

Zhen Yang

New Haven, CT | zhen.yang.zy372@yale.edu | Website | GitHub | LinkedIn

Summer 2027: Eligible to apply for CPT authorization through Yale

Education

Yale University

M.S. in Biostatistics, Data Science Pathway

Nanjing University of Posts and Telecommunications

B.Eng. in Data Science and Big Data Technology

New Haven, CT

Sep 2026 – May 2028 (Expected)

Nanjing, China

Sep 2022 – Jun 2026

Selected Research

Full publication list: personal website

Accuracy Is Not Fidelity: Decision Churn in Low-Rank Compressed Language Models *1st Author · AAAI 2027 ·*

Under Review

- Reframed compression as paired decision fidelity; introduced Prediction Flip, Damage/Repair/Wrong-to-Wrong decomposition and HiddenChurn, then derived BAR linking perturbation direction, score sensitivity and margin; $\text{BAR} < 1$ certifies prediction preservation for a unique dense winner under the stated analysis.
- Audited 235 compressed checkpoints from 6 dense LLMs, 14 low-rank variants and 7 benchmarks; 39.22% of predictions changed on average, and among 216 cross-method pairs matched within 0.5 pp accuracy drop, PredFlip still differed by 2.27 pp on average and up to 6.97 pp.

When Production Outpaces Evaluation: AI-Assisted Scientific Authorship as a Portfolio Game *1st Author ·*

NeurIPS 2026 · Under Review

- Modeled bounded peer review as a fixed-budget portfolio game with heterogeneous authors/reviewers, latent vs. surface quality, capacity constraints and noisy assignment; derived a portfolio-pressure boundary and analyzed how capacity, viability floors, setup/reputation costs and prestige-weighted rewards shift strategy choice.
- Calibrated the observation layer using public ICLR/NeurIPS review statistics; on 6,010 ICLR 2024 papers, 27.6% lay within the simulated threshold ± 0.5 , a $+0.25$ score perturbation replaced 14.5% of the top- k set, and 3 sampled reviews produced 20.1% selection-set change.

EviE: Evidential Exploration for Value-Based Reinforcement Learning *1st Author · NeurIPS 2026 · Under Review*

- Used Dirichlet vacuity as a state-adaptive exploration rate; proved $1/N(s)$ decay and GLIE/Q-learning convergence, then validated across Chain/Cliff/RiverSwim, DeepSea, 4 Classic-Control and 5 MinAtar tasks with strongest gains in hard-exploration settings.

Industry Experience

Baidu

Algorithm Intern - Search Retrieval

Beijing, China

Jan 2026 – Jun 2026

- Owned post-training, data construction and evaluation for a dual-encoder query-URL retrieval pipeline; diagnosed term omission/mismatch, semantic misinterpretation and relevant-but-not-answerable failures.
- Designed hard-negative mining, higher loss weights for confusing false recalls and staged optimization, improving average Recall by 10.5% across multiple Recall cutoffs and evaluation sets.
- Built four LLM-driven hard-case routes covering answer-preserving paraphrase, condition perturbation, confusion-pair reconstruction and URL-to-query generation, expanding scarce boundary cases for a further +3.2% sequential average Recall.
- Added LLM+rule teacher scoring over candidate URLs and listwise post-training/distillation for finer ranking supervision, delivering another +6.0% sequential average Recall improvement.

HONOR

Algorithm Intern - Generative Engine Optimization

Nanjing, China

May 2025 – Dec 2025

- Independently built the core GEO optimization loop for LLM product recommendation: mined demand from Weibo/Xiaohongshu, allocated weekly content by keyword heat, generated articles, measured live retrieval/product visibility and iterated strategy from weekly reports.
- Built an automated posterior evaluator tracking both retrieved-source inclusion and final HONOR product recommendation, plus an 8B LoRA+contrastive prior evaluator trained on SEO-ranked but LLM-uncited negatives versus LLM-selected references for pre-publication scoring.
- Reached >70% visibility on 14 core prompts with 67 initial articles across 13 online outlets; later supported 1,000+ articles and sustained >65% visibility for six months, while a targeted five-article strategy moved visibility from 0.8% to 63.5%.

Bosch

Algorithm Intern - Automotive Ethernet / Developer Tools

Suzhou, China

Jul 2024 – Dec 2024

- Designed a rule-based gateway generator that transformed structured Excel requirements into complete C++ parsing, message/port adaptation and test code, reducing repetitive implementation from days/weeks to under 30 minutes.
- Built differential validation for a lightweight Protocol Buffer implementation against Google's reference library on real vehicle signals and randomized legal-range cases; isolated cross-language rounding discrepancies and raised output consistency from 98.4% to 100%.
- Implemented RTP/RTSP/RTCP components in Python and studied/implemented IEEE 1588 and gPTP synchronization logic, documenting protocol behavior and integration details for the team.

Industrial and Commercial Bank of China (ICBC)

Algorithm Intern - Credit Risk Modeling

Changzhou, China

Jan 2024 – Feb 2024

- Standardized multi-source customer records and engineered features for an internal model-comparison pipeline; implemented KNN, Naive Bayes and MLP loan-risk baselines, with the best lightweight model reaching 72% accuracy.
- Developed an LSTM-Transformer over three months of behavioral data, reaching 0.89 AUC and identifying 85% of high-risk customers two weeks in advance as an experimental early-warning signal.