

# Kevin Liantono

Seattle, WA 98107, (408) 886-1217, kevin.liantono@gmail.com, [www.linkedin.com/in/kevinliantono](http://www.linkedin.com/in/kevinliantono)

---

## AREA OF EXPERTISE

**Analytics:** Regression (LR, SVR, GPR, Ensemble, Decision Trees), Classification (SVM, Discriminant Analysis, Naive Bayes, KNN), Clustering (K-Means, Hierarchical, Gaussian Mixture, Hidden Markov Model), Anomaly Detection, T-test, A/B/N testing

**Languages:** Python (Pandas, NumPy, Scikit-Learn, Seaborn, Matplotlib, NLTK, Tensorflow II, Keras), R(ggplot2), SQL

**Data Visualization:** Tableau, Sisense, Plotly, Looker, Grafana, Baremetrics, Facebook Analytics, Google Analytics

**Technologies:** Spark, Snowflake, AWS S3 Datalake, PostgreSQL, DBT, Stitch, Airflow, Bugsnag, BigQuery, GCP/GCS, Github

## PROFESSIONAL EXPERIENCE

### Within Unlimited LLC, Supernatural (VR/AR)/ Los Angeles, CA

October 2020 – Present

#### Senior Data Scientist (Applied)

- Machine Learning Projects:
  - Recommendation Engine | Algorithm: RBM, KNN, SVD | Evaluation Metrics: Cosine Similarity, MAE, & MRR
  - Customer Profiles | Algorithm: K-Means | Evaluation Metrics: Convergence Rate, Sum of Squared Error (SSE)
  - Anomaly Detection in Data Quality | Algorithm: Local Outlier Factor | Evaluation Metrics: F1 Score, Precision, Recall
- Segmentation & Clustering analysis to build out customer cohorts for marketing, content, & product.

#### Data Scientist

- Help hire, train, and onboard two Senior Data analysts.
- Pioneer a standardized process for planning & deploying of AB tests, along with other causal inference methods.
- Build a centralized Tableau dashboard to monitor key metrics & KPI's. Evaluate new BI tools & migrate dashboards into Sisense.
- Create a democratized data store to gather data from disparate sources (Airflow, s3 datalake, stitch/dbt, Snowflake)
- Establish standards and best practices for maintaining an efficient data engineering workflow, adopt JIRA sprint system, Github version control, and Confluence to document different segments of users, onboarding plan, schemas & tables & their relationships, company-wide terminology & definitions, analysis catalogue, experimentation results, and research.

### Porch Inc./ Seattle, WA

December 2018 – October 2020

#### Data Scientist

- Design & build a Random Forest classification model to predict the probability of leads monetizing for high value products based on a set of features & using the probability score to create a managed queue system to reduce operational costs.
- Build reliability (log loss, Chi-Squared, P-value, CAP, AUC) plots in Plotly to evaluate model performance & accuracy over time.
- Conduct hypothesis testing & retraining of our XGBoost classification model to better predict the probability of pro coverage.
- Build a time-series model to track & forecast the volume of new customers and create cohort analyses to determine customer attrition rates over time, resulting in calculated action that lead to the increased retention rates of customers.
- Exploratory investigation resulting in discovery and ingestion of missing data, increasing net profit by ~200% for 3<sup>rd</sup> party leads.

### Commercial Industrial Waste Applications Inc./ San Jose, CA

September 2016 – November 2018

#### Senior Data Scientist

- Build monitoring, forecasting, & pricing models for management & supply partners leading to a 126% growth in revenue.
- Create bidding models in Python and contract deals in order to broker business and manage client's waste stream.
- Manage SQL database for accounts and relationships with partners, vendors, and clients at the firm.
- Lead project to implement a native payment system in website portal; increase turnaround on payments by 30%.

### Travix International, Vayama / San Jose, CA

August 2014 – August 2016

#### Senior Data Analyst / Data Analyst

- Create a 100+ structured airline contract portfolio & monitored local, global, and backend contracts to push sales for the North American and Asian Regions in order to increase commission levels & receivables by 30-200%.
- Lead marketing campaign to improve partner optimization using Google Analytics to identify key metrics & alter bidding tiers, increase click-through on meta channels due to ads by 25%, leading to a 9% increase in bookings.
- Increase conversion by 50% by examining the conversion funnel/ user demographics; restricted negative margin routes, analyzed strategic pricing modules, tracked booking patterns & rule evaluation.

## MACHINE LEARNING | PROGRAMMING | EDA PROJECTS

Housing Prices | Credit Fraud | Fake vs Real News | Netflix Shows | Amazon Reviews | Google Play Store Apps | Global Terrorism | Police Fatal Shootings | M5 Forecasting | Nifty 50 Stock Prices | Video Game Sales | Trending YouTube Videos | Campus Placement | Titanic Dataset | Novel COVID-19 | World University Rankings | Brazilian E-Commerce | Multi-Armed Bandit Testing | Starcraft AI

## EDUCATION

### Machine Learning / Deep Learning / Programming Coursework

- Programming Bootcamp: London App Brewery May 2022
- Machine Learning Bootcamp: Open Data Science - ods.ai April 2020
- Machine Learning Course: Andrew Ng CS229 Machine Learning, with Mathematical Proofs – Stanford University March 2019
- Data Science Bootcamp: Machine Learning A-Z: Hands on Learning with Python & R in Data Science – Udemy February 2019
- Data Science Bootcamp: The Data Science Course: The Complete Data Science Bootcamp – Udemy February 2018

### University of California, Davis / Davis, CA

Summer 2014

Bachelor of Arts in Economics