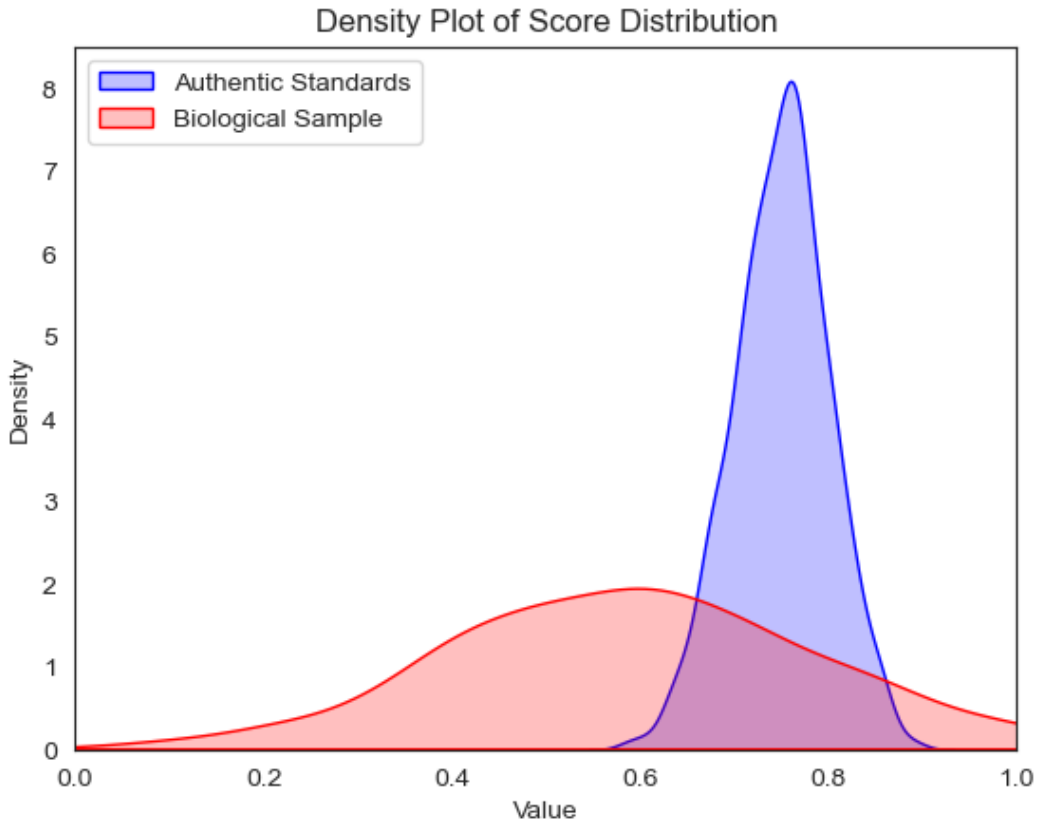
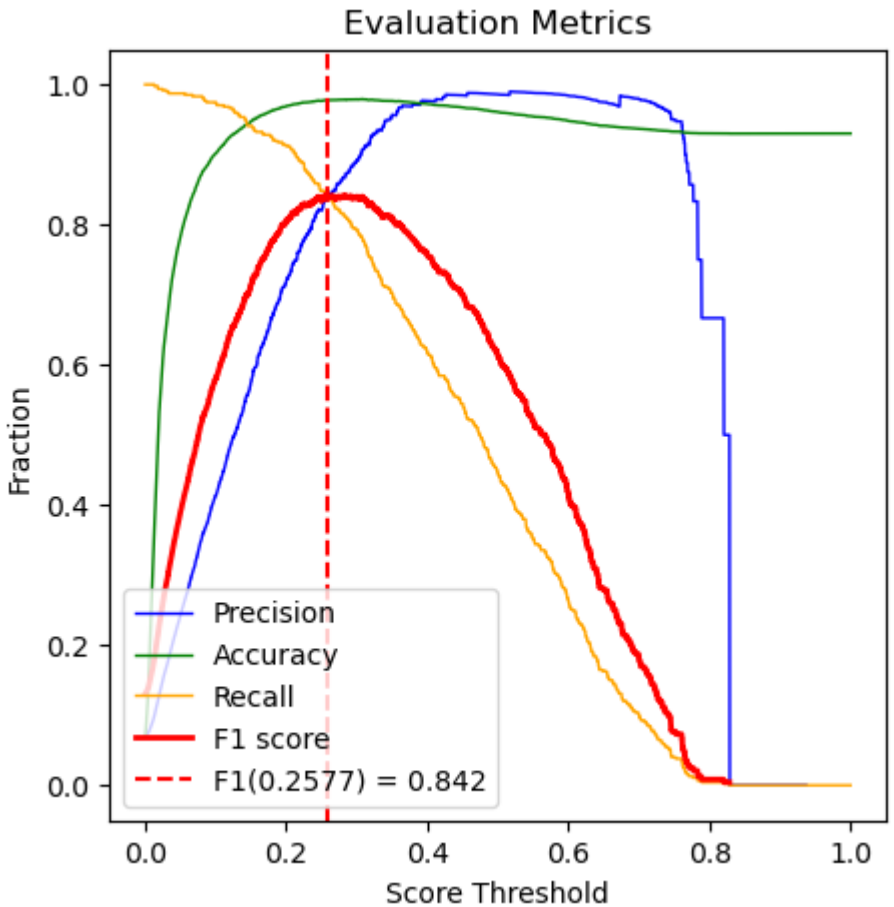
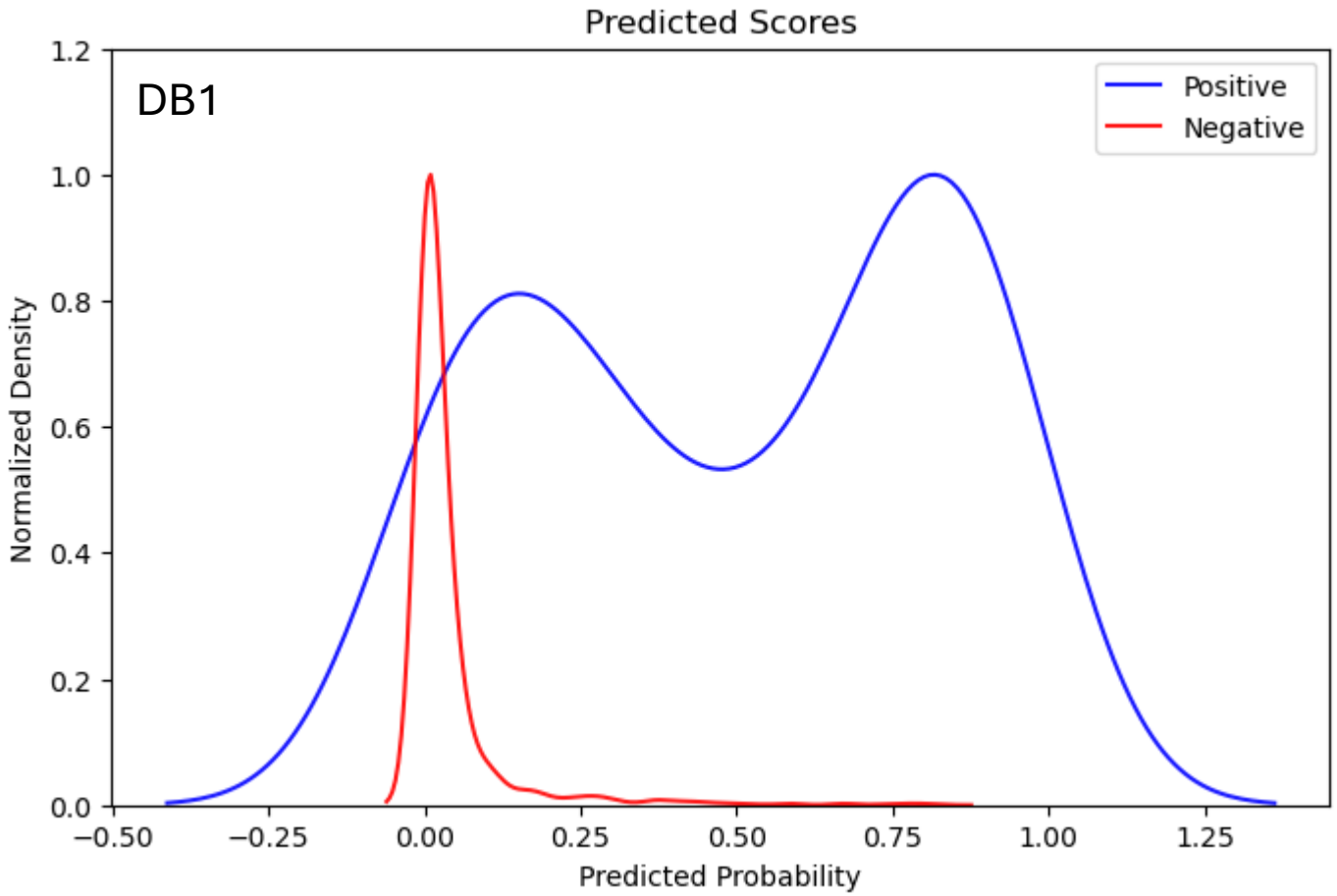


