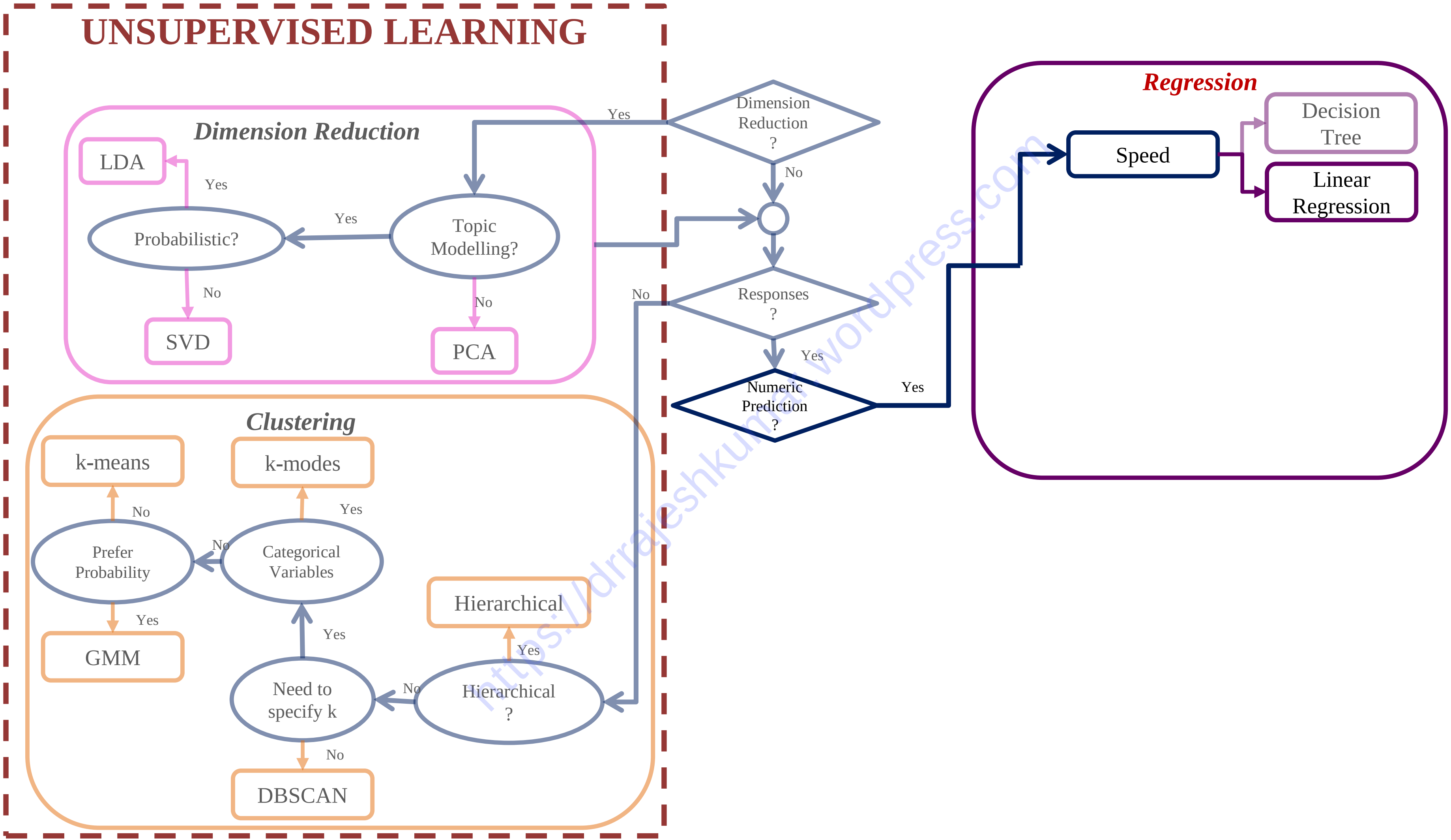
Machine Learning

Dr. Rajesh Kumar

PhD, PDF (NUS, Singapore)
SMIEEE, FIET (UK), FIETE, FIE (I),SMIACSIT, LMISTE, MIAENG
Professor, Department of Electrical Engineering
Professor, Centre of Energy and Environment
Malaviya National Institute of Technology, Jaipur, India, 302017
Tel: (91) 9549654481
http://drrajeshkumar.wordpress.com
rkumar.ee@mnit.ac.in, rkumar.ee@gmail.com



