

Universidade Federal de São Carlos
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
Programa de Pós-Graduação em Estruturas e Construção Civil

**MÉTODO PARA PROGNÓSTICO DO
CONSUMO DE MATERIAIS EM
INSTALAÇÕES PREDIAIS
ELÉTRICAS UTILIZANDO
SISTEMAS INTELIGENTES**

Raphael Negri Milion

**MÉTODO PARA PROGNÓSTICO DO
CONSUMO DE MATERIAIS EM
INSTALAÇÕES PREDIAIS
ELÉTRICAS UTILIZANDO
SISTEMAS INTELIGENTES**

Raphael Negri Milion

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Estruturas e Construção Civil da Universidade Federal de São Carlos como parte dos requisitos para a obtenção do Título de Mestre em Estruturas e Construção Civil.

Área de concentração: Sistemas Construtivos

Orientador: Prof. Dr. José Carlos Paliari

São Carlos

2014

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

M644mp Milion, Raphael Negri.
Método para prognóstico do consumo de materiais em
instalações prediais elétricas utilizando sistemas inteligentes
/ Raphael Negri Milion. -- São Carlos : UFSCar, 2014.
182 p.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2014.


1. Construção civil. 2. Materiais - consumo. 3. Edifícios -
instalações elétricas. 4. Sistemas inteligentes. I. Título.

CDD: 690 (20^a)

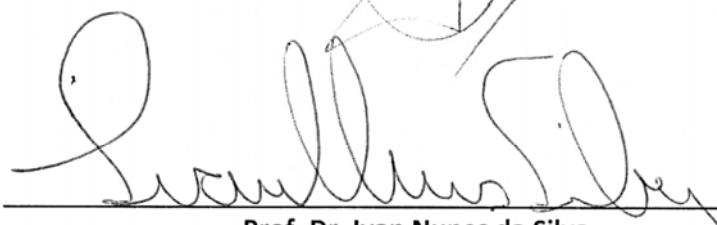
RAPHAEL NEGRI MILION

Dissertação de Mestrado defendida e aprovada em 11 de agosto de 2014.

Banca Examinadora constituída pelos membros:



Prof. Dr. José Carlos Paliari
Departamento de Engenharia Civil/PPGECiv/UFSCar
Orientador



Prof. Dr. Ivan Nunes da Silva
Departamento de Engenharia Elétrica/PPGEE//EESC-USP



Profª Drª Fernanda Fernandes Marchiori
Departamento de Engenharia Civil/Universidade Federal de Santa Catarina/UFSC

Agradecimentos

Primeiramente à minha esposa Luisa pelo companheirismo e amor durante todo o tempo, pelo incentivo ao ingresso no programa de mestrado, por sempre estar disponível para ajudar com o que fosse preciso e até mesmo por cobrar que eu fosse persistente. Sem você esse trabalho jamais existiria. Também não posso deixar de agradecê-la pela revisão do texto.

Agradeço também o Jacques que foi um companheiro, fonte de lazer e distração, sempre que o cansaço chegava ao limite.

À minha mãe, minha avó e a toda minha família pela educação e apoio que tive.

À família da minha esposa que sempre me apoia desde que nos conhecemos.

Ao Paliari pela orientação, incentivo e confiança durante todo o programa de mestrado.

À Mércia por me ajudar a chegar até o PPGECIV.

Aos engenheiros Leandro Rocha, Eliana Guadalupe (Stéfani Nogueira), Vladimir Viola (V&V Engenharia), Marcelo Matsusato e Rafael Rossi (Tecnisa) pela disponibilização de informações.

À todos os colegas da Stéfani Nogueira que de alguma forma apoiaram esse trabalho, em especial aos engenheiros Eduardo Nogueira e Márcio de la Corte que apoiaram e disponibilizaram os recursos necessários para que a realização do trabalho fosse possível.

“Se enxerguei mais longe foi por estar sobre os ombros de gigantes.”
(Isaac Newton)

Resumo

Devido à importância da previsão de custos nas fases iniciais de empreendimentos, nas quais é possível intervir no produto de forma a obter impactos nos custos de produção; e também, devido à dificuldade de prognóstico do consumo de materiais das instalações prediais elétricas, que ainda não contam com ferramentas eficazes para tal tarefa, este trabalho propõe modelos para prognóstico do consumo de materiais de instalações prediais elétricas a partir de informações disponíveis na fase do estudo de viabilidade e estudo preliminar de empreendimentos. São utilizadas para o prognóstico sistemas inteligentes, mais especificamente, redes neurais artificiais, além de métodos convencionais, como regressão linear e índices de consumo. O método de pesquisa consta das etapas: a) constituição de um banco de dados com informações coletadas em quantitativos utilizados para orçamentos; b) análise e preparação dos dados para a devida utilização dos sistemas inteligentes e convencionais c) seleção de atributos para identificação das características com melhor capacidade de realizar o prognóstico e d) desenvolvimento dos modelos de prognóstico e análise de desempenho por meio de comparações entre os valores prognosticados e os valores reais. Os modelos desenvolvidos melhoram significativamente o prognóstico de consumo destes sistemas em relação a prática atual, que consiste em levantar quantitativos de projeto e multiplicá-los por índices de consumo. Os modelos propostos também permitem tomadas de decisões com maior consciência dos impactos nos custos, principalmente em fases iniciais dos empreendimentos. Espera-se também que a metodologia apresentada possa ser extrapolada para outros serviços de construção.

Palavras-chave: Consumo de materiais. Instalações prediais elétricas. Sistemas inteligentes.

Abstract

Given the importance of forecasting costs in early stages of architectural projects, when it is possible to make changes in the product design and therefore obtain changes in the production costs, and also due to the difficulty of electrical-material consumption prognosis, this research proposes models for predicting electrical-material consumption used in buildings' electrical installations. It was used artificial neural networks, an intelligent system, and conventional methods, such as linear regression and consumption rates for the prognostic models. The available data were collected from projects' feasibility study and draft design. The research method includes the following steps: a) creation of a database with information collected in quantitatives used for estimates, b) data analysis and preprocessing for use in intelligent and conventional systems, c) attribute selection for best feature identification, i.e, for identifying features with high ability to influence the prognosis and d) development of the models and performance analysis, comparing the predicted values with the actual values. The developed models improves the consumption prognosis performance when compared with common prognostic tools. Current tools consists in multiplying quantitatives by a consumption rate. Also, the novel models allows more cautious decision-making in projects' early design phases, allowing greater awareness of costs impacts. It is expected that this methodology could be used for predicting other building materials.

Keywords: Materials consumption. Electrical building installations. Intelligent Systems.

