Johns Hopkins Engineering

Applied Machine Learning for Mechanical Engineers

Multi-Paradigm Machine Learning Models, Part 1, B

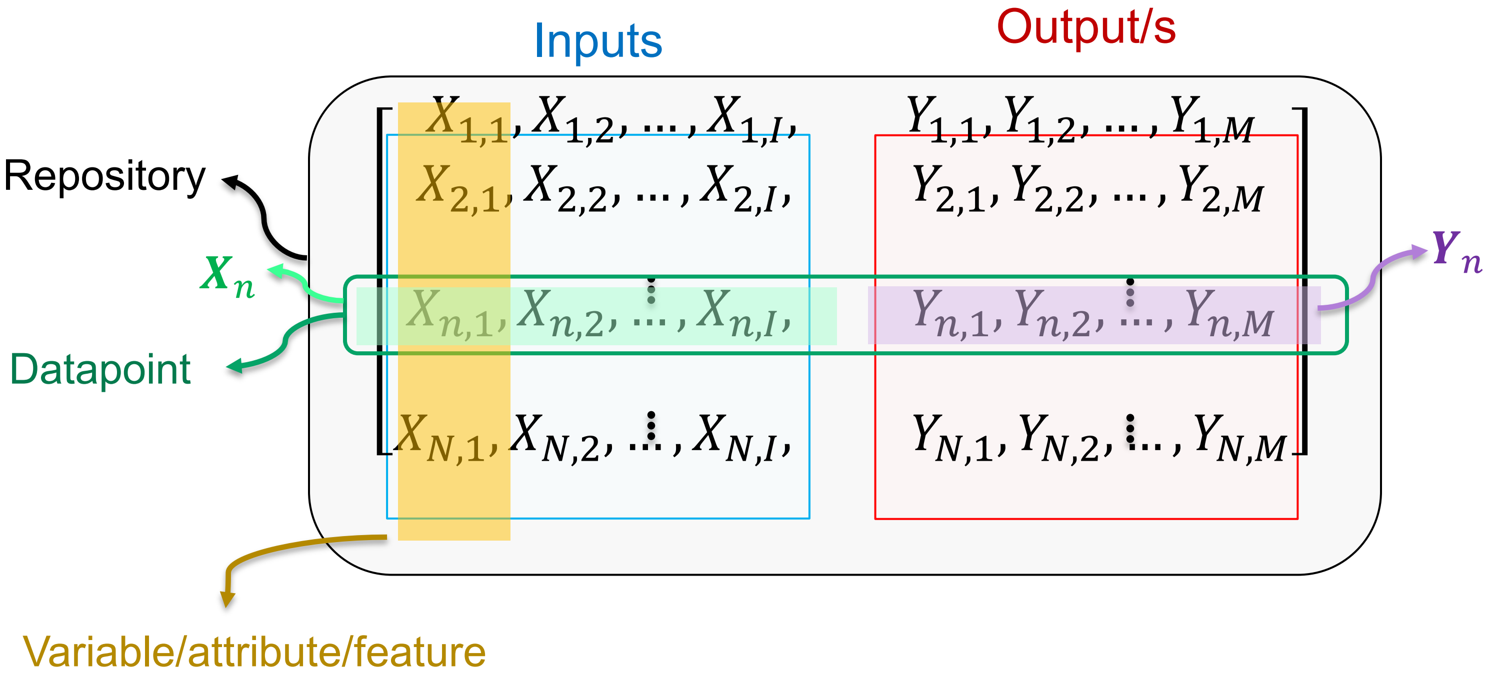
JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Combinatory Pattern Recognition

- By the end of this lecture you will be able to:
  - Describe combinatory pattern recognition computational models

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models



Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models
  - Are all the attributes needed for an accurate estimation?
  - Is there extraneous attributes?
  - How to find the best combination of attributes to maximize our estimation accuracy?

