

Arnav Malhotra

703-973-8096 | arnavmalhotra@gmail.com | Brooklyn, NY

Data expert with 5+ years of experience leveraging advanced analytics and machine learning using Python, SQL, and data visualization tools across a diverse set of industries including real estate, healthcare, and health insurance. Proven track record of solving key business challenges by developing innovative solutions for many stakeholders—creating new metrics, data-driven features, and insights leading to increased efficiencies and customer retention.

EDUCATION

JOHNS HOPKINS

BS IN BIOMEDICAL ENGINEERING

May 2018 | Baltimore, MD

Conc. in Computational Biology

College of Engineering &

School of Medicine

INDUS INTL. SCHOOL

May 2014 | Bangalore, India

International Baccalaureate

SKILLS

LANGUAGES

- Python • SQL (Postgres & TSQL)
- SAS • R • \LaTeX

LIBRARIES

- pandas • shapely • GeoPandas
- scikit-learn • fastai
- matplotlib • folium • ggplot

SERVICES

- AWS RedShift & Sagemaker
- Snowflake • Redash • JIRA
- Looker • Tableau • PowerBI

PROJECTS

KAGGLE COMPETITIONS

Rainforest Species Audio, Bundesliga Highlights, iMet Collection

Built multi-classification deep learning models, in Python using fastai

CURRY IN A HURRY

Co-Founder

Started as a small-batch meal prep service specializing in Asian curries. Now popping up at wine bars across New York City, serving modern Indian-inspired cuisine (@curryinahurrynyc)

LINKS

Github:// [amalhot5](#)

LinkedIn:// [arnavmalhotra](#)

EXPERIENCE

DATA ANALYST | PERCHWELL

May 2023 - Present | New York, NY

- Joined as a founding member of Data team
- Overhauled Street Search by utilizing GIS analysis in Python, allowing users to search for listings based on which streets they were on, by finding out what street(s) each building was on and identify every intersection in NYC
- Led development on Brokerage Similarity Score, which is an algorithm (built with scikit-learn) that finds the most similar brokerage based on the many factors including locations, pricing, number, and performance of their listings
- Built prototypes in Python and defined logic for the more complex portions of the client and third party ingestion process; including finding off-market sales in public deeds records, creating & updating buildings/units and their geometries from public assessments and parcels data for four different markets (NYC, Cincinnati, Iowa, California)
- Performed triage for high priority client complaints, finding bugs, and working with engineers to implement fixes
- Built internal tools to automate manual processes, most notably creating and updating neighborhood geographies, quarterly reporting for brokerages, and many data quality fixes
- Worked with Engineering and Customer Service teams to create metrics around data quality and customer satisfaction, as well as dashboards to track those metrics and reveal trends in Looker and Redash
- Gave presentations to the whole company, executives, and clients demonstrating new features and prototypes
- Interviewed prospective hires/contractors and mentored intern

HEALTHCARE DATA ANALYST | CAREJOURNEY

May 2021 - Apr 2023 | Arlington, VA

- Led development on the Provider Cost Score, which uses Machine Learning to rank providers, by building a model in Python (using AWS Sagemaker) to find the expected cost of a medical episode while taking into account many risk factors related to the procedure, the provider, and their patient population
- Developed medical appropriateness measures and provider-to-provider referral metrics using commercial health insurance claims data in SQL along with algorithms to get around limitations of that data
- Supported the Platform team's new product launch by interpreting wireframes and providing the data and metrics
- Worked with Data Science team to implement multi-classification models to categorize doctors into sub-specialties
- Developed new quality measures for medical procedures, in SAS, using Medicare claims data

INFORMATICIST II | HUMANA INSURANCE COMPANY

Jun 2020 - April 2021 | Arlington, VA

INFORMATICIST I

Feb 2019 - Jun 2020 | Louisville, KY

- Worked on the Business Intelligence Advanced Analytics Team to create and monitor metrics for other departments in order to create benchmarks, aid in goal-setting, and show their impact on the company's OKRs
- Built predictive models with scikit-learn to discover what factors influenced our members' impression of Agent Knowledge
- Created dashboards in Tableau and PowerBI for real-time monitoring of metrics and trends by business partners
- Performed sentiment analysis and categorization of agent case comments, web chat transcripts, and free-text responses from surveys using Python

