Test Results

Experimental comparison of k-NN and linear classification on the Iris data-set

---

Scenario 1.1 Outer ratio: 60/90 Inner ratio: 15/45					
Test ID	**ML Classification Model**	**k**	**No of random "folds" (Bagging)**	**Average Validation Error**	**Average Test Error (on new data)**
1.1	Linear	N/A	1,000	0.0480	0.1111
1.2	Linear	N/A	1,000	0.0481	0.0444
1.3	Linear	N/A	1,000	0.0471	0.0889
1.4	Linear	N/A	1,000	0.0495	0.0667
1.5	Linear	N/A	1,000	0.0451	0.0556
1.6	Linear	N/A	1,000	0.0459	0.0333
1.7	Linear	N/A	1,000	0.0497	0.0667
1.8	Linear	N/A	1,000	0.0468	0.0444
1.9	Linear	N/A	1,000	0.0451	0.0889
1.10	Linear	N/A	1,000	0.0496	0.1111
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.0475** • Average Tests Error = **0.0711** • **Time: 174.282664 sec**					
2.1	kNN	1	1,000	0.0586	N/A
2.2	kNN	2	1,000	0.0889	N/A
2.3	kNN	3	1,000	0.0960	N/A
2.4	kNN	4	1,000	0.1473	N/A
2.5	kNN	5	1,000	0.1660	N/A
2.6	kNN	6	1,000	0.2177	N/A
2.7	kNN	7	1,000	0.2777	N/A
2.8	kNN	8	1,000	0.3264	N/A

2.9	kNN	9	1,000	0.4018	N/A
2.10	kNN	10	1,000	0.4480	N/A
2.11	kNN	11	1,000	0.4997	N/A
2.19	kNN	19	1,000	0.7044	N/A
2.25	kNN	25	1,000	0.7051	N/A
—–> The kNN model is optimum for the value **k = 1**					
2.1.1	kNN	1	1,000	0.0561	0.0889
2.1.2	kNN	1	1,000	0.0595	0.0667
2.1.3	kNN	1	1,000	0.0568	0.0444
2.1.4	kNN	1	1,000	0.0573	0.0889
2.1.5	kNN	1	1,000	0.0576	0.0889
2.1.6	kNN	1	1,000	0.0552	0.0778
2.1.7	kNN	1	1,000	0.0601	0.0778
2.1.8	kNN	1	1,000	0.0579	0.0667
2.1.9	kNN	1	1,000	0.0550	0.0333
2.1.10	kNN	1	1,000	0.0542	0.0889
| (**k=1**) | Average Validations Error = **0.0570** Average Tests Error = **0.0722** **Time: 46.368117 sec** —–> The kNN model is optimum for the value **k = 1**.					

Scenario 1.2 **Outer ratio: 60/90** **Inner ratio: 30/30**					
Test ID	**ML Classification Model**	**k**	**No of random "folds" (Bagging)**	**Average Validation Error**	**Average Test Error (on new data)**
1.1	Linear	N/A	1,000	0.0217	0.0222
1.2	Linear	N/A	1,000	0.0219	0.0333
1.3	Linear	N/A	1,000	0.0236	0.0222
1.4	Linear	N/A	1,000	0.0206	0.0667

1.5	Linear	N/A	1,000	0.0216	0.0333
1.6	Linear	N/A	1,000	0.0210	0.0667
1.7	Linear	N/A	1,000	0.0227	0.0778
1.8	Linear	N/A	1,000	0.0225	0.0556
1.9	Linear	N/A	1,000	0.0223	0.0556
1.10	Linear	N/A	1,000	0.0213	0.0333
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.0219** • Average Tests Error = **0.0467** • **Time: 173.770359 sec**					
2.1	kNN	1	1,000	0.0376	N/A
2.2	kNN	2	1,000	0.0413	N/A
2.3	kNN	3	1,000	0.0356	N/A
2.4	kNN	4	1,000	0.0529	N/A
2.5	kNN	5	1,000	0.0475	N/A
2.6	kNN	6	1,000	0.0625	N/A
2.7	kNN	7	1,000	0.0630	N/A
2.8	kNN	8	1,000	0.0822	N/A
2.9	kNN	9	1,000	0.0831	N/A
2.10	kNN	10	1,000	0.0982	N/A
2.11	kNN	11	1,000	0.1119	N/A
2.19	kNN	19	1,000	0.4018	N/A
2.25	kNN	25	1,000	0.5474	N/A
——> The kNN model is optimum for the value **k = 3**					
2.3.1	kNN	3	1,000	0.0337	0.0444
2.3.2	kNN	3	1,000	0.0342	0.0778
2.3.3	kNN	3	1,000	0.0329	0.0667
2.3.4	kNN	3	1,000	0.0326	0.0444
2.3.5	kNN	3	1,000	0.0355	0.1000
2.3.6	kNN	3	1,000	0.0342	0.0444
2.3.7	kNN	3	1,000	0.0338	0.0667

2.3.8	kNN	3	1,000	0.0335	0.0667
2.3.9	kNN	3	1,000	0.0317	0.0667
2.3.10	kNN	3	1,000	0.0350	0.0667

| (**k=3**) | Average Validations Error = **0.0337** Average Tests Error = **0.0644**
Time: 45.395600 sec

——> The kNN model is optimum for the value **k = 3**.

