	Accuracy	F1	Model	Batch	Epochs	Learning rate	Initialization
0	1,0%	0,1%	4 layers	32	500	1,00E-04	Default algorithm
1	77,9%	77,5%	4 layers	32	500	1,00E-05	Default algorithm
2	68,7%	65,9%	4 layers	64	500	1,00E-04	Default algorithm
3	79,1%	78,0%	4 layers	64	500	1,00E-05	Default algorithm
4	1,0%	0,1%	4 layers	32	1000	1,00E-04	Default algorithm
5	78,2%	78,9%	4 layers	32	1000	1,00E-05	Default algorithm
6	69,8%	67,8%	4 layers	64	1000	1,00E-04	Default algorithm
7	72,5%	69,8%	4 layers	64	1000	1,00E-05	Default algorithm
8	55,1%	52,8%	5 layers	32	500	1,00E-04	Default algorithm
9	72,9%	72,4%	5 layers	32	500	1,00E-05	Default algorithm
10	66,6%	62,2%	5 layers	64	500	1,00E-04	Default algorithm
11	74,5%	74,3%	5 layers	64	500	1,00E-05	Default algorithm
12	56,1%	53,8%	5 layers	32	1000	1,00E-04	Default algorithm
13	72,2%	72,2%	5 layers	32	1000	1,00E-05	Default algorithm
14	67,7%	64,4%	5 layers	64	1000	1,00E-04	Default algorithm
15	74,8%	74,2%	5 layers	64	1000	1,00E-05	Default algorithm
16	2,2%	0,1%	4 layers	32	500	1,00E-04	Pre-trained CAE
17	71,1%	69,1%	4 layers	32	500	1,00E-05	Pre-trained CAE
18	72,0%	71,4%	4 layers	64	500	1,00E-04	Pre-trained CAE
19	71,4%	69,0%	4 layers	64	500	1,00E-05	Pre-trained CAE
20	2,2%	0,1%	4 layers	32	1000	1,00E-04	Pre-trained CAE
21	73,5%	71,6%	4 layers	32	1000	1,00E-05	Pre-trained CAE
22	75,1%	74,7%	4 layers	64	1000	1,00E-04	Pre-trained CAE
23	72,1%	69,1%	4 layers	64	1000	1,00E-05	Pre-trained CAE
24	76,6%	74,9%	5 layers	32	500	1,00E-04	Pre-trained CAE
25	80,5%	79,4%	5 layers	32	500	1,00E-05	Pre-trained CAE
26	63,1%	61,3%	5 layers	64	500	1,00E-04	Pre-trained CAE
27	80,0%	79,0%	5 layers	64	500	1,00E-05	Pre-trained CAE
28	77,7%	77,6%	5 layers	32	1000	1,00E-04	Pre-trained CAE
29	78,2%	77,5%	5 layers	32	1000	1,00E-05	Pre-trained CAE
30	75,7%	76,2%	5 layers	64	1000	1,00E-04	Pre-trained CAE
31	79,9%	79,6%	5 layers	64	1000	1,00E-05	Pre-trained CAE

Supplementary Table 2 - Classification performance (Accuracy and F1-score) of models trained on BrainPedia dataset with different hyperparameters (batch size, epochs, architecture). Best model for the Default algorithm initialization is a 4-layers architecture, a batch size of 32 for 500 epochs and a learning rate of 1e-05.
For the pre-trained CAE, best model was the one with a 5-layers architecture, a batch size of 64, 500 epochs and a learning rate of 1e-05.