



THE
DEVELOPER'S
CONFERENCE



Detecção de anomalias em sensor veicular a partir de classificadores one-class

Eronides da Silva Neto

efsn@cesar.org.br

- **Introdução**
- **O sistema de detecção de anomalias veiculares**
- **Experimentos realizados**
- **Resultados**





Engenheiro da Computação no C.E.S.A.R.



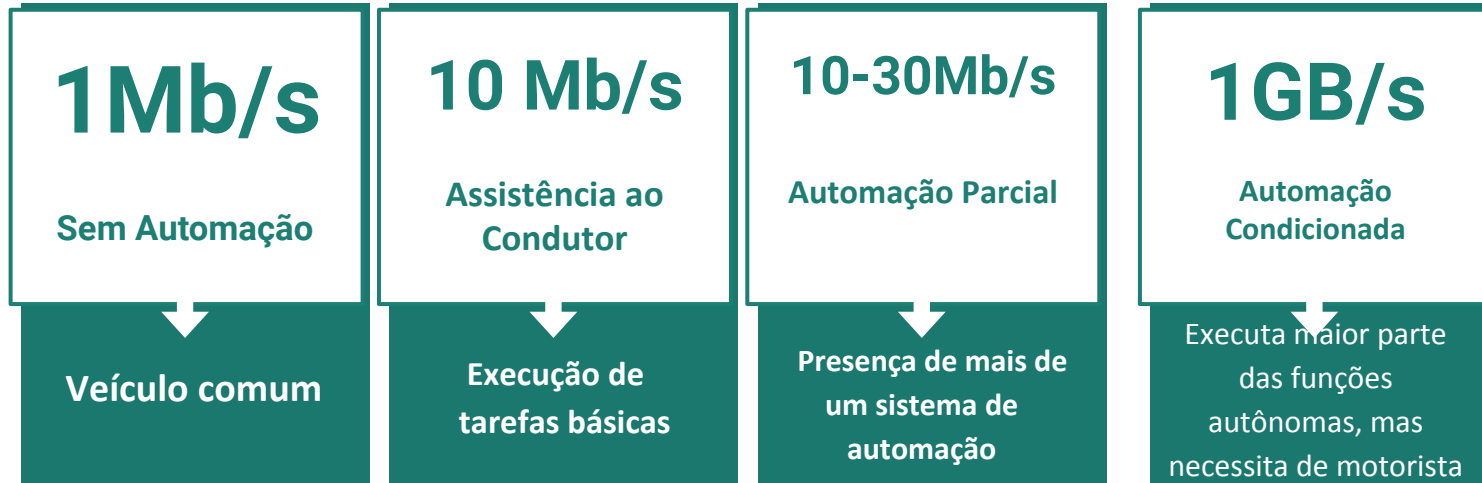
- **Mestre em Ciência da Computação (CIn-UFPE)**
- **Engenheiro Eletrônico (UFPE)**



