

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
DEPARTAMENTO ACADÊMICO DE INFORMÁTICA**

**NATANAEL EVANGELISTA DE FREITAS**

**ESTIMATIVA DO ÍNDICE DE IRREGULARIDADE INTERNACIONAL  
A PARTIR DE SINAIS OBTIDOS POR SENSORES DE SMARTPHONES  
EMBARCADOS EM VEÍCULOS**

**TRABALHO DE CONCLUSÃO DE CURSO**

**PATO BRANCO**

**2021**

**NATANAEL EVANGELISTA DE FREITAS**

**ESTIMATIVA DO ÍNDICE DE IRREGULARIDADE  
INTERNACIONAL A PARTIR DE SINAIS OBTIDOS POR  
SENSORES DE SMARTPHONES EMBARCADOS EM VEÍCULOS**

**Estimate of international roughness index by signals from  
accelerometer of smartphones embedded in vehicles**

Trabalho de Conclusão de Curso apresentado(a) como requisito para obtenção do título(grau) de Bacharel em Engenharia de Computação, do Departamento Acadêmico de Informática, da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. Dalcimar Casanova  
Coorientador: Prof. Dr. Danilo Rinaldi Bisconsini

**PATO BRANCO**

**2021**



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es).

Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**NATANAEL EVANGELISTA DE FREITAS**

**ESTIMATIVA DO ÍNDICE DE IRREGULARIDADE INTERNACIONAL A PARTIR DE  
SINAIS OBTIDOS POR SENSORES DE SMARTPHONES EMBARCADOS EM  
VEÍCULOS**

Trabalho de Conclusão de Curso de Graduação  
apresentado como requisito para obtenção do título de Bacharel em Engenharia de Computação da Universidade Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 01 de Dezembro de 2021

---

Dalcimar Casanova  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Danilo Rinaldi Bisconsini  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Marco Antonio De Castro Barbosa  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Giovanni Bratti  
Doutorado  
Universidade Tecnológica Federal do Paraná

**PATO BRANCO  
2021**

*Trabalhe. Sirva. Seja forte e não encha o saco*  
(MARSILI, 2019).

## RESUMO

No Brasil, o transporte de cargas e passageiros ocorre majoritariamente por meio de rodovias. Para proporcionar uma qualidade de rolamento satisfatória é essencial que o país disponha de instrumentos de avaliação que permitam o monitoramento constante dessa infraestrutura. Nos últimos anos que *smartphones* podem ser utilizados como instrumentos de avaliação pois seus sensores captam os sinais de aceleração vertical que podem ser relacionados ao Índice de Irregularidade Internacional (IRI), parâmetro objetivo mais utilizado para a caracterização da qualidade de rolamento de pavimentos. Portanto, o objetivo deste trabalho foi desenvolver modelos inteligentes para relacionar o sinal de aceleração vertical com o índice de irregularidade internacional. Os modelos desenvolvidos são redes neurais artificiais (RNA) sendo elas: redes neurais convolucionais e redes neurais recorrentes. A análise dos resultados deste trabalho considera o resultado de uma regressão linear simples apresentado em trabalhos encontrados na literatura e é utilizado para comparação com o resultado obtido com o sistema desenvolvido neste trabalho, utilizando-se redes neurais. Ao comparar os resultados nota-se que as redes neurais apresentam uma melhora significativa na correlação entre índice de irregularidade internacional e aceleração vertical mostrando-se uma excelente alternativa para a solução deste problema com coeficiente de determinação acima de 0,95.

**Palavras-chave:** Índice de Irregularidade Internacional. Aprendizado de máquina. Redes Neurais Artificiais. Irregularidade Longitudinal. Rede Neural Recorrente.

## ABSTRACT

FREITAS, Natanael Evangelista. **Estimate of international roughness index by signals from accelerometer of smartphones embedded in vehicles**. 2021. 53 p. Bachelor Thesis (Bachelor's Degree in Computer Engineering) – Universidade Tecnológica Federal do Paraná. Pato Branco, 2021.

In Brazil, the transport of cargo and passengers takes place mainly through highways. In order to provide a satisfactory bearing quality, it is essential that the country has assessment instruments that allow constant monitoring of this infrastructure. In recent years, smartphones can be used as evaluation instruments because their sensors capture the vertical acceleration signals that can be related to the International Roughness Index (IRI), the most used objective parameter for the characterization of the bearing quality of floors. Therefore, the objective of this work was to develop intelligent models to relate the vertical acceleration signal to the International Roughness Index. The models developed are artificial neural networks (ANN) namely: convolutional neural networks and recurrent neural networks. The analysis of the results of this work considers the result of a simple linear regression presented in works found in the literature and is used for comparison with the result obtained with the system developed in this work of neural networks. When comparing the results, it is noted that the neural networks show a significant improvement in the correlation between the International Roughness Index and vertical acceleration, proving to be an excellent alternative for solving this problem. TNeural networks obtained a coefficient of determination above 0.95.

**Keywords:** International Roughness Pavement. Machine Learning. Artificial Neural Network. Longitudinal Pavement. Recurrent Neural Network.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Ilustração da irregularidade longitudinal . . . . .	12
Figura 2 – Representação do modelo de um quarto de carro. . . . .	16
Figura 3 – Resultado do polinômio interpolador . . . . .	19
Figura 4 – Modelo de regressão linear. . . . .	21
Figura 5 – (a) Para $SSE = 0$ onde não há variação; (b) Um valor pequeno de SSE; (c) Uma variação que não apresenta relacionamento linear e um valor de SSE elevado. . . . .	23
Figura 6 – (a) Soma do Quadrado dos Erros em relação a reta de regressão; (b) Soma total dos quadros em relação a reta horizontal na altura $\bar{y}$ . . . . .	23
Figura 7 – Modelo matemático simples de um neurônio. A saída é o resultado da função de ativação, $a_i$ é a saída da função de ativação da unidade $i$ e $w_{i,j}$ seu peso associado ao link da unidade $i$ para a unidade atual $j$ . . . . .	27
Figura 8 – Rede neural em camadas. . . . .	29
Figura 9 – Rede neural multicamadas. . . . .	29
Figura 10 – Rede convolucional para processamento de imagem. . . . .	31
Figura 11 – Representação das camadas, $x(t)$ a camada de entrada, $h(t)$ a camada oculta e $o(t)$ a camada de saída. . . . .	32
Figura 12 – Representação do neurônio <i>LSTM</i> . . . . .	34
Figura 13 – Acelerações verticais medidas no trecho PR. . . . .	40
Figura 14 – Acelerações verticais medidas no trecho BB. . . . .	41
Figura 15 – Sinal de aceleração com picos de menor amplitude. . . . .	42
Figura 16 – Regressão linear simples com $R^2 = 0,73$ com filtro aplicado de 0,2Hz a 25Hz para velocidade de 40km/h. . . . .	44
Figura 17 – Regressão linear simples com $R^2 = 0,80$ com filtro aplicado de 0,2Hz a 25Hz para velocidade de 60km/h. . . . .	44
Figura 18 – Modelo da linha 6 da Tabela 3 com $R^2$ de 0,97. . . . .	48
Figura 19 – Modelo da linha 10 da Tabela 4 com $R^2$ de 0,99. . . . .	48

## LISTA DE TABELAS

Tabela 1 – Classificação dos trechos obtida com o equipamento Nível e Mira. Retirado de Bisconsini <i>et al.</i> (2019). . . . .	17
Tabela 2 – Classificação dos trechos obtida com o equipamento Nível e Mira. Retirado de Bisconsini <i>et al.</i> (2019). . . . .	37
Tabela 3 – Resultado dos modelos para o conjunto de dados de 40 km/h. . . . .	45
Tabela 4 – Resultado dos modelos para o conjunto de dados de 60 km/h. . . . .	47



## LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS

### SIGLAS

CNT	Confederação Nacional de Transporte
VSA	Valor de Serventia Atual
IRI	<i>International Roughness Index</i>
GPS	<i>Global Positioning System</i>
DNIT	Departamento Nacional de Infraestrutura de Transportes
IRRE	<i>International Road Roughness Experiment</i>
SQE	Soma dos Quadrados dos Erros
SSE	<i>Error Sum of Squares</i>
SQT	Soma Total dos Quadrados
EQM	Erro Quadrático Médio
MSE	<i>Mean Squared Error</i>
RMSE	<i>Root Mean Squared Error</i>
LASSO	<i>Least Absolute Shrinkage and Selection Operator</i>
RNA	Rede Neural Artificial
ANN	<i>Artificial Neural Network</i>
CNN	<i>Convolutional Neural Network</i>
ASCE	<i>American Society of Civil Engineers</i>
RNN	<i>Recurrent Neural Network</i>
LSTM	<i>Long Short-Term Memory</i>
LOO	<i>Leave-One-Out</i>
K	<i>Kernel</i>
F	<i>Filters</i>
R <sup>2</sup>	Coeficiente de determinação
RMSVA	Raíz do valor de aceleração quadrático médio

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>11</b>
1.1	OBJETIVO GERAL . . . . .	13
1.2	OBJETIVOS ESPECÍFICOS . . . . .	13
<b>2</b>	<b>REVISÃO DE LITERATURA . . . . .</b>	<b>14</b>
2.1	IRREGULARIDADE LONGITUDINAL . . . . .	14
2.2	TRABALHOS CORRELATOS . . . . .	16
2.2.1	ANÁLISE DE FATORES RELACIONADOS AO USO DE SMARTPHONES PARA A AVALIAÇÃO DA IRREGULARIDADE LONGITUDINAL DE PAVIMENTOS . . . . .	16
2.3	INTERPOLAÇÃO . . . . .	18
2.3.1	INTERPOLAÇÃO POLINOMIAL . . . . .	19
2.4	REGRESSÃO LINEAR . . . . .	20
2.4.1	MODELO DE REGRESSÃO LINEAR SIMPLES . . . . .	20
2.4.2	MEDIDAS DE DESEMPENHO . . . . .	22
2.4.2.1	COEFICIENTE DE DETERMINAÇÃO . . . . .	22
2.4.2.2	ERRO MÉDIO QUADRÁTICO . . . . .	24
2.4.3	REGULARIZAÇÃO . . . . .	24
2.5	REDES NEURAIAS ARTIFICIAIS . . . . .	25
2.5.1	ESTRUTURA . . . . .	26
2.5.2	PROCESSO DE TREINAMENTO . . . . .	27
2.5.3	REDES MULTICAMADAS . . . . .	28
2.6	REDE NEURAL CONVOLUCIONAL . . . . .	30
2.7	REDE NEURAL LSTM . . . . .	31
2.7.1	DADOS SEQUENCIAIS . . . . .	32
2.7.2	REDE NEURAL RECORRENTE . . . . .	32
2.7.3	<i>LONG SHORT-TERM MEMORY</i> . . . . .	33
2.7.4	AVALIANDO A QUALIDADE DA REDE . . . . .	35
<b>3</b>	<b>ESTIMANDO O ÍNDICE DE IRREGULARIDADE INTERNACIONAL</b>	<b>37</b>
3.1	MATERIAIS . . . . .	37
3.2	METODOLOGIA . . . . .	38
3.3	VISUALIZAÇÃO DOS DADOS . . . . .	39
<b>4</b>	<b>ANÁLISE E DISCUSSÃO DOS RESULTADOS . . . . .</b>	<b>43</b>
4.1	RESULTADOS DO ARTIGO ANÁLISE DE FATORES RELACIO- NADOS AO USO DE SMARTPHONES PARA A AVALIAÇÃO DA IRREGULARIDADE LONGITUDINAL DE PAVIMENTOS . . . . .	43
4.2	DADOS NÃO INTERPOLADOS . . . . .	45
4.3	DADOS INTERPOLADOS . . . . .	49
<b>5</b>	<b>CONCLUSÃO . . . . .</b>	<b>50</b>
5.1	TRABALHOS FUTUROS . . . . .	50

**REFERÊNCIAS . . . . . 51**

## 1 INTRODUÇÃO

A maior parte da malha rodoviária pavimentada brasileira possui revestimento asfáltico. O Brasil possui 1.720.700 km de rodovia, dos quais aproximadamente 12% são pavimentadas. O revestimento é a camada do pavimento visível aos usuários da via, é a superfície sobre a qual trafegam os veículos e responsável por prover o conforto e a segurança no rolamento (CNT, 2019a).

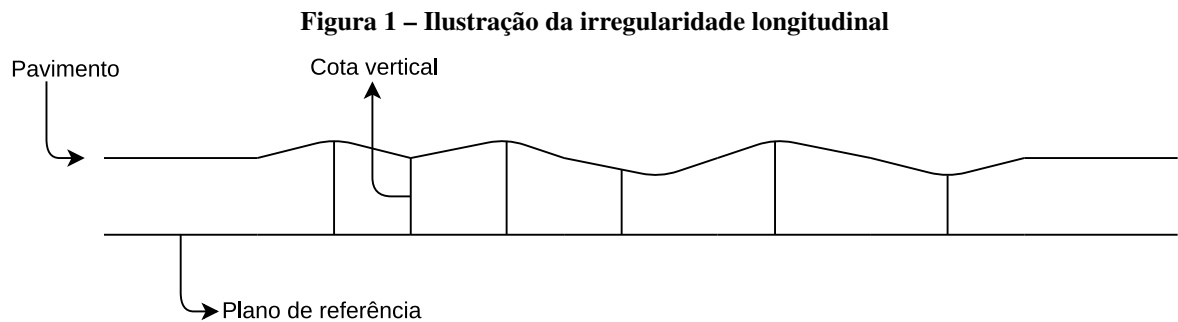
O revestimento asfáltico está presente em 99% das rodovias federais pavimentadas, desse modo, é fundamental para o desenvolvimento da infraestrutura rodoviária do Brasil tê-lo em especificações adequadas. A Confederação Nacional de Transporte (CNT), avalia todo ano 100% das rodovias federais e as principais rodovias estaduais, essas pesquisas têm demonstrado que a maioria dessas rodovias se encontram em condições precárias (CNT, 2019a).

As deficiências decorrem da manutenção insuficiente e do aumento da demanda devido ao crescente número de veículos em circulação, tanto leves quanto pesados. A ausência de investimentos, na manutenção e na conservação, contribui cada vez mais para depreciação da malha rodoviária. Portanto, para proporcionar uma infraestrutura rodoviária de qualidade, além de amplos e contínuos investimentos, é essencial que o país disponha de instrumentos de avaliação que permitam o monitoramento constante das condições da infraestrutura existente para que se tenha uma previsão de quais intervenções são necessárias (CNT, 2019b).

