YOSUA SAPUTRA

Los Angeles, CA | (626) 940-6974 | yosua.jsaputra@gmail.com | linkedin.com/in/yosua-saputra

EXPERIENCE

Data Annotation.Tech | Remote

Data Annotator Specialist   **02/2024 - Present**

- Debugged over 90 tasks involving Data Analysis and Visualization in Python, achieving 100% accuracy and reliability.
- Trained 50+ AI models by analyzing coding responses and conducting AI conversations using Python and SQL.
- Developed and implemented Python scripts and SQL queries, enhancing AI model performance, and optimizing SQL query efficiency for ML models and Exploratory Data Analysis.

Scale Media Inc | Los Angeles, CA

Data Analyst   **03/2022 – 06/2023**

- Analyzed AB test results for subscription-based features, contributing to a 10% reduction in customer cancellation rates.
- Increased query runtime performance by over 300% for tables with 700+ million rows using Amazon Redshift query scheduler, leading to more efficient data analysis and data workflow.
- Revamped SQL queries to provide real-time updates on subscription-based customer data.
- Debugged and optimized SQL queries to maintain seamless data pipelines and eliminate latency issues.
- Implemented and quarterbacked Tableau dashboards for comprehensive visualizations, improving business KPIs and facilitating data-driven insights for stakeholders.
- Collaborated with analytics managers and cross-functional teams to automate A/B testing Excel model, saving 1–2 hours daily and supporting a major campaign launch.

UCSB Arts & Lectures | Santa Barbara, CA

Data Entry Specialist   **03/2022 – 06/2023**

- Synchronized archive data using WIX platform to successfully measure its discrepancies.
- Constructed templates for UCSB A&L website and generated content for weekly emails.
- Collaborated with cross-functional teams to input archive data imported from Excel into pre-launched websites.

SKILLS

- SQL, Python, Pandas, NumPy, Matplotlib, Scikit-learn, Microsoft Excel, Google Sheets, XLOOKUP, Pivot Tables, Tableau, AWS - Amazon Redshift, Regular Expressions, Microsoft Office, Statistics, Data Science, Exploratory Data Analysis (EDA), Data Analysis, Data Visualizations, Data Cleaning, Quality Assurance, A/B Testing, Database Management, Hive, User Engagement, Customer Segmentation

PROJECTS

Predictive Modeling: Insurance Call Lead Conversion

- Handled imbalanced data utilizing over-sampling techniques to create a random forest classifier model. Achieved a F1 Score of 94.8% and an AUC score of 94.9%, indicating strong performance in predicting conversion.
- Analyzed post-modeling results to derive insights into types of health calls that are most likely to convert.

Recommender System & Exploratory Data Analysis

- Performed data analysis to gather insights on user engagement, genres, series, and categories.
- Utilized python libraries and machine learning techniques such as collaborative filtering and k-Nearest Neighbors for data cleaning, pre-processing, visualizing insights, and building the recommender model.
- Presented dynamic recommendations of ten anime series based on user input, with a random generator default option.

EDUCATION

University of California, Santa Barbara | Santa Barbara, CA

Bachelor of Arts: Statistics & Data Science, Economics