

Selected Techniques in Machine Learning

Goals

- Overview
 - Know where to look
 - Terminology
 - How to ask/answer the "right" questions
 - The value of data
 - Opportunities (in astronomy)
 - Challenges
-
- How to answer questions with data?
 - How to use data to come up with good questions?

Historical Context

- Data science is **NOT NEW!**
- Data-intensive/driven science:
 - Science that we do based on data, e.g., astronomy
- Data science:
 - Emergent discipline motivated by data-intensive/driven science.
 - The science of how to: collect, handle, clean, store, visualize, and model with data.

Tycho Brahe

- 1546–1601
- Collected rich data set of motion of planets
- Noticed inconsistencies of model at the time with what he observed
- Realized as a teenager to collect data to make a science of data
- Had a "science island" between Copenhagen and Sweden
- Knew the value of his data, didn't let other people (Kepler) have full access!



Johannes Kepler

- 1571–1630
- An assistant to the Tycho Brahe
- Discovered three (geometrical) laws of planetary motion:
 - Planets move in elliptical orbits with the Sun as a focus
 - A planet covers the same area of space in the same amount of time no matter where it is in its orbit
 - A planet's orbital period is proportional to the size of its orbit



Pars Tertia.

89

TABULARUM

RUDOLPHI

ASTRONOMI-
CARUM

PARS TERTIA,

DE ECLIPSIBUS SOLIS ET LUNÆ, ALIISQUE
PLANETARUM CONGRUENTIIS ET CON-
figurationibus.

