Machine Learning
(6CS4-02)

Unit-3 Notes

Vision of the Institute

To become a renowned centre of outcome based learning and work towards academic, professional, cultural and social enrichment of the lives of individuals and communities.

Mission of the Institute

M1- Focus on evaluation of learning outcomes and motivate students to inculcate research aptitude by project based learning.
M2- Identify, based on informed perception of Indian, regional and global needs, the areas of focus and provide platform to gain knowledge and solutions.
M3- Offer opportunities for interaction between academia and industry.
M4- Develop human potential to its fullest extent so that intellectually capable and imaginatively gifted leaders can emerge in a range of professions.

Vision of the Department

To become renowned Centre of excellence in computer science and engineering and make competent engineers & professionals with high ethical values prepared for lifelong learning.

Mission of the Department

M1-To impart outcome based education for emerging technologies in the field of computer science and engineering.
M2-To provide opportunities for interaction between academia and industry.
M3- To provide platform for lifelong learning by accepting the change in technologies
M4- To develop aptitude of fulfilling social responsibilities.

Program Outcomes (PO)

1. **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem analysis**: Identify, formulate, research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/development of solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern tool usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The engineer and society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and team work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-long learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Program Educational Objectives (PEO)

1. To provide students with the fundamentals of Engineering Sciences with more emphasis in Computer Science &Engineering by way of analyzing and exploiting engineering challenges.
2. To train students with good scientific and engineering knowledge so as to comprehend, analyze, design, and create novel products and solutions for the real life problems.
3. To inculcate professional and ethical attitude, effective communication skills, teamwork skills, multidisciplinary approach, entrepreneurial thinking and an ability to relate engineering issues with social issues.
4. To provide students with an academic environment aware of excellence, leadership, written ethical codes and guidelines, and the self-motivated life-long learning needed for a successful professional career.
5. To prepare students to excel in Industry and Higher education by Educating Students along with High moral values and Knowledge

Program Specific Outcomes (PSO)

PSO1: Ability to interpret and analyze network specific and cyber security issues, automation in real word environment.

PSO2: Ability to Design and Develop Mobile and Web-based applications under realistic constraints.

Course Outcome:

CO1: Understand the concept of machine learning and apply supervised learning techniques.
CO2: Illustrate various unsupervised leaning algorithm for clustering, and market basket analysis.
CO3: Analyze statistical learning theory for dimension reduction and model evaluation in machine learning.
CO4: Apply the concept of semi supervised learning, reinforcement learning and recommendation system.

CO-PO Mapping:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
Understand the concept of machine learning and apply supervised learning techniques.	3	3	3	3	2	1	1	1	1	2	1	3
Illustrate various unsupervised leaning algorithm for clustering, and market basket analysis.	3	3	3	2	2	1	1	1	1	1	1	3
Analyze statistical learning theory for dimension reduction and model evaluation in machine learning.	3	3	3	3	2	2	2	2	1	2	2	3
Apply the concept of semi supervised learning, reinforcement learning and recommendation system.	3	3	3	3	2	1	1	1	1	2	1	3

SYLLABUS:

