# Komal Sairam Reddy Bhimireddy

(669)-204-6802 | komalbhimireddy@gmail.com | Portfolio | LinkedIn | GitHub | San Francisco, USA

## SUMMARY

Data Scientist skilled in Python, SQL, and ML, applying statistical analysis and predictive modeling to solve business problems. Experienced in data visualization and extracting insights to support data-driven decisions. Focused on building and optimizing ML models that enhance business outcomes and support strategic growth.

## EDUCATION

**San Francisco State University, MS in Statistical Data Science, GPA: 3.58** **May 2025**

**Coursework:** Data Mining, Probability and Statistics, Advanced Probability Models, Computational Statistics, Statistical and Machine Learning, Multivariate Statistical Methods

## SKILLS

**Programming:** Python (Pandas, Numpy, Scikit-learn, Scipy, Matplotlib, Seaborn), R, SQL
**Machine Learning:** Regression, Classification, Clustering, Ensemble Methods, Feature Engineering
**Statistical Analysis:** Hypothesis Testing, A/B Testing, Time-Series Forecasting, Statistical Modeling
**Data Handling & Visualization:** Data Wrangling, Data Analysis, Power BI, Tableau, Excel
**Tools & Platforms:** Jupyter Notebook, VS Code, PyCharm, Anaconda Navigator, RStudio, Google Colab
**Version Control:** Git, GitHub

## PROJECTS

**Enhancing E-commerce Experience through Sentiment Analysis** | Code **February 2024 – April 2024**

- Developed a sentiment analysis model for 205,052 e-commerce reviews, leveraging TF-IDF vectorization and machine learning models (SVM, Random Forest, Gradient Boosting) for automated sentiment classification.
- Experimented with data balancing techniques (Tomek Links, SMOTE, Random Under-Sampling) and found that SVM naturally handled class imbalance, making additional resampling unnecessary.
- Achieved an F1 Score of 0.9360 (+3%) and Recall of 0.9419 (+4.6%) using SVM, proving sampling was unnecessary and reinforcing the need to test pre-processing techniques before applying them.

**Music Recommendation System** | Code **October 2023 – December 2023**

- Engineered a music recommendation system using Spotify API data, applying machine learning models to predict song popularity and enhance recommendation accuracy.
- Developed and fine-tuned regression & classification models (Random Forest Regressor: $R^2 = 0.269$, Random Forest Classifier: Accuracy = 0.671) to estimate song popularity and improve ranking of recommended tracks.
- Implemented a content-based recommendation engine using Cosine Similarity, analyzing 5,333 songs to personalize suggestions, demonstrating practical application of machine learning in real-world music analytics.

**SAR Image Classification** | Code **February 2023 – April 2023**

- Applied advanced SAR image filtering techniques (Mean, Median, Gaussian, Refined Lee, Boxcar) to reduce speckle noise, improving image clarity and interpretability for remote sensing applications.
- Implemented unsupervised (K-Means) and supervised learning models (SVM, Decision Tree, Random Forest) to segment SAR images into urban, water, and vegetation classes, ensuring efficient land cover classification.
- Leveraged specialized libraries (GDAL, SciPy, scikit-learn) to develop a remote sensing pipeline for image filtering, classification, and visualization, showcasing expertise in SAR image processing.

## EXTRACURRICULAR ACTIVITIES

**Math Tutor | Tutoring & Academic Support Center** **Feb 2024 – Present**

- Provided personalized instruction in Calculus, Statistics, and Probability to enhance student learning.

**Graduate Teaching Associate | San Francisco State University** **Aug 2024 – Present**

- Delivered Calculus 2 lectures to 40 students, supporting exam prep and conceptual understanding.