Regularization Terms: L1, L2, and Elastic Net

Preventing Overfitting through Model Complexity Constraints

🎯 **Regularization prevents overfitting by constraining model complexity**

L2 Regularization

Ridge

$$\lambda \sum w^2$$

- ✓ Penalizes sum of squared weights
- ✓ Encourages small weights
- ✓ Smooth solutions
- ○ Does not enforce sparsity

L1 Regularization

Lasso

$$\lambda \sum |w|$$

- ✓ Penalizes sum of absolute weights
- ✓ Promotes sparsity
- ✓ Enables automatic feature selection
- ○ Sets some weights to zero

Elastic Net

L1 + L2 Combined

$$\lambda_1 \sum |w| + \lambda_2 \sum w^2$$

- ✓ Combines advantages of L1 and L2
- ✓ Balances sparsity and grouping effect
- ✓ Handles correlated features
- ★ **Best of Both Worlds**

λ

Lambda (λ): Regularization Strength

Hyperparameter controlling regularization strength - larger values lead to stronger regularization effect