Taro Shibusawa Software Engineer

tarose.cs0112@outlook.com | Morioka, Iwate, Japan | Japanese | Linkedin

PROFESSIONAL SUMMARY

Distinguished AI/ML Engineer and Full-Stack Architect with 8+ years of expertise spanning artificial intelligence, machine learning, data engineering, and full-stack development across web and mobile platforms. Proven track record of architecting and deploying production-grade AI systems that process petabytes of data, building intelligent automation solutions, and creating scalable applications used by millions. Specialized in bridging the gap between cutting-edge AI research and practical business applications through end-to-end product development—from data pipeline architecture and deep learning model development to intuitive UI/UX implementation. Expert in transforming complex technical challenges into elegant, high-performance solutions that deliver measurable ROI.

Core Competencies: AI/ML Model Development & Deployment | NLP & Computer Vision | AI Agents & Automation | Predictive Analytics | Full-Stack Web Development | Cross-Platform Mobile Apps (iOS/Android) | MLOps & Data Pipeline Engineering | Microservices Architecture | Cloud Infrastructure (AWS/GCP/Azure)

EXPERIENCE

Lead Software Engineer Nov 2023 – Apr 2025 | Morioka, Japan

BASE SYSTEM CORP.

In this company: Leading AI/ML initiatives for enterprise-scale predictive analytics and intelligent automation solutions.

Predictive Analytics Platform Development:

- Architected and engineered a company-wide predictive analytics platform processing **2TB+ of daily user event data** across 5 million+ active users, implementing Apache Airflow orchestration on AWS infrastructure
- Designed and deployed distributed ETL pipelines using Apache Spark and Airflow, automating feature engineering workflows that process 50+ data sources in real-time
- Built and optimized XGBoost ensemble models for customer churn prediction with hyperparameter tuning via Optuna, achieving **18% YoY reduction in customer attrition** and **$2.3M annual revenue retention**
- Implemented MLflow for experiment tracking and model versioning, establishing CI/CD pipeline for automated model retraining and deployment using GitHub Actions and AWS SageMaker

Full-Stack Data Visualization & Business Intelligence:

- Led end-to-end development of enterprise data visualization dashboard serving 200+ internal stakeholders, using React 18, TypeScript, and D3.js for interactive frontend visualizations
- Engineered FastAPI backend with PostgreSQL database optimization, implementing connection pooling and query optimization that reduced average response time from 3.2s to 0.4s
- Integrated real-time data streaming using Kafka and WebSocket connections, enabling live metric updates and collaborative data exploration
- **Impact:** Eliminated **20+ hours per week** of manual reporting work, enabling analytics team to focus on strategic initiatives

NLP & Intelligent Automation:

- Pioneered adoption of transformer-based NLP models to automate categorization and priority scoring of **500K+ annual customer support tickets**
- Fine-tuned DistilBERT model using custom training loop with mixed-precision training, achieving **92% accuracy** and **89% F1-score** across 15 ticket categories
- Deployed model using TorchServe on AWS ECS with auto-scaling, handling 10K+ inference requests per minute with <100ms latency
- Developed intelligent routing system that reduced average ticket resolution time by 35% and improved customer satisfaction scores by 22%

Cross-Platform Mobile Development:
- Architected and developed customer-facing mobile applications using Flutter for iOS/Android with offline-first architecture and local ML model inference
- Implemented TensorFlow Lite models for on-device predictions, enabling real-time recommendations without network dependency
- Built React Native components with native module integration for performance-critical features on both platforms

Leadership & Mentorship:
- Mentored 3 junior data scientists and engineers on MLOps best practices, code review standards, and software engineering principles
- Established team coding standards, documentation practices, and knowledge-sharing sessions that improved code quality metrics by 45%
- Led weekly architecture review meetings and contributed to technical hiring by conducting 20+ technical interviews

Project Manager Feb 2022 – Oct 2023
Sakura Data System Hanamaki, Iwate, Japan

In this company: Drove AI-powered innovation in agricultural technology through computer vision and predictive modeling.