Hypothesized Reasons

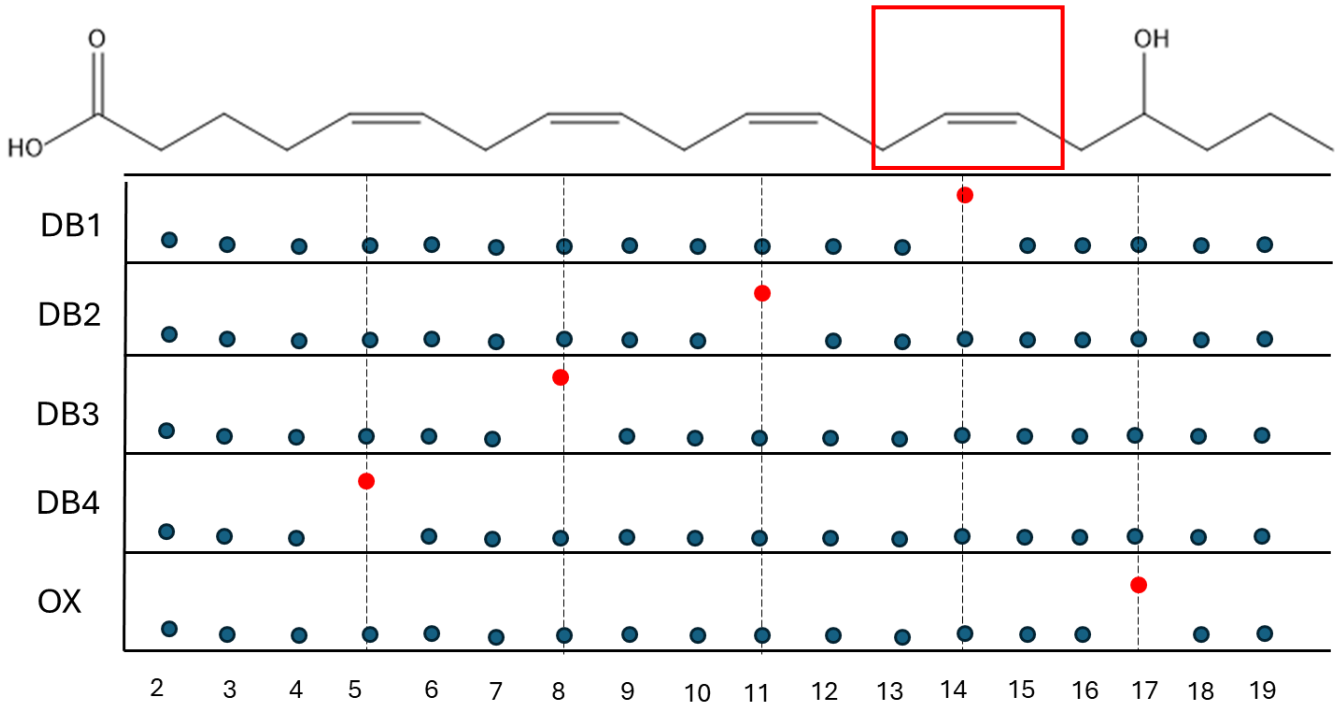
- Chimeric Spectra (expected in nature)
- Spectral Quality
- Overconfidence (Standards not covering the whole space)