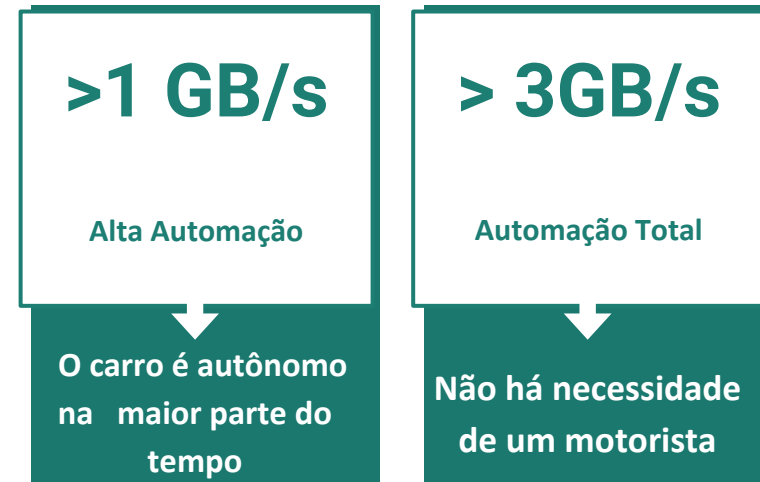
Open-source hardware enthusiast



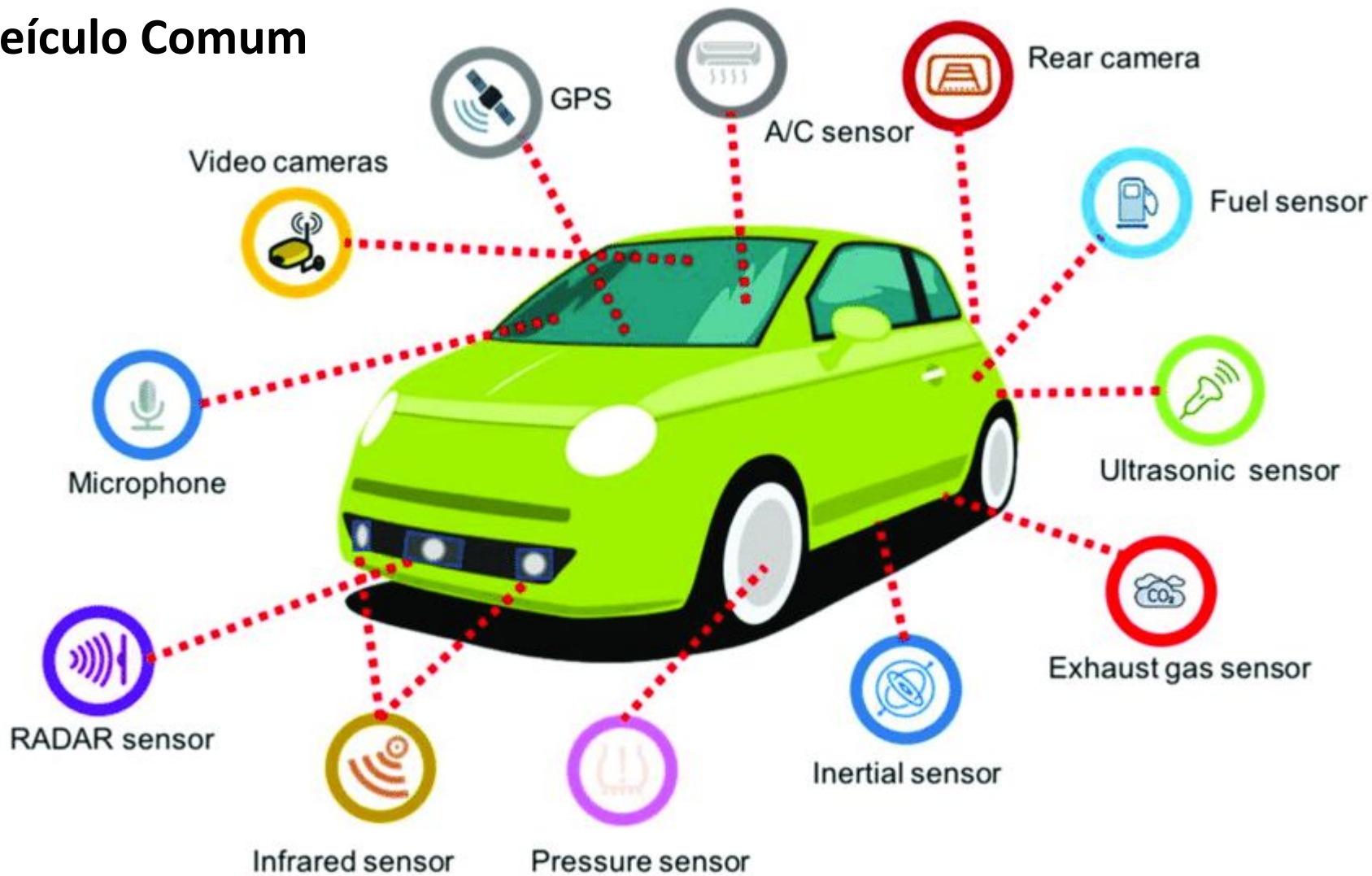
**Quantos MB de dados produzimos
diariamente em um carro?**



De onde vem essa quantidade de dados toda?



Veículo Comum



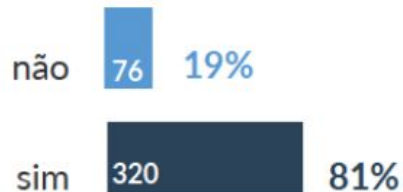
Fonte: Sensor Technologies for Intelligent Transportation Systems

O que fazer então com dados de um veículo comum?



- O setor de telemetria veicular possui um mercado estimado em **45 bilhões** de dólares neste ano (WAHLSTRÖM; SKOG; HÄN-DEL, 2017).
- **Pesquisa: Como é a experiência do consumidor ao usar um carro conectado?**

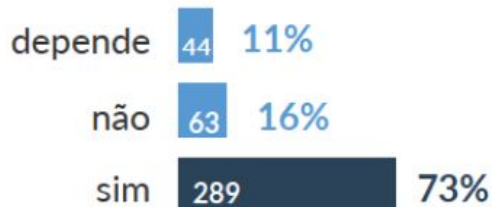
acha importante avaliar parâmetros?



gostaria de ter informações sobre manutenção?



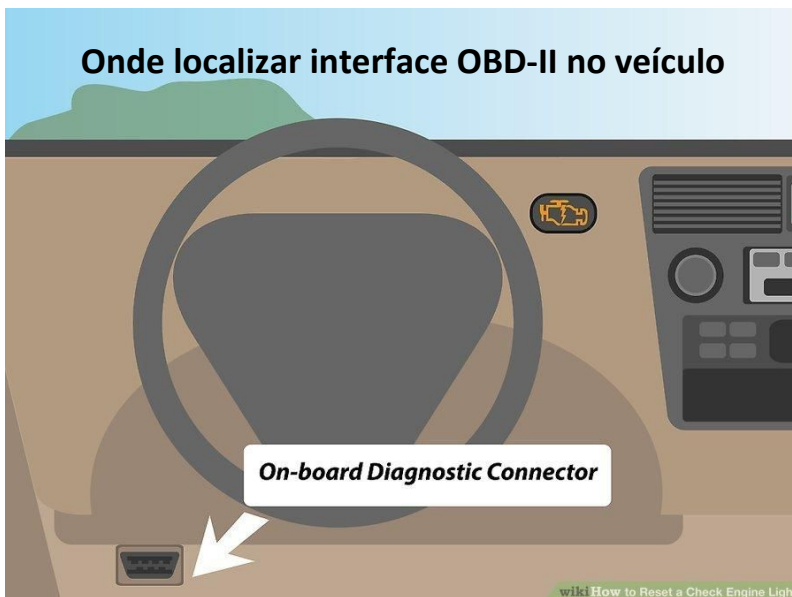
permitiria a coleta de dados?



qual melhor forma para coleta de dados?



