



# Desenvolvimento de um sistema embarcado para classificação de movimentos do usuário utilizando técnicas de aprendizado de máquina

Aluno: Paulo César Vasco  
Orientador: Prof. Dr. Rafael Ferrari

## 1 Resumo

Este projeto consistiu no desenvolvimento de um sistema embarcado utilizando a placa Arduino e o uso de sensores de movimento, para mensurar e capturar variações de aceleração e velocidade angular que posteriormente foram janeladas e algumas características pertinentes foram extraídas para o treinamento de modelos matemáticos lineares e não lineares, que se mostraram capazes de classificar diferentes tipos de movimentos que uma pessoa realiza no dia a dia com sucesso.

**Palavras-Chave:** Sistema Embarcado. Classificação de Movimento. Machine Learning.

## 2 Introdução

Com o envelhecimento da população, tecnologias para melhoria da qualidade de vida de idosos vem se tornando uma importante área multidisciplinar de pesquisa. Um problema importante na velhice são os traumas e lesões provocados por quedas, que podem levar a sérias complicações do estado de saúde do idoso. Nesse sentido, o monitoramento e a detecção automática de quedas pode ser uma ferramenta interessante para acelerar o socorro e reduzir os danos e possíveis sequelas ao acidentado.

Existem várias abordagens para o monitoramento de quedas que se distinguem, por exemplo, quanto à natureza da informação utilizada na detecção. Destacam-se três tipos principais: detecção baseada, em monitoramento de impacto, em processamento de imagens e em informações inerciais. A primeira consiste da instalação de sensores de impacto no cômodo em que a pessoa a ser monitorada se encontra e as quedas são detectadas a partir da intensidade dos sinais provenientes dos sensores porém apresentando como desvantagem a grande ocorrência de falsos positivos que podem ser causados por diversos fatores, como, por exemplo, acionamento dos sensores por quedas de objetos, atividade de animais no recinto, etc [3]. A técnica baseada em processamento de imagens faz uso de câmeras instaladas no ambiente para monitorar as atividades do indivíduo. Com base na orientação do corpo obtida a partir das imagens, o sistema detecta a ocorrência ou não de uma queda [1]. A grande desvantagem desse método é a limitação espacial já que o monitoramento se restringe ao alcance das câmeras [1]. A terceira abordagem se baseia em informações inerciais obtidas a partir de acelerômetros e giroscópios presos ao corpo do indivíduo monitorado [4][3][1], apresentando a vantagem da portabilidade do sistema que permite o monitoramento do indivíduo independentemente do local em que se encontra.

Diante desses aspectos, o projeto foi desenvolvido focando nessa abordagem. Foram estudados meios de se utilizar informações inerciais (aceleração e velocidade angular) registradas a partir de sensores fixados ao corpo para detectar quedas. Foi desenvolvido utilizado um aparato de aquisição onde os sinais registrados serão processados utilizando técnicas de classificação e detecção baseadas em aprendizado de máquina.

### 3 Desenvolvimento e Resultados

Para captura de dados foi projetado um sistema embarcado baseado na placa de desenvolvimento Arduíno integrada com um sensor capaz de medir variações de aceleração e velocidade angular do sistema em que o sensor se encontra fixado. Os dados foram capturados e armazenados em cartão de memória para posteriores análises, foram analisados as seguintes ações básicas do dia a dia: **andar, subir e descer escadas, deitar e levantar, sentar e levantar, pular.**

Como o único movimento de interesse é do corpo em relação a atividade que está sendo exercida no momento, o tronco foi escolhido como um ponto de fixação, por se tratar de uma parte mais estável do corpo, além disso tomou-se o cuidado de sempre colocar o sensor na mesma posição para que o sistema referencial não fosse alterado, uma vez que o sensor considera seu próprio corpo como referencia dos eixos. Obtendo-se assim três medidas de variação de aceleração e três da variação da velocidade angular, correspondendo aos eixos  $x$ ,  $y$ , e  $z$ .

