

# Ted Zhao

Melbourne, VIC, Australia | [ted.zhao.au@gmail.com](mailto:ted.zhao.au@gmail.com) | 0416 551 622 | [linkedin.com/in/ted-zhao](https://linkedin.com/in/ted-zhao)  
[github.com/tade0726](https://github.com/tade0726)

## Summary

---

- Machine Learning Engineer with 6+ years of experience delivering end-to-end AI solutions across fintech, travel, and retail sectors.
- Expertise in RAG systems, computer vision, and reinforcement learning, improving business metrics.
- Strong MLOps background with Databricks, MLflow, and ZenML, focusing on production-grade model deployment and monitoring.
- Track record of innovative solutions from credit scoring to LLM-powered chatbots, consistently driving business value through AI.

## Experience

---

**AI Engineer**, Lyrebird Health – Melbourne, VIC July 2025 – Aug 2025

- Built LLM evaluation in Langfuse and LangSmith to measure and improve healthcare documentation quality, with A/B testing and user feedback for continuous improvement.
- Built an LLM-as-a-judge system to assess the authenticity, accuracy, and completeness of clinical documentation.
- Optimized prompt design to increase prompt cache hit rate.
- Provided Redaction solution for sensitive information to comply with HIPAA/GDPR.

**AI Engineer**, Department for Education, South Australia – Adelaide, SA June 2025 – July 2025

- Developed LLM-powered chatbot using agentic and workflow paradigms with PDF data sources, implementing metadata-filtered RAG for efficient document search and analytical query processing.
- Built cross-referencing systems for PDF documents and implemented sensitive information redaction workflows to support LLM-powered human-in-the-loop application approval processes.
- Enhanced PDF parsing accuracy through integration of OCR tools and vision-language models (VLMs), improving data extraction completeness for government efficiency initiatives.

**Data/AI Engineer**, Rainmakr.ai – Adelaide, SA (Remote) May 2023 – July 2024

- Built GPT-4 theme extraction pipeline for 10K+ monthly company profiles, achieving 92% accuracy and reducing review time by 75% through optimized prompt engineering.
- Simplified 30+ table data models with DBT, reducing schema size by 50% and redundant data by 15%.
- Optimised queries through join refinement and re-indexing, cutting runtime by 50%.
- Integrated DBT with Dagster Cloud for mixed orchestration of notebook and DBT jobs.

**ML Engineer/Data Scientist**, Skyscanner – Shenzhen, China Nov 2019 – July 2021

- Performed end-to-end ML pipeline in Databricks/MLflow, engineering user behaviour features and iteratively improving CTR by 5-10%.
- Deployed ensemble of LightGBM and Deep Learning models with A/B testing framework for continuous model refinement.
- Developed ResNet-based classifier for 20 hotel categories with hierarchical labelling system, achieving 99% Top-5 accuracy.
- Enhanced image selection using MobileNet with multi-dimensional aesthetic scoring (composition, lighting, detail), lifting CTR by 3%.
- Built T+1 batch processing pipeline using AWS Lambda/PySpark for feature extraction, storing in Parquet format.

**Data Scientist**, DJI – Shenzhen, China Apr 2018 – Aug 2018

- Built Scrapy-based web crawler to gather city demographics and competitor store distribution data for market analysis.

- Developed quantitative ranking algorithm incorporating economic indicators and business factors for city prioritisation.
- Analysed post-launch user demographics (location, income, age) to create detailed customer segments.
- Monitored sales patterns and provided data-driven insights for campaign strategy optimisation.
- Developed Django/SQLAlchemy-based API for Jupyter extension, enabling efficient SQL querying and data model management.

**Data Scientist**, Smart Decision Technology Ltd – Shenzhen, China

Aug 2015 – Mar 2018

- Implemented credit scorecard system with feature selection, WOE calculation, and logistic model conversion in Python.
- Built SimPy-based logistics simulation with Flask/SQLAlchemy backend, securing 2-year contract with S.F. Express.
- Applied K-Means clustering for HKJC member segmentation, delivering lifecycle-based targeting strategies.
- Developed Chinese text sentiment analysis using Word2Vec and K-means for thematic clustering.
- Built supervised learning model with custom Chinese tokenization for app review classification.

## Projects

**ATO Chatbot**

2023

- **Live Demo:** [ato-chat.streamlit.app](https://ato-chat.streamlit.app) | **Source:** [github.com/tade0726/ato\\_chatbot](https://github.com/tade0726/ato_chatbot)
- Built production-grade RAG system with LlamaIndex and Qdrant vector store, enabling context-aware tax information retrieval.
- Implemented MLOps pipeline with ZenML for data ingestion, fine-tuning, and automated model evaluation.
- Developed Streamlit-based UI with real-time query rephrasing and source citations for enhanced user trust.

**Quantitative Trading with Reinforcement Learning** – University of Adelaide

Sept 2022

- Researched and developed crypto trading strategy using Reinforcement Learning.
- Designed multi-asset portfolio optimisation with Sharpe Ratio reward system for buy/sell quantity decisions.
- Built dual Deep Learning modules for price signal prediction and market sentiment analysis.
- Implemented comprehensive backtesting framework evaluating Sharpe Ratio and Maximum Drawdown (MDD).
- Developed profit analysis system incorporating trading fees and compound returns visualisation.

## Skills

**ML/AI:** Reinforcement Learning, Deep Learning, Clustering, LightGBM, XGBoost, SVM, Word2Vec, A/B Testing

**MLOps/Cloud:** Azure ML, AWS, SageMaker, MLflow

**Data Engineering:** DBT, Databricks, Snowflake, Dagster

**Programming:** Python, VBA

**Web Development:** Django, Flask, Asyncio, Scrappy

**ETL/Data Tools:** SQL, SQLAlchemy, Spark, Pandas, Tableau

**Version Control & Containers:** Git, Docker

## Education

**University of Adelaide**, Master in Machine Learning – Adelaide, SA

June 2021 – May 2023

**Shenzhen University**, Bachelor in Mathematical Sciences – Shenzhen, China

Sept 2010 – July 2014

## Languages

**English:** Full professional proficiency **Mandarin:** Native **Cantonese:** Native

## Certifications

---

<b>Coursera</b> , Machine Learning - <a href="https://www.coursera.org/account/accomplishments/verify/Q89DMB9RYYTB">https://www.coursera.org/account/accomplishments/verify/Q89DMB9RYYTB</a>	Oct 2016
<b>Udacity</b> , Machine Learning Nanodegree - <a href="https://graduation.udacity.com/confirm/PKAQPLU2">https://graduation.udacity.com/confirm/PKAQPLU2</a>	Dec 2018