- Coleta de 27 parâmetros de um veículo Toyota Etios a uma taxa de uma amostra por segundo (1Hz) a partir da interface OBD-II;
- Uso da plataforma open-source hardware Particle Electron com a placa Carloop.



Hardware utilizado



- Criação de uma base de dados com dois modos veiculares. Modo em movimento possui **quatro cenários distintos**, considerando diferentes tipos de trajetória e o modo IDLE representa o veículo parado.
- Ciclo veicular (*trip*): uma trajetória completa, para ambos modos.
- Trajeto em Áreas Urbanas
- Trajeto em Vias Expressas
- Trajeto em Universidade
- Trajeto Fixo

Modo de operação	Trajeto	Duração total (s)	Distância percorrida	Número de ciclos (<i>trips</i>)
<i>Idle</i>	-	35309	-	47
Movimento	Áreas urbanas	29958	717 km	12
Movimento	Via expressa	19493	458 km	13
Movimento	Universidade	15873	387 km	18
Movimento	Fixo	12561	1104 km	37

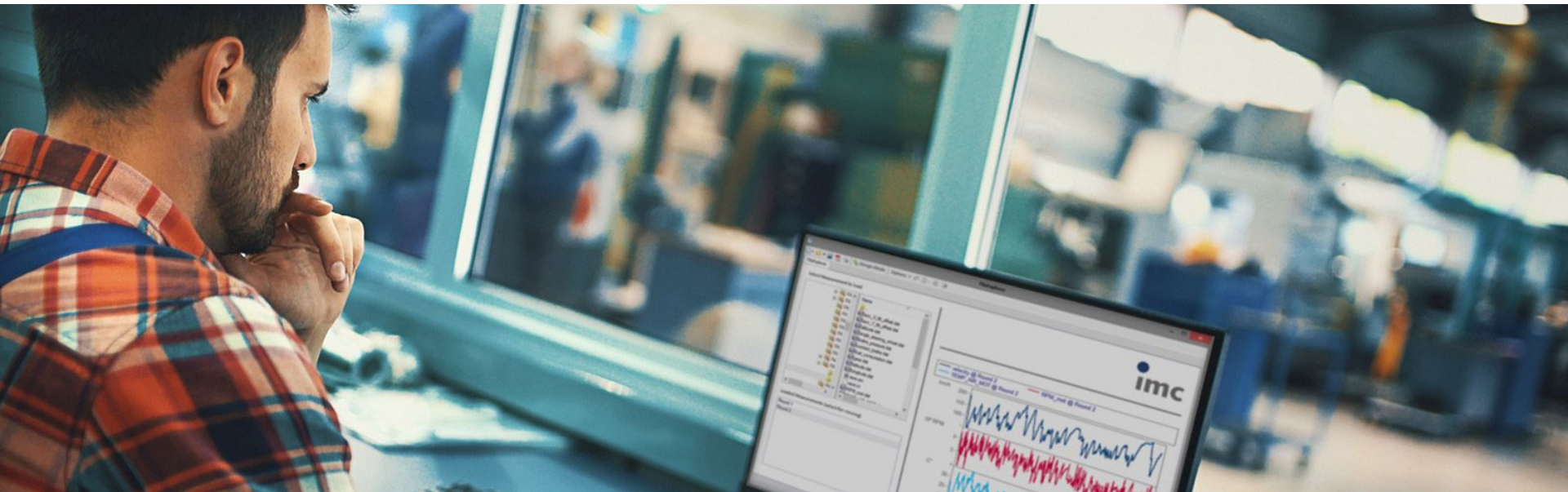
Dados disponíveis em github.com/eron93br/carOBD

Mas quais foram os dados obtidos?

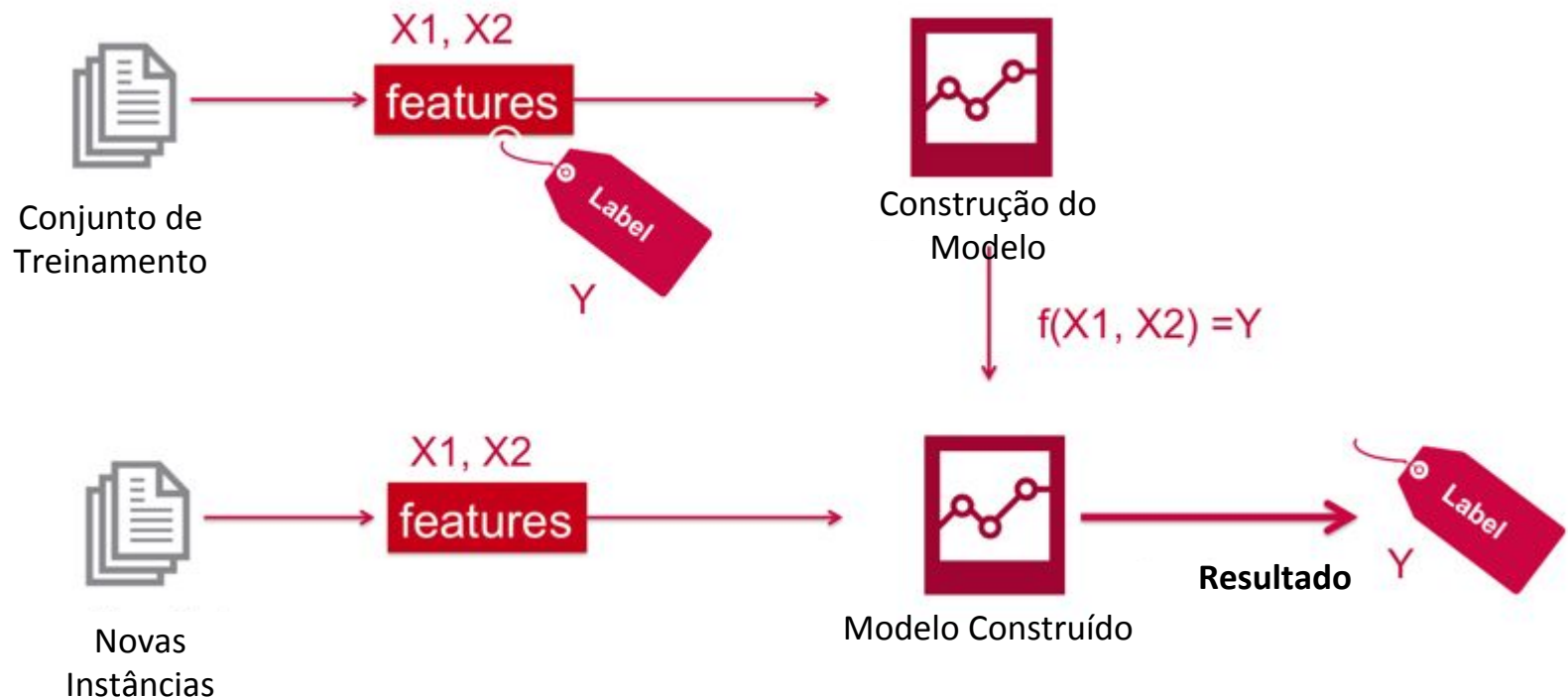
- Alguns dos dados obtidos eram:
 - Carga do Motor
 - Velocidade
 - Rotações por Minuto (RPM)
- Entre os problemas mais comuns enfrentados pelos motoristas está o superaquecimento do motor.