Considerando que o movimento humano encontra-se na ordem dos milissegundos, foi setado uma frequência de amostragem,  $F_s$  de 500 Hz. Os dados então foram janelados a cada 100 amostras, resultando em passos de 200ms, além de haver uma sobreposição de 50 amostras entre as janelas para garantir redundância nos padrões estudados. Essas janelas foram utilizadas para extração de algumas características estatísticas e harmônicas durante cada passo, seguindo [4]:

$$média_i(s) = \mu_i = \frac{1}{N} \sum_{n=1}^N s_n \quad (1)$$

$$variância_i(s) = \sigma^2 = \frac{1}{N} \sum_{n=1}^N (s_n - \mu)^2 \quad (2)$$

$$assimetria_i(s) = \frac{1}{N\sigma^3} \sum_{n=1}^N (s_n - \mu)^3 \quad (3)$$

$$curtose_i(s) = \frac{1}{N\sigma^4} \sum_{n=1}^N (s_n - \mu)^4 \quad (4)$$

$$RF_i(s) = \frac{\sum_{n=1}^3 \|s_n e^{-\frac{j2\pi qn}{N}}\|}{\sum_{n=1}^{N-1} \|s_n e^{-\frac{j2\pi qn}{N}}\|} \quad q = 1, \dots, N-1 \quad (5)$$

Sendo  $s$  correspondente a aceleração ou a velocidade angular da  $n$ -ésima amostra, os valores calculados dessas características para cada uma das janelas corresponde a entrada para cada um dos classificadores testados.

Na classificação dos dados foram implementados classificadores lineares e não lineares no modo supervisionado, com os dados recebendo uma representação numérica, sendo 1 para o movimento a ser classificado e -1 para o demais.

Nos classificadores lineares, os quais são definidos pelo uso de funções discriminantes lineares, com a seguinte característica:

$$\begin{aligned} g(\mathbf{x}(i)) &= \mathbf{w}^T \mathbf{x}(i) + w_0 \\ &= \mathbf{W}^T \Phi(\mathbf{x}(i)) \end{aligned} \quad (6)$$

em que  $\mathbf{W} = [w_0 \ \mathbf{w}^T]^T$  é o vetor de parâmetros do classificador linear e  $\Phi(\mathbf{x}(i)) = [1 \ \mathbf{x}^T(i)]^T$ . O vetor de parâmetros do classificador linear pode ser obtido a partir da solução de quadrados mínimos

$$\mathbf{W} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{Y} \quad (7)$$

em que  $\mathbf{Y} = [y(0) \ y(1) \ \dots \ y(N-1)]^T$  é o vetor de rótulos dos padrões de entrada e

$$\Phi = \begin{bmatrix} \Phi(\mathbf{x}(0))^T \\ \vdots \\ \Phi(\mathbf{x}(N-1))^T \end{bmatrix}$$

é a matriz de padrões de entrada.

Para abordagem de classificadores não lineares, foi usado redes neurais artificiais (RNA), que definem-se como modelos matemáticos inspirados no cérebro, fazendo-se uso de unidades simples de processamento denominados neurônios. O neurônio artificial é a menor unidade que compõe uma rede neural e, quando conectado a outros neurônios, forma sinapses capazes de armazenar a informação que foi-lhes apresentada. Assim, a influência de uma sinapse é matematicamente representada pelo peso sináptico associado a esta. A saída de um neurônio artificial pode ser modelada pela equação:

$$h_k(i) = \varphi(\mathbf{w}_k^T \mathbf{x}(i) + b_k) [2] \quad (8)$$

Em que  $\mathbf{x}(i) = [x_1(i) \ x_2(i) \ \dots \ x_M(i)]^T$  é o vetor de entradas,  $\mathbf{w}_k = [w_{k,1} \ w_{k,2} \ \dots \ w_{k,M}]^T$ , é o vetor de pesos sinápticos,  $b_k$  é o *bias* (um termo constante de entrada que não depende dos pesos sinápticos, usado para ajuste da rede) e  $\varphi(\cdot)$  é a função de ativação, que neste trabalho será a tangente hiperbólica. Podendo afirmar que a saída do neurônio é, portanto, uma função não linear das entradas, que são combinadas linearmente através do vetor de pesos sinápticos[2].