Computer Vision & Quality Assurance Systems:
- Designed and deployed CNN-based defect detection system for agricultural quality control using YOLOv8 architecture, achieving **99.2% accuracy** and **97.8% precision** on 100K+ labeled images
- Implemented real-time image processing pipeline with OpenCV and CUDA acceleration, processing 30 frames per second on edge devices
- Trained custom object detection models using transfer learning from COCO-pretrained weights, with data augmentation strategies including geometric transforms, color jittering, and mixup
- **Impact:** Replaced manual inspection process, saving 1,200+ labor hours monthly and reducing defect-related losses by 85%

Mobile Application Development:
- Built cross-platform mobile application using Flutter with BLoC state management, serving 5,000+ active users across iOS and Android
- Integrated on-device ML inference using TensorFlow Lite with model quantization, reducing model size from 127MB to 18MB while maintaining 98% accuracy
- Implemented offline-first architecture with SQLite for local storage and background sync, ensuring functionality in remote areas with poor connectivity
- Designed intuitive camera interface with real-time overlay feedback, achieving 4.8/5.0 app store rating

Supply Chain Optimization & Forecasting:
- Developed time-series forecasting models using PySpark and Prophet to predict supply chain demand across 200+ distribution points
- Implemented distributed data processing pipelines with Apache Airflow, processing 50GB+ of IoT sensor data daily from field devices
- Built ensemble forecasting models combining LSTM, ARIMA, and seasonal decomposition, improving forecast accuracy by 34%
- **Impact:** Reduced supply chain waste by **20%** and optimized inventory levels, saving **$480K annually**

AI-Powered Decision Support:
- Created GPT-3-powered conversational AI chatbot providing expert agricultural advice to 3,000+ farmers
- Implemented RAG (Retrieval-Augmented Generation) architecture with ChromaDB vector database for domain-specific knowledge
- Fine-tuned language model on 50K+ historical support tickets and agricultural documentation
- **Result:** Reduced crop losses by **15%** through timely intervention recommendations

IoT & Data Engineering:
- Architected IoT data collection infrastructure connecting 500+ field sensors with MQTT protocol and AWS IoT Core
- Built real-time streaming analytics pipeline using Kinesis and Lambda for anomaly detection and alerts
- Implemented data lake architecture on S3 with Glue ETL jobs for automated data cataloging and quality checks

Software Developer Jun 2021 – Nov 2021
Forest Innovations Shizukuishi, Iwate, Japan

In this company: Delivered enterprise-grade full-stack applications with integrated AI/ML capabilities for Fortune 500 clients.

AI-Enhanced CRM & Customer Intelligence Platform:
- Architected predictive CRM system for major automotive manufacturer serving 2M+ customers, using React.js, Node.js, and Python microservices
- Implemented ML-powered customer segmentation using K-means clustering and RFM analysis, identifying 15 distinct customer personas
- Built churn prediction models using gradient boosting (XGBoost) with SHAP explainability, achieving **0.89 AUC-ROC**
- Developed recommendation engine using collaborative filtering and content-based algorithms, increasing upsell conversion by 28%
- **Impact:** Improved client management efficiency by **40%** and increased customer lifetime value by **$850 per customer**

Conversational AI & NLP Solutions:
- Deployed production NLP chatbot system handling 50K+ daily conversations across web and mobile channels
- Fine-tuned BERT models for intent classification and entity extraction, achieving 94% accuracy on domain-specific queries
- Implemented context management system with Redis for multi-turn conversations, maintaining conversation history across sessions
- Integrated with Twilio, Slack, and WhatsApp APIs for omnichannel support
- **Result:** Reduced customer support response time by **50%** and handled 65% of queries without human escalation

