Nischay Sai Cherukuri | Data Analyst

cherukurinischaysai@gmail.com | +1 (571) 336-2321| LinkedIn | GitHub | Portfolio

SKILLS

Programming: Python (Pandas, NumPy, Scikit-learn), R
Data Analysis & Visualization: Excel (Pivot Tables, Macros, VLOOKUP), Tableau, Power BI, QuickSight, Looker
Databases & System Administration: Oracle SQL, MySQL, PostgreSQL, Database Performance Tuning
Big Data & Data Warehousing: Snowflake, AWS (Athena, Redshift, S3), Apache Spark, Hive, Data Lake Management
ETL & Data Processing: Alteryx, Talend, Informatica, Airflow, Apache Spark, Data Pipeline Optimization
Statistical Analysis & Machine Learning: Regression Analysis, Predictive Modeling, A/B Testing, Machine Learning, TensorFlow, PyTorch, Feature Engineering, Time Series Analysis, Natural Language Processing, RAG, LLMs

EXPERIENCE

DC Water, Washington, D.C – Data Consultant   *Oct 2024 – Present*
- Engineered Azure-based anomaly detection and forecasting for pump and motor performance using Isolation Forest and ARIMA/SARIMAX, reducing unplanned maintenance by 22%.
- Developed predictive-maintenance ML workflows for pump performance data, generating station-level reports that reduced pump downtime by 15%.
- Created SQL-powered Power BI dashboards for DC Water's internal finance expenditures and $10M+ contracts, providing insights to ensure budget compliance.

GW, Environmental and Energy Management Institute, Washington, DC – Analyst   *Jun 2024 – Jun 2025*
- Analyzed and researched data on renewable energy patterns, developing ML models like Random Forest and Gradient Boosting to forecast trends with 85% accuracy.
- Assisted in professor-led research on landfill data to identify renewable energy opportunities and optimize waste utilization, contributing to enhanced research outcomes and potential policy recommendations.

The World Bank Group, Washington, D.C. – Data Science Consultant   *Jan 2024 – May 2024*
- Achieved comprehensive insights into regional freight and passenger flows by analyzing EU traffic patterns across 27 countries using Eurostat data, producing 30+ map visualizations with Plotly to inform transport policy decisions.
- Supported 10+ strategic transport decisions with data-driven insights by leveraging machine learning and econometrics to forecast railway share price trends, achieving a 90% forecast accuracy.
- Automated a Python data-cleaning pipeline processing 100K+ records, reducing manual errors by 90%, halving runtime, and saving 11+ hours weekly, and deployed a Streamlit web app for interactive data selection, viewing, and downloading.

The George Washington University, Washington, D.C. – Data Analyst   *Mar 2023 – May 2024*
- Designed a Tableau dashboard visualizing demographics of 8,000+ students for Career Services and Admissions, driving data-informed counseling and recruitment decisions.
- Achieved a 35% increase in user engagement by analyzing user interaction data with site improvement analytics, identifying high-performing content, and optimizing website strategy.
- Developed a Siteimprove dashboard for 6 engineering schools, delivering KPIs on website performance and SEO, with actionable recommendations that increased site quality scores by 25% through content optimization.

Mphasis, India – Data Analyst   *Jun 2021 – Jul 2022*
- Automated end-to-end validation and ETL for 1.5M records with Python, SQL, and Airflow, standardizing data workflows, reducing reporting errors by 30%, and saving 20+ manual hours weekly.
- Designed Power BI and Tableau dashboards for cross-functional teams to track KPIs, reducing the turnaround time by 40%.
- Developed a machine learning-based churn prediction model using Python and XGBoost, achieving 92% accuracy in identifying at-risk customers and enabling targeted retention campaigns that reduced churn by 15%.

PROJECTS

Revenue Prediction and Customer Analytics for Supermarket Data | Time Series Analysis   *Aug 2023*
Predicted revenue for UK supermarkets using XGBoost, Random Forest, LSTM, and ARIMA models. Achieved a 30% accuracy improvement with ARIMA and an R-squared value of 0.91 with XGBoost.

Seasonality and Consumer Behavior Analysis in Brazilian E-Commerce | Python, Tableau   *Mar 2023*
Analyzed the Olist dataset using Python and Tableau to uncover seasonal trends, product popularity, and consumer behavior; delivered insights that optimized inventory, boosted marketing effectiveness, and improved customer experience.

Smart Health Predictor: A Multifaceted Disease Analysis and Forecast Model | Python   *Mar 2022*
Engineered ML models predicting cancer, diabetes, and heart disease risk with 85% accuracy and designed efficient ETL pipelines for real-time disease analytics.

EDUCATION

The George Washington University   *Washington, DC*
Master of Science in Data Science

Malla Reddy College of Engineering   *Hyderabad, India*
Bachelor of Technology in Computer Science and Engineering

CERTIFICATIONS

- AWS Certified Solutions Architect – Associate (Link)   *Oct 2023 – Oct 2026*