David Ly

Data Science/Analytics

P: (408) 892-6862 **E: davidtnly@gmail.com**

LI: www.linkedin.com/in/davidtly **Web: https://davidtnly.github.io/**

Technical Skills & Knowledge

- **Languages/Frameworks:** Python (sci-kit learn, matplotlib, seaborn, pandas, numpy, Keras, TensorFlow, OpenCV), R (dplyr, ggplot2, caret, glmnet), SQL
- **Platforms:** Tableau, SSRS, Microsoft SQL Server, Alteryx, Power BI, Jupyter, Basic Google Cloud Platform
- **Topics:** Advanced data visualization, data cleaning & storytelling, statistical analysis, predictive analytics (machine learning, random forest, gradient boosting), feature engineering, basic deep learning / cloud, neural networks

Experience

Business Analyst, Data Science & Reporting – loanDepot, Wholesale (October 2016 – Present)

- Oversaw development in analytics and reporting with hands-on ownership of almost all business intelligence matters including data quality assurance, control of reporting infrastructure, and created new dashboards for 300+ users
- Led cross-functional teams of 6+ in providing guidance from concept to production in multiple projects including building new ETL solutions using Alteryx/R
- Solved complex data integrity issues with our third-party data, which involved analyzing ~10M rows of data and presented it in multiple views in Tableau for senior management and the sales team
- Decreased average loan process time by 1-2 days by creating SQL stored procedures visualized in Tableau & SSRS
- Increased average team production of loans by 42% through measuring risks and analyzing productivity data which also reduced the average team turn times by 0.5 – 3 hours
- Slashed average execution time for all SQL stored procedures and queries by 1.5x – 10x using SQL best practices
- Trained loan prediction models in development using XGBoost & Elastic Net and forecasted monthly funded units to track goal within ~7% MAPE using additive methods such as ARIMA, ETS, and Prophet in R
- Optimized document gathering process up to 45% by identifying inconsistencies in our loan processes by stages
- Conducted quantitative analyses to construct new risk scoring, propensity to fallout, and loan lifecycle models in R and SQL that is implemented in Tableau with interactive weighting parameters
- Automated reports that reduced 12+ hours of work every month for operational efficiency using SQL/Alteryx/Tableau

Business Analyst, Marketing Analytics – Twelve Springs, LLC. (July 2015 – July 2016)

- Saved company over ~$9,000 by driving brand performance, customer acquisition and digital strategies internally
- Increased average revenue by 1.5x and metrics by 2x in 2 quarters by implementing consumer psychology strategies
- Grew website traffic by 7x by utilizing cluster analyses of consumers to email marketing and connecting with bloggers
- Led team in reporting to executives and presented new marketing methodology which was adopted to improve online presence, monthly rate of business inquiries for revenue growth, and understand our customer base better
- Created all SQL queries and collaborated to define data structures and sources required for the data pipeline
- Worked on customer segmentation projects involving lead generation, acquisition, and estimated chances of leads booking using logistic regression, propensity scoring scorecards, and visualizations in Tableau

Projects

- **SIIM-ACR Pneumothorax Image Classification** – (Aug 2019 – Present)
  - Competition to develop a model to classify & segment pneumothorax from a set of chest radiographic images
- **Pneumonia Detection Using Chest X-Ray Images** – (Aug 2019)
  - Trained a 5-layer separable convolutional neural network model from scratch to classify and detect the presence of pneumonia from a collection of chest X-ray image samples
  - Resulted in a 93% training accuracy and 87% testing accuracy along with a 87% F1-score
- **Apple Health & Personal Data Exploration / Visualizations** – (July 2019 – Aug 2019)
  - Analyses on personal health activity (~900k rows) and study topics data using Python and Tableau

Education & Relevant Coursework

- **Independent Deep Learning Research –** Python, Keras, TensorFlow, & OpenCV
  - **Stanford, Coursera, Lynda** // Deep Learning & Computer Vision – (July 2019 – Present)
    - Coursework covering applications like facial recognition, image processing, object detection, computer vision, deep learning networks including YOLOv3, OpenCV, and cloud computing
- **University of California, Irvine** // Bachelor of Arts in Economics and Minor in Psychology (June 2015)