Ao avaliar um pavimento são definidos as condições funcionais, estruturais e operacionais dos pavimentos. Neste trabalho será dado mais ênfase às condições funcionais. A avaliação funcional de um pavimento está relacionado ao estado de sua superfície e de como este influencia no conforto e no rolamento. Assim, existem dois parâmetros de avaliação, o Valor de Serventia Atual (VSA), uma medida subjetiva baseada em notas dadas por técnicos avaliadores e o Índice de Irregularidade Internacional, parâmetro determinado por meio de medições de irregularidade longitudinal, efetuadas por meio de aparelhos especificamente projetados para este fim (DNIT, 2011).

A irregularidade longitudinal é o somatório dos desvios da superfície de um pavimento em relação a um plano de referência, como mostra a cota vertical da Figura 1, que afeta a dinâmica dos veículos, a qualidade do rolamento e a drenagem superficial da via. O parâmetro utilizado para a medida da irregularidade é o índice de irregularidade internacional, *International Roughness Index* (IRI) em inglês, um índice estatístico e objetivo, expresso em m/km, que

quantifica os desvios da superfície do pavimento. O IRI tem sido utilizado como ferramenta de controle de obras e aceitação de serviços em alguns países (DNIT, 2011).



Fonte: Autor (2019)

A irregularidade longitudinal é medida ao longo de uma linha imaginária, paralela ao eixo da estrada e, em geral, coincidente com as regiões de trilhas de roda, podendo, em alguns casos, haver o interesse de melhor detalhar o perfil, levantando-o em diversas linhas paralelas imaginárias. A irregularidade pode ser determinada com medidas topográficas ou por equipamentos medidores do perfil longitudinal com ou sem contato, ou ainda indiretamente avaliada por equipamentos do tipo resposta, que fornece um somatório de desvios do eixo de um veículo em relação à sua suspensão (DNIT, 2011).

Diversos sistemas de gerência de pavimentos utilizam perfilômetros e ferramentas sofisticadas, que têm alto custo de aquisição e operação. Em outros casos são utilizados equipamentos estáticos de precisão, mas que não são práticos para levantamentos com muitos quilômetros de malha rodoviária, conseqüentemente, ainda não há uma forma de medir a irregularidade longitudinal com precisão, exatidão, alto rendimento e baixo custo (BISCONSINI, 2016). Portanto, o objetivo deste trabalho é construir um modelo inteligente para estimar o índice de irregularidade internacional. Com a estimativa do IRI é possível, por meio do uso otimizado de equipamentos sofisticados, manter o controle da qualidade da malha rodoviária. A validação desse modelo se dará por meio da comparação dos resultados obtidos com o valor real do índice.

Os modelos inteligentes foram a rede neural convolucional para obtenção de filtros e a rede recorrente. Para o treino foram realizadas coletas de dados de alguns pavimentos do município de Pato Branco - PR utilizando sistema de posicionamento global, *global positioning system* (GPS) em inglês, e acelerômetro dos smartphones para estimar as acelerações verticais decorrentes das vibrações do chassi provocadas pelas irregularidades superficiais do pavimento. Diversas coletas foram feitas variando os modelos de veículos e de smartphones, sendo também

realizadas coletas com um equipamento de referência para o cálculo do IRI e calculado o IRI verdadeiro de cada trecho de estudo.

As redes neurais mostraram-se uma excelente alternativa para a resolução deste problema. Obteve-se um coeficiente de determinação,  $R^2$  acima de 0,95. A rede convolucional mostrou-se importante neste trabalho devido a sua característica de aprender filtros o que permite, como trabalho futuro, estudos mais detalhados no processo de seleção dos filtros.

## 1.1 OBJETIVO GERAL

Construir um modelo inteligente para estimar o índice de irregularidade internacional com os dados obtidos por meio do acelerômetro e do GPS de smartphones embarcados em veículos.

## 1.2 OBJETIVOS ESPECÍFICOS

Para atingir o objetivo geral proposto têm-se os objetivos específicos como guia:

- a Realizar o pré processamento dos dados;
- b Analisar o comportamento estatístico dos dados;
- c Construir um modelo inteligente de aprendizado de máquina;
- d Avaliar a qualidade da estimativa do valor de IRI.

## 2 REVISÃO DE LITERATURA

Neste capítulo é apresentada a fundamentação teórica para os seguintes tópicos: irregularidade longitudinal, o índice de irregularidade internacional, interpolação, regressão linear, redes neurais, trabalhos correlatos e os dados da avaliação da condição dos pavimentos com smartphones.

### 2.1 IRREGULARIDADE LONGITUDINAL

A irregularidade longitudinal é definida, de acordo com o Departamento Nacional de Infraestrutura de Transportes (DNIT), como: "A irregularidade longitudinal é o somatório dos desvios da superfície de um pavimento em relação a um plano de referência ideal de projeto geométrico, que afeta a dinâmica dos veículos, o efeito dinâmico das cargas, a qualidade ao rolamento e a drenagem superficial da via" (DNIT, 2011, p. 48). No entanto, Barella (2007) apresenta uma definição para pessoas não familiarizadas com os conceitos da área de transportes: "A irregularidade longitudinal de um pavimento representa o conjunto dos desvios indesejáveis de sua superfície, que atrapalha o rolamento rápido e suave dos veículos, gera insegurança e onera seus usuários, além de acelerar a degradação do pavimento" (BARELLA, 2007).

A irregularidade está relacionada a diversos fatores como segurança, conforto e economia. Pavimentos com elevada irregularidade dificultam a drenagem de água, acarretam no aumento do consumo de combustível, aumento no custo da manutenção dos veículos, contribui para a degradação do pavimento, diminui a dirigibilidade, a velocidade de operações dos veículos e, por fim, o conforto (BARELLA, 2007).

A importância dada a irregularidade longitudinal aumentou, principalmente após a adesão ao modelo de concessão de rodovias devido a exigência de pagamentos para a sua utilização. Se comparado com outros índices de desempenho a irregularidade conta com equipamentos confiáveis, de baixo custo e que proporcionam medições em um curto período de tempo (CAMPOS, 2004). Entretanto, não significa que os equipamentos e o custo de operação para medições com frequência são baixos, pois, ainda não há uma forma de medir, periodicamente, a irregularidade longitudinal com exatidão, precisão e preço acessível (BISCONSINI, 2016).

Foram desenvolvidas diversas técnicas e equipamentos para medir a irregularidade e, conseqüentemente, existiam muitas medidas de irregularidade não equivalentes. A solução desse

problema foi a criação de um índice padrão de irregularidade (SAYERS *et al.*, 1986a).

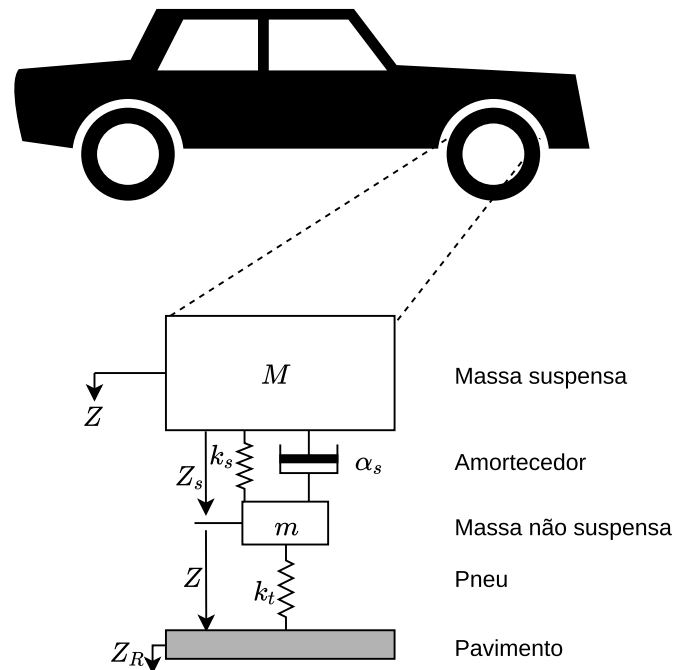
O Banco Mundial, *World Bank*, propôs o Experimento Internacional de Irregularidade dos Pavimentos, *International Road Roughness Experiment (IRRE)*, para desenvolver um índice padrão de irregularidade. O experimento foi realizado no Brasil no ano de 1982, porém padronizado em 1986, e conduzido por pesquisadores do Brasil, da Inglaterra, da França, do Estados Unidos e da Bélgica, após diversas medidas com variados equipamentos e análises de dados foi definido um índice, o Índice de Irregularidade Internacional (IRI). O IRI apresentou uma boa correlação com as medidas dos diversos equipamentos utilizados no experimento e com grande parte dos disponíveis no mundo (SAYERS *et al.*, 1986b).

O Índice de Irregularidade Internacional é uma escala de irregularidade, medida em metro por quilômetro (m/km), baseada na resposta de um veículo motorizado e genérico em relação a irregularidade do pavimento. O valor de IRI é determinado obtendo medidas precisas de irregularidade longitudinal ao longo de um perfil de pavimento e, por meio de uma simulação, com um quarto de carro com valores de massa, mola e amortecedor definidos no experimento. Esses valores são parâmetros padrões de massa, mola e amortecedor e foi denominado parâmetros ouro de carro (GILLESPIE, 1992). O modelo de um quarto de carro é representado por uma roda com estes parâmetros, conforme ilustrado na Figura 2. O IRI resume os impactos que a irregularidade causa na resposta do veículo e, portanto, é um índice indicador da qualidade do pavimento em geral (SAYERS; KARAMIHAS, 1998).

A seção a seguir apresenta como foi realizada a coleta dos dados.



Figura 2 – Representação do modelo de um quarto de carro.



Fonte: Adaptado de Bisconsini (2016)

## 2.2 TRABALHOS CORRELATOS

### 2.2.1 ANÁLISE DE FATORES RELACIONADOS AO USO DE SMARTPHONES PARA A AVALIAÇÃO DA IRREGULARIDADE LONGITUDINAL DE PAVIMENTOS

O trabalho desenvolvido por Bisconsini *et al.* (2019) foi utilizado como fonte de dados para este trabalho, o artigo apresenta a metodologia da coleta dos dados. O resultado apresentado pelo artigo foi utilizado para realizar comparações com o resultado obtido neste trabalho. Os resultados são apresentados no Capítulo 4.

A coleta dos dados foi dividida em duas partes, sendo a primeira: a coleta dos dados de aceleração vertical com o acelerômetro junto com a coleta de coordenadas do GPS. A segunda parte consiste em obter as medidas de cotas do pavimento que foram realizadas com o equipamento Nível e Mira. E Com a assistência de um programa foram calculados os IRI a partir dos valores de cotas. Esses valores de IRI são apresentados na Tabela 1.

**Tabela 1 – Classificação dos trechos obtida com o equipamento Nível e Mira. Retirado de Bisconsini *et al.* (2019).**

Segmento	Trecho (m)	IRI (m/km)	Classificação funcional DNIT (2006)
SG	0 - 100	3,8	Ruim
	100 - 200	4,0	Ruim
	200 - 300	3,8	Ruim
	300 - 400	2,8	Regular
	400 - 500	3,1	Regular
BB	0 - 100	4,4	Ruim
	100 - 200	3,8	Ruim
SF	0 - 100	1,9	Muito Bom
IP	0 - 100	4,8	Muito Ruim
	100 - 200	6,7	Muito Ruim
PR	0 - 100	4,7	Muito Ruim
	100 - 200	5,1	Muito Ruim
	200 - 300	6,6	Muito Ruim
	300 - 400	4,5	Ruim
	400 - 500	4,1	Ruim

Foram realizadas dez coletas em cada trecho somente em um sentido, devido as diferenças na quantidade de amostras em cada coleta a teoria de interpolação será discutida na próxima seção.

A seguir são apresentados três trabalhos relacionados ao tema deste trabalho, sendo o primeiro desenvolvido no Brasil e os outros dois nos Estados Unidos da América.

- **Técnicas de Aprendizado de Máquina Aplicadas a Classificação da Qualidade de Pavimentos Asfálticos utilizando Smartphones:** Este é um trabalho de conclusão de curso, apresentado ao curso de Engenharia de Computação da Universidade Federal de Ouro Preto em 2018 (NEVES, 2018). O trabalho consiste em uma melhoria para o aplicativo *RoadScan*, que realiza classificação do estado dos pavimentos com dados obtidos por smartphones, por meio de redes neurais. A diferença entre o trabalho atual e o apresentado neste tópico é que o objetivo do presente trabalho não é a classificação do estado do pavimento, mas a estimativa do índice de irregularidade internacional. O índice foi desenvolvido para quantificar a irregularidade longitudinal dos pavimentos. Os projetos de estradas devem garantir que o pavimento a ser construído tenha um requisito de IRI e ao cumprir o requisito garante-se que o pavimento apresenta uma qualidade de rolamento satisfatória

- **International Roughness Index Prediction of Flexible Pavements Using Neural Networks:** Este trabalho, publicado na *American Society of Civil Engineers (ASCE)* faz uso de redes neurais para prever o IRI de pavimentos flexíveis sujeitos a condições climáticas específicas. Os dados foram extraídos da base de dados do governo federal dos Estados Unidos da América (HOSSAIN *et al.*, 2018). Comparando o trabalho atual com o deste tópico pode-se observar que a diferença está nos dados dispostos para a construção do modelo. Estes dados foram obtidos por smartphones e independentes da condição climática. As técnicas de aprendizado de máquina realizaram classificações baseadas nos dados em que foram treinadas. O projeto de estradas nos Estados Unidos são diferentes dos projetos de estradas do Brasil, portanto, os dados do governo federal dos Estados Unidos não apresentam correlação com os dados dessa proposta.
- **Application of a Machine Learning Method to Evaluate Road Roughness from Connected Vehicles:** Publicado na ASCE e desenvolvido na Universidade do Estado de Luisiana, neste trabalho foi desenvolvido um método que utiliza redes neurais para estimar o IRI com dados obtidos por meio de sensores inerciais em veículos conectados independentemente do modelo do veículo (ZHANG *et al.*, 2018). A forma como os dados foram obtidos diferem do atual trabalho. Os veículos conectados são veículos que apresentam protocolos de comunicação e, conseqüentemente, transmitem dados, de velocidade, aceleração, aceleração vertical, da condição da suspensão e outros dados, pela internet. No atual trabalho os dados foram obtidos por meio dos sensores de acelerômetro e GPS dos smartphones dentro dos veículos.

### 2.3 INTERPOLAÇÃO

Essa seção apresenta uma introdução à interpolação polinomial e tem como base Franco (2006).

