



1 2 9 0

UNIVERSIDADE D  
COIMBRA

UNIVERSIDADE DE COIMBRA  
MESTRADO EM CIÊNCIA E ENGENHARIA DE DADOS  
TÓPICOS DE CIÊNCIA DOS DADOS

---

# Classificação de Atividades Humanas

---

Componente A

João Cardoso, 2017247458  
Leonor Coelho, 2017254561  
Mariana Coelho, 2020207009

# Contents

<b>1</b>	<b>Densidade de <i>outliers</i></b>	<b>1</b>
1.1	Através dos quartis . . . . .	1
1.2	Através do Z-score . . . . .	3
1.3	Através do k-means . . . . .	5
<b>2</b>	<b>Deteção e previsão de outliers</b>	<b>6</b>
<b>3</b>	<b>Features</b>	<b>7</b>
3.1	PCA . . . . .	10
<b>4</b>	<b>Referências</b>	<b>11</b>

# 1 Densidade de *outliers*

## 1.1 Através dos quartis

Label		Label Name	Acc N Total	Acc N Outliers	Acc Density	Gyr N Total	Gyr N Outliers	Gyr Density	Mag N Total	Mag N Outliers	Mag Density
1	1	Stand	11012	762	6.919724	11012	902	8.191064	11012	12	0.108972
2	2	Sit	5760	201	3.489583	5760	368	6.388889	5760	0	0.000000
3	3	Sit and Talk	5760	357	6.197917	5760	324	5.625000	5760	45	0.781250
4	4	Walk	8706	363	4.169538	8706	26	0.298645	8706	106	1.217551
5	5	Walk and Talk	9474	192	2.026599	9474	117	1.234959	9474	12	0.126662
6	6	Climb Stairs (CS)	6529	209	3.201103	6529	76	1.164037	6529	0	0.000000
7	7	CS and Talk	2817	50	1.774938	2817	45	1.597444	2817	0	0.000000
8	8	Stand to Sit	384	116	30.208333	384	51	13.281250	384	55	14.322917
9	9	Sit to Stand	384	63	16.406250	384	63	16.406250	384	43	11.197917
10	10	Stand to Sit and Talk	384	99	25.781250	384	70	18.229167	384	69	17.968750
11	11	Sit to Stand and Talk	384	88	22.916667	384	23	5.989583	384	96	25.000000
12	12	Stand to Walk	254	8	3.149606	254	0	0.000000	254	6	2.362205
13	13	Walk to Stand	508	21	4.133858	508	13	2.559055	508	0	0.000000
14	14	Stand to CS/CS and Talk	254	6	2.362205	254	0	0.000000	254	0	0.000000
15	15	CS to Walk	127	4	3.149606	127	0	0.000000	127	15	11.811024
16	16	CS and Talk to Walk and Talk	127	3	2.362205	127	0	0.000000	127	3	2.362205

Figure 1: Tabela de valores das densidades de *outliers* calculada através dos quartis nos 3 sensores

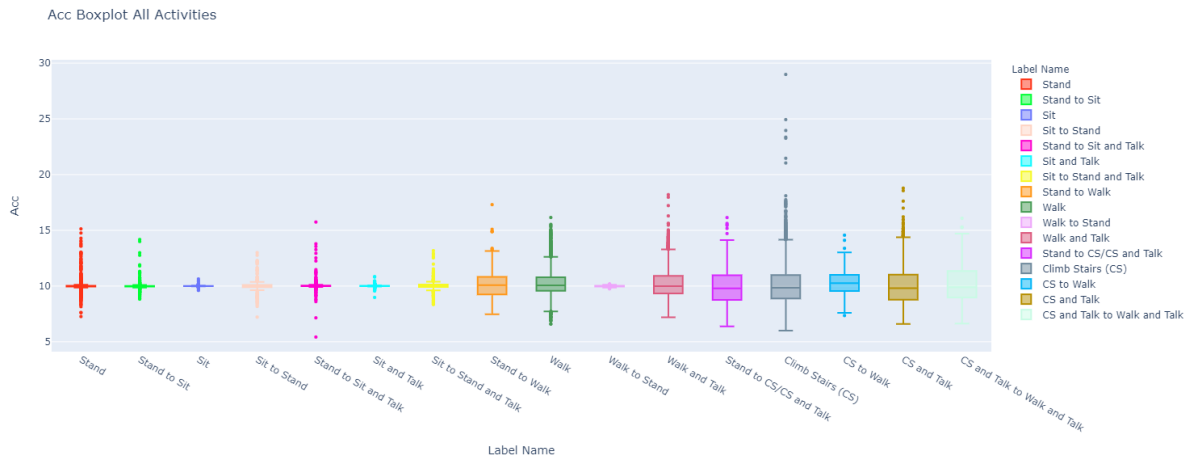


Figure 2: Densidade de *outliers* por atividade do **acelerômetro**

O que mais chama atenção na Tabela 1 é que as atividades *Stand*, *Sit*, *Walk*, são atividades que mais têm representatividade quantitativa de dados, porém não possuem um número elevado de *outliers*, a relação *outliers*/número de dados, o que a faz ser baixa. É perceptível o aumento de densidade de *outliers* quando se adiciona a atividade *Talk* a outra atividade, como *Sit and Talk*. Já temos um aumento percentual de quase 100% se compararmos o vetor de aceleração. Na atividade *Sit* e *Sit and Talk*, o intervalo de valores é pequeno pois, quando uma pessoa está sentada, não se movimenta pelo que a aceleração no eixo  $yy'$  e  $zz'$  é perto de 0. A pessoa que fala enquanto caminha, *Walk and Talk*, possui muitas variações de movimentos, como também as atividades *Climb Stairs*, *Stand to Walk*, *Walk to Stand*, *Climb Stairs and Talk* durante as atividades, o indivíduo apresenta uma aceleração inconstante, gerando os *outliers* e uma maior janela de valores, o que faz com

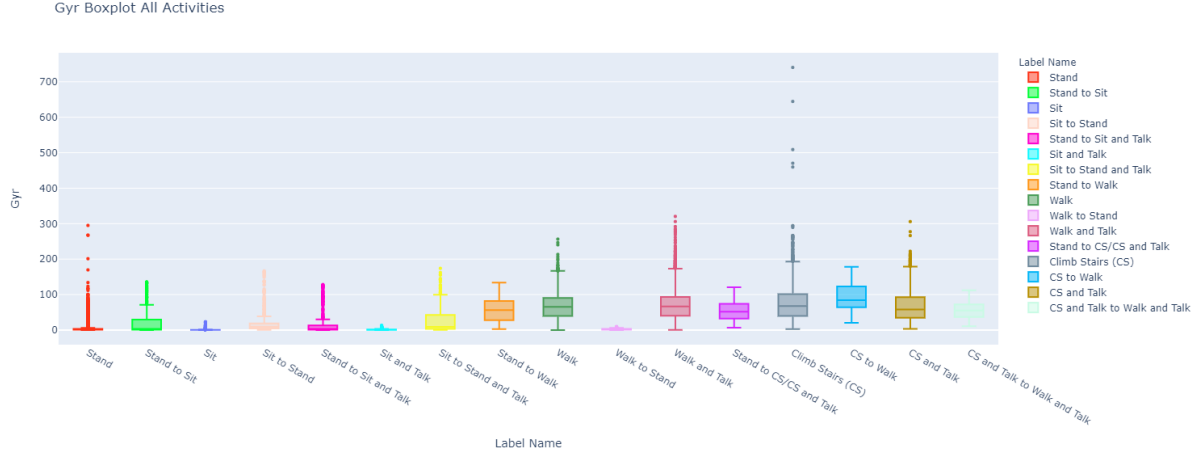


Figure 3: Densidade de *outliers* por atividade do **giroscópio**

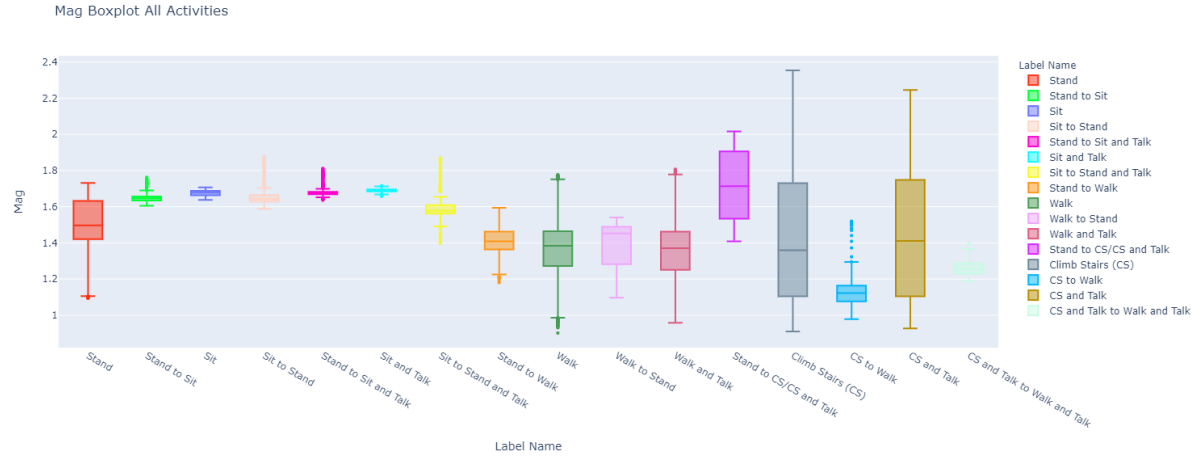


Figure 4: Densidade de *outliers* por atividade do **magnetómetro**

que a densidade dos outliers seja baixa. As atividades 8 a 11 (*Stand* a *Sit*) que consistem na mudança de posição entre estar em pé e sentado, ou vice-versa, enquanto o indivíduo realiza outras atividades (*Talk*). É perceptível que a densidade de outliers nestes casos é alta. Isto pode dever-se a 2 coisas:

