



**UFOP**

Universidade Federal  
de Ouro Preto

**Universidade Federal de Ouro Preto  
Instituto de Ciências Exatas e Aplicadas  
Departamento de Computação e Sistemas**

**Uso de técnicas de aprendizado de  
máquina para identificar a melhor  
disposição de sensores vestíveis para  
detecção de quedas**

**Arthur Bernardo Assumpção Pinto**

**TRABALHO DE  
CONCLUSÃO DE CURSO**

**ORIENTAÇÃO:**

Gilda Aparecida de Assis

**COORIENTAÇÃO:**

Luiz Carlos Bambirra Torres

**Dezembro, 2019  
João Monlevade–MG**

**Arthur Bernardo Assumpção Pinto**

**Uso de técnicas de aprendizado de máquina  
para identificar a melhor disposição de sensores  
vestíveis para detecção de quedas**

Orientador: Gilda Aparecida de Assis

Coorientador: Luiz Carlos Bambirra Torres

Monografia apresentada ao curso de Engenharia da Computação do Instituto de Ciências Exatas e Aplicadas, da Universidade Federal de Ouro Preto, como requisito parcial para aprovação na Disciplina “Trabalho de Conclusão de Curso II”.

**Universidade Federal de Ouro Preto**

**João Monlevade**

**Dezembro de 2019**

P659u

Pinto, Arthur Bernardo Assumpção.

Uso de técnicas de aprendizado de máquina para identificar a melhor disposição de sensores vestíveis para detecção de quedas [manuscrito] / Arthur Bernardo Assumpção Pinto. - 2019.

46f.: il.: color; tabs.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Gilda Aparecida de Assis.

Coorientador: Prof. Dr. Luiz Carlos Bambera Torres.

Monografia (Graduação). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Aplicadas. Departamento de Computação e Sistemas de Informação.

1. Quedas (Acidentes) em idosos - Detectores. 2. Aprendizado do computador. 3. Automação. I. Assis, Gilda Aparecida de. II. Torres, Luiz Carlos Bambera. III. Universidade Federal de Ouro Preto. IV. Título.

Catálogo: [ficha.sisbin@ufop.edu.br](mailto:ficha.sisbin@ufop.edu.br) CDU: 681.5



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DE OURO PRETO  
REITORIA  
INSTITUTO DE CIÊNCIAS EXATAS E APLICADAS  
DEPARTAMENTO DE COMPUTAÇÃO E SISTEMAS



**FOLHA DE APROVAÇÃO**

**Arthur Bernardo Assumpção Pinto**

**Uso de técnicas de aprendizado de máquina para identificar a melhor disposição de sensores vestíveis para detecção de quedas**

Membros da banca

Harlei Miguel de Arruda Leite - Doutor - UFOP  
Vicente José Peixoto de Amorim - Doutor - UFOP

Versão final  
Aprovado em 16 de dezembro de 2019

De acordo

Gilda Aparecida de Assis (orientador), Luiz Carlos Bambirra Torres (coorientador)



Documento assinado eletronicamente por **Gilda Aparecida de Assis, PROFESSOR DE MAGISTERIO SUPERIOR**, em 22/12/2019, às 19:25, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [http://sei.ufop.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **0030189** e o código CRC **C4349C89**.

**Referência:** Caso responda este documento, indicar expressamente o Processo nº 23109.204265/2019-91

SEI nº 0030189

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35400-000  
Telefone: - www.ufop.br

# Agradecimentos

Agradeço muito às duas pessoas mais importantes para mim, minha mãe Regina por seu amor incondicional e que nunca mediu esforços para me apoiar no que fosse preciso, e ao meu pai José Carlos apoio e inspiração. À Deus, pois até aqui o Senhor me ajudou, fazendo-se presente em todos os momentos. Agradeço aos meus tios, tias e primos por todo incentivo e ajuda em todos os momentos. A Bianca por fazer meus dias melhores com seu amor, companheirismo e compreensão. Aos demais familiares, minha gratidão por toda força e carinho. Agradeço também aos amigos que fiz nessa jornada, e que tornaram cada dia de luta mais leve, nunca me esquecerei de vocês. Minha gratidão aos professores por todo aprendizado, apoio e pela dedicação para transmitir seus ensinamentos da melhor forma possível . Guardarei sempre uma profunda gratidão por cada um que me ajudou nessa fase da vida, sem vocês concluir essa etapa não seria possível!

*“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê.”*

Arthur Schopenhauer

# Resumo

Com o constante crescimento do número de idosos na população atual, tornou-se necessário repensar e propor novas maneiras de se lidar com as demandas específicas da faixa etária, uma vez que as mudanças corporais e psico-sociais nessa idade exigem cuidado. Nesse sentido, as quedas dos idosos estão entre as principais causas de situações de difícil recuperação e mortes. Diante disso, o uso de aprendizagem de máquina junto com a captura de movimentos através de sensores tem sido cada vez mais utilizado para o desenvolvimento de sistemas detectores de quedas. Com isso, o presente trabalho busca utilizar um algoritmo baseado em aprendizagem de máquina para, a partir de dados de sensores vestíveis posicionados em diferentes regiões anatômicas, identificar as melhores posições do corpo humano para uso dos mesmos. Os resultados mostram que a maioria das regiões apresentam um bom desempenho para identificar as atividades diárias. Entretanto, quando trata-se de diferenciar as atividades das quedas, as maiores acurácias foram obtidas com os sensores posicionados no quadril direito e na cintura.

**Palavras-chaves:** quedas, sensores vestíveis, idosos, aprendizagem de máquina, regiões anatômicas.

# Abstract

With the growing number of elderly people in the current population, it has become necessary to rethink and propose new ways to deal with the specific demands of the age group, since body and psycho-social changes at this age require care. In this sense, the falls of the elderly are among the main causes of difficult recovery situations and deaths. Given this, the use of machine learning along with motion capture through sensors has been increasingly used for the development of fall detection systems. Thus, the present work seeks to use a machine learning algorithm to identify, from wearable sensors data placed in different anatomical regions, to identify the best positions of the human body for their use. Results show that most regions perform well to identify daily activities. However, when it comes to differentiating activities from falls, the highest accuracy was obtained with sensors positioned on the right hip and waist.

**Key-words:** falls, wearable sensors, elderly, machine learning, anatomical regions.



# Lista de ilustrações

Figura 1 – Localização dos sensores utilizados. . . . .	18
Figura 2 – Hierarquia da Aprendizagem de Máquina. . . . .	20
Figura 3 – Matriz de correlação entre os dados de todos os sensores. . . . .	25
Figura 4 – Matriz de correlação dos sensores do tornozelo. . . . .	26
Figura 5 – Matriz de correlação dos sensores da cintura. . . . .	27
Figura 6 – Matriz de correlação dos sensores do pescoço. . . . .	28
Figura 7 – Matriz de correlação dos sensores do quadril direito. . . . .	29
Figura 8 – Matriz de correlação dos sensores do pulso. . . . .	30
Figura 9 – Matriz de Confusão normalizada para os dados do tornozelo utilizando o <i>Random Forest</i> . . . . .	31
Figura 10 – Matriz de Confusão normalizada para os dados do tornozelo utilizando o <i>XGBoost</i> . . . . .	32
Figura 11 – Matriz de Confusão normalizada para os dados da cintura utilizando o <i>Random Forest</i> . . . . .	33
Figura 12 – Matriz de Confusão normalizada para os dados da cintura utilizando o <i>XGBoost</i> . . . . .	34
Figura 13 – Matriz de Confusão normalizada para os dados do pescoço utilizando o <i>Random Forest</i> . . . . .	35
Figura 14 – Matriz de Confusão normalizada para os dados do pescoço utilizando o <i>XGBoost</i> . . . . .	36
Figura 15 – Matriz de Confusão normalizada para os dados do quadril direito utilizando o <i>Random Forest</i> . . . . .	37
Figura 16 – Matriz de Confusão normalizada para os dados do quadril direito utilizando o <i>XGBoost</i> . . . . .	38
Figura 17 – Matriz de Confusão normalizada para os dados do pulso utilizando o <i>Random Forest</i> . . . . .	39
Figura 18 – Matriz de Confusão normalizada para os dados do pulso utilizando o <i>XGBoost</i> . . . . .	40

# Lista de tabelas

Tabela 1 – Estatísticas dos Indivíduos . . . . .	17
Tabela 2 – Atividades realizadas pelos indivíduos . . . . .	17
Tabela 3 – Exemplo de matriz de confusão . . . . .	22
Tabela 4 – Desempenho geral dos algoritmos para a região do tornozelo . . . . .	32
Tabela 5 – Desempenho geral dos algoritmos para a região da cintura . . . . .	34
Tabela 6 – Desempenho geral dos algoritmos para a região do pescoço . . . . .	36
Tabela 7 – Desempenho geral dos algoritmos para a região do quadril direito . . . . .	38
Tabela 8 – Desempenho geral dos algoritmos para a região do pulso . . . . .	40
Tabela 9 – Localização com melhor acurácia . . . . .	41
Tabela 10 – Localização com melhor acurácia . . . . .	42

# Lista de abreviaturas e siglas

cm	Centímetros
CPU	<i>Central Processing Unit</i>
EEG	Eletroencefalógrafo
FN	Falsos Negativos
FP	Falsos Positivos
GB	<i>Gigabyte</i>
IA	Inteligência Artificial
Kg	Quilograma
m	Metros
OMS	Organização Mundial da Saúde
RAM	<i>Random Access Memory</i>
VN	Verdadeiros Negativos
VP	Verdadeiros Positivos
<i>XGBoost</i>	<i>Extreme Gradient Boost</i>

# Lista de símbolos

$R^+$	Classificações com diferenças positivas
$R^-$	Classificações com diferenças negativas
$d_i$	Diferença de desempenho entre dois classificadores
$\alpha$	Letra grega Alfa

# Sumário

	<b>Lista de abreviaturas e siglas</b> . . . . .	<b>10</b>
<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>13</b>
1.1	O problema de pesquisa . . . . .	14
1.2	Objetivos . . . . .	15
1.3	Organização do trabalho . . . . .	15
<b>2</b>	<b>MATERIAIS E MÉTODOS</b> . . . . .	<b>16</b>
2.1	Metodologia . . . . .	16
2.2	<i>UP-Fall Detection Dataset</i> . . . . .	16
2.2.1	Indivíduos e Atividades . . . . .	16
2.2.2	Sensores e Distribuição . . . . .	17
<b>3</b>	<b>REVISÃO BIBLIOGRÁFICA</b> . . . . .	<b>19</b>
3.1	<b>Aprendizagem de Máquina</b> . . . . .	<b>19</b>
3.1.1	Aprendizado Supervisionado . . . . .	19
3.1.2	Aprendizado Não-Supervisionado . . . . .	19
3.2	<i>XGBoost</i> . . . . .	20
3.3	<i>Random Forest</i> . . . . .	21
3.4	<b>Métodos de avaliação</b> . . . . .	<b>21</b>
3.4.1	Matriz de Confusão . . . . .	21
3.4.2	Teste de Wilcoxon . . . . .	22
3.4.3	Métricas . . . . .	23
<b>4</b>	<b>RESULTADOS E ANÁLISES</b> . . . . .	<b>24</b>
4.1	<b>Análise Exploratória</b> . . . . .	<b>24</b>
4.2	<b>Resultados</b> . . . . .	<b>30</b>
4.2.1	Tornozelo . . . . .	31
4.2.2	Cintura . . . . .	32
4.2.3	Pescoço . . . . .	34
4.2.4	Quadril Direito (Bolso) . . . . .	36
4.2.5	Pulso . . . . .	38
4.3	<b>Análise</b> . . . . .	<b>40</b>
<b>5</b>	<b>CONCLUSÃO</b> . . . . .	<b>43</b>

**REFERÊNCIAS** ..... 44

# 1 Introdução

O envelhecimento é um processo natural da vida do ser humano, no qual ocorrem mudanças físicas, psicológicas e sociais que afetam de forma particular cada indivíduo. Atualmente, esse processo atinge a população mundial em larga escala, tornando-se um dos grandes desafios a serem enfrentados no século XXI, já que esse grupo demanda atenções sociais especiais. Segundo a OMS (OMS, 2002), são definidos como idosos aqueles indivíduos que possuem 60 anos ou mais. De acordo com o Censo Populacional de 2010, os brasileiros com 60 anos ou mais já somam pouco mais de 20,5 milhões de indivíduos, representando 10,8% da população brasileira (IBGE, 2011).

Nesse contexto, deve-se ressaltar que o aumento da expectativa de vida deve ser acompanhado de precauções contínuas nos quesitos de saúde, segurança e participação social. Um dos fatores que exigem tais precauções advém do fato de aproximadamente 30% das pessoas com 65 anos ou mais terem vivenciado um evento de queda a cada ano, e metade dos idosos que caem repetem o evento (PEREIRA et al., 2001). As lesões decorrentes dessas quedas são responsáveis pela sexta maior causa de morte nessa faixa etária, sendo que cerca de 5% das quedas resultam em lesões graves e ocasionam mais de 200 mil hospitalizações por fratura de quadril a cada ano (PEREIRA et al., 2001).

