LONG H LE

Sydney, Australia
(61)493-908-641
lehoanglong95@gmail.com
https://www.linkedin.com/in/hoang-long-le-713b41111/
https://github.com/lehoanglong95

Skills

Languages: Python, SQL, Golang

Data Analyst: Airflow, Numpy, Pandas, Superset, Matplotlib, PowerBI

Backend Frameworks: Flask, FastAPI

Databases: PostgreSQL, MongoDB, Trino, Redis

AWS Service: EC2, S3, SQS, SNS, EBS, Lambda, RDS

Transferable Skills: Strong problem-solving and analytical skills, Excellent attention to detail and organization, Effective at meeting deadlines, multitasking, and prioritizing workload, Proactive in fast-paced environments, Quickly adapt to changes in products, processes, and procedures

Domain Knowledge: Media buying platforms, E-Commerce

Work Experience

Data Analyst   *Oct 2023 – now*
PTE Master   Sydney, Australia

- Utilize SQL queries to extract and organize student data, including PTE scores, study duration, mock test performances, and other progression metrics. Identify performance trends, areas of struggle, and propose interventions for improved outcomes, resulting in a 10% increase in student satisfaction.
- Gather marketing data from platforms like Google Ads, Facebook Ads, and email campaigns. Use Python and Pandas to clean and merge data, conduct advanced analytics, optimize budgets, and identify target demographics, leading to a 5% reduction in marketing budget while maintaining or increasing effectiveness.
- Develop automated reporting using Power BI for insightful reports on metrics like PTE scores, pass rates, and study habits. Update and distribute reports to inform decision-making.
- Implement Python scripts for data validation, audit databases for consistency, and collaborate with IT to maintain data integrity.

Data Analyst   *Jan 2023 – Oct 2023*
Superfine   Hanoi, Vietnam

- Orchestrated the migration of a Web3 user database from MongoDB to PostgreSQL, enhancing data consistency and slashing query time by 30%, alongside a 20% reduction in operational costs.
- Built ETL pipeline to (1) ingest spend and revenue raw data from multi data-source, (2) clean, validate and transform data, (3) load processed data into data warehouse using Airflow, Pandas, BigQuery and Google Cloud Storage.
- Revamped the organization's spend and revenue table layout by adopting a snowflake schema methodology, a meticulous process designed to fortify data integrity and consistency. This restructuring initiative not only optimizes data storage but also lays a robust foundation for streamlined data retrieval, empowering stakeholders with reliable insights for informed decision-making.
- Formulated an optimized SQL query to compute Customer Lifetime Value (LTV) from the revamped spend and revenue table, resulting in significant performance enhancements for reporting: accelerated overall user query performance by 250%, reduced the calculation time for Lifetime Value (LTV) by 250%, and slashed retention rate calculation time by 300%.
- Engineered an advanced data alert system that autonomously detects anomalies within the dataset, promptly issuing notifications to the dedicated data engineering team

AI Engineer   *May 2022 – November 2022*
TI KI CO., LTD.   Hanoi, Vietnam

- Researched and implemented deep learning algorithms for image comparison, yielding a commendable 72% F1-score on the test set.
- Developed robust product classifiers leveraging product information, showcasing exceptional performance with a 95% F1-score on the test set.
- Led the successful deployment of a Production Management System (PMS) in the production environment, resulting in significant improvements. This included reducing system latency by 200% and optimizing resource efficiency by consolidating nodes by 400%.

EDUCATION

Master in Computer Science   *Feb 2022*
Sejong University
Bachelor in Information Technology   *July 2017*
Hanoi University of Science and Technology, Vietnam

CERTIFICATIONS

AWS Certified Developer – Associate   *Jan 2024*
Datacamp Data Processing in Shell   *May 2020*
Datacamp Writing Efficient Python Code   *May 2020*
Coursera Using Databases with Python   *Jan 2020*