Dylan Dsouza

San Diego, California

dydsouza@ucsd.edu | dylandsouza.com | linkedin.com/in/dsouza-dylan | github.com/dsouza-dylan

Summary

Undergraduate data science student specializing in machine learning and artificial intelligence. Seeking entry-level roles to apply data analytics and machine learning to enable data-driven business decisions.

Education

Bachelor of Science, Data Science | University of California San Diego — Expected Jun 2027
Relevant Coursework: Data Structures and Algorithms, Machine Learning Theory, Applied Data Science

Technical Skills

Languages: Java, Python, C/C++, HTML/CSS, JavaScript, R, SQL
Libraries/Frameworks: NumPy, pandas, scikit-learn, Keras, TensorFlow, OpenCV, Plotly, Streamlit, Xarray, Joblib
Core Competencies: Data Cleaning, Visualization, EDA, Feature Engineering, Machine Learning, Predictive Modeling

Experience

Undergraduate Researcher – Marine Data Science — May 2025 – Present
Scripps Institution of Oceanography (Dr. Jennifer Smith Lab) — *San Diego, CA*

- Using Viscore, a UC San Diego-developed photogrammetry and data annotation tool, to generate scaled and oriented photo quadrats for downstream spatial and temporal analysis of 15 coral reef sites in the Central Pacific.
- Actively labeling 9,000+ georeferenced points (400 per site) across multiple time points, classifying coral, algae, and invasive corallimorphs to produce a structured dataset for reef health assessment.

Board Member – Online Content — Oct 2024 – Present
Data Science Student Society at UC San Diego — *San Diego, CA*

- Research and author data-driven articles on machine learning applications, translating complex technical topics into accessible narratives for a student audience comprising 2,350+ followers and 500+ community members.
- Featured how predictive modeling, photomosaic imaging, and machine learning advance coral reef restoration and wildfire detection, highlighting projects from NASA, UC Berkeley, and UC San Diego's WIFIRE Lab.

Marketing Intern — Jul 2023 – Aug 2023
Bennett Coleman & Co. Ltd. (The Times of India) — *Mumbai, India*

- Analyzed 56,000+ newspaper headlines using TF-IDF vectorization and multi-output linear regression, training a model to predict relevance scores across 5 categories: Business, Politics, Technology, Social Issues, and Lifestyle.
- Built an interpretable classification system using scikit-learn to assist editorial teams in headline categorization; exported results for downstream analysis and reporting.

Student Ambassador and Leadership Fellow — Sep 2021 – Nov 2021
Inspirit AI — *Remote*

- Developed and optimized a skin cancer diagnosis model using Keras CNNs and MobileNet transfer learning on a dataset of 10,000+ images, achieving 93% accuracy and an ROC AUC score of 0.97.
- Improved model generalization and efficiency by applying OpenCV image augmentation, Grid Search hyperparameter tuning, and dimensionality reduction (PCA, t-SNE) with feature extraction (SIFT) to mitigate bias.

Projects

ENSOcast | Python, pandas, NumPy, scikit-learn, Streamlit, Plotly, Xarray, Joblib — View Project

- Trained a Random Forest model (82% accuracy) on 40+ years of climate data using feature engineering techniques including lagged variables and temporal encodings to classify El Niño–Southern Oscillation (ENSO) phases.
- Built a production-ready Streamlit dashboard with interactive Plotly visualizations, enabling users to explore ocean temperature trends, evaluate model performance, and train custom prediction models.

Geriasphere | HTML, CSS, User Interface (UI), User Experience (UX) — View Project

- Designed a responsive website to enhance digital literacy among middle-aged and senior citizens in India by providing step-by-step video tutorials for 9 essential mobile apps, including Uber, Google Maps, Zoom, and Gmail.
- Conducted an in-person workshop for 15+ senior citizens to introduce the platform and assist with hands-on learning, embedding Google Forms for continuous community-driven updates to tutorial content.