

Abhin Gude

Phone: 980-267-7254 | Email: abhingude14@gmail.com | <https://abhingude.com/> | LinkedIn | Ashburn, VA

TECHNICAL SKILLS

Python, SQL, Java, Scala, C, C++, R, MySQL, PostgreSQL, DynamoDB, MongoDB, Oracle, SQLite, Apache Spark, Kafka, Hadoop, Hive, AWS (EC2, S3, Lambda, Glue, Athena, AWS Pipeline, Redshift, CloudWatch), GCP (BigQuery), Azure (Azure Data Factory, Data Lake, Synapse Analytics, DevOps, Blob Storage), Talend, Informatica, Pentaho, Tableau, PowerBI, Google Analytics, Scikit-Learn, TensorFlow, Keras, PyTorch, OpenCV, Numpy, Pandas, Matplotlib, Seaborn, NLTK, PyTorch, Docker, Github, Gitlab, Jenkins, Kubernetes, Databricks, Airflow, Jira

EXPERIENCE

- **Universal Service Administrative Company** Washington, DC
Data Engineer *Jan 2023 - Present*
 - Designed and implemented a cost-efficient data lake on AWS S3, utilizing lifecycle policies and Athena for efficient querying of Parquet data, reducing storage costs by 20% to support USAC's Affordable Connectivity Program.
 - Engineered and maintained scalable data pipelines using Python, AWS Lambda, S3, and Glue, streamlining data flow and boosting processing efficiency by 25%.
 - Developed serverless ETL pipelines using AWS Lambda, S3, and Amazon Redshift, enabling real-time streaming data processing via Amazon Kinesis and improving processing speeds by 35%.
 - Integrated AWS Step Functions and Apache Airflow for advanced workflow orchestration, automating processes and enhancing data pipeline reliability by 30%, significantly reducing operational errors.
 - Streamlined infrastructure provisioning with Terraform and CloudFormation, reducing manual intervention by 40% and deployment time by 50%, ensuring scalability for the growing Affordable Connectivity Program.
 - Collaborated with data scientists to develop a data architecture supporting machine learning models, ensuring data accuracy for reimbursement eligibility and financial disbursements.
- **MTA New York City Transit** New York, NY
Data Research Associate Intern *June 2022 - Aug 2022*
 - Designed and implemented an Azure Data Factory pipeline for the real-time extraction, transformation, and loading (ETL) of live bus and train data, resulting in a 35% reduction in manual data processing efforts.
 - Leveraged Azure Databricks to perform complex data transformations on large datasets, achieving a 40% reduction in processing time compared to traditional methods.
 - Utilized Kubernetes and Docker for containerizing data processing environments, enabling scalable and efficient management of data workflows.
 - Developed customized Python scripts to analyze and enhance subway and bus performance metrics, along with essential customer-centric indicators. Additionally, designed interactive dashboards in Azure Synapse Analytics to visualize key performance indicators (KPIs), resulting in a 30% improvement in operational efficiency and service reliability, and facilitating the publication of data for open access.
- **Bajaj Finserv Health Limited** Pune, MH, India
Data Scientist Intern *Jan 2021 - July 2021*
 - Conducted exhaustive customer data analysis to fuel strategic business development and targeted marketing initiatives, focusing on health insurance sales trends and customer segmentation for personalized outreach campaigns.
 - Revamped Search API functionality, implementing automated zero search results diagnosis using Python script, Microsoft LUIS and language processing API. Successfully slashed zero search results from 18.9% to an impressive 2.13% during the migration from Azure to Elastic search, ensuring enhanced user experience and search relevance.
 - Leveraged the power of Power BI and Google Analytics Dashboards for comprehensive insights into search personalization, health insurance sales dynamics, and doctor-patient booking trends, empowering data-driven decision-making and fostering agile business growth strategies.
 - Propelled user engagement and platform usage by introducing a cutting-edge medical articles search feature, resulting in a remarkable 29% surge in search bar hits and a notable 13.3% increase in customer acquisition, exemplifying a commitment to innovation and user satisfaction.

EDUCATION

- **New York University** New York, NY
Master of Science in Computer Science; GPA: 3.93 *Aug 2021 - May 2023*
- **SRM University, AP** Amaravati, India
Bachelor of Engineering in Computer Science; GPA: 3.82 *Aug 2017 - May 2021*
- **University Of Wisconsin - Madison** Madison, WI
Semester Exchange Student in Computer Science; GPA: 3.25 *Jan 2020 - May 2020*

PROJECTS

- **Recommendation System:** Advanced movie recommendation system using Apache Spark and Machine Learning. Collaborative filtering, content-based and hybrid recommendation techniques are implemented for personalisation.
- **Facial Snapchat Filters:** CNN model for facial landmark detection and alpha blending to fuse image and filters.
- **Hospital Management System:** Web application for hospital management using Streamlit and PostgreSQL.