# Deepak Deokar

Long Beach, CA | +1 (562) 336-2862 | deepakdeokar2626@outlook.com | Portfolio | LinkedIn | GitHub

## SUMMARY

- Data Scientist with 1 year of experience in machine learning, predictive modeling, and AI fairness, specializing in Python, SQL, TensorFlow, Scikit-learn, & Flask. Skilled in time-series forecasting, anomaly detection, & bias mitigation, leveraging generative models (GANs) & Explainable AI (XAI) for improved model transparency. Experienced in building scalable ML pipelines, deploying models on AWS, & collaborating with cross-functional teams to drive data-driven solutions.

## EDUCATION

California State University Long Beach, CA          ***GPA - 3.8***
Master's in Information Systems          ***Graduated Mar 2025***

## SKILLS

- **Programming & Scripting:** Python, R, SQL, MATLAB, Linux
- **Machine Learning & Deep Learning:** TensorFlow, Keras, PyTorch, Scikit-learn, MLFlow
- **Data Science & Analytics:** Data Analysis, Statistics, Data Modeling, Feature Engineering, Pattern Recognition
- **Big Data & Cloud Computing:** Spark, Apache Kafka, Sagemaker, Lambda, Microsoft Azure
- **Version Control:** Git, GitHub
- **Data Engineering & ETL:** ETL, Apache Airflow, Data cleaning & preprocessing, Data Transformation
- **Data Visualization & Business Intelligence:** Power BI, Tableau, Matplotlib, Seaborn, MS-Excel
- **Databases & Data Manipulation:** NumPy, Pandas, Alteryx
- **Cloud:** AWS, Microsoft Azure
- **MLOps & Model Deployment:** Docker

## WORK EXPERIENCE

### *Programmer Analyst (Cognizant Tech Soln)*          *Jul 2022 – Aug 2023*

- Spearheaded an anomaly detection framework (DBSCAN, Isolation Forest, Prophet), minimizing sensor failures by 40% and preventing $100K+ in downtime losses annually.
- Optimized ETL pipelines, refining data ingestion for 10M+ time-series sensor records daily with Python (Pandas, NumPy), SQL, and Apache Airflow, boosting processing speed by 60%.
- Implemented a real-time analytics solution in Power BI, automating anomaly detection and trend monitoring, reducing manual data analysis efforts by 70%.
- Collaborated with ML, Data, and Cloud Engineers to enhance predictive analytics deployment on AWS SageMaker, accelerating model rollout by 50% and strengthening operational reliability.

### *Data Scientist Intern (AFour Technologies)*          *Dec 2021 – May 2022*

- Refined crew scheduling by 30% supporting the creation of ML-driven workforce planning models using linear programming and K-Means clustering.
- Accelerated data ingestion by 40%, processing 5M+ pilot scheduling records with Python, Pandas, and SQL, improving scheduling accuracy.
- Developed an interactive Flask-based visualization tool, enabling real-time insights into flight allocations, workload balancing, and scheduling conflicts, improving decision-making speed by 20%.
- Deployed scheduling models onto AWS SageMaker, reducing deployment time by 35% and improving operational efficiency.

## PROJECTS

### *Responsible AI Toolkit for Finance*

- Architected a Responsible AI Toolkit, incorporating GANs, AIF360, and OpenAI APIs, lowering bias in AI-driven financial models by 20%, aligning with ECOA compliance.
- Devised a Flask-based web application, incorporating automated bias detection, synthetic data generation and Explainable AI (XAI), enhancing AI model transparency by 35% for financial institutions.
- Built a GAN-powered synthetic data synthesizer in TensorFlow & Keras, rectifying demographic disparities, improving fairness scores by 30% and advancing ethical deployments.
- Embedded GPT for bias interpretation, delivering human-readable fairness insights, reducing bias misinterpretation by 40%, and making AI decisions 50% more explainable for non-technical stakeholders.

## CERTIFICATIONS

Alteryx Designer Core Certification 2023