

# HAOYU (VINCENT) TIAN

Cambridge, MA

(857) 320-8907 | vinnyt@mit.edu | [LinkedIn](#) | [Portfolio](#) | Australia Citizen (E3)

## EDUCATION

### MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Cambridge, MA

*Master of Business Analytics, Operations Research Center, GPA: 5.00/5.00*

Graduating: Aug 2024

- **Deep Learning Project:** Utilized Amazon product images to generate Google Gemini image captions, constructed 3K DALL-E 2 images and fine-tuned VGG19, Resnet50 and EffNet architectures, improved baseline error rate by 96% (*Python*)
- **Leadership: VP of Blockchain Club**, organized mixer event with MIT, Harvard Business School and Harvard Law School

### UNIVERSITY OF MELBOURNE

Melbourne, Australia

*Bachelor of Science in Mathematics and Statistics, GPA: 3.92/4.00*

2020

- **Awards:** Study Abroad Scholarship in Statistics - University College London (Spring 2019)

## TECHNICAL SKILLS

- **Languages:** Python (NumPy, Pandas, scikit-learn, PyTorch, TensorFlow, Keras), R, SQL, Julia/JuMP (Gurobi), Git, Java
- **Tools:** Snowflake, GCP (BigQuery, VertexAI), AWS (SageMaker, S3), MS SQL Server, Azure Data Factory, PowerBI
- **Skills:** Statistical/Data Analysis, Data Visualization, ETL, Clustering, LLMs, Data Quality Assurance, A/B Testing, MLOps

## PROFESSIONAL EXPERIENCE

### QUANTIUM

Melbourne, Australia

*Data Scientist, Product Analytics*

Jan 2022 – Jun 2023

- Leveraged OOP to integrate customer segmentations and reports into analytical testing process, capturing \$42M of revenue
- Pioneered customer engagement metrics and built Random Forest, LightGBM models in R/Python for churn identification
- Collaborated with Product Managers and Software Engineers to address client inquiries, plan tasks and facilitate debugging
- Mentored 10 Associate Data Scientists team members in Python, Snowflake, SQL, Git over one year

*Associate Data Scientist, Health & Government*

Feb 2021 – Dec 2021

- Elevated XGBoost models by 2% in Python, introducing the state government's first daily population estimation method, enhancing risk planning across all suburban areas during Covid-19
- Led a 5 person PowerBI workstream on a \$400K project to develop an end-to-end tool that analyzes trends in education and health, aiding in identification of at-risk communities; led to a \$2M contract extension
- Engineered automated decision support tool pipeline (Microsoft Power Automate) used in daily meetings by Emergency Management Commissioner and Victorian Chief Medical Officer, reduced weekly labor by 14 hours

## OTHER EXPERIENCE

### MACY'S | MIT Capstone

New York City, NY

*Data Scientist*

Feb 2024 - Present

- Performing exploratory data analysis, clustering (t-SNE, PCA, UMAP) and constructing ETL pipelines on 500M rows of transaction, loyalty, customer and product data
- Building customer price brand models and refining recommender systems for personalized e-commerce/in-store experiences
- Presenting to non-technical audiences, discussing product use cases with business stakeholders and directors

### WEB3 AI STARTUP

Cambridge, MA

Featured as an innovative AIxWeb3 project on the largest Web3 platform in the Asia-Pacific region

Jan 2024 - Present

*Co-Founder & AI Researcher*

- Leading a team of 5 to build the first end-to-end AI-powered Web3 analytics tool that scrapes and clusters text data from Telegram, Discord and Twitter to generate personalized market and sentiment analysis insights using open-source LLMs
- 1/17 finalists selected from 238 teams at Edge Matrix Computing (EMC) DeAI Hackathon; Presented to audience of 30 consisting of EMC CEO, AI Lead of EMC Lab, AI and ML Researchers, Professors, VC and Web3 founders

### HANDLE GLOBAL | MIT Analytics Lab

Cambridge, MA

*Data Scientist*

Sep 2023 – Dec 2023

- Proved medical asset prices have correlation of 4% with PPI inflation index through analysis of price changes on 900K rows
- Developed ensemble of ARIMA and Prophet time series models in Python, improved current model by 38% for \$380M worth of annual medical equipment expenditure
- Presented results and actionable next steps to 200 attendees, including Data Analytics team and Senior Directors

## ADDITIONAL INFORMATION

- Mandarin (fluent), Interests: Novice Film Reviewer, Dedicated NBA/Soccer fan (English Premier League)