

Supplementary Materials for paper:

A Multi-View Learning-Based Rule Extraction Algorithm For Accurate Hepatotoxicity Prediction

I. ADDITIONAL RESULTS

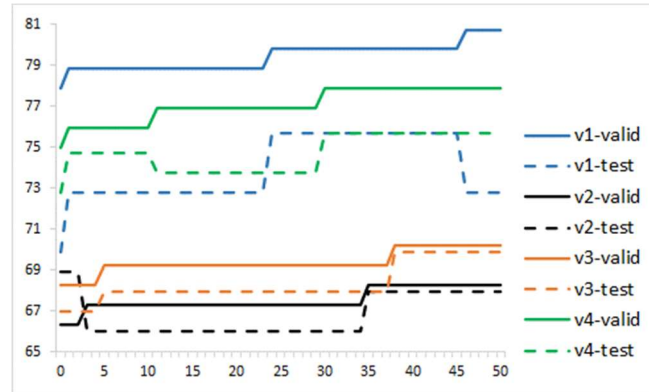


Fig. S1. The evolution curves of MVR-GA

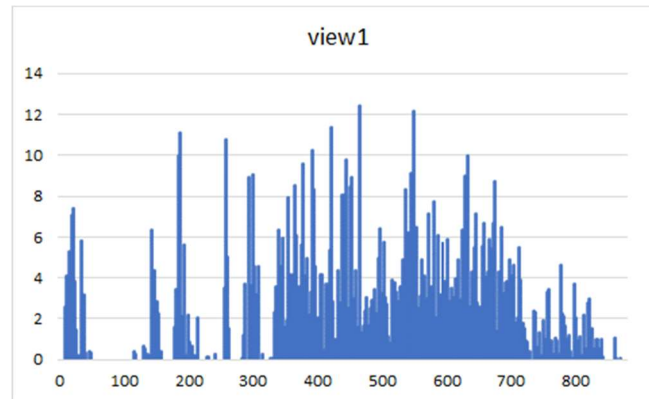


Fig. S2. The selection frequency of each feature on view1

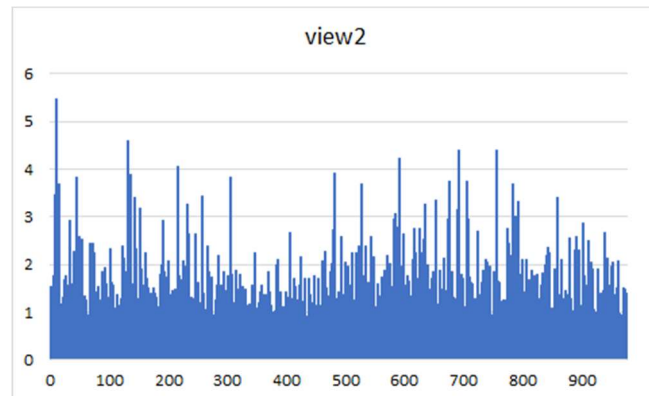


Fig. S3. The selection frequency of each feature on view2

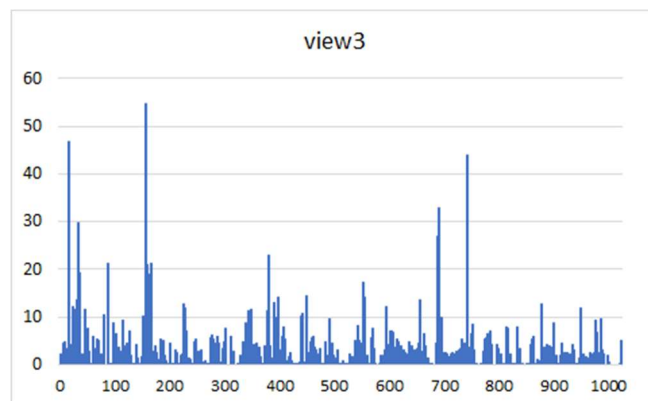


Fig. S4. The selection frequency of each feature on view3

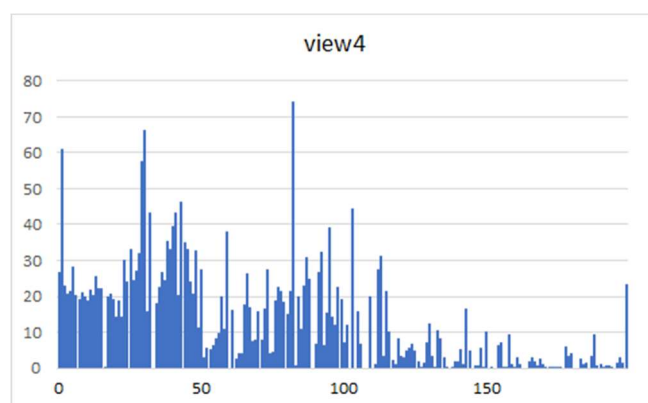


Fig. S5. The selection frequency of each feature on view4

TABLE SI. DATASETS OF DRUG INDUCED HEPATOTOXICITY AND BINARIZATION RULES OF LABELS

ID	Source	Type of Data	No. of Compound (Positive/Negative)	DILI Categories
1	Greene et al.,2010	Literature reviews and medical monographs	487(331/156)	HE, WE represented positives and NE represented negatives
2	Xu et al.,2015	Medical monographs and FDA-approved drug labeling	475(236/239)	Authors definition
3	Mulliner et al.,2017	Clinical data and drug labeling	1370(932/438)	Authors definition
4	DILrank	Drug labeling and clinical data	504(192/312)	Most concern as 1; no concern as 0
5	Livertox[1][1]	Scientific literature and public database	343(119/224)	Categories A and B were combined into positives, and Category E was considered as negatives
6	LTKB	FDA-approved drug labeling	195(113/82)	Most concern as 1;no concern as 0

TABLE SII. FEATURES OF HEPATOTOXICITY DATASET

Feature	View	Dimensionality	Type
finger print (fp)	V1	881	Boolean type