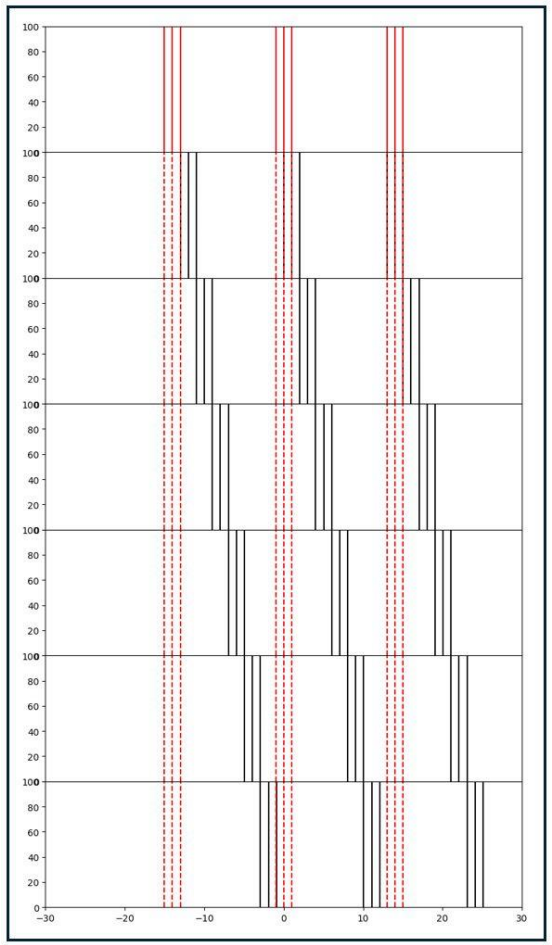
Threshold overfitting
Weird Shape of curves

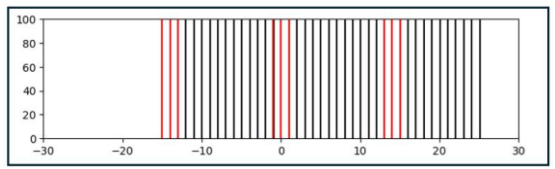


Epoch 100/500, Loss: 0.2016 Epoch 200/500, Loss: 0.1691
Epoch 300/500, Loss: 0.1479 Epoch 400/500, Loss: 0.1376
Epoch 500/500, Loss: 0.1264
Test Accuracy : 0.9654
Test Precision: 0.8200
Test Recall : 0.5256



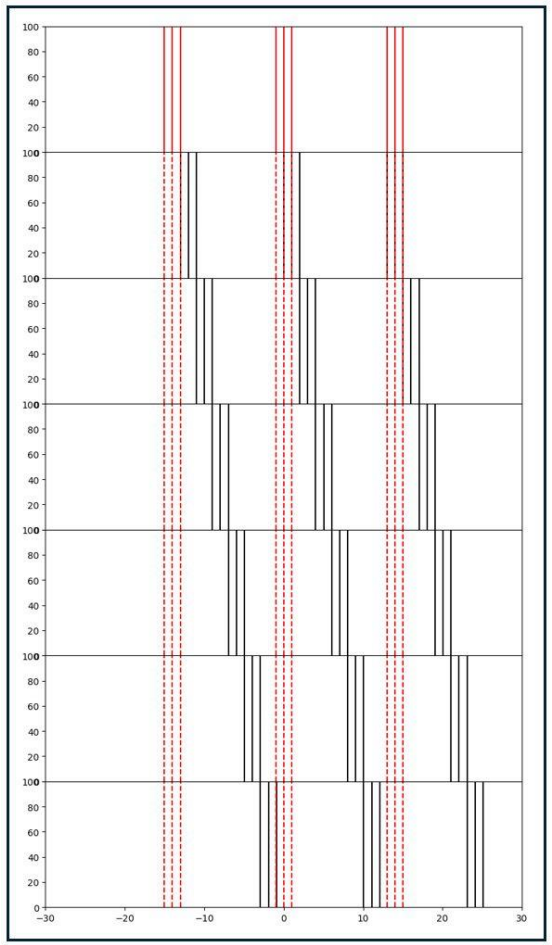
This approach is more intensity pattern based than mz based for DB1

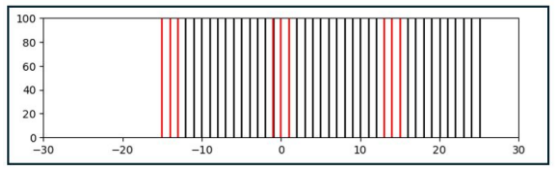




C: 13-22 and DB:0-6

9





DB: 1, Clusters: 1

•Cluster 1.) Pattern 1-1-1-1-1-1-1-1-1-1: 270919/943070 [28.727 %]

DB: 2, Clusters: 1

•Cluster 1.) Pattern 1-1-1-1-1-1-1-1-1-1: 269324/269324 [100.0 %]

DB: 3, Clusters: 1

•Cluster 1.) Pattern 1-1-1-1-1-1-1-1-1-1: 259589/259589 [100.0 %]

DB: 4, Clusters: 1

•Cluster 1.) Pattern 1-1-1-1-1-1-1-1-1-1: 226699/226699 [100.0 %]