Microservices Architecture & System Modernization:
- Led migration of monolithic PHP application to microservices architecture using Node.js, Python FastAPI, and Spring Boot
- Designed 12 independent microservices for financial platform handling **$5M+ daily transaction volume**
- Implemented AI-augmented fraud detection service using Random Forest and anomaly detection algorithms, catching 98% of fraudulent transactions
- Built real-time risk assessment engine processing 200+ features per transaction with sub-100ms latency
- Deployed services on Kubernetes with Istio service mesh, implementing circuit breakers, rate limiting, and distributed tracing
- **Impact:** Reduced deployment time by **60%**, improved system reliability to 99.97% uptime, and prevented **$1.2M in fraud losses**

Real-Time Analytics & AI-Powered Insights:
- Engineered real-time financial dashboard using WebSocket, D3.js, and React, processing **100K+ data points per minute**
- Implemented streaming anomaly detection using Isolation Forest and LSTM autoencoders for fraud detection
- Built time-series forecasting models (LSTM, GRU, Prophet) for market trend prediction with 89% directional accuracy
- Developed automated alert system using predictive models that identified market opportunities 2-3 hours ahead of traditional indicators
- **Result:** Dashboard achieved **95% anomaly detection accuracy** and enabled traders to act on insights **3x faster**

Security & AI-Driven Threat Intelligence:
- Designed OAuth2 authentication service with behavioral biometrics using TensorFlow for anomaly detection
- Implemented ML-based login pattern analysis using autoencoders to detect account takeover attempts
- Built automated penetration testing framework using reinforcement learning agents for vulnerability discovery
- Created security monitoring dashboard with SIEM integration and automated threat response
- **Impact:** Reduced unauthorized access attempts by **95%** and identified vulnerabilities **3x faster** than manual testing

Mobile Application Development:
- Developed iOS application using SwiftUI and Combine framework with MVVM architecture
- Built Android application using Kotlin, Jetpack Compose, and Coroutines for reactive programming
- Implemented shared business logic using Kotlin Multiplatform Mobile (KMM) reducing code duplication by 40%
- Integrated Core ML and TensorFlow Lite for on-device AI inference on mobile platforms

Data Scientist Feb 2019 – Sep 2021
Artiza Networks, Inc. Morioka, Iwate, Japan

In this Comapany: Applied statistical analysis and machine learning to pharmaceutical research and clinical trials.

Statistical Analysis & Clinical Trial Optimization:

- Performed advanced statistical analysis on Phase III clinical trial data involving 10,000+ patients across multiple treatment arms
- Conducted rigorous A/B testing, survival analysis, and multivariate regression using Python (Pandas, SciPy, statsmodels) and R
- Implemented Bayesian statistical methods for adaptive trial design, enabling early stopping rules and sample size recalculation
- Identified statistically significant factors ($p < 0.05$) influencing patient outcomes through propensity score matching and causal inference techniques
- **Impact:** Findings directly contributed to successful refinement of 2 Phase III drug trials, accelerating FDA approval timeline by 4 months

Predictive Modeling for Clinical Operations:

- Developed and deployed time-series forecasting models in PyTorch to predict patient enrollment rates across **100+ clinical sites globally**
- Implemented LSTM and Transformer-based models with attention mechanisms for multi-step forecasting
- Built feature engineering pipeline incorporating historical enrollment data, demographic factors, site characteristics, and seasonal patterns
- Created ensemble model combining statistical and deep learning approaches, achieving 87% forecast accuracy
- **Result:** Improved resource allocation efficiency and reduced operational costs by **15%** (**$2.8M annual savings**)

Data Infrastructure & Analytics Platform:

- Architected scalable data models in Snowflake supporting 50+ concurrent analysts and 200+ daily queries
- Designed star schema data warehouse with slowly changing dimensions (SCD Type 2) for historical tracking
- Authored complex SQL queries with window functions, CTEs, and query optimization reducing execution time by 75%
- Implemented dbt for data transformation and testing, establishing data quality checks and documentation
- Built automated data pipeline using Airflow for daily ETL processing of clinical trial data from 15+ source systems

Business Intelligence & Reporting:

