

BRIAN CHI YAN LI

Results-driven data scientist with a strong background in data cleansing, feature engineering and machine learning. Proficient in scripting and process automation using SQL, Python and R. Skilled in cross-functional collaboration, Agile project management, and communication.



EDUCATION

- | | | |
|-------------------------|--|----------------------|
| 05/2019

08/2017 | • Master of Science in Statistics (GPA: 3.9)
North Carolina State University | 📍 Raleigh, NC |
| 05/2014

08/2012 | • Master of Science in Engineering Management CO-OP (GPA: 3.5)
Purdue University | 📍 West Lafayette, IN |
| 05/2012

08/2007 | • Bachelor of Science in Chemical Engineering (GPA: 3.6)
Purdue University

• Minor in Mathematics and Economics | 📍 West Lafayette, IN |

WORKING EXPERIENCES

- | | | |
|-------------------------|---|------------------------|
| Current

01/2023 | • Data Scientist
Levi's

• Engineer time lag/product features and apply Gradient Boosting algorithms to forecast shipping demand and sizing distribution for inventory planning
• Develop an Airflow sellout data pipeline integrated with Github CI/CD deployment and build test cases to alarm product assortment changes
• Migrate Europe forecast from AWS SageMaker to GCP Vertex AI and refactor pyspark/pandas data pipeline to reduce ~15% runtime
• Enhance model hyperparameter tuning process by replacing grid search with Optuna framework, reducing effort from 5 to 2 days.
• Conduct EDA/visualization on time series data and apply prophet model on a subset of products to reduce WMAPE by ~12%
• Build an analytical chatbot coupling code and text-bison models from GCP PaLM2 suite, enabling business planners to self inquiry without writing SQL
• Maintain a feature store by ensuring data consistency and monitor prediction accuracy by performing model backtesting | 📍 Orlando, FL (Remote) |
| 01/2023

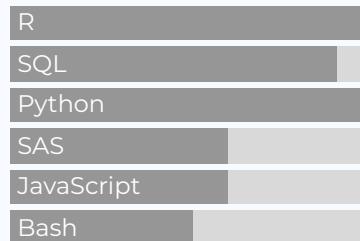
08/2021 | • Principal Data Scientist
Verizon

• Applied ML models including Logistic Regression, Generalized Linear Models with Regularization (Lasso/ Ridge), Random Forests, XGB Tree to develop optimized credit strategy for wireless consumer to promote customer growth and reduce default rate
• Leveraged random forest model with payment behavioral features and survival model to build churn forecasting to provide key inputs for profit/loss financial report
• Built a Qlik dashboard visualization with REST API to automatically update churn forecast trending vs budget for client, eliminating manual update work
• Performed 5G Home credit modeling and optimization to determine deposit schemes that encourage market share growth while accounting for voluntary churn
• Cross trained other team members on fraud scorecut modeling/deployment with new EFX Fraud Superscore and Neustar score
• Acted as subject matter expert on Oxford Economics macroeconomic data correlation analysis for SOX compliance | 📍 Lake Mary, FL |

CONTACT

- ✉ 8787 Lookout Pointe Dr., Windermere, FL 34786
✉ ziyong68@gmail.com
📞 765-413-3280
🔗 github.com/ziyong68

PROGRAMMING SKILLS



OTHER SKILLS

Tools: Teradata, Hadoop, Spark, Docker, Airflow, GCP Vertex, AWS SageMaker

Reporting Systems: Qlik, SAP Business objects, Eclipse BIRT, Jira, Confluence

Languages: English, Mandarin, Cantonese

CERTIFICATIONS

Lean and Six Sigma Green Belt Certificate, Purdue University

Managing Big Data with MySQL, Duke University

Python Programming Certification, University of Michigan

FCRA Data User Certification program through CDIA

Made w/ R [pagedown](#).

Source code:

github.com/ziyong68/data-driven-resume

Last updated on 2024-04-18.

08/2021
|
01/2019

● Data Scientist

Verizon

📍 Lake Mary, FL

- Developed a credit tightening policy for Iphone prelaunch to reduce bad debt while ensuring activations from top customers
- Implemented Multi Adaptive Regression Spline (MARS) model to predict exposure at default for device payment plan loans with 6, 30, 36 month terms
- Optimized FraudIQ score threshold of red flag policies to stop identity frauds from entering credit check and prevent losses (est. with 700K - 1 million per month)
- Analyzed Ignite credit bureau data assets using Impala SQL in Hadoop clusters and perform quality control & data deduplication on transactions and archives
- Initiated the use of R markdown (knitr) during CECL auditing to automatically generate dynamic documents that integrates model statistics, interpretation and validation plots with inline code
- Developed a new scorecut automation to generate credit policy adapting post NITP v2 score launch under both volume neutral and risk neutral scenarios
- Created new R functions for team library, such as Teradata mload API and aggregated empirical probability density/ cumulative distribution functions

01/2019
|
08/2014

● Planning Analyst

Walt Disney Parks & Resorts

📍 Lake Buena Vista, FL

- Project lead for strategic facilities planning of Disney Springs and ESPN Wide World of Sports (\$15M annual)
- Developed BIRT user reports using SQL to query data and create custom functionalities with JavaScript under Eclipse IDE
- Performed data cleansing and migration from a project based (Maximo) to an equipment-based database application (Tririga)
- Initiated projects to replace sport field light fixtures, with a total savings of 44% of the original energy consumption

08/2014
|
09/2013

● Professional Intern

Walt Disney Parks & Resorts

📍 Lake Buena Vista, FL

- Surveyed conditions of various property assets (e.g. roofing, light poles, HVAC systems) and predict the timing of next major maintenance or upgrade
- Collaborated with an architectural consulting firm to transform Disney roof survey results into quantifiable data, used for plotting roof material degradation curves
- Collected data of approximately 920 projectors/screens across locations and created an inventory with comprehensive specifications and warranty documentation



RESEARCH & PROJECTS

Current

● R Shiny App for Forest Fire analysis

Github

📍 <https://brianli.shinyapps.io/Forest-Fire-Investigation/>

05/2012
|
01/2012

● Study of Biomass Torrefaction

Purdue University

📍 West Lafayette, IN

- Applied linear regression to determine the kinetic order of Fatty Acid Methyl Ester (FAME) production and optimize yield

05/2012
|
01/2012

● AIChE Chem-E-Car Design

Purdue University

📍 West Lafayette, IN

- Drove a Lego built pneumatic engine with CO₂ released in an acid-base reaction under controllable conditions. The team was awarded a second place (2/10) recognition in the regional competition held in the University of Akron