ML
LECTURE-8

BY
Dr. Ramesh Kumar Thakur
Assistant Professor (II)
School Of Computer Engineering

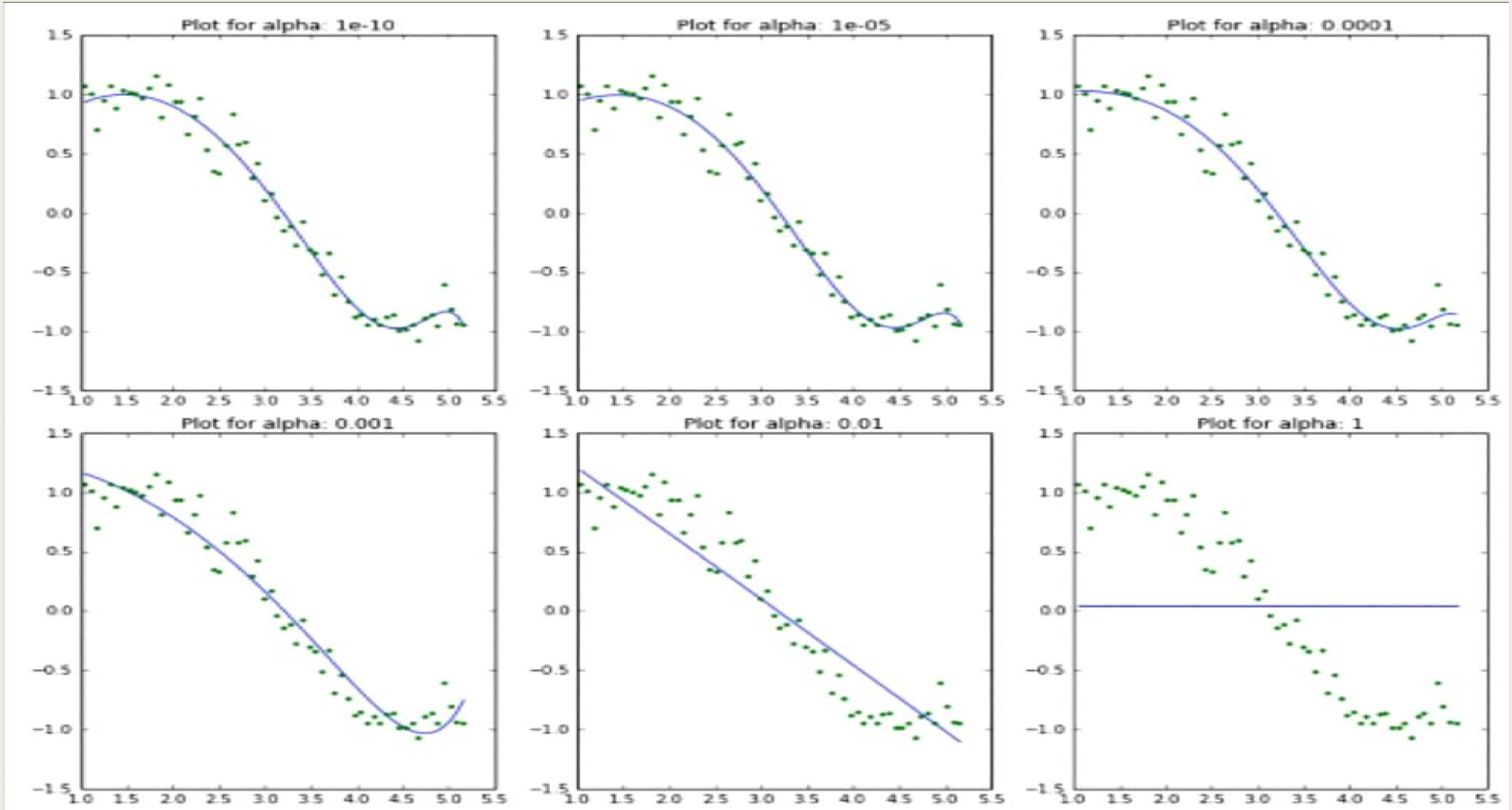
Regularization

- Regularization is a technique used in machine learning and statistical modelling to **prevent overfitting and improve the generalization ability of models.**
- When a model is **overfitting**, it has **learned the training data too well and may not perform well on new, unseen data.**
- Regularization introduces additional constraints or penalties to the model during the training process, aiming to control the complexity of the model and avoid over-reliance on specific features or patterns in the training data.
- By doing so, **regularization helps to strike a balance between fitting the training data well and generalizing it well to new data.**
- The most common regularization techniques used are **L1 regularization (Lasso), L2 regularization (Ridge), and Elastic Net regularization.**
- **L1 regularization adds the sum of the absolute values of the model’s coefficients to the loss function,** encouraging sparsity and feature selection.
- **L2 regularization adds the sum of the squared values of the model’s coefficients**, which enables smaller but non-zero coefficients.
- Finally, **Elastic Net regularization combines both L1 and L2 regularization.**