- O indivíduo pode ter-se movimentado mais rápido nuns momentos do que noutros, gerando uma grande variedade nos valores.
- Como há menos momentos destas atividades (como podemos observar na tabela na coluna *Acc N Total*), e dada a situação anterior, os momentos que se tornam *outliers* tornam-se demasiados em comparação com o número de momentos da atividade em causa.

Atividades semelhantes, como *Climb Stairs to Walk*, *Climb Stairs and Talk to Walk and Talk* e *Stand -> Climb Stairs*, *Stand -> Climb Stairs and Talk* têm densidades semelhantes. Conclui-se, assim, que adicionar a atividade *Talk* não modifica de forma significativa o número de outliers. Tal acontece também nas outras atividades que apenas é adicionado o *Talk*, a diferença da densidade de *outliers* é pequena.

## 1.2 Através do Z-score

	Label Name	N Total	Density Quartiles	Density Z-score k=3	Density Z-score k=3.5	Density Z-score k=4
1	Stand	11012	6.919724	1.162368	0.989829	0.889938
2	Sit	5760	3.489583	1.232639	0.902778	0.746528
3	Sit and Talk	5760	6.197917	1.319444	1.024306	0.815972
4	Walk	8706	4.169538	1.206065	0.436481	0.218240
5	Walk and Talk	9474	2.026599	0.517205	0.137218	0.073886
6	Climb Stairs (CS)	6529	3.201103	0.597335	0.367591	0.183795
7	CS and Talk	2817	1.774938	0.390486	0.141995	0.106496
8	Stand to Sit	384	30.208333	1.562500	1.562500	1.041667
9	Sit to Stand	384	16.406250	3.125000	2.343750	0.781250
10	Stand to Sit and Talk	384	25.781250	2.343750	2.343750	1.822917
11	Sit to Stand and Talk	384	22.916667	2.343750	1.562500	1.302083
12	Stand to Walk	254	3.149606	1.968504	1.968504	0.393701
13	Walk to Stand	508	4.133858	0.393701	0.196850	0.000000
14	Stand to CS/CS and Talk	254	2.362205	0.787402	0.000000	0.000000
15	CS to Walk	127	3.149606	1.574803	0.000000	0.000000
16	CS and Talk to Walk and Talk	127	2.362205	0.787402	0.000000	0.000000

Figure 5: Comparação da densidade de *outliers* calculada através dos quartis e do Z-score para o **acelerómetro**

Ao trabalharmos com uma base de dados sem tratamento inicial, estamos sujeitos a erros de análises, logo é imprescindível a remoção dos *outliers* e aplicação das técnicas de pré-processamento.

Ao analisarmos as Figuras 5, 5 e 5, temos a imagem da tabela que avalia acelerómetro, giroscópio e magnetómetro respectivamente. Temos os valores numéricos que cada atividade foi repetida, a densidade de *outliers* pelo método dos quartis, e também por z-score com  $k = 3, 3.5, 4$ . Percebe-se que a quantidade de dados que seriam removidos de dados ao aplicarmos o método dos quartis é muito mais alta, visto que usamos  $k = 1.5$ .

Nota-se que as atividades identificadas com maior densidade de *outliers* quando aplicamos o método dos quartis, também seriam as com maior número de *outliers* aplicando o método z-score. Porém no z-score utilizamos  $k = 3, 3.5, 4$ , temos que o valor mais próximo ao método dos quartis seria o  $k=3$ , porém com esse tamanho de  $k$  faz-nos ter uma maior densidade de outliers.

Nas Figura 8 e 10, é perceptível identificarmos uma maior uniformidade da densidade dos *outliers*, do que tínhamos ao aplicarmos o método dos quartis, visto que a escala máxima é de 4.0, o que faz com que os dados fiquem mais evidentes, mesmo que com uma baixa densidade de outliers. Já nas Figuras 9 e 11, a qual temos a adição de mais um elemento, que é o tratamento dos dados por método dos quartis, vemos que temos atividades com uma densidade de *outliers* que se destoam das outras. É visível que quanto menor o  $k$ , mais amostras que se destoam da análise, são removidas.

Na Figura 12, avaliação do magnetómetro, percebe-se que o número de *outliers* ao aplicarmos o método de z-score nos fornece pouquíssimos dados,  $k = 3$  temos somente 6 ativi-

	Label Name	N Total	Density Quartiles	Density Z-score k=3	Density Z-score k=3.5	Density Z-score k=4
1	Stand	11012	8.191064	1.916092	1.698147	1.534689
2	Sit	5760	6.388889	1.145833	0.937500	0.746528
3	Sit and Talk	5760	5.625000	2.013889	1.493056	1.024306
4	Walk	8706	0.298645	0.172295	0.080404	0.045945
5	Walk and Talk	9474	1.234959	0.759975	0.369432	0.200549
6	Climb Stairs (CS)	6529	1.164037	0.658600	0.444172	0.291009
7	CS and Talk	2817	1.597444	0.780973	0.141995	0.106496
8	Stand to Sit	384	13.281250	1.041667	0.000000	0.000000
9	Sit to Stand	384	16.406250	3.645833	2.343750	0.260417
10	Stand to Sit and Talk	384	18.229167	2.864583	0.000000	0.000000
11	Sit to Stand and Talk	384	5.989583	2.343750	0.520833	0.000000
12	Stand to Walk	254	0.000000	0.000000	0.000000	0.000000
13	Walk to Stand	508	2.559055	1.574803	1.377953	0.787402
14	Stand to CS/CS and Talk	254	0.000000	0.000000	0.000000	0.000000
15	CS to Walk	127	0.000000	0.000000	0.000000	0.000000
16	CS and Talk to Walk and Talk	127	0.000000	0.000000	0.000000	0.000000

Figure 6: Comparação da densidade de *outliers* calculada através dos quartis e do Z-score para o **giroscópio**

dades com *outliers*, enquanto na Figura 13, identificamos que, na aplicação do método dos quartis temos 11 atividades *outliers*. Logo, ao aplicarmos o z-score k=3, temos aproximadamente 50% menos atividades com *outliers* nos dados.

	Label Name	N Total	Density Quartiles	Density Z-score k=3	Density Z-score k=3.5	Density Z-score k=4
1	Stand	11012	0.108972	0.000000	0.000000	0.0
2	Sit	5760	0.000000	0.000000	0.000000	0.0
3	Sit and Talk	5760	0.781250	0.347222	0.052083	0.0
4	Walk	8706	1.217551	0.000000	0.000000	0.0
5	Walk and Talk	9474	0.126662	0.000000	0.000000	0.0
6	Climb Stairs (CS)	6529	0.000000	0.000000	0.000000	0.0
7	CS and Talk	2817	0.000000	0.000000	0.000000	0.0
8	Stand to Sit	384	14.322917	0.260417	0.000000	0.0
9	Sit to Stand	384	11.197917	4.427083	1.302083	0.0
10	Stand to Sit and Talk	384	17.968750	2.864583	0.781250	0.0
11	Sit to Stand and Talk	384	25.000000	0.000000	0.000000	0.0
12	Stand to Walk	254	2.362205	0.393701	0.000000	0.0
13	Walk to Stand	508	0.000000	0.000000	0.000000	0.0
14	Stand to CS/CS and Talk	254	0.000000	0.000000	0.000000	0.0
15	CS to Walk	127	11.811024	1.574803	0.000000	0.0
16	CS and Talk to Walk and Talk	127	2.362205	0.000000	0.000000	0.0

Figure 7: Comparação da densidade de *outliers* calculada através dos quartis e do Z-score para o magnetómetro

### 1.3 Através do k-means

Ao aplicarmos a clusterização com o algoritmo k-means usamos a métrica de que definimos  $k$  grupos (clusters) onde cada observação pertence ao grupo que tenha média mais próxima ao seu valor. A coordenada de cada ponto se aproxima as semelhantes para formar um cluster, e assim, ficando mais afastada dos outros pontos que fazem parte de outros clusters. Ao contrário deste algoritmo, o z-score avalia os *outliers* pela sua distância ao centro. Pode acontecer de um ponto ter uma coordenada muito mais alta que as outras e é considerado *outlier* pelo z-score mas não pelo k-means pois pode estar próximo de outros pontos e formar um cluster.