Um sistema para detecção de anomalias no sensor do líquido de arrefecimento do motor?

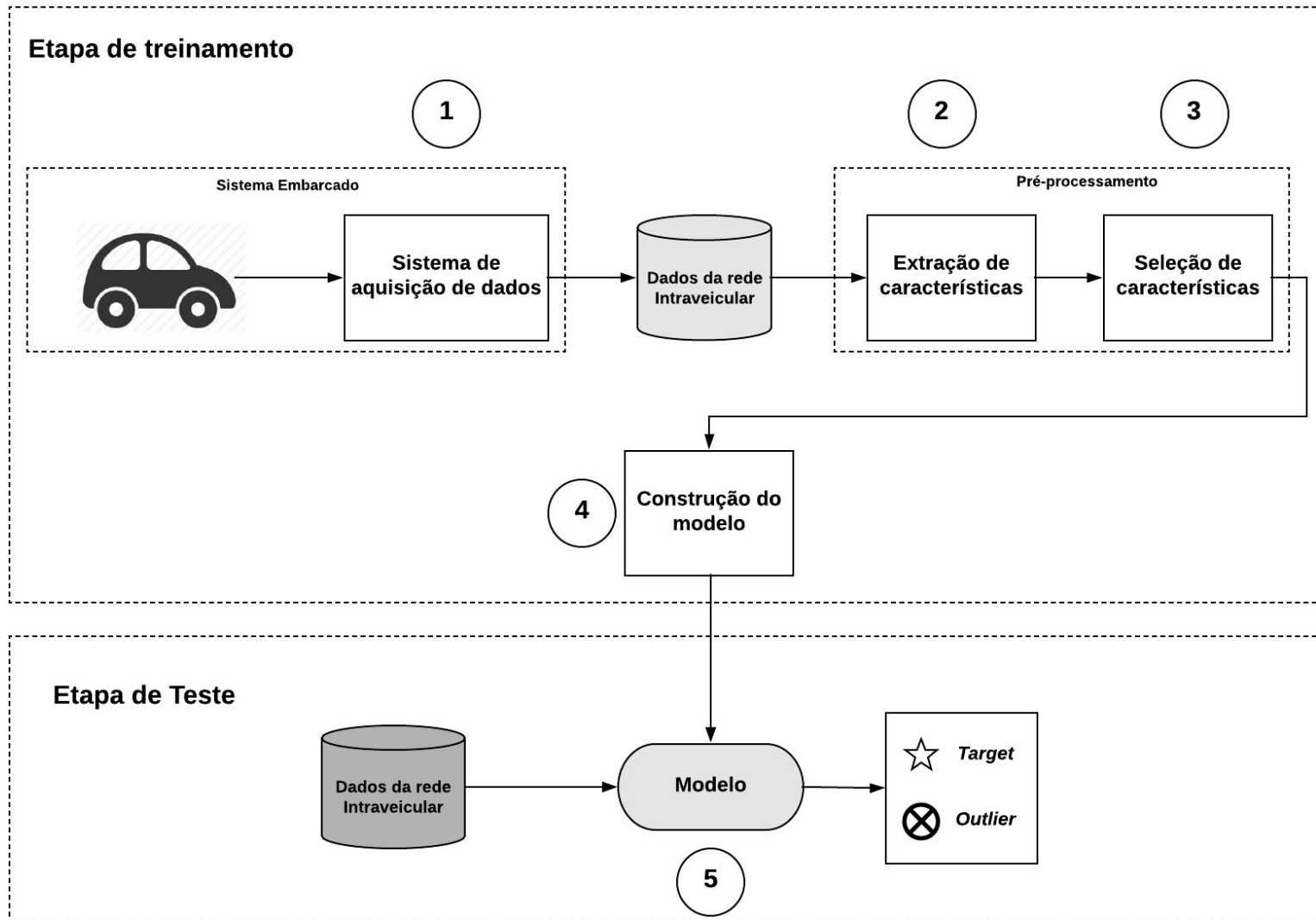


- Sistemas aprendem a partir de dados fornecidos e conseqüentemente, identificam padrões a partir de decisão própria;



Fonte: adaptado de (MAPR Technologies, 2019)

- Um classificador one-class é uma técnica de aprendizado de máquina adaptada para o problema.