Segundo (OLIVEIRA et al., 2011) os episódios de queda podem estar relacionados a processos intrínsecos ao envelhecimento, como por exemplo, a redução da acuidade visual e diminuição de força muscular. Pode-se citar também a relação com fatores extrínsecos ao utilizarem mobiliário doméstico inadequado e mal posicionado, ou mesmo ao adotarem hábitos perigosos como subir em escadas móveis para alcançar armários altos. Segundo (OMS, 2005), a maioria das quedas ocorre no ambiente doméstico e devem ser evitadas, uma vez que as consequências das lesões sofridas em uma idade mais avançada são mais graves do que entre pessoas jovens.

Entretanto, discriminar eventos de quedas e atividade do dia a dia não é uma tarefa fácil visto que algumas atividades diárias, como abaixar ou sentar, podem ser facilmente confundidas com quedas (IGUAL; MEDRANO; PLAZA, 2013). Dessa forma, para implementar e testar um sistema computacional que detecta quedas em tempo real se faz necessário a coleta de dados de atividades diárias e quedas provenientes de vários sensores. Dentre estes câmeras, sensores infravermelhos, sensores de pressão plantar e dispositivos vestíveis os quais podem estar espalhados por um ambiente controlado ou, como no caso dos vestíveis, pelo corpo.

Além disso, quando tratamos da detecção de quedas um cenário real, devem ser levadas em consideração as dificuldades física e ambientais e o alto custo de se ter uma

pessoa totalmente monitorada por câmeras, sensores e dispositivos vestíveis. Visto que, este último, apesar de mais barato que os outros dispositivos, possivelmente trará um desconforto, principalmente para idosos. Sendo assim, este trabalho visa realizar um estudo direcionado aos sensores vestíveis com o objetivo de identificar uma ou mais regiões do corpo humano onde estes sensores possam ser melhores utilizados, com a finalidade de implementar um detector de quedas de baixo custo e mais confortável para os usuários. Os resultados mostram que as regiões da cintura e quadril direito apresentam os melhores resultados para diferenciar quedas das demais atividades.

## 1.1 O problema de pesquisa

De acordo com a OMS (OMS, 2005), em 2025 haverá aproximadamente um bilhão e duzentos milhões de pessoas idosas, e até 2050 serão dois bilhões, com 80% destes nos países em desenvolvimento, como o Brasil. Dessa forma, o rápido crescimento do número de idosos no país nos últimos anos gera um aumento da necessidade da atenção e de cuidados com a saúde deste grupo de faixa etária avançada. Uma vez que, alterações crônicas e degenerativas próprias da idade contribuem para o aumento de quedas e dificuldades de locomoção.

Algumas condições comuns nesses indivíduos são a presença de desequilíbrio postural, alterações sensoriais, motoras e dificuldade de locomoção. Dentre as consequências desses problemas, destacam-se as quedas, que representam a primeira causa de acidentes em pessoas com mais de 60 anos de idade (MORAIS et al., 2012). Dessa forma, detectores de quedas têm sido amplamente estudados por pesquisadores.

Segundo (IGUAL; MEDRANO; PLAZA, 2013), um sistema de detecção de queda pode ser definido como um dispositivo assistivo cujo objetivo principal é alertar quando um evento de queda ocorre. Entretanto, discriminar eventos de quedas entre as atividades do dia a dia não é uma tarefa fácil, visto que algumas atividades diárias, como abaixar ou sentar, podem ser facilmente confundidas como quedas. Atualmente, de acordo com (MUBASHIR; SHAO; SEED, 2013), abordagens para detecção de quedas podem ser divididas em três categorias diferentes: baseadas em dispositivos vestíveis, baseadas sensores ambientais e baseadas em câmeras.

Contudo, quando trata-se de um cenário real, o uso de câmeras e sensores ambientais na maioria das vezes se torna inviável fisicamente e financeiramente. Do mesmo modo, a utilização de múltiplos dispositivos vestíveis espalhados pelo corpo durante o dia a dia seria um inconveniente, embora os mesmos tenham preços mais acessíveis.



## 1.2 Objetivos

O objetivo principal do trabalho é realizar um estudo direcionado ao uso de sensores vestíveis para detecção de eventos de quedas com o objetivo de identificar um ou mais locais onde esses possam ser melhor utilizados, com a finalidade de implementar um detector de quedas com boa acurácia de baixo custo e mais confortável aos usuários. A partir disso, pretende-se contribuir para o monitoramento e acompanhamento de idosos e/ou pessoas mais debilitadas.

Os objetivos específicos do trabalho são:

- Definir um conjunto de dados relevante que será utilizado como entrada para o classificador.
- Selecionar a técnica de aprendizagem de máquina que melhor se adéque à natureza dos dados utilizados.
- Validar e apresentar o desempenho do classificador por meio de técnicas de validação e de medida de desempenho existentes.
- Identificar a melhor posição anatômica a ser usada em um detector de quedas.

## 1.3 Organização do trabalho

O restante deste trabalho é organizado como se segue. No Capítulo 2 são apresentados os materiais utilizados e uma breve descrição da base de dados utilizada. O Capítulo 3 apresenta alguns conceitos básicos sobre aprendizagem de máquina, bem como a descrição dos algoritmos de classificação utilizados e trabalhos correlatos. No Capítulo 4 são apresentados os resultados obtidos e por fim, no Capítulo 5 são apresentadas as conclusões acerca do trabalho.

## 2 Materiais e Métodos

### 2.1 Metodologia

As etapas para a execução desse trabalho podem ser definidas como:

1. Revisar a literatura sobre a aprendizagem de máquina, bem como trabalhos correlatos na identificação de padrões em biomecânica.
2. Realizar uma análise exploratória para visualizar o comportamento dos dados.
3. Selecionar as técnicas de aprendizagem de máquina que melhor se adéquem à natureza dos dados e implementar o classificador.
4. Realizar a validação do classificador implementado.
5. Realizar os testes de desempenho e comparativos entre as técnicas.
6. Analisar e discutir os resultados obtidos, além de identificar possíveis melhorias.

Para realizar este trabalho, foi utilizado a base de dados *UP-Fall Detection Dataset* (MARTÍNEZ-VILLASEÑOR et al., 2019) disponibilizada em um artigo publicado na revista *Sensors*, revista internacional sobre a ciência e tecnologia de sensores e bio-sensores.

Os testes foram feitos no ambiente computacional em nuvem da plataforma Kaggle, a qual conta com um Intel<sup>®</sup> Xeon<sup>®</sup> CPU @ 2.30GHz, 16 GB de memória RAM e 5 GB de capacidade de disco. Os algoritmos de classificação foram implementados na linguagem *Python* (versão 3.7) utilizando a biblioteca *Scikit-learn* (PEDREGOSA et al., 2011).

### 2.2 *UP-Fall Detection Dataset*

A base de dados contém dados brutos e atributos adquiridos de 17 indivíduos jovens e saudáveis que realizaram 6 tipos atividades de atividades diárias e 5 tipos de quedas, cada uma repetida três vezes. A base de dados compreende mais de 850 GB de informações de sensores vestíveis, sensores ambientais e dispositivos de visualização.

#### 2.2.1 Indivíduos e Atividades

Para a coleta de dados, 17 indivíduos, dentre eles 9 homens e 8 mulheres, com idades entre 18 e 24 anos, peso médio de 66,8 Kg e altura média de 1,66 m foram convidados para realizarem as 11 diferentes atividades. A Tabela 1 sumariza os dados dos indivíduos.

As atividades realizadas são seis do dia a dia (andar, levantar, pegar um objeto, sentar, pular e deitar) e cinco tipos de quedas (cair para frente usando as mãos, cair para frente usando os joelhos, cair para trás, cair sentando em uma cadeira vazia e cair de lado). Foram selecionadas as atividades mais comuns do dia a dia e em particular, “pegar um objeto” foi incluída devido ao fato de que a mesma é frequentemente confundida com uma queda. A Tabela 2 resume os dados das atividades.

Tabela 1 – Estatísticas dos Indivíduos.

ID do Indivíduo	Idade	Altura (m)	Peso (Kg)	Gênero	Lado Dominante
1	18	1.70	99	Masculino	Destro
2	20	1.70	58	Masculino	Destro
3	19	1.57	54	Feminino	Canhoto
4	20	1.62	71	Feminino	Destro
5	21	1.71	69	Masculino	Destro
6	22	1.62	68	Masculino	Destro
7	24	1.74	70	Masculino	Destro
8	23	1.75	88	Masculino	Destro
9	23	1.68	70	Feminino	Destro
10	19	1.69	63	Masculino	Destro
11	20	1.65	73	Feminino	Destro
12	19	1.60	53	Feminino	Destro
13	20	1.64	55	Masculino	Destro
14	19	1.70	73	Feminino	Destro
15	21	1.57	56	Feminino	Destro
16	20	1.70	62	Masculino	Destro
17	20	1.66	54	Feminino	Destro

Fonte: [Martínez-Villaseñor et al. \(2019\)](#)

Tabela 2 – Atividades realizadas pelos indivíduos.

Atividade ID	Descrição
1	Cair para frente usando as mãos
2	Cair para frente usando os joelhos
3	Cair para trás
4	Cair de lado
5	Cair sentado em uma cadeira vazia
6	Andar
7	Levantar
8	Sentar
9	Pegar um objeto
10	Pular
11	Deitar

Fonte: [Martínez-Villaseñor et al. \(2019\)](#)

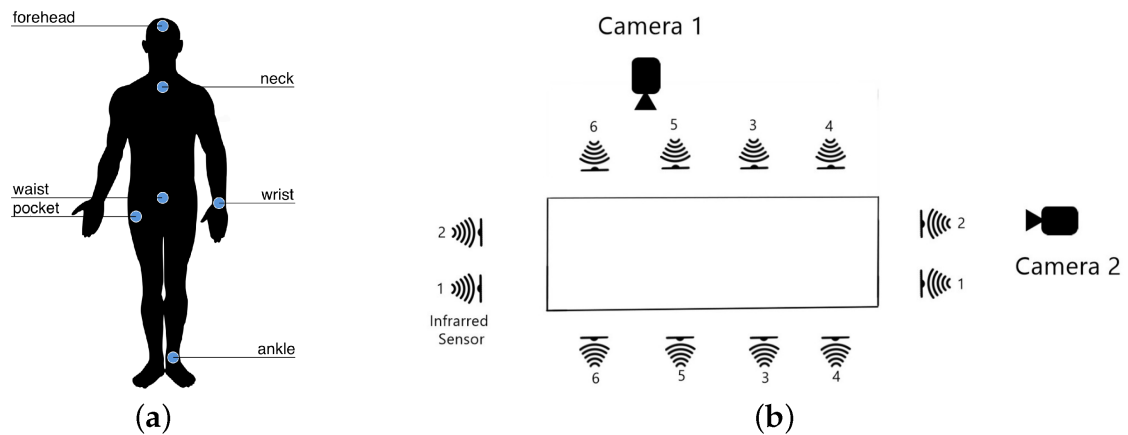
## 2.2.2 Sensores e Distribuição

Para adquirir os dados foi adotada uma abordagem multimodal utilizando três tipos diferentes de sensores: vestíveis, câmeras e sensores infravermelhos. Foram usados

cinco sensores vestíveis *Mbientlab MetaSensor* coletando dados brutos de acelerômetro de 3 eixos, giroscópio de 3 eixos e sensor de luz ambiente. Os vestíveis estavam localizados no pulso esquerdo, sob o pescoço, no bolso direito da calça, no meio da cintura (no cinto) e no tornozelo esquerdo. Além disso, um eletroencefalógrafo (EEG) *NeuroSky MindWave*, localizado na testa, foi encarregado de medir sinais de ondas cerebrais.

No caso dos sensores infravermelhos foram instalados seis, posicionados a 40 cm do chão. Por fim, duas câmeras *Microsoft LifeCam Cinema* foram posicionadas a 1,82 m acima do chão, uma para captar uma visão frontal e outra para captar uma visão lateral. A Figura 1(a) mostra as posições anatômicas utilizadas na aquisição dos dados e a Figura 1(b) esquematiza as câmeras e sensores infravermelho dispostos no ambiente.

Figura 1 – Localização dos sensores utilizados.



Fonte: (MARTÍNEZ-VILLASEÑOR et al., 2019)

## 3 Revisão bibliográfica

### 3.1 Aprendizagem de Máquina

O termo Aprendizagem de Máquina (*Machine Learning*), caracteriza um ramo da Inteligência Artificial (IA), cujo objetivo é o desenvolvimento de técnicas para que computadores possam aprender ou melhorar seu desempenho de decisão e classificação baseada em dados. Com isso, há a possibilidade de construção de sistemas capazes de adquirir conhecimento de forma automática através de padrões significativos em dados (MONARD; BARANAUSKAS, 2003).