RAJASTHAN TECHNICAL UNIVERSITY, KOTA

Syllabus

III Year-VI Semester: B.Tech. Computer Science and Engineering

6CS4-02:Machine Learning

Credit: 3 **Max. Marks: 150(IA:30, ETE:120)**

3L+0T+0P **End Term Exam: 3 Hours**

SN	Contents	Hours
1	**Introduction:** Objective, scope and outcome of the course.	01
2	**Supervised learning algorithm:** Introduction, types of learning, application, Supervised learning: Linear Regression Model, Naive Bayes classifier Decision Tree, K nearest neighbor, Logistic Regression, Support Vector Machine, Random forest algorithm	09
3	**Unsupervised learning algorithm:** Grouping unlabelled items using k-means clustering, Hierarchical Clustering, Probabilistic clustering, Association rule mining, Apriori Algorithm, f-p growth algorithm, Gaussian mixture model.	08
4	**Introduction to Statistical Learning Theory**, Feature extraction - Principal component analysis, Singular value decomposition. Feature selection – feature ranking and subset selection, filter, wrapper and embedded methods, Evaluating Machine Learning algorithms and Model Selection.	08
5	**Semi supervised learning, Reinforcement learning:** Markov decision process (MDP), Bellman equations, policy evaluation using Monte Carlo, Policy iteration and Value iteration, Q-Learning, State-Action-Reward-State-Action (SARSA), Model-based Reinforcement Learning.	08
6	**Recommended system,** Collaborative filtering, Content-based filtering Artificial neural network, Perceptron, Multilayer network, Backpropagation, Introduction to Deep learning.	08
	Total	42

LECTURE PLAN:

Unit No./ Total Lecture Reqd.	Topics	Lect. Reqd.	Lect. No.
Unit-I (10)	1. Introduction to subject and scope	1	1
	2. Introduction to learning, Types of learning and Applications	1	2
	3. Supervised Learning	1	3
	4. Linear Regression Model	1	4
	5. Naïve Bayes Classifier	1	5
	6. Decision Tree	1	6
	7. K-nearest Neighbor	1	7
	8. Logistic Regression	1	8
	9. Support Vector Machine	1	9
	10. Random Forest Algorithm	1	10
BC-1	**Gradient Descent**	1	11
Unit-II (8)	1. Introduction to clustering, K-mean clustering	2	12
	2. Hierarchical Clustering	1	14
	3. Probabilistic Clustering	1	15
	4. Association Rule Mining	1	16
	5. Apriori Algorithm	1	17
	6. f-p Growth Algorithm	1	18
	7. Gaussian Mixture Model	1	19
Unit-III (8)	1. Feature Extraction- PCA and SVD	3	22
	2. Feature Selection- Feature Ranking and Subset Selection	2	24
	3. Filter, Wrapper and Embedded Methods	1	25
	4. Evaluating Machine Learning Algorithms	1	26
	5. Evaluating Model Selection	1	27
Unit- IV (8)	1. Semi supervised learning: Markov Decision Process (MDP)	2	29
	2. Bellman Equations	1	30
	3. Policy Evaluation using Monte Carlo	1	31
	4. Policy iteration and Value iteration	1	32
	5. Q-Learning	1	33
	6. State-Action-Reward-State-Action (SARSA)	1	34
	7. Model-based Reinforcement Learning	1	35

Unit- V (8)	1. Recommendation system: Collborative Filtering	1	36
	2. Content based filtering	1	37
	3. Artificial neural network	1	38
	4. Perceptron	1	39
	5. Multilayer network	1	40
	6. Backpropagation	1	41
	7. Introduction to Deep learning.	2	42
BC-2	**Genetic Algorithms**	1	44

Text Book:
Machine learning- Tom M Mitchell

Feature Selection →

feature Selection is the Process of reducing the input variable to the Model by using only relevant data.

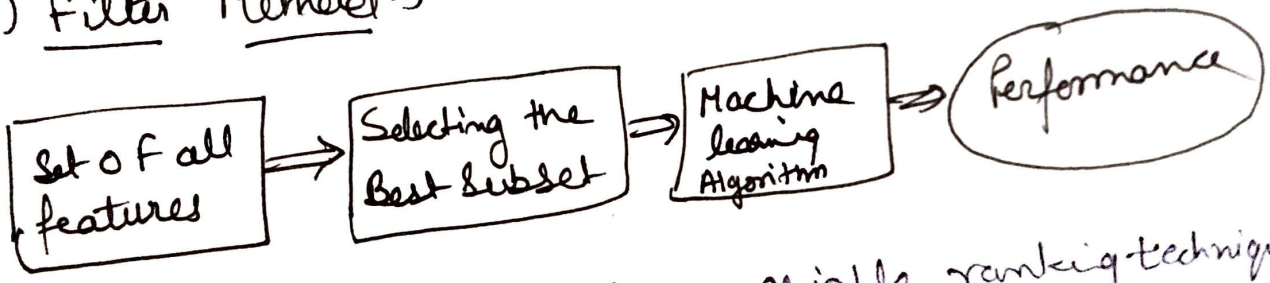
The goal of feature Selection in Machine learing is to find the best Set of features that allows one to build Models