- Como parte do pré-processamento dos dados, todos os atributos são normalizados.

$$y_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

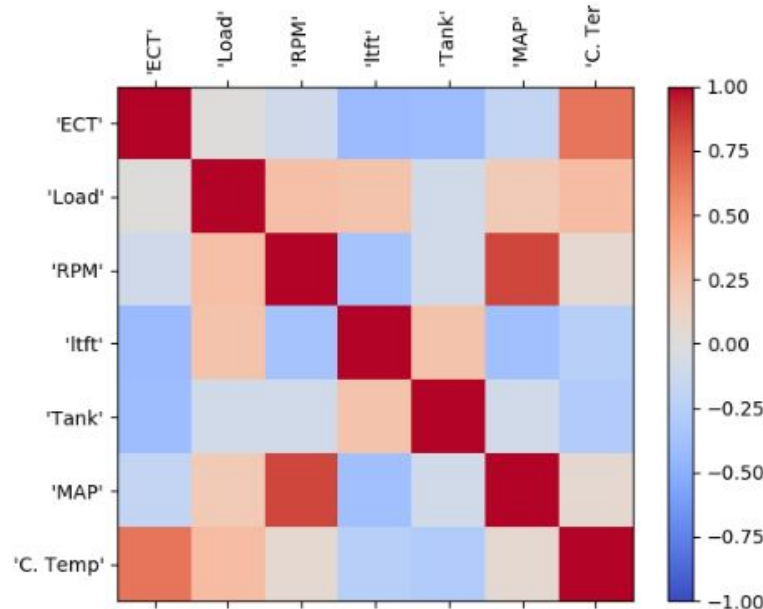
- O sinal de uma *trip* (ciclo veicular) de k segundos pode ser representado então por:

$$X_{trip1} = \begin{pmatrix} x_{1,t_1} & x_{2,t_1} \\ x_{1,t_2} & x_{2,t_2} \\ \vdots & \vdots \\ x_{1,t_k} & x_{2,t_k} \end{pmatrix}$$

- Torna-se interessante analisar intervalos de operação do sensor. Assim, é usada uma **janela de tamanho N** para extração da média, variância e desvio padrão de cada atributo:

$$Y_{t_j \dots t_{j+N-1}} = \left(\bar{x}_1 \quad s_{x_1}^2 \quad s_{x_1} \quad \bar{x}_2 \quad s_{x_2}^2 \quad s_{x_2} \right)$$

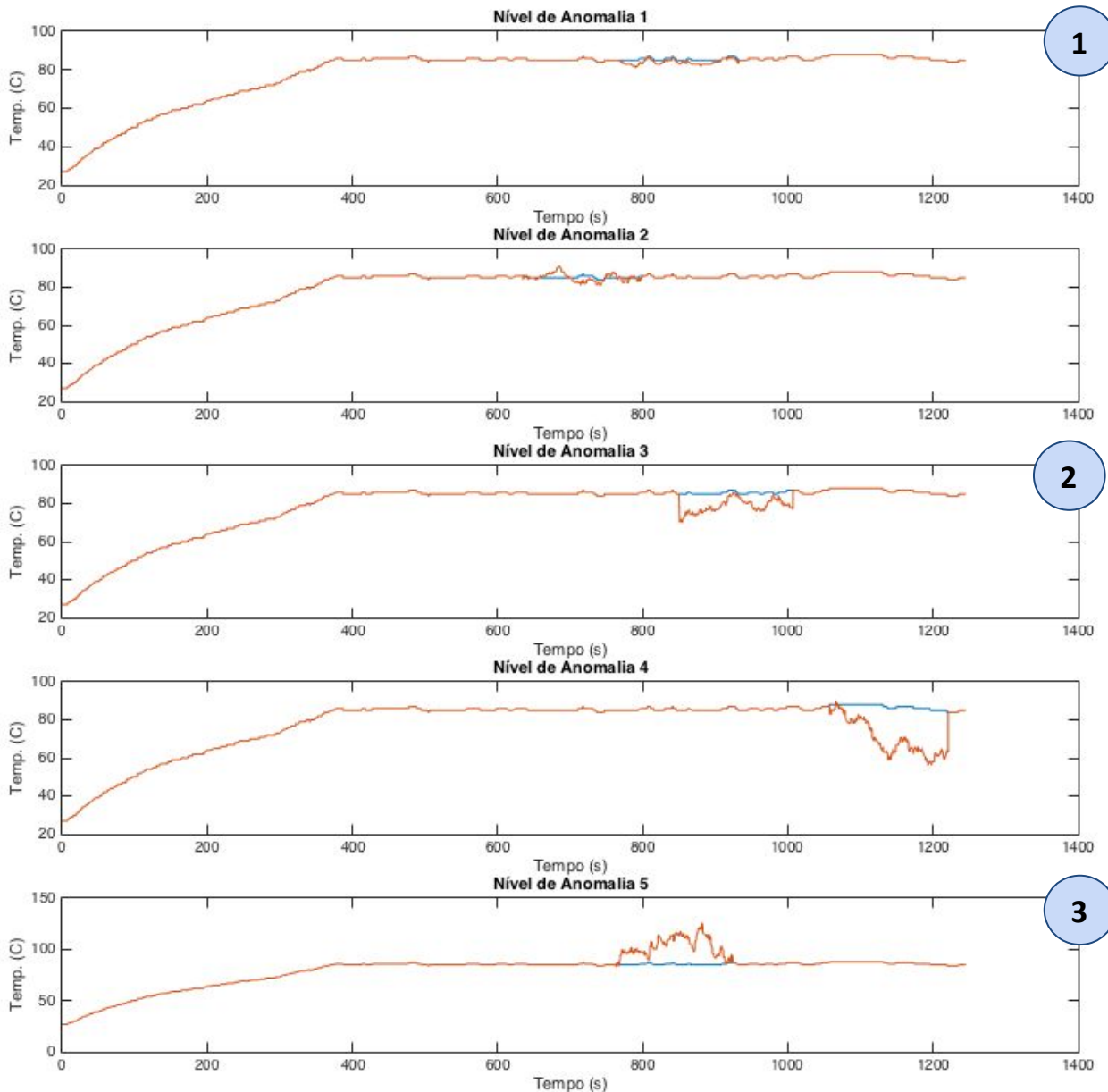
- A partir de análise dos 27 parâmetros coletados pelo sistema, apenas a carga do motor (*Load*) e número de rotações (RPM) por minuto são diretamente relacionados ao sistema do sensor ECT.
- Análise da correlação média destes parâmetros com outros quatro (*Long Term Fuel Trim*, Nível do tanque, *Manifold Absolute Pressure* e Temperatura do Catalisador).



