Machine Learning Basics

Nicole Schanche

March 14, 2018

Cake & Computing

Goals

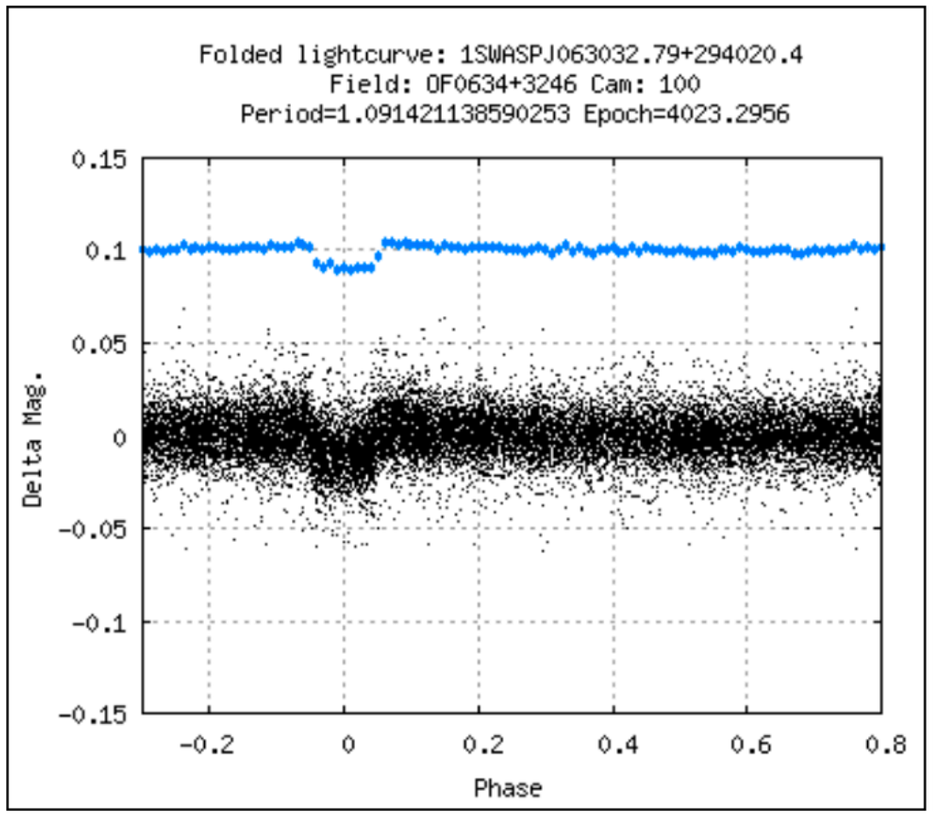
- find planets
- reduce follow-up time (ie reduce false-positives)
- quickly and consistently look through large database

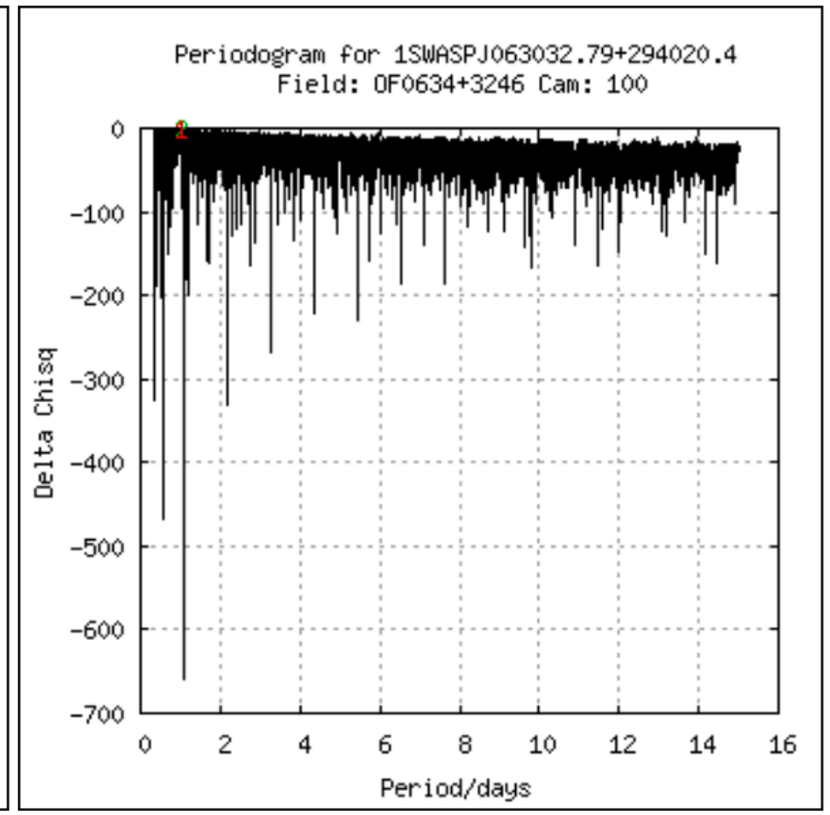
  - CLASSIFICATION PROBLEM

WASP-12

i	Teff JH	Teff VK	R_{star} JH	R_{star} VK	V mag	J mag	V-K	J-H	MU_RA (mas/yr)	MU_DEC (mas/yr)	RPMJ	RPMJ diff	Giant?	Dw:Gi	Dil. V	Dil. R
	6100 = **F9**	6078 = **F9**	1.17	1.16	11.58	10.48	1.39	0.25	-0.6 (± 1.3)	-7.8 (± 0.9)	-0.06	-9.32	0	199:1	1 %	1 %

i	N_{harm}	T_0	Period [d]	a_2 [mag]
Plot				





Period: 1.0914211 Epoch i: 4023.295 Re-plot

Pk1 | **Pk2** | Pk3 | Pk4 | Pk5 | MCMC | Updated
Px2 | P/2 | Zoom | Transits | Opti-fold

P1=**1.091** P2=**1.091**
harmonics on | lime line = selected period P
magenta lines = harmonics = {0.5, 1.5, 2, 3, 4} x P
Detrended Data | Full archive LC

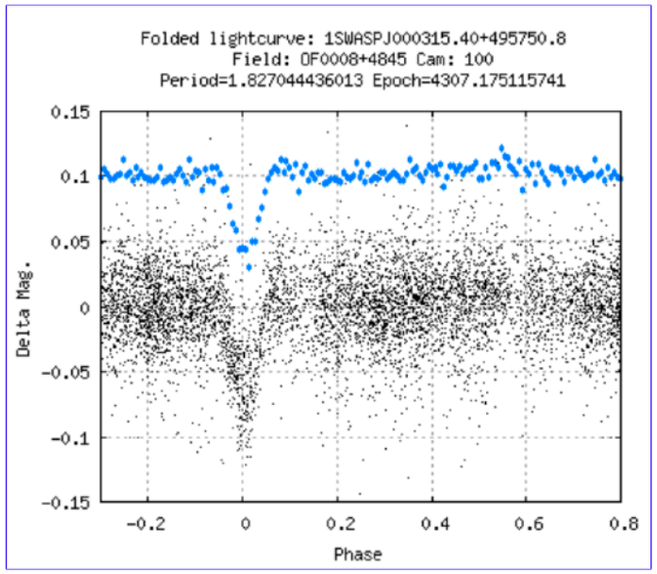
i	Period	Epoch	Width (hr)	Depth (mag)	Del Chisq	S/N Red	N_tr	Ellip Var	Ellip S/N	Frac. in Tr.	SN Anti	R_{Pl} VK	R_{Pl} JH	sig det eff	H_match i	Clmp	All
	1.091421138590253	4023.2956	2.7134	-0.0103	1159.3779	-21.2823	65	0.0004	2.7513	0.0719	1159.3779	1	1.01	94.5713	232	0.4	