DB: 5, Clusters: 1

•Cluster 1.) Pattern 1-1-1-1-1-1-1-1-1-1: 158862/158862 [100.0 %]

DB: 6, Clusters: 1

•Cluster 1.) Pattern 1-1-1-1-1-1-1-1-1-1: 70785/70785 [100.0 %]

42

DB: 1, Clusters: 3

•Cluster 1.) Pattern 0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-1-1-1-0-1-0-0-0-0-0-0-0-0-0-0-1-1-1-1-1: 245760/245760 [100.0 %]

•Cluster 2.) Pattern 0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0: 24938/24938 [100.0 %]

•Cluster 3.) Pattern 0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-1-1-1-0-1-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0: 221/221 [100.0 %]

DB: 2, Clusters: 4

•Cluster 1.) Pattern 0-0-0-0-0-0-0-0-1-1-1-1-1-0-0-0-0-0-0-0-0-1-1-1-1-1-0-1-0-0-0-0-0-0-0-1-1-1-0-1-1: 238722/238722 [100.0 %]

•Cluster 2.) Pattern 0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0: 26798/26798 [100.0 %]

•Cluster 3.) Pattern 0-0-0-0-0-0-0-0-1-1-1-1-1-0-0-0-0-0-0-0-0-1-1-1-1-1-0-1-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0: 2298/2298 [100.0 %]

•Cluster 4.) Pattern 0-0-0-0-0-0-0-0-1-1-1-1-1-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0: 1506/1506 [100.0 %]

DB: 3, Clusters: 4

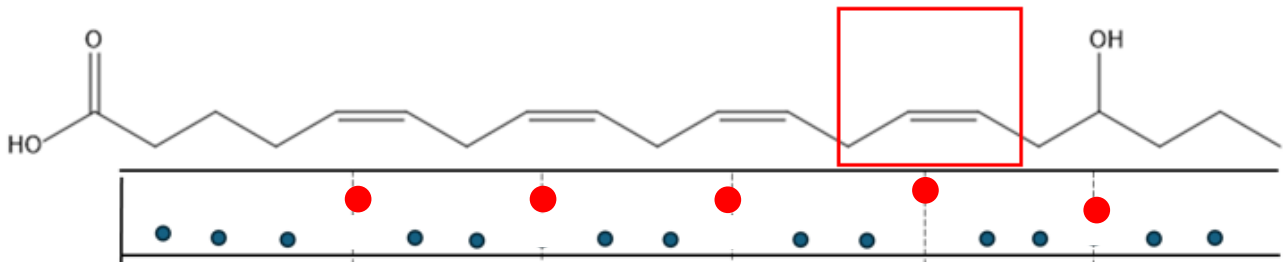
•Cluster 1.) Pattern 0-0-0-0-0-0-1-1-1-0-1-1-1-0-0-0-0-0-0-1-1-1-0-0-1-1-0-1-0-0-0-0-0-0-1-1-1-0-0-0-1-1-1: 208267/208267 [100.0 %]

•Cluster 2.) Pattern 0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0: 35764/35764 [100.0 %]

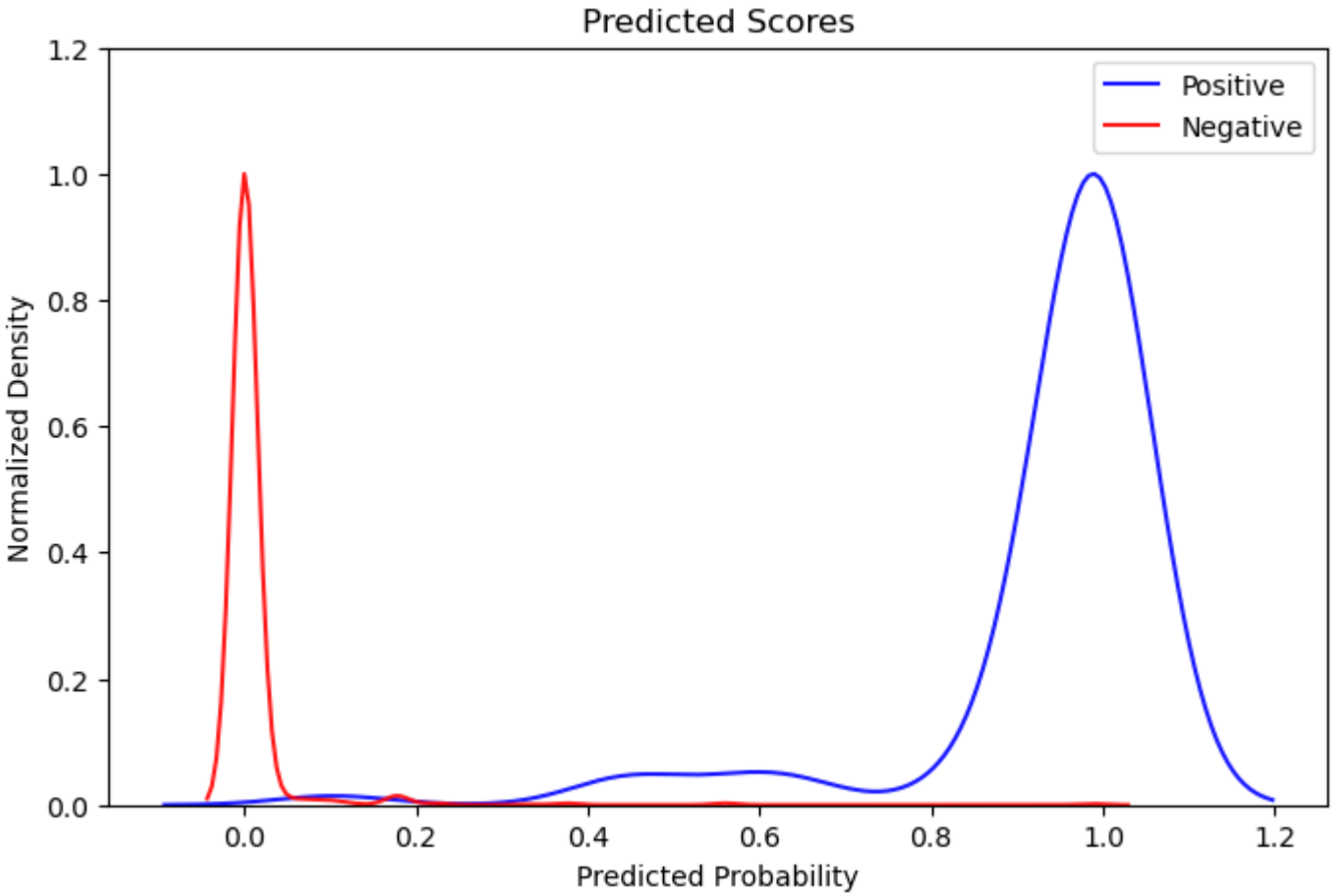
•Cluster 3.) Pattern 0-0-0-0-0-0-1-1-1-0-1-1-1-0-0-0-0-0-0-1-1-1-0-0-1-1-0-1-0-1-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0: 8997/8997 [100.0 %]

