Lasso Regression algorithm

Name: Saloni Thete

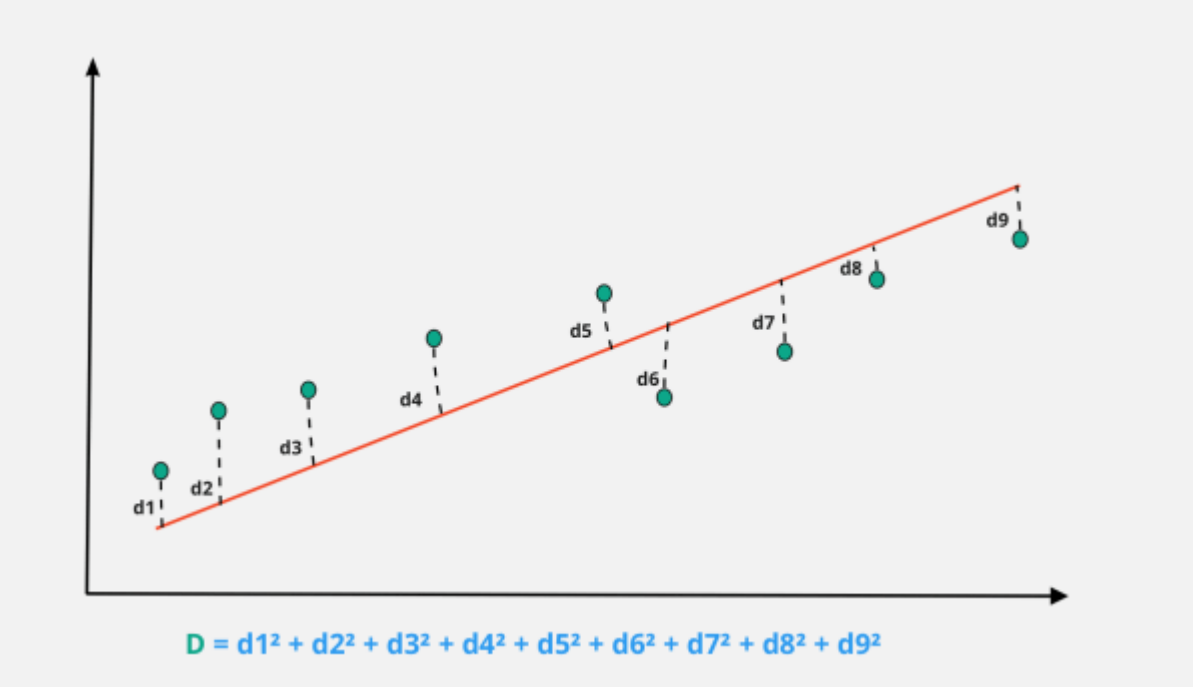
Batch: 19

Introduction

Lasso regression is described as a regularization technique. it is used over regression techniques for a greater correct prediction. This version makes use of shrinkage. Shrinkage is wherein data values are gotten smaller in the direction of a relevant factor due to the fact the mean. The lasso method encourages simple, sparse models. This precise style of regression is well-appropriate for models displaying excessive levels of multicollinearity or once you need to automate positive components of model selection, like variable selection/parameter elimination.

Lasso regression is like linear regression, but it uses a technique **"shrinkage"** where the coefficients of determination are shrunk towards **zero**.

The Statistics Of Lasso Regression



When To Use Lasso Regression?

Choosing a model depends on the dataset and the problem statement you are dealing with. It is essential to understand the dataset and how features interact with each other.

Lasso regression penalizes less important features of your dataset and makes their respective coefficients zero, thereby eliminating them. Thus it provides you with the benefit of feature selection and simple model creation.

So, if the dataset has high dimensionality and high correlation, lasso regression can be used.

Workflow

1. **Import the Libraries**

2. **Load the Dataset**
3. **Data Cleaning and Preprocessing**
4. **Split the Data into train and test sets**
5. **Apply Lasso Regression algorithm on the training set.**
6. **Evaluate the performance of a Lasso regression for different regularization parameters using 5-fold cross validation on the training set**
7. **Finally, store your test data results in a DataFrame and identify the lambda**
8. **Furthermore, obtain the mean squared error for the test data of this model.**

Conclusion

- Regression is a popular statistical technique used in machine learning to predict an output.
- **Overfitting** happens while doing regression due to the irrelevant noises in the training dataset.
- Lasso regression is a regularization algorithm which can be used to eliminate irrelevant noises and do feature selection and hence regularize a model.
- Evaluation of the lasso model can be done using metrics like RMSE and R-Square.
- Alpha is a hyper-parameter in the lasso model which can be tuned using lassoCV to control the regularization.