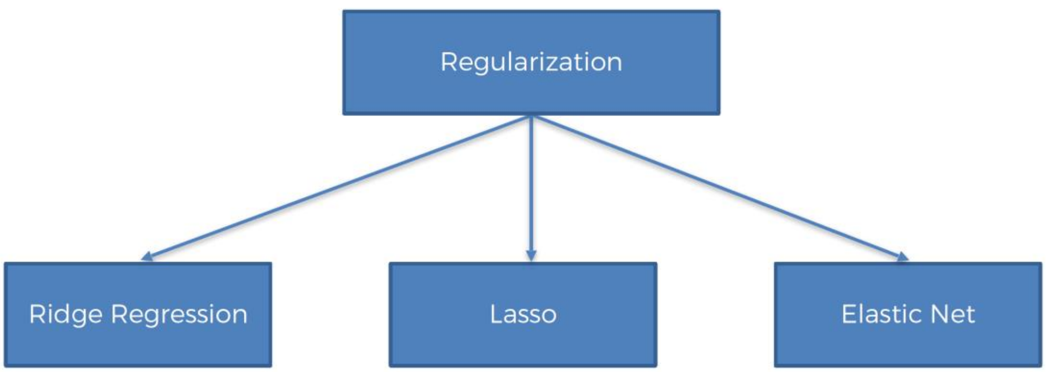
Regression

Regression Model	Pros	Cons
Linear Regression	Works on any size of dataset, gives informations about relevance of features	The Linear Regression Assumptions
Polynomial Regression	Works on any size of dataset, works very well on non linear problems	Need to choose the right polynomial degree for a good bias/variance tradeoff
SVR	Easily adaptable, works very well on non linear problems, not biased by outliers	Compulsory to apply feature scaling, not well known, more difficult to understand
Decision Tree Regression	Interpretability, no need for feature scaling, works on both linear / nonlinear problems	Poor results on too small datasets, overfitting can easily occur
Random Forest Regression	Powerful and accurate, good performance on many problems, including non linear	No interpretability, overfitting can easily occur, need to choose the number of trees

Examples of Regularization



$$\text{Minimize} \sum_{i=1}^{n} \left(y^i - (b_0 + b_1 x_1^i + \ldots + b_m x_m^i)\right)^2$$

Ridge Regression:

$$\text{Minimize} \sum_{i=1}^{n} \left(y^i - (b_0 + b_1 x_1^i + \ldots + b_m x_m^i)\right)^2 + \lambda(b_1^2 + \ldots + b_m^2)$$

Lasso Regression:

$$\text{Minimize } \sum_{i=1}^{n} \left(y^i - (b_0 + b_1 x_1^i + \ldots + b_m x_m^i)\right)^2 + \lambda(|b_1| + \ldots + |b_m|)$$

Elastic Net:

$$\text{Minimize} \sum_{i=1}^{n} \left(y^i - (b_0 + b_1 x_1^i + \ldots + b_m x_m^i)\right)^2 + \lambda_1(|b_1| + \ldots + |b_m|) + \lambda_2(b_1^2 + \ldots + b_m^2)$$