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models
  - Solution 1: investigate all possible combinations of attributes and check which one yields the highest accuracy.
  - Solution 2: define an optimization model to find an optimum combination that maximizes the accuracy of estimation

Pick solution 1 or 2 based on available resources

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models
  - The total number of exclusive combinations of I attributes is $2^I - 1$
  - We can show if an attribute is selected in a combination using logical expressions or binary configurations

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models
  - Example: $I = 3$, totally $2^{I=3} - 1 = 7$ combinations

1st Combination:	1	0	0
2nd Combination:	0	1	0
3rd Combination:	0	0	1
4th Combination:	1	1	0
5th Combination:	1	0	1
6th Combination:	0	1	1
7th Combination:	1	1	1

1 for select
0 for no select

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models

  - Example: $I = 3$, totally $2^{I=3} - 1 = 7$ combinations

1st Combination:	1	0	0	→	Score 01 (e.g., accuracy or MSE)
2nd Combination:	0	1	0	→	Score 02 (e.g., accuracy or MSE)
3rd Combination:	0	0	1	→	Score 03 (e.g., accuracy or MSE)
4th Combination:	1	1	0	→	Score 04 (e.g., accuracy or MSE)
5th Combination:	1	0	1	→	Score 05 (e.g., accuracy or MSE)
6th Combination:	0	1	1	→	Score 06 (e.g., accuracy or MSE)
7th Combination:	1	1	1	→	Score 07 (e.g., accuracy or MSE)

The one with the best score

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models
  - Example: sort combinations based on their scores (high to low)
    - Your may compute selection rates

	Combinations			Selection rates		
5th Combination:	1	0	1	1.00	0.00	1.00
7nd Combination:	1	1	0	1.00	0.50	0.50
4rd Combination:	1	1	1	1.00	0.67	0.67
1th Combination:	1	0	0	1.00	0.50	0.50
6th Combination:	0	1	1	0.80	0.6	0.6
3th Combination:	0	0	1	0.67	0.5	0.67
2th Combination:	0	1	0	0.57	0.57	0.57

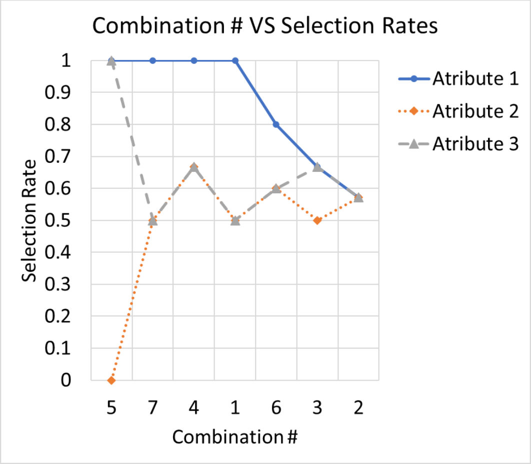
$S_{i=3,c=2}$

$S_{i=1,c=6}$

$B_{i=2,j=3}$

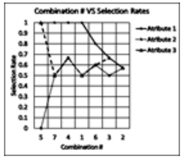
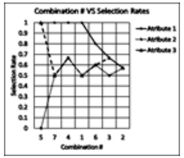
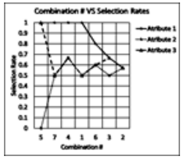
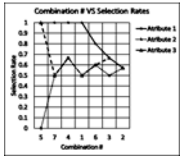
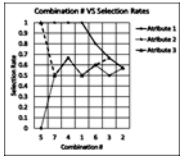
$$S_{i,c} = \frac{\sum_{j=1}^{j=c} B_{i,j}}{c}$$

$B_{i=1,j=5}$



Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models
  - Example: for each machine learning technique, use multiple RTTs and multiple RRS for each combination

RTT				
0.1	0.2	0.3	0.4	0.5
				

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models

  - Example: $I\ = 25$, totally $2^{I=25} - 1 = 33{,}554{,}431$ combinations.
    - You may want to go over all combinations using supercomputers, distributed computer nodes, and/or parallel computing techniques
    - You might want to define and optimization technique to find the optimum or near optimum combination at a lower computation cost.

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models
  - Example: $I\ = 100$, totally $2^{I=100} - 1 = 1{,}267{,}650{,}600{,}228{,}229{,}401{,}496{,}703{,}205{,}375$ combinations. an astronomical number of combinations!
    - You cannot go over all combinations even by using all available supercomputers
    - You might want to define and optimization technique to find some near optimum combination at a lower computation cost but practically investigating an extremely small portion the feasible domain
    - One solution is to reduce the number of attributes by identifying the ones that, perhaps, have been emphasized in the previous literature

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models

Example: George, et al. (2017a):

Table 5
Different combinations of 18 inputs for EPNN resulting in an average accuracy of 100% (1: selected; 0: not-selected).

