

# Deepak Deokar

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

## SUMMARY

Entry-level Data Scientist with experience delivering machine learning solutions that enhanced operational efficiency, reduced manual work, and improved data-driven decision-making. Contributed to initiatives that resulted to a cumulative 30-50% improvement in key performance metrics, across analytics, automation, and deployment workflows. Focused on building responsible, transparent, and scalable systems that align with business goals.

## SKILLS

- **Programming & Scripting:** Python, R, SQL, MATLAB, Linux
- **Machine Learning & Analytics:** TensorFlow, Keras, PyTorch, Scikit-learn, Random Forest, XGBoost, LightGBM, Time-Series forecasting, Predictive modeling, Feature Engineering, Pattern Recognition, AI Fairness, Explainable AI, Statistical Analysis.
- **Data Engineering:** Pandas, NumPy, Apache Spark, ETL Pipelines, Apache Airflow, Data cleaning & Transformation
- **Visualization & BI Tools:** Power BI, Tableau, Matplotlib, Seaborn, Excel, Alteryx
- **Cloud & MLOps:** AWS (SageMaker, Lambda), Microsoft Azure, Docker, MLFlow, Kafka
- **Version Control & Collaboration:** Git, GitHub

## WORK EXPERIENCE

### Programmer Analyst (Cognizant Tech Soln)

*Jul 2022 – Aug 2023*

- Engineered anomaly detection (DBSCAN, IF, Prophet), slashing sensor failures by 40% and averting \$100K+ in system downtime.
- Streamlined ETL for over 10M records per day with Python, SQL, Airflow; amplified data throughput and reduced latency by 40%.
- Designed power BI dashboards for 3 ops teams, automating alerts and minimizing reporting effort by 70%, boosting agility by 30%.
- Launched predictive maintenance models, reducing breakdowns 25% and elevating uptime 20% via stakeholder-aligned feedback.
- Orchestrated model deployment with AWS SageMaker; accelerated production rollout by 50% through ML-cloud team collaboration.

### Data Scientist Intern (AFour Technologies)

*Dec 2021 – May 2022*

- Developed AI-Powered pilot management system leveraging TensorFlow, PyTorch and scikit-learn to optimize bidding, flight assignments, and seniority-based scheduling, improving operational efficiency by 20% and enhancing decision-making.
- Engineered large-scale ETL pipelines with Pandas, Spark and AWS Lambda improving scheduling forecast accuracy by 28%.
- Delivered insights on scheduling efficiency and workforce utilization with SQL & Tableau, increasing model adoption by 30%.
- Collaborated with cross-functional teams to minimize disruptions by 25% and sharpen forecast prediction.

## PROJECTS

### Responsible AI Toolkit for Finance

*Sep 2024 – Dec 2024*

- Piloted a Responsible AI Toolkit using GANs and AIF360, reducing model bias by 20% in ECOA-compliant lending systems.
- Crafted a GAN-based data generator to enhance dataset balance, boosting fairness scores by 30% in credit scoring pipelines.
- Embedded XAI for interpretable outputs, elevating AI decision clarity by 50% for non-technical financial stakeholders.
- Automated bias control workflow, trimming manual intervention by 40% while maintaining consistent mitigation accuracy.
- Released a scalable audit platform for financial teams, improving AI risk evaluation efficiency by 35%.

### Diabetes and Retinopathy Prediction Using Machine Learning

*Sep 2024 – Dec 2024*

- Trained a ResNet-50 model for retinopathy severity, achieving 82% accuracy and strong separation across all diagnostic stages.
- Formulated a Random Forest model for diabetes detection, achieving 91% accuracy, while mitigating class imbalance via SMOTE.
- Applied feature scaling, SMOTE and image augmentation to elevate model generalization, enhancing model performance by 15%.
- Conducted EDA to extract key health features (BMI, BP) identifying top 5 indicators aiding early diagnosis and patient risk profiling.
- Refined model performance via hyperparameter tuning, boosting recall by 18% and enhancing diagnostic precision across both tasks.

## EDUCATION

### Master of Science in Information Systems

*Aug 2023 – May 2025*

California State University, Long Beach, CA

### B. Tech in Instrumentation & Control

*Aug 2018 – May 2022*

COEP Technological University, Pune, India

## CERTIFICATIONS

Alteryx Designer Core Certification | **Alteryx**

*2024*

AWS Academy Cloud Foundation | **AWS**

*2023*

Machine learning for Data Science | **Udemy**

*2021*