

The Mythos of Model Interpretability

Dr. Bean - Stat 5100

Questions from Canvas Quiz

- According to Section 3, in what ways could linear models be considered “interpretable”?
- According to Section 4, why are linear models not necessarily more interpretable than complex neural networks?
- Please articulate at least one thing you learned as a result of studying this article.

Additional Discussion Questions

In what ways is a linear model “interpretable”? In what ways is it NOT interpretable?

Why should we care about interpretability?

How might a statistical model be racist/prejudiced? Are linear models also prone to prejudice? If so, how?

How might statistics be used to ensure that the predictive models of tomorrow don't suffer from the prejudices of yesterday?

Interesting Articles

- <https://www.npr.org/sections/live-updates-protests-for-racial-justice/2020/06/24/883107627/boston-lawmakers-vote-to-ban-use-of-facial-recognition-technology-by-the-city>
- <https://www.npr.org/2020/06/24/882678392/man-says-he-was-falsely-arrested-after-facial-recognition-mistake>
- <https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-ar-kansas-cerebral-palsy>