•Cluster 4.) Pattern 0-0-0-0-0-0-1-1-1-0-1-1-1-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0-0-0-0-0-0-0-1-1-1-0-0-0-0-0-0: 6561/6561 [100.0 %]

Etc.....



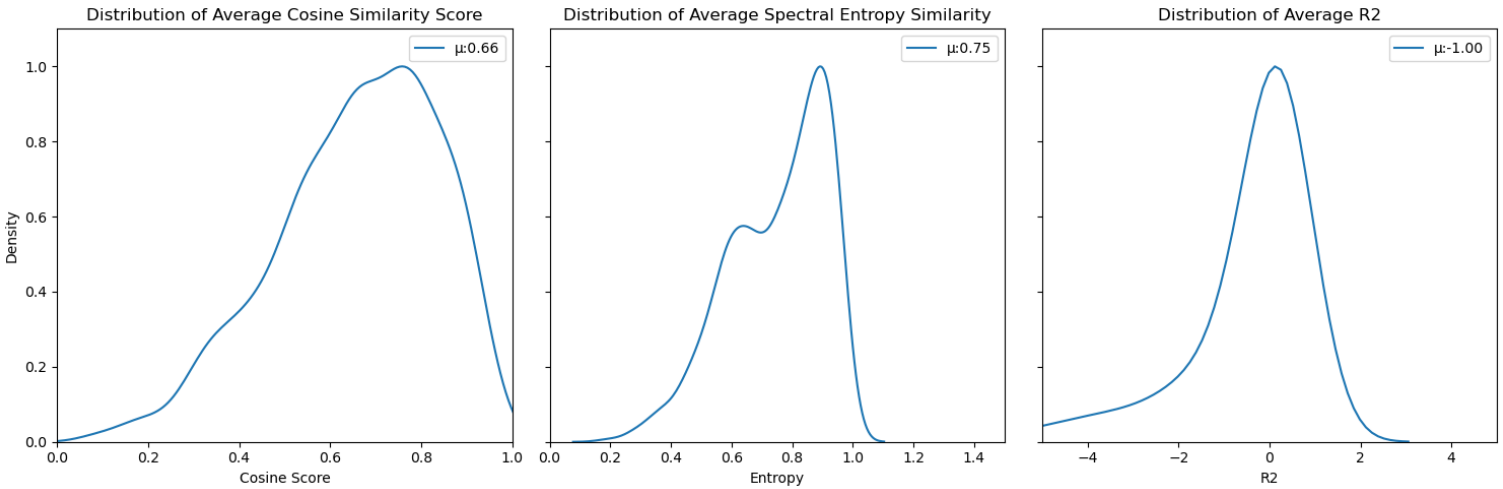
C: 13-22 and DB:0-6



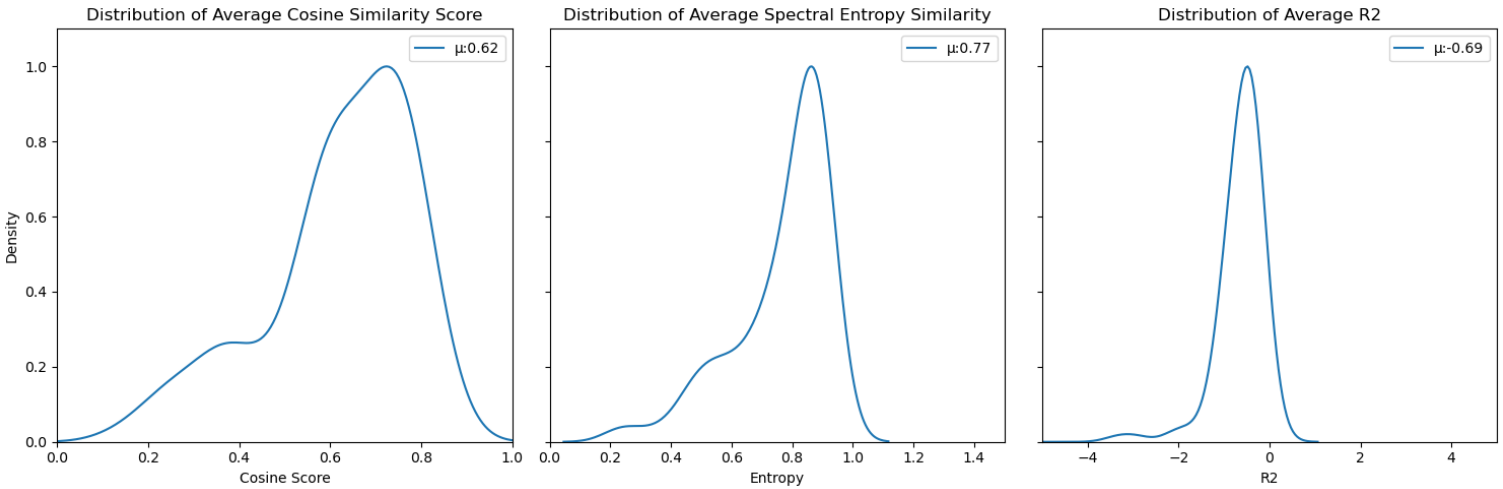

```
Epoch 100/500, Loss: 0.1647 Epoch 200/500,
Loss: 0.0831 Epoch 300/500, Loss: 0.0491 Epoch
400/500, Loss: 0.0361 Epoch 500/500, Loss:
0.0311
Test Accuracy : 0.9929
Test Precision: 0.9405
Test Recall : 0.9518
```

