**MOOC 2-MODULE 4**

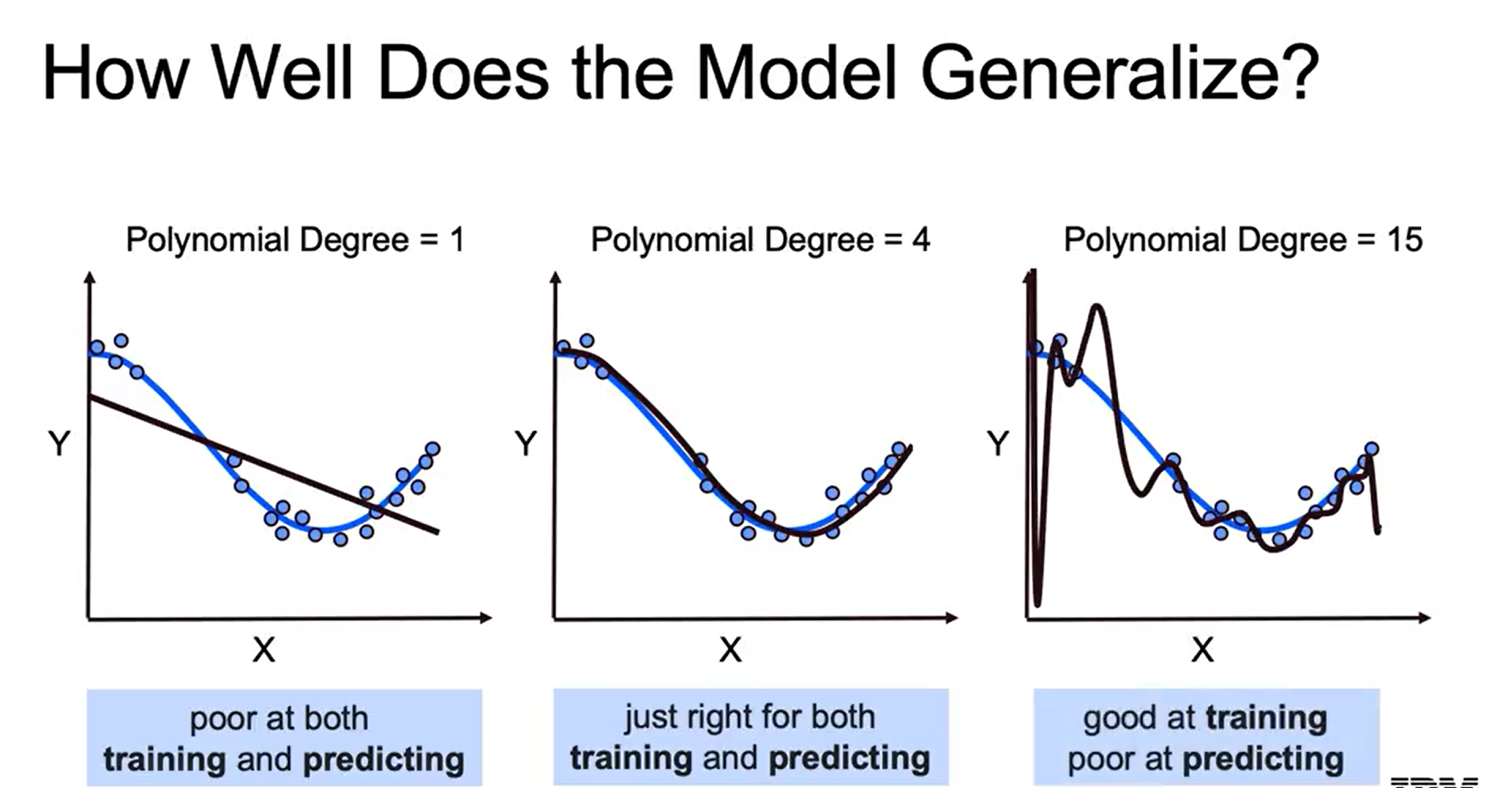
[**Supervised Machine Learning: Regression**](https://www.coursera.org/learn/supervised-machine-learning-regression/home/welcome)

