

Actividad: Detección de anomalías (“clásico” o “serie temporal”):

En mi caso el problema que vamos a analizar para detectar anomalías es detectar anomalías en la forma de conducir de un usuario, para así poder reducir el riesgo de accidentes o identificar posibles hábitos peligrosos en su conducción.

Las características que tendremos en cuenta para detectar anomalías son:

- Tipo de Vida (TipoDeVia), variable categórica.
- Velocidad, en kilómetros hora (Velocidad_km_h).
- Aceleración longitudinal, en metros segundos al cuadrado, para detectar frenazos y acelerones bruscos (AceleracionLongitudinal_m_s2).
- Aceleración lateral, en metros segundos al cuadrado, para detectar giros agresivos (AceleracionLateral_m_s2).
- Sacudidas, en metros segundos al cubo, para cuanto de rápido se acelera o se frena (Jerk_m_s3).
- Velocidad de giro angular, en grados segundo, para detectar giros muy bruscos (VelocidadGiroYaw_deg_s).
- Porcentaje de freno, en porcentaje, para medir la intensidad de frenado (PorcentajeFreno).
- Porcentaje de aceleración, en porcentaje, para medir la posición del acelerador (PorcentajeAcelerador).
- Precisión estimada del GPS, en metros, para corroborar si los datos anteriores son fiables o no (PrecisionGPS_m).

Una vez generado el programa capaz de la detección de anomalías, usando el csv, que se encuentra en el proyecto, he obtenido los siguientes resultados.

Threshold 0,30: 36 anomalías detectadas

Threshold 0,45: 15 anomalías detectadas

Threshold 0,60: 7 anomalías detectadas

Threshold 0,50: 13 anomalías detectadas






AU ROC: 0,9891






























DR at FP=5: 1,0000

















Id TipoDeVia Velocidad_km_h AceleracionLongitudinal_m_s2

AceleracionLateral_m_s2 Jerk_m_s3 VelocidadGiroYaw_deg_s

PorcentajeFreno PorcentajeAcelerador PrecisionGPS_m Anomaly Score

1	urbana	39,2	-0,6	-1,1	0,3	12,4	13,0	15,0	1,1		0,4901
2	autopista	37,5	0,1	-0,5	3,0	1,0	30,0	24,0	5,0		0,5137
3	carretera	76,9	2,9	-0,2	0,2	15,7	40,0	43,0	3,6		0,2581
4	autopista	110,5	1,2	-0,7	1,8	4,5	18,0	49,0	4,2		0,2054
5	urbana	95,0	-6,2	0,6	9,8	12,0	96,0	5,0	5,8		0,7953