Scenario 1.3

Outer ratio: 60/90
Inner ratio: 45/15

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0196	0.0333
1.2	Linear	N/A	1,000	0.0164	0.0444
1.3	Linear	N/A	1,000	0.0193	0.0444
1.4	Linear	N/A	1,000	0.0198	0.0222
1.5	Linear	N/A	1,000	0.0208	0.0444
1.6	Linear	N/A	1,000	0.0184	0.0444
1.7	Linear	N/A	1,000	0.0191	0.0556
1.8	Linear	N/A	1,000	0.0184	0.0444
1.9	Linear	N/A	1,000	0.0186	0.0444
1.10	Linear	N/A	1,000	0.0170	0.0556

Total averages for test runs **1.1 to 1.10**:

- Average Validations Error = **0.0187**
- Average Tests Error = **0.0433**
- **Time = 160.260601 sec**

2.1	kNN	1	1,000	0.0428	N/A
2.2	kNN	2	1,000	0.0328	N/A
2.3	kNN	3	1,000	0.0244	N/A
2.4	kNN	4	1,000	0.0319	N/A

2.5	kNN	5	1,000	0.0300	N/A
2.6	kNN	6	1,000	0.0440	N/A
2.7	kNN	7	1,000	0.0365	N/A
2.8	kNN	8	1,000	0.0452	N/A
2.9	kNN	9	1,000	0.0416	N/A
2.10	kNN	10	1,000	0.0522	N/A
2.11	kNN	11	1,000	0.0507	N/A
2.19	kNN	19	1,000	0.0910	N/A
2.25	kNN	25	1,000	0.2018	N/A
——> The kNN model is optimum for the value **k = 3**					
2.3.1	kNN	3	1,000	0.0216	0.0667
2.3.2	kNN	3	1,000	0.0213	0.0556
2.3.3	kNN	3	1,000	0.0248	0.0333
2.3.4	kNN	3	1,000	0.0236	0.0667
2.3.5	kNN	3	1,000	0.0223	0.0778
2.3.6	kNN	3	1,000	0.0216	0.0778
2.3.7	kNN	3	1,000	0.0224	0.0333
2.3.8	kNN	3	1,000	0.0233	0.0667
2.3.9	kNN	3	1,000	0.0230	0.0556
2.3.10	kNN	3	1,000	0.0240	0.0778
| (**k=3**) | Average Validations Error = **0.0228** Average Tests Error = **0.0611** **Time: 43.927965 sec**					
——> The kNN model is optimum for the value **k = 3**.					

Scenario 2.1 Outer ratio: 75/75 Inner ratio: 25/50					
Test ID	**ML Classification Model**	**k**	**No of random "folds" (Bagging)**	**Average Validation Error**	**Average Test Error (on new data)**
1.1	Linear	N/A	1,000	0.0416	0.0667
1.2	Linear	N/A	1,000	0.0411	0.0267
1.3	Linear	N/A	1,000	0.0416	0.0800
1.4	Linear	N/A	1,000	0.0418	0.0400
1.5	Linear	N/A	1,000	0.0429	0.1067
1.6	Linear	N/A	1,000	0.0425	0.0267
1.7	Linear	N/A	1,000	0.0429	0.0400
1.8	Linear	N/A	1,000	0.0405	0.0533
1.9	Linear	N/A	1,000	0.0416	0.0133
1.10	Linear	N/A	1,000	0.0399	0.0400
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.0416** • Average Tests Error = **0.0493** • **Time: 168.756263 sec**					
2.1	kNN	1	1,000	0.0560	N/A
2.2	kNN	2	1,000	0.0718	N/A
2.3	kNN	3	1,000	0.0553	N/A
2.4	kNN	4	1,000	0.0773	N/A
2.5	kNN	5	1,000	0.0725	N/A
2.6	kNN	6	1,000	0.0957	N/A
2.7	kNN	7	1,000	0.0927	N/A
2.8	kNN	8	1,000	0.1261	N/A
2.9	kNN	9	1,000	0.1373	N/A
2.10	kNN	10	1,000	0.1557	N/A
2.11	kNN	11	1,000	0.1890	N/A
2.19	kNN	19	1,000	0.4933	N/A

2.25	kNN	25	1,000	0.7072	N/A
—–> The kNN model is optimum for the value **k = 3**					
2.3.1	kNN	3	1,000	0.0584	0.1067
2.3.2	kNN	3	1,000	0.0576	0.0800
2.3.3	kNN	3	1,000	0.0568	0.0667
2.3.4	kNN	3	1,000	0.0590	0.0267
2.3.5	kNN	3	1,000	0.0588	0.0533
2.3.6	kNN	3	1,000	0.0592	0.0533
2.3.7	kNN	3	1,000	0.0605	0.0267
2.3.8	kNN	3	1,000	0.0582	0.0400
2.3.9	kNN	3	1,000	0.0580	0.0533
2.3.10	kNN	3	1,000	0.0609	0.0933
| (**k=3**) | Average Validations Error = **0.0587** Average Tests Error = **0.0600** **Time: 48.192874 sec** —–> The kNN model is optimum for the value **k = 3**.					

