🎓 ElasticNet Regression | Intuition

ElasticNet Regression is a **powerful hybrid** of **Ridge** and **Lasso Regression**. If you're just stepping into regularization techniques, this model offers **flexibility** and **balance** by combining both **L1 (Lasso)** and **L2 (Ridge)** penalties.

🧠 What Is ElasticNet?

ElasticNet adds both **L1** and **L2 regularization terms** to the linear regression cost function:

$$\text{Loss} = \text{RSS} + \lambda_1 \sum |w_j| + \lambda_2 \sum w_j^2$$

- `RSS` : Residual Sum of Squares (regular loss)
- `λ1 * sum(|w|)` : Lasso penalty (L1) – encourages sparsity (drives some weights to 0)
- `λ2 * sum(w²)` : Ridge penalty (L2) – reduces magnitude of weights (shrinks them)

ElasticNet helps when:

- **Features are correlated** (Lasso fails here)
- **You want both sparsity and shrinkage**

🔢 ElasticNet Parameters

- **alpha**: Total strength of regularization (like λ)
- **l1_ratio**: Controls mix of Lasso and Ridge:
  1. l1_ratio = 0 → Ridge only
  2. l1_ratio = 1 → Lasso only
  3. l1_ratio = 0.5 → 50% Ridge + 50% Lasso

🧪 Key Takeaways

✅ Why Use ElasticNet?

- **Better when features are correlated**: Lasso may randomly select one of them; ElasticNet distributes weights better.
- **Sparsity + Shrinkage**: It gives us both the feature selection of Lasso and the stability of Ridge.
- **Flexibility**: You can **tune the l1_ratio** depending on the problem.

Visual Intuition

Think of the ElasticNet penalty as a **mix between a diamond (Lasso)** and **a circle (Ridge)** — it creates a **smooth but still sparse model**, unlike Lasso which is aggressive in zeroing coefficients.
