Baistan Zhyldyzbekov

+86-186-2922-4737 | bz529@connect.hkust-gz.edu.cn | linkedin.com/in/baistanzhyldyzbekov | github.com/baistangz

Education

The Hong Kong University of Science and Technology (Guangzhou) — Guangzhou, China
B.S. in Data Science and Big Data Technology; ***Full Tuition Scholarship*** — *Sept. 2025 – May 2029*

Tamchy Gymnasium School — Tamchy, Kyrgyzstan
High School Diploma; GPA: 5.0/5.0 — *Sept. 2013 – June 2024*

- National Main Test (ORT): **218/245** (Top **0.7%** nationwide); achieved the **highest score in school history** since the exam's introduction in 2003.
- Bronze Medalist, 2024 Republican Math Olympiad (Issyk-Kul Region); secured the school's first mathematics victory since 2014.

Experience

GenAI Research Assistant — Oct. 2025 – Present
HKUST(GZ) Computational Media and Arts (CMA) Lab — *Guangzhou, China*

- Building a research pipeline for affective 3D scene generation by combining 3D Gaussian Splatting with VLM-guided, emotion-conditioned editing of panoramic scenes.
- Developing and evaluating visual editing workflows with **Qwen2.5-VL**, **InstructPix2Pix**, and **PyTorch**, with emphasis on scene-wide emotional control, content preservation, and hallucination analysis.

Mathematics Instructor & Mentor — Sept. 2023 – May 2024
Tamchy Gymnasium School — *Tamchy, Kyrgyzstan*

- Selected to teach advanced mathematics while mentoring students on university applications, national exams, and Olympiad preparation.

Projects

lectureGZ | **GitHub** — Mar. 2026 – Present
Full-Stack Developer

- Built an AI-powered lecture study platform with **Next.js**, **React**, **TypeScript**, **Tailwind CSS**, **PostgreSQL**, and **pgvector** for authenticated PDF upload, indexing, and retrieval.
- Implemented a lecture-grounded **RAG** pipeline with PDF parsing, chunking, embeddings, vector retrieval, and page-aware grounding to generate source-anchored explanations.
- Developed quiz generation, free-response evaluation, and interactive PDF-selection workflows, and integrated **Clerk**, **OpenRouter**, and **Render** for authentication, model access, and deployment.

TradeMaster Cup 2025 Kaggle Competition | **GitHub** — Dec. 2025 – Mar. 2026
Data Science / Quant Modeling Project

- Placed **4th/57 teams** in a university-wide Kaggle competition by identifying deterministic leakage in the target-generation process and reconstructing most predictions exactly from future `feature_16` values.
- Derived exact target-reconstruction formulas and reduced the forecasting problem to tail-only prediction on the final unknown **10/60/240** rows for short-, medium-, and long-horizon targets.
- Built custom boundary backtests to simulate the hidden final trading day and developed reproducible pipelines with **Python**, **Pandas**, **NumPy**, **scikit-learn**, **XGBoost**, and **PyTorch**.

Technical Skills

Languages: Python, TypeScript, JavaScript, SQL, HTML5, CSS3
Libraries: Pandas, NumPy, scikit-learn, XGBoost, PyTorch, React, Next.js
Frameworks: FastAPI, Tailwind CSS, Pydantic
Databases: PostgreSQL, pgvector
Tools: Git, GitHub, Render, Vercel, Railway, Clerk, OpenRouter
Concepts: Feature Engineering, Backtesting, Time-Series Modeling, Model Evaluation, RAG, Vector Search, Authentication, REST APIs, PDF Ingestion Pipelines, Reproducible Research