Allan Cheerakunnil Alex

Mumbai, India | +91 9930703439 | Email | LinkedIn | GitHub | Portfolio | Blog

Professional Summary

Results-driven Data Scientist with a Master's in Computer Science (Specialization: Data Science), leveraging a strong foundation in Python, SQL, and machine learning to transform complex data into actionable insights. Proficient in the end-to-end data science lifecycle, from predictive modeling and statistical analysis to creating impactful BI dashboards in Tableau and Power BI. Eager to apply disciplined problem-solving abilities to drive business value.

Technical Skills

Languages
Python, SQL

Libraries & Frameworks
Pandas, NumPy, Scikit-Learn, TensorFlow, Keras, Matplotlib, SciPy

Databases
PostgreSQL, MySQL, SQL Server

BI & Visualization
Tableau, Power BI, MS Excel

Core Methodologies
Predictive Modeling, Statistical Analysis, NLP, Machine Learning, Deep Learning, Computer Vision, EDA, Data Mining

Web Technologies
Astro.js, Tailwind CSS, JavaScript, HTML/CSS

Deployment (CI/CD)
Vercel

Developer Tools
Git, GitHub, VS Code, Jupyter Notebook, Google Colab, Node.js, npm, conda

Projects

Retail Sales & Revenue Optimization

Technologies: Python (Pandas, Matplotlib), SQL, EDA

- Conducted end-to-end EDA on a transactional dataset to validate the 80/20 principle, identify top revenue drivers, and map peak sales seasons (Nov/Dec) to optimize inventory and marketing strategy.

Customer Segmentation with K-Means & RFM

Technologies: Python (Scikit-learn), K-Means Clustering, RFM Modeling

- Developed an RFM model and used K-Means clustering (optimized with Log Transformation) to profile 4,300+ customers into 5 distinct segments for tailored marketing.

Kindle Book Recommendation System

Technologies: Python (Surprise), SVD, Collaborative Filtering

- Built a Matrix Factorization model using SVD to solve data sparsity, achieving a low RMSE of 3.49 to generate highly personalized book recommendations.

Kaggle Competitions

Digit Recognizer - Handwritten Digit Classification

Designed and trained a Convolutional Neural Network (CNN) with Keras to classify handwritten digits, achieving **98.92% accuracy** on the Kaggle test set (Tech: Python, TensorFlow, Scikit-learn).

Publications / Articles

The Latent Factors: Building a High-Accuracy Recommendation Engine with SVD

Authored a technical deep-dive explaining the role of latent factors and the strategic choice of SVD for sparse data, published on my blog.