

Temuujin Erdene

Ulaanbaatar, Mongolia · temuujin.razy@gmail.com · +976-9914-7798
github.com/TemuujinE · linkedin.com/in/temuujin-erdene · kaggle.com/temuujinerdene

RESEARCH INTERESTS

My interests center on optimization for multimodal machine learning, especially efficient and robust training for structured document understanding. I am particularly interested in the convergence and stability of learning algorithms, entropy-guided reinforcement learning and image processing.

EDUCATION

B.Sc. in Applied Mathematics

2020–2024

National University of Mongolia — Ulaanbaatar, Mongolia — GPA: 3.2/4.0

Concentration GPA: **3.8/4.0** across eight mathematics courses in optimization, PDEs, dynamical systems, and analysis.

- **Concentration coursework:** ODE, PDE, Complex Analysis, Applied Functional Analysis, Dynamical Systems, Mathematical Modelling, Linear Optimization, and Nonlinear Optimization.
- **Additional mathematical coursework:** Measure Theory, Probability Theory, Mathematical Statistics, Numerical Methods, and Monte Carlo Simulation.

B.Sc. thesis: *Applying Image Processing and Machine Learning Approaches for Analyzing Grassland Degradation*. Advisor: Prof. Galtbayar Artbazar.

- **Problem and data:** Developed an image-based alternative to extrapolating pasture condition from a small number of one-square-meter field plots. Used multispectral drone imagery from five sites in the mountain-steppe landscape around Khustai Nuruu National Park.
- **Image analysis:** Estimated bare-soil cover with the MSAVI vegetation index and Otsu thresholding. Detected clump-forming weeds with a bank of Gabor filters, using Gaussian-windowed sinusoidal kernels at selected directions and scales to measure recurring texture, then refined the masks with erosion, dilation, and connected-component analysis.
- **Mathematical formulation and evaluation:** Represented each site by the pair $(p_{\text{soil}}, p_{\text{weed}})$ and defined a two-variable degradation criterion that reproduced the relative order of the field assessments. Trained U-Net-based segmentation models for comparison and tested the measurements across three image resolutions using Welch's t -tests and 10,000-sample permutation tests.

ADVANCED MATHEMATICAL STUDY

Directed Reading Seminars

November 2021–December 2024

Center of Mathematics for Applications, National University of Mongolia

- **Format:** Served as a rotating presenter for three years, presenting assigned material and complete proofs under Prof. Artbazar Galtbayar and Prof. Majig Mend-Amar.
- **Analysis and optimization:** Studied measure and Lebesgue integration, L^p spaces, Hilbert and Banach spaces, Hahn–Banach theory, bounded operators, distributions, and Fourier analysis through Stein and Shakarchi. Through Bertsekas, studied gradient methods, convergence rates, convexity, KKT theory, and Lagrangian duality.

ACADEMIC COLLABORATION

Hepatic Steatosis from Histopathology Images

April–June 2024

National University of Mongolia and Mongolia–Japan Hospital

- **Team leadership:** Worked in a multidisciplinary collaboration involving six faculty and researchers and directed a three-student modeling team. Planned and assigned its data-labeling, model-training, testing, and architecture-comparison work.
- **Technical contribution:** Jointly defined the annotation protocol with academic and clinical collaborators; built the annotated dataset and the preprocessing, training, and evaluation pipeline; trained and compared several models for estimating liver-fat percentage; and presented the work at the National University of Mongolia.

INDUSTRY R&D

Document Image Text Detection

April 2025

- **Method and result:** Independently applied knowledge distillation to replace a slower document-text detector. Swept input scale and decision thresholds rather than comparing the models at only one operating point, reducing latency from 0.579 to 0.073 seconds per image ($7.9\times$ faster) while increasing recall from 0.847 to 0.931 and F_1 from 0.870 to 0.897.
- **Analysis and current use:** Examined the errors corrected by the distilled model, including false detections on stamps and figures and words merged across separate rows or columns. The model now serves as the text-detection stage of the

production tender system.

Technical-Specification Extraction with Vision-Language Models

September 2025

- **Problem:** Found that the OCR-to-language-model approach made structured-output quality dependent on upstream recognition quality. This motivated a shift to direct image-to-structure prediction for tables with inconsistent layouts, merged or nested cells, and widely varying attribute formats.
- **Benchmark:** On 669 specifications from 71 products, the in-house model reached 98.1% recall, 97.9% precision, and 83.1% product-name accuracy, outperforming GPT-4.1, GPT-5, and Gemini Pro on those measures.
- **Finding and current use:** Reported a split result because Gemini Pro remained stronger on exact specification fields. Identified loss of contextual grouping in nested specifications as the main structural failure mode. The resulting model now performs technical-specification extraction in the production tender system.

Entropy-Gated Reinforcement Learning

October 2025

- **Method:** Studied the progression from PPO to GRPO and DAPO and the use of token entropy to restrict policy updates to the most uncertain parts of a generated sequence. Replicated the method on vision-language document classification to test whether a result developed for text reasoning transferred across modality and domain.
- **Current use:** Adapted the same entropy-guided training idea to hierarchical call-tag classification. It now informs reinforcement-learning experiments for the production tag classifier.

Training VLMs for Mongolian Document Understanding

April 2026

- **Research question:** Investigated whether one VLM could jointly detect page layout and transcribe handwritten and typed Cyrillic documents and typed Traditional Mongolian vertical script.
- **Training and analysis:** Compared supervised fine-tuning, multi-token prediction, and GRPO. Used token entropy to analyze instability under visual noise.
- **Efficiency and deployment:** Studied MTP as both a training objective and a built-in draft model for speculative decoding, measuring a 26.5% inference speedup in the experimental setting. This work now underlies the National Security Council OCR system, where the same VLM predicts exact bounding boxes for page elements and transcribes them.

