ITTOOP SHINU SHIBU

shinu.shibu1798@gmail.com | YOUR-DOMAIN | linkedin.com/in/ittoop-shinu-shibu | github.com/shinushibu17

SUMMARY

Data scientist and applied ML engineer with nearly six years in data science and analytics, including three years building and operating production ML at Singtel, Singapore's largest telecom. Recent work on hybrid retrieval with cited generation and on the reliability of LLM-based evaluation.

PROFESSIONAL EXPERIENCE

Data Scientist | Singtel, Singapore — Aug 2022 - Aug 2025

- Built and deployed propensity and customer-product alignment models (XGBoost, LightGBM) on joint Singtel and Etiqa insurance customer data in a Snowflake data clean room during a $2M BCG engagement; drove +80% YoY lead conversion and +210% product-line revenue growth.
- Designed the rollout measurement framework: A/B tests against a random 10% no-touch holdout held for three months, isolating incremental conversion from targeting effects; lead propensity model reached 0.81 AUC, and consumer-mobile propensity models lifted cross-sell conversion +10% MoM.
- Diagnosed a label pipeline gap through error analysis, fixed the write-back, and rebuilt the affected model on corrected labels.
- Piloted a Llama 4 text-to-plot RAG system for analyst reporting during testing: 86% analyst agreement, 50% less reporting time.
- Operated models as Databricks jobs (Spark, MLflow): run-cost optimization, scheduled retraining, drift monitoring, holdout validation.

Business Analyst, promoted to Manager | Singtel, Singapore — Aug 2020 - Aug 2022

- Built a customer lifetime value model in Python and R, benchmarking regularized regression and survival models; improved segment profitability by 15%.
- Maintained Airflow pipelines on a Hadoop cluster and automated reporting workflows, cutting delivery timelines ~20%.
- Managed four analysts and set delivery processes, raising team productivity by 30%.

GRADUATE AND RESEARCH EXPERIENCE

Special Initiatives Intern, Applied AI Prototyping | BU Spark!, Boston University — Jun - Aug 2026

- Scoped 6+ client ML projects; prototyped an AI Chinese-character learning app and an IATI data pipeline (Python ETL, Neo4j).

Graduate Research Assistant, Deepfake Forensics | Boston University — Mar - Jun 2026

- Built frame-extraction and dataset-standardization pipelines for DF40, FaceForensics++, Celeb-DF v2, and DeepSpeak v2, and evaluated deepfake detection models on A40 and H100 GPUs; co-author on a paper accepted at IDFC 2026.

ML Technical Project Manager | BU Spark!, Boston University — Jan - May 2026

- Led technical delivery of four client ML projects with teams of 4 to 6. For the Boston Public Library: hybrid retrieval with cited generation (BGE-M3, pgvector, Neo4j, RRF, GPT-4o), a 51-query gold benchmark, and MRR 0.049 to 0.209 and Hit@10 0.029 to 0.429 on 35 held-out queries.

SELECTED PROJECTS

LLM-as-Judge Reliability Audit | github.com/shinushibu17/judge-reliability — 2026

- Pre-registered replication on MT-Bench human judgments: across three judges and 100 pairwise comparisons, judge-human agreement ran 0.17 to 0.21 below the human-human baseline (bootstrap CIs exclude zero); judges chose ties on 3 to 9% of pairs versus 23 to 25% for humans.

Generative Augmentation for Hand Osteoarthritis | CS790, BU: three generators at five augmentation ratios on 41,051 radiographs (135:1 imbalance); none improved KL4 recall.

Wearable IMU Exercise Classification | BU: five architectures on 164 Apple Watch recordings, recording-level CV; CNN-Transformer 90.0% ± 5.7%.

PUBLICATIONS

Akarsu Z., Rizinski M., Zhang M., Choi K., Shah D., Shibu I. S., Shin C. W., Chitkushev L. Detecting Deepfakes for the Courtroom: A Multi-Modal Forensic Output Framework. International Conference on Intelligent Digital Forensics and Cybersecurity (IDFC 2026). Accepted.

EDUCATION

MS, Computer Science | Boston University, GPA 3.67 — Aug 2026

BS, Data Science | University of California, San Diego, Provost Honors — 2020

Awards: Singtel Spot Award (2024) | Singtel Star Award (2023) | Winner, Civic Hacks 2026 | 4th place, BU DS+X Hackathon 2025

SKILLS

Programming and Data: Python, R, SQL, Spark and PySpark, Databricks, Snowflake, Airflow, Hadoop, Docker, Git
ML and MLOps: XGBoost, LightGBM, scikit-learn, PyTorch, survival models, MLflow, retraining, drift monitoring, holdout design
LLM Systems: RAG, GraphRAG, pgvector, Neo4j, LangChain, LangGraph, vLLM, LLM evaluation