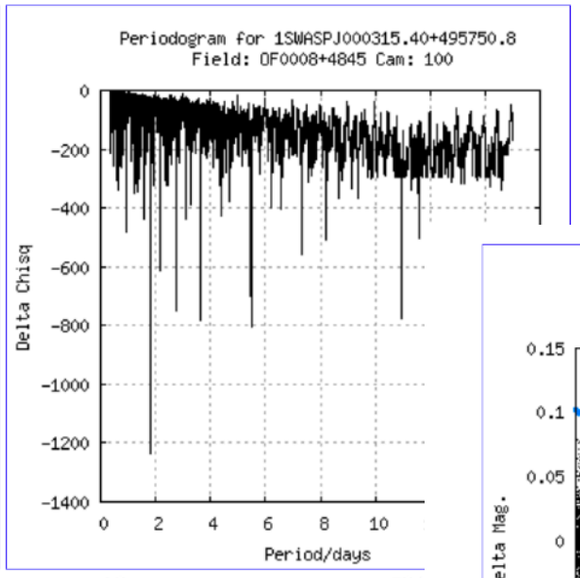
i	Period	Epoch	Width (hr)	Depth	Impact	R_{star}	M_{star}	R_{Pl}	Prob pl	Prob MS	Prob imp	Chisq_cs	Chisq_ucs	Q	All	
	1.0914222 ±	5110.34895 ±	2.7032	0.0090	0.03	1.350	1.065	1.248	0.998	0.065	1	16887.3	16882.8	5.33	Plot 1	Plot 2

Updated ephemeris: Period = 1.09141345 Epoch = 4500.24396 Source = Update

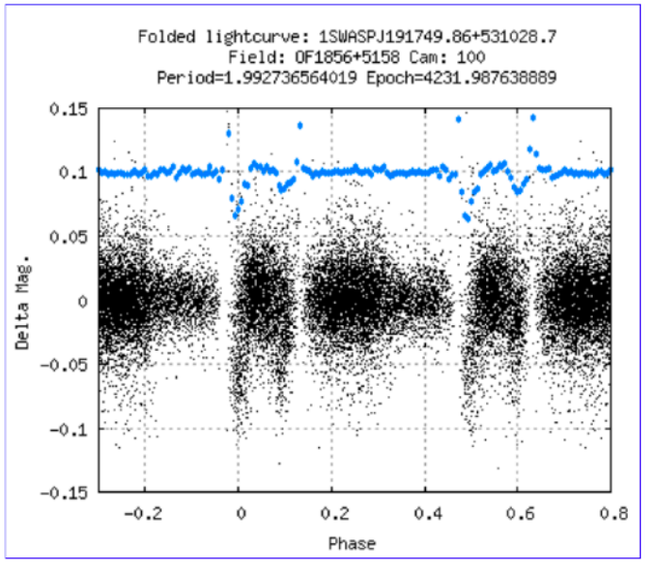
Includes revised LT z + previous

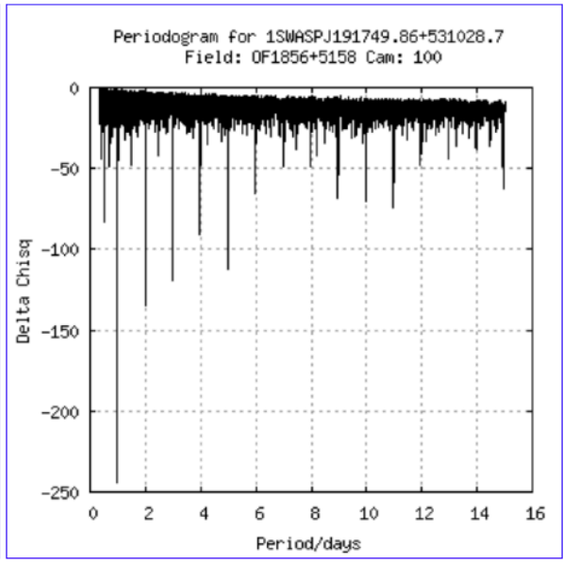
Non-planets



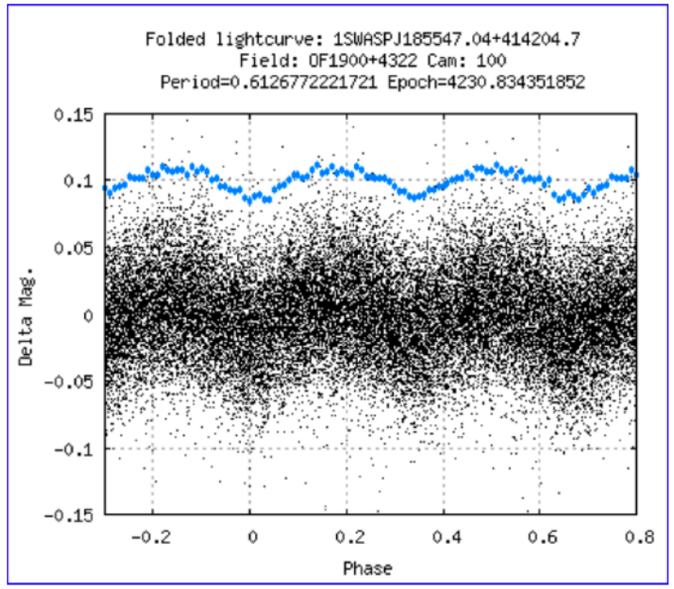


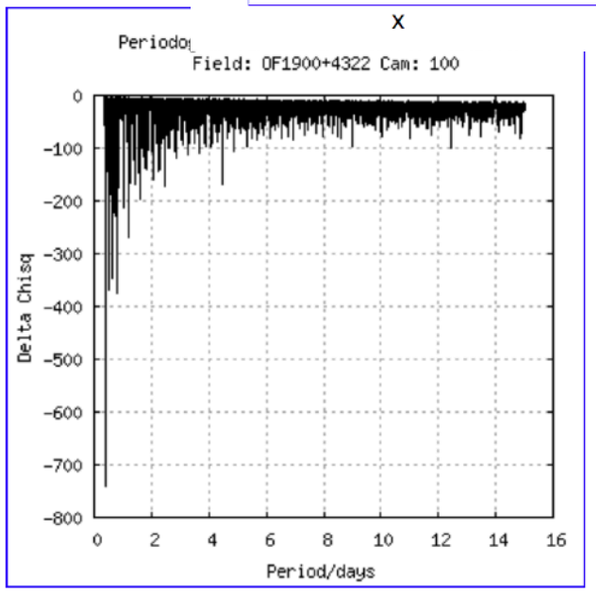
EB 16 12 F9/





X 12.2 F7/F1





V 48 12.9 G1/G6

Decision Trees, cont.

