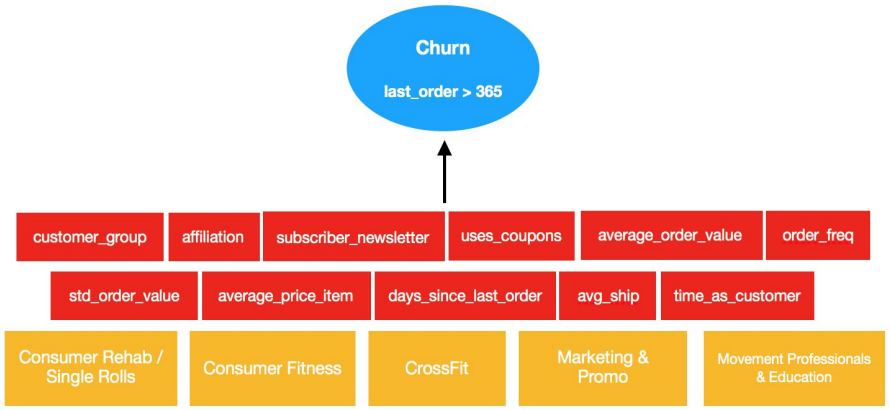
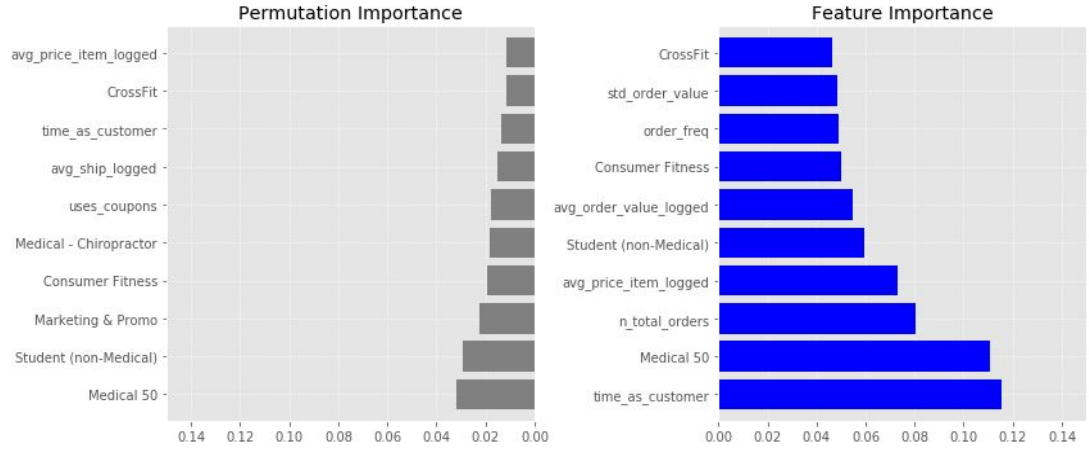
Dominique Vanden Dries (domvandendries@gmail.com)
GitHub: @domdomdidom

Analyzing Customer Lifetimes with eCom Data

It's much more cost-effective to keep existing customers than to recruit new ones, so wouldn't it be great if we had a way of knowing which customers are at risk of churning? I defined churn as not having ordered in the past 365 days, and I was able to correctly classify a customer around 80% of the time (relative to ~50/50 split). I used linear algebra (**Non-negative matrix factorization**) to extract latent features, and a **Random Forest Classifier** to model.





The real business use case of predicting churn is when we have *new* customers and we'd like to know *for how long they will be customers*. We don't have access to lengthy order histories for this task, so I masked my features to only consider a person's first order. By using a **Gradient Boosted Regressor**, I was able reduce the RMSE by 15%. I also created a webapp for this!

Strongest predictors of churn	Strongest predictors of lifespan
➔ The strongest predictors of churn **came from customer type**, not their purchasing behavior ➔ **Students** churned at higher rates than most other types ➔ People who receive **testers/samples** often don't follow through as customers	➔ New **chiropractors** are most likely to stick around longer ➔ **Using a coupon** with a first order positively correlates with lifespan ➔ **Buying expensive items** on a first purchase contributes negatively to lifespan

Dominique Vanden Dries
GitHub
Linkedin

Bay Area, CA
408.628.2315//domvandendries@gmail.com

PROFESSIONAL OVERVIEW

Excels in learning new technologies & implementing digital solutions. Data-driven, innovative worker who puts a premium on performance & is fully committed to contributing to the success of your marketing or analytics team.

SKILLS

Python	Git & GitHub	Email marketing	KNN
SQL	Some HTML and JavaScript	Wordpress	A/B Testing
ML Algorithms	Regression Modeling	Decision Trees & Forests	NLP

EDUCATION

Galvanize DSI
University of California - Santa Cruz, CA
B.S., Planetary Sciences, Graduated with Honors Thesis
Undergraduate Intern – Fisher Hydrogeology Lab

EXPERIENCE

Data Science Immersive, Jan 2019 – April 2019 Galvanize, San Francisco, CA
12 weeks of immersive data science coursework, case studies and projects that covered classification and regression modeling, machine learning, visualization, linear algebra and more.

Capstone Project

Analyzing customer lifetimes and predicting churn with eCommerce data. I used NMF, Random Forest Classification and Gradient Boost Regression to improve lifetime predictions by 15% and correctly predict churn with 78% accuracy and 88% precision. View my repo here. I also developed a locally-hosted webapp for this project.

Case Studies

Used Gradient Boost to identify parameters that contribute to churn for ride share data
Modeled tractor sales data with Linear Regression algorithm and identified factors that increased sale prices
Developed a movie recommender system based on Non-negative matrix factorization
Used NLP and text vectorization to classify events as fraudulent with event-hosting data

Marketing Analytics, Information Systems & Web Design, June 2017- Oct 2018 ROCKTAPE, Campbell, CA

- Oversaw data collection through our eCommerce platform: wrote scripts that chart sales statistics, warranty replacement rates, coupon redemption
- Served as the second lead developer for our WordPress site
- Executed all SEO, familiar with page ranking algorithms
- Created and A/B tested email campaigns in MailChimp and facebook
- Developed and carried out two customer satisfaction surveys resulting in an NPS score of 87
- Oversaw three BigCommerce stores. Built two BigCommerce stores, and one Shopify store
- Served as lead technical consultant to all departments - evaluated new IT systems, eCom platforms, plugins, data analytics tools, etc.

Undergraduate Research Assistant, Aug 2014 - June 2017 Fisher Hydrology Lab, Santa Cruz, CA