Typus Anni Numeri, neque Politicus, neque Ecclesiasticus usualis, sed more Astronomicus, serviens indagandis Mensibus Eclipticis in Methodo Anni Juliani.													Periodos Cy- clorum magna.	
Numerus Anni.	Janu- arii	Febru- arii	Mar- tii	Aprili- is	Maii	Iunii	Iulii	Augu- sti	Septem- bris	Octo- bris	Novem- bris	Decem- bris	Anni	Hec.
III	1	31	1	30	29	28	27	26	25	24	23	22	76	5.10
XI	2	1	2	31	30	29	28	27	26	25	24	23	152	11.40
XIX	3	2	3	1	30	29	28	27	26	25	24	23	228	17.11
VIII	4	3	4	2	29	28	27	26	25	24	23	22	304	22.81
XVI	5	4	5	3	28	27	26	25	24	23	22	21	380	28.52
V	6	5	6	4	27	26	25	24	23	22	21	20	456	34.23
XIII	7	6	7	5	26	25	24	23	22	21	20	19	532	40.00
II	8	7	8	6	25	24	23	22	21	20	19	18	608	45.75
X	9	8	9	7	24	23	22	21	20	19	18	17	684	51.50
XVIII	10	9	10	8	23	22	21	20	19	18	17	16	760	57.25
VII	11	10	11	9	22	21	20	19	18	17	16	15	836	63.00
XV	12	11	12	10	21	20	19	18	17	16	15	14	912	68.75
IV	13	12	13	11	20	19	18	17	16	15	14	13	988	74.50
XII	14	13	14	12	19	18	17	16	15	14	13	12	1064	80.25
IX	15	14	15	13	18	17	16	15	14	13	12	11	1140	86.00
XVII	16	15	16	14	17	16	15	14	13	12	11	10	1216	91.75
VI	17	16	17	15	16	15	14	13	12	11	10	9	1292	97.50
XIV	18	17	18	16	15	14	13	12	11	10	9	8	1368	103.25
III	19	18	19	17	16	15	14	13	12	11	10	9	1444	109.00
XII	20	19	20	18	17	16	15	14	13	12	11	10	1520	114.75
I	21	20	21	19	18	17	16	15	14	13	12	11	1596	120.50
IX	22	21	22	20	19	18	17	16	15	14	13	12	1672	126.25
XVII	23	22	23	21	20	19	18	17	16	15	14	13	1748	132.00
VI	24	23	24	22	21	20	19	18	17	16	15	14	1824	137.75
XIV	25	24	25	23	22	21	20	19	18	17	16	15	1900	143.50
III	26	25	26	24	23	22	21	20	19	18	17	16	1976	149.25
XII	27	26	27	25	24	23	22	21	20	19	18	17	2052	155.00
I	28	27	28	26	25	24	23	22	21	20	19	18	2128	160.75
IX	29	28	29	27	26	25	24	23	22	21	20	19	2204	166.50
XVII	30	29	30	28	27	26	25	24	23	22	21	20	2280	172.25
VI	31	30	31	29	28	27	26	25	24	23	22	21	2356	178.00
XIV	32	31	32	30	29	28	27	26	25	24	23	22	2432	183.75
III	33	32	33	31	30	29	28	27	26	25	24	23	2508	189.50
XII	34	33	34	32	31	30	29	28	27	26	25	24	2584	195.25
I	35	34	35	33	32	31	30	29	28	27	26	25	2660	201.00
IX	36	35	36	34	33	32	31	30	29	28	27	26	2736	206.75
XVII	37	36	37	35	34	33	32	31	30	29	28	27	2812	212.50
VI	38	37	38	36	35	34	33	32	31	30	29	28	2888	218.25
XIV	39	38	39	37	36	35	34	33	32	31	30	29	2964	224.00
III	40	39	40	38	37	36	35	34	33	32	31	30	3040	229.75
XII	41	40	41	39	38	37	36	35	34	33	32	31	3116	235.50
I	42	41	42	40	39	38	37	36	35	34	33	32	3192	241.25
IX	43	42	43	41	40	39	38	37	36	35	34	33	3268	247.00
XVII	44	43	44	42	41	40	39	38	37	36	35	34	3344	252.75
VI	45	44	45	43	42	41	40	39	38	37	36	35	3420	258.50
XIV	46	45	46	44	43	42	41	40	39	38	37	36	3496	264.25
III	47	46	47	45	44	43	42	41	40	39	38	37	3572	270.00
XII	48	47	48	46	45	44	43	42	41	40	39	38	3648	275.75
I	49	48	49	47	46	45	44	43	42	41	40	39	3724	281.50
IX	50	49	50	48	47	46	45	44	43	42	41	40	3800	287.25
XVII	51	50	51	49	48	47	46	45	44	43	42	41	3876	293.00
VI	52	51	52	50	49	48	47	46	45	44	43	42	3952	298.75
XIV	53	52	53	51	50	49	48	47	46	45	44	43	4028	304.50
III	54	53	54	52	51	50	49	48	47	46	45	44	4104	310.25
XII	55	54	55	53	52	51	50	49	48	47	46	45	4180	316.00
I	56	55	56	54	53	52	51	50	49	48	47	46	4256	321.75
IX	57	56	57	55	54	53	52	51	50	49	48	47	4332	327.50
XVII	58	57	58	56	55	54	53	52	51	50	49	48	4408	333.25
VI	59	58	59	57	56	55	54	53	52	51	50	49	4484	339.00
XIV	60	59	60	58	57	56	55	54	53	52	51	50	4560	344.75
III	61	60	61	59	58	57	56	55	54	53	52	51	4636	350.50
XII	62	61	62	60	59	58	57	56	55	54	53	52	4712	356.25
I	63	62	63	61	60	59	58	57	56	55	54	53	4788	362.00
IX	64	63	64	62	61	60	59	58	57	56	55	54	4864	367.75
XVII	65	64	65	63	62	61	60	59	58	57	56	55	4940	373.50
VI	66	65	66	64	63	62	61	60	59	58	57	56	5016	379.25
XIV	67	66	67	65	64	63	62	61	60	59	58	57	5092	385.00
III	68	67	68	66	65	64	63	62	61	60	59	58	5168	390.75
XII	69	68	69	67	66	65	64	63	62	61	60	59	5244	396.50
I	70	69	70	68	67	66	65	64	63	62	61	60	5320	402.25
IX	71	70	71	69	68	67	66	65	64	63	62	61	5396	408.00
XVII	72	71	72	70	69	68	67	66	65	64	63	62	5472	413.75
VI	73	72	73	71	70	69	68	67	66	65	64	63	5548	419.50
XIV	74	73	74	72	71	70	69	68	67	66	65	64	5624	425.25
III	75	74	75	73	72	71	70	69	68	67	66	65	5700	431.00
XII	76	75	76	74	73	72	71	70	69	68	67	66	5776	436.75
I	77	76	77	75	74	73	72	71	70	69	68	67	5852	442.50
IX	78	77	78	76	75	74	73	72	71	70	69	68	5928	448.25
XVII	79	78	79	77	76	75	74	73	72	71	70	69	6004	454.00
VI	80	79	80	78	77	76	75	74	73	72	71	70	6080	459.75
XIV	81	80	81	79	78	77	76	75	74	73	72	71	6156	465.50
III	82	81	82	80	79	78	77	76	75	74	73	72	6232	471.25
XII	83	82	83	81	80	79	78	77	76	75	74	73	6308	477.00
I	84	83	84	82	81	80	79	78	77	76	75	74	6384	482.75
IX	85	84	85	83	82	81	80	79	78	77	76	75	6460	488.50
XVII	86	85	86	84	83	82	81	80	79	78	77	76	6536	494.25
VI	87	86	87	85	84	83	82	81	80	79	78	77	6612	500.00
XIV	88	87	88	86	85	84	83	82	81	80	79	78	6688	505.75
III	89	88	89	87	86	85	84	83	82	81	80	79	6764	511.50
XII	90	89	90	88	87	86	85	84	83	82	81	80	6840	517.25
I	91	90	91	89	88	87	86	85	84	83	82	81	6916	523.00
IX	92	91	92	90	89	88	87	86	85	84	83	82	6992	528.75
XVII	93	92	93	91	90	89	88	87	86	85	84	83	7068	534.50
VI	94	93	94	92	91	90	89	88	87	86	85	84	7144	540.25
XIV	95	94	95	93	92	91	90	89	88	87	86	85	7220	546.00
III	96	95	96	94	93	92	91	90	89	88	87	86	7296	551.75
XII	97	96	97	95	94	93	92	91	90	89	88	87	7372	557.50
I	98	97	98	96	95	94	93	92	91	90	89	88	7448	563.25
IX	99	98	99	97	96	95	94	93	92	91	90	89	7524	569.00
XVII	100	99	100	98	97	96	95	94	93	92	91	90	7600	574.75