Scenario 2.2

Outer ratio: 75/75
Inner ratio: 45/30

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0305	0.0533
1.2	Linear	N/A	1,000	0.0297	0.0533
1.3	Linear	N/A	1,000	0.0293	0.0533
1.4	Linear	N/A	1,000	0.0278	0.0400
1.5	Linear	N/A	1,000	0.0283	0.0667
1.6	Linear	N/A	1,000	0.0286	0.0533
1.7	Linear	N/A	1,000	0.0286	0.0667
1.8	Linear	N/A	1,000	0.0292	0.0400
1.9	Linear	N/A	1,000	0.0289	0.0667

1.10	Linear	N/A	1,000	0.0302	0.0667
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.0291** • Average Tests Error = **0.0560** • **Time: 184.449941 sec**					
2.1	kNN	1	1,000	0.0525	N/A
2.2	kNN	2	1,000	0.0534	N/A
2.3	kNN	3	1,000	0.0464	N/A
2.4	kNN	4	1,000	0.0504	N/A
2.5	kNN	5	1,000	0.0451	N/A
2.6	kNN	6	1,000	0.0544	N/A
2.7	kNN	7	1,000	0.0501	N/A
2.8	kNN	8	1,000	0.0592	N/A
2.9	kNN	9	1,000	0.0558	N/A
2.10	kNN	10	1,000	0.0664	N/A
2.11	kNN	11	1,000	0.0618	N/A
2.19	kNN	19	1,000	0.1063	N/A
2.25	kNN	25	1,000	0.2165	N/A
——> The kNN model is optimum for the value **k = 5**					
2.5.1	kNN	5	1,000	0.0448	0.0267
2.5.2	kNN	5	1,000	0.0447	0.0533
2.5.3	kNN	5	1,000	0.0437	0.0400
2.5.4	kNN	5	1,000	0.0435	0.0667
2.5.5	kNN	5	1,000	0.0456	0.0800
2.5.6	kNN	5	1,000	0.0452	0.0800
2.5.7	kNN	5	1,000	0.0427	0.0667
2.5.8	kNN	5	1,000	0.0446	0.0533
2.5.9	kNN	5	1,000	0.0437	0.0267
2.5.10	kNN	5	1,000	0.0448	0.0800
| (**k=5**) | Average Validations Error = **0.0443** Average Tests Error = **0.0573** **Time: 48.422533 sec** ——> The kNN model is optimum for the value **k = 5.**					

Scenario 2.3 Outer ratio: 75/75 Inner ratio: 65/10					
Test ID	**ML Classification Model**	**k**	**No of random "folds" (Bagging)**	**Average Validation Error**	**Average Test Error (on new data)**
1.1	Linear	N/A	1,000	0.0251	0.0133
1.2	Linear	N/A	1,000	0.0244	0.0533
1.3	Linear	N/A	1,000	0.0238	0.0267
1.4	Linear	N/A	1,000	0.0240	0.0400
1.5	Linear	N/A	1,000	0.0249	0.0533
1.6	Linear	N/A	1,000	0.0248	0.0400
1.7	Linear	N/A	1,000	0.0253	0.0133
1.8	Linear	N/A	1,000	0.0232	0.0400
1.9	Linear	N/A	1,000	0.0219	0.0667
1.10	Linear	N/A	1,000	0.0247	0.0533
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.0242** • Average Tests Error = **0.0400** • **Time: 170.392563**					
2.1	kNN	1	1,000	0.0541	N/A
2.2	kNN	2	1,000	0.0525	N/A
2.3	kNN	3	1,000	0.0478	N/A
2.4	kNN	4	1,000	0.0431	N/A
2.5	kNN	5	1,000	0.0341	N/A
2.6	kNN	6	1,000	0.0452	N/A
2.7	kNN	7	1,000	0.0452	N/A
2.8	kNN	8	1,000	0.0490	N/A
2.9	kNN	9	1,000	0.0403	N/A
2.10	kNN	10	1,000	0.0484	N/A
2.11	kNN	11	1,000	0.0457	N/A
2.19	kNN	19	1,000	0.0657	N/A
2.25	kNN	25	1,000	0.0764	N/A

——> The kNN model is optimum for the value **k = 5**					
2.3.1	kNN	5	1,000	0.0310	0.0400
2.3.2	kNN	5	1,000	0.0316	0.0800
2.3.3	kNN	5	1,000	0.0316	0.0533
2.3.4	kNN	5	1,000	0.0340	0.0800
2.3.5	kNN	5	1,000	0.0315	0.0800
2.3.6	kNN	5	1,000	0.0325	0.0800
2.3.7	kNN	5	1,000	0.0322	0.0800
2.3.8	kNN	5	1,000	0.0321	0.0800
2.3.9	kNN	5	1,000	0.0353	0.0667
2.3.10	kNN	5	1,000	0.0306	0.0800

| (**k=5**) | Average Validations Error = **0.0323** Average Tests Error = **0.0720**
Time: 45.395600 sec
——> The kNN model is optimum for the value **k = 5**.