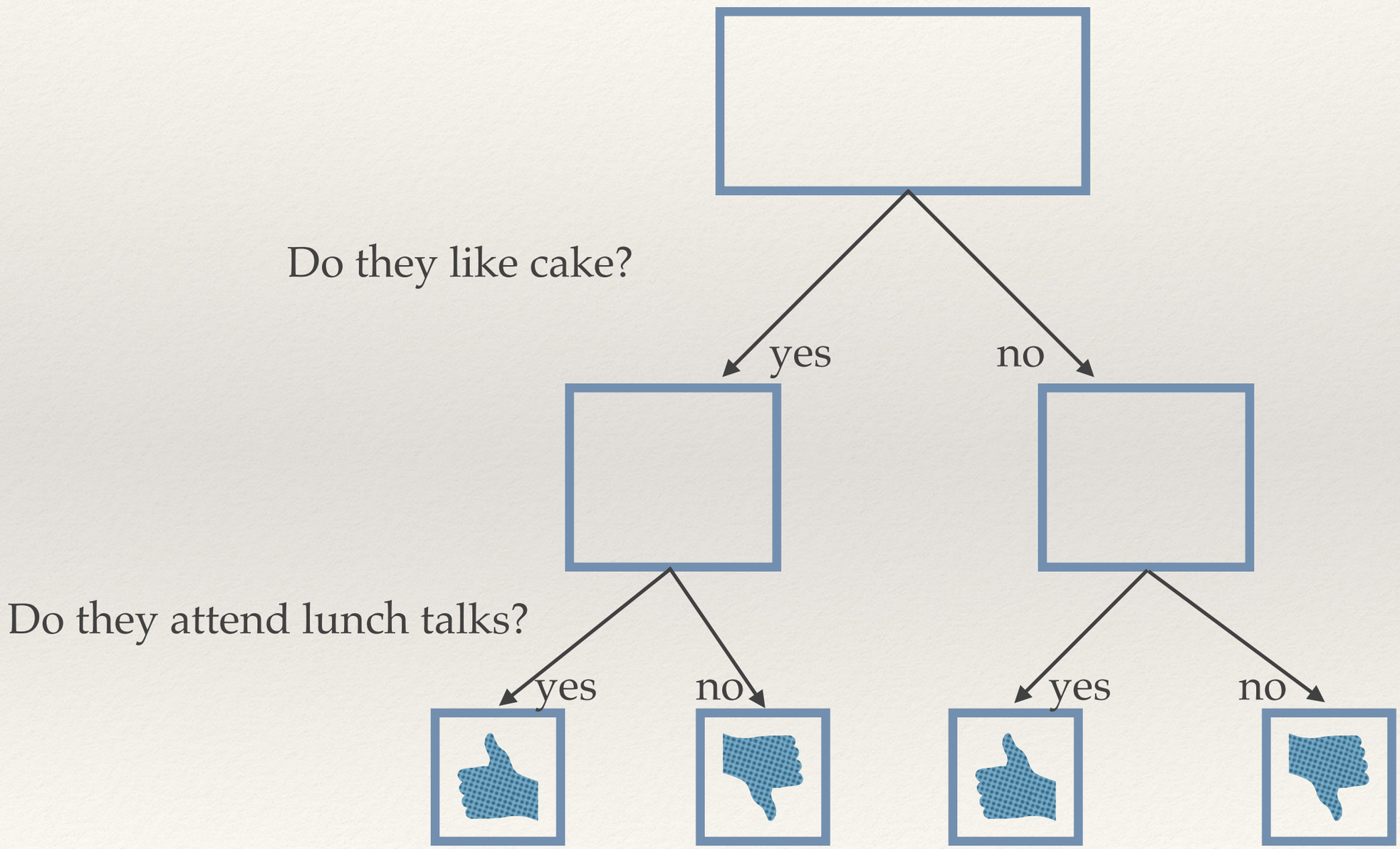
SWASP ID	period	flag	npts_good	ntrans	delta_chisq	sn_red	sn_ellipse	epoch	depth	width
1SWASP J000127.75+212725.4	0.57270379	EB	15379	80	-484.303	-12.0973	10.6992	3152.83972	-0.008053	0.134444
1SWASP J000052.49+183134.0	0.60913973	EB	6921	36	-29867	-35.4662	30.7823	3153.28653	-0.171516	0.127369
1SWASP J000109.42+183608.7	4.37296324	EBLM	6907	6	-1283.6	-14.7324	6.38205	3155.72369	-0.027165	0.037079
1SWASP J000212.21+232307.5	0.56673416	V	7122	72	-100.159	-9.29942	1.92871	3153.15242	-0.005445	0.103688
1SWASP J001135.44+540355.0	0.92167482	Blend	6485	27	-181.714	-9.99203	3.71982	4307.17068	-0.003319	0.083448
1SWASP J000701.59+550520.3	0.63060637	Blend	5448	34	-136.11	-3.98514	2.79027	4306.63332	-0.005286	0.09863
1SWASP J002353.18+531300.5	2.22481918	Blend	6531	15	-242.354	-8.08972	2.48781	4306.69642	-0.007514	0.068174
1SWASP J002405.36+540006.3	2.05844717	Blend	6206	17	-5956.71	-36.2975	11.5314	4307.0797	-0.050111	0.100834
1SWASP J000457.76+542324.1	0.57129588	Blend	4722	45	-238.721	-8.26127	8.84423	4306.84582	-0.009953	0.162668
1SWASP J003207.69+552639.7	1.12251859	Blend	6015	21	-787.015	-13.4128	2.59525	4306.99942	-0.009391	0.08948
1SWASP J000840.84+564830.8	1.82434197	EB	3348	11	-2007.84	-21.2804	2.38881	4307.00293	-0.039114	0.084493
1SWASP J001639.36+554303.8	2.0537463	EB	5867	12	-1648.57	-28.743	4.79367	4308.07747	-0.021158	0.055003
1SWASP J000956.14+514645.6	1.5053803	EB	5413	11	-314.363	-20.0333	3.77417	4307.75583	-0.038489	0.050412
1SWASP J002846.61+540216.0	3.17539268	EB	6472	12	-1590.91	-12.9015	4.5351	4307.03821	-0.011209	0.059312
1SWASP J002934.16+544536.4	1.26287809	EB	6518	25	-220.696	-15.4754	1.61745	4306.70776	-0.009897	0.091285
1SWASP J000204.38+512929.6	3.72707556	EB	6999	14	-1089.02	-12.2154	8.34517	4308.94264	-0.014242	0.079246
1SWASP J001228.10+540237.2	3.1331131	EB	5646	13	-1843.43	-19.8631	5.55422	4307.11694	-0.019075	0.060498
1SWASP J003437.56+560154.1	0.69715818	EB	5759	39	-1127.88	-12.085	0.721712	4306.85041	-0.007045	0.145606
1SWASP J002408.89+535943.7	2.05838382	EB	6128	17	-7024.54	-41.7953	11.8323	4307.08272	-0.061103	0.100837
1SWASP J002750.87+540447.9	0.58153074	EB	6462	55	-4432.94	-19.2669	5.22847	4307.00931	-0.021377	0.179943
1SWASP J002817.59+532241.9	0.9389853	EB	6488	26	-1770.88	-14.934	21.5897	4307.3819	-0.024853	0.118089
1SWASP J235851.80+511900.4	0.77472453	EB	8833	58	-3042.83	-48.0042	7.90219	4307.22617	-0.011358	0.276186
1SWASP J003815.99+530859.0	1.19035094	EB	6504	24	-1294.1	-14.6479	10.2932	4307.62457	-0.010775	0.084756
1SWASP J000336.89+514410.9	0.74190145	EB	4930	39	-1521.89	-12.903	14.7643	4307.28932	-0.008023	0.166382
1SWASP J003716.48+530200.6	0.66151918	EB	6575	35	-7157.41	-35.6172	7.0645	4307.24146	-0.053143	0.115081
1SWASP J002358.45+531324.2	2.22464684	EBLM	6519	15	-927.645	-14.4261	1.21588	4306.70208	-0.012678	0.075975
1SWASP J001926.76+535259.9	1.56928672	EBLM	6540	13	-1370.74	-23.3096	2.21113	4307.43982	-0.023608	0.048399
1SWASP J002133.89+522600.4	1.52861973	EBLM	6651	16	-1820.22	-24.468	1.17333	4307.62561	-0.041866	0.065737
1SWASP J000727.39+540630.3	0.71484219	V	5379	32	-116.536	-7.61723	2.52675	4307.07126	-0.007756	0.128044
1SWASP J003325.08+522101.2	0.88839718	V	6228	40	-623.997	-28.3799	3.68646	4307.13057	-0.003393	0.23012
1SWASP J000037.94+542337.0	0.65320118	V	6284	41	-316.463	-17.2368	2.40522	4306.65463	-0.014533	0.076134
1SWASP J001001.61+514931.5	0.79919425	V	5517	31	-227.946	-10.9239	4.13269	4307.33421	-0.006371	0.066201

etc.

