Govind Pillai

+1(408)601-7560 | govindpillai2002@gmail.com | linkedin.com/in/pillai-govind

EDUCATION

University of California, Santa Barbara — Santa Barbara, CA
Bachelor of Science in Data Science, Minor in Computer Science — *Graduated: Jun 2024*

EXPERIENCE

West Coast Financial: Data Engineer — Jul 2024 - Mar 2025
- Automated recurring data analysis tasks using R, Excel Macros, and Bash scripting, reducing report generation time by 50% while improving data integrity.
- Applied SQL queries and regression analysis to transform complex data into actionable insights, guiding retirement planning and investment strategies for 25+ clients.
- Worked with data engineering to implement scalable data pipelines using Azure Synapse and dbt, enabling faster access to investment performance metrics and historical client trends.
- Created clear documentation and dashboards to communicate technical results to clients and internal stakeholders with varying technical knowledge.

Payactiv: Data Science Intern — Jun 2023 - Sep 2023
- Developed interactive Power BI and Tableau dashboards to visualize Visa Card user behavior and third-party transaction data, tracking engagement trends and key performance indicators across 150,000+ users.
- Used predictive analytics and stochastic techniques with R and SQL to identify 500+ at-risk user accounts, applying machine learning algorithms to drive targeted engagement strategies and improve retention.
- Used Git and Linux-based environments to manage analytics codebase and perform exploratory analysis on production logs and transaction data.
- Optimized data retrieval using SQL and dplyr in Snowflake and R; collaborated on analytics pipelines using Databricks.

Juni Learning: Computer Science Instructor — May 2021 - May 2023
- Delivered engaging curriculum of Python and Java through diverse methods of computer lab activities and online learning systems.
- Assessed, documented and reported on student progress as a liaison between company and client.

PROJECTS

Baseball Predictive Model
- Engineered and optimized a Random Forest model for scoring prediction, using advanced hyperparameter tuning (grid search, random search) to maximize AUC, precision, and recall while ensuring stability and interpretability.
- Conducted extensive model evaluation using ROC curves, precision-recall curves, and confusion matrix analysis, alongside implementing out-of-sample testing to validate the model's robustness and reduce overfitting.
- Streamlined data aggregation, cleaning, and transformation across three data sources using SQL, R, and ETL pipelines.

Field Goal Prediction Model
- Developed a Random Forest model to predict field goal success, incorporating ensemble learning techniques and optimizing feature interactions, while evaluating model efficacy through precision-recall and F1-score metrics.
- Prototyped neural network alternatives using TensorFlow to capture nonlinear feature interactions and evaluate deep learning model viability.
- Applied robust validation techniques, including stratified k-fold cross-validation and SMOTE for class imbalance, to ensure high generalization and minimize bias in predictions.

SKILLS

Languages: C/C++, HTML/CSS, Java, JavaScript, Linux, LaTeX, MATLAB, Python, R, SQL
Infrastructure: Azure Synapse, Databricks, AWS, Docker, Firebase, Kubernetes, MongoDB, Snowflake Cloud Data
Libraries: Keras, Material-UI, Matplotlib, NumPy, OpenCV, Pandas, PyTorch, React
General: Excel Macros/Visual Basic, Git, Kubernetes, MS Office, Power BI, Tableau
Relevant Coursework: Principles of Data Science in R, Regression Analysis, Data Visualizations, Assembly, Data Mining and Machine Learning, Data Structures, Discrete Math, Linear Algebra, System Design and Architecture
Certifications: AWS Cloud Practitioner, Coursera Supervised Machine Learning, Coursera Advanced Learning Algorithms