Os neurônios são agrupados em camadas e a forma como estão conectados define a arquitetura da rede, neste trabalho foi utilizada apenas a chamada rede *feedforward*, que têm por característica propagar o sinal de entrada em um único sentido, da entrada para a saída. Outro exemplo de arquitetura é a recorrente, na qual laços de realimentação transmitem as saídas para camadas neurais anteriores.

Foi escolhida uma técnica de aprendizado particular nomeada de Extreme Learning Machine (ELM), a qual consiste do uso de uma RNA, com uma única camada intermediária na qual seus pesos de entrada são escolhidos de forma aleatória e o problema assume uma característica linear para obtenção dos parâmetros, os vetores de pesos sinápticos da camada de entrada podem ser agrupados em uma matriz  $\mathbf{W}_{in}$ , em que cada linha corresponde ao vetor de pesos associado a um dos neurônios da camada. Desse modo, a  $i$ -ésima linha equivale ao vetor de pesos do  $i$ -ésimo neurônio:

$$\mathbf{W}_{in} = \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_N^T \end{bmatrix} = \begin{bmatrix} w_{10} & w_{11} & \dots & w_{1M} \\ w_{20} & w_{21} & \dots & w_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N0} & w_{N1} & \dots & w_{NM} \end{bmatrix}$$

Em que o primeiro termo de cada linha da matriz  $\mathbf{W}_{in}$  corresponde ao bias do neurônio. Nesse trabalho, cada elemento dessa matriz foi obtido a partir do resultado de uma realização de uma variável aleatória com densidade de probabilidade uniforme no intervalo  $[-1,1]$ .

As saídas dos neurônios da camada intermediária podem ser organizadas em um vetor em um vetor  $\mathbf{h}(i)$  na forma

$$\mathbf{h}(i) = \begin{bmatrix} \mathbf{h}_1(i) \\ \mathbf{h}_2(i) \\ \vdots \\ \mathbf{h}_N(i) \end{bmatrix} = \begin{bmatrix} \varphi(\mathbf{w}_1^T \mathbf{x}(i)) \\ \varphi(\mathbf{w}_2^T \mathbf{x}(i)) \\ \vdots \\ \varphi(\mathbf{w}_N^T \mathbf{x}(i)) \end{bmatrix} = \varphi(\mathbf{W}_{in}^T \mathbf{x}(i)).$$

A saída da rede produzida pelo padrão de entrada  $\mathbf{x}(i)$  é obtida a partir da combinação linear das saídas dos neurônios, ou seja,

$$y(i) = \mathbf{w}_{out}^T \mathbf{h}(i) \quad (9)$$

em que  $\mathbf{w}_{out}$  é o vetor dos pesos sinápticos da camada de saída.

Por conta de não ocorrer o ajuste dos pesos sinápticos da camada intermediária, o único parâmetro a ser adaptado na ELM é o vetor  $\mathbf{w}_{out}$ , tornando o treinamento da ELM equivalente a um problema de regressão linear. O valor ótimo desse vetor pode ser obtido a partir da solução de quadrados mínimos de maneira análoga à mostrada na equação 7, também foi avaliado o desempenho da ELM variando-se o número de neurônios na camada intermediária.

As entradas  $\Phi$  e  $\mathbf{x}$ , dos classificadores, correspondem a concatenação das características extraídas de cada janela de amostras, esse conjunto de entrada é dividido em dois subconjuntos, treinamento e validação, o primeiro consiste nos dados utilizados para obtenção dos parâmetros dos classificadores e o outro na comparação se os rótulos calculados para esses dados equivalem aos rótulos extraídos, além disso para testar a capacidade de generalização dos classificadores, o processo de classificação foi rodado 200000, alternando sempre de forma aleatória o conjunto de validação dentro do conjunto total de dados, ao final foram extraídas a média e o desvio padrão dos resultados. A qualidade dos modelos foi mensurada pela sua taxa de acerto (acurácia), definida por:

$$Acurácia = \frac{TP + TN}{TP + TN + FP + FN}$$

Onde: TP = Verdadeiro Positivo, TN = Verdadeiro Negativo, FP = Falso Positivo, FN = Falso Negativo.