expression	V2	978	Continuity type
target	V3	1023	Boolean type
phychem	V4	200	Continuity type

TABLE SIII. ACCURACY AND COVERAGE OF OUTSTANDING RULES

No.	Accuracy	Coverage
Rule 536	71.43	74.04
Rule 561	72.00	72.00
Rule 690	70.00	67.31
Rule 805	67.94	75.00
Rule 815	75.71	67.31

TABLE SIV. MODEL COMPARISON

View	MVR-GA			Random Forest		
	Rule Num	Rule Length	Feature Num	Rule Num	Rule Length	Feature Num
V ₁	249.18	5.29	355.04	10,469.32	10.89	536.76
V ₂	349.42	3.64	484.94	4,521.26	8.55	964.44
V ₃	246.06	8.93	320.96	15,513.64	54.55	633.24
V ₄	410.32	6.22	149.58	5,742.06	8.41	169.56
V ₁ +V ₂	598.60	4.33	839.98	4,722.38	8.49	1,165.72
V ₁ +V ₃	495.24	7.10	676.00	10,927.86	12.18	963.34
V ₁ +V ₄	659.50	5.87	504.62	6,665.82	8.85	623.54
V ₂ +V ₃	595.48	5.83	805.90	4,759.88	8.77	1,057.54
V ₂ +V ₄	759.74	5.03	634.52	4,370.24	8.12	1,078.68
V ₃ +V ₄	656.38	7.23	470.54	7,236.40	9.67	487.50
V ₁ +V ₂ +V ₃	844.66	5.67	1160.94	4,892.38	8.65	1,251.74
V ₁ +V ₂ +V ₄	1008.92	5.10	989.56	4,531.16	8.17	1,252.22
V ₁ +V ₃ +V ₄	905.56	6.70	825.58	7,184.32	9.43	865.44
V ₂ +V ₃ +V ₄	1005.80	5.99	955.48	4,588.86	8.35	1,166.92
V ₁ +V ₂ +V ₃ +V ₄	1254.98	5.85	1310.52	4,679.52	8.34	1,329.38

TABLE SV. IMPORTANT FEATURES

View	Feature_index
V ₁	464, 549, 420, 186, 257, 392, 633, 185, 443, 376, 545, 299, 628, 451, 293, 391, 673, 365, 449, 393
V ₂	11, 131, 755, 692, 591, 216, 480, 135, 304, 45, 706, 675, 527, 784, 14, 7, 257, 142, 859, 653
V ₃	156, 16, 741, 689, 157, 34, 688, 379, 165, 87, 158, 159, 166, 35, 164, 552, 33, 32, 162, 449
V ₄	82, 30, 1, 29, 43, 103, 41, 32, 40, 95, 59, 38, 44, 39, 45, 25, 48, 92, 28, 113