* Est fides Bifexti, qui tamen more Romano non Augu-
stus, sed bis 24. nunquam 29. pronuntiatur.

M

Cyclo. Obv

86

Tabularum Rudolphi

Tabula Latitudinis LVNÆ simplicis, una cum Reductione loci Orbitæ ad
Eclipticam, quæ valent, Nodo ♄ in Quadris existente.

Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "
0	0. 0. 0	0. 0. 0	10	2.29.52	5.40. 150	20	4.59.49	8.00. 100	30	7.28.13	10.00. 50
1	0. 5.14	0.14. 179	11	2.34.23	5.47. 149	21	4.54.43	7.55. 99	31	7.23.13	9.55. 49
2	0.10.28	0.28. 178	12	2.38.50	5.54. 148	22	4.49.43	7.50. 98	32	7.18.13	9.50. 48
3	0.15.42	0.42. 177	13	2.43.15	6. 0. 147	23	4.44.43	7.45. 97	33	7.13.13	9.45. 47
4	0.20.55	0.55. 176	14	2.47.37	6. 6. 146	24	4.39.43	7.40. 96	34	7.08.13	9.40. 46
5	0.26. 8	1. 9. 175	15	2.51.56	6.10. 145	25	4.34.43	7.35. 95	35	7.03.13	9.35. 45
6	0.31.20	1.22. 174	16	2.56.12	6.14. 144	26	4.29.43	7.30. 94	36	6.58.13	9.30. 44
7	0.36.32	1.36. 173	17	3. 0.25	6.18. 143	27	4.24.43	7.25. 93	37	6.53.13	9.25. 43
8	0.41.43	1.49. 172	18	3. 4.34	6.21. 142	28	4.19.43	7.20. 92	38	6.48.13	9.20. 42
9	0.46.53	2. 2. 171	19	3. 8.40	6.24. 141	29	4.14.43	7.15. 91	39	6.43.13	9.15. 41
10	0.52. 2	2.14. 170	20	3.12.43	6.26. 140	30	4.09.43	7.10. 90	40	6.38.13	9.10. 40
11	0.57.10	2.27. 169	21	3.16.43	6.29. 139	31	4.04.43	7.05. 89	41	6.33.13	9.05. 39
12	1. 2.15	2.40. 168	22	3.20.37	6.30. 138	32	3.59.43	7.00. 88	42	6.28.13	9.00. 38
13	1. 7.24	2.52. 167	23	3.24.28	6.32. 137	33	3.54.43	6.55. 87	43	6.23.13	8.55. 37
14	1.12.29	3. 5. 166	24	3.28.16	6.33. 136	34	3.49.43	6.50. 86	44	6.18.13	8.50. 36
15	1.17.33	3.17. 165	25	3.32. 0	6.33. 135	35	3.44.43	6.45. 85	45	6.13.13	8.45. 35
16	1.22.36	3.28. 164	26	3.35.46	6.33. 134	36	3.39.43	6.40. 84	46	6.08.13	8.40. 34
17	1.27.37	3.40. 163	27	3.39.16	6.32. 133	37	3.34.43	6.35. 83	47	6.03.13	8.35. 33
18	1.32.36	3.51. 162	28	3.42.43	6.30. 132	38	3.29.43	6.30. 82	48	5.58.13	8.30. 32
19	1.37.34	4. 2. 161	29	3.46.17	6.29. 131	39	3.24.43	6.25. 81	49	5.53.13	8.25. 31
20	1.42.30	4.12. 160	30	3.49.41	6.26. 130	40	3.19.43	6.20. 80	50	5.48.13	8.20. 30
21	1.47.24	4.23. 159	31	3.53. 2	6.22. 129	41	3.14.43	6.15. 79	51	5.43.13	8.15. 29
22	1.52.16	4.33. 158	32	3.56.18	6.21. 128	42	3.09.43	6.10. 78	52	5.38.13	8.10. 28
23	1.57. 6	4.42. 157	33	3.59.30	6.18. 127	43	3.04.43	6.05. 77	53	5.33.13	8.05. 27
24	2. 0. 0	4.52. 156	34	4. 2.37	6.14. 126	44	2.59.43	6.00. 76	54	5.28.13	8.00. 26
25	2. 6.40	5. 1. 155	35	4. 5.40	6.10. 125	45	2.54.43	5.55. 75	55	5.23.13	7.55. 25
26	2.11.23	5.10. 154	36	4. 8.28	6. 5. 124	46	2.49.43	5.50. 74	56	5.18.13	7.50. 24
27	2.16. 4	5.17. 153	37	4.11.33	6. 0. 123	47	2.44.43	5.45. 73	57	5.13.13	7.45. 23
28	2.20.42	5.26. 152	38	4.14.21	5.54. 122	48	2.39.43	5.40. 72	58	5.08.13	7.40. 22
29	2.25.18	5.34. 151	39	4.17. 5	5.47. 121	49	2.34.43	5.35. 71	59	5.03.13	7.35. 21
30	2.29.52	5.40. 150	40	4.19.44	5.40. 120	50	2.29.43	5.30. 70	60	4.58.13	7.30. 20

