Peng Chen

Data Science | Machine Learning | Management Consulting

814 810 8589 | pfc5098@gmail.com | pfc5098 | pfc5098

Skills

- Python - general programming, pandas, sklearn, keras, pyspark
- R - dplyr, ggplot2, tidyverse, tidymodels, torch, shiny, devtools
- SQL - logical execution order, temp tables/views, CTEs, window functions, performance improvement
- Experience with cloud services and big data platforms - AWS, Databricks (Apache Spark), H2O

Work Experience

Axtria - Ingenious Insights

Berkeley Heights, NJ

Senior Decision Scientist Supporting Pharma B

2021 - Present

- Developed a one-stop patient analytics solution that incorporates business funnel analysis, patient journey visualization, and machine learning based root cause discovery, using Databricks SQL, tidymodels, and flexdashboard, saving the client from hiring another machine learning team.
- Utilized network models to improve the accuracy of a hospital affiliation analysis by 33%.
- Collaborated with client and other consulting teams to continuously improve the accuracy of the market definition and sizing for a new product.

Axtria - Ingenious Insights

Berkeley Heights, NJ

Senior Decision Scientist Supporting Pharma A

2020 - 2021

- Improved the sensitivity of a new-customer discovery model from 60% to 90% while remaining its specificity above 90% by applying regularization, cross-validation, and downsampling techniques. Made the model run 20 times faster while reducing its total memory usage from 60% to 5%, by simplifying overcomplicated classification models and removing improper statistical tests, saving AWS EC2 cost.
- End-to-end automated several payer-stream standard reports using DBI and R Shiny, saving a five-person global team 10-day worth of work every month and enabling clients to make quicker data-driven decisions.
- Led a payer-stream consulting team of five. Timely responded to big volumes of client requests. Effectively communicated with clients to understand underlying business problems, proactively cleared roadblocks, and assigned tasks to offshore colleagues to form a seamless global workflow. Quickly took feedback from clients, iterated, and delivered satisfactory results.
- Developed a set of R packages, facilitating new members to pick up work quickly. These packages have been helping us stay away from repetitive tasks and focus on more creative ones.

Axtria - Ingenious Insights

Berkeley Heights, NJ

Senior Decision Scientist Supporting Internal Teams

2020 - Present

- Led the efforts to help our group migrate to GitHub for better code collaboration and intellectual property formation.
- Secured a new brand supporting project from Pharma B by demonstrating extensive analytical skills and experiences.
- Built Natural Language Processing (NLP) capacity for Axtria to help clients automatically detect adverse events.
- Helped to cultivate a data-centric culture within Axtria by recommending learning resources on data science and software engineering.

The Pennsylvania State University

University Park, PA

College Instructor of Stochastic Models in Operations Research (IE 425)

2019 - 2020

- One of the three Ph.D. instructors in this world-renowned IE department, among which I was in charge of the only core course.
- Provided 132 senior undergraduates with a solid analytical foundation in conditional probability, Poisson processes, Markov chains, queuing theory, inventory theory, dynamic programming, and basic Bayesian statistics via MCMC, by collaborating with two teaching assistants and utilizing independently developed lectures, computer simulations, homework assignments, and exams.

Education

The Pennsylvania State University

University Park, PA

PhD in Operations Research with minors in Statistics and Mathematics (GPA 3.9/4.0)

2015 - 2020

- Advisor: Enrique Del Castillo, Distinguished Professor of Industrial Engineering and Professor of Statistics
- Dissertation: Algorithms for Statistical Inference on the Optima of Nonparametric and High-dimension Regression Functions
- Co-authored statistical software: OptimaRegion, https://CRAN.R-project.org/package=OptimaRegion

Technical Training

- Mathematics: linear algebra, real analysis, functional analysis, numerical computations at Penn State University (PSU)
- Optimization: linear programming, convex optimization, nonlinear programming, statistical process optimization at PSU
- Statistics: linear models, experimental design, probability theory, statistical theory, stochastic process and MCMC, spatial statistics, time series at PSU
- Artificial Intelligence and Machine Learning: machine learning at PSU, deep learning specialization at deeplearning.ai