Mrityunjay Kumar

Email : mjay.cse@gmail.com

linkedin.com/in/mj2030 github.com/mrityunjaykumar911 Mobile : +1-631-710-1058

Core ML engineer focused on the reward and evaluation side of RL for tool-using agents: reward signal correctness, LLM-as-judge grading, multi-turn evaluation, and multimodal tool integration into live RL rollouts. Co-author **EuroSys 2022** (Rolis). Patent holder in storage metadata verification.

Skills

- **Core ML:** RL reward & evaluation infrastructure, RLVR/RLHF/SFT, LLM-as-judge reward design, multi-turn agentic evaluation, multimodal grading, RL environment & tool integration, async rollout integration, LLM fine-tuning pipelines, prompt optimization
- **Distributed Systems & Infra:** Distributed systems, consensus protocols (Paxos, Raft), transactional databases, container orchestration, cloud-native microservices, Redis, high-throughput pipeline design
- **Cloud & DevOps:** Azure (ACI, Blob, ML, AAD), AWS, Docker, Prometheus, Grafana, OpenTelemetry/Jaeger, Nginx, gRPC, REST
- **Languages:** Python, C#, C++, C, Java · PyTorch, Git

Experience

Microsoft 05/2022 – Present

Senior ML Engineer, RL Reward & Evaluation Infrastructure · PowerPoint document agent *Mountain View, CA*

- **RL Reward Signal & Evaluation Methodology**
  - Diagnosed a corrupt reward signal in the document agent's RL training loop: single-turn evaluation was structurally blind to turn-level error propagation, tool-call convergence failure, and stale-context degradation. Designed a multi-turn *simulation* methodology to fix it, replaying real agent trajectories against live tool execution in a controlled environment to validate reward-signal quality before it enters training.
  - Built *LLM-as-judge* reward grading for multi-turn agent trajectories.
  - Built a document-render-to-image visual grading pipeline; `ETag/IfNotModified` optimistic concurrency across parallel graders eliminates duplicate rewards at trajectory scale.
  - Drove an org-wide checkpoint pre-screening gate: RL candidates simulated before end-to-end integration, eliminating wasted compute on non-viable checkpoints; simulation-grounded comparisons directly informed which checkpoints advanced to full RL training.
  - Established the first quantified per-slide latency profile (33 to 75s/slide) by isolating LLM reasoning (7–30s) from tool execution (5–8s/turn) under a 5-turn cap, enabling model-versus-tool regression attribution invisible to end-to-end metrics.
- **Multimodal RL Environment & Tool Execution**
  - Integrated an image-generation engine as an agent-callable tool in the RL rollout, extending the document agent's action space to a new modality.
  - Owned the full tool-execution stack for RL training on the document agent: client-to-daemon architecture, per-trajectory VM provisioning, and an asynchronous multimodal pipeline. Scaled to 7,000 containers while holding SLOs (P50 under 60s, P95 under 120s, P99 under 180s), unblocking concurrent RL pipelines across teams.
  - Unified Anthropic Claude, Azure OpenAI (incl. O-series reasoning models), and internal production serving endpoints behind a common provider-adapter interface for the rollout/eval loop.
- **Production Platform, Scale & Observability**
  - Engineered the platform running the simulation at scale: nginx-balanced Python orchestration tier over a .NET/COM PowerPoint execution tier, Redis metadata store + distributed queue, artifact storage, and a cluster management UI/API; deployed as a 20+ container system on ACI with per-service health checks.
  - Achieved **5.7x throughput** (0.70 to 4.01 files/s) and **7.8x worker utilization** (12% to 94%) by resolving a three-stage bottleneck chain: base64 upload elimination (+73%), COM-pool redistribution (util 17% to 72%), and CPU quotas to prevent cache thrashing (+24%). 200-file batch 285s to 49.8s. Fleet projection: 100 VMs at 301 files/s, approx. 1M files/hr.
  - Resolved two RL-scale blockers (platform throttling at 1.4M+ requests/10 min; VM-lifecycle resource leak blocking 20k-VM training+grading parallelism); instrumented end-to-end with OpenTelemetry/Jaeger tracing that makes multi-turn eval trajectories attributable per turn (queue wait, worker execution, and artifact emission), alongside Prometheus + Grafana.

- **ML Platform & Fine-Tuning Data Pipeline**
  - Authored a model-first platform architecture (ingestion, grading/tagging, eval, observability) adopted as org blueprint; adjacent teams directed to use rather than rebuild.
  - Designed a multi-turn SFT data strategy: three-phase curation (programmatic tagging to expert audit to automation), full training schema, DAG execution optimization cutting agent latency from approx. 100s to 37s.

VMware Inc. 07/2020 – 05/2022
Member of Technical Staff III *Palo Alto, CA*

- **Distributed File System: Storage Layer**
  - Designed and implemented distributed FSCK for post-crash metadata repair: cross-consistency matrices across B+ trees, bitmaps, and segment usage tables.
  - Built cloud-native microservice for on-demand/scheduled integrity validation of the file system k-v store; extended to snapshot, unmap, and segment-cleaning checks.
  - Implemented aggregated snapshot capacity statistics collection from scratch; p99 latency 30ms.

Distributed Systems Lab, Stony Brook University 06/2019 – 05/2020
Graduate Research Assistant, Prof. Shuai Mu *EuroSys 2022*

- **Rolis: Replicated Multi-core Transactional Database Engine**
  - Implemented core engine mechanisms (see Publications): async multi-process Paxos replication and multi-core log truncation for checkpointing. Profiled CPU/heap/memory via gperftools and mutrace; built a header-only C++ logging library with aligned allocators and auto-flush.

Talentica Software 04/2016 – 01/2019
Senior Software Engineer *Pune, India*

- Built real-time ML inference pipeline (Apache Storm, Kafka, Cassandra, AWS) for BLE indoor location prediction; feedback loop improved accuracy by 95% and a live calibration portal improved site correctness by 70%.

MediaTek 08/2014 – 04/2016
Software Engineer *Noida, India*

- Audio module (C/C++): reduced playlist overload by 25%, integrated BT stack into MMI layer, authored GPS-WiFi-BT combo stress-test tooling.

Publications & Patents

- **Rolis: a software approach to efficiently replicating multi-core transactions [EuroSys 2022]:** Designed and implemented core mechanisms for a distributed multi-core transactional database engine: asynchronous Paxos replication with minimal throughput loss, multi-core log truncation for checkpointing enabling read-only transactions at 10M ops/sec, and a replay protocol guaranteeing serializability across in-memory and disk logs.
- **Verification of metadata consistency across snapshot COW B+ tree logical maps [US11573860B1]:** Methodology for verifying consistency of snapshot metadata across a hierarchy of ordered data structures; covers cross-tree node identification, consistency matrix construction, and repair logic.
- **Learning to Fingerprint the Latent Structure in Question Articulation [publication]:** Demonstrated that latent question patterns can be modeled as a cost-function-maximizing system approximated by a trained neural auto-encoder.

Education

- **Stony Brook University:** M.S. in Computer Science (2019–2020)
- **NIT Bhopal:** B.Tech. in Computer Science & Engineering (2010–2014)