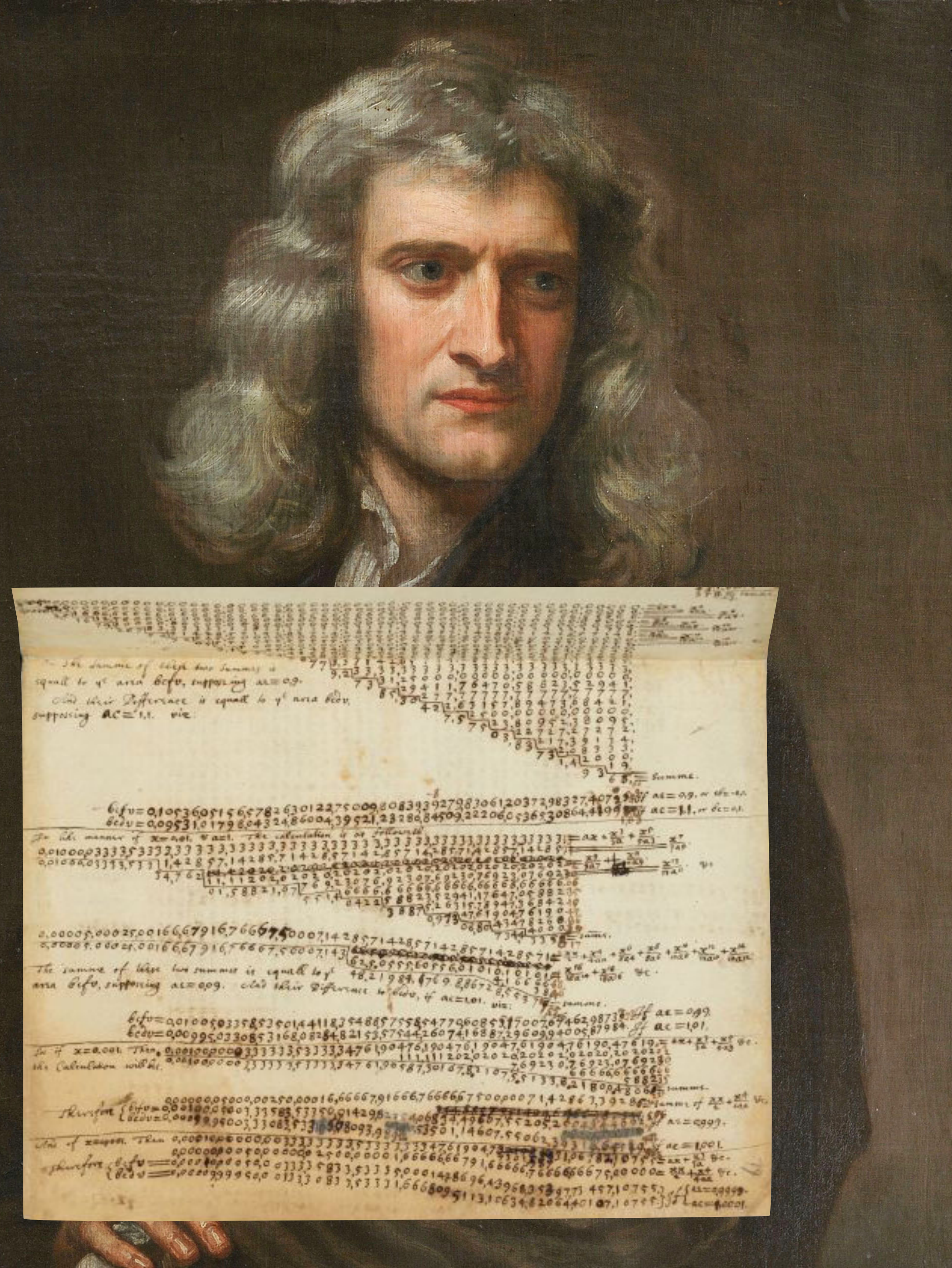
Tabula exhibens portionem ipsam Latitudinis Menstruam.