Scenario 3.1

Outer ratio: 90/60
Inner ratio: 60/30

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	10,000	0.0330	0.0333
1.2	Linear	N/A	10,000	0.0328	0.0333
1.3	Linear	N/A	10,000	0.0337	0.0167
1.4	Linear	N/A	10,000	0.0335	0.0500
1.5	Linear	N/A	10,000	0.0330	0.0167
1.6	Linear	N/A	10,000	0.0328	0.0333
1.7	Linear	N/A	10,000	0.0336	0.0167
1.8	Linear	N/A	10,000	0.0329	0.0333
1.9	Linear	N/A	10,000	0.0326	0.0167

1.10	Linear	N/A	10,000	0.0333	0.0500
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.03312** • Average Tests Error = **0.0300** • **Time: 181.457877 sec**					
2.1	kNN	1	10,000	0.0553	0.0333
2.2	kNN	2	10,000	0.0626	0.0667
2.3	kNN	3	10,000	0.0514	0.0167
2.4	kNN	4	10,000	0.0555	0.0333
2.5	kNN	5	10,000	0.0534	0.0167
2.6	kNN	6	10,000	0.0561	0.0333
2.7	kNN	7	10,000	0.0550	0.0333
2.8	kNN	8	10,000	0.0609	0.0333
2.9	kNN	9	10,000	0.0578	0.0333
2.10	kNN	10	10,000	0.0638	0.0500
2.11	kNN	11	10,000	0.0601	0.0500
2.12	kNN	12	10,000	0.0662	0.0333
2.13	kNN	13	10,000	0.0614	0.0333
2.14	kNN	14	10,000	0.0676	0.0667
2.15	kNN	15	10,000	0.0640	0.0333
2.16	kNN	16	10,000	0.0730	0.0333
2.17	kNN	17	10,000	0.0706	0.0333
2.18	kNN	18	10,000	0.0826	0.0333
2.19	kNN	19	10,000	0.0795	0.0333
2.20	kNN	20	10,000	0.0909	0.0333
2.25	kNN	25	10,000	0.1095	0.1000
2.31	kNN	31	10,000	0.1542	0.0333
2.41	kNN	41	10,000	0.4308	0.1333
2.51	kNN	51	10,000	0.5370	0.6667
2.59	kNN	59	10,000	0.6906	0.5167
Further testing to find optimum k. Best results for the range: **3 < k < 7**					

2.3.1	kNN	3	10,000	0.0513	0.0167
2.3.2	kNN	3	10,000	0.0517	0.0167
2.3.3	kNN	3	10,000	0.0516	0.0167
2.3.4	kNN	3	10,000	0.0511	0.0500
2.3.5	kNN	3	10,000	0.0520	0.0167
2.3.6	kNN	3	10,000	0.0516	0.0167
2.3.7	kNN	3	10,000	0.0517	0.0167
2.3.8	kNN	3	10,000	0.0521	0.0333
2.3.9	kNN	3	10,000	0.0515	0.0167
2.3.10	kNN	3	10,000	0.0519	0.0333
2.4.1	kNN	4	10,000	0.0562	0.0500
2.4.2	kNN	4	10,000	0.0559	0.0500
2.4.3	kNN	4	10,000	0.0554	0.0167
2.4.4	kNN	4	10,000	0.0558	0.0167
2.4.5	kNN	4	10,000	0.0553	0.0167
2.4.6	kNN	4	10,000	0.0551	0.0500
2.4.7	kNN	4	10,000	0.0557	0.0500
2.4.8	kNN	4	10,000	0.0561	0.0167
2.4.9	kNN	4	10,000	0.0564	0.0500
2.4.10	kNN	4	10,000	0.0557	0.0333
2.5.1	kNN	5	10,000	0.0536	0.0500
2.5.2	kNN	5	10,000	0.0530	0.0167
2.5.3	kNN	5	10,000	0.0531	0.0167
2.5.4	kNN	5	10,000	0.0530	0.0333
2.5.5	kNN	5	10,000	0.0534	0.0500
2.5.6	kNN	5	10,000	0.0528	0.0167
2.5.7	kNN	5	10,000	0.0531	0.0500
2.5.8	kNN	5	10,000	0.0530	0.0333
2.5.9	kNN	5	10,000	0.0529	0.0333
2.5.10	kNN	5	10,000	0.0532	0.0500

2.6.1	kNN	6	10,000	0.0569	0.0667
2.6.2	kNN	6	10,000	0.0567	0.0500
2.6.3	kNN	6	10,000	0.0568	0.0167
2.6.4	kNN	6	10,000	0.0560	0.0167
2.6.5	kNN	6	10,000	0.0571	0.0500
2.6.6	kNN	6	10,000	0.0570	0.0333
2.6.7	kNN	6	10,000	0.0568	0.0500
2.6.8	kNN	6	10,000	0.0576	0.0500
2.6.9	kNN	6	10,000	0.0573	0.0333
2.6.10	kNN	6	10,000	0.0564	0.0167
2.7.1	kNN	7	10,000	0.0545	0.0333
2.7.2	kNN	7	10,000	0.0552	0.0333
2.7.3	kNN	7	10,000	0.0543	0.0333
2.7.4	kNN	7	10,000	0.0549	0.0500
2.7.5	kNN	7	10,000	0.0547	0.0333
2.7.6	kNN	7	10,000	0.0541	0.0500
2.7.7	kNN	7	10,000	0.0543	0.0167
2.7.8	kNN	7	10,000	0.0552	0.0500
2.7.9	kNN	7	10,000	0.0556	0.0333
2.7.10	kNN	7	10,000	0.0542	0.0167

