

# **Model, Method and Feature (i.e., variable) Selection**

# What yields the best predictive model?

- 4 • The most **sophisticated** and **complex** model?
- 4 • The most **parsimonious** (i.e., simple) model?
- 3 • The least **biased** model?
- 3 • The model that **better fits** the **data**?
- 2 • The most **accurate** model?
- 2 • The most **interpretable** model?
- 1 • The model that **fulfills** the data and model **assumptions**?
- 1 • The highest **data quality** possible?
- 1 • The most appropriate (business) **selection of predictors**?
- 4 • Quadratic and other **transformations** that fit data the best?
- 5 • **All of the above?**

# Justifying the Selected Model/Method

1. **Classification** vs. **quantitative** prediction?
2. Is **goal** hypotheses testing, interpretation or prediction?
3. **Tree** vs. **regression** approach?
4. Best method based on model **assumptions**
5. Best method based on predictive **accuracy** (cross-validation → lowest MSE)
6. Best model based on **explanatory power** (e.g., **R<sup>2</sup>**, **MSE**)
7. Model **feature** (i.e., variable) **selection** based on:
  - **Business** domain knowledge
  - Model **diagnostics** (e.g. **R<sup>2</sup>**, collinearity, serial correlation)
  - Variable **selection** (e.g., stepwise) vs. **shrinkage** (e.g., ridge)
  - **Variance** (over-identification) vs. **Bias** (under-identification)
  - **Sophistication** (complexity) vs. **Parsimony** (simplicity)

# Final Notes

- The main **learning goal** in this course is for you to learn how to select, justify and employ the most appropriate model to answer a predictive analytics question
- There are so **many modeling options** that we could not possibly cover all of them in a class (or even in a full masters program).
- But you will develop a good **understanding** of the modeling **options** available.
- We will cover some methods **in depth**
- And provide a **higher level overview** for others
- With the knowledge gained in this class you should be able to **learn** several other modeling approaches **on your own**
- You will also learn how to use **R** for modeling methods covered in the course, but you will also be able to explore other modeling methods in **R on your own**.





KOGOD SCHOOL  
*of*  
BUSINESS

