1.Regularized regression

2. Random Forest

3. kNN (Assume k as 5):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Brightness | 40 | 50 | 60 | 10 | 70 | 60 | 25 | 20 |
| Saturation | 20 | 50 | 90 | 25 | 70 | 10 | 80 | 35 |
| Class | Red | Blue | Blue | Red | Blue | Red | Blue | ? |

4. kNN (Assume k as 3):

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Name | Ajay | Mark | Sara | Zaira | Sachin | Rahul | Pooja | Smith | Laxmi | Mitchell | Angelina |
| Age | 32 | 40 | 16 | 34 | 55 | 40 | 20 | 15 | 55 | 15 | 5 |
| Gender | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| Game | Football | Neither | Cricket | Cricket | Neither | Cricket | Neither | Cricket | Football | Football | ? |

5. Simple Linear Regression:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X | 2 | 4 | 6 | 8 |
| Y | 3 | 7 | 5 | 10 |

6. Multiple Linear Regression:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X1 | 60 | 62 | 67 | 70 | 71 | 72 | 75 | 78 |
| X2 | 22 | 25 | 24 | 20 | 15 | 14 | 14 | 11 |
| Y | 140 | 155 | 159 | 179 | 192 | 200 | 212 | 215 |

7. Polynomial Regression (y=a0+a1\*x+a2\*x2):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X | 3 | 4 | 5 | 6 | 7 |
| Y | 2.5 | 3.2 | 3.8 | 6.5 | 11.5 |

8. SVD:

A =

9. Find the maximum, minimum, median, first quartile, third quartile for the given data set: 23, 42, 12, 10, 15, 14, 9. Plot the outliers on the Box Plot.

10 Find the Outliers in Every Field and Plot on a Box Plot and Scatter Plot:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Height | 2.3 | 2.7 | 3.1 | 3.6 | 3.8 | 4 | 4.3 | 4.5 |

11. Replace missing values in the given dataset using Mean or Median or Mode (As per the data in the field)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | Gender | Salary | Country | Company |
| 1 | Male | 15000 | India | Google |
| 2 | Female | 45000 | China | NaN |
| 3 | Female | 25000 | India | Google |
| 4 | NaN | NaN | Australia | Google |
| 5 | Male | NaN | India | Google |
| 6 | Male | 54000 | NaN | Alibaba |
| 7 | NaN | 74000 | China | NaN |
| 8 | Male | 14000 | Australia | NaN |
| 9 | Female | 15000 | NaN | NaN |
| 10 | Male | 33000 | Australia | NaN |

12. ) Draw the Histogram for the following data:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bulb Life | 300-400 | 400-500 | 500-600 | 600-700 | 700-800 | 800-900 | 900-1000 |
| Number of Lamps | 14 | 56 | 60 | 86 | 74 | 62 | 48 |

13. PCA:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 2.5 | 0.5 | 2.2 | 1.9 | 3.1 | 2.3 | 2 | 1 | 1.2 | 1.1 |
| Y | 2.4 | 0.7 | 2.9 | 2.2 | 3.0 | 2.7 | 1.6 | 1.1 | 1.6 | 0.9 |

14. Find the Outliers and draw Box Plot for every field and Scatter Plot for the below data:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 3 | 0 | 12 | 0 | 2 | 0 | 26 | 0 | 7 | 5 | 5 | 2 | 1 | 1 | 2 |
| Y | 12 | 0 | 9 | 4 | 15 | 2 | 6 | 10 | 27 | 15 | 5 | 9 | 1 | 14 | 2 |