(**k=3**)	Average Validations Error = **0.05165**	Average Tests Error = **0.02335**
(k=4)	Average Validations Error = 0.05576	Average Tests Error = 0.03501
(k=5)	Average Validations Error = 0.05311	Average Tests Error = 0.035
(k=6)	Average Validations Error = 0.05686	Average Tests Error = 0.03834
(k=7)	Average Validations Error = 0.0547	Average Tests Error = 0.03499

Time: 62.301395 sec

——> The kNN model is optimum for the value **k = 3**.

Scenario 3.2

Outer ratio: 90/60
Inner ratio: 45/45

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0394	0.0333
1.2	Linear	N/A	1,000	0.0373	0.0500
1.3	Linear	N/A	1,000	0.0385	0.0333
1.4	Linear	N/A	1,000	0.0377	0.0167
1.5	Linear	N/A	1,000	0.0370	0.0333
1.6	Linear	N/A	1,000	0.0395	0.0500
1.7	Linear	N/A	1,000	0.0384	0.0500
1.8	Linear	N/A	1,000	0.0377	0.0167
1.9	Linear	N/A	1,000	0.0396	0.0500
1.10	Linear	N/A	1,000	0.0388	0.0333
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.03839** • Average Tests Error = **0.06672** • **Time: 185.630999 sec**					
2.1	kNN	1	1,000	0.0606	0.0167
2.2	kNN	2	1,000	0.0689	0.0500
2.3	kNN	3	1,000	0.0549	0.0333
2.4	kNN	4	1,000	0.0640	0.0333
2.5	kNN	5	1,000	0.0557	0.0167
2.6	kNN	6	1,000	0.0642	0.0333
2.7	kNN	7	1,000	0.0611	0.0333
2.8	kNN	8	1,000	0.0692	0.0833
2.9	kNN	9	1,000	0.0647	0.0167
2.10	kNN	10	1,000	0.0726	0.0500
2.11	kNN	11	1,000	0.0697	0.0667
2.19	kNN	19	1,000	0.1268	0.1000
2.25	kNN	25	1,000	0.2398	0.0667
——> The kNN model is optimum for the value **k = 3 and k = 5**					
2.3.1	kNN	3	1,000	0.0582	0.0167

2.3.2	kNN	3	1,000	0.0570	0.0167
2.3.3	kNN	3	1,000	0.0558	0.0333
2.3.4	kNN	3	1,000	0.0562	0.0500
2.3.5	kNN	3	1,000	0.0587	0.0167
2.3.6	kNN	3	1,000	0.0569	0.0667
2.3.7	kNN	3	1,000	0.0570	0.0167
2.3.8	kNN	3	1,000	0.0539	0.0500
2.3.9	kNN	3	1,000	0.0572	0.0333
2.3.10	kNN	3	1,000	0.0570	0.0667
2.5.1	kNN	5	1,000	0.0584	0.0167
2.5.2	kNN	5	1,000	0.0565	0.0500
2.5.3	kNN	5	1,000	0.0574	0.0333
2.5.4	kNN	5	1,000	0.0582	0.0500
2.5.5	kNN	5	1,000	0.0575	0.0333
2.5.6	kNN	5	1,000	0.0572	0.0333
2.5.7	kNN	5	1,000	0.0578	0.0333
2.5.8	kNN	5	1,000	0.0572	0.0333
2.5.9	kNN	5	1,000	0.0596	0.0167
2.5.10	kNN	5	1,000	0.0574	0.0333

| (**k=3**) | Average Validations Error = **0.05679** Average Tests Error = **0.03668**
| (**k=5**) | Average Validations Error = **0.05772** Average Tests Error = **0.03332**
Time: 60.828525 sec

——> The kNN model is optimum for the value **k = 3** because the average validation error is a better representation of the model performance as more tests are performed overall, and it is lower for k = 3.