### **I.Bias Variance Trade off**

### Understanding Model Complexity and Error

### The goal is to minimize both training and test errors by finding the right model complexity.

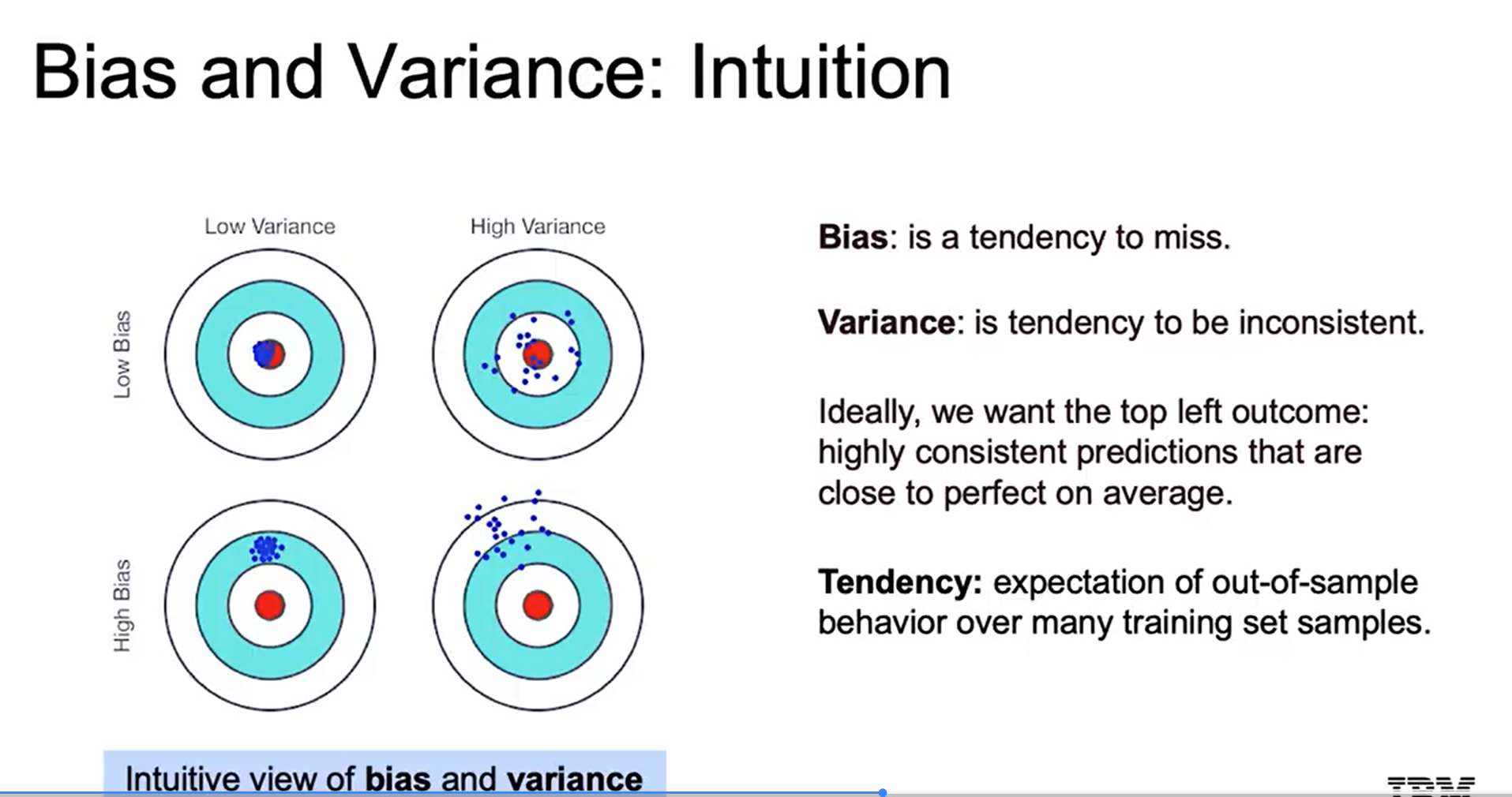
### A model that is too complex may perform well on training data but poorly on test data, while a model that is too simple will have high errors on both.



