



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TRANSPORTES
CURSO DE ENGENHARIA CIVIL

NATHANIEL SILVA FONSECA

**ESTIMATIVAS DO NÍVEL DE SEVERIDADE DOS ACIDENTES ENVOLVENDO
MOTOCICLISTAS EM VIAS URBANAS: UMA ANÁLISE DAS VARIÁVEIS
EXPLICATIVAS COM REDES NEURAS ARTIFICIAIS**

FORTALEZA

2013

NATHANIEL SILVA FONSECA

**ESTIMATIVAS DO NÍVEL DE SEVERIDADE DOS ACIDENTES ENVOLVENDO
MOTOCICLISTAS EM VIAS URBANAS: UMA ANÁLISE DAS VARIÁVEIS
EXPLICATIVAS COM REDES NEURASIS ARTIFICIAIS**

Monografia apresentada ao Curso de Engenharia Civil do Centro de tecnologia da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Engenheiro Civil.

Orientador: Carlos Augusto Uchôa da Silva

Co-orientador: Flávio José Craveiro Cunto

FORTALEZA

2013

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

F745e Fonseca, Nathaniel Silva.

Estimativas do nível de severidade dos acidentes envolvendo motociclistas em vias urbanas : uma análise das variáveis explicativas com redes neurais artificiais / Nathaniel Silva Fonseca. – 2013.
58 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia, Curso de Engenharia Civil, Fortaleza, 2013.

Orientação: Prof. Dr. Carlos Augusto Uchôa da Silva.

Coorientação: Prof. Dr. Flávio José Craveiro Cunto.

1. Níveis de severidade. 2. Modelos de previsão de acidentes. 3. Motocicletas. 4. Redes neurais artificiais. I. Título.

CDD 620

NATHANIEL SILVA FONSECA

**ESTIMATIVAS DO NÍVEL DE SEVERIDADE DOS ACIDENTES ENVOLVENDO
MOTOCICLISTAS EM VIAS URBANAS: UMA ANÁLISE DAS VARIÁVEIS EXPLICATIVAS
COM REDES NEURAIS ARTIFICIAIS**

Monografia apresentada ao Curso de Engenharia Civil do Centro de tecnologia da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Engenheiro Civil.

Aprovado em __/__/_____.

BANCA EXAMINADORA

Prof. Carlos Augusto Uchôa da Silva, Dr. (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Flávio José Craveiro Cunto, Dr. (Co-orientador)
Universidade Federal do Ceará (UFC)

Profa. Maria Elisabeth Pinheiro Moreira, Dra. (Examinadora)
Universidade Federal do Ceará (UFC)

A Deus.

A meus pais, Roque e Claudenice.

A minha esposa, Ana Raquel.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por está sempre ao meu lado, guiando os meus passos e me fortalecendo perante a sua sabedoria.

Ao Prof. Uchôa pela orientação neste trabalho, pela sua confiança e por sua amizade.

Ao Prof. Flávio Cunto pela sua dedicação e companheirismo.

A minha mãe, Claudenice e ao meu pai, Roque por todas as palavras de incentivo dadas a mim e por acreditarem no meu potencial.

A minha irmã, Nayara pelo apoio e por acreditar em mim desde sempre.

A minha esposa Ana Raquel por seu amor, carinho e dedicação.

Ao meu tio Kléber Augustopelo incentivo e por toda a confiança na minha capacidade.

Aos amigos Daniel Oliveira, Carlos Bruno e tantos outros do Curso de Engenharia Civil pela ajuda, amizade e companheirismo.

A minha tia Ana Paula por estar sempre ao meu lado em todos os momentos da minha vida e pelas palavras de conforto dirigidas a mim.

A minha tia Cleane por acreditar nos meus sonhos e por me ajudar a realizá-los.

A toda a minha família e amigos que me deram total apoio e conforto nos momentos de tribulação que apareceram no decorrer dessa jornada.

Ao Pastor José Roberto e pela sua esposa Diaconisa Natália, que me trouxeram palavras vindas de Deus para tocar ao meu coração e que fizeram eu não desistir dos meus planos e meus projetos.

RESUMO

Os acidentes de trânsito em vias urbanas são considerados como fenômenos de grande variabilidade e, têm chamado a atenção de pesquisadores da área de segurança viária tentativa, tanto de se compreender quais são as variáveis que concorrem para sua ocorrência, quanto para estimar a probabilidade desses eventos. Normalmente utiliza-se modelos de regressão neste tipo de predição, geralmente não linear, conhecidos como Modelos de Previsão de Acidentes (MPA), que tendem a correlacionar os acidentes com os atributos da via. O principal objetivo deste trabalho é construir um modelo baseado em inteligência artificial, dedicado à estimação das variáveis que mais contribuem na ocorrência de diferentes níveis de severidade dos acidentes de trânsito com motocicletas, nas vias arteriais da cidade de Fortaleza. O modelo desenvolvido foi calibrado, validado e testado usando uma amostra de 1137 observações e como variáveis explicativas; foram utilizados atributos geométricos, climáticos, operacionais e funcionais da via. Embora os MPA apresentem bons resultados em suas predições, o modelo aqui implementado, baseado em Redes Neurais Artificiais (RNAs) obteve resultados consideravelmente promissores.

Palavras-chave: Níveis de severidade. Modelos de Previsão de Acidentes. Motocicletas. Redes Neurais Artificiais.

ABSTRACT

Traffic accidents on urban roads are considered as phenomena of great variability and, have drawn attention of researchers in the field of road safety in an attempt, both to understand what are the variables that contribute to its occurrence, and to estimate the likelihood of these events. Normally are used regression models, generally not linear; known as Accident Prediction Models (MPA), which tend to correlate the accident with the attributes of the road. The main objective of this work is to build a model based on artificial intelligence dedicated to the estimation of the variables that contribute most in the occurrence of different levels of severity of traffic accidents with motorcycles, the arterial roads in the city of Fortaleza. The model was calibrated, validated and tested using a sample of 1137 observations and explanatory variables; geometric attributes were used, climatic, operational and functional attributes of the route. Although the MPA are performing well in their predictions, the model implemented here, based on Artificial Neural Networks (ANN) gave considerably promising results.

Keywords: Levels of severity. Accident Prediction Models. Motorcycles. Artificial Neural Networks.

LISTA DE ILUSTRAÇÕES

| | |
|---|----|
| Figura 1 - Neurônio Biológico..... | 25 |
| Figura 2 - Representação do modelo de um neurônio artificial. | 26 |
| Figura 3 - Desenho esquemático da estrutura de uma Rede Neural..... | 27 |
| Figura 4 - Interface gráfica NNTOOL..... | 34 |
| Figura 5 - Fluxograma da metodologia adotada..... | 35 |
| Figura 6 - Topologia adotada para as 5 variáveis base (5:11:5)..... | 47 |
| Figura 7 - Topologia adotada para 8 variáveis (8:17:5) | 47 |
| Figura 8 - Topologia adotada para 10 variáveis (10:11:5) | 48 |
| Figura 9 - Comparação entre os algoritmos TRAINLM x TRAINGDX (5:11:5)..... | 48 |
| Figura 10 - Comparação entre os Algoritmos TRAINLM x TRAINGDX (8:17:5) | 49 |
| Figura 11 - Comparação entre os Algoritmos TRAINLM x TRAINGDX (10:11:5) | 49 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1: Consumo de tempo na mobilidade (bilhões de horas/ano) – 2011..... | 18 |
| Tabela 2: Consumo de combustível líquido na mobilidade | 18 |
| Tabela 3 - Tipos de Funções de ativação..... | 28 |
| Tabela 4 - Codificação de dados das variáveis..... | 40 |
| Tabela 5 - Variáveis envolvidas na modelagem..... | 41 |
| Tabela 6 - Organização das 5 variáveis base..... | 42 |
| Tabela 7 - Organização das 8 variáveis explicativas..... | 42 |
| Tabela 8 - Organização das 10 variáveis explicativas..... | 43 |
| Tabela 9 - Desempenho das principais Topologias com 5, 8 e 10 variáveis..... | 45 |
| Tabela 10 - Desempenho das RNAs e tempo de processamento | 50 |
| Tabela 11 - Percentuais de acertos da RNA para os conjuntos de teste de cada grupo de variáveis..... | 50 |

LISTA DE ABREVIATURAS E SIGLAS

Abraciclo – Associação Brasileira dos Fabricantes de Motocicletas, Ciclomotores, Motonetas, Bicicletas e Similares.

ANTP – Associação Nacional de Transporte Público

CEFTRU – Centro Interdisciplinar de Estudos em Transportes

CTB – Código de Trânsito Brasileiro

DENATRAN – Departamento Nacional de Trânsito

DETRAN – Departamento Estadual de Trânsito

IPCVs – Indicadores Primários de Segurança Viária

ISSVs – Indicadores Secundários de Segurança Viária

MATLAB – *Matrix Laboratory*

MLG – Modelos Lineares Generalizados

MLP – *Multilayer Perceptron*

MPA – Modelos de Previsão de Acidentes

MSE – *Mean of Squared Error* ou Quadrado Médio dos Erros

NNTOOL – *Neural Network Tool*

RNAs – Redes Neurais Artificiais

SIAT-FOR – Sistema de Informação de Acidentes de Trânsito de Fortaleza

SNT – Semana Nacional de Trânsito

SUMÁRIO

| | |
|---|----|
| AGRADECIMENTOS | 5 |
| RESUMO | 6 |
| ABSTRACT | 7 |
| LISTA DE ILUSTRÇÕES..... | 8 |
| LISTA DE TABELAS | 9 |
| 1 INTRODUÇÃO..... | 13 |
| 1.1 Contextualização..... | 13 |
| 1.2 Problema e questões de pesquisa | 14 |
| 1.2.1 Problema de pesquisa | 14 |
| 1.2.2 Questões de pesquisa | 15 |
| 1.3 Objetivos..... | 15 |
| 1.3.1 Objetivo geral | 15 |
| 1.3.2 Objetivos específicos..... | 15 |
| 1.4 Estrutura do trabalho..... | 16 |
| 2 UTILIZAÇÃO DE MOTOCICLETAS EM MEIO URBANO..... | 17 |
| 2.1 Considerações iniciais..... | 17 |
| 2.1.1 Características geométricas do veículo..... | 17 |
| 2.1.2 Tempo de viagem, custo e consumo..... | 18 |
| 2.1.3 Perfil do condutor | 19 |
| 2.2 Acidentes de trânsito como medida de desempenho da segurança viária | 19 |
| 2.2.1 Classificação e conceituação dos tipos de acidentes de trânsito | 19 |
| 2.2.2 Níveis de severidade dos acidentes de trânsito..... | 20 |
| 3 REDES NEURAIS ARTIFICIAIS (RNAs) | 23 |
| 3.1 Considerações iniciais..... | 23 |
| 3.1.1 Histórico das RNAs | 24 |
| 3.1.2 O neurônio biológico | 24 |
| 3.1.3 O neurônio artificial..... | 25 |
| 3.1.4 Estrutura das RNAs | 26 |
| 3.1.5 Topologia da Rede..... | 27 |
| 3.1.6 Funções de ativação..... | 28 |
| 3.1.7 Aprendizagem em RNAs..... | 29 |
| 3.1.8 Perceptron Multi Camadas (MLP) e o Algoritmo Backpropagation..... | 30 |

| | | |
|-------|---|----|
| 4 | MATERIAIS E MÉTODO..... | 31 |
| 4.1 | Considerações iniciais..... | 31 |
| 4.2 | Dados Utilizados para Elaboração do Trabalho | 32 |
| 4.3 | Aplicativos Computacionais | 32 |
| 4.3.1 | Microsoft Excel | 33 |
| 4.3.2 | MATLAB | 33 |
| 4.4 | Metodologia Adotada | 34 |
| 4.4.1 | Revisão de literatura | 35 |
| 4.4.2 | Pesquisa e obtenção dos dados | 36 |
| 4.4.3 | Tratamento dos dados | 36 |
| 4.4.4 | Pré-processamento..... | 36 |
| 4.4.5 | Modelagem neural (RNAs) | 37 |
| 4.4.6 | Análise dos resultados | 37 |
| 4.4.7 | Modelo proposto..... | 38 |
| 5 | RESULTADOS E DISCUSSÕES..... | 39 |
| 5.1 | Considerações iniciais..... | 39 |
| 5.2 | Análise exploratória dos dados | 39 |
| 5.3 | Árvore de teste e desempenho da modelagem..... | 41 |
| 5.4 | Considerações finais | 51 |
| 6 | CONCLUSÕES E RECOMENDAÇÕES | 52 |
| 6.1 | Considerações iniciais..... | 52 |
| 6.2 | Principais Conclusões | 52 |
| 6.2.1 | Quanto ao método e a modelagem neural | 53 |
| 6.2.2 | Quanto às incertezas dos modelos gerados..... | 53 |
| 6.2.3 | Quanto ao modelo proposto para o grupo com 5 variáveis | 54 |
| 6.2.4 | Quanto aos modelos de 8 variáveis e 10 variáveis explicativas | 54 |
| 6.3 | Recomendações e sugestões para trabalhos futuros..... | 55 |
| | REFERÊNCIAS | 56 |

1 INTRODUÇÃO

1.1 Contextualização

O aumento dos acidentes de trânsito em vias urbanas no estado do Ceará vem provocando uma grande perda de vidas humanas nos últimos anos. Adicionalmente, o número total de acidentes ao final do ano de 2010 foi de 26.507, desse total, 1.703 foram vítimas fatais, em média cinco mortes/dia. Na maioria dos casos as vítimas são jovens e adultos que se encontram em plena produtividade (DETRAN, 2010a).

Anualmente, mais de três quartos de milhão de pessoas são mortas no mundo, enquanto o número de vítimas lesionadas e as que apresentam alguma deficiência superam em conjunto a marca de 40 milhões (ALI, 2010). Segundo pesquisa realizada na Região Metropolitana do Ceará, mais de 50% dos acidentes aconteceram em interseções da capital, e destas, mais da metade foram consideradas severas (DETRAN, 2010b).

Esse cenário é bastante preocupante quando o veículo em questão é uma motocicleta, pois o passageiro ou até mesmo o próprio condutor, estão expostos e muito vulneráveis. Mesmo com essa insegurança as motocicletas vêm sendo o transporte escolhido por grande parte da população devido ao custo, praticidade e agilidade em ambientes com pouco espaço.

Vários programas e ações das autoridades competentes têm a finalidade de reduzir a quantidade e a gravidade dos acidentes viários em Fortaleza, bem como a ação educativa da Semana Nacional de Trânsito (SNT) de 2007, tendo participação dos mais diversos segmentos da sociedade. No que se refere à eficácia e à eficiência, esses programas e ações, dependem de uma avaliação cuidadosa das informações dos acidentes ocorridos, pois trabalhar com a análise dos acidentes é uma tarefa complexa e demanda bastante tempo, mesmo se os dados já estiverem coletados, armazenados e disponíveis em um banco de dados.

Os acidentes de trânsito possuem diferentes níveis de severidade e são considerados como eventos aleatórios raros (CUNTO e NODARI, 2011). Devido a esses fatores, faz-se necessário uma análise minuciosa de suas variáveis explicativas com o intuito de determinar dentre elas, quais possuem maior importância na explicação da ocorrência dos diferentes níveis de severidade.

A maior parte dos estudos relacionados à segurança viária usa modelos estatísticos tradicionais. Poucas são as pesquisas que utilizam modelos estatísticos mais sofisticados, tais como aquelas realizadas por distribuição de Poisson e Binominal Negativa (BONNISON e McCOY, 1993; PERSAUD e MUCSI, 1995).

Recentemente, a técnica de redes neurais artificiais, do inglês *artificial neural networks*, nesta investigação também tratada por RNAs, tem sido utilizada na modelagem e na análise de variáveis tanto na infraestrutura quanto no planejamento e operações dos sistemas de transportes, pois esta técnica permite o uso de dados: não lineares; incompletos e até mesmo contraditórios, sendo estas características também encontradas nas análises em questão. Dentre os estudos relacionados à segurança viária podem ser citados os trabalhos de AWAD e JANSON (1998), que utilizaram RNAs na previsão de acidentes de caminhões em acessos de vias expressas, e MUSSONE *et al.* (1999) cujo objetivo foi identificar os parâmetros mais significativos que determinam a possibilidade da ocorrência de acidentes em uma interseção urbana.

O processo de escolha das variáveis para a realização de um modelo de previsão em segurança viária não possui critérios objetivos, na maioria dos casos depende muito da experiência do modelador, dos objetivos do estudo e também da disponibilidade de informações sobre a situação da via (CUNTO *et al.*, 2012). Devido a esses fatores, torna-se bastante complexa a definição de quais variáveis explicativas devem ser empregadas em um modelo com esta finalidade.

1.2 Problema e questões de pesquisa

1.2.1 Problema de pesquisa

Existe uma dificuldade em se determinar quais são as variáveis que mais contribuem para explicar a ocorrência de diferentes níveis de severidade em acidentes de trânsito com motocicletas em interseções urbanas.

1.2.2 Questões de pesquisa

Considerando o problema em questão, algumas questões aparecem, sobretudo com o intuito de obter um maior entendimento do problema e da busca de possíveis alternativas de solucioná-lo. Tais questões são as seguintes:

- a) Quais são os principais tipos de acidentes encontrados em interseções?
- b) De que maneira a técnica de Redes Neurais Artificiais (RNAs) pode facilitar na identificação de variáveis explicativas/preditoras de nível de severidade em acidentes de trânsito em interseções urbanas?
- c) Qual função de ativação melhor se adapta a modelagem neural?

1.3 Objetivos

1.3.1 Objetivo geral

O objetivo geral deste trabalho é implementar (calibrar, validar e testar) um modelo baseado em redes neurais artificiais dedicado à predição de níveis de severidade em acidentes de trânsito envolvendo motocicletas em interseções urbanas e, com isso identificar quais são as variáveis preditoras que melhor explicam a ocorrência desses diferentes níveis de severidade.

1.3.2 Objetivos específicos

Através do desenvolvimento das questões de pesquisa pode-se chegar a certos objetivos específicos de cada questão, sendo eles:

- a) Identificar os principais tipos de acidentes, envolvendo motocicletas, encontrados em interseções urbanas;

- b) Calibrar, validar e testar um modelo neural que permita identificar, de forma rápida e prática, quais as variáveis que mais contribuem na determinação do nível de severidade dos acidentes de trânsito nas interseções urbanas;
- c) Identificar a melhor função de ativação para a modelagem neural desse fenômeno comparando os desempenhos obtidos por cada função.

1.4 Estrutura do trabalho

A monografia está dividida em cinco capítulos, conforme a seguinte estrutura:

Capítulo 1 (introdução): é composto pela problemática do tema de estudo, pelas questões de pesquisa, pelo objetivo geral e pelos objetivos específicos.

Capítulo 2 (Utilização de motocicletas em meio urbano): neste capítulo apresenta-se uma descrição sobre o uso das motocicletas em vias urbanas.

Capítulo 3 (Redes Neurais Artificiais): neste capítulo apresenta-se uma descrição sobre a fundamentação histórica das redes neurais artificiais, suas principais vantagens, sobretudo quando aplicadas ao estudo de fenômenos não lineares, bem como de seus principais componentes.

Capítulo 4 (Materiais e Métodos): neste capítulo apresentam-se descrições sobre a área de estudo utilizada para promover a estimação dos níveis de severidade, os aplicativos computacionais utilizados e a metodologia proposta e aplicada no desenvolvimento desta investigação.

Capítulo 5 (Resultados e Discussões): neste capítulo apresentam-se os resultados obtidos no processo da calibração, validação e teste do modelo que melhor se adequou ao fenômeno estudado e algumas discussões relativas aos resultados encontrados.

Capítulo 6 (Conclusões e Recomendações) neste capítulo são apresentadas as conclusões geradas a partir do estudo realizado, bem como as sugestões e recomendações para trabalhos futuros.

2 UTILIZAÇÃO DE MOTOCICLETAS EM MEIO URBANO

2.1 Considerações iniciais

A utilização de motocicletas em meio urbano tem aumentado bastante no decorrer dos anos. Segundo a Associação Brasileira dos Fabricantes de Motocicletas, Ciclomotores, Motonetas, Bicicletas e Similares (Abraciclo) em 2013 houve um aumento 22,3% nas vendas de motocicletas no mês de agosto em relação ao mês anterior.

Segundo HOLZ (2013), isso ocorre devido a duas variáveis principais: a atração ao uso e aos aspectos de operação. Um dos fatores que estimulam a atração dos usuários é o custo de obtenção do veículo, facilidade de estacionamento e circulação no trânsito que, hoje em dia, está cada vez mais congestionada. Outro fator bastante comum que incentiva a utilização das motocicletas é a falta de qualidade ofertada pelo transporte público.

No que se refere ao uso da motocicleta, pode-se apontar vários atributos, bem como, as características geométricas do veículo, o tempo de viagem, o custo, o consumo de combustível e também o perfil do condutor. Cada um desses atributos será explicitado nos tópicos a seguir.

2.1.1 *Características geométricas do veículo*

Uma grande vantagem da motocicleta em relação ao automóvel está vinculada à suas dimensões, cujos valores, para um modelo mais popular, são (2,00 x 0,75 x 1,05) metros. A área e o volume ocupado pela motocicleta são aproximadamente 78% menores se comparados àqueles característicos de um automóvel também popular (3,8 x 1,77 x 1,35) metros possibilitando, assim, uma maior facilidade em manobras, maior fluidez no sistema viário e apresentando mais opções para estacionar em lugares de difícil acesso (HOLZ, 2013).

As características geométricas das motocicletas são importantes fatores que promovem a utilização desse meio de transporte. Características essas, que resultam em uma maior facilidade em manobras perante os veículos com maiores proporções (MUSSO *et al.*, 2010).

2.1.2 Tempo de viagem, custo e consumo

O tempo de viagem gasto pelo transporte coletivo é duas vezes maior se comparado ao tempo gasto por veículos individuais. Em se tratando de motocicletas, esse número cai ainda mais como mostra a Tabela 1.

Tabela 1: Consumo de tempo na mobilidade (bilhões de horas/ano) – 2011

| Modo | Tempo (Bilhões de horas /ano) | Participação % |
|---------------------------------------|--------------------------------------|-----------------------|
| Ônibus municipal | 7,0 | 32 |
| Ônibus metropolitano | 2,8 | 13 |
| Trilhos | 0,8 | 4 |
| <i>Subtotal coletivo</i> | 10,7 | 49 |
| Auto | 4,4 | 20 |
| Moto | 0,4 | 2 |
| <i>Subtotal individual motorizado</i> | 4,8 | 22 |
| Bicicleta | 0,8 | 4 |
| A pé | 5,6 | 26 |
| <i>Subtotal não motorizado</i> | 6,4 | 29 |
| Total | 21,9 | 100 |

Fonte: ANTP (2012)

Em relação ao custo, as motocicletas possuem um valor de mercado relativamente baixo quando comparadas aos automóveis. Devido a esse fato, muitos usuários tendem a ser atraídos por este modal (HSU *et al.*, 2003). Outro fator que chama bastante a atenção é que o consumo médio de combustível das motocicletas é cerca de 80 % menor que o combustível gasto pelos automóveis para um período de um ano em uma faixa de população de 60 a 100 mil habitantes, segundo o relatório feito pela Associação Nacional dos Transportes Públicos (ANTP) em dezembro de 2012 referentes ao ano de 2011, como mostra a Tabela 2.

Tabela 2: Consumo de combustível líquido na mobilidade

| Faixa de População | Diesel (l) | Gasolina – Auto (l) | Gasolina – Moto (l) |
|---------------------------|-------------------|----------------------------|----------------------------|
| >1 milhão | 2,42 | 8,18 | 0,32 |
| 500-1000 mil | 0,27 | 1,96 | 0,09 |
| 250-500 mil | 0,33 | 1,19 | 0,10 |
| 100-250 mil | 0,30 | 0,83 | 0,11 |
| 60-100 mil | 0,10 | 0,21 | 0,04 |
| Total | 3,43 | 12,37 | 0,66 |

Fonte: ANTP (2012)

2.1.3 Perfil do condutor

A Abraciclo, que é a associação que representa as indústrias de motocicletas no Brasil, vem realizando pesquisas junto aos compradores nos últimos anos. A maioria das compras de motocicletas, cerca de 75%, são feitas por homens, sendo que desse total, 70% são homens com idade inferior a 40 anos (HOLZ, 2013). O principal motivo de o usuário optar em comprar uma motocicleta em 40% dos casos, é para substituir o transporte coletivo, seguido de lazer com 19%, como instrumento de trabalho tem-se 16%, substituir o automóvel são 10%, e outros motivos 15% (Abraciclo, 2013).

2.2 Acidentes de trânsito como medida de desempenho da segurança viária

O fenômeno acidente de trânsito ocorre quando há o conflito envolvendo um ou mais veículos, motorizados ou não, em movimento em uma determinada via, provocando ferimentos em pessoas e/ou danos físicos em veículos e/ou objetos de outra natureza (FERRAZ *et al.*, 2008).

Segundo o Centro Interdisciplinar de Estudos em Transporte (CEFTRU) em 2002, os fatores que contribuem para a ocorrência de acidentes de trânsito estão relacionados tanto ao comportamento do homem, à via e ao meio ambiente quanto às condições operacionais do veículo. Esse tipo de classificação toma por base a relação que ocorre entre as condições geométricas, ambientais, temporais e do tráfego além do estado em que se encontram os veículos, dentre outras questões.

2.2.1 Classificação e conceituação dos tipos de acidentes de trânsito

Com base nos estudos realizados por CEFTRU (2002), os acidentes de trânsito são classificados da seguinte forma:

- a) Colisão: acidentes em que há o impacto de veículos em movimento, podendo ser subdividida em:

- Lateral: colisão lateral entre veículos que transitem na mesma via, podendo estar no mesmo sentido ou em sentidos opostos;
 - Transversal: colisão entre veículos que transitam em direções que se cruzam, perpendicularmente ou não;
 - Frontal: colisão entre veículos que transitam na mesma via em sentidos opostos;
 - Traseira: colisão entre veículos que transitam na mesma via no mesmo sentido;
- b) Choque: é definido como o impacto de um veículo em movimento contra qualquer obstáculo fixo, seja ele, um poste, uma árvore, um veículo estacionado, um muro ou um objeto de outra natureza;
- c) Capotamento: é definido quando o veículo rotaciona sobre si mesmo em qualquer direção, ficando com as rodas para cima;
- d) Tombamento: quando o veículo tomba sobre sua lateral, imobilizando-se;
- e) Engavetamento: é uma colisão tipo traseira, que envolve três ou mais carros;
- f) Atropelamento: é definido quando há a presença de pedestres, ou animais, atingidos por um veículo motorizado ou não.

2.2.2 Níveis de severidade dos acidentes de trânsito

Os acidentes de trânsito possuem diferentes níveis de severidade e esses níveis podem ser classificados de várias maneiras. Para tornar essa investigação mais simplificada será utilizada a caracterização dos níveis de severidade com base nos dados estatísticos levantados pelo DETRAN e pelo Departamento Nacional de Trânsito (DENATRAN).

O DENATRAN (1987) instituiu a Unidade Padrão de Severidade (UPS), cujo valor, expresso em UPS, é resultado da soma dos produtos do número de ocorrências por severidade pelo peso atribuído a respectiva severidade. Os pesos atribuídos para cada nível de severidade é demonstrado a seguir:

- a) Acidentes com danos materiais – Peso 1;
- b) Acidentes com vítimas feridas – Peso 5;
- c) Acidentes com vítimas fatais – Peso 13.

Desta forma o grau de severidade de um acidente e conseqüentemente de um determinado local, pode ser expresso pela Equação 1.

$$N^{\circ} \text{ de UPS} = \text{ADM} \times 1 + \text{AVF} \times 5 + \text{AVFa} \times 13 \quad (1)$$

Onde: ADM= acidentes com danos materiais;

AVF = acidentes com vítimas feridas;

AVFa = acidentes com vítimas fatais;

A utilização dessa técnica possui as mesmas vantagens da técnica utilizada na obtenção do número de acidentes, com o adicional de priorizar aqueles pontos ou segmentos cujos resultados apresentam maior gravidade em termos de vítimas (BERNARDINO, 2007).

Os níveis de severidade seguem uma hierarquia, começando pelo nível mais baixo encontra-se: (i) acidentes com danos materiais (ii) acidentes com vítimas feridas (iii) acidentes com vitimas fatais. A seguir será caracterizado cada nível de severidade com base nos dados levantados pelo DETRAN.

- Acidentes com danos materiais

Com base nos dados estatísticos, esse tipo de acidente possui uma maior frequência e é caracterizado quando existe uma pequena colisão entre os veículos envolvidos, suficiente para obter prejuízos financeiros, mas sem que haja a presença de vítimas feridas no local do acidente.

Muitas vezes esse tipo de ocorrência não é registrado, pois os envolvidos no acidente já realizam uma negociação e se dispersam do local antes da chegada da perícia para compreensão de como aconteceu o acidente.

- Acidentes com vítimas feridas

Esse tipo de acidente pode ser definido quando ocorre a presença de vítimas feridas no local envolvendo somente os condutores dos veículos.

- Acidentes com vítimas fatais

É caracterizado com a presença de vítimas fatais envolvendo os condutores ou até mesmo os pedestres.

Uma maneira de determinar o risco de ocorrência e/ou de severidade de acidentes de trânsito; verificar se os veículos atendem aos padrões de segurança; quantificar o nível de segurança de um via e monitorar a mudança de comportamento da sociedade, é a utilização de indicadores de desempenho viário (DIÓGINES e LINDAU, 2004).

Esses indicadores podem ser classificados em indicadores primários e secundários. Os Indicadores Primários de Segurança Viária (IPCVs), também conhecidos como indicadores dos impactos ou indicadores estatísticos, são definidos como medidas que estão relacionadas diretamente aos problemas da insegurança nas vias, sendo representados pela frequência de acidentes com variáveis de exposição (LÖTTER, 2000). Já os Indicadores Secundários de Segurança Viária (ISSVs) medem fatores que estão correlacionados aos acidentes, bem como, o comportamento dos usuários, aspectos legais e tecnologia veicular (LÖTTER, 2000).

3 REDES NEURAIS ARTIFICIAIS (RNAs)

3.1 Considerações iniciais

Redes Neurais Artificiais (RNAs), são técnicas computacionais que apresentam um modelo matemático inspirado no funcionamento da estrutura neural de organismos biológicos, inteligentes e que desenvolvem conhecimento/aprendizagem através da experiência adquirida. As RNAs são capazes de simular conexões sinápticas que acontecem no cérebro humano, onde cada neurônio do cérebro humano tem a capacidade de criar até 10.000 sinapses com outros neurônios.

Por analogia ao neurônio biológico, o neurônio artificial recebe sinais de entrada e distribui para a saída ou para outros neurônios da camada seguinte. O sinal de entrada é então multiplicado pelos pesos sinápticos, gerando um somatório de todas as entradas ponderadas. Esse somatório é considerado a função de ativação para que se possa transmitir o sinal para outros neurônios ou para saída(DÂNGELO *et al.*, 2011).

As redes neurais podem ser consideradas como sistemas paralelos distribuídos compostos por unidades de processamento simples, através do qual é possível armazenar conhecimento experimental adquirido e torná-lo disponível para uso posterior (HAYKIN, 2001). Algumas características importantes em uma rede neural artificial são:

- a) Robustez e tolerância a falhas: eliminação de alguns neurônios não afeta substancialmente o seu desempenho global;
- b) Flexibilidade: pode ser ajustada a novos ambientes por meio de um processo de aprendizagem, sendo capaz de aprender novas ações com base na informação contida nos dados de treinamento;
- c) Processamento de informações contraditórias: mesmo que a informação fornecida esteja incompleta, com ruído, ainda é possível adquirir um raciocínio correto;
- d) Paralelismo: grande número de neurônios está ativo ao mesmo tempo;
- e) Trabalha muito bem com dados de entrada e saída não lineares;
- f) Adaptável mediante a uma algoritmo de aprendizado.

Muitos pesquisadores vêm utilizando essa técnica em diversas áreas de atuação para resolver diferentes problemas através de diversas aplicações (RIBEIRO, 2013).

3.1.1 Histórico das RNAs

A primeira pesquisa relacionada ao desenvolvimento de um neurônio artificial biológico teve início na década de 40 do século passado, mais precisamente no ano de 1943, o neurofisiologista Warren McCulloch e o matemático Walter Pitts fizeram uma analogia entre células vivas e o processo eletrônico, onde o neurônio possuía apenas uma saída, que era função do valor de suas diversas entradas. Ao longo do tempo, pesquisadores começaram a desenvolver estudos relacionados com a forma de aprendizagem das redes biológicas e artificiais devido à experiência adquirida e sobre sua capacidade de executar determinadas funções (BRAGA *et al.*, 2000).

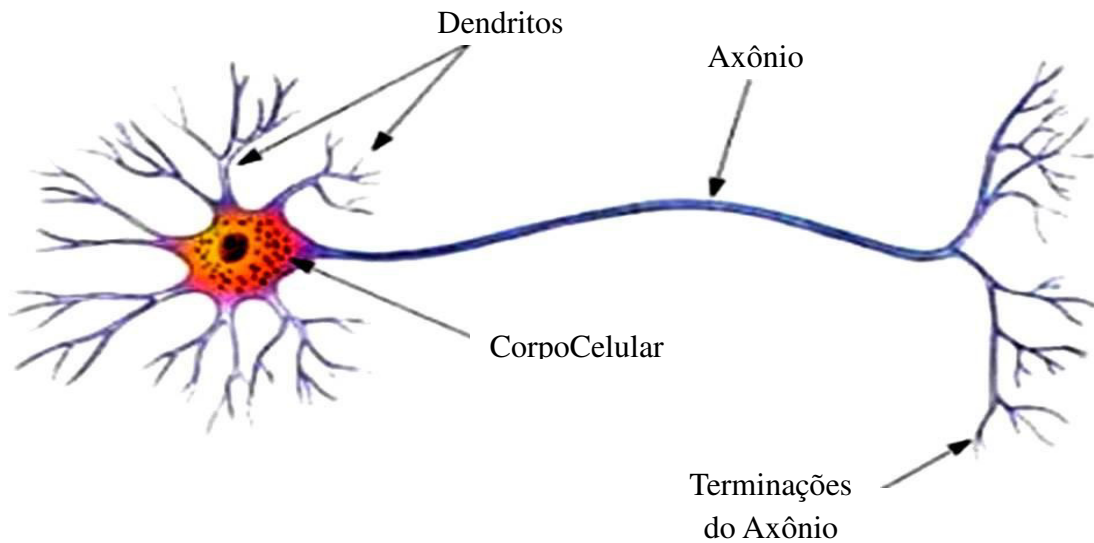
3.1.2 O neurônio biológico

O neurônio biológico é composto pelo corpo celular ou soma, pelos dendritos e axônios. O corpo celular do neurônio é responsável por coletar e combinar informações vindas de outros neurônios. Já os dendritos recebem estímulos e o axônio transmite esses estímulos. Um neurônio recebe informações através dos dendritos, onde os sinais são ponderados e enviados ao axônio, cada sinal pode ser ampliado ou reduzido, por que a cada passagem por um neurônio está associada a um peso que é multiplicado ao sinal (LUDWIG e COSTA, 2007).

O corpo celular pode realizar operações de grandes complexidades em relação aos sinais que chegam até ele. Todavia, essas operações podem ser aproximadas por um somatório simples de sinais vindos dos neurônios vizinhos (HAYKIN, 2001).

O processo de ligação entre um axônio e um dendrito é conhecido como sinapse; portanto para que o sinal possa ser passado para o próximo neurônio, deve ocorrer sinapse entre as células nervosas. Mas para que isso aconteça, esse sinal deve ser superior a certo limite. Segundo LUDWIG e COSTA (2007), o sinal deve ser superior a 50 mV (limiar de disparo), caso isso não ocorra, o sinal é bloqueado. A Figura 1 demonstra como é composto um neurônio biológico.

Figura 1 - Neurônio Biológico.

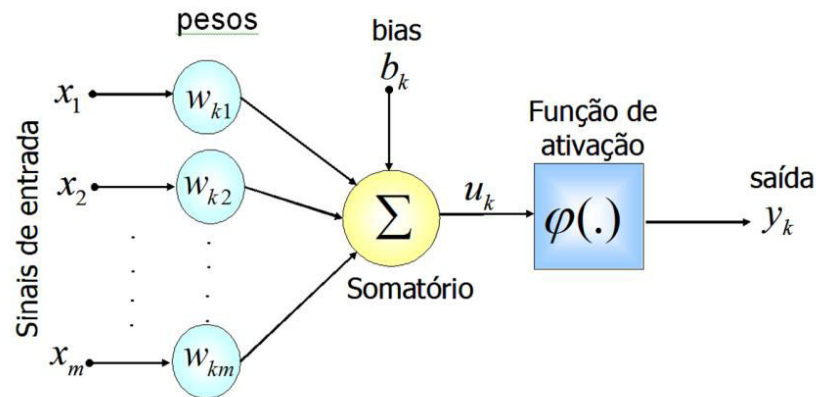


Fonte: (Packter, 2005)

3.1.3 O neurônio artificial

O modelo matemático que representa o funcionamento de um neurônio biológico de forma artificial foi denominado de *Perceptron*, esse neurônio representa uma aproximação simplificada com base na neurofisiologia, porém possui várias características básicas de um neurônio biológico e proporciona sua implementação computacional (DA SILVA, 2003). O neurônio artificial apresenta os seguintes componentes: X_m são as entradas da rede ou as variáveis de interesse, W_k são os pesos associados a cada entrada, b_k é denominado de bias, u_k é a combinação linear dos sinais de entrada, $\varphi(.)$ é a função de ativação, Y_k é a saída do neurônio, bem como mostra a Figura 2.

Figura 2 - Representação do modelo de um neurônio artificial.



Fonte: (DA SILVA, 2003)

RIBEIRO (2013), afirma que é nos pesos que reside todo o conhecimento adquirido pela rede, e que eles são parâmetros ajustáveis que mudam e se adaptam à medida que o conjunto de treinamento é apresentado à rede neural. Segundo DA SILVA (2003), o processo de aprendizado supervisionado em uma RNA com pesos, resulta em sucessivos ajustes dos pesos sinápticos de tal modo que a saída da rede seja mais próxima possível da resposta desejada.

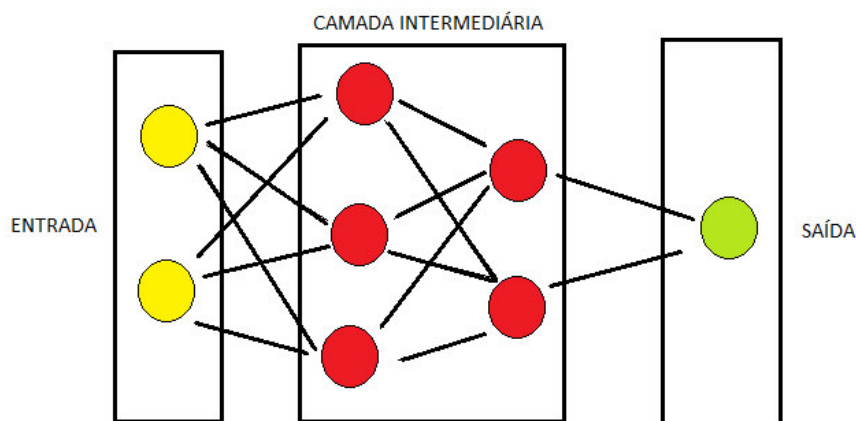
Normalmente, a amplitude da saída do neurônio está no intervalo $[0,1]$ ou alternativamente $[-1,1]$. Está presente no modelo neural um termo denominado de *bias*, valor esse aplicado externamente, representado pelo símbolo b_k . Esse termo possui um efeito de acréscimo ou de decréscimo da função de ativação na entrada da rede, dependendo se esta é positiva ou negativa, respectivamente (DA SILVA, 2003).

3.1.4 Estrutura das RNAs

A estrutura das RNAs pode ser definida como várias unidades de cálculo, onde as informações são processadas em paralelo (RIBEIRO, 2013). Nela também se encontram presentes unidades de processamento chamadas de “neurônios”, que estão conectados entre si através de ligações que determinam o fluxo das informações que passam pela rede, do mesmo

modo que acontece com a rede neural biológica. A Figura 3 apresenta um esquema da estrutura de uma rede neural artificial.

Figura 3 - Desenho esquemático da estrutura de uma Rede Neural.



Fonte: Adaptado (RIBEIRO, 2013).

3.1.5 Topologia da Rede

A topologia da Rede Neural está relacionada com a quantidade de camadas intermediárias, entradas, saídas e a quantidades de neurônios em cada camada. Para cada situação pode ser atribuída uma determinada topologia, procura-se, então, determinar uma topologia ótima que possa obter resultados de melhor qualidade oriundas do seu treinamento.

O desempenho da rede neural está relacionado também, com a quantidade de camadas de neurônios, pois uma maior quantidade de camadas aumentaria a capacidade de aprendizado, com isso melhorando a precisão com que a rede delimita regiões de decisão (DA SILVA, 2003).

Na literatura, encontram-se diversos trabalhos de pesquisadores dando sugestões para a determinação de uma topologia ótima, onde pode ser citado o trabalho de HECTH e NIELSEN (1989) que sugerem para uma rede neural, apresentando uma camada intermediária, o número de neurônios nesta camada será conforme apresentado na Equação 2.

$$N = 2i + 1 \quad (2)$$

Onde: N é o número de neurônio na camada;

i é o número de variáveis de entrada.

LIPPMANN (1987) defende a ideia de que quando existir uma segunda camada intermediária na rede, esta deverá apresentar o dobro de neurônios da camada de saída. E no caso de haver somente uma camada intermediária, essa seguirá a Equação 3.

$$N = s(i + 1) \quad (3)$$

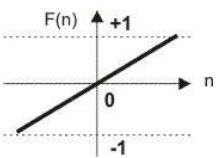
Onde: s é o número de neurônios da camada de saída

3.1.6 Funções de ativação

Em uma RNA, cada “neurônio”, ou unidade de processamento, está associado a um determinado estado de ativação que é imposto pela função de ativação, continua ou discreta, no qual se quer representar o modelo neural (RIBEIRO, 2013). A função de ativação é responsável pelo sinal a ser transmitido por um determinado neurônio e repassado aos demais que estão conectados a ele. As funções mais utilizadas geralmente são a linear, a rampa, a degrau, a tangente hiperbólica e a sigmoideal.

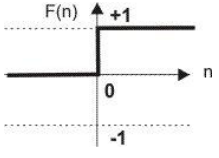
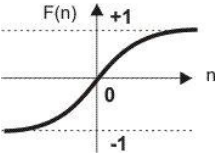
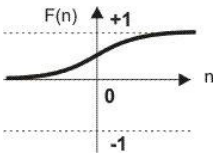
A saída de um neurônio é definida pela função de ativação que está relacionada com a atividade de entrada (HAYKIN, 2001). A Tabela 3 mostra as funções mais utilizadas.

Tabela 3- Tipos de Funções de ativação.

| NOME DA FUNÇÃO | REPRESENTAÇÃO GRÁFICA | CARACTERÍSTICA | FUNÇÃO |
|----------------|---|---|---------------------------|
| Linear |  | Função que não limita a saída da RNA e é usada para armazenar entrada e saída de dados. Os neurônios que possuem esta função atuam como aproximadores lineares. | $\varphi_i(t+1) = v_i(t)$ |

Continua

Conclusão

| NOME DA FUNÇÃO | REPRESENTAÇÃO GRÁFICA | CARACTERÍSTICA | FUNÇÃO |
|-----------------------|---|---|--|
| Degrau ou Limiar |  | Função utilizada para neurônios que tomam decisões binárias, limitando a saída do neurônio somente a dois valores, sendo assim, classificadora. | $\varphi_i(t+1) = \begin{cases} 1, & \text{se } v_i(t) \geq \theta \\ 0, & \text{se } v_i(t) < \theta \end{cases}$ |
| Tangente Hiperbólica |  | Função limitada, definida no intervalo de -1 e +1. | $\varphi_i(t+1) = (1 - e^{-v_i(t)}) / (1 + e^{-v_i(t)})$ |
| Sigmóide ou Logística |  | Função limitada que assume valores entre um limite superior e um inferior (0 e 1), sem jamais atingi-los. | $\varphi_i(t+1) = 1/(1 + e^{-v_i(t)})$ |

Fonte: Adaptado de RIBEIRO (2013)

3.1.7 Aprendizagem em RNAs

Para que uma rede possa ter condições de aprender sobre o problema a ser solucionado é necessário que ela retenha informações relevantes de padrões de informação apresentados, criando assim uma representação própria para o problema (BRAGA *et al.*, 2000). Apesar de existirem diversos métodos de treinamento de redes neurais, estes podem ser agrupados em dois paradigmas principais que são: aprendizado supervisionado e não supervisionado.

O aprendizado supervisionado é o método mais comum no treinamento de RNAs, tanto de neurônio com peso como de neurônios sem peso (BRAGA *et al.*, 2000). Também conhecida como aprendizagem com professor, consiste em que o professor tenha o

conhecimento do ambiente, e forneça um conjunto de exemplos de entrada e de saída desejada. O objetivo é ajustar os parâmetros da rede, de modo a encontrar uma ligação entre os pares de entrada e de saída fornecidos (RIBEIRO, 2013).

No caso de aprendizado não supervisionado não há a presença de um “professor” para supervisionar o processo de aprendizagem, ou seja, não há pares de entrada e de saída, havendo somente entradas. Nesse paradigma, são dadas condições para realizar uma medida da representação que a rede deve aprender, e os parâmetros livres da rede são otimizados em relação a essa medida. A RNA procura então, características estatísticas relevantes entre os dados apresentados, sendo que essa técnica só se torna aplicável quando há um grande número de dados (SILVA *et al.* 2004 *apud* RIBEIRO, 2013).

3.1.8 Perceptron Multi Camadas (MLP) e o Algoritmo Backpropagation

Segundo RIBEIRO (2013), as RNAs com *perceptrons* de múltiplas camadas, do inglês *Multilayer Perceptron* (MLP) constituem o tipo de arquitetura mais comum, consistindo em um conjunto de unidades de processamento com unidades de entrada, unidades intermediárias (ou camadas de escondidas) e camada de saída.

Será utilizado nessa monografia um algoritmo denominado como retropropagação de erro, mais conhecido como *backpropagation*. Segundo HAYKIN (1999), o *backpropagation* é o algoritmo de RNAs mais usado em aplicações práticas de previsão, classificação e de reconhecimento de padrões em geral.

O funcionamento desse algoritmo é caracterizado por duas fases, sendo elas:

- a) Fase *forward*: o sinal que parte da entrada se desloca em direção a saída, os pesos são mantidos e a saída fornecida é então comparada com a saída observada. Os dados observados são subtraídos dos dados fornecidos gerando assim um erro;
- b) Fase *backward*: o erro gerado na primeira fase retorna pelo caminho inverso, ou seja, da saída em direção à entrada, e esse retorno é denominado como “retropropagação do erro”. Dessa forma os pesos são ajustados e o sinal é propagado novamente pela rede, de modo a encontrar um sinal de erro menor a cada propagação (HAYKIN, 1999).

4 MATERIAIS E MÉTODO

4.1 Considerações iniciais

Para o desenvolvimento deste trabalho considerou-se, como área de estudo, o município de Fortaleza e dele, visando-se uma certa homogeneidade no conjunto de dados disponível, usou-se somente os acidentes com motocicletas ocorridos nas vias classificadas funcionalmente como arteriais. O município de Fortaleza foi escolhido como área de estudo por se tratar de uma das capitais do país que apresentam uma grande quantidade de circulações de veículos em transporte tanto de pessoas quanto de carga. Levando em consideração o percentual de acidentes envolvendo motocicletas pôde-se notar que boa parte, cerca de 30% das ocorrências, foi nas vias arteriais do município de Fortaleza. Encontra-se em Anexo um mapa da região de estudo.

Segundo o Código de Trânsito Brasileiro (CTB) – Lei de número 9503/97. Art. 60. No anexo I. A via é uma superfície por onde transitam veículos, pessoas e animais, compreendendo a pista, a calçada, o acostamento, ilha e canteiro central, sendo classificadas da seguinte maneira, de acordo com sua utilização:

- a) Vias urbanas: ruas, avenidas, vielas, ou caminhos e similares abertos à circulação pública, situados na área urbana, caracterizados principalmente por possuírem imóveis edificadas ao longo de sua extensão, subdividida em:
 - Via de trânsito rápido: aquela caracterizada por acessos especiais com trânsito livre, sem interseções em nível, sem acessibilidade direta aos lotes lindeiros e sem travessia de pedestres em nível;
 - Via arterial: aquela caracterizada por interseções em nível, geralmente controlada por semáforo, com acessibilidade aos lotes lindeiros e às vias secundárias e locais, possibilitando o trânsito entre as regiões da cidade;
 - Via coletora: aquela destinada a coletar e distribuir o trânsito que tenha necessidade de entrar ou sair das vias de trânsito rápido ou arteriais, possibilitando o trânsito dentro das regiões da cidade;
 - Via local: aquela caracterizada por interseções em nível não semaforizadas, destinadas apenas ao acesso local ou a áreas restritas;
- b) Vias rurais: estradas e rodovias:

- Estradas: via rural não pavimentada;
- Rodovias: via rural pavimentada.

4.2 Dados utilizados para elaboração do trabalho

Os dados brutos dos acidentes, envolvendo motocicletas, utilizados neste trabalho foram coletados pelo Sistema de Informação de Acidentes de Trânsito de Fortaleza (SIAT-FOR). O SIAT-FOR é um banco de dados informatizado e georeferenciado que reúne informações sobre os acidentes de trânsito registrado por 10 órgãos responsáveis pelo atendimento às ocorrências de trânsito do município de Fortaleza (CUNTO *et al.*, 2012). Dentre as informações coletadas e disponibilizadas pelo SIAT-FOR, destacam-se: o local da ocorrência (nome da via e número/referência), a data, hora, tipo de acidente, veículos envolvidos, severidade do acidente (sem vítimas, com vítimas feridas e vítimas fatais), além de possuir informações sobre as vítimas, bem como o gênero, idade, o tipo (condutor, passageiro, pedestre etc.).

Os dados utilizados nesta investigação referem-se ao ano de 2011 e tem como variável de interesse os diferentes níveis de severidades obtidas em cada informação, e como possíveis variáveis explicativas a jurisdição da via, o tipo de acidente, iluminação, condições climáticas, superfície da pista, tipo de cruzamento, controle de tráfego, manobra realizada, gênero do condutor, uso do capacete e idade.

4.3 Aplicativos Computacionais

Os aplicativos computacionais Excel e MATLAB foram os principais aplicativos utilizados no desenvolvimento desse trabalho. O Excel foi utilizado na manipulação dos dados oriundos do banco de dados do SIAT-FOR e no pré-processamento dos dados. O MATLAB foi utilizado para treinar, validar e teste a rede neural.

4.3.1 Microsoft Excel

O MS-Excel 2007® é um aplicativo desenvolvido pela Microsoft® para elaboração de planilhas eletrônicas possuindo diversos recursos para cálculo de funções matemáticas básicas e avançadas, bem como cálculos financeiros e estatísticos. Neste trabalho ele foi utilizado para fazer o pré-processamento dos dados para a modelagem, etapa esta que precede ao desenvolvimento do modelo neural.

4.3.2 MATLAB

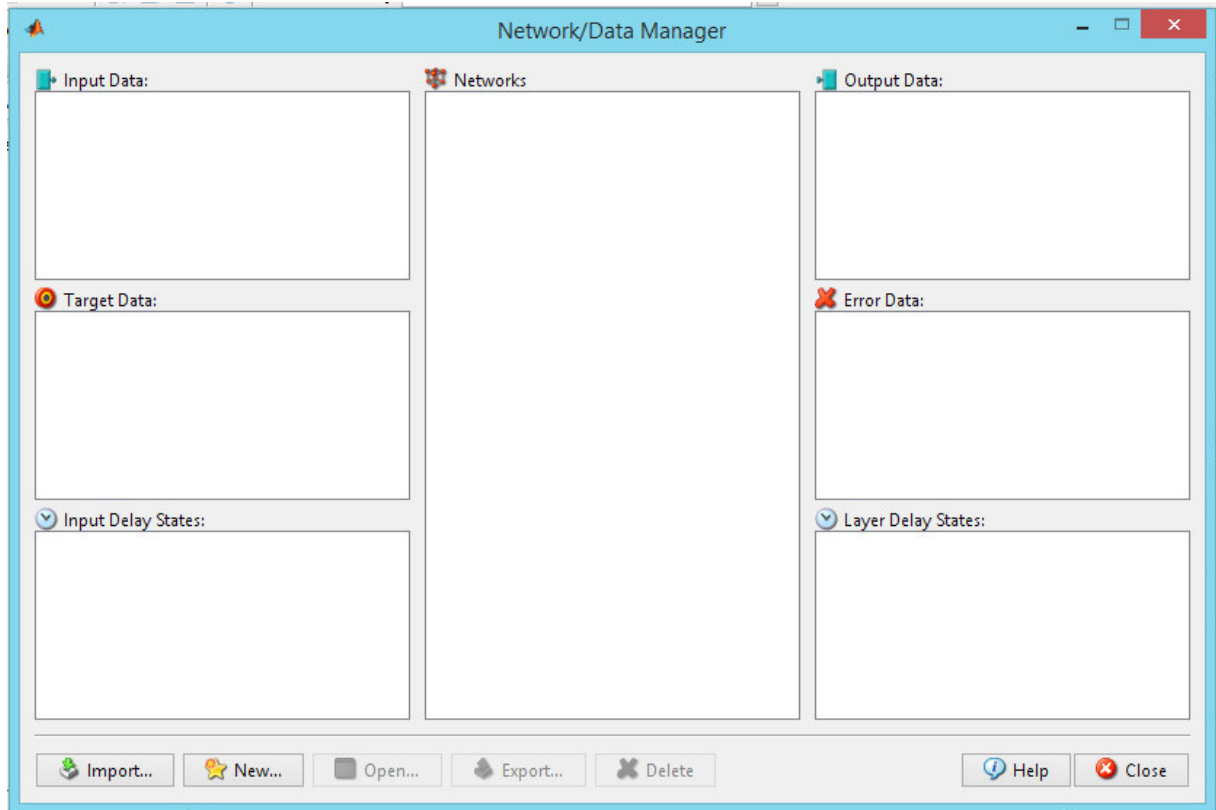
O MATLAB (*MATrixLABoratory*) 2009 desenvolvido pela *Mathworks, Inc*, é um sistema interativo e uma linguagem de programação baseados em matrizes, onde as soluções e os problemas são expressos em linguagem matemática. Com a utilização do MATLAB é possível realizar análise numérica, cálculo com matrizes, processamento de sinais e construção de gráficos em ambiente fácil de usar onde problemas e soluções são expressos somente como eles são escritos matematicamente, ao contrário da programação tradicional.

A principal vantagem desse software é a facilidade de escrever e depurar um programa, comparado a outras linguagens de programação, bem como o C, Basic, Pascal ou o Fortran. Além disso, possui diversas funções matemáticas, matriciais e gráficas que simplificam e minimizam a estrutura do programa.

O aplicativo computacional MATLAB foi utilizado para o treinamento da RNA proposta neste trabalho, sendo escolhido por possuir uma interface gráfica de treinamento de redes neurais. Esse aplicativo suporta os mais diversos tipos de algoritmos de treinamento de redes neurais, permitindo que se possa executar uma extensa árvore de testes na tentativa de definir o modelo neural mais adequado.

A interface gráfica utilizada para criar as redes neurais é denominada de *Neural Network Toolbox* (NNTOOL) podendo ser acionado apenas digitando seu nome na barra de comando do MATLAB, essa interface permite importar, criar, utilizar e exportar dados de redes neurais. A Figura 4 mostra como é a interface do NNTOOL.

Figura 4 - Interface gráfica NNTOOL

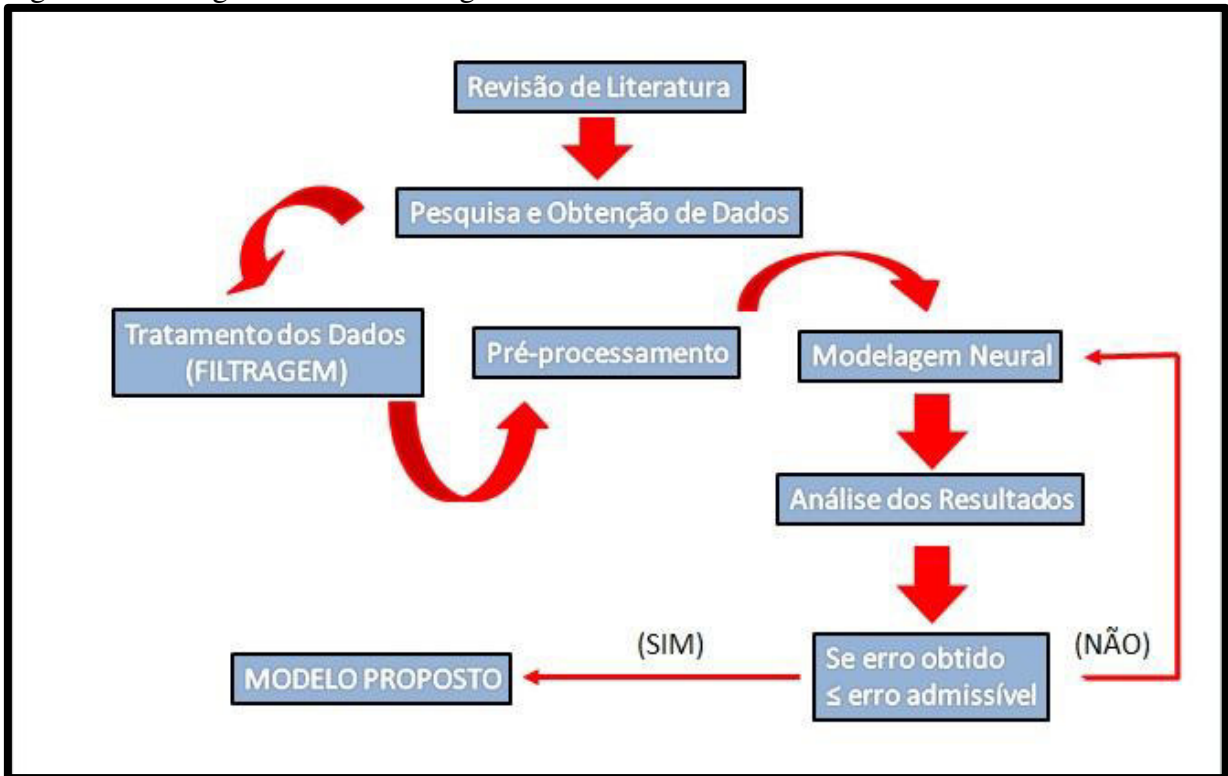


Fonte: Próprio autor.

4.4 Metodologia Adotada

A metodologia empregada para o desenvolvimento deste trabalho está descrita nos tópicos seguintes e suas etapas são: revisão de literatura, pesquisa e obtenção de dados, tratamentos dos dados, pré-processamento dos dados, modelagem neural, análise dos resultados e por último a obtenção do modelo calibrado e testado. A Figura 5 representa a metodologia adotada para o desenvolvimento desta pesquisa, seguida por uma descrição de cada etapa identificada no fluxograma.

Figura 5 - Fluxograma da metodologia adotada



Fonte: Próprio autor

4.4.1 Revisão de literatura

Nesta etapa foram realizadas pesquisas relacionadas à modelagem estatística, RNAs, segurança viária, sendo as principais fontes: artigos, trabalhos acadêmicos, manuais dentre outros meios. Essa revisão permitiu suprir as necessidades de aprofundamento teórico, bem como conhecer o estado da arte dos estudos sobre segurança viária e tentou responder algumas das questões propostas.

4.4.2 Pesquisa e obtenção dos dados

Com a finalidade de obter dados para o desenvolvimento do modelo neural, conforme descrito no item 4.2 deste capítulo, foram realizadas pesquisas junto aos órgãos detentores dessas informações. Esses dados vieram organizados em formato tabular, digital garantindo assim uma maior facilidade de manipulação dos mesmos.

4.4.3 Tratamento dos dados

Os dados obtidos através dos órgãos responsáveis possuem diferentes formatações por isso há uma necessidade de uma uniformização destes, de modo que todos possuam o mesmo formato. Nesta etapa fez-se necessário a realização de uma filtragem nos dados disponíveis, a fim de eliminar aqueles que são inconsistentes para a modelagem, ou seja, dados que forneceram informações incompletas.

Outra etapa que ocorreu durante esse processo foi a seleção dos dados que realmente interessavam à modelagem, dados esses relativos somente às vias arteriais e que apresentavam como vítimas dos acidentes apenas os motociclistas.

4.4.4 Pré-processamento

Essa etapa metodológica garantiu que os dados obtidos fossem utilizados na modelagem neural, para isso foram necessárias duas subetapas:

- a) Numerização dos dados: a numerização permitiu a atribuição de valores decimais aos dados que possuem valores alfanuméricos (variáveis qualitativas);
- b) Dummyzação: esse procedimento teve a finalidade de converter os dados de saída da rede neural para uma forma em que os dados numéricos possuam valores 0 e 1. Essa medida foi necessária devido ao fato de que a estrutura de modelagem neural escolhida exigiu esse padrão.

4.4.5 Modelagem neural (RNAs)

Para essa etapa da metodologia foi proposta e implementada uma árvore de testes visando encontrar uma topologia adequada, dentre as potencialmente possíveis. O Resultado dessa busca foi a topologia que melhor se adequou ao problema em questão, ou seja, aquela cujos erros das saídas estimadas, para o conjunto de testes, eram menores quando comparados aos valores reais, dentre todas aquelas que foram testadas.

A determinação de uma topologia ótima está relacionada a três parâmetros: (i) escolha do número de camadas intermediárias; (ii) escolha do número de neurônios nas camadas; (iii) parâmetros de treinamento (adaptações importantes que envolvem a escolha dos algoritmos de treinamento, funções de ativação, taxas de aprendizagem).

- a) Escolha do número de camadas: segundo CYBENKO (1989), começa utilizando uma camada e variando até duas camadas no máximo;
- b) Escolha do número de neurônios nas camadas: segundo HECH e NIELSEN (1989), para redes neurais com uma única camada intermediária utilizando-se a Equação 2, mencionada anteriormente.

LIPPMANN (1987) sugere que no caso da existência de uma segunda camada, ela deverá possuir o dobro de neurônios da camada de saída e se houver uma única camada intermediária o número de neurônios será como descrito na Equação 3.

Outras adaptações importantes envolvem a escolha: do algoritmo de treinamento, função de ativação e taxas de aprendizagem que possibilitem uma taxa de acerto maior.

4.4.6 Análise dos resultados

A análise dos resultados será realizada através de cálculos estatísticos que permitam verificar e comprovar qual o grau de aderência entre os valores estimados pela RNA e os reais valores dos dados observados em campo.

4.4.7 Modelo proposto

Essa última etapa caracterizou-se pela obtenção de um modelo calibrado, validado e testado que atenda aos requisitos exigidos no problema de pesquisa, onde o modelo demonstra a sua capacidade de identificar qual agrupamento de variáveis preditoras, melhor explica a ocorrência dos diferentes níveis de severidade com motocicletas em interseções urbanas.

5 RESULTADOS E DISCUSSÕES

5.1 Considerações iniciais

Objetivo principal desse trabalho foi desenvolver um método para estimação dos níveis de severidade dos acidentes envolvendo motocicletas, nas vias arteriais da cidade de Fortaleza de forma eficaz, prática e rápida. Espera-se que o método, proposto aqui, possa contribuir para que os pesquisadores e tomadores de decisão da área de segurança viária, a respeito de soluções ou até mesmo intervenções no sistema, visando à mitigação dos impactos causados pelos acidentes com motocicletas.

Este capítulo descreve a aplicação da metodologia proposta (ver capítulo 4) para estimar os diferentes níveis de severidade (ileso, leve, moderado, grave e fatal) dos acidentes de trânsito nas vias arteriais de Fortaleza. Então, são apresentados e discutidos os resultados obtidos através da modelagem neural das variáveis que foram utilizados neste trabalho, bem como, a determinação de quais delas melhor explicam esse fenômeno.

5.2 Análise exploratória dos dados

As informações referentes aos dados coletados dos acidentes de trânsito foram adquiridas por intermédio do SIAT-FOR e compilados em uma planilha referente a todas as vias de Fortaleza, discriminando os acidentes, os veículos e as pessoas envolvidas em cada acidente. O número total de observações disponíveis era de 7607 acidentes, sendo que desse total havia mais de 14320 veículos envolvidos e um total de 15687 pessoas que sofreram algum tipo de ferimento devido ao acidente.

Com o objetivo de sistematizá-los; foram então separados, dentre todos os dados, somente aqueles em que os veículos envolvidos eram motocicletas, ficando com um total de 7939 motocicletas. Através dessa análise pôde-se perceber que há um maior número de motocicletas do que o número total de acidentes observados, levando a entender que havia mais de uma motocicleta envolvida num mesmo acidente.

Então houve a necessidade de filtrar os dados para remover aqueles que eram duplicados e que estavam relacionados ao mesmo acidente, em seguida foram retiradas do

banco de dados todas as observações nas quais não havia informação em pelo menos uma das variáveis (incompletas), reduzindo ainda mais o número de observações. Após esse processo foram então selecionadas no banco de dados somente às classificadas funcionalmente como vias arteriais, e compiladas essas observações em uma única planilha que contém todas as informações sobre os acidentes, totalizando 1137 observações.

Na literatura consultada identificou-se as principais variáveis utilizadas por outros pesquisadores para servir de base aos treinamentos da rede. Segundo ALI (2004), as variáveis empregadas em sua pesquisa foram divididas nas categorias: condutor, veículo, estrada e meio ambiente, como mostra a Tabela 4.

Tabela 4 - Codificação de dados das variáveis

| Fator | Binário/Código numérico |
|---------------------------------------|---|
| <u>Variáveis de Entrada</u> | |
| Condutor | |
| Idade do Condutor | Valor Numérico |
| Gênero do Condutor | 1 Feminino; 0 Masculino |
| Envolvimento com Álcool | 1 Sim, 0 Não |
| Uso de Cinto de Segurança | 1 Não usou; 0 Usou |
| Veículo | |
| Tipo de Veículo | (1,0) Carro; (0,1) Van; (0,0) Pick-up. |
| Ponto de Impacto | 1 no lado do motorista, 0 no lado do passageiro |
| Relação de Velocidade | Valor Numérico |
| Estrada/ Meio Ambiente | |
| Tipo de Área | 1 Rural; 0 Urbano |
| Iluminação | 1 Luz do dia; 0 Escuro |
| Período de Pico | 1 Período de Pico, 0 Fora do Período de Pico |
| Tempo | 1 Limpo, 0 Outro |
| Característica da pista | 1 Curva, 0 reta |
| <u>Variáveis de Saída</u> | |
| Gravidade da Lesão do Condutor | (1,0,0,0) ileso, (0,1,0,0) leve, (0,0,1,0) grave, (0,0,0,1) fatal |

Fonte: Adaptado de ALI, 2004.

Através da análise das variáveis foi então possível extrair dentre as disponíveis no banco de dados, algumas variáveis preditoras que serviram de base para a modelagem neural proposta aqui, essas variáveis foram também divididas em condutor e informações sobre o acidente, retirando o item veículo do banco de dados, pois o veículo envolvido nesta modelagem referem-se essencialmente as motocicletas, bem como mostra a Tabela 5.

Tabela 5 - Variáveis envolvidas na modelagem

| | Fator | Binário/ Código Numérico |
|----------------------|-------------------------------------|--|
| Variáveis de Entrada | Condutor | |
| | Idade | Valor Numérico |
| | Gênero | 1 Feminino, 2 Masculino |
| | Informações sobre o Acidente | |
| | Jurisdição | 1 Municipal, 2 Estadual, 3 Federal |
| | Iluminação | 1 Amanhecer, 2 Anoitecer, 3 Luz do Dia, 4 Mal Iluminada, 5 Via Iluminada |
| | Tempo | 1 Bom, 2 Chuva, 3 Nublado |
| Variáveis de Saída | Nível de Severidade | (1,0,0,0,0) ileso, (0,1,0,0,0) leve, (0,0,1,0,0) moderado, (0,0,0,1,0) grave, (0,0,0,0,1) fatal. |

Fonte: Próprioautor

5.3 Árvore de teste e desempenho da modelagem

Para realizar a modelagem neural proposta, os dados foram então organizados de maneira que as variáveis operacionais, funcionais e pessoais de cada condutor fossem usadas como dados de entrada das RNAs (Jurisdição, Iluminação, Tempo, Idade e Gênero) sendo estas as variáveis consideradas básicas para compor o modelo, além destas, também foram adicionadas mais cinco variáveis (Natureza do acidente, Superfície da pista, Tipo de Cruzamento, Controle de Tráfego e manobra) para que possa ser estimado, dentre estas

variáveis, as que melhor explicam as saídas. A Tabela 6 descreve como os dados relativos às cinco variáveis base foram organizados.

Tabela 6 - Organização das 5 variáveis base

| Dados de Entrada | | | | | Dados de Saída |
|--------------------|--------------------|--------------------|---------------------|--------------------|---------------------|
| Jurisdição | Iluminação | Tempo | Idade | Gênero | Nível de Severidade |
| Informação Nominal | Informação Nominal | Informação Nominal | Informação Numérica | Informação Nominal | Informação Nominal |

Fonte: Próprio autor

Para possibilitar a estimação de maneira mais rápida foram então criados mais dois grupos de treinamento, adicionando-se a este, as demais variáveis disponíveis no banco de dados do SIAT-FOR. Para o primeiro grupo foram adicionadas como possíveis variáveis explicativas: a natureza do acidente, superfície da pista e tipo de cruzamento totalizando 8 variáveis explicativas, bem como mostra a Tabela 7.

Tabela 7 - Organização das 8 variáveis explicativas

| Dados de Entrada | | | | | | | | Dados de Saída |
|--------------------|--------------------|--------------------|----------------------|---------------------|--------------------|---------------------|--------------------|---------------------|
| Jurisdição | Iluminação | Tempo | Natureza do Acidente | Superfície da Pista | Tipo de Cruzamento | Idade | Gênero | Nível de Severidade |
| Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Numérica | Informação Nominal | Informação Nominal |

Fonte: Próprio autor

Para o segundo grupo foi adicionado mais duas variáveis, sendo elas: controle de Tráfego e manobra, totalizando assim 10 variáveis explicativas, que possivelmente apresentam alguma contribuição para a ocorrência de diferentes níveis de severidade, como mostra a Tabela 8.

Tabela 8 - Organização das 10 variáveis explicativas

| Dados de Entrada | | | | | | | | | | Dados de Saída |
|--------------------|--------------------|--------------------|--------------------|---------------------|--------------------|---------------------|--------------------|---------------------|--------------------|---------------------|
| Jurisdição | Iluminação | Tem | Natureza Acidente | Superfície da Pista | Tipo de Cruzamento | Controle de Tráfego | Manobra | Idade | Gênero | Nível de Severidade |
| Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Nominal | Informação Numérica | Informação Nominal | Informação Nominal |

Fonte: Próprio autor

Para estimar as variáveis explicativas que mais contribuem para a ocorrência de diferentes níveis de severidade, foram realizados diversos testes utilizando vários algoritmos do tipo *backpropagationfeed-forward*. Pôde-se verificar que duas variações obtiveram melhores resultados dentre os modelos construídos: TRAINGDX (Gradiente descendente *backpropagation* com *momentum* e taxa de aprendizagem adaptativa) e TRAINLM (Levenberg-Marquardt). Sendo que dentre os dois o primeiro obteve melhores resultados na predição dos níveis de severidade.

Apesar de o TRAINLM ser considerado um algoritmo de treinamento rápido que utiliza método de Newton (BEALE *et al.*, 2010) o algoritmo TRAINGDX identificou o padrão das informações e à medida que passava o processamento, mais aumentava o desempenho da modelagem neural. Com a necessidade de se encontrar qual era o grupo de variáveis explicativas que melhor explicam a ocorrência de diferentes níveis de severidade nesses acidentes, foi montado uma árvore de testes para identificar à(s) melhor (es) topologia(s):

- a) redes com diferentes algoritmos de treinamento, sendo principalmente utilizados o TRAINGDX e TRAINLM;
- b) topologias com uma e com duas camadas escondidas no máximo;
- c) topologias com quantidades de neurônios diferentes nas camadas intermediárias seguindo a indicação de NIELSEN (1989) para a determinação do número de neurônios na camada intermediária. Também foi seguida a indicação sugerida por LIPPMAN (1987) (ver no capítulo 4);

- d) topologias com função de ativação Logística Sigmoidal (LOGSIG) e Tangente Hiperbólica (TANSIG) nos neurônio(s) das camada(s) intermediária(s) e função identidade (PURELINE) nos neurônios de saída;
- e) topologias que apresentaram diferentes parâmetros de treinamento: taxa de aprendizagem variando de 0,01 e 1,05; e número de épocas variando de 10, 100, 1000 ciclos.

A estratégia utilizada afim de se alcançar a topologia ótima foi obtida seguindo-se as indicações fornecidas por RIBEIRO (2013), que buscou definir uma topologia mais apropriada que convergisse para um erro mínimo em um menor tempo de processamento. A melhor topologia observada e identificada como a mais adequada nestes resultados, foi aquela apresentou melhor desempenho dentro das limitações de tempo disponíveis para a execução deste estudo.

O período de testes demorou por volta de um mês. Este tempo foi considerado suficiente para verificar todas as demais possibilidades, mas permitiu que fossem testadas cerca de 100 topologias diferentes para cada um dos três grupos de dados utilizadas para estimar as variáveis que mais contribuem para a ocorrência de diferentes níveis de severidade. DA SILVA (2003), afirma que a determinação de uma topologia ótima requer um tempo considerável e mesmo assim, é possível que haja uma combinação de parâmetros não testada que apresente melhor resultados em generalizar o fenômeno e com menor tempo de convergência.

Para demonstrar o desempenho gerado dentre as topologias verificadas, foi então criada a Tabela 9, que apresenta as 10 melhores topologias testadas para cada grupo de variáveis explicativas, sendo ressaltada aquela que obteve o melhor desempenho, ou seja, apresentaram erros bem menores para o conjunto de teste, na estimação das variáveis em questão. A variável “tempo de processamento” não foi considerada para comparar as topologias, devido ao fato de que todos os processos obtiveram tempos aproximados não havendo então diferença nos resultados, apesar da utilização do algoritmo de treinamento (Levenberg – Marquardt) que mostrou uma maior rapidez na identificação do menor MSE (*Mean of Squared Error*) ou (Quadrado Médio os Erros), parâmetro esse utilizado para medir o desempenho das topologias testadas. Esta medida de desempenho foi encontrada em DA SILVA (2003), sendo a mesma calculada através da Equação 4.

$$MSE = \frac{SSE}{n - p} \quad (4)$$

Onde:

SSE = Diferenças quadráticas entre os valores de saída da RNA e os valores esperados como ideais (BOCANEGRA, 2002);

n = Número de exemplos do conjunto de treinamento (quantidade de padrões);

p = Número de parâmetros a serem estimados (quantidade de pesos da rede).

Foi então confirmada durante essa etapa de determinação da topologia ótima a afirmação de CYBENKO (1988), onde diz que “uma camada intermediária é suficiente para aproximar qualquer função contínua e duas camadas intermediárias aproximam qualquer função matemática”. Devido a isso, não foram consideradas topologias com mais de duas camadas ocultas.

Tabela 9 - Desempenho das principais Topologias com 5, 8 e 10 variáveis

| 5 Variáveis Base | | | | | | | | | | |
|--------------------------|--------|---------------|--------|--------|--------|---------|---------------|---------|-------|-------|
| Entrada e Saída | (5/5) | (5/5) | (5/5) | (5/5) | (5/5) | (5/5) | (5/5) | (5/5) | (5/5) | (5/5) |
| Camadas ocultas | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Nº Neurônios | 6 | 11 | 20 | 50 | 100 | 6 | 11 | 20 | 50 | 100 |
| Algoritmo | Trainl | Trainl | Trainl | Trainl | Trainl | Traingd | Traing | Traingd | Train | Train |
| | m | m | m | m | m | x | dx | x | gdx | gdx |
| MSE | 0,109 | 0,110 | 0,110 | 0,107 | 0,109 | 0,109 | 0,104 | 0,108 | 0,110 | 0,110 |
| 8 Variáveis Explicativas | | | | | | | | | | |
| Entrada e Saída | (8/5) | (8/5) | (8/5) | (8/5) | (8/5) | (8/5) | (8/5) | (8/5) | (8/5) | (8/5) |
| Camadas ocultas | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Nº Neurônios | 9 | 17 | 20 | 50 | 100 | 9 | 19 | 20 | 50 | 100 |
| Algoritmo | Trainl | Trainl | Trainl | Trainl | Trainl | Traingd | Traingd | Traingd | Train | Train |
| | m | m | m | m | m | x | x | x | gdx | gdx |
| MSE | 0,112 | 0,109 | 0,110 | 0,115 | 0,110 | 0,110 | 0,111 | 0,114 | 0,111 | 0,110 |

(Continua)

(Conclusão)

| 10 Variáveis Explicativas | | | | | | | | | | |
|---------------------------|-------------|-------------|-------------|-------------|-------------|----------------------|--------------|--------------|--------------|--------------|
| Entrada e Saída | (10/5) | (10/5) | (10/5) | (10/5) | (10/5) | (10/5) | (10/5) | (10/5) | (10/5) | (10/5) |
| Camadas ocultas | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Nº Neurônios | 11 | 15 | 23 | 50 | 100 | 11 | 15 | 23 | 50 | 100 |
| Algoritmo | Trainl m | Trainl m | Trainl m | Trainl m | Trainl m | Traing dx | Traingd x | Traingd x | Train gdx | Train gdx |
| MSE | 0,115 | 0,110 | 0,110 | 0,111 | 0,110 | 0,107 | 0,112 | 0,110 | 0,115 | 0,113 |

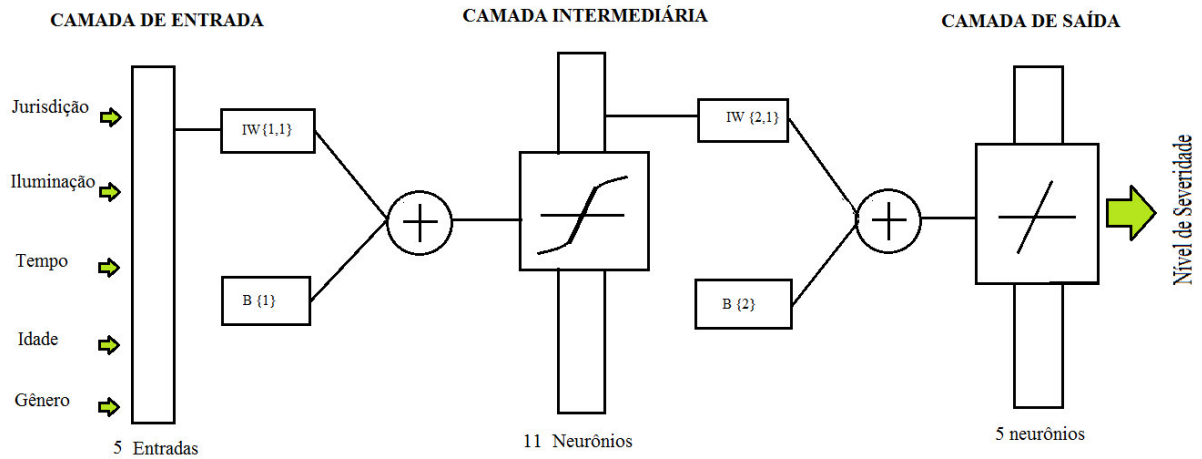
Fonte: Próprio autor

Após a implantação da árvore de testes relacionada aos resultados descrita acima, verificou-se que em todos os grupos a quantidade de neurônio nas camadas intermediárias não interferia muito no desempenho das RNAs, mas tornava evidente que quanto maior o número de neurônios o tempo de processamento também aumentava, mas esse aumento não foi significativo. Determinou-se assim, três topologias ótimas, uma para cada grupo de treinamento. As três topologias tinham um fator interessante, o melhor desempenho foi obtido quando o número de neurônios era igual ao que indicava HECTH e NIELSEN (1989), (ver capítulo 4) em exceção ao grupo de treinamento com 11 variáveis que com apenas 11 neurônios apresentou o melhor desempenho. As topologias ótimas encontradas representadas pelas Figuras 6, 7 e 8, obtidas individualmente foram as seguintes:

- Para os grupos com 5 e 10 variáveis, algoritmo TRAINGDX (*Gradiente descendente com momentum e taxa de aprendizagem adaptativa*);
- Para o grupo com 8 variáveis, algoritmo TRAINLM (*Levenberg-Marquardt*);
- Uma camada de entrada com 5, 8 e 10 neurônios representados pelas variáveis:
 - Jurisdição, Iluminação, Tempo, Idade, Gênero, Uso de Capacete, para o grupo com 6 variáveis;
 - Jurisdição, Iluminação, Tempo, Natureza do acidente, Superfície da pista, Tipo de cruzamento, Idade, Gênero, Uso de Capacete, para o grupo de 9 variáveis;
 - Jurisdição, Iluminação, Tempo, Natureza do acidente, Superfície da pista, Tipo de cruzamento, Controle de tráfego, Manobra, Idade, Gênero, Uso de Capacete, para o grupo com 11 variáveis;
- Uma camada intermediária (escondida) com 13, 19 e 11 neurônios, para os grupos 5, 8 e 10 respectivamente;
- Uma camada de saída com 5 neurônios para classificação dos níveis de severidade;

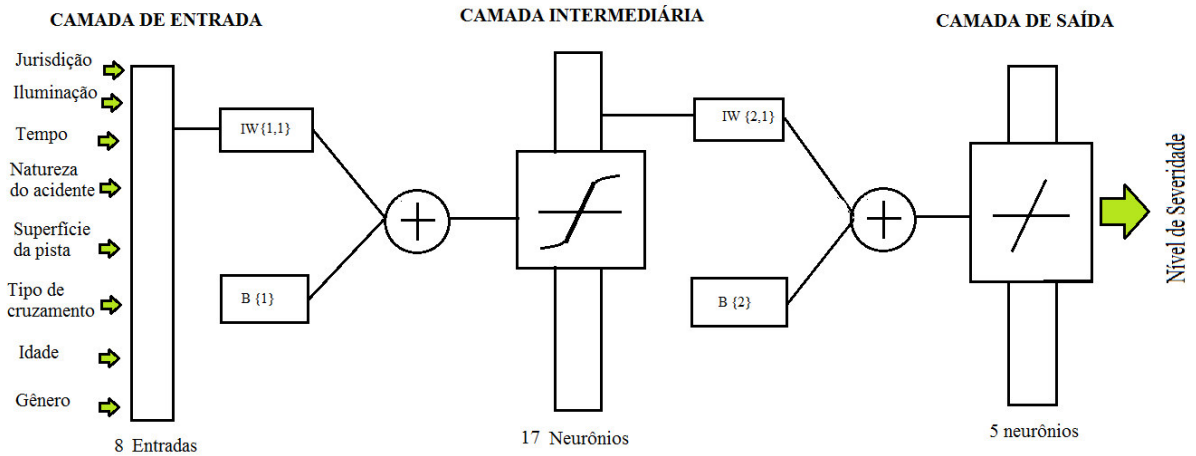
- f) Uso da função de ativação identidade na camada de entrada, função de ativação linear (*TANSIG*) com intervalo de ativação $[-1; 1]$ para as camadas intermediárias e função identidade na camada de saída.

Figura 6 - Topologia adotada para as 5 variáveis base (5:11:5)



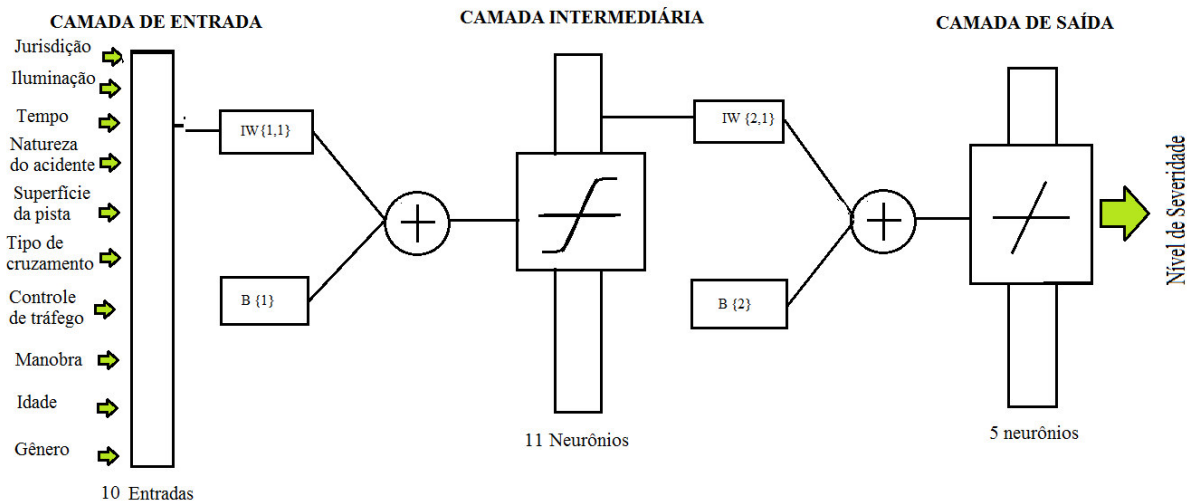
Fonte: Adaptado de RIBEIRO (2013).

Figura 7 - Topologia adotada para 8 variáveis (8:17:5)



Fonte: Adaptado de RIBEIRO (2013).

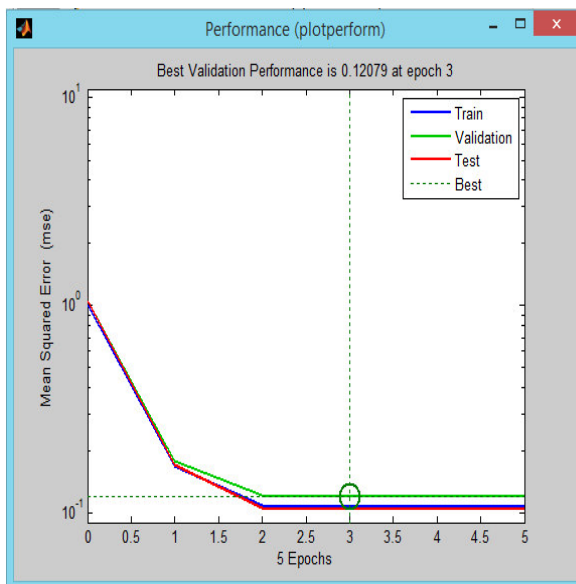
Figura 8 - Topologia adotada para 10 variáveis (10:11:5)



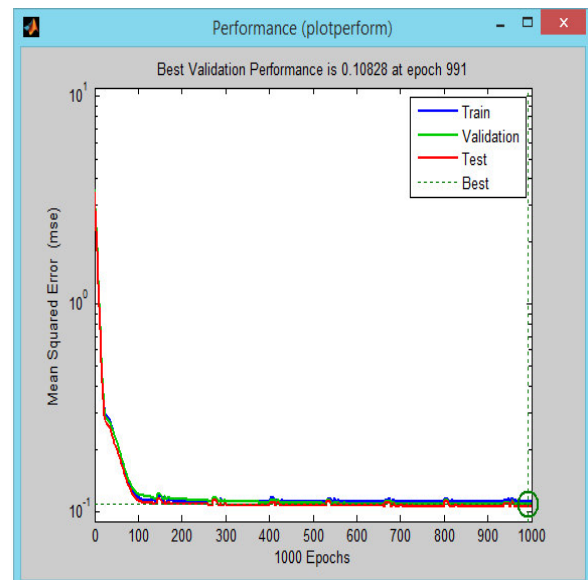
Fonte: Adaptado de RIBEIRO (2013).

As Figuras 9, 10 e 11 retiradas do NNTOOL, mostram uma comparação dos desempenhos obtidos utilizando os principais algoritmos utilizados neste trabalho, sendo eles: TRAINLM e TRAINGDX. É possível perceber que dentre os três grupos de variáveis, a rede que utilizou apenas 5 variáveis demonstrou um melhor desempenho em comparação com os outros dois, mesmo possuindo uma quantidade menor de variáveis.

Figura 9 - Comparação entre os algoritmos TRAINLM x TRAINGDX (5:11:5)



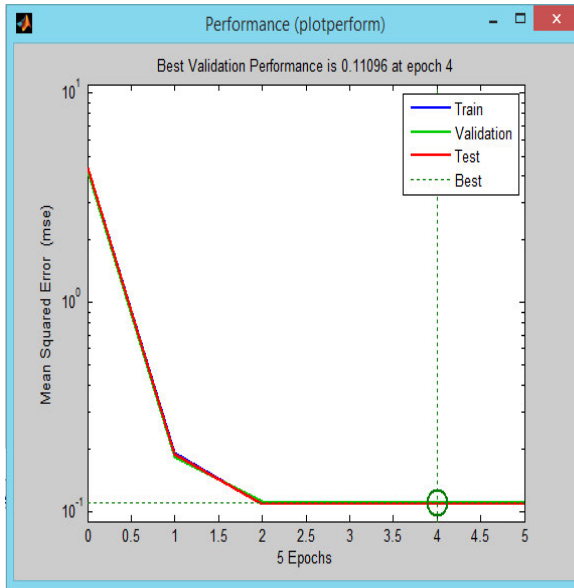
Desempenho com o algoritmo TRAINLM



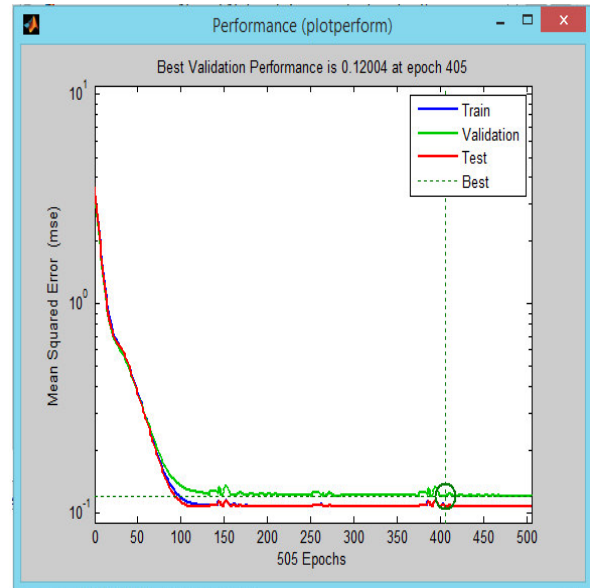
Desempenho com o algoritmo TRAINGDX

Fonte: Próprio autor.

Figura 10 - Comparação entre os Algoritmos TRAINLM x TRAINGDx (8:17:5)



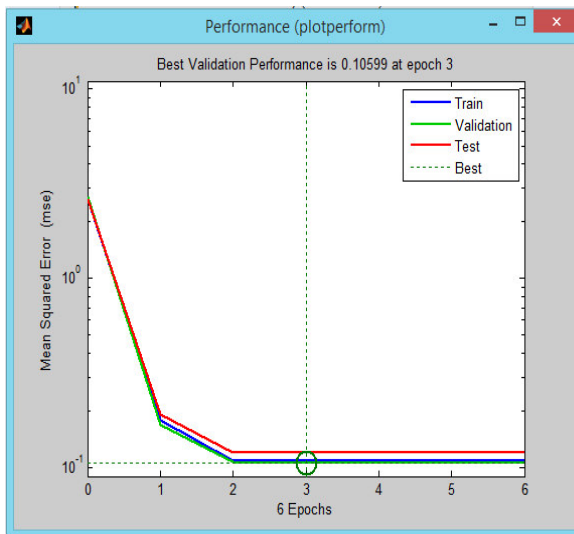
Desempenho com o algoritmo TRAINLM



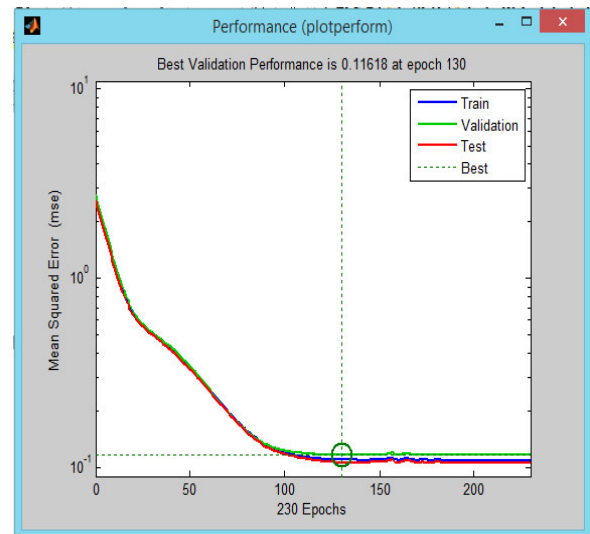
Desempenho com o algoritmo TRAINGDx

Fonte: Próprio autor.

Figura 11 - Comparação entre os Algoritmos TRAINLM x TRAINGDx (10:11:5)



Desempenho com o algoritmo TRAINLM



Desempenho com o algoritmo TRAINGDx

Fonte: Próprio autor

Observando-se os gráficos das Figuras 8, 9 e 10, percebe-se que os MSE do conjunto de treinamento, de teste e validação de todos os modelos eram bem aproximados. Esse comportamento denota que os modelos possuem uma relação bem próxima não havendo divergência entre os resultados. A Tabela 10 apresenta os valores do MSE para todos os conjuntos e para cada modelo neural escolhido, bem como os intervalos de tempo de processamento gasto em cada modelo.

Tabela 10 - Desempenho das RNAs e tempo de processamento

| GRUPOS | 5 variáveis | | | 8 variáveis | | | 10 variáveis | | |
|-----------------|--------------------|-----------|-------|--------------------|-----------|-------|---------------------|-----------|-------|
| RNA | (5:11:5) | | | (8:17:5) | | | (10:11:5) | | |
| Conjunto | Treinamento | Validação | Teste | Treinamento | Validação | Teste | Treinamento | Validação | Teste |
| MSE | 0,106 | 0,108 | 0,104 | 0,110 | 0,112 | 0,109 | 0,107 | 0,110 | 0,107 |
| Tempo | 20 segundos | | | 25 segundo | | | 23 segundos | | |

Fonte: Adaptado de RIBEIRO (2013).

Analisando-se a Tabela 10, nota-se que os tempos de convergência dos modelos deram muito baixo, menores até que meio minuto, um ciclo muito bom devido à quantidade de dados fornecidos para a modelagem neural. Em relação aos MSE dos conjuntos de treinamento para todos os modelos, é nítido que eles estão bem próximos aos do conjunto de validação e teste.

O percentual de acerto dos modelos foi medido levando em consideração apenas o conjunto de teste, pelo fato de que a rede neural ainda não conhecia as saídas, sendo então as saídas da RNA comparadas com os dados reais apresentados à rede após os testes. Com isso determinou-se o desempenho percentual de acerto dos modelos calibrados para a generalização dos níveis de severidades dos acidentes de trânsito a fim de estimar quais as variáveis que mais contribuem para a sua ocorrência. A Tabela 11 apresenta os valores percentuais de acertos para os três grupos de variáveis em relação aos conjuntos de testes.

Tabela 11 - Percentuais de acertos da RNA para os conjuntos de teste de cada grupo de variáveis

| GRUPOS | 5 variáveis | 8 variáveis | 10 variáveis |
|----------------|--------------------|--------------------|---------------------|
| RNA | (5:11:5) | (8:17:5) | (10:11:5) |
| Acertos | 89,6% | 89,1% | 89,3% |

Fonte: Próprio Autor

A partir dos resultados obtidos neste trabalho, podem-se levantar comparações com outros modelos de predição de acidentes que envolvem a determinação das variáveis que contribuem para os níveis de severidade. ABDEL-ATY (2004), conseguiu alcançar, utilizando RNAs, um resultado de 73,5% de acerto para seu conjunto de teste na predição do nível de severidade das lesões nos acidentes de tráfegos na região do Sudão. BOCANEGRA (2006), utilizou as técnicas de redes neurais com a finalidade de observar a variabilidade da velocidade dos veículos antes e depois da implantação de lombadas eletrônicas na cidade de Brasília, com uma taxa de acerto de 95,2%.

Através dessas informações, percebe-se que o modelo aqui desenvolvido atendeu basicamente as expectativas no que se refere à predição dos níveis de severidade dos acidentes de trânsito ocorridos com motocicletas nas vias arteriais da cidade de Fortaleza com um percentual de acerto de 89,4%. Pode-se verificar também que no grupo de 8 variáveis houve uma pequena redução, mas praticamente insignificante, e isso se repete com o grupo de 10 variáveis, mostrando que com apenas 5 variáveis explicativas é possível generalizar de forma satisfatória os diferentes níveis de severidade.

5.4 Considerações finais

Com relação à análise dos resultados obtidos no processo de modelagem neural das variáveis explicativas, constatou-se que muito embora a quantidade de variáveis de entrada da rede fossem mínimas, estas conseguiram obter um desempenho melhor na predição dos níveis de severidade dos acidentes de trânsito em vias arteriais. Não havendo assim necessidade de acrescentar outras variáveis para explicar o fenômeno e talvez nem mesmo de coletar essas variáveis.

MONMAR (2005), afirma que os acidentes em geral são mais frequentes em dias com boas condições meteorológicas, em interseções firmadas com pelo menos uma via arterial situadas em áreas com alto adensamento do solo, ou seja, em locais onde a uma maior concentração de atividades comerciais e conseqüentemente mais propício a haver conflitos entre veículos devido ao grande número de tráfego.

6 CONCLUSÕES E RECOMENDAÇÕES

6.1 Considerações iniciais

Este Projeto de Graduação teve como principal objetivo a proposição de um modelo dedicado à estimação dos diferentes níveis de severidade em acidentes envolvendo motocicletas em vias arteriais na cidade de Fortaleza e de identificar quais variáveis preditoras melhor explicam essa severidade, com o intuito de proporcionar, aos pesquisadores de segurança viária, meios de prever as severidades dos acidentes e tentar mitigá-las. A implantação dos modelos neurais usou apenas dados secundários, não havendo a necessidade de pesquisas em campo adicionais, limitando-se o estudo às vias arteriais da cidade de Fortaleza, onde ocorre uma maior concentração de veículos circulando e, segundo a literatura, uma maior probabilidade da ocorrência de acidentes.

Procurou-se então, com este trabalho, deixar como contribuição principal a determinação da importância de diferentes grupamentos de variáveis explicativas para a ocorrência de diferentes níveis de severidade. Através da utilização de técnicas de inteligência artificial, de forma rápida e de fácil compreensão, pode-se perceber a grande vantagem proporcionada pela RNA comparada ao que é empregado nos modelos estocásticos, bem como os MPA.

Espera-se que a monografia aqui desenvolvida venha a auxiliar na facilitação do reconhecimento prévio das variáveis explicativas dos acidentes de trânsito envolvendo motocicletas, de maneira a proporcionar a construção de rodovias que garantam, através de seus componentes operacionais, funcionais e até mesmo por fatores humanos, a segurança dos que a utilizam no seu dia a dia.

6.2 Principais Conclusões

A seguir são apresentadas as principais conclusões obtidas a partir dos três modelos desenvolvidos através dos grupos de variáveis explicativas que foram calibrados, validados e testados nesta pesquisa.

6.2.1 *Quanto ao método e a modelagem neural*

É fato que houve um grande empenho do modelador na tentativa da obtenção de uma topologia ótima que oferecesse um menor erro possível, em pequeno período de processamento, muito embora não fosse certamente possível esgotar todas as possibilidades de combinações de parâmetros para o desenvolvimento de novas topologias. Dessa maneira, pôde-se usar, para a modelagem das variáveis explicativas, as topologias obtidas no Capítulo 5, entretanto, deve-se notar que pode existir outra topologia ainda não testada que demonstre melhores resultados em seu processo de modelagem, apresentando erros ainda menores para o conjunto de teste.

A modelagem neural demonstrou um índice de acerto bastante promissor, apresentando desempenhos por volta de 90% de acerto para todos os três grupos de variáveis explicativas (5 variáveis base, 8 variáveis e com 10 variáveis). Com isso, mostrou-se possível a estimação das variáveis explicativas a partir da previsão dos níveis de severidade dos acidentes de trânsito de uma determinada via, neste caso as vias arteriais de Fortaleza, de forma rápida e precisa.

Através desse estudo, pôde-se perceber que é possível obter as informações sobre os diferentes níveis de severidade, como dados de saída da rede, com modelagem neural a partir de dados secundários operacionais, funcionais, climáticos e pessoais, como dados de entrada. Esses dados foram disponibilizados pelo SIAT-FOR, que compilou todas as informações fornecidas por diversos órgãos que fiscalizam os acidentes nas rodovias.

6.2.2 *Quanto às incertezas dos modelos gerados*

Os três grupos previsão de diferentes níveis de severidades dos acidentes de trânsito envolvendo motocicletas (5 variáveis base, 8 variáveis e 10 variáveis) modelados neste trabalho tiveram os níveis de acertos considerados satisfatórios e relevantes, sendo: 89,4% para o grupo com 5 variáveis tidas como básicas, 89,3% para o grupo com 8 variáveis, sendo composta pelas seis variáveis base adicionada de outras três variáveis explicativas e 89,4% para o grupo com 10 variáveis. Isto se deu em função da quantidade de dados de acidentes de trânsito disponíveis no banco de dados e a quantidade de variáveis utilizadas.

A medida de desempenho utilizada foi o MSE do conjunto de teste que tem 20% do conjunto de dados disponível para cada grupo estudado. Quando comparado estes resultados com os de outros modelos neurais de igual proposição, observou-se que os três modelos gerados têm desempenho igual e em certos casos até mesmo superiores.

6.2.3 Quanto ao modelo proposto para o grupo com 5 variáveis

Pôde-se perceber que de acordo com o modelo proposto para o grupo constituído de seis variáveis, mesmo com uma quantidade reduzida de variáveis o seu desempenho continua sendo maior comparado com os demais grupos aqui também estudados. Ou seja, essas variáveis garantem de fato corresponder bastante na ocorrência de diferentes níveis de severidade em acidentes de trânsito, não havendo a necessidade de acrescentar outras variáveis para compreender o fenômeno.

A respeito das variáveis utilizadas neste grupo, pode-se destacar uma variável bastante importante que é a idade do condutor, que na maioria das vezes, nos acidentes envolvendo motociclistas, tinham idade por volta de 20 a 40 anos, período entre a adolescência e a maturidade, onde os jovens tendem a ter atitudes inconsequentes, inclusive no trânsito. Outra variável que descreve basicamente esses níveis de severidade em acidentes envolvendo motocicletas é o tempo/clima que muito embora os devidos acidentes ocorressem em boas condições climáticas gerando até mesmo uma contradição entre o que realmente ocorre e o que era esperado.

6.2.4 Quanto aos modelos de 8 variáveis e 10 variáveis explicativas

O modelo de previsão do nível de severidade dos grupos com 8 e 10 variáveis por apresentarem em seus componentes as seis variáveis tidas como básicas pela literatura consultada, demonstraram em seus desempenhos para o grupo de teste, índices de acerto muito aproximados ao grupo com 6 variáveis, cerca de 89% de acerto em ambos os modelos. Esse fato leva a crer que a partir da verificação das 6 primeiras variáveis, foi possível então manter a generalização das observações.

6.3 Recomendações e sugestões para trabalhos futuros

A partir dos resultados obtidos e com base no conhecimento adquirido ao longo do desenvolvimento deste trabalho, percebeu-se que existe a necessidade de trabalhos que venham a complementar este ou ainda melhorar e expandir a área de estudo da malha viária. Assim, recomenda-se realização das seguintes pesquisas:

- a) Sugere-se que haja uma mesclagem maior e com mais combinações de entre as variáveis estudadas para que possa ser possível observar se há alguma diferença entre os resultados obtidos;
- b) Recomenda-se para trabalhos futuros, a implantação de mais variáveis de entrada da rede, bem como a hierarquização das vias , para contemplar também todas as vias da cidade de Fortaleza não somente as arteriais;
- c) Recomenda-se também um estudo de correlação entre as variáveis de entrada e delas com os respectivos níveis de severidades para obtenção de um modelo ainda mais preciso com um índice de acerto mais elevado;
- d) Recomenda-se finalmente com a utilização desses modelos de previsão dos níveis de severidade, o desenvolvimento de estudo onde possa ser verificado quais locais estão ocorrendo um maior número de vítimas graves ou fatais, a fim de promover soluções para a diminuição desses acidentes, baseado na análise dessas 5 variáveis preditoras.

REFERÊNCIAS

- ALI, G. A. **Traffic Accidents and Road Safety Management: A Comparative Analysis and Evaluation in Industrial, Developing and Rich-Developing Countries.** the 29th Southern African Transport Conference (SATC 2010). Pretoria. 2010. p. 16-19.
- AWAD, W. H.; JANSON, B. N. Prediction Models for Truck Accidents at Freeway Ramps in Washington State Using Regression and Artificial Intelligence Techniques. **Transportation Research Record**, n. 1635, p. 30-36, 1998.
- BÉLANGER, C. Estimation of safety of four-legged unsignalized intersections. **TRR: Journal of the Transportation Research Board**, Washington, D.C., n. 1467, p. 23-29, 1994.
- BERNARDINO, A. R. **Especialização dos acidentes de trânsito em Uberlândia (MG): Técnicas de Geoprocessamento como instrumento de análise - 2000 a 2004.** Tese de Doutorado, Universidade de São Paulo. São Paulo. 2007.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e aplicações.** Rio de Janeiro: LTC, 2000.
- CEFTRU. **Procedimentos para o tratamento de locais críticos de acidentes de trânsito.** Brasília, 2002.
- CUNTO, F. J. C. **Assessing Safety Performance of Transportation Systems using Microscopic Simulation.** Universidade de Waterloo. Ontário. 2008.
- CUNTO, F. J. C.; NODARI, C. T. **Aspectos conceituais sobre o desenvolvimento de modelos de previsão de acidentes de trânsito.** XXV ANPET - Congresso de Pesquisa e Ensino em Transportes. Belo Horizonte: ANPET. 2011. p. 1854-1865.
- CYBENKO. Approximation by Superpositions of a Sigmoidal Function. **Mathematics of Control, Signal, and Systems**, New York, Inc., n. 2, p. 303-314, 1989.
- DA SILVA, C. A. U. **Um método para estimar observáveis GPS usando Redes Neurais Artificiais.** Tese de Doutorado, Departamento de Transportes, EESC-USP. São Paulo. 2003.
- DENATRAN. **Manual de identificação, análise e tratamento de pontos negros.** Brasília, 1987.

DETRAN. **Acidentes de trânsito no estado segundo o tipo - jurisdição da via - até dezembro de 2010.** Fortaleza. Disponível em:<http://portal.detran.ce.gov.br/index.php/estatisticas>, 2010.

DETRAN. **Comparativo total de vítimas fatais e não fatais - mensal - 2009 -2010.** Fortaleza. Disponível em:<http://portal.detran.ce.gov.br/index.php/estatisticas>, 2010.

DETRAN. **Estatísticas gerais de acidentes - até dezembro de 2010.** Disponível em:<http://portal.detran.ce.gov.br/index.php/estatisticas>, 2010.

DETRAN. **Índice de acidentes por 10.000 veículos - 2002 - 2010.** Disponível em:<http://portal.detran.ce.gov.br/index.php/estatisticas>, 2010.

DIÓGENES, M. C.; LINDAU, L. A. Avaliando ações de segurança viária através de indicadores. **Transportes**, v. XII, p. 29-36, Dezembro 2004.

FERRAZ, A. C. P.; PRAI JUNIOR, A. A.; BEZERRA, B. S. **Segurança no Trânsito.** Grupo São Francisco. São Paulo. 2008.

HAUER, E. **Observational Before-after Studies in Road Safety.** Pergamon, 2002.

HAUER, E.; JERRY, C. N. N.; LOVELL, J. Estimation of safety at signalized intersections. **Transportation Research Record**, Washington, D.C., n. 1185, p. 48-61, 1988.

HAYKIN. **Neural Networks, A Comprehensive Foundation.** 2ª. ed. Ontário, 2001.

LÖTTER, S. **Road Safety Diagnostic System for South Africa.** Institute of Technology, Linköping University. Linköping. 2000.

MCCULLAGH, P.; NELDER, J. A. **Generalized Linear Models.** 2ª. ed. Chapman & Hall/CRC, 1989.

MUSSONE, L.; FERRARI, A.; ONETA, M. An Analysis of Urban Collisions Using an Artificial Intelligence Model. **Accident Analysis & Prevention**, n. 31, p. 705-718, 1999.

PERSAUD, B.; MUCSI, K. Microscopic Accident potential models for two-lane rural roads. **TRR. Journal of the Transportation Research Board**, Washington, D.C., n. 1485, p. 134-139, 1995.

RIBEIRO, A. J. A. Um método para localização e estimação das características geotécnicas dos solos da região metropolitana de Fortaleza-Ce para fins de pavimentação. Dissertação de Mestrado, Universidade Federal do Ceará. 2013.

