbot using NLP/IR methods), [Splash](#) (3rd place at Hoboken Climate Hackathon), [QuizMe](#) (Best Potential Hack at Hack NYU)

SKILLS

Infra: large-scale experimentation systems for ranking, distributed data pipelines, CI/CD automation

ML Systems: model iteration, content ranking, offline/online evaluation, feature engineering

Core: Python, SQL, JavaScript, React, Flask