L1 - Regularization (Lasso)

- L1 regularization, also known as **Lasso** (**Least Absolute Shrinkage and Selection Operator**) regularization, **adds the sum of the absolute values of the model's coefficients to the loss function.**
- It encourages sparsity in the model by **shrinking some coefficients to precisely zero.**
- This has the effect of performing **feature selection, as the model can effectively ignore irrelevant or less important features.**
- L1 regularization is particularly useful when dealing with **high-dimensional datasets with desired feature selection.**
- Mathematically, the L1 regularization term can be written as:
- L1 regularization $= \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda * \Sigma|\beta_i|$
- Here, λ **is the regularization parameter that controls the strength of regularization,** β_i **represents the individual model coefficients** and the sum is taken over all coefficients.

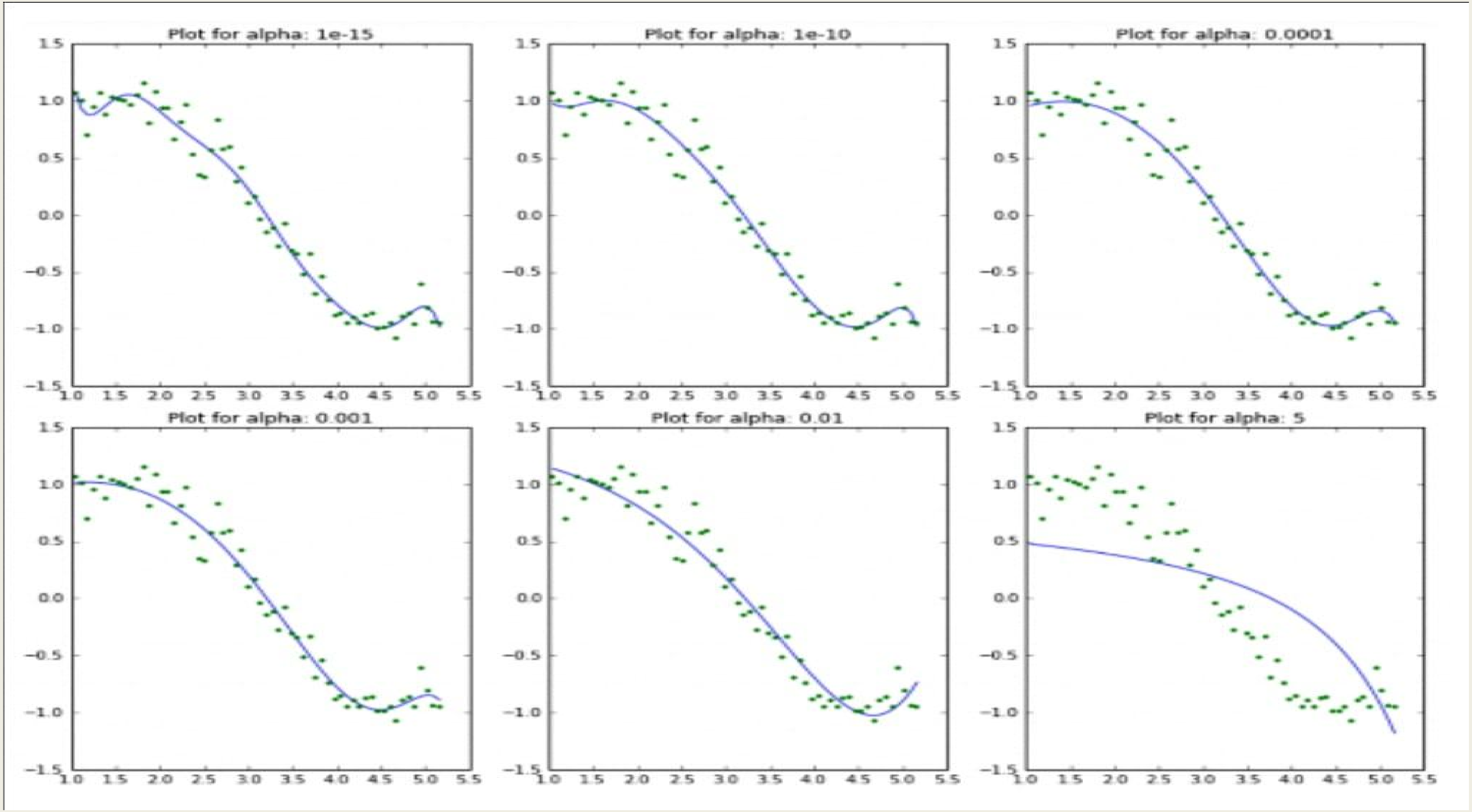
L1 - Regularization (Lasso)

