

HW2 – Machine Learning

Leonard Christopher Limanjaya

1a. Perform Shapiro-Wilk test on the differences of two datasets using R.

File: 1a.R

Result:

algo_a: w -> 0.88731, p-value -> 0.009812

algo_b: w -> 0.83833, p-value -> 0.001066

1b. Perform paired t test by hand

Data	Algorithm A	Algorithm B	diff	diff ²
anneal	93,99	76,17	17,82	317,5524
Balancing-scale	91,36	91,04	0,32	0,1024
Breast-cancer	71,68	72,38	-0,7	0,49
Breast-w	97,28	95,99	1,29	1,6641
Credit-a	81,94	84,49	1,45	2,1025
Dermatology	97,81	93,17	4,64	21,5296
Diabetes	77,47	69,92	7,55	57,0025
Glass	77,1	71,96	5,14	26,4196
Heart-c	85,81	84,49	1,32	1,7424
Heart-h	86,39	83,33	3,06	9,3636
Heart-hatlog	83,7	81,85	1,85	3,4225
Hepatitis	89,03	79,31	9,68	93,7024
Hypothyroid	91,68	92,29	3,39	11,4921
Ionosphere	92,02	90,31	1,71	2,9241
Kr-vi-kp	87,89	79,72	8,17	66,7489
Labor	91,23	91,23	0	0
Letter	72,84	77,71	-4,87	23,7169
Lymph	55,81	81,08	4,73	22,3729
Primary-tumor	50,15	29,15	20,65	426,4225
Sick	97,48	93,88	4,6	21,16
Sonar	99,04	88,94	10,1	102,01
Splice	95,36	82,04	43,32	1876,6224
Vehicle	66,67	69,62	-2,95	8,7025
Vote	90,11	90,17	-0,46	0,2116
Wormform-few	64,04	63,92	0,12	0,0144

141,93 7097,492

$\alpha = 0,05$

$\sum D = 141,93$

$\sum D^2 = 7097,492$

$$t = \frac{\sum D}{\sqrt{\frac{n(\sum D^2 - (\sum D)^2/n)}{n-1}}} = \frac{141,93}{\sqrt{\frac{25 \cdot 7097,492 - 141,93^2}{24}}}$$

$t_{stat} = 2,9$ $t_{crit} = 2,064$

$t_{stat} > t_{crit} \therefore \text{Reject } H_0$

$$= \frac{141,93}{\sqrt{\frac{2591,271}{24}}} = \frac{141,93}{10,277} = 2,9 //$$

The H_0 is Both of the algorithm is equal, but from the calculation we can know that t_{stat} is bigger than t_{crit} so we reject the H_0

1c. Perform paired t test using R

File: 1c.R

Result:

$t = 2.9049$, $df = 24$, $p\text{-value} = 0.007777$

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval: 1.643587, 9.710813

Mean of the differences: 5.6772

1d. Perform Wilcoxon Signed-Rank Test by hand

Data	Algo A	Algo B	diff	diff	Rank
animal	93,99	76,17	17,82	17,82	23
Bolton-scale	91,36	91,04	0,32	0,32	3
Breast-center	71,68	72,38	-0,7	0,7	5
Breast-w	97,28	95,99	1,29	1,29	6
Credit-Q	85,94	84,49	1,45	1,45	8
Dermatology	92,81	93,17	-0,36	0,36	15
Diabetes	77,47	69,92	7,55	7,55	19
Glass	77,1	71,86	5,24	5,24	18
Heart-c	85,81	84,49	1,32	1,32	7
Heart-h	86,39	83,33	3,06	3,06	12
Heart-hollog	83,7	81,85	1,85	1,85	10
Hepatitis	89,03	79,35	9,68	9,68	21
Hypothyroid	93,62	92,29	1,33	1,33	13
Ionosphere	92,02	90,31	1,71	1,71	9
Kr-vs-Kp	87,89	79,72	8,17	8,17	20
Lobar	91,23	91,23	0	0	1
Letter	72,84	77,71	-4,87	4,87	17
Lymph	55,81	81,08	-25,27	25,27	16
Primary-tumor	30,15	29,5	0,65	0,65	24
Sick	97,43	93,83	3,6	3,6	14
Solar	99,04	98,94	0,1	0,1	22
Spect	95,36	92,04	3,32	3,32	25
Vehicle	66,67	69,62	-2,95	2,95	11
Vote	90,11	90,17	-0,06	0,06	4
Worms-Proc	64,04	63,92	0,12	0,12	2

$$T^- = 5 + 17 + 11 + 4 = 37$$

$$T^+ = 1 + 2 + 3 + 6 + 7 + 8 + 9 + 10 + 12 + 13 + 14 + 15 + 16 + 18 + 19 + 20 + 21 + 22 + 23 + 24 + 25 = 288$$

$$W_{stat} = 37$$

$$W_{stat} < W_{crit} \Rightarrow \text{Reject the } H_0$$

$$W_{crit} = 89$$

The H_0 is both of the algorithm is equal, but from the calculation we can know that w_{stat} is smaller than w_{crit} so we reject the H_0

1e. Perform Wilcoxon Signed-Rank Test using R

File: 1e.R

Result:

$V = 267$, $p\text{-value} = 0.0008729$

Alternative hypothesis: true location shift is not equal to 0

Because of the $p\text{-value}$ is below the crit value, we can say that H_0 is rejected