Como dito acima, temos na Figura ?? a densidade de outliers para a aplicação do k-means nas atividades, e conseguimos identificar que as atividades de *Sit to Stand* a *Sit to Stand and Talk*, que são as que mais tínhamos dados a serem removidos se usássemos a métrica dos quartis ou z-score. Vemos que provavelmente algumas delas, como a 9 que ficou com 0.26 de densidade de *outliers*, provavelmente teve a formação de um cluster com os *outliers*. O que poderia afetar negativamente na análise dos dados, a não remoção correta dos *outliers*.

Vemos na Figura 15, que a atividade *Stand -> CS*, *Stand -> CS and Talk* apresentou um valor muito maior de densidade de outliers 13.77, do que a aplicação dos quartis e ao z-score que deram a essa atividade a não presença de *outliers*.

Nas Figuras 18, 20, ?? chama atenção que ao aplicarmos o método k-means temos remoção de outliers em todas as atividades, o que não foi verificado ao aplicarmos o método z-score, além de termos valores de densidade de outliers muito mais elevadas.

A atividade *Stand* chama atenção da presença de densidade de *outliers* muito mais alta

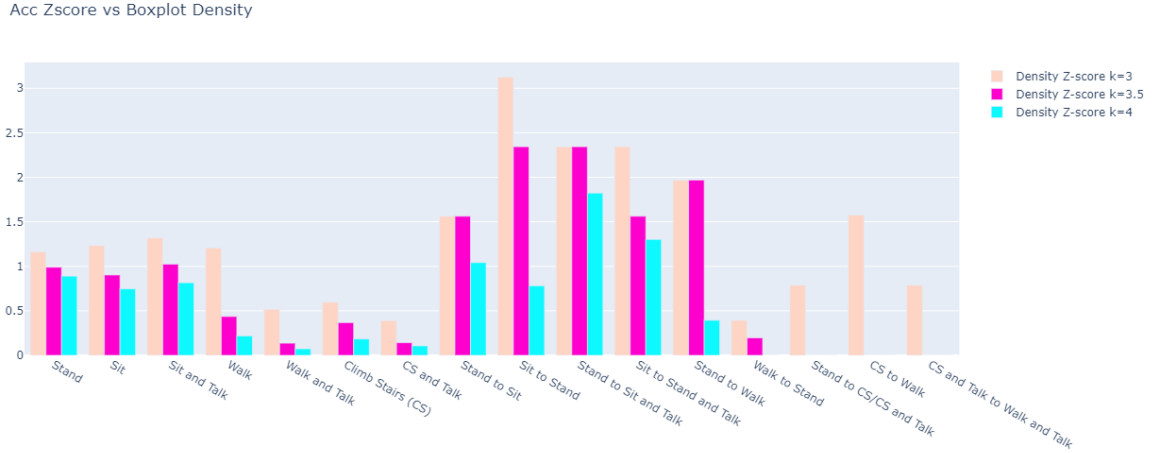


Figure 8: Densidade de *outliers* calculada através do Z-score com k igual a 3, 3.5 e 4 para o **acelerómetro**

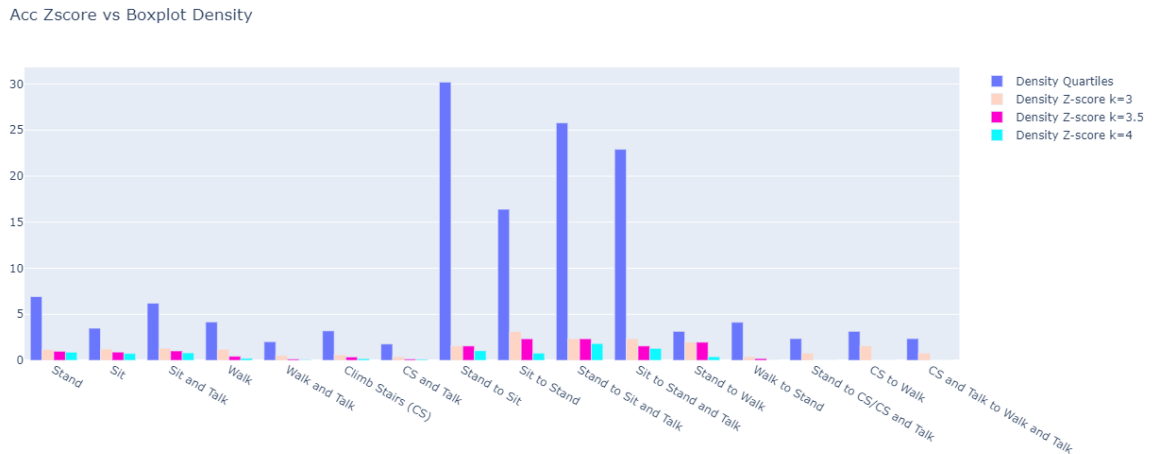


Figure 9: Gráfico de comparação da densidade de *outliers* calculada através dos quartis e do Z-score para o **acelerómetro**

do que na avaliação do z-score. Visto que não é uma atividade com grandes variações de movimentos humanos então, esperava-se que não existissem muitos *outliers*. Além disso, se avaliarmos os gráficos de giroscópio e magnetómetro, temos que é a atividade com maior valor de densidade de *outliers*, para serem removidos utilizando a métrica do k-means.

## 2 Detecção e previsão de outliers

Nos gráficos apresentados são usadas amostras do Acelerómetro.

Como podemos observar, olhando para as tabelas 24 e 28, o erro de predição é pequeno. No entanto entre os dois conseguimos ver que quando fazemos a predição centrada no ponto a prever os resultados são muito melhores.

Enquanto que na predição baseada nos valores anteriores o erro se concentra no intervalo



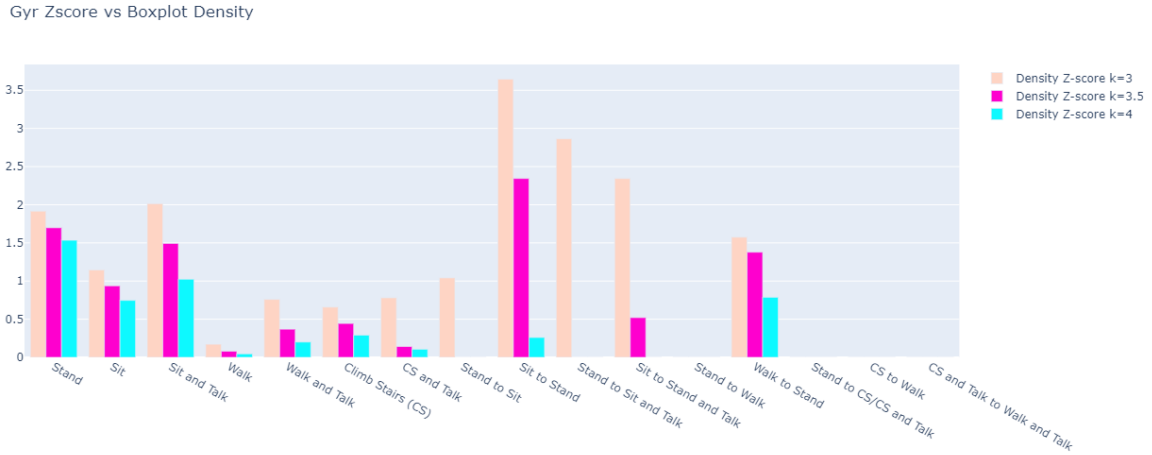


Figure 10: Densidade de *outliers* calculada através do Z-score com k igual a 3, 3.5 e 4 para o **giroscópio**

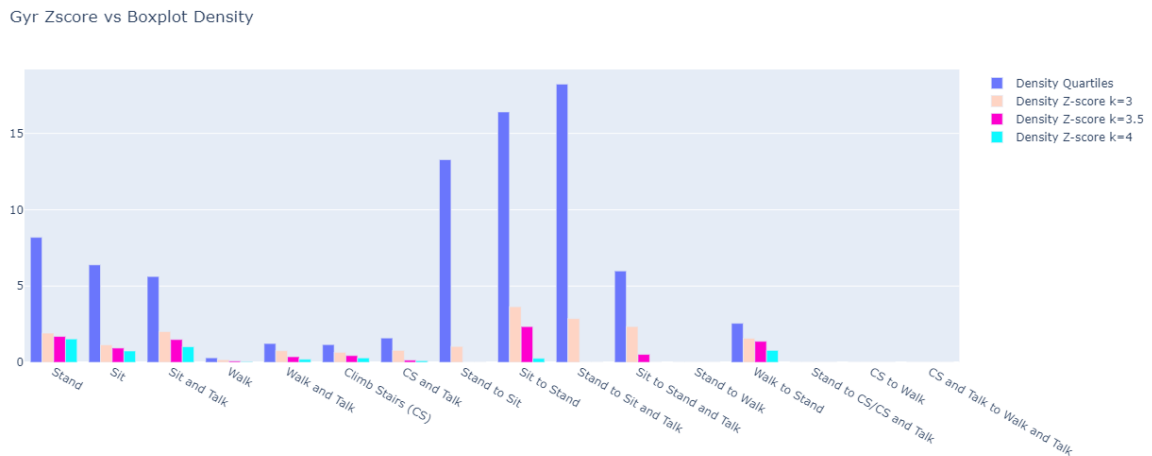


Figure 11: Gráfico de comparação da densidade de *outliers* calculada através dos quartis e do Z-score para o **giroscópio**

de -0.5 a 0.5, com um erro máximo perto de 2.5, na predição centrada temos os valores do erro concentrados no intervalo de -0.2 a 0.2 com um erro máximo de 1. Um erro muito menor.

