

Heb dataset

- KNN
 - euclidean function

	Test Accuracy	Train Accuracy
K = 1	0.833333	1.000000
K = 2	0.625000	0.892857
K = 3	0.916667	0.892857
K = 4	0.833333	0.875000
K = 5	0.916667	0.857143
K = 6	0.875000	0.875000
K = 7	0.958333	0.839286
K = 8	0.958333	0.857143
K = 9	0.958333	0.803571
K = 10	0.958333	0.821429
K = 11	0.958333	0.785714
K = 12	0.958333	0.803571
K = 13	0.958333	0.785714
K = 14	0.958333	0.785714

- manhattan function

	Test Accuracy	Train Accuracy
K = 1	0.833333	1.000000
K = 2	0.625000	0.892857
K = 3	0.916667	0.892857
K = 4	0.833333	0.875000
K = 5	0.916667	0.857143
K = 6	0.875000	0.875000
K = 7	0.958333	0.839286
K = 8	0.958333	0.857143
K = 9	0.958333	0.803571
K = 10	0.958333	0.821429
K = 11	0.958333	0.785714
K = 12	0.958333	0.803571
K = 13	0.958333	0.785714
K = 14	0.958333	0.785714

- DT
 - entropy function

	Test Accuracy	Train Accuracy
Max_depth = 1	0.875000	0.821429
Max_depth = 2	0.958333	0.928571
Max_depth = 3	0.958333	0.964286
Max_depth = 4	0.916667	1.000000
Max_depth = 5	0.916667	1.000000
Max_depth = 6	0.916667	1.000000
Max_depth = 7	0.916667	1.000000
Max_depth = 8	0.916667	1.000000
Max_depth = 9	0.916667	1.000000

- misclassification function

	Test Accuracy	Train Accuracy
Max_depth = 1	0.916667	0.857143
Max_depth = 2	0.958333	0.875000
Max_depth = 3	0.916667	0.892857
Max_depth = 4	0.916667	0.946429
Max_depth = 5	0.958333	0.982143
Max_depth = 6	0.958333	1.000000
Max_depth = 7	0.958333	1.000000
Max_depth = 8	0.958333	1.000000
Max_depth = 9	0.958333	1.000000

- gini index function

	Test Accuracy	Train Accuracy
Max_depth = 1	0.875000	0.857143
Max_depth = 2	0.875000	0.910714
Max_depth = 3	0.833333	1.000000
Max_depth = 4	0.833333	1.000000
Max_depth = 5	0.833333	1.000000
Max_depth = 6	0.833333	1.000000
Max_depth = 7	0.833333	1.000000
Max_depth = 8	0.833333	1.000000
Max_depth = 9	0.833333	1.000000

Dr dataset

- KNN
 - euclidean function

	Test Accuracy	Train Accuracy
K = 1	0.628986	1.000000
K = 2	0.594203	0.805211
K = 3	0.620290	0.827543
K = 4	0.626087	0.777916
K = 5	0.631884	0.764268
K = 6	0.657971	0.754342
K = 7	0.669565	0.743176
K = 8	0.657971	0.730769
K = 9	0.649275	0.719603
K = 10	0.643478	0.705955
K = 11	0.637681	0.702233
K = 12	0.649275	0.700993
K = 13	0.646377	0.704715
K = 14	0.655072	0.688586

- manhattan function

	Test Accuracy	Train Accuracy
K = 1	0.623188	1.000000
K = 2	0.628986	0.794045
K = 3	0.643478	0.827543
K = 4	0.646377	0.758065
K = 5	0.631884	0.785360
K = 6	0.666667	0.761787
K = 7	0.675362	0.758065
K = 8	0.689855	0.733251
K = 9	0.692754	0.727047
K = 10	0.678261	0.719603
K = 11	0.663768	0.719603
K = 12	0.666667	0.715881
K = 13	0.669565	0.715881
K = 14	0.672464	0.699752

- DT

- entropy function

	Test Accuracy	Train Accuracy
Max_depth = 20	0.652174	0.967742
Max_depth = 25	0.660870	0.995037
Max_depth = 30	0.657971	1.000000

- misclassification function

```

                Test Accuracy  Train Accuracy
Max_depth = 20      0.594203      0.933002
Max_depth = 25      0.585507      0.970223
Max_depth = 30      0.585507      0.998759
Using cost gini index
```

- gini index function

```

                Test Accuracy  Train Accuracy
Max_depth = 20      0.646377      0.996278
Max_depth = 25      0.646377      1.000000
Max_depth = 30      0.646377      1.000000
```

Sorted by accuracy diff

Heb

- KNN

Modeling Hepdata set:

Using euclidean function:

	Test Accuracy	Train Accuracy	Accuracy Diff
K = 8	0.833333	0.839286	0.005952
K = 6	0.791667	0.821429	0.029762
K = 5	0.916667	0.839286	0.077381
K = 3	0.791667	0.892857	0.101190
K = 4	0.750000	0.857143	0.107143
K = 7	0.958333	0.821429	0.136905
K = 9	0.958333	0.785714	0.172619
K = 11	0.958333	0.785714	0.172619
K = 13	0.958333	0.785714	0.172619
K = 10	0.958333	0.767857	0.190476
K = 14	0.958333	0.767857	0.190476
K = 1	0.791667	1.000000	0.208333
K = 12	0.958333	0.750000	0.208333
K = 2	0.666667	0.910714	0.244048

