Quiz#2

Given the dataset below, predict Andrew’s default status using KNN with k =5.

| **Customer** | **Age** | **Loan** | **Default** |
| --- | --- | --- | --- |
| John | 25 | 40000 | N |
| Smith | 35 | 60000 | N |
| Alex | 45 | 80000 | N |
| Jade | 20 | 20000 | N |
| Kate | 35 | 120000 | N |
| Mark | 52 | 18000 | N |
| Anil | 23 | 95000 | Y |
| Pat | 40 | 62000 | Y |
| George | 60 | 100000 | Y |
| Jim | 48 | 220000 | Y |
| Jack | 33 | 150000 | Y |
| Andrew | 48 | 142000 | ? |

John: 58000

Smith: 78000

Alex: 103000

Jade: 22000

Kate: 98000

Mark: 122000

Anil: 71000

Pat: 8000

George: 58000

Jim: 78000

Jack: 28000

Nearest neighbors: Pat, Jade, Jack, George, John

Euclidean Distance:

For Andrew (48, 142000):

Distance to Pat (40, 62000): (48-40)^2 + (142000-62000)^2 = 82000

Distance to Jade (20, 20000):(48-20)^2 + (142000-20000)^2 = 122000

Distance to Jack (33, 150000): (48-33)^2 + (142000-150000)^2 = 11000

Distance to George (60, 100000): (48-60)^2 + (142000-100000)^2= 45000

Distance to John (25, 40000):(48-25)^2 + (142000-40000)^2 = 102000

Selecting K Nearest Neighbors:

Nearest neighbors: Pat, Jade, Jack, George, John

Majority Voting:

Among these neighbors, the majority class is N (non-default).

Prediction:

Predict that Andrew's default status is N (non-default).

So, based on the KNN algorithm with k = 5 and Euclidean distance, Andrew's default status is predicted to be N (non-default).

The 5 nearest neighbors to Andrew are Pat, Jade, Jack, George, and John, with default status of N, N, Y, Y, and N respectively. Among these 5 neighbors, the majority class is N (non-default), so we predict that Andrew's default status is also N.