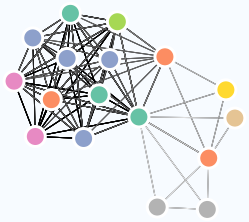
JOSHUA GOLDBERG

Data scientist proficient in statistics, machine learning, and software engineering. Proficient in python, R, SQL, and C++ using functional programming and object-oriented design. Examples of my work include machine learning models to optimize sales and marketing programs with an estimated impact of $1 million net revenue annually at a financial institution; models to detect risky behavior across millions of third-party sellers on amazon.com[1], and forecasting models to predict demand of 500,000+ products on Amazon's Marketplace. Outside of work, I enjoy distance running, reading, and developing/implementing algorithms[2] in C++.



INDUSTRY EXPERIENCE

Current | 2020

Data Scientist

Amazon — Seattle, WA

- Conceptualize and implement machine learning models for forecasting and supply-chain use-cases; own and maintain python production code in AWS / Amazon SageMaker that predicts customer demand of 500,000+ Amazon products worldwide
- Built, enhanced, and maintained natural-language machine learning models to detect fraud and abuse from third-party sellers' communication with customers on Amazon's marketplace; models detected and prevented thousands of fraud and abuse cases on a quarterly basis.
- Create software tools to monitor machine learning models in production

2020 | 2017

AVP, Lead Data Scientist

Nuveen — Chicago, IL

- Pioneered end-to-end (execution and experimental design) deep learning time series model for client onboarding; estimated impact of the model was $1 million net revenue annually that maximized client journey (improvement in client retention, client growth, etc.)
- Built recommendation engine for 150,000 clients in 50+ products
- Presented model/analysis to executive management; results included model adoption by 100+ sales people and a significant increase sales for clients treated by the model
- Conceptualized and created simulation engine that isolated, detected and measured the ROI impact of company sales events

2017 | 2014

Senior Equity Research Associate, Financial Services

Raymond James Financial, Inc. — Chicago, IL

- Built company and industry models using finance and statistical techniques, including regression and discounted cash flows (DCF)

View this CV online with links at https://goldbergdata.github.io/cv

CONTACT

josh.goldberg1@outlook.com

Twitter

Github

https://joshuagoldberg.name

LinkedIn

LANGUAGE SKILLS

- Python
- R
- SQL
- C++
- Bash
- HTML/CSS
- Javascript

EDUCATION

Current | 2021

STEM Continuing Education

Various Institutions

- I actively take STEM courses at different universities to enhance, revisit, or refine my technical skillset. Most of the courses are in computer science or mathematics.
- Harvard: Calculus 2 with Series and Differential Equations; Linear Algebra and Differential Equation
- University of Illinois Urbana-Champaign (UIUC): Calculus 1: First course in Calculus and Analytic Geometry
- Edmonds College: CS I, II, II; Courses in C/C++ covering Data structures & algorithms and object-oriented design and programming

2020

M.S. in Applied Data Science

University of Chicago — Chicago, IL

- Coursework in statistics, linear algebra, machine learning, and deep learning

2013

B.S. in Accounting and Finance

University of South Florida — Tampa, FL

SELECTED CODE REPOSITORIES

2021

Machine learning decision tree and data frame implementation in C++[3]

Github — Seattle, WA

- Authored with John Nguyen

2020

Generative adversarial network used to generate musical samples[4]

University of Chicago — Chicago, IL

- Capstone project and paper[5] authored with Terry Wang and Rima Mittal. Supervised by Yuri Balasanov[6]

In my free time, I enjoy working with friends, peers, and colleagues on algorithm designs/implementations. Recently, we built data frame and decision tree classes in C++.

TEACHING EXPERIENCE

Current | 2022

Python for Data Science

University of Chicago — Remote

- Instructor
- Topics include introductory and advanced topics in python: variables, logical operators, containers, loops, conditionals, comprehensions, functions, object oriented (basics), advanced data analysis and manipulation with numpy and pandas, model evaluation, parallel computation, and APIs

I am passionate about teaching and helping others. It brings me joy and satisfication to teach others new skills.

Current | 2020

Various courses

University of Chicago

Remote

- TA
- Intro statistics, machine learning, time series analysis

2022 | 2020

MastersTrack Statistics for Machine Learning and Machine Learning Courses

University of Chicago

Remote

- Instructor and TA
- Statistics course: topics include simple and multiple regression, logistic regression, hypothesis testing, variable transformations. Machine learning course: topics include a survey of machine learning algorithms: kNN, support vector machine, decision tree, random forest, boosted trees, and clustering algorithms

2021 | 2020

Data Understanding via SQL, Databases, and R

University of Chicago

Remote

- Instructor and TA
- Topics include introduction to databases, mySQL, and R