MACHINE LEARNING

1.A

2.A

3.C

4.B

5.C

6.B

7.D

8.D

9.D

10.B

11.B

12.A,B

13. The word regularize means to make things regular or acceptable. This is exactly why we use it for. Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting. Now to get a clear picture of what the above definition means, let's get into the details.

Let's try fitting a polynomial on the given data.

$$Y(x, w) = w0 + w1x + w2x^2 + \cdots . wnx^n.$$

- **14**Ridge Regression.
- LASSO (Least Absolute Shrinkage and Selection Operator) Regression.
- Elastic-Net Regression.

.

Ridge Regression

Ridge regression is a method for analyzing data that suffer from multi-collinearity.

$$Loss = \sum_{i=1}^{n} (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^{n} w_i^2$$

Loss Function for Ridge Regression

Ridge regression adds a penalty ***(L2 penalty)*** to the loss function that is equivalent to the square of the magnitude of the coefficients.

The regularization parameter *(λ)* regularizes the coefficients such that if the coefficients take large values, the loss function is penalized.

- ***λ → 0***, the penalty term has no effect, and the estimates produced by ridge regression will be equal to least-squares i.e. the loss function resembles the loss function of the Linear Regression algorithm. Hence, a lower value of λ will resemble a model close to the Linear regression model.
- ***λ → ∞,*** the impact of the shrinkage penalty grows, and the ridge regression coefficient estimates will **approach zero** (coefficients are close to zero, but not zero).

LASSO Regression

LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model.

$$Loss = \sum_{i=1}^{n} (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^{n} |w_i|$$

Loss Function for LASSO Regression

LASSO regression adds a penalty ***(L1 penalty)*** to the loss function that is equivalent to the magnitude of the coefficients.

In LASSO regression, the penalty has the effect of forcing some of the coefficient estimates to be **exactly equal to zero** when the regularization parameter λ is sufficiently large.

To sum up, **LASSO regression converts coefficients of less important features to zero, which indeed helps in feature selection, and it shrinks the coefficients of remaining features to reduce the model complexity, hence avoiding overfitting.**

Elastic-Net Regression

Elastic-Net is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

$$Loss = \sum_{i=0}^{n} (y_i - (w_i x_i + c))^2 + \lambda_1 \sum_{i=0}^{n} |w_i| + \lambda_2 \sum_{i=0}^{n} w_i^2$$

15. Within a linear regression model tracking a stock's price over time, the error term is **the difference between the expected price at a particular time and the price that was actually observed**.

Regression is a maximum likelihood estimation where we find parameters of the relation between independent and dependent variables (which is in the form of an equation often times) which maximize the likelihood of getting such samples from the population.

Since regression is an estimation, we cannot be completely correct at it. So, the error term is a catch-all for what we miss out in this estimation because in reality

-The true relation may not be what we assume(linear relation in case of linear regression)

-There may be other variables not included in the model that cause variation in response variable

-There may be measurement errors in the observations