TABLE SVI. MVR-GA IMPORTANT FEATURES TOP 200

Index	V1 (fingerprint)		V2 (expression)		V3 (target)		V4 (phychem)	
	feature name	frequency	feature name	frequency	feature name	frequency	feature name	frequency
1	464	12.4	3108	5.48	CHRM1	54.74	82	74.2
2	549	12.14	5058	4.6	ACHE	46.76	30	66.1
3	420	11.38	29978	4.4	SCN10A	44.04	1	61
4	186	11.14	10318	4.4	PTGS2	32.82	29	57.4
5	257	10.78	55011	4.22	CHRM2	32.2	43	46.18
6	392	10.24	5440	4.06	ADRB2	29.76	103	44.28

Index	V1 (fingerprint)		V2 (expression)		V3 (target)		V4 (phychem)	
	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>
7	633	10	3482	3.92	PTGS1	26.88	41	43.1
8	185	9.98	5223	3.88	HRH1	22.96	32	43.04
9	443	9.8	7168	3.84	CHRNA7	21.16	40	39.48
10	376	9.6	695	3.82	BCHE	21.12	95	39.02
11	545	9.1	9924	3.76	CHRM3	20.84	59	37.86
12	299	9.06	54807	3.74	CHRM4	20.24	38	35.24
13	628	9	9961	3.7	CHRNA2	20.22	44	34.9
14	451	8.96	4846	3.68	ADRB3	19.28	39	33.16
15	293	8.94	1633	3.68	CHRNA4	19.06	45	32.98
16	391	8.76	5898	3.46	NR1I2	17.32	25	32.9
17	673	8.72	27032	3.44	ADRB1	17.16	48	32.8
18	365	8.56	652	3.42	ADRA2C	16.18	92	32.08
19	449	8.44	5982	3.4	CHRNA2	15.24	28	31.86
20	393	8.36	10245	3.34	KCNJ8	14.54	113	31.1
21	535	8.36	23223	3.32	HTR2C	14.24	87	30.6
22	437	8.06	7849	3.28	NR3C1	14.06	23	30.18
23	439	8.06	1452	3.26	ADRA2A	13.54	5	27.96
24	353	7.92	9128	3.26	PPARG	13.46	50	27.54
25	579	7.76	6464	3.18	ADRA2B	13.14	112	27.5
26	377	7.54	23131	3.16	HTR1A	13.08	73	27.32
27	20	7.44	5438	3.06	TOP2A	12.82	27	26.86
28	571	7.16	55604	3.02	DRD1	12.8	0	26.64
29	645	7.12	5255	2.98	ADORA2A	12.26	91	26.6
30	19	7.06	89910	2.96	PCNA	12.12	36	26.58
31	657	6.7	2778	2.96	DRD2	11.96	66	26.2
32	672	6.7	5480	2.96	gyrA	11.94	13	25.32
33	553	6.5	5290	2.94	GRIN3A	11.72	88	24.86
34	684	6.46	1029	2.94	PPARA	11.56	26	24.32
35	495	6.44	960	2.86	AKR1C1	11.5	37	24.2
36	143	6.38	9124	2.82	ADRA1A	11.48	24	23.94
37	624	6.36	55111	2.8	HPGDS	11.4	46	23.84
38	339	6.34	22841	2.76	GRIN1	11.24	199	23.28
39	540	6.22	10270	2.76	KCNH7	10.74	86	22.88
40	366	6.08	5747	2.76	GRIN2D	10.62	2	22.72
41	585	6.08	1994	2.72	TOP1	10.52	77	22.58
42	345	5.94	7866	2.7	ATP6V1A	10.4	35	22.42
43	665	5.9	5827	2.68	ADRA1B	10.34	98	22.28
44	600	5.86	5427	2.68	KCNH2	10.18	14	21.96
45	33	5.84	8480	2.66	CDIPT	10.06	15	21.92
46	502	5.78	4836	2.64	KCNH6	9.96	11	21.82
47	346	5.76	64746	2.64	NR1I3	9.96	4	21.48
48	580	5.76	22827	2.64	PTPRS	9.94	115	21.26
49	671	5.72	1956	2.6	HTR2A	9.82	78	21.18

