

AMEY SADANAND BHILEGAONKAR

☎ [480-616-3980](tel:480-616-3980) ✉ abhilega@asu.edu 🌐 github.com/ameygoes in [linkedin.com/in/amey-bhilegaonkar](https://www.linkedin.com/in/amey-bhilegaonkar)

SUMMARY

Expert in Data Lake and Cloud Solutions | Proficient in Python, Apache Spark, and Advanced SQL | Skilled in Data Security, Streaming, and DataOps | Strong Background in Data Integration, Migration Performance, and Optimization and Monitoring

WORK EXPERIENCE

BigCommerce

Data Engineer Intern

June 2023 - August 2023

- Spearheaded and managed a **large-scale Snowflake** data retrieval pipeline for efficient data warehousing.
- Implemented **logistic regression**, and predictive models, achieving **customer retention** prediction accuracy of 86%.
- Leveraged advanced **data-mining** techniques to process and analyze 3 million data points, extracting critical features for search indexing and ranking.

Publicis Sapient

Data Engineer - II

June 2019 - July 2022

- Engineered complex **ETL** pipelines using **Apache Spark**, optimizing data extraction, transformation, and loading from diverse sources, such as **Redshift**, **S3**, **Kinesis Streams**, and **Kafka**.
- Utilized distributed computing frameworks such as Apache Spark to manage **large-scale data processing** tasks on **NoSQL Cassandra DB**, enhancing performance and resource utilization by 15%.
- Utilized SQL for data extraction, querying, and validation, contributing to software testing and release processes.
- Utilized analytical software and database management systems to collect, manage, and analyze diverse datasets, ensuring data integrity and quality.
- Independently built data validation module to simplify and automate reduced operation toil by 15%.

DevOps Engineer - I

- Implemented Jenkins **CI/CD** pipelines for automated deployment, enhancing efficiency and reliable software delivery.
- Orchestrated **Docker** containerization pipeline and deployment on **AWS ECR**, optimizing resource utilization and enhancing scalability

PROJECTS

Data Engineering - Apache Airflow, Docker, Kafka, Spark

- Orchestrated the setup of **Apache Airflow**, Zookeeper, **Kafka**, and **Cassandra**, within **Docker** containers.
- Implemented a comprehensive data pipeline, integrating random user data generation, Kafka streaming, and Cassandra storage.

Search Engine for All file types - Opportunity Hackathon - Meta Sponsored

- Spearheaded **Elasticsearch** implementation for blazing-fast search responses, with millisecond response times.
- Converted and stored every file type data as **vector embeddings**, ensuring **low-latency search** capabilities.
- Led Python **FAST API** development, providing efficient data access and retrieval.
- Examined Machine Learning techniques such as BERT, OCR, ResNet50, and Image Captioning to parse Image features.

Data Pipeline for Email Automation

- Initiated and completed developing a robust email automation project, streamlining job application outreach and networking efforts with individuals of similar interests.
- Employed **RESTful** API integration to acquire and manage a comprehensive contact database from CRM tools, demonstrating prowess in **full-stack development** and **API design**.

EDUCATION

Arizona State University, Tempe, USA

Expected May 2024

Masters of Science in Computer Science ([GPA: 4/4](#))

Pune Institute of Computer Technology, Pune, India

May 2019

Bachelors of Engineering in Electronics and Telecommunications ([GPA: 8.78/10](#))

TECHNICAL SKILLS

Programming Languages : Python, Unix / Linux Scripting, Java
Cloud Platforms & Databases : GCP, AWS, Big Query, Cassandra, No SQL, SQL, PostgreSQL
Data Engineering : PySpark, Snowflake, Airflow, Spark, Kafka, Pandas, PySpark
DevOps / SRE : CI/CD, Git, Jenkins, Docker, Kubernetes
Certified Google Cloud Platform : [Associate Cloud Engineer](#).