- Após análise dos atributos, é definida a instância de entrada para o sistema:

$$F_i = \left(\bar{x}_1 \quad s_{x_1}^2 \quad s_{x_1} \quad \bar{x}_2 \right)$$

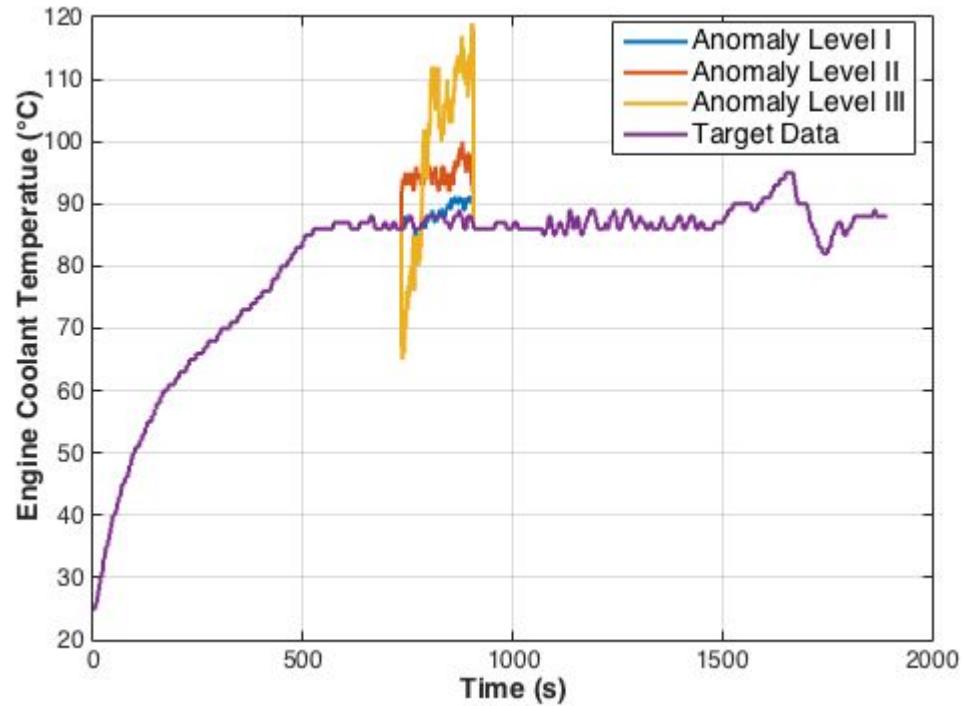
Inserção de Anomalias



- Com a coleta de dados do funcionamento normal do veículo, anomalias são inseridas de maneira artificial.

- Inserção a partir de análise do mau funcionamento do sensor, baseado em níveis de ruído presente no sinal.

Comparação Sinal x sinal com anomalia



- Avaliação de um classificador pode ser realizada a partir de sua matriz de confusão.
- Objetivo de avaliar a detecção de maneira conjunta, tanto *outliers* quanto *target*.
- Uso da métrica *F2-score*, que considera tanto a precisão quanto a sensibilidade em sua fórmula.