Chimeric Spectra and Spectral Quality based on Simulated data

Authentic Standard Reference Similarity Cleavage Intensity

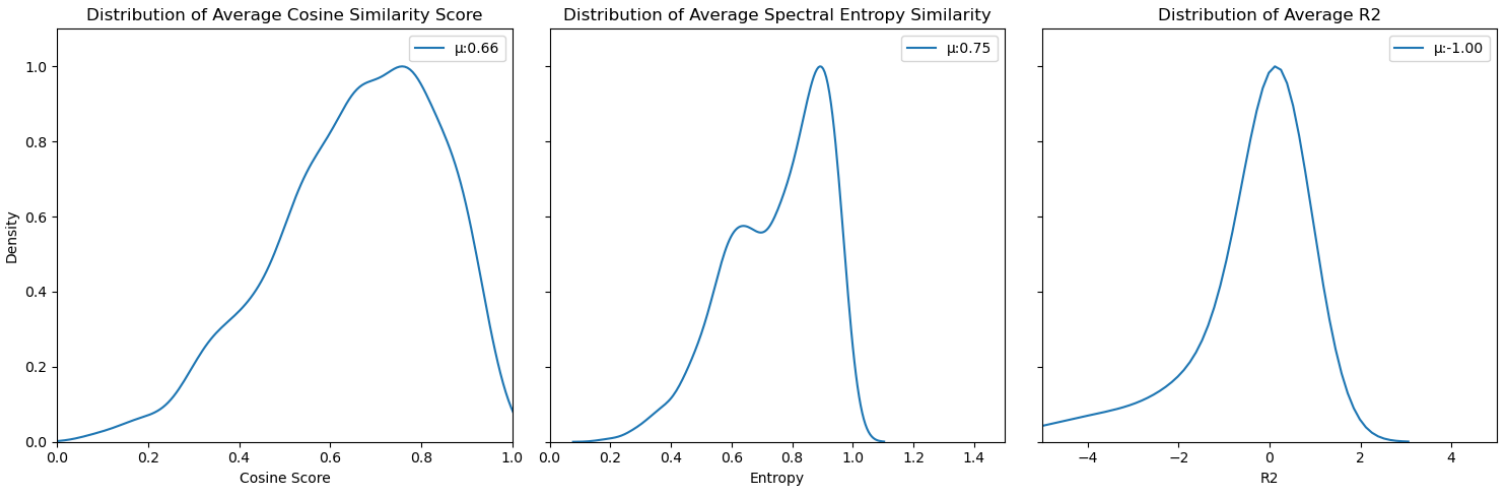


NZ

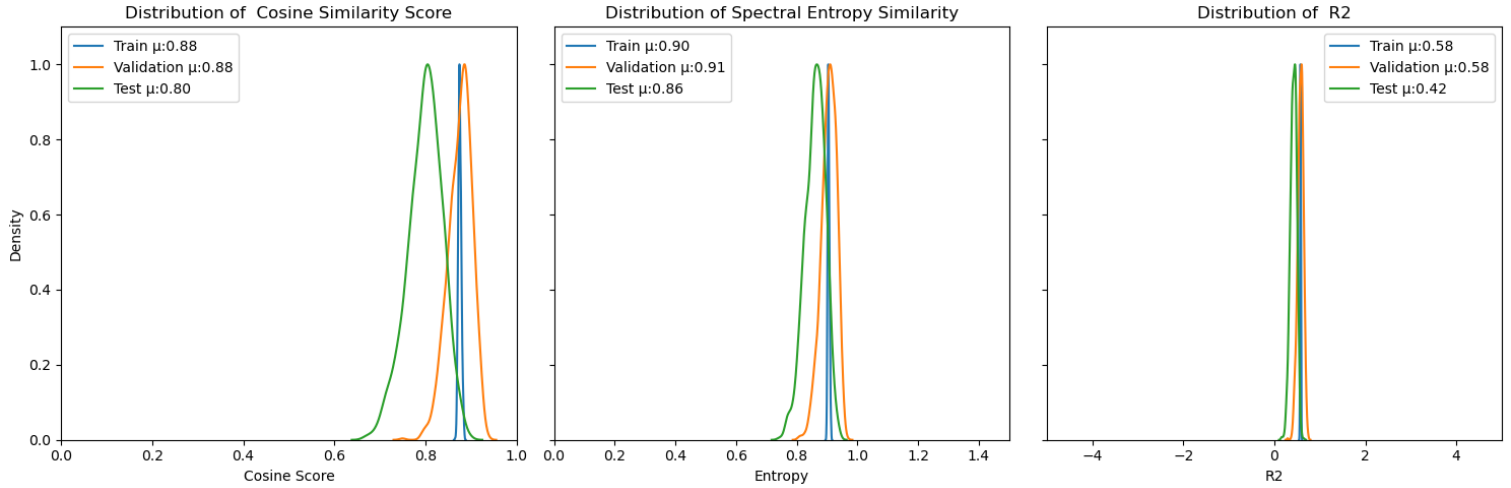


Chimeric Spectra and Spectral Quality based on Simulated data

Authentic Standard Reference Similarity Cleavage Intensity



Decision Tree



Chimeric Spectra and Spectral Quality based on Simulated data

I can even add the Noise from real spectra at the “unexpected m/z”

