Brian Stockman

Hollywood, CA - Email me on Indeed: indeed.com/r/Brian-Stockman/cbd5e4473d8e2f4f

Willing to relocate: Anywhere
Authorized to work in the US for any employer

WORK EXPERIENCE

Data Analyst/Scientist

Oxford Road - Hollywood, CA - November 2016 to Present

Design and implement analytic processes in regards to measuring success and guiding strategy in offline media, as well as systematically identifying areas for optimization.
- Minimum Spend Analysis (Regression, Confidence Intervals)
- Lift Analysis (Regression, ARIMA Forecasting, Test & Control Groups)
- Market Entry Selection and Testing (Clustering, Correlation Matrices)
- Minute Level TV Attribution (Spike Level Model)
- Automated client reporting workflow from scratch in R (Google Analytics API, SQL database, data processing)
- Monte Carlo forecasting given variable spend, media mix, and other parameters.
- Built an optimization tool for any media plan including input variables such as budget, length of campaign, and flighting strategy. (optimized using bounded knapsack algorithm, powered by gradient boosted decision trees)
- Granular analysis of critical KPIs, with easily digestible and informative visuals (ggplot2, heatmaps, plotly)
- Converted all work/scripts mentioned above into easily accessible toolkits for analytical scale (RShiny)
- Frequently perform ad-hoc analysis upon request
- A/B Testing (Chi-Squared Test, Statistical Significance)
- Update Tableau dashboards for client facing reporting

Data Science Competitor

Kaggle.com - August 2015 to Present

Performed exploratory data analysis on several large data sets and leverage valuable insights to create accurate predictive models.
Competition Results in order of entry –
1) Avito - Duplicate Ads Detection - top 25% 11 million image files, 3 million ads w/text and metadata. Created a model that detects duplicate ad listings with 90% accuracy. NLP, text mining, regex, image hashing (dhash), gradient boosted decision trees.
2) Facebook - Predicting Check Ins - top 20% ; 7 million (time,coordinates,ID) tuples. Predicted check-in ID. Statistical modeling, grid slicing, K-nearest neighbors, gradient boosted linear regression.
3) Grupo Bimbo - Inventory Demand - top 17% ; 8 million deliveries and returns, thousands of clients, products, and routes. Predicted next weeks' product demand. Set-up and run cluster on Amazon Redshift, gradient boosted linear regression.
4) Talking Data - User Demographics - top 15% ; 3 million rows of cellular phone activity. Used apps and time to predict user demographics. Tm processing, Tf-idf, gradient boosted decision trees, model ensembling.
5) Redhat - Predicting User Business Value - top 3%; use anonymous user data to predict a binary business value. Heavy feature engineering and visual EDA, sparse matrices, gradient boosted linear regression.

EDUCATION

BS in Applied Mathematics

University of California San Diego - La Jolla, CA
September 2013 to June 2015

Engineering

Orange Coast College - Costa Mesa, CA
2009 to 2013

SKILLS

Excel (10+ years), Microsoft Office (10+ years), Microsoft Powerpoint (10+ years), R (3 years), Python (2 years), Javascript (4 years), C++ (2 years), SQL (2 years), HTML (4 years), Tableau (1 year)

LINKS

https://www.kaggle.com/brassmonkey381

CERTIFICATIONS/LICENSES

Machine Learning, Stanford Online

August 2016 to Present
This course provides a broad introduction to machine learning, datamining, and statistical pattern recognition. Topics include: (i) Supervised learning (parametric/non-parametric algorithms, support vector machines, kernels, neural networks). (ii) Unsupervised learning (clustering, dimensionality reduction, recommender systems, deep learning). (iii) Best practices in machine learning (bias/variance theory; innovation process in machine learning and AI). The course will also draw from numerous case studies and applications, so that you'll also learn how to apply learning algorithms to building smart robots (perception, control), text understanding (web search, anti-spam), computer vision, medical informatics, audio, database mining, and other areas.

Big Data in Education, Columbia University

September 2015 to Present
Introduction to EDM
Regressors
Classifiers
Classifying in RapidMiner 5.3
Case study in Prediction Modeling
Detector Confidence
Diagnostic Metrics: Kappa, Accuracy, ROC, A', Correlation, RMSE
Cross-Validation
Over-Fitting
Model/Detector Validity
Ground Truth for Behavior Detection
Data Synchronization and grain-sizes
Automated Feature Generation
Automated Feature Selection
Knowledge Engineering and Data Mining
Case study in Behavior Detection
Knowledge Inference
Bayesian Knowledge Tracing
Performance Factors Analysis
Item Response Theory

Q-Matrices
Knowledge Spaces
Advanced Knowledge Structure Inference
Relationship Mining
Correlation Mining
Causal Mining
Association Rule Mining
Sequential Pattern Mining
Network Analysis
Visualization of Educational Data
Learning Curves
Moment-by-Moment Learning Graphs
Learnograms
Heat Maps
Scatterplots
Parameter Space Maps
Flow Diagrams
Basic Clustering Algorithms
Advanced Clustering Algorithms
Validation of Clustering
Case Study in Clustering
Factor Analysis
Discovery with Models
Uses of Discovery with Models
Risks of Discovery with Models
Data Mining Data Mining
Future Directions in EDM

Explore Statistics with R, Karolinska Institutet

August 2015 to Present
How to import and clean data in R.
Find, import, visualize and clean data in R.
Statistics under the hood: distributions and tests.
Account for how statistical distributions can be used to design hypothesis tests.
Select and calculate a few common parametric tests with R.
Non-parametric tests.
Account for how resampling can be used when distributions are unknown.
Perform non-parametric tests with R.
Use published workflows to make statistical analysis in R.

ADDITIONAL INFORMATION

I am an aspiring data scientist with several years of education and experience in mathematics and computer science. My time spent studying Applied Mathematics at UCSD provides me the foundation to quickly comprehend and apply many of the most up-to-date data analysis and predictive modelling techniques available today. I am seeking a career position to further develop expertise and experience in data analytics, while helping to solve real-world business problems.

COMPUTER EXPERTISE:

• Built and maintained websites from scratch including systems incorporating PHP for data collection and SQL for data storage and access.
• Grew up using MS Office Suite. Proficient in MS Office, Excel (pivot tables/charts, vlookup, ...), and PowerPoint.
• Intermediate to advanced scripting skills in R, Python, and Javascript