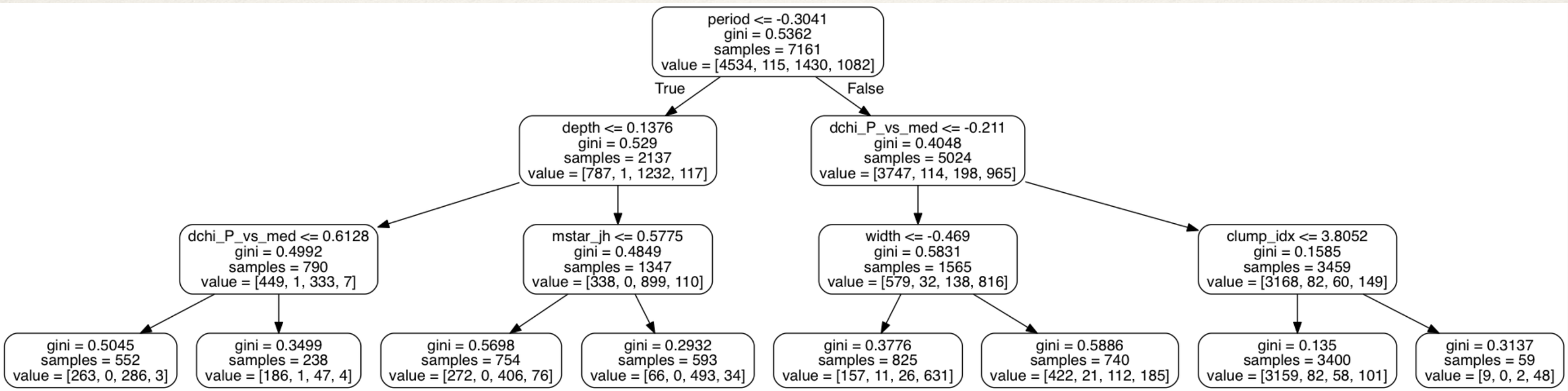
Now for the Machine Learning!

Decision Trees

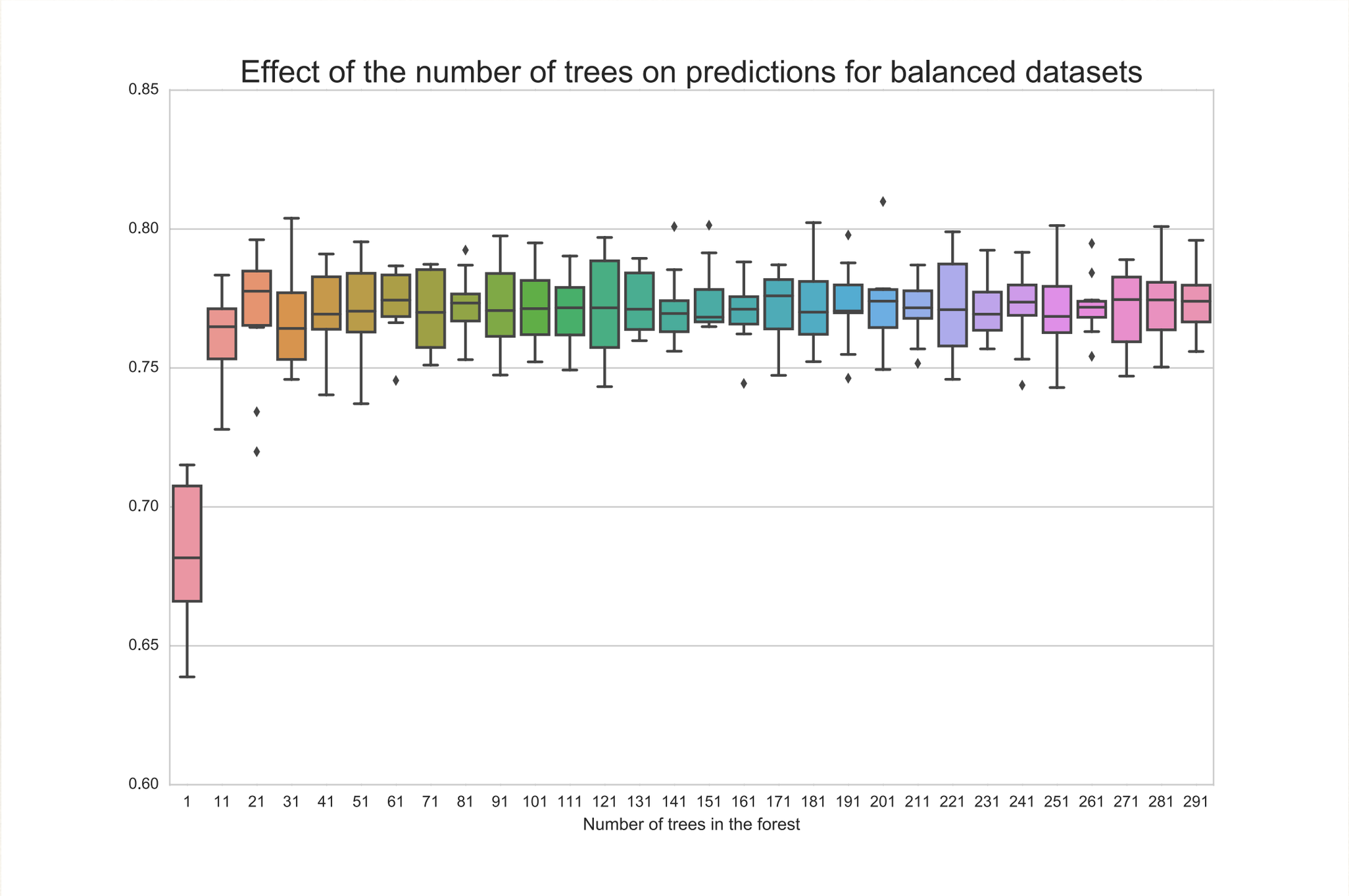
Is the person a good scientist?



Decision Trees, cont.

