KAT DESAI

khyatee.d@gmail.com • New York, NY

PROFESSIONAL SUMMARY

Seasoned Data Scientist with over 5 years of industry expertise in leveraging data and machine learning to drive business solutions. Recognized for keen attention to detail and customer obsession. Always learning more.

TECHNICAL STACK

LANGUAGES
Python, SQL, HTML, CSS

TOOLS
Pandas
NumPy
Sklearn
PyTorch
Matplotlib
PySpark
Github

AWARDS & CERTIFICATIONS

AWS Big Data Specialty
Fall 2019

AWS Solutions Architect
Fall 2019

Dean's List
Winter 2015, 2017

EMPLOYMENT

Data Scientist • Amazon Web Services • 2022–2024

Co-owner of ML model that predicts the likelihood of an AWS sales opportunity to launch

- scaled out the initial model from coverage of 3 AWS services, to 200+ services
- led R&D initiatives to improve model performance by 10%+ (absolute)
- developed a new evaluation metric to measure the revenue impact of the model's predictions
- performed ablation studies for 20+ new model features from multiple data sources, and evaluated feature efficacy using hypothesis testing
- optimized the model lifecycle, speeding up the end-to-end process by 3+ hours

Solutions Architect • Amazon Web Services • 2018–2022

- provided prescriptive guidance of the AWS platform, encompassing 200+ services, to thousands of customers via calls, whiteboarding, and product demonstrations
- led workshops for customers groups of 5-100+ at Re:Invent & other conferences
- authored official AWS technical content (ppts, workshops, documentation, etc)

Data Science Consultant • Math Action • 2021

Developed a highschool math curriculum module using the data science life cycle (data collection, EDA, etc,) as the underlying framework, piloted on 5+ student cohorts

EDUCATION

MS in Data Science (GPA 3.9) • CUNY Graduate Center • 2022–2024 (Expected)

PROJECTS

- Lung Cancer Segmentation – fine-tuned a binary segmentation model which identifies the tumor within a lung CT scan, yielding a dice score of 0.89
- Deep Reinforcement Learning – trained an AI agent using DQN & DDQN models to successfully complete a level of Super Mario Bros
- Transfer Learning with GPT-2 – fine-tuned an LLM to generate new song lyrics that emulate the musical style of the Beatles
- Album Review Sentiment Analysis – built an ML model to estimate the sentiment of a written music album review using NLP methods

Data Science Bootcamp • Flatiron School of NY • 2020–2021

PROJECTS

- Food Insecurity Projection – completed a final capstone project to predict & map future food insecurity rates in the US
- Music Label Song Recommender – built a recommendation system using cosine similarity, k-means clustering, and PCA to suggest similar songs from a music label
- Spotify Song Classifier – built a classifier to categorize songs into musical eras based on their attributes, using logistic regression, KNN, and ensemble methods

BS in Information Science • University of Michigan • 2014–2018

COURSEWORK

- Data Manipulation – web-scraping, API protocols, code optimization
- Data Exploration – data analysis & visualization
- Statistics Fundamentals – probability distributions, statistical testing
- Calculus 2 – Integrals, differential equations, parametric and vector functions, limits