

Pattern Recognition

Question 3:

A bank wants to create a relationship between the amount of loan granted to a client, average delay in payments, and the client's gender. The following data is available.

1. Find out if there is any relationship between the amount of loan and delay, and gender. Consider Amount and Delay as features to assign each feature vector to one the classes 'Male' or 'Female'. Randomly choose 5 feature vectors for testing your method and use the remaining feature vectors as your data set.
2. Use K-Nearest Neighbor to implement your solution. Change parameter K (use 1,3,5). Discuss your results for each value.
3. Discuss the effect of un-equal number of data items from each class on your results.

Data

Loan Amount	Delay (Days)	Gender
87	2	F
89	2	F
90	2	F
90	2	F
101	2	F
110	2	F
115	2	F
115	2	F
115	2	F
117	2	F
127	2	F
149	2	F
150	2	F
157	2	F
166	2	F
170	2	F
183	2	F
183	2	F
190	2	F
201	2	F
202	2	F
207	2	F
212	2	F
220	2	F
310	2	F

322	2	F
399	2	F
415	2	F
326	3	F
12	4	F
12	4	F
38	4	F
200	4	F
302	4	F
30	5	F
44	5	F
131	5	F
23	7	F
152	7	F
89	10	F
24	12	F
41	12	F
151	14	F
25	17	F
48	18	F
12	19	F
267	20	F
18	22	F
43	25	F
15	28	F
60	28	F
86	28	F
12	29	F
77	29	F
38	30	F
46	30	F
64	30	F
77	30	F
132	30	F
150	30	F
12	31	F
24	31	F
12	32	F
27	32	F
36	34	F
26	46	F

26	46	F
116	46	F
88	53	F
27	60	F
84	60	F
162	60	F
277	60	F
51	66	F
12	11	F
100	26	M
78	27	M
61	48	M
343	28	M
24	49	M
43	53	M
417	13	M
275	54	M
42	57	M
26	58	M
42	59	M
47	59	M
170	59	M
12	60	M
49	62	M
36	64	M
48	65	M
17	66	M
85	69	M
350	39	M
207	70	M
229	70	M
65	72	M
37	76	M
73	80	M