

SHERMAN PANG

Singapore

m: +65 97545106

shermanpang@gmail.com

shermanpang.com

[LinkedIn](#)



PROFESSIONAL SUMMARY

Versatile data scientist with a strong track record in delivering high-impact solutions across fintech, consumer goods, and e-commerce. Proven expertise in statistical modeling, machine learning, and Generative AI, including GPT-4-based chatbot development with LangGraph, Langchain, and ChromaDB. Successfully led projects on marketing mix modeling, credit scoring, pricing optimization, and sentiment analysis—driving measurable gains in revenue, engagement, and operational efficiency. Skilled in deploying scalable ML and GenAI systems using GCP Vertex AI, Kubeflow, Docker, and Kubernetes. Adept at translating complex analytics into business value and fostering data-driven cultures.

CORE COMPETENCIES

- Segmentation/Clustering
- Time Series Forecasting
- Large Language Models
- Generative AI
- Optimization
- A/B Testing

PROFESSIONAL EXPERIENCE

L'ORÉAL, SINGAPORE

DATA SCIENTIST – CENTER OF EXCELLENCE, OCT 2022 TO PRESENT

- Developed a Python library for Marketing Mix Modelling (MMM) using Bayesian Regression to estimate and optimize return on advertising spend (ROAS) in Southeast Asia.
- Led MMM A/B testing in Thailand for Garnier, Lancôme, and La Roche-Posay, resulting in a 17% improvement in ROAS.
- Developed a Python package for Revenue Growth Management (RGM) using CatBoost to forecast revenue and analyze uplift, cannibalization, and pull-forward effects of pricing and promotional strategies for Garnier, L'Oréal Paris, and Maybelline in Southeast Asia.
- Utilized Google OR-Tools to optimize strategic pricing discounts on Shopee and Lazada across Thailand, Vietnam, and Indonesia, achieving a 15% average Gross Merchandise Value (GMV) uplift, validated through A/B testing.
- Deployed machine learning models using GCP Vertex AI and Kubeflow pipelines, streamlining the deployment process and automating end-to-end workflows for MMM and RGM across Southeast Asia.
- Led weekly data science lessons for L'Oréal management trainees, enhancing their skills in data analysis, machine learning, and statistical modeling, fostering a data-driven culture.

ADVANCE INTELLIGENCE GROUP – ATOME FINANCIAL, SINGAPORE

DATA SCIENTIST – GROWTH & OPERATIONS, MAY 2021 TO OCT 2022

- Led the development and modelling of the Pintar Score credit scoring and loyalty system for Kredit Pintar, achieving over 10M downloads and significantly enhancing user engagement.
- Optimized targeted marketing campaigns by segmenting 4M Kredit Pintar users, resulting in a 30% reduction in acquisition costs and a 15% increase in conversion rates.
- Executed A/B testing using causal inference methodologies to evaluate and improve campaign effectiveness, driving more precise marketing strategies.
- Developed a highly accurate treasury cash flow forecasting model (95% accuracy), enabling better financial planning and resource allocation.
- Deployed and managed scalable machine learning models using Kubernetes and Docker, and data pipelines with PySpark and Airflow, improving reliability and efficiency in production environments.

IMPACT CREDIT SOLUTIONS, SINGAPORE

DATA SCIENTIST – CREDIT RISK, SEPTEMBER 2019 TO MAY 2021

- Developed microloan credit risk models for Esta Dana Ventura, utilizing Logistic Regression and Gradient Boosting techniques, enhancing credit assessment accuracy and reducing default rates.
- Enhanced risk assessment for Indonesian SME loans at Bank Central Asia by implementing K-means clustering, improving loan approval processes and reducing risk.
- Designed and deployed an anomaly detection model for bank loan applications, decreasing manual review requirements and increasing operational efficiency.

- Mentored and trained junior data scientists in advanced data science techniques including Regression, Classification, and Unsupervised Learning, fostering a skilled and knowledgeable team.
- Designed a loan management system using Python, SQL, and AWS, managing 500K+ loans.
- Built a data pipeline for analyzing transaction data, supporting risk modeling with 1M+ rows.
- Established a Tableau dashboard to monitor a loan portfolio worth \$10M USD.
- Produced a 1-year forecast for a \$10M loan portfolio using delinquency transition matrices, simulating best- and worst-case scenarios to ensure risk tolerance.

EDUCATION AND CREDENTIALS

MASTER OF SCIENCE (M.Sc.) IN ANALYTICS, 2025

Georgia Institute of Technology

GPA: 4.0/4.0

HONOURS BACHELOR OF SCIENCE (B.Sc.) IN PHYSICS, MINOR IN FINANCIAL MATHEMATICS, 2019

National University of Singapore

CAP: 4.6/5.0 (First-Class Honours)

PROJECT EXPERIENCE

MAY 2025 TO JUNE 2025

COLLEGE COACH CHATBOT

<https://college-coach.dev/>

<https://github.com/shermanpch/college-chatbot>

- Architected an AI-driven "College Coach" chatbot using Chainlit + LangGraph, GPT-4o retrieval-augmented generation, and ChromaDB, delivering SAT-based Safety/Target/Reach recommendations with hybrid semantic-keyword search and natural-language filtering by US state, major, tuition, financial aid, athletics, and other preferences.
- Containerized the stack with Docker and one-click scripts; deployed behind an Nginx reverse proxy on Vercel with automatic SSL.

JUNE 2020 TO AUGUST 2020

SHOPEE SENTIMENT ANALYSIS

<https://github.com/shermanpch/shopee-sentiment-analysis>

- Completed product review sentiment classification with top-1 accuracy as the evaluation metric.
- Optimized the pretrained transformer BERT models using TensorFlow, Keras, PyTorch, and GPU/TPU.
- Ranked in the top 19% on a private leaderboard using an ensemble of fine-tuned transformer models.

JUNE 2020 TO AUGUST 2020

SHOPEE PRODUCT DETECTION

<https://github.com/shermanpch/shopee-product-detection>

- Performed e-commerce product image classification with top-1 accuracy as the evaluation metric.
- Enhanced the pretrained EfficientNetB4 and EfficientNetB5 using TensorFlow, Keras, and GPU.
- Achieved top 10% status on a private leaderboard using an ensemble of fine-tuned CNN models.

ADDITIONAL INFORMATION

Technical Proficiencies:

- Programming Languages: Python, R
- Machine Learning Libraries/Framework: Scikit-Learn, Keras, TensorFlow, PyTorch
- LLM & GenAI: LangGraph, Langchain, ChromaDB
- Data Analysis/Visualization: SQL, Tableau, Power BI
- Cloud & DevOps: AWS, GCP (Vertex AI, Kubeflow), Docker, Kubernetes, Airflow
- Big Data Tools: Hadoop, PySpark