### Defining Bias and Variance

### Bias refers to the tendency of a model to consistently miss the true values, often due to oversimplification (underfitting).

### Variance indicates the model's sensitivity to fluctuations in the training data, leading to inconsistent predictions (overfitting).



### Sources of Model Error

### Errors can arise from three main sources: incorrect model assumptions (bias), instability in predictions (variance), and unavoidable randomness in data.

### High bias models miss real patterns, while high variance models capture noise, leading to poor generalization on new data.

### Bias and Variance

### High bias occurs with overly simple models that fail to capture underlying patterns, leading to underfitting.

### High variance arises from overly complex models that are sensitive to fluctuations in the training data, resulting in overfitting.

### 

### **II.Regularization and Model Selection**

### Regularization Techniques

### Key methods include Ridge and Lasso regression, which add penalties to the cost function to manage model complexity.

### Elastic Net combines features of both Ridge and Lasso, providing a balanced approach.

### Feature Selection and Elimination

### Regularization also aids in feature selection by reducing the impact of less important features, which can improve model generalization.

### Recursive feature elimination can be used to systematically remove features that do not contribute significantly to the model's predictive power.

### 

### **III.Ridge Regression**

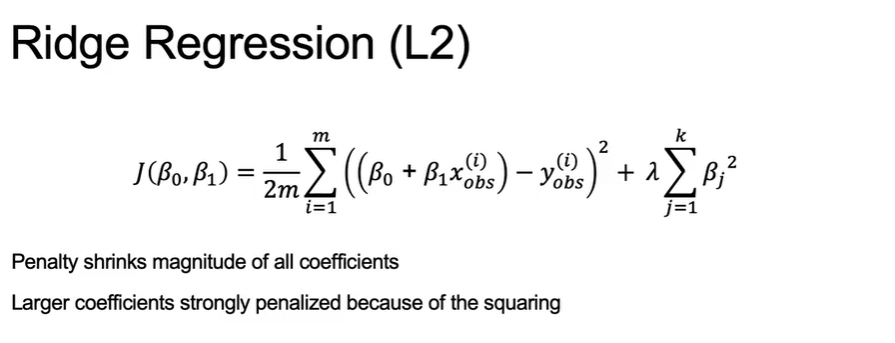
### 

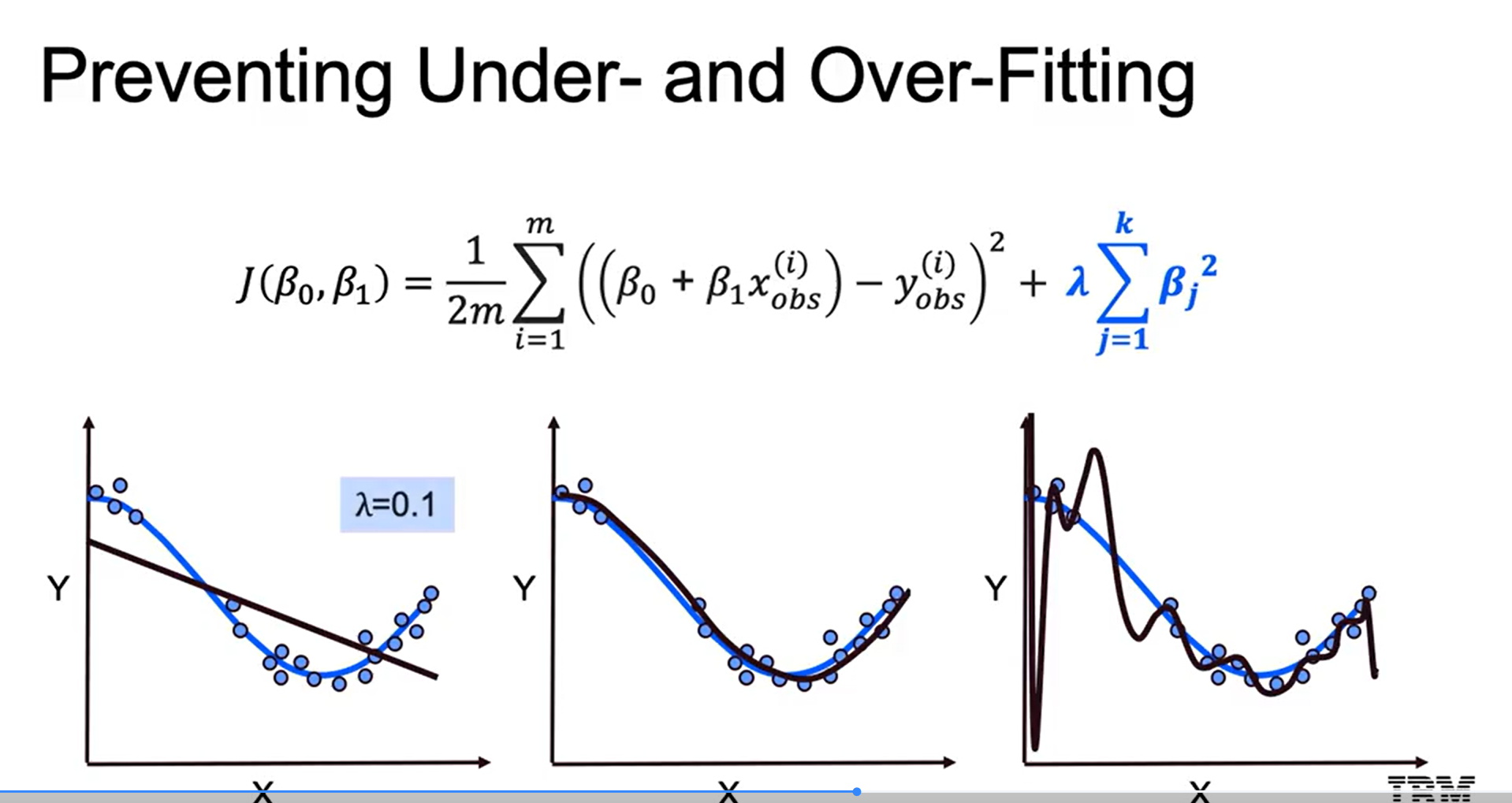
### Cost Function and Penalty

### The cost function includes the residual sum of squares and a penalty term, which is the square of the coefficients multiplied by a chosen lambda value.