A maior vantagem da interpolação polinomial é o fato de que trabalhar com polinômios pode diminuir a complexidade da análise matemática em relação a outras formas de interpolação, inclusive, funções avançadas de cálculo, como integrais e derivadas, quando aplicadas em polinômios, resultam em polinômios. Levando em conta a vantagem descrita anteriormente, temos o Teorema de Weirstrass que afirma: *toda função contínua pode ser arbitrariamente*

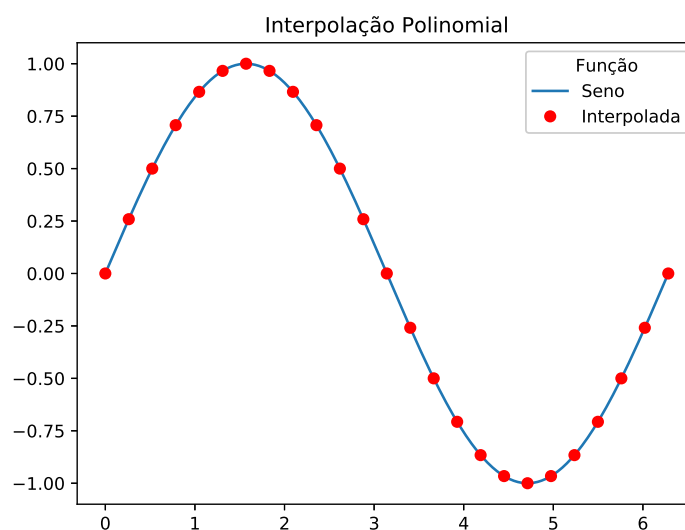
*aproximada por um polinômio.* Portanto, o próximo tópico apresenta o método de interpolação polinomial.

### 2.3.1 INTERPOLAÇÃO POLINOMIAL

O método de interpolação polinomial tem grande importância nos casos em que dispõem-se de dados que não apresentam uma função analítica ou que não tem uma função que os descreva. Outra aplicação importante deste método é quando uma função  $f(x)$  apresenta um elevado grau de complexidade e, por meio da interpolação, pode-se definir uma função aproximada mais simplificada que permite realizar uma análise com menor dificuldade, contudo a desvantagem é que a precisão, dos dados e da análise, é sacrificada.

A Figura 3 ilustra a interpolação polinomial da função trigonométrica seno, os pontos vermelhos são os valores resultantes do polinômio interpolador e a linha azul valores da função seno.

**Figura 3 – Resultado do polinômio interpolador**



**Fonte: Autor (2019)**

A interpolação polinomial, de funções ou dados, é uma tarefa trabalhosa se realizada manualmente, portanto, neste trabalho a interpolação foi construída com a assistência de computador por meio da biblioteca *SciPy* que pertence à linguagem de programação *Python*.

Após a construção do polinômio interpolador foi realizada uma análise estatística dos dados interpolados, o modelo de análise escolhido foi a regressão linear. A seção a seguir apresenta os conceitos de construção de um modelo de regressão linear.

## 2.4 REGRESSÃO LINEAR

Essa seção é baseada em Devore (2016).

Análise de regressão é a parte da estatística que estuda a relação entre duas ou mais variáveis que não possuem uma correlação bem definida. Entende-se como uma correlação bem definida uma variável que é perfeitamente descrita em relação a outra. Por exemplo a equação da posição:

$$y = v_o x + \frac{1}{2} a x^2 \quad (1)$$

sendo que:  $y$  é a distância percorrida,  $x$  o tempo,  $v_o$  velocidade inicial e  $a$  aceleração. Quando dispõem-se de dados que aparentemente não apresentam um padrão, aplica-se algumas técnicas estatísticas para tirar conclusões sobre os dados. Nesta seção é descrita a técnica de regressão linear.

### 2.4.1 MODELO DE REGRESSÃO LINEAR SIMPLES

A maneira mais simples de descrever uma correlação bem definida entre duas variáveis é por meio de uma equação linear  $y = \beta_0 + \beta_1 x$  onde cada ponto  $(x, y)$  determina uma reta com coeficiente angular  $\beta_1$  e coeficiente linear  $\beta_0$ .

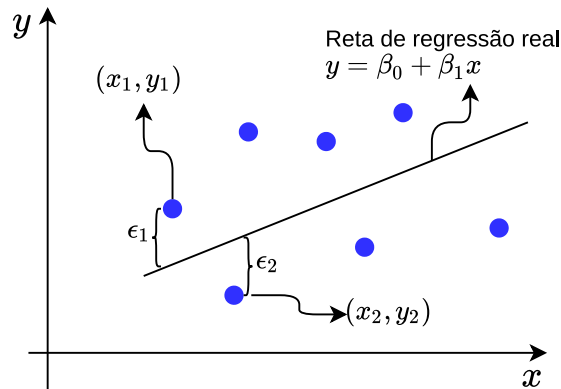
Quando o relacionamento de duas variáveis não é linear, normalmente um valor fixo de  $x$  apresenta vários valores de  $Y$ , sendo  $Y$  uma variável aleatória e  $y$  seu valor observado. Desta forma, define-se que  $x$  é a variável independente e a segunda variável,  $Y$  e  $y$ , como variável dependente. Portanto a definição generalizada de um modelo linear para a relação  $x$  e  $Y$  é descrita como:

$$Y = \beta_0 + \beta_1 x + \epsilon \quad (2)$$

Onde  $\beta_0, \beta_1$  são parâmetros que para qualquer valor fixo da variável independente  $x$ , a variável aleatória  $Y$  é relacionada por meio da Equação 2 e  $\sigma^2$  é a variação intrínseca ao modelo de regressão. A variável  $\epsilon$  é chamada de erro aleatório ou desvio aleatório e a ausência dessa variável implica que qualquer par  $(x, y)$  observado corresponderia a um ponto exato na reta  $y = \beta_0 + \beta_1 x$  que é denominada reta real de regressão. A adição da variável  $\epsilon$  faz com que o resultado  $y$  fique acima da reta se  $\epsilon > 0$  ou abaixo da reta se  $\epsilon < 0$ , portanto, os pares  $(x, y)$  ficarão dispostos em

torno da reta como mostra a Figura 4.

**Figura 4 – Modelo de regressão linear.**



**Fonte: Adaptado de Devore (2016)**

Os parâmetros  $\beta_0$ ,  $\beta_1$  e  $\sigma^2$ , quase sempre não são conhecidos, porém as amostras dos dados são conhecidas e por meio dessas amostras é estimado os parâmetros e a reta real de regressão. Segundo o modelo descrito na Equação 2 os pontos dos dados observados devem estar distribuídos em torno da reta de regressão e os parâmetros são estimados por meio do princípio dos mínimos quadrados. De acordo com o princípio, uma reta representa uma boa relação com os dados se as distâncias verticais (desvios), dos dados com a reta, são pequenas, deste modo, a melhor reta é aquela que dispõe a menor soma dos desvios.

A estimativa do coeficiente angular  $\beta_1$ , por meio do princípio dos mínimos quadrados, da reta real de regressão é:

$$\beta_1 = \hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \quad (3)$$

no qual

$$S_{xy} = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} \quad S_{xx} = \sum (x_i)^2 - \frac{(\sum x_i)^2}{n}$$

e a estimativa para o coeficiente linear é:

$$\beta_0 = \hat{\beta}_0 = \frac{\sum y_i - \hat{\beta}_1 \sum x_i}{n} = \bar{y} - \hat{\beta}_1 \bar{x} \quad (4)$$

em que  $\hat{\beta}_1$  e  $\hat{\beta}_0$  são estimativas para  $\beta_1$  e  $\beta_0$  respectivamente.

O parâmetro  $\sigma^2$  da Equação 2 determina a variação intrínseca ao modelo de regressão, desta forma, um valor elevado de  $\sigma^2$  implica que os pares  $(x_i, y_i)$  estão distantes da reta de

regressão e, seguindo essa mesma lógica, se  $\sigma^2$  é pequeno os valores dos pares tendem a ficar próximos da reta de regressão. Devido a reta real de regressão ser desconhecida a estimativa é baseada no desvio das amostras em relação a reta estimada. Desvios elevados sugerem grandes valores de  $\sigma^2$  e desvios pequenos implicam em  $\sigma^2$  menor, esses desvios são denominado resíduos.

Os valores estimados (ou previstos)  $\hat{y}_1, \hat{y}_2, \hat{y}_3 \dots \hat{y}_n$  são obtidos substituindo  $x_1, x_2, x_3, \dots, x_n$ , respectivamente, na reta de regressão estimada  $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ , sendo  $\hat{y}$  o valor estimado de  $y$ . O resíduo é a diferença do valor observado para o valor estimado  $y_1 - \hat{y}_1, y_2 - \hat{y}_2, \dots, y_n - \hat{y}_n$ . Portanto a estimativa de  $\sigma^2$  é dada como sendo a soma do quadrado dos resíduos.

A Soma do Quadrado dos Erros (resíduos), denominado pelo acrônimo SQE ou do inglês *Error Sum of Squares (SSE)*, é:

$$SSE = \sum (x_i - \hat{y}_i)^2 = \sum [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2 \quad (5)$$

e a estimativa para  $\sigma^2$  é:

$$\sigma^2 = \frac{SSE}{n - 2} = \frac{\sum (y_i - \hat{y}_i)^2}{n - 2} \quad (6)$$

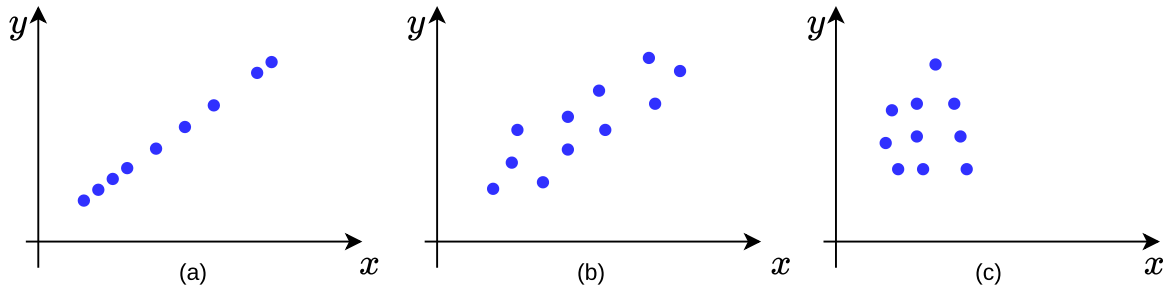
## 2.4.2 MEDIDAS DE DESEMPENHO

Após a construção do modelo apresentado, é necessário analisar a qualidade da regressão e se o modelo é suficiente para descrever os dados, portanto, nessa subseção é apresentado medidas de desempenho para o modelo de regressão linear.

### 2.4.2.1 COEFICIENTE DE DETERMINAÇÃO

A soma do quadrado dos erros (SSE) pode ser interpretada como uma medida de quantidade de variação que não pode ser atribuída a uma relação linear, ou seja, é uma variação que o modelo de regressão não explica e pode-se observar essa medida na Figura 5. Na Figura 5(a) o modelo explica toda variação existente, porém, na figura Figura 5(b) existe uma pequena variação que não é explicada, contudo, na Figura 5(c) há uma elevada variação inexplicada.

Figura 5 – (a) Para  $SSE = 0$  onde não há variação; (b) Um valor pequeno de SSE; (c) Uma variação que não apresenta relacionamento linear e um valor de SSE elevado.



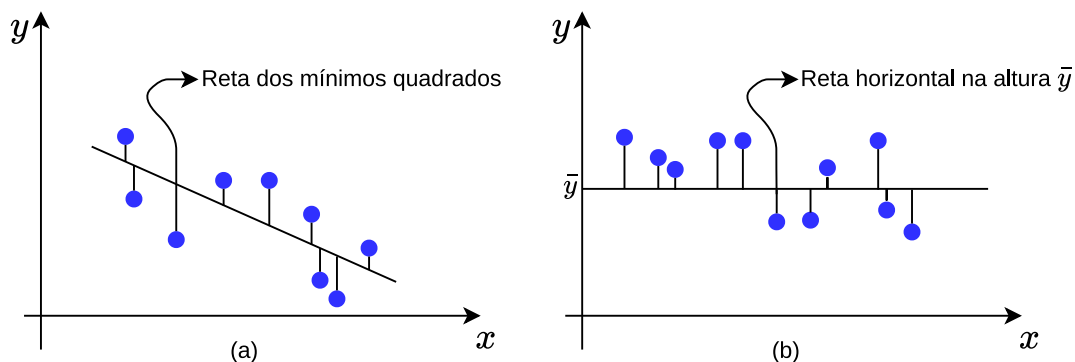
Fonte: Adaptado de Devore (2016)

Uma medida quantitativa do total de variação observada nos valores de  $y$  é dada pela soma total dos quadrados (SQT) do inglês *total sum of squares* ( $SST$ ), dada por:

$$SST = S_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n} \quad (7)$$

Como SSE é a soma do quadrado dos desvios em relação a reta estimada de regressão, analogamente, SST é soma do quadrado da variação da altura horizontal em relação ao valor estimado  $\bar{y}$  como mostrado na Figura 6. Contudo, a soma dos desvios quadrados em relação a reta dos mínimos quadrados é menor que a soma dos desvios quadrados em relação a qualquer outra reta, portanto  $SSE < SST$ , implica que, a razão  $SSE/SST$  é menor que 1. A razão  $SSE/SST$  é a proporção da variação total que o modelo de regressão não consegue explicar, consequentemente,  $1 - SSE/SST$  é a proporção da variação  $y$  observada explicada pelo modelo.

Figura 6 – (a) Soma do Quadrado dos Erros em relação a reta de regressão; (b) Soma total dos quadrados em relação a reta horizontal na altura  $\bar{y}$ .



Fonte: Adaptado de Devore (2016)

O Coeficiente de determinação é denotado por  $r^2$  e é dado por

$$r^2 = 1 - \frac{SSE}{SST} \quad (8)$$



o qual é interpretado como a proporção da variação de  $y$  observada que pode ser explicada pelo modelo de regressão linear simples.

Quanto maior o valor de  $r^2$  melhor é o modelo de regressão ao descrever a variação de  $y$ , se o valor de  $r^2$  é pequeno, é um bom indicativo de que o modelo de regressão não descreve bem a relação entre as variáveis e portanto é recomendado utilizar outro modelo mais eficiente.

#### 2.4.2.2 ERRO MÉDIO QUADRÁTICO

O Erro Quadrático Médio (EQM) do inglês *Mean Squared Error (MSE)* é uma medida quantitativa da performance do modelo de regressão linear. É a média da soma dos quadrados dos erros SSE. O MSE é importante para comparar diferentes modelos de regressão e para aprimorar os parâmetros da reta de regressão, o qual é dado por:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

#### 2.4.3 REGULARIZAÇÃO

O *MSE*, também, serve para indicar quando há um sobre-ajuste (*overfitting*) do modelo em relação aos dados, isto é, quando o modelo apresenta bons resultados para prever os dados conhecidos, porém, é ineficaz para novos resultados (RASCHKA; MIRJALILI, 2017).

Quando ocorre o problema do *overfitting* a regularização é uma abordagem para a solução do problema adicionando termos matemáticos no modelo que minimizam os valores dos parâmetros do modelo de regressão e, conseqüentemente, induzem uma penalidade. As abordagens populares de regularização para regressão linear são a *Ridge Regression* e *Least Absolute Shrinkage and Selection Operator (LASSO)* (RASCHKA; MIRJALILI, 2017).