- Created 20+ interactive Tableau dashboards visualizing key trial metrics for C-suite executives and clinical operations teams
- Implemented calculated fields, LOD expressions, and advanced visualizations including cohort analysis and funnel charts
- Designed executive summary reports with drill-down capabilities and real-time data refresh
- Established KPI tracking system for trial milestones, patient recruitment, and safety metrics
- **Impact:** Enabled data-driven decision-making that improved trial efficiency and stakeholder communication

TECHNOLOGY & SKILLSET

Software Engineering

Programming Languages:

- **Expert:** Python, JavaScript/TypeScript, Java
- **Proficient:** Swift, Kotlin, Go, R, C++, SQL

Frontend Development:

- **Frameworks:** React/Next.js, Vue/Nuxt.js, Angular, Svelte
- **Mobile:** Flutter, React Native, SwiftUI, UIKit, Jetpack Compose
- **State Management:** Redux, MobX, Zustand, Riverpod, BLoC
- **Styling:** Tailwind CSS, Material-UI, Styled Components, CSS3, SASS

Backend Development:

- **Frameworks:** Node.js/Express/NestJS, Django, Flask, FastAPI, Spring Boot, ASP.NET Core
- **APIs:** REST, GraphQL, gRPC, WebSocket, Server-Sent Events
- **Authentication:** OAuth 2.0, JWT, Auth0, Firebase Auth

Databases & Data Storage:

- **SQL:** PostgreSQL, MySQL, SQLite, SQL Server
- **NoSQL:** MongoDB, Redis, Cassandra, DynamoDB
- **Vector DB:** Pinecone, ChromaDB, Weaviate, Milvus
- **Data Warehouses:** Snowflake, BigQuery, Redshift

Architecture & Design Patterns:

- Microservices, Monolithic, Event-Driven, Serverless, Domain-Driven Design (DDD)
- Design Patterns: MVC, MVVM, BLoC, Clean Architecture, Repository Pattern
- System Design: Load Balancing, Caching, Message Queues, API Gateway

Cloud & DevOps

Cloud Platforms:

- **AWS:** EC2, Lambda, S3, ECS/EKS, SageMaker, RDS, DynamoDB, CloudFormation, API Gateway, Step Functions
- **Google Cloud Platform:** Cloud Run, Cloud Functions, GKE, BigQuery, Cloud Storage, Vertex AI
- **Azure:** Azure Functions, App Service, AKS, Cosmos DB, Azure ML

Infrastructure & Orchestration:

- **Containers:** Docker, Docker Compose, Podman
- **Orchestration:** Kubernetes, Helm, Istio, AWS ECS
- **IaC:** Terraform, AWS CDK, CloudFormation, Pulumi

CI/CD & Automation:

- **CI/CD:** GitHub Actions, GitLab CI, Jenkins, CircleCI
- **Testing:** Jest, Pytest, Cypress, Selenium, JUnit, Mocha
- **Monitoring:** Prometheus, Grafana, ELK Stack, Datadog, New Relic, Sentry

Version Control & Collaboration:

- Git, GitHub, GitLab, Bitbucket
- Agile/Scrum methodologies, Jira, Trello, Asana, Linear

AI/ML & Data Science

Machine Learning:

- **Frameworks:** Scikit-learn, XGBoost, LightGBM, CatBoost
- **Techniques:** Supervised Learning (Classification, Regression), Unsupervised Learning (Clustering, Dimensionality Reduction), Ensemble Methods, Transfer Learning
- **Optimization:** Hyperparameter Tuning (Optuna, Ray Tune), AutoML, Model Compression, Quantization

Deep Learning:

- **Frameworks:** PyTorch, TensorFlow/Keras, JAX, ONNX
- **Architectures:** CNNs (ResNet, EfficientNet, YOLO), RNNs/LSTMs, Transformers, GANs, Autoencoders
- **Deployment:** TensorFlow Lite, TorchServe, ONNX Runtime, TensorRT, Core ML

Natural Language Processing:

- **Libraries:** Hugging Face Transformers, spaCy, NLTK, Gensim
- **Models:** BERT, GPT, T5, RoBERTa, DistilBERT, Sentence Transformers
- **Techniques:** Fine-tuning, RAG, Prompt Engineering, Text Classification, NER, Sentiment Analysis, Topic Modeling, Question Answering
- **Frameworks:** LangChain, LlamaIndex, Semantic Kernel

Computer Vision:

- **Libraries:** OpenCV, PIL/Pillow, Albumentations, imgaug
- **Models:** YOLO (v5/v8), ResNet, EfficientNet, ViT, Mask R-CNN
- **Applications:** Object Detection, Image Classification, Segmentation, OCR, Face Recognition

Data Analysis & Engineering:

- **Analysis:** Pandas, NumPy, SciPy, Polars, Dask
- **Statistics:** Statistical Testing, A/B Testing, Hypothesis Testing, Bayesian Analysis, Time Series Analysis
- **Visualization:** Matplotlib, Seaborn, Plotly, D3.js, Tableau, Power BI
- **Big Data:** Apache Spark (PySpark), Apache Kafka, Apache Airflow, Databricks
- **Databases:** PostgreSQL, MySQL, MongoDB, Redis, Snowflake, BigQuery, DynamoDB

MLOps & Production ML:

- **Experiment Tracking:** MLflow, Weights & Biases, Neptune.ai
- **Model Serving:** TorchServe, TensorFlow Serving, FastAPI, Flask
- **Monitoring:** Evidently, Grafana, Prometheus, Kibana
- **Orchestration:** Apache Airflow, Kubeflow, MLflow, Prefect
- **Feature Stores:** Feast, Tecton

EDUCATION

Master of Science (MSc) / Computational Data Science Apr 2014 – Oct 2018 | Tokyo, Japan

University of Tokyo

Paper : "Synthetic Data Augmentation via GANs for Enhanced Coronary Heart Disease (CHD) Multi-Class Classification using Convolutional Neural Networks"

This paper presents a novel hybrid GAN-CNN framework designed to address the challenge of limited data in medical diagnostics by generating high-fidelity synthetic patient data for robust model training. Our model achieved high classification accuracy (>85%) on a 4-class CHD dataset, demonstrating that GAN based augmentation significantly improves CNN performance and generalization for predicting complex disease categories.

Bachelor of Science (BSc) / Computer Science Apr 2014 – Oct 2018 | Tokyo, Japan

University of Tokyo

Paper: *"Edge-AI for Sustainable Agriculture: A Federated Learning Approach to Pest Detection on Low-Power Devices"*

This research developed a sustainable federated learning framework that trains CNN models directly on distributed, solar-powered edge devices, eliminating the need to centralize sensitive agricultural data; by implementing TensorFlow Lite optimization with quantization and pruning, the system achieved an 87% reduction in model size while maintaining 94% accuracy, and its deployment across 15 farms successfully reduced pesticide usage by 30% through targeted early pest detection, as published in the Journal of Sustainable Computing.

Academic Achievements Apr 2014 – Oct 2018 | Tokyo, Japan

University of Tokyo

- **1st Prize Winner** - Japan's National Green Computing Challenge (2017) for energy-efficient model training methodology
- GPA: 3.9/4.0 | Dean's List all semesters
- Teaching Assistant for "Machine Learning Fundamentals" and "Deep Learning Applications"

Relevant Coursework: Advanced Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Distributed Systems, Big Data Analytics, Cloud Computing, Statistical Learning Theory

CERTIFICATES

AWS Certified Solutions Architect - Associate
Amazon Web Services, (2022)

Microsoft Certified: Azure Developer Associate
Microsoft Corporation, (2023)

LangChain for LLM Application Development
DeepLearning.AI (2023)

Deep Learning Specialization
DeepLearning.AI / Coursera (2022)

TensorFlow Developer Certificate
TensorFlow (2021)

LANGUAGES

Japanese

English