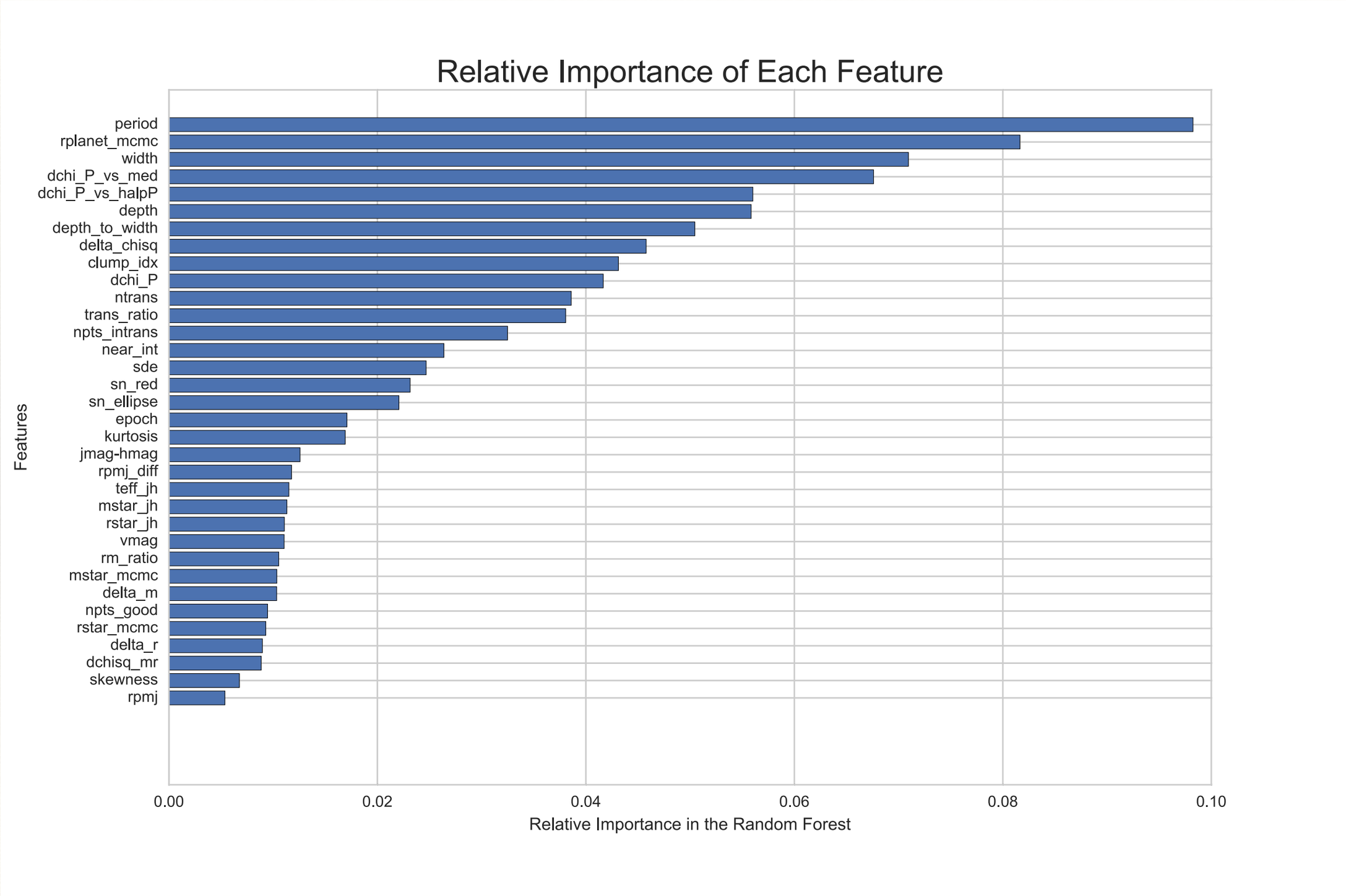


Random Forest, cont.



More trees = better predictions (up to a point)

Random forest, cont.