Scenario 3.3

Outer ratio: 90/60
Inner ratio: 70/20

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0317	0.0167

1.2	Linear	N/A	1,000	0.0309	0.0333
1.3	Linear	N/A	1,000	0.0308	0.0333
1.4	Linear	N/A	1,000	0.0304	0.0167
1.5	Linear	N/A	1,000	0.0309	0.0167
1.6	Linear	N/A	1,000	0.0307	0.0500
1.7	Linear	N/A	1,000	0.0301	0.0333
1.8	Linear	N/A	1,000	0.0284	0.0333
1.9	Linear	N/A	1,000	0.0305	0.0333
1.10	Linear	N/A	1,000	0.0302	0.0500
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.03046** • Average Tests Error = **0.03166** • **Time: 189.429026 sec**					
2.1	kNN	1	1,000	0.0540	0.0333
2.2	kNN	2	1,000	0.0682	0.0500
2.3	kNN	3	1,000	0.0476	0.0167
2.4	kNN	4	1,000	0.0597	0.0167
2.5	kNN	5	1,000	0.0532	0.0167
2.6	kNN	6	1,000	0.0515	0.0500
2.7	kNN	7	1,000	0.0493	0.0500
2.8	kNN	8	1,000	0.0585	0.0167
2.9	kNN	9	1,000	0.0542	0.0500
2.10	kNN	10	1,000	0.0594	0.0333
2.11	kNN	11	1,000	0.0587	0.0167
2.19	kNN	19	1,000	0.0663	0.0667
2.25	kNN	25	1,000	0.0909	0.0500
——> The kNN model is optimum for the value **k = 3**					
2.3.1	kNN	3	1,000	0.0493	0.0167
2.3.2	kNN	3	1,000	0.0476	0.0167
2.3.3	kNN	3	1,000	0.0476	0.0167
2.3.4	kNN	3	1,000	0.0464	0.0167

2.3.5	kNN	3	1,000	0.0479	0.0167
2.3.6	kNN	3	1,000	0.0506	0.0167
2.3.7	kNN	3	1,000	0.0499	0.0333
2.3.8	kNN	3	1,000	0.0492	0.0500
2.3.9	kNN	3	1,000	0.0492	0.0167
2.3.10	kNN	3	1,000	0.0480	0.0167

| (**k=3**) | Average Validations Error = **0.04857** Average Tests Error = **0.02169**
Time: 49.352516 sec
—–> The kNN model is optimum for the value **k = 3**.

Scenario 3.4

Outer ratio: 90/60
Inner ratio: 80/10

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0308	0.0167
1.2	Linear	N/A	1,000	0.0276	0.0500
1.3	Linear	N/A	1,000	0.0302	0.0167
1.4	Linear	N/A	1,000	0.0301	0.0333
1.5	Linear	N/A	1,000	0.0291	0.0167
1.6	Linear	N/A	1,000	0.0307	0.0167
1.7	Linear	N/A	1,000	0.0305	0.0333
1.8	Linear	N/A	1,000	0.0288	0.0167
1.9	Linear	N/A	1,000	0.0301	0.0333
1.10	Linear	N/A	1,000	0.0296	0.0167

Total averages for test runs **1.1 to 1.10**:

- Average Validations Error = **0.02975**
- Average Tests Error = **0.02501**
- **Time: 179.249740 sec**

2.1	kNN	1	1,000	0.0472	0.0333
2.2	kNN	2	1,000	0.0564	0.0667

2.3	kNN	3	1,000	0.0462	0.0167
2.4	kNN	4	1,000	0.0500	0.0500
2.5	kNN	5	1,000	0.0482	0.0167
2.6	kNN	6	1,000	0.0461	0.0167
2.7	kNN	7	1,000	0.0500	0.0500
2.8	kNN	8	1,000	0.0547	0.0500
2.9	kNN	9	1,000	0.0516	0.0500
2.10	kNN	10	1,000	0.0573	0.0500
2.11	kNN	11	1,000	0.0570	0.0333
2.19	kNN	19	1,000	0.0551	0.0333
2.25	kNN	25	1,000	0.0746	0.0333
——> The kNN model is optimum for the value **k = 3 and k = 6**					
2.3.1	kNN	3	1,000	0.0468	0.0167
2.3.2	kNN	3	1,000	0.0465	0.0167
2.3.3	kNN	3	1,000	0.0460	0.0167
2.3.4	kNN	3	1,000	0.0506	0.0167
2.3.5	kNN	3	1,000	0.0508	0.0167
2.3.6	kNN	3	1,000	0.0451	0.0167
2.3.7	kNN	3	1,000	0.0425	0.0167
2.3.8	kNN	3	1,000	0.0465	0.0167
2.3.9	kNN	3	1,000	0.0466	0.0167
2.3.10	kNN	3	1,000	0.0476	0.0167
2.6.1	kNN	6	1,000	0.0448	0.0333
2.6.2	kNN	6	1,000	0.0481	0.0167
2.6.3	kNN	6	1,000	0.0458	0.0333
2.6.4	kNN	6	1,000	0.0429	0.0500
2.6.5	kNN	6	1,000	0.0495	0.0500
2.6.6	kNN	6	1,000	0.0522	0.0333
2.6.7	kNN	6	1,000	0.0475	0.0500
2.6.8	kNN	6	1,000	0.0499	0.0333
2.6.9	kNN	6	1,000	0.0476	0.0500

2.6.10	kNN	6	1,000	0.0477	0.0333

| (**k=3**) | Average Validations Error = **0.0469** Average Tests Error = **0.0167**
| (**k=6**) | Average Validations Error = 0.0476 Average Tests Error = 0.03832
Time: 58.006714 sec

——> The kNN model is optimum for the value **k = 3**.

