Prinjal Dave

San Jose, CA | dave.pri@northeastern.edu | +1 (669) 308-1017 | https://prinjal.github.io | linkedin.com/in/prinjal-dave

EDUCATION

Master of Science in Computer Science   **Apr 2024**
Northeastern University • 3.9 GPA   San Jose, CA

Bachelor of Information and Communication Technology   **May 2018**
L.J. University • 3.8 GPA   Ahmedabad, GJ

SKILLS

- **Programming Languages:** Python, Java, R, JavaScript, HTML, CSS, PL/SQL
- **Frameworks:** Apache Spark, Airflow, MapReduce, Hive, NumPy, Pandas, TensorFlow, Matplotlib, Scikit-Learn, Spring Boot
- **Databases:** PostgreSQL, MySQL, Oracle 11g, AWS Redshift, Redis, SQLite
- **Tools & Technologies:** Node.js, Terraform, Docker, Kafka, Kubernetes, Kibana, Git, Jenkins, Tableau, Jupyter

EXPERIENCE

Addgene   **Jan 2023 - Aug 2023**
Data Engineer Co-Op   Boston, MA
- Orchestrated orders ETL data pipeline with Airflow, executing Spark jobs to ingest data in BigQuery, achieving 33% faster runtime.
- Boosted plasmid retrieval with optimized SQL/Django queries using CTEs and indexing, enhancing PostgreSQL efficiency by 40%.
- Integrated NCBI's PubMed data, employing PCA for feature engineering, yielding 2.5x quicker training and a 2% accuracy boost.
- Upgraded UPS shipment pipeline from credential-based to OAuth 2.0 with bearer tokens, reducing identified security risk by 10%.

KPMG   **May 2021 - Dec 2021**
Software Engineer, Data Platform   Bengaluru, KA
- Facilitated the implementation of a CI/CD pipeline for running payroll transaction microservice on k8s using GitHub Actions, resulting in 40% fewer job failures, enhancing SLA reliability and data quality.
- Dockerized a data processing framework to gather user website activity, generating metrics displayed on the Kibana dashboard.

Flip Robo Technologies   **Oct 2020 - May 2021**
Machine Learning Engineer Intern   Bengaluru, KA
- Collaborated with 3 interns to deploy an AI loan model on AWS SageMaker for Risk Assessment, enhancing the presales demos.
- Built pre-processing and exploratory analysis framework, achieving 89% accuracy and 0.83 AUC in loan default prediction.
- Automated Selenium-driven data extraction from top news outlets, enriching News Aggregation with diverse, real-time content.

Evosys   **Jan 2018 - Oct 2020**
Software Engineer   Ahmedabad, GJ
- Architected a scalable Oracle HRMS solution on OCI's Kubernetes Engine using Spring Boot microservices and Hibernate, successfully scaling to manage 14k+ employees while achieving three nines uptime.
- Led on-site support for a $5B+ media conglomerate in Qatar, overseeing initial 5k+ staff appraisal cycle post-system implementation.
- Managed a team of 5 junior engineers and testers to develop EvoGlide, a Data Migration tool for migrating legacy HRMS data onto the cloud platform, achieving 75% automation and reducing migration time by 30%.

PROJECTS

GitTone
- Developed and hosted a Java-based sentiment analysis tool on AWS EMR, efficiently processing 92.53 GB of GitHub pull requests.
- Enhanced a two-phase MapReduce job with Hive and HBase, integrating AWS Kinesis for streamlined real-time data handling, employing a custom comparator and partitioner, enhancing data sorting and partitioning efficiency by over 60%.
- Analyzed programming languages over time, identifying key negative comment patterns, ready for terabyte-scale data expansion.

Twitter Clone
- Deployed Node.js REST APIs on AWS EBS for tweet actions, user interactions, admin features with 45+ TDD-driven commits.
- Implemented secure login with cookies, optimized CRUD operations using DAO design pattern, boosting data efficiency by 30%.
- Cut deployment time by 73%, conducted tests ensuring 99% automated test coverage for improved application quality.