JONATHAN YU

DATA SCIENTIST

jonathanyu90@gmail.com | runjon90.github.io | (847) 756-0169 | Greater Chicagoland Area | /in/jonyu231/ | runjon90

EXPERIENCE

METIS · Chicago, IL — June 2021 to Aug. 2021

Data Scientist

Completed an immersive data science & engineering bootcamp and built predictive models using regression and classification to develop strategies for various domains, including rental bike ridership, visa applications, and police staffing. Created interactive visualizations and presented findings to peer, senior, and chief data scientists.

Selected work:

Alternative Airbnb Rating System: Employed natural language processing to extract topics and conduct sentiment analysis using 600k+ Airbnb reviews from Hawaii. Utilized these in conjunction with dimensionality reduction and clustering to discover hidden features which drive guest ratings. Implemented the finalized topic filter to derive topic weight adjusted guest ratings which were visualized in a Tableau interactive dashboard.

Predicting Divvy Bike Rental Demand: Built and optimized a regression model to predict daily Divvy (Chicago's bike rental operated by Lyft) demand by location and time slot. Selected numerical predictors included 1.2+ million prior ridership entries, crime count for the preceding 7 days, web-scraped weather information (precipitation, temperature, visibility, wind speed, sunlight), and CTA transit data while the categorical predictors were day of the week and holidays. Final Ridge model achieved 0.904 R^2 on test set.

Interactive Dashboard for Chicago Food Safety Inspections: Built an end-to-end data pipeline for Chicago food inspection results. Obtained raw inspection data through Socrata Open API and processed for relevance. Launched a web app through Streamlit to empower local consumers with an interactive map with ability to adjust for risk level, actual inspection results, and violations as well as search for specific business entries.

Correlating NYC Subway Traffic with Adjacent Crime Counts: Evaluated MTA turnstile and overlapping NYPD crime data to discover correlation while adjusting for noise. Deployed model to recommend staffing levels by station, season, and time of day.

Exploring U.S. Permanent Labor Visa Applications: Built and compared competing classification models to gauge risk of applications being denied to better prepare employees and employers alike. Ultimately selected the best-performing XGBoost model with features considering salary, employer prominence, currently held visa class, education level/background, and various metrics pertaining to the country of origin (literacy rate, education expenditure as % of GDP, etc.).

BRAINLAB · Chicago, IL — Aug. 2013 to June 2021

Remote Support Manager & Radiation Safety Officer (Jan 2019 - Jun 2021)

- Promoted to lead and mentor team of six application support engineers responsible for front-line external product support, back-end licensing structure maintenance, and remote deployment of software applications per client design
- Conducted regular trend analysis using Salesforce to SQL pipeline to identify targets for sale of service contracts, clinical software subscriptions, and system upgrades
- Concurrently served as Radiation Safety Officer and advised senior leadership on radiation shielding and regulatory requirements prior to new product (Loop-X Mobile Intraoperative Imaging Robot) introduction
- Kept organization's registration as service provider of radiation equipment up-to-date and compliant with individual U.S. states and Canadian Nuclear Safety Commission. The covered radiation-producing products were ExacTrac, Vero, Airo (acquired by Stryker), and Loop-X.

Senior Solutions Engineer (Jul 2016 - Dec 2018)

- Used natural language processing on comments of closed tickets to derive resolution details and construct a constantly updating solutions database
- Used AQL/DQL sampling to automate internal audits of service reports and preemptively detect violations

Application Support Engineer (Aug 2013 - Jun 2016)

- Supported 25+ software products (Class I medical devices) in radiation therapy and image-guided surgery, successfully resolving over 6,500 tickets and ensuring inclusion of sufficient information for FDA vigilance
- Primarily worked with products in spatial tracking algorithms for patient positioning
- Diagnosed DICOM transfer/integration issues with PACS, QA, record verify, and third-party treatment planning software in radiation oncology

DMC AMERICA · South Elgin, IL — Apr. 2013 to Aug. 2013

Management Associate

- Served as liaison between Korean headquarters and dealers in Canada, Mexico, Brazil, and the US by setting model prices and forecasting purchase orders for corresponding specs of CNC machines
- Managed and consolidated databases covering inventory, payment/invoice accounts, and dealer performance
- Ensured quality control and information accuracy from the marketing department by verifying correct translation and machine specifications

SKILLS

LANGUAGES + TOOLS

Python
C/C++
JavaScript
HTML/CSS
R
SQL
Selenium
pandas
Numpy
NLTK
Scikit-learn

MACHINE LEARNING

Classification
Logistic / Linear Regression
Web Scraping
Natural Language Processing
Dimensionality Reduction
Clustering
Unsupervised Learning
Feature Engineering
LASSO / Ridge
XGBoost
Naive Bayes
Random Forest
k-Nearest Neighbors

DATA MANAGEMENT

AWS
Google Cloud Platform
mongoDB
Hadoop
SQLite
Docker

VISUALIZATION

Tableau
Streamlit
Folium
Seaborn
MatPlotLib
Plotly
CartoDB
PowerPoint

EDUCATION

University of Illinois at Urbana-Champaign

BS Bioengineering 2012

Coursework: Discrete Structures, Analysis of Data, C Programming, Symbolic Computation Lab, Microcomputer Applications, Calculus 3, Modeling Human Physiology (MATLAB), Differential Equations, Analog Signal Processing, Statistics