As tarefas de aprendizado de máquina podem ser divididas em três categorias principais: supervisionado, não-supervisionado e semi-supervisionado, Figura 2. Tais grupos se diferenciam pela presença ou ausência do atributo de classe que rotulam os exemplos. Na categoria supervisionada é fornecido ao algoritmo um conjunto de exemplos nos quais o atributo de classe associado é conhecido. De forma contrária, no aprendizado não-supervisionado essa informação sobre a classe é desconhecida (MONARD; BARANAUSKAS, 2003). Já o aprendizado semi-supervisionado engloba fatores das duas categorias anteriores.

#### 3.1.1 Aprendizado Supervisionado

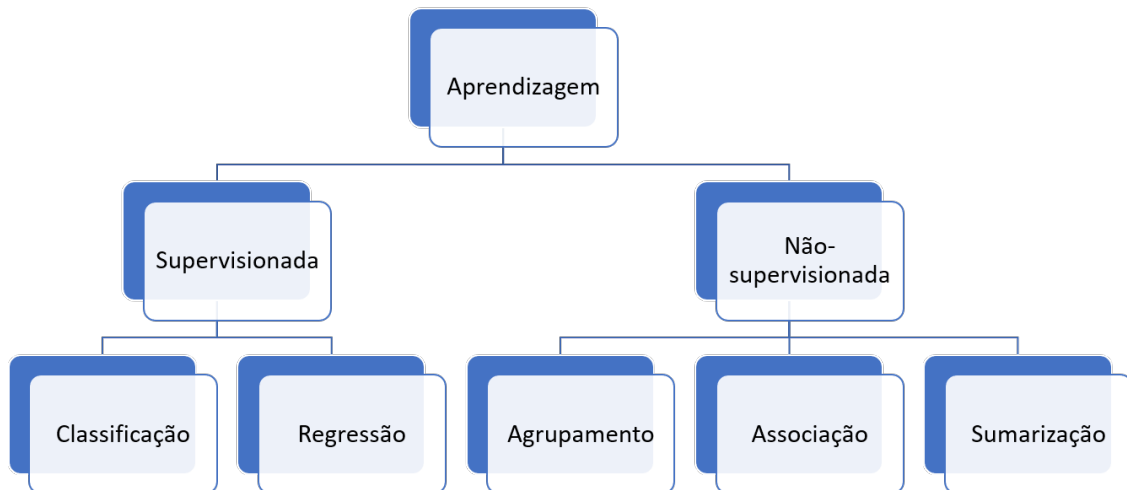
Na aprendizagem supervisionada, juntamente com o conjunto de dados obtidos temos também o resultado esperado e, de maneira implícita, uma relação entre entrada e saída. Um conjunto de dados para um algoritmo de aprendizagem supervisionada pode ser escrito na forma  $(x_i, y_i)$ , onde  $x_i$  é um conjunto de características de uma dada observação e  $y_i$  é o seu rótulo. Através desse conjunto de dados é gerado um classificador capaz de prever o rótulo de novas entradas. Os rótulos, também chamados de classes, são uma representação dos fatos sobre os quais desejamos realizar algum tipo de previsão. Quando os rótulos assumem valores discretos temos um problema de classificação. De maneira similar, quando os rótulos são valores contínuos temos uma regressão (LORENA; CARVALHO, 2007).

#### 3.1.2 Aprendizado Não-Supervisionado

De forma contrária, o aprendizado não supervisionado, não nos permite dizer previamente qual resultado final devemos esperar, já que o mesmo não possui rótulos. Dessa forma o algoritmo irá avaliar os dados e buscar por grupos que apresentem características semelhantes. Tal tarefa de agrupar os dados similares é conhecida como clusterização. Essa

análise nos possibilita extrair informações importantes de dados que anteriormente não apresentavam explicitamente qualquer relação entre si (HARRINGTON, 2012).

Figura 2 – Hierarquia da Aprendizagem de Máquina.



Fonte: (MONARD; BARANAUSKAS, 2003)

## 3.2 *XGBoost*

O *Extreme Gradient Boost* (*XGBoost*) (CHEN; GUESTRIN, 2016) é uma implementação eficiente e escalável do *framework Gradient Boosting* (FRIEDMAN, 2001). Seu desempenho é amplamente reconhecido em vários desafios de aprendizado de máquina e mineração de dados hospedados pelo site de competições *Kaggle*. Dentre 29 soluções vencedoras de desafios publicadas no blog da *Kaggle* em 2015, 17 soluções utilizaram o *XGBoost*.

Segundo (CHEN; GUESTRIN, 2016), o sucesso do *XGBoost* muito se deve a sua escalabilidade em todos os cenários. O sistema é capaz de executar cerca de 10 vezes mais rápido que soluções populares existentes em uma única máquina e escala para bilhões de exemplos em configurações distribuídas ou com limite de memória. De acordo com (CHEN; GUESTRIN, 2016), a escalabilidade do *XGBoost* se deve a vários sistemas importantes e otimizações algorítmicas. Essas inovações incluem:

- Um novo algoritmo de aprendizado em árvore é para lidar com dados esparsos;
- Um procedimento de esboço quantil ponderado teoricamente justificado permite lidar com pesos de instância no aprendizado aproximado das árvores.

A computação paralela e distribuída torna o aprendizado mais rápido, o que permite uma exploração mais rápida do modelo.

### 3.3 *Random Forest*

O classificador *Random Forest* consiste em uma combinação de classificadores de árvores, onde cada classificador é gerado usando um vetor aleatório amostrado independentemente do vetor de entrada, e cada árvore vota na classe mais popular para classificar um vetor de entrada (BREIMAN, 1999). Quaisquer exemplos são classificados pela classe votada mais popular de todos os preditores de árvores da floresta (BREIMAN, 1999). Uma árvore de decisão exige a escolha de uma medida de seleção de atributo e um método de remoção.

Existem muitas abordagens para a seleção de atributos usados para a indução de uma árvore de decisão, e a maioria delas atribui uma medida de qualidade diretamente ao atributo. As medidas de seleção de atributos usadas com mais frequência na indução de árvore de decisão são o critério *Information Gain Ratio* e o Índice de Gini. O *Random Forest* usa o Índice Gini como uma medida de seleção de atributo, que mede a impureza de um atributo em relação às classes. Segundo (BREIMAN, 1999), à medida que o número de árvores aumenta, o erro de generalização sempre converge mesmo sem podar a árvore.

Em cada nó, apenas os recursos selecionados são pesquisados para a melhor divisão. Dessa forma, o *Random Forest* consiste em  $N$  árvores, onde  $N$  é o número de árvores a serem cultivadas, que pode ser qualquer valor definido pelo usuário. Para classificar um novo conjunto de dados, cada caso do conjunto de dados é passado para cada uma das  $N$  árvores e assim a floresta escolhe uma classe com o máximo de  $N$  votos, para esse caso.

### 3.4 Métodos de avaliação

Nessa seção são apresentadas os métodos de avaliação utilizados para analisar o desempenho dos classificadores.

#### 3.4.1 Matriz de Confusão

Um método comum para a avaliação de modelos classificadores é a matriz de confusão. Esta ferramenta consiste em uma matriz dos valores preditos pelo classificador e dos

valores reais de cada classificação. É utilizada para mostrar a quantidade de classificações corretas e incorretas de um modelo.

Tabela 3 – Exemplo matriz de confusão.

Previsto	Real	
	Classe 1	Classe 2
Classe 1	5	0
Classe 2	2	3

A matriz funciona da seguinte forma: os índices das linhas indicam a classe prevista pelo classificador enquanto os índices da coluna indicam qual a classe o classificador deveria indicar. Na Tabela 3 tem-se de exemplo que um classificador previu cinco vezes a classe 1 corretamente. Já para a classe 2, o classificador previu de forma errônea 2 observações como sendo da classe 1 e de forma correta 3 observações como sendo da classe 2. A matriz de confusão apresenta quatro categorias de valores:

- Verdadeiros Positivos (VP): é a quantidade de valores que o modelo previu corretamente para cada classe, ou seja, a quantidade de valores que o modelo previu que realmente pertenciam a classe prevista.
- Verdadeiros Negativos (VN): é a quantidade de valores que o modelo previu corretamente para as outras classes, ou seja, a quantidade de valores que o modelo previu que realmente pertenciam a classe as outras classes.
- Falsos Positivos (FP): é a quantidade de valores que o modelo previu de forma errônea para cada classe, ou seja, a quantidade de valores que o modelo previu como sendo de uma classe mas que na realidade não pertenciam a ela.
- Falsos Negativos (FN): é a quantidade de valores que o modelo previu como sendo de outras classes de forma, ou seja, a quantidade de valores que o modelo previu como sendo de outras classes mas que na realidade não pertenciam a elas.

### 3.4.2 Teste de Wilcoxon

O teste de Wilcoxon ([WILCOXON, 1992](#)) classifica a diferença de dois classificadores ignorando os sinais e comparando a classificação para as diferenças negativas e positivas. É utilizado para testar a hipótese de que as diferenças entre 2 conjunto de valores são bem próximas de zero, ou seja, não existem diferenças entre eles. Seja  $d_i$  a  $i$ -ésima diferença entre desempenho de dois classificadores entre  $N$  *datasets*. As diferenças são classificadas de acordo com seus valores absolutos. Seja  $R^+$  a soma das classificações para os conjuntos de dados nos quais o segundo algoritmo superou o primeiro, e  $R^-$  a soma das classificações



para o contrario. Classificações de  $d_i = 0$  são divididos igualmente entre as somas; se houver um número ímpar deles, um será ignorado:

$$R^+ = \sum_{d_i > 0} rank(d_i) + \frac{1}{2} \sum_{d_i = 0} rank(d_i) \quad R^- = \sum_{d_i < 0} rank(d_i) + \frac{1}{2} \sum_{d_i = 0} rank(d_i)$$

Seja  $T$  a menor da somas,  $T = \min(R^+, R^-)$  (DEMŠAR, 2006).

### 3.4.3 Métricas

Para avaliar o desempenho geral dos classificadores foram utilizadas as seguintes métricas:

- **Acurácia:** A acurácia é uma métrica que informa a performance geral do modelo. Ela indica dentre todas as classificações, quantas o modelo classificou corretamente.

$$acurácia = \frac{VP+VN}{VP+VN+FP+FN}$$

- **Precisão:** A precisão nos indica dentre todas as classificações de classe Positivo que o modelo fez, quantas estão corretas.

$$precisão = \frac{VP}{VP+FP}$$

- **Sensibilidade:** A sensibilidade indica dentre todas as situações de classe Positiva como valor esperado, quantas estão corretas.

$$sensibilidade = \frac{VP}{VP+FN}$$

- **F-1 Score:** O *F-1 Score* é uma média harmônica entre precisão e a sensibilidade.

$$F-1 \text{ score} = 2 * \frac{precisão * sensibilidade}{precisão + sensibilidade}$$

## 4 Resultados e Análises

### 4.1 Análise Exploratória

Antes da utilização dos algoritmos de aprendizagem de máquina para classificar os dados, foi feita uma breve análise dos dados com o objetivo de conhecer melhor o comportamento dos mesmos. Sendo assim, foram feitos testes de correlação entre os dados dos sensores de todas as localidades e também testes de correlação entre os dados de cada localidade separadamente.

Observa-se na Figura 3 que dois pares de dados apresentam valores de correlação muito próximos de 1, sendo o primeiro par os dados do eixo Z do acelerômetro localizado na cintura e no quadril direito e o segundo os dados do eixo Z do acelerômetro localizado no pescoço e na cintura. Valores próximos a 1 indicam uma correlação positiva, ou seja, os dados em cada par crescem na mesma proporção.

Verifica-se também que os dados do eixo Y do acelerômetro localizado no pulso e o eixo X do acelerômetro localizado no pescoço apresentam um valor de correlação próximo a -1. Valores próximos a -1 indicam uma correlação negativa, ou seja, os dados decrescem na mesma proporção. A Figura 3 apresenta a matriz de correlação da totalidade dos dados. Valores próximos de 1 tendem ao azul escuro, valores próximos a -1 tendem ao vermelho e valores próximos a 0 tendem ao branco.



Entretanto, ao analisarmos os dados agrupados pela localização de cada sensor, nota-se que a correlação entre eles é muito próxima de zero, ou seja, os dados não apresentam crescimento e decréscimo semelhantes. As Figuras 4, 5, 6, 7 e 8 apresentam as matrizes de correlação para os sensores de cada localidade. Valores próximos de 1 tendem ao azul escuro, valores próximos a -1 tendem ao vermelho e valores próximos a 0 tendem ao branco.

Figura 4 – Matriz de correlação dos sensores do tornozelo.

