Nicholas Spyrison, Ph.D.

Machine Learning & High Dimensional Visualization

Parker, CO 80134
(720) 765-1836
spyrison@gmail.com
nspyrison.netlify.com/
github.com/nspyrison
linkedin.com/in/nspyrison/

Career Summary

A published data scientist with years of experience performing advanced analytics, statistical forecasting techniques, and building predictive models on diverse data including healthcare, manufacturing, accommodations, and information technology. Conducted end-to-end development on multiple R packages delivering free, open-source solutions to the community. Start to finish machine learning modeling has been has been core in recent years leading to several peer-reviewed articles. Nick is a frequently acts as a mentor to more junior technical personal leading to continuous improvement for the team sharing best practices, improving code performance, and driving project completion.

Employment History

06/2022 – present **Research Data Scientist**, *University of Queensland*, Clayton, VIC.
- Delivered batteries of models varying cohorts, model formulas, and hyperparameters tuning on patient-level health data
- Lead end-to-end analysis of the largest global COVID-19 patient-level data resulting in several submitted journal articles on kidney health outcomes

07/2018 – 10/2022 **Research Officer**, *Monash University*, Clayton, VIC.
- *Data Fluency Instructor*: Lead workshops (10-30 people) on statistical modeling, R, Python, PowerBI, & Git
- *Teaching Assistant*: Teaching tutorials on postgrad high dimensional data analysis, postgrad data visualization, business & economic statistics
- *Research Officer*: Conducted extract, transform, load, analysis, and reporting on census and text data

12/2016 – 1/2018 **Business Intelligence Developer**, *CPI Card Group*, Denver, CO.
- SQL reporting; across 4 reporting platforms and 4 database instances
- Mentor junior teammates and improve query performance
- Triage and transfer 30+ legacy reports to modern reporting solution
- Developed supply chain report tracking a cost reduction of more than $4M USD

07/2015 – 10/2016 **Business Intelligence Developer**, *UDR Real Estate Investment Trust*, Denver, CO.
- Transition enterprise footprint from WebFOCUS to SSRS – 10+ reports, 5 tables, 2 servers
- On-call for enterprise-wide ETL process

11/2014 – 04/2015 **Business Intelligence Analyst**, *Heartland Financial USA*, Dubuque, IA.
- Financial SQL reporting, adhoc and scheduled

02/2013 – 05/2014 **Delivery Analyst**, *IBM – International Business Machines*, Dubuque, IA.
- KPI reporting to teams and management
- Step-wise model selection with dozens of variables
- Automated ETL with VBA macros for about 0.5 FTE savings across the team
- IBM's lean/six sigma green belt, root cause analysis (RCA)

01/2010 – 08/2012 **Laboratory Technician**, *The Ames Laboratory & Iowa State University*, Ames, IA.

Authored Software (R Packages)

09/2022 **cheem** – Novel analysis and interactive visual for exploring ML interpretability *Link*

10/2020 **spinifex** – ggplot2-like API for composing & exporting animated projections *Link*

Education

02/2018 – 06/2022 **Ph.D. Information Technology & Statistics**, *Monash University*, Clayton, VIC,
Dimension reduction, interactive and animated data visualization, machine learning
Thesis: *Interactive and Dynamic Visualization of High-dimensional Data*.

- Novel interoperability visualization methods for **Fit random forest models on several contemporary datasets** also shown to generalize to other models with **SVM and XGboost**
- Utilized **SHAP and LIME local explanations** to explore local linear variable importance to understand large residuals in the models
- Created **novel interactive animation** facilitating this exploration in a web app
- **Authored an R packages** on CRAN as free and open source work in improve reproducibility and contribute back to the community

08/2008 – 05/2012 **B.Sc. Statistics**, *Iowa State University*, minors in Physics and Mathematics, Ames, IA.

