

Aditya Bhat

New Brunswick, NJ | [LinkedIn](#) | [Github](#) | adityacvbat@gmail.com | +1 (732) 763-7131

SUMMARY

Machine Learning and Data Engineering Professional with over 3.5 years of experience, adept in Python, SQL, **Azure, AWS, Snowflake, Tableau**, specializing in real-time data processing, advanced analytics, interactive visualization and backend API development.

SKILLS

Language: Python, SQL, C++, C#, UNIX Shell scripting

Data Orchestration: Azure Databricks, Apache Airflow, Apache Kafka

Tools: Git, Docker, Unity3D

Libraries: PySpark, Flask, Numpy, Pandas, Matplotlib, scikit-learn.

Cloud: AWS S3, EC2,RDS, Azure Data Lake Storage

Database: Snowflake Postgres SQL, MS SQL Server, OracledB

ETL Tools: Azure data factory

Editors: Pycharm, Visual Studio Code.

EDUCATION

MSc In Computer Science | Rutgers University, New Brunswick, NJ | CGPA: 4.0

Sep 2021 - May 2023

BE In Information Science | BMS Institute of Technology Bangalore, India

Aug 2013 - Jun 2017

PERSONAL PROJECTS

Real-Time Food Ordering System Using Apache Kafka with ML-Powered Recommendation

- Hosted **Apache Kafka** and **Apache Zookeeper** locally on a Docker container to develop a real-time food ordering system, integrating HTML/CSS front-end for user-friendly order processes with efficient **RESTful API backend** communication.
- Developed **Kafka producers** send order details and payment information to distinct **Kafka topics**, incorporating a unique identifier to accurately match orders with payments.
- Designed **Kafka consumers** for an **automated email system** for order and payment confirmations and data-driven **analytics generation**, providing insights into customer preferences, peak ordering times, and popular menu items to guide strategic business decisions.
- Built **collaborative filtering recommender system** utilizing historical order data and customer preferences from Kafka streams, to personalize menu suggestions for enhanced customer engagement.

Real Estate Market Dynamics Analysis: A Snowflake and Tableau Integrated Project

- Executed extensive EDA and data cleaning on a housing dataset with Python's Pandas and efficiently uploaded it to AWS S3 using AWS CLI.
- Utilized **Snowpipe** in **Snowflake** for automated data ingestion from S3, with event-driven notifications and integration objects for real-time data access.
- Implemented **Snowflake's Streams** for efficient incremental data loading, coupled with advanced data modeling and robust security measures including role-based access control and secure data sharing protocols.
- Used **Tableau for storytelling**, creating a series of **dashboards** that guided viewers from market overviews to in-depth trend explorations, enhanced by interactive filters and geospatial mapping for an engaging and insightful narrative.

WORK EXPERIENCE

Nokia Bell Labs

New Jersey, US

Applied AI/ML Data Science Intern

Jun 2022 - Aug 2022

- Created and **incorporated APIs** for the efficient retrieval and management of report templates, improving the report generation system for **B2B clients** and facilitating the handling of varied data queries.
- Implemented server-side APIs** using **Flask** for easy data access, improving client engagement and custom report requests.
- Built a robust ETL pipeline for media agencies with advanced error handling and logging for in-depth data analysis.
- Generated reports in **Python** and loaded them into a **PostgreSQL** database, ensuring a scalable, flexible storage solution with easy query capabilities for precise marketing strategy adjustments.
- Engaged actively with the analytics team, by providing insightful reports, aiding clients in sharpening their marketing strategies and contributing to successful product launches.

Thinking Stack

Bangalore, India

AI/ML Data Engineer

Oct 2020 - Jun 2021

- Developed an **end-to-end ETL pipeline using Azure Data Factory** to process COVID-19 data, generating insightful reports for key clients including doctors and health officials, enabling them to make informed decisions.
- Implemented event-triggered pipelines, linked services and datasets for transferring data seamlessly from **Azure Blob to Azure Data Lake Storage Gen2**.
- Collaborated with clients to tailor COVID-19 reports, in **Azure Databricks using PySpark**, ensuring meaningful insights and actionable data.
- Migrated transformed reports to **Azure SQL** and leveraged **PowerBI** to craft dashboards, enhancing the Health Department's COVID-19 strategy and resource allocation.
- Developed an Apache Airflow DAG for an end-to-end pipeline, extracting and loading new papers from the National Library of Medicine to **AWS RDS, hosted on EC2**.
- Utilized Python scripts to download the latest papers to an **S3 bucket** and transformed XML data into a clean dataframe, automating the extraction and transformation process.
- Connected to **AWS RDS MySQL** database for iterative data loading and fully automated the ETL process, with daily execution controlled through a cron job in **Apache Airflow**, ensuring timely and consistent data updates.

DHS Informatics

Bangalore, India

Machine Learning Engineer

Feb 2020 - Jul 2020

- Developed 5 custom **YOLOv5** models for detecting safety PPE kit by capturing data from direct client location, and trained with up to 4 augmentation techniques achieving an overall mAP(.50) value of 0.81.
- Built an image augmentation pipeline that enhanced model performance through data processing and increased lower-class image counts, resulting in a 10% accuracy boost.
- Developed a POC for on-screen logo detection and analysis, leveraging YOLOv5, pandas, and matplotlib to assess attributes like screen time, size, and frequency.
- Constructed an interactive **Flask** interface for streamlined data visualization in advertisement logo analytics.

Oracle Financial Service Software

Bangalore, India

Technical Analyst-QA

Jun 2017 - Mar 2019

- Implement the OFSBD application on client's systems, providing guidance on product features and functionalities, and conducted knowledge transfer sessions.
- Personally investigated and documented product bugs for reporting through internal tools and facilitated meetings with cross-functional teams including development, project management, and QA to address client's service requests.
- Communicated effectively with clients on their issues on the product features and interacted with respective consultants to ensure **client's satisfaction.**
- Extracted and organized customer transaction data from Oracle Financial Crime and Compliance Management using complex SQL queries and joins.