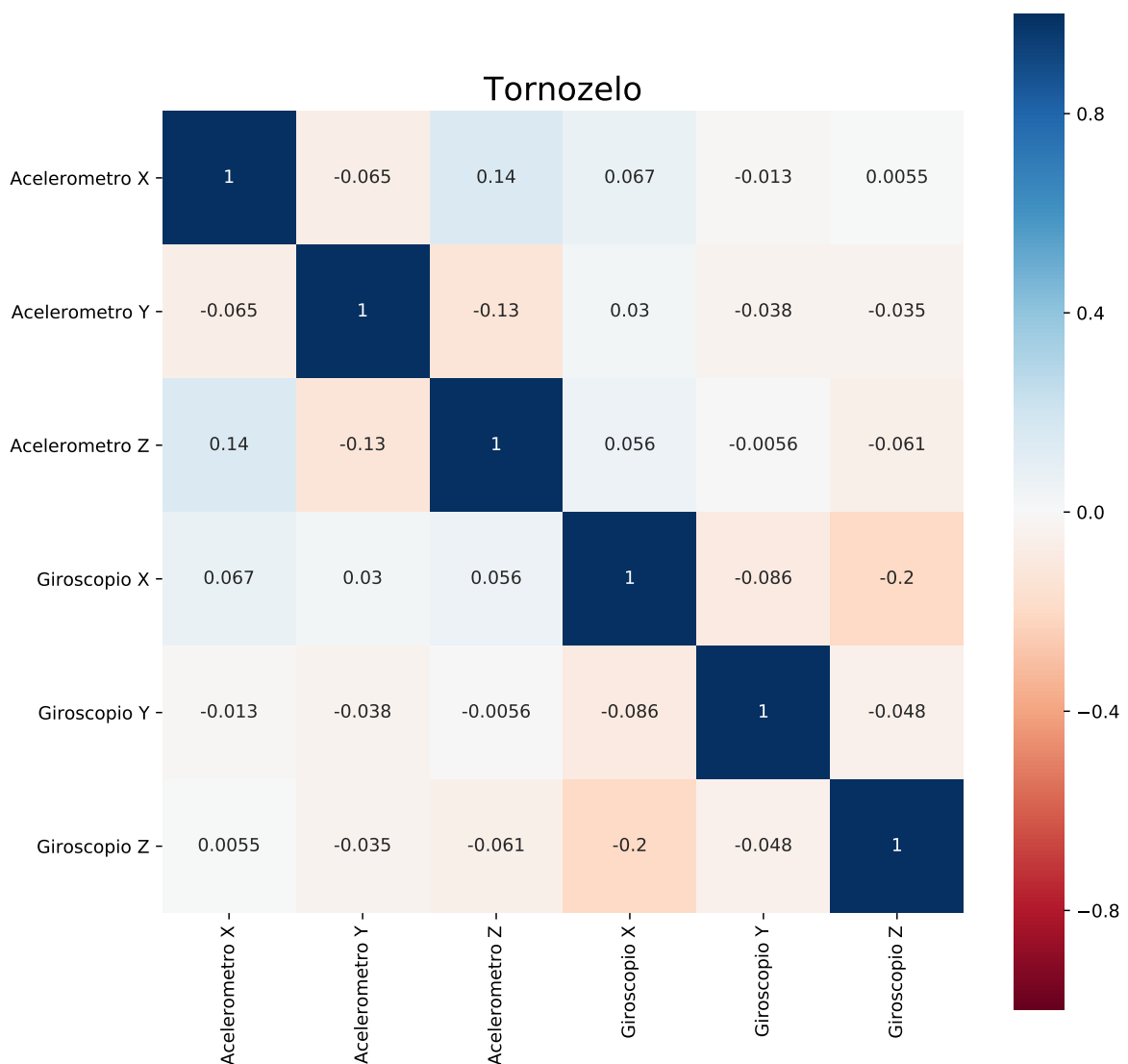


Figura 5 – Matriz de correlação dos sensores da cintura.

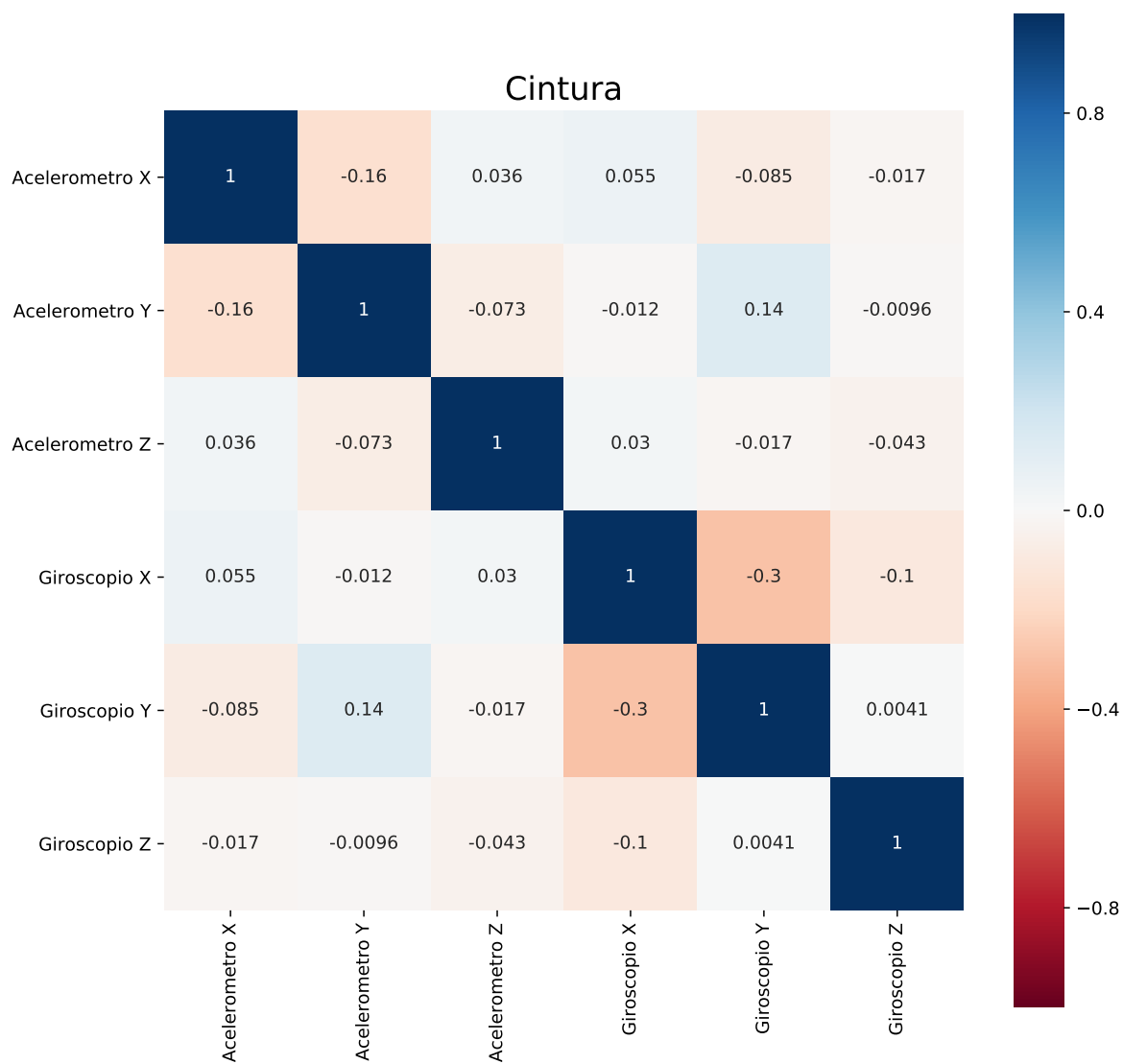


Figura 6 – Matriz de correlação dos sensores do pescoço.

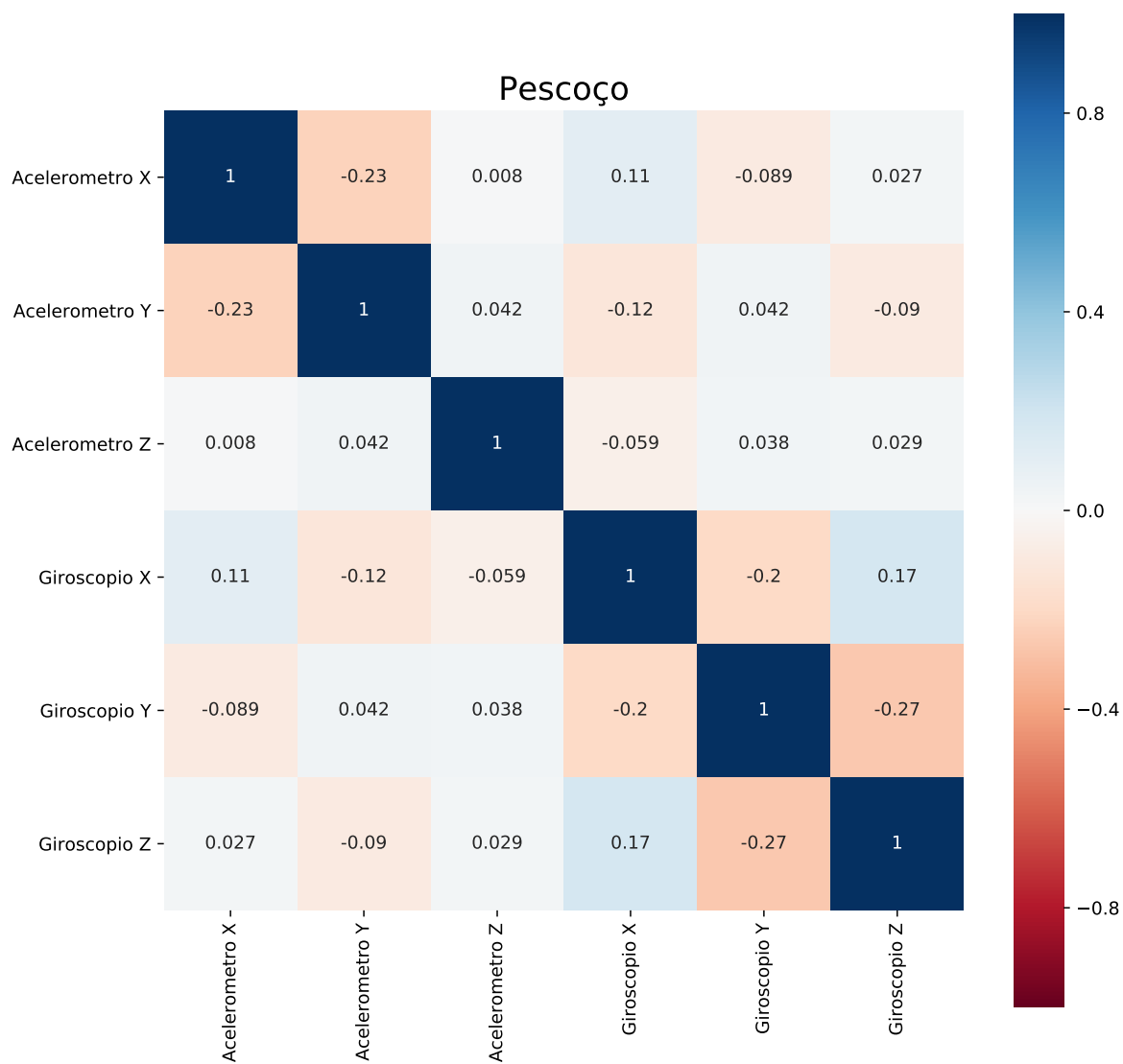


Figura 7 – Matriz de correlação dos sensores do quadril direito.

