

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GEOGRAFIA
ENGENHARIA DE AGRIMENSURA E CARTOGRÁFICA

APLICAÇÃO DE REDES NEURAS ARTIFICIAIS NA PREDIÇÃO DA
ACIDENTALIDADE VIÁRIA NA RODOVIA BR-365

Willian dos Santos Cruz

Monte Carmelo

2023

WILLIAN DOS SANTOS CRUZ

**APLICAÇÃO DE REDES NEURAS ARTIFICIAIS NA PREDIÇÃO DA
ACIDENTALIDADE VIÁRIA NA RODOVIA BR-365**

Trabalho de Conclusão de Curso apresentado ao Instituto de Geografia, da Universidade Federal de Uberlândia, Campus Monte Carmelo, como parte dos requisitos para obtenção do Título de bacharel em Engenharia de Agrimensura e Cartográfica.

Orientadora: Prof.^a Dra. Luciany Oliveira Seabra
Coorientador: Me. Vinícius Ferreira Sales

Monte Carmelo

2023

WILLIAN DOS SANTOS CRUZ

**APLICAÇÃO DE REDES NEURAIIS ARTIFICIAIS NA PREDIÇÃO DA
ACIDENTALIDADE VIÁRIA NA RODOVIA BR-365**

Trabalho de Conclusão de Curso apresentado ao Instituto de Geografia, da Universidade Federal de Uberlândia, Campus Monte Carmelo, como parte dos requisitos para obtenção do Título de bacharel em Engenharia de Agrimensura e Cartográfica.

Orientadora: Prof.^a Dra. Luciany Oliveira Seabra
Coorientador: Me. Vinícius Ferreira Sales

Aprovado em: 21 / 09 / 2023

Banca Examinadora:

Prof.^a Dra. Luciany Oliveira Seabra (Orientadora)

Universidade Federal de Uberlândia – Campus Monte Carmelo

Me. Vinícius Ferreira Sales (Coorientador)

Universidade Federal de Uberlândia – Campus Monte Carmelo

Prof.^o Dr. Wagner Carrupt Machado (Membro convidado)

Universidade Federal de Uberlândia – Campus Monte Carmelo

À Deise, dedico este trabalho.

AGRADECIMENTOS

À Deus, por ter me permitido chegar até aqui com saúde, e por ter me fornecido luz nos momentos mais difíceis, em que, apesar de todos os empecilhos e dificuldades enfrentadas, sempre me abençoou e orientou, por isso, à Ele toda honra e toda glória.

À minha família, que sempre me apoiou e incentivou ao longo de todas as etapas da minha vida, estando sempre presente comigo mesmo quando estávamos fisicamente distantes. Além disso, eterna gratidão a minha mãe Ivone Rosa dos Santos e ao meu pai Joilson dos Santos Cruz que sempre fizeram tudo por mim, sem a motivação e o sacrifício deles nada disso seria possível, à eles também dedico este trabalho. Agradeço também aos meus irmãos em especial meu irmão Joanderson dos Santos Cruz, que sempre foi meu parceiro durante esta e muitas outras etapas da minha vida, servindo de grande inspiração a cada passo e decisão tomada.

À minha excelente orientadora Prof.^a Luciany Oliveira Seabra, que sempre foi muito atenciosa e compreensiva na elaboração deste estudo, me auxiliando em cada fase e me orientando nas melhores escolhas possíveis, à ela minha gratidão por tudo.

Ao meu coorientador Vinícius Ferreira Sales, que sempre foi extremamente prestativo e disposto a me guiar e ensinar tudo a respeito do tema, servindo de grande inspiração e estímulo para a conclusão do trabalho, à ele minha gratidão por todos os ensinamentos.

À todos os professores da UFU pelos ensinamentos, em especial os professores Wagner Carrupt Machado e Diego de Oliveira Martins pela contribuição e conselhos dados para a melhoria deste estudo.

À Mariana Alves Oliveira, por ter sido a principal inspiração deste trabalho e moldado a forma como este trabalho foi executado.

À todos os amigos formados durante os 5 anos de estudos na UFU, por todos os momentos marcantes, trabalhos, festas e comemorações.

“Se a educação sozinha não transforma a sociedade, sem ela tampouco a sociedade muda”.

(Paulo Freire)

RESUMO

A segurança viária brasileira possui atualmente diversos problemas, sendo a acidentalidade viária responsável por ser uma das principais causas de mortes no Brasil. Em Minas Gerais, a BR-365, importante via do triângulo mineiro, constitui numa das rodovias federais que mais desponta no ranking de vítimas fatais neste estado. Para a atenuação deste cenário, técnicas de aprendizado de máquina como as Redes Neurais Artificiais (RNA's) que se destacam no ramo da predição, tem sido utilizadas para modelar abordagens que busquem mitigar o nível de ocorrências viárias. Nesse sentido, o objetivo desse estudo consistiu em desenvolver um modelo de rede neural artificial para predição da frequência de acidentes da BR-365, com trecho do município de Patos de Minas ao município de Ituiutaba, no estado de Minas Gerais. Para construção da metodologia do trabalho e visando selecionar a melhor arquitetura de RNA, a metodologia consistiu em analisar a área de estudo a partir de três métodos diferentes: por margens de erros, variando o tamanho dessas margens de 5 a 30 km, em seguida, por trechos fixados segmentando a área de estudo em diferentes tamanhos de extensões variando de 5 a 30 km, e, por trechos situados entre seis municípios escolhidos contidos dentro da área estudada. A partir dos resultados obtidos, é possível compreender que a assertividade das RNA's é diretamente proporcional ao aumento do trecho analisado, sendo o último método aplicado uma forma de análise mais aprimorada, demonstrando que o trecho com os segmentos mais críticos estão localizados entre os municípios de Uberlândia e Monte Alegre de Minas.

Palavras-chave: Redes Neurais Artificiais, MLP, modelos de predição, acidentalidade viária, rodovias federais.

LISTA DE GRÁFICOS

Gráfico 1 – Tipos de sinistros de trânsito nas rodovias federais	12
Gráfico 2 – Causas associadas aos sinistros de trânsito nas rodovias federais.....	12
Gráfico 3 – Treinamento de uma rede neural artificial.....	29
Gráfico 4 – Comparação entre os métodos.....	54

LISTA DE FIGURAS

Figura 1 – Estrutura de um neurônio humano	17
Figura 2 – Estrutura de um neurônio artificial	18
Figura 3 – Rede neural de camada única (Rede de <i>perceptrons</i>)	22
Figura 4 – Rede neural <i>perceptron</i> de múltiplas camadas (MLP).....	24
Figura 5 – Pontos de mínimo locais e menor erro global associados à função erro.....	26
Figura 6 – Mapa de localização da área de estudo	32
Figura 7 – Fluxograma dos processos metodológicos.....	34
Figura 8 – Divisão do conjunto de dados	40
Figura 9 – Validação cruzada com três subconjuntos de treinamento e três subconjuntos de validação	41
Figura 10 – Configuração da arquitetura da rede neural 9	50
Figura 11 – Segmento crítico do trecho entre Patrocínio – MG e Monte Carmelo - MG.....	57
Figura 12 – Segmento crítico do trecho Monte Carmelo – MG e Uberlândia – MG.....	58
Figura 13 – Segmentos críticos do trecho entre Uberlândia – MG e Monte Alegre de Minas – MG.....	59
Figura 14 – Segmentos críticos do trecho entre Monte Alegre de Minas – MG e Ituiutaba – MG.....	60
Figura 15 – Segmentos críticos da área de estudo da BR-365	61

LISTA DE TABELAS

Tabela 1 – Definição dos termos utilizados na base de dados da PRF.....	10
Tabela 2 – Tipos de sinistros de trânsito mais ocorridos em Minas Gerais em 2022	13
Tabela 3 – Principais funções de ativação	20
Tabela 4 – Atribuição de pesos para cada parâmetro nominal (0 a 1).....	36
Tabela 5 – Atribuição de vetores binários para o parâmetro “Severidade”	37
Tabela 6 – Parâmetros utilizados na arquitetura das redes	39
Tabela 7 – Classificação de Hopkins (2000)	43
Tabela 8 – Subdivisão da área de estudo em trechos entre municípios.....	44
Tabela 9 – Topologias das RNA’s.....	45
Tabela 10 – Resultados do método “Margens de erro” da validação	45
Tabela 11 – Cinco melhores redes pelo método “Margens de erro” da validação	46
Tabela 12 – Melhor topologia pelo método “Margens de erro” da validação.....	47
Tabela 13 – Resultados do método “Trechos fixados” da validação.....	47
Tabela 14 – Cinco melhores redes pelo método “Trechos fixados” da validação.....	48
Tabela 15 – Melhor topologia pelo método “Trechos fixados”	49
Tabela 16 – Resumo dos parâmetros empregados.....	51
Tabela 17 – Resultado médio da arquitetura 9	51
Tabela 18 – Resultados médio da rede neural 9 com o subconjunto de teste.....	52
Tabela 19 – Resultados da rede neural 9 na análise de trechos entre municípios	53
Tabela 20 – Número de sinistros de trânsito ocorridos ao longo de 2017 a 2019 do município de Patos de Minas - MG a Ituiutaba – MG.....	56
Tabela 21 – Número de sinistros de trânsito ocorridos em 2022 entre os municípios de Patos de Minas - MG e Ituiutaba – MG	63
Tabela 22 – Normalização da faixa etária	74
Tabela 23 – Normalização dos quilômetros	76
Tabela 24 – Valores Alvos e Valores Preditos pela rede neural 9.....	78

LISTA DE ABREVIATURAS E SIGLAS

(ABNT)	Associação Brasileira de Normas Técnicas
(ABRACD)	Associação Brasileira de Ciências de Dados
(ABRAMET)	Associação Brasileira de Medicina de Tráfego
(ATT)	Acidentes de Transportes Terrestres
(CNT)	Confederação Nacional de Transportes
(CTB)	Código de Trânsito Brasileiro
(DNIT)	Departamento Nacional de Infraestrutura de Transporte
(EAM)	Erro Absoluto Médio
(IA)	Inteligência Artificial
(IBGE)	Instituto Brasileiro de Geografia e Estatística
(IBM)	<i>International Business Machines Corporation</i>
(IPEA)	Instituto de Pesquisa Econômica Aplicada
(MESV)	Modelagem Estatística de Segurança Viária
(MLP)	Perceptron de Múltiplas Camadas (<i>Multilayer Perceptron</i>)
(MPA)	Modelos de Previsão de Acidentes
(PIB)	Produto Interno Bruto
(PRF)	Polícia Rodoviária Federal
(RBF)	Redes de Função Base Radial
(REQM)	Raiz do Erro Quadrático Médio
(RNA's)	Redes Neurais Artificiais
(SNV)	Sistema Nacional de Viação
(WHO)	Organização das Nações Unidas (<i>World Health Organization</i>)

SUMÁRIO

1	INTRODUÇÃO	3
2	JUSTIFICATIVA	5
3	OBJETIVOS	7
3.1	OBJETIVO GERAL	7
3.2	OBJETIVOS ESPECÍFICOS	7
4	REFERENCIAL TEÓRICO	8
4.1	ACIDENTALIDADE VIÁRIA.....	8
4.1.1	Caracterização dos acidentes	9
4.2	MODELAGEM DA SEGURANÇA VIÁRIA.....	13
4.3	INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA	15
4.3.1	Redes Neurais Artificiais (RNA's)	16
4.3.1.1	<i>Perceptron</i>	22
4.3.1.2	<i>Multilayer Perceptron (MLP)</i>	23
4.3.1.3	Treinamento de uma rede neural artificial	25
4.3.1.4	Métricas estatísticas para avaliação de um modelo de predição	30
5	MATERIAL E MÉTODOS	31
5.1	ÁREA DE ESTUDO	31
5.2	MATERIAL	33
5.2.1	Dados de acidentalidade viária	33
5.2.2	Dados espaciais	33
5.2.3	<i>Hardware e Software</i>	33
5.3	MÉTODOS.....	33
5.3.1	Escolha do trecho da BR-365 (Fase 1)	34
5.3.2	Dados (Fase 2)	35
5.3.2.1	Obtenção dos dados (Etapa 2.1).....	35
5.3.2.2	Definição dos parâmetros a serem empregados (Etapa 2.2)	35
5.3.3	Normalização dos dados (Fase 3)	36
5.3.4	Treinamento das RNA's (Fase 4)	38
6	RESULTADOS E DISCUSSÕES	44

7	CONCLUSÃO	64
7.1	LIMITAÇÕES.....	65
7.2	RECOMENDAÇÕES	66
	REFERÊNCIAS	68
	APÊNDICE	74

1 INTRODUÇÃO

De acordo com a Organização Mundial da Saúde (WHO, 2018), por meio da publicação do “*Global Status Report on Road Safety 2018*”, os acidentes de trânsito resultam na morte de aproximadamente 1,35 milhões de pessoas no mundo a cada ano, constituindo-se a oitava principal causa de morte em todo o globo. Ainda conforme esta organização, quando se trata de feridos não fatais, os acidentes viários deixam entre 20 a 50 milhões de pessoas lesionadas. Além disso, mais da metade de todas as mortes e lesões causadas pelo trânsito envolvem usuários mais vulneráveis das vias, como pedestres, ciclistas, motociclistas e seus passageiros.

Em relação ao Brasil, a cada ano, mais de 40.000 vidas são perdidas em função dos Acidentes de Transportes Terrestres (ATT), tornando o país um dos mais destacados no ranking mundial de países com maior número de vítimas de trânsito (Carvalho, 2016). Este cenário se intensifica ainda mais quando se trata das rodovias federais brasileiras, onde, somente em 2022, foi registrado um total de 64.447 acidentes, sendo 52.948 com vítimas (feridas ou fatais), o que totalizou uma média de 7 acidentes com vítimas a cada 10 quilômetros de rodovia em 2022 (CONFEDERAÇÃO NACIONAL DO TRANSPORTE – CNT, 2023).

No estado de Minas Gerais, esse contexto de acidentalidade viária pode ser observado na BR-365, que conforme os registros de dados da Polícia Rodoviária Federal – PRF no ano de 2022 ocorreram 577 acidentes rodoviários, onde, 474 destes acidentes resultaram em vítimas feridas ou fatais (Polícia Rodoviária Federal - PRF, 2023). A BR-365 juntamente com uma parte da BR-364 constitui uma das mais importantes rotas para o escoamento de grãos da região Centro-Oeste até o porto de Santos-SP, além disso, esse trecho é essencial para o abastecimento de Minas Gerais e da região sul de Goiás, principalmente de produtos da indústria, materiais de construção e alimentos (ECOVIAS DO CERRADO, 2022).

Devido ao fato de envolverem maior energia, e, portanto, maior velocidade, os acidentes em rodovias federais são os que apresentam maior letalidade do que os acidentes urbanos, correspondendo a mais de 70% do total dos custos dos acidentes de trânsito no Brasil, sendo este valor total estimado em média a R\$ 50 bilhões a cada ano (Carvalho, 2016). Carvalho (2016) ainda afirma que, além do alto custo à sociedade brasileira, a maior fatia desse montante é destinada aos componentes de custos associados às vítimas, principalmente, àquelas que sofrem graves conseqüências, ocasionando perda de produção ou invalidez dessas vítimas, o que gera impactos diretos sobre a previdência social e também ao governo federal,

demonstrando cada vez mais que a acidentalidade viária é uma grave questão de saúde pública.

Lord e Mannering (2010), afirmam que dado todo o contexto em relação à acidentalidade viária, pesquisadores têm estruturado suas pesquisas e abordagens analíticas focadas em compreender e estudar os possíveis fatores que influenciam no número de acidentes ocorridos em um determinado espaço geográfico (geralmente um segmento de estrada ou interseção), sobre um período de tempo específico (semana, mês, ano).

Segundo o Instituto de Pesquisa Econômica Aplicada – IPEA (2008), a partir de experiências acumuladas ao longo das décadas, os ATT quase sempre resultam de causas multifatoriais, sendo 3 os principais fatores: viário que refere-se às características físicas das vias ou condições inadequadas de segurança como falta de passarelas, curvas com superelevação negativa ou insuficiente, fluxos veiculares de sentido duplo; humano que envolve o comportamento inadequado por parte de condutores dos veículos, que trafegam com excesso de velocidade, excesso de álcool, drogas e tensão nervosa; e veicular que refere-se às condições dos veículos e falta de manutenção como pneus carecas e problemas mecânicos.

Nodari (2003) afirma que a forma mais eficiente de tratar o problema da segurança viária, é através de um programa abrangente de gerenciamento que consiste em dois métodos: reativo, onde se utiliza de informações sobre o histórico de acidentes a fim de identificar os locais com os pontos mais críticos e, a partir da identificação desses locais, adotar possíveis medidas corretivas; e o pró-ativo, que se baseia no conceito de promover melhorias no sistema viário visando aprimorar as condições de segurança para evitar potenciais acidentes.

Um dos principais obstáculos iniciais para o estudo de predição da frequência de acidentes é o critério escolhido para determinar a segmentação da rodovia, ou seja, da área de estudo (Silva, 2019). Na literatura existem diversas abordagens diferentes de segmentação, cujos trabalhos focam na relação entre os acidentes e a geometria da estrada, em que foi concluído que o modelo que contém geometria e *design* com variáveis consistentes foram mais confiáveis do que outras formas de segmentação (Cafiso; Agostino; Persuad, 2018).

Durante anos, a modelagem da segurança viária consistia tradicionalmente em metodologias estatísticas, contudo, os limites deste tipo de abordagem como a suposição sobre a distribuição dos dados e os pressupostos estatísticos existentes entre as variáveis dependentes e explicativas, abriram margem para a necessidade de se buscar por novas metodologias (Mussone et al., 1999; Chang, 2005; Li et al., 2012 citados por Silva, 2019).

Nesse contexto que as Redes Neurais Artificiais (RNA's), também conhecidas como Modelos Conexionistas se destacam e tornam-se cada vez mais atrativas, por se tratarem de modelos de aprendizado de máquinas que buscam se assemelhar às redes neurais biológicas, objetivando solucionar problemas em que hoje os seres humanos possuem bom desempenho, podendo ser citado o reconhecimento de imagens, como, por exemplo, o rosto de pessoas (Martineli, 1999).

Martineli (1999) ainda destaca que o interesse crescente em RNA's ocorre principalmente devido ao seu bom desempenho em reconhecimento de padrões, sua capacidade de adaptação, e , ao contrário dos modelos estatísticos, da ausência da necessidade de uma suposição inicial de uma hipótese, dispensando as premissas estatísticas para se basear em suposto modelo de comportamento. Dessa forma, a suposição do modelo fica sobre responsabilidade apenas das conexões e dos neurônios da rede neural artificial.

Dito isso, a hipótese deste trabalho é avaliar se o emprego de técnica de Aprendizado de Máquina tem uso potencial na modelagem da segurança viária.

2 JUSTIFICATIVA

O trânsito constitui uma das principais causas de morte não natural no Brasil, país que é apontado como um dos mais violentos do mundo, o trânsito brasileiro apresenta um quadro permanente de epidemia com mortes diárias ocasionadas pelo tráfego rodoviário (Associação Brasileira de Medicina de Tráfego - ABRAMET, 2021).

De acordo com o Ipea (2015), segundo dados do Departamento de Informática do Sistema Único de Saúde (Datasus) do Ministério da Saúde (MS), os acidentes de transportes terrestres ceifam a vida de aproximadamente 43 mil pessoas por ano no país, onde, 20% dessas mortes ocorrem em rodovias federais, correspondendo só no ano de 2014 a 8.227 mortes decorrentes de acidentes nessas vias.

Quando se fala de acidentalidade rodoviária, além dos grandes traumas e sofrimentos físicos e psicológicos causados tanto às vítimas como também aos familiares e amigos, bem como outros efeitos a longo prazo que não são passíveis de se mensurar, os acidentes de trânsito também são consequências graves para o governo federal, visto que quando se trata principalmente de acidentes com vítimas fatais ou graves, essas tragédias tornam-se muito mais onerosas não só para os cofres públicos como também para a sociedade de forma geral (IPEA, 2015). De acordo com o instituto, cada acidente, em média, custou cerca de R\$

72.705,31 à sociedade brasileira, no qual, quando se trata de uma vítima fatal o custo médio sobe para R\$ 646.762,94.

Apenas em 2022, estima-se que os custos totais referentes apenas aos acidentes ocorridos nas rodovias federais brasileiras geraram um valor de R\$ 12,92 bilhões, onde, 58% desta quantia foram destinadas aos acidentes com vítimas feridas e 38% aos acidentes com vítimas fatais (CNT, 2023).

Campos e Ferreira (2013) constata que mais de 90% das mortes causadas por acidentes de trânsito no mundo ocorrem em países subdesenvolvidos e ainda em desenvolvimento, no entanto, estes por sua vez possuem apenas 48% da frota de veículos mundiais. Além disso, segundo os autores, é a partir das comparações entre indicadores socioeconômicos como o Produto Interno Bruto (PIB) e a taxa de motorização de um determinado país, que é possível analisar a relação entre o nível de segurança viária e o grau de desenvolvimento de uma sociedade, onde, no Brasil, devido o seu desenvolvimento socioeconômico e as políticas econômicas de incentivo à obtenção de transportes pessoais adotadas ocasionou num incremento em larga escala no aumento da frota de veículos no país.

Posto isso, de acordo com dados da PRF fornecidos pelo Ipea (2015), o estado de Minas Gerais no ano de 2014 foi o que apresentou mais acidentes e mortes nas rodovias federais, juntamente com Paraná, Espírito Santo, Santa Catarina e São Paulo. Segundo o instituto, isso ocorre por se tratarem de estados localizados no sul e no sudeste do país, regiões onde se localizam os maiores índices de PIB e por concentrarem a maior parte de riquezas do Brasil, o que ocasiona em uma maior quantidade de viagens e também de automóveis nas estradas, de modo a proporcionar um maior volume de tráfego e consequentemente de acidentes.

Segundo aponta o relatório da CNT (2023) a respeito de Minas Gerais, o número de óbitos decorrentes dos sinistros de trânsito na BR-365 foi de 69 pessoas, tornando essa rodovia federal a quinta no ranking de vítimas fatais. Além disso, a partir de informações da PRF (2023), o trecho pertencente entre os municípios de Patos de Minas e Ituiutaba destaca-se pelo alto número de ocorrências viárias, configurando no segmento que mais acarreta em vítimas de toda a BR-365, contabilizando apenas entre os anos de 2017 a 2019 um total de 1260 ocorrências, em que, destas, 982 ocasionaram em vítimas feridas ou fatais.

Os acidentes ocorrem na medida em que inevitavelmente os seres humanos cometem falhas. No planejamento de transportes, utilizam-se ferramentas e técnicas que permitam aos projetistas de infraestrutura viária prover um sistema de transporte que minimize as

consequências da participação do erro humano (Nodari, 2003). Para tanto, é essencial a investigação da ocorrência dos acidentes de trânsito, onde a modelagem de segurança viária tem sido realizada com o emprego de RNA's. Verifica-se na literatura sobre o tema, que este é um método promissor no ramo da análise de padrões, de classificação e predição de acidentes viários, como mostram os estudos de Lord e Mannering (2010), Zeng et al.(2016), Sameen e Pradhan (2017), Amorim (2019), Silva (2019), Chuerubim et al.(2020) e Oliveira (2021).

3 OBJETIVOS

Neste tópico são abordados os objetivos gerais e específicos deste trabalho.

3.1 OBJETIVO GERAL

Desenvolver um modelo de rede neural artificial para predição da frequência de acidentes da BR-365, com trecho do município de Patos de Minas ao município de Ituiutaba, no estado de Minas Gerais.

3.2 OBJETIVOS ESPECÍFICOS

- Identificar quais são os trechos com os segmentos mais críticos da BR-365 contidos na área de estudo analisada;
- Testar diferentes arquiteturas de redes neurais artificiais visando encontrar a de melhor desempenho e representatividade dos dados;
- Avaliar os resultados da arquitetura de melhor desempenho quanto às influências e efeitos das variáveis propostas na ocorrência da acidentalidade viária;
- Analisar a influência da extensão do segmento nos resultados obtidos;

4 REFERENCIAL TEÓRICO

Neste tópico é apresentada uma contextualização a respeito da acidentalidade viária no Brasil e uma caracterização dos acidentes de transportes terrestres. Além disso, são abordados temas como modelagem de segurança viária, inteligência artificial, aprendizado de máquina, e, as Redes Neurais Artificiais ou RNA's.

4.1 ACIDENTALIDADE VIÁRIA

Retallack e Ostendorf (2019) afirmam que existem diversos precedentes para a ocorrência de acidentes de trânsito, entre alguns dos principais que podem ser citados são: as condições meteorológicas (clima e visibilidade), características das vias, o próprio *design* dos veículos, erros humanos e além de tudo, a quantidade de veículos circulando nas estradas. Em alguns casos, a acidentalidade também pode ser causada pela combinação de dois ou mais fatores, onde, a probabilidade de acidentes viários aumenta com o volume de veículos nas rodovias, principalmente quando alinhado a determinadas características das estradas (Cadar et al., 2017; Martin et al., 2002; Wang et al., 2013 citados por Saladié, Bustamante, Gutiérrez, 2020). Quando se trata da caracterização de acidentes em rodovias de veículos mais pesados como tratores, por exemplo, destacam-se outros possíveis agentes que podem ser a causa das tragédias, como a velocidade incompatível da máquina com a via, a falta de amortecimentos nos tratores e as suas configurações técnicas que às vezes podem não ser proporcionais aos outros veículos presentes nas estradas (Gkritza et al., 2010 apud Macedo et al., 2015).

Para Andrade e Antunes (2019), embora não seja de fato possível de se evitar todas as consequências em sua plenitude e em todas as situações, os ATT são causados também por falhas humanas, dito isso, eles são passíveis em alguns casos de serem prevenidos considerando determinadas legislações e políticas de comportamentos no trânsito. Os autores afirmam que o CTB instituído pela Lei nº 9.503 de 23 de setembro de 1997 aponta aspectos ligados aos cinco principais riscos comportamentais que influenciam na ocorrência de acidentes, sendo eles: a velocidade, a falta de uso do componentes básicos de proteção como capacete e cinto de segurança, alcoolemia e a falta de uso de dispositivos de retenção para o transporte de crianças. Para a OMS, há a necessidade também de incorporar citações em relação ao uso celulares e de psicotrópicos enquanto o motorista estiver na direção, no

entanto, estes influenciadores já são contemplados pelo CTB, apesar de serem elementos de difícil fiscalização (Andrade; Antunes, 2019).

4.1.1 Caracterização dos acidentes

O CTB dispõe de atribuições tanto dos agentes como também dos órgãos de trânsito que regem a funcionalidade do tráfego terrestre, esse código serve para tornar o trânsito um ambiente mais seguro para veículos, condutores e pedestres, assim como também estabelece todas as infrações e suas gravidades, incluindo as penalidades e as normas de conduta para se obter uma mobilidade organizada (DOK Despachante, 2022). No entanto, quanto à conceituação dos tipos de acidentes o CTB não apresenta e nem determina os termos associados à acidentalidade viária, papel esse que é desempenhado pela NBR 10697, que fornece os conceitos ligados aos acidentes de trânsito e suas terminologias.

Publicado pela Associação Brasileira de Normas Técnicas (ABNT) e recentemente revisada em 2020, a norma NBR 10697/2020 redefiniu os termos técnicos utilizados na preparação e execução de pesquisas relativas e também na elaboração de relatórios estatísticos e operacionais em relação aos acidentes de trânsito (ABRAMET, 2021).

Devido à falta de consenso e às divergências existentes na literatura, os autores possuem diversos termos e definições a respeito do que é um acidente viário, porém, além de estabelecer uma definição concisa, essa norma foi elaborada também com o intuito de substituir a expressão “acidente de trânsito” por “sinistro de trânsito”, isso porque a adoção do novo termo foi para enfatizar que os acidentes de transportes terrestres, em sua grande maioria, não são acidentais ou do acaso, mas sim causados (ABRAMET, 2021). Logo, conforme esta entidade, o termo acidente trazia a conotação de algo que era imprevisível ou incontrolável, sem nenhum motivo ou razão de causalidade, isso contrariava todo o conhecimento obtido e acumulado até então a respeito da acidentalidade viária. Dessa forma, diante deste contexto, se os ATT em sua grande maioria são provocados, eles podem, portanto, serem passíveis de prevenção.

Por se tratar de um tema recorrente e muito presente no cotidiano brasileiro e também do mundo, a literatura a respeito de acidentalidade viária possui diversas definições quanto ao que é acidente de trânsito. Para Alves e Raia Junior (2012), acidente de trânsito pode ser interpretado como um evento não intencional, que envolve um ou mais veículos, motorizados

ou não, que circulam por uma via para trânsito e que resulta em algum tipo de dano, seja ele material ou lesões em pessoas e/ou animais.

Já conforme a NBR 10697/2020 “Pesquisa de sinistros de trânsito - Terminologia”, é definido sinistro de trânsito como:

Todo evento que resulte em dano ao veículo ou à sua carga e/ou em lesões a pessoas e/ou animais, e que possa trazer dano material ou prejuízos ao trânsito, à via ou ao meio ambiente, em que pelo menos uma das partes está em movimento nas vias terrestres ou em áreas abertas ao público (ABNT, 2020).

Para Schuster (2012), os acidentes podem ser categorizados como acidentes sem vítimas, com vítimas feridas e com vítimas fatais. Segundo o autor, os acidentes com vítimas classificam-se quanto ao estado físico dos acidentados como fatais (acidentes onde há pelo menos um morto), graves (quando as vítimas possuem ferimentos que exigem tratamento prolongado), leves (onde houve apenas acidentados com lesões leves) e ilesos (quando as vítimas não sofrem nenhuma lesão ou sintoma). Além disso, os acidentes são também classificados quanto ao tipo, sendo eles: colisão traseira, colisão frontal, colisão transversal, colisão lateral, choque, atropelamento, tombamento, capotamento e engavetamento.

Todos os acidentes supracitados são caracterizados na **Tabela 1** a seguir, estando em consonância com os termos estabelecidos pela norma NBR 10697/2020.

Tabela 1 – Definição dos termos utilizados na base de dados da PRF

		Definição
Tipo de acidente	Atropelamento de Pessoa/Animal	Sinistro de trânsito em que animal(is) e/ou pessoa(s) sofre(m) o impacto de um veículo em movimento.
	Capotamento/Tombamento	Sinistro de trânsito em que o veículo gira sobre si mesmo, em qualquer sentido, ficando em algum momento com as rodas para cima, imobilizando-se em qualquer posição.
	Colisão	Sinistro de trânsito em que um veículo em movimento sofre o impacto de outro veículo ou contra qualquer objeto fixo podendo estar também em movimento. Esta colisão pode acontecer com veículos que transitam em sentidos opostos, na mesma direção, em direções que se cruzam e na frente contra traseira ou na traseira contra traseira.
	Saída de Leito	Ação ou efeito de sair da pista; passagem por onde sai, pela superfície aplainada de caminho, rua, estrada, etc. onde o veículo pode sair ou escapar.

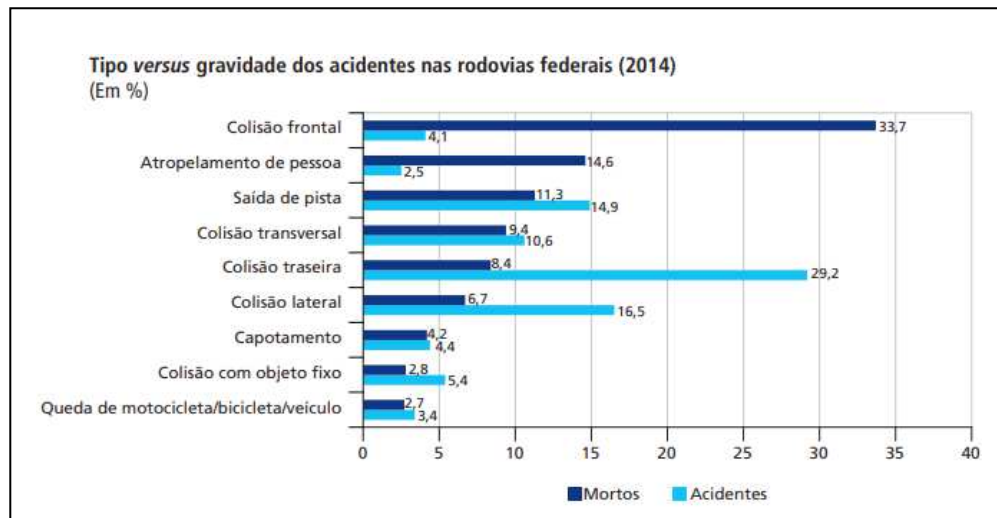
Continua...

		Conclusão
Definição		
Classificação	Sem vítimas	Aquele que não resulte em vítima, mas que traga dano material ou prejuízos ao trânsito, à via ou ao meio ambiente.
	Com vítimas feridas	Aquele que não resulte em vítima fatal.
	Com vítimas fatais	Aquele que resulte em vítima morrendo imediatamente ou em até 30 dias, como resultado do acidente ou suas implicações.
Tipo de veículo	Automóvel	Veículo automotor destinado ao transporte de passageiros, com capacidade para até oito pessoas, exclusive o condutor.
	Caminhão	Veículo automotor dotado de chassi e cabine, destinado ao transporte de carga, com peso bruto total superior a 3500 kg.
	Motocicletas	Veículo motorizado dotado de duas rodas, dirigido por condutor em posição montada, com ou sem carrinho lateral, cuja velocidade máxima de fabricação exceda 50 km/h.
	Ônibus	Veículo motorizado de transporte coletivo, com capacidade para mais de 20 passageiros, ainda que, em virtude de adaptações com vista à maior comodidade destes, transporte um número menor.

Fonte: Adaptado de Mendes (2021).

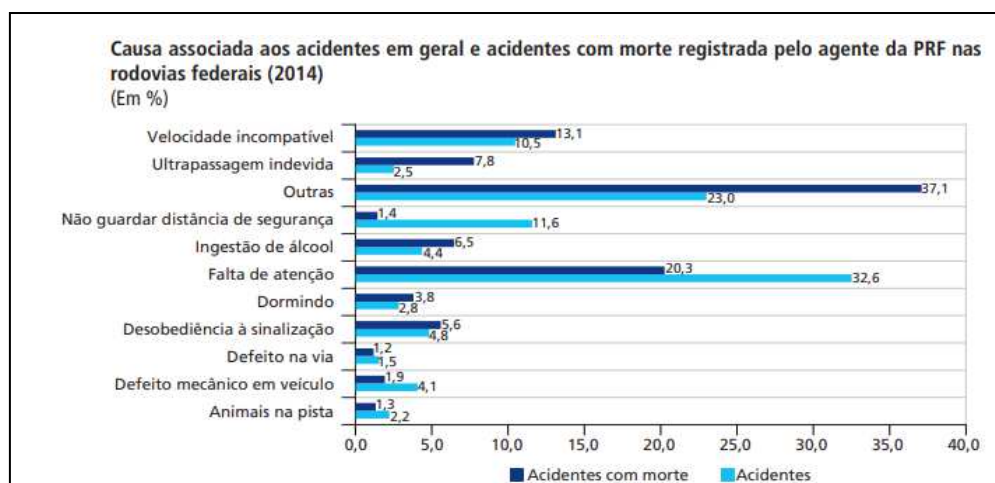
Em 2014, a colisão frontal foi responsável por 33,7% das mortes de trânsito dos sinistros totais ocorridos no Brasil naquele ano, sendo que dessas colisões, 89,71% ocorreram em pista simples, ocasionando em 93,91% dos mortos neste tipo de acidente. Os atropelamentos, (segunda maior causa de óbito no ano), foram responsáveis por 14,6% das mortes ocorridas do total de sinistros de trânsito, sendo que em relação aos atropelamentos de pessoas, 71,73% correspondem a ocorrências em trechos urbanos, que resultaram em 58,47% de mortes nestes ocorridos. Ao realizar análises mais aprofundadas, percebeu-se que as mortes consequentes dessas ocorrências se concentravam principalmente nos períodos noturnos, no qual equivaliam a 63,12% dos mortos decorrentes deste tipo de ocorrência (IPEA, 2015).

No **Gráfico 1** é possível ver toda essa conjuntura descrita acima.

Gráfico 1 – Tipos de sinistros de trânsito nas rodovias federais

Fonte: IPEA (2015).

Ainda segundo estudos do IPEA (2015), quando se analisa a principal causa dos sinistros de trânsito registrada pela PRF, a que mais se destaca é a falta de atenção, fator responsável por 32,6% dos casos, já nos casos em que envolvem vítimas fatais, esse valor corresponde a 20,3%. Outras causas existentes que englobam o fator humano e que são os principais responsáveis pela ocorrência dos acidentes são: a velocidade incompatível (13,1%), a ultrapassagem indevida (7,8%) e a ingestão de álcool (6,5%). Todo esse cenário mostra a necessidade da intensificação de campanhas educativas focadas tanto no comportamento defensivo como também no respeito às regras de trânsito por parte dos motoristas, para que, dessa forma, a acidentalidade viária no Brasil reduza de forma significativa. No **Gráfico 2** é possível visualizar as causas associadas aos acidentes.

Gráfico 2 – Causas associadas aos sinistros de trânsito nas rodovias federais

Fonte: IPEA (2015).

Quando se trata do estado de Minas Gerais, o tipo principal de acidente e também de mortes no ano de 2022, foi ocasionado pelas colisões (**Tabela 2**), seguida pela saída de pista e pelos capotamentos e tombamentos (CNT, 2023).

Tabela 2 – Tipos de sinistros de trânsito mais ocorridos em Minas Gerais em 2022

Acidentes com vítimas e mortes por Tipo de Acidente - 2022				
Tipo de acidente	Acidentes	%	Mortes	%
Colisão	3.682	51,8%	415	59,3%
Saída de Pista	1.491	21,0%	113	16,1%
Capotamento/Tombamento	1.111	15,7%	66	9,4%
Atropelamento	455	6,4%	91	13,0%
Queda de Ocupante	305	4,3%	13	1,9%
Eventos atípicos	33	0,5%	2	0,3%
Incêndio	11	0,2%	0	0,0%
Derramamento de Carga	7	0,1%	0	0,0%
Total	7.095	100,0%	700	100,0%

Fonte: CNT (2023).

Além disso, segundo o relatório do CNT (2023) ainda sobre este estado brasileiro, a rodovia federal mineira com maior número de mortes no ano de 2022 foi a BR-381, contabilizando um total 151 óbitos relatados.

4.2 MODELAGEM DA SEGURANÇA VIÁRIA

A modelagem da segurança viária é uma das principais formas de atenuação e prevenção da ocorrência de acidentes rodoviários, de modo que as análises dos resultados dos modelos gerados podem viabilizar possíveis medidas e soluções mitigadoras (Silva, 2019). Lord e Mannering (2010) apontam que a modelagem viária surgiu como consequência da falta de detalhes a respeito das informações do condutor (idade, resposta do motorista a estímulos), e dos veículos (aceleração, frenagem e outras informações relevantes dos veículos a partir, por exemplo, da implantação de caixas-pretas), que, uma vez disponíveis, ajudariam na identificação das causas e dos efeitos envolvidos na relação entre estes dois elementos (condutor e veículo) com a probabilidade de ocorrer um acidente.

Conforme Wang, Quddus e Ison (2011), “Modelos de Previsão de Acidentes (MPA) são largamente utilizados para estimar a frequência de acidentes dado uma unidade de espaço sobre um determinado período de tempo.”

A Modelagem Estatística de Segurança Viária (MESV) é o ajuste de um modelo estatístico aos dados. Estes dados consistem nas informações a respeito dos acidentes que ocorreram anteriormente, bem como nas características de um conjunto de segmentos de estrada, interseções ou de outros elementos existentes na infraestrutura (Hauer, 2004). De modo geral, de acordo com este autor, os propósitos da MESV consistem em duas alternativas: estimar a frequência esperada de acidentes a partir de um elemento da infraestrutura baseada em suas características; ou estimar a mudança na frequência esperada de acidentes causada por uma alteração em uma característica de um elemento da infraestrutura.

Lord e Mannering (2010) mostram em seus estudos a evolução nos últimos anos e as inovações metodológicas que introduziram algumas abordagens estatísticas muito promissoras, como: Modelos de parâmetros aleatórios; Modelos de mistura finita; Modelos de computação de Markov. Os autores ainda salientam que também existem outros modelos que são promissores para melhorar a compreensão dos motivos que afetam a frequência da acidentalidade viária, e, é esperado que nos próximos anos as variações e refinamentos desses modelos mais avançados possam auxiliar a revelar novas informações úteis para atenuar o cenário das tragédias rodoviárias. Apesar disso, ainda assim são reconhecidas as limitações neste tipo de abordagem, uma vez que cada modelo estatístico tem pressupostos próprios e contém uma relação pré-definida entre variáveis dependentes e independentes (Zeng et al., 2016).

Dessa forma, o emprego de áreas da inteligência artificial como o aprendizado de máquina se mostra eficiente diante deste contexto, uma vez que não exige este tipo de relação pré-definida entre as variáveis, sendo um campo de grande destaque atualmente as redes neurais artificiais, que, em suma, apresentaram desempenho superior à modelagem tradicional de acordo com os resultados obtidos (Chang, 2005 apud Silva; Andrade; Weigang, 2017).

4.3 INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA

Conforme Costa et al. (2021) “máquinas que se assemelham a humanos, em sua aparência física e seu comportamento intelectual, sempre foram um sonho da humanidade”. A Inteligência Artificial (IA) é um conceito que se originou nos anos 50, com o objetivo principal de executar funções e tarefas que no momento são melhores desempenhados pelo homem (Machado, 2011).

De acordo com seus estudos, Machado (2011) classifica as IA's da seguinte forma:

- Sistemas que raciocinam de forma semelhante a dos seres humanos;
- Sistemas que pensam de forma racional;
- Sistemas que agem como os seres humanos;
- Sistemas que agem de forma racional;

O teste de Alan Turing desenvolvido durante a segunda guerra mundial foi a base para que todos os estudos a respeito do tema fossem desenvolvidos posteriormente nas décadas seguintes. Sendo considerado o pai da inteligência artificial e da ciência computacional, o teste de Turing consistia em avaliar a interação entre três elementos (entrevistador, homem e máquina), no qual, uma máquina e um homem seriam interrogados por uma terceira pessoa de modo que esse interrogador não soubesse com quem estaria dialogando, uma vez que cada um se encontrava isolado um do outro. Se o entrevistador não fosse capaz de distinguir qual resposta era de um humano e qual era da máquina, dessa forma, a IA passaria no teste (Machado, 2011).

O Aprendizado de Máquina (AM) ou *machine learning*, caracteriza-se por ser uma subárea da inteligência artificial, e por conta disso, torna-se uma ferramenta muito poderosa para resolver problemas em vários domínios de aplicação (Silva, 2019).

Para Machado (2011), a área de AM é responsável por pesquisar métodos computacionais adequados para a obtenção de novos conhecimentos, novas habilidades e novas formas de organização do conhecimento já existente, de forma que esse aprendizado possibilita que o sistema faça uma tarefa ou tarefas sobre uma mesma população de modo mais eficiente a cada execução.

Kopiler et al. (2019) destacam que devido à capacidade de simular a forma de pensar e de resolver problemas, fornecendo soluções iguais ou semelhantes aos de um especialista da área, torna atraente a utilização de técnicas de inteligência computacional nas resoluções de problemáticas que envolvem um nível maior de complexidade. Logo, mediante a este cenário

que se destacam as Redes Neurais Artificiais, um campo de grande proeminência na área da inteligência computacional.

4.3.1 Redes neurais artificiais (RNA's)

As Redes Neurais Artificiais ou RNA's são modelos computacionais criados com base na idéia de simular o cérebro humano em sua funcionalidade e potencialidade, emulando, dessa forma, o comportamento do cérebro e sua estrutura, fazendo com que seja análoga a fisiologia humana (Kopiler et al., 2019).

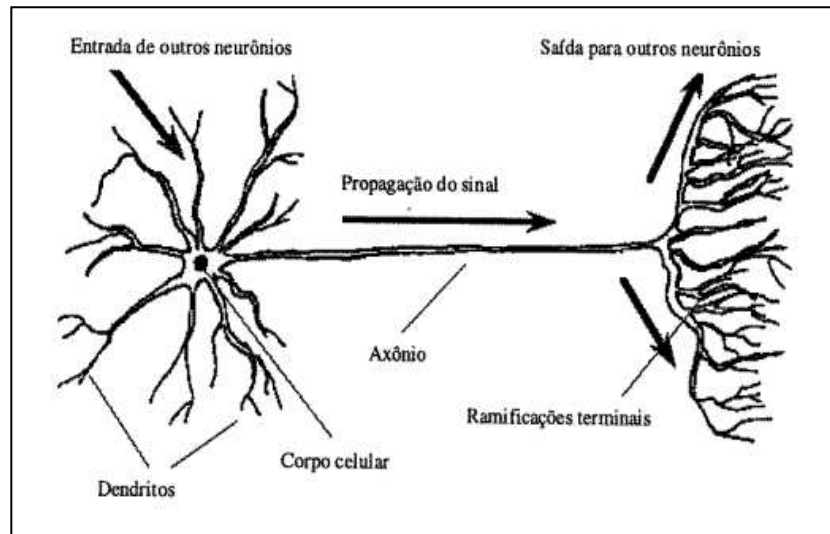
Assim, de modo formal, para Haykin (2001, p.28),

Uma rede neural é um processador maciçamente paralelamente distribuído constituído de um grande número de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso.

Machado (2011) destaca que uma grande rede neural artificial pode ter centenas ou milhares de unidades de processamento, enquanto o cérebro de um mamífero pode ter muitos bilhões de neurônios. Dito isso, o autor aponta que o papel principal das RNA's consiste em solucionar problemas de inteligência artificial, construindo um sistema que tenha circuitos que simulem o cérebro humano, inclusive seu comportamento, ou seja, aprendendo, errando e fazendo descobertas.

A **Figura 1** mostra o modelo do neurônio humano.

Figura 1 – Estrutura de um neurônio humano



Fonte: Silva (1998).

A partir da **Figura 1** é possível observar como é estruturado o neurônio humano, no qual, os neurônios são conectados uns aos outros através de sinapses, e a partir disso, formam uma grande rede denominada de rede neural. As sinapses transmitem estímulos através de diferentes concentrações de Na^+ (Sódio) e K^+ (Potássio), e o resultado disto pode ser estendido por todo o corpo humano, onde essa grande rede proporciona a capacidade de processar e armazenar informações (Machado, 2011).

Haykin (2001, p.36) ressalta que para um bom funcionamento de uma RNA um neurônio artificial é primordial. Este elemento nada mais é do que uma unidade de processamento de informação fundamental para a operação dessa rede. Segundo o autor, são três os elementos básicos principais de um modelo de neurônio artificial:

- Um conjunto de sinapses ou elos de conexão, onde cada uma delas é caracterizada por um peso ou força própria. Assim, quando chega um sinal $x(j)$ na entrada da sinapse ele é conectado ao neurônio k e em seguida multiplicado pelo peso sináptico $w(kj)$. Quanto à notação dos pesos sinápticos, o primeiro índice indica qual é o neurônio em questão, e o segundo se refere ao terminal de entrada da sinapse do qual o peso se refere.
- Um somador que realiza a soma dos sinais de entrada, no qual é ponderada pelas respectivas sinapses do neurônio.
- Uma função de ativação que serve para restringir a amplitude da saída de um neurônio. Essa função também é chamada de restritiva, pois, limita o intervalo

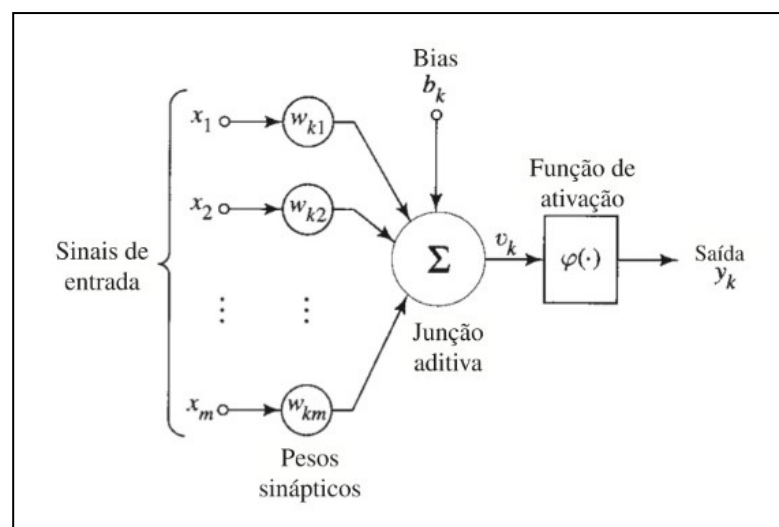
permitido da amplitude do sinal de saída a um valor finito. Esse intervalo é normalizado como um intervalo unitário $[0,1]$ ou alternativamente $[-1,1]$.

O modelo neural possui também um *bias* aplicado externamente que possui o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele for positivo ou negativo, respectivamente (Haykin, 2001, p.37).

Machado (2011) aponta que uma RNA é composta por várias unidades de processamento cujo funcionamento consiste nessas unidades uma vez conectadas por canais de comunicação e associados a determinado peso realizarem operações apenas sobre seus dados locais, que são entradas recebidas pelas suas conexões. O comportamento inteligente de uma rede neural artificial vem das interações entre as unidades de processamento da rede.

A **Figura 2** mostra o modelo de um neurônio artificial.

Figura 2 – Estrutura de um neurônio artificial



Fonte: Haykin (2001).

O funcionamento de um neurônio artificial, proposto por McCulloch e Pitts em 1943, consiste em sinais que são apresentados à entrada onde cada sinal é multiplicado por um número ou peso, que irá indicar a sua influência na saída da unidade. Após isso, é realizada a soma ponderada dos sinais que produz um nível de atividade, no qual, se este nível exceder um determinado limite denominado de *threshold*, o neurônio irá produzir uma determinada resposta de saída (Machado, 2011).

Kopiler et al. (2019) salientam que a inteligência dos neurônios de McCulloch e Pitts está toda concentrada nos pesos, pois o neurônio opera sobre a soma ponderada de várias entradas. Em termos matemáticos, um neurônio k é descrito pelas seguintes equações:

$$u_k = \sum_{j=0}^m w_{kj} \cdot x_j \quad (1)$$

$$y_k = \varphi(u_k + b_k) \quad (2)$$

sendo:

- x_j = sinais de entrada;
- w_{kj} = pesos sinápticos do neurônio k e entrada j ;
- u_k = saída do combinador linear devido aos sinais de entrada;
- b_k = bias;
- φ = função de ativação;
- y_k = sinal de saída do neurônio;

A **Equação 1** significa apenas que as entradas são apresentadas ao neurônio k e são multiplicadas cada um por um peso específico, onde cada parcela desta multiplicação é somada, resultando em u_k . Campo local induzido ou potencial de ativação do neurônio é o nome dado do resultado desta operação. Já a **Equação 2**, indica que é aplicada uma função (denominada função de ativação) a u_k , resultando na saída do neurônio (denominada y) (Kopiler, 2019).

Outro elemento muito importante se tratando de redes neurais artificiais é a função de ativação, que esta relacionada à capacidade representativa que as RNA's possuem (Silva, 2019). Kopiler et al (2019) complementa que as funções de ativação mais utilizadas são em sua grande maioria não lineares e de comportamento relativamente simples. Estas têm a propriedade de transformar não linearmente o sinal de entrada, assim, a saída transformada é enviada para próxima camada de neurônios como entrada, mediante a isto, relaciona-se à decisão de ativar ou não o neurônio com base na informação recebida (Silva, 2019).

Kopiler et al. (2019) afirmam que as funções não lineares procuram simular a ideia do comportamento básico de um neurônio, onde gera a saída ou disparo (nos termos dos neurocientistas) ocorrendo o *thresold* de ativação. Assim, de modo geral, uma função de

ativação tem o cargo de decidir a forma como o neurônio responde ao nível de excitação, limitando e definindo a saída da rede neural (Cintra, 2018). “O estado de ativação de um neurônio pode assumir valores binários (0 para inativo e 1 para ativo), bipolares (-1 e 1) ou valores reais nos intervalos de 0 a 1 ou de -1 a 1.” (Martineli, 1999).

Em suma, as principais funções de ativação são mostradas na **Tabela 3**:

Tabela 3 – Principais funções de ativação

Função	Equação com Polarização	Sem polarização	Com polarização
Degrau	$y = \begin{cases} 1, & x > -b \\ 0, & x < -b \end{cases}$		
Degrau Simétrico	$y = \begin{cases} 1, & x > -b \\ 0, & x < -b \end{cases}$		
Linear	$y = x + b$		
Logística Sigmoidal	$y = \frac{1}{1 + e^{-(n+b)}}$		
Tangente Sigmoidal	$y = \frac{e^{(x+b)} - e^{-(x+b)}}{e^{(x+b)} + e^{-(x+b)}}$		

Fonte: Adaptado de Moreira (2007).

Contudo, as funções de ativação mais usadas são do tipo sigmóide, sendo a tangente hiperbólica (tanh) e a sigmóide logística, isto, devido ao fato que além de serem diferenciáveis, essas funções apresentam também um comportamento mais gradual nas regiões de transição ativação/não ativação, o que oferece maior nuance nas suas respostas (Kopiler et al., 2019).

As expressões estrutura, arquitetura ou topologia de uma RNA, referem-se à forma como os neurônios artificiais são organizados e distribuídos em uma rede neural, logo, estes termos apresentam como estão interligados os nós entre si e como a informação é transmitida

através da rede (Cintra, 2018). Souza (2015) acrescenta que a maneira pelo qual os neurônios são estruturados esta intimamente ligada com o algoritmo de aprendizagem usado para treinar a rede, deste modo, o desempenho dela depende não apenas da função de ativação utilizada pelos neurônios como também dos algoritmos de aprendizado usados na etapa de treinamento.

Silva (2019) evidencia que há dois tipos de aprendizagem de uma RNA: supervisionado e não supervisionado. No primeiro caso, os valores de resposta das amostras do conjunto de treinamento já são conhecidos, assim, o objetivo baseia-se em aprender um mapeamento de x para y , dado um conjunto de treinamentos com os pares (x_i, y_i) . Nesta condição, a resposta do modelo ao padrão atual das entradas é avaliada, permitindo que as alterações que levem a resposta do modelo aproximem-se da resposta esperada ou já conhecida. Quanto ao segundo caso, a aprendizagem não supervisionada, o autor afirma que as classes são desconhecidas e o objetivo consiste em encontrar uma estrutura com relações ou padrões entre os dados (x_1, \dots, x_n) dos n exemplos que foram dispostos.

O aprendizado supervisionado pode ser aplicado em dois tipos de situações: para classificação e para a regressão. Os problemas de classificação se referem a aqueles onde se pretende prever um valor discreto, ou seja, classificar um dado segundo uma categoria, enquanto os problemas de regressão são aqueles que visam prever um valor contínuo (Oliveira, 2021).

Haykin (2001, p.75) enfatiza que a propriedade que é de importância essencial para uma rede neural é a sua habilidade de aprender a partir de seu ambiente, e, além disso, de melhorar o seu desempenho com a aprendizagem ocorrida. O autor explica ainda que esta melhora citada só ocorre com o tempo e de acordo com alguma medida já preestabelecida, no qual, uma rede neural aprende cada vez mais a partir do seu ambiente e através de um processo iterativo de ajustes aplicados aos seus pesos sinápticos e níveis de *bias*. Dessa forma, com o decorrer do processo de aprendizagem, a rede se torna mais instruída sobre o seu ambiente após a cada nova iteração e responde de uma maneira nova devido às modificações ocorridas em sua estrutura interna.

Silva (2019) constata que quanto à arquitetura de uma rede neural, há dois tipos: as *feedforwards* ou as recorrentes. As redes do tipo *feedforwards* (mais comuns e mais utilizadas) são conhecidas pela propagação dos sinais sempre das camadas anteriores para as posteriores em direção à camada de saída, enquanto que as redes do tipo recorrentes contém, pelo menos, um laço realimentando a saída de neurônios para outros neurônios da rede de forma cíclica.

Dentre as redes neurais do tipo *feedforwards*, se destacam aquelas constituídas por um elemento muito conhecido no campo das RNA', o *perceptron*.

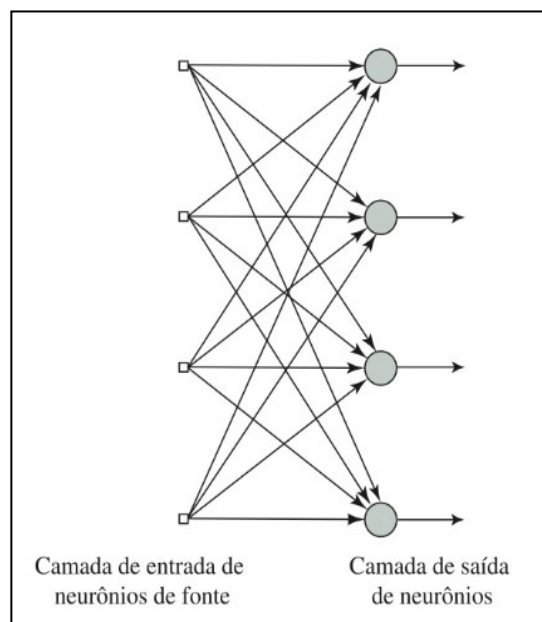
4.3.1.1 *Perceptron*

O *perceptron*, elemento idealizado por Rosenblatt, em 1958, é a forma mais simples de configuração de uma rede neural artificial, cujo propósito focava em criar um modelo computacional inspirado na retina humana, em que o objetivo era obter um elemento de percepção eletrônica de sinais (Silva; Spatti; Flauzino, 2016, p.57).

Construído em torno de um único neurônio artificial, o *perceptron* é limitado a realizar classificações de padrões com apenas duas classes ou hipóteses. Uma vez que se expande a camada de saída do *perceptron* para incluir mais de um neurônio, possibilita-se realizar classificações com mais de duas classes correspondentemente. Contudo, tais classes devem ser linearmente separáveis para que o *perceptron* atue e funcione de forma adequada (Haykin, 2001, p.143).

A **Figura 3** apresenta a arquitetura de uma rede de camada única.

Figura 3 – Rede neural de camada única (Rede de *perceptrons*)



Fonte: Haykin (2001).

Como o *perceptron* pertence à arquitetura *feedforward* de camada única (simples), o fluxo de informações em sua estrutura se desloca sempre no sentido da camada de entrada em direção à camada neural de saída, inexistindo-se qualquer tipo de realimentação de valores

produzidos pelo seu único neurônio. Já quanto ao ajuste dos pesos e o limiar do *perceptron*, ambos são efetuados utilizando processo de treinamento supervisionado, no qual, para cada amostra dos sinais de entrada se tem a respectiva saída (resposta) desejada. Como o *perceptron* é tipicamente usado em problemas de reconhecimento de padrões, sendo que sua saída pode assumir somente dois valores possíveis, então cada um dos valores será associado a uma das duas classes que o *perceptron* identificará (Silva; Spatti; Flauzino, 2016, p.59).

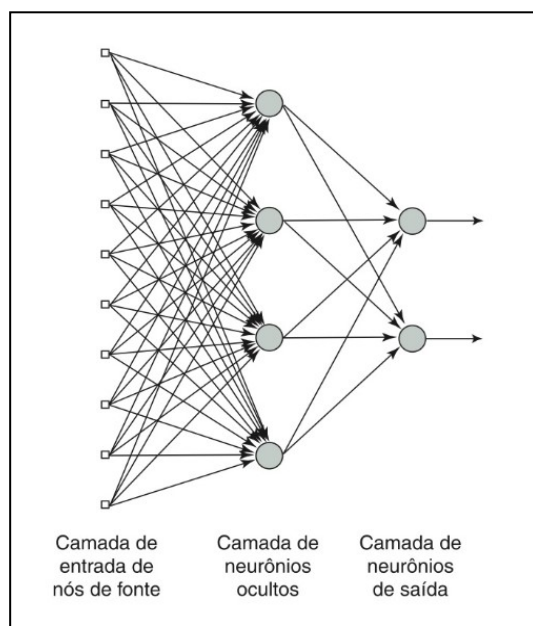
Kopiler et al. (2019) apontam que por se tratar de uma rede neural de camada única cujo as capacidades são extremamente limitadas, este tipo de rede também chamada de rede de *perceptrons*, praticamente já não é mais utilizada, visto que Minsky e Papert provaram no final da década de 60 que este tipo de rede só conseguia classificar corretamente problemas linearmente separáveis. Tal conjuntura só mudou a partir do início da década de 80, quando Rummelhart desenvolveu o algoritmo do *Backpropagation* para o treinamento de redes de múltiplas camadas, problema esse que era considerado sem solução até então.

4.3.1.2 *Multilayer Perceptron (MLP)*

As redes *perceptrons* de múltiplas camadas ou também denominadas pela literatura de *Multilayer perceptron (MLP)*, são caracterizadas pela presença de pelo menos uma camada intermediária (oculta) de neurônios, situada entre a camada de entrada e a respectiva camada de saída, no qual essas camadas ocultas podem possuir pesos distintos. Devido a isso, as redes MLP possuem no mínimo duas camadas de neurônios, os quais estarão distribuídos entre as camadas intermediárias e a camada de saída (Silva; Spatti; Flauzino, 2016, p.91).

A **Figura 4** apresenta a arquitetura de uma rede de múltiplas camadas.

Figura 4 – Rede neural *perceptron* de múltiplas camadas (MLP)



Fonte: Haykin (2001).

Silva, Spatti e Flauzino (2016, p.93) salientam que em contraste da rede de *perceptrons* ou de outros tipos, além da camada oculta em sua topologia, as redes MLP podem ser compostas por diversos neurônios na camada de saída, sendo que cada um destes neurônios representaria uma das saídas (classes) do processo a ser mapeado. Assim, se tal processo consistisse em m saídas, esta rede teria também m neurônios em sua última camada neural.

Os *perceptrons* de múltiplas camadas têm sido empregados com sucesso para a resolução de diversos problemas complexos, através do seu treinamento de forma supervisionada com o algoritmo de retropropagação de erro ou *backpropagation* (Haykin, 2001, p.182). Este algoritmo, segundo o autor, consiste basicamente na aprendizagem por meio da retropropagação de erro através das diferentes camadas da rede, fundamentado em duas fases: no passo para frente (propagação) e no passo para trás (retropropagação). No primeiro, um vetor de entrada é aplicado aos nós sensoriais da rede e o efeito desse fenômeno se propaga pela rede camada por camada até um conjunto de saídas onde é produzido como resposta real da rede, durante esse passo os pesos sinápticos da rede são todos fixos (Haykin, 2001, p.182). Já durante o segundo, o passo para trás, o autor destaca que os pesos sinápticos são todos ajustados de acordo com uma regra de correção de erro, no qual a resposta real da rede é subtraída de uma resposta desejada (alvo) para assim produzir um sinal de erro que é propagado para trás através da rede, na direção contrária das conexões sinápticas.

Silva (2019) aponta que apesar de existir outras arquiteturas de redes *feedforwards*, como Redes de Função Base Radial (RBF), é o MLP o tipo mais popular de rede supervisionada, tendo a fama de ser uma espécie de aproximador universal e robusto para problemas de previsão e classificação. Em suas observações, ele afirma que é, também, o mais utilizado para análise de dados de acidentes.

Silva (2019) complementa que quando se trata de acidentalidade viária, para o caso de modelos de previsão de frequência de acidentes totais em um determinado trecho, irá existir apenas um neurônio na camada de saída (y_{k1}), ao passo que, nos modelos de classificação de acidentes por severidade o número de saídas corresponderá ao número de classes de severidade pré-estabelecidos. O autor exemplifica que em um problema em que essas classes fossem classificadas em três tipos: sem vítimas, com vítimas e com vítimas fatais, o número de neurônios na camada de saída seria três (y_{k1}, y_{k2}, y_{k3}).

4.3.1.3 Treinamento de uma rede neural artificial

Especificar a topologia de uma rede MLP mais apropriada para mapear um problema específico é normalmente efetuado de modo empírico (Silva; Spatti; Flauzino, 2016, p.146). Haykin (2001, p.51) complementa que a forma de representação de um conhecimento advindo das redes neurais artificiais depende e é definido pelos valores de parâmetros utilizados, como função de ativação, pesos sinápticos, *bias*, número de épocas, taxa de aprendizagem, número de camadas ocultas e o número de neurônios da camada oculta.

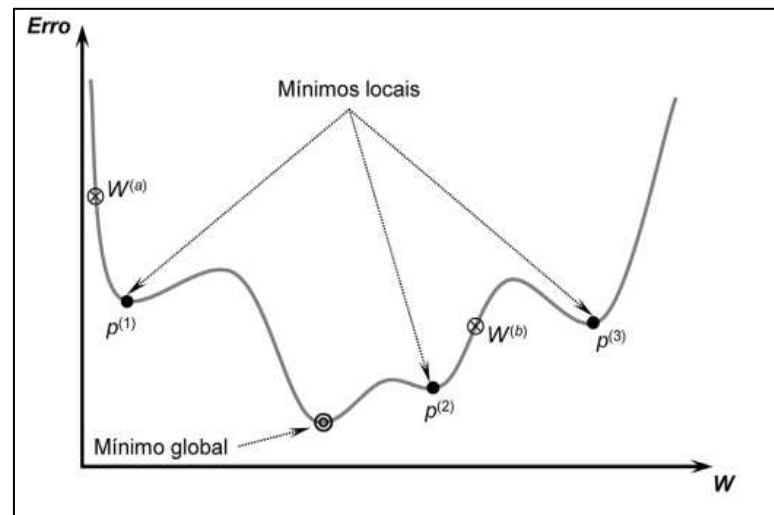
O número de ciclos de treinamento ou épocas consiste na quantidade de vezes que todo o conjunto de treinamento foi apresentado à RNA durante o processo de aprendizagem (Silva, 1998). Logo, quanto maior for o número de épocas utilizadas, maior será o tempo de processamento que a rede neural irá demandar para encerrar o treinamento.

A taxa de aprendizagem dita o quão rápido a rede neural atualiza os pesos sinápticos, ou seja, o quão rápido o processo de treinamento da rede será conduzido à sua convergência ou estabilização (Silva; Spatti; Flauzino, 2016, p.64). Os autores ainda complementam que a escolha da taxa de aprendizado deve ser realizada com cautela visto que deve ser evitado a instabilidade no processo de treinamento de uma rede, sendo adotado a essa taxa de aprendizagem valores pertencentes ao intervalo compreendido de 0 a 1.

Outro parâmetro de grande importância que também varia dentro do intervalo de 0 a 1 e é aplicável no algoritmo *backpropagation* é o *momentum*. Este parâmetro consiste numa constante que determina o efeito das mudanças anteriores de pesos na direção do movimento

atual no espaço dos pesos (Moreira, 2007). De modo geral, quando é encontrado um mínimo local durante o treinamento o erro do conjunto de treinamento para de diminuir e estaciona em um valor maior que o aceitável, o *momentum* tem como objetivo fazer com que a rede não fique estagnada nesse mínimo local, fazendo com que a rede encontre o menor erro global da função. Logo, com a inserção do termo *momentum*, o processo de convergência da rede se torna bem mais eficiente, uma vez que é levado em consideração o critério de quão afastada está a solução atual da solução final (ótima) (Silva; Spatti; Flauzino, 2016, p.113). A **Figura 5** exemplifica este cenário:

Figura 5 – Pontos de mínimo locais e menor erro global associados à função erro



Fonte: Silva; Spatti; Flauzino (2016).

No âmbito de RNA's, o conceito de generalização refere-se à capacidade de uma rede de mapear os valores de entrada-saída que não foram usados na amostra de treino, gerando valores aproximados daqueles esperados (Silva, 2018). Há dois cenários extremos relacionados ao conceito de generalização que um modelo de excelente eficácia deve evitar: o *overfitting* (superajuste) e o *underfitting* (sub-ajuste). O *overfitting* é quando o modelo aprende demais sobre os dados, fazendo com que o modelo fique adequado somente para os dados de treinamento, memorizando excessivamente estes dados e tornando o modelo incapaz de generalizar frente a novos dados (Associação Brasileira de Ciências de Dados – ABRACD, 2022). Já o segundo, o *underfitting*, é quando o modelo não conseguiu aprender de forma suficiente a partir dos dados de treino, ou seja, não conseguiu encontrar a relação entre as variáveis, levando a um erro elevado tanto no treinamento como no teste (ABRACD, 2022).

Souza (2015) afirma que usualmente técnicas de pré-processamento são aplicadas nos dados como método de acelerar a convergência. Martineli (1999) destaca que quando se trata do pré-processamento de uma RNA, o processo é um pouco mais complexo, visto que há a necessidade de normalização das variáveis (quantitativas e qualitativas), garantindo com que todos os atributos estejam dentro de uma mesma escala e intervalo de valores. Não há um consenso em como deve ser efetuado a normalização, porém, tal procedimento deve ser feito em consonância com o intervalo de valores dinâmicos da função de ativação utilizada nas camadas ocultas, comumente representadas pela função logística ou tangente hiperbólica, com o intuito de evitar a saturação dos neurônios que compõe a rede (Silva; Spatti; Flauzino, 2016, p.159).

As variáveis ou atributos quantitativos com valores reais e inteiros são geralmente normalizados por algum operador que os mapeiam no intervalo de $[0,1]$ ou $[-1,1]$ (SILVA, 1998). Já quanto às variáveis simbólicas ou nominais Martineli (1999) complementa que os valores não numéricos devem ser codificados para valores numéricos, o que pode resultar em um vetor binário ou de valores inteiros, no qual, uma característica codificada em valores inteiros deve ser associada à um único elemento de entrada da rede. O autor ainda afirma que a codificação dos dados não numéricos deve levar em consideração a presença ou não de ordenação, onde, quando se trata de dados não ordenados como informações sobre sexo, por exemplo, o sexo feminino pode ser representado pelo vetor $(0,1)$ e o sexo masculino pelo vetor $(1,0)$. Já quando se trata de dados ordenados como classes sociais A, B e C, devem ser codificados em uma seqüencia, onde, as classes sociais corresponderiam respectivamente aos vetores binários $(1,1)$, $(1,0)$ e $(0,0)$, ou aos números inteiros 3,2,1 (Martineli, 1999).

Martineli (1999) também salienta que em relação a toda base de dados, é necessário fazer o devido pré-processamento, dividindo a base de dados em subconjuntos distintos de treinamento, validação e teste. O autor complementa que se tratando de RNA's o subconjunto de validação pode ser utilizado como um critério de parada do processo de aprendizagem. O subconjunto de teste assim como o de validação, também é constituído por exemplos nunca vistos durante a etapa de treinamento e caracteriza-se por ser a última de todo o processo, no qual, é a partir dele que é verificado se a extração de conhecimento foi bem sucedida e absorvida pela rede neural, permitindo a avaliação do modelo criado quanto à capacidade de generalização de dados ainda desconhecidos para a RNA (Martineli, 1999).

Ao efetuar o processo de normalização, cada variável de entrada deve ser pré-processada com o objetivo de que a média calculada sobre todo o conjunto de treinamento

esteja próxima de zero, ou, ao menos que seja um valor pequeno em relação ao desvio-padrão, visto que sem essa condição o vetor peso de determinado neurônio pode ter dificuldades para convergir, tornando lento o treinamento (Lecun, 1993 apud Manzan, 2016).

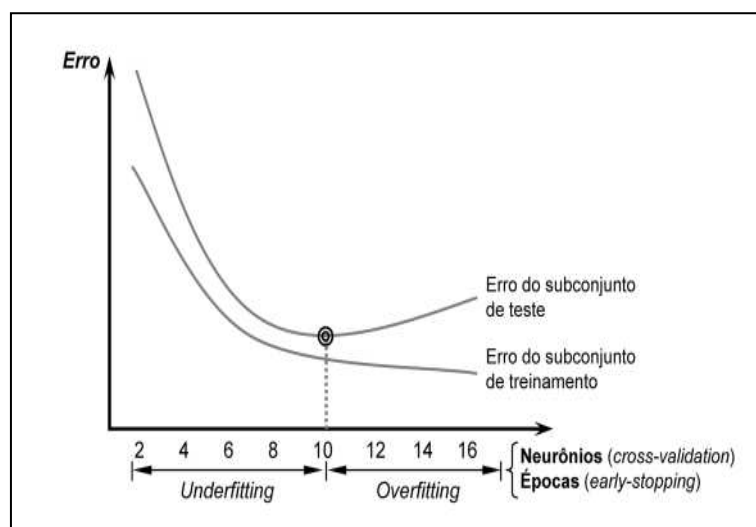
Quando se treina uma RNA busca-se a obtenção de uma rede com a melhor capacidade de generalização possível, ou seja, com a maior capacidade de responder corretamente a dados que não foram utilizados na etapa de treinamento (Silva, 1998). Uma das técnicas estatísticas mais utilizadas para a seleção das melhores topologias candidatas é a validação cruzada (*cross – validation*), cujo objetivo é avaliar a aptidão de cada uma quando aplicadas a um conjunto de dados diferente daquele usado no ajuste de seus hiperparâmetros (Silva; Spatti; Flauzino, 2016, p.147). Os autores ainda complementam que existem três métodos aplicados à técnica de validação cruzada, são eles:

- Validação cruzada por amostragem aleatória (*random subsampling cross-validation*), no qual consiste em dividir aleatoriamente todo o conjunto de dados em diferentes partes, como subconjunto de treinamento e de teste, onde normalmente de 60% a 90% são destinados ao subconjunto de treino, e o restante ao de teste. Essa sistemática de partição deve ser repetida várias vezes durante o processo de aprendizado das topologias candidatas, visando em cada ensaio a possibilidade de maior contemplação de amostras diferentes em cada subconjunto, efetuando-se a média dos desempenhos individuais em cada experimento e obtendo-se o desempenho global da arquitetura proposta.
- Validação cruzada por k -partições (*k-fold cross-validation*), este método baseia-se em realizar a divisão da base de dados total em amostras em k partições, sendo que $(k-1)$ delas serão utilizadas para compor o subconjunto de treino, enquanto que a partição restante irá compor o subconjunto de teste. Dessa forma, o processo de aprendizado se repete k vezes até que todas as partições tenham sido usadas como subconjunto de teste.
- Validação cruzada por unidade (*leave-one-out cross-validation*), este método é o menos usado e consiste na utilização de uma única amostra para o subconjunto de teste sendo as restantes alocadas para compor o subconjunto de treino, assim como o método k -partições, o processo de aprendizado é repetido até que todas as amostras sejam de forma individual usadas como subconjunto de teste, consistindo numa espécie de caso particular da validação cruzada por k -partições.

Silva (1998) afirma que quando se trata de procedimentos envolvendo a validação cruzada o erro do conjunto de validação deve ser avaliado constantemente a cada k iterações (épocas), onde deve ser interrompido o treinamento quando o erro deste subconjunto for maior que o que era há k iterações atrás. Este processo descrito anteriormente é amplamente usado como critério de parada do treino de uma RNA, e é denominado de parada antecipada (*early stopping*). O autor salienta que a abordagem de validação cruzada geralmente utiliza o subconjunto de validação como forma de antecipar o comportamento em situações reais (ou em um conjunto de teste), assumindo que o erro em ambos os casos será semelhante, sendo o erro de validação geralmente uma estimativa do erro de generalização. O uso de um subconjunto de teste nem sempre é empregado, principalmente quando o número de amostras de treinamento é pequeno (Silva, 1998).

De modo geral, o treinamento de uma rede neural artificial envolve diversas problemáticas e exige uma grande dedicação. O **Gráfico 3** exemplifica de forma resumida todo o contexto envolvendo o treinamento de uma RNA:

Gráfico 3 – Treinamento de uma rede neural artificial



Fonte: Silva; Spatti; Flauzino (2016).

Conforme demonstra o **Gráfico 3**, um ótimo nível de generalização é no momento em que os erros do subconjunto de treinamento e de teste (ou validação) se encontrem nos seus níveis mais aceitáveis, ou seja, mais próximos possíveis do menor erro capaz de ser concebido em cada subconjunto, isto, tendo em vista que a RNA deve levar em consideração o compromisso de superar o *underfitting* e evitar a condição de *overfitting* (Silva; Spatti; Flauzino, 2016, p.156).

4.3.1.4 Métricas estatísticas para avaliação de um modelo de predição

Existem diversas métricas estatísticas usadas para a avaliação do desempenho de um modelo de predição, visto que o propósito de um modelo preditivo é o de poder ser aplicado novamente em novos dados ou instâncias (LINKEDIN, 2018). O Coeficiente de Correlação de Pearson (r , **Equação 3**), o Erro Absoluto Médio (EAM, **Equação 4**), a Raiz do Erro Quadrático Médio (REQM, **Equação 5**) e a Acurácia (AC, **Equação 6**) foram métricas utilizadas para mensurar a eficácia e o erro na estimativa dos dados de acidentalidade viária das RNA's, Equações de 3 a 6, respectivamente:

$$r = \frac{\sum_{i=1}^n (O_i - \bar{O}) * (P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 * \sum_{i=1}^n (P_i - \bar{P})^2}} \quad (3)$$

$$EAM = \frac{\sum_{i=1}^n |O_i - P_i|}{n} \quad (4)$$

$$REQM = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (5)$$

$$AC = \frac{VP + VN}{VP + VN + FP + FN} \quad (6)$$

sendo:

- O_i = são os valores observados;
- P_i = são os valores preditos;
- n = número de observações;
- \bar{O} = valor médio dos valores observados;
- \bar{P} = valor médio dos valores preditos;
- VP = número de verdadeiros positivos;
- VN = número de verdadeiros negativos;
- FP = número de falsos positivos;
- FN = número de falsos negativos;

O Coeficiente de Correlação de Pearson (**Equação 3**) é uma medida que indica a associação linear entre variáveis, ela varia de -1 a 1, cujo sinal indica a direção positiva ou negativa do relacionamento e o valor sugere a força de relação entre as variáveis (Figueiredo Filho; Silva Júnior, 2009).

O Erro Médio Absoluto (**Equação 4**) mede o quão próximas as previsões são dos resultados reais, ou seja, consiste na média de todos os erros do modelo, em que o erro do modelo é a distância absoluta entre o valor de rótulo previsto e o valor de rótulo correto, onde, quanto mais próximo de 0 melhor é a qualidade (MICROSOFT, 2023).

A Raiz do Erro Quadrático Médio (**Equação 5**) é uma métrica que mede a diferença entre valores previstos por um modelo e os valores realmente observados no ambiente que esta sendo modelado, é semelhante ao EAM mas dando mais peso para as diferenças maiores (MICROSOFT, 2023). Uma característica do REQM é que os erros (reais – previsões) são elevados ao quadrado antes de ter a média calculada, fazendo com que os *outliers* no conjunto de dados tenham um peso maior para o cálculo do REQM (LINKEDIN, 2018).

A Acurácia (**Equação 6**) é uma métrica que indica a proporção de previsões corretas contidas dentro de seu modelo, no qual, é definido como a soma do número de positivos e negativos verdadeiros divididos pela soma dos positivos e negativos verdadeiros e falsos positivos e negativos (*International Business Machines Corporation - IBM*, 2021).

5 MATERIAL E MÉTODOS

Neste tópico são abordados os elementos necessários para desenvolvimento da metodologia do trabalho, como o material que foi utilizado e os métodos que foram empregados, além disso, foi realizada a caracterização e localização da área de estudo.

5.1 ÁREA DE ESTUDO

A rodovia federal BR-365 possui 868,2 quilômetros de extensão, tendo seu início no município de Montes Claros - MG e seu término no município de Santa Vitória - MG (Departamento Nacional de Infraestrutura de Transporte - DNIT, 2022).

A BR-365 é a principal via de ligação entre o Triângulo Mineiro e Alto Paranaíba com o nordeste do país, sendo esta rodovia marcada fortemente pela movimentação intensa de tráfego pesado, com grande predominância tanto de caminhões como também de carretas (PATOSHOJE, 2017).

A área de estudo deste presente trabalho contém uma extensão de 333 quilômetros, onde apenas do Km 608 ao Km 610,8 e do Km 614,7 ao Km 708,9, constituem trechos de pista dupla, sendo todo o restante do segmento escolhido composto por pista simples. A **Figura 6** apresenta o mapa de localização do trecho supracitado.

Figura 6 – Mapa de localização da área de estudo



Fonte: Autoria própria (2023).

5.2 MATERIAL

Neste tópico é apresentado o material utilizado para o desenvolvimento da metodologia deste trabalho.

5.2.1 Dados de acidentalidade viária

- Dados de acidentalidade viária obtidos a partir das planilhas da PRF – (dados abertos), disponibilizados no site do órgão (<https://www.gov.br/prf/pt-br/aceso-a-informacao/dados-abertos/dados-abertos-acidentes>)

5.2.2 Dados espaciais

- A base cartográfica das rodovias federais utilizada foi obtida a partir do Sistema Nacional de Viação - SNV disponibilizado pelo DNIT, cuja *shapefile* esta referenciada no sistema de referência geodésico SIRGAS2000 disponíveis no *site* (<https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/dnit-geo>). A base cartográfica do Brasil utilizada é disponibilizada pelo Instituto Brasileiro de Geografia e Estatística – IBGE no *site* do órgão (<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais>).

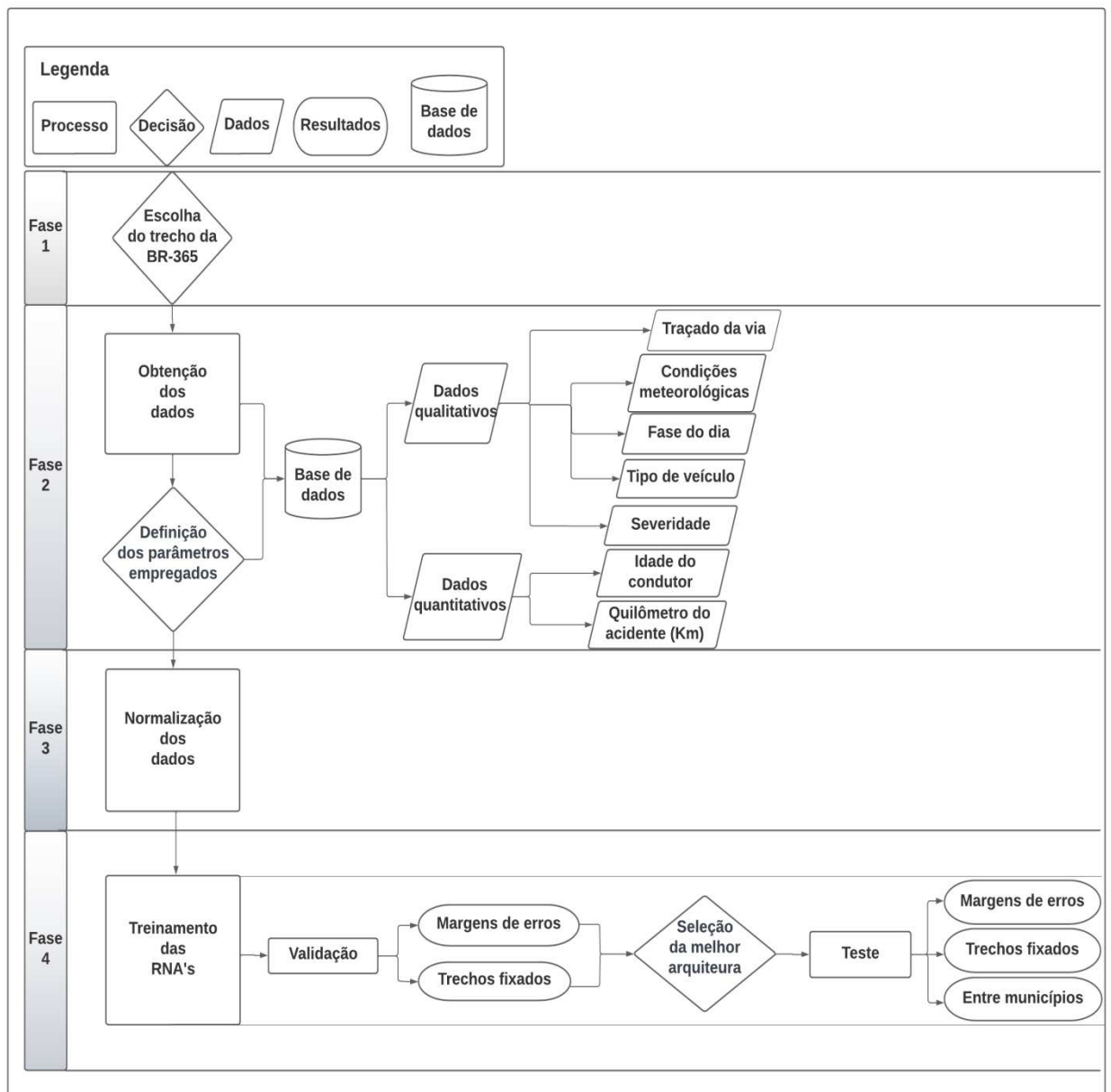
5.2.3 Hardware e Software

- Notebook Asus Intel Core i5-11300H, 8 GB RAM, SSD 512GB, GTX 1650 Max;
- *Software Excel 2007* (Normalização dos dados e cálculos);
- *Software ArcGIS* (Mapa de localização da área de estudo e análises cartográficas dos resultados);
- *Software Weka 3.8.6* (Treinamento das RNA's) disponibilizado no *site* (<https://www.cs.waikato.ac.nz/ml/weka/>)

5.3 MÉTODOS

Para melhor compreensão do desenvolvimento da metodologia que foi aplicada, foi elaborado um fluxograma (**Figura 7**) que aponta as principais etapas deste processo.

Figura 7 – Fluxograma dos processos metodológicos



Fonte: Autoria própria (2023).

5.3.1 Escolha do trecho da BR-365 (Fase 1)

Primeiramente, o início do processo metodológico consistiu em definir a área de estudo, ou seja, delimitar o trecho da BR-365 que foi analisado no presente estudo, sendo o escolhido o segmento que vai do Km 420,1 até o Km 753,1, localizado entre as cidades de Patos de Minas e Ituiutaba em Minas Gerais.

5.3.2 Dados (Fase 2)

5.3.2.1 Obtenção dos dados (Etapa 2.1)

Os dados de acidentalidade viária da área de estudo foram obtidos a partir da base de dados abertos disponíveis no site do governo federal que são agrupados em três categorias diferentes: “Agrupados por ocorrência”, “Agrupados por pessoa” e “Agrupados por pessoa – Todas as causas e tipos de acidentes (a partir de 2017 até o presente)”. O presente trabalho utilizou as planilhas fornecidas pela segunda opção de agrupamento citada anteriormente. Uma vez feito o *download*, a planilha vem compactada e ao extraí-la é fornecida na extensão (csv), ao abri-la é possível visualizar tanto variáveis quantitativas quanto qualitativas.

A planilha fornece informações importantes como a data, o dia da semana, a Unidade Federativa (UF), a BR em que a ocorrência aconteceu bem como o Km, a causa e o tipo de acidente, o tipo de veículo, a severidade do sinistro de trânsito e muitos outros dados extremamente relevantes, como a geolocalização do ocorrido, fornecendo a latitude e longitude geográfica.

5.3.2.2 Definição dos parâmetros a serem empregados (Etapa 2.2)

Em posse dos dados obtidos uma breve análise preliminar foi feita para a escolha de quais parâmetros (variáveis) seriam utilizados como entrada das RNA's, assim como também qual seria a janela temporal escolhida. As variáveis selecionadas como propostas para o estudo são: traçado da via, período do dia, tipo de veículo, condições meteorológicas, severidade do sinistro de trânsito, a idade do condutor e o quilômetro (Km) em que ocorreu o acidente.

A escolha janela temporal foi baseada em estudos anteriores como da Oliveira (2021) e do Silva (2019), foi determinado um intervalo de três anos (2017 a 2019), visto que o ano mais recente (2022) encontrava-se com informações incompletas, sendo divulgadas em sua totalidade apenas em abril de 2023, período em que a execução do presente projeto já estava em andamento. Já em relação aos anos de 2020 e 2021 foram anos atingidos fortemente pela pandemia, o que pode ter influenciado diretamente na acidentalidade viária, tornando esse período possivelmente uma amostra atípica do que normalmente ocorria não só nesse trecho da BR-365, mas como também em outras rodovias federais. Este contexto pode ser justificado devido às consequências das políticas de restrição e de distanciamento do COVID-19, que fez com que as pessoas passassem a se deslocar menos e a se manterem mais isoladas.

5.3.3 Normalização dos dados (Fase 3)

Com os parâmetros definidos e a janela temporal de dados estabelecida, foi então efetuado o processo de normalização das variáveis qualitativas e quantitativas.

Para os parâmetros qualitativos foram atribuídos pesos que variam num intervalo numérico de 0 até 1, respeitando os limites da função de ativação sigmóide, onde, esse peso varia conforme a menor ou maior influência na ocorrência de sinistros de trânsito. Todos os valores dos pesos atribuídos foram justificados conforme dados do IPEA, PRF e DNIT, baseado nas características que essas variáveis possuem ou promovem na acidentalidade viária federal de acordo com os registros ao longo dos anos, buscando uma associação entre estas características e os sinistros de trânsito, conforme demonstra a **Tabela 4**. Além disso, ainda a respeito das variáveis qualitativas, o atributo “Severidade” por se tratar de um parâmetro nominal, foi normalizado a partir de vetores binários conforme a **Tabela 5**, ordenados de acordo com a severidade da ocorrência.

Além disso, para os parâmetros cujos valores não foram informados ou eram desconhecidos foi arbitrado o símbolo de interrogação, sendo uma alternativa a este contexto conforme explicitado por Martineli (1999).

Tabela 4 – Atribuição de pesos para cada parâmetro nominal (0 a 1)

Atribuição de pesos para cada parâmetro (de 0 a 1)		
Parâmetro	Peso	Justificativa
Traçado da rodovia		
Não Informado	?	-
Túnel	0,1	Menor incidência
Viaduto	0,2	Pouca incidência
Ponte	0,3	Pouca incidência
Retorno Regulamentado	0,4	Pouca incidência
Rotatória	0,5	Exige mais atenção do que em túneis, retornos, viadutos e pontes
Desvio temporário	0,6	Desvios não previstos pelo condutor
Interseção de vias	0,7	Cruzamentos tendem a ser mais perigosos

Reta	0,8	Maior incidência
Curva	0,9	Comprometimento da visibilidade
Condições meteorológicas		
Ignorado	0	Não levado em consideração
Neve	0,1	Improvável de acontecer na região
Granizo	0,2	Improvável de acontecer na região
Vento	0,3	Pouco frequente na região
Nevoeiro/Neblina	0,4	Comprometimento da visibilidade
Sol	0,5	Comprometimento da visibilidade e maior incidência
Garoa/Chuvisco	0,6	Comprometimento da visibilidade e pista escorregadia
Nublado	0,7	Maior incidência
Céu claro	0,8	Maior incidência
Chuva	0,9	Comprometimento da visibilidade e pista escorregadia
Fase do dia		
Anoitecer	0,3	Comprometimento da visibilidade
Amanhecer	0,5	Maior razão de casos com óbito em relação ao anoitecer
Plena noite	0,7	Comprometimento da visibilidade
Pleno dia	0,9	Volume de tráfego ser maior
Tipo de veículo		
Outros	0,1	Menor incidência
Caminhões pesados e Especiais	0,2	Menor incidência
Caminhões leves e Ônibus	0,5	Grande fluxo em rodovias (BR)
Carros e veículos pequenos	0,7	Maior incidência
Motos	0,9	Maior vulnerabilidade do condutor

Fonte: Adaptado de Oliveira (2021).

Tabela 5 – Atribuição de vetores binários para o parâmetro “Severidade”

Atribuição de vetores binários para o parâmetro severidade	
Severidade	Vetores binários
Sem vítimas (Ilesas)	[0,0]
Com vítimas feridas (Feridas)	[1,0]
Com vítimas fatais (Fatais)	[1,1]

Fonte: Autoria própria (2023).

Quanto aos parâmetros quantitativos, ambas as variáveis foram normalizadas para o intervalo numérico entre 0 a 1 na tentativa de atenuar o esforço computacional. Para a normalização das variáveis quantitativas foi utilizada a **Equação 7**:

$$X_{saída} = \frac{X - X_{min}}{X_{máx} - X_{min}} \quad (7)$$

onde:

- $X_{saída}$ = valor gerado após a normalização dentro do intervalo 0 a 1;
- X = valor do banco de dados a ser normalizado;
- $X_{máx}$ = maior valor do intervalo de dados ainda não normalizados;
- X_{min} = menor valor do intervalo de dados ainda não normalizados;

No processo de normalização das variáveis quantitativas foi adotada uma precisão de 4 casas decimais. Dito isto, todos os valores normalizados da variável “idade do condutor” e da variável “quilômetro” podem ser vistos nas **Tabelas 22 e 23** do **APÊNDICE**.

5.3.4 Treinamento das RNA's (Fase 4)

Após a normalização dos parâmetros, foi utilizado o software *Weka 3.8.6* para a implementação da base de dados no *software*, dando início aos processos de especificação e de treinamento das redes neurais artificiais. Dessa forma, foram importados todos os parâmetros, sendo selecionado como parâmetro de saída, o quilômetro.

É importante ressaltar que os dados de entrada da rede foram os parâmetros qualitativos e a idade do condutor, sendo os dados de saída que a rede fornecerá os quilômetros, ou seja, a predição dos Km's da área de estudo onde a frequência de sinistros de trânsito são maiores.

Para o processo de especificação das arquiteturas das redes neurais artificiais foi utilizado à variação dos parâmetros, conforme mostra a **Tabela 6**:

Tabela 6 – Parâmetros utilizados na arquitetura das redes

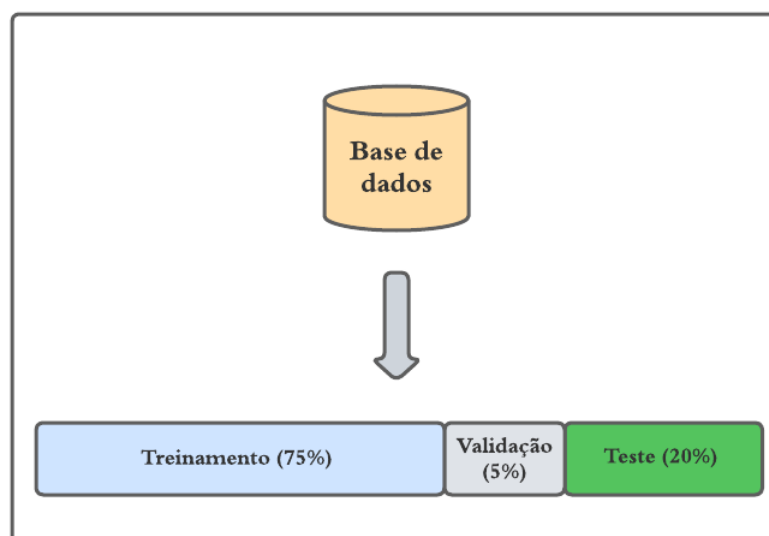
Tipo de rede	Nº de neurônios e Nº de camadas ocultas	Algoritmo de treinamento	Função de ativação	Taxa de aprendizado e Momentum	Número de Épocas	Critério de generalização	Critério de parada
<i>Multilayer Perceptron</i> (MLP)	Varia de 5 a 40 neurônios e varia de 1 a 4 camadas	<i>Backpropagation</i>	Sigmóide Logística	Varia de 0,001 a 1	Varia de 500 a 10000	Validação cruzada por amostragem aleatória	Parada antecipada (<i>Early stopping</i>)

Fonte: Adaptado de Silva (2019).

A definição da variação da taxa de aprendizado e de *Momentum* foi baseada no estudo de Binoti, Binoti e Leite (2014), enquanto que os números de neurônios e camadas ocultas foram baseados no estudo de Oliveira (2021). O critério de parada antecipada foi inspirado no critério de convergência usado por Silva (2019), em que consistiu em definir o número máximo de ciclos como 10.000 (iniciando em 500 épocas e incrementando 1000 épocas a cada novo treino) ou a média do REQM da validação igual ou menor que 1×10^{-3} entre iterações, finalizando o treinamento quando um dos dois critérios foi atingido. Por fim, o critério de generalização foi determinado conforme os autores Silva, Spatti e Flauzino (2016), e levando em consideração os estudos de Souza (2015), Reis (2018), Wolfrann et al. (2018).

A variação dos parâmetros descritos na **Tabela 6** foi utilizada na especificação de 10 RNA's, cada uma com uma arquitetura distinta, visando encontrar aquela que alcançasse melhor desempenho na predição a partir do subconjunto de validação e aplicando o método de validação cruzada, para, posteriormente, utilizar esta RNA de melhor desempenho no subconjunto de teste.

Todo o conjunto de dados foi subdividido em três partes, conforme a **Figura 8**:

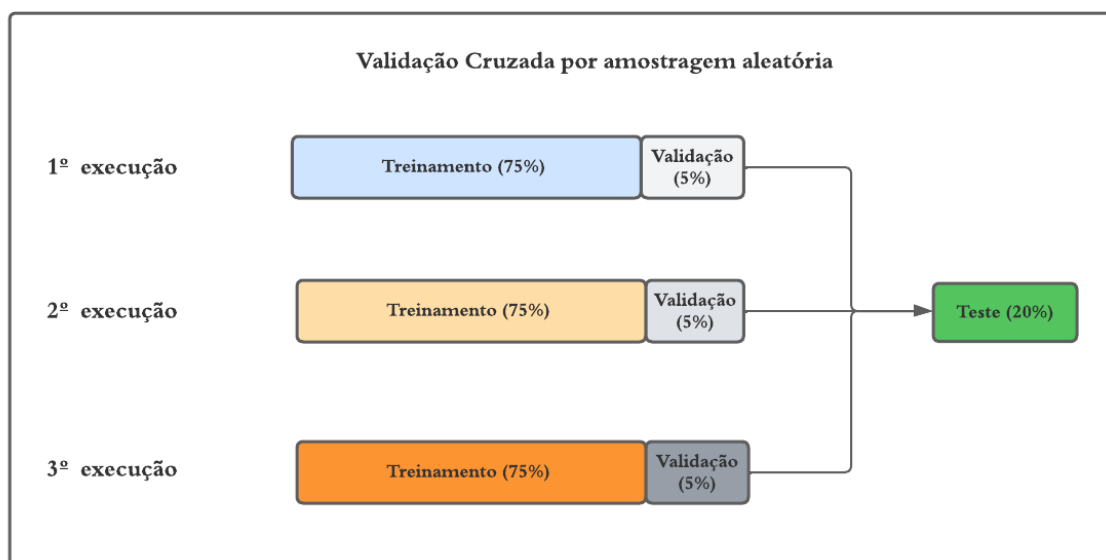
Figura 8 – Divisão do conjunto de dados

Fonte: Adaptado de Silva (2019).

Foi trabalhado com um total de 2025 dados de sinistros de trânsito ocorridos na área de estudo ao longo dos três anos da janela temporal (2017 – 2019), no qual, 405 dados (20%) foram separados para compor o exclusivo subconjunto de teste, enquanto que 1520 dados (75%) para compor os subconjuntos de treinamentos, e 100 dados (5%) para compor os subconjuntos de validação. É importante salientar que toda a divisão do conjunto de dados foi feita de forma aleatória através do *Excel 2007*.

Para a análise de assertividade foi empregado o método de validação cruzada por amostragem aleatória (*random subsampling cross-validation*), em que o subconjunto de validação foi utilizado como critério para a avaliação da capacidade de generalização. Ao executar a validação cruzada, foi permutado o subconjunto de treinamento e o de validação três vezes, ou seja, uma vez separado e isolado o subconjunto exclusivo de teste que equivale a 20% de toda a base de dados, os outros 80% foram aleatoriamente divididos três vezes utilizando o *Excel 2007*, para dessa forma, compor três subconjuntos distintos de treinamento (75%) e três subconjuntos distintos de validação (5%), conforme mostra a **Figura 9**:

Figura 9 – Validação cruzada com três subconjuntos de treinamento e três subconjuntos de validação



Fonte: Adaptado de Silva (2019).

Após o ajuste dos hiperparâmetros, treinada cada RNA com os três subconjuntos de treinamento e aplicado a regressão nos três respectivos subconjuntos de validação, foi calculada a média global de cada correlação de treino obtida e a média do EAM e REQM de cada validação, compilando os resultados. Dois métodos foram empregados para análise de assertividade entre os dados observados e os preditos:

- **Primeiro método – Análise por margens de erros**

O primeiro método consistiu em analisar toda a área de estudo considerando seis margens de erros:

- 5 km;
- 10 km;
- 15 km;
- 20 km;
- 25 km;
- 30 km;

Uma vez estabelecido o tamanho de cada uma das seis margens, foi analisada a assertividade de cada rede neural quanto a estes valores propostos. Para a primeira análise da assertividade de cada rede neural, foi utilizada a **Equação 8**:

$$A = |\text{Valor alvo} - \text{Valor predito}| \quad (8)$$

O valor alvo mostrado na **Equação 8** consiste nos valores da variável quantitativa de saída normalizados pela **Equação 7**, ou seja, são os valores em quilômetros variando no intervalo de 0 até 1. Quanto ao valor predito, nada mais é do que a predição fornecida pela rede neural em relação ao valor alvo. Para a efetuação dos cálculos da **Equação 8** tanto os valores alvo como os preditos foram transformados para quilômetro novamente, a partir da aplicação inversa da **Equação 7**, utilizando o *Excel 2007*.

Portanto, para a análise da margem de erro de 5 km se a diferença em módulo do valor predito em relação ao valor alvo estivesse dentro desta margem (para mais ou para menos) foi computado como um acerto da rede neural, o mesmo foi feito para as margens de erro de 10 km, de 15 km, de 20 km, de 25 km e de 30 km. Assim, uma vez efetuada a análise para cada valor de margem de erro e para cada um dos três subconjuntos de validação, foram contabilizados os acertos dentro de cada amostra, obtendo a porcentagem de acerto para cada subconjunto e feito a média global dessas porcentagens.

- **Segundo método – Análise por trechos fixados**

O segundo método consistiu dessa vez em segmentar toda a área de estudo a cada:

- 5 km;
- 10 km;
- 15 km;
- 20 km;
- 25 km;
- 30 km;

Deste modo, semelhante ao primeiro método, os valores preditos em relação aos valores alvos foram analisados de forma que se o valor predito se encontrasse exatamente dentro do mesmo intervalo que o valor alvo, seria computado como um acerto da rede neural, isto é, levando em consideração o estudo para cada tamanho de extensão proposta. Assim

como no primeiro método, uma vez efetuada a análise para cada um dos 6 tamanhos de extensão de trechos fixados e para cada um dos três subconjuntos de validação, foram contabilizados os acertos dentro de cada amostra, obtendo a porcentagem de acerto para cada subconjunto e feito a média global dessas porcentagens.

Como saída, o software *Weka* disponibiliza além dos valores preditos, importantes métricas estatísticas para avaliar o desempenho do modelo como o índice de correlação de Pearson (r), o EAM e o REQM.

Para selecionar qual foi a melhor topologia, foi observado qual delas teve melhor desempenho frente a dados novos, ou seja, melhor capacidade de generalização nos subconjuntos de validação, sendo levado em consideração três critérios em ordem decrescente de importância, no qual, foi selecionada a RNA que:

- Possuir melhor média global de acurácia (taxa de assertividade) para cada método de análise (Margem de erro e Trechos fixados) e para cada tamanho de extensão de quilometragem;
- Possuir menor média global do EAM e do REQM, no caso de empate no critério descrito acima, e;
- Possuir melhor índice de classificação da média da correlação de Pearson (r) dos subconjuntos de treinamento, caso haja empate nos critérios descritos anteriormente;

Cada média do coeficiente de correlação linear de Pearson dos subconjuntos de treinamento foi classificada conforme a **Tabela 7**:

Tabela 7 – Classificação de Hopkins (2000)

Coefficiente de correlação (r)	Descrição
0.0 - 0.1	Muito Baixa
0.1 - 0.3	Baixa
0.3 - 0.5	Moderada
0.5 - 0.7	Alta
0.7 - 0.9	Muito Alta
0.9 - 1	Quase ou Praticamente Perfeita

Fonte: Adaptado de Hopkins (2000).

Selecionada a arquitetura de RNA que teve melhor capacidade generalização, este modelo de RNA foi empregada para a análise de acidentalidade viária agora no subconjunto de teste, em que, semelhante aos subconjuntos de validação, foi aplicado novamente os mesmos dois métodos nesse novo subconjunto de dados, ou seja, por margens de erros e por trechos fixados. Além disso, para essa amostra exclusiva de teste foi empregado uma terceira forma de análise, sendo o foco de estudo trechos pertencentes entre seis municípios da área de estudo, sendo estes trechos constituídos de diferentes extensões de tamanho entre eles.

- **Terceiro método – Análise entre municípios**

A assertividade neste terceiro e último método foi considerada de modo que se o valor predito se encontrar no mesmo trecho entre municípios que o valor alvo foi computado como um acerto da rede neural. Os municípios foram escolhidos com base no número populacional, visto que se trata de um fator relevante e que interfere na acidentalidade viária, como explicitado anteriormente. A **Tabela 8** traz os municípios e os tamanhos dos trechos:

Tabela 8 – Subdivisão da área de estudo em trechos entre municípios

Município	População (hab)	Localização na BR-365 (km)	Distância entre municípios (km)
Patos de Minas	159.235	420,1	50
Patrocínio	89.826	470,1	53,2
Monte Carmelo	47.689	523,3	84,7
Uberlândia	713.232	608	47,8
Monte Alegre de Minas	20.170	655,8	97,3
Ituiutaba	102.217	753,1	-

Fonte: Autoria própria (2023).

6 RESULTADOS E DISCUSSÕES

Todas as 10 RNA's treinadas são apresentadas na **Tabela 9**, onde primeiramente é informado o número de tentativa da rede em questão, em seguida é especificada a topologia de cada rede informando, respectivamente, o número de entradas, o número de camadas e neurônios ocultos e o número de neurônios de saída. Por fim, é informada a função de ativação utilizada nas camadas ocultas.

Tabela 9 – Topologias das RNA's

Rede	Número de entradas	Número de Neurônios					Função de ativação				
		1ª camada	2ª camada	3ª camada	4ª camada	5ª camada	1ª camada	2ª camada	3ª camada	4ª camada	5ª camada
1	8	15	15	5	1	-	LOGSIG	LOGSIG	LOGSIG	-	-
2	8	30	15	5	1	-	LOGSIG	LOGSIG	LOGSIG	-	-
3	8	10	10	40	1	-	LOGSIG	LOGSIG	LOGSIG	-	-
4	8	10	15	10	15	1	LOGSIG	LOGSIG	LOGSIG	LOGSIG	-
5	8	10	40	1	-	-	LOGSIG	LOGSIG	-	-	-
6	8	20	25	25	20	1	LOGSIG	LOGSIG	LOGSIG	LOGSIG	-
7	8	20	30	1	-	-	LOGSIG	LOGSIG	-	-	-
8	8	30	1	-	-	-	LOGSIG	-	-	-	-
9	8	40	1	-	-	-	LOGSIG	-	-	-	-
10	8	40	40	1	-	-	LOGSIG	LOGSIG	-	-	-

Fonte: Adaptado de Oliveira (2021).

Conforme mostra a **Tabela 9** foi aplicado a mesma quantidade de entradas para cada topologia de rede, no qual a variação do número de camadas ocultas variou de 1 a 4 camadas.

Efetuada a média global das correlações obtidas nos três treinamentos, da porcentagem de acerto de cada rede neural em ambos os métodos e das métricas estatísticas (EAM e REQM) nos três subconjuntos de validação, foi obtido os resultados do primeiro método na **Tabela 10**:

Tabela 10 – Resultados do método “Margens de erro” da validação

Rede	Correlação de treino	Classificação da correlação	EAM da validação	REQM da validação	Porcentagem de acerto (%)					
					5 km	10 km	15 km	20 km	25 km	30 km
1	0,531	ALTA	0,2612	0,3165	5,3	13	19,6	23	26,3	29,3
2	0,592	ALTA	0,2656	0,3319	4,6	12,3	16,3	20	22,6	26
3	0,4388	MODERADA	0,2455	0,3011	5	10,3	13,3	17,6	23,3	28
4	0,4136	MODERADA	0,2322	0,2891	5,6	11,6	19	25	28,3	32,3
5	0,4968	MODERADA	0,2515	0,308	4,3	9	13	16,6	20,6	24,3
6	0,6087	ALTA	0,2624	0,3297	5,6	9,6	14	17	22,6	27,3
7	0,6477	ALTA	0,2788	0,3482	6	11	15,6	18,3	21,6	23,6
8	0,5142	ALTA	0,267	0,3682	5,3	13	20	24,6	28	30,3
9	0,528	ALTA	0,2705	0,369	6,6	13,3	19	25	28,6	31
10	0,7154	MUITO ALTA	0,2933	0,3848	5	8,3	14	17	22,3	23,6

Fonte: Autoria própria (2023).

Em sua grande maioria as redes treinadas alcançaram uma média acima de 0,50 de índice de correlação de treino como mostra a **Tabela 10**, sendo classificadas como uma correlação “ALTA” conforme a metodologia de Hopkins (2000) descrita na **Tabela 7**. Assim, após breve análise das 10 RNA’s foram selecionadas quais foram as cinco redes que mais se destacaram na média da porcentagem de acerto nos subconjuntos de validação em cada extensão de quilometragem, conforme a **Tabela 11** a seguir:

Tabela 11 – Cinco melhores redes pelo método “Margens de erro” da validação

Para 5 km			
REDE NEURAL	9	% de acerto	6,6
REDE NEURAL	7	% de acerto	6
REDE NEURAL	4	% de acerto	5,6
REDE NEURAL	6	% de acerto	5,6
REDE NEURAL	1	% de acerto	5,3
Para 10 km			
REDE NEURAL	9	% de acerto	13,3
REDE NEURAL	1	% de acerto	13
REDE NEURAL	8	% de acerto	13
REDE NEURAL	2	% de acerto	12,3
REDE NEURAL	4	% de acerto	11,6
Para 15 km			
REDE NEURAL	8	% de acerto	20
REDE NEURAL	1	% de acerto	19,6
REDE NEURAL	4	% de acerto	19
REDE NEURAL	9	% de acerto	19
REDE NEURAL	2	% de acerto	16,3
Para 20 km			
REDE NEURAL	9	% de acerto	25
REDE NEURAL	4	% de acerto	24,6
REDE NEURAL	8	% de acerto	23
REDE NEURAL	1	% de acerto	22
REDE NEURAL	2	% de acerto	20,3
Para 25 km			
REDE NEURAL	9	% de acerto	28,6
REDE NEURAL	4	% de acerto	28,3
REDE NEURAL	8	% de acerto	28
REDE NEURAL	1	% de acerto	26,3
REDE NEURAL	3	% de acerto	23,3
Para 30 km			
REDE NEURAL	4	% de acerto	32,3
REDE NEURAL	9	% de acerto	31
REDE NEURAL	8	% de acerto	30,3
REDE NEURAL	1	% de acerto	29,3
REDE NEURAL	3	% de acerto	28

Fonte: Aatoria própria (2023).

De acordo com a **Tabela 11**, três topologias se destacam quando observado os seis tamanhos de extensão, aparecendo como as três melhores em três situações. Entretanto, apenas uma arquitetura de fato se sobressaiu sendo a melhor em quatro ocasiões, como mostra a **Tabela 12**:

Tabela 12 – Melhor topologia pelo método “Margens de erro” da validação

Rede Neural	Correlação de treino	Frequência	Melhor
9	0,528	6	4
4	0,4136	6	1
8	0,5142	5	1

Fonte: Autoria própria (2023).

Quando observado o segundo método aplicado aos subconjuntos de validação (por “Trechos fixados”), foi obtido os seguintes resultados nas porcentagens de assertividade conforme mostra a **Tabela 13**:

Tabela 13 – Resultados do método “Trechos fixados” da validação

Rede	Correlação de treino	Classificação da correlação	EAM da validação	REQM da validação	Porcentagem de acerto (%)					
					5 km	10 km	15 km	20 km	25 km	30 km
1	0,531	ALTA	0,2612	0,3165	2	4,6	8,3	11,3	13,6	18
2	0,592	ALTA	0,2656	0,3319	3	5,3	9,3	10,6	12	16,6
3	0,4388	MODERADA	0,2455	0,3011	2,3	4,3	7,3	8,3	9,6	17
4	0,4136	MODERADA	0,2322	0,2891	4,3	6,3	9	10,6	12,6	20,3
5	0,4968	MODERADA	0,2515	0,308	2,3	3	7,6	6,3	11,3	15
6	0,6087	ALTA	0,2624	0,3297	3	4	8,6	9,3	11	16,3
7	0,6477	ALTA	0,2788	0,3482	4	5	9,3	8,3	10,3	12,6
8	0,5142	ALTA	0,267	0,3682	1,3	6	8,6	12,3	12,6	21,6
9	0,528	ALTA	0,2705	0,369	2,6	7,6	10	10	12,3	21,6
10	0,7154	MUITO ALTA	0,2933	0,3848	4,6	5,6	6,3	8,6	11	13,6

Fonte: Autoria própria (2023).

Assim como no primeiro método foram selecionadas quais foram as cinco redes que mais se destacaram na média da porcentagem de acerto nos subconjuntos de validação em cada extensão de quilometragem, conforme a **Tabela 14**:

Tabela 14 – Cinco melhores redes pelo método “Trechos fixados” da validação

Para 5 km			
REDE NEURAL	10	% de acerto	4,6
REDE NEURAL	4	% de acerto	4,3
REDE NEURAL	7	% de acerto	4
REDE NEURAL	2	% de acerto	3
REDE NEURAL	6	% de acerto	3
Para 10 km			
REDE NEURAL	9	% de acerto	7,6
REDE NEURAL	4	% de acerto	6,3
REDE NEURAL	8	% de acerto	6
REDE NEURAL	10	% de acerto	5,6
REDE NEURAL	2	% de acerto	5,3
Para 15 km			
REDE NEURAL	9	% de acerto	10
REDE NEURAL	2	% de acerto	9,3
REDE NEURAL	7	% de acerto	9,3
REDE NEURAL	4	% de acerto	9
REDE NEURAL	6	% de acerto	8,6
Para 20 km			
REDE NEURAL	8	% de acerto	12,3
REDE NEURAL	1	% de acerto	11,3
REDE NEURAL	2	% de acerto	10,6
REDE NEURAL	4	% de acerto	10
REDE NEURAL	9	% de acerto	10
Para 25 km			
REDE NEURAL	1	% de acerto	13,6
REDE NEURAL	4	% de acerto	12,6
REDE NEURAL	8	% de acerto	12,6
REDE NEURAL	9	% de acerto	12,3
REDE NEURAL	2	% de acerto	12
Para 30 km			
REDE NEURAL	9	% de acerto	21,6
REDE NEURAL	8	% de acerto	21,6
REDE NEURAL	4	% de acerto	18,6
REDE NEURAL	1	% de acerto	18
REDE NEURAL	3	% de acerto	17

Fonte: Autoria própria (2023).

As três topologias que se destacaram no segundo método são apresentadas na **Tabela 15**:

Tabela 15 – Melhor topologia pelo método “Trechos fixados”

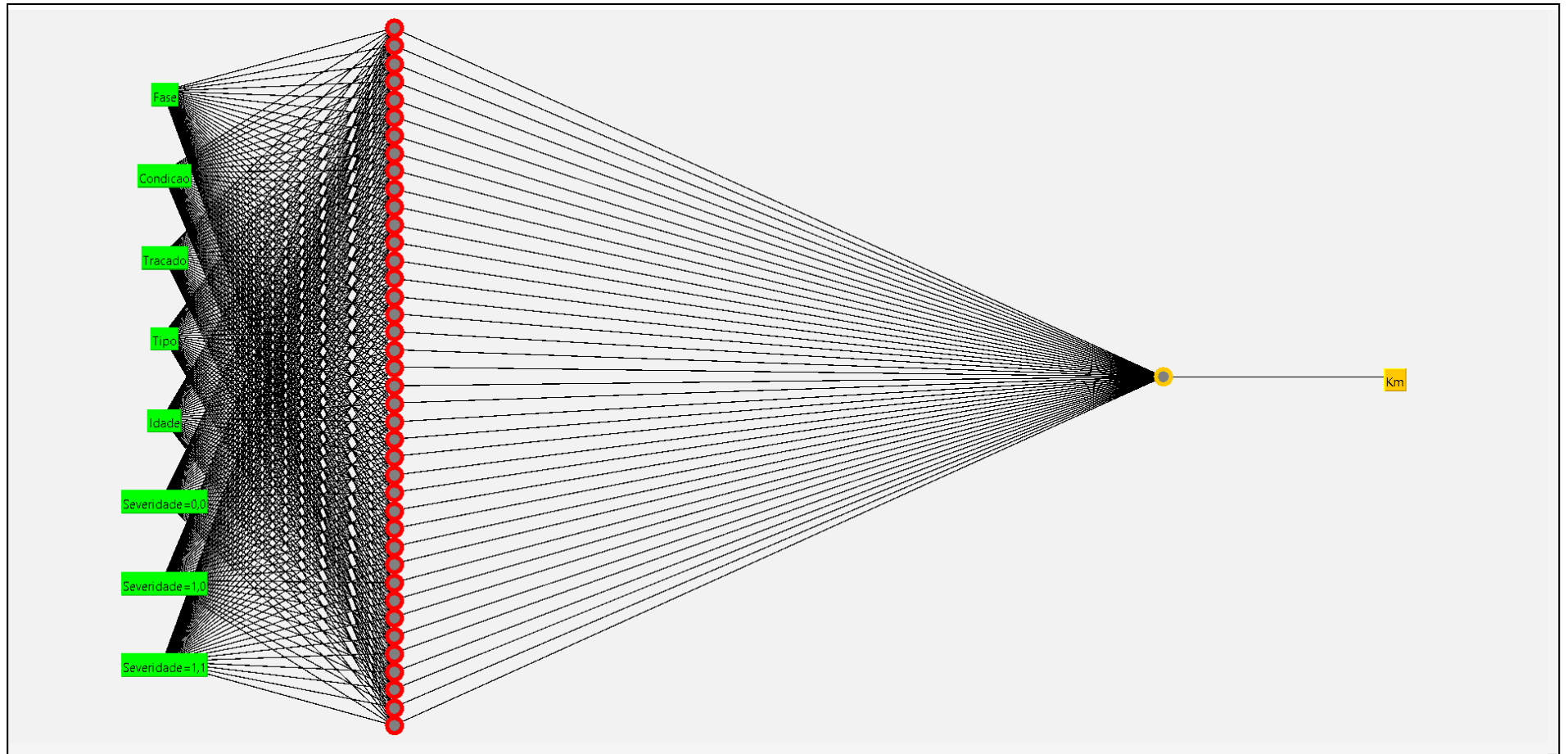
Rede Neural	Correlação de treino	Frequência	Melhor
9	0,528	5	3
8	0,5142	4	2
1	0,531	3	1

Fonte: Autoria própria (2023).

Logo, mediante os resultados obtidos a partir da análise da generalização das RNA's frente á validação, a rede neural 9 (8-40-1) composta por 8 entradas, uma camada oculta contendo 40 neurônios ocultos e 1 neurônio na camada de saída, foi a que mais se destacou em ambos os métodos, no qual, por “Margens de erro” aparece 4 vezes como a arquitetura de melhor desempenho, já por “Trechos fixados” essa mesma rede aparece como melhor em 3 ocasiões, sendo elas: com 7,6% para 10 km, 10% para 15 km e 21,6% para 30 km.

É importante salientar que apesar de outras redes terem uma classificação de correlação de Pearson mais qualificada, como a rede neural 10, cuja correlação foi de 0,7154 considerada como “MUITO ALTA” por Hopkins (2000), foi respeitado os critérios estabelecidos anteriormente cuja ordem de importância de forma decrescente se iniciava com a análise da acurácia (porcentagem de acerto) para cada extensão de quilômetro e para cada método aplicado. Dessa mesma forma vale a justificativa para o segundo critério, em que apesar de outras redes também possuírem índices menores de EAM e de REQM como a rede 4, foi a rede neural 9 a selecionada justamente devido a sua capacidade de assertividade frente a dados não vistos anteriormente na etapa de treinamento. Assim, a **Figura 10** evidencia a configuração da arquitetura da RNA selecionada para a análise final, a do subconjunto de teste.

Figura 10 – Configuração da arquitetura da rede neural 9



Fonte: Software *Weka* (2023).

Além disso, a **Tabela 16** apresenta os parâmetros e os resultados obtidos com a configuração topológica da rede neural 9 replicada em cada execução da validação cruzada:

Tabela 16 – Resumo dos parâmetros empregados

Rede	Execução	I	TA	M	NO	O	Ep	F	(r)	EAM (v)	REQM (v)	TP (s)
MLP 8-40-1	1	8	0,09	0,095	40	1	10000	Logsig	0,5265	0,2644	0,3452	56,58
MLP 8-40-1	2	8	0,09	0,095	40	1	10000	Logsig	0,5293	0,2528	0,3465	55,14
MLP 8-40-1	3	8	0,09	0,095	40	1	10000	Logsig	0,5283	0,2945	0,4154	55,31

Legenda

I - Inputs (entradas) **TA** - Taxa de aprendizado **M** - Momentum **NO** - Neurônios ocultos **O** - Outputs (saídas) **Ep** - Épocas **F** - Função de ativação **(r)** - Correlação de treino **EAM** - Erro absoluto médio da validação **REQM** - Raiz do erro quadrático médio da validação **TP (s)** - Tempo de processamento

Fonte: Adaptado de Silva, Andrade, Weigang (2019).

A **Tabela 17** mostra a compilação do desempenho da rede selecionada obtidos na validação em ambos os métodos aplicados:

Tabela 17 – Resultado médio da arquitetura 9

Resultado médio	
REDE NEURAL	9
CORRELAÇÃO de TREINO	0,528
CLASSIFICAÇÃO DA CORRELAÇÃO	ALTA
EAM da VALIDAÇÃO	0,2705
REQM da VALIDAÇÃO	0,369
Acertos por margens de erro (%)	
5 km	6,6
10 km	13,3
15 km	19
20 km	25
25 km	28,6
30 km	31
Acertos por trechos fixados (%)	
5 km	2,6
10 km	7,6
15 km	10
20 km	10
25 km	12,3
30 km	21,6

Fonte: Autoria própria (2023).

Selecionada a melhor arquitetura de RNA, foi empregada esta rede para a análise do subconjunto de teste, que foi totalmente separado dos outros subconjuntos e distintos dos dados da validação, onde, foram aplicados os mesmos métodos de análise vistos anteriormente com o acréscimo do terceiro método “Análise entre municípios”. Assim, diferentemente dos subconjuntos de validação em que cada um dos três eram constituídos por dados distintos onde foi feita a média das porcentagens, ou seja, do desempenho, no subconjunto de teste foi efetuado a média das predições de cada valor alvo das três execuções da rede neural 9, uma vez que o subconjunto de teste foi repetido nas três execuções conforme demonstra a **Figura 9**. Os valores alvos e os valores preditos do subconjunto de teste podem ser vistos na **Tabela 24** do **APÊNDICE**. Deste modo, a **Tabela 18** mostra o resultado médio da rede neural 9 a partir dos métodos “Margens de erro” e “Trechos fixados” no subconjunto de teste:

Tabela 18 – Resultados médio da rede neural 9 com o subconjunto de teste

Resultado médio	
REDE NEURAL	9
CORRELAÇÃO de TREINO	0,528
CLASSIFICAÇÃO DA CORRELAÇÃO	ALTA
EAM do TESTE	0,2913
REQM do TESTE	0,3994
Acertos por margens de erro (%)	
5 km	4,9
10 km	12,6
15 km	17,2
20 km	20,7
25 km	24,2
30 km	26,6
Acertos por trechos fixados (%)	
5 km	2,7
10 km	4,9
15 km	9,6
20 km	9,3
25 km	10,3
30 km	17,2

Fonte: Autoria própria (2023).

Nota-se que ao comparar a **Tabela 17** e **Tabela 18**, os valores obtidos de EAM e de REQM nos subconjuntos de validação são relativamente menores que o do subconjunto de teste, isto, devido ao fato que este último se trata de um subconjunto com um volume maior de dados (405) em relação ao primeiro (100 dados). Toda essa conjuntura evidencia o que já era pressuposto por Silva (1998) em seus estudos, no qual, a validação pode ser utilizada como uma espécie de estimativa do teste.

Já na análise do terceiro método aplicado, observando os trechos contidos entre os seis municípios distribuídos ao longo da BR-365, os resultados em porcentagem são mostrados na **Tabela 19**:

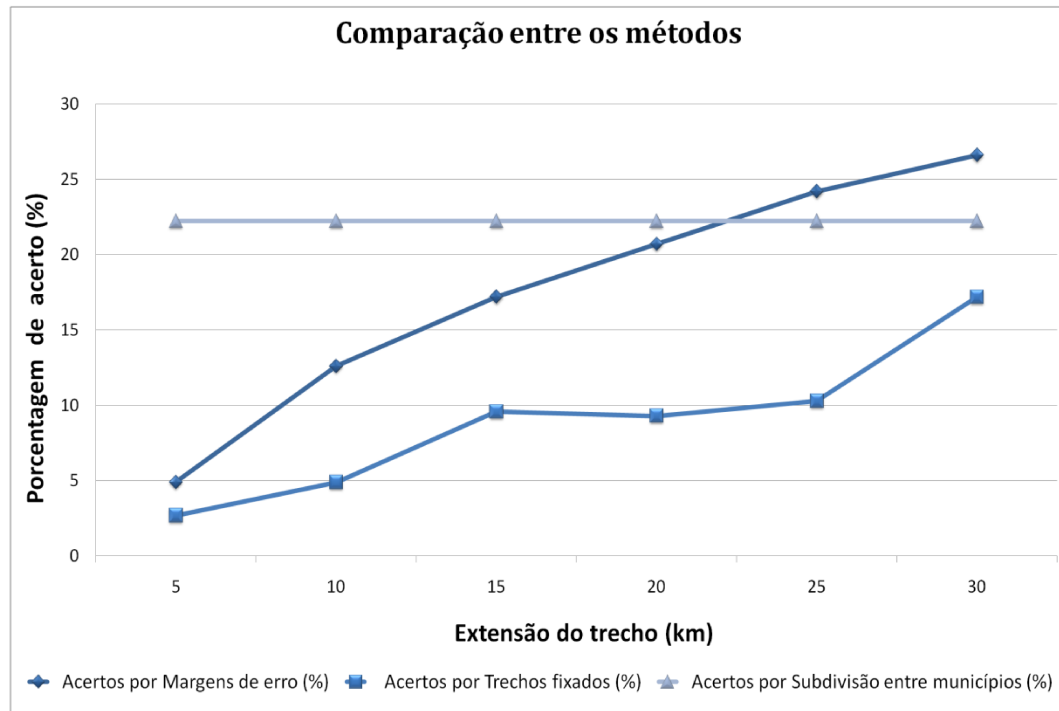
Tabela 19 – Resultados da rede neural 9 na análise de trechos entre municípios

Trecho	Acertos por trecho (%)	Comprimento do trecho (Km)
Patos de Minas - Patrocínio	0,24	50
Patrocínio - Monte Carmelo	1,48	53,2
Monte Carmelo - Uberlândia	4,69	84,7
Uberlândia - Monte Alegre de Minas	14,32	47,8
Monte Alegre de Minas - Ituiutaba	1,48	97,3
Acertos totais (%)	22,21	

Fonte: Autoria própria (2023).

No **Tabela 19** é notório a maior frequência de sinistros de trânsito no trecho pertencente entre o município Monte Alegre de Minas e o município de Uberlândia, com uma alta frequência de 14,32%. Tal conjuntura pode ser explicada em virtude de Uberlândia se tratar do município com o maior número populacional da região, contabilizando 713.232 habitantes (IBGE, 2023). Outro fator relevante é que em sua totalidade o trecho pertencente entre o Km 608 e o Km 655,8 é constituído de pista dupla, o que consiste numa maior capacidade de fluxo de veículos simultaneamente, o que diferencia este trecho dos outros existentes entre os municípios escolhidos.

No **Gráfico 4** é possível observar o desempenho da rede 9 quando comparado os três métodos aplicados:

Gráfico 4 – Comparação entre os métodos

Fonte: Autoria própria (2023).

Ao analisar a área de estudo com os três diferentes métodos de análise é possível perceber que pelo método “Margens de erro” a rede neural 9 obteve uma taxa maior de assertividade quando comparada aos outros métodos, alcançando 26,6% de acerto no subconjunto de teste, isto, quando observada a área de estudo em margens de erro de 30 km em relação ao valor alvo.

Já pelo segundo método, por “Trechos fixados”, por se tratar de uma análise mais específica que o primeiro método, a rede obteve uma acurácia de 17,2% para a segmentação da área de estudo a cada 30 km.

O terceiro método, entretanto, apesar de não ter alcançado uma assertividade maior que por “Margens de erro”, atingiu um valor próximo, cuja taxa de acerto foi de 22,21%. À primeira vista, apesar de próximos os resultados obtidos, ambos os valores constituem numa acurácia ainda relativamente baixa. No entanto, estes resultados também evidenciam a especificidade do problema tratado neste trabalho, no qual, prever corretamente os quilômetros com maiores probabilidades de ocorrência de acidentalidade, a partir de poucas variáveis de entrada empregadas, ainda constitui numa tarefa de grande complexidade para um modelo de RNA.

Não obstante, no último método de análise aplicado é nítido que a área de estudo pode ser melhor examinada visto que as frequências de acidentalidade viária são visualizadas de acordo com a sua localização entre os municípios, fornecendo um olhar mais detalhado de quais trechos são os mais críticos. Além disso, é importante salientar que mesmo possuindo diferentes valores de frequência em cada trecho, o terceiro método consiste na soma de todas estas porcentagens de frequência, totalizando em apenas um valor conforme mostra a **Tabela 19**.

Quando observados a **Tabela 18** e o **Gráfico 4** nota-se que se tratando de segmentos menores como de 5 km, o desempenho da rede neural nos dois primeiros métodos é menor, isso devido a rigorosidade requerida em ambas as análises, visto que quando observada a assertividade por “Margens de erros” é muito mais provável a ocorrência do acerto, já que consiste em um método mais amplo em relação ao segundo. Além disso, é perceptível que o acerto da rede cresce em ambas as análises concomitantemente ao aumento do segmento de trecho estudado, denotando a influência do tamanho dos segmentos quanto aos ajustes dos dados e na obtenção de melhores desempenhos.

Os resultados obtidos no presente trabalho nas três formas de análises convergem aos verificados por Oliveira (2021), onde, a maior porcentagem de acerto obtida nos dois primeiros métodos foi uma assertividade de 18,3% e de 15%, respectivamente, isto, quando observado os resultados obtidos no segmento de 30 km para ambas as análises. Já para a frequência de acertos de trechos entre os municípios, Oliveira (2021) obteve um total de 26%, em que também foi trabalhado com trechos entre 6 municípios em sua forma final.

Contudo, diferentemente do ocorrido no estudo de Oliveira (2021), o desempenho da rede neural selecionada não alcançou a maior taxa de assertividade no terceiro método aplicado, isso pode ser justificado devido a uma falta de melhor ajuste da rede neural quanto aos dados na etapa de treinamento, o que prejudicou a obtenção de uma capacidade de generalização mais eficaz da RNA, e na extração de resultados mais aprimorados nos subconjuntos de validação e de teste. Além disso, o melhor desempenho conseguido a partir do método “Margens de erros” pode ser fundamentado no conjunto de dados trabalhado e na janela temporal, cuja influência na etapa de treinamento reside no grande volume de ocorrências viárias nos trechos de pista dupla, que constituem em quase metade dos sinistros de trânsitos estudados, conforme mostra a **Tabela 20**:

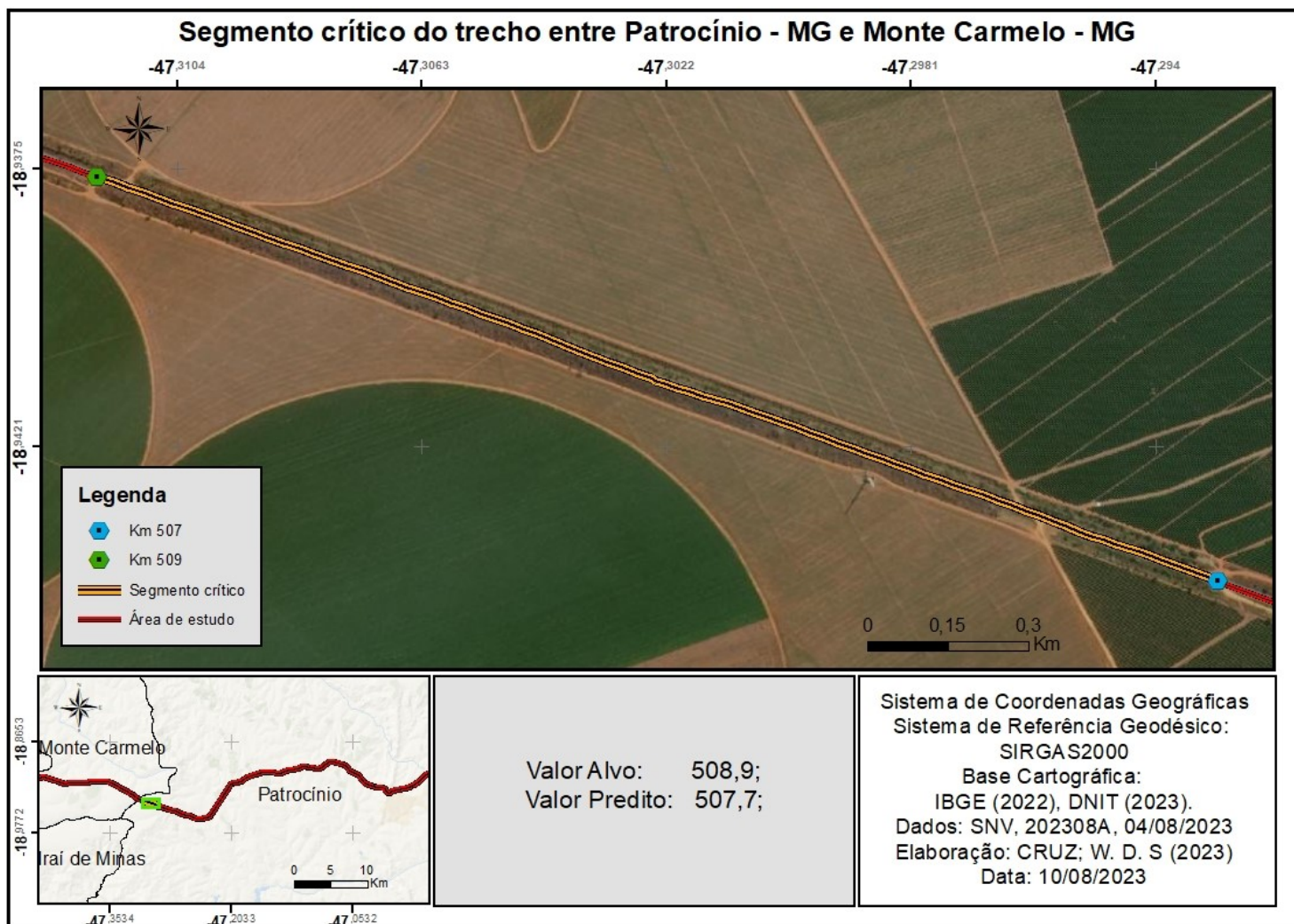
Tabela 20 – Número de sinistros de trânsito ocorridos ao longo de 2017 a 2019 do município de Patos de Minas - MG a Ituiutaba – MG

Trecho	Número de sinistros de trânsito			
	2017	2018	2019	Total
Patos de Minas - Patrocínio	47	40	30	117
Patrocínio - Monte Carmelo	51	41	40	132
Monte Carmelo - Uberlândia	49	61	68	178
Uberlândia - Monte Alegre de Minas	219	170	205	594
Monte Alegre de Minas - Ituiutaba	96	65	78	239

Fonte: Autoria própria (2023).

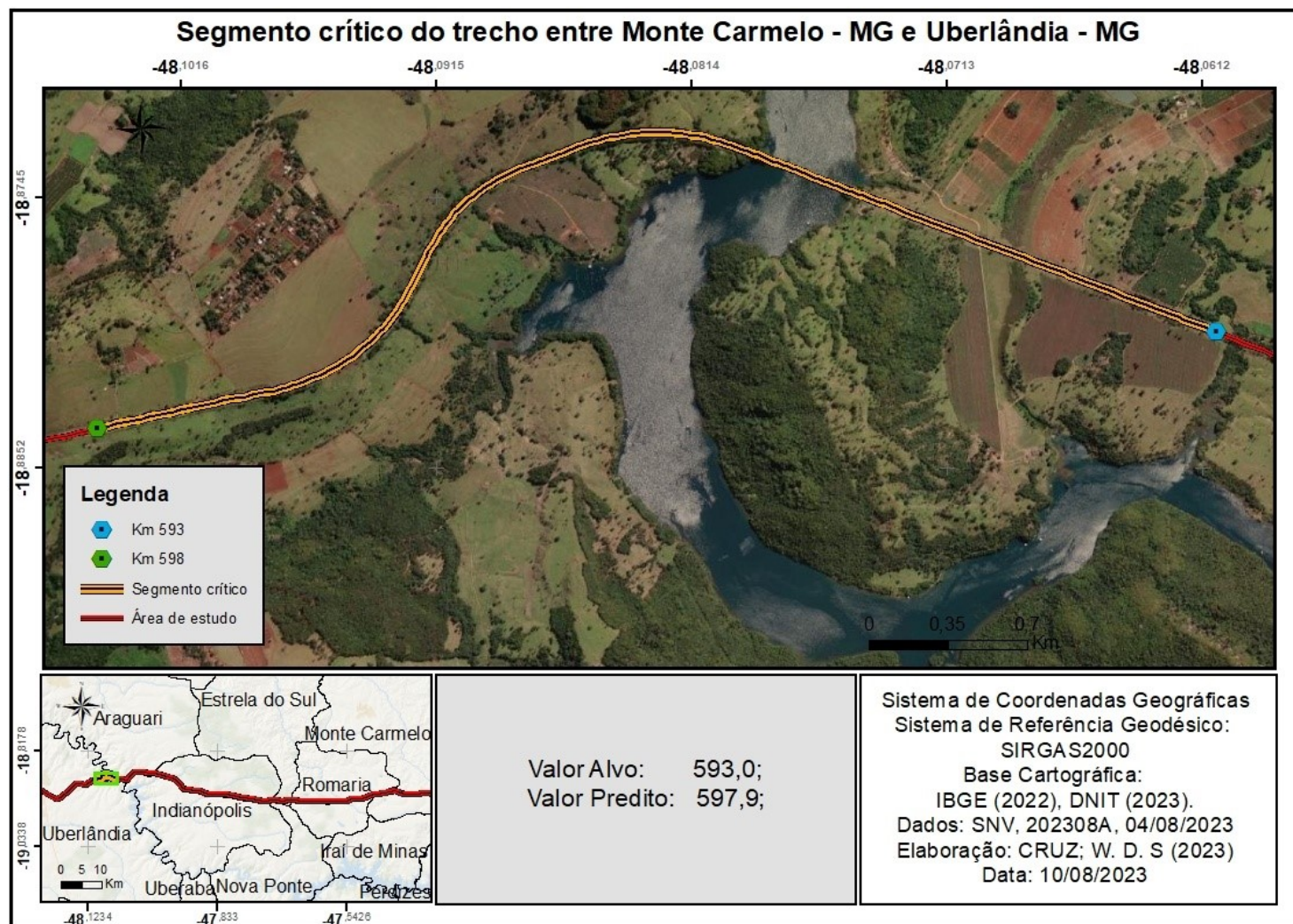
Portanto, de acordo com a **Tabela 20** é notória a quantidade expressiva de sinistros de trânsito ocorridos no trecho entre os municípios Monte Alegre de Minas e Uberlândia, convergindo com os valores de frequência de acidentalidade viária obtidos nas predições da rede neural 9.

Para melhor análise e visualização dos resultados obtidos no terceiro método aplicado, mapas foram elaborados visando identificar os trechos com os segmentos mais críticos de cada trecho pertencente entre os municípios, como mostra a **Figura 11, Figura 12, Figura 13, Figura 14 e Figura 15**.

Figura 11 – Segmento crítico do trecho entre Patrocínio – MG e Monte Carmelo - MG

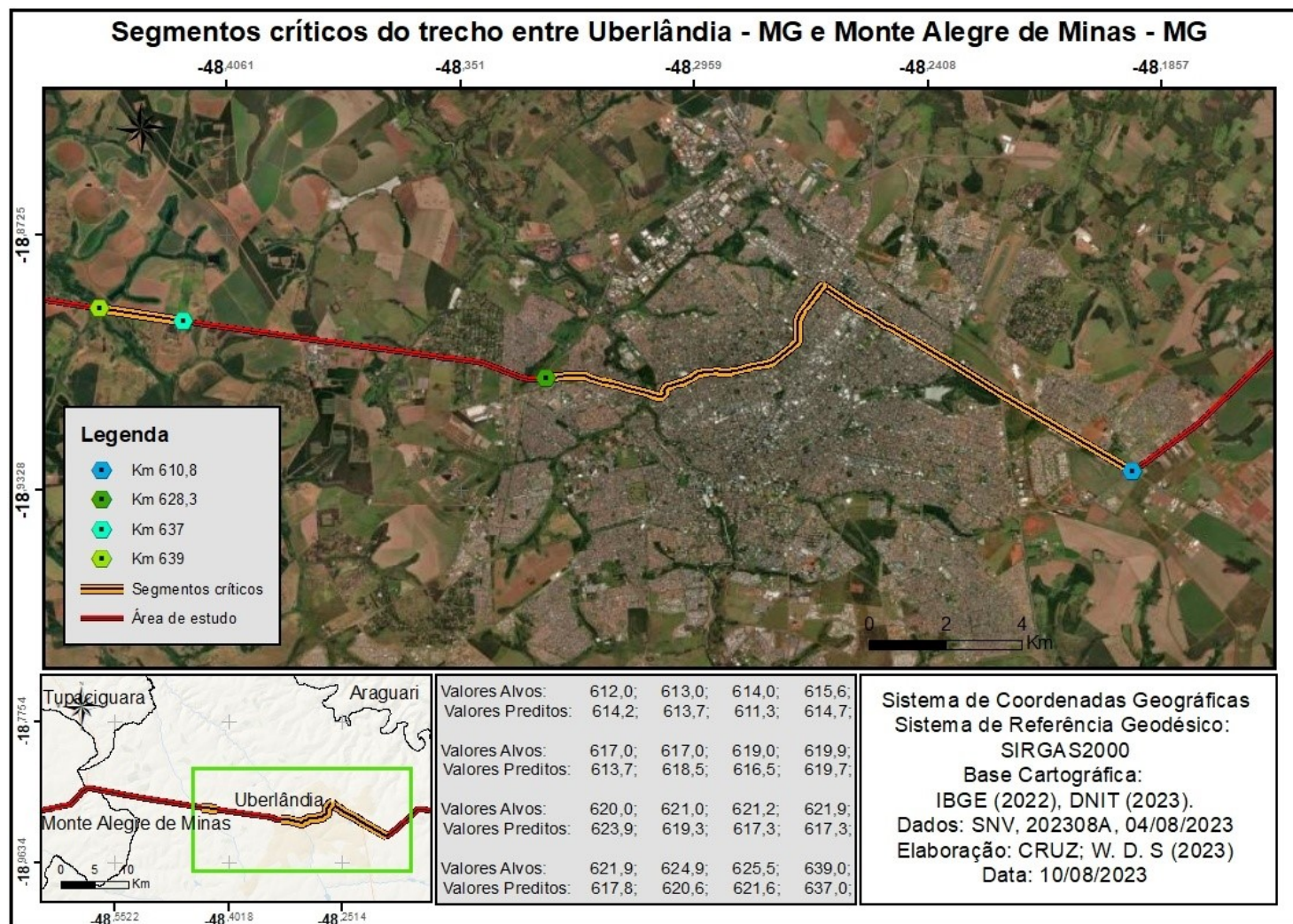
Fonte: Autoria própria (2023).

Figura 12 – Segmento crítico do trecho entre Monte Carmelo – MG e Uberlândia – MG



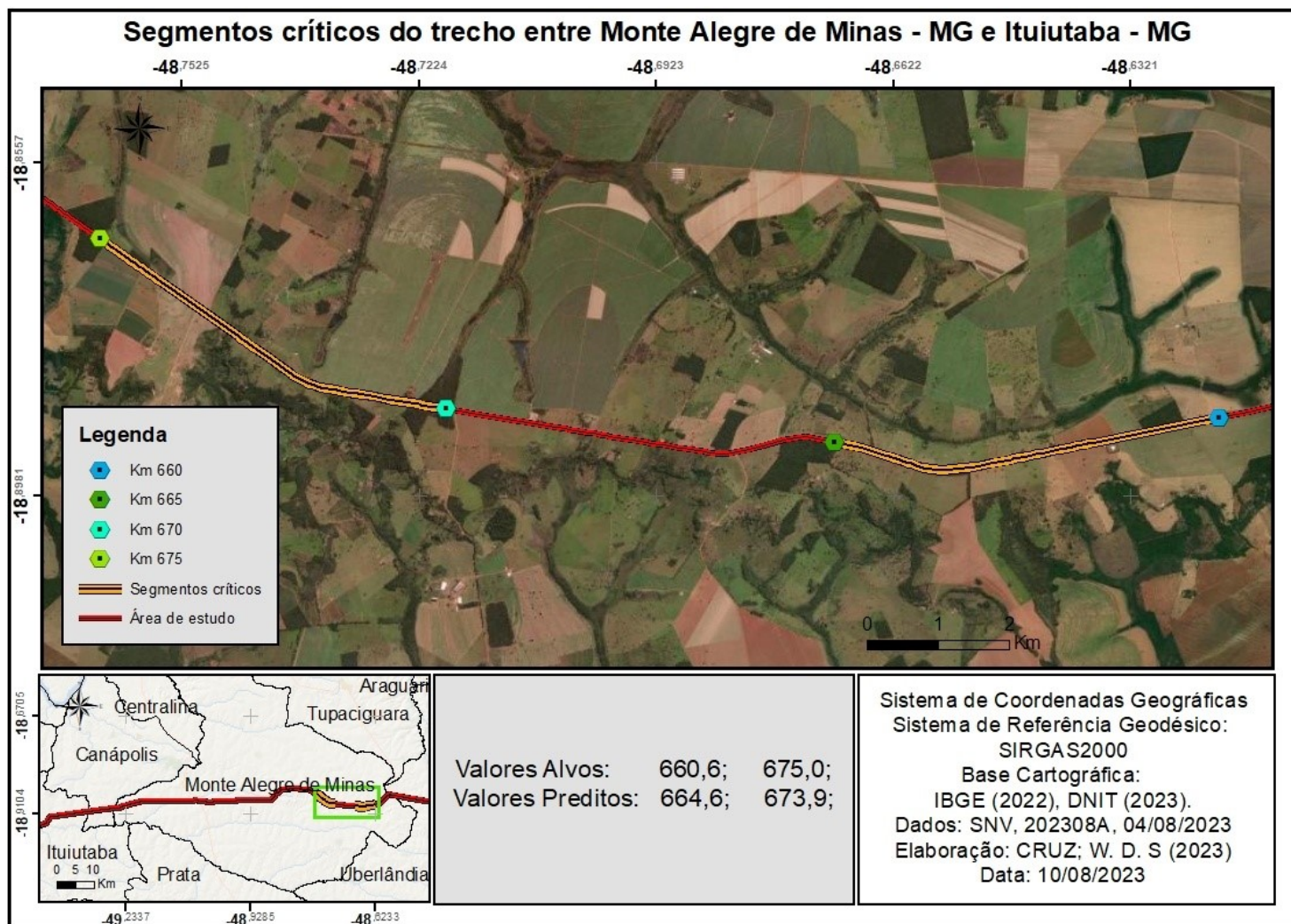
Fonte: Autoria própria (2023).

Figura 13 – Segmentos críticos do trecho entre Uberlândia – MG e Monte Alegre de Minas – MG



Fonte: Autoria própria (2023).

Figura 14 – Segmentos críticos do trecho entre Monte Alegre de Minas – MG e Ituiutaba – MG



Fonte: Autoria própria (2023).

Figura 15 – Segmentos críticos da área de estudo da BR-365



Fonte: Autoria própria (2023).

Para melhor interpretação dos mapas elaborados, foram considerados dentro dos acertos contidos no terceiro método de análise do subconjunto de teste, as predições que se encontravam numa margem de erro de até 5 km (para mais ou para menos) em relação ao seu valor alvo, visto que este foi o menor tamanho de segmentação trabalhado neste estudo. Logo, foi devido a esta acurácia exigida que o trecho contido entre Patos de Minas e Patrocínio não foi representado, uma vez que não houve acerto numa margem de erro de até 5 km entre o km 420,1 e o km 470,1.

Quando examinado o trecho pertencente entre Patrocínio e Monte Carmelo (**Figura 11**), foi identificado apenas um trecho crítico dentro da margem de acerto, variando do km 507 ao km 509.

Em relação ao trecho contido entre Monte Carmelo e Uberlândia (**Figura 12**), foi também identificado pela rede neural 9 apenas um segmento crítico, dessa vez, variando do km 593 ao km 598.

Ao ponderar o resultado obtido no mapa do trecho situado entre Uberlândia e Monte Alegre de Minas (**Figura 13**), é expressiva a quantidade de segmentos críticos e de predições corretas da RNA, totalizando 16 valores alvos preditos corretamente considerando o critério exigido, distribuídos em dois segmentos, do km 610 ao km 628,3 e do km 637 ao km 639. Quando analisadas as circunstâncias dos sinistros de trânsito ocorridos neste trecho a partir das planilhas da PRF, é observado que a geometria ou traçado da via na grande maioria das ocorrências é uma reta, constituídas em quase sua totalidade de pista dupla. Além disso, a fase do dia e a condição climática predominante nestes sinistros caracterizam-se, respectivamente, por ser em pleno dia e céu claro, no qual, foi observado esse cenário em pelo menos 239 ocorrências registrado pela PRF entre os anos de 2017 a 2019. Quanto ao tipo de veículo, os mais presentes nos sinistros de trânsito neste trecho analisado foram os automóveis (189 sinistros), seguidos das motocicletas (160 sinistros). No que tange a severidade, a grande maioria das ocorrências entre Uberlândia e Monte Alegre de Minas ocasionaram em vítimas feridas ou fatais, totalizando em 460 sinistros de trânsito, onde o tipo de acidente mais registrado foi a colisão traseira (145 sinistros) seguida da colisão lateral (50 sinistros).

Quanto ao último trecho analisado localizado entre o município de Monte Alegre de Minas e Ituiutaba (**Figura 14**), foram identificados dois segmentos críticos, situados do km 660 ao km 665, e do km 670 ao km 675.

Assim, são observados numa escala menor, todos os segmentos críticos apontados pela rede neural 9, cuja concentração na área de estudo reside nos municípios de maior população (**Figura 15**).

Levando a análise da área de estudo ao último ano mais recente e com os dados completos, o ano de 2022, é possível observar que o trecho com segmentos onde mais ocorrem sinistros de trânsito permanece sendo o apontado pela rede neural 9 (**Tabela 21**):

Tabela 21 – Número de sinistros de trânsito ocorridos em 2022 entre os municípios de Patos de Minas - MG e Ituiutaba – MG

Trecho	Número de sinistros de trânsito no ano 2022
Patos de Minas - Patrocínio	31
Patrocínio - Monte Carmelo	31
Monte Carmelo - Uberlândia	47
Uberlândia - Monte Alegre de Minas	126
Monte Alegre de Minas - Ituiutaba	82

Fonte: Autoria própria (2023).

7 CONCLUSÃO

O principal objetivo do trabalho “Desenvolver um modelo de rede neural artificial para predição da frequência de acidentes da BR-365, com trecho do município de Patos de Minas ao município de Ituiutaba, no estado de Minas Gerais” foi atingido. No que concerne ao desempenho das redes neurais, a rede neural com 8 entradas, 40 neurônios na camada oculta e 1 neurônio na camada de saída alcançou uma média satisfatória do índice de correlação de Pearson, onde, as predições fornecidas pela RNA selecionada condizem com os dados atestados na realidade, identificando corretamente os segmentos mais críticos dos trechos situados entre os municípios analisados.

Em primeira instância, apesar de o primeiro método ter atingido uma porcentagem de acerto maior na margem de 30 km em relação aos outros métodos aplicados, os resultados se mostraram enviesados, já que essa porcentagem foi alcançada principalmente em função do acúmulo de segmentos críticos pertencentes a apenas um intervalo contido na área de estudo, trecho esse em que se localizavam grande parte dos sinistros de trânsito analisados entre os anos 2017 a 2019. Dessa forma, o terceiro método de análise constitui numa forma de investigação mais aprimorada, na medida em que integra ao estudo um olhar mais atento ao identificar quais são os municípios que compõe o início e o término de trecho analisado.

Além disso, os parâmetros de entrada utilizados “Traçado da via”, “Condições meteorológicas”, “Fase do dia”, “Tipo de veículo” e “Severidade” e “Idade do condutor”, se mostraram eficientes na modelagem da segurança viária ao serem implementados nas redes neurais artificiais, comprovando-se variáveis importantes na determinação de segmentos críticos de uma rodovia federal.

Em contrapartida, embora a rede neural 9 tenha alcançado o melhor desempenho em ambos os primeiros métodos e em quase todas as segmentações no subconjunto de validação, sendo selecionada para analisar o subconjunto de teste, os resultados obtidos ainda são insuficientes para constatar que a capacidade de generalização é de fato eficiente, dado que as taxas de porcentagens de acerto obtidas nos três métodos aplicados no teste ainda são baixas, prejudicando na corroboração de como e de quais fatores interferem diretamente na ocorrência da acidentalidade viária na área de estudo.

Paralelo a isso, quanto à influência da estratégia de segmentação, foi constatado que as RNA's apresentam resultados mais satisfatórios na medida em que o tamanho do segmento

analisado é aumentado, sendo esse fator, portanto, diretamente proporcional ao acerto. Este cenário comprovado é verificado também nos estudos de Silva (2019) e de Oliveira (2021).

Desse modo, no que concerne a avaliação do emprego de algoritmos de aprendizado de máquina na modelagem da segurança viária, percebe-se que mesmo diante de resultados que ainda precisam ser aprimorados, as RNA's possuem uso potencial e constituem numa técnica extremamente revolucionária e promissora, em virtude de suas capacidades de solucionar problemas não linearmente separáveis e de possuírem validade replicativa e preditiva, sendo capazes de ajustarem-se aos dados de treinamento, e, mediante a isso, realizarem previsões razoáveis a partir de dados desconhecidos, como aponta Silva (2019).

7.1 LIMITAÇÕES

Ainda que o objetivo geral tenha sido alcançado neste trabalho, houve limitações que podem ser aperfeiçoadas e ampliadas. Sendo elas:

- A ausência de mais postos de coleta de dados na área de estudo para a utilização de informações de contagem volumétrica. Embora exista o “Plano Nacional de Contagem de Tráfego” ou PNCT iniciado em 1975 e retomado em 2014 pelo DNIT, o PNCT apresenta 10 postos de coleta na BR-365, em que, destes, apenas 4 estão contidos na área de estudo deste trabalho, sendo necessária a ampliação de mais pontos para a aquisição de mais informações.
- A indisponibilidade de dados abertos contendo informações da sinalização vertical da área de estudo, cuja aplicação como parâmetro de entrada poderia trazer resultados mais promissores, bem como um melhor entendimento dos fatores contribuintes na ocorrência de sinistros de trânsito.
- No que diz respeito à função de ativação, o *Weka* possui como *default*, ou seja, como configuração padrão, a função de ativação do tipo sigmóide logística, sendo, portanto, limitado a liberdade de escolha para a utilização de outras funções como, por exemplo, a tangente hiperbólica.

- Outro fator existente é quanto à especificação das topologias das RNA's para a solução de um problema específico, que consiste num longo processo empírico, cuja metodologia não tem uma padronização nesse quesito, dependente inteiramente da variação de valores dos hiperparâmetros impostos pelo modelador.

7.2 RECOMENDAÇÕES

Alguns pontos podem ser aprimorados em trabalho futuros para a continuidade da base empírica referente ao tema proposto neste estudo. Sendo eles:

- Utilização de dados de contagem volumétrica, caso seja possível a aquisição desses dados para a realização do estudo de acidentalidade viária na área de estudo a ser analisada.
- Aumentar o limite para cada camada oculta de 40 para 100 neurônios ocultos, o que pode trazer resultados mais satisfatórios, ao se evitar, concomitantemente, a saturação dos neurônios e a ocorrência das condições de *overfitting* e *underfitting*.
- Alterar o valor do critério de parada assim como o número máximo de épocas, buscando melhores médias do índice de correlação de treino e das taxas de assertividade, evitando as condições extremas durante o treinamento das redes.
- Tornar mais robusta a validação cruzada, utilizando dessa vez dez subconjuntos distintos de treinamento e dez subconjuntos de validação, realizando, posteriormente, a média dos treinos e a média da validação.
- Ampliar a janela temporal da área de estudo a ser analisada para cinco anos, objetivando uma maior variedade de situações ocorridas, incrementando o processo de absorção de conhecimento pelas RNA's e suas capacidades preditivas.

- Acrescentar mais parâmetros de entrada, uma vez que isso eleva a capacidade de generalização da rede neural artificial, acelerando o processo de convergência. Recomenda-se das planilhas da PRF, o “Tipo de acidente”, o “Sentido da via” e o “Uso do solo”.

REFERÊNCIAS

- ABRACD.** Overfitting e underfitting , 2022. Disponível em: <https://abracd.org/overfitting-e-underfitting-em-machine-learning/>. Acesso em: 30 abr. 2023.
- ALVES, Priscilla; JUNIOR, Archimedes Azevedo Raia. **Análise de correlação entre acidentes de trânsito, uso e ocupação do solo, polos geradores de viagens e população em Uberlândia-MG.** Revista dos Transportes Públicos - ANTP - Ano 34, 2012. Disponível em: http://files-server.antp.org.br/_5dotSystem/download/dcmDocument/2013/01/10/0F9D9C11-F1D4-49B0-9A5F-C43F4079C0EB.pdf. Acesso em: 7 out. 2022
- AMORIM, Brunna de Sousa Pereira. **Uso de aprendizado de máquina para classificação de risco de acidentes em rodovias.** 2019. 106 p. Pós-Graduação (Mestre em Ciência da Computação) - Universidade Federal de Campina Grande, Campina Grande, Paraíba, 2019. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/xmlui/handle/riufcg/7340>. Acesso em: 29 set. 2022.
- ANDRADE, Flávia Reis; ANTUNES, José Leopoldo Ferreira. **Tendência do número de vítimas em acidentes de trânsito nas rodovias federais brasileiras antes e depois da Década de Ação pela Segurança no Trânsito.** 2019. Disponível em: <https://doi.org/10.1590/0102-311X00250218>. Acesso em: 9 out. 2022.
- ASSOCIAÇÃO BRASILEIRA DE MEDICINA DE TRÁFEGO (ABRAMET). **ABNT muda terminologia e adota a expressão sinistro de trânsito para qualificar incidentes no tráfego:** Pesquisa de sinistros de trânsito — Terminologia. 2021. Disponível em: <https://www.abramet.com.br/noticias/abnt-muda-terminologia-e-adota-a-expressao-sinistro-de-transito-para-qualificar-incidentes-no-trafego/>. Acesso em: 28 out. 2022.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **ABNT NBR 10697:** Pesquisa de sinistros de trânsito — Terminologia. 2020. Disponível em: <https://www.abramet.com.br/repo/public/commons/ABNT%20NBR10697%202020%20Acidentes%20de%20Transito%20Terminologia.pdf>. Acesso em: 28 out. 2022.
- BINOTI, Daniel Henrique Breda; BINOTI, Mayra Luiza Marques da Silva; LEITE, Helio Garcia. CONFIGURAÇÃO DE REDES NEURAS ARTIFICIAIS PARA ESTIMAÇÃO DO VOLUME DE ÁRVORES. **Ciência da Madeira (Braz. J. Wood Sci.)**, 2014. Disponível em: <https://periodicos.ufpel.edu.br/index.php/cienciadamadeira/article/view/4067>. Acesso em: 27 abr. 2023.
- CAFISO, Salvatore; AGOSTINO, Carmelo D'; PERSAUD, Bhagwant. Investigating the influence of segmentation in estimating safety performance functions for roadway sections. **Journal of Traffic and Transportation Engineering**, 2018. DOI:10.1016/j.jtte.2017.10.001. Disponível em: <https://www.researchgate.net/publication>

/324123675_ Investigating_the_influence_of_segmentation_in_estimating_safety_performance_functions_for_roadway_sections. Acesso em: 13 out. 2022.

CAMPOS, Cintia Isabel; FERREIRA, José Leopoldo. **POLÍTICAS PÚBLICAS DE SEGURANÇA VIÁRIA PARA O TRANSPORTE DE CRIANÇAS: O CASO BRASILEIRO**. Porto, PORTUGAL, 2013. Disponível em: DOI: 10.14684/SHEWC.13.2013.244-248. Acesso em: 18 out. 2022.

CARMO, Cássio Leandro do. **Segurança viária em trechos urbanos de rodovias federais**. 2019. Tese (Doutorado em Engenharia Urbana) – Universidade Federal de São Carlos, São Carlos, 2019. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/11500>. Acesso em: 16 out. 2022.

CARVALHO, Carlos Henrique Ribeiro. **Acidentes de transportes terrestres no Brasil: uma tragédia anunciada**. 2016. Disponível em: <https://repositorio.ipea.gov.br/handle/11058/9603>. Acesso em: 28 set. 2022.

CHUERUBIM, M. L.; FERREIRA, L. N.; VALEJO, A. D.; BEZERRA, B. S.; MAROTTA, G. S.; DA SILVA, I. Avaliação de técnicas de balanceamento de bases de dados para classificação da severidade de acidentes rodoviários empregando redes neurais artificiais. **TRANSPORTES**, v. 28, n. 5, p. 252–266, 2020. DOI: 10.14295/transportes.v28i5.2271. Disponível em: <https://www.revistatransportes.org.br/anpet/article/view/2271>. Acesso em: 13 dez. 2022.

CINTRA, Rosangela. **Introdução à neurocomputação**. INPE, 2018. Disponível em: http://www.inpe.br/elac2018/arquivos/ELAC2018_MC3_apostila.pdf. Acesso em: 3 nov. 2022.

CONFEDERAÇÃO NACIONAL DO TRANSPORTE (CNT). **Painel CNT de Acidentes Rodoviários - Principais dados - 2022**. 2023. Disponível em: <https://www.cnt.org.br/painel-acidente>. Acesso em: 11 maio. 2023.

COSTA, Anna Helena Reali *et al.* Trajetória acadêmica da Inteligência Artificial no Brasil. In: COZMAN, Fabio G.; PLONSKI, Guilherme Ary; NERI, Hugo. **Inteligência artificial: avanços e tendências**. Universidade de São Paulo, São Paulo: 2021. cap. Introdução, p. 28-65. ISBN 978-65-87773-13-1. Disponível em: <https://doi.org/10.11606/9786587773131>. Acesso em: 5 nov. 2022.

DNIT. **Plano Nacional de Viação e Sistema Nacional de Viação**. 2022. Disponível em: <https://www.gov.br/dnit/pt-br/assuntos/atlas-e-mapas/pnv-e-snv>. Acesso em: 10 dez. 2022.

DOK DESPACHANTE. <https://www.despachantedok.com.br/codigo-de-transito-brasileiro>. 2022. Disponível em: <https://www.despachantedok.com.br/codigo-de-transito-brasileiro>. Acesso em: 25 out. 2022.

ECOVIAS DO CERRADO. **Ecovias do Cerrado**. 2022. Disponível em: <https://www.ecoviasdocerrado.com.br/institucional/ecovias-do-cerrado>. Acesso em: 4 dez. 2022.

FILHO, Dalson Britto Figueiredo; JÚNIOR, José Alexandre da Silva. Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). **Revista Política Hoje**, 2009. Disponível em: http://dirin.s3.amazonaws.com/drive_materias/1666287394.pdf. Acesso em: 10 maio 2023.

HAUER, Ezra. **Statistical Road Safety Modeling**. 2004. Disponível em: https://www.researchgate.net/publication/239438865_Statistical_Road_Safety_Modeling. Acesso em: 14 nov. 2022.

HAYKIN, Simon. **Redes neurais princípios e prática**. Grupo A, 2001. *E-book*. ISBN 9788577800865. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788577800865/>. Acesso em: 30 out. 2022

HOPKINS, W.G. **A New View Statistics: Correlation Coefficient**. Internet Society for Sport Science, 2000. Disponível em: <http://www.sportsci.org/resource/stats/correl.html>. Acesso em: 7 dez. 2022.

IBGE. 2023. Disponível em: <https://www.ibge.gov.br/>. Acesso em: 5 agos. 2023.

IBM. Exatidão, 2021. Disponível em: < <https://www.ibm.com/docs/pt-br/cloud-paks/cp-data/3.5.0?topic=overview-accuracy>>. Acesso em: 18 maio. 2023.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). **FATORES CONDICIONANTES DA GRAVIDADE DOS ACIDENTES DE TRÂNSITO NAS RODOVIAS BRASILEIRAS**. Brasília, 2008. Disponível em: <https://repositorio.ipea.gov.br/handle/11058/1597>. Acesso em: 28 set. 2022.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). **ACIDENTES DE TRÂNSITO NAS RODOVIAS FEDERAIS BRASILEIRAS: CARACTERIZAÇÃO, TENDÊNCIAS E CUSTOS PARA A SOCIEDADE**. Brasília, 2015. Disponível em: <https://repositorio.ipea.gov.br/handle/11058/7493>. Acesso em: 28 set. 2022.

KOPIER, Alberto A. *et al.* **Redes Neurais Artificiais e suas aplicações no setor elétrico**. Revista de Engenharias da Faculdade Salesiana n. 9 pp. 27-33, 2019. Disponível em: http://www.fsma.edu.br/RESA/Edicao9/FSMA_RES_2019_1_04.pdf. Acesso em: 29 out. 2022.

LINKEDIN. RMSE ou MAE? Como avaliar meu modelo de machine learning?, 2018. Disponível em: <https://www.linkedin.com/pulse/rmse-ou-mae-como-avaliar-meu-modelo-de-machine-learning-rezende/?originalSubdomain=pt>. Acesso em: 30 abr. 2023.

LORD, Dominique; MANNERING, Fred. **The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives**. 2010. Disponível em: https://www.researchgate.net/publication/222659783_The_Statistical_Analysis_of_Crash-Frequency_Data_A_Review_and_Assessment_of_Methodological_Alternatives. Acesso em: 9 out. 2022.

MACEDO, Deivielison Ximenes Siqueira *et al.* **Caracterização dos acidentes com máquinas agrícolas em rodovias federais no estado do Rio Grande do Sul**. Ciência Rural,

Santa Maria, v.45, n.1, p.43-46, 2015. Disponível em: <https://doi.org/10.1590/0103-8478cr20140333>. Acesso em: 17 out. 2022.

MACHADO, Vinicius Ponte. **INTELIGÊNCIA ARTIFICIAL**. Universidade Federal do Piauí, 2011. Disponível em: <https://sigaa.ufpi.br/sigaa/verProducao?idProducao=3005620&key=3e434d2c6ca0839057b4f3f7a795d48e>. Acesso em: 23 out. 2022.

MANZAN, JOSÉ RICARDO G. **ANÁLISE DE DESEMPENHO DE REDES NEURAIS ARTIFICIAIS DO TIPO MULTILAYER PERCEPTRON POR MEIO DO DISTANCIAMENTO DOS PONTOS DO ESPAÇO DE SAÍDA**. Tese (Doutorado em Ciências) – Universidade Federal de Uberlândia, Uberlândia, 2016. DOI • <https://doi.org/10.14393/ufu.te.2016.133>. Disponível em: <https://repositorio.ufu.br/handle/123456789/17967>. Acesso 21 out. 2022.

MARTINELLI, Edmar. **Extração de conhecimento de redes neurais artificiais**. 1999. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 1999. Disponível em: [doi:10.11606/D.55.1999.tde-19102001-100256](https://doi.org/10.11606/D.55.1999.tde-19102001-100256). Acesso em: 10 nov. 2022

MENDES, Eduardo Augusto de Oliveira. **ANÁLISE EXPLORATÓRIA DA ACIDENTALIDADE EM RODOVIAS FEDERAIS BRASILEIRAS DURANTE A PANDEMIA DO COVID-19**. 2021. 59 p. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Civil) - Instituto Federal Goiano, Rio Verde, Goiás, 2021. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/xmlui/handle/riufcg/7340>. Acesso em: 29 set. 2022.

MICROSOFT. Avaliar o modelo ML.NET com métricas, 2022. Disponível em: <https://docs.microsoft.com/pt-br/dotnet/machine-learning/resources/metrics> > Acesso em: 30 abr. 2023.

MOREIRA, Samuel de Carvalho. **Detecção e Diagnóstico de Faltas via Estatística Multivariada e Redes Neurais**. 2007. Dissertação de Pós-graduação (Mestre em Engenharia Elétrica) – Universidade Católica de Minas Gerais, 2007. Disponível em: http://www.biblioteca.pucminas.br/teses/EngEletrica_MoreiraSC_1.pdf. Acesso em: 27 abr. 2023.

NODARI, Christine Tessele. **MÉTODO DE AVALIAÇÃO DA SEGURANÇA POTENCIAL DE SEGMENTOS RODOVIÁRIOS RURAIS DE PISTA SIMPLES**. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003. Disponível em: <https://lume.ufrgs.br/handle/10183/3675>. Acesso em: 8 nov. 2022.

OLIVEIRA, Mariana Alves. **Aplicação de redes neurais artificiais na previsão de acidentes rodoviários em pistas simples: estudo do trecho da BR-230 pertencente ao estado da Paraíba**. 2021. 66f. Trabalho de Conclusão de Curso (Monografia), Curso de Bacharelado em Engenharia Civil, Centro de Tecnologia e Recursos Naturais, Universidade Federal de Campina Grande - Campina Grande - Paraíba - Brasil, 2021. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/23714>. Acesso em: 30 set. 2022.

PATOSHOJE. **BR 365 recebe asfalto novo no trecho que vai de Patos de Minas até a cidade de Uberlândia**. 2017. Disponível em: <https://patoshoje.com.br/noticias/br-365-recebe->

asfalto-novo-no-trecho-que-vai-de-patos-de-minas-ate-a-cidade-de-uberlandia-49872.html. Acesso em: 4 dez. 2022.

POLÍCIA RODOVIÁRIA FEDERAL - PRF. **Acidentes**. 2023. Disponível em: <https://www.gov.br/prf/pt-br/aceso-a-informacao/dados-abertos/dados-abertos-acidentes>. Acesso em: 30 set. 2022.

REIS, Carlos Henrique. **Otimização de Hiperparâmetros em Redes Neurais Profundas**. 2018. Trabalho de Conclusão de Curso (Ciências da computação) - Universidade Federal de Itajubá, 2018. Disponível em: https://carlos-henreis.github.io/files/Monografia_TFG.pdf. Acesso em: 27 abr. 2023.

RESTALLACK, Angus Eugene; OSTENDORF, Bertram. **Current Understanding of the Effects of Congestion on Traffic Accidents**. 2019. Disponível em: <https://doi.org/10.3390/ijerph16183400>. Acesso em: 29 out. 2022.

SALADIÉ, Òscar; BUSTAMANTE, Edgar; GUTIÉRREZ, Aaron. **COVID-19 lockdown and reduction of traffic accidents in Tarragona province, Spain**. 2020. Disponível em: <https://doi.org/10.1016/j.trip.2020.100218>. Acesso em: 16 out. 2022.

SAMEEN, Maher Ibrahim; PRADHAN, Biswajeet. **Severity Prediction of Traffic Accidents with Recurrent Neural Networks**. Appl. Sci. 2017, 7(6), 476, 2017. Disponível em: <https://doi.org/10.3390/app7060476>. Acesso em: 11 out. 2022..

SCHUSTER, Fernanda Pivato. **O uso adequado de rotatórias como agente redutor da accidentalidade no trânsito**. 2012. Disponível em: http://files-server.antp.org.br/_5dotSystem/download/dcmDocument/2013/10/07/96FB1A2D-43D2-4261-866C-F1A334AE7C2A.pdf. Acesso em: 24 out. 2022.

SILVA, Daniel Souza da. **Redes neurais aplicadas a estimativa das perdas de transformadores trifásicos de distribuição**. 2018. Trabalho de Conclusão de Curso (Engenharia Elétrica) - Escola Superior de Tecnologia de da Universidade do Estado do Amazonas, 2018. Disponível em: <http://repositorioinstitucional.uea.edu.br//handle/riuea/1532>. Acesso em: 25 abr. 2023.

SILVA, Ivan Nunes; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes Neurais Artificiais para engenharia e ciências aplicadas: fundamentos teóricos e aspectos práticos**. 2ª edição. ed. São paulo: Artliber, 2016. Disponível em: https://artliber.com.br/amostra/redes_neurais.pdf. Acesso em: 29 out. 2022.

SILVA, Leandro Nunes Castro. **Análise e síntese de estratégias de aprendizado para redes neurais artificiais**. 1998. Dissertação de Pós-graduação (Mestre em Engenharia Elétrica - Área de concentração: Ciência da computação) - Universidade Estadual de Campinas, 1998. DOI: <https://doi.org/10.47749/T/UNICAMP.1998.136083>. Disponível em: <https://repositorio.unicamp.br/acervo/detalhe/136083>. Acesso em: 20 abr. 2023.

SILVA, Philippe Barbosa. **Modelos de previsão de acidentes em rodovia brasileira de pista dupla**. 2019. xxii, 320 f., il. Tese (Doutorado em Transportes) — Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2019. Disponível em: <https://repositorio.unb.br/handle/10482/38140>. Acesso em: 28 set. 2022.

SILVA, Philippe Barbosa; ANDRADE, Michelle; WEIGANG, Li. **ANÁLISE DE ACIDENTES EM RODOVIAS FEDERAIS BRASILEIRAS POR MEIO DE REDES NEURAIAS ARTIFICIAIS**. 2017. Disponível em: http://146.164.5.73:30080/tempsite/anais/documentos/2017/Trafego%20Urbano%20e%20Rodoviario/Seguranca%20Viaria%20I/2_232_AC.pdf. Acesso em: 28 set. 2022.

SOUZA, Raquel Machado de. **Avaliação da Predição de Algoritmos de Treinamento Supervisionado de Redes Neurais Artificiais Aplicado a Qualidade de Biodiesel**. 2015. Dissertação (Mestrado em Engenharia de Eletricidade – Área de concentração: Ciência da computação) - Universidade Federal de Itajubá, 2015. Disponível em: <http://tedebc.ufma.br:8080/jspui/handle/tede/292>. Acesso em: 27 abr. 2023.

WANG, Chao; QUDDUS, Mohammed; ISON, Stephen. **Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model**. 2011. Disponível em: https://www.researchgate.net/publication/51550327_Predicting_accident_frequency_at_their_severity_levels_and_its_application_in_site_ranking_using_a_two-stage_mixed_multivariate_model. Acesso em: 30 out. 2022.

WOLFRANN, Joice *et al.* Simulação da irrigação por aspersão utilizando redes neurais artificiais treinadas com e sem termo momentum. **Revista de Computação aplicada ao Agronegócio**, 2018. DOI 10.3895/rcaa.v1i2.9501. Disponível em: <https://periodicos.utfpr.edu.br/rcaa/article/view/9501>. Acesso em: 26 abr. 2023.

World Health Organization – WHO, 2018. **Global status report on road safety 2018: summary**. Genebra: Organização Mundial da Saúde (OMS). Acesso em: 3 out. 2022.

ZENG, Qiang *et al.* **Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks**. 2016. Disponível em: https://www.researchgate.net/publication/301241972_Modeling_nonlinear_relationship_between_crash_frequency_by_severity_and_contributing_factors_by_neural_networks. Acesso em: 30 out. 2022.

APÊNDICE

Tabela 22 – Normalização da faixa etária

Normalização da faixa etária	
Idade	Normalizado
16	0
17	0,0135
18	0,0270
19	0,0405
20	0,0541
21	0,0676
22	0,0811
23	0,0946
24	0,1081
25	0,1216
26	0,1351
27	0,1486
28	0,1622
29	0,1757
30	0,1892
31	0,2027
32	0,2162
33	0,2297
34	0,2432
35	0,2568
36	0,2703
37	0,2838
38	0,2973
39	0,3108
40	0,3243
41	0,3378
42	0,3514
43	0,3649
44	0,3784
45	0,3919
46	0,4054
47	0,4189
48	0,4324

49	0,4459
50	0,4595
51	0,4730
52	0,4865
53	0,5000
54	0,5135
55	0,5270
56	0,5405
57	0,5541
58	0,5676
59	0,5811
60	0,5946
61	0,6081
62	0,6216
63	0,6351
64	0,6486
65	0,6622
66	0,6757
67	0,6892
68	0,7027
69	0,7162
70	0,7297
71	0,7432
72	0,7568
73	0,7703
74	0,7838
75	0,7973
76	0,8108
77	0,8243
78	0,8378
79	0,8514
80	0,8649
86	0,9459

Fonte: Autoria própria (2023).

Tabela 23 – Normalização dos quilômetros

Normalização dos quilômetros									
Valor	Normalizado	Valor	Normalizado	Valor	Normalizado	Valor	Normalizado	Valor	Normalizado
420,7	0,0018	493,0	0,2189	592,9	0,5189	625,0	0,6153	669,1	0,7477
421,0	0,0027	493,5	0,2204	593,0	0,5192	625,1	0,6156	669,5	0,7489
421,4	0,0039	494,0	0,2219	593,1	0,5195	625,2	0,6159	670,0	0,7505
421,5	0,0042	494,2	0,2225	594,0	0,5222	625,3	0,6162	671,0	0,7535
423,0	0,0087	495,0	0,2249	594,5	0,5237	625,4	0,6165	671,2	0,7541
425,0	0,0147	497,0	0,2309	595,0	0,5252	625,5	0,6168	672,0	0,7565
425,3	0,0156	497,2	0,2315	595,5	0,5267	625,6	0,6171	673,0	0,7595
426,0	0,0177	497,8	0,2333	595,7	0,5273	625,7	0,6174	673,3	0,7604
426,4	0,0189	498,5	0,2354	595,8	0,5276	625,8	0,6177	673,8	0,7619
428,0	0,0237	498,6	0,2357	596,0	0,5282	625,9	0,6180	674,0	0,7625
429,0	0,0267	498,8	0,2363	596,8	0,5306	626,0	0,6183	675,0	0,7655
429,5	0,0282	499,0	0,2369	597,0	0,5312	626,1	0,6186	675,7	0,7676
429,6	0,0285	501,0	0,2429	597,5	0,5327	626,2	0,6189	677,0	0,7715
429,8	0,0291	502,3	0,2468	598,0	0,5342	626,5	0,6198	677,2	0,7721
430,0	0,0297	503,0	0,2489	598,9	0,5369	626,7	0,6204	677,7	0,7736
432,0	0,0357	503,9	0,2517	599,0	0,5372	626,8	0,6207	677,8	0,7739
432,2	0,0363	504,0	0,2520	599,4	0,5384	626,9	0,6210	678,0	0,7745
432,5	0,0372	505,0	0,2550	599,7	0,5393	627,0	0,6213	679,4	0,7787
432,6	0,0375	505,9	0,2577	599,9	0,5399	627,1	0,6216	679,5	0,7790
434,0	0,0417	508,9	0,2667	600,0	0,5402	627,2	0,6219	679,8	0,7799
435,0	0,0447	509,7	0,2691	600,5	0,5417	627,5	0,6228	679,9	0,7802
436,0	0,0477	510,7	0,2721	601,0	0,5432	627,7	0,6234	681,2	0,7841
436,5	0,0492	511,9	0,2757	601,5	0,5447	627,8	0,6237	682,3	0,7874
437,2	0,0514	512,0	0,2760	601,8	0,5456	627,9	0,6240	682,9	0,7892
437,5	0,0523	513,0	0,2790	602,0	0,5462	628,0	0,6243	683,0	0,7895
438,0	0,0538	513,2	0,2796	602,1	0,5465	628,1	0,6246	683,6	0,7913
440,8	0,0622	514,0	0,2820	602,2	0,5468	628,6	0,6261	684,0	0,7925
441,0	0,0628	515,7	0,2871	603,0	0,5492	628,7	0,6264	684,5	0,7940
442,0	0,0658	515,8	0,2874	603,2	0,5498	628,9	0,6270	686,0	0,7985
443,0	0,0688	516,0	0,2880	603,6	0,5511	629,0	0,6273	686,5	0,8000
443,1	0,0691	516,5	0,2895	604,0	0,5523	629,2	0,6279	687,0	0,8015
443,3	0,0697	517,0	0,2910	605,0	0,5553	629,3	0,6282	688,0	0,8045
443,4	0,0700	517,4	0,2922	605,2	0,5559	629,7	0,6294	688,3	0,8054
443,8	0,0712	518,0	0,2940	605,3	0,5562	629,9	0,6300	689,0	0,8075
445,0	0,0748	518,5	0,2955	606,0	0,5583	630,0	0,6303	689,1	0,8078
446,0	0,0778	520,7	0,3021	606,5	0,5598	630,5	0,6318	689,8	0,8099
446,1	0,0781	521,2	0,3036	607,0	0,5613	630,6	0,6321	689,9	0,8102
446,2	0,0784	522,0	0,3060	607,1	0,5616	631,0	0,6333	690,0	0,8105

447,0	0,0808	522,2	0,3066	607,7	0,5634	631,1	0,6336	690,2	0,8111
448,2	0,0844	523,0	0,3090	608,0	0,5643	631,3	0,6342	690,8	0,8129
448,3	0,0847	523,2	0,3096	609,0	0,5673	631,5	0,6348	691,0	0,8135
448,5	0,0853	523,3	0,3099	610,0	0,5703	631,8	0,6357	691,8	0,8159
449,0	0,0868	524,0	0,3120	610,5	0,5718	631,9	0,6360	692,0	0,8165
449,5	0,0883	524,7	0,3141	610,7	0,5724	632,0	0,6363	693,0	0,8195
449,7	0,0889	525,0	0,3150	611,0	0,5733	633,0	0,6393	693,3	0,8204
450,5	0,0913	525,6	0,3168	611,1	0,5736	633,5	0,6408	693,7	0,8216
451,0	0,0928	525,8	0,3174	611,7	0,5754	634,0	0,6423	695,0	0,8255
451,7	0,0949	527,0	0,3210	611,8	0,5757	634,3	0,6432	696,1	0,8288
452,0	0,0958	533,0	0,3390	612,0	0,5763	635,4	0,6465	698,0	0,8345
453,6	0,1006	534,0	0,3420	612,3	0,5772	635,5	0,6468	699,0	0,8375
454,5	0,1033	535,0	0,3450	612,5	0,5778	635,8	0,6477	700,0	0,8405
454,9	0,1045	538,1	0,3544	612,6	0,5781	636,0	0,6483	700,3	0,8414
456,3	0,1087	538,3	0,3550	612,9	0,5790	636,1	0,6486	700,4	0,8417
456,5	0,1093	539,3	0,3580	613,0	0,5793	636,3	0,6492	702,0	0,8465
457,5	0,1123	543,4	0,3703	613,2	0,5799	636,9	0,6511	702,5	0,8480
458,0	0,1138	544,0	0,3721	613,3	0,5802	637,0	0,6514	704,0	0,8526
458,9	0,1165	544,7	0,3742	613,5	0,5808	637,1	0,6517	705,0	0,8556
459,5	0,1183	545,2	0,3757	613,6	0,5811	637,3	0,6523	705,9	0,8583
460,8	0,1222	545,3	0,3760	613,8	0,5817	637,5	0,6529	706,0	0,8586
461,0	0,1228	546,0	0,3781	613,9	0,5820	638,0	0,6544	708,0	0,8646
461,3	0,1237	548,0	0,3841	614,0	0,5823	638,2	0,6550	708,2	0,8652
461,6	0,1246	549,0	0,3871	614,2	0,5829	638,3	0,6553	708,9	0,8673
461,8	0,1252	550,0	0,3901	614,3	0,5832	638,5	0,6559	709,0	0,8676
462,0	0,1258	552,0	0,3961	614,4	0,5835	638,7	0,6565	709,7	0,8697
462,6	0,1276	553,0	0,3991	614,5	0,5838	639,0	0,6574	709,8	0,8700
463,0	0,1288	553,1	0,3994	615,0	0,5853	639,7	0,6595	710,9	0,8733
463,3	0,1297	554,0	0,4021	615,1	0,5856	640,0	0,6604	711,0	0,8736
463,8	0,1312	557,0	0,4111	615,3	0,5862	640,1	0,6607	712,0	0,8766
464,0	0,1318	558,0	0,4141	615,6	0,5871	640,4	0,6616	712,2	0,8772
464,6	0,1336	559,0	0,4171	616,0	0,5883	641,0	0,6634	713,0	0,8796
464,9	0,1345	560,0	0,4201	616,2	0,5889	641,1	0,6637	715,0	0,8856
465,0	0,1348	561,0	0,4231	616,4	0,5895	642,0	0,6664	715,5	0,8871
466,0	0,1378	561,7	0,4252	617,0	0,5913	642,5	0,6679	716,0	0,8886
467,0	0,1408	563,0	0,4291	617,1	0,5916	642,7	0,6685	717,6	0,8934
468,0	0,1438	565,0	0,4351	617,3	0,5922	642,8	0,6688	718,2	0,8952
468,5	0,1453	566,0	0,4381	618,0	0,5943	643,0	0,6694	719,0	0,8976
469,0	0,1468	567,0	0,4411	618,1	0,5946	643,1	0,6697	720,0	0,9006
469,3	0,1477	567,1	0,4414	618,2	0,5949	644,0	0,6724	720,8	0,9030
469,4	0,1480	567,3	0,4420	618,5	0,5958	645,0	0,6754	721,0	0,9036
469,8	0,1492	567,8	0,4435	618,6	0,5961	645,1	0,6757	726,0	0,9186
470,0	0,1498	568,0	0,4441	618,8	0,5967	645,5	0,6769	726,2	0,9192

471,0	0,1529	568,1	0,4444	618,9	0,5970	645,6	0,6772	726,3	0,9195
471,1	0,1532	569,8	0,4495	619,0	0,5973	645,8	0,6778	726,5	0,9201
471,2	0,1535	570,0	0,4502	619,2	0,5979	645,9	0,6781	727,0	0,9216
472,0	0,1559	571,0	0,4532	619,3	0,5982	646,0	0,6784	729,2	0,9282
473,0	0,1589	572,0	0,4562	619,5	0,5988	647,0	0,6814	729,5	0,9291
473,7	0,1610	572,7	0,4583	619,8	0,5997	648,0	0,6844	729,6	0,9294
474,6	0,1637	573,0	0,4592	619,9	0,6000	649,0	0,6874	730,0	0,9306
475,1	0,1652	573,5	0,4607	620,0	0,6003	649,9	0,6901	731,0	0,9336
475,5	0,1664	574,0	0,4622	620,2	0,6009	650,0	0,6904	731,5	0,9351
476,0	0,1679	575,0	0,4652	620,3	0,6012	650,7	0,6925	731,8	0,9360
476,8	0,1703	575,2	0,4658	620,5	0,6018	651,0	0,6934	732,5	0,9381
476,9	0,1706	577,1	0,4715	620,7	0,6024	652,9	0,6991	734,0	0,9426
477,0	0,1709	577,8	0,4736	620,9	0,6030	653,0	0,6994	735,0	0,9456
477,9	0,1736	578,0	0,4742	621,0	0,6033	653,5	0,7009	737,2	0,9523
478,0	0,1739	578,4	0,4754	621,1	0,6036	654,0	0,7024	738,0	0,9547
479,0	0,1769	579,0	0,4772	621,3	0,6042	654,8	0,7048	740,8	0,9631
479,2	0,1775	580,0	0,4802	621,4	0,6045	655,8	0,7078	741,0	0,9637
480,0	0,1799	580,2	0,4808	621,5	0,6048	656,0	0,7084	741,4	0,9649
480,2	0,1805	580,4	0,4814	621,6	0,6051	656,1	0,7087	741,5	0,9652
481,0	0,1829	580,7	0,4823	621,7	0,6054	657,0	0,7114	741,8	0,9661
481,2	0,1835	581,0	0,4832	621,8	0,6057	658,0	0,7144	742,0	0,9667
481,5	0,1844	581,1	0,4835	621,9	0,6060	660,0	0,7204	742,2	0,9673
481,6	0,1847	581,2	0,4838	622,0	0,6063	660,7	0,7225	742,9	0,9694
482,0	0,1859	581,4	0,4844	622,5	0,6078	660,8	0,7228	744,8	0,9751
482,5	0,1874	583,0	0,4892	622,8	0,6087	661,0	0,7234	745,0	0,9757
482,8	0,1883	583,4	0,4904	622,9	0,6090	662,0	0,7264	745,5	0,9772
483,0	0,1889	583,5	0,4907	623,0	0,6093	662,2	0,7270	746,0	0,9787
485,9	0,1976	584,0	0,4922	623,1	0,6096	663,0	0,7294	747,0	0,9817
486,0	0,1979	585,0	0,4952	623,5	0,6108	664,0	0,7324	747,4	0,9829
487,0	0,2009	585,3	0,4961	623,6	0,6111	664,2	0,7330	748,0	0,9847
487,2	0,2015	586,5	0,4997	623,8	0,6117	665,0	0,7354	748,7	0,9868
488,2	0,2045	587,0	0,5012	623,9	0,6120	665,9	0,7381	749,0	0,9877
488,5	0,2054	588,0	0,5042	624,0	0,6123	666,0	0,7384	749,5	0,9892
488,7	0,2060	589,0	0,5072	624,2	0,6129	666,3	0,7393	750,0	0,9907
489,0	0,2069	589,1	0,5075	624,4	0,6135	666,6	0,7402	751,0	0,9937
489,8	0,2093	590,0	0,5102	624,5	0,6138	667,0	0,7414	751,6	0,9955
490,0	0,2099	590,6	0,5120	624,6	0,6141	667,2	0,7420	752,0	0,9967
491,0	0,2129	591,8	0,5156	624,7	0,6144	667,5	0,7429	752,1	0,9970
491,1	0,2132	592,0	0,5162	624,8	0,6147	668,0	0,7444	752,8	0,9991
492,0	0,2159	592,8	0,5186	624,9	0,6150	669,0	0,7474	753,0	0,9997

Tabela 24 – Valores Alvos e Valores Preditos pela rede neural 9

Valor alvo (Km)	Valor Predito (Km)	Valor alvo (Km)	Valor Predito (Km)	Valor alvo (Km)	Valor Predito (Km)
567,0	616,1593	567,1	616,3480	471,0	773,5795
420,7	599,3095	495,0	667,4635	613,0	613,6951
625,6	617,9020	516,5	525,9052	504,0	617,7577
559,0	608,4226	603,2	625,3390	693,3	599,2984
753,0	617,6134	553,0	603,2722	481,2	615,4711
727,0	611,2198	570,0	616,9696	660,7	664,6441
498,8	322,8973	610,0	783,9025	554,0	592,7716
430,0	599,0098	745,5	562,3243	627,0	618,0574
626,7	613,5952	605,3	610,5316	665,0	608,8000
429,0	613,1512	737,2	619,9000	560,0	547,4614
523,0	617,1250	522,0	537,4270	613,0	626,8153
436,0	625,2391	575,2	585,4567	421,0	612,8737
497,0	615,8263	456,5	685,4788	726,0	670,9711
445,0	614,7274	704,0	631,0000	618,0	484,3912
538,3	624,9949	751,6	617,1805	575,0	611,5861
636,0	611,4196	557,0	721,7536	467,0	613,0957
626,0	618,3571	675,0	629,5126	689,1	617,5024
608,0	679,0297	595,0	615,6265	596,8	559,2274
630,5	618,9898	595,5	613,7839	617,0	613,7506
626,5	667,6411	518,5	437,4826	711,0	615,1381
625,3	599,0098	744,8	670,3828	428,0	687,4546
600,0	616,9918	446,0	596,6233	670,0	610,9423
620,5	634,3411	623,0	611,5861	590,0	880,3615
466,0	630,1897	552,0	613,4176	501,0	605,3368
425,0	615,1492	731,0	608,4226	464,6	605,3035
626,0	605,6698	482,5	612,2743	470,0	615,1159
495,0	616,9363	546,0	615,4045	423,0	617,5579
548,0	659,7712	602,1	616,3147	624,9	600,8968
552,0	537,2494	620,0	543,8317	675,0	673,9348
603,6	618,1240	624,9	616,8919	620,5	627,0151
640,0	620,8435	453,6	828,7465	481,0	577,6201
523,0	595,8463	548,0	634,8073	622,0	617,8465
631,1	610,9423	490,0	615,3934	420,7	542,7994
489,0	623,4298	583,4	628,7800	497,8	616,0150
510,7	590,7403	570,0	596,0350	661,0	617,2138
740,8	602,2732	517,0	615,1492	434,0	617,1583
474,6	513,4510	626,5	640,1353	742,2	616,7920
625,0	618,8011	461,8	605,9029	477,0	554,8429

481,2	613,1512	517,0	608,3560	590,0	615,9484
592,9	503,3611	625,9	615,1492	625,0	618,8344
616,0	608,1784	625,6	621,6871	472,0	617,7799
645,5	581,7493	673,8	616,7476	494,0	616,7476
464,9	545,4967	621,0	614,2279	688,0	617,9353
599,0	568,4182	482,0	617,9464	517,4	593,5708
704,0	614,7940	489,8	619,7668	442,0	615,7153
495,0	571,1155	592,0	613,2844	575,2	616,6144
619,3	559,0720	730,0	608,9998	469,8	599,3872
584,0	612,2410	660,0	601,5073	612,3	622,0756
432,0	609,3661	645,6	614,0725	609,0	617,7799
470,0	614,3056	483,0	632,8870	469,3	580,5838
436,5	548,0497	461,0	687,8431	683,0	620,1109
480,2	544,3090	595,8	602,3287	625,0	618,0352
473,7	612,5407	688,0	618,1906	718,2	619,3006
678,0	606,3469	614,0	611,3197	622,5	615,3601
731,5	644,4532	596,0	622,9747	504,0	609,0886
644,0	619,6447	493,5	559,6603	621,1	619,3561
448,5	600,7747	505,0	598,6213	610,7	640,3795
623,0	614,3944	459,5	616,5922	713,0	637,3492
599,0	566,3203	471,0	610,9978	620,0	614,4277
508,9	608,4004	593,1	616,0150	693,7	653,1334
545,3	618,1795	749,5	592,0390	560,0	617,2138
621,0	795,3910	677,0	614,8717	447,0	626,3824
446,2	619,7557	663,0	541,2454	481,5	616,3258
503,0	471,1822	729,5	617,1139	429,0	614,8717
665,0	595,5688	425,0	613,8394	612,0	602,9614
646,0	625,6498	753,0	616,2481	432,2	609,4660
432,5	618,0796	493,5	553,9216	523,0	554,6209
491,0	596,0905	601,8	614,8384	473,0	604,1158
523,2	594,9694	448,5	614,8606	491,0	614,6275
469,3	355,1206	513,0	615,0937	464,0	613,1623
492,0	355,2649	713,0	599,1874	621,3	616,7254
746,0	738,1927	472,0	616,5922	466,0	630,1120
521,2	613,7062	690,2	615,6709	749,5	619,0786
577,8	615,9484	503,0	608,4448	672,0	620,3218
552,0	624,0292	468,5	613,3954	598,0	606,1027
624,9	616,7254	709,0	725,0614	625,0	612,5962
463,0	614,4277	494,0	521,7427	621,4	894,8026
731,0	565,9651	612,0	614,2501	712,0	626,9818
516,0	618,1795	695,0	618,1351	436,0	631,4884
731,0	584,9128	579,0	615,5155	603,0	625,9828
448,3	607,9564	520,7	601,3630	467,0	618,1795

706,0	646,1515	715,0	569,3617	706,0	595,5799
423,0	642,1777	592,0	598,0885	560,0	543,9316
625,0	620,6881	463,3	617,5912	634,0	616,6144
430,0	602,5618	504,0	597,8332	688,0	616,1260
578,0	616,2259	595,5	165,4216	470,0	597,8332
625,9	613,1401	593,0	622,4086	622,0	617,3803
475,5	718,6234	505,0	537,9709	458,9	616,9918
448,2	596,6344	683,0	618,6568	573,5	520,7992
523,3	477,8866	445,0	708,5557	423,0	539,3362
752,8	539,1364	558,0	618,8455	446,0	488,1985
622,0	616,7476	614,0	605,1592	607,1	641,0677
429,8	580,9279	740,8	584,8129	595,7	489,6082
437,5	611,1865	640,0	612,3520	515,8	634,2856
737,2	612,3964	502,3	534,5854	589,1	612,5518
741,8	769,2838	688,0	615,5377	695,0	616,5922
623,5	604,8595	575,0	795,0913	518,0	554,8318
581,1	549,9811	432,6	553,9549	524,7	615,6376
561,0	616,0483	600,5	674,4787	605,0	617,7577
493,0	615,7375	593,1	597,9664	671,2	616,1593
523,3	616,0483	513,0	599,0098	454,5	608,3893
671,0	607,6900	642,0	627,9808	429,6	620,6104
454,5	616,9807	746,0	500,3863	488,7	617,3470
450,5	610,9978	432,2	609,5992	513,0	597,7444
670,0	615,4045	611,0	631,8436	432,2	674,3233
560,0	595,9684	626,5	609,2995	464,0	618,3349
429,0	613,9393	623,0	582,9037	750,0	627,0706
505,9	609,5881	667,5	630,7336	636,0	480,8725
508,9	507,7012	709,7	626,5600	620,0	623,9404
605,0	613,1512	456,3	614,2501	731,8	615,3490
595,8	612,3964	482,5	636,9718	629,0	598,2883
601,8	593,0269	497,2	590,1742	463,8	608,8111
448,3	607,9453	476,9	617,5024	434,0	615,2713
523,0	634,7629	712,2	587,5768	546,0	332,2879
622,0	611,0533	619,9	619,7890	666,6	617,6356
559,0	608,4448	702,0	630,0121	464,6	572,0146
617,0	618,5347	626,9	597,0118	581,1	667,2304
457,5	617,4025	570,0	615,3490	538,1	609,0886
614,0	595,0582	429,0	614,3056	432,2	492,9382
548,0	630,9556	621,0	615,5821	638,2	617,5024
469,4	534,7741	479,2	628,1806	619,0	616,5145
615,6	614,7829	675,0	600,6526	523,0	616,9918
449,7	612,4408	463,0	469,6060	508,9	606,3247
693,0	628,5802	468,0	656,9518	595,7	548,4604

470,0	620,4328	517,0	629,2906	693,0	600,7192
451,0	616,0150	570,0	618,7789	643,0	658,8610
690,8	619,3117	747,0	622,8193	667,0	629,0686
581,2	612,7516	488,2	602,3953	474,6	613,2400
515,7	616,9807	432,2	598,3438	684,0	614,8273
517,0	507,2794	683,0	615,7930	488,2	604,8817
602,1	618,3682	688,0	616,8253	442,0	610,9423
454,5	606,5911	639,0	637,0273	464,0	615,4156
629,0	622,0534	656,0	622,0090	581,0	611,2309
609,0	594,1924	688,0	616,3147	482,8	618,7012
590,0	609,3772	550,0	596,8675	602,0	693,7594

Fonte: Autoria própria (2023).