LASSO Regression

- **LASSO** stands for **Least Absolute Shrinkage and Selection Operator**.

- The algorithm is another variation of linear regression like ridge regression.

- We use lasso regression when we have large number of predictor variables.

- Lasso regression is a parsimonious model which performs L1 regularization.

- The L1 regularization adds a penality equivalent to the absolute of the magnitude of regression coefficients and tries to minimize them.

- The equation of lasso is similar to ridge regression and looks like as given below.
  ```
  LS Obj + λ (sum of the absolute values of coefficients)
  ```

  If $\lambda = 0$, We get same coefficients as linear regression

  If λ = vary large, All coefficients are shrinked towards zero

- The two models, lasso and ridge regression are almost similar to each other.
- In lasso the coefficients which are responsible for large variance are converted to *zero*.
- On the other hand, coefficients are only shrinked but are never made zero.
- **Lasso regression analysis is also used for variable selection as the model imposes coefficients of some variables to shrink towards zero.**
-

What does large number of variables mean?

- Large number here means that model tend to over-fit.

- Theoretically, minimum ten variables can cause overfitting problem.

- When you face computational challenges due to presence of n number of variables.

The training of lasso regression model is exactly same as that of ridge regression.
We need to identify the optimal lambda value and then use that value to train the model.
To achieve this, we can use the same `glmnet` function by passing `alpha = 1` argument.
When we pass `alpha = 0, glmnet()` runs a ridge regression and
when we pass `alpha = 0.5` the glmnet runs another kind of model which is called as **elastic net**
and is a combination of ridge and lasso regression.