Juntamente com as tabelas temos 3 exemplos para cada previsão e um gráfico com todos os outliers e os seus valores previstos.

### 3 Features

No Exercício 4.2., extraímos as *features* estatísticas, temporais e espectrais referidas em [Artigo 4] com uma janela de 2s com 50% de sobreposição, ou seja, cada janela tem 102 amostras. Optámos por juntar todos os momentos dos *datasets* de todos os indivíduos por atividade e só depois fazer a análise por janela (tendo cada janela apenas um tipo de atividade). Ao todo, acabámos com 101 *features* relacionadas com o acelerómetro e o giroscópio. O magnetómetro não foi analisado por não ser referido no artigo.

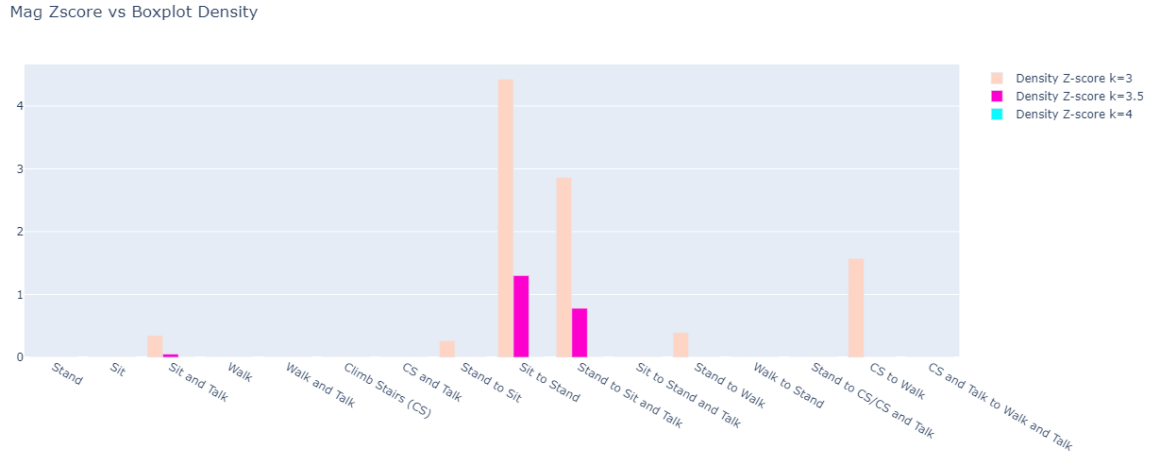


Figure 12: Densidade de *outliers* calculada através do Z-score com k igual a 3, 3.5 e 4 para o **magnetômetro**

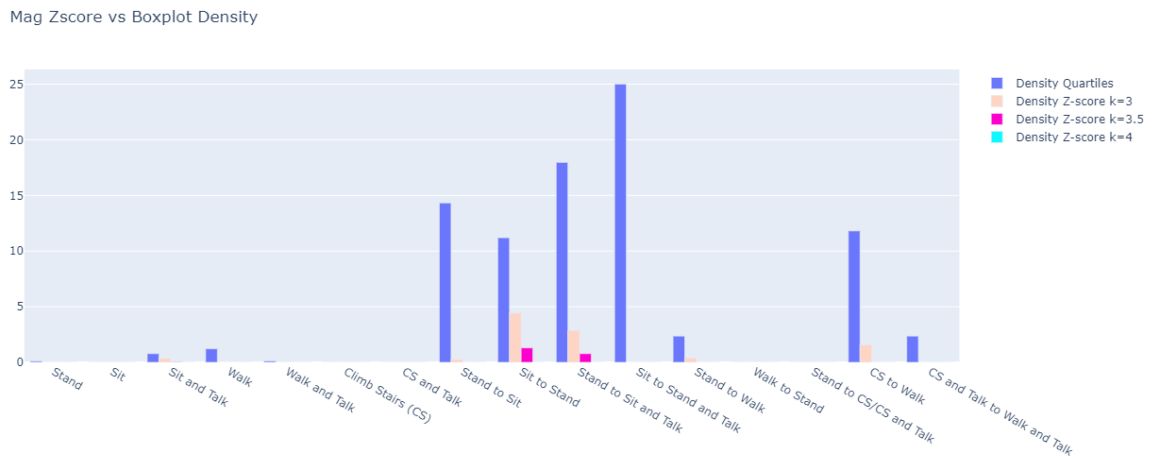


Figure 13: Gráfico de comparação da densidade de *outliers* calculada através dos quartis e do Z-score para o **magnetômetro**

Depois de calcularmos estes dados, procedemos à análise da dependência e proporcionalidade entre features. Encontrámos certos conjuntos que dispõe as várias atividades de forma distinta, sendo possível diferenciá-las visualmente nos gráficos.

Destes gráficos apresentados, conseguimos detetar alguns padrões:

- As ações com menos medições, como a *Stand to Sit*, não nos permitem retirar conclusões dessas atividades. Por isso, iremos falar mais das que têm mais momentos.
- A ação *Talk* não altera com grande significância as outras ações.
- As ações que impliquem mais movimento, como *Walk* e *Climb Stairs* são as menos certas e mais dispersas, pois são mais propícias a erros de medição por serem ações com maior amplitude de movimento.
- As ações que envolvam estar sentado (*Sit* e *Sit and Talk*) diferenciam-se das outras em praticamente todos os gráficos, sendo difícil de distinguir as duas pelo que foi

	Label Name	N Clusters	K-means Density	Density Quartiles	Density Z-score k=3	Density Z-score k=3.5	Density Z-score k=4
1	Stand	6	7.237559	6.919724	1.162368	0.989829	0.889938
2	Sit	4	6.197917	3.489583	1.232639	0.902778	0.746528
3	Sit and Talk	4	1.527778	6.197917	1.319444	1.024306	0.815972
4	Walk	4	18.883529	4.169538	1.206065	0.436481	0.218240
5	Walk and Talk	5	9.446907	2.026599	0.517205	0.137218	0.073886
6	Climb Stairs (CS)	5	4.870577	3.201103	0.597335	0.367591	0.183795
7	CS and Talk	5	7.099752	1.774938	0.390486	0.141995	0.106496
8	Stand to Sit	4	13.541667	30.208333	1.562500	1.562500	1.041667
9	Sit to Stand	7	0.260417	16.406250	3.125000	2.343750	0.781250
10	Stand to Sit and Talk	4	3.645833	25.781250	2.343750	2.343750	1.822917
11	Sit to Stand and Talk	5	4.427083	22.916667	2.343750	1.562500	1.302083
12	Stand to Walk	5	11.811024	3.149606	1.968504	1.968504	0.393701
13	Walk to Stand	4	4.133858	4.133858	0.393701	0.196850	0.000000
14	Stand to CS/CS and Talk	5	9.448819	2.362205	0.787402	0.000000	0.000000
15	CS to Walk	4	21.259843	3.149606	1.574803	0.000000	0.000000
16	CS and Talk to Walk and Talk	4	4.724409	2.362205	0.787402	0.000000	0.000000

Figure 14: Comparação da densidade de *outliers* calculada através dos quartis, do Z-score e do k-means para o **acelerómetro**

referido acima.