SELECTED PROFESSIONAL EXPERIENCE

AI Engineer

July 2024–Present

Egune AI — Ulaanbaatar, Mongolia

Automated Tender Evaluation

Erdenet Mining Corporation

- **Role and responsibility:** Working with the team lead, redesigned and extended the initial release (version 1) into the current production system. Performed all model training and designed the end-to-end pipeline; implementation was divided between us.
- **Problem framing:** Formulated tender evaluation as traceable constraint checking: extract requirements from the invitation, find supporting evidence in each bidder's documents, and produce compliance decisions that procurement specialists can review and correct.
- **Production system:** Built staged processing for rasterization, orientation correction, OCR, layout and table detection, classification of more than 75 document types, structured specification extraction, and requirement validation. Trained task-specific language and vision-language models and separated the stages into independently scalable services. The system has been in production since July 2025.

Mongolian Document OCR

Sole AI engineer

National Security Council of Mongolia

- **End-to-end responsibility:** Designed, trained, implemented, and continue to maintain the complete production system.
- **Core model:** A single VLM locates page elements with exact bounding boxes and transcribes handwritten and typed Cyrillic documents and typed Traditional Mongolian vertical script.
- **Training and script handling:** Incorporated the supervised fine-tuning, MTP, GRPO, and entropy-analysis experiments described above. Added orientation handling that corrects rotated pages without disturbing naturally vertical Traditional Mongolian text.

Mongolian Speech Recognition

Golomt Bank, Erdenet Mining Corporation, and Chimege

- **Production maintenance:** Maintain speech-to-text models and services spanning bank and mining call centers, long-form dictation, protocol transcription, and subtitle generation.
- **Subtitle alignment:** Independently migrated the subtitle system from Whisper to faster-whisper and developed alignment between generated text and character offsets so subtitle boundaries retain exact audio timestamps.
- **Call-center alignment:** Diagnosed a dual-channel failure caused by character offsets spanning long intra-channel silences. Replaced a one-directional heuristic with direction-aware cluster correction while preserving a text-to-offset length invariant through every transcript-normalization rule.

Speaker Diarization

Production transcription systems

- **Method:** Maintain and improve the diarization system. Implemented spectral clustering on the graph Laplacian of a pruned, symmetrized cosine-affinity matrix, estimated the unknown speaker count from its eigenvalue spectrum, and compared the method with agglomerative clustering on identical embeddings.
- **Result and integration:** Improved the speaker-detection score by approximately 4% and integrated the output with the shared speech-recognition and transcript-alignment components.

Call-Center Quality Analytics

Golomt Bank

- **Scoring system:** Maintain and extend a production service covering 19 call-quality criteria as the rules evolve. Compute timing, overlap, and lexicon-based measures deterministically from word timestamps and use language-model judgments only for context-dependent criteria.
- **Model training and evaluation:** Prepare training data and train hierarchical call-tag classifiers through supervised fine-tuning and GRPO. Used a structured evaluation protocol and the entropy-guided study above to inform further tag-classification experiments.

Conversational Assistant

Egune app

- Maintain the assistant, prepare and curate model-training data, and run fine-tuning experiments across chat, push-to-talk, and free-form voice channels.

Data Analyst (Part-time)

February–September 2023

Zeely — Ulaanbaatar, Mongolia

- **Data preparation:** Integrated and standardized loan histories, application records, and customer data, expanding 19 source fields into 60 model-ready variables through aggregation and feature engineering.
- **Risk modeling:** Compared random forests, gradient-boosted models, and voting ensembles for predicting loan delinquency. The selected random forest reached $F_1 = 0.866$ on an 80/20 held-out split; analyzed the precision–recall trade-off in terms of missed risk and unnecessary customer intervention.
- **Additional experiments:** Trained models to count faces in loan-application images and match applicants with their identity-document photographs. Also tested a genetic-algorithm formulation for selecting relevant earlier loans and reported the negative result when it did not improve prediction.

Data Analyst Intern

June–October 2022

Golomt Bank — Ulaanbaatar, Mongolia

- Developed intent classification, slot filling, and dialogue management for a Mongolian banking chatbot using Rasa. Prepared and refined training examples and conversation flows for common banking requests.
- Implemented custom actions and database-backed responses and tested multi-turn flows to inspect predictions and improve how the chatbot collected required information.

Data Analyst Intern

June–September 2021

Golomt Bank — Ulaanbaatar, Mongolia

- **Data preparation:** Corrected spelling, converted transliterated text to Cyrillic, verified transcripts, and applied consistent intent, entity, slot, and sentiment labels to Mongolian text and speech data.
- **Modeling and analysis:** Trained a 35-tag HMM part-of-speech tagger on 140,000 tokens (88% accuracy) and a CRF named-entity recognizer ($F_1 = 0.770$). Studied spectral features, HMM acoustic models, and Viterbi decoding, focusing on errors caused by repetitions, prolongations, and pauses.

HONORS AND AWARDS

Kaggle Gold Medal — JPX Tokyo Stock Exchange Prediction

2022

- Placed 11th among 1,039 scored teams on unseen market data. Independently reimplemented the Sharpe-ratio metric and evaluated 10,000 random submissions, showing that many leaderboard differences over the short final window were statistically indistinguishable from sampling noise.

SKILLS

Programming and ML: Python, PyTorch, NumPy, SciPy, scikit-learn, CUDA.

Training and systems: SFT, LoRA, GRPO, MTP, Knowledge Distillation, Docker, Kafka, REST, vLLM, MongoDB, PostgreSQL, Git.