Deepak B Deokar

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

Work Experience

Student Assistant | *Beach Building Services* *Mar 2024 - Oct 2024*

- Evaluated building performance data (energy usage, maintenance costs, occupancy rates) using Excel and Python, uncovering trends that optimized resource allocation efficiency by 15%.
- Created and automated reports in Excel and Power BI, reducing manual reporting time by 25% and providing management with real-time insights on operational performance.
- Worked with teams to analyze operational data and uncovered inefficiencies, resulting in a 10% reduction in maintenance costs through better resource planning and scheduling.

Programmer Analyst | *Cognizant Technology Solutions* *Jul 2022 - Aug 2023*

- Examined over 5 million data points from sensor data in pharmaceutical manufacturing, utilizing Python and SQL to monitor and optimize conditions like temperature and humidity, securing 99% optimal conditions and minimizing production disruptions.
- Developed machine learning models (e.g., anomaly detection) that identified equipment issues early, leading to a 15% reduction in unplanned maintenance and a 20% increase in equipment uptime, significantly advancing operational efficiency.
- Constructed interactive dashboards in Power BI to visualize 50+ key metrics, facilitating stakeholders to make faster, data-driven decisions, which refined maintenance scheduling efficiency by 25% and minimized downtime.
- Collaborated with cross-functional teams to design and implement insight-based solutions that transformed workflows, accelerating a 30% improvement in overall production performance, eliminating bottlenecks and enhancing productivity.

Intern | *AFour Technologies* *Dec 2021 - May 2022*

- Streamlined data collection and reporting processes for clients in the software development sector, cutting manual effort by 30%, which saved over 200 hours/month and boosted data accuracy, enabling 20% faster decision-making.
- Investigated large datasets (over 1 million records) to identify trends, patterns, and anomalies, highlighting critical areas for improvement, which led to a 10% reduction in system downtimes and a 15% increase in overall system reliability.
- Exploited SQL and Python to clean, manipulate, and process data, supporting the development of predictive analytics models, improving forecasting accuracy by 12% for system performance and trends.
- Partnered with senior analysts on data quality assessments, optimizing the overall data pipeline efficiency by 25%, and delivered actionable insights to transform client systems.

Technical Skills

Programming & Scripting: Python, R, SQL, MATLAB, Linux, Git

AI/ML: TensorFlow, Keras, PyTorch, Scikit-learn, NLP, Time Series Analysis, Pattern Recognition, Feature Engineering

Data Manipulation: NumPy, Pandas, Data Cleaning & Preprocessing, Data Transformation

Data Visualization: Matplotlib, Seaborn, Power BI, Tableau

Big Data & Cloud: Spark, Apache Kafka, AWS (Sagemaker, Lambda), Azure

Skills: Statistics, Data Analysis, Data Modelling, ETL, MS-Excel, Alteryx

Education

Master's in Information Systems *Aug 2023 - May 2025*

California State University, Long Beach

B. Tech in Instrumentation and Control *Aug 2018 - May 2022*

College of Engineering, Pune, India

Projects

Responsible AI Toolkit for Finance

- Engineered a toolkit to assess and mitigate bias in AI-powered financial decision-making, upgrading model fairness in credit scoring, loan approvals, and fraud detection, impacting over 1 million financial transactions.
- Utilized AIF360 for Bias Detection & Fairness Analysis, expanding the fairness of financial AI models by 30% and confirming ethical compliance. Additionally, deployed GANs to generate synthetic data, addressing dataset bias and boosting model accuracy and generalization by 25%.
- Integrated AI Explainability with OpenAI's GPT models and created visualizations/reports, increasing transparency in financial AI systems and decreasing model interpretation time by 40%.

Diabetes and Retinopathy Classification: A Multi-Modal Approach

- Built machine learning models (Random Forest, Gradient Boosting, Neural Networks) to classify diabetes status, achieving 91% accuracy with Random Forest and augmenting prediction reliability for over 10,000 patient records.
- Applied ResNet50 for classifying diabetic retinopathy into five severity levels, achieving 82% accuracy across 5,000+ retinal images. Applied SMOTE for class balancing, enhancing model performance by 15% and addressing label imbalance in datasets with over 50% skewed labels.
- Evaluated models using Confusion Matrices, ROC Curves, and Class-wise Accuracy, maintaining 98% consistency in model predictions across various data subsets.

Certifications

Alteryx Designer Core Certification 2023

Machine Learning in Data Science 2022

SQL for Data Science 2021