Hall Ticket Number: 1703PH0651
Title : Privacy Preserving Data Mining
Status: InProgress

Privacy Preserving Data Mining

Abstract:

In today's databases, there is a significant amount of sensitive data present in the attributes of database in structured/unstructured form. As a result, it's critical to design data structures that limit the importance or access of personal data. Various government and organization foundations are naturally gathering personal records of persons for the purposes of information analysis. These organizations encourage information evaluation in order to disseminate "adequately private" ideas about the data acquired. Privacy could be a double-edged sword: there should be enough safeguards in place to ensure that sensitive information about the general public is kept private. isn't revealed by the views, and there should be enough data to carry out the inquiry in a reasonable amount of time. Furthermore, an adversary attempting to obtain sensitive information from unprotected perspectives may contain information about the general population.

In micro data distribution, information security is a critical concern. Anonymity methods are frequently used to assure single security while having no impact on the substance of the data released. Recently, a couple of approaches have become well-known for maintaining data security or perhaps minimizing data loss to the greatest extent possible. That is, they improve the anonymous system's adaptability to make it more realistic, and then to suit the diverse needs of the general public. In the meantime, other propositions and estimates have been made for them.

Different approaches has been proposed by the reserachers in different aspects like some have worked on identifying proper ‘k’ value in k-anonimity, worked on time complexity when we are grouping them using supervised learning approaches and some have used optimization techniques like Dragon Particle Swarm Optimization. Inspite of have this there is a scope of improving k-anonymity by addressing in different aspects like time complexity, attacks , in terms of data utility along with privacy concern and not the last but also can be improved by using better group while handling with dynamic datasets.