

# Jacob Baum

New York, NY • jacoblbaum@gmail.com • (925) 457-1941 • [linkedin.com/in/JacobLundinBaum](https://www.linkedin.com/in/JacobLundinBaum)

---

## Education

New York University, New York, NY  
Bachelor of Arts in Computer Science

Dec 2022

Honors/Positions: Magna cum laude, Dean's List (GPA: 3.87/4.0), Executive Board Tech@NYU

---

## Professional Experience

Software/Machine Learning Engineer

Feb 2023 – Present

Capital One, New York, NY

- Developed Kubeflow components and tooling performing various machine learning tasks at large-scale—RFE, XGBoost, Optuna—on multiple datasets over 500 million rows and >5,000 features, to train essential fraud, credit, and other ML models
- Led implementation of GPU in Kubeflow components along with benchmarking resulting in ~7.6X speedup of component runtime
- Researched novel feature elimination methods and presented findings at internal conference. Found ~5X speedup with comparable model performance creating roadmap future enterprise tooling
- Tuned Dask clusters, optimizing performance and resource utilization for massive workflows
- Designed and implemented stress testing suite interacting with AWS S3, DynamoDB and Lambda

Assistant Research Engineer

Jan 2022 – Jun 2023

NYU Center for Social Media and Politics, New York, NY

- Programmatically collecting data from 94 news websites collecting ~74K articles, using technologies: BeautifulSoup and Selenium ran on Digital Ocean
- Manipulated collected datasets in python notebooks and performed comparative analysis for research focused on the spread of misinformation through news sources
- Worked with research lead to perform write up and documentation of our collection process

Software Engineer Intern

Jun 2022 – Aug 2022

Capital One, New York, NY

- Created front end using React integrating API calls into existing user interface in JavaScript
  - Integrated document extraction software into associate workflow saving estimated 3K hours annually
- 

## Skills

- |                |          |                    |               |
|----------------|----------|--------------------|---------------|
| • Scikit-learn | • AWS    | • Dask             | • Python      |
| • Kubernetes   | • Docker | • Machine Learning | • Open Source |
- 

## Certifications

- [AWS Certified Solutions Architect](#) - Amazon Web Services
  - [Unsupervised Learning, Recommenders, Reinforcement Learning](#) - DeepLearning.AI, Stanford Online
  - [Supervised Machine Learning: Regression and Classification](#) - DeepLearning.AI, Stanford Online
- 

## Personal Projects

Hark AI

Created an AI podcast service that creates daily podcasts with up to date news. The app uses Perplexity's completions API to generate updated podcast scripts which are then used to generate a podcast using ElevenLabs text to speech API. The app is hosted on Cloudflare [here](#). Uses Cloudflare's built database service R2 and scheduled workers to populate the DB daily with podcasts.