Homework: Identifying and Explaining Predictive Fallacies.

Total Points = 25 points

1. You train a model with 50 predictors on 500 observations and achieve 99% accuracy on training data. (points 5)
   * Identify the fallacy(2 points)
     1. Overfitting
   * Explain why this model will likely fail on new data (3 points)
     1. The model looks great on training data because it basically memorized the patterns, but that doesn’t mean it learned anything useful. On new data, the accuracy will likely drop because the model can’t generalize well
2. A researcher tests 100 variables against customer churn and finds 3 “statistically significant” predictors at p < 0.05. (points 5)
   * Identify the fallacy (2 points)
     1. Publication Bias
   * Explain why this is misleading (3 points)
     1. When you test a huge number of variables, some will look significant by random chance. Highlighting only the three “significant” ones and ignoring the other 97 makes the results look more meaningful than they actually are.
3. A VC fund studies only successful startups to design a “success formula.” (points 7)
   * Explain the fallacy (2 points)
     1. Survivorship Bias
   * Give one real-world predictive analytics context where ignoring “failures” would distort insights ( 5 points)
     1. If hospitals only study patients who recovered, they’ll overlook the traits of patients who didn’t make it, which would make predictions way too optimistic.
4. Your Example (8 pts)
   * Choose one fallacy ( 2 points)
     1. McNamara Fallacy
   * Define it (2 points)
     1. Relying solely on metrics in complex situations and losing sight of the bigger picture.
   * Give a real-world predictive analytics example (2 points)
     1. A college evaluates professors only based on student test scores. This overlooks things like mentoring, research quality, and classroom engagement, which are just as important for education.
   * In ≤60 words, explain how you would avoid it ( 2 points)
     1. I’d combine both quantitative measures, like grades, with qualitative feedback, such as peer reviews and student surveys. This gives a fuller picture and avoids relying only on numbers.