

TBOS

LTC Mike Powell, LTC Jon Paynter, and Dr. Diana Thomas

January 23, 2024

Abstract

We are pursuing a multifaceted faculty development opportunity for AY24. Interested parties may wish to participate for any number of reasons: Machine Learning Journal Club (Teach me some cool math!) / ORSA Preparation (Where do I start?) / Coding Exposure and Practice (I want more!) / Collaborative Writing (Sure, I'd be willing to contribute to a paper.). If any or all of these resonate with you, there's a place for you in TBOS.

1 Welcome to TBOS! (tē-bôs or TEE-BOSS)

What is TBOS? You might know it as “two birds one stone,” but in this case, it's more like three birds...or perhaps 10 birds.



Figure 1: Why get just one when you can get two...or more?

Collectively, we have several research and professional development objectives that happen to align quite nicely for this academic year.

2 What does TBOS do?

We plan to meet 7 (more) times during the semester in pursuit of the following objectives:

1. Write an invited review paper for *Annual Review of Nutrition* covering common machine learning methods and their applicability to nutrition research. (due November 2024)
2. March through the chapters of *Introduction to Statistical Learning* to familiarize ourselves with the underlying mathematics of common machine learning algorithms.
3. Follow along with the math seminars by doing the coding tutorials that accompany every book chapter (the books exists in R and Python versions - choose your favorite! ...or consider choosing your weakest?).
4. Write spin-off papers and publish nutrition-related machine learning tutorials as desired.

3 Where do I fit into TBOS?

There are lots of ways you can be part of TBOS.

1. You can simply come and go as you please. Drop in for a meeting and absorb whatever is presented.
2. You can volunteer to present one or more chapters of *Introduction to Statistical Learning* to the group.
3. You can help do some exploratory data analysis on candidate data sets for the proposed review paper.
4. You can write one or more sections of the review paper detailing a machine learning method and its utility in nutrition research.
5. You can prepare/work through code tutorials in R and/or Python to accompany the methods we cover in the seminar sessions. This would be adapting the existing R/Python tutorials in the textbook to utilize a nutrition data set.
6. You can serve as an editor to help consolidate and standardize the various prose or code contributions as part of the review paper.
7. You can help conduct a literature review documenting usage of the machine learning methods we're studying in a nutrition context.

4 What's the endstate?

Depending on the size and interests of the group, we aim to produce the following:

1. *A Review of Machine Learning for Nutrition Research* (the invited paper, which is the application area for the seminar series)
2. *Using Your Journal Club to Write a Paper* (this is what we're proposing we do...maybe we write up how it goes?)
3. *Upskilling Your Research Team During the Project* (this is what LTC Powell was intending to do for himself through this seminar series...maybe we write up the group's experiences in some way?)

In the end, we'll have covered a lot of ground:

1. Delivered a high-quality invited review paper on time.
2. Provided a nutrition-relevant set of code tutorials for a variety of machine learning methods.
3. Learned a bunch of new techniques (or at least at a deeper level).
4. Gained proficiency in one or more programming languages.
5. Practiced the discipline of academic writing.

5 What's next?

If you're interested in being part of this group, let LTC Powell know via email with subject line, "I'm a TBOS." This will add you to the distro for all TBOS events. You can already find "*TBOS Brown Bag*" appearing on the Math.All calendar.

Proposed schedule for the AY24-1 semester:

1. 12 SEP 23 - Introductory Meeting - Discuss group goals, ISLR/P resources, how to get set up with R/Python on your laptop, data sets (NHANES by default), and sign up to lead future meetings.
2. 19 SEP 23 - ISLR/P Chapter 2: Statistical Learning - Powell

3. 26 SEP 23 - ISLR/P Chapter 3: Linear Regression - Reder
 4. 04 OCT 23 - ISLR/P Chapter 4: Classification - Siekman
 5. 20 OCT 23 - ISLR/P Chapter 5: Resampling Methods - Bingman & Starling
 6. 01 NOV 23 - ISLR/P Chapter 6: Linear Model Selection and Regularization - Moore
 7. 15 NOV 23 - ISLR/P Chapter 7: Moving Beyond Linearity - Withenbury
 8. 08 DEC 23 - ISLR/P Chapter 8: Tree-Based Methods - Sherrell
-
9. 16 JAN 24 - ISLR/P Chapter 9: Support Vector Machines - Deverill
 10. 30 JAN 24 - ISLR/P Chapter 10: Deep Learning - Kassa
 11. 13 FEB 24 - ISLR/P Chapter 11: Survival Analysis and Censored Data - Moore
 12. 27 FEB 24 - ISLR/P Chapter 12: Unsupervised Learning - Lee
 13. 12 MAR 24 - ISLR/P Chapter 13: Multiple Testing - Powell
 14. 09 APR 24 - Writing Workshop I
 15. 23 APR 24 - Writing Workshop II
 16. 07 MAY 24 - Writing Workshop III

What does a typical meeting look like? That's up for discussion, but as a starting point, let's assume the presenter will do the following:

1. Provide the motivation for the topic. What kinds of problems is this technique suitable for addressing?
2. How is this model typically written in mathematical notation? What do all those symbols mean? What is the model going to look like if we encounter it in a textbook or journal article?
3. What are the assumptions or validity conditions we should care about when using this model?
4. How do I know if this model worked well? How do we quantify fit/performance for this model?
5. Show us an example (could be the textbook example) of this technique in action.
6. (BONUS) Show us an example of this technique using nutrition data.
7. (BONUS) Provide some working R/Python code to demonstrate this technique using nutrition data.

You can't cover everything in a textbook chapter in a single lunch seminar. Fortunately, there's no expectation that you do so. Your job as a presenter is to help the audience gain enough familiarity with the topic that they can confidently go learn the rest of what's needed to correctly employ the presented technique. You want to knock down the first wall they face by demystifying the topic a little bit and pointing them to good resources to learn more.