Septentrionalis.												Meridiana.											
Distantia SOLIS a Nodo Lunæ Ascendente ♄.												Distantia SOLIS a Nodo Lunæ Descendente ♄.											
Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "						
0	0. 0. 0	0. 0. 0	10	2.29.52	5.40. 150	20	4.59.49	8.00. 100	30	7.28.13	10.00. 50	40	9.56.43	12.00. 10	50	12.21.13	14.00. 0						
1	0. 5.14	0.14. 179	11	2.34.23	5.47. 149	21	4.54.43	7.55. 99	31	7.23.13	9.55. 49	41	9.51.43	11.55. 0	51	12.16.13	13.55. 0						
2	0.10.28	0.28. 178	12	2.38.50	5.54. 148	22	4.49.43	7.50. 98	32	7.18.13	9.50. 48	42	9.46.43	11.50. 0	52	12.09.13	13.50. 0						
3	0.15.42	0.42. 177	13	2.43.15	6. 0. 147	23	4.44.43	7.45. 97	33	7.13.13	9.45. 47	43	9.41.43	11.45. 0	53	12.02.13	13.45. 0						
4	0.20.55	0.55. 176	14	2.47.37	6. 6. 146	24	4.39.43	7.40. 96	34	7.08.13	9.40. 46	44	9.36.43	11.40. 0	54	11.55.13	13.40. 0						
5	0.26. 8	1. 9. 175	15	2.51.56	6.10. 145	25	4.34.43	7.35. 95	35	7.03.13	9.35. 45	45	9.31.43	11.35. 0	55	11.48.13	13.35. 0						
6	0.31.20	1.22. 174	16	2.56.12	6.14. 144	26	4.29.43	7.30. 94	36	6.58.13	9.30. 44	46	9.26.43	11.30. 0	56	11.41.13	13.30. 0						
7	0.36.32	1.36. 173	17	3. 0.25	6.18. 143	27	4.24.43	7.25. 93	37	6.53.13	9.25. 43	47	9.21.43	11.25. 0	57	11.34.13	13.25. 0						
8	0.41.43	1.49. 172	18	3. 4.34	6.21. 142	28	4.19.43	7.20. 92	38	6.48.13	9.20. 42	48	9.16.43	11.20. 0	58	11.27.13	13.20. 0						
9	0.46.53	2. 2. 171	19	3. 8.40	6.24. 141	29	4.14.43	7.15. 91	39	6.43.13	9.15. 41	49	9.11.43	11.15. 0	59	11.20.13	13.15. 0						
10	0.52. 2	2.14. 170	20	3.12.43	6.26. 140	30	4.09.43	7.10. 90	40	6.38.13	9.10. 40	50	9.06.43	11.10. 0	60	11.13.13	13.10. 0						
11	0.57.10	2.27. 169	21	3.16.43	6.29. 139	31	4.04.43	7.05. 89	41	6.33.13	9.05. 39	51	9.01.43	11.05. 0	61	11.06.13	13.05. 0						
12	1. 2.15	2.40. 168	22	3.20.37	6.30. 138	32	3.59.43	7.00. 88	42	6.28.13	9.00. 38	52	8.56.43	11.00. 0	62	10.59.13	13.00. 0						
13	1. 7.24	2.52. 167	23	3.24.28	6.32. 137	33	3.54.43	6.55. 87	43	6.23.13	8.55. 37	53	8.51.43	10.55. 0	63	10.52.13	12.55. 0						
14	1.12.29	3. 5. 166	24	3.28.16	6.33. 136	34	3.49.43	6.50. 86	44	6.18.13	8.50. 36	54	8.46.43	10.50. 0	64	10.45.13	12.50. 0						
15	1.17.33	3.17. 165	25	3.32. 0	6.33. 135	35	3.44.43	6.45. 85	45	6.13.13	8.45. 35	55	8.41.43	10.45. 0	65	10.38.13	12.45. 0						
16	1.22.36	3.28. 164	26	3.35.46	6.33. 134	36	3.39.43	6.40. 84	46	6.08.13	8.40. 34	56	8.36.43	10.40. 0	66	10.31.13	12.40. 0						
17	1.27.37	3.40. 163	27	3.39.16	6.32. 133	37	3.34.43	6.35. 83	47	6.03.13	8.35. 33	57	8.31.43	10.35. 0	67	10.24.13	12.35. 0						
18	1.32.36	3.51. 162	28	3.42.43	6.30. 132	38	3.29.43	6.30. 82	48	5.58.13	8.30. 32	58	8.26.43	10.30. 0	68	10.17.13	12.30. 0						
19	1.37.34	4. 2. 161	29	3.46.17	6.29. 131	39	3.24.43	6.25. 81	49	5.53.13	8.25. 31	59	8.21.43	10.25. 0	69	10.10.13	12.25. 0						
20	1.42.30	4.12. 160	30	3.49.41	6.26. 130	40	3.19.43	6.20. 80	50	5.48.13	8.20. 30	60	8.16.43	10.20. 0	70	10.03.13	12.20. 0						
21	1.47.24	4.23. 159	31	3.53. 2	6.22. 129	41	3.14.43	6.15. 79	51	5.43.13	8.15. 29	61	8.11.43	10.15. 0	71	9.56.13	12.15. 0						
22	1.52.16	4.33. 158	32	3.56.18	6.21. 128	42	3.09.43	6.10. 78	52	5.38.13	8.10. 28	62	8.06.43	10.10. 0	72	9.49.13	12.10. 0						
23	1.57. 6	4.42. 157	33	3.59.30	6.18. 127	43	3.04.43	6.05. 77	53	5.33.13	8.05. 27	63	8.01.43	10.05. 0	73	9.42.13	12.05. 0						
24	2. 0. 0	4.52. 156	34	4. 2.37	6.14. 126	44	2.59.43	6.00. 76	54	5.28.13	8.00. 26	64	7.56.43	10.00. 0	74	9.35.13	12.00. 0						
25	2. 6.40	5. 1. 155	35	4. 5.40	6.10. 125	45	2.54.43	5.55. 75	55	5.23.13	7.55. 25	65	7.51.43	9.55. 0	75	9.28.13	11.55. 0						
26	2.11.23	5.10. 154	36	4. 8.28	6. 5. 124	46	2.49.43	5.50. 74	56	5.18.13	7.50. 24	66	7.46.43	9.50. 0	76	9.21.13	11.50. 0						
27	2.16. 4	5.17. 153	37	4.11.33	6. 0. 123	47	2.44.43	5.45. 73	57	5.13.13	7.45. 23	67	7.41.43	9.45. 0	77	9.14.13	11.45. 0						
28	2.20.42	5.26. 152	38	4.14.21	5.54. 122	48	2.39.43	5.40. 72	58	5.08.13	7.40. 22	68	7.36.43	9.40. 0	78	9.07.13	11.40. 0						
29	2.25.18	5.34. 151	39	4.17. 5	5.47. 121	49	2.34.43	5.35. 71	59	5.03.13	7.35. 21	69	7.31.43	9.35. 0	79	8.59.13	11.35. 0						
30	2.29.52	5.40. 150	40	4.19.44	5.40. 120	50	2.29.43	5.30. 70	60	4.58.13	7.30. 20	70	7.26.43	9.30. 0	80	8.52.13	11.30. 0						