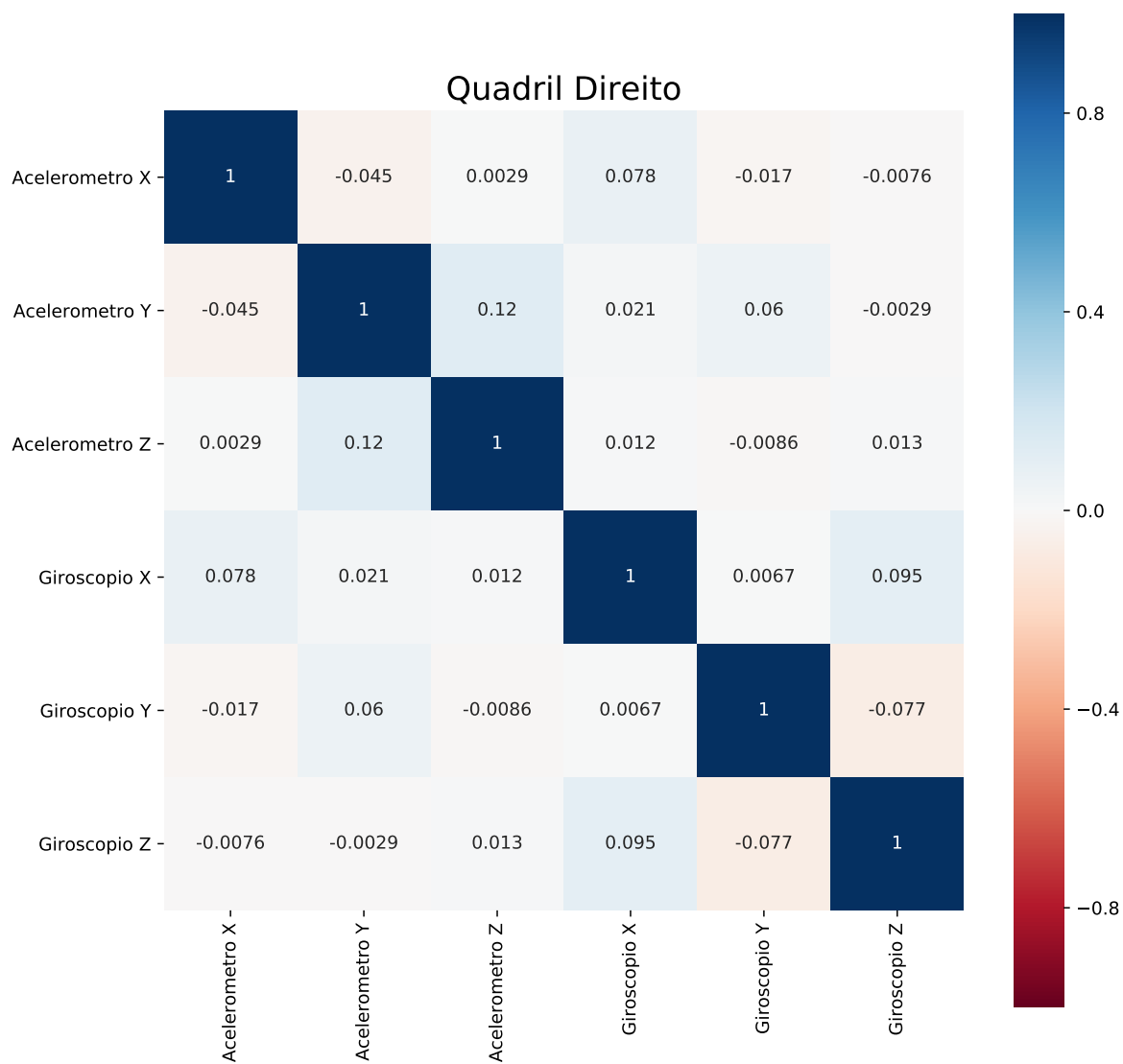
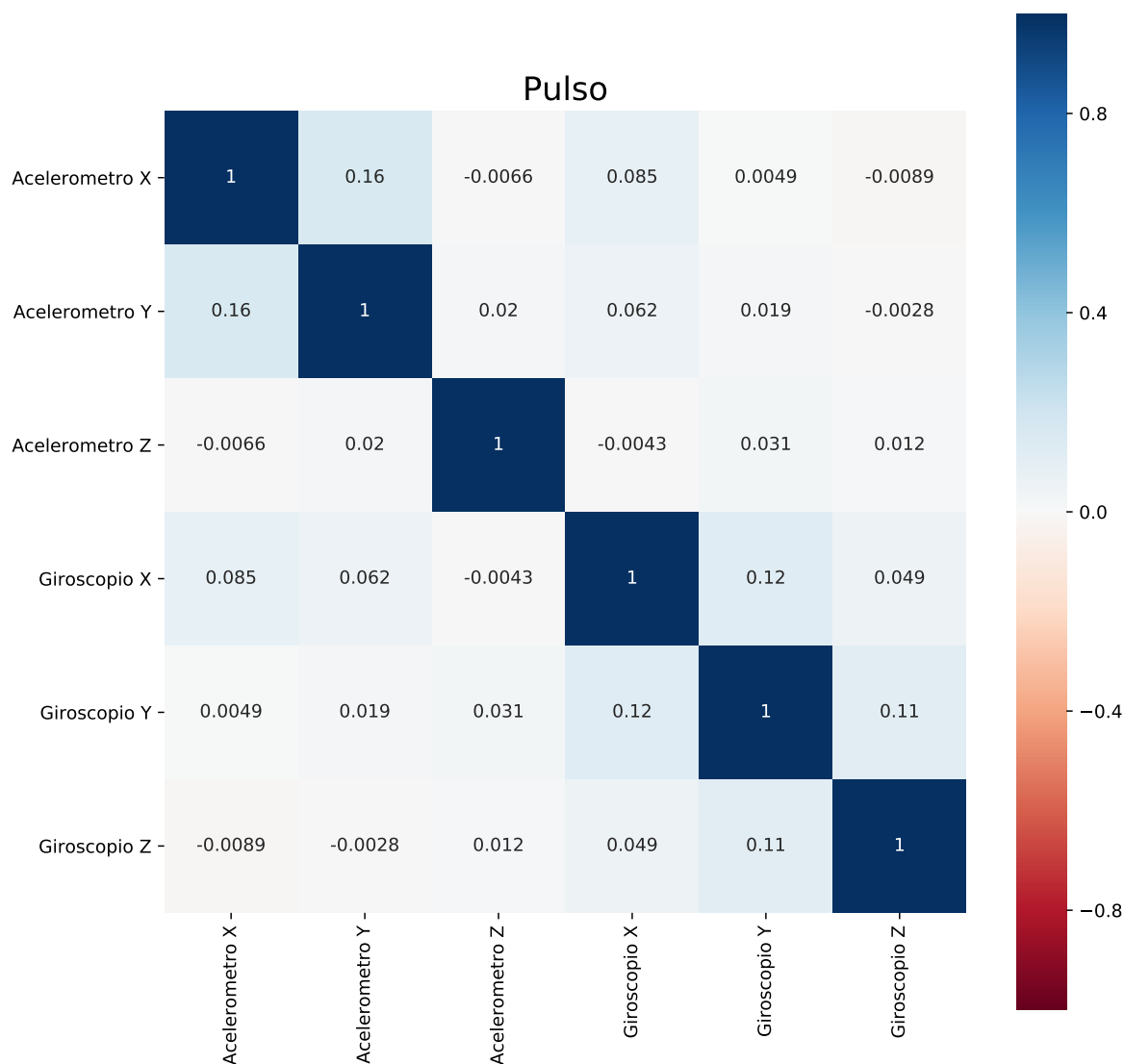


Figura 8 – Matriz de correlação dos sensores do pulso.



## 4.2 Resultados

Para a implementação dos classificadores foram utilizados os algoritmos *Random Forest* e *XGBoost*, onde 70% dos dados foram reservados para o treinamento e 30% para a validação.



### 4.2.1 Tornozelo

Para a região do tornozelo foram utilizados como entrada dos classificadores os dados dos três eixos do acelerômetro e giroscópio. Em seguida, com as previsões dos dois algoritmos utilizados, foi gerada uma matriz de confusão normalizada que nos mostra os valores em porcentagem. Na Figura 9 é apresentada a matriz de confusão para o *Random Forest* e na Figura 10 é apresentada a matriz de confusão para o *XGBoost*.

Figura 9 – Matriz de Confusão normalizada para os dados do tornozelo utilizando o *Random Forest*.

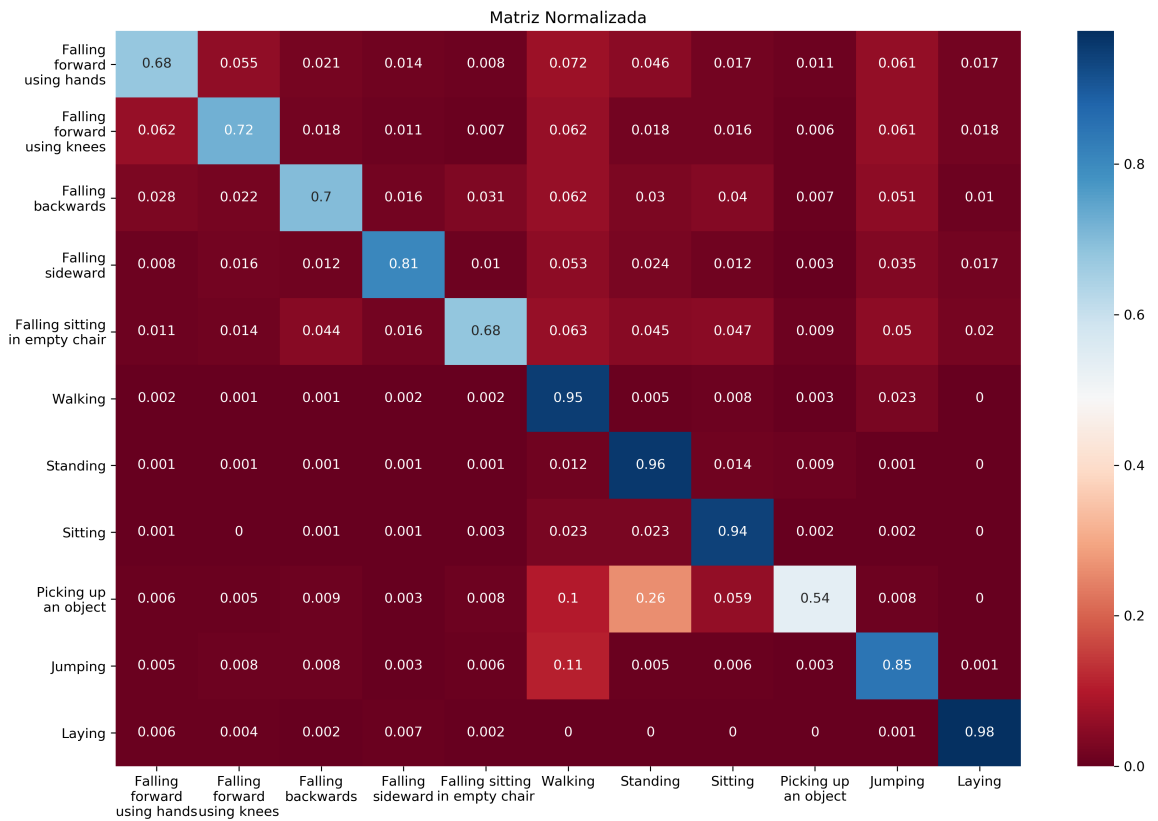
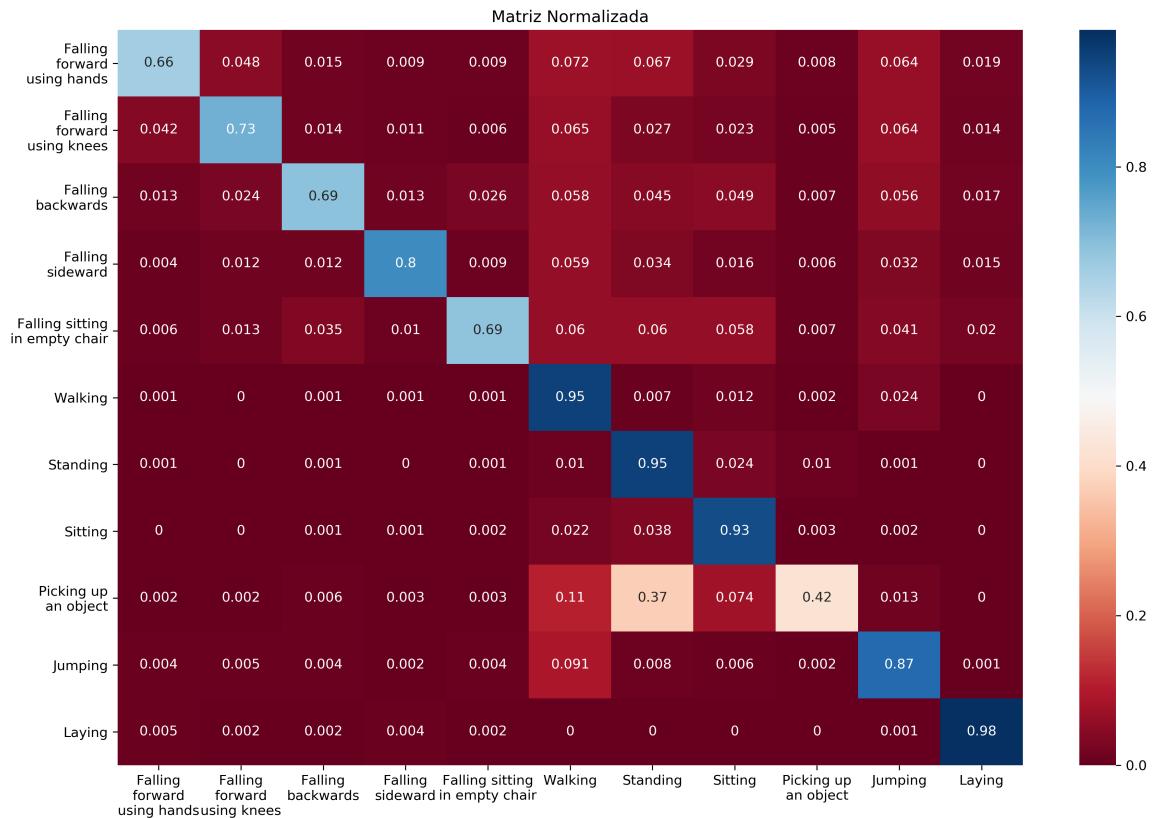


Figura 10 – Matriz de Confusão normalizada para os dados do tornozelo utilizando o *XGBoost*.



A Tabela 4 apresenta os dados de acurácia, precisão, sensibilidade e *F-1 score* dos dois algoritmos utilizados.

Tabela 4 – Desempenho geral dos algoritmos para a região do tornozelo.