Technical Skills

Models	Random forest, xgboost, neural networks, SVM, logistic regression, parameter tuning
Programs	R, SQL, Python, Git, GitHub, Excel, VBA
Packages	ggplot2, tidyverse, plotly, shiny, tidymodels, caret, keras, pandas, scikit-learn
Domains	Data visualization, dimension reduction, machine learning, interpretability
Reporting	SSRS, Tableau, PowerBI, Rmarkdown, Shiny

Awards and Honors

06/2022 Invited article for INFORM's OR/MS Tomorrow *Link*

04/2022 RStudio: March "Top 40" New CRAN Packages *Link*

05/2022 Monash Postgraduate Publication Award

10/2020 Melbourne Datathon 2020 – **1st place of several hundred entries** *Link*

- *Changes to Victorian Intraday Electricity Demand Following COVID-19 Restrictions*

11/2018 ACEMS Impact and Engagement Award

07/2018 UseR2018 Datathon – 3rd place

- Spatio temporal animated visualization of over 7 million avian sightings

2018 – 2022 Faculty Graduate Research Scholarship

2018 – 2022 Co-funded Monash Graduate Scholarship

05/2012 Nominated for Student Employee of the Year

05/2011 Nominated for Student Employee of the Year

05/2008 American Legion Award

Publications

Doctoral thesis

[1] Nicholas S Spyrison. "Interactive and dynamic visualization of high-dimensional data". en. PhD thesis. Monash University, Mar. 2022. URL: https://github.com/nspyrison/thesis_ns/blob/master/docs/thesis_ns.pdf.

Journal articles

[2] Nicholas Spyrison and Dianne Cook. "spinifex: an R Package for Creating a Manual Tour of Low-dimensional Projections of Multivariate Data". en. In: *The R Journal* 12.1 (2020), p. 243. ISSN: 2073-4859. DOI: 10.32614/RJ-2020-027. URL: https://journal.r-project.org/archive/2020/RJ-2020-027/index.html (visited on 10/16/2020).

[3] Stuart Lee et al. "The state-of-the-art on tours for dynamic visualization of high-dimensional data". en. In: *WIREs Computational Statistics* (Dec. 2021), p. 21. ISSN: 1939-0068. DOI: 10.1002/wics.1573. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wics.1573 (visited on 12/10/2021).

[4] Nicholas Spyrison and Dianne Cook. "Exploring Local Explanations of Nonlinear Models Using Animated Linear Projections". In: *arXiv preprint JMLR* (2022). DOI: 10.48550/ARXIV.2205.05359. URL: https://arxiv.org/abs/2205.05359.

Software (R packages)

[5] Nicholas Spyrison and Dianne Cook. *spinifex: Manual Tours, Manual Control of Dynamic Projections of Numeric Multivariate Data*. 2021. URL: https://CRAN.R-project.org/package=spinifex.

[6] Nicholas Spyrison. *cheem: Interactively Explore the Support of Local Explanations of a Model*. 2022. URL: https://CRAN.R-project.org/package=cheem.

[11] Nicholas Spyrison, Dianne Cook, and Kim Marriott. *A Study on a User-Controlled Radial Tour for Variable Importance in High-Dimensional Data*. arXiv:2301.00077 [stat]. Dec. 2022. DOI: 10.48550/arXiv.2301.00077. URL: http://arxiv.org/abs/2301.00077 (visited on 01/06/2023).

Conference proceedings

[7] Nicholas S. Spyrison. "spinifex: visualizing local structure of higher dimensions". en. In: Brisbane, Australia, 2018. URL: https://user2018.r-project.org/poster/ (visited on 07/28/2020).

[8] Madeleine Barrow, Jieyang Chong, and Nicholas Spyrison. "Changes to Victorian Intraday Electricity Demand Following COVID-19 Restrictions". en. In: *Melbourne Datathon 2020, Insights category*. Oct. 2020, p. 4. URL: https://github.com/nspyrison/melb_datathon2020/blob/master/_paper/paper.pdf (visited on 02/09/2022).

[9] Nicholas Spyrison, Benjamin Lee, and Lonni Besançon. ""Is IEEE VIS *that* good?" On key factors in the initial assessment of manuscript and venue quality". In: *IEEE AltVis Workshop*. type: article. July 2021. DOI: 10.31219/osf.io/65wm7. URL: https://osf.io/65wm7/ (visited on 08/21/2021).

[10] Nicholas Spyrison. "Animated Linear Projections". In: *Summer 2022*. INFORMS, May 2022.

References and publications available on request.