|  |  |
| --- | --- |
| Semester | T.E. Semester VI – 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 | 01 | |
| Experiment Title | Naive Bayes classifier approach | |
| Resources / Apparatus Required | Hardware:  Computer system | Software:  Python |
| Description | Naive Bayes is a popular and simple machine learning classification algorithm that is based on Bayes' theorem. It's particularly useful for text classification and is known for its efficiency and effectiveness in various applications. Here's some key information about the Naive Bayes classifier approach:  **1. Bayes' Theorem:**   * The Naive Bayes classifier is built on Bayes' theorem, which is a fundamental probability theorem used to calculate conditional probabilities. * Bayes' theorem calculates the probability of an event based on prior knowledge of conditions that might be related to the event.   **2. Independence Assumption:**   * The "Naive" in Naive Bayes refers to the assumption that all features (attributes) used to predict the class label are independent of each other. In reality, this assumption is often not true, but it simplifies the calculations significantly and still produces good results in many cases. * Despite this simplification, Naive Bayes can perform surprisingly well, especially for text classification tasks.   **3. Classification Task:**   * Naive Bayes is primarily used for classification tasks, where the goal is to categorize data into predefined classes or labels. * It's commonly used for text classification problems, such as spam email detection, sentiment analysis, and document categorization.   **4. Probability Calculation:**   * The Naive Bayes classifier calculates the probability of an instance belonging to each possible class and assigns the instance to the class with the highest probability. * It uses prior probabilities (based on training data) and conditional probabilities of each feature given the class to make these calculations. | |
| Program |  | |
| Output |  | |
| Conclusion: | Naive Bayes classifiers are a family of probabilistic classifiers that are particularly useful for text classification tasks and other situations where the independence assumption is a reasonable approximation. They are easy to implement, computationally efficient, and can deliver good results with proper data preprocessing and handling of feature independence. | |