Nas tabelas abaixo estão exibidos os resultados obtidos para o modelo linear e para a ELM variando-se o número de neurônios na camada intermediária:

Tabela 1: Classificador Linear

Movimento	Acurácia	
	Média	Std
ANDANDO	82%	0,02%
DEITAR	90,2%	0,02%
ESCADA	78,6%	0,03%
SENTAR	91,6%	0,02%
PULAR	97,7%	0,01%

Tabela 2: ELM com 30 neurônios

Movimento	Acurácia	
	Média	Std
ANDANDO	68,7%	0,04%
DEITAR	83,1%	0,03%
ESCADA	77,1%	0,03%
SENTAR	88,5%	0,02%
PULAR	97,1%	0,01%

Tabela 3: ELM com 150 neurônios

Movimento	Acurácia	
	Média	Std
ANDANDO	83%	0,02%
DEITAR	91,1%	0,02%
ESCADA	81,3%	0,03%
SENTAR	91,8%	0,01%
PULAR	96,8%	0,01%

Tabela 4: ELM com 300 neurônios

Movimento	Acurácia	
	Média	Std
ANDANDO	83,1%	0,03%
DEITAR	92%	0,02%
ESCADA	82,5%	0,02%
SENTAR	93%	0,02%
PULAR	96,3%	0,02%

Observando os resultados obtidos é possível notar que ambos classificadores apresentaram uma alta taxa de acerto (acima de 75%), com o modelo linear apresentando dificuldades em distinguir movimentos com uma natureza parecida, como subir e descer escadas o qual é relativamente parecido com o ato de caminhar.

Para o modelo não linear nota-se que o número de neurônios na camada intermediária impacta diretamente na qualidade do classificador, com poucos neurônios o modelo apresentou um desempenho pior que o linear, porém aumentando-se essa quantidade apresentou uma taxa de acerto bem alta até mesmo para os movimentos mais difíceis de distinguir, a custo de um poder computacional maior, esse modelo se mostrou bem mais lento que o linear, outro ponto que foi possível notar é que um aumento no número de neurônios na camada intermediária não tem um impacto significativo no desempenho do modelo a partir de determinada quantidade (redes com 150 e 300 neurônios, praticamente obtiveram o mesmo desempenho).

## 4 Conclusão

A proposta inicial do projeto foi desenvolver um sistema capaz de detectar quando o indivíduo viesse a sofrer uma queda, através da análise dos dados coletados. Devido a problemas no sistema e à impossibilidade de repará-lo durante a pandemia, foi necessário alterar um pouco o foco do projeto, utilizando os dados que já haviam sido coletados. Foi possível projetar classificadores que identificaram com elevado grau de sucesso o tipo de atividade que estava sendo realizada no momento da coleta.

Os sistema mostra-se promissor, para ambos os classificadores, e podem ser ainda melhorados aumentando-se a base de dados com mais medições, a fim de obter-se mais informações sobre cada movimento. Além de aumentar o conjunto de treinamento, em trabalhos futuros, o cenário de queda deverá ser incluído após o reparo do sistema de aquisição e implementação de melhorias para aumentar sua resistência a impactos sofridos durante as medições, porém é possível esperar um elevado grau de sucesso com relação a classificação de quedas, pois o movimento **pular** estudado, ambos os classificadores mostraram uma alta taxa de acerto se aproximando dos 100%, mostrando o potencial dos modelos para conseguir distinguir a queda dos demais movimentos.

## Referências

- [1] Ling Shao Mubashir, Muhammad and Luke Seed. A survey on fall detection: Principles and approaches. *Neurocomputing 100*, (1), 2013.
- [2] Maiane Junqueira Teixeira Neto. Uso de redes neurais na desconvolução preditiva de traços sísmicos. *Exame de Qualificação*, (1), 2018.
- [3] et al Noury, Norbert. Fall detection-principles and methods. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (1), 2007.
- [4] Ahmet Özdemir and Billur Barshan. Detecting falls with wearable sensors using machine learning techniques. *Sensors 14.6*, (1), 2014.