Kobe Bryant Basketball Shot Selection

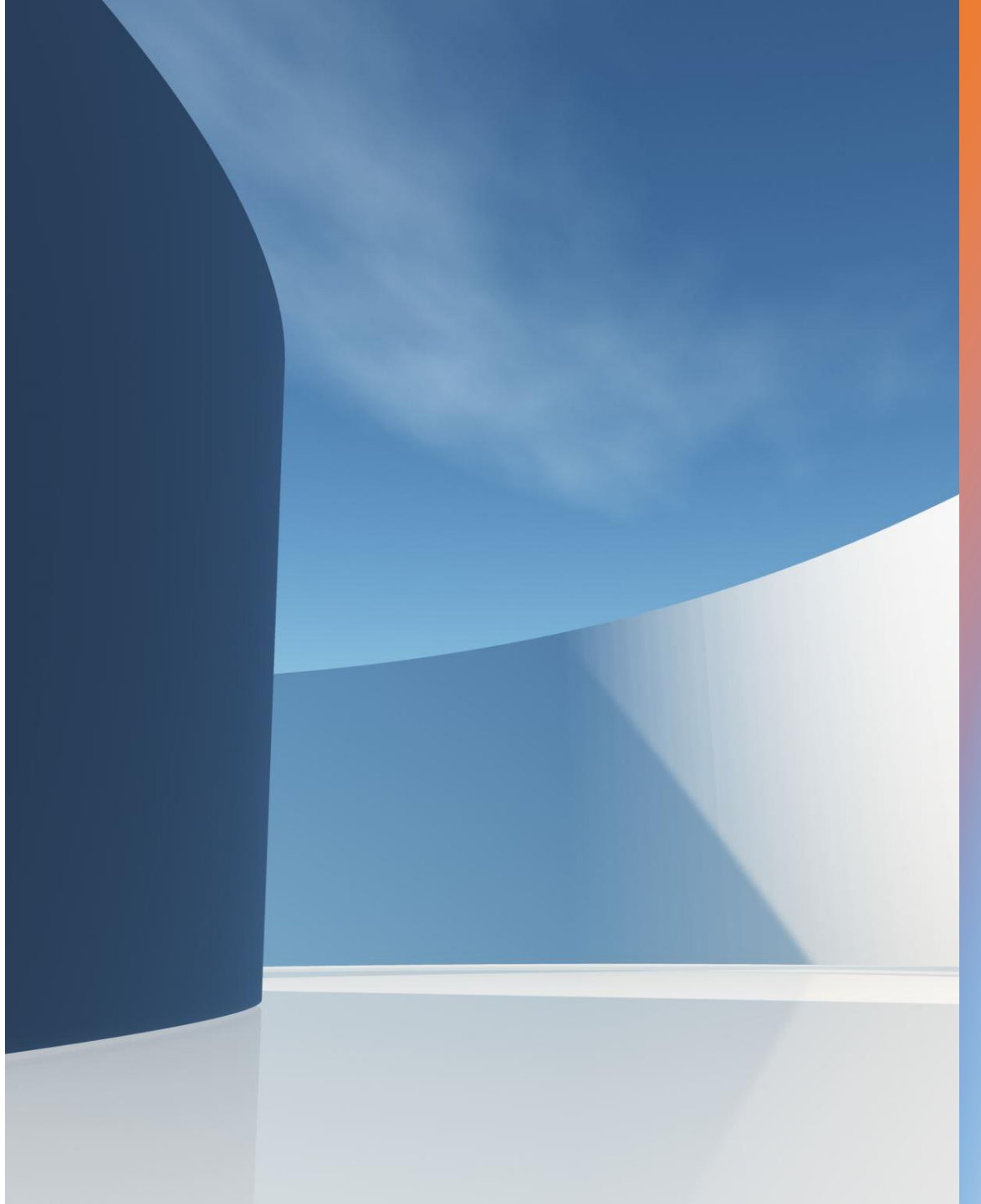
Joseph Bradley

Problem Background

The problem is simple and provided by Kaggle: Using 20 years of data on Kobe's swishes and misses, can you predict which shots will find the bottom of the net? This competition is well suited for practicing classification basics, feature engineering, and time series analysis. Practice got Kobe an eight-figure contract and 5 championship rings. What will it get you?