A *Ridge Regression* é uma regularização L2 que adiciona uma soma quadrada de pesos à função de mínimos quadrados, dado por:

$$J(w)_{ridge} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \|w\|_2^2 \quad (10)$$

sendo:

$$L2 : \lambda \|w\|_2^2 = \lambda \sum_{j=1}^m w_j^2$$

Aumentando o parâmetro  $\lambda$  aumenta a força da regularização e diminuí os parâmetros do nosso modelo (RASCHKA; MIRJALILI, 2017).

O LASSO, uma regularização L1, dependendo da força da regularização pode levar o valor de alguns parâmetros para zero, sendo útil como uma técnica de seleção de característica (RASCHKA; MIRJALILI, 2017), dada por:

$$J(w)_{LASSO} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \|w\|_1 \quad (11)$$

em que

$$L1 : \lambda \|w\|_1 = \lambda \sum_{j=1}^m |w_j|^2$$

As regularizações L1 e L2, minimizam os valores dos parâmetros dos modelos e realiza uma seleção de características, respectivamente, e são aplicadas para algoritmos de aprendizado de máquina.

As regularizações tornam os modelos de regressão linear mais eficazes, porém, dados com muitas variáveis e com pouca correlação a regressão linear simples torna-se um modelo limitado. Outra técnica para regressão são as redes neurais artificiais, as RNAs podem aprender modelos de regressão mais robustos, que aprendem complexas estruturas de classificação. A seção a seguir discutirá sobre as redes neurais.

## 2.5 REDES NEURAIIS ARTIFICIAIS

As Redes Neurais Artificiais - RNA são um dos métodos de aprendizado mais robustos na aproximação de funções reais, discretas e vetoriais. Para algumas classes de problemas as redes neurais são um dos métodos mais eficazes atualmente. Há excelentes resultados em problemas como o reconhecimento da escrita manual, reconhecimento facial e reconhecimento de palavras faladas (MITCHELL, 1997).

As redes neurais foram inspiradas por meio da observação biológica do aprendizado humano e pela observação das complexas e interconectadas redes de neurônios do cérebro. A construção de redes neurais artificiais é dada por meio de um conjunto de unidades densamente interconectadas, em que as unidades recebem um número de várias entradas e respondem apenas um valor que pode ser entradas de outras unidades (MITCHELL, 1997).

O cérebro humano contém uma rede densamente interconectada de, aproximadamente,  $10^{11}$  neurônios em que cada um é conectado, em média, a  $10^4$  neurônios. Através das conexões com outros neurônios, um neurônio é excitado ou inibido e a transição de estado mais rápida conhecida é da ordem de  $10^{-3}$  segundos. Comparando-se com o tempo de um computador, que é

da ordem de  $10^{-10}$  segundos, a troca do neurônio é lenta, entretanto os humanos são capazes de tomar decisões complexas em velocidade surpreendente, apesar de as RNAs serem inspiradas no sistema biológico humano, ainda existem muitas complexidades biológicas que não são modeladas pelas RNAs e muitas características não têm relações com o funcionamento biológico (MITCHELL, 1997).

### 2.5.1 ESTRUTURA

As redes neurais são compostas de nós, ou unidades, diretamente conectadas por uma ligação, do inglês, *link*. O *link* que parte da unidade  $i$  para a unidade  $j$  serve para propagar a ativação  $a_i$  de  $i$  para  $j$  como mostra a Figura 7. Cada link tem um valor numérico associado, chamado de peso, do inglês *weight*  $w_{i,j}$ , que determina o quão significativa é a conexão. Cada unidade tem uma entrada, *bias*,  $a_0 = 1$  associada a um peso  $w_{0,j}$  e cada unidade  $j$  realiza a soma dos pesos multiplicados pelos valores de suas entradas (RUSSELL; NORVING, 2009):

$$in = \sum_{i=0}^n w_{i,j} a_i \quad (12)$$

E então para a saída da unidade é aplicado a função de ativação  $g$ :

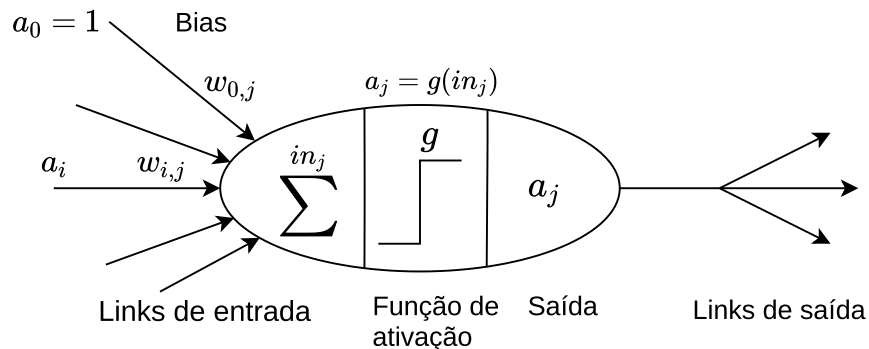
$$a_j = g(in_j) = g\left(\sum_{i=0}^n w_{i,j} a_i\right) \quad (13)$$

Quando a saída da função de ativação é maior que um valor de limiar a saída é  $a_j = 1$ , se não, a saída é  $a_j = -1$ , esse tipo de unidade é chamado de *perceptron*. Precisamente, sejam as entradas  $x_1, x_2, \dots, x_n$  a saída  $o$  pelo perceptron é:

$$o(x_1, \dots, x_n) = \begin{cases} 1 & \text{se } w_0 + w_1 x_1 + \dots + w_n x_n > 0 \\ -1 & \text{caso contrário} \end{cases}$$

De forma que o peso  $w_i$  determina a contribuição das entradas  $w_1 x_1 + \dots + w_n x_n$  para a saída do perceptron. Observa-se que  $w_0$  (bias) é o limiar que a saída  $o(x_1, \dots, x_n)$  deve superar para ser 1 (MITCHELL, 1997).

**Figura 7 – Modelo matemático simples de um neurônio. A saída é o resultado da função de ativação,  $a_i$  é a saída da função de ativação da unidade  $i$  e  $w_{i,j}$  seu peso associado ao link da unidade  $i$  para a unidade atual  $j$ .**



Fonte: Adaptado de Russell e Norving (2009)

O perceptron simples (apenas uma unidade) é útil para representar funções booleanas, por exemplo, a porta E (*AND*), pode ser representada por uma unidade perceptron de duas entradas onde a saída, se for 1 então é verdadeiro (*true*), se  $-1$  então falso (*false*). Porém, para as saídas corresponderem aos valores corretos de entrada é necessário definir os pesos, esse processo é chamado de treinamento (MITCHELL, 1997).

### 2.5.2 PROCESSO DE TREINAMENTO

O objetivo é compreender o processo de aprendizagem das redes que contém muitas unidades conectadas. A partir de agora as unidades são denominados neurônios. Para entender este processo começa-se o estudo do aprendizado da rede para o neurônio *perceptron* simples. O processo de aprendizado consiste em determinar um vetor de pesos que produz a resposta correta  $\pm 1$  para cada entrada, no qual pertence ao conjunto de treino da rede (MITCHELL, 1997).

Uma forma de aprendizado dos pesos é começar com um valor aleatório no vetor de pesos, aplicar o neurônio *perceptron* e corrigir o valor do peso a cada vez que o neurônio classifica o conjunto de treinamento de forma errada. Esse processo é repetido até que todo o conjunto de treino seja classificado corretamente. Os pesos são modificados a cada passo de acordo com a regra:

$$w_i = w_i + \Delta w_i \quad (14)$$

Em que é:

$$\Delta w_i = \eta(t - o)x_i \quad (15)$$

O parâmetro  $\eta$  é uma constante positiva chamada de taxa de aprendizado,  $t$  é a resposta correta

que o neurônio deveria gerar e  $o$  a resposta que o neurônio gera. O objetivo da taxa de aprendizado é moderar a intensidade que os pesos são modificados a cada iteração. Caso o neurônio classifique corretamente, a diferença  $t - o$  será igual a zero e portanto não haverá mudanças no valor dos pesos (MITCHELL, 1997). A regra pode ser composta por uma função denominada custo ou erro.

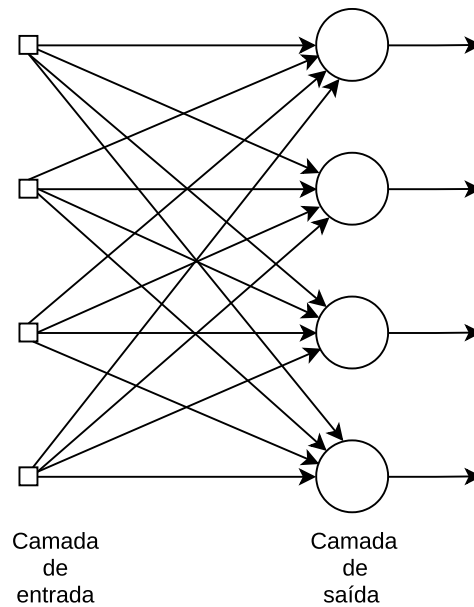
As RNAs estão sujeitas ao *overfitting*, as regularizações apresentada na subseção 2.4.3 são ferramentas poderosas para o tratamento do *overfitting* tanto em RNAs ou em regressões lineares. Ambas as regularizações adicionam termos extras a função de custo e objetivam diminuir o valor dos pesos, na regularização L1 os pesos diminuem de forma constante até zero, em contrapartida, a regularização L2 diminui os pesos proporcionalmente ao peso  $w$ . A implicação disso é que a regularização L1 apresenta impactos maiores em pesos pequenos, enquanto, a regularização L2 impacta pesos maiores (ACADEMY, 2019).

Existem vários algoritmos de aprendizado, o apresentado acima é relevante para redes simples, que contém apenas uma camada de neurônios, e para compreender o processo de aprendizado. A próxima subseção tratará das redes multicamadas e apresentará um algoritmo de aprendizado de redes multicamadas.

### 2.5.3 REDES MULTICAMADAS

Uma rede neural em camadas é estruturada em uma camada de entrada e a camada de saída, este tipo de rede é classificado como rede de alimentação para a frente, ou direta, do inglês *feedforward*. A rede é classificada como rede de camada única porque a camada de entrada não contém neurônios, apenas a camada de saída, como ilustra a Figura 8 (HAYKIN, 2001).

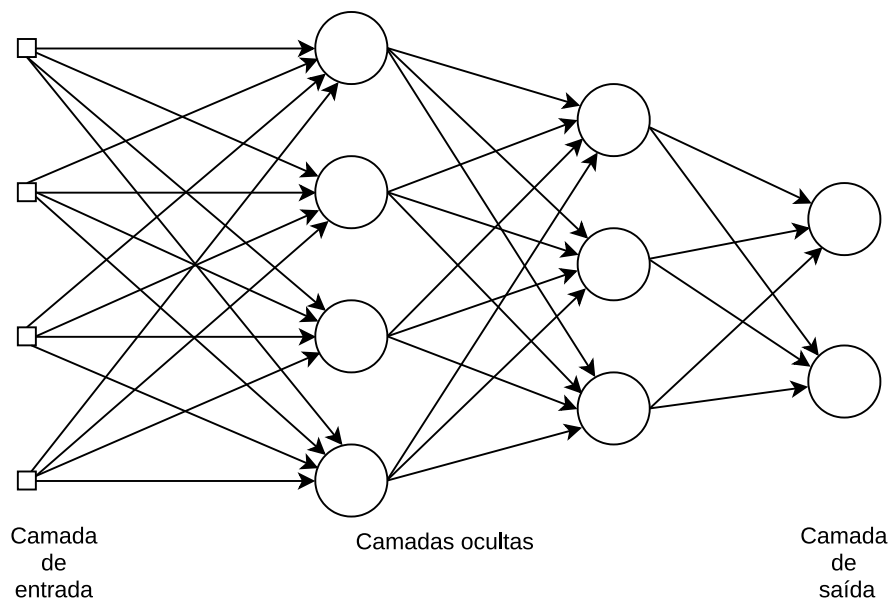
**Figura 8 – Rede neural em camadas.**



**Fonte: Adaptado de Haykin (2001).**

A outra classe de rede neural *feedforward* é a rede com camadas entre as camadas de entrada e de saída, essas camadas são denominadas camadas ocultas, em inglês *hidden layer*, os neurônios dessas camadas são denominados neurônios ocultos, adicionando-se camadas ocultas a rede torna-se capaz de aprender estatísticas de ordem elevada. A camada de entrada fornece as entradas para a primeira camada oculta, os sinais de saída dessa camada alimenta a segunda camada e assim por diante para o resto da rede. O conjunto de sinais de saída da rede constitui a resposta da rede como mostra a Figura 9 (HAYKIN, 2001).

**Figura 9 – Rede neural multicamadas.**



**Fonte: Adaptado de Haykin (2001).**

O algoritmo de retropropagação de erros, *backpropagation* em inglês, é um algoritmo amplamente utilizado para treinamento de redes multicamadas *feedforward*. O algoritmo consiste na aplicação da regra da cadeia para computar a influência de cada peso na rede em relação a uma função de erro  $E$ .

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_i} \frac{\partial o_i}{\partial in_i} \frac{\partial in_i}{\partial w_{ij}}$$

Em que  $w_{ij}$  é o peso do neurônio  $j$  para o neurônio  $i$ ,  $o_i$  é a saída e  $in_i$  é o somatório das entradas do neurônio  $i$  multiplicadas pelo seu peso. Portanto, a atualização dos pesos acontece por meio de um gradiente descendente:

$$w_{ij}(t + 1) = w_{ij}(t) - \eta \frac{\partial E}{\partial w_{ij}} \quad (16)$$

o algoritmo é aplicado na camada de saída e ocasiona a atualização dos pesos fazendo uma retropropagação (RIEDMILLER; BRAUN, 1993).

Um tipo de rede neural *feedforward* é a rede neural convolucional muito utilizada no processamento e análises de imagens. O próximo tópico apresentará os conceitos de rede neural convolucional.

## 2.6 REDE NEURAL CONVOLUCIONAL

As redes neurais convolucionais, *Convolutional neural networks* (CNN) em inglês, foram inspiradas no funcionamento do cortex visual do cérebro humano no reconhecimento de objetos. As CNNs trouxeram grandes avanços em aprendizado de máquina e nas aplicações de visão computacional (RASCHKA; MIRJALILI, 2017). Uma característica marcante dessa rede é a extração de característica, *feature extraction*.