UNSUPERVISED LEARNING
Dimension Reduction
LDA
Yes
Probabilistic?
Yes
Topic Modelling?
No
SVD
No
PCA
Clustering
k-means
k-modes
No
Yes
Prefer Probability
No
Categorical Variables
Yes
GMM
Yes
Hierarchical
Yes
Need to specify k
No
Hierarchical ?
No
DBSCAN
Yes
Dimension Reduction ?
No
No
Responses ?
Yes
Numeric Prediction ?
Yes
Regression
Speed
Decision Tree
Linear Regression

Linear Regression

- Model linear relation of dependent variables with one or more independent variable
  - Simple Linear Regression (one independent variable)
  - Multiple Linear Regression (multi independent variable)
- Most applications fall into one of the following two broad categories:
  - Prediction, forecasting, error reduction
    Fit predictive model and than use for Prediction
  - Variation in the response variable
    Quality of strength between the responses

Linear Regression

- A linear regression model assumes that the linear relationship between the dependent variable *y* and the *p*-vector of regressors **x** (independent variable)

$$y_i = \beta_0 1 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i = \mathbf{x}_\mathbf{i}^\mathbf{T} \boldsymbol{\beta} + \varepsilon_i \qquad i = 1, \ldots, n$$

  - y_i is dependent variable for observation *i*
  - x_i is independent variable for observation *i*
  - ε_i is error term for the observation *i*
  - β_0 is intercept coefficient
  - β_1 is regression coefficient for independent variable

Linear Regression

- *n* equations are stacked together and written in matrix notation as

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}}_{\mathbf{X}} \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}_{\boldsymbol{\varepsilon}}$$

Linear Regression

- Estimated Model $\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ for given $\boldsymbol{\beta}$
- Objective of the regression model is to find the $\boldsymbol{\beta}$ to have the least possible error of estimated model $\hat{\mathbf{y}}$ from the data $\mathbf{y}$
- Error **e** used for fitting

$$\mathbf{e} = \hat{\mathbf{y}} - \mathbf{y}$$

Some Approaches	Objective Function
Least Square	$\mathbf{e}^T\mathbf{e}$
Ridge Regression	$\left(\sum_{i=1}^{n} e_i^2\right)^{1/2} = \left(\mathbf{e}^T\mathbf{e}\right)^{1/2}$
Lasso (least absolute shrinkage and selection operator)	$\left(\sum_{i=1}^{n} \lvert e_i \rvert\right)$

Linear Regression(least square)

- **Simple Linear Regression** model fitting with least square

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

- Objective is

$$\min_{\beta_0, \beta_1} \sum_{i=1}^{n} [y_i - \beta_0 - \beta_1 x_i]^2$$

- The solution can be obtain

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2} ; \hat{\beta}_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

Linear Regression(least square)

- **Multiple Linear Regression** model fitting with least square

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

- Objective is

$$= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left[y_i - \hat{y}_i \right]^2$$

$$= \min_{\boldsymbol{\beta}} (\mathbf{y} - \hat{\mathbf{y}})^{\mathrm{T}} (\mathbf{y} - \hat{\mathbf{y}}) = \min_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

- The solution can be obtain

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\mathrm{T}} \mathbf{X} \right)^{-1} \mathbf{X}^{\mathrm{T}} \mathbf{y}$$

Note: $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ should be non singular / invertible matrix

Linear Regression(least square)

- Correlation coefficient *r*

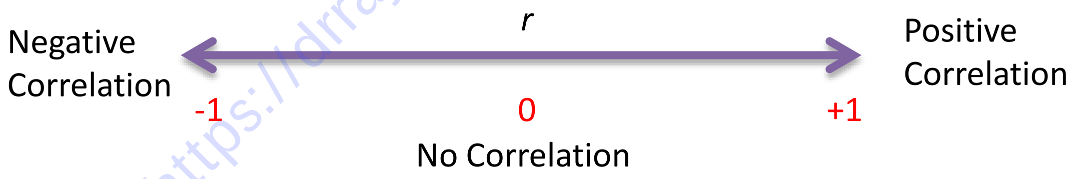
$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

$$s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}}$$

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

Linear Regression(least square)

- Correlation coefficient *r*
  - Positive Correlation
    Direct relationship with independent variable
  - Negative Correlation
    Inverse relationship with independent variable
  - Correlation relationship strength is proportional to magnitude of *r*

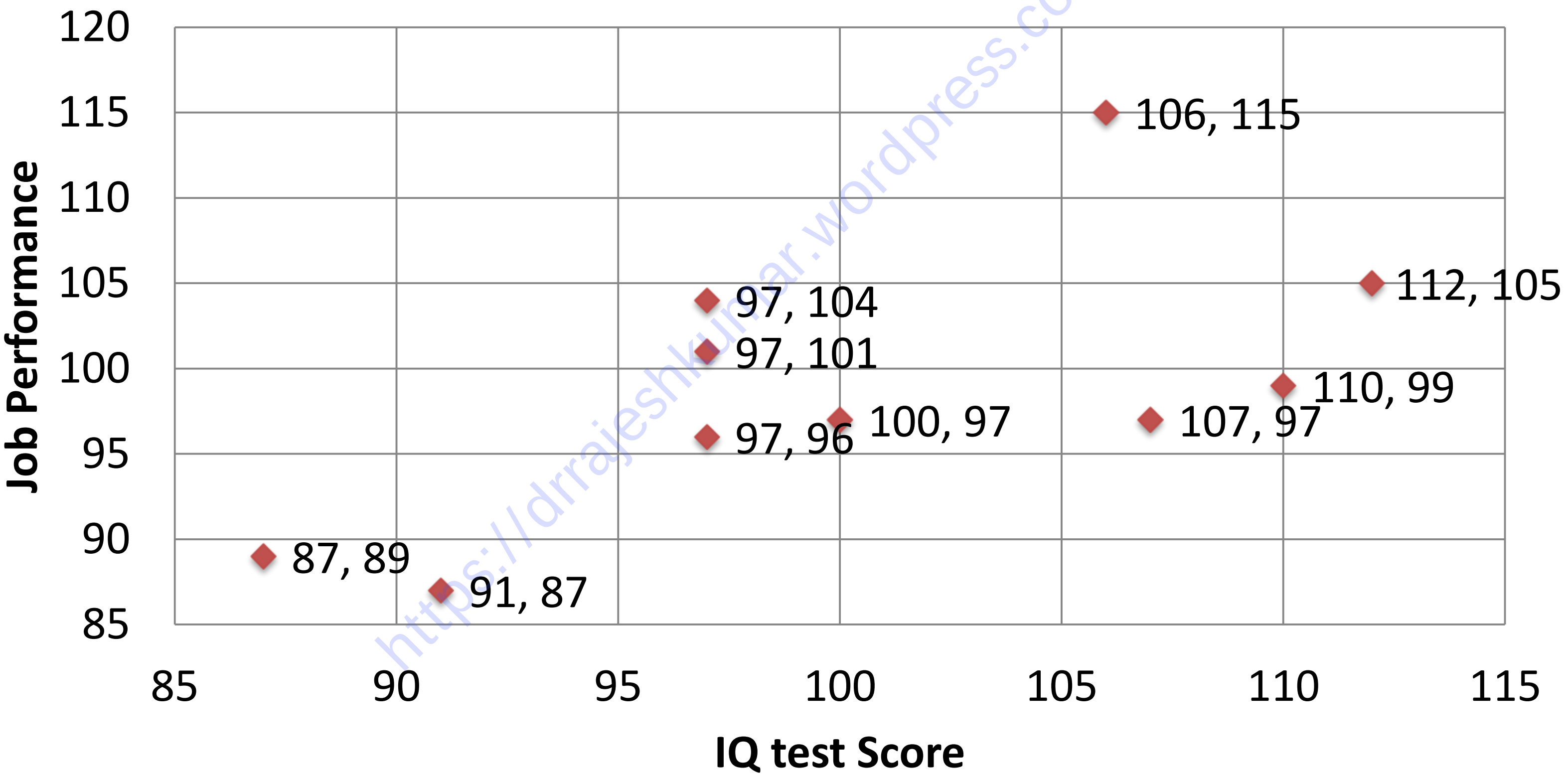


Example of Linear Regression

- Objective is to predict Job Performance from IQ
- Data comprising of 10 person information

Id	IQ	Performance
1	106	115
2	97	104
3	107	97
4	97	101
5	110	99
6	112	105
7	97	96
8	87	89
9	100	97
10	91	87

Example of Linear Regression



Example of Linear Regression

- Coefficients for linear regression

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{990 \times 101406 - 1004 \times 99776}{10 \times 101406 - (1004)^2} = 35.8762$$

$$\hat{\beta}_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{10 \times 99776 - 1004 \times 990}{10 \times 101406 - (1004)^2} = 0.6287$$

- Simple linear regression model

$$\hat{y} = 35.8762 + 0.6287x$$

Example of Linear Regression