

RECENT WORK EXPERIENCE

Senior/Principle Data Analyst

April 2017 – present

Capital One Commercial Bank, Financial Institutions Group

- Developed risk models in R and Python to assess the loss distribution for leverage lending contracts. I developed the models for both our portfolio resiliency reporting and for underwriters to use when assessing the risk of a potential deal structure.
- Built machine learning models to predict Moody's risk ratings on unrated leverage lending collateral with a 93% accuracy, which was 3x more accurate than previous method.
- Wrote a data validation toolkit in Python to automate receiving, validating, and uploading manually captured data to a database. The package's modular structure allows users to apply a simple methodology to build out data pipelines. Users can switch out new connections and data input formats easily. This project is used in the FIG group to manage the monthly collateral information given to us by our clients across several books.
- Automated several reports that normally took 5 days into a daily automated service.

Data Science Intern

May 2016 - August 2016

Zurich Financial Services Group

- Created a machine learning model for corporate auto insurance contracts that is 12% more predictive than the standard industry model.
- Used Hive and SQL for data summarization, querying, and analysis
- Created web applications for managers to track timelines and costs of analytics projects

Marketing Analyst I

May 2014 - August 2014 & May 2015 - July 2015

Management Science Associates

- Created an automated process to generate Interactive Dashboards(link) with R and Javascript

Open Source Memberships

Stan: Core Team member, MLR: developer, quantmod: contributor

SELECTED PROJECTS

stanMathCL: GPU Routines for the Stan Programming Language

To be presented at Stancon 2018

This work is the first known open source implementation of the Cholesky decomposition using a GPU in a Hamiltonian Monte Carlo setting. Results show that large matrices can see speedups of 8x while retaining the same precision as CPU methods.

mlrfc: An R package for forecasting with machine learning models

Presented at R in Finance 2017: (click here for video, I am the first speaker)

R package combining methodologies of machine learning with classical forecasting techniques. Combining these models with machine learning methods for tuning hyper-parameters makes for more robust models. This package can also use machine learning methods for forecasting.

Columbus: GPU Routines for solving the traveling salesman problem

Presented at Columbia Electrical Engineering Student Conference (2016)

Columbus uses CUDA to implement a multi-core simulated annealing routine for solving large ($N > 100K$) Euclidian TSPs. This software can come within 3% of the optimal path with a 30x speed increase.

PANICr: Tests for Nonstationarity Within Large Panels of Time Series

Panel Analysis of Nonstationarity in Idiosyncratic and Common Components (PANIC) is a methodology for testing large panels of time series for pervasive or variable specific nonstationarity.







> SKILLS

git, R (5+ years), Python (5+ years), C++, unix, SQL, Fortran, C, OpenCL, CUDA, VBA, Excel



**Steve
Bronder**

Data Science





-  New York City, NY
-  sab2287@columbia.edu
-  +4127191282
-  stevebronder.com
-  linkedin.com/stevebronder
-  github.com/stevebronder

EDUCATION

M.A. in Quantitative Methods
Columbia University,
2015–2016
GPA: 3.6

BSBA in Economics
Duquesne University,
2010–2014
GPA: 3.2

INTERESTS

-  Open Source
-  Bicycling
-  Cooking
-  Dungeons and Dragons