Scenario 4.1

Outer ratio: 120/30
Inner ratio: 30/90

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0576	0.0333
1.2	Linear	N/A	1,000	0.0556	0.0000
1.3	Linear	N/A	1,000	0.0565	0.0000
1.4	Linear	N/A	1,000	0.0555	0.0000
1.5	Linear	N/A	1,000	0.0556	0.0000
1.6	Linear	N/A	1,000	0.0575	0.0333
1.7	Linear	N/A	1,000	0.0569	0.1000
1.8	Linear	N/A	1,000	0.0563	0.0000
1.9	Linear	N/A	1,000	0.0556	0.0000
1.10	Linear	N/A	1,000	0.0553	0.0000

Total averages for test runs **1.1 to 1.10**:

- Average Validations Error = **0.0563**
- Average Tests Error = **0.0167**
- **Time 167.119241 sec**

2.1	kNN	1	1,000	0.0717	N/A
2.2	kNN	2	1,000	0.0837	N/A
2.3	kNN	3	1,000	0.0719	N/A
2.4	kNN	4	1,000	0.0811	N/A
2.5	kNN	5	1,000	0.0776	N/A
2.6	kNN	6	1,000	0.0856	N/A

2.7	kNN	7	1,000	0.0893	N/A
2.8	kNN	8	1,000	0.1000	N/A
2.9	kNN	9	1,000	0.1076	N/A
2.10	kNN	10	1,000	0.1155	N/A
2.11	kNN	11	1,000	0.1280	N/A
2.19	kNN	19	1,000	0.3814	N/A
2.25	kNN	25	1,000	0.5185	N/A
——> The kNN model is optimum for the value **k = 1**					
2.1.1	kNN	1	1,000	0.0731	0.0000
2.1.2	kNN	1	1,000	0.0714	0.0000
2.1.3	kNN	1	1,000	0.0719	0.0333
2.1.4	kNN	1	1,000	0.0728	0.0000
2.1.5	kNN	1	1,000	0.0722	0.0333
2.1.6	kNN	1	1,000	0.0721	0.0000
2.1.7	kNN	1	1,000	0.0728	0.0000
2.1.8	kNN	1	1,000	0.0731	0.0000
2.1.9	kNN	1	1,000	0.0735	0.0667
2.1.10	kNN	1	1,000	0.0726	0.1000
| (**k=1**) | Average Validations Error = **0.0725** Average Tests Error = **0.0233** **Time: 48.562371 sec** ——> The kNN model is optimum for the value **k = 1**.					

Scenario 4.2 **Outer ratio: 120/30** **Inner ratio: 60/60**					
Test ID	**ML Classification Model**	**k**	**No of random "folds" (Bagging)**	**Average Validation Error**	**Average Test Error (on new data)**
1.1	Linear	N/A	1,000	0.0430	0.0000
1.2	Linear	N/A	1,000	0.0431	0.0000
1.3	Linear	N/A	1,000	0.0427	0.0000

1.4	Linear	N/A	1,000	0.0436	0.0000
1.5	Linear	N/A	1,000	0.0428	0.0000
1.6	Linear	N/A	1,000	0.0449	0.0000
1.7	Linear	N/A	1,000	0.0436	0.0000
1.8	Linear	N/A	1,000	0.0440	0.0000
1.9	Linear	N/A	1,000	0.0419	0.0000
1.10	Linear	N/A	1,000	0.0423	0.0000
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.0432** • Average Tests Error = **0.0** • **Time: 175.945882**					
2.1	kNN	1	1,000	0.0622	N/A
2.2	kNN	2	1,000	0.0740	N/A
2.3	kNN	3	1,000	0.0571	N/A
2.4	kNN	4	1,000	0.0611	N/A
2.5	kNN	5	1,000	0.0574	N/A
2.6	kNN	6	1,000	0.0634	N/A
2.7	kNN	7	1,000	0.0590	N/A
2.8	kNN	8	1,000	0.0659	N/A
2.9	kNN	9	1,000	0.0623	N/A
2.10	kNN	10	1,000	0.0678	N/A
2.11	kNN	11	1,000	0.0674	N/A
2.19	kNN	19	1,000	0.0810	N/A
2.25	kNN	25	1,000	0.1088	N/A
——> The kNN model is optimum for the value **k = 3**					
2.3.1	kNN	3	1,000	0.0572	0.0000
2.3.2	kNN	3	1,000	0.0572	0.0000
2.3.3	kNN	3	1,000	0.0574	0.0000
2.3.4	kNN	3	1,000	0.0574	0.0667
2.3.5	kNN	3	1,000	0.0578	0.0000
2.3.6	kNN	3	1,000	0.0570	0.0333

2.3.7	kNN	3	1,000	0.0574	0.0000
2.3.8	kNN	3	1,000	0.0568	0.0000
2.3.9	kNN	3	1,000	0.0569	0.0000
2.3.10	kNN	3	1,000	0.0565	0.0000

| (**k=3**) | Average Validations Error = **0.0572** Average Tests Error = **0.0100**
Time: 57.908479 sec

——> The kNN model is optimum for the value **k = 3**.