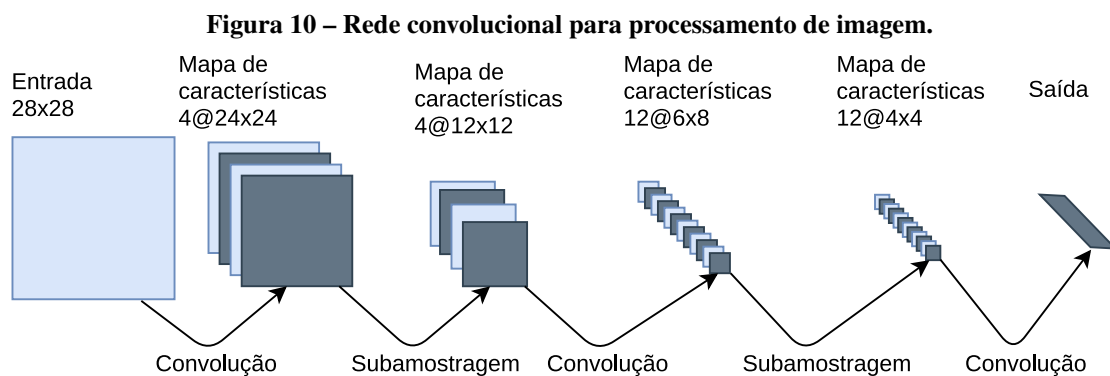
A extração de características é a chave do aprendizado de máquina e as redes neurais são capazes de extraírem características a partir dos dados em sua forma mais bruta (sem filtros). Portanto, é comum definir que as primeiras camadas, após a camada de entrada, são para extraírem características de baixo nível, *low-level features* (RASCHKA; MIRJALILI, 2017).

As redes neurais convolucionais profundas (redes que contêm muitas camadas) definem uma hierarquia de características - *feature hierarchy*, combinando camadas de extração de características das mais baixas até as mais altas, (*high-level features*). No caso de uma extração de características de uma imagem, uma característica de baixo nível seria a borda, extraída nas primeiras camadas e combinadas com outras camadas para formar características de alto nível,

como o formato de um objeto (carro, cachorro), formando um mapa de características, *feature maps* (RASCHKA; MIRJALILI, 2017).

As camadas computacionais são formadas por vários mapas de características, em que cada mapa está na forma de um plano no qual os neurônios, que compõem o mapa, estão restritos a compartilhar o mesmo conjunto de pesos. Essa restrição garante a invariância ao deslocamento, que é feita por um mapa de características através do uso de convolução com um filtro/kernel pequeno seguido de uma função sigmóide, e garante a redução do número de variáveis livres por meio do compartilhamento dos pesos (HAYKIN, 2001).

As camadas convolucionais são compostas de vários mapas de características, cada mapa com conjunto de pesos diferentes, em que várias características são extraídas. Cada camada convolucional é seguida por uma camada que realiza a média local e subamostragem, *subsampling* em inglês, reduzindo a resolução do mapa de características. Esta operação contribui para a redução da sensibilidade de saída do mapa em relação aos deslocamentos e outras formas de distorção, a Figura 10 ilustra uma rede CNN aplicada em processamento de imagem (LECUN; BENGIO, 1995).



**Fonte: Adaptado de LeCun e Bengio (1995).**

A CNN é um perceptron de multicamada projetado para reconhecer formas bidimensionais (HAYKIN, 2001). A criação de filtros, por meio das camadas convolucionais, é a justificativa para a utilização dessa rede neste trabalho.

## 2.7 REDE NEURAL LSTM

Nesta seção é introduzido os conceitos das redes neurais utilizadas em dados sequenciais.



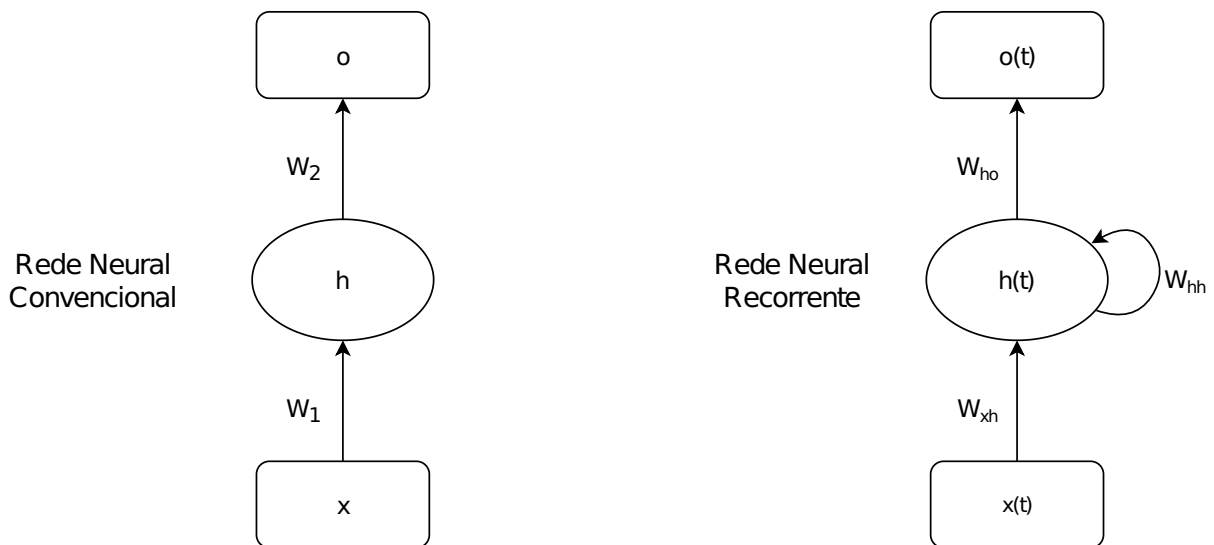
### 2.7.1 DADOS SEQUENCIAIS

A diferença entre os dados sequenciais para os dados não sequenciais é que os elementos que formam uma sequência são ordenados e dependentes entre si, ou seja, o elemento atual precisa do elemento anterior para ter algum significado. O valor de uma ação no mercado financeiro, uma frase ou um texto são exemplos de dados sequenciais. Para treinar redes com este tipo de dado foi desenvolvida a rede neural recorrente – *recurrent neural network* - RNN (RASCHKA; MIRJALILI, 2019).

### 2.7.2 REDE NEURAL RECORRENTE

A principal diferença entre a RNN e a rede neural convencional está na arquitetura. As redes recorrentes apresentam uma componente recursiva no neurônio como mostra a Figura 12 (RASCHKA; MIRJALILI, 2019).

**Figura 11 – Representação das camadas,  $x(t)$  a camada de entrada,  $h(t)$  a camada oculta e  $o(t)$  a camada de saída.**



Fonte: Adaptado de Raschka e Mirjalili (2019).

Na rede convencional a informação flui da camada de entrada para a camada de saída através das camadas ocultas, todavia, na RNN as camadas ocultas recebem duas informações, a informação do tempo atual e a informação do instante anterior. Essa componente recursiva recebe o nome de *recurrent edge*. Com essa modificação, a rede ganha a capacidade do efeito memória dos eventos do passado (RASCHKA; MIRJALILI, 2019).

O processo de ativação dos neurônios é semelhante aos da rede multicamadas padrão, a

função de ativação da primeira camada oculta é computada por meio da combinação linear do vetor de pesos e da unidade de *bias*. Segue abaixo a função de ativação da primeira camada no tempo  $t$ .

$$Z_h(t) = x(t)W_{xh} + h(t-1)W_{hh} + b_h \quad (17)$$

em que  $b_h$  é o *bias* da unidade oculta.

A função de ativação da camada de saída no tempo  $t$  é calculada como mostra a equação a seguir.

$$o(t) = W_{ho}h(t) + b_o \quad (18)$$

em que  $b_o$  é o *bias* da unidade de saída (RASCHKA; MIRJALILI, 2019).

### 2.7.3 LONG SHORT-TERM MEMORY

Para treinar as redes RNN utiliza-se o algoritmo de *backpropagation* modificado. Como resultado da alteração surge um fator multiplicativo no cálculo do gradiente da função de erro. Esse fator, quando muito pequeno gera um problema chamado de *vanishing gradient* que consiste em zerar alguns pesos e fazer com que a rede perca algumas informações, por outro lado, quando muito grande, *exploding gradient*, faz com que o treino seja interrompido com erro de execução devido a um valor de peso muito grande. Esses problemas são comuns quando a sequência é muito grande (RASCHKA; MIRJALILI, 2019).

Seguindo a ideia do algoritmo de *backpropagation* apresentado na seção 2.5 a seguir é apresentado a equação que calcula a influência de cada peso em relação a uma função de erro,  $E$ , para as redes recorrentes. O objetivo é todos valores de erros da função  $E$  no tempo  $t = 1$  até  $t = T$ , ou seja:

$$E = \sum_{t=1}^T E(t) \quad (19)$$

Desde que o erro no tempo  $t$  depende das camadas ocultas de todos os tempos anteriores, a derivada será computada como:

$$\frac{\partial E(t)}{\partial W_{hh}} = \frac{\partial E(t)}{\partial O(t)} \times \frac{\partial o(t)}{\partial h(t)} \times \left( \sum_{k=1}^t \frac{\partial h(t)}{\partial h(k)} \times \frac{\partial h(k)}{\partial W_{hh}} \right)$$

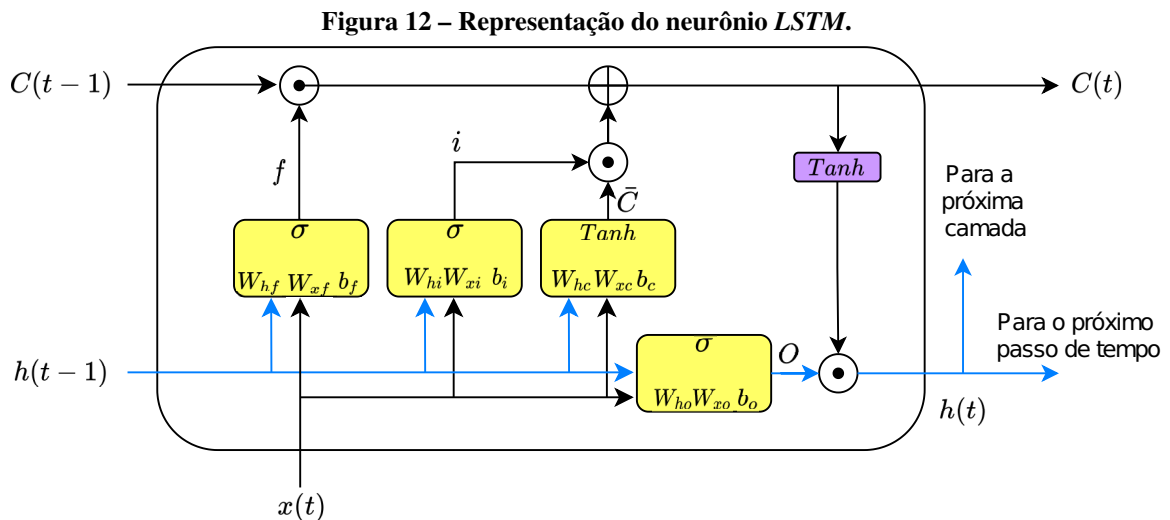
em que  $\frac{\partial h(t)}{\partial h(k)}$  é computado como a multiplicação de tempo adjacente:

$$\frac{\partial h(t)}{\partial h(k)} = \prod_{i=k+1}^t \frac{\partial h(i)}{\partial h(i-1)}$$

Sendo este o fator multiplicativo citado no primeiro parágrafo desta subseção (RASCHKA; MIRJALILI, 2019).

Uma solução para este problema foi a criação de um neurônio chamado *Long Short-Term Memory - LSTM*. Este neurônio é uma célula de memória que substitui as camadas ocultas de uma RNN convencional. Cada célula tem um *recurrent edge* e o valor de peso desejado, evitando os problemas de *exploding* e *vanishing*, além disso, a célula armazena as informações que passam pela rede gerando a possibilidade de treinarmos longas sequências de dados (RASCHKA; MIRJALILI, 2019).

O neurônio *LSTM* é apresentado na imagem abaixo



Fonte: Adaptado de Raschka e Mirjalili (2019).

em que  $C(t - 1)$ : é a matriz do estado da célula no tempo anterior, é a variável que armazena a memória do neurônio,  $C(t)$ : é a matriz de estado atual da célula,  $h(t - 1)$ : é a matriz da camada oculta do tempo anterior,  $h(t)$ : é a matriz de saída da camada oculta do tempo atual,  $x(t)$ : é a entrada do neurônio,  $f$ : *forget gate*,  $i$ : *input gate*,  $\bar{C}$ : *input node*,  $O$ : *output*,  $\sigma$ : função de ativação logística,  $\tanh$ : função de ativação tangente.

O neurônio LSTM é formado por outros neurônios convencionais e pelas matrizes de estado  $C(t - 1)$  e  $C(t)$ . As matrizes servem para dar o efeito memória ao neurônio *LSTM*, os neurônios convencionais são chamados de portões (*gates*) e estão identificados na figura pelas letras:  $f$ ,  $i$  e  $\bar{C}$ . O primeiro *gate* é o *Forget Gate*,  $f$ . Este *gate* serve para selecionar quais informações serão esquecidas pela célula  $C(t - 1)$ , este neurônio recebe a saída da camada oculta do tempo anterior  $h(t - 1)$ , a entrada atual do neurônio  $x(t)$  sendo que cada entrada é multiplicada por sua matriz de peso correspondente  $W_{hf}$  e  $W_{xf}$  e somada a um *bias*. Para

finalizar é aplicada a função de ativação logística, como mostra a equação abaixo, nota-se a semelhança com a Equação 13:

$$f_t = \sigma(W_{hf}h(t-1) + W_{xf}x(t) + b_f) \quad (20)$$

A saída dessa equação fará uma multiplicação elemento a elemento,  $\odot$ , com a matriz de estado  $C(t-1)$ , se o valor resultante de  $f_t$  é zero, a informação da matriz de estado será esquecida, caso contrário, permanecerá na memória do neurônio.

Os portões  $i$  e  $\bar{C}$  servem para definir quais novas informações serão adicionados na rede, ou seja, o quanto que a célula de estado  $C(t)$  será modificada. Eles são calculados da seguinte forma:

$$i_t = \sigma(W_{hi}h(t-1) + W_{xi}x(t) + b_i)$$

$$\bar{C}_t = \tanh(W_{hc}h(t-1) + W_{xc}x(t) + b_c)$$

O estado de célula atual é computado como:

$$C(t) = (C(t-1) \odot f_t) \oplus (i_t \odot \bar{C}_t)$$

E por fim, o neurônio de saída da rede,  $O$ , decide como atualizar os valores das camadas ocultas, da seguinte forma:

$$O_t = \sigma(W_{ho}h(t-1) + W_{xo}x(t) + b_o)$$

Portanto, a saída do neurônio *LSTM* no tempo atual é computada como:

$$h(t) = O_t \odot \tanh C(t)$$

e  $h(t)$  serve como entrada para a componente recursiva do neurônio e como entrada para a próxima camada oculta (RASCHKA; MIRJALILI, 2019).