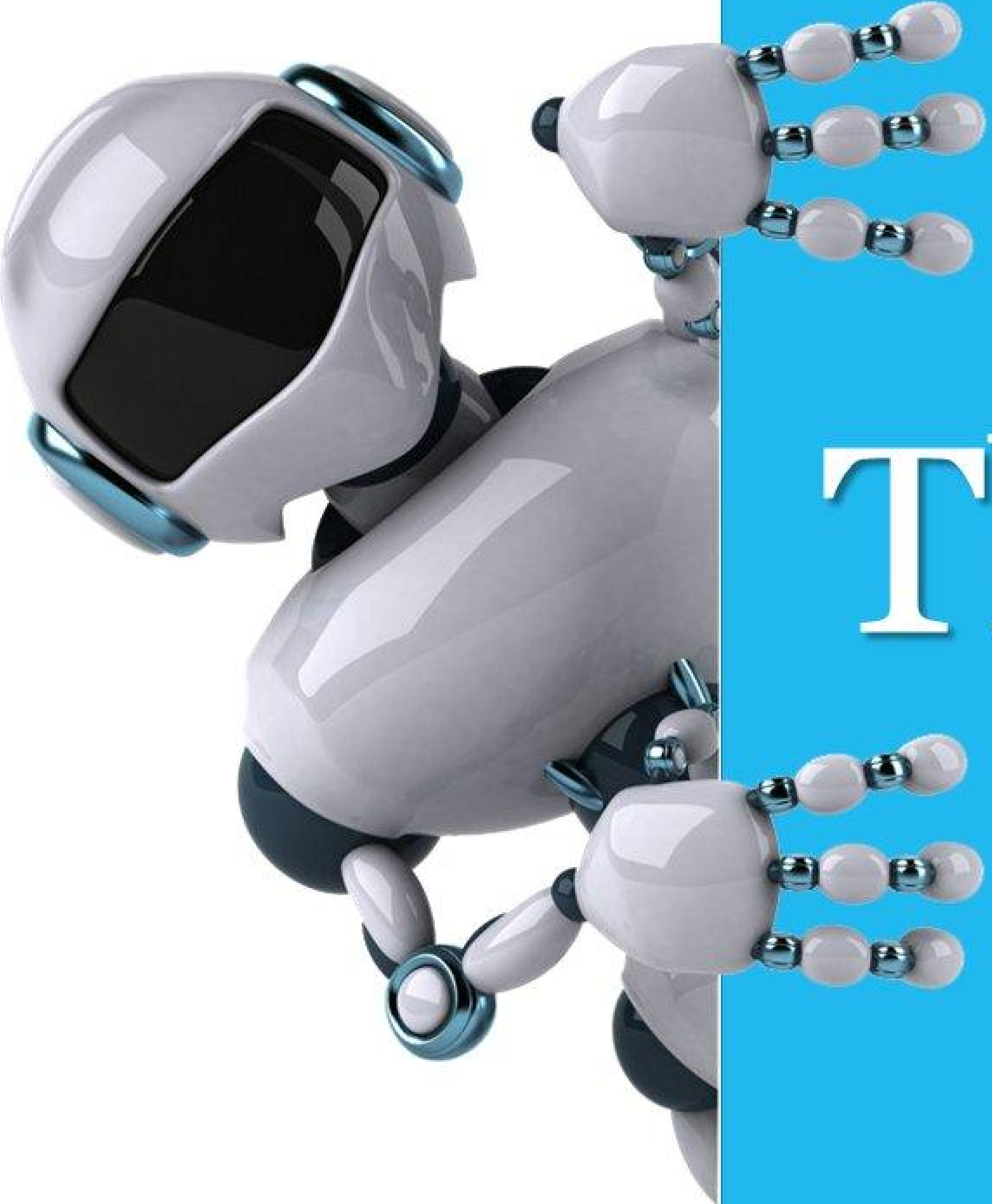


L2 - Regularization (Ridge)

- L2 regularization, also known as **Ridge regularization, adds the sum of the squared values of the model's coefficients to the loss function.**
- Unlike L1 regularization, L2 regularization does not force the coefficients to be exactly zero but instead **encourages them to be small.**
- L2 regularization can **prevent overfitting** by spreading the influence of a single feature across multiple features.
- It is **advantageous when there are correlations between the input features.**
- Mathematically, the L2 regularization term can be written as:
- L2 regularization $= \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda * \Sigma (\beta_i^2)$
- Similar to L1 regularization, **λ is the regularization parameter, and β_i represents the model coefficients.**
- The sum is taken over all coefficients, and the squares of the coefficients are summed.

L2 - Regularization (Ridge)





Thank you