Overfitting — Model learn training data too well, including noise → high training accuracy, low test accuracy.

Fix - Simplify model, regularization, more data.

High variance, low bias — Train dataset

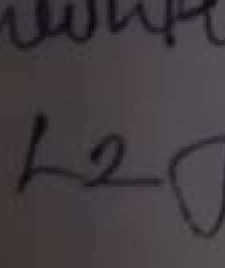
↳ Test dataset.

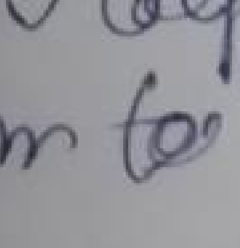
Underfitting — Model too simple, misses patterns → low training & test accuracy.

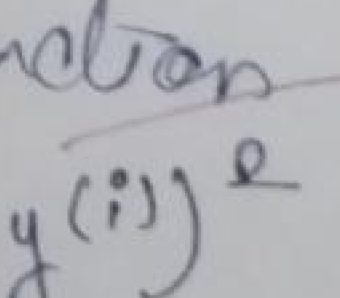
Fix — Use more complex model, train longer, add features

High bias, low variance.

Over -	Under	Generalized Model
Train Accuracy - 90%	Train Accuracy 60%	Train — 90%
Test Accuracy - 70%	Test Accuracy 62%.	Test - 89%.
Low - bias, High Variance	Low - Variance High bias	Low bias, Low Variance

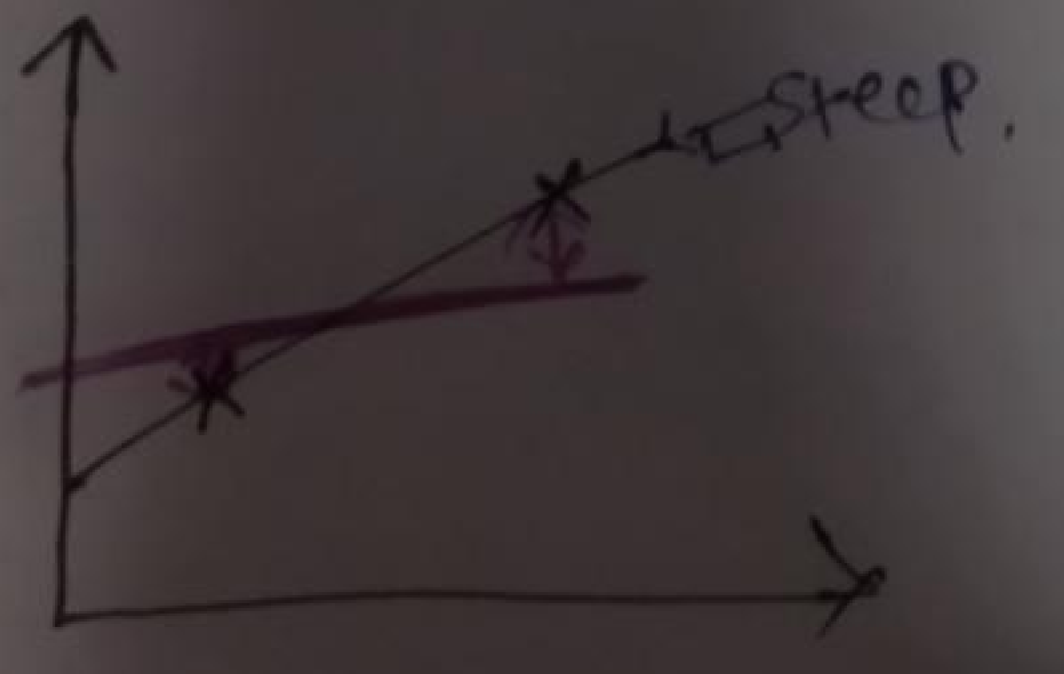






Ridge Regression — (L2 Regularization)

- Purpose — prevent overfitting in linear regression by shrinking large coefficient.
- How — Add L2 Penalty term to cost function — MSE



$$\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^{(i)}-y^{(i)}\right)^2$$

Penalizing

$$\left(h_\theta(x)^{(i)}-y^{(i)}\right)^2+\lambda(\text{slope})^2$$

$$= 0 + 1(2)^2$$

4

$$\left\{\text{Small value}\right\} + 1(1.3)^2$$

2.05

Lasso Regression - L1 Regularization

Purpose - prevent overfitting and perform features selection in linear Regression.

How - Add L1 penalty term of cost function

$$(h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda\,|slope|$$

$$\hookrightarrow \lambda\,|m_1 + m_2 + m_3 + m_4|$$

Elastic Net - is a regularization technique used in regression model.

It combine the strength of Ridge Regression (L_2 Penalty) and Lasso Regression (L_1 Penalty)

Formula -

$$\text{cost fun} - \frac{1}{2n}\sum_{i=1}^{\omega}(y-\hat{y}_i)^2 + \underbrace{\boxed{\lambda\sum_{i=1}^{n}(slope)^2}}_{\text{Prevent overfitting}} + \underbrace{\boxed{\lambda_2\sum_{i=1}^{n}|slope|}}_{\text{feature Selection}}$$