MATT KAYE

🌐 mrkaye97.github.io   ⬛ mrkaye97   @ mrkaye97@gmail.com   📞 (646) 853-5997

EXPERIENCE

Data Scientist
CollegeVine

📅 Sept 2020 – Present

- First member of CollegeVine's data team. Responsible for a variety of data projects, maintaining the codebase, integrating new tooling, and onboarding new team members.
- Promoted a culture of shipping data scientific products with software engineering best practices, including unit and integration testing and a CI/CD process.
- Devised and implemented CollegeVine's Chancing model, used by hundreds of thousands of students and families every year to understand their chances of admission at their favorite schools.
- Created a recommender system to suggest new colleges to students. School recommendations also powers affinity scoring, a major component of our college product and sales process driving hundreds of thousands of dollars in revenue.
- Built a sequential testing toolkit to enable our team to call A/B tests more quickly without sacrificing statistical rigor, often cutting A/B test runtimes by half or more.
- Built out our first data warehouse for analytics, helping teammates uncover insights into all aspects of our business more quickly and correctly.
- Build and optimized data pipelines moving billions of rows of data from raw sources into our data warehouse to be consumed by downstream users.
- Introduced new tooling (dbt, Airflow, MLFlow, AWS stack) and owned its adoption by the team and integration into the codebase.
- Owned the deployment and monitoring of machine learning models in production.

Baseball Operations Fellow
Baltimore Orioles

📅 Mar 2020 – Sept 2020

- Created a fully Bayesian, simulation-based projection system for MLB player performance.
- Implemented a hurdle regression framework to model free agent salaries.
- Devised a Markov Chain Monte Carlo approach to determining optimal shifts against opposing hitters.

PROJECTS & OPEN-SOURCE

- Contributor to MLFlow, an open-source platform for managing the machine learning lifecycle.
- Current author and maintainer of slackr, an R package for connecting R to Slack with 350k+ downloads.
- Contributor to recmetrics, a python library for evaluating recommender systems.
- Author and creator of lightMLFlow, a lightweight, opinionated R wrapper for the MLFlow REST API.
- Author and creator of fitbitr, an opinionated R wrapper for the Fitbit API.

EDUCATION

Bachelor of Arts
Economics, Mathematics
Carleton College

📅 Sept 2016 – Nov 2019

Choate Rosemary Hall

📅 Sept 2013 – June 2016

SKILLS

Programming:

Bash   Python   R   SQL

Frameworks, Software, and Tools:

AWS Tooling (Redshift, S3, Batch, etc.)

Airflow   CI/CD (Circle, GHA)

dbt   Docker   Git   Heroku

MLFlow   {{plumber}}   {{shiny}}

HOBBIES

Distance running

Learning Haskell

Low and slow cooking

Nature & architecture photography

Reading fantasy novels

Skiing chop and powder

Solo traveling + hostel hopping