- O gráfico 33 não deu o resultado esperado em comparação com o Artigo [Artigo 4]. Isto pode-se dever ao facto de as atividades presentes no Artigo serem diferentes das analisadas neste trabalho. Mas mesmo o resultado não estando semelhante, conseguimos detetar uma concentração das atividades *Sit*, *Stand* e *Sit and Talk*.
- O gráfico 34 mostra que as ações com menor amplitude de movimento (*Stand* e *Sit*) formam uma curva semelhante a uma parábola, mostrando-nos o tipo de dependência que há entre a velocidade no sentido do movimento e no sentido gravítico. Neste caso, conseguimos igualmente distinguir os movimentos *Sit* e *Sit and Talk*, que nem sempre é possível pois a ação *Talk*, como já referido, não altera muito os dados. Também conseguimos notar que as restantes atividades geram também uma curva, sendo esta menos acentuada dada a dispersão dos dados em torno da mesma.
- O gráfico 35 deu um resultado muito semelhante ao esperado (embora haja uma leve dispersão dos pontos). Nota-se claramente uma curva do tipo quadrática na distribuição dos dados. Esta relação também se nota no gráfico 36.
- Nas imagens 37 e 38, notamos novamente as atividades *Sit*, *Stand* e *Sit and Talk* distanciadas das outras e bem definidas. As atividades *Climb Stairs* e *Climb Stairs and Talk* também estão mais concentradas numa zona do gráfico assim como as ações *Walk and Talk* e *Walk*, mostrando mais uma vez a relação entre as atividades simples e as que têm a atividade *Talk*.
- No gráfico 39 notamos uma distinção clara entre *Sit* e *Sit and Talk* das restantes atividades, estando estas mais dispersas entre -2 e 1.
- A relação que notamos em 40 é clara: os valores de atividades com pouco movimento seguem quase uma reta com declive próximo de 0. Isto era de esperar dado que o desvio padrão para estas atividades é baixo, logo não é um grande desvio dos valores

	Label Name	N Clusters	K-means Density	Density Quartiles	Density Z-score k=3	Density Z-score k=3.5	Density Z-score k=4
1	Stand	4	19.115510	8.191064	1.916092	1.698147	1.534689
2	Sit	5	6.111111	6.388889	1.145833	0.937500	0.746528
3	Sit and Talk	5	8.802083	5.625000	2.013889	1.493056	1.024306
4	Walk	5	8.074891	0.298645	0.172295	0.080404	0.045945
5	Walk and Talk	5	9.837450	1.234959	0.759975	0.369432	0.200549
6	Climb Stairs (CS)	5	10.905192	1.164037	0.658600	0.444172	0.291009
7	CS and Talk	5	6.638268	1.597444	0.780973	0.141995	0.106496
8	Stand to Sit	5	5.468750	13.281250	1.041667	0.000000	0.000000
9	Sit to Stand	4	10.156250	16.406250	3.645833	2.343750	0.260417
10	Stand to Sit and Talk	5	9.375000	18.229167	2.864583	0.000000	0.000000
11	Sit to Stand and Talk	5	8.072917	5.989583	2.343750	0.520833	0.000000
12	Stand to Walk	5	8.661417	0.000000	0.000000	0.000000	0.000000
13	Walk to Stand	5	8.858268	2.559055	1.574803	1.377953	0.787402
14	Stand to CS/CS and Talk	4	13.779528	0.000000	0.000000	0.000000	0.000000
15	CS to Walk	4	3.149606	0.000000	0.000000	0.000000	0.000000
16	CS and Talk to Walk and Talk	5	8.661417	0.000000	0.000000	0.000000	0.000000

Figure 15: Comparação da densidade de *outliers* calculada através dos quartis, do Z-score e do k-means para o **giroscópio**

da média. Os restantes pontos concentram-se de uma forma diferente, criando um aglomerado mais perto da origem. A outra concentração mais perto de -2.5 pode dever-se aos dados de um indivíduo que teve um comportamento diferente dos restantes.

A partir da análise descrita acima podemos retirar algumas conclusões:

- As atividades *Sit*, *Sit and Talk* e *Stand* são as mais fáceis de detetar, permitindo-nos analisá-las melhor.
- As atividades que impliquem a ação *Talk* não têm valores muito diferentes das que não têm esta ação.
- As atividades com grande amplitude de movimento têm valores mais dispersos por serem mais propícias a erros de medição.
- Através das relações analisadas, conseguimos utilizar algumas para fazer uma boa deteção de qual a atividade presente naquela janela.

### 3.1 PCA

Para explicar 75% do feature set são necessárias 11 dimensões, para a amostra utilizada.

Para obter as features para um instante, basta selecionar no dataset do PCA a linha respectiva.

A vantagem desta abordagem é a grande redução do espaço de features que vai aumentar mais tarde a performance do nosso algoritmo de machine learning. Também existem desvantagens como por exemplo o erro associado ao uso deste método e a perda de interpretabilidade dos dados.

	Label Name	N Clusters	K-means Density	Density Quartiles	Density Z-score k=3	Density Z-score k=3.5	Density Z-score k=4
1	Stand	4	6.538322	0.108972	0.000000	0.000000	0.0
2	Sit	5	7.708333	0.000000	0.000000	0.000000	0.0
3	Sit and Talk	5	16.145833	0.781250	0.347222	0.052083	0.0
4	Walk	5	9.648518	1.217551	0.000000	0.000000	0.0
5	Walk and Talk	4	7.177539	0.126662	0.000000	0.000000	0.0
6	Climb Stairs (CS)	4	11.487211	0.000000	0.000000	0.000000	0.0
7	CS and Talk	5	7.632233	0.000000	0.000000	0.000000	0.0
8	Stand to Sit	4	17.708333	14.322917	0.260417	0.000000	0.0
9	Sit to Stand	4	8.072917	11.197917	4.427083	1.302083	0.0
10	Stand to Sit and Talk	4	11.197917	17.968750	2.864583	0.781250	0.0
11	Sit to Stand and Talk	4	13.281250	25.000000	0.000000	0.000000	0.0
12	Stand to Walk	6	4.724409	2.362205	0.393701	0.000000	0.0
13	Walk to Stand	4	10.433071	0.000000	0.000000	0.000000	0.0
14	Stand to CS/CS and Talk	4	14.566929	0.000000	0.000000	0.000000	0.0
15	CS to Walk	4	11.023622	11.811024	1.574803	0.000000	0.0
16	CS and Talk to Walk and Talk	5	7.874016	2.362205	0.000000	0.000000	0.0

Figure 16: Comparação da densidade de *outliers* calculada através dos quartis, do Z-score e do k-means para o **magnetômetro**

## 4 Referências

Mi Zhang, Alexander A Sawchuk, A. Sawchuk, A feature selection-based framework for human activity recognition using wearable multimodal sensors, : BodyNets '11: Proceedings of the 6th International Conference on Body Area Networks, November 2011 Pages 92–98. <https://pdfs.semanticscholar.org/8522/ce2bfce1ab65b133e411350478183e79fae7.pdf>

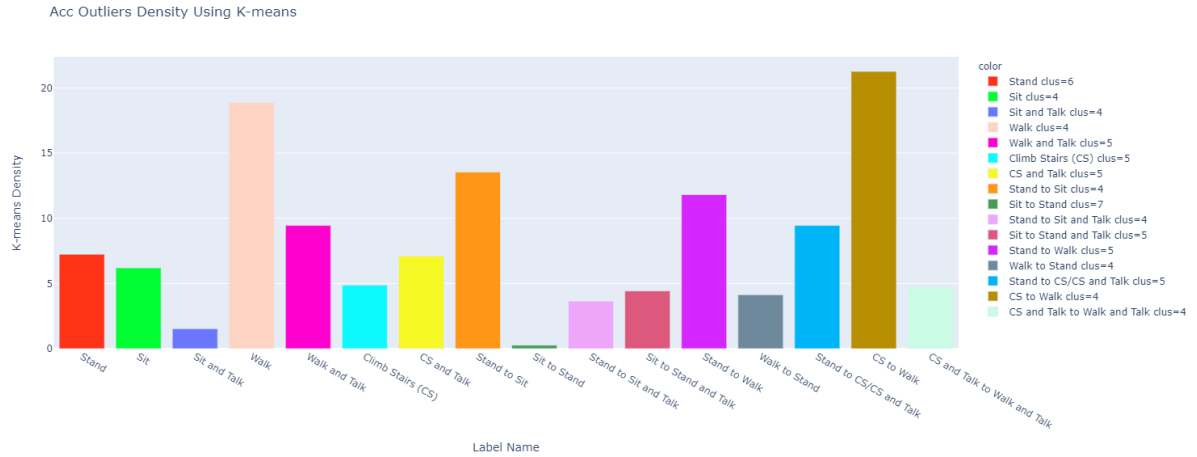


Figure 17: Gráfico de comparação da densidade de *outliers* calculada através do k-means para o **acelerómetro**

