

Omkar Pathak

+918087996634 | omkarpathak27@gmail.com | [LinkedIn](#) | [Github](#) | [Personal Website](#)

EXPERIENCE

Lead Consultant

May 2022 – Present

Emergys Solutions Pvt. Ltd.

- Led the end-to-end design, development, and deployment of three enterprise-grade SaaS platforms:
 - * A Supply Chain Risk Intelligence platform enabling real-time risk detection and mitigation through live data pipelines and third-party integrations.
 - * A Cost Intelligence and Procurement Optimization system offering dynamic dashboards and insights into supplier performance, spend behavior, and contract efficiency.
 - * A Marketing Mix Modeling (MMM) platform built using Meta's Robyn and integrated with client data sources to measure channel-wise ROI, optimize media spend, and support data-driven marketing decisions.
- Architected cloud-native solutions on Microsoft Azure, leveraging services such as Azure App Services, Functions, Blob Storage, Key Vault, SQL Server DB, and Data Factory, ensuring secure, scalable, and high-availability infrastructure.
- Managed two agile teams (11 and 9 developers), leading sprint planning, architecture reviews, backlog grooming, and release management across projects.
- Built real-time data ingestion and transformation pipelines using Python, Azure Data Factory, Event Grid, and custom connectors, powering dynamic risk scores and operational insights.
- Developed modular microservices and REST APIs using .NET Core and FastAPI, enabling interoperability across systems and external clients.
- Collaborated with UI/UX teams to design and implement responsive interfaces using ReactJS, Tailwind CSS, and Figma, tailored for business users across supply chain, procurement, and marketing domains.
- Designed and delivered predictive analytics modules using LightGBM, Scikit-learn, and Time Series models for forecasting demand, detecting anomalies, and simulating test scenarios in the MMM and cost platforms.
- Conducted technical workshops on Python, ML workflows, and modern SDLC best practices, improving internal capability and knowledge sharing.
- Led development of multiple PoCs in areas such as cost anomaly detection, test design and market simulation, and agent-based resume screening, many of which evolved into long-term product features.
- Implemented DevOps practices using Azure Pipelines, Git, Docker, and automated test frameworks, resulting in faster deployment cycles and consistent code quality.
- Provided technical mentorship, code review leadership, and guidance on clean architecture, security best practices, and scalable design patterns across teams.

Python ETL Developer

Nov. 2020 – May 2022

Numerator

- Designed and implemented robust ETL pipelines using Apache Airflow to enable seamless data movement across diverse systems and ensure data integrity.
- Developed a scalable Airflow framework to ingest and transform data from emails, PostgreSQL, Oracle, and CSV files, with final integration into Snowflake data warehouse.
- Authored and maintained CI/CD pipelines using tools like Git, Docker, and Jenkins, facilitating smooth and reliable deployments across staging and production environments.
- Developed and maintained RESTful APIs to support data access and integration for Numerator's Data Lake products, ensuring high availability and performance.
- Collaborated with cross-functional teams to optimize data workflows, improve pipeline resilience, and ensure alignment with business analytics goals.

Data Engineer

July 2018 – Nov. 2020

Schlumberger

- Led the development of the Source to Hub pipeline, enabling direct data ingestion from source systems to Google BigQuery, eliminating intermediate Hadoop stages to enhance efficiency, reliability, and processing speed.
- Designed and implemented a fully automated CI/CD pipeline with custom validations for Informatica migration workflows—reducing deployment time from 9 hours to under 1.5 hours with zero manual intervention.

- Maintained and enhanced a custom data ingestion framework handling over 1TB of daily data across 70+ business units, ensuring consistency, auditing, and performance optimization.
- Collaborated with the L3 engineering team to deliver high-priority Scrum PBIs related to ingestion and transformation pipelines, ensuring adherence to sprint timelines and quality standards.
- Oversaw and executed QA and production release cycles, coordinating across teams to ensure stable deployments and minimal downtime.

Full Stack Python Developer Intern

June 2018 – July 2018

Truso

- Contributed to the development of a social media platform by building RESTful APIs and backend services using Django and Django REST Framework.
- Implemented user authentication, content handling, and API endpoints to support scalable interactions across the platform.
- Explored frontend development with Angular 5/6, collaborating with the UI team to integrate responsive components and optimize user experience.
- Gained hands-on experience in full-stack development within an agile startup environment, working closely with the founding tech team on rapid feature iterations.

Data Engineering Intern

Oct. 2017 – Jan. 2018

Propellus

- Developed automation scripts in Python to scrape structured and unstructured data from diverse web sources, enabling scalable data collection for startup analytics.
- Applied Natural Language Processing (NLP) techniques to extract named entities and key information from web articles using custom entity recognition pipelines.
- Built machine learning models to classify articles and tag startups with relevant themes based on their activities and descriptions.
- Implemented a hybrid approach using K-Nearest Neighbors (KNN) with Latent Semantic Analysis (LSA) to recommend meaningful tags and categories for startup profiles.

EDUCATION

MMCOE, Savitribai Phule Pune University

Pune, India

Bachelor of Engineering, Computer

June 2014 – May 2018

PROJECTS

Pyresparser | *Python, Docker, NLP, Model Training, SpaCy*

June 2018 – Present

- A simple resume parser used for extracting information from resumes
- Extract information from thousands of resumes in just a few seconds
- Author and maintainer of this project
- 450k+ downloads on Github

Pygorithm | *Python, DSA*

June 2016 – Present

- Author and maintainer of this project
- An educational library to learn all the major algorithms in native Python
- Got covered in Fosstack, FullStackFeed, Kleiber and Tagged under Hotest Github Project on ITCODEMONKEY
- 40k+ downloads on Github

TECHNICAL SKILLS

Programming Languages: Python, PHP, Shell Script

Frontend, Backend Technologies: Django, Flask, Streamlit, FastAPI

Databases: MySQL, PostgreSQL, SQLite, MongoDB, Snowflake

Developer Tools: CI/CD (Travis CI, Github Actions, Azure DevOps), Docker, Kubernetes, Apache Airflow, Git, NLP, Scikit-Learn, Cloud (GCP, Azure, DigitalOcean)

HONOURS AND AWARDS

High Flyer Award

Awarded for best quarterly performance

Emergys

Q4 2023

Winner at Paintball competition

Organized by Commerzone IT Park

Schlumberger

Jan 2020

Quora Top Writer

Awarded for active contribution for technical QnA

Quora

2018

Best Outgoing Student Award

Awarded for best overall performance in final year

MMCOE

2018