C#	Inputs																		LR
	1, SA	2, BT	3, CN	4, EE	5, EW	6, FC	7, FB	8, FT	9, HB	10, HT	11, KY	12, LC	13, PL	14, SC	15, RR	16, TL	17, BK	18, TM	
1	0	0	0	0	1	1	1	0	0	0	0	1	1	0	1	1	1	0	0.000
2	0	0	1	1	1	0	1	0	1	0	0	1	1	0	1	1	0	0	0.118
3	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	0	1	1	0.174
4	1	0	1	0	1	0	0	1	0	1	0	0	1	0	1	0	0	1	0.178
5	0	0	1	1	1	0	1	0	1	1	0	1	1	0	1	1	0	0	0.218
6	0	0	0	0	1	0	1	1	0	0	1	1	1	0	0	1	1	0	0.228
7	0	0	0	0	1	0	1	1	0	0	0	1	1	0	0	1	1	0	0.233
8	0	0	0	0	1	0	1	1	0	0	0	1	1	0	1	1	0	0	0.238
9	0	0	0	0	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0.252
10	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	1	1	1	0.256
11	0	0	1	0	1	0	1	1	0	0	0	1	1	0	0	1	1	0	0.263
12	0	0	0	0	1	0	1	0	1	0	0	1	1	0	1	1	0	0	0.276
13	0	1	0	0	0	0	1	0	0	0	0	1	1	0	1	1	1	1	0.29
14	1	0	1	0	1	0	1	1	0	0	1	0	1	0	0	0	0	0	0.293
15	1	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	0	1	0.302
16	0	0	0	0	1	0	1	1	1	0	1	1	1	0	0	1	1	0	0.314
17	0	0	0	0	1	0	0	1	1	1	0	1	1	0	0	1	1	0	0.334
18	0	0	0	0	1	0	1	1	1	0	0	1	1	0	0	1	0	0	0.339
19	0	0	1	0	1	1	1	1	0	1	1	0	1	0	0	1	0	0	0.342
20	1	0	0	0	0	1	0	1	1	1	0	0	1	0	1	1	1	1	0.344
21	1	0	0	0	0	0	0	1	1	0	0	0	1	0	1	1	1	1	0.346
22	0	0	0	0	1	0	1	1	0	1	1	1	1	0	0	1	1	0	0.351
23	1	1	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	1	0.352
24	0	0	1	0	1	0	1	1	1	0	0	1	1	0	0	1	1	0	0.356
25	0	0	0	0	1	0	1	1	0	0	0	1	1	0	1	1	1	0	0.358
26	0	0	0	0	1	0	1	0	0	1	0	1	1	0	1	1	1	0	0.358
27	0	0	0	0	1	0	1	1	0	1	0	1	1	0	0	1	1	0	0.358
28	0	0	0	0	1	0	0	1	1	0	0	1	1	0	1	1	1	0	0.36
29	0	0	0	0	1	0	1	1	0	1	0	1	1	0	1	1	0	0	0.36
30	0	0	0	0	1	0	1	1	1	0	0	1	1	0	0	1	1	0	0.361
31	0	0	1	0	1	0	1	1	0	1	0	1	1	0	0	1	1	0	0.362
32	0	1	0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1	0.384
33	0	0	0	0	1	0	1	0	1	1	0	1	1	0	1	1	0	0	0.399
34	0	0	0	0	0	0	0	0	1	1	0	1	0	1	1	1	1	1	0.404
35	0	0	0	0	0	0	0	1	1	0	0	1	0	1	1	1	1	1	0.407
36	0	1	0	0	0	0	1	1	0	0	0	1	1	0	1	1	1	1	0.436
37	0	0	0	0	0	0	1	0	0	1	1	1	0	1	1	1	1	1	0.442
38	0	0	1	0	1	1	1	1	0	0	1	0	1	0	0	1	0	0	0.456
39	0	0	0	1	0	1	1	0	1	0	0	1	1	0	1	1	1	0	0.469
40	0	1	0	0	0	0	0	0	1	1	0	1	0	0	1	1	1	1	0.47
41	0	0	0	0	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0.47
42	0	0	0	0	1	0	1	1	0	1	0	1	1	0	1	1	1	0	0.479
43	0	0	0	1	0	0	1	0	1	0	1	1	1	0	1	1	1	0	0.484
44	1	1	0	0	0	0	0	1	1	1	0	0	0	0	1	1	0	1	0.49
45	0	1	0	0	0	0	1	1	0	1	0	1	1	0	1	1	1	1	0.515
46	0	1	0	0	0	0	1	0	1	1	0	0	0	0	1	1	1	1	0.553
47	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	1	1	1	0.559
48	0	0	0	1	0	1	1	0	1	1	0	1	1	0	1	1	1	0	0.568
49	0	0	0	1	0	1	1	0	1	0	1	0	1	1	1	1	1	0	0.623
50	0	0	0	1	0	0	1	0	1	1	1	1	1	0	1	1	1	0	0.634
51	0	0	0	1	0	1	1	0	1	1	1	0	1	1	1	1	1	0	0.736
52	1	1	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	1	1.000
TS	8	10	9	8	28	9	38	30	30	27	12	39	41	6	34	49	37	19	N/A