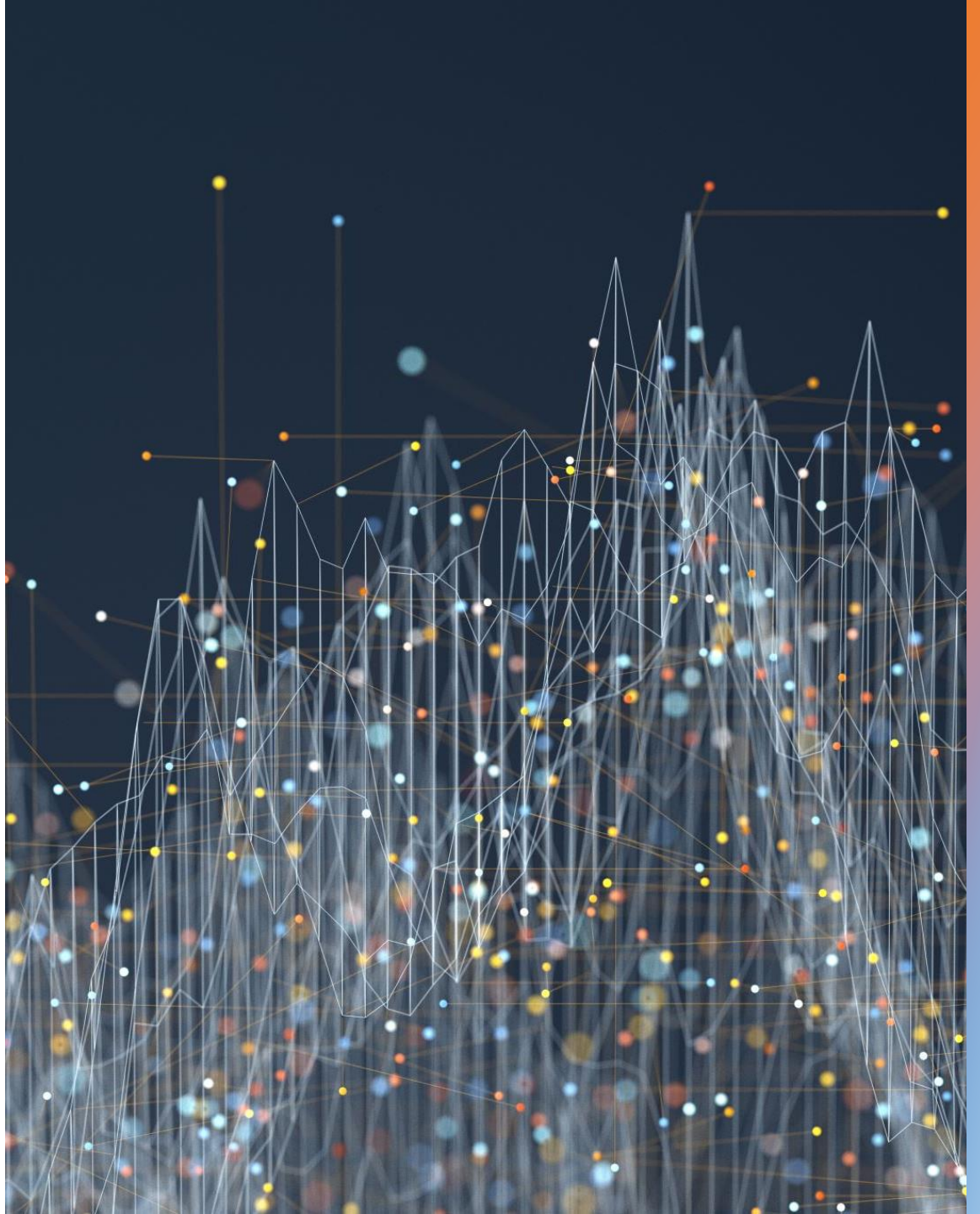
Description of the Feature Engineering

- Make Training and Data Sets
- Made 4 Variables
- Dropped 12 variables
- The final model used 15 variables out of the original 23
- Recipe made dummy variables for all non-numeric variables