Neural Networks

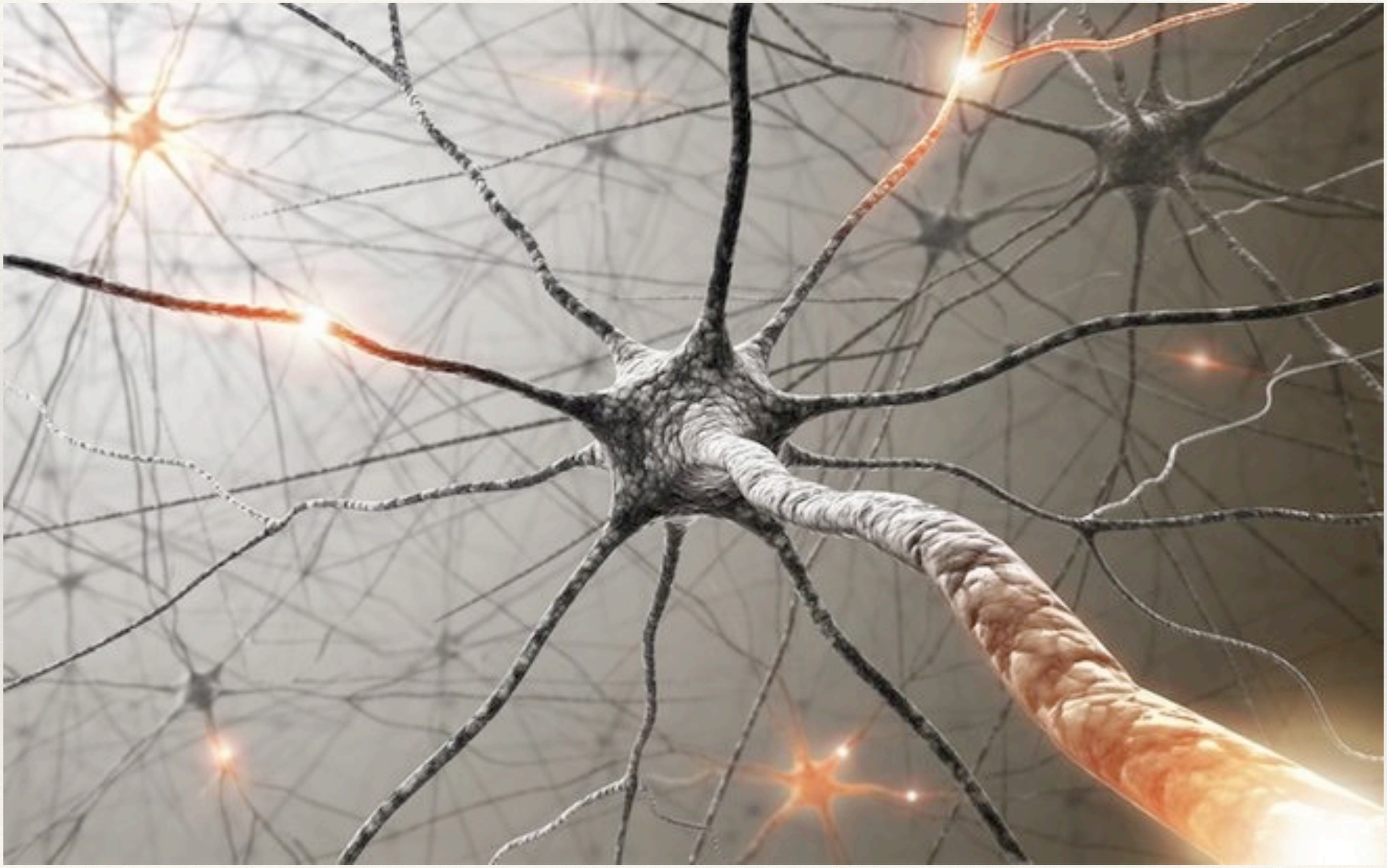
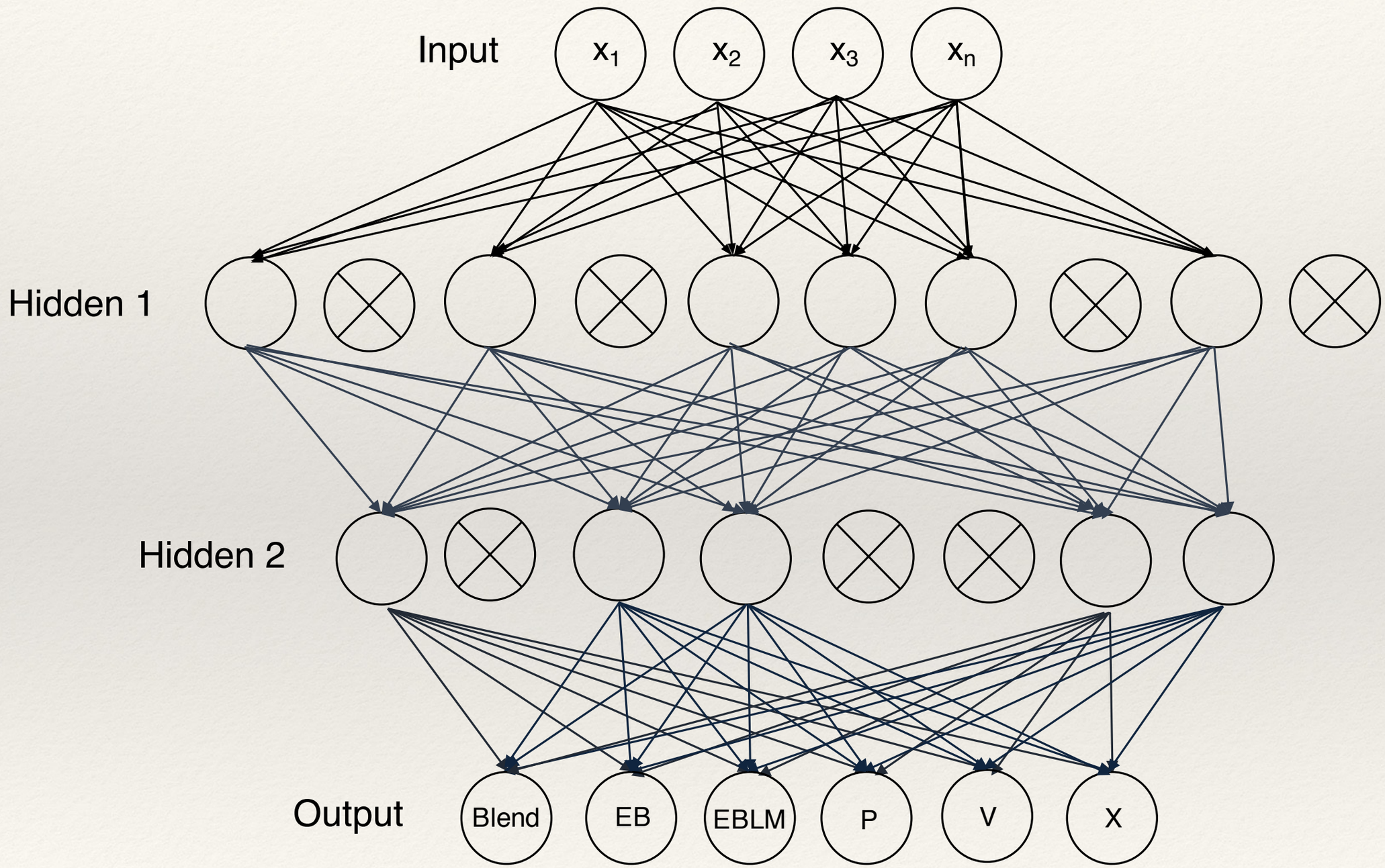
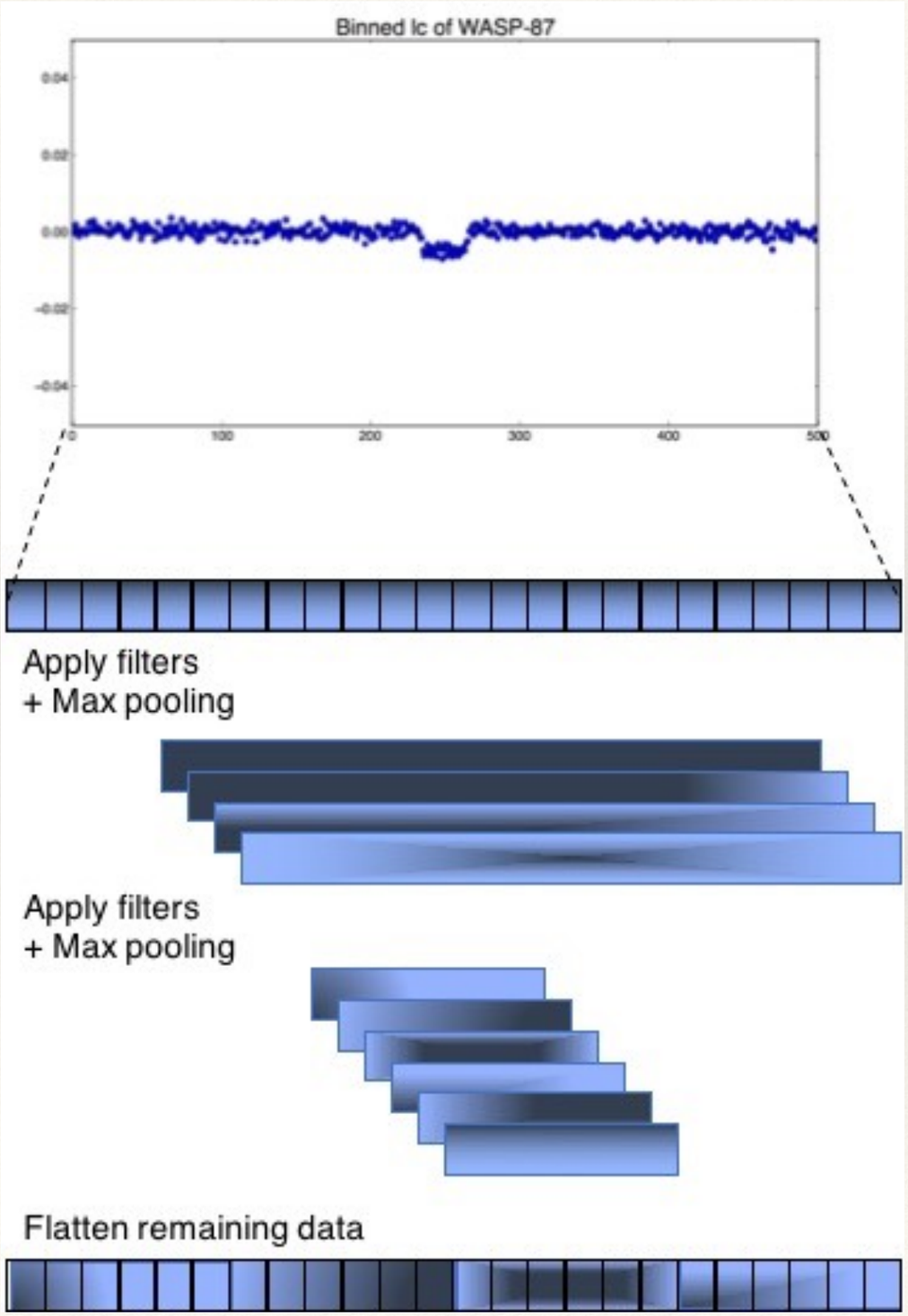


