



Pratik Wadekar
Computer Science & Engineering
Indian Institute of Technology Bombay

24M0843
M.Tech.
Gender: Male
DOB: 18/12/2000

Examination	University	Institute	Year	CPI / % Credits
Post Graduation	IIT Bombay	IIT Bombay	2027	8.85 Core: 12 Total: 54
Graduation	Shivaji University	RIT, Rajaramnagar	2023	8.98
Graduation Specialization: Computer Engineering				
Intermediate	Maharashtra State Board	The New College, Kolhapur	2019	85.85%
Matriculation	Maharashtra State Board	New English School, Bahireshtar	2017	95.80%

M.TECH THESIS & SEMINAR

- Workload-Aware Index Selection for PostgreSQL** (May'26–Present)
(M.Tech Project | Guide: Prof. Suraj Shetiya)
 - Developed a custom **PostgreSQL extension** using **Executor Hooks** and shared-memory hash tables to collect column-level UPDATE statistics, achieving **<2% runtime overhead** across 50K concurrent DML queries.
 - Designed a **B-tree write-penalty model** using planner costs and UPDATE statistics for index update overhead.
 - Implemented the **ICDE'19 recursive multi-attribute index selection** algorithm in PostgreSQL via **HypoPG**.
 - Integrating the cost model into an **index advisor framework**, balancing query performance and update costs.
- Automatic Index Selection & Database Tuning** (Jan'26–May'26)
(M.Tech Seminar | Guide: Prof. Om Damani)
 - Surveyed **automatic index selection**, comparing heuristic, ILP, and what-if approaches and runtime trade-offs.
 - Investigated workload-aware database optimization and **self-driving DBMS** adapting to workload forecasts.
 - Evaluated hypothetical indexes using **HypoPG** to estimate workload cost prior to physical deployment.

WORK EXPERIENCE & INTERNSHIPS

- Associate Developer (AWS/DevOps) | Ignitiv Technologies Pvt. Ltd.** (Jul'23–Jul'24)
 - Automated application deployments via **Bitbucket CI/CD pipelines**, reducing deployment time by **70%**.
 - Implemented **IAM-based access control** using **least-privilege policies** to secure multi-team access.
 - Resolved **PostgreSQL bottlenecks** via worker configuration tuning, reducing load times from **5m to 10s**.
 - Configured automated **S3 backup mechanisms** and lifecycle policies to ensure zero-loss data recovery.
- Developer Intern | KONE** (Jan'23–Jun'23)
 - Designed an **AWS resource monitoring pipeline**, reducing deletion identification time from **hours to 10s**.
 - Developed serverless logic in **Python and AWS Lambda** to parse **CloudTrail events** and dispatch SNS alerts.
 - Automated serverless deployment via **AWS SAM and CloudFormation** with secure environment variables.

KEY PROJECTS

- Bare-Metal KVM Hypervisor & Virtualization Engine** (Jan'25–Apr'25)
(Course Project | CS 695: Topics in Virtualization and Cloud Computing | Instructor: Prof. Purushottam Kulkarni)
 - Engineered a userspace hypervisor using **Linux KVM API** and **mmap** to manage isolated guest vCPU contexts.
 - Implemented guest-to-host virtual address translation by handling **KVM I/O-port exit traps** in userspace.
 - Orchestrated time-sliced **vCPU scheduling** using a circular buffer to synchronize **Producer/Consumer VMs**.
- Conductor: Docker-Lite Container Runtime & Image Engine** (Jan'25–Apr'25)
(Course Project | CS 695: Topics in Virtualization and Cloud Computing | Instructor: Prof. Purushottam Kulkarni)
 - Engineered a container runtime using **namespaces** and **cgroups v2**, achieving **73x faster startup than Docker**.
 - Accelerated image builds by **85x** via **SHA-256 content-addressable storage** and **OverlayFS** layer caching.
 - Implemented container networking using **veth pairs** and **iptables** routing for a C/Flask microservice.
- Attention Mechanisms & LLM Inference Optimization** (Jan'26–Apr'26)
(Course Project | CS 794: Systems for Machine Learning | Instructor: Prof. Mythili Vutukuru)
 - Accelerated the **decode phase** by **29.1x GPU and 9.8x CPU** using **KV caching** to avoid redundant computation.
 - Designed a **radix-tree shared-prefix KV cache** to eliminate redundant **prefill**, gaining a **1.25x GPU speedup**.
 - Developed **FlashAttention in CUDA** with **online-softmax tiling**, scaling VRAM footprint from $O(N^2)$ to $O(N)$.

OTHER PROJECTS

- **CRISP: Critical Slice Prefetching** (Aug'25–Nov'25)
(Course Project | CS 683: Advanced Architecture | Instructor: Prof. Biswabandan Panda)
 - Architected critical-slice prefetching in **ChampSim** to identify and prioritize latency-critical dependency chains.
 - Redesigned the OoO scheduler with **priority issue queues** to issue critical slices while preserving ROB ordering.
 - Improved Limit Order Book matching engine IPC by **4.98%** evaluated on Intel PIN-generated traces.
- **Preemptive Multitasking OS Kernel** (Aug'24–Nov'24)
(Course Project | CS 744: Design and Engineering of Computing Systems | Instructor: Prof. Purushottam Kulkarni)
 - Built a **64-bit x86_64 Rust kernel**, configuring the **GDT** and **IDT** for interrupt and exception handling.
 - Implemented **physical frame allocation** and **heap allocation** using bootloader-provided memory mappings.
 - Orchestrated **Round-Robin multitasking** via **PCBs**, context switching, and **Mutex/Semaphore** synchronization.
 - Developed a **VGA driver**, an **interactive shell**, and a **hierarchical in-memory VFS** with **CRUD support**.
- **MyRedis: In-Memory Key-Value Database** (Jan'26–Apr'26)
(Self Project | C++, Linux Sockets, Multithreading)
 - Engineered a **non-blocking event-driven Redis-like server** using **multiplexed I/O** and TCP pipelining.
 - Designed an **incremental-rehashing key-value store** and **dual-indexed sorted sets** for $O(\log N)$ range queries.
 - Implemented **active TTL key eviction**, **LRU connection pruning**, and asynchronous memory cleanup.
- **Linux Kernel Modules & Character Drivers** (Jan'25–Apr'25)
(Course Project | CS 695: Topics in Virtualization and Cloud Computing | Instructor: Prof. Purushottam Kulkarni)
 - Implemented 5-level page walks for VA-to-PA translation and **2 MiB Transparent Huge Page (THP)** accounting.
 - Developed character drivers using **ioctl interfaces** for physical-memory writes and process tree reparenting.
 - Exposed page-fault & memory statistics via **procfs/sysfs**, comparing **$O(1)$ RSS accounting** with $O(N)$ page walks.
- **RAG Documentation Assistant: Hybrid Search & Reranking Pipeline** (May'26–Jun'26)
(Self Project | Python, PostgreSQL, pgvector, Gemini API)
 - Built an ingestion pipeline chunking technical docs into PostgreSQL using **pgvector (HNSW)** and GIN indexes.
 - Designed **hybrid retrieval** combining vector search and full-text search with **cross-encoder reranking**.
 - Generated **citation-grounded LLM responses** with source references to reduce hallucinations.
- **GPU-Optimized Matrix Multiplication & MLP Forward Pass** (Jan'26–Apr'26)
(Course Project | CS 794: Systems for Machine Learning | Instructor: Prof. Mythili Vutukuru)
 - Profiled **MLP forward-pass performance** across SIMD, cache tiling, prefetching, and CUDA shared memory.
 - Applied **cache tiling** (26.1x) and **AVX2 SIMD** (55.8x) to CPU matmul, reducing L1/L2 cache misses.
 - Deployed **shared-memory tiling** in CUDA matmul (1.7x speedup), boosting GPU utilization from **13.6% to 64.5%**.

POSITIONS OF RESPONSIBILITY

- **System Administrator | CSE, IIT Bombay** (Jul'24–Present)
(Linux Server Management | LDAP, DNS, NFS Infrastructure)
 - Managed **Linux servers, LDAP, DNS, NFS, and mail infrastructure**, providing core lab IT services for **350+ students** daily and **700+ candidates** during department-wide lab examinations.
 - Migrated legacy **PHP inventory to GLPI**, integrating **1,000+ LDAP users** and importing **500+ asset records**.
 - Created **Ansible playbooks** for automated user management, software installation, and **iptables** access control.

TECHNICAL SKILLS

- **Languages:** C++, C, Rust, Python, Bash / Shell Scripting, CUDA, SQL
- **Core CS:** Data Structures & Algorithms (DSA), Operating Systems (OS), DBMS, Computer Networks
- **Cloud, DevOps & Tools:** AWS (EC2, Lambda, Amplify), Docker, Kubernetes, CI/CD Pipelines, Linux perf, Git

COURSES TAKEN

- Design and Engineering of Computing Systems
- Advanced Architecture
- Design & Re-engineering of OOP
- Software Lab
- Topics in Virtualization and Cloud Computing
- Systems For Machine Learning
- Blockchains, Cryptocurrencies & Smart Contracts
- Web Security

COMPETITIVE PROGRAMMING & INTERESTS

- **LeetCode:** Peak 2014 (Knight, Top 3%, 950+ solved) | **Codeforces:** Peak 1366 (Pupil, 250+ solved).
- **Hobbies & Interests:** Trekking, Playing Cricket, Fitness & Gym, Listening to Music.

CERTIFICATIONS

- **Machine Learning Specialization** by DeepLearning.AI & Stanford Online on Coursera.
- **Deep Learning Specialization** by DeepLearning.AI on Coursera.