Index	V1 (fingerprint)		V2 (expression)		V3 (target)		V4 (phychem)	
	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>
50	593	5.7	226	2.58	MAOA	9.58	81	21.16
51	192	5.64	1429	2.58	plcA	9.54	8	20.88
52	375	5.64	8204	2.58	CACNA1C	9.36	47	20.74
53	442	5.56	5287	2.56	parC	9.28	3	20.6
54	654	5.54	1050	2.52	ADRA1D	8.96	18	20.56
55	643	5.52	1738	2.52	CA1	8.76	12	20.34
56	712	5.48	23597	2.5	GRIA2	8.74	42	20.34
57	546	5.46	30849	2.46	TYR	8.66	6	20.26
58	667	5.46	4791	2.46	SCN5A	8.54	84	19.98
59	418	5.36	22905	2.46	NOS2	8.32	9	19.86
60	15	5.3	1870	2.4	SLC6A2	8.06	109	19.76
61	259	5.06	9813	2.4	HTR7	8.02	17	19.72
62	383	4.96	22823	2.4	SOD1	7.86	57	19.7
63	493	4.96	11007	2.38	MAOB	7.84	19	19.22
64	531	4.92	4651	2.38	OPRM1	7.78	99	19.02
65	697	4.9	7511	2.36	ALB	7.64	7	18.9
66	560	4.88	3486	2.34	SLC6A4	7.54	21	18.86
67	617	4.88	7099	2.34	GBA	7.5	76	18.68
68	532	4.78	27244	2.32	ADORA1	7.48	10	18.54
69	776	4.66	998	2.3	PDE4C	7.1	79	18.32
70	704	4.62	93594	2.3	PDE4A	7.08	34	17.92
71	308	4.6	51569	2.3	CHRNA4	7.06	65	17.46
72	395	4.58	3091	2.28	SLC15A1	7.04	67	16.82
73	301	4.56	8974	2.28	CACNB2	7.02	143	16.3
74	341	4.54	5715	2.26	CACNA1D	6.96	72	16.28
75	342	4.4	51422	2.26	DRD5	6.94	61	16.08
76	146	4.38	8900	2.26	PDE4D	6.92	31	15.74
77	459	4.38	6284	2.26	pbp3	6.82	105	15.68
78	432	4.36	2356	2.26	DRD3	6.76	70	15.62
79	681	4.34	3725	2.24	SCN4A	6.62	94	15.24
80	663	4.32	10007	2.24	PRKAA1	6.6	80	14.8
81	638	4.3	51001	2.24	GRIN3B	6.54	20	14.34
82	538	4.22	54512	2.24	SLC12A3	6.5	96	14.08
83	358	4.18	1829	2.24	ADORA2B	6.48	22	14
84	374	4.18	3028	2.2	CA2	6.42	130	12.1
85	403	4.18	6253	2.2	HTR1B	6.4	101	12.06
86	362	4.16	51021	2.2	CHRNA3	6.38	97	11.96
87	406	4.16	2887	2.2	GABRA3	6.34	49	11.32
88	11	4.14	8444	2.18	SLC6A3	6.08	85	10.82
89	379	4.14	5873	2.16	ALOX5	6	58	10.58
90	380	4.14	813	2.16	GABRA4	5.96	133	10.36
91	699	4.14	4775	2.14	KCNN3	5.94	116	9.86
92	566	4.12	11011	2.14	GABRG2	5.92	150	9.84

