Machine Learning in the News

(or, Why is SLG Important?)

Todd Wilson

Statistical Learning Group
Department of Statistics
North Carolina State University

September 29, 2017

Presentation Outline

- What is SLG?
- What types of topics will be covered at SLG?
- Two roads that you can take...
- Examples of machine learning in the news
- With each example, some critical thinking
- ...so which road will you take?

SLG

- Statistical Learning Group founded in Fall 2014
  - Founding students have graduated
  - I ran the group last year
  - This year, Kara and I are running it together
- What are our goals?
  - Space to learn about learning
  - Presentation skills, CV builder, etc.
  - "Low-risk student seminar"

SLG FAQ

- What can I talk about? A: Anything!
- Clarification: We'd like to trace your topic back to machine learning (or more generally, computation) if we can.
- What do I wear/use/say? A: Anything!
- Does SLG count for seminar credit? A: ...for your soul, yes.

SLG, Fall 2017

- Three speakers are lined up
- We hope for six speakers each semester
- Be prepared for me to recruit you
- Final projects, prelims, interests...

Yes. But *why?*

- The most exciting point of your week happens: Kara emails you reminding you of an SLG talk!
- Frost: "Two roads diverged in a wood..."
- In one world, you decide to "do work"
- In another, you attend the SLG talk

The Wonders of World Two

- Suppose SLG has a talk about neural networks
- Mathematical formalization, example code, etc.
- Historical persepective: literally about *neurons*
- In a short hour, you have an awesome grasp of a new subject

The Roads Converge Again

- You go home and unwind
  - Social media
  - YouTube
  - Evening news

- **Important Question:** Who is better prepared for understanding what they might see next on a deep level?

Case Study I: @DeepDrumpf

- Can this technology be used for good?
- How might this technology serve to confuse people? Can we afford this confusion?

Case Study II: YouTube Demonetization

- We don't know the exact algorithm that Google uses to demonetize videos or even place ads on videos in the first place.

- But, we do know the algorithm that they use to prioritize videos in the first place.

- Would a knowledge of this algorithm help to inform content creators about how to retain monetization of their content?

Case Study III: Facial Recognition

- iPhone X will use facial recognition technology
- Kosinski and Wang (2017) claim to use facial recognition to predict sexual orientation with great accuracy
- What are the limitations of this study? How important is it for us as statisticians to make sure these limitations are reported?
- Is this application of facial recognition ethical? Is it discriminatory?

Summary

- It is clear that neural networks (and, more broadly, machine learning) are becoming an increasing part of our lives.
- What is the role of statistics, and of the statistician, in the wake of news stories like these?
- So, which version of you was better prepared for the news? For life?
- ...take the SLG road. It could make all the difference.

References

- Covington, P., et. al. "Deep neural networks for YouTube recommendations." *Proceedings of the 10th ACM Conference on Recommender Systems*, ACM, New York, NY, USA, 2016.
- Kosinski, M. "Authors' Note: Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images." Google Docs, Google, docs.google.com/document/d/11oGZ1Ke3wK9E3BtOFfGfUQuuaSMR8AO2WfWH3aVke6U/edit#.
- Kosinski, M., and Wang, Y. "Deep neural networks are more accurate than humans at detecting sexual orientation from facial images." *Journal of Personality and Social Psychology*, in press.
- Murphy, M. "MIT Built a Donald Trump AI Twitter Bot That Sounds Scarily like Him." Quartz, 4 Mar 2016, qz.com/631497/mit-built-a-donald-trump-ai-twitter-bot-that-sounds-scarily-like-him/.
- Think Tank. "AI Can Probably Tell If You're Gay." YouTube, 12 Sept 2017, www.youtube.com/watch?v=khTzoJi8hbk.