using feature Selection we can Optimize our Models in Several ways.
1) Prevent learing from overfitting
2) Improved accuracy
3) Reduce training time

Feature Selection techniques

(i) filter Methods
(ii) Wrapper Methods
(iii) Embedded Method.

(1) filter Method →



This Method uses the variable ranking technique in order to Select the variables for ordering and the Selection of features is independent of the classifier used.

Ranking Means How Much useful and important each features is. Expected to be for classification.

it basically Select the Subsets of variables as a Pre Processing step Independently of the chosen Predictor

In this Method features are dropped based on their relation to the output or How they are correlating to the output.

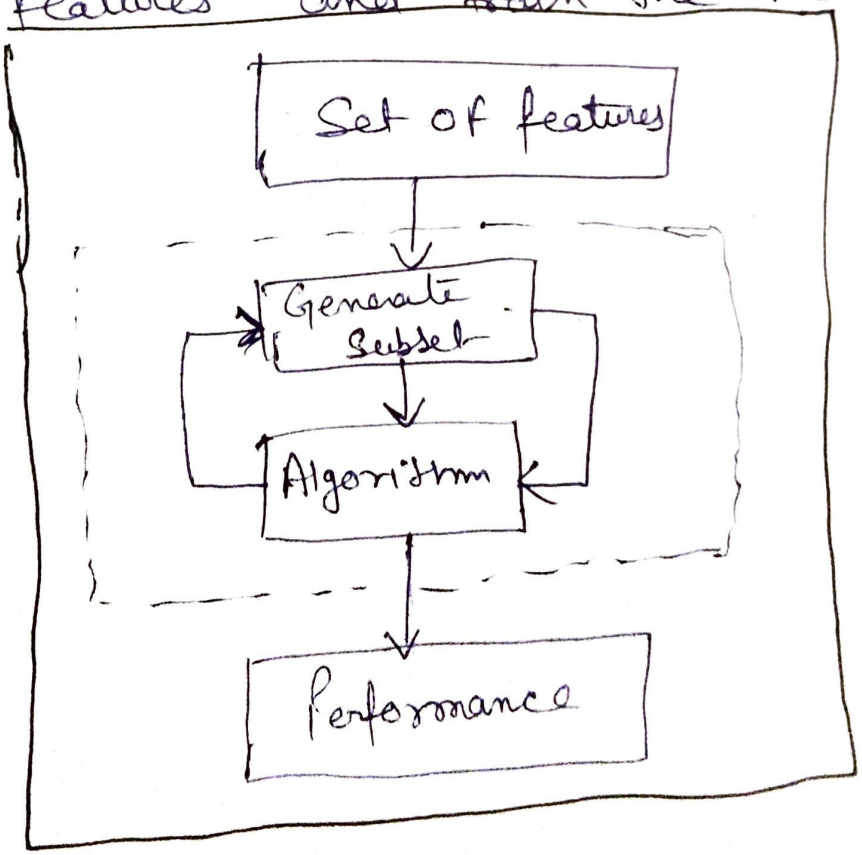
Example

Name	No of times read	Condition of Book	Color

In Book classifier we dropped the color Column based on a simple deduction.