EPNN: enhanced probabilistic neural networks; C#: combination number; SA: stroke affected side; BT: WMFT basket; CN: WMFT can; EE: WFMT extend elbow; EW: WMFT extend elbow weight; FC: WMFT flip cards; FB: WMFT forearm to box; FT: WMFT forearm to table; HB: WMFT hand to box; HT: WMFT hand to table; KY: WMFT key; LC: WMFT lift paper clip; PL: WMFT pencil; SC: WMFT stack checkers; RR: WMFT reach retrieve; TL: WMFT towel; BK: BKT; WMFT: Wolf motor function test; TM: touch monofilament; OT: output; TS: times selected; LR: Likelihood ratio.

George, S. H., Rafiei, M. H., Gauthier, L., Borstad, A., Buford, J. A., & Adeli, H. (2017a). Computer-aided prediction of extent of motor recovery following constraint-induced movement therapy in chronic stroke. Behavioural Brain Research, 329, 191-199.

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models

Example: George, et al. (2017b):

George, S. H., Rafiei, M. H., Borstad, A., Adeli, H., & Gauthier, L. V. (2017b). Gross motor ability predicts response to upper extremity rehabilitation in chronic stroke. *Behavioural brain research*, *333*, 314-322.

Table 6
Rates of selection of 18 predictors in the prognosis model for the gaming and CI therapies, and the combined approach

	Predictors																			
R#	1. SA	2. BT	3. CN	4. EE	5. EW	6. FC	7. FB	8. FT	9. HB	10. HT	11. KY	12. LC	13. PL	14. SC	15. RR	16. TL	17. BK	18. TM	#C	AC
Gaming																				
1	75.0	75.0	12.5	0.0	75.0	0.0	25.0	100.0	37.5	75.0	0.0	0.0	0.0	0.0	25.0	25.0	12.5	37.5	8	94.7
2	58.8	71.3	12.5	45.0	46.3	28.8	35.0	78.8	36.3	48.8	0.0	23.8	3.8	0.0	22.5	67.5	46.3	57.5	80	89.5
A	66.9	73.1	12.5	22.5	60.6	14.4	30.0	89.4	36.9	61.9	0.0	11.9	1.9	0.0	23.8	46.3	29.4	47.5	N/A	N/A
CI therapy																				
1	15.4	19.2	17.3	15.4	53.8	17.3	73.1	57.7	57.7	51.9	23.1	75.0	78.8	11.5	65.4	94.2	71.2	36.5	52	100
2	29.1	56.6	44.2	30.1	34.9	29.9	56.1	37.5	67.9	52.6	43.3	55.4	67.6	26.6	76.2	82.7	61.2	33.9	1014	97.1
3	27.7	60.0	49.8	32.7	36.0	31.0	53.5	45.7	61.2	50.7	43.2	56.6	62.0	30.7	66.9	75.3	62.0	35.4	4828	94.3
4	30.4	56.8	54.0	39.3	40.7	37.8	50.7	48.0	54.7	49.8	43.7	56.3	55.3	32.8	60.9	69.5	58.9	35.3	15052	91.4
A	25.6	48.2	41.3	29.3	41.4	29.0	58.3	47.2	60.4	51.3	38.3	60.8	65.9	25.4	67.3	80.4	63.3	35.3	N/A	N/A
Combined																				
1	100	100	0	0.0	100	0.0	100	100	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	1	94.5
2	100	100	0.	0.0	50.0	50.0	100	100	0.0	50.0	0.0	0.0	0.0	0.0	0.0	50.0	0.0	100	1	93.3
3	100	100	0.0	0.0	66.7	33.3	66.7	100	33.3	66.7	0.0	0.0	0.0	0.0	0.0	33.3	0.0	100	1	93.1
4	66.7	66.7	0.0	0.0	33.3	33.3	50.0	66.7	66.7	66.7	0.0	33.3	16.7	0.0	33.3	66.7	50.0	100	3	92.1
5	60.0	50.0	0.0	10.0	30.0	40.0	50.0	50.0	60.0	40.0	0.0	30.0	20.0	0.0	20.0	80.0	50.0	100	4	91.9
6	66.7	58.3	0.0	8.3	41.7	33.3	41.7	58.3	58.3	50.0	0.0	25.0	16.7	0.0	16.7	66.7	41.7	91.7	2	91.7
7	59.7	29.0	6.5	9.7	37.1	35.5	61.3	51.6	54.8	41.9	0.0	54.8	35.5	0.0	48.4	83.9	62.9	74.2	50	90.7
8	59.2	33.8	5.6	9.9	38.0	33.8	60.6	53.5	54.9	42.3	0.0	53.5	33.8	0.0	43.7	84.5	64.8	73.2	9	90.5
9	58.3	33.3	6.9	9.7	37.5	33.3	59.7	54.2	55.6	41.7	0.0	52.8	33.3	0.0	44.4	84.7	63.9	72.2	1	90.2
A	74.5	63.5	2.1	5.3	48.3	32.5	65.5	70.5	42.6	55.5	0.0	27.7	17.3	0.0	22.9	61.1	37.0	90.1	N/A	N/A

EPNN: Enhanced probabilistic neural networks; R#: Row number; SA: Stroke affected side; BT: WMFT basket; CN: WMFT can; EE: WFMT extend elbow; EW: WMFT extend elbow weight; FC: WMFT flip cards; FB: WMFT forearm to box; FT: WMFT forearm to table; HB: WMFT hand to box; HT: WMFT hand to table; KY: WMFT key; LC: WMFT lift paper clip; PL: WMFT pencil; SC: WMFT stack checkers; RR: WMFT reach retrieve; TL: WMFT towel; BK: BKT; WMFT: Wolf motor function test; TM: Touch monofilament; OT: Output; TS: Times selected; AC: Accuracy percentage; #C: Number of combinations associated with AC; A: Average.

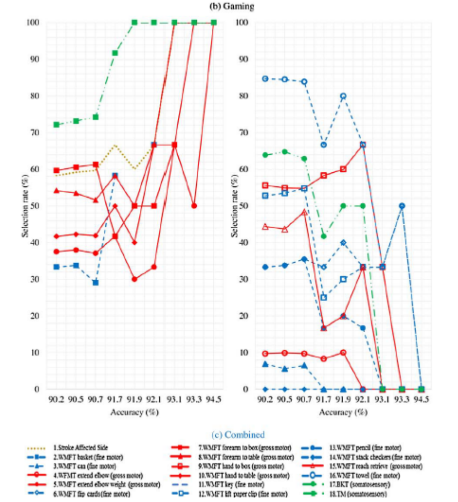


Fig. 4. a) The selection rates of predictors that were not selected in the best combination in Table 6 versus accuracy including the combinations with at least that accuracy or more and b) The selection rates of predictors that were selected in the best combination in Table 6 versus accuracy including the combinations with at least that accuracy or more.

Combinatory Pattern Recognition

- Combinatory Pattern Recognition Computational Models

Example: Rafiei and Adeli (2017):

Table 3
The best combination of seismicity indicators (step 6, Fig. 1) corresponding to various magnitude thresholds-lag combinations using NDC (1: selected, 0: not selected).

$N_{lag}-G_{threshold}$ Combination Number T_{char}	N_{lag} C	$G_{threshold}$	Seismicity indicators ΔT	G_{mean}		$dE^{1/2}$	b --	va-lue	D	ΔG	Accuracy of NDC (%)
1	1	4.5	0	0	0	1	0	1	1	1	100.0
2	1	5.0	1	1	1	1	1	0	1	1	100.0
3	1	5.5	1	1	1	1	1	1	1	1	100.0
4	1	6.0	1	1	1	1	1	1	1	1	100.0
5	1	6.5	1	1	1	1	1	1	1	1	100.0
6	6	4.5	0	0	1	1	0	1	1	1	97.3
7	6	5.0	0	1	0	0	0	1	0	1	99.3
8	6	5.5	1	0	1	1	1	1	0	1	100.0
9	6	6.0	1	1	1	1	1	1	1	1	99.3
10	6	6.5	1	1	1	1	1	1	0	1	99.3
11	7	4.5	0	0	0	0	0	0	1	1	95.3
12	7	5.0	0	0	0	0	0	1	1	1	100.0
13	7	5.5	1	0	1	1	0	1	1	1	100.0
14	7	6.0	1	1	1	1	1	1	1	1	100.0
15	7	6.5	1	1	1	0	1	1	0	1	99.3
16	8	4.5	0	0	1	1	1	0	0	0	99.3
17	8	5.0	1	0	0	0	1	1	1	1	100.0
18	8	5.5	0	0	0	1	1	0	0	1	100.0
19	8	6.0	1	1	1	1	1	1	1	1	99.3
20	8	6.5	1	1	1	1	1	1	1	0	100.0
21	9	4.5	0	0	0	0	0	0	1	1	96.7
22	9	5.0	0	1	1	1	0	1	1	1	99.3
23	9	5.5	1	0	0	1	1	0	1	1	100.0
24	9	6.0	1	1	1	1	1	1	0	1	100.0
25	9	6.5	0	0	1	1	0	1	1	1	100.0
26	10	4.5	1	0	0	0	0	0	0	1	99.3
27	10	5.0	0	1	0	1	1	1	1	1	98.7
28	10	5.5	0	0	1	1	1	0	1	0	100.0
29	10	6.0	0	1	1	1	1	0	1	1	100.0
30	10	6.5	0	1	1	1	1	1	1	1	100.0
31	11	4.5	1	0	1	0	1	0	0	0	97.3
32	11	5.0	0	0	1	0	1	1	1	1	98.7
33	11	5.5	0	1	1	0	1	1	1	1	98.7
34	11	6.0	0	1	0	0	1	1	1	1	100.0
35	11	6.5	1	1	1	1	1	1	1	1	99.3
36	12	4.5	1	0	1	1	0	0	0	0	100.0
37	12	5.0	0	1	0	1	1	0	1	0	99.3
38	12	5.5	1	1	1	1	0	0	0	1	99.3
39	12	6.0	1	1	1	1	1	0	1	1	99.3
40	12	6.5	1	1	1	1	1	1	1	1	100.0

$G_{threshold}$: Earthquake magnitude threshold; N_{lag}: Lag number; ΔT: Time elapsed (days); G_{mean}: The mean magnitude; $dE^{1/2}$: the rate of the square root of seismic energy; b-value: The slope of the Gutenberg-Richter inverse power law; D: Summation of the mean squared deviation from the regression line based on the Gutenberg-Richter inverse power law; ΔG: The difference between the observed maximum magnitude among the last N events and the largest expected; T_{char}: The mean time between characteristic events; C: the coefficient of variation of the mean time between characteristic events; NDC: Neural dynamic classification

Rafiei, M. H., & Adeli, H. (2017). NEEWS: A novel earthquake early warning model using neural dynamic classification and neural dynamic optimization. *Soil Dynamics and Earthquake Engineering*, *100*, 417-427.

Combinatory Pattern Recognition

- In this lecture, you learned about:
  - Combinatory pattern recognition computational models
- In the next lecture, we will talk about multi-paradigm smart systems

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING