Prasad Chandrashekhar Kanade

+1 6174125390 | kanade.pra@northeastern.edu | linkedin/prasad-kanade | github/prasad0411 | portfolio-website | Boston, MA

EDUCATION

Masters of Science in Computer Science *September 2025 - May 2027*
Northeastern University, Boston, USA
Courses: Data Analysis, Database Management Systems, Distributed Systems, Algorithms, Machine Learning

Bachelor of Technology in Computer Science and Engineering *October 2020 - November 2023*
Maharashtra Institute of Technology - World Peace University, Pune, India
Courses: Data Structures, Probability & Statistics, Software Engineering, Computer Networks, Artificial Intelligence

TECHNICAL SKILLS

- **Languages:** Python, SQL, Shell/Bash, R, Java, JavaScript, C++
- **Data Engineering:** Apache Spark (PySpark), Apache Airflow, DBT, ETL/ELT Pipelines, Data Modeling, Data Warehousing, Data Cleaning, Data Wrangling, REST APIs
- **Visualization & BI:** Tableau, Power BI, Matplotlib, Seaborn, Dashboard Development, Reporting, Google Sheets API
- **Databases:** PostgreSQL, MySQL, Oracle, MongoDB, Firestore, NoSQL
- **Cloud & DevOps:** AWS (EC2, ECR, ECS, SageMaker), Docker, Kubernetes, Git, Linux/Unix, CI/CD, Agile/Scrum
- **Libraries & Frameworks:** Pandas, NumPy, scikit-learn, SHAP, Selenium, Flask, FastAPI, REST APIs, Spring Boot

WORK EXPERIENCE

Amdocs Corporation *December 2023 - May 2025*
Software Engineering Associate Pune, India

- Designed and deployed ETL pipelines using Spring Boot and Oracle for JCOM's telecom platform, ingesting and transforming 2M+ daily subscriber events across 10M+ accounts for analytics, reporting, and business intelligence.
- Optimized Oracle database queries powering operational dashboards by analyzing execution plans and implementing composite indexing, reducing data retrieval latency by 60% and eliminating 15 hours/week of processing delays.
- Implemented automated data validation and anomaly detection workflows using Python and Shell scripts on Linux servers, proactively catching pipeline quality issues early and cutting production data incidents by 30%.

ISKCON Organization *July 2022 - December 2022*
Software Engineer Intern Pune, India

- Streamlined infrastructure monitoring using Bash scripts and cron jobs on Linux/Unix servers, implementing alerting for system metrics (CPU, memory, disk), saving 4+ hours weekly and reducing unplanned downtime by 25%.
- Deployed cloud-native applications on AWS EC2 with Docker containers via ECR/ECS orchestration, building automated pipelines that reduced deployment time by 78% and maintaining 93% service uptime.

Simba Developers Organization *March 2021 - October 2021*
Software Developer Intern Pune, India

- Architected distributed microservices using Node.js and Express with PostgreSQL for authentication and payment data modules, enabling independent scaling and lowering failures by 65% through fault-tolerant design patterns.

PROJECTS

Automated Job Data Pipeline *GitHub* *January 2026 - March 2026*

- Designed an ETL pipeline extracting 8,000+ job postings weekly from 6+ sources, applying a 12-stage validation chain across title, location, and sponsorship signals, loading structured records into Google Sheets with 95% accuracy.
- Built a persistent data quality layer across 247+ domains using DNS-cached MX inference, city-to-state normalization, and Bayesian pattern scoring — reducing bad-data rate by 85% and eliminating manual corrections.
- Implemented source quality monitoring with 7-run rolling average decay detection and per-API ROI tracking; automated alerts trigger when valid-rate drops below 50% of baseline, enabling proactive pipeline maintenance.
- Engineered a circuit-breaker system with bounce rate thresholds at 15% and 30%, per-domain confidence scoring, and automatic pattern retry via Microsoft Graph API — replacing static blacklists with a self-correcting delivery pipeline.

Thyroid Disease Classification System *GitHub* | Springer Paper *December 2024 - June 2025*

- Developed a classification model achieving 97.6% accuracy across 7,200 patient records for 3 thyroid conditions; applied SMOTE and feature engineering (19→12 features) to improve minority recall from 68% to 93%.
- Created an explainable AI system using the SHAP framework for model interpretability, quantifying individual feature contributions (65% combined importance for TSH, T3, T4) for transparent, data-driven clinical decision-making.