1f. Perform Sign Test (Binomial test) using R

File: 1f.R

Result:

$S=20$, $p\text{-value} = 0.001544$

95 percent confidence interval: 1.29, 5.1

Median of x : 3.06

Achieved and Interpolated Confidence Intervals:

	Conf.Level	L.E.pt	U.E.pt
Lower Achieved CI	0.8922	1.3200	4.7300
Interpolated CI	0.9500	1.2931	5.0973
Upper Achieved CI	0.9567	1.2900	5.1400

2a. Perform Friedman Test by hand

Data	Algo A	Algo B	Algo C	Algo D	R _A	R _B	R _C	R _D
anneal	93,99	76,17	97,77	97,44	2	1	4	3
Balancing scale	91,36	91,04	89,6	89,6	4	3	1,5	1,5
Breast-cancer	71,68	72,38	71,68	72,03	1,5	4	1,5	3
Breast-w	97,28	95,99	97	97	2	1	2,5	2,5
Credit-Q	85,94	84,48	87,54	87,1	2	1	4	3
Dermatology	97,81	93,17	98,09	97,81	2,5	1	4	2,5
Diabetes	77,47	69,92	76,04	76,04	4	3	1,5	1,5
Glass	77,1	71,96	78,97	78,97	2	1	3,5	3,5
Heart-c	85,81	84,49	84,82	85,48	4	1	2	3
Heart-h	86,39	83,33	86,07	85,37	3	1	2	4
Heart-holog	83,7	81,85	84,07	83,7	2,5	1	4	2,5
Hepatitis	89,03	79,35	89,68	89,68	2	1	3,5	3,5
Hypothyroid	95,68	92,29	95,26	95,29	2	1	3	4
Ionosphere	92,02	90,71	93,45	93,45	2	1	3,5	3,5
Kr-vs-hp	87,89	79,72	91,24	91,27	2	1	3	4
Libras	91,23	91,23	82,46	85,96	1,5	1,5	3	4
Letter	72,84	77,71	81,2	81,2	1	2	3,5	3,5
Lymph	87,81	81,08	87,21	86,49	1	2	3	4
Primary-tumor	80,15	29,05	49,16	80,15	3,5	1	2	3,5
Sick	97,48	93,88	97,51	97,38	3	1	4	2
Solar	99,04	88,94	99,04	99,04	3	1	3	3
Spectra	95,36	52,04	96,21	95,57	2	1	4	3
Vehicles	66,67	69,62	73,12	73,52	1	2	3,5	3,5
Vow	90,11	90,17	94,25	94,25	1	2	3,5	3,5
Waveform-1000	64,04	63,92	64,02	64,02	4	1	2,5	2,5

60,5 36,5 69,5 77,5
77,5

$$Fr = \frac{12}{25 \cdot 4 \cdot 4} (60,5^2 + 36,5^2 + 69,5^2 + 77,5^2) - 3 \cdot 25 \cdot 5$$

$$= 500,97 - 375 = 125,97$$

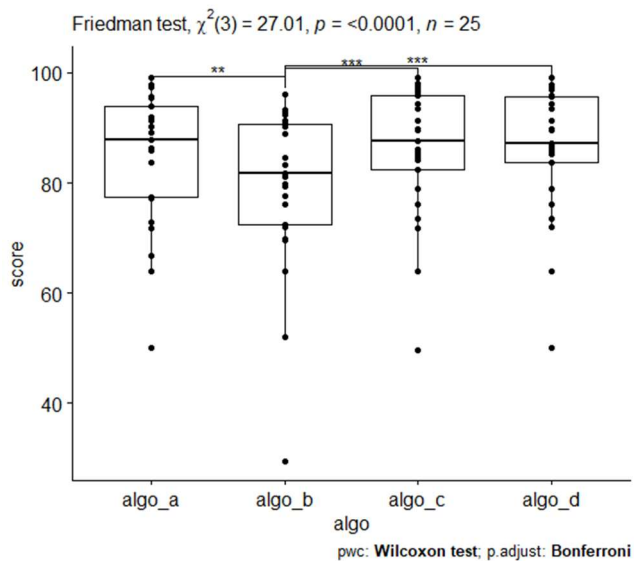
$$Fr > X_{0,05}^2 \Rightarrow \text{Reject } H_0$$

$$X^2 \quad 37,6525$$

We have H_0 that say that all of them are equal, but from the calculation we can know that Fr is higher than X^2 , so the H_0 is rejected

2b. Perform Friedman Test using R

File: 2b.R



From the result we can know that the H_0 is rejected, because of p value and X^2 value

2c. If needed, perform Nemenyi post-hoc test by hand.

$$CD = \frac{q \sqrt{h(h+1)}}{\sqrt{6N}} = \frac{2.569 \sqrt{4(4+1)}}{\sqrt{6 \cdot 25}} = \frac{2.569 \cdot 0.361}{1} = 0.938$$

avg of	R_a	R_b	R_c	R_d
	2.42	1.46	2.15	3.1

the R_b have more than CD from the others
so the H_0 is rejected

2d. If needed, perform Nemenyi post-hoc test using R. Draw Friedman test graph using R

File: 2d.R

```

      algo_a algo_b algo_c
algo_b 0.00675 -      -
algo_c 0.69233 8.9e-05 -
algo_d 0.72549 0.00011 0.99994

```

3. Run the Friedman test for the following result.

File: 3.R

