

# Week 3-5

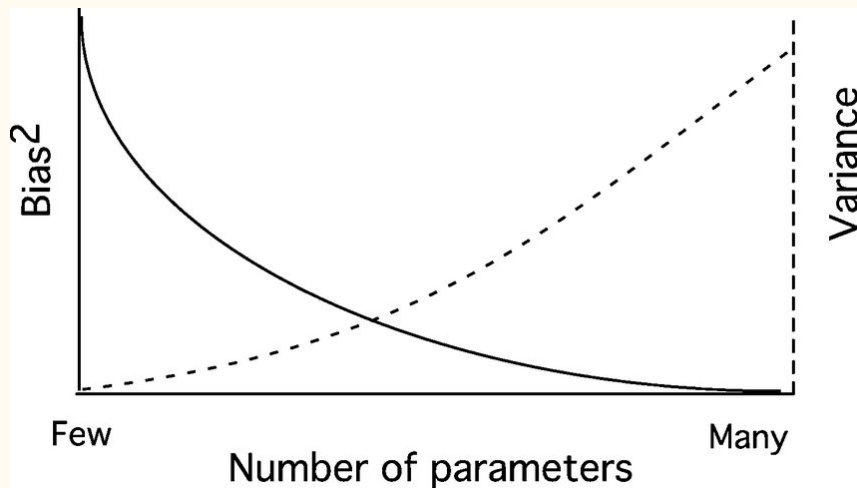
---

Regression Models, Bias- Variance Tradeoff, Feature Selection and Model Parsimony

# Linear Regression

# Linear Regression

- Example in Python.
- Bias Variance Tradeoff
- Parsimony : Simple models with less parameters



# Feature Selection

# Why do it and How?

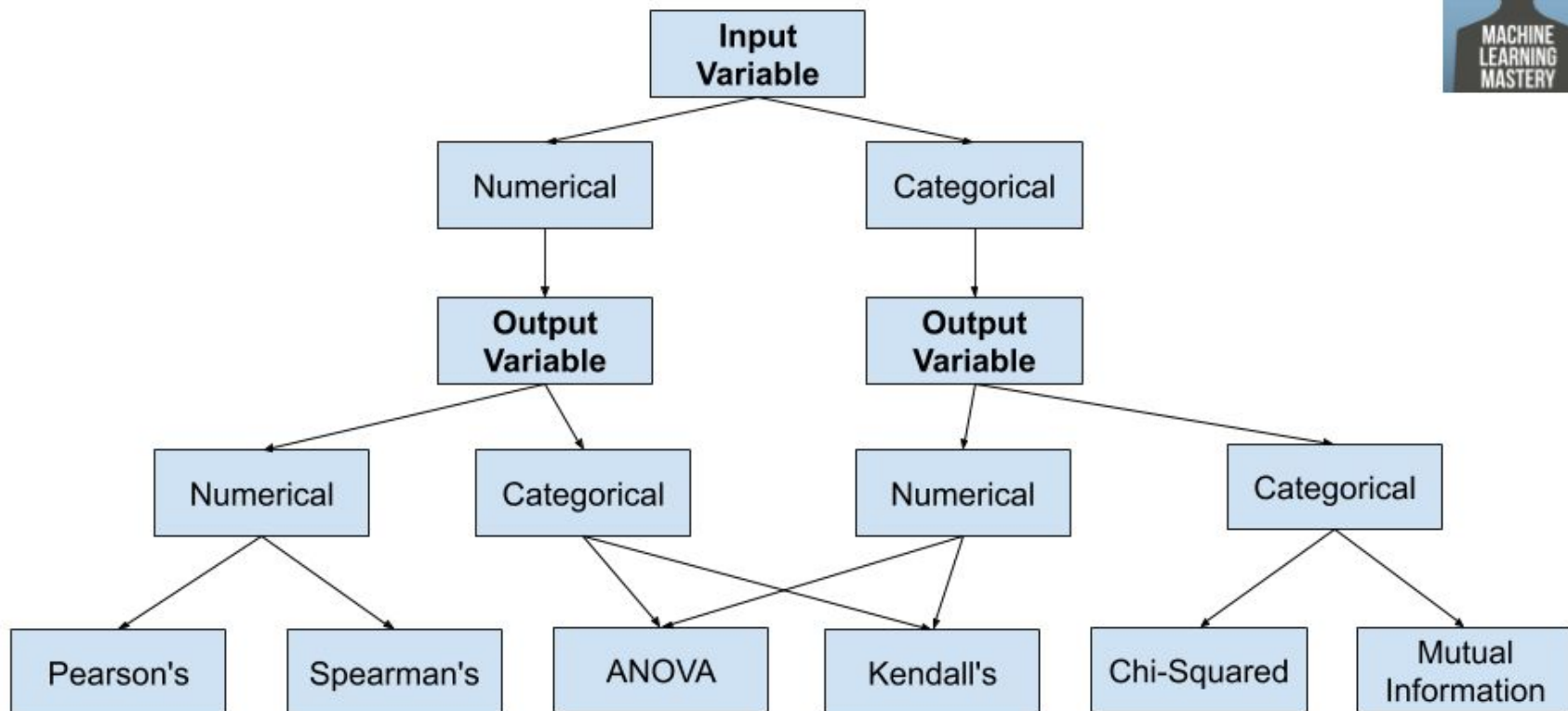
## Why?

- Curse of Dimensionality
- Occam's Razor: We lose explainability with too many features.
- Overfitting

## How?

- Filter Based
- Wrapper Based

## How to Choose a Feature Selection Method



# Scikit Learn Libraries

- Pearson's Correlation Coefficient: `f_regression()`
- ANOVA: `f_classif()`
- Chi-Squared: `chi2()`
- Mutual Information: `mutual_info_classif()` and `mutual_info_regression()`

## Selection:

- Select the top k variables: `SelectKBest`
- Select the top percentile variables: `SelectPercentile`

# Feature Selection to Model Parsimony