Piyush Badule
Data Engineer

✉ piyush.badule30@gmail.com

☎ +91- 9322479517

📍 Pune, Maharashtra

🔗 piyushbadule.github.io

🐙 github.com/piyushbadule

in linkedin.com/in/piyushbadule

AWS Certified Data Engineer and AWS Certified Developer with 5 years of experience, specializing in Python, Django, Flask, and FASTAPI for backend development. Proficient in PostgreSQL and AWS and AZURE cloud deployments, with expertise in data manipulation using Pandas, Numpy, PySpark, and Databricks. Skilled in managing data pipelines with Apache Airflow, ensuring efficient ETL processes, and experienced in version control with GIT and containerization with Docker.

🧠 Skills

Skill	Level
Python	●●●●●
AWS	●●●●○
Pandas & Numpy	●●●●○
Pysparks	●●●●○
FastApi	●●●●○
Flask	●●●●○
Django	●●●●○
HTML, CSS & JS	●●●○○
PostgreSQL	●●●●○
GIT	●●●●○
Airflow	●●●○○
Databricks	●●●○○
Microsoft AZURE	●●○○○

💼 Professional Experience

T - Systems ICT India Pvt Ltd, *Software Engineer*
Jun 2022 – present | Pune, India

∅ **OneBI- Lakehouse:**
- **Description:** OneBI is designed to provide a unified Business Intelligence (BI) solution for T-Systems, supporting both internal and external customers. The project focuses on consolidating data into a single enterprise data platform, enabling efficient data preparation, orchestration, and seamless integration of data for reporting, analytics, and planning. It supports business processes across multiple departments including finance, sales, and human resources. The OneBI platform leverages state-of-the-art BI tools and techniques to deliver reliable and scalable solutions that enhance decision-making and business performance.
- **Technologies:** Azure, DataBricks, DBT, PySparks, SQL
- **Responsibilities:**
  - **Data Engineer**: Worked on developing scalable data pipelines and automated data workflows using PySpark, Azure, and Databricks.
  - **Dashboard Development**: Created interactive dashboards and visualizations in Databricks to support business reporting and data analysis needs.
  - **Notebook Development**: Wrote and maintained PySpark notebooks for data ingestion, transformation, and orchestration across various data sources.
  - **Data Consolidation**: Integrated data from multiple sources into a centralized platform to ensure consistency and availability for reporting and analytics.
  - **Collaboration with Stakeholders**: Collaborated with teams across finance, sales, portfolio, and human resources to develop custom data solutions and ensure business needs were met.
  - **Consultation and Support**: Provided technical support and consultation for the development of predefined reports, dashboards, and semantic models.

∅ **Aqua Future - Mercedes-Benz (Europe)**
- **Description:** The Aqua Future project is a data engineering initiative undertaken by Mercedes-Benz for its European market. The project focuses on managing and analyzing post-purchase car data, including service history, part replacements, vendor interactions, and vehicle status such as accidents or damages. The primary goal was to build a robust and scalable data pipeline that could efficiently process large volumes of customer and vehicle data, enabling accurate reporting and insights into car maintenance and customer behavior.
- **Technologies:** Python, Pysparks, AWS, GIT.
- **Responsibilities:**
  - **Data Pipeline Development**: Designed and implemented an end-to-end data pipeline to handle the ingestion, transformation, and storage of large-scale datasets. This involved leveraging AWS services such as S3, Glue, Step Functions, and Athena to ensure high performance, scalability, and cost efficiency.
  - **Data Processing and Analysis**: Utilized PySpark to process and analyze complex datasets, ensuring data accuracy and integrity. Focused on optimizing data workflows to meet the project's performance requirements.
  - **Collaboration with Stakeholders**: Worked closely with stakeholders to understand data requirements and ensure that only the necessary data was provided for analysis and reporting, maintaining data privacy and efficiency.
  - **Scalability and Performance Tuning**: Implemented best practices for data pipeline scalability and performance tuning to accommodate the growing data needs of the project.
  - **AWS Integration**: Integrated various AWS services into the data pipeline to automate workflows, reduce manual intervention, and streamline data processing tasks.
  - **Support and Maintenance**: Provided ongoing support and maintenance for the data pipeline, addressing any issues related to data ingestion, transformation, or storage.