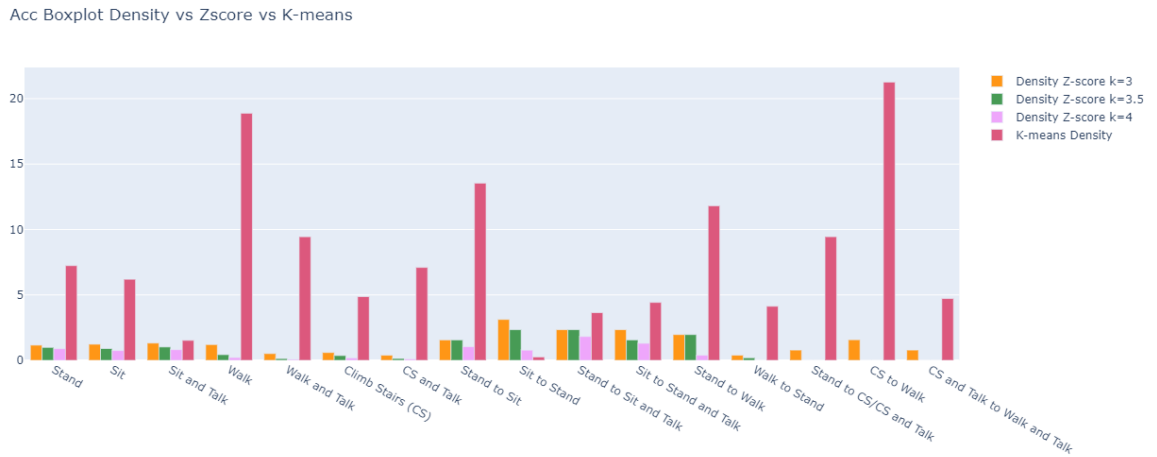


Figure 18: Gráfico de comparação da densidade de *outliers* calculada através dos quartis, do z-score e do k-means para o **acelerómetro**

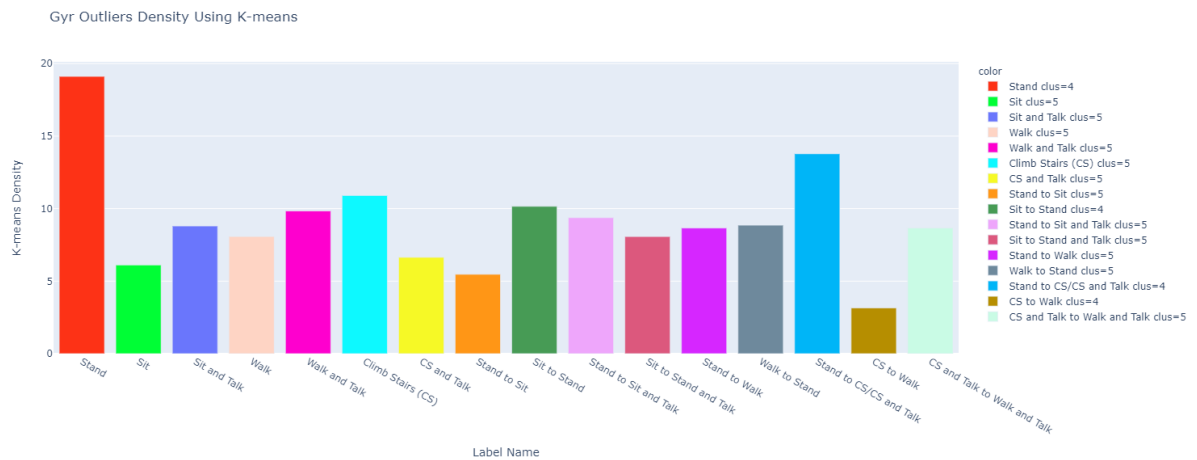


Figure 19: Gráfico de comparação da densidade de *outliers* calculada através do k-means para o **giroscópio**

Gyr Boxplot Density vs Zscore vs K-means

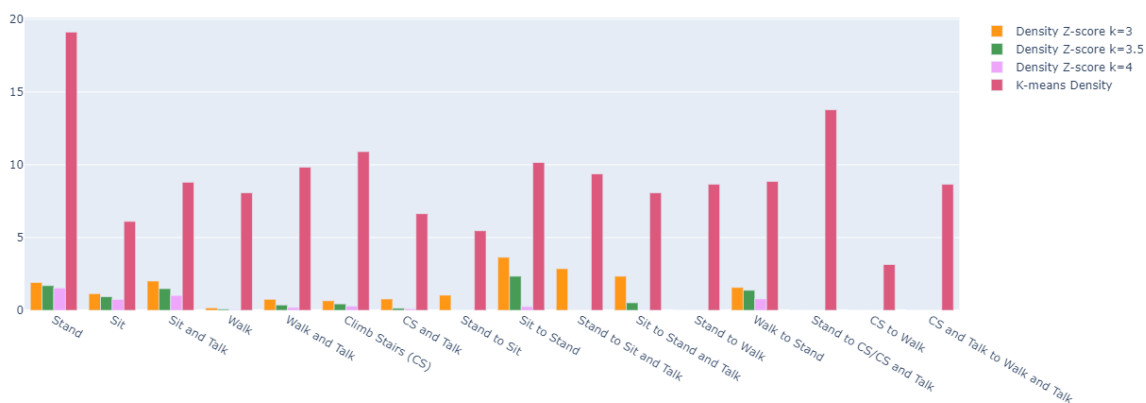


Figure 20: Gráfico de comparação da densidade de *outliers* calculada através dos quartis, do z-score e do k-means para o **giroscópio**

Mag Outliers Density Using K-means

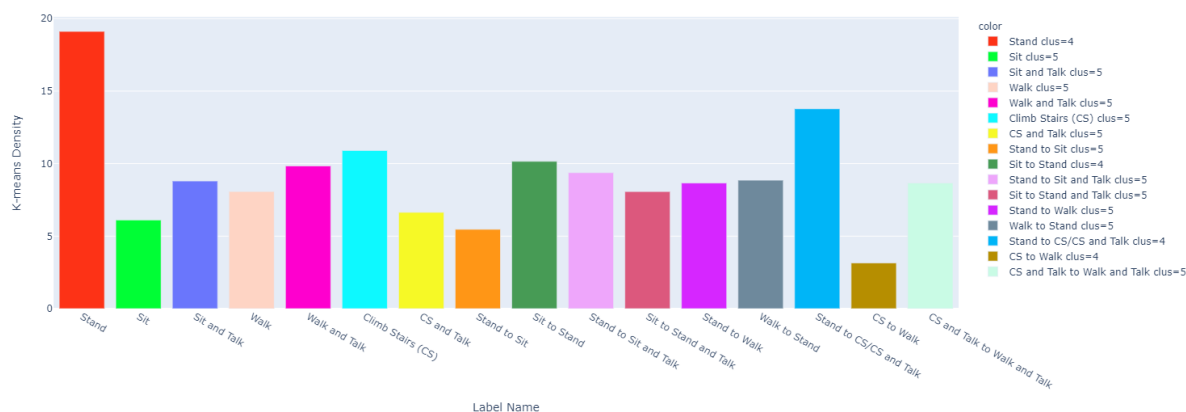


Figure 21: Gráfico de comparação da densidade de *outliers* calculada através do k-means para o **magnetômetro**

Mag Boxplot Density vs Zscore vs K-means

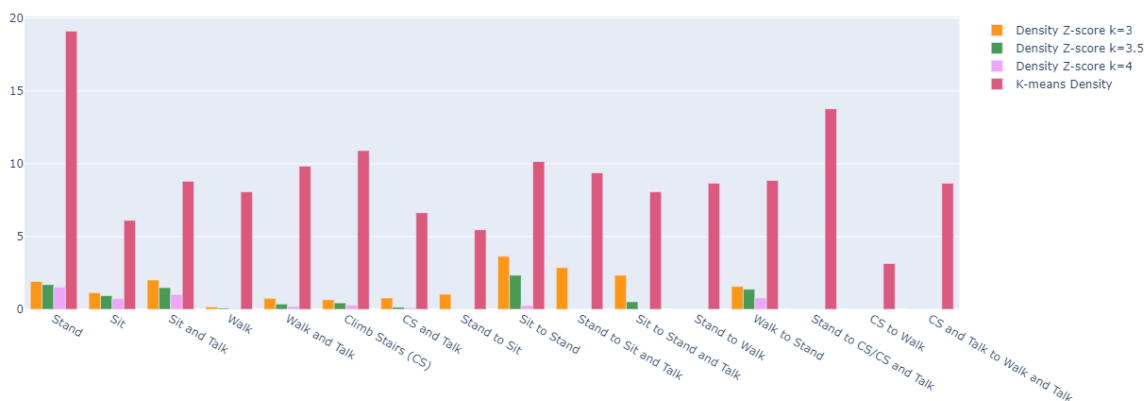


Figure 22: Gráfico de comparação da densidade de *outliers* calculada através dos quartis, do z-score e do k-means para o **magnetômetro**

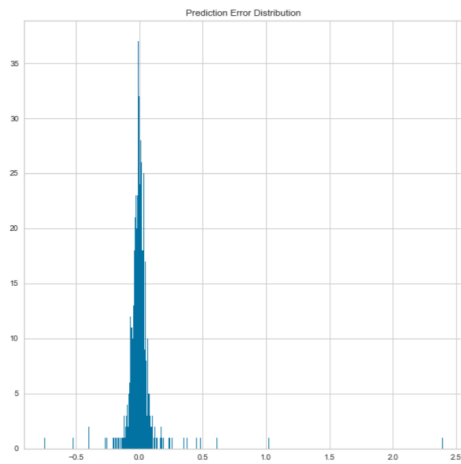


Figure 23: Distribuição de erros dos valores previstos com os  $p$  valores anteriores

