

TIFFANY DINH

linkedin.com/in/tiffndinh | dinhtechdata.github.io | tiffndinh@gmail.com | (714) 837-1849

SUMMARY

Motivated data professional equipped with an MS in Business Analytics, 2 years of hands-on experience, and proven skills in statistical analysis, data visualization, and data storytelling. Demonstrated competency in Python, SQL, and R to translate complex data into actionable insights to enable data-informed decision-making.

PROFESSIONAL EXPERIENCE

Warner Music Group *Data Analyst, Quality Control*

Aug. 2022 – Aug. 2023

- Instituted data governance standards and protocols to reinforce data quality and integrity across 100B+ streaming and subscription data points
- Championed adoption of data governance best practices – define standards, build validation checks, and monitor KPIs to ensure data quality for data products
- Drove development and implementation of a centralized data catalog, increasing data visibility by 30% and enabling self-service data discovery and data lineage across the organization
- Developed an automated data quality checks pipeline using statistical profiling and data testing, leading to a 20% reduction in the workflow for escalating data issues for in-production products
- Conducted regular audits and data profiling to determine pattern and trend anomalies in data quality, resulting in a 35% decrease in data anomalies and outliers
- Managed cross-functional teams aligned on organizational data policies and regulatory requirements, driving projects forward to meet deadlines

UCI Capstone Project – Ingram Micro *Student Data Scientist*

Jan. 2022 – June 2022

- Designed predictive models to rank vendors' purchase propensity to identify targeted sales campaign
- Investigated relationship and correlation between variables and analyzed vendors' purchase behavior and operational trends
- Evaluated models' performance using learning-to-rank and other metrics to determine optimal algorithm-based on robustness, accuracy, and longevity
- Maximized models' performance by approximately 5% using hyperparameter tuning for models such as Logistic Regression, XGBoost, LightGBM, and Random Forest
- Presented findings and recommendations to the VP of Data Science to improve current propensity models and enhance sales targeting strategies

RELEVANT PROJECTS

Detecting Cells Infected with Malaria (<https://dinhtechdata.github.io/tabs/project1.html>)

June 2022

- Leveraged deep learning algorithms to develop an image classification model that identifies malaria infections in cell images with a 94% accuracy
- Improved models' optimization by regularly assessing accuracy and binary cross-entropy loss during training

Flight Delay Prediction at JFK Airport (<https://dinhtechdata.github.io/tabs/project4.html>)

May 2022

- Conducted exploratory data analysis and investigated relationships between variables that increased the likelihood of a JetBlue flight delay such as peak hours, day of the week, and month for flight delays
- Identified NY airport with greater flight delays through hypothesis testing leveraging statistical techniques
- Forecasted flight delays with 34% error rate by training machine learning models including KNN, Logistic Regression, Random Forest, Decision Trees, and Naïve Bayes

EDUCATION

University of California, Irvine, Paul Merage School of Business *Irvine, CA*

Master of Science in Business Analytics

June 2022

Bachelor of Chemistry with a Concentration in Chemical Biology

June 2019

TECHNICAL SKILLS

Programming Languages: Python, R/RStudio, SQL, Mathematica, Git, HTML, CSS

Tools: Tableau, Alteryx, Snowflake, Jira, GitHub, Jupyter Notebook, Google Colab, Excel/ Google Sheets

Certification: Data Science Team Lead (DSTL) Certificate