Education

Bachelor of Engineering - Electronics & Telecommunication,

G. H. Raisoni College of Engineering

2016 – 2019 | Nagpur

CGPA: 8.73

Diploma in Electronics & Telecommuniction,

Government Polytechnic, Nagpur

2013 – 2016 | Nagpur

Percentage: 72.31

Grade - X, *K. John Public School*

2012 – 2013 | Nagpur

CBSE
CGPA: 8.8

Certificates

- AWS Certified Data Engineer - Associate
- AWS Certified Developer - Associate

Awards

Certification Of Appreciation

Jun 2024

T - Systems ICT India Pvt. Ltd

Pat on the Back

Mar 2024

T - Systems ICT India Pvt. Ltd

Pat on the Back

Jan 2023

T - Systems ICT India Pvt. Ltd

Stay Curious & Grow

May 2023

T - Systems ICT India Pvt. Ltd

Languages

English ● ● ● ● ●

Hindi ● ● ● ● ●

Marathi ● ● ● ○ ○

Ø **Space-ECB (European Central Bank):**
- **Description:** The SPACE European Central Bank (ECB) development and Migration Project is a comprehensive initiative aimed at efficiently managing and analyzing economic data from various European countries. The project involves collecting, processing, and migrating data related to key economic indicators, including interest rates, income, GDP, and more. This data is organized as time series, and its availability varies based on different frequencies (D: Daily, M: Monthly, W: Weekly, Q: Quarterly, A: Annually).
- **Technologies:** Python, Object-Oriented Programming (OOP's), JupyterHub, Pandas, Numpy.
- **Responsibilities:**
  - **Script Migration:** Transition legacy FAME scripts to Python for enhanced efficiency and maintainability.
  - **Function Development:** Create custom Python functions for specific data processing needs.
  - **Quality Assurance:** Write test cases to ensure data processing accuracy and reliability and doing code reviews.
  - **Data Processing:** Clean, validate, and harmonize data to maintain quality and consistency.

Ø **DTSE Translation:**
- **Description:** DTSE Translation is an internal project to facilitate document translation in various languages. It encompasses translation services for text, images, websites, and live screen translation. The project supports translation of multiple images and documents, including embedded images within documents.
- **Technologies:** Python, Object-Oriented Programming (OOP's), FastApi, Docker, Azure.
- **Responsibilities:**
  - **Coding and Development:** Led the coding efforts alongside four backend colleagues, implementing text, image, website, and live screen translation functionalities using Python Fastapi and integrating Azure cognitive services.
  - **Code Review and Quality Assurance:** Conducted thorough code reviews, ensuring adherence to high coding standards and best practices.
  - **Unit Testing:** Developed comprehensive unit test cases to validate the functionality and robustness of the translation services.
    - **Documentation and Guidance:** Authored Docker files, test cases, and provided guidance to junior developers, fostering their growth and ensuring project consistency.

MindNerves Technology Services Pvt. Ltd., *Software Engineer*

Dec 2019 – Jun 2022 | Pune, India

Ø **Vistabee**
- **Description:** Vistabee is a digital marketing platform catering to real estate agents and businesses, enabling them to effectively convey their product or service narratives through compelling digital content. The platform empowers clients to create high-quality digital content that showcases their offerings.
- **Technologies:** Python, Object-Oriented Programming (OOP's), HTML, CSS, JavaScript, Django, PostgreSQL, AWS.
- **Responsibilities:**
  - **Designing Models, Views and Templates as per requirements.**
  - **Requirement Analysis:** Thoroughly understanding and analyzing project requirements, then translating them into code.
  - **Business Logic Development:** Developing the core business logic by comprehending the complete project flow and API development.
  - **Technology Expertise:** Proficiency in utilizing technologies like Celery and Django Debug Toolbar for efficient development.

Ø **Trumov**
- **Description:** Trumov is a versatile video creation platform designed to assist users in producing engaging videos tailored for various industries such as real estate, news, automotive, and more.
- **Technologies:** Python, Object-Oriented Programming (OOP's), HTML, CSS, JavaScript, Django, PostgreSQL, AWS.
- **Responsibilities:**
  - **Video, Audio, & Image Processing:** Spearheaded the implementation of critical video, audio, and image processing functionalities, enhancing the platform's multimedia capabilities.
  - **Email Configuration:** Successfully implemented email configuration to streamline communication and user engagement.

- **Responsive Design Resolution:** Collaborated with team members to resolve responsive design issues by leveraging HTML, CSS, and Python, ensuring a seamless user experience. Trumov's commitment to providing a user-friendly and feature-rich video creation solution empowers businesses across different sectors to create captivating video content for their specific needs.

Ø Neeri Job Portal
- **Description:** Neeri, a government organization, embarked onthe development of a web-based Job Portal dedicated torecruitment activities.
- **Technologies:**Python, Object-Oriented Programming (OOP's),HTML, CSS, JS, Django, PostgreSQL, AWS.
- **Responsibilities:**
  - Team Guidance, Established the foundational structure ofthe project, Spearheaded the end-to-end development