2) Wrapper Method →

we Split data in to Subsets and train Model using this. Based on the output of the Model we add and Substract features and train the Model again.



for.

for Example

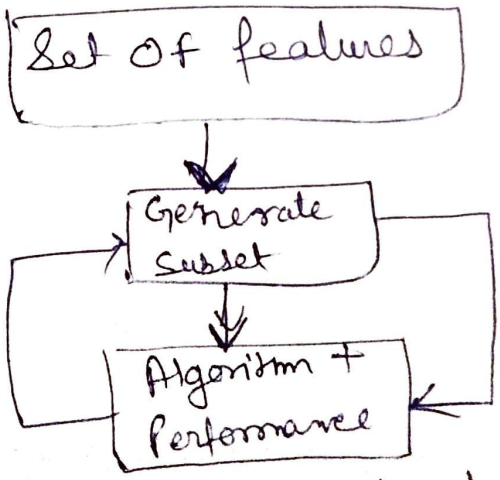
By using wrapper Method we would use a subset of different features to train the machine and adjust the subset according to output.

Name	No of time read	Condition of Book	Color

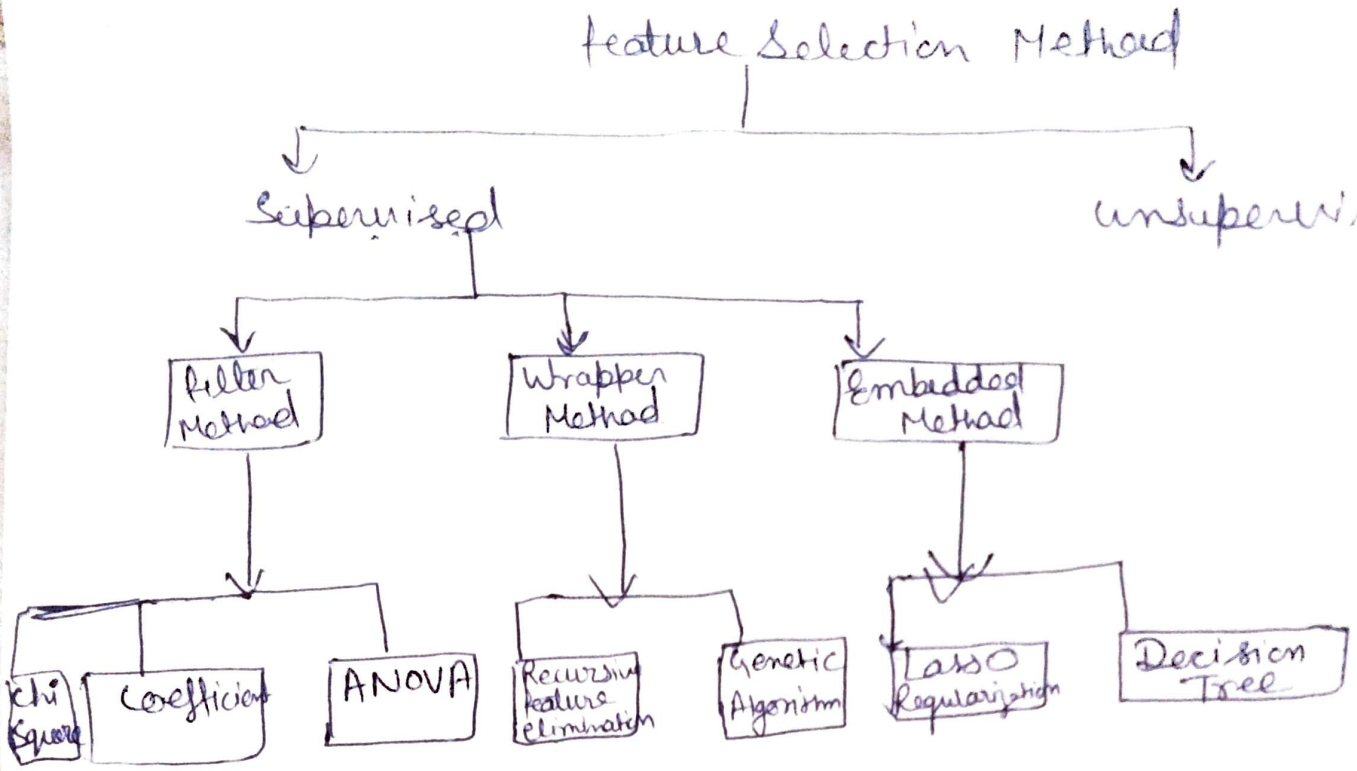
Name and No of times read.

Name, No of times read and Conditions after this check output.

(3) Embedded Method ⇒ This Method combines the qualities of both filter and Wrapper Method to Create the best subset.



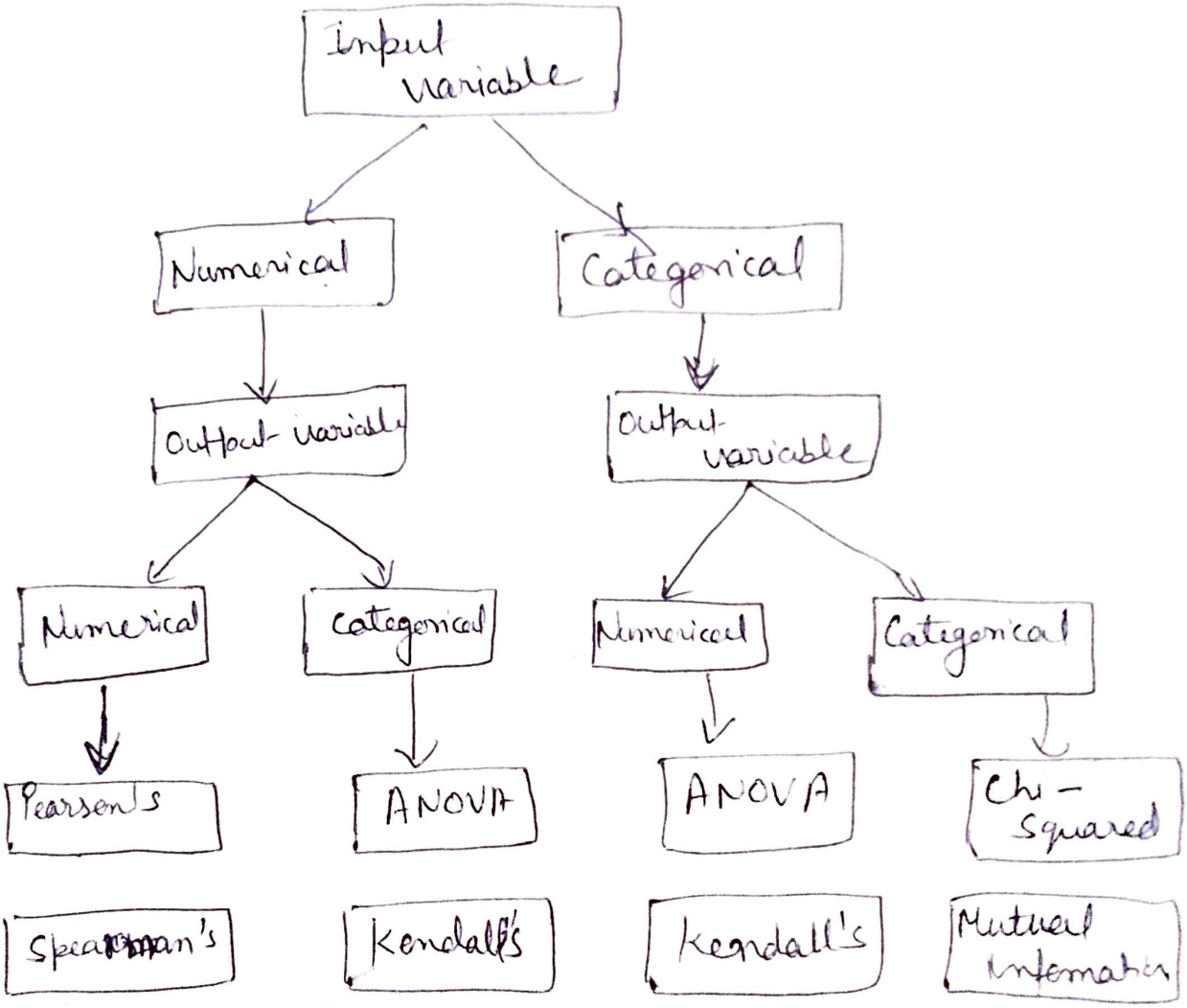
The Model will train and check the accuracy of different subsets and select the best among them.



* filter Method
  (i) Chi Square
  (ii) Coefficient - (~~Pearson~~ Correlation Coefficient)
  (iii) ANOVA
* Wrapper Method
  (i) Recursive feature elimination
  (ii) Genetic Algorithm
* Embedded Method
  (i) Lasso Regularization
  (ii) Decision Tree

Based on Input and output variables we can choose feature selection Model.

1. Numerical input & Numerical output
2. Numerical input & categorical output
3. categorical input & Numerical output
4. categorical input & Categorical output



Chi Square Test (χ^2)

Chi-Square Test is used to find the two variables are these related to each other or there is NO relationship.

For Q1

Table of observed value

Qualification / Marital Status	Middle class	High class	Bachelor's	Masters	Ph.D	Total
Never Married	18	36	21	9	6	90
Married	12	36	45	36	21	150
Divorced	6	9	9	3	3	30
Widowed	3	9	9	6	3	30
Total	39	90	84	54	33	300

Table of Expected value

Qualification / Marital Status	Middle class	High Class	Bachelor's	Masters	Ph.D	Total
Never Married	$\frac{90 \times 39}{300} \Rightarrow 11.7$	$\frac{90 \times 90}{300} \Rightarrow 27$	25.2	26.2	9.9	
Married	29.5	45	42	27	16.5	
Divorced	3.9	9	8.4	5.4	3.3	
Widowed	3.9	9	8.4	5.4	3.3	
Total						

Chi Square (χ^2) = $\frac{(\text{Observed value} - \text{Expected value})}{\text{Expected value}}$

Observed value	Expected value	(O−E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
18	11.7	6.3	39.69	3.39
36	27	.9	81	3
21	25.2	−4.2	17.64	0.7
9	16.2	−7.2	51.84	3.2
6	9.9	−3.9	15.21	1.53
12	19.5	−7.5	56.25	2.88
36	45	−9	81	1.8
45	42	3	9	4.5
36	27	9	81	3
21	16.5	4.5	20.25	1.22
6	3.9	2.1	4.41	1.13
9	9	0	0	0
9	8.4	0.6	0.36	0.04
3	5.4	−2.4	5.76	1.06
3	3.3	−0.3	0.09	0.027
3	3.9	−0.9	0.81	0.207
9	9	0	0	0
9	8.4	0.6	0.36	0.04
6	5.4	−2.4	5.76	1.06
3	3.3	−0.3	0.09	0.02

$\chi^2_{\text{Calculated}} = 23.57$

$\chi^2 = \sum \frac{(O-E)^2}{E}$

$\chi^2 = 23.57$

Degree of freedom = (Cols-1)(Rows-1)

$\Rightarrow (5-1)(4-1)$

$\Rightarrow 12$

from tabular value of Chi Square.

Significance level (α) = 0.05

$\chi^2_{tabular} = 21.03$

$\chi^2_{Calculated} = 23.57$

$\chi^2_{Calculated} > \chi^2_{tabular}$ (or $\chi^2_{Critical}$)

then we reject Null hypothesis and accept alternate hypothesis,

Alternate hypothesis there is significant relationship between Marital Status and Qualification.

The Principal Component are

$$z_1 = a_{11}x_1 + a_{12}x_2$$

$$\boxed{z_1 = 0.61x_1 + 0.79x_2}$$

$$z_2 = a_{21}x_2 + a_{22}x_2$$

$$\boxed{z_2 \Rightarrow 0.79x_2 + 0.61x_2}$$

Singular Value Decomposition ⇒

SVD is a widely used technique to decompose a matrix into several component matrices, exposing many of the useful and interesting properties of the original matrix.

The purpose of singular value decomposition is to reduce a dataset containing a large number to a data set containing significantly fewer values. but which still contains a large fractions of the variability present in the original data.

For Example

$A_{m \times n}$ → it will devide this matrix

2 unitary matries which are orthogonal and a rectangular diagonal matrix of singular values.

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V^T_{n \times n}$$

$$\boxed{\begin{aligned} U \cdot U^T &= I \\ V \cdot V^T &= I \end{aligned}} \rightarrow \text{Orthogonal Property}$$

$A^T \cdot A$

if $A = U \Sigma V^T$

then $A^T = V \Sigma^T U^T$

$A^T \cdot A = (V \Sigma^T U^T) U \Sigma V^T$ because $U^T \cdot U = I$

$A^T \cdot A \rightarrow V \Sigma^T \Sigma V^T$

Now $A \cdot A^T = (U \Sigma V^T)(V \Sigma^T U^T)$ because $V^T \cdot V = I$

$A A^T = U \Sigma \Sigma^T U^T$

$$A = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}_{2\times 3}$$

$A^T \cdot A = V$

$A \cdot A^T = U$

$$A^T = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}_{3\times 2}$$

$$A \cdot A^T = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3\times 3 + 1\times 1 + 1\times 1 & 3\times -1 + 1\times 3 + 1\times 1 \\ -1\times 3 + 3\times 1 + 1\times 1 & -1\times -1 + 3\times 3 + 1\times 1 \end{bmatrix}_{2\times 2}$$

$$\Rightarrow \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}_{2\times 2}$$

$A - \lambda I = 0$

$$\begin{bmatrix} (11-\lambda) & 1 \\ 1 & (11-\lambda) \end{bmatrix} = 0$$

$(11-\lambda)^2 - 1^2 = 0$

$(11-\lambda+1)(11-\lambda-1) = 0$

$(12-\lambda) = 0 \qquad (10-\lambda) = 0$

$$\boxed{\lambda_1 = 12, \quad \lambda_2 = 10}$$

first we put $\lambda_1 = 12 \quad \lambda_2 = 10$

$$\begin{bmatrix} (11-\lambda) & 1 \\ 1 & (11-\lambda) \end{bmatrix} \Rightarrow \begin{bmatrix} (11-10) & 1 \\ 1 & (11-12) \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \sqrt{1^2+1^2} = \sqrt{1^2+(-1)^2}$$

orthogenalization using Gram Schmidt orthogenalization process:-

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = U$$

$$A^T A = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}_{3\times 3}$$

$$(A-\lambda I)x = 0$$

$$\Rightarrow \begin{bmatrix} (10-\lambda) & 0 & 2 \\ 0 & (10-\lambda) & 4 \\ 2 & 4 & 2-\lambda \end{bmatrix}$$

Short cut Method

① $\quad \lambda^3 \quad \lambda^2 \quad \lambda$

② $\quad + \quad - \quad + \quad -$

③ $\quad S_1 \quad S_2 \quad S_3$

$$\lambda^3 - S_1\lambda^2 - S_2\lambda - S_3$$

S_1 = trace (A), S_2 = minors of diagonals, S_3 = determinant of (A)

trace(A) = Sum of diagonal value

$\Rightarrow 10+10+2$

$\Rightarrow 2$

determinant(A) = $10\begin{bmatrix}10 & 4\\ 4 & 2\end{bmatrix}+0\begin{bmatrix}0 & 4\\ 2 & 2\end{bmatrix}+2\begin{bmatrix}0 & 10\\ 2 & 4\end{bmatrix}$

$$\begin{bmatrix}10 & 0 & 2\\ 0 & 10 & 4\\ 2 & 4 & 2\end{bmatrix}$$

$\Rightarrow 10[20-16]+0+2[0-20]$

$\Rightarrow 10[4]+2(-20)$

$\Rightarrow 40-40=0$

minors of diagonals:

