

ANJU PADIA

Syracuse, NY | 📞 +1(315)-952-7534 | 🔗 [linkedin.com/in/anju-padia](https://www.linkedin.com/in/anju-padia) | ✉ apadia@syr.edu

EDUCATION

Syracuse University

Expected - May 2024

Master of Science in Computer Science (GPA: 3.73/4.0)

Syracuse, NY

- Coursework: Intro to Machine Learning, Data Structures and Algorithms, Natural Language Processing
- Teaching Assistant - CSE 283: Object Oriented Design

Fall 23

University of Mumbai - Thadomal Shahani Engineering College

Oct 2020

Bachelor of Engineering in Information Technology (GPA: 3.66/4.0)

Mumbai, India

- Coursework: DBMS, Cloud Computing, Data Structures, Web Development, Object-Oriented Design

TECHNICAL SKILLS

Web: React.js, Flask, Django, Angular JS

Databases: MySQL, NoSQL, Cassandra, Postgres

BigData Tools: Hadoop, Spark, Hive, dbt, Databricks

Visualization: PowerBI, Tableau, Looker

Languages: Python, Java, R, PHP, C/C++, Scala

Open Source: Apache Airflow, Kafka, Hbase

Cloud: AWS S3, Azure, GCP, Snowflake

Frameworks: Pandas, Pytorch, TensorFlow, Matplotlib

EXPERIENCE

Delaware North

Summer 2023

Data Engineer Intern

Buffalo, NY

- Designed and implemented modular Python code to address data completeness and integrity within an enterprise data platform which covered 90% of the locations
- Collaborated with cross-functional teams to analyze user data, fostering a 40% increase in user engagement
- Utilized SQL to query and manipulate large datasets, improving data accuracy by 20%
- Maintained and developed data pipelines for the extraction, transformation, cleaning, pre-processing, aggregation, and loading of data from diverse sources using Python, SQL, DBT, and other data technologies
- Conceptualized and developed data models, enhancing the efficiency of data pipeline processes.
- Conducted code testing, created test cases, and deployed software to prod servers for reliable data pipeline efficiency
- Leveraged data visualization techniques through Tableau to communicate analytical results effectively

Larsen and Toubro Infotech

Aug 2020 - Aug 2022

Data Engineer

Mumbai, India

- Explored, examined, and interpreted large volumes of structured and unstructured data to solve complex problems
- Contributed to the development of highly scalable Big Data solutions and ETL ecosystems, optimizing data processing
- Designed and modernized the Data Ingestion automation Framework for a client that loads data into a specific target warehouse such as BigQuery from various on-prem source systems such as Oracle, Teradata and HANA
- Created PySpark jobs for extracting data from RDBMS, file sources and load it into a data warehouse
- Wrote basic hive queries and Shell Scripts for automation and conducted in-depth data analyses using SQL and Python, ensuring accuracy and reliability of insights.
- Demonstrated proficiency in version control and CI/CD pipelines

PROJECTS

Global Health Data Analysis and Visualization | R, Tableau, Statistical Analysis

Oct 2023

- Conducted in-depth exploratory data analysis and statistical analysis using R, revealing significant trends and patterns in global health indicators.
- Developed interactive and visually compelling Tableau dashboards that effectively communicated complex global health trends to a diverse audience.

College Enrollment Data ETL System | Python, Pandas, Apache Airflow, SQL

Apr 2023

- Develop an automated ETL pipeline that could handle data from different sources and adapt to changes in enrollment.
- Reduced data inaccuracies and inconsistencies by 95%, leading to a 70% reduction in data-related errors in reports.

Machine Learning-Based Chatbot for Customer Support | NLTK, Apache Kafka, Apache Spark, Kubernetes

Aug 2022

- Developed a chatbot with advanced natural language understanding capabilities, allowing it to comprehend and respond to complex user queries effectively
- Reduced average response time to customer queries by 40%, leading to improved customer satisfaction

Sentiment Analysis of E-Commerce site product reviews | Web, SQL, Python, Machine Learning

Mar 2020

- Conducted statistical analyses, providing actionable recommendations to improve product features and user experience.
- Validated and executed algorithms and predictive models to detect patterns and recommend solutions