Compare on Simulation:

- 42 vs 9
- Mixture Chimeric Spectra at different ratio (1:1; 10:1 etc.)
- Robustness to missing peaks
- Robustness to Noise level

Name	Adduct	Chains	Carbons	Double Bonds	Position	Is DB	Is +1 DB	Is +2 DB	Is +3 DB	Is -1 DB	Is -2 DB	Is -3 DB	Number DB before	Number DB after	Number of Carbons before	Number of Carbons after	Number of DB in +-5	Overlap	Overlap X	Overlap X+H	Overlap X-H	Triplet Position	Linkage	MZ	ID_id	Intensity	Mean	Median	Sum	Max	Final Linkage
PC 15:0_18:1(	[M+Na]+	2	15	0	14	0	0	0	0	0	0	0	0	0	1	14	0	1	1	0	0	X	0,42	753.52844	5	386.9856	200.475798	138.470871	12429.49947	484.4069489	[450 492]
PC 15:0_18:1(	[M+Na]+	2	15	0	13	0	0	0	0	0	0	0	0	0	2	13	0	1	1	0	0	X	1,43	739.5128	5	484.40695	200.475798	138.470871	12429.49947	484.4069489	[451 493]
PC 15:0_18:1(	[M+Na]+	2	15	0	12	0	0	0	0	0	0	0	0	0	3	12	0	1	1	0	0	X	2,44	725.49713	5	484.40695	200.475798	138.470871	12429.49947	484.4069489	[452 494]
PC 15:0_18:1(	[M+Na]+	2	15	0	11	0	0	0	0	0	0	0	0	0	4	11	0	1	1	0	0	X	3,45	711.4815	5	484.40695	200.475798	138.470871	12429.49947	484.4069489	[453 495]
PC 15:0_18:1(	[M+Na]+	2	15	0	10	0	0	0	0	0	0	0	0	0	5	10	0	1	1	0	0	X	4,46	697.4659	5	482.91223	200.475798	138.470871	12429.49947	484.4069489	[454 496]
PC 15:0_18:1(	[M+Na]+	2	15	0	9	0	0	0	0	0	0	0	0	0	6	9	0	1	1	0	0	X	5,47	683.4502	5	267.49304	200.475798	138.470871	12429.49947	484.4069489	[455 497]
PC 15:0_18:1(	[M+Na]+	2	15	0	8	0	0	0	0	0	0	0	0	0	7	8	0	1	1	0	0	X	6,48	669.4346	5	484.40695	200.475798	138.470871	12429.49947	484.4069489	[456 498]
PC 15:0_18:1(	[M+Na]+	2	15	0	7	0	0	0	0	0	0	0	0	0	8	7	0	1	1	0	0	X	7,49	655.4189	5	484.40695	200.475798	138.470871	12429.49947	484.4069489	[457 499]
PC 15:0_18:1(	[M+Na]+	2	15	0	6	0	0	0	0	0	0	0	0	0	9	6	0	1	0	1	0	X	8,84	641.40326	5	294.77356	200.475798	138.470871	12429.49947	484.4069489	[458 534]
PC 15:0_18:1(	[M+Na]+	2	15	0	5	0	0	0	0	0	0	0	0	0	10	5	0	0	0	0	0	X	9	627.3876	5	187.88457	200.475798	138.470871	12429.49947	484.4069489	[459]
PC 15:0_18:1(	[M+Na]+	2	15	0	4	0	0	0	0	0	0	0	0	0	11	4	0	0	0	0	0	X	10	613.37195	5	214.1291	200.475798	138.470871	12429.49947	484.4069489	[460]
PC 15:0_18:1(	[M+Na]+	2	15	0	3	0	0	0	0	0	0	0	0	0	12	3	0	0	0	0	0	X	11	599.3563	5	134.27486	200.475798	138.470871	12429.49947	484.4069489	[461]
PC 15:0_18:1(	[M+Na]+	2	15	0	2	0	0	0	0	0	0	0	0	0	13	2	0	0	0	0	0	X	12	585.34064	5	93.26389	200.475798	138.470871	12429.49947	484.4069489	[462]
PC 15:0_18:1(	[M+Na]+	2	15	0	1	0	0	0	0	0	0	0	0	0	14	1	0	0	0	0	0	X	13	571.325	5	127.50757	200.475798	138.470871	12429.49947	484.4069489	[463]
PC 15:0_18:1(	[M+Na]+	2	15	0	14	0	0	0	0	0	0	0	0	0	1	14	0	1	0	0	1	X-H	14,59	754.53625	5	15.419261	200.475798	138.470871	12429.49947	484.4069489	[464 509]
PC 15:0_18:1(	[M+Na]+	2	15	0	13	0	0	0	0	0	0	0	0	0	2	13	0	1	0	0	1	X-H	15,60	740.5206	5	198.09793	200.475798	138.470871	12429.49947	484.4069489	[465 510]
PC 15:0_18:1(	[M+Na]+	2	15	0	12	0	0	0	0	0	0	0	0	0	3	12	0	1	0	0	1	X-H	16,61	726.505	5	34.14337	200.475798	138.470871	12429.49947	484.4069489	[466 511]
PC 15:0_18:1(	[M+Na]+	2	15	0	11	0	0	0	0	0	0	0	0	0	4	11	0	1	0	0	1	X-H	17,62	712.4893	5	19.783321	200.475798	138.470871	12429.49947	484.4069489	[467 512]
PC 15:0_18:1(	[M+Na]+	2	15	0	10	0	0	0	0	0	0	0	0	0	5	10	0	1	0	0	1	X-H	18,63	698.4737	5	15.419261	200.475798	138.470871	12429.49947	484.4069489	[468 513]