#### 2.7.4 AVALIANDO A QUALIDADE DA REDE

Uma forma convencional de analisar a qualidade de predição da rede é o método de *HoldOut*. Este método consiste em dividir o conjunto dos dados em outros dois conjuntos: o conjunto de treino e o conjunto de teste. O conjunto de treino é utilizado para o aprendizado da rede e o conjunto de teste para avaliar a qualidade de predição da rede, ou seja, o quão bem ela aprendeu e consegue generalizar os dados (RASCHKA; MIRJALILI, 2019).

O problema desta abordagem é que ela é pouco eficiente quando há poucos dados disponíveis. Para contornar este problema é utilizado o método de validação cruzada, *Leave-One-Out*, que separa aleatoriamente e sem reposição um elemento do conjunto de treino e o utiliza para teste. Este processo é repetido até que todos os dados sejam utilizados para teste e a cada interação é avaliado e armazenado o resultado da rede durante o teste, uma vez obtido todos os resultados é calculado a média e este valor define a qualidade do modelo treinado (RASCHKA; MIRJALILI, 2019).

### 3 ESTIMANDO O ÍNDICE DE IRREGULARIDADE INTERNACIONAL

A partir dos dados de aceleração vertical obtidos com acelerômetro dos smartphones e das informações de tempo do GPS foi construída uma rede neural artificial para estimar o índice de irregularidade internacional.

#### 3.1 MATERIAIS

Para a realização deste trabalho as seguintes ferramentas foram utilizadas:

- Linguagem de programação *Python*: é uma linguagem de alto nível com bibliotecas ricas em algoritmos de inteligência artificial, processamento e visualização de dados. Cita-se como exemplo as bibliotecas: *Pandas*, *Sklearn*, *Keras*, *TensorFlow*, *Matplotlib*.
- *Google Colaboratory*: é um ambiente gratuito de desenvolvimento virtual para escrita, compartilhamento e análise de código que provê acesso a um computador de alta performance. Este ambiente foi utilizado para o treinamento das redes neurais.

Entretanto, o material mais importante e que fundamentou este trabalho são os dados do artigo apresentados pelos autores Bisconsini *et al.* (2019), representados também na Tabela 2

**Tabela 2 – Classificação dos trechos obtida com o equipamento Nível e Mira. Retirado de Bisconsini *et al.* (2019).**

Segmento	Trecho (m)	IRI (m/km)	Classificação funcional DNIT (2006)
SG	0 - 100	3,8	Ruim
	100 - 200	4,0	Ruim
	200 - 300	3,8	Ruim
	300 - 400	2,8	Regular
	400 - 500	3,1	Regular
BB	0 - 100	4,4	Ruim
	100 - 200	3,8	Ruim
SF	0 - 100	1,9	Muito Bom
IP	0 - 100	4,8	Muito Ruim
	100 - 200	6,7	Muito Ruim
PR	0 - 100	4,7	Muito Ruim
	100 - 200	5,1	Muito Ruim
	200 - 300	6,6	Muito Ruim
	300 - 400	4,5	Ruim
	400 - 500	4,1	Ruim

## 3.2 METODOLOGIA

Para atingir o objetivo deste trabalho foi necessário realizar um pré-processamento dos dados. Os dados estão divididos em trechos de 100 metros e cada trecho apresenta uma classificação de IRI. Para melhor compreensão desta etapa são expostas as etapas do pré-processamento do o segmento SG. Conforme observado na Tabela 1 este segmento tem 500 metros de distância e está dividido em trechos de 100 metros, ou seja, dentro desse segmento há 5 trechos.

Os trechos, quando ordenados, formam uma sequência e essa é uma característica que não pode ser perdida por que afeta o significado do resultado. Na subseção 2.2.1 foi apresentado que cada segmento têm dez coletas e cada coleta está dividida em trechos de 100 metros, portanto, as coletas foram agrupadas de forma que formem uma sequência.

Os dados foram organizados na estrutura de dados *DataFrame* que é definida na biblioteca *Pandas*. Após organizados em sequência foi construída a função que faz a interpolação dos dados. A interpolação foi feita a partir do método *interp1d* da classe *interpolate* que pertence a biblioteca *scipy*. Para padronizar a quantidade de pontos de cada trecho foi definido um limiar de 2000 pontos. A justificava para este valor é que a média da quantidade de pontos que cada trecho tem está em torno de 2000 pontos.

Após a interpolação, obteve-se a ter dois conjuntos de dados: os dados interpolados e os dados sem interpolação. A próxima etapa foi a construção das redes neurais, os modelos de rede utilizados nesse trabalho encontram-se na biblioteca *Keras* que atua nos algoritmos de aprendizado de máquina implementados na plataforma *TensorFlow*. As redes da biblioteca *Keras* são organizadas em camadas, como mostra a Figura 9. A camada de entrada é onde a dimensão dos dados é definida e logo em seguida são adicionadas as camadas ocultas. Os tipos de camadas utilizadas nesse trabalho foram:

- *Dense*: essa camada pode ter um ou vários neurônios e ela implementa o neurônio apresentado na seção 2.5.
- *Conv1D*: essa camada implementa todos os conceitos apresentados na seção 2.6, para diminuir a complexidade e manter a coerência com os dados foi utilizado a rede neural convolucional de uma dimensão.
- *LSTM*: implementa os conceitos apresentados na seção 2.7, a justificativa de escolha desta

rede é a característica sequencial dos dados.

Todas as camadas acima apresentam parâmetros e que podem ser modificados, por exemplo a quantidade de neurônios, a função de ativação e a função *loss*. No Capítulo 4 são apresentados os parâmetros utilizados nesse trabalho.

Para avaliar o resultado foi feita a comparação do resultado dos modelos de rede neural com o resultado da regressão linear obtido por Bisconsini *et al.* (2019). Para enriquecer os resultados foram treinadas várias redes neurais utilizando as camadas citadas anteriormente e para amenizar o problema da disponibilidade dos dados foi utilizado a técnica *Leave-One-Out*, introduzida na subseção 2.7.4, durante o treinamento das redes neurais. A comparação foi feita utilizando as medidas de desempenho apresentadas na seção 2.4.

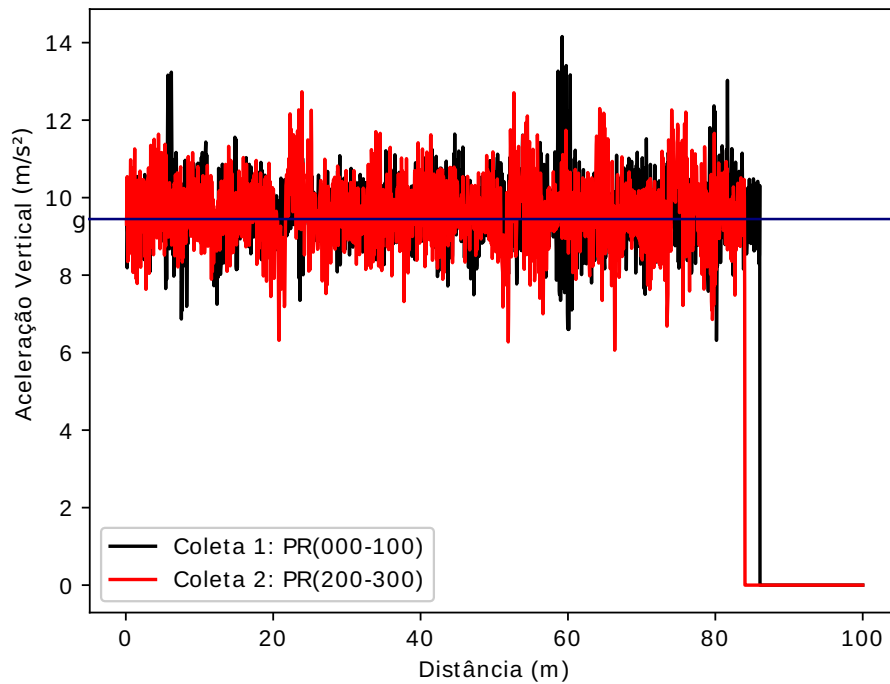
### 3.3 VISUALIZAÇÃO DOS DADOS

Esta seção apresenta os sinais medidos pelo acelerômetro dos smartphones dos trechos apresentados na subseção 2.2.1.

A Figura 13 apresenta o sinal do acelerômetro de duas coletas, sendo a primeira coleta do trecho contabilizada de 0 a 100m e a segunda coleta entre 200 a 300m. A escolha desses trechos serve para ilustrar a variação dos sinais de aceleração verticais coletados. De acordo com a Tabela 1 o trecho PR(0-100), coleta 1, apresenta um IRI de 4,7 m/km enquanto que, o trecho PR(200-300), coleta 2, apresenta o índice de 6,6 m/km. Observa-se na Figura 13 que, apesar de, a coleta 1, apresentar dois picos de aceleração vertical, a coleta 2, apresenta os maiores, e os menores, valores de aceleração vertical na maior parte do trecho, implicando uma variação maior da irregularidade, estando de acordo com os índices apresentados em que o trecho PR(200-300) é mais irregular que o trecho PR(0-100). De acordo com Tabela 1 os dois trechos estão classificados em muito ruins.



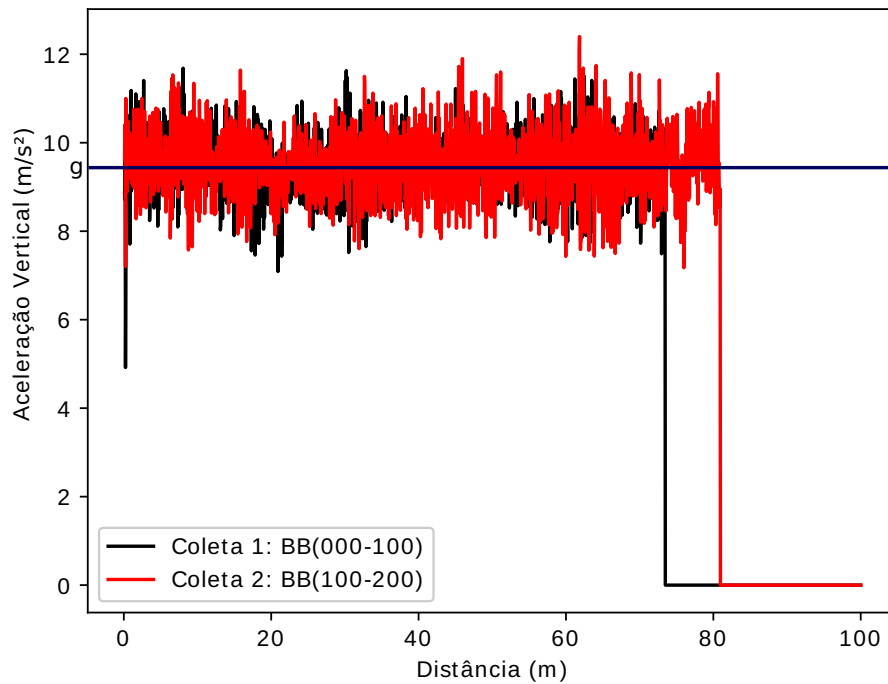
**Figura 13 – Acelerações verticais medidas no trecho PR.**



**Fonte: Autor (2019).**

A Figura 14 apresenta os trechos de 0 a 100m, coleta 1, e de 100 a 200m, coleta 2, BB(0-100) e BB(100-200), respectivamente. O trecho BB(0-100) apresenta o índice de 4,4 m/km e o segundo trecho de 3,8 m/km, portanto, a Figura 14 ilustra o sinal de acelerômetro de trechos com IRI semelhante, i.e, a diferença do índice entre os trechos é baixa. Comparando-se a Figura 14 com a Figura 13, observa-se que os trechos BB apresentam maior semelhança entre os sinais e menos picos. Entretanto, os trechos estão classificados como ruins.

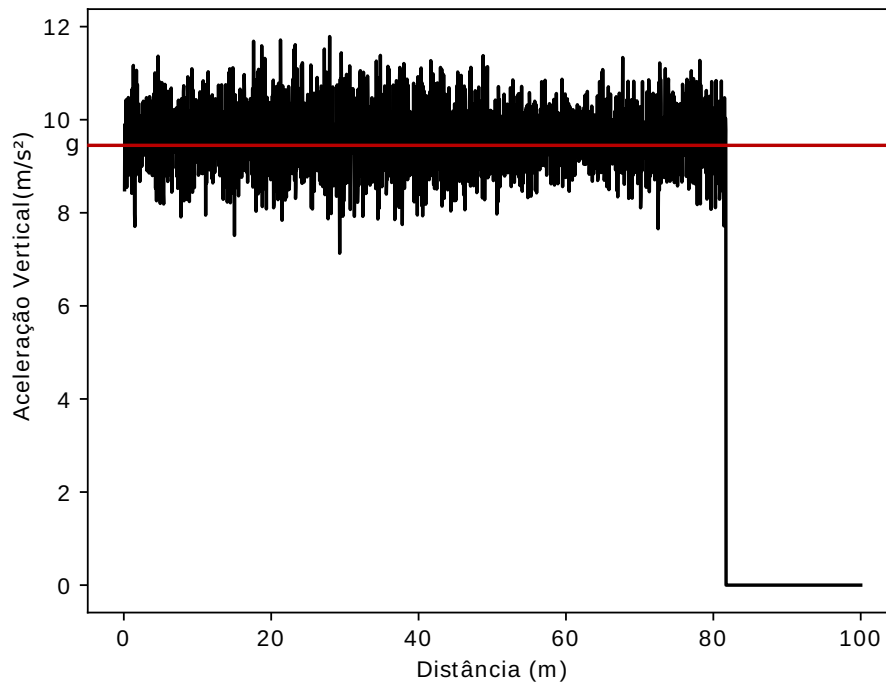
**Figura 14 – Acelerações verticais medidas no trecho BB.**



**Fonte: Autor (2019).**

A Figura 15 apresenta o sinal de acelerômetro do trecho SF(0-100), com índice de 1,9 m/km. De acordo com os dados apresentados na Tabela 1 é o único trecho com a classificação muito boa. Portanto, comparando-se com a Figura 13 e Figura 14, o sinal apresentado contém menos picos e não apresenta variações muito abruptas.

**Figura 15 – Sinal de aceleração com picos de menor amplitude.**



**Fonte: Autor (2019).**

As figuras apresentam os sinais sem a aplicação de filtros, em que a amplitude da aceleração vertical é elevada devido a influência gravitacional no acelerômetro. Como as coletas foram interrompidas antes dos 100 metros, as coletas apresentam quantidades de pontos diferentes, observa-se que na Figura 14. E problema foi resolvido aplicando-se uma interpolação polinomial, conforme apresentada na seção 2.3.

## 4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo são apresentadas duas tabelas de resultados, a primeira com o conjunto de dados com velocidade de 40 km/h e a segunda com os dados referentes a velocidade de 60 km/h. A análise é feita primeiro para os dados não interpolados e depois para os dados interpolados. É apresentado, também, o resultado obtido pelo artigo introduzido na subseção 2.2.1.

### 4.1 RESULTADOS DO ARTIGO ANÁLISE DE FATORES RELACIONADOS AO USO DE SMARTPHONES PARA A AVALIAÇÃO DA IRREGULARIDADE LONGITUDINAL DE PAVIMENTOS

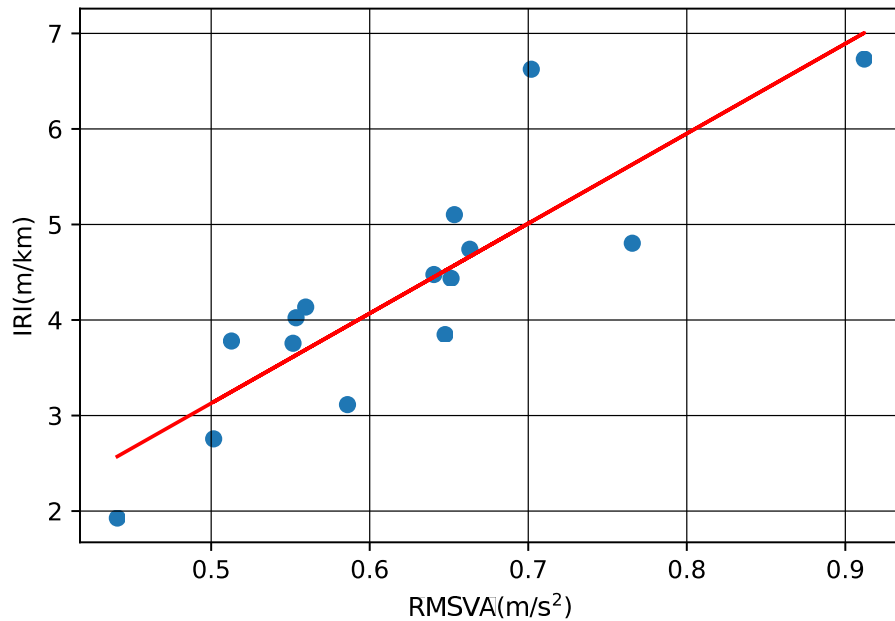
O resultado apresentado nesta seção foi reconstruído a partir dos dados do artigo desenvolvido pelos autores Bisconsini *et al.* (2019). Foram desenvolvidos dois modelos de regressão linear, um para os dados obtidos na velocidade de 40 km/h e o outro para os dados obtidos a 60 km/h. Ambos os modelos tiveram a mesma metodologia e o mesmo filtro, passa-banda, de frequência na faixa de 0,2 Hz até 25 Hz. O objetivo destes modelos foi de relacionar o valor de IRI com a raiz quadrática média da aceleração - RMSVA. O valor de RMSVA foi obtido subtraindo o valor do sinal de aceleração pela sua média como mostra a equação abaixo.

$$RMSVA = \sqrt{\frac{1}{N} \sum_{i=1}^N (Acel_i - \bar{Acel})^2} \quad (21)$$

$Acel_i$  valor de aceleração no instante  $i$ ,  $\bar{Acel}$  valor médio do sinal de aceleração,  $N$  quantidade de pontos ou de instantes do sinal de aceleração.

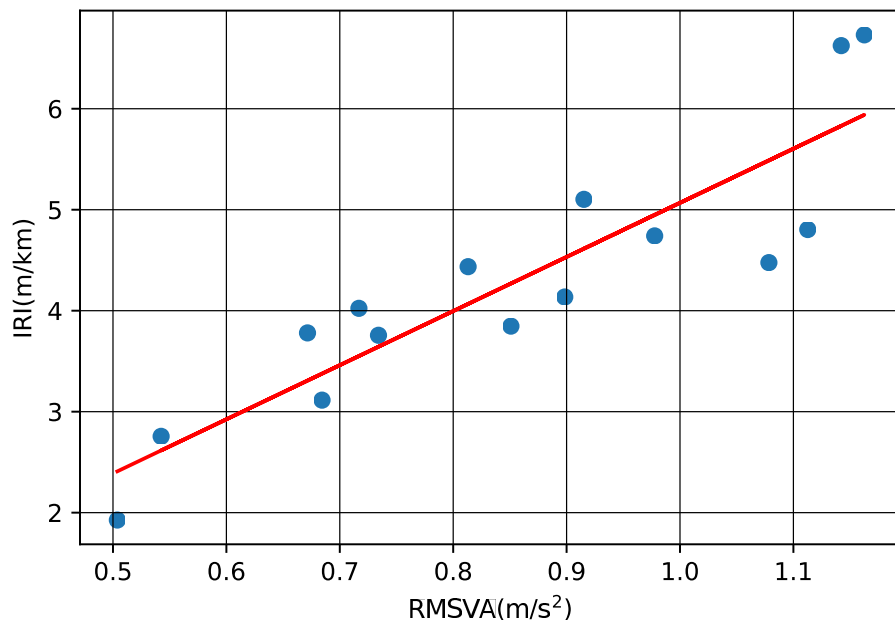
Antes de calcular o valor de RMSVA, o sinal foi filtrado por um filtro passa-banda com frequências de corte de 0,2 Hz até 25 Hz e, após o cálculo do RMSVA, foi aplicada a técnica de regressão linear simples, seção 2.4, utilizando o coeficiente de determinação, subseção 2.4.2.1. Na Figura 16 é apresentada a reta de regressão linear para o conjunto dos dados obtidos a 40 km/h e na Figura 17 para os dados obtidos a 60 km/h.

Figura 16 – Regressão linear simples com  $R^2 = 0,73$  com filtro aplicado de 0,2Hz a 25Hz para velocidade de 40km/h.



Fonte: Autor (2019) reconstruído de Biconsini *et al.* (2019).

Figura 17 – Regressão linear simples com  $R^2 = 0,80$  com filtro aplicado de 0,2Hz a 25Hz para velocidade de 60km/h.



Fonte: Autor (2019) reconstruído de Biconsini *et al.* (2019).

É importante frisar que para reconstruir os modelos acima não houve o uso de técnicas de validação cruzada. Essas técnicas são importantes para avaliar a qualidade de predição do modelo em dados desconhecidos e revelar se o conjunto possui dados enviesados. Além disso,

serve também para comparar modelos com diferentes parâmetros de construção, onde a ausência destes métodos de validação não permite enxergar possíveis irregularidades no modelo (KOHAVI, 1995).

Nas tabelas abaixo são apresentados os modelos em que as regressões foram construídas. O conjunto de dados em que os modelos foram aplicados e as estatísticas que foram utilizadas para avaliar e comparar os modelos. Na descrição do modelo algumas informações importantes são exibidas: a quantidade de camadas e qual o tipo da camada utilizada. Os parâmetros F e K são acrônimos para *Filters* e *Kernel*, respectivamente, da rede neural convolucional discutida na seção 2.6. Os dois primeiros modelos identificados como regressão linear com filtro ou sem filtro são referentes ao artigo desenvolvido pelos autores Bisconsini *et al.* (2019).

#### 4.2 DADOS NÃO INTERPOLADOS

Nas Tabela 3 e Tabela 4 há quatro modelos de redes neurais desenvolvidos com uma camada convolucional, com variações em seus parâmetros, e uma camada *Dense*. Em seguida, há duas redes neurais com as camadas convolucional e *LSTM*, com variações nos parâmetros F e K. Por fim, há três RNAs com três camadas: a camada convolucional, a camada *LSTM* e a camada *Dense*. Neste caso variou apenas a quantidade de camadas *LSTM*. Os melhores resultados estão destacados em negrito.

**Tabela 3 – Resultado dos modelos para o conjunto de dados de 40 km/h.**

MODELO		40 km/h					
		DADOS SEM INTERPOLAÇÃO			DADOS INTERPOLADOS		
		MSE	RMSE	R2	MSE	RMSE	R2
1	Regressão linear sem filtro	1,30	1,14	0,15			
2	Regressão linear com filtro	0,40	0,63	0,74			
3	1 Conv1d (2F, 2K) - 1 Dense	0,44	0,67	0,70	52,71	7,26	-2,71
4	1 Conv1d (2F, 4K) - 1 Dense	0,40	0,63	0,86	23,40	4,84	-0,06
5	1 Conv1d (4F, 4K) - 1 Dense	0,47	0,69	0,92	39,28	6,27	-0,41
6	1 Conv1d (4F, 8K) - 1 Dense	<b>0,32</b>	<b>0,56</b>	<b>0,97</b>	2,36	1,54	<b>0,69</b>
7	1 Conv1d (2F, 2K) - 1 LSTM	10,21	3,20	-0,44	9,70	3,11	-0,09
8	1 Conv1d (8F, 64K) - 1 LSTM	4,05	2,01	0,94	19,77	4,45	-12,19
9	1 Conv1d (8F, 64K) - 1 LSTM - 1 Dense	2,45	1,57	0,83	17,11	4,14	-8,30
10	1 Conv1d (8F, 64K) - 2 LSTM - 1 Dense	5,44	2,33	0,93	<b>1,58</b>	<b>1,26</b>	0,06
11	1 Conv1d (8F, 64K) - 4 LSTM - 1 Dense	3,69	1,92	0,05	1,84	1,35	-0,03

Para o desenvolvimento das redes neurais, o primeiro desafio foi o pré-processamento dos dados para deixá-los na dimensão que a biblioteca *Keras* utiliza para a camada convolucional, que neste caso é a camada de entrada. A interpolação foi desenvolvida para observar se este processamento extra contribui em um bom resultado em relação a regressão linear simples,

podendo ser uma alternativa em casos em que há dados com muita irregularidade na quantidade de amostras, ou seja, o número de pontos que compõe um trecho apresenta muita variação em relação a outro trecho.

Após o pré-processamento, iniciou-se o treinamento das redes. O segundo desafio foi encontrar os parâmetros da camada convolucional em que melhor se adequassem aos dados. Como introduzidos na seção 2.6 a rede convolucional aprende filtros de *Fourier*, ou seja, filtros passa-baixa, passa-banda ou passa-alta. A seleção correta desses parâmetros influencia muito no resultado. Para os dados utilizados neste trabalho uma pequena mudança gerou uma variação alta no valor de R<sup>2</sup>. O coeficiente de determinação, R<sup>2</sup>, é a principal métrica para analisar a performance do modelo, as demais estatísticas serviram de suporte para selecionar os parâmetros das redes.

Observa-se que essa pequena variação nos parâmetros comparando o modelo da linha 3 da Tabela 3 com o modelo da linha 4 da mesma tabela. A única mudança entre eles é a quantidade de *kernel* na camada convolucional. O modelo da linha três apresentou um coeficiente de determinação de 0,70 que é inferior a regressão linear simples com filtro. Em contrapartida, o modelo da linha 4. Apresentou um aumento significativo no R<sup>2</sup> e com apenas um aumento de duas unidades na quantidade de *kernel*. Vale ressaltar que essa mudança não gerou um aumento significativo no uso de recursos computacionais o que gera um resultado melhor que a regressão linear da linha 2 e, portanto, melhorando o resultado encontrado no artigo dos autores Bisconsini *et al.* (2019).

O motivo que levou o modelo da linha 4 ser superior a regressão linear simples é o filtro que a camada convolucional aprende, é um filtro mais robusto e mais adequado as características intrínsecas do conjunto de dado, ou seja, a rede neural adequá-se aos dados. Se for mudado o conjunto de dados obtém-se um filtro diferente, essa diferença é observada na Tabela 3 se for feita a comparação entre os modelos que tem os mesmos parâmetros da Tabela 4. O modelo da linha 4 da Tabela 4 apresenta melhor R<sup>2</sup> que o modelo da linha 4 da Tabela 3 devido a rede se adaptar melhor as características intrínsecas do conjunto de dados da Tabela 4. Entretanto, o modelo da linha 4 da Tabela 3 não é o melhor resultado obtido, os melhores são os modelos das linhas 5 e 6, considerando apenas as redes com camadas convolucionais, nota-se um R<sup>2</sup> acima de 0,90, que significa que o modelo consegue justificar muito bem a variação do conjunto de dados.

As análises feitas à Tabela 3 também se aplicam à Tabela 4. Nota-se que o conjunto de dados da Tabela 4 é bem justificado pelas regressões lineares simples e pelo modelo da linha 3,

que é o modelo mais simplificado de rede neural deste trabalho e que gerou um bom resultado. Observa-se que ao aumentar a complexidade da rede os modelos podem não encontrar padrões dentro do conjunto de dados e não gerar um bom resultado como é o caso dos modelos das linhas 6, 7 e 11 da Tabela 4 e das linhas 7 e 11 da Tabela 3.

**Tabela 4 – Resultado dos modelos para o conjunto de dados de 60 km/h.**

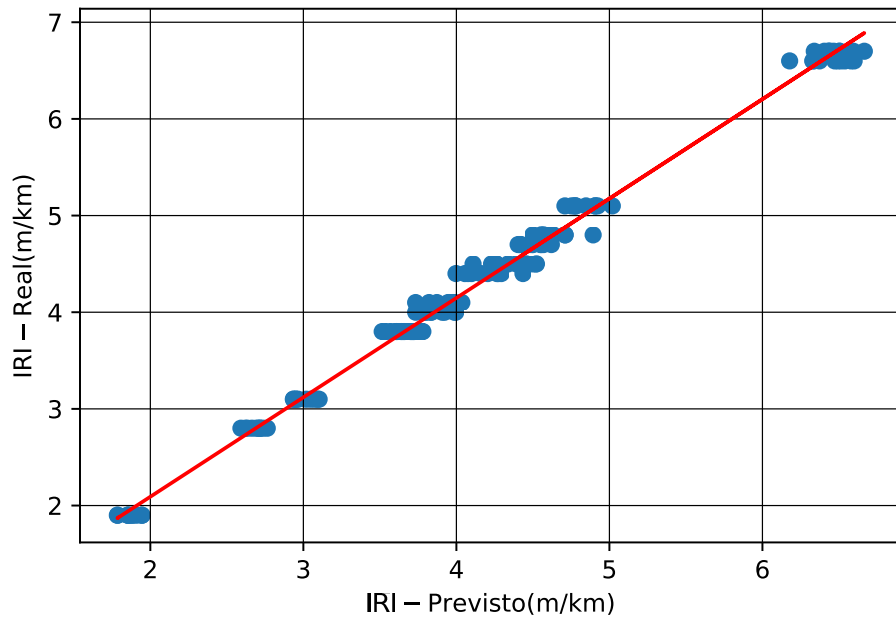
	MODELO	60 km/h					
		DADOS SEM INTERPOLAÇÃO			DADOS INTERPOLADOS		
		MSE	RMSE	R2	MSE	RMSE	R2
1	Regressão linear sem filtro	0,40	0,63	0,74			
2	Regressão linear com filtro	0,30	0,55	0,80			
3	1 Conv1d (2F, 2K) - 1 Dense	0,25	0,50	0,96	16,82	4,10	-8,26
4	1 Conv1d (2F, 4K) - 1 Dense	<b>0,17</b>	<b>0,42</b>	0,93	2,22	1,49	0,81
5	1 Conv1d (4F, 4K) - 1 Dense	0,29	0,54	0,96	1,10	1,05	0,55
6	1 Conv1d (4F, 8K) - 1 Dense	3,16	1,78	0,58	<b>0,08</b>	<b>0,28</b>	<b>0,99</b>
7	1 Conv1d (2F, 2K) - 1 LSTM	12,03	3,47	-1,27	31,61	5,62	-0,10
8	1 Conv1d (8F, 64K) - 1 LSTM	4,94	2,22	0,98	1,78	1,33	0,08
9	1 Conv1d (8F, 64K) - 1 LSTM - 1 Dense	3,35	1,83	0,90	3,14	1,77	0,36
10	1 Conv1d (8F, 64K) - 2 LSTM - 1 Dense	2,72	1,64	<b>0,99</b>	1,64	1,28	0,01
11	1 Conv1d (8F, 64K) - 4 LSTM - 1 Dense	22,73	4,77	0,54	1,82	1,35	0,14

Com os resultados apresentados, observa-se que a regressão com redes neurais são melhores que a regressão linear simples. A grande vantagem está na capacidade da rede de extrair características e aprender filtros que permitem trabalhar com pouco pré-processamento nos dados e manter o dado na forma em que foi coletado. Conforme o conjunto de dados cresce, a sua complexidade também aumenta. Por exemplo, os dados deste trabalho sendo adquiridos com velocidade variável, a rede neural mantém essas características de extração e aprendizado. Mas, a desvantagem é o aumento da complexidade do modelo pois será necessário uma rede neural mais robusta e o aumento de recursos computacionais para treiná-la.

A Figura 18 e a Figura 18 mostram o gráfico da regressão formada pelas redes neurais da linha 6 da Tabela 3 e da linha 10 da Tabela 4. As regressões não são lineares, portanto, para representar graficamente foi necessário plotar no eixo das ordenadas o valor real do IRI com o valor do IRI estimado pelos modelos no eixo das abscissas. A linha vermelha no gráfico é a reta de tendência. Os gráficos mostram que a regressão feita por rede neural são mais precisas que a regressão linear da Figura 16 e da Figura 17.

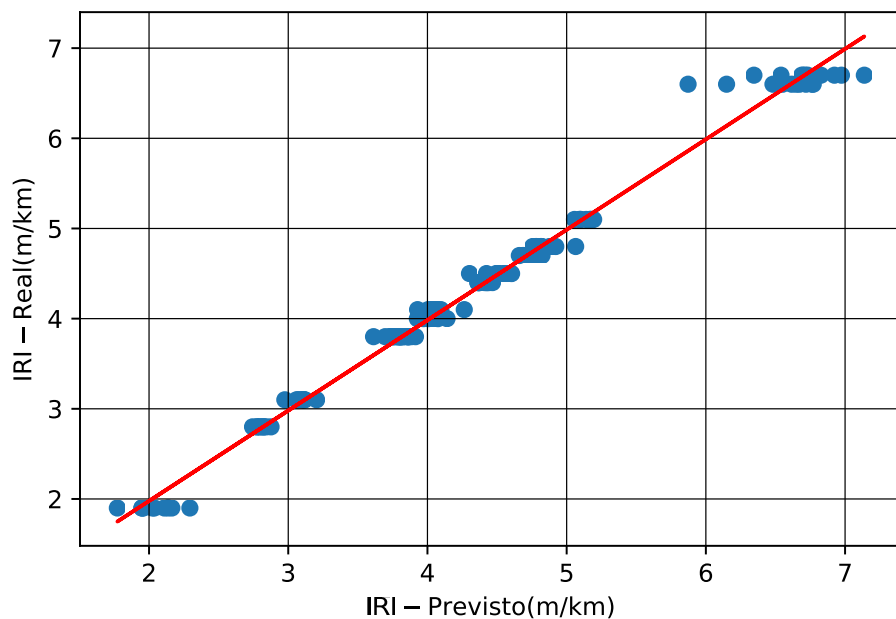


Figura 18 – Modelo da linha 6 da Tabela 3 com  $R^2$  de 0,97.



Fonte: Autor (2021).

Figura 19 – Modelo da linha 10 da Tabela 4 com  $R^2$  de 0,99.



Fonte: Autor (2021).

Aproveitando a característica sequencial dos dados foram desenvolvidas redes com camada *LSTM*. As RNAs das linhas 7 e 8 apresentam uma camada com neurônio *LSTM* na sua camada de saída. A rede 7 apresentou resultados ruins para os dados da Tabela 3 e da Tabela 4. O coeficiente de determinação está negativo porque a biblioteca *scikit-learn* foi implementado de forma que admita valor negativo para o  $R^2$ , mas matematicamente este valor não é admissível

e, portanto, considerado zero, isto implica que o modelo da linha 7 de ambas as tabelas não conseguiu explicar a variação do conjunto de dados. A rede *LSTM* não conseguiu adequar-se ao conjunto de dados mas aumentando os parâmetros da camada convolucional e conseqüentemente aumentando o seu poder de extração de característica a rede *LSTM* conseguiu explicar a variação no conjunto de dados, como mostra a linha 8 das duas tabelas. Este aumento de complexidade se dá devido a característica sequencial do conjunto de dados e para extrair essa característica fez-se necessário o uso de uma camada convolucional mais robusta.

Esta característica sequencial é altamente improvável que seja explicada pela regressão linear simples. Com as análises realizadas pode-se concluir que o uso de redes neurais é mais adequado para realizar estudos com este tipo de dados. As redes neurais permitem utilizar os dados em sua forma bruta, ou seja, sem pré-processamento e, também, aproveitar as características mais complexas e únicas do conjunto de dados. Conseqüentemente é gerado um aumento no leque de possibilidades de realizar a coleta dos dados podendo simplificar o processo e deixá-lo mais acessível financeiramente.

Nos últimos três modelos da Tabela 3 e Tabela 4 foi alterada a camada de saída adicionando a camada *Dense* e transformando a camada *LSTM* em uma camada oculta. Nestes casos variou-se apenas a quantidade de camadas *LSTM*, nas linhas 9 e 10, este aumento de complexidade gerou bons resultados de  $R^2$ , sendo que o modelo 10 obteve um dos melhores resultados deste trabalho em ambos os conjuntos de dados em termos de  $R^2$ . Mas ao aumentar para 4 camadas *LSTM* a rede não conseguiu explicar a variação do conjunto de dados. Há uma infinidade de possibilidades na seleção destes parâmetros, quanto mais camadas e neurônios mais recursos computacionais é exigido e maior o tempo de execução. Se aumentar a complexidade da camada convolucional da linha 11 provavelmente será obtido um bom resultado em termos de  $R^2$ .

### 4.3 DADOS INTERPOLADOS

Para os dados interpolados apenas o modelo 6 apresentou um resultado razoável para os dois conjuntos de dados. Mas para os dados da Tabela 3 o resultado foi inferior a regressão linear simples com filtro. A interpolação aplicada neste trabalho foi a interpolação polinomial. Para estes dados este tipo de interpolação não deve ser a mais adequada, portanto, outra abordagem de interpolação é recomendada para fazer este estudo.

## 5 CONCLUSÃO

O objetivo deste trabalho foi a construção de um modelo inteligente para estimar o índice de irregularidade internacional. Este objetivo foi cumprido para os conjuntos de dados utilizados neste trabalho e é perfeitamente plausível um modelo inteligente que estime os valores de IRI com dados de aceleração obtidos em diferentes velocidades por meio de smartphones, lembrando que a velocidade dos dados deste trabalho foi aproximadamente constante.

Foi obtido mais de um modelo que apresentou resultados relevantes, mas os modelos diferenciaram-se em sua complexidade. Os modelos de menor complexidade merecem destaque pois é necessário menos recurso computacional e intelectual para treiná-los e compreendê-los.

Isto permite compreender melhor o comportamento destes dados, apesar de ser uma sequência, uma hipótese inicial seria a necessidade de modelos complexos para estimar o IRI, mas esta hipótese não se provou verdadeira. Para trabalhar com esses dados é preciso definir bem a camada que aprenderá os filtros e extrair as principais características, que é a camada convolucional.

Pode-se concluir que para o aumento da complexidade dos dados, por exemplo velocidade variável durante a aquisição dos dados, o filtro será a parte mais importante na complexidade do modelo e conseqüentemente, nos recursos utilizados.

### 5.1 TRABALHOS FUTUROS

A continuação dos estudos deste trabalho será desenvolver novos modelos utilizando uma nova abordagem para interpolação e a utilização de uma amostra de dados maior, sendo os dados em novos trechos, trechos maiores e com velocidades variáveis. Também estão disponíveis dados com aquisições realizadas com outros smartphones e com outros veículos, o sensor dos smartphones e as características do motor, e da suspensão afetam a qualidade do sinal obtido e, conseqüentemente, as suas características.

## REFERÊNCIAS

ACADEMY, Data Science. **Deep Learning Book**. 2019. Disponível em: <http://www.deeplearningbook.com.br>. Acesso em: 21 de novembro de 2019.

BARELLA, Rodrigo Maluf. **Contribuição para a avaliação da irregularidade longitudinal de pavimentos com perfilômetros inerciais**. 2007. 362 p. Tese (Doutorado) — Universidade de São Paulo, São Paulo, 2007.

BISCONSINI, Danilo Rinaldi. **Avaliação da Irregularidade Longitudinal dos Pavimentos com Dados Coletados por Smartphones**. 2016. 176 p. Dissertação (Mestrado) — Universidade de São Paulo, São Carlos, 2016.

BISCONSINI, Danilo R.; LOUREIRO, Jakeline; JUNIOR, José Leomar F. Análise de fatores relacionados ao uso de smartphones para a avaliação da irregularidade longitudinal de pavimentos. **33º congresso de pesquisa e ensino em transporte da ANPET**, v. 2019, p. 1351–1362, 2019.

CAMPOS, Alexandre Conti Ribeiro de. **Métodos de previsão de desempenho de irregularidade longitudinal para pavimentos asfálticos: aplicação e proposição de critérios de ajuste**. 2004. 252 p. Dissertação (Mestrado) — Universidade Estadual de Campinas, Faculdade de Engenharia Civil, Arquitetura e Urbanismo, Campinas, SP, 2004. Disponível em: <http://www.repositorio.unicamp.br/handle/REPOSIP/258208>. Acesso em: 11 de outubro de 2019.

CNT. Transporte rodoviário: impactos da qualidade do asfalto sobre o transporte rodoviário. 2019. Disponível em: <https://cnt.org.br/impactos-qualidade-asfalto-transporte-rodoviario>. Acesso em: 5 de dezembro de 2021.

CNT, SEST SENAT. Pesquisa cnt de rodovias 2019: Relatório gerencial. 2019. Disponível em: <https://pesquisarodovias.cnt.org.br/downloads/ultimaversao/gerencial.pdf>. Acesso em: 5 de dezembro de 2021.

DEVORE, Jay L. **Probability and Statistics for Engineering and the Sciences**. 9. ed. California Polytechnic State University, San Luis Obispo: Cengage Learning, 2016.

DNIT. **Manual de Gerência de Pavimentos**. Rio de Janeiro, 2011.

FRANCO, Neide Maria Bertoldi. **Cálculo Numérico**. [S.l.]: Prentice Hall Brasil, 2006.

GILLESPIE, Thomas D. Everything you always wanted to know about the iri, but were afraid to ask! **Road Profile Users Group Meeting**, September 1992.

HAYKIN, Simon. **Redes neurais: princípios e prática**. 2. ed. [S.l.]: Bookman, 2001.

HOSSAIN, M. I.; GOPISETTI, L. S. P.; MIAH, M. S. International roughness index prediction of flexible pavements using neural networks. **Journal of Transportation Engineering, Part B: Pavements**, ASCE, 2018.

KOHAVI, Ron. A study of cross-validation and bootstrap for accuracy estimation and model selection. **International Joint Conference on Artificial Intelligence (IJCAI)**, 1995.

LECUN, Yann; BENGIO, Yoshua. Convolutional networks for images, speech, and time series. **The handbook of brain theory and neural networks**, v. 3361, n. 10, p. 1995, 1995.

MARSILI, Italo. **Terapia de Guerrilha**. [S.l.]: Auster, 2019.

MITCHELL, Tom M. **Machine Learning**. [S.l.]: McGraw-Hill Science/Engineering/Math, 1997.

NEVES, Samuel Antônio das. **Técnicas de Aprendizado de Máquina Aplicadas a Classificação da Qualidade de Pavimentos Asfálticos utilizando Smartphones**. 2018. 48 p. Monografia (Trabalho de conclusão de curso) — Universidade Federal de Ouro Preto, João Monlevade, MG, 2018.

RASCHKA, Sebastian; MIRJALILI, Vahid. **Python Machine Learning**. 2. ed. [S.l.]: Packt Publishing Ltd., 2017.

RASCHKA, Sebastian; MIRJALILI, Vahid. **Python Machine Learning**. 3. ed. [S.l.]: Packt Publishing Ltd., 2019.

RIEDMILLER, Martin; BRAUN, Heinrich. A direct adaptive method for faster backpropagation learning: The rprop algorithm. v. 1993, p. 586–591, 1993.

RUSSELL, Stuart; NORVING, Peter. **Artificial Intelligence: A modern approach**. 3. ed. [S.l.]: Prentice Hall, 2009.

SAYERS, Michael W.; GILLESPIE, Thomas D.; QUEIROZ, Cesar A. V. The international road roughness experiment: A basis for establishing a standard scale for road roughness measurements. **65th Annual Meeting of the Transportation Research Board**, 1986.

SAYERS, Michael W.; GILLESPIE, Thomas D.; QUEIROZ, Cesar A. V. The international road roughness experiment: Establishing correlation and a calibration standard for measurements. The World Bank, 1986.

SAYERS, Michael W.; KARAMIHAS, Steven M. **The Little Book of Profiling**. Ann Arbor: University of Michigan, 1998.

ZHANG, Zhiming; SUN, Chao; BRIDGELALL, Raj; SUN, Mingxuan. Application of a machine learning method to evaluate road roughness from connected vehicles. **Journal of Transportation Engineering, Part B: Pavements**, ASCE, 2018.