a Lunæ a SOLE.												a Lunæ a SOLE.											
Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "	Gr.	Latitudo Gr. ° ' "	Reductio Gr. ° ' "						
0	0. 0. 0	0. 0. 0	10	2.29.52	5.40. 150	20	4.59.49	8.00. 100	30	7.28.13	10.00. 50	40	9.56.43	12.00. 10	50	12.21.13	14.00. 0						
1	0. 5.14	0.14. 179	11	2.34.23	5.47. 149	21	4.54.43	7.55. 99	31	7.23.13	9.55. 49	41	9.51.43	11.55. 0	51	12.16.13	13.55. 0						
2	0.10.28	0.28. 178	12	2.38.50	5.54. 148	22	4.49.43	7.50. 98	32	7.18.13	9.50. 48	42	9.46.43	11.50. 0	52	12.09.13	13.50. 0						
3	0.15.42	0.42. 177	13	2.43.15	6. 0. 147	23	4.44.43	7.45. 97	33	7.13.13	9.45. 47	43	9.41.43	11.45. 0	53	12.02.13	13.45. 0						
4	0.20.55	0.55. 176	14	2.47.37	6. 6. 146	24	4.39.43	7.40. 96	34	7.08.13	9.40. 46	44	9.36.43	11.40. 0	54	11.55.13	13.40. 0</						

Sir Isaac Newton

- 1642–1726
- Formulated the laws of motion and universal gravitation
- Derived Kepler's laws of planetary motion
- Explained "why"
- Generalizable; enabled us to put people on the moon
- "Higher order" theory
- In data-driven science, in astrophysics, maybe an ultimate goal is push **from Kepler to Newton**