$4 \quad + \quad 16+100 \Rightarrow 120$

$$\begin{vmatrix}10 & 4\\ 4 & 2\end{vmatrix} \quad \begin{vmatrix}10 & 2\\ 2 & 2\end{vmatrix} \quad \begin{vmatrix}10 & 0\\ 0 & 10\end{vmatrix}$$

for diagonal value (10) (10) (2)

$$\boxed{\lambda^3-22\lambda^2+120\lambda=0}$$

$\lambda(\lambda^2-22\lambda+120)=0 \qquad \lambda_1=0$

$\lambda^2-22\lambda+120=0$

$\lambda^2-12\lambda-10\lambda+120=0$

$\lambda(\lambda-12)-10(\lambda-12)=0$

$$\boxed{\begin{aligned}\lambda_2&=12\\ \lambda_3&=10\end{aligned}}$$

$d_1 = 12 \quad d_2 = 10 \quad d_3 = 0$

$d_1 = 12$

$$\begin{bmatrix} (10-d) & 0 & 2 \\ 0 & (10-d) & 4 \\ 2 & 4 & (2-d) \end{bmatrix}$$

$$\begin{bmatrix} V_1^T \\ V_2^T \\ V_3^T \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -2 & 0 & 2 \\ 0 & -2 & 4 \\ 2 & 4 & -10 \end{bmatrix}$$

$$(A - dI)X = 0$$

So we have to find $x_1\ x_2\ x_3$

$$\begin{bmatrix} -2 & 0 & 2 \\ 0 & -2 & 4 \\ 2 & 4 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

By Cramer's rule.

$$\frac{x_1}{\begin{bmatrix} -2 & 4 \\ 4 & -10 \end{bmatrix}} = \frac{-x_2}{\begin{bmatrix} 0 & 4 \\ 2 & -10 \end{bmatrix}} = \frac{x_3}{\begin{bmatrix} 0 & -2 \\ 2 & 4 \end{bmatrix}}$$

$$\frac{x_1}{4} = \frac{-x_2}{8} = \frac{x_3}{4}$$

minimum we divide it by 4

$$\frac{x_1}{1} \quad 2 \quad 1$$

When $d_2 = 10$

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 2 & 4 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$V = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & -5 \end{bmatrix}$$

$$\frac{x_1}{-16} = \frac{-x_2}{-8} = \frac{x_3}{0}$$

divide by -8 $\Rightarrow$

$d_3 = 0$

$$\begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\frac{x_1}{4} = \frac{-x_2}{-8} = \frac{x_3}{-20}$$

divide by 4

$$[1, 2, -5]$$

$$\Rightarrow V^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & -5 \end{bmatrix}$$

In $A = U \Sigma V^T$

They both are orthogonal vector.

So for orthogonalization we apply

$$V^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & -5 \end{bmatrix}$$

$\sqrt{1^2+2^2+1^2} = \sqrt{6}$

$\sqrt{2^2+1^2+0^2} = \sqrt{5}$

$\sqrt{1^2+2^2+(-5)^2} = \sqrt{30}$

$$V^T = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{0}{\sqrt{5}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{30}} & \frac{-5}{\sqrt{30}} \end{bmatrix}$$

$A_{2\times 3} = \Sigma_{2\times 3}$

$d_1 = 12$ $d_2 = 10$ are repeating. these are unique value

→ Higher value.

$$\Sigma = \begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{bmatrix}$$

Singular matrix. is smaller.

Increasing to Decreasing way.