Modelo	Acurácia (%)	Precisão (%)	Sensibilidade (%)	F-1 Score (%)
Random Forest	90.0	86.0	80.0	83.0
XGBoost	89.0	87.0	79.0	82.0

## 4.2.2 Cintura

Para a região da cintura foram utilizados como entrada dos classificadores os dados dos três eixos do acelerômetro e giroscópio. Em seguida, com as predições dos dois

algoritmos utilizados, foi gerada uma matriz de confusão normalizada que nos mostra os valores em porcentagem. Na Figura 11 é apresentada a matriz de confusão para o *Random Forest* e na Figura 12 é apresentada a matriz de confusão para o *XGBoost*.

Figura 11 – Matriz de Confusão normalizada para os dados da cintura utilizando o *Random Forest*.

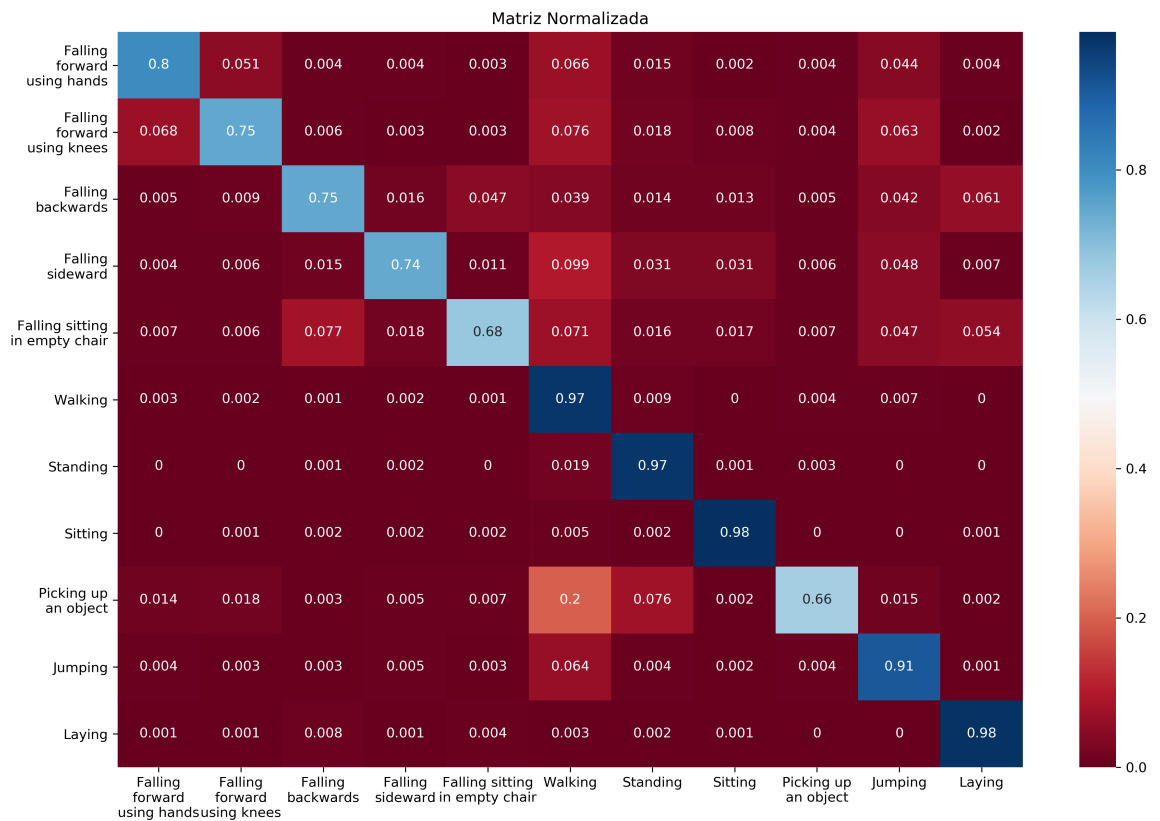
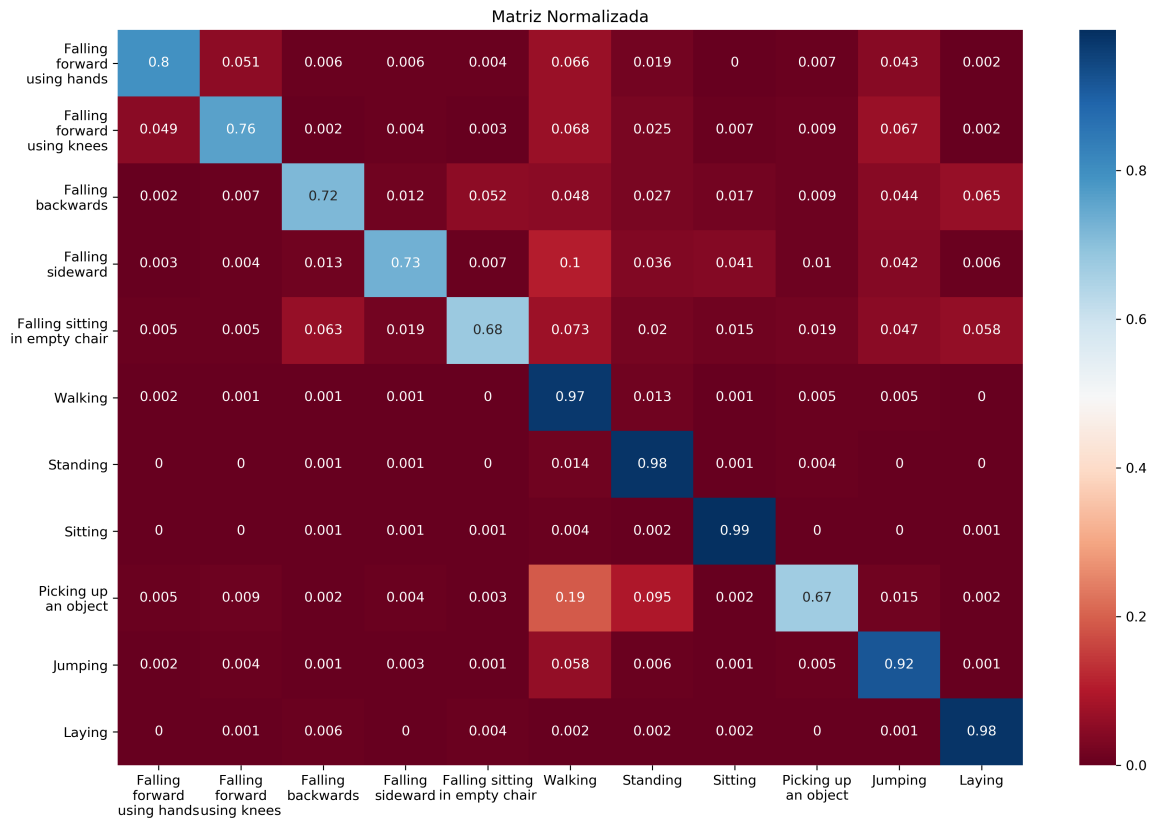


Figura 12 – Matriz de Confusão normalizada para os dados da cintura utilizando o *XG-Boost*.

A Tabela 5 apresenta os dados de acurácia, precisão, sensibilidade e *F-1 score* dos dois algoritmos utilizados.

Tabela 5 – Desempenho geral dos algoritmos para a região da cintura.

Modelo	Acurácia (%)	Precisão (%)	Sensibilidade (%)	F-1 Score (%)
Random Forest	92.0	89.0	84.0	86.0
XGBoost	93.0	90.0	84.0	86.0

### 4.2.3 Pescoço

Para a região do pescoço foram utilizados como entrada dos classificadores os dados dos três eixos do acelerômetro e giroscópio. Em seguida, com as predições dos dois

algoritmos utilizados, foi gerada uma matriz de confusão normalizada que nos mostra os valores em porcentagem. Na Figura 13 é apresentada a matriz de confusão para o *Random Forest* e na Figura 14 é apresentada a matriz de confusão para o *XGBoost*.

Figura 13 – Matriz de Confusão normalizada para os dados do pescoço utilizando o *Random Forest*.

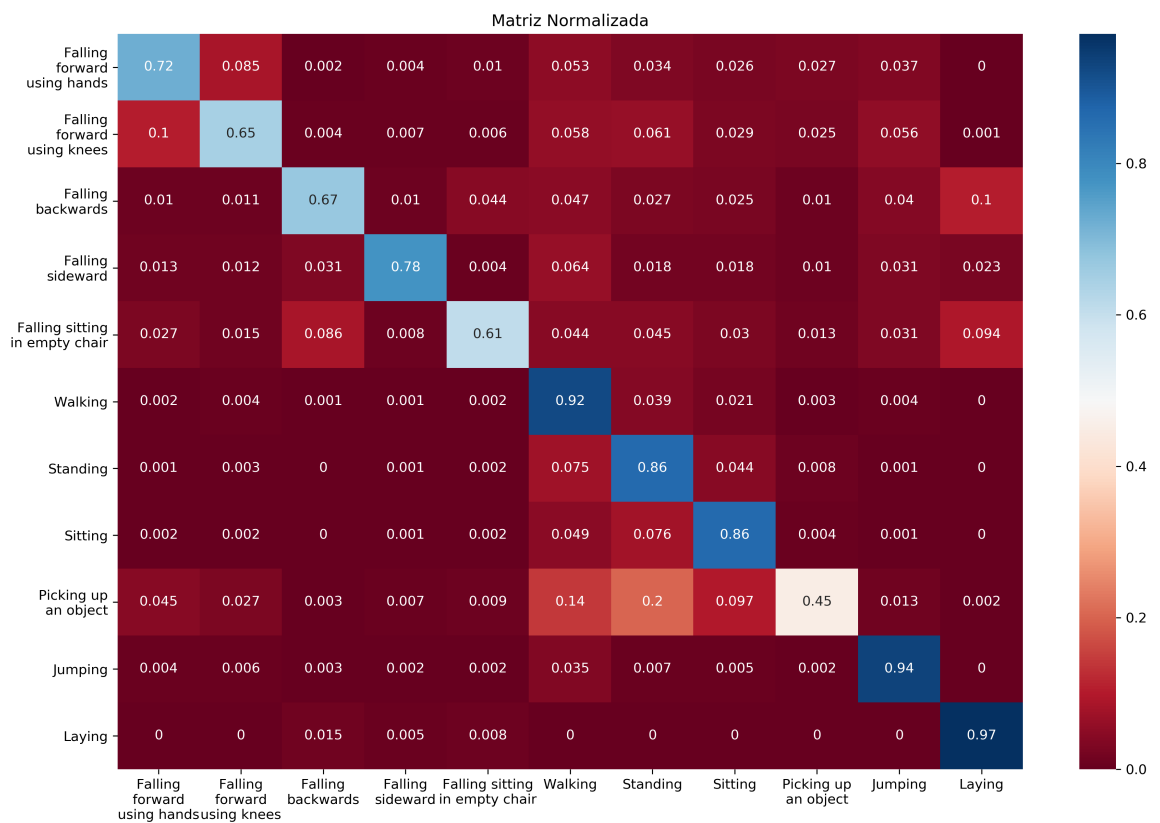
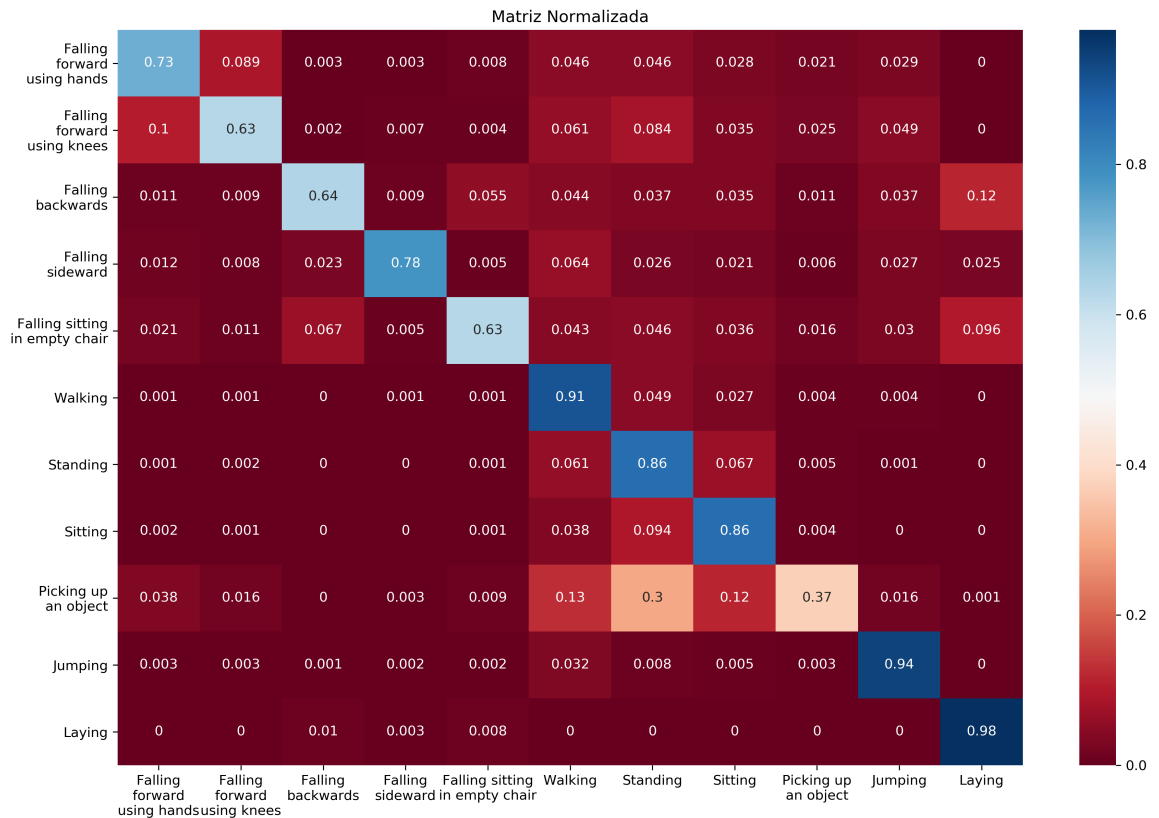


Figura 14 – Matriz de Confusão normalizada para os dados do pescoço utilizando o *XGBoost*.