6	carretera	60,6	-0,4	-1,4	1,4	0,4	11,0	22,0	5,1		0,3098
7	urbana	104,4	-0,8	0,0	0,1	7,0	22,0	82,0	3,4		0,1717
8	urbana	74,3	0,6	-1,6	2,1	6,6	16,0	71,0	4,7		0,4464
9	carretera	71,8	0,7	-0,2	0,2	6,9	0,0	31,0	1,4		0,6739
10	autopista	74,5	-1,5	0,6	0,2	10,6	21,0	0,0	4,2		0,5077
11	urbana	67,6	0,8	1,0	1,0	8,4	29,0	34,0	3,7		0,3479
12	carretera	58,1	-0,7	1,1	1,5	9,8	15,0	73,0	3,2		0,4168
13	urbana	101,4	-0,8	-0,7	1,1	14,7	12,0	26,0	2,4		0,3607
14	carretera	119,8	0,7	0,2	1,7	14,7	5,0	6,0	4,2		0,3975
15	urbana	94,4	-1,4	-0,3	0,6	7,4	0,0	20,0	1,0		0,4269
16	urbana	122,8	0,1	-0,5	0,7	12,8	16,0	1,0	2,9		0,5661
17	autopista	123,4	0,5	0,6	2,8	3,0	37,0	18,0	1,9		0,5489
18	autopista	117,2	0,2	1,2	2,1	12,2	18,0	30,0	3,8		0,1690
19	urbana	82,8	-1,3	0,9	1,8	8,6	4,0	38,0	3,0		0,2960
20	urbana	65,0	0,3	4,3	6,9	88,0	22,0	40,0	1,2		0,6523
21	urbana	63,1	-1,3	0,4	1,7	16,8	8,0	23,0	3,3		0,2635
22	carretera	54,0	0,4	1,0	0,8	8,4	11,0	21,0	4,8		0,3821
23	autopista	88,0	1,5	-0,5	1,2	5,0	14,0	55,0	2,1		0,2786
24	carretera	120,0	-0,9	0,3	0,5	12,0	30,0	70,0	3,5		0,4448
25	carretera	64,6	-0,6	-1,0	1,1	5,0	9,0	25,0	4,2		0,1682
26	carretera	48,8	2,3	-0,7	1,8	11,8	22,0	0,0	1,0		0,2768
27	urbana	97,0	-0,3	-0,7	1,4	7,9	26,0	46,0	2,5		0,3020
28	carretera	152,0	1,1	0,4	2,2	10,0	5,0	82,0	1,4		0,7288
29	carretera	72,0	1,0	-0,1	1,1	11,0	7,0	37,0	3,9		0,3898
30	autopista	85,5	-0,3	0,1	0,4	9,2	12,0	50,0	2,7		0,3592
31	urbana	97,1	0,1	-0,9	0,5	11,6	25,0	35,0	1,8		0,3677
32	autopista	66,2	1,1	-0,1	1,3	13,6	17,0	14,0	3,0		0,3616
33	urbana	76,0	0,8	0,9	1,9	4,6	0,0	9,0	1,3		0,6120
34	carretera	83,2	0,6	0,5	0,7	5,4	32,0	56,0	3,5		0,3394

35	carretera	88,7	-0,7	0,3	1,3	12,3	14,0	40,0	2,9		
		0,2752									
36	urbana	55,4	0,4	-1,5	2,1	9,8	45,0	20,0	4,1		0,5076
37	autopista	112,3	1,2	0,2	1,0	3,2	10,0	60,0	2,4		
		0,3520									
38	carretera	78,5	-1,1	0,4	0,5	12,0	22,0	0,0	1,0		
		0,5874									
39	urbana	49,6	0,3	-0,9	1,2	7,5	15,0	25,0	3,3		0,2657
40	carretera	91,2	0,8	0,2	1,5	6,0	18,0	45,0	2,8		
		0,3743									
41	autopista	105,5	-0,5	0,1	0,3	4,0	12,0	55,0	2,6		
		0,2602									
42	urbana	59,4	-2,0	-1,7	2,2	7,5	1,0	8,0	2,3		0,1378
43	urbana	64,6	0,3	-1,2	2,2	10,8	23,0	34,0	4,4		0,3488
44	carretera	93,3	-1,4	0,7	1,0	1,4	1,0	32,0	4,0		0,3804
45	urbana	45,0	-5,7	0,2	8,4	8,0	93,0	2,0	4,0		0,7947
46	carretera	82,0	0,6	1,0	1,5	13,3	16,0	33,0	2,9		
		0,1564									
47	carretera	79,9	-0,6	0,2	1,2	11,0	39,0	38,0	2,9		
		0,6936									
48	carretera	82,2	0,8	-0,7	0,2	7,8	5,0	49,0	4,7		0,4867
49	autopista	69,8	-0,4	1,0	2,3	15,7	1,0	19,0	6,8		
		0,4009									
50	urbana	68,6	-1,1	-0,7	1,3	2,3	10,0	30,0	2,3		0,3143

De los cuales podemos decir que el modelo es capaz de separar muy bien entre datos que son normales de aquellos que podemos considerar como anomalías, AU ROC: 0,9891, al utilizar un rank de 5 el modelo es capaz de recuperar todas las anomalías reales, y en cuanto a los diferentes umbrales cada vez que se aumenta el valor del Threshold se reduce el número de anomalías.