Index	V1 (fingerprint)		V2 (expression)		V3 (target)		V4 (phychem)	
	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>
93	659	4.06	11031	2.14	HTR6	5.86	56	9.62
94	613	4	10892	2.12	GLRA1	5.82	188	9.28
95	340	3.98	2353	2.12	THRB	5.82	158	9.06
96	453	3.94	55129	2.1	CHRM5	5.76	119	8.18
97	514	3.94	10150	2.1	KCNMB3	5.66	134	8.12
98	550	3.92	9710	2.1	KCNN1	5.64	55	7.94
99	713	3.9	11000	2.1	SLC12A1	5.62	69	7.82
100	23	3.86	4482	2.1	GABRA2	5.58	71	7.58
101	692	3.86	11182	2.08	OPRD1	5.52	68	7.3
102	623	3.84	23378	2.08	RXRA	5.5	100	7.08
103	596	3.82	25874	2.08	THRA	5.5	129	7.08
104	10	3.78	23325	2.08	CRYZ	5.44	155	6.84
105	519	3.76	11142	2.08	HTR2B	5.42	124	6.74
106	691	3.76	3930	2.04	ESR1	5.4	106	6.5
107	414	3.74	9112	2.04	IKBKB	5.4	90	6.46
108	797	3.74	57178	2.04	AGTR1	5.38	93	6.34
109	412	3.72	9842	2.04	SIGMAR1	5.38	154	6.26
110	287	3.68	23670	2.02	GABRA6	5.34	54	6.18
111	405	3.66	7158	2.02	PDK1	5.34	178	6.04
112	294	3.62	9270	2.02	GRIN2A	5.28	123	5.44
113	334	3.6	51160	2.02	AR	5.24	148	5.4
114	364	3.58	23013	2.02	GABRA1	5.24	52	5.3
115	370	3.58	2264	2	yedY	5.18	141	5.24
116	440	3.56	55746	2	HTR3A	5.08	53	5.02
117	528	3.56	51375	2	ARG1	5.06	125	4.7
118	576	3.56	57192	2	CTNNB1	5.04	144	4.64
119	256	3.52	55012	2	NPY	5.02	122	4.56
120	714	3.52	664	1.98	HRH2	4.98	75	4.44
121	607	3.5	10813	1.98	NOLC1	4.98	64	3.98
122	145	3.48	51170	1.98	KCNN2	4.92	63	3.96
123	335	3.46	9467	1.96	KCNMB2	4.86	180	3.9
124	487	3.44	10776	1.96	ABCC8	4.84	74	3.84
125	758	3.44	2542	1.96	GAST	4.82	179	3.36
126	180	3.42	1019	1.94	KCNMB1	4.82	187	3.12
127	450	3.42	3162	1.94	ERG1	4.8	120	3.1
128	598	3.4	1398	1.94	PDXK	4.8	114	3.08
129	614	3.4	1454	1.92	GPR55	4.76	131	3.08
130	656	3.4	2770	1.92	LPA	4.74	135	3
131	601	3.38	1153	1.92	PLAT	4.7	121	2.94
132	674	3.34	22934	1.92	PTGER1	4.66	166	2.9
133	521	3.32	10557	1.92	bla	4.66	161	2.88
134	389	3.3	7852	1.9	NQO1	4.64	197	2.84
135	755	3.3	6657	1.88	HBA1	4.6	51	2.7

Index	V1 (fingerprint)		V2 (expression)		V3 (target)		V4 (phychem)	
	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>
136	597	3.28	55556	1.88	POMC	4.58	183	2.6
137	337	3.26	57147	1.88	OPRK1	4.54	62	2.54
138	689	3.24	50813	1.88	GABRA5	4.5	169	2.32
139	390	3.22	50810	1.88	GABRB2	4.5	117	2
140	501	3.22	8996	1.88	MAPK1	4.5	167	1.86
141	655	3.22	2064	1.88	CYP2A6	4.48	140	1.84
142	37	3.2	8727	1.88	KCNMB4	4.48	165	1.82
143	591	3.2	4893	1.86	CACNA1S	4.46	126	1.76
144	698	3.2	5696	1.86	RYR1	4.46	139	1.64
145	396	3.18	55847	1.86	GABRQ	4.44	198	1.46
146	548	3.18	8985	1.86	SI00P	4.42	185	1.44
147	305	3.16	9143	1.86	ABCC2	4.4	196	1.32
148	639	3.16	847	1.86	TRPV3	4.36	128	1.22
149	572	3.14	10190	1.86	fabI	4.36	111	1.08
150	357	3.12	1385	1.86	CALM1	4.32	162	1.08
151	452	3.1	23224	1.86	SLC18A2	4.3	190	1
152	558	3.1	56924	1.84	SLC15A2	4.28	142	0.9
153	577	3.1	2109	1.84	ANXA1	4.24	118	0.88
154	615	3.1	94239	1.84	GPT2	4.24	184	0.88
155	476	3.06	54915	1.84	PDE3A	4.22	159	0.84
156	611	3.06	80758	1.84	GSTZ1	4.2	170	0.8
157	181	3.02	831	1.84	DRD4	4.18	168	0.7
158	504	3.02	1534	1.82	TGM2	4.18	146	0.66
159	454	2.98	5321	1.8	GSS	4.14	189	0.58
160	696	2.98	5347	1.8	ACY1	4.12	147	0.54
161	255	2.96	2523	1.8	GABRG1	4.1	193	0.46
162	569	2.96	392	1.8	GRIN2C	4.1	192	0.44
163	821	2.96	79170	1.8	GRIN2B	4.06	83	0.42
164	621	2.94	1052	1.8	PPARD	4.06	132	0.34
165	144	2.92	102	1.8	TOP1MT	4.06	157	0.32
166	484	2.92	4690	1.78	HTR1F	4.04	191	0.32
167	680	2.92	9533	1.78	PRKAB1	4.04	16	0.3
168	573	2.9	7398	1.78	NQO2	3.98	194	0.28
169	423	2.88	4016	1.78	CNR1	3.96	127	0.24
170	150	2.86	5529	1.78	ACE	3.94	136	0.16
171	604	2.86	7088	1.78	CACNA1H	3.92	171	0.16
172	523	2.84	51719	1.78	CACNA2D1	3.92	160	0.12
173	641	2.84	3553	1.76	GABRP	3.92	172	0.12
174	381	2.82	5289	1.76	HRH4	3.92	173	0.08
175	483	2.82	5359	1.76	TK	3.9	176	0.06
176	818	2.82	572	1.76	PGR	3.88	149	0.04
177	373	2.8	54442	1.76	HLCS	3.86	152	0.04
178	435	2.8	10921	1.76	PLG	3.86	138	0.02