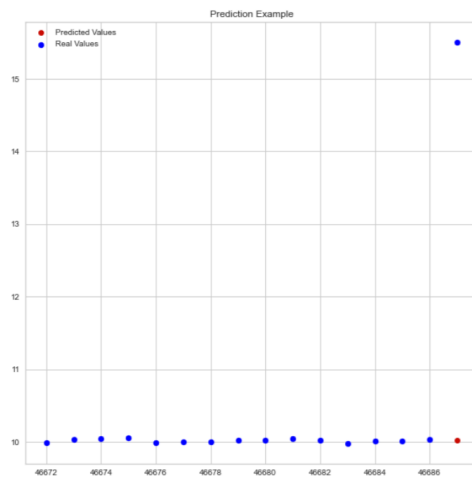


Figure 24: Exemplo de previsão 1 com os  $p$  valores anteriores

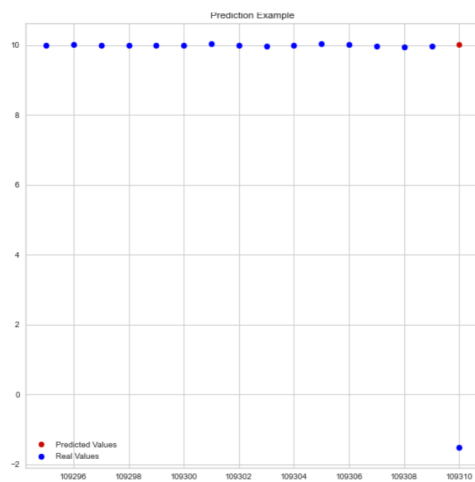


Figure 25: Exemplo de previsão 2 com os  $p$  valores anteriores



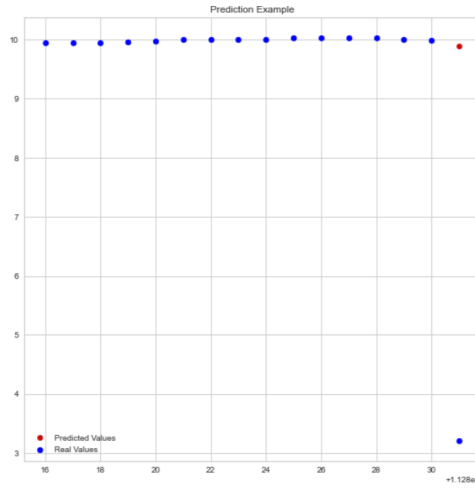


Figure 26: Exemplo de previsão 3 com os  $p$  valores anteriores

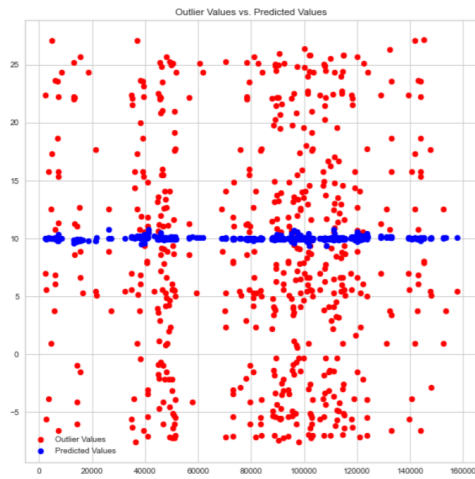


Figure 27: Previsão dos outliers com os  $p$  valores anteriores

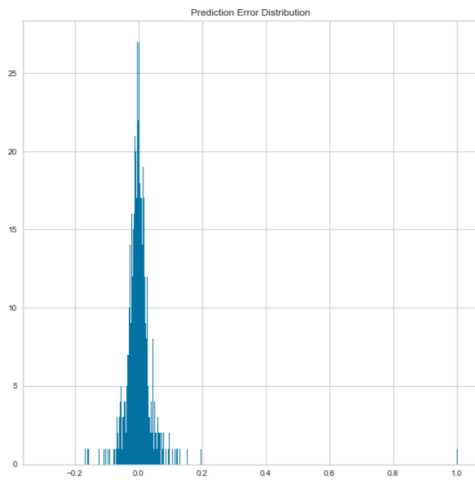


Figure 28: Distribuição de erros dos valores previstos com os  $p$  valores anteriores