A Tabela 6 apresenta os dados de acurácia, precisão, sensibilidade e *F-1 score* dos dois algoritmos utilizados.

Tabela 6 – Desempenho geral dos algoritmos para a região do pescoço.

Modelo	Acurácia (%)	Precisão (%)	Sensibilidade (%)	F-1 Score (%)
Random Forest	86.0	82.0	77.0	79.0
XGBoost	86.0	83.0	76.0	79.0

#### 4.2.4 Quadril Direito (Bolso)

Para a região do quadril direito foram utilizados como entrada dos classificadores os dados dos três eixos do acelerômetro e giroscópio. Em seguida, com as predições dos dois

algoritmos utilizados, foi gerada uma matriz de confusão normalizada que nos mostra os valores em porcentagem. Na Figura 15 é apresentada a matriz de confusão para o *Random Forest* e na Figura 16 é apresentada a matriz de confusão para o *XGBoost*.

Figura 15 – Matriz de Confusão normalizada para os dados do quadril direito utilizando o *Random Forest*.

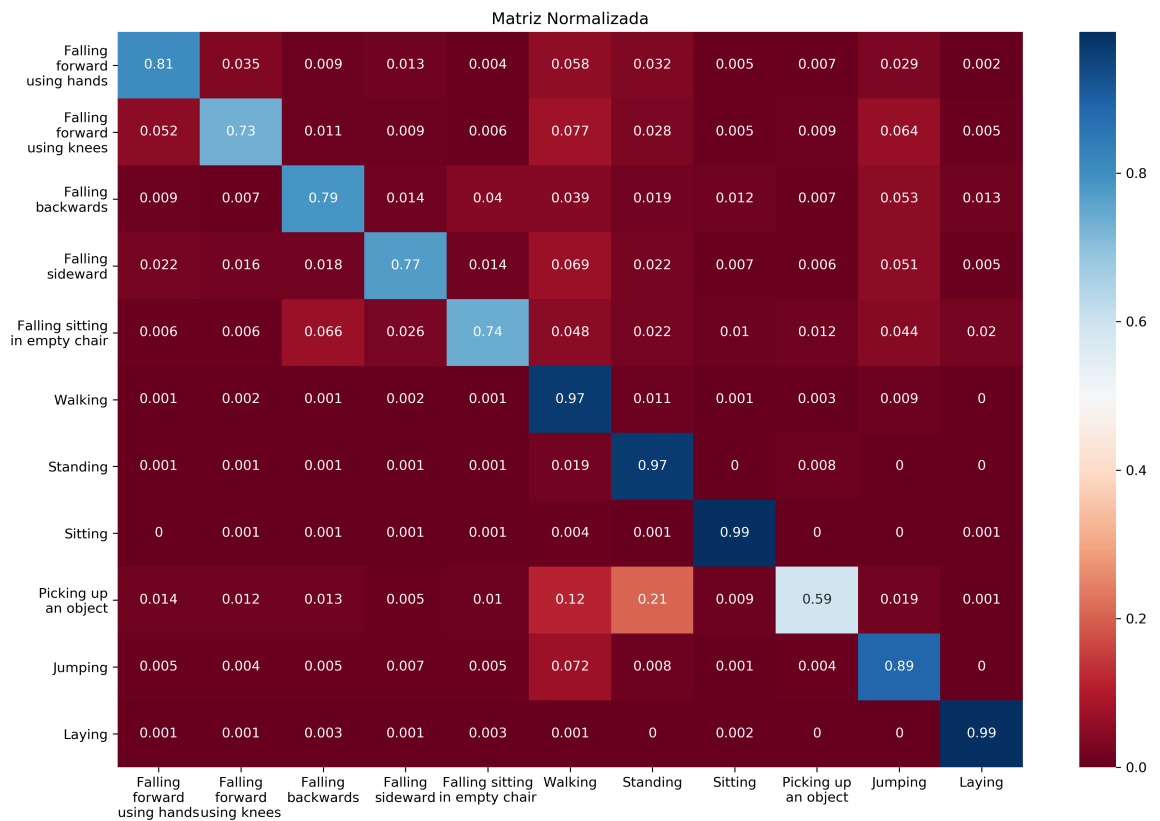
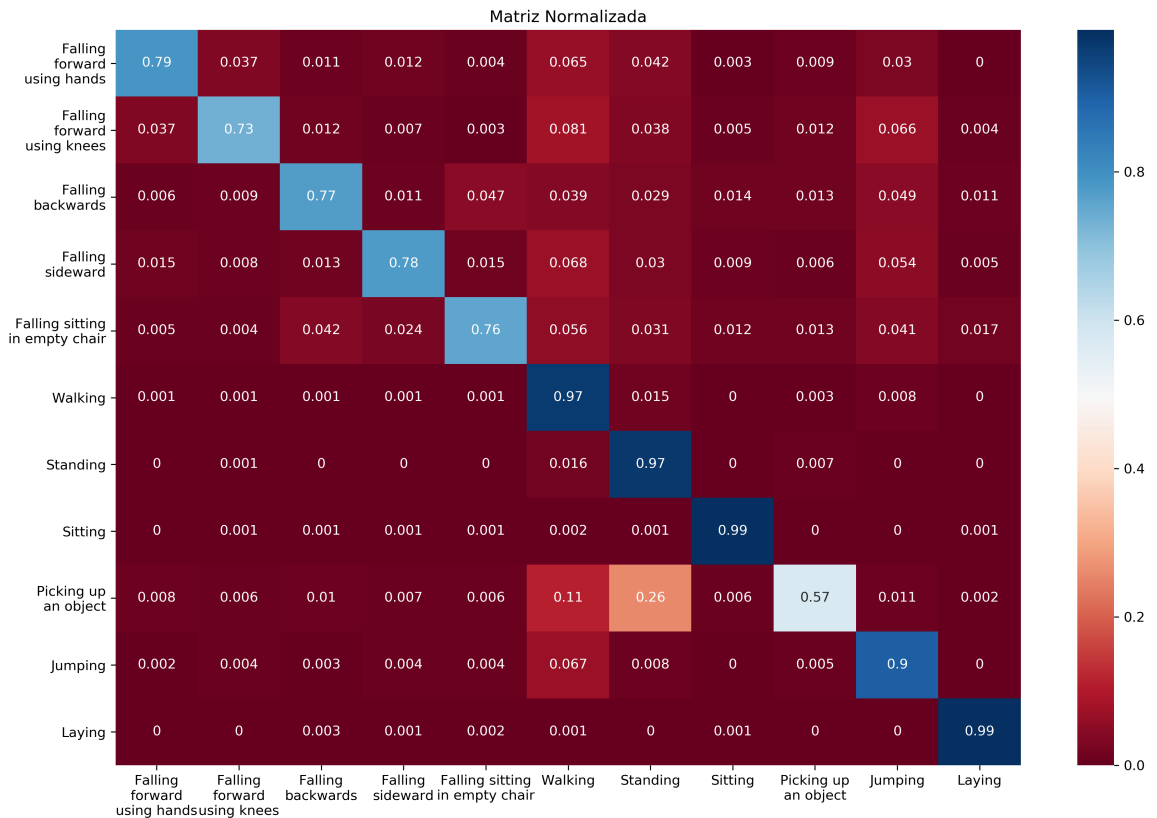


Figura 16 – Matriz de Confusão normalizada para os dados do quadril direito utilizando o *XGBoost*.



A Tabela 7 apresenta os dados de acurácia, precisão, sensibilidade e *F-1 score* dos dois algoritmos utilizados.

Tabela 7 – Desempenho geral dos algoritmos para a região do quadril direito.

Modelo	Acurácia (%)	Precisão (%)	Sensibilidade (%)	F-1 Score (%)
Random Forest	93.0	89.0	84.0	86.0
XGBoost	93.0	90.0	84.0	87.0

#### 4.2.5 Pulso

Para a região do pulso foram utilizados como entrada dos classificadores os dados dos três eixos do acelerômetro e giroscópio. Em seguida, com as predições dos dois algoritmos



utilizados, foi gerada uma matriz de confusão normalizada que nos mostra os valores em porcentagem. Na Figura 17 é apresentada a matriz de confusão para o *Random Forest* e na Figura 18 é apresentada a matriz de confusão para o *XGBoost*.

Figura 17 – Matriz de Confusão normalizada para os dados do pulso utilizando o *Random Forest*.

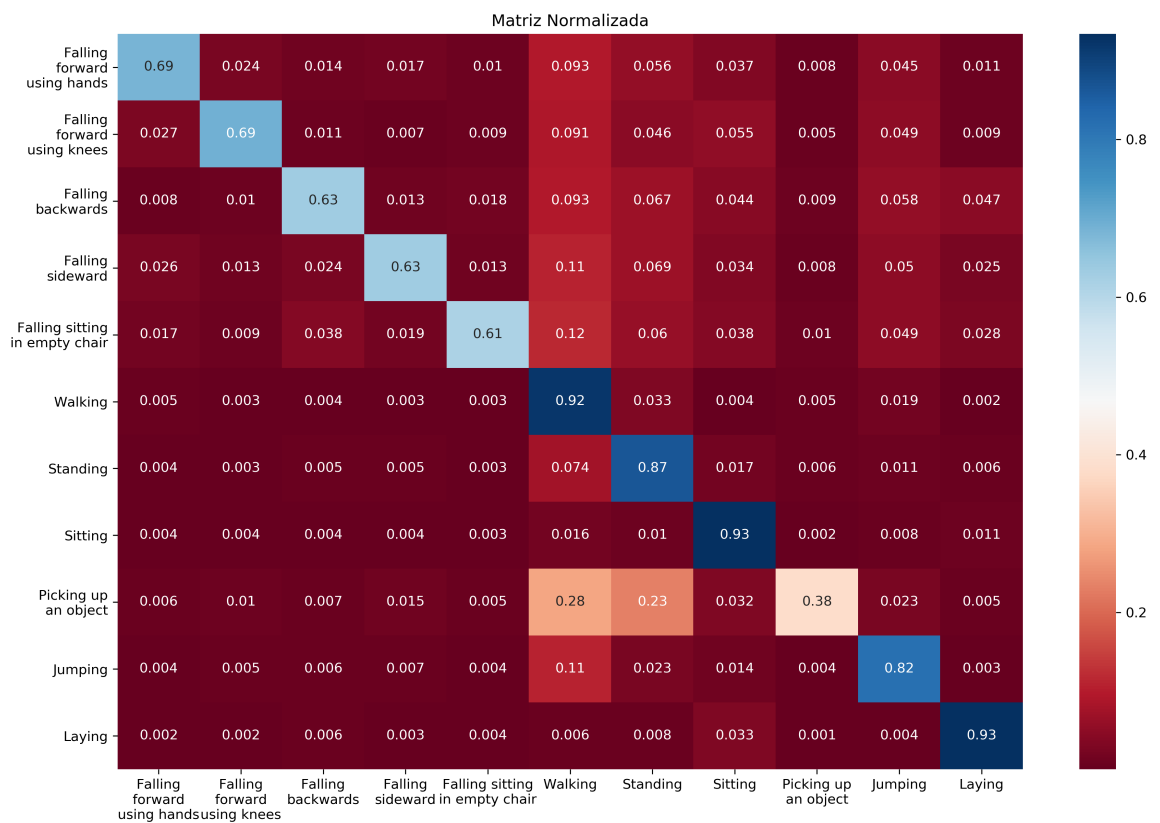
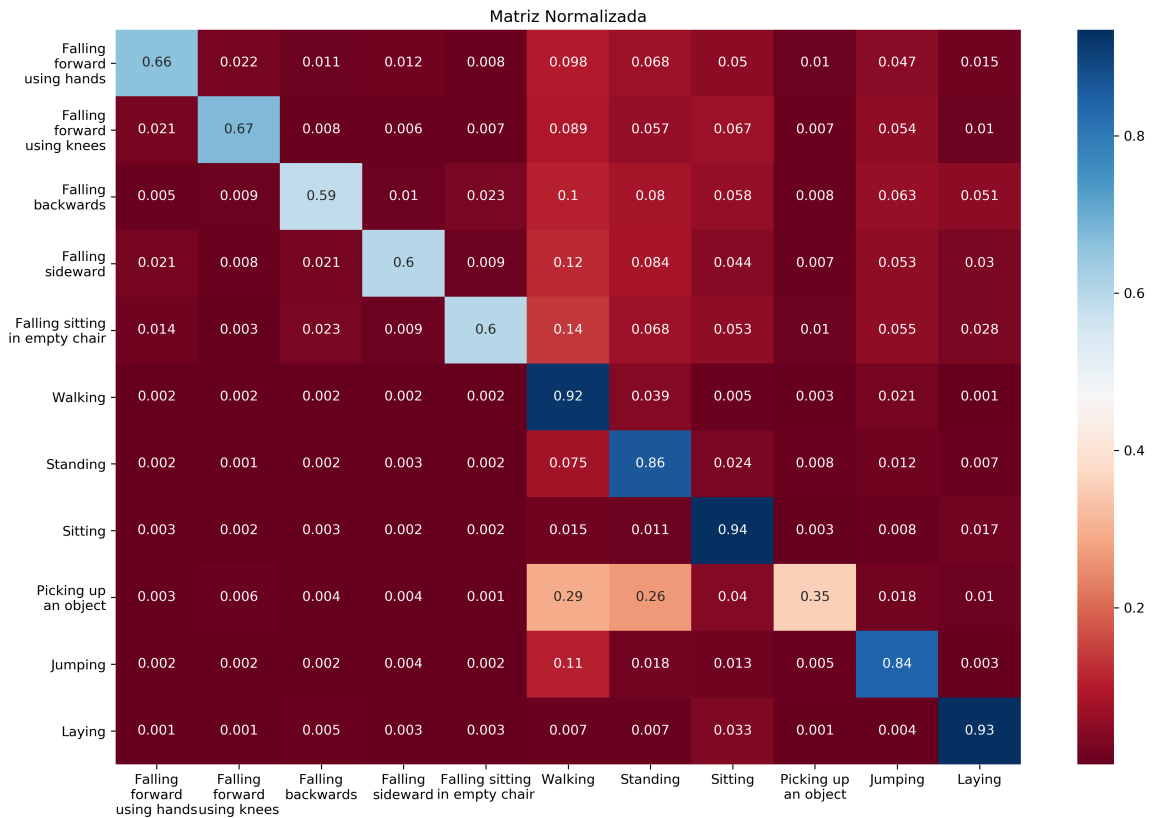


