Modeling Pipeline

Modeling Tradeoffs

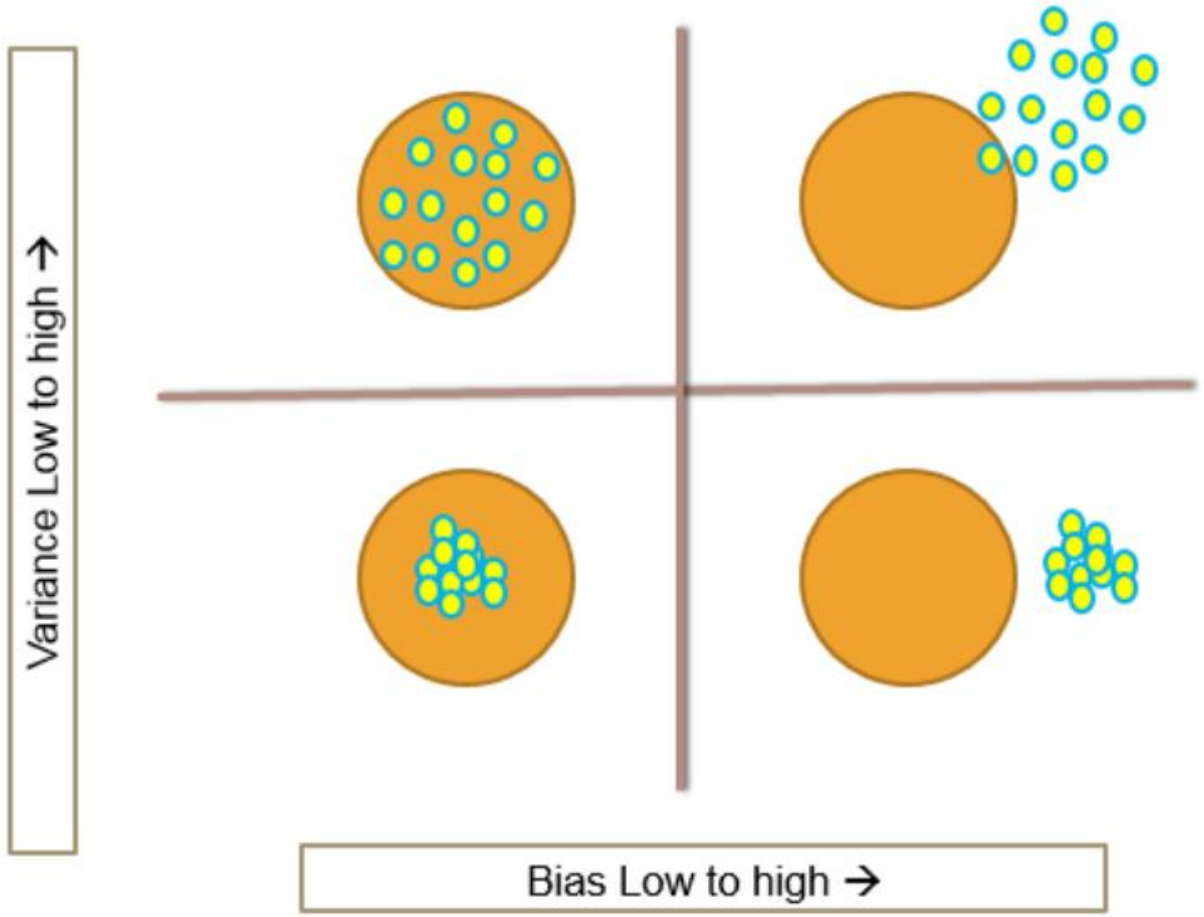
- Bias vs. Variance
- Inference vs. prediction
- Interpretability vs. flexibility
- Accuracy vs. performance (time)
- Optimization vs generalization

Bias Variance Decomposition

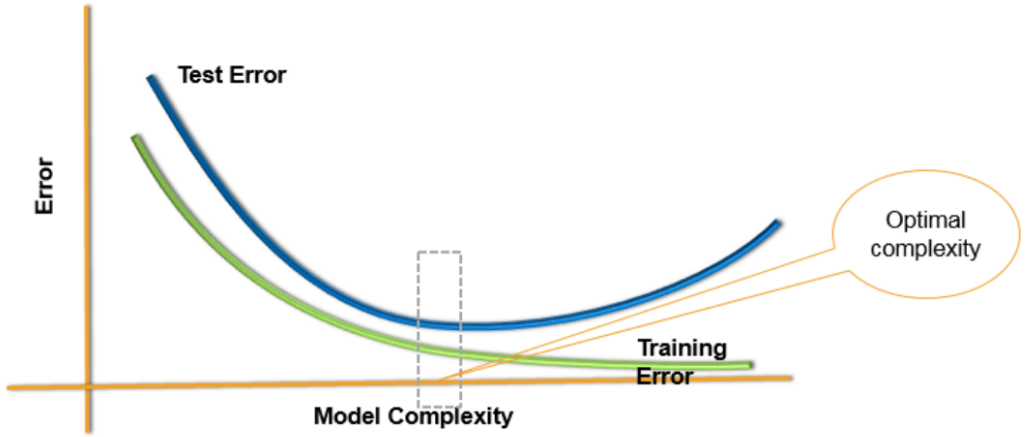
$$
\begin{aligned}
\mathrm{MSE}(x_0) &= \mathrm{E}_{\mathcal{T}}[f(x_0) - \hat{y}_0]^2 \\
&= \mathrm{E}_{\mathcal{T}}[\hat{y}_0 - \mathrm{E}_{\mathcal{T}}(\hat{y}_0)]^2 + [\mathrm{E}_{\mathcal{T}}(\hat{y}_0) - f(x_0)]^2 \\
&= \mathrm{Var}_{\mathcal{T}}(\hat{y}_0) + \mathrm{Bias}^2(\hat{y}_0).
\end{aligned}
$$

https://artofsoftware.org/2012/09/13/derivation-of-bias-variance-decomposition/

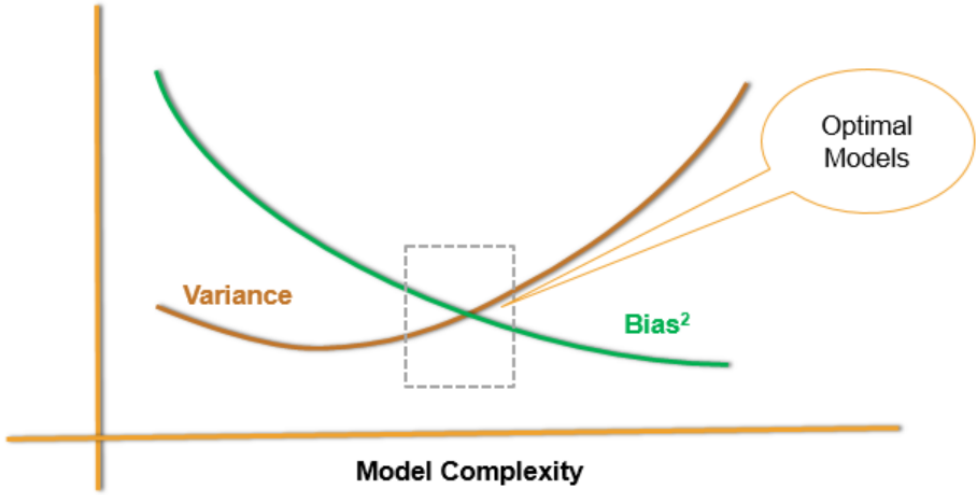
Bias and Variance



Test and Training Error

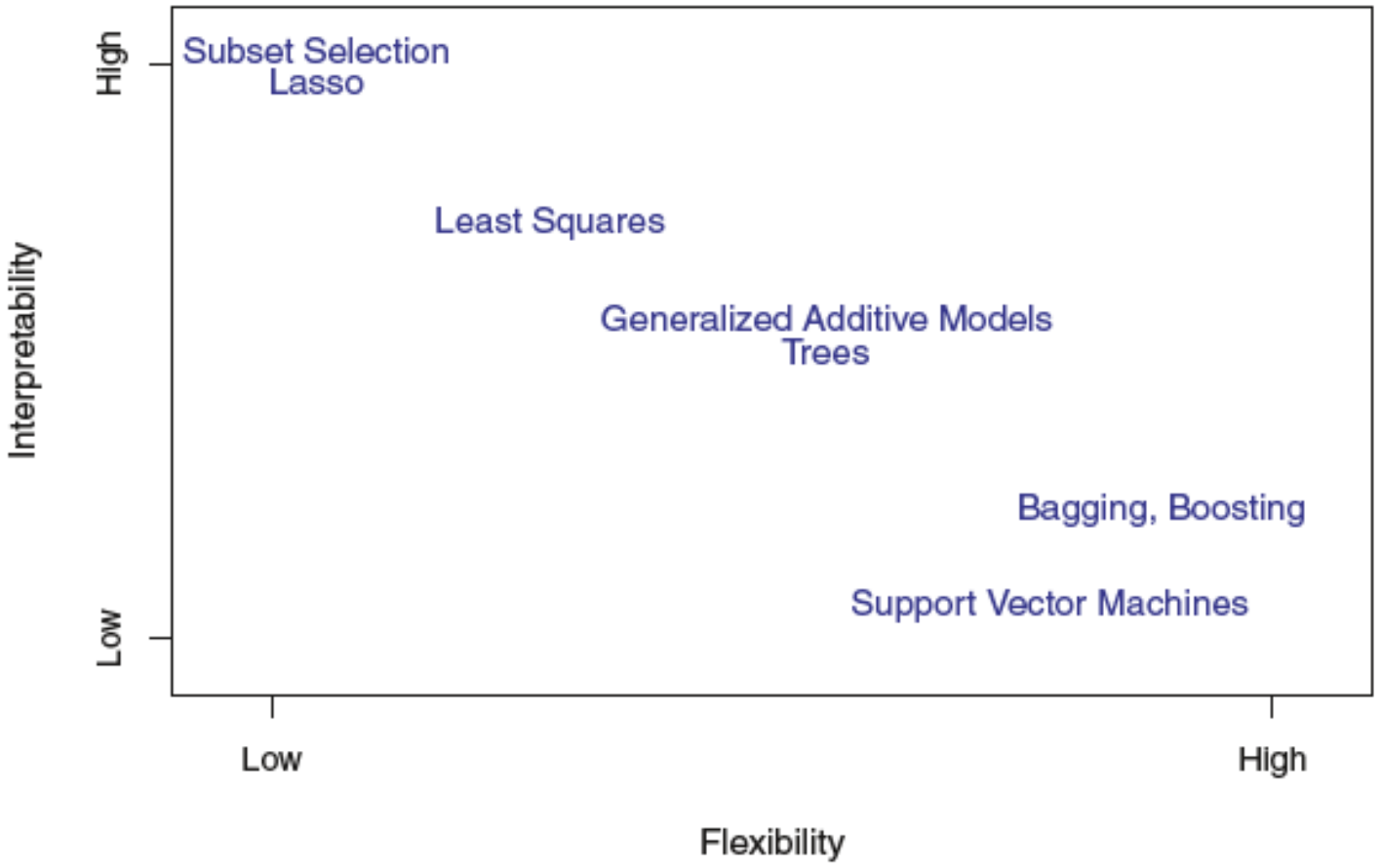


Bias Variance Tradeoff



Optimal Bias-Variance Tradeoff

Interpretability vs Flexibility



Statistics vs Machine Learning

Traditional Statistics **Machine Learning**

White-box modelling
simpler computation, emphasis on introspection, form, causal effects and processes, finding a 'correct' model

Black-box modelling
high computational complexity, emphasis on speed and quality of prediction, finding a 'performant' model