UNIT - V

Regression & Applications

Regression :

* Main task of analytics is Prediction
+ induction model is used to predict and assign labels to a new, unlabeled, object
* Prediction is used to
  - reduce cost
  - Increase profits
  - Improve Product & service quality
  - Improve customer satisfaction
  - Reduce environmental damage.

→ Test data
  ↳ used to test the performance of the induced model.

→ deduction
  → Predict correct label.

Regression task:

A predictive task whose aim is to assign a quantitative value to a new, unlabeled object, given the values of its predictive attributes.

Regression methods used in many different domains:

1. stock market
2. Transport
3. Higher Education
4. Survival analysis
5. Macro economics

Types:

Linear Regression

Ridge "

Lasso "

Principal Components Regression

Partial Least squares "

Linear Regression:

- oldest & simplest regression alg.
- Induce good regression models, which are easily interpretable

$$height = 128.017 + 0.611 \times weight.$$

- instance 'x' associated with only one attribute

- 'y' is associated with the height

- 2 parameters $\hat{\beta}_0$ & $\hat{\beta}_1$ ↳ associated with x (weight)

- 2 types
  - Simple linear reg
  - Multiple " "

$$\underset{\hat{\beta}_0, \hat{\beta}_1}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \underset{\hat{\beta}_0, \hat{\beta}_1}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 \times x_i))^2$$

- Multivariate linear regression - the linear model generalized for any no p of predictive attributes

No of predictive attributes

$$\hat{y} = \beta_0 + \sum_{j=1}^{p} \beta_j \times x_j$$

slope of linear model

value of $\hat{y}$, when $x_j = 0$ ~~$\hat{\beta}_1$~~

x_j - j^{th} attribute of some object x represented as tuple $x = (x_1 \ldots x_j, \ldots x_p)$.

$\beta_0, \beta_1, \ldots \beta_p$ - estimated using an appropriate optimization method to minimize objective function

$$\underset{\hat{\beta}_1 \ldots \hat{\beta}_p}{\text{argmin}} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 =$$

$$\underset{\hat{\beta}_0 \ldots \hat{\beta}_1}{\text{argmin}} \sum_{i=1}^{n} \left[y_i - \left(\hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j \times x_{j_i}\right)\right]^2$$

Lasso Regression:

- Least absolute shrinkage & selection operator ~~r~~regression alg.
- deal efficiently with high dimensional data sets.
- Does not predict just performs attribute selection.
- Complexity measured by predictive attributes.
- Usually produces spare solutions.
- sparse means - Large no of predictive attributes have zero weight.
- Performs shrinkage.

$$\underset{\hat{\beta}_0 \ldots \hat{\beta}_j}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} \text{error}\,(y_i, \hat{y}_i) + \lambda * \sum_{j=1}^{p} |\hat{\beta}_j| \right\}$$

Principal Components Regression

- PCR Creates linear Combinations of predictive attributes.
- 1st principal Component: the first linear Combination, is the most variance of all possible linear Combinations.
- Just evaluate attribute without considering target attributes.
- Used to predictive attributes is the formulation of the multivariate linear regression problem.

Partial Least Squares Regression:

- PLS starts by evaluating the Correlation of each predictive attribute with the target attribute.

- 1st Component is linear Combination of the predictive attributes.

- gives similar result as PCR.