Lista de ilustrações

Figura 1	Fluxograma das atividades que compõem a estratégia de pesquisa. . . .	26
Figura 2	“Valores de metros de tubulação e número de conexões por área de apartamento tipo – prognóstico: tubos de quedas de águas pluviais”. .	29
Figura 3	Proposta de processo para melhoria contínua do prognóstico de consumo de materiais na construção civil.	30
Figura 4	Composição da TCPO para tomadas.	31
Figura 5	Faixa de consumo unitário de argamassa para revestimento.	32
Figura 6	Modelo geral de neurônio artificial.	39
Figura 7	Exemplo de topologia utilizada para problemas de aproximação de função.	44
Figura 8	Ilustração do funcionamento de um PMC, como aquele apresentado na Figura 7, aplicado a aproximação de funções, com a superposição de funções logísticas.	44
Figura 9	Gráfico entradas x saídas de uma RNA sem sobreajuste.	45
Figura 10	Gráfico entradas x saídas de uma RNA com sobreajuste.	46
Figura 11	Exemplo de curva de aprendizado.	48
Figura 12	Curva de aprendizagem para os modelos que utilizam RNA para prognóstico de eletrodutos de $\frac{1}{2}$ ” e $\frac{3}{4}$ ”.	83
Figura 13	Curva de aprendizagem para os modelos que utilizam regressão linear para prognóstico de eletrodutos de $\frac{1}{2}$ ” e $\frac{3}{4}$ ”.	84
Figura 14	Quantidade de eletroduto $\frac{1}{2}$ ” e $\frac{3}{4}$ ” por ponto de luz. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	86
Figura 15	Quantidade de eletroduto $\frac{1}{2}$ ” e $\frac{3}{4}$ ” por ponto de luz + interruptores + tomadas + chuveiros elétricos + condicionadores de ar. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	87
Figura 16	Curva de aprendizagem para os modelos que utilizam RNA para prognóstico de condutores de $1,5 \text{ mm}^2$	93

Figura 17	Curva de aprendizagem para os modelos que utilizam regressão linear para prognóstico de condutor de $1,5 \text{ mm}^2$	94
Figura 18	Quantidade de condutor de $1,5 \text{ mm}^2$ por ponto de luz. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	96
Figura 19	Quantidade de condutor de $1,5 \text{ mm}^2$ por “Interruptor”. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	96
Figura 20	Curva de aprendizagem para os modelos que utilizam RNA para prognóstico de condutores de $2,5 \text{ mm}^2$	102
Figura 21	Curva de aprendizagem para os modelos que utilizam regressão linear para prognóstico de condutor de $2,5 \text{ mm}^2$	103
Figura 22	Curva de aprendizagem para os modelos que utilizam RNA e regressão linear para prognóstico de condutores de $2,5, 4$ e 6 mm^2	103
Figura 23	Quantidade de condutor de $2,5 \text{ mm}^2$ por tomadas. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	105
Figura 24	Quantidade de condutor de $2,5 \text{ mm}^2$ por “Área”. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	106
Figura 25	Quantidade de condutor de $2,5, 4$ e 6 mm^2 por “Tomadas + Chuveiros Elétricos + Condicionadores de Ar”. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	107
Figura 26	Curva de aprendizagem para os modelos que utilizam RNA para prognóstico de condutores de 4 e 6 mm^2	113
Figura 27	Curva de aprendizagem para os modelos que utilizam regressão linear para prognóstico de condutor de 4 e 6 mm^2	114
Figura 28	Quantidade de condutor de 4 e 6 mm^2 por área. Prognóstico utilizando RNA e regressão linear e dados de treinamento.	115
Figura 29	Quantidade de condutor de 4 e 6 mm^2 por área. Prognóstico utilizando RNA e regressão linear e dados de treinamento. Sem dados de treinamento referente às instâncias 15, 22, 26 e 27.	116

Lista de tabelas

Tabela 1	Composições do TCPO que podem ser utilizadas no prognóstico do CUM de IPE	33
Tabela 2	Banco de dados. Obras 1 a 30. Parte 1 de 3.	64
Tabela 3	Banco de dados. Obras 1 a 30. Parte 2 de 3.	65
Tabela 4	Banco de dados. Obras 1 a 30. Parte 3 de 3.	66
Tabela 5	Banco de dados. Obras 31 a 60. Parte 1 de 3.	67
Tabela 6	Banco de dados. Obras 31 a 60. Parte 2 de 3.	68
Tabela 7	Banco de dados. Obras 31 a 60. Parte 3 de 3.	69
Tabela 8	Banco de dados após análise preliminar. Obras 1 a 30. Parte 1 de 3.	73
Tabela 9	Banco de dados após análise preliminar. Obras 1 a 30. Parte 2 de 3.	74
Tabela 10	Banco de dados após análise preliminar. Obras 1 a 30. Parte 3 de 3.	75
Tabela 11	Banco de dados após análise preliminar. Obras 31 a 60. Parte 1 de 3.	76
Tabela 12	Banco de dados após análise preliminar. Obras 31 a 60. Parte 2 de 3.	77
Tabela 13	Banco de dados após análise preliminar. Obras 31 a 60. Parte 3 de 3.	78
Tabela 14	Resultado dos algoritmos de seleção de atributos para eletrodutos com informações de pontos de utilização.	79
Tabela 15	Resultado dos algoritmos de seleção de atributos para eletrodutos sem informações de pontos de utilização.	80
Tabela 16	Análise de desempenho dos modelos de prognóstico da quantidade de eletrodutos de $\frac{1}{2}$ " e $\frac{3}{4}$ "	82
Tabela 17	Resultado dos algoritmos de seleção de atributos para condutores de $1,5 \text{ mm}^2$ com informações de pontos de utilização.	89
Tabela 18	Resultado dos algoritmos de seleção de atributos para condutores de $1,5 \text{ mm}^2$ sem informações de pontos de utilização.	90
Tabela 19	Análise de desempenho dos modelos de prognóstico da quantidade de condutores de $1,5 \text{ mm}^2$	92

Tabela 20	Resultado dos algoritmos de seleção de atributos para condutores de 2,5 mm^2 com informações de pontos de utilização.	98
Tabela 21	Resultado dos algoritmos de seleção de atributos para condutores de 2,5 mm^2 sem informações de pontos de utilização.	98
Tabela 22	Análise de desempenho dos modelos de prognóstico da quantidade de condutores de 2,5 mm^2	101
Tabela 23	Análise de desempenho dos modelos de prognóstico da quantidade de condutores de 2,5, 4 e 6 mm^2	101
Tabela 24	Resultado dos algoritmos de seleção de atributos para condutores de 4 e 6 mm^2 com informações de pontos de utilização.	109
Tabela 25	Resultado dos algoritmos de seleção de atributos para condutores de 4 e 6 mm^2 sem informações de pontos de utilização.	110
Tabela 26	Análise de desempenho dos modelos de prognóstico da quantidade de condutores de 4 e 6 mm^2	112
Tabela 27	Amostras inéditas. Parte 1 de 2.	117
Tabela 28	Amostras inéditas. Parte 2 de 2.	117
Tabela 29	Teste em dados inéditos	118

Lista de siglas

BIM *Building Information Modeling*

CfsSubsetEval Seleção de Subconjunto de Atributos Baseada em Correlação para Aprendizado de Máquinas

CPOS Companhia Paulista de Obras e Serviços

CUM Consumo Unitário de Materiais

FDE Fundação para o Desenvolvimento da Educação

IPE Instalações Prediais Elétricas

MCMV Minha Casa Minha Vida

PAC Programa de Aceleração do Crescimento

PBQP-h Programa Brasileiro da Qualidade e Produtividade do Habitat

PMC *Perceptron* Multicamadas

PMSP Prefeitura Municipal de São Paulo

PRJ Prefeitura do Rio de Janeiro

RNA Redes Neurais Artificiais

SINAPI Sistema Nacional de Pesquisa de Custos e Índices da Construção Civil

Sinduscon Sindicato da Indústria da Construção

TCPO Tabelas de Composição de Preços para Orçamento

TUE Tomadas de Uso Específico

TUG Tomadas de Uso Geral

Sumário

1	Introdução	21
1.1	Justificativa	21
1.2	Objetivos	24
1.2.1	Objetivos específicos	24
1.3	Delimitações	24
1.4	Estratégia de Pesquisa	25
2	Revisão Bibliográfica	27
2.1	Consumo de Materiais	27
2.1.1	Tabelas de Composição de Preços para Orçamentos	30
2.1.2	Outros índices de consumo	35
2.2	Instalações Prediais Elétricas	35
2.3	Redes Neurais Artificiais	37
2.3.1	Arquitetura e topologia	39
2.3.2	Treinamento	40
2.3.3	Aproximador universal de funções	43
2.3.4	Sobreajuste	44
2.3.5	Curvas de Aprendizagem	47
2.3.6	Aplicações de Redes Neurais Artificiais na área de gestão de recursos da construção civil	48
2.3.7	Redes Neurais Artificiais para o prognóstico do consumo unitário de materiais	49
2.4	Seleção de Atributos	50
3	Método de Pesquisa	55
3.1	Coleta de dados	55
3.2	Análise preliminar	57
3.3	Seleção de Atributos	58

3.4	Modelos de Prognóstico	59
3.4.1	Curvas de aprendizagem	61
3.4.2	Gráficos dos Modelos Simplificados	61
3.4.3	Teste em Dados Inéditos	61
4	Desenvolvimento dos Modelos	63
4.1	Dados coletados	63
4.2	Análise preliminar	70
4.2.1	Eletrodutos de 1/2" e de 3/4"	70
4.2.2	Condutores de 4 e 6 mm^2	70
4.2.3	Ano, construtora, cidade e projetista	70
4.2.4	Combinação de atributos	71
4.2.5	Banco de dados após análise preliminar	72
4.3	Eletrodutos	79
4.3.1	Preparação dos dados	79
4.3.2	Seleção de atributos	79
4.3.3	Modelos de prognóstico	80
4.3.4	Gráficos dos modelos simplificados unidimensionais	86
4.4	Condutor de 1,5 mm^2	88
4.4.1	Preparação dos dados	88
4.4.2	Seleção de atributos	88
4.4.3	Modelos de prognóstico	91
4.4.4	Gráficos dos modelos simplificados unidimensionais	95
4.5	Condutor de 2,5 mm^2	97
4.5.1	Preparação dos dados	97
4.5.2	Seleção de atributos	97
4.5.3	Modelos de prognóstico	99
4.5.4	Gráficos dos modelos simplificados unidimensionais	105
4.6	Condutores de 4 e 6 mm^2	108
4.6.1	Preparação dos dados	108
4.6.2	Seleção de atributos	108
4.6.3	Modelos de prognóstico	110
4.6.4	Gráficos dos modelos simplificados unidimensionais	115
4.7	Teste em Dados Inéditos	116
5	Considerações Finais	121
5.1	Sugestões para trabalhos futuros	123
	Referências	125

.1	Relatórios do Weka - Eletrodutos	133
.1.1	Com Pontos de Utilização - Regressão linear	133
.1.2	Com Pontos de Utilização - Rede Neural Artificial	134
.1.3	Sem Pontos de Utilização - Regressão linear	139
.1.4	Sem Pontos de Utilização - Rede Neural Artificial	140
.1.5	Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condi- cionadores de Ar - Regressão linear	142
.1.6	Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condi- cionadores de Ar - Rede Neural Artificial	143
.1.7	Luz - Regressão linear	145
.1.8	Luz - Rede Neural Artificial	146
.2	Relatórios do Weka - Condutores de $1,5\text{ mm}^2$	147
.2.1	Com Pontos de Utilização - Regressão linear	147
.2.2	Com Pontos de Utilização - Rede Neural Artificial	149
.2.3	Sem Pontos de Utilização - Regressão linear	151
.2.4	Sem Pontos de Utilização - Rede Neural Artificial	152
.2.5	Luz - Regressão linear	154
.2.6	Luz - Rede Neural Artificial	155
.2.7	Interruptores - Regressão linear	156
.2.8	Interruptores - Rede Neural Artificial	157
.3	Relatórios do Weka - Condutores de $2,5\text{ mm}^2$	159
.3.1	Com Pontos de Utilização - Regressão linear	159
.3.2	Com Pontos de Utilização - Rede Neural Artificial	160
.3.3	Sem Pontos de Utilização - Regressão linear	163
.3.4	Sem Pontos de Utilização - Rede Neural Artificial	164
.3.5	Tomadas - Regressão linear	166
.3.6	Tomadas - Rede Neural Artificial	167
.3.7	Área - Regressão linear	169
.3.8	Área - Rede Neural Artificial	170
.4	Relatórios do Weka - Condutores de $2,5, 4$ e 6 mm^2	172
.4.1	Tomadas + Chuveiros Elétricos + Condicionadores de Ar - Re- gressão linear	172
.4.2	Tomadas + Chuveiros Elétricos + Condicionadores de Ar - Rede Neural Artificial	173
.5	Relatórios do Weka - Condutores de 4 e 6 mm^2	175
.5.1	Com Pontos de Utilização - Regressão linear	175
.5.2	Com Pontos de Utilização - Rede Neural Artificial	176
.5.3	Sem Pontos de Utilização - Regressão linear	177

.5.4	Sem Pontos de Utilização - Rede Neural Artificial	178
.5.5	Área - Regressão linear	180
.5.6	Área - Rede Neural Artificial	181

Introdução

1.1 Justificativa

O mercado de construção civil brasileiro encontra-se aquecido. Embora a crise financeira mundial de 2008 tenha afetado a oferta de capital de giro, derrubado as vendas, feito com que lançamentos fossem cancelados, desacelerando todo o setor (GALA; CASTELO, 2009), as medidas tomadas pelo governo federal para combater seus efeitos, não só reverteram a situação como proporcionaram o crescimento do setor (MONTEIRO FILHA et al., 2010). Cita-se como principais medidas:

- ❑ Redução do IPI sobre materiais de construção;
- ❑ Implantação do Programa Minha Casa Minha Vida (MCMV), com grande ampliação do crédito imobiliário;
- ❑ Aporte de investimentos no Programa de Aceleração do Crescimento (PAC).

As perspectivas futuras também são otimistas, já que o país tem sediado a Copa do Mundo de Futebol em 2014 e sediará as Olimpíadas em 2016, eventos que geram investimentos em arenas, hotéis, infraestrutura e outros, aquecendo a economia como um todo (MONTEIRO FILHA et al., 2010). Dividem esse mercado cerca de 64 mil empresas de construção civil (IBGE, 2009).

Nos países desenvolvidos, até as décadas de 1960 e 1970 a única preocupação estratégica das indústrias era o aumento da produção. Tudo que era produzido era vendido e as margens de lucro eram grandes. Com o aumento do número de empresas no mercado e com a globalização, onde empresas passam a ter atuação internacional, o cenário mudou e o aumento da oferta levou a compressão dos preços e das margens de lucro, devido a competição entre os fabricantes. Também houve mudança no perfil dos consumidores, que passaram a exigir mais qualidade e produtos com projetos inovadores. Essas preocupações ocorrem desde o fim do século passado e vem se intensificando até hoje (BOLWIJN; KUMPE, 1990).

As mesmas mudanças ocorreram cada vez com mais intensidade nos países em desenvolvimento (MOREIRA, 1991). Segundo Picchi (1993) e Cardoso (1996) as principais causas apontadas para o acirramento da competição entre as empresas e o aumento da exigência dos consumidores na indústria brasileira de construção, desde a década de 1990, são, entre outras:

- ❑ Aumento da construção para o mercado privado;
- ❑ Abertura das importações e entrada de agentes estrangeiros e
- ❑ Entrada em vigor do Código de Defesa do Consumidor.

Acrescenta-se a popularização do uso da internet, ampliando o nível de informação dos consumidores, os quais agora tem mais facilidade de buscar e comparar produtos disponíveis no mercado. Também o uso das redes sociais pelos consumidores para comentar suas satisfações, e, principalmente, suas insatisfações. Essas redes passam a ser um canal de comunicação com o consumidor e a reputação das empresas nas redes uma importante ferramenta de *marketing* (KÄRKKÄINEN; JUSSILA; LEINO, 2012).

Com o acirramento da concorrência e aumento da oferta, o preço de venda é definido pelo mercado e o lucro de um empreendimento passa a ter como variável seu custo (CARDOSO, 2009). Desta forma, é necessário um estudo mais aprofundado da formação dos custos e suas relações com materiais e mão de obra. O conceito de Consumo Unitário de Materiais (CUM) é utilizado para medir a eficiência do uso de materiais e componentes (SOUZA, 2001).

Uma série de programas desenvolvidos pelo setor produtivo, por entidades patronais, governamentais e por universidades foi implementada com o objetivo de estudar a produtividade de mão de obra, CUM e a qualidade da construção. Citam-se:

- ❑ Programa da Qualidade e da Produtividade na Construção Civil do Rio Grande do Sul;
- ❑ Sistema de Gestão para Empresas Construtoras do Sindicato da Indústria da Construção (Sinduscon) de São Paulo;
- ❑ Programa da Qualidade da Construção Habitacional do Estado de São Paulo;
- ❑ Programa Brasileiro da Qualidade e Produtividade do Habitat (PBQP-h) (PALIARI, 2008) e
- ❑ Revisão das Composições de Serviços do Sistema Nacional de Pesquisa de Custos e Índices da Construção Civil (SINAPI) realizada pela Fundação para Desenvolvimento Tecnológico da Engenharia por encomenda da Caixa Econômica Federal.

Um objetivo comum a todos esses programas é a promoção do melhor gerenciamento do sistema produtivo e da qualidade dos produtos e serviços, tanto dos próprios quanto o de fornecedores (PALIARI, 2008).

Seja qual for o caminho adotado pelas empresas de construção civil, torna-se imprescindível o conhecimento dos fatores que influenciam o consumo de materiais e como, quantitativamente, essa influência ocorre (SOUZA, 2001). Para tanto, também é preciso conhecer o desempenho atual das empresas referente ao consumo de materiais e, para isso, são necessários métodos de coleta e análise de dados que ofereçam precisão para embasar propostas de melhorias (PALIARI, 1999), assim como, métodos para a realização de prognósticos do consumo de materiais considerando os fatores que nele interferem (PALIARI, 2008). Com a redução das margens de lucro dos empreendimentos, em especial aqueles destinados a habitação de interesse social financiadas pelo Estado, errar no prognóstico do consumo de materiais significa comprometer significativamente o retorno financeiro do empreendimento (MARCHIORI, 2009).

A necessidade de prever os custos de construção nas fases iniciais de um empreendimento também decorre da necessidade de estudos de viabilidade. De acordo com Mascaro (1998) é a concepção de uma edificação que define seu custo. Alterações de produto com objetivo de reduzir custos são eficientes nas fases iniciais dos empreendimentos. Alterações em fases adiantadas de projeto, durante ou após a execução, implicam em retrabalho, que custam e comprometem a sua qualidade (MELHADO; AGOPYAN, 1995). Além disso, após venda na planta, alterações de produto e do preço de venda são inviáveis na prática, portanto, para garantir a lucratividade prevista em estudo de viabilidade, a precisão na previsão de custos feita nas fases iniciais de projeto é essencial. Por exemplo, a previsão de custos é de extrema importância na fase de viabilidade, na qual, o empreendimento não possui projeto executivo, o qual permite realizar levantamentos para orçamentação (OLIVEIRA; BELCHIOR; JUNGLES, 2007). Também é importante a previsão de custos na fase de estudo preliminar, a qual consiste na fase dos projetos iniciais do produto. Falhas nessa previsão podem provocar redução da margem de lucro esperada, redução no padrão de acabamento (promovendo a insatisfação dos clientes) ou, em último caso, interrupção do empreendimento por falta de recursos (CARDOSO, 2009) (GONCALVES, 2001).

Especificamente para serviços de instalações prediais, ainda são comuns, empresas que utilizam “verbas” em seu orçamento, abstendo-se de levantar quantitativos ou de usar prognósticos de CUM em função da sua terceirização e/ou da sua complexidade, ou mesmo devido a atribuição de importância secundária a esse serviço por parte dos construtores (PALIARI, 2008). Quando essas “verbas” são estimadas por meio de comparações simples entre diferentes empreendimentos, sem o entendimento dos fatores que influenciam as variáveis estimadas, podem gerar grandes erros (STOY; SCHAULCHER, 2007).

Um método capaz de realizar o prognóstico do consumo dos componentes de Instalações Prediais Elétricas (IPE) é necessário. Além da utilização de ferramentas estatísticas con-

vencionais, como médias, faixas de consumo e regressão linear, diante da complexidade do problema em questão o uso de Redes Neurais Artificiais (RNA) pode ser útil, já que, esta é capaz de trabalhar com diversas variáveis de entrada (fatores que influenciam o consumo) e captar características não lineares do prognóstico do consumo e, vem sendo utilizada para estimativas de custos e consumo (problemas não lineares que envolvem grande número de variáveis) (MARCHIORI, 2009).

As RNA são sistemas inteligentes que podem ser utilizados tanto para problemas de regressão quanto de classificação de padrões. Elas fazem parte de um conjunto de ferramentas computacionais chamadas sistemas inteligentes, que têm o objetivo de simular mecanismos da inteligência humana (SILVA; SPATTI; FLAUZINO, 2010).

O contexto apresentado é aquele no qual o presente trabalho se desenvolve.

1.2 Objetivos

Desenvolver um método para prognóstico do consumo de materiais empregados em IPE, utilizando sistemas inteligentes, mais especificamente, RNA, a partir de informações de projeto disponíveis na fase de estudo de viabilidade e de estudo preliminar de empreendimentos.

1.2.1 Objetivos específicos

Os objetivos específicos são:

- ❑ Realizar prognósticos utilizando sistemas convencionais, mais especificamente a regressão linear, e comparar os resultados obtidos com aqueles obtidos com RNA;
- ❑ Aplicar o método proposto para gerar modelos de prognósticos de consumo de materiais em IPE. Esses modelos também são um produto do trabalho e poderão ser utilizados para realização de prognósticos por agentes que não possuam dados próprios para a aplicação do método proposto e
- ❑ Gerar modelos simplificados e apresentá-los de forma gráfica, facilitando sua utilização pelos usuários em potencial destes, como, por exemplo, orçamentistas e projetistas.

1.3 Delimitações

O presente trabalho limita-se a estudar o consumo teórico, levantado em projetos, de eletrodutos e condutores de apartamentos tipo de edifícios residenciais de múltiplos pavimentos,.

Os eletrodutos e condutores foram escolhidos para a realização do trabalho devido aos seguintes fatores:

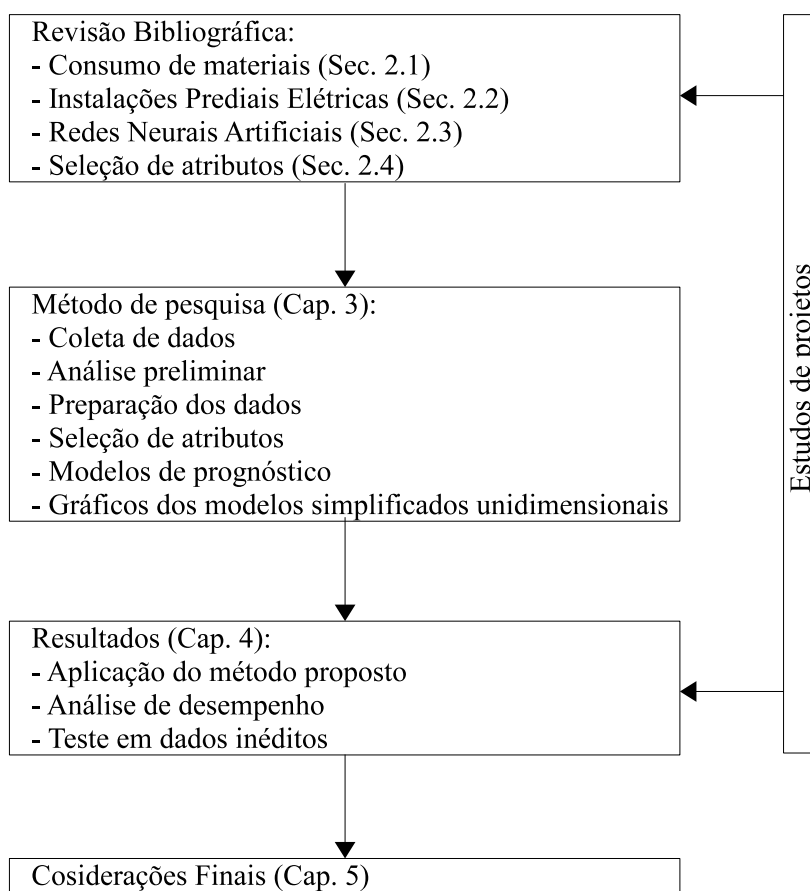
- ❑ Devido ao grande número de componentes existentes em uma IPE, uma seleção prévia de componentes foi realizada, de forma a garantir a aplicação do método exposto. A extensão do método à outras variáveis será discutida em sessões futuras deste texto;
- ❑ O levantamento de eletrodutos e condutores em projeto é trabalhoso, já que seus caminhamentos apresentam curvas e ocorrem em três dimensões, enquanto o levantamento dos demais componentes resume-se a uma contagem, mais simples, e
- ❑ Grande parte dos demais componentes, como caixas, interruptores, quadros, entre outros, podem ser levantados a partir do projeto de arquitetura, antes mesmo do início do projeto de IPE.

1.4 Estratégia de Pesquisa

A estratégia utilizada no trabalho foi a de modelagem e simulação, ou seja, o desenvolvimento de um modelo de prognóstico de um sistema construtivo real com objetivo de conduzir experimentos para entender o comportamento desse sistema, dentro das condições de contorno impostas (SHANNON, 1998). Um modelo de simulação captura o comportamento do sistema real, sendo capaz de representar sistemas complexos, de natureza dinâmica e aleatória (CHWIF; MEDINA, 2006).

Para o desenvolvimento de um método que gere modelos capazes de prognosticar o consumo teórico de materiais de IPE, será realizado um conjunto de atividades, apresentadas de forma resumida, assim como suas conexões, na Figura 1. O detalhamento de cada uma das atividades apresentadas encontra-se no capítulo ou seção indicados na mesma figura.

Figura 1 – Fluxograma das atividades que compõem a estratégia de pesquisa.



Fonte: elaborado pelo autor.

Revisão Bibliográfica

A partir do contexto apresentado anteriormente e dos objetivos, delimitações e estratégias de pesquisa, desenvolve-se a revisão bibliográfica. O capítulo atual aborda o objeto de estudo e as ferramentas utilizadas. Para tanto, é feita uma revisão sobre consumo de materiais e sistemas prediais elétricos e também sobre RNA e algoritmos de seleção de atributos.

2.1 Consumo de Materiais

Segundo Souza (2001), consumo é a relação entre entradas demandadas e saídas de um processo. Mais especificamente, CUM é a relação entre materiais utilizados para realizar uma unidade de um serviço ou de parte de uma obra. O consumo ainda pode ser dividido em consumo teórico, no qual são consideradas as quantidades teoricamente necessárias de materiais, e em consumo real, no qual são consideradas as quantidades de materiais efetivamente consumidas nas obras, as quais incluem perdas, incorporadas ou não. Enquanto a quantidade teoricamente necessária é levantada por meio de projetos e especificações, tanto do produto, quanto dos processos; as perdas são calculadas a partir da quantidade de material utilizada na obra, subtraída da quantidade teoricamente necessária.

Melhorar a eficiência da construção envolve, então, reduzir o consumo de materiais, tanto com melhorias de projetos, ou seja, reduzir a quantidade teoricamente necessária, quanto com melhorias na execução dos serviços, ou seja, reduzir as perdas (PALIARI, 2008).

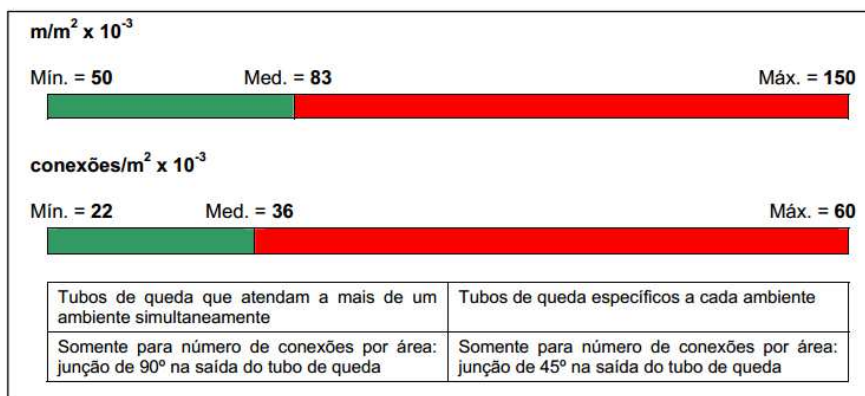
O Programa de Pós-Graduação do Departamento de Engenharia Civil da Escola Politécnica da Universidade de São Paulo, sob a coordenação do Prof. Dr. Ubiraci Espineli Souza, tem atuação destacada no estudo de produtividade e consumo de materiais. Como trabalhos mais relevantes, cita-se os realizados por Souza (1996), Paliari (1999 e 2008), Carraro (1998), Andrade (1999), Araujo (2000) e Librais (2001). Dentre os serviços estudados, encontram-se:

- ❑ Formas para estrutura de concreto armado;
- ❑ Armadura para estruturas de concreto armado;
- ❑ Concreto para estruturas de concreto armado;
- ❑ Blocos para alvenaria;
- ❑ Argamassa para assentamento de alvenaria;
- ❑ Argamassa para revestimento interno de paredes (emboço ou massa única);
- ❑ Argamassa para revestimento de fachada (emboço ou massa única);
- ❑ Argamassa para contrapiso interno;
- ❑ Gesso para revestimento;
- ❑ Placas cerâmicas para revestimento interno e
- ❑ Instalações prediais hidráulicas.

Os trabalhos de Paliari (1999 e 2008) são os únicos que abordam sistemas prediais, especificamente, instalações prediais hidráulicas. O erro obtido com dados disponíveis no estudo de viabilidade foi de 21,6% para o comprimento de tubos e 35,5% para o número de conexões. Quando utiliza dados disponíveis no projeto de arquitetura, os erros passam para 18,17% para o comprimento das tubulações e 19,83% para o número de conexões. Não foram encontrados na literatura trabalhos que abordem consumo de materiais em IPE.

A metodologia de prognóstico de materiais utilizada nos trabalhos citados anteriormente baseiam-se em estatística convencional, com levantamento de dados, em projetos ou em obras, e na definição de faixas de valores, incluindo mínimos, medianas e máximos. Também são indicados os fatores que deslocam o consumo no sentido do valor mínimo ou do valor máximo. Embora verifique-se grande avanço na realização dos prognósticos de consumo de materiais, observa-se que nem sempre a intensidade da influência de cada fator analisado em relação ao consumo é apresentada (PALIARI, 2008). Na Figura 2 é apresentado uma faixa de valores para prognóstico de tubos de quedas de águas pluviais.

Figura 2 – “Valores de metros de tubulação e número de conexões por área de apartamento tipo – prognóstico: tubos de quedas de águas pluviais”.



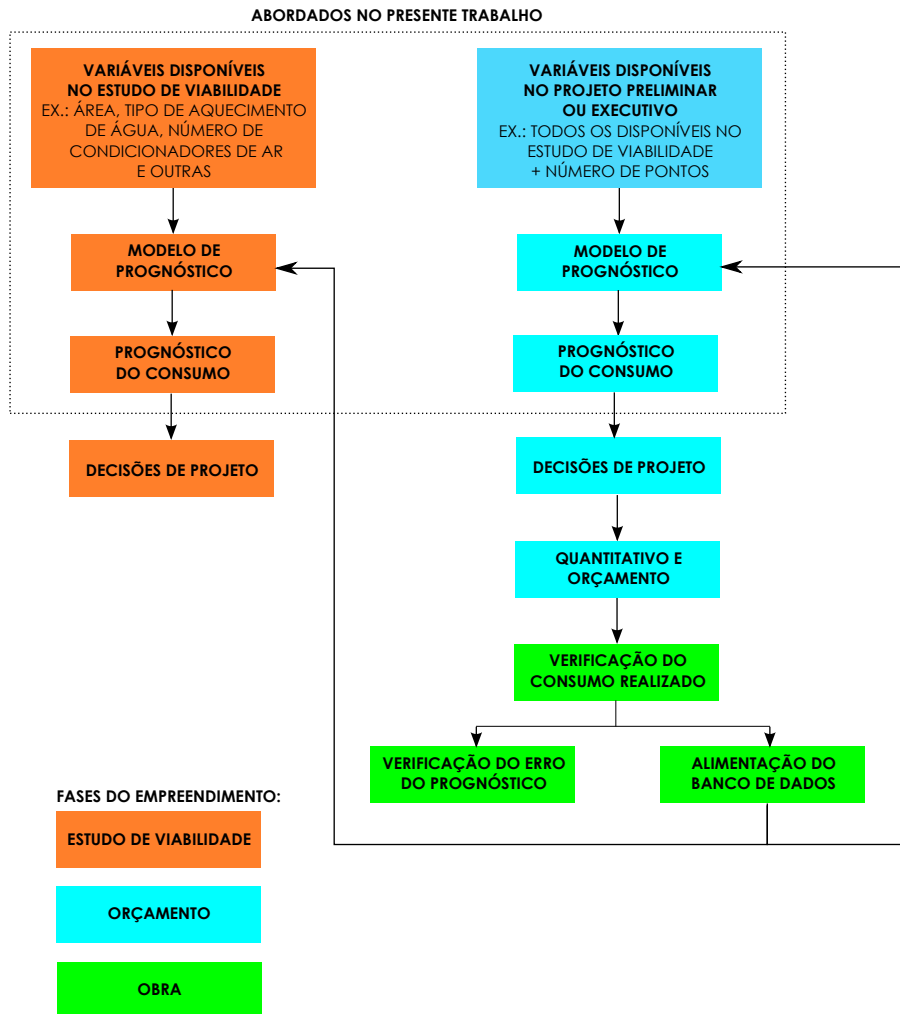
Fonte: Paliari (2008).

Os seguintes passos são utilizados de forma genérica por Souza (2001) para estimar o CUM:

- ❑ Definição de faixas (máximos e mínimos) de consumos associados a diferentes caracterizações de serviços;
- ❑ Indicação de fatores que auxiliem na estimativa de um valor dentro das faixas definidas anteriormente e
- ❑ Em alguns casos, haverá expressões matemáticas para a estimativa de um valor dentro das faixas de consumo.

Segundo Souza (2001), embora sua pesquisa sobre produtividade de mão de obra e de materiais seja muito extensa, o número de dados obtidos ainda é aquém do necessário para uma análise estatística de grande confiabilidade quando é levado em conta o número de construções do Brasil. Mesmo que isto fosse realizado para algum caso particular, este se desatualizaria com o tempo, a não ser que o produto desse trabalho seja um sistema que possa ser abastecido continuamente com informações. O processo apresentado na Figura 3 é uma alternativa para aperfeiçoar continuamente o prognóstico de consumo na construção civil, onde as empresas construtoras conseguiriam produzir seus próprios modelos de prognósticos e melhorá-los continuamente, em contrapartida ao processo corrente, que se limita a consulta de um manual de orçamentação.

Figura 3 – Proposta de processo para melhoria contínua do prognóstico de consumo de materiais na construção civil.



Fonte: elaborado pelo autor.

2.1.1 Tabelas de Composição de Preços para Orçamentos

Cotidianamente, nas empresas de construção e projeto, em âmbito nacional e internacional, informações sobre produtividade e consumo são obtidas por meio de manuais de orçamentação. No Brasil, o mais relevante, que detêm praticamente todo o mercado, é a publicação Tabelas de Composição de Preços para Orçamento (TCPO) da Editora Pini. (SOUZA, 2001)

Na indústria da construção civil, embora haja diversas formas de estimativa de custos da construção (custo unitário básico, regressões estatísticas, entre outros menos difundidos) (MARCHIORI, 2009) o processo de orçamentação é baseado na quantificação de serviços em projetos executivos e na multiplicação desses quantitativos por índices de consumo, retirados de tabelas de orçamentação correntes e por preços unitários. Muitas vezes, não é revelado o método de obtenção destes índices publicados, os quais são, consequentemente, contestados pelas empresas construtoras. Na falta de índices, são utilizadas

ainda médias do setor. Esse processo gera estimativas falhas (ANDRADE et al., 2001).

A primeira edição da TCPO foi lançada em 1955 e desde então tornou-se referência como manual de orçamentação no Brasil. Em seu início, esse manual era uma compilação do periódico semanal “Construção em São Paulo: Informador Profissional”. A partir de 2003 suas novas edições passaram a contar com informações de pesquisas acadêmicas sobre produtividade e CUM (PINI, 2008).

Alguns dos índices da TCPO, como é o caso daqueles que referem-se as IPE, são apresentados na forma de um único valor (Figura 4). Na Figura 4 observa-se que o consumo unitário de eletrodutos e condutores é fixa, independente de qualquer característica da edificação. Índices de outros serviços (como formas, armação, concreto, alvenaria e revestimentos) são apresentados por meio de faixas de valores, incluindo máximo, mediana e mínimo, de forma similar à metodologia dos trabalhos acadêmicos citados anteriormente (Figura 5).

Figura 4 – Composição da TCPO para tomadas.

16143.8.11.1 PONTO de tomada com eletroduto de PVC rígido, sem placa, Ø 3/4” – unidade: un

CÓDIGO	COMPONENTES	UNID.	CONSUMOS
01270.0.1.13	Ajudante de eletricista	h	4,50
01270.0.22.1	Eletricista	h	3,50
16120.3.7.1	Fio isolado em PVC (encordoamento: classe 1 / seção transversal: 1,50 mm ² / tensão: 750,00 V)	m	33,00
16132.3.1.2	Curva 90° de PVC rígido rosqueável para eletroduto (diâmetro da seção: 3/4”)	un	1,00
16132.3.3.2	Eletroduto de PVC rígido rosqueável (diâmetro da seção: 3/4”)	m	15,00
16132.3.4.2	Luva de PVC rígido rosqueável para eletroduto (diâmetro da seção: 3/4”)	un	1,00
16136.3.2.2	Caixa estampada em chapa de aço esmaltada de embutir 4” x 2” (formato da seção transversal: retangular / Chapa: 18)	un	1,00
16143.3.4.2	Tomada de embutir universal dois pólos redonda (corrente elétrica: 10,00 A / tensão: 250,00 V)	un	1,00

CONTEÚDO DO SERVIÇO

Esta composição destina-se a fase onde o orçamentista ainda não possui os projetos hidráulicos para levantamento das quantidades reais a serem consumidas na obra, porém já existe o projeto básico arquitetônico, sendo assim, é possível quantificar a quantidade de pontos. A partir de orçamentos realizados determinamos as quantidades usuais de material e mão de obra consumidas para execução do serviço.

CRITÉRIO DE MEDIÇÃO

Por unidade.

Fonte: PINI (2008).

Figura 5 – Faixa de consumo unitário de argamassa para revestimento.



Fonte: PINI (2008).

A metodologia utilizada para a geração e atualização dos dados da TCPO não é informada pela Editora Pini, o que reduz a confiabilidade dos dados. Há, também, uma frequente reclamação dos usuários da TCPO com a desatualização dos dados apresentados e muitas divergências entre dados apresentados por estudos acadêmicos e os apresentados pela publicação (SOUZA, 2001).

Há críticas também quanto à falta de regionalização dos dados oferecidos pela TCPO. Para Padilha Junior, Oliveira e Araújo (2007) e também para Araújo e Alves (2010) as composições da TCPO refletem as práticas das regiões Sul e Sudeste, não sendo aplicáveis satisfatoriamente às técnicas e materiais utilizados na Paraíba, por exemplo. É possível que mesmo dentro das regiões Sul e Sudeste essa afirmação seja válida, pois, há grande heterogeneidade nos materiais e técnicas utilizadas nessas regiões.

Para Paliari (2008) a possibilidade da utilização de dados próprios de cada empresa construtora é uma forma de conseguir estimar com mais precisão consumos unitários utilizados nos orçamentos dessas empresas. A utilização de dados locais ou regionais, levantados por meio de organizações como Sinduscon ou universidades, também seria uma alternativa para melhorar a precisão dos prognósticos.

2.1.1.1 Consumo Unitário de Materiais de Instalações Prediais Elétricas na Tabela de Composição de Preços para Orçamentos

Até a versão 13 da TCPO, publicada em 2010, as IPE não contavam com informações de CUM variáveis, ou seja, aquelas que apresentam valores medianos, mínimos e máximos,

além dos fatores influenciadores.

Na versão 13 da TCPO, para os diversos serviços e materiais das IPE são apresentados dados de produtividade e perdas. Esses dados são úteis quando já se foi levantado o quantitativo a partir de projetos executivos, quando precisa-se estimar as quantidades reais de materiais que serão utilizados e dimensionar equipes. Porém, o objetivo deste trabalho é prever o consumo de materiais a partir de informações disponíveis no estudo de viabilidade e no estudo preliminar, quando ainda não são disponíveis projetos das IPE.

As informações disponíveis na TCPO relacionadas ao objeto deste trabalho estão na Tabela 1, na qual é apresentada uma compilação de todas as composições que poderiam ser utilizadas para estimativas de consumo no início do empreendimento, quando ainda não há projeto de IPE.

Tabela 1 – Composições do TCPO que podem ser utilizadas no prognóstico do CUM de IPE

Código	Serviço	Eletroduto PVC rígido 3/4" (m)	Eletroduto PVC flexível 3/4" (m)	Condutor 1,5 mm² (m)
16100.8.3.1	Ponto seco com eletroduto de PVC rígido	6	-	-
16100.8.3.2	Ponto seco com eletroduto de PVC flexível	-	10	-
16132.8.18.1	Ponto de luz com eletroduto de PVC rígido	15	-	33
16143.8.11.1	Ponto de tomada com eletroduto de PVC rígido	15	-	33
16143.8.12.1	Ponto de interruptor com eletroduto de PVC rígido	15	-	33

Fonte: adaptado de PINI (2008)

Em todos esses itens, há a seguinte nota:

“Esta composição destina-se à fase onde o orçamentista ainda não possui os projetos para levantamento das quantidades reais a serem consumidas na obra, porém, já existe o projeto básico arquitetônico, sendo assim, é possível quantificar a quantidade de pontos. A partir de orçamentos realizados determinamos as quantidades usuais de material e mão-de-obra consumidas para execução dos serviços”.

A partir dessas informações pode-se observar que o denominador do CUM utilizado é número de pontos de utilização. Também é importante notar que há previsão de consumo

de um único diâmetro de eletroduto e de uma única bitola de condutor. No caso dos pontos de tomada, a TCPO contraria o item 6.2.5.1.1 da NBR 5410 - Instalações Elétricas de Baixa Tensão (ABNT, 2004), que especifica seção mínima de 2,5 mm² para condutores de cobre utilizados em circuitos de força.

2.1.1.2 Evolução das Instalações Prediais Elétricas na Tabela de Composição de Preços para Orçamentos

Desde sua primeira edição, de 1955, a TCPO vem incorporando cada vez mais serviços, e se adaptando à evolução da indústria da construção, tanto em relação a técnicas e materiais disponíveis, quanto ao conhecimento produzido.

Avaliando a evolução das IPE na TCPO, nota-se que nas versões iniciais (PINI, 1962), as composições não consideravam perdas de materiais. Elas abrangiam a quantidade teoricamente necessária de materiais e a mão de obra necessária para executar os serviços. Os eletrodutos e condutores deveriam ser quantificados por meio de um projeto executivo.

Na 6ª edição (PINI, 1977) as composições de interruptores e tomadas contavam com uma estimativa de quantidade de condutores e eletrodutos. Na composição dos interruptores, por exemplo, constavam 3 m de eletroduto de 1/2". Ainda não eram consideradas perdas de materiais nas IPE.

A 7ª edição (PINI, 1980) trouxe como novidade custos horários de equipamentos, auxiliando os trabalhos relacionados com construção pesada e também a informação da mão de obra necessária para a instalação de tubulações incluindo suas conexões, tanto para instalações prediais hidráulicas como para instalações prediais elétricas. Contudo, as quantidades de eletrodutos e condutores são removidas das composições dos pontos de utilização. Infelizmente não há nenhuma explicação no conteúdo da TCPO do motivo dessa exclusão.

Na 10ª edição (PINI, 1996) as IPE já contam com índices de perdas. Por exemplo, na composição de eletrodutos flexíveis de PVC, já eram consideradas perdas de 10%. Ainda não constam as previsões de consumo de eletrodutos e condutores nas composições dos pontos de utilização. Essa edição inaugura a integração da TCPO com o software de orçamentação da Pini, o Volare.

Na 12ª edição do TCPO (PINI, 2003) foram introduzidas as informações de consumo e produtividade variáveis, conforme descrito anteriormente, com indicação de máximos, mínimos e medianas. Essas informações estão disponíveis para os serviços de:

- Forma;
- Armação;
- Concretagem;
- Alvenaria;

- ❑ Revestimentos de paredes internas e externas com argamassa;
- ❑ Revestimento de parede com gesso;
- ❑ Revestimentos cerâmicos de parede e piso e
- ❑ Contrapiso.

Não houve mudança nas IPE.

2.1.2 Outros índices de consumo

Embora a TCPO seja o principal manual de orçamentos do mercado brasileiro, existem outras fontes disponíveis. Cita-se a Prefeitura Municipal de São Paulo (PMSP), Prefeitura do Rio de Janeiro (PRJ), SINAPI, Fundação para o Desenvolvimento da Educação (FDE) e Companhia Paulista de Obras e Serviços (CPOS). Essas composições, de forma geral, exigem o levantamento de quantitativo, oferecendo informações de perdas e produtividade.

O FDE oferece o consumo unitário de forma similar a TCPO, ou seja, com valores estáticos. Cabe destacar que a Tabela de Composição da FDE é mais detalhada que a TCPO, oferecendo diversos serviços, como por exemplo, tomadas simples, conjunto de tomada simples mais interruptor, entre outros. Como a tipologia é específica, direcionada para escolas, os dados não podem ser extrapolados para o objeto em estudo neste trabalho. Os pontos de luz, por exemplo, são alimentados por condutores de $2,5 \text{ mm}^2$, segundo a Tabela de Composição da FDE, prática atípica para residências.

Todos os índices e informações das organizações citadas são disponibilizadas em suas respectivas páginas eletrônicas.

2.2 Instalações Prediais Elétricas

As IPE de uma edificação residencial, as quais são normatizadas pela NBR5410 - Instalações Elétricas de Baixa Tensão, são sistemas responsáveis por disponibilizar energia elétrica às cargas em diversos pontos de utilização: pontos de iluminação, de tomada, tanto Tomadas de Uso Geral (TUG) quanto Tomadas de Uso Específico (TUE) e de ligação direta (ABNT, 2004).

Tipicamente, as edificações residenciais são alimentadas diretamente em baixa tensão pela concessionária, ou por meio de um transformador exclusivo, embora condomínios de grande porte possam ser alimentados em alta tensão e contar com transformador próprio.

A concessionária é responsável pelas instalações (equipamentos e condutores) até o ponto de entrega e, a partir desse ponto, o investimento inicial e a manutenção das instalações correm por conta do usuário. Após o ponto de entrega, um ramal de entrada conduz a energia até a denominada entrada consumidora, na qual se localizam os equipamentos de

proteção, medição individualizada e distribuição. Dos equipamentos de medição partem linhas elétricas exclusivas para cada unidade de consumo, as quais chegam até o quadro de distribuição da unidade. O quadro de distribuição conta com dispositivos de proteção e seccionamento, além de distribuir a energia nos diversos circuitos que compõem as IPE da unidade (SOUZA; MORENO, 2001).

Embora ainda não sejam comuns, já estão disponíveis no Brasil, equipamentos de medição eletrônicos, como o *CashPower*, os quais substituem os medidores eletromagnéticos e, por permitir leituras a distância, não exