PALASH SURYAWANSHI

619-719-6196 | psuryawanshi2178@sdsu.edu | linkedin.com/in/palash | github.com/palash | San Diego, CA

EDUCATION

San Diego State University — *San Diego, CA*
Master of Science in Big Data Analytics — *Aug 2024 – May 2026*

Savitribai Phule Pune University — *Pune, MH*
Bachelor of Engineering in Computer Engineering — *Aug 2018 – May 2022*

PROFESSIONAL EXPERIENCE

Research Assistant — *Nov 2024 – Present*
San Diego State University — *San Diego, CA*

- Extracted and analyzed 133,000+ financial articles from Seeking Alpha using Python (Scrapy), improving data retrieval efficiency by 30%.
- Processed and modeled data for 1,000+ stock tickers (e.g., AMZN, AAPL) using Pandas, NumPy, and SQL, leading to 15% faster query execution for financial analytics.
- Integrated datasets from CRSP, Compustat, and I/B/E/S, optimizing financial data aggregation and improving stock trend prediction accuracy by 20%.
- Applied NLP and statistical transformations, increasing data structuring efficiency and ensuring 99% data accuracy through outlier handling and Winsorization.

Graduate Assistant — *Jan. 2024 – Present*
San Diego State University — *San Diego, CA*

- Assisted faculty in designing and evaluating course materials for the Marketing Research program, increasing student engagement by 40% through improved instructional design.
- Conducted comprehensive data analysis on sustainability in online grocery shopping, synthesizing insights from Deloitte, PwC, McKinsey, and other industry reports, identifying three key trends that increased market efficiency.

Software Engineering Intern — *Aug. 2022 – Feb. 2023*
Cognizant Technology Solutions — *Pune, MH*

- Developed 'Bank Lending Portal,' a responsive web application leveraging Angular, Spring Boot, and REST APIs, resulting in a 25% improvement in user interaction.
- Integrated front-end and back-end systems, enhancing functionality and reducing load times by 15%.
- Oversaw MySQL databases with SQL queries, maintaining 99% data accuracy and integrity.

TECHNICAL SKILLS

Programming & Data Analysis: Python, R, SQL, Excel
Machine Learning & Predictive Analytics: Regression, Classification, Clustering, A/B Testing, Time Series Analysis
Data Visualization & BI Tools: Tableau, Power BI, ArcGIS, QGIS, Highcharts
Databases & Query Optimization: MySQL, MongoDB, Google BigQuery, SQL Query Optimization
Big Data & Cloud Platforms: Hadoop, Spark, AWS (Redshift, S3, Athena), Microsoft Azure (Data Factory)
ETL & Data Cleaning: Alteryx, Apache NiFi, Data Wrangling, Data Transformation, OpenRefine
Version Control: Git, GitHub
Libraries & Tools: Pandas, NumPy, Scikit-learn, TensorFlow, Seaborn, Matplotlib, Plotly

PROJECTS

TrafficSensAI | *Python, scikit-learn, Highcharts*

- Built ML models (Random Forest, DBSCAN) to predict traffic accident severity, improving prediction accuracy by 20%, reducing false positives by 15%, and supporting urban traffic safety optimization.
- Developed interactive dashboards using Highcharts, allowing city planners to identify high-risk accident zones, leading to a reduction in accident-prone areas over six months.

EduDB: Personalized SQL Learning through Generative AI | *MySQL, LLM, Python, Flask*

- Designed an AI-powered SQL learning platform, generating custom SQL queries for 1,000+ students, improving learning efficiency by 35% through interactive modules.
- Engineered a theme-based database system, dynamically populating tables using LLM-based APIs, reducing query-building errors by 40% and enhancing student engagement.

Medi-Vu | *R, Tableau, ArcGIS Pro, Snowflake, Highcharts*

- Analyzed disease incidence data & healthcare facility locations, identifying 20+ high-risk regions and enabling better resource allocation in outbreak-prone areas.
- Built interactive Tableau dashboards, providing real-time epidemiological tracking, reducing decision-making time for disease outbreak predictions.

A Smart Blogging Site with Hybrid Recommendation System | *Python, Pandas, Flask*

- Developed a blogging platform with an integrated chatroom, fostering real-time collaboration among 50+ authors, leading to a 30% increase in article engagement.
- Implemented NLP-driven recommendation systems, improving personalized content discovery, increasing reader retention rates by 30%, and boosting user satisfaction scores.