Peter Yushun Li

peteryushunli@gmail.com | (630) 706- 0569 | linkedin.com/in/peteryushunli | medium.com/@pysli
github.com/peteryushunli | huggingface.co/peteryushunli | peteryushunli.github.io

SUMMARY

Experienced data scientist with a proven track record of designing Generative AI and Large Language Models (LLMs) tools for internal and external stakeholders. Deeply curious and passionate about exploring the boundaries of AI to change the world

Areas of Expertise:

Generative AI ▪ LLM ▪ RAG ▪ GPT ▪ BERT ▪ Llamaindex ▪ LangChain ▪ Pinecone ▪ Machine Learning ▪ Python ▪ Azure ▪ AWS

RELEVANT EXPERIENCE

KNOWN

Data Science Manager   May 2023 – Present (New York)

- Created and deployed a productionized LLM application hosted on Slack that automated the creation of ad campaigns and digital audiences, which reduced time spent on tasks by over 80%. New York Times article mention of the "Big LeBotski"
- Developed custom LLM-driven recommendation engines; one for a large book publisher using GPT-4 and text embeddings on the publisher catalogue, and another for Charity Navigator by fine-tuning a pre-trained BERT model
- Led a cross-functional team to design an internal Retrieval-Augmented Generation (RAG) application using Pinecone (vector database) and Llamaindex (retrieval and generation) to accelerate the creation of client deliverables
- Designed an NLP tool that analyzes the quality of survey responses to removes bots, reducing time spent by 70%

Data Scientist   Dec 2020 – April 2023 (New York)

- Managed the paid search account for a major cancer hospital client with successful year-over-year growth in budget (over $6M in 2023) and patient leads (11%), while supervising a team of 2 junior analysts
- Designed custom bid and budget optimization algorithms that outperformed Google's standard bid strategies in both conversion efficiency (by over 40%) and brand awareness (by over 25%)
- Lead a team of data scientists on a social listening project that pulls data from Twitter, Spotify, Reddit, and LinkedIn APIs to build audiences based on user affinities for social, programmatic, audio, video, and influencer marketing

ANNALECT

Data Scientist   Aug 2019 – Nov 2020 (Chicago)

- Leveraged user data from 1P data vendors to develop customer segmentation models (k-means clustering) and look-alike models (logistic regression) to build custom audiences that can be targeted for digital media
- Built an eCommerce topic modeling tool to understand shopper behaviors and product propensity

HILTI (GRAD SCHOOL PRACTICUM PROGRAM)

Team Lead / Data Analyst   Jan – Jul 2019 (San Francisco)

- Developed a standardized pricing algorithm and implemented a Tableau dashboard for the regional sales teams to calculate discount thresholds on incoming sales, which helped reduce unnecessary price leakage by over 15%

PRIOR EXPERIENCE

DELOITTE, CAPGEMINI & CAPCO   Apr 2014 – Jul 2018 (Chicago)

Management Consultant

- Worked on change management and org redesign projects that prepared companies for large-scale transformations, identified processes and roles facing change, and developed new role descriptions, trainings and org charts

EDUCATION & CERFTIFICATIONS

NORTHWESTERN UNIVERSITY, Evanston, IL
BA., Economics & Chinese, June 2013

UNIVERSITY OF CALIFORNIA, DAVIS – GRADUATE SCHOOL OF MANAGEMENT, San Francisco, CA
M.S., Data Science & Business Analytics, July 2019

AMAZON WEB SERVICES: Machine Learning Specialty, Dec. 2019 & Cloud Practitioner Certification, Dec. 2018
DEEPLEARNING.AI: Deep Learning Specialization, May 2021
TABLEAU: Desktop Certified Associate, June 2019