

CHENFENG LI

Chicago, IL 60637 | (872)-215-0270 | cfli@chenfengli.com | <https://ChenfengLi.com>

SUMMARY

Data Scientist / ML Engineer with 1-2 years of industry experience. MS in Statistics from UChicago, skilled in Python, R, SQL, and Excel. Certified AWS Cloud Practitioner, Azure Data Engineer, and Power BI Data Analyst. Expertise in Machine Learning, Deep Learning, and Statistical Analysis (GLM, Bayesian, Time Series). Proficient in cloud computing with AWS and Azure. Experienced in creating interactive dashboards and reports using Power BI.

SKILLS and CERTIFICATIONS

Programming: Python, R, SQL, SAS, DAX, C, C++

Technical Skills: Machine Learning (TensorFlow, PyTorch, sklearn), NLP and LLM, Databases, Data Processing (Excel, Pandas), Visualization (Matplotlib, Power BI), Cloud Services (AWS, Azure), Statistical Analysis, Algorithm

Language: English (Fluent), Mandarin Chinese (Native), Cantonese (Native)

Certifications: Azure Data Engineer Associate, AWS Certified Cloud Partitioner, Excel 2019 Associate, Power BI Data Analyst Associate, Google Advanced Data Analytics Certificate, Deep Learning Specialization

WORK & RELEVANT EXPERIENCE

Data Scientist | SynergisticIT, Fremont, CA, Remote Jun 2023 - Present

- Developed an Azure-based system with a data scientist partner for sentiment analysis on 16,000 post-sale reviews for an online clothing store on social media. Ingested data with Data Factory pipeline, cleaned data using SQL in Synapse Analytic, and implemented NLP techniques with a BERT-based model in Azure Databricks for sentiment analysis. Concluded a satisfactory rate of 96% and trend upward over time.
- Performed a machine learning model to predict total sales for each product and shop for a retail chain using 2.9 million daily historical sales data from 22,169 kinds of items and 59 shops using NLP and time series techniques, enhancing inventory optimization and driving data-driven decision-making.
- Implemented a hybrid cloud architecture that connects an on-premises data center to an AWS VPC using a Site-to-Site VPN with configuring multiple AWS services. Solved the challenge of seamless, secure, and scalable integration between two environments improving operational efficiency and data protection.
- Constructed two interactive Power BI dashboards for sales and transaction analysis of a sport corporation and a retail chain. Cleaned and transformed data with SQL and Power Query, created measures with DAX, and designed one-page dashboards with key metrics, enabling real-time sales insights and facilitating discount analysis, Conducted stakeholder meeting for interpretation and training.

Statistical Researcher and Consultant | UChicago, Chicago, IL Oct 2022 – May 2024

- Led an analysis to provide actionable insights on media bias related to politician facial expression by analyzing 842,232 images from 8 medias. Built an 81%-accuracy name recognition model with politician images from Wikimedia. Extracted and analyzed facial expression logits from influential politicians with dimensionality reduction and classification models.
- Evaluated GPT-2 model performance by experimenting with Gaussian, polynomial, and periodic kernels using few-shot learning. Gained hands-on experience with LLM development and model fine-tuning.
- Directed and mentored a team of three researchers, promoting collaboration and effective communication while facilitating knowledge sharing and problem-solving to successfully apply Distributionally Robust Optimization (DRO) on the MNIST dataset with pre-generated spurious correlations.
- Managed the project execution of three client-facing consulting projects within two months, working with a team of five consultants to analyze data issues, communicate insights, and provide tailored statistical recommendations through client meetings and comprehensive reports.

EDUCATION

MS in Statistics | University of Chicago Sep 2022 – Jun 2024

BSc in Mathematics | Chinese University of Hong Kong (CUHK) Sep 2018 – Jul 2022