PRACHI MEHTA

Skilled backend engineer with over 4 years of experiences in developing, deploying, and architecting highly scalable data platforms and distributed systems to deliver impactful projects at LinkedIn.

prachi.mehta@gmail.com · (646) 642-1350 · www.linkedin.com/in/prachinmehta

TECHNICAL SKILLS

Languages & Methodologies: Java, Python, Scala, SQL, YAML, JQ, OOP, R Programming

Frameworks & Platforms: Airflow, ETLs, HDFS, Hadoop, Presto, Samza, Spark, Kafka, EspressoDB, Distributed Systems

Technologies & Tools: Azure cloud resources (ADF, ADL, BlobStorage, CosmosDB, EventHubs, KeyVault), Bash, Git, Jira, Jupyter, Postman

INDUSTRY EXPERIENCES

LINKEDIN, SUNNYVALE CA — SOFTWARE ENGINEER, MARCH 2020 - OCTOBER 2023

- Developed the next-generation data platform on Azure that hosts LinkedIn's Economic Graph data to enable privacy compliant integrations with third-parties and Microsoft products like Outlook365.
- Integrated new datasets into the data pipeline, employing configuration-driven ETLs involving batch and real-time data processing, utilizing Samza and Kafka-based streams.
- Led data validation procedures to guarantee data precision between on-premises data sources & Azure CosmosDB collections supported by alerting, leading to data accuracy exceeding 99.9% for users.
- Onboarded new hires to assimilate into the team and make them efficient engineers.
- Resolved all system alerts within 60 minutes per team's internal SLA and was an active contributor to team's on-call rotation.

LINKEDIN, SUNNYVALE CA — SOFTWARE ENGINEER INTERN

JUNE - SEPTEMBER 2019

- Re-architected steps to the data transfer pipeline to build offline and nearline pipelines for transferring and processing datasets into Azure CosmosDB from LinkedIn on-premises datasets.
- Spearheaded migration effort for facilitating future batch data transfers through manual verification.
- Developed Python scripts to automate team reports, featuring data accuracy percentages and nearline stream latency data, reducing developer workload and enhancing data quality visibility during team on-call rotations, which reduced developer cost by 97%.

JUNE - SEPTEMBER 2018

- Designed and implemented SQL indexing algorithm to enhance categorization of ad-hoc dashboarding solutions, reducing load time of landing page by 75% and enabled 100% of LinkedIn developers to reach data insights faster, which substantially boosted engineering productivity.
- Delivered design documents and presented indexed metrics reporting tool to organization leaders and sister teams.

GAINE SOLUTIONS, SAN LUIS OBISPO, CA — DATA INTERN, NOVEMBER 2017 - APRIL 2018

- Refined data quality within Gaine Healthcare's enterprise dataset by applying rules to detect anomalies, which enhanced the overall data accuracy and increased efficiency for health care providers.

PLATFORM9 SYSTEMS, SUNNYVALE CA — SOFTWARE ENGINEER INTERN, JUNE - SEPT 2017

- Implemented web platform for customers to provision their servers on the Platform9 management system using Node.js, AngularJS, and CockroachDB.
- Designed and implemented an intuitive web interface, offering customers an effortless experience to configure and oversee their resources, leading to improved user satisfaction and streamlined server management.

EDUCATION

BS IN COMPUTER SCIENCE & MINOR IN STATISTICS, CALIFORNIA POLYTECHNIC STATE UNIVERSITY - SAN LUIS OBISPO, 2019

System Programming, Computer Architecture, Design & Analysis of Algorithm, Fundamentals of computer Science, Data Structures and Algorithms, Intro to Computing (Security), Operating Systems, Compilers, Theory of Computation, Discrete Structures, Linear Analysis