Model Comparison

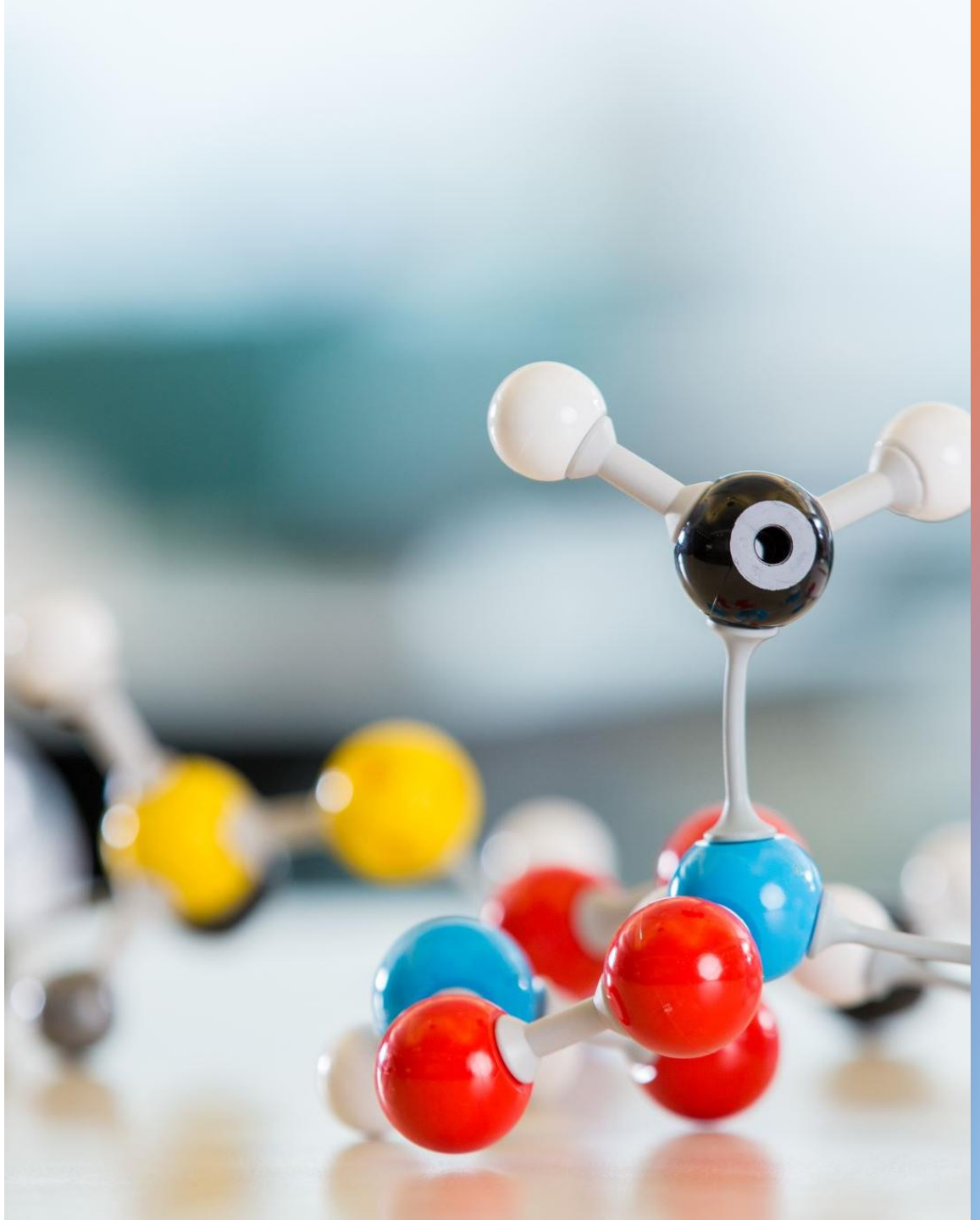
- Random Forest (Best)
- Penalized Regression (Third Best)
- BART (Second Best)



Description of the Best Model

- How do Random Forests work?
  - Regression Tree
  - Multiple Trees

- Used 500 trees
- Tuned using expand.grid() function



Final Score

0.60053

Threshold: 0.602