|  |  |
| --- | --- |
| Semester | T.E. Semester V – Computer Engineering |
| Subject | Data Warehousing and Mining |
| Subject Professor In-charge | Prof. Kavita Shirsat |
| Assisting Teachers | Prof. Kavita Shirsat |
| Laboratory | Lab 312 A |

|  |  |  |
| --- | --- | --- |
| Student Name | Deep Salunkhe | |
| Roll Number | 21102A0014 | |
| Grade and Subject  Teacher’s Signature |  |  |

|  |  |  |
| --- | --- | --- |
| Experiment Number | 07 | |
| Experiment Title | Implementation of Apriori algorithm. | |
| Resources / Apparatus Required | Hardware:  Computer system | Software:  Python |
| Description | The Apriori algorithm is a widely used algorithm in data mining and association rule learning. It is used to discover frequent itemsets in a transaction database and generate association rules based on these frequent itemsets. Here's a theoretical overview of the Apriori algorithm:   * **Introduction:** The Apriori algorithm is designed to find associations or relationships between items in a transactional database. * **Objective:** Its primary goal is to identify frequent itemsets, which are sets of items that often appear together in transactions. * **Support:** The algorithm uses a user-defined threshold called "support" to determine the minimum frequency required for an itemset to be considered frequent. * **Apriori Principle:** The algorithm is based on the Apriori principle, which states that if an itemset is frequent, then all of its subsets must also be frequent. * **Algorithm Steps:**  1. Start with individual items as 1-itemsets. 2. Count the support of each 1-itemset by scanning the database. 3. Prune infrequent 1-itemsets. 4. Generate 2-itemsets from the remaining 1-itemsets. 5. Count the support of each 2-itemset. 6. Prune infrequent 2-itemsets.   Continue this process, incrementally generating larger itemsets and pruning until no more frequent itemsets can be found.   * **Termination:** The algorithm terminates when no more frequent itemsets can be generated. * **Association Rule Generation:** After identifying frequent itemsets, the Apriori algorithm can generate association rules. These rules express relationships between items, e.g., "If A and B are bought, then C is also bought." | |
| Program |  | |
| Output | A screenshot of a computer program  Description automatically generated | |
| Conclusion: | The Apriori algorithm is a fundamental tool in data mining for discovering association rules among items in a dataset. It uses support and confidence measures, as well as the Apriori principle, to efficiently identify frequent itemsets and generate meaningful association rules. | |