

# ABHIJIT CHAKRABORTY

United States

4089212573

abhijit.chak@gmail.com

abhijitchak (LinkedIn)

## EXPERIENCE

**MongoDB** 06/2022 – present  
Palo Alto, CA Principle Solutions Architect - Partner

- Design scalable and efficient machine learning and generative AI architectures that integrate seamlessly with our database solutions.
- Develop and propose end-to-end solutions that address complex business challenges using advanced ML and AI techniques.
- Ensure the architectural integrity and scalability of solutions, aligning with industry best practices and emerging technologies.
- Collaborate with clients to understand their business needs, technical requirements, and objectives.
- Present and demonstrate ML and AI solutions to stakeholders, articulating the value and potential impact on their business.
- Provide technical guidance and support throughout the project lifecycle, from initial consultation to deployment and post-implementation.
- Lead the development and deployment of ML models and generative AI applications, ensuring optimal performance and accuracy.
- Work closely with data engineers to ensure proper data collection, preprocessing, and integration with ML pipelines.
- Utilize our database technologies to optimize data storage, retrieval, and processing for ML applications.
- Partner with cross-functional teams, including data scientists, software engineers, and product managers, to drive project success.
- Mentor and guide junior team members, fostering a collaborative and knowledge-sharing environment.
- Stay abreast of the latest advancements in ML, AI, and database technologies, and incorporate new techniques and tools into solutions.
- Perform rigorous testing and validation of ML models and AI systems to ensure reliability and robustness.
- Continuously monitor and fine-tune solutions to enhance performance, scalability, and accuracy.
- Implement best practices for model versioning, documentation, and reproducibility.

**Oracle** 06/2015 – 05/2022  
Redwood City, CA Principle Data Engineer

- Automated ETL process across millions of rows of data which reduced manual workload by 33% monthly.
- Ingested data from disparate sources using a combination of SQL, Oracle Analytics API using Python to create views to be used in BI tools like OBIEE, Tableau.
- Designed & implemented BI data warehouse for the division, segregating the adhoc querying through Data Vault modeling while customizing Kimball modeling for reporting.

## SKILLS

**SCRIPTS** : SQL, SHELL, PYTHON  
**VISUALIZATION** : Tableau, Oracle Analytics Cloud, D3.JS.OBIEE  
**DB** : Oracle, Vertica, Teradata, MongoDB  
**VERSION CONTROL** : SVN, GIT  
**PROJECT MANAGEMENT** : JIRA, PRIMAVERA, INSTANTIS  
**CONTENT MANAGEMENT** : Confluence  
**MESSAGING** : Oracle Advance Queue Stream, Java Messaging Services, KAFKA  
**ETL/ELT** : Oracle Data Integrator, AbInitio, Informatica Power Builder, Apache Nifi  
**Cloud** : Amazon, GCP, Oracle Cloud & AWS Infrastructure.  
**CRM** : Oracle Subscription Mgmt Cloud Fusion Cloud (mainly OSMC), RMCS/OIC/CPQ, Oracle R12 Service Contract. Oracle Order Management  
**Domain** : Manufacturing, Telecom, Finance & Leasing, Rail Road Supply Chain(Freight).

## CERTIFICATIONS

**Oracle Certified Associate** : Database Management  
**Oracle Certified Specialist** : Database Development and Optimization  
**Oracle Partner Network Certified Specialist** : BI Technologies

## EDUCATION

**Arizona State University**

**PhD** Computer Science (AI/ML)  
**Master of Science** in Computer Science (Big Data Systems)

**West Bengal University of Technology**  
Bachelors of Technology EE

- Leads the BI delivery effort through requirement gathering from stakeholders, documents in FDD,FSD, overseeing the junior engineers and also accessing the completeness of the work from data and best practices.
- Maintained data pipeline up time of 99.9% while ingesting transactional data from 5 different primary data sources using Oracle Advance Queues, JMS queue and Python.
- Demonstrated experience as a senior resource assisting in the definition of systems strategy, gathering and documenting business requirements, leading fit-gap analysis, as-is and to- be business process designs, conference room pilots (CRPs) functional configuration, testing, client user training
- Architected and developed real time reporting on leads pipeline data and closure using Oracle stream, Golden Gate and Oracle Data Integrator.
- Built tools using APEX to provide real-time data around international currency exchange, reducing latency by 12%. Implemented the user level and data level security at data base and reporting layer through Samuel token authentication and OATH services.
- Build the cloud based data integration framework to ingest the data from enterprise data platform (Vertica cloud database hosted on AWS) to Salesforce marketing cloud, using extensive AWS technologies and services like Lambda (Python) EC2, S3, SQS. Also design and build API integration between Informatica MDM and Salesforce Marketing cloud for capturing real time customer attribute updates. This application facilitates several automated campaign run by marketing business team. translate those business needs into actionable reports in Siebel CRM analytics saving 16 hours of manual work each work.
- Led a team of three members working offshores IDC. Designed the data model for BI warehouse reporting for Orange Wholesale Calls that scaled up from 100 to 115,000 daily.

**Union Pacific** 10/2014 – 05/2015

Omaha, NE Senior Data Engineer (Contract)

- Worked as a BI lead developer converting the cold fusion reports to Oracle Sales CRM reporting.
- Ingested logging (Freight) events data regarding railroad driver feedback, linesman logs and track logs generated in real-time and batch using Oracle Advance Queuing, change data capture in Oracle Golden Gate. with a throughput of 2000 events/seconds stored in Vertica.
- Built a front end UI using sparkline.JS and D3.JS to capture and monitor feedback in real time.

**Cisco** 05/2011 – 09/2014

San Jose, CA Lead Data Engineer

- Created internal scripts(library) to parse and reformat data external vendors like (AC Nielson, D&B), reducing the error rate in the data pipeline by 7%.
- Tuned query performance of the BI reports to improve the execution time by 10%.
- Built the balanced scorecard deriving data directly from transactional systems data using ELT method on column wise data capturing technique.
- Implemented the local language preferences in the reporting server providing data in local languages which improved the user conversion rate from manual excel based to BI by 7%.
- Designed and developed while leading a team for Oracle to Teradata migration. To improve the analytical cube query performance and reporting needs. Writing BTECH scripts and redeveloping the BI(Oracle) reports and dashboard to Tableau. Improved the ETL and data availability leading to more user adaptability and reducing maintenance cost of \$250,000 a year.

**British Telecom** 11/2007 – 04/2011

Cardiff, Wales Data Engineer (Contract)

- Built basic ETL that ingested transactional data and event data from web apps with 10,000 daily active users that saved over \$85000 annually in external vendor costs.
- Worked with the client to understand business needs and translate those business needs into actionable reports in Siebel CRM analytics saving 16 hours of manual work each work.
- Led a team of three members working offshores IDC.
- Designed the data model for BI warehouse reporting for Orange Wholesale Calls that scaled up from 0 to 115,000 daily active users.
- Led the migration from Business Objects reporting space to Oracle Business Intelligence and CRM resulting in an annual cost savings of \$750,000 and an increase in performance of 15%.
- Automated the ETL process PL/SQL scripts to Ab-Initio that improve the data availability, governance and increased reliability of customer by reducing the development efforts and timeline.