ASSIGNMENT-2

MACHINE LEARNING

1. D) OPTION 2 AND 4
2. B) OPTION 1 AND 2
3. B)FALSE
4. REOMOVAL OF OUTLIERS
5. A) 0
6. B) NO
7. A) YES
8. D) ALL OF THE ABOVE
9. A) K MEANS CLUSTERING ALGORITHM
10. D) ALL OF THE ABOVE
11. D) ALL OF THE ABOVE
12. YES
13. K MEANS IS A POPULAR CLUSTERING ALGORITHM IN MACHINE LEARNING AND HAS SEVERAL ADVANTAGES THAT MAKES IT A BETTER CHOICE FOR CERTAIN APPLICATION:
    - SIMPLICITY: K MEANS IS A SIMPLE AND EASY TO IMLEMENT ALGORITHM. IT IS BASED ON A SIMPLE MATHEMATICAL PRINCIPLE OF MINIMIZING THE SUM OF SQUARED DISTANCES BETWEEN DATA POINTS AND THEIR ASSIGNED CLUSTER CENTERS
    - SCALABILITY: K MEANS IS COMPUTATIONALLY EFFICIENT AND CAN HANDLE LARGE DATA SET WITH HIGH DIMENSIONS. IT CAN ALSO BE PARALLELIZED, MAKING IT SCALABLE TO BIG DATA PROBLEMS.
    - FLEXIBILITY: K MEANS CAN CAN HANDLE A VARIETY OF DATA TYPES, INCLUDING CONTINOUS AND CATEGORICAL VARIABLES. IT CAN ALSO BE ADAPTED FOR DIFFERENT DISTANCE METRICS AND SIMILARITY MEASURES, DEPENDING ON THE DATA AND THE PROBLEM AT HAND
    - INTERPRETABLE: THE RESULTING CLUSTERS IN K MEANS ARE EASILY INTERPRETABLE AND CAN PROVIDE INSIGHTS INTO THE UNDERLYING STRUCTURES OF THE DATA. THIS CAN HELP IN DECISION MAKING AND PROBLEM SOLVING IN VAIOUS DOMAINS.
14. YES , KMEANS IS A DETERMINISTIC ALGORITHM, MEANING THAT GIVEN THE SAME INPUT DATA AND PARAMETERS, IT WILL RPODUCE THE SAME RESULTS EVERY TIME IT IS RUN.