The
F O U R T H
P A R A D I G M

DATA-INTENSIVE SCIENTIFIC DISCOVERY

EDITED BY TONY HEY, STEWART TANSLEY, AND KRISTIN TOLLE

Experiments/
Observations

Data Science

Science

Theory

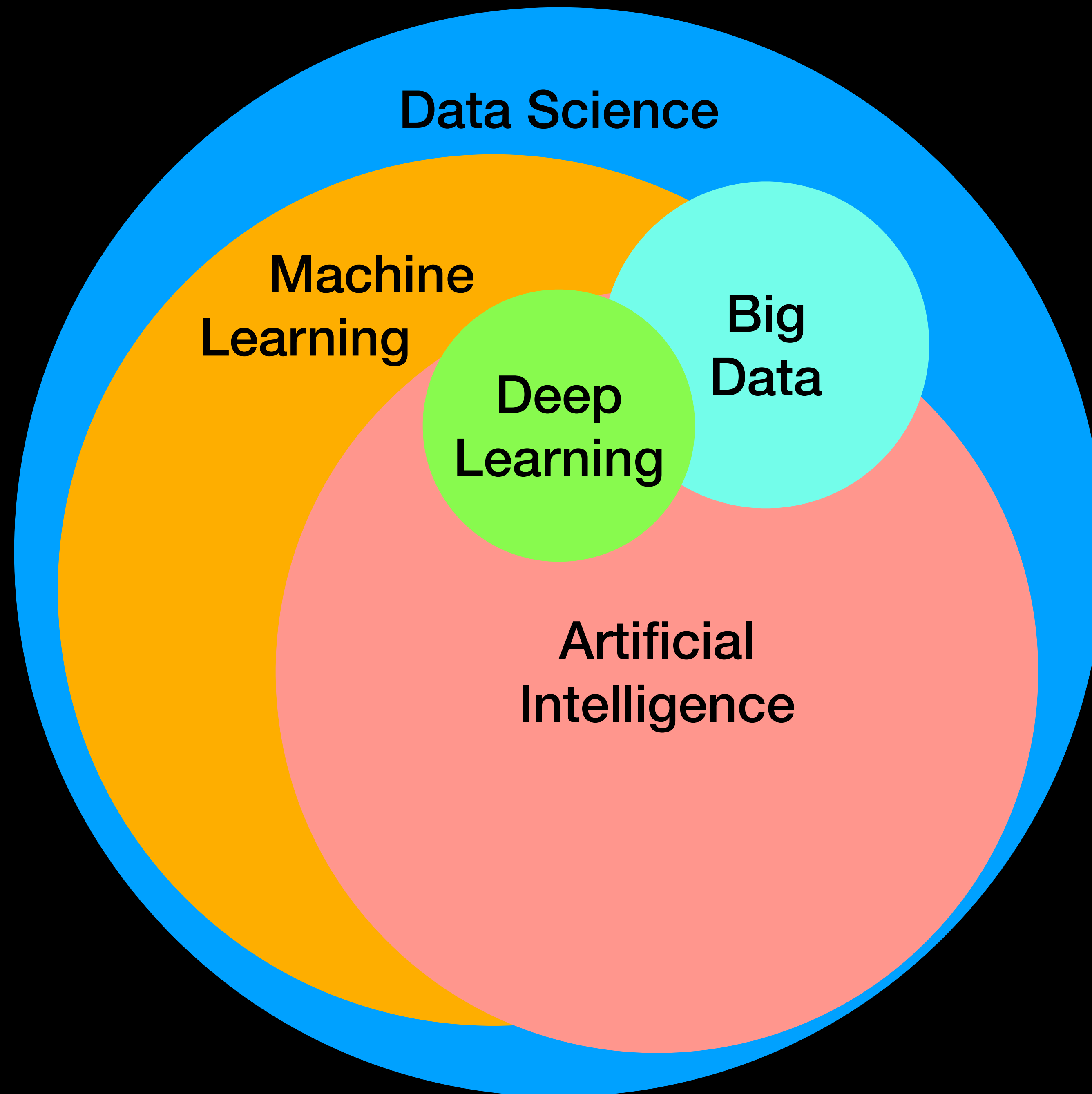
Computational
Science

Computational Science

- “Traditional” simulation
- Forward problem:
 - model $\rightarrow$ solution
- Equations is (relatively) simple
- Algorithms are well understood, often designed to preserve equation properties
- Complexity is in the many DoF and non-linearity
- Codes are usually developed by a small team in University or national labs

Data Science

- “Cool” machine learning
- Inverse problem:
 - data $\rightarrow$ model
- Data is usually messy
- The most powerful algorithms are less well understood, as long as “they work”
- Complexity is in the coefficients learned from the data
- Codes are developed by large communities, many support by big tech (Tensorflow, PyTorch, JAX)



Advancement of Machine Learning

- Some of the most exciting applications of deep learning
 - [DQN Breakout](#)
 - [AlphaGo](#) (core algorithm [mctx](#), [documentary](#))
 - [AlphaZero](#)
 - [AlphaFold3](#)
 - [DALL-E3](#)
 - [ChatGPT](#)
 - [Segmenting Anything 2](#)
 - [AlphaGeometry](#)
 - [GraphCast](#)
- What will the “ChatGPT moment” be like in astronomy?



