

Youngmin Paul Cho

www.chopaul.com | choym92@gmail.com | (201) 870-3161 | linkedin.com/in/ympcho

Creative and intuitive problem solver with a passion for breaking down complex problems with end-to-end AI and ML systems

EXPERIENCE

HSAD America (LG Electronics)

Englewood Cliffs, New Jersey

Sr. Manager - Data Science

Jan 2022 - Present

- Led development and end-to-end productionization of a transformer-based behavioral intent model on LG.com events, to predict revisit and purchase propensity, driving a 10% increase in website revisit rate via targeted remarketing.
- Built an LLM-powered product review intelligence engine that ingests 50K+ reviews, generates abstractive summaries, and clusters themes using BERT embeddings + UMAP + HDBSCAN, surfacing high-signal UX issues and feature requests and contributing to a 14% increase in on-site engagement.
- Operationalized psychological segmentation at enterprise scale by mapping attitudinal survey data to a 20M+ customer CDP and building look-alike models and segment classifiers, then used these segments to drive a brand-voice GenAI content chatbot with agentic intent routing and promotion-aware responses, powering personalized campaigns that achieved a 2x lift in CTR and CVR.
- Designed and deployed a AI call-center analytics pipeline: automated speech-to-text transcription, then used a fine-tuned LLM for caller intent and root cause extraction, enabling faster triage, better routing, and significantly improving resolution time and agent utilization.
- Developed a Marketing Mix Modeling (MMM) tool by using Google's lightweight Bayesian framework to evaluate media channel effectiveness and optimize budget allocation, data-driven decisions that improved ROI by 15%.

Mars

Newark, New Jersey

Data Scientist

Sept 2019 – Jan 2022

- Optimized factory production processes by solving a convex optimization problem using cvxpy and Gurobi, balancing scheduling, raw material availability, and sales forecasts, resulting in a 15% improvement in factory throughput.
- Built a scalable time-series clustering algorithm using dynamic time warping and neural network model to group products by forecasted sales behavior, reducing forecast error by 22%, which improved the speed of inventory planning.

NYC Data Science Academy

NYC, New York

Teaching Assistant

Feb 2019 – Sept 2019

Specialization in Deep Learning and Big Data

- Designed convolutional neural network, computer vision model for classifying people's age, gender, and race
Partnered with MotionFlow, a smart AIoT ad company, delivering inference under 0.5 seconds with over 92% accuracy.

Samsung Electronics

Ridgefield Park, New Jersey

Supply Chain S&OP/Forecast Analyst

Jan. 2017 – Feb. 2019

- Reduced daily reschedule rates to under 2% YTD by automating demand forecasting in VBA with Holt-Winters and designing a delivery-date-driven prioritization procedure and improving forecast accuracy and speed in demand planning.

PERSONAL PROJECTS

- Developed an AI-driven equity research and valuation platform that ingests SEC filings and uses RAG to power interactive financial statement Q&A, comparable screening, and scenario modeling for fundamental investors.
- Developed a reinforcement learning-based drafting bot to uncover fantasy football strategies by learning from patterns in ADP, fantasy points, and other features, enabling data-driven recommendations for optimal team selection
- Built an automated audio briefing system for financial news using web scraping, APIs, and Whisper API to convert daily headlines into concise audio summaries optimized for passive listening.

EDUCATION AND CERTIFICATES

University of Texas at Austin - Bachelor of Science in Mathematical Statistics

Austin, TX

Element of Computing: Computer Science Certificate: Department of Computer Science

Dec 2016

Certificates in Machine Learning from University of Washington

Sept 2019

Certificates in Machine Learning from Stanford University

Nov 2019

SKILLS

Languages: Python • SQL • R

Frameworks & Tools: PyTorch • TensorFlow • Scikit-learn • Transformers • PySpark • FastAPI • Vector DBs • Airflow • MLOps (MLflow, Docker, Kubernetes) • Cloud: GCP (Vertex AI, BigQuery), Azure (Azure ML, Databricks)

Domain & Concepts: Data Science • Machine Learning • Deep Learning • AI • NLP • Retrieval-Augmented Generation (RAG)