Index	V1 (fingerprint)		V2 (expression)		V3 (target)		V4 (phychem)	
	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>	<i>feature name</i>	<i>frequency</i>
179	338	2.78	9702	1.76	PTPRE	3.86	156	0.02
180	646	2.78	868	1.74	RYR2	3.86	174	0.02
181	457	2.74	9517	1.74	TUBB	3.86	175	0.02
182	594	2.74	93487	1.74	GABRE	3.84	33	0
183	622	2.74	57149	1.74	PTPN4	3.82	60	0
184	503	2.7	6909	1.74	PGD	3.78	89	0
185	507	2.7	56889	1.74	CA6	3.76	102	0
186	527	2.7	10320	1.72	ABCC1	3.74	104	0
187	574	2.7	6622	1.72	PDE7B	3.74	107	0
188	619	2.7	6919	1.72	HTR1D	3.72	108	0
189	711	2.7	10058	1.72	KCNMA1	3.7	110	0
190	590	2.68	9688	1.72	TUBG1	3.7	137	0
191	382	2.66	8678	1.72	HDAC2	3.68	145	0
192	421	2.64	64422	1.72	TUBE1	3.66	151	0
193	438	2.64	128	1.72	GABRB1	3.64	153	0
194	682	2.64	5867	1.72	KCNN4	3.64	163	0
195	637	2.62	375346	1.72	TPK1	3.64	164	0
196	651	2.62	9805	1.7	CACNB1	3.62	177	0
197	695	2.62	57048	1.7	CHRNA10	3.62	181	0
198	626	2.6	63933	1.7	GRIK2	3.62	182	0
199	9	2.58	3315	1.68	GABRD	3.6	186	0
200	536	2.58	207	1.68	GABRB3	3.56	195	0

Rule 536

```

1.  if X['1626'] > -0.6336418986320496:
2.      if X['889'] > -0.29730215668678284:
3.          if X['1676'] <= 0.7438333928585052:
4.              if X['1587'] > -0.34194841980934143:
5.                  return 1

```

Rule 561

```

1.  if X['1721'] > -0.7491983473300934:
2.      if X['1445'] <= 0.18207036703824997:
3.          if X['1692'] > -0.6236386001110077:
4.              if X['1558'] > -0.5988142788410187:
5.                  return 1

```

Rule 690

```

1.  if X['1225'] <= 0.06973003223538399:
2.      if X['941'] > -0.18589357286691666:
3.          if X['1430'] > -0.5958110690116882:
4.              if X['1479'] > -0.4400264173746109:
5.                  return 1

```

Rule 805

```
1.  if X['1828'] <= 0.6469834744930267:
2.      if X['1154'] <= 0.4267232120037079:
3.          if X['1044'] > -0.36726604402065277:
4.              if X['927'] <=
0.5427587032318115:
5.                  return 1
```

Rule 815

```
1.  if X['892'] > -0.23688504099845886:
2.      if X['942'] <= 0.2453337162733078:
3.          if X['1573'] <= 0.7088802456855774:
4.              if X['1853'] > -
0.1360762044787407:
5.                  return 1
```
