Experiment 3

Shashwat Shah
60004220126
Div B C2-2

Aim: To implement Classification using Decision tree ID3 and Naive Bayes Algorithm.

Theory: It is a supervised learning technique that involves training a model to predict the class or category of instanse based on their attributes.
The Model is consumed using a labelled dataset where instance are paired with their corresponding class labels.

During trainning it learns decision boandaries rule that distingui between different classes. Common models include decision tree naive bayes, SVMs and k nearest ~~neo~~ neighbours.
Evaluation metrics sum as accuracy, precision and the confusion matrix are used.
It finds application in diverse fields such as spam detection, image recognition, medical dignosis, audit survey.

Decision trees - It is a structured tree like flowchart whis consists of nodes representing decision points, branches denoting possible outcomes, and leaves signifying final predicted classes.

These are susceptible to overfitting which maybe mitigated by techniques like privilege.
They are valued for their transparency.

Naive Bayes : It is a probabilistic machine learning algorithm based on Bayes theorum. It operates on assumptions of features independence where the impact of a particular feature of classification is considered independently of other features. It performs well in text classification and spam filtering. It calculates the probability of an instance belonging to a particular class given its features values.

Conclusion : We compared 5 datasets and observed that Naive Bayes classification has a higher accuracy than decision tree algorithm. We also used k-cross validation & ensemble methods to enhance the accuracy.