Figura 18 – Matriz de Confusão normalizada para os dados do pulso utilizando o *XGBoost*.



A Tabela 8 apresenta os dados de acurácia, precisão, sensibilidade e *F-1 score* dos dois algoritmos utilizados.

Tabela 8 – Desempenho geral dos algoritmos para a região do pulso.

Modelo	Acurácia (%)	Precisão (%)	Sensibilidade (%)	F-1 Score (%)
Random Forest	85.0	82.0	74.0	77.0
XGBoost	84.0	83.0	73.0	77.0

### 4.3 Análise

A Tabela 9 apresenta a melhor acurácia para a identificação de cada atividade e sua respectiva localização referente a cada algoritmo.

Tabela 9 – Localização com melhor acurácia para cada atividade.

Atividade	Localização		Acurácia (%)	
	<i>Random Forest</i>	<i>XGBoost</i>	<i>Random Forest</i>	<i>XGBoost</i>
Cair para frente usando as mãos	Quadril Direito	Cinto	81.0	80.0
Cair para frente usando os joelhos	Cinto	Cinto	75.0	76.0
Cair para trás	Quadril Direito	Quadril Direito	79.0	77.0
Cair de Lado	Tornozelo	Tornozelo	81.0	80.0
Cair sentando em uma cadeira vazia	Quadril Direito	Quadril Direito	74.0	76.0
Andar	Cinto/Quadril Direito	Cinto/Quadril Direito	97.0	97.0
Levantar	Cinto/Quadril Direito	Cinto	97.0	98.0
Sentar	Quadril Direito	Cinto/Quadril Direito	99.0	99.0
Pegar um objeto	Cinto	Cinto	66.0	67.0
Pular	Abaixo do Pescoço	Abaixo do Pescoço	94.0	94.0
Deitar	Quadril Direito	Quadril Direito	99.0	99.0

Ao analisar-se a Tabela 10 observa-se que a região do quadril direito apresentou a melhor acurácia para 6 das 11 atividades a serem identificadas, enquanto a região do cinto apresentou a melhor acurácia para 4 das 11 atividades. Ao observar o desempenho geral dos algoritmos (Tabelas 5 e 7) para essas duas regiões, percebe-se que ambas apresentaram uma acurácia acima de 92% e precisão acima de 89%.

Tabela 10 – Melhor localização e acurácia para cada atividade.

Atividade	Localização	Modelo	Acurácia (%)
Cair para frente usando as mãos	Quadril Direito	<i>Random Forest</i>	81.0
Cair para frente usando os joelhos	Cinto	<i>XGBoost</i>	76.0
Cair para trás	Quadril Direito	<i>Random Forest</i>	79.0
Cair de Lado	Tornozelo	<i>Random Forest</i>	81.0
Cair sentando em uma cadeira vazia	Quadril Direito	<i>XGBoost</i>	76.0
Andar	Cinto/Quadril Direito	<i>Random Forest/XGBoost</i>	97.0
Levantar	Cinto	<i>XGBoost</i>	98.0
Sentar	Quadril Direito	<i>Random Forest/XGBoost</i>	99.0
Pegar um objeto	Cinto	<i>XGBoost</i>	67.0
Pular	Abaixo do Pescoço	<i>Random Forest/XGBoost</i>	94.0
Deitar	Quadril Direito	<i>Random Forest/XGBoost</i>	99.0

A região do Tornozelo apresentou acurácia e precisão de 90% e 86% para o *Random Forest* e de 89% e 87% para o *XGBoost* (Tabela 4), entretanto, destacou-se como a melhor para identificar apenas uma atividade, “Cair de lado”, com 81% de acurácia. De forma similar, a região do Pescoço apresentou-se como a melhor para identificar apenas a atividade “Pular” com 94% de acurácia.

Entretanto, apesar da alta taxa de acerto para a atividade citada, essa região obteve um dos piores desempenhos gerais com 86% de acurácia para ambos os algoritmos e precisão de 82% para o *Random Forest* e 83% para o *XGBoost* (Tabela 6).

Já a região do Pulso foi a única que não se apresentou como destaque para a identificação das atividades. Ao analisar-se as matrizes de confusão (Figuras 17 e 18) percebe-se que essa região apresentou as mais baixas taxas de acerto para as quedas, tanto para o *Random Forest* quanto para o *XGBoost*. Ao observar o desempenho geral na Tabela 8 constata-se que a região do Pulso, simultaneamente com a região do Pescoço apresentam os piores desempenhos gerais.

Ao realizar o teste de Wilcoxon (WILCOXON, 1992) comparando o valores de acurácia apresentados na Tabela 9 dos dois classificadores obteve-se um  $p - \text{valor} = 0.79$ . Dessa forma utilizando significância  $\alpha = 5\%$  constata-se que o desempenho dos classificadores não apresentam diferenças estatisticamente significantes.

Nota-se também que a atividade de “Pegar um objeto” apresentou a região do cinto como melhor local para sua identificação, entretanto a acurácia foi de apenas 67%. Além disso, ao analisar as matrizes de confusão vemos que para as atividades andar, levantar, sentar, pular e deitar todas as localizações apresentaram uma acurácia acima de 85% independente do modelo utilizado. Percebe-se também que a região do Quadril Direito apresenta os melhores valores de acurácia para detectar as quedas de forma geral.

## 5 Conclusão

Este trabalho apresentou um estudo para identificar, a partir de cinco regiões anatômicas comumente utilizadas com tecnologias vestíveis, qual delas seria a melhor para um detector de queda em tempo real. As regiões anatômicas investigadas foram: o pulso, pescoço, quadril direito, no meio da cintura (cinto) e no tornozelo. Os sensores considerados no estudo, que podem ser facilmente incorporados em dispositivos vestíveis, foram o acelerômetro e o giroscópio.

Estudos anteriores relataram dificuldades para que os classificadores distinguíssem a atividade de “Pegar um objeto” de outras atividades como “Quedas”, etc. A melhor configuração de algoritmo de aprendizado de máquina e região anatômica para diferenciar o movimento de pegar um objeto de queda foi *XGBoost*-Cintura, com acurácia de 67%. Evidenciando assim, a dificuldade que os classificadores apresentam para diferenciar essa atividade em particular das demais.

Baseado nas métricas de avaliação apresentadas, percebe-se que no geral a região mais indicada para diferenciar tanto as quedas quanto as atividades diárias foi o quadril direito. Essa região apresentou-se como a melhor para identificar 6 das 11 atividades e o algoritmo obteve uma acurácia de 93% e precisão de 89%. Sendo assim, a confirmada como a melhor região para distinguir atividades diárias e quedas. Observa-se também que a região da cintura apresentou resultados satisfatórios, sendo esta a segunda melhor localização para a distinção das atividades analisadas.

Quando uma queda acontece, ocorre também o deslocamento do centro de massa do corpo humano. Dessa forma, acredita-se que os bons resultados apresentados tanto pela região do quadril direito quanto pela região da cintura, se devem ao fato de que ambas as localizações estão próximas ao centro de massa do corpo humano, fazendo com que os sensores não captem pequenas variações das extremidades do corpo as quais não estariam relacionadas ao deslocamento da massa corporal por completo.

Surpreendentemente a região Pescoço foi a melhor para identificar a atividade “Pular” apresentando uma acurácia de 94% e a região do Pulso não obteve destaque entre as demais na identificação das atividades sendo umas das localizações com pior desempenho geral, Tabela 8, juntamente com o Pescoço. Entretanto, as taxas de acerto da região do Pulso para as atividades andar, levantar, sentar, pular e deitar são altas (entre 82% e 93%) evidenciando assim o motivo de *smart watches* identificarem tão bem essas atividades e outras semelhantes, como corrida, pular corda e etc.

## Referências

- BREIMAN, L. 1 random forests–random features. Citeseer, 1999. Citado na página 21.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: ACM. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. [S.l.], 2016. p. 785–794. Citado na página 20.
- DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research*, v. 7, n. Jan, p. 1–30, 2006. Citado na página 23.
- FRIEDMAN, J. H. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, JSTOR, p. 1189–1232, 2001. Citado na página 20.
- HARRINGTON, P. *Machine learning in action*. [S.l.]: Manning Publications Co., 2012. Citado na página 20.
- IBGE, I. Sinopse do censo demográfico 2010. *Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística*, p. 261, 2011. Citado na página 13.
- IGUAL, R.; MEDRANO, C.; PLAZA, I. Challenges, issues and trends in fall detection systems. *Biomedical engineering online*, Springer, v. 12, n. 1, p. 66, 2013. Citado 2 vezes nas páginas 13 e 14.
- LORENA, A. C.; CARVALHO, A. C. de. Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 43–67, 2007. Citado na página 19.
- MARTÍNEZ-VILLASEÑOR, L. et al. Up-fall detection dataset: a multimodal approach. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 19, n. 9, p. 1988, 2019. Disponível em: <<http://sites.google.com/up.edu.mx/har-up/>>. Citado 3 vezes nas páginas 16, 17 e 18.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. *Sistemas inteligentes-Fundamentos e aplicações*, v. 1, n. 1, p. 32, 2003. Citado 2 vezes nas páginas 19 e 20.
- MORAIS, H. C. C. et al. Identificação do diagnóstico de enfermagem “risco de quedas em idosos com acidente vascular cerebral”. *Revista Gaúcha de Enfermagem*, v. 33, n. 2, p. 117–124, 2012. Citado na página 14.
- MUBASHIR, M.; SHAO, L.; SEED, L. A survey on fall detection: Principles and approaches. *Neurocomputing*, Elsevier, v. 100, p. 144–152, 2013. Citado na página 14.
- OLIVEIRA, H. et al. Perfil epidemiológico de idosos frequentadores de grupos de convivência no município de iguatu, ceará. *Revista Brasileira de Geriatria e Gerontologia*, Universidade do Estado do Rio de Janeiro, v. 14, n. 1, p. 123–133, 2011. Citado na página 13.
- OMS. A contribution of the world health organization to the second united nations world assembly on aging. *Madrid, Spain, April*, 2002. Citado na página 13.

OMS. *Envelhecimento ativo: uma política de saúde*. [S.l.]: Organização Pan-Americana da Saúde, 2005. Citado 2 vezes nas páginas 13 e 14.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011. Citado na página 16.

PEREIRA, S. et al. Sociedade brasileira de geriatria e gerontologia. *Projeto Diretrizes. Quedas em idosos*. São Paulo: Associação Médica Brasileira, Conselho Federal de Medicina, 2001. Citado na página 13.

WILCOXON, F. Individual comparisons by ranking methods. In: *Breakthroughs in statistics*. [S.l.]: Springer, 1992. p. 196–202. Citado 2 vezes nas páginas 22 e 42.