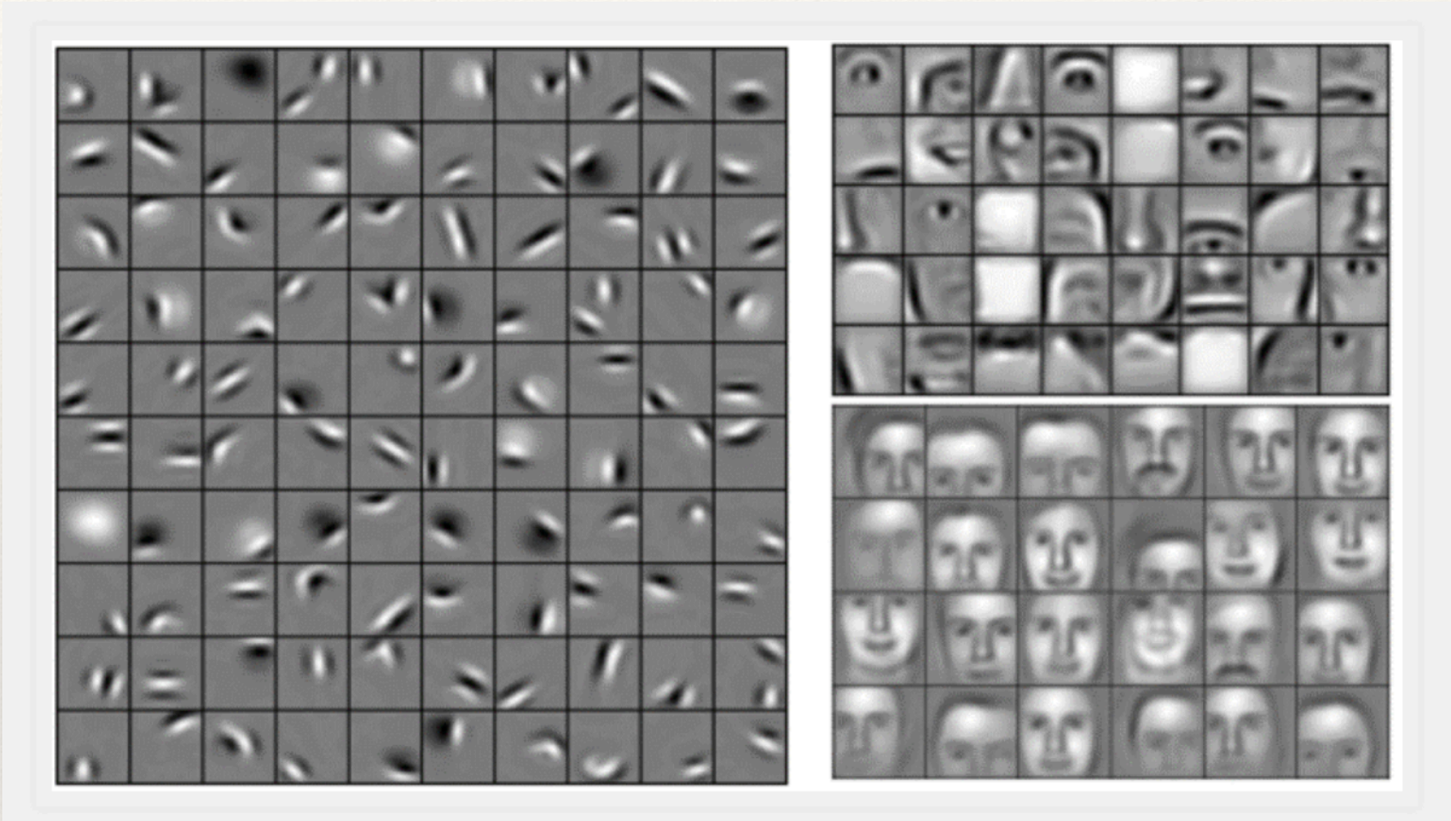
image: https://www.livescience.com/18749-human-brain-cell-number.html

Neural networks, cont.



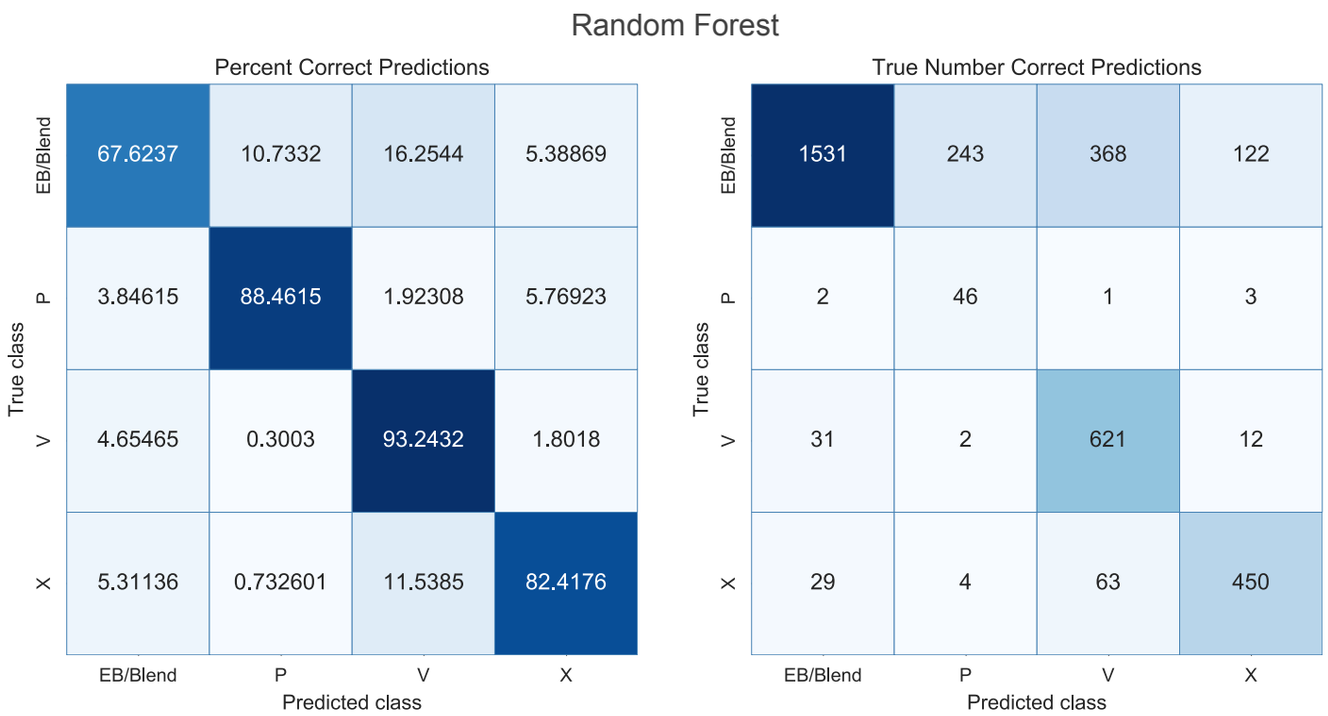
Convolution

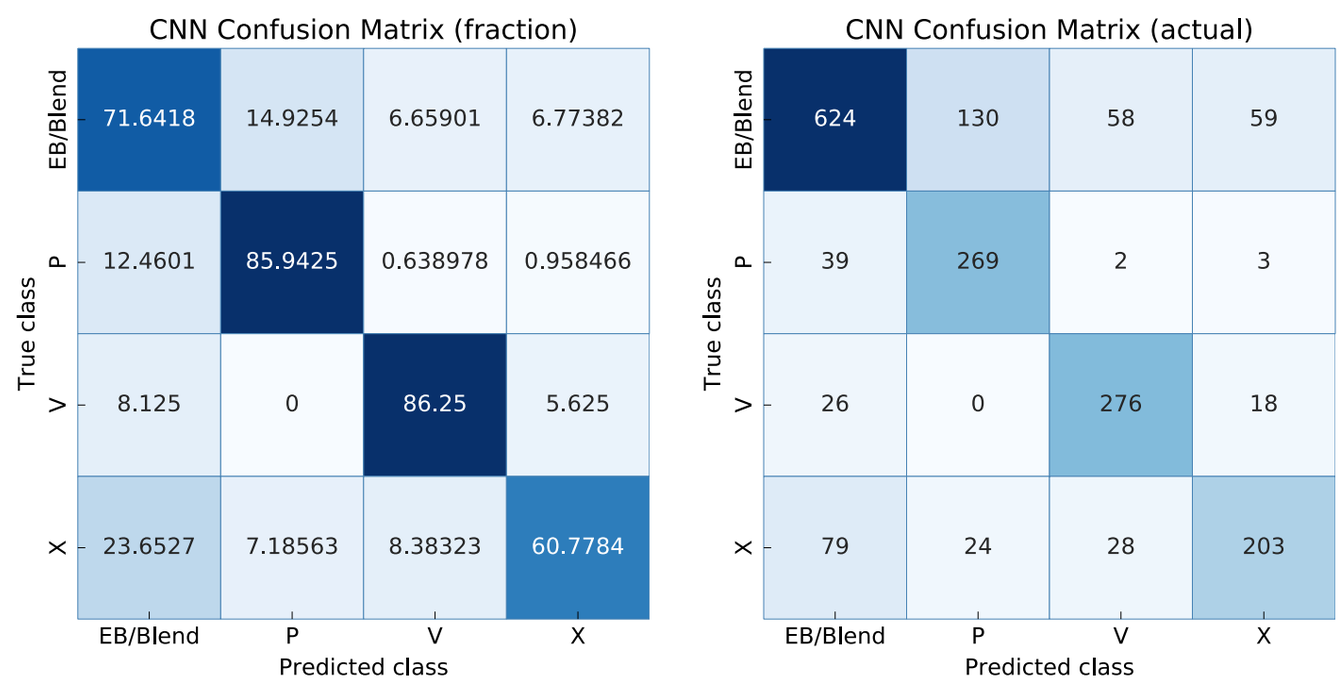




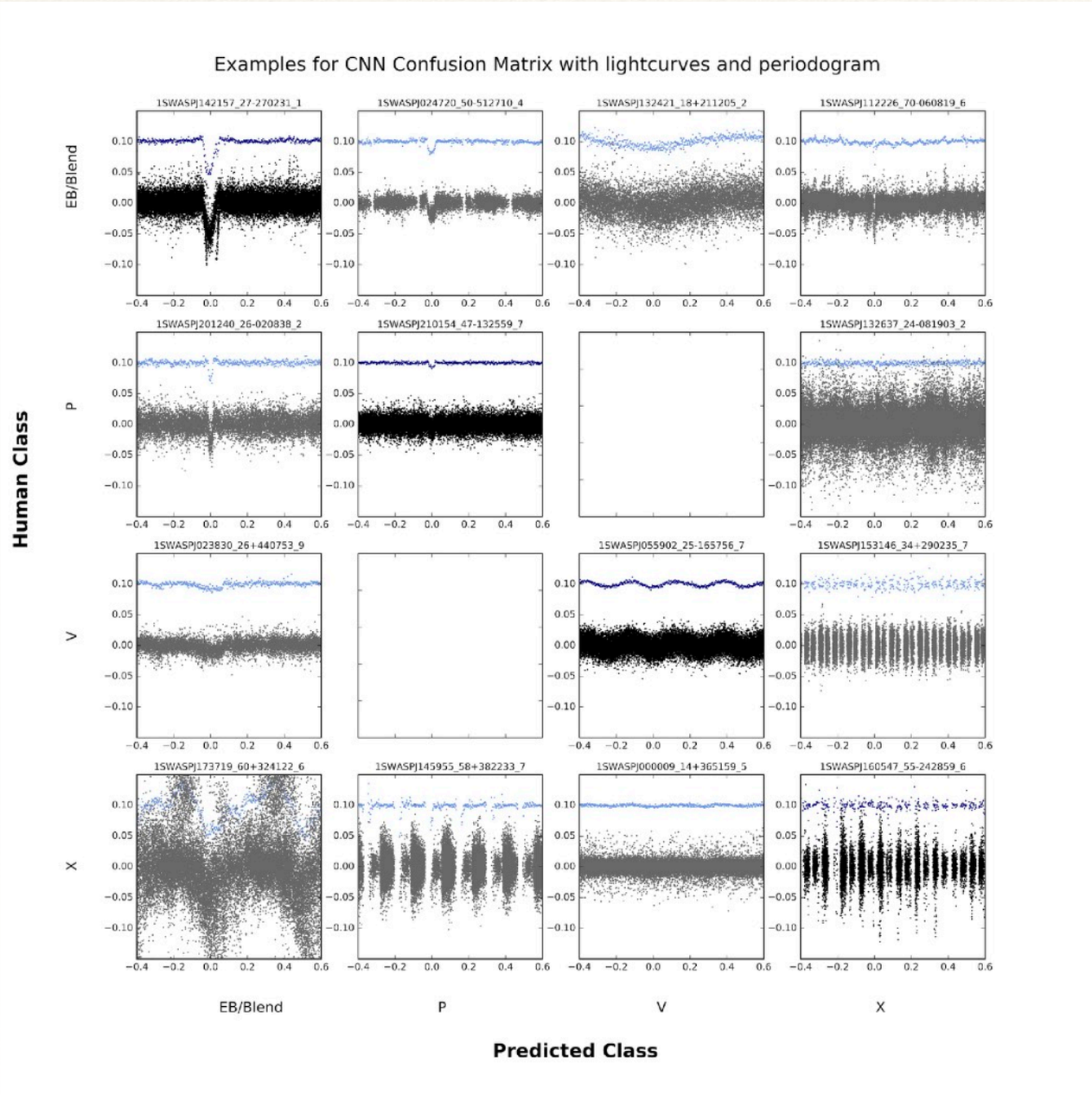
above:https://brohrer.github.io/how_convolutional_neural_networks_work.html
original image from Lee et al, 2009 proceedings

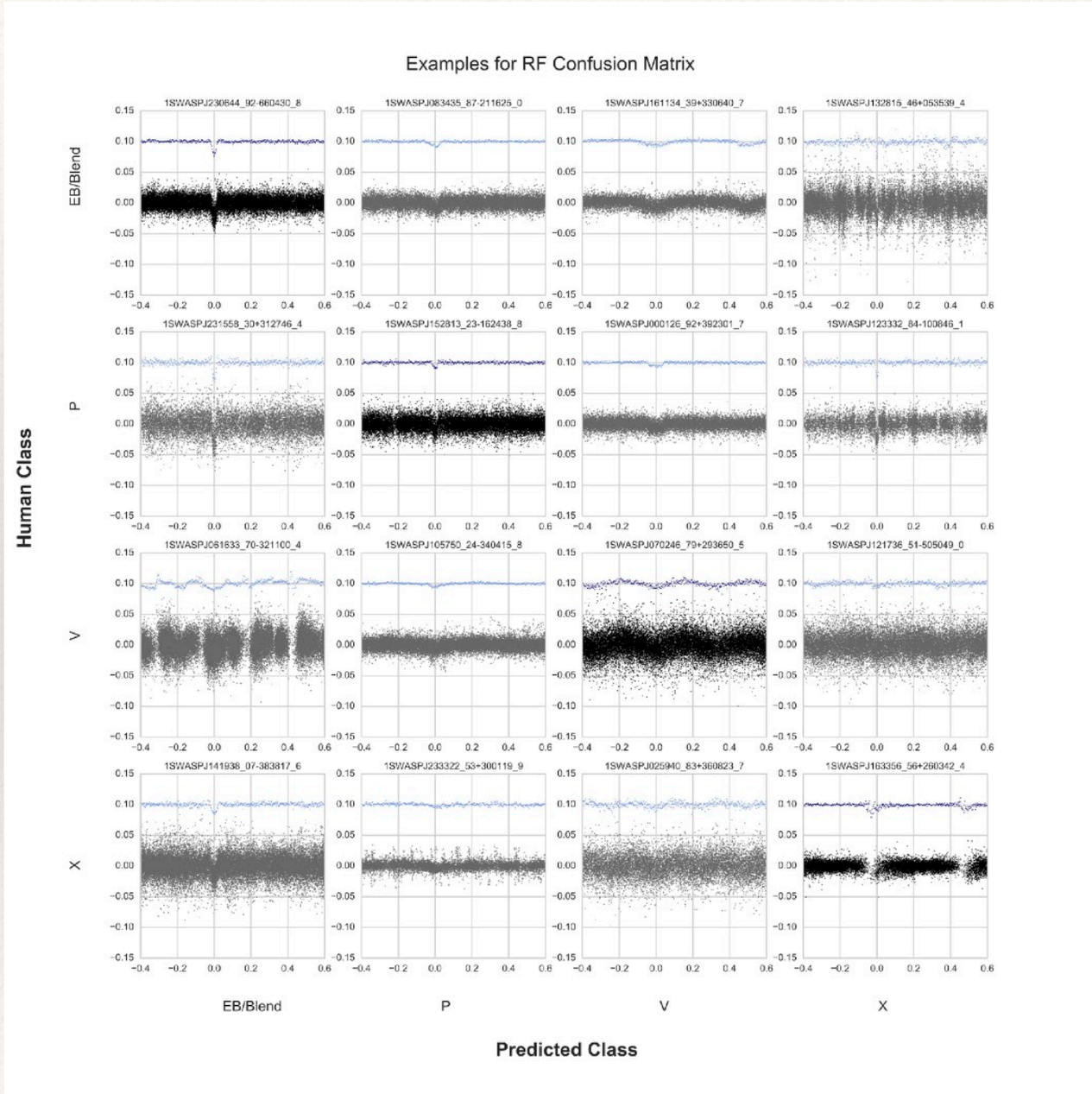
Measuring Results





Measuring results, cont.





Questions?

Further reading

Random Forest original paper: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

Machine learning review: LeCun, Bengio, & Hinton, *Nature* **volume 521**, pages 436–444 (28 May 2015)

Online course for machine learning: https://www.coursera.org/learn/machine-learning