|  |  |  |  |
| --- | --- | --- | --- |
| **Sinhgad Technical Education Society’s**  **Sinhgad Academy of Engineering, KONDHWA (BK),PUNE**  **COMPUTER ENGINEERING DEPARTMENT**  **ASSIGNMENT NO-04**  3  **T.E (Computer) (Semester –V) A.Y.2021-2022**  **Subject: Data Science And Big Data Analytics Division –ALL**   |  | | --- | |  | | | |
|  | | |
| |  |  | | --- | --- | | Que. No | Questions | | Q1) | Explain a few data quality issues. | | Q2) | Write a short note on association rules | | Q3) | |  |  |  |  | | --- | --- | --- | --- | | Transaction | Item1 | Item 2 | Item 3 | | 1 | Bread | Butter | Milk | | 2 | Bread | Butter | Milk | | 3 | Bread | Butter |  | | 4 | Bread | Butter | Milk | | 5 | Bread | Butter |  |   For the given dataset,calculate support for all the possible itemsets. | | Q4) | What is the market basket analysis? How can you use it to define a business strategy? | | Q5) | Write a short note on regression analysis. | | Q6) | Describe some of the application of Logistic Regression. | | Q7) | Define apriori algorithm. Why is it used? | | Q8) | Write a short note on decision trees. | | Q9) | Explain Naïve Bayes Classifier | | Q10) | Explain the various approaches to improve the efficiency of apriori algorithm | |