### This penalty discourages overly complex models by increasing the cost as the coefficients grow larger.

### 



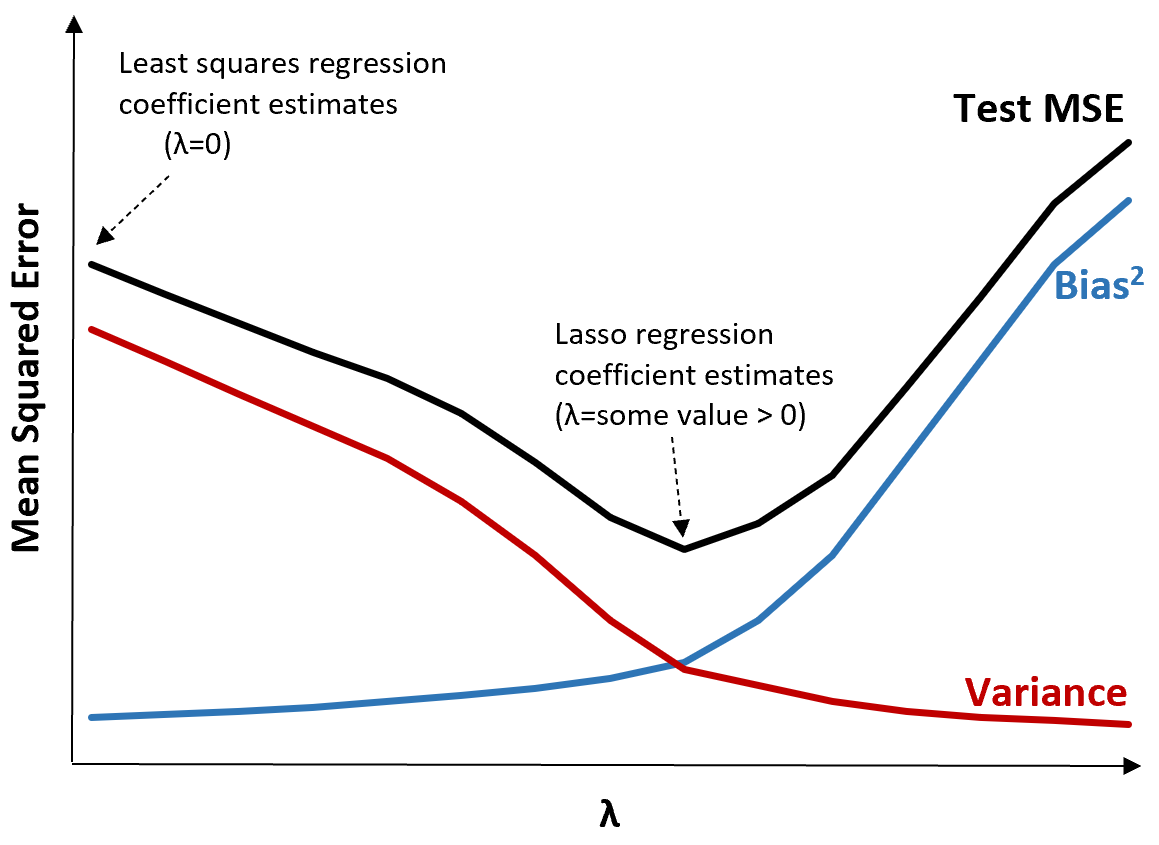


Effect of Lambda on Coefficients

* The penalty term shrinks coefficients towards zero, reducing model complexity and variance, while introducing some bias.
* The relationship between lambda and coefficients is inverse; as lambda increases, the coefficients generally decrease.

### **IV.Lasso Regression**

### LASSO regression is a regularization technique used in linear regression that differs from Ridge regression primarily in how it penalizes the cost function.



### LASSO vs. Ridge

### LASSO uses the absolute values of coefficients for regularization, while Ridge uses the squared values.

### LASSO is more effective in feature selection, often zeroing out certain coefficients, whereas Ridge tends to shrink all coefficients.

### Regularization Effects

### Increasing the lambda parameter in LASSO raises bias but reduces variance, helping to manage model complexity.

### LASSO can eliminate features, leading to a simpler model that may enhance interpretability.

### Elastic Net

### Elastic Net combines LASSO and Ridge, allowing for a hybrid approach to regularization.

### It uses two lambda values to balance the penalties from both L1 (LASSO) and L2 (Ridge) norms, optimizing model performance based on specific needs.