Using manhattan function:

	Test Accuracy	Train Accuracy	Accuracy Diff
K = 6	0.875000	0.875000	0.000000
K = 3	0.916667	0.892857	0.023810
K = 4	0.833333	0.875000	0.041667
K = 5	0.916667	0.857143	0.059524
K = 8	0.958333	0.857143	0.101190
K = 7	0.958333	0.839286	0.119048
K = 10	0.958333	0.821429	0.136905
K = 9	0.958333	0.803571	0.154762
K = 12	0.958333	0.803571	0.154762
K = 1	0.833333	1.000000	0.166667
K = 11	0.958333	0.785714	0.172619
K = 13	0.958333	0.785714	0.172619
K = 14	0.958333	0.785714	0.172619
K = 2	0.625000	0.892857	0.267857

- DT

Modeling Hepdata set:

Using cost entropy

	Test Accuracy	Train Accuracy	Accuracy Diff
Max_depth = 3	0.958333	0.964286	0.005952
Max_depth = 2	0.958333	0.928571	0.029762
Max_depth = 1	0.875000	0.821429	0.053571
Max_depth = 4	0.916667	1.000000	0.083333
Max_depth = 5	0.916667	1.000000	0.083333
Max_depth = 6	0.916667	1.000000	0.083333
Max_depth = 7	0.916667	1.000000	0.083333
Max_depth = 8	0.916667	1.000000	0.083333
Max_depth = 9	0.916667	1.000000	0.083333

Using cost misclassification

	Test Accuracy	Train Accuracy	Accuracy Diff
Max_depth = 3	0.916667	0.892857	0.023810
Max_depth = 5	0.958333	0.982143	0.023810
Max_depth = 4	0.916667	0.946429	0.029762
Max_depth = 6	0.958333	1.000000	0.041667
Max_depth = 7	0.958333	1.000000	0.041667
Max_depth = 8	0.958333	1.000000	0.041667
Max_depth = 9	0.958333	1.000000	0.041667
Max_depth = 1	0.916667	0.857143	0.059524
Max_depth = 2	0.958333	0.875000	0.083333

Using cost gini index

	Test Accuracy	Train Accuracy	Accuracy Diff
Max_depth = 1	0.875000	0.857143	0.017857
Max_depth = 2	0.875000	0.910714	0.035714
Max_depth = 3	0.833333	1.000000	0.166667
Max_depth = 4	0.833333	1.000000	0.166667
Max_depth = 5	0.833333	1.000000	0.166667
Max_depth = 6	0.833333	1.000000	0.166667
Max_depth = 7	0.833333	1.000000	0.166667
Max_depth = 8	0.833333	1.000000	0.166667
Max_depth = 9	0.833333	1.000000	0.166667

Dr dataset

- KNN

Modeling Drdata set:

Using euclidean function:

	Test Accuracy	Train Accuracy	Accuracy Diff
K = 14	0.655072	0.688586	0.033513
K = 12	0.649275	0.700993	0.051717
K = 13	0.646377	0.704715	0.058338
K = 10	0.643478	0.705955	0.062477
K = 11	0.637681	0.702233	0.064552
K = 9	0.649275	0.719603	0.070328
K = 8	0.657971	0.730769	0.072798
K = 7	0.669565	0.743176	0.073611
K = 6	0.657971	0.754342	0.096371
K = 5	0.631884	0.764268	0.132384
K = 4	0.626087	0.777916	0.151829
K = 3	0.620290	0.827543	0.207254
K = 2	0.594203	0.805211	0.211008
K = 1	0.628986	1.000000	0.371014

Using manhattan function:

	Test Accuracy	Train Accuracy	Accuracy Diff
K = 14	0.672464	0.699752	0.027288
K = 9	0.692754	0.727047	0.034294
K = 10	0.678261	0.719603	0.041342
K = 8	0.689855	0.733251	0.043396
K = 13	0.669565	0.715881	0.046316
K = 12	0.666667	0.715881	0.049214
K = 11	0.663768	0.719603	0.055835
K = 7	0.675362	0.758065	0.082702
K = 6	0.666667	0.761787	0.095120
K = 4	0.646377	0.758065	0.111688
K = 5	0.631884	0.785360	0.153476
K = 2	0.628986	0.794045	0.165059
K = 3	0.643478	0.827543	0.184065
K = 1	0.623188	1.000000	0.376812

- DT

Modeling Drdata set:

Using cost entropy

	Test Accuracy	Train Accuracy	Accuracy Diff
Max_depth = 20	0.652174	0.967742	0.315568
Max_depth = 25	0.660870	0.995037	0.334168
Max_depth = 30	0.657971	1.000000	0.342029

Using cost misclassification

	Test Accuracy	Train Accuracy	Accuracy Diff
Max_depth = 20	0.594203	0.933002	0.338800
Max_depth = 25	0.585507	0.970223	0.384716
Max_depth = 30	0.585507	0.998759	0.413252

Using cost gini index

	Test Accuracy	Train Accuracy	Accuracy Diff
Max_depth = 20	0.646377	0.996278	0.349901
Max_depth = 25	0.646377	1.000000	0.353623
Max_depth = 30	0.646377	1.000000	0.353623