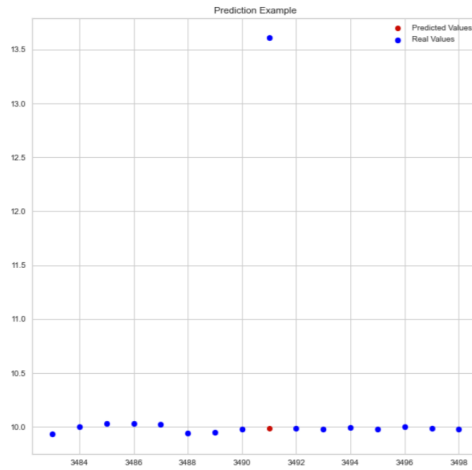


Figure 29: Exemplo de previsão 1 com os  $p$  valores anteriores

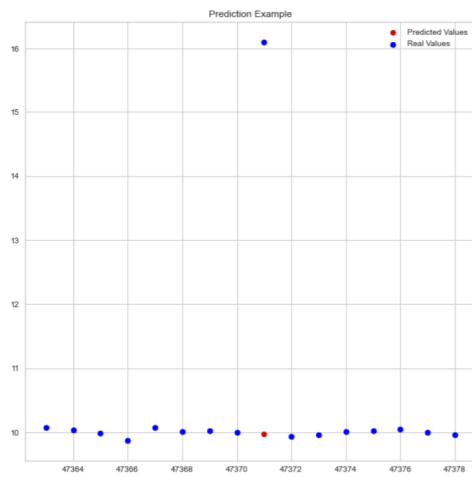


Figure 30: Exemplo de previsão 2 com os  $p$  valores anteriores

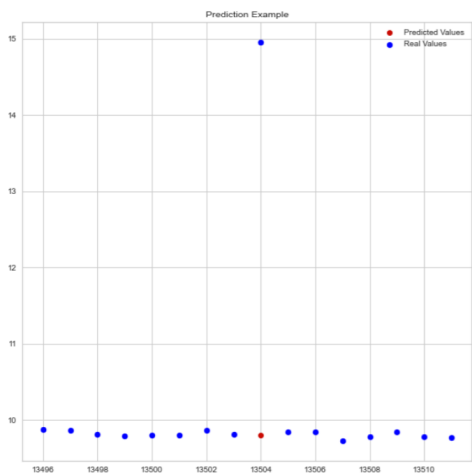


Figure 31: Exemplo de previsão 3 com os  $p$  valores anteriores

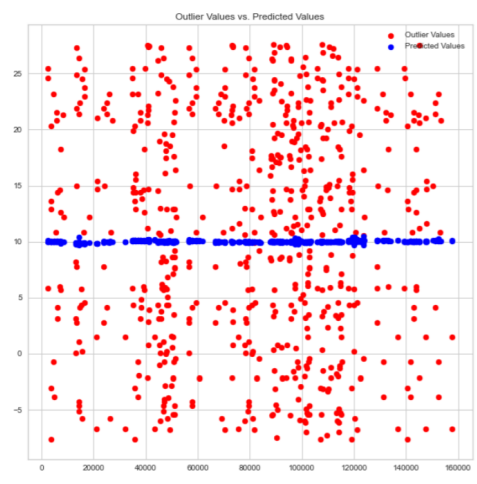


Figure 32: Previsão dos outliers com os  $p$  valores anteriores

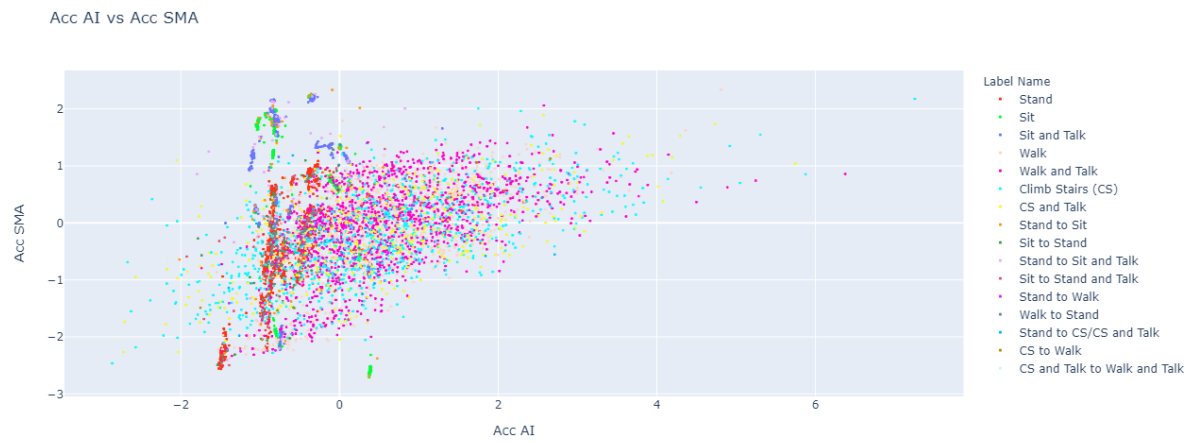


Figure 33: **Acelerómetro: AI vs. SMA**



Figure 34: **Acelerómetro: AVG vs. AVH**

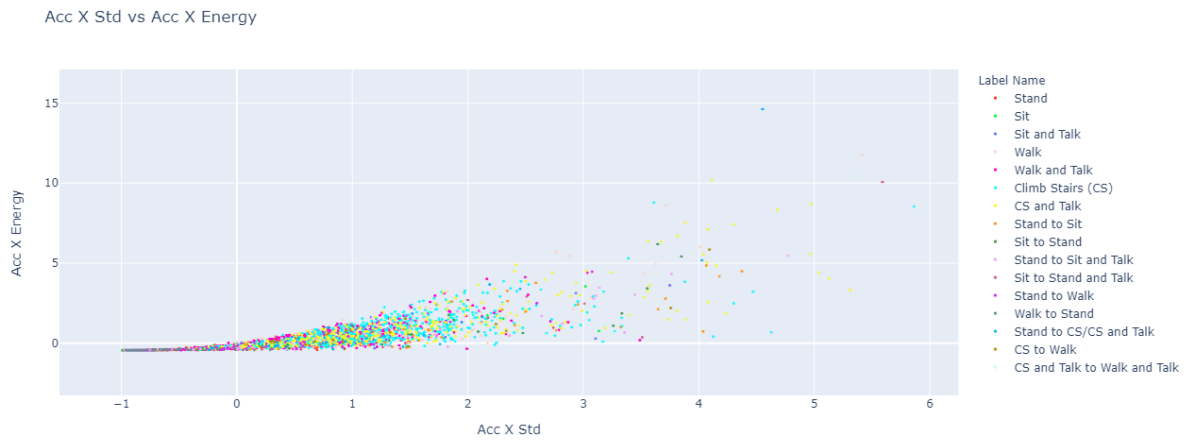


Figure 35: **Acelerómetro: X Std vs. X Energy**

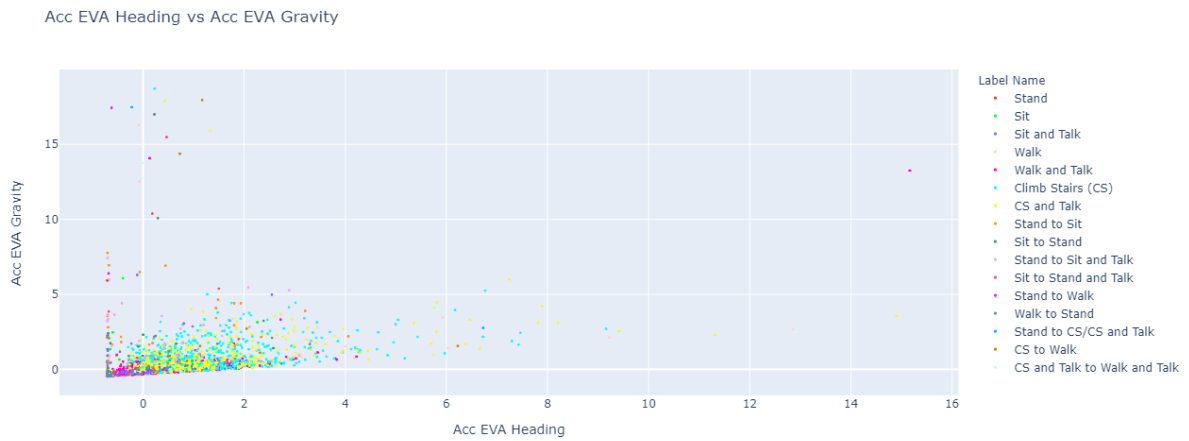


Figure 36: **Acelerómetro: EVA Heading vs. EVA Gravity**

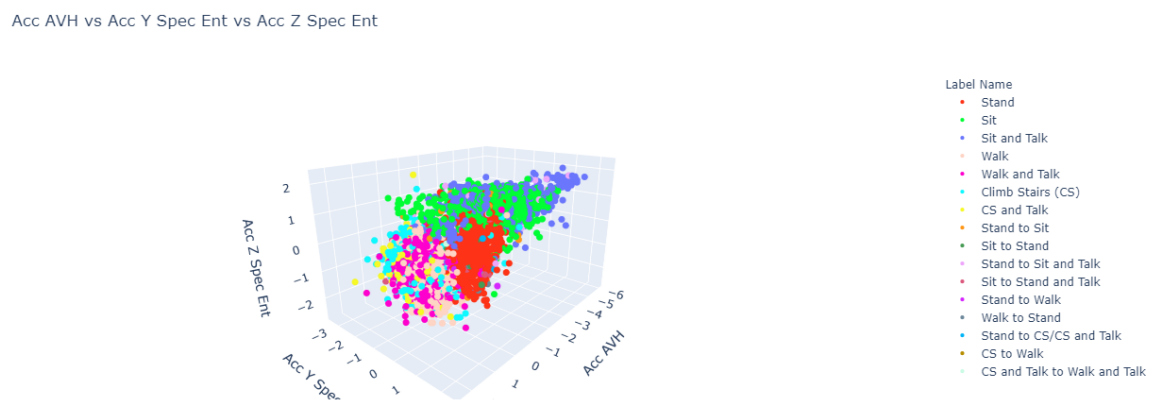


Figure 37: **Acelerómetro: AVH vs. Y Spectral Entropy vs. Z Spectral Entropy**

Acc Y Z Corr vs Acc Y Spec Ent vs Acc Z Spec Ent

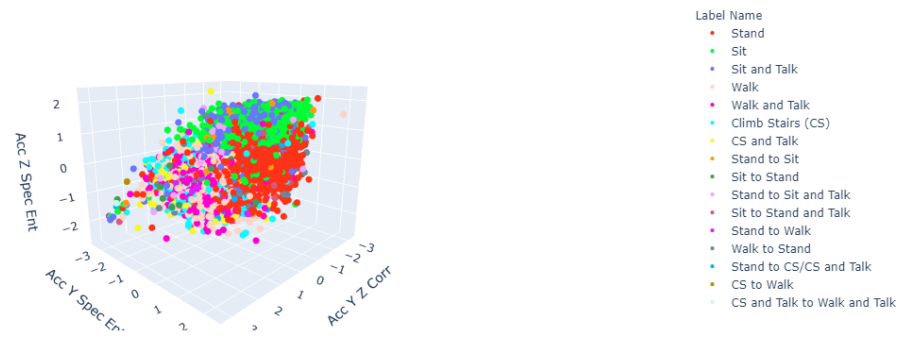


Figure 38: **Acelerómetro:** Y Z Correlation vs. Y Spectral Entropy vs. Z Spectral Entropy

Acc X MCR vs Gyr X MCR

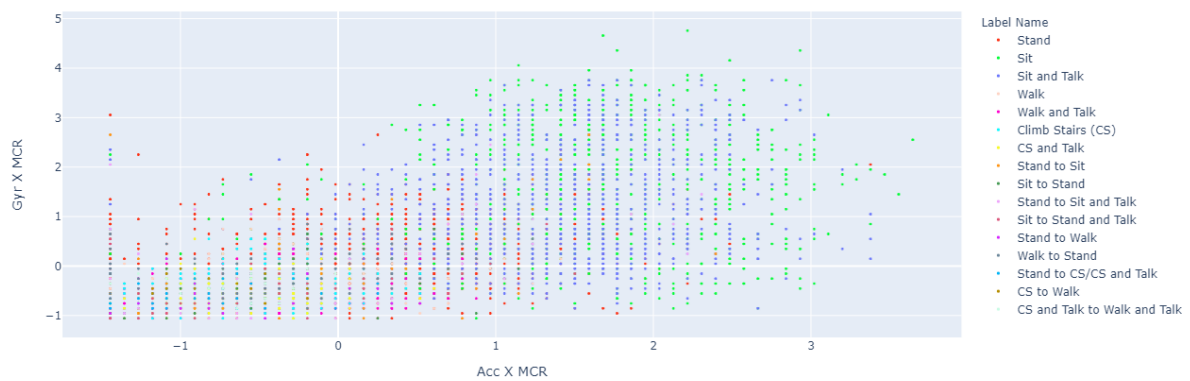


Figure 39: **Acelerómetro e Girómetro:** MCR

Acc X Mean vs Gyr X Mean



Figure 40: **Acelerómetro e Girómetro:** Mean