PC 15:0_18:1(	[M+Na]+	2	15	0	9	0	0	0	0	0	0	0	0	0	6	9	0	1	0	0	1	X-H	19,64	684.458	5	17.114552	200.475798	138.470871	12429.49947	484.4069489	[469 514]
PC 15:0_18:1(	[M+Na]+	2	15	0	8	0	0	0	0	0	0	0	0	0	7	8	0	1	0	0	1	X-H	20,65	670.4424	5	112.46323	200.475798	138.470871	12429.49947	484.4069489	[470 515]
PC 15:0_18:1(	[M+Na]+	2	15	0	7	0	0	0	0	0	0	0	0	0	8	7	0	1	0	0	1	X-H	21,66	656.4267	5	327.29495	200.475798	138.470871	12429.49947	484.4069489	[471 516]
PC 15:0_18:1(	[M+Na]+	2	15	0	6	0	0	0	0	0	0	0	0	0	9	6	0	1	1	0	0	X-H	22,50	642.4111	5	68.42294	200.475798	138.470871	12429.49947	484.4069489	[472 500]
PC 15:0_18:1(	[M+Na]+	2	15	0	5	0	0	0	0	0	0	0	0	0	10	5	0	1	0	1	0	X-H	23,85	628.39545	5	89.21259	200.475798	138.470871	12429.49947	484.4069489	[473 535]
PC 15:0_18:1(	[M+Na]+	2	15	0	4	0	0	0	0	0	0	0	0	0	11	4	0	1	0	1	0	X-H	24,86	614.37976	5	484.40695	200.475798	138.470871	12429.49947	484.4069489	[474 536]
PC 15:0_18:1(	[M+Na]+	2	15	0	3	0	0	0	0	0	0	0	0	0	12	3	0	1	0	1	0	X-H	25,87	600.36414	5	107.29666	200.475798	138.470871	12429.49947	484.4069489	[475 537]
PC 15:0_18:1(	[M+Na]+	2	15	0	2	0	0	0	0	0	0	0	0	0	13	2	0	1	0	1	0	X-H	26,88	586.34845	5	192.51895	200.475798	138.470871	12429.49947	484.4069489	[476 538]
PC 15:0_18:1(	[M+Na]+	2	15	0	1	0	0	0	0	0	0	0	0	0	14	1	0	1	0	1	0	X-H	27,89	572.3328	5	99.47438	200.475798	138.470871	12429.49947	484.4069489	[477 539]
PC 15:0_18:1(	[M+Na]+	2	15	0	14	0	0	0	0	0	0	0	0	0	1	14	0	1	0	1	0	X+H	28,76	752.5206	5	484.40695	200.475798	138.470871	12429.49947	484.4069489	[478 526]
PC 15:0_18:1(	[M+Na]+	2	15	0	13	0	0	0	0	0	0	0	0	0	2	13	0	1	0	1	0	X+H	29,77	738.505	5	162.58075	200.475798	138.470871	12429.49947	484.4069489	[479 527]
PC 15:0_18:1(	[M+Na]+	2	15	0	12	0	0	0	0	0	0	0	0	0	3	12	0	1	0	1	0	X+H	30,78	724.4893	5	343.22537	200.475798	138.470871	12429.49947	484.4069489	[480 528]
PC 15:0_18:1(	[M+Na]+	2	15	0	11	0	0	0	0	0	0	0	0	0	4	11	0	1	0	1	0	X+H	31,79	710.4737	5	390.47504	200.475798	138.470871	12429.49947	484.4069489	[481 529]
PC 15:0_18:1(	[M+Na]+	2	15	0	10	0	0	0	0	0	0	0	0	0	5	10	0	1	0	1	0	X+H	32,80	696.458	5	379.99203	200.475798	138.470871	12429.49947	484.4069489	[482 530]
PC 15:0_18:1(	[M+Na]+	2	15	0	9	0	0	0	0	0	0	0	0	0	6	9	0	1	0	1	0	X+H	33,81	682.4424	5	324.35583	200.475798	138.470871	12429.49947	484.4069489	[483 531]
PC 15:0_18:1(	[M+Na]+	2	15	0	8	0	0	0	0	0	0	0	0	0	7	8	0	1	0	1	0	X+H	34,82	668.4267	5	473.17566	200.475798	138.470871	12429.49947	484.4069489	[484 532]
PC 15:0_18:1(	[M+Na]+	2	15	0	7	0	0	0	0	0	0	0	0	0	8	7	0	1	0	1	0	X+H	35,83	654.4111	5	400.69186	200.475798	138.470871	12429.49947	484.4069489	[485 533]
PC 15:0_18:1(	[M+Na]+	2	15	0	6	0	0	0	0	0	0	0	0	0	9	6	0	0	0	0	0	X+H	36	640.39545	5	223.46967	200.475798	138.470871	12429.49947	484.4069489	[486]
PC 15:0_18:1(	[M+Na]+	2	15	0	5	0	0	0	0	0	0	0	0	0	10	5	0	0	0	0	0	X+H	37	626.37976	5	123.81663	200.475798	138.470871	12429.49947	484.4069489	[487]
PC 15:0_18:1(	[M+Na]+	2	15	0	4	0	0	0	0	0	0	0	0	0	11	4	0	0	0	0	0	X+H	38	612.36414	5	142.66689	200.475798	138.470871	12429.49947	484.4069489	[488]
PC 15:0_18:1(	[M+Na]+	2	15	0	3	0	0	0	0	0	0	0	0	0	12	3	0	0	0	0	0	X+H	39	598.34845	5	334.62576	200.475798	138.470871	12429.49947	484.4069489	[489]
PC 15:0_18:1(	[M+Na]+	2	15	0	2	0	0	0	0	0	0	0	0	0	13	2	0	0	0	0	0	X+H	40	584.3328	5	86.63219	200.475798	138.470871	12429.49947	484.4069489	[490]
PC 15:0_18:1(	[M+Na]+	2	15	0	1	0	0	0	0	0	0	0	0	0	14	1	0	0	0	0	0	X+H	41	570.3172	5	89.48101	200.475798	138.470871	12429.49947	484.4069489	[491]