$$sens = \frac{TP}{TP + FN}$$

$$prec_{out} = \frac{TN}{TN + FN}$$

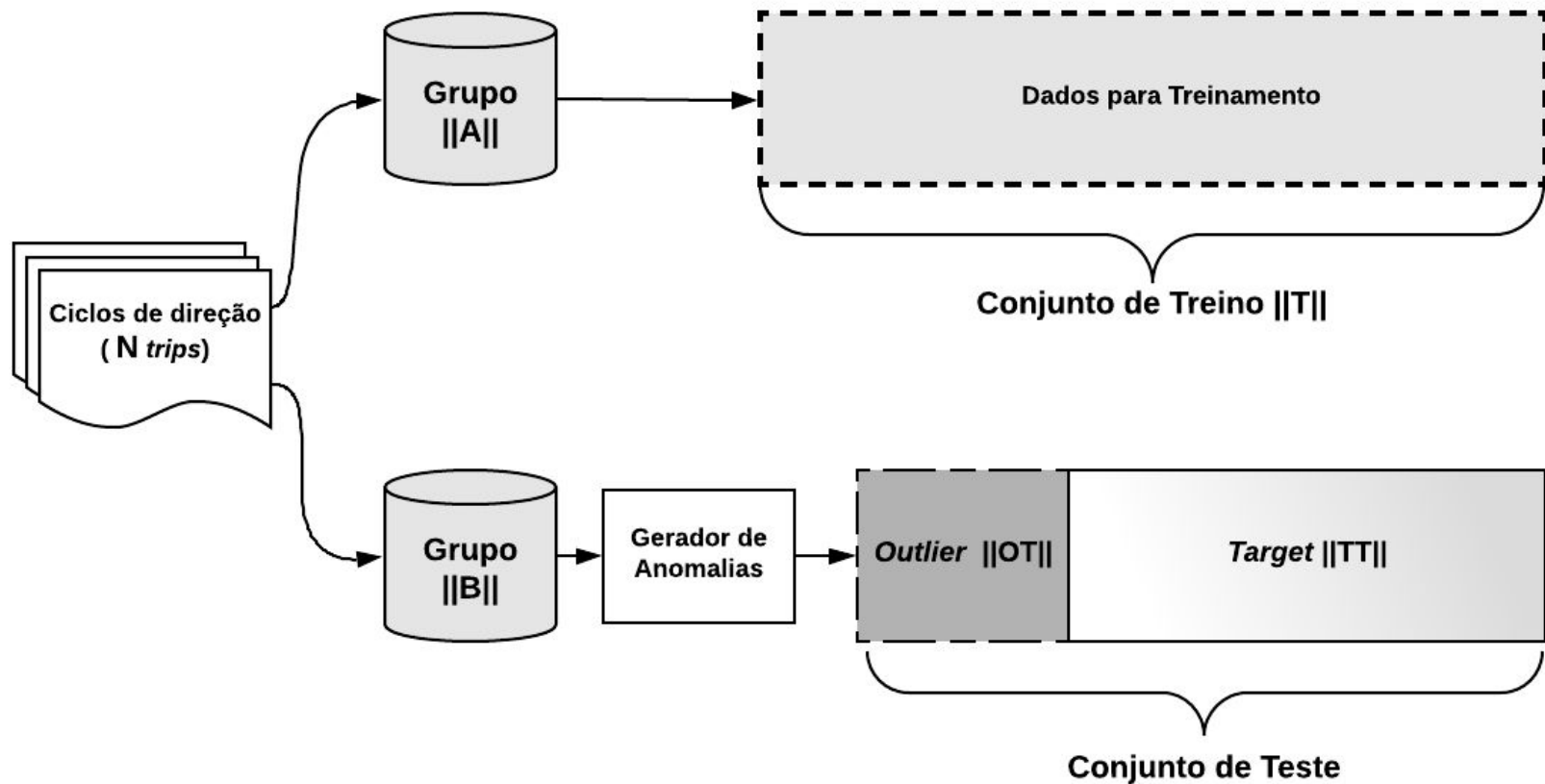
$$F_2 = \frac{5 \times prec \times sens}{4 \times prec + sens}$$

Rótulo da classe	Resultado da classificação	
	Normal	Anomalia
Normal	TP	FN
Anomalia	FP	TN

- *Gauss data description*
- *k-NN data description*
- *Mahalanobis classifier*
- *Minimum Spanning Tree (MST)*
- *Naive Parzen data description*
- *Self-organized map (SOM)*
- *Parzen data description*
- *Support Vector Data Description (SVDD)*
- *Extreme Value*
- *One Class SVM (OC-SVM)*

- Uso da toolbox DDTools (desenvolvida por David Tax) versão 2.1.3 no *software* MATLAB.
- Processo de treinamento apenas com dados do funcionamento normal do veículo.
- Separação de dois grupos , $||\mathbf{A}||$ e $||\mathbf{B}||$ para treinamento e teste.
- Para fins estatísticos, são treinados e testados 30 diferentes modelos.

Divisão dos grupos de treinamento e teste



Modo	T	TT	OT	Número de <i>trips</i>
Idle	16430	8400	1875	25
Movimento	58321	19840	3375	45

- O processo de treinamento e teste foram executados trinta vezes e então os resultados analisados:

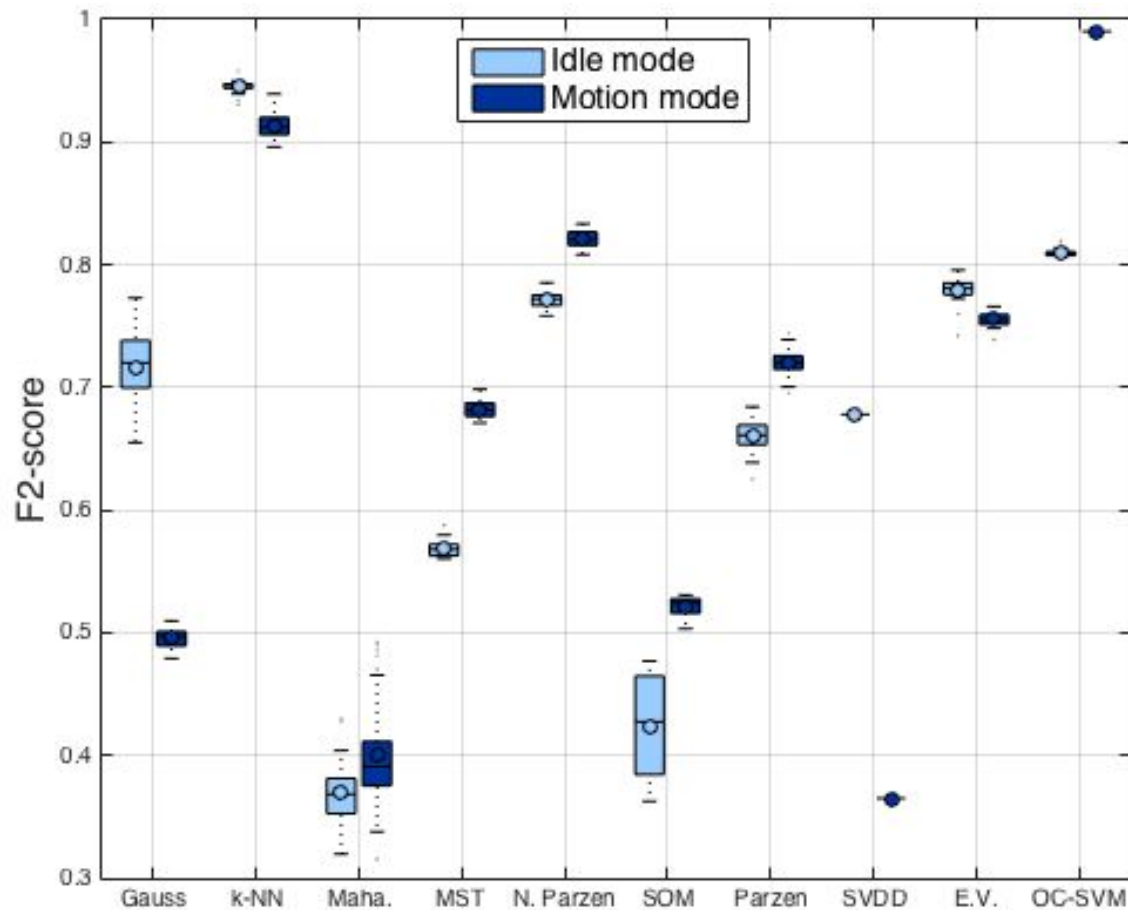
ANOMALY DETECTION SYSTEM RESULTS FOR THE VEHICLE ON IDLE MODE

	Selected window size	Level I			Level II			Level III		
		TPR	Prec	F2	TPR	Prec	F2	TPR	Prec	F2
Gauss	10	0.920	0.225	0.568	0.935	0.375	0.720	0.935	0.375	0.720
k-NN	5	0.957	0.972	0.960	0.941	0.983	0.949	1.000	1.000	1.000
Mahalanobis	10	0.487	0.017	0.008	0.532	0.142	0.343	0.641	0.474	0.598
MST	6	0.687	0.109	0.333	0.738	0.318	0.583	0.869	0.672	0.820
Naive Parzen	6	0.717	0.325	0.577	0.821	0.617	0.770	0.927	0.859	0.912
SOM	7	0.698	0.217	0.483	0.682	0.172	0.428	0.692	0.196	0.459
Parzen	6	0.692	0.278	0.533	0.747	0.462	0.665	0.898	0.813	0.879
SVDD	4	0.718	0.383	0.611	0.814	0.649	0.774	0.832	0.724	0.807
Extreme Value	10	0.669	0.075	0.259	0.749	0.663	0.729	0.803	0.942	0.827
OC-SVM	3	0.972	0.488	0.811	0.972	0.485	0.809	0.977	0.587	0.862

ANOMALY DETECTION SYSTEM RESULTS FOR THE VEHICLE ON MOTION MODE

	Selected window size	Level I			Level II			Level III		
		TPR	Prec	F2	TPR	Prec	F2	TPR	Prec	F2
Gauss	10	0.654	0.189	0.438	0.670	0.246	0.498	0.799	0.616	0.754
k-NN	6	0.939	0.873	0.925	0.927	0.847	0.909	0.934	0.868	0.920
Mahalanobis	6	0.685	0.005	0.024	0.717	0.129	0.375	0.823	0.550	0.748
MST	7	0.680	0.109	0.332	0.763	0.318	0.596	0.873	0.718	0.836
Naive Parzen	10	0.722	0.336	0.587	0.848	0.617	0.788	0.941	0.873	0.926
SOM	6	0.639	0.185	0.428	0.672	0.187	0.442	0.653	0.227	0.474
Parzen	10	0.669	0.301	0.537	0.771	0.448	0.673	0.884	0.842	0.875
SVDD	3	0.704	0.379	0.600	0.771	0.659	0.745	0.900	0.724	0.858
Extreme Value	10	0.831	0.071	0.264	0.872	0.319	0.648	0.921	0.816	0.897
OC-SVM	9	0.998	0.957	0.989	0.998	0.957	0.989	0.998	0.943	0.986

- Uma análise gráfica dos resultados para o segundo nível de anomalia



- **A análise realizada neste trabalho confirma a necessidade da análise de anomalias e falhas em sensores veiculares em diferentes níveis de degeneração.**
- **O procedimento utilizado para a inserção de anomalias no sensor veicular pode ser usado para análise de outros sensores resistivos.**

Obrigado!



efsn@cesar.org.br



Eronides da Silva Neto



github.com/eron93br