Scenario 4.3

Outer ratio: 120/30
Inner ratio: 90/30

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0399	0.0000
1.2	Linear	N/A	1,000	0.0430	0.0000
1.3	Linear	N/A	1,000	0.0412	0.0000
1.4	Linear	N/A	1,000	0.0407	0.0000
1.5	Linear	N/A	1,000	0.0402	0.0000
1.6	Linear	N/A	1,000	0.0418	0.0000
1.7	Linear	N/A	1,000	0.0419	0.0000
1.8	Linear	N/A	1,000	0.0395	0.0000
1.9	Linear	N/A	1,000	0.0398	0.0000
1.10	Linear	N/A	1,000	0.0398	0.0000

Total averages for test runs **1.1 to 1.10**:

- Average Validations Error = **0.0**
- Average Tests Error = **0.0408**
- **Time: 170.496823 sec**

2.1	kNN	1	1,000	0.0556	N/A
2.2	kNN	2	1,000	0.0717	N/A
2.3	kNN	3	1,000	0.0507	N/A
2.4	kNN	4	1,000	0.0554	N/A

2.5	kNN	5	1,000	0.0513	N/A
2.6	kNN	6	1,000	0.0556	N/A
2.7	kNN	7	1,000	0.0518	N/A
2.8	kNN	8	1,000	0.0593	N/A
2.9	kNN	9	1,000	0.0520	N/A
2.10	kNN	10	1,000	0.0555	N/A
2.11	kNN	11	1,000	0.0530	N/A
2.19	kNN	19	1,000	0.0665	N/A
2.25	kNN	25	1,000	0.0713	N/A
——> The kNN model is optimum for the value **k = 3**					
2.3.1	kNN	3	1,000	0.0505	0.0000
2.3.2	kNN	3	1,000	0.0494	0.0000
2.3.3	kNN	3	1,000	0.0501	0.0000
2.3.4	kNN	3	1,000	0.0497	0.0000
2.3.5	kNN	3	1,000	0.0515	0.0000
2.3.6	kNN	3	1,000	0.0496	0.0000
2.3.7	kNN	3	1,000	0.0498	0.0000
2.3.8	kNN	3	1,000	0.0504	0.0000
2.3.9	kNN	3	1,000	0.0519	0.0000
2.3.10	kNN	3	1,000	0.0516	0.0000

| (**k=3**) | Average Validations Error = **0.0504** Average Tests Error = **0.0000**
Time: 51.698454 sec
——> The kNN model is optimum for the value **k = 3**.

Scenario 4.4

Outer ratio: 120/30
Inner ratio: 110/10

Test ID	ML Classification Model	k	No of random "folds" (Bagging)	Average Validation Error	Average Test Error (on new data)
1.1	Linear	N/A	1,000	0.0443	0.0000
1.2	Linear	N/A	1,000	0.0468	0.0000
1.3	Linear	N/A	1,000	0.0432	0.0000
1.4	Linear	N/A	1,000	0.0451	0.0000
1.5	Linear	N/A	1,000	0.0404	0.0000
1.6	Linear	N/A	1,000	0.0425	0.0000
1.7	Linear	N/A	1,000	0.0397	0.0000
1.8	Linear	N/A	1,000	0.0442	0.0000
1.9	Linear	N/A	1,000	0.0428	0.0000
1.10	Linear	N/A	1,000	0.0419	0.0000
Total averages for test runs **1.1 to 1.10**: • Average Validations Error = **0.0431** • Average Tests Error = **0.0** • **Time: 181.836209 sec**					
2.1	kNN	1	1,000	0.0494	N/A
2.2	kNN	2	1,000	0.0763	N/A
2.3	kNN	3	1,000	0.0500	N/A
2.4	kNN	4	1,000	0.0476	N/A
2.5	kNN	5	1,000	0.0444	N/A
2.6	kNN	6	1,000	0.0545	N/A
2.7	kNN	7	1,000	0.0450	N/A
2.8	kNN	8	1,000	0.0548	N/A
2.9	kNN	9	1,000	0.0470	N/A
2.10	kNN	10	1,000	0.0589	N/A
2.11	kNN	11	1,000	0.0466	N/A
2.19	kNN	19	1,000	0.0617	N/A
2.25	kNN	25	1,000	0.0658	N/A
——> The kNN model is optimum for the value **k = 5**					
2.5.1	kNN	5	1,000	0.0421	0.0000

2.5.2	kNN	5	1,000	0.0481	0.0000
2.5.3	kNN	5	1,000	0.0437	0.0000
2.5.4	kNN	5	1,000	0.0434	0.0000
2.5.5	kNN	5	1,000	0.0426	0.0000
2.5.6	kNN	5	1,000	0.0459	0.0000
2.5.7	kNN	5	1,000	0.0434	0.0000
2.5.8	kNN	5	1,000	0.0448	0.0000
2.5.9	kNN	5	1,000	0.0450	0.0000
2.5.10	kNN	5	1,000	0.0440	0.0000

| (**k=5**) | Average Validations Error = **0.0443** Average Tests Error = **0.0000**
Time: 47.811813 sec

——> The kNN model is optimum for the value **k = 5**.