Machine Learning

- Building useful models from data
- Models hopefully improve with increase volume of data
- Misconcept 1: ML is magic
- Reality: built on foundational advances in math and computer science in decades:
 - optimization & regression
- Most useful background fields: linear algebra, optimization, statistics
- Misconcept 2: ML is just Neural Network (Deep Learning)
- Reality: outside NN/DL, there are many ML techniques that are:
 - effective, efficient, and very often understandable/interpretable

Machine Learning Goals

- Model be "useful" in the future
- Accurate
- Fast
- Generalizable ($f = ma$)
- Interpretable/explainable
- Certify/guarantee

Cross-Validation

- Past: training data
- Future: testing data
- Split available to 80% for training and 20% for "testing"
- K-fold validation for increase statistics
- **Never trust/publish any ML results without cross-validation!!!**

Supervised

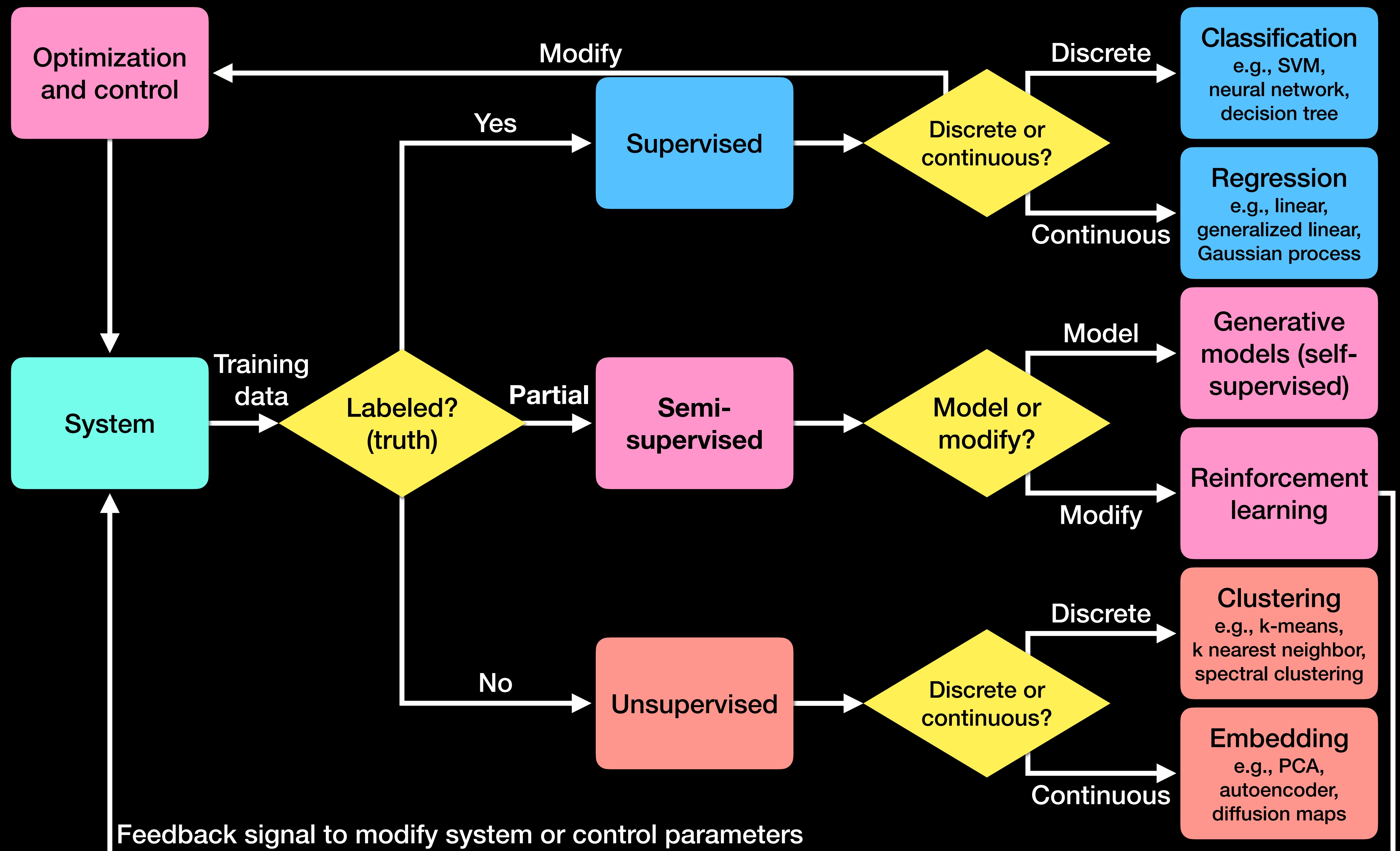
- With labels
- Classification (discrete):
classify type I and II supernovae
 - SVM
 - neural network
 - decision tree
- Regression (continue): light curve of AGNs
 - linear,
 - generalized linear,
 - Gaussian process
- Optimization and control

Semi-supervised

- Some feedbacks
- Generative models (self-supervised)
 - Generative adversarial networks (GAN)
- Reinforcement learning
 - Deep reinforcement learning

Unsupervised

- No labels, "data mining"
- Cluster (discrete)
 - k-means
 - k nearest neighbor
 - spectral clustering
- Embedding, dimensionality reduction, pattern extraction
 - PCA
 - autoencoder
 - diffusion maps



Resources

- Scikit-learn: <https://scikit-learn.org>
- Keras: <https://keras.io>
- AstroML: <https://www.astroml.org>