### RFE Overview

### RFE is a tool provided by sklearn for recursive and automatic feature selection.

### It requires the user to select a model and define the number of features to retain.

### RFE Process

### The model chosen must have coefficients or feature importances for RFE to work.

### RFE measures feature importances and recursively removes the least important features.

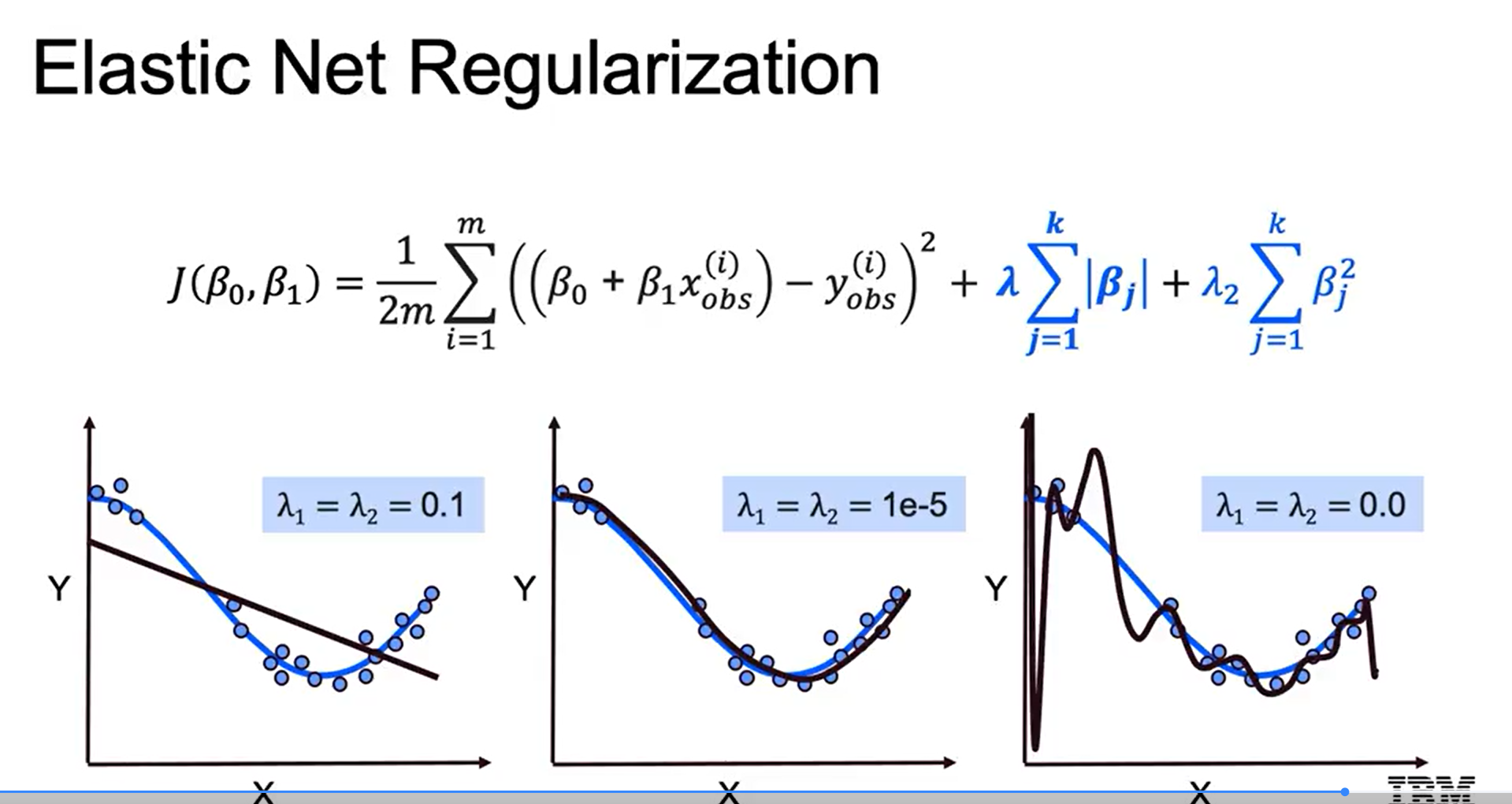
### RFECV Class

### RFECV incorporates cross-validation to evaluate the impact of feature elimination on model performance.

### It helps ensure that the model generalizes well to new data while managing complexity.

### **V.Elastic Net**

### 



### The Elastic Net, a hybrid approach that combines Ridge and LASSO regression techniques for model selection in supervised machine learning.

### Understanding Model Selection

### The choice between models can be based on prediction accuracy or interpretability.

### LASSO helps eliminate less important features, while Ridge regression is more computationally efficient.

### Elastic Net Overview

### Elastic Net introduces a combination of Ridge and LASSO, using two parameters: Lambda (for penalizing coefficients) and Alpha (for weighting L1 and L2 penalties).

### This method aims to leverage the strengths of both Ridge and LASSO to optimize model performance.

### Effects of Elastic Net

### The correct choice of Lambda values can lead to underfitting, perfect fitting, or overfitting.

### The balance between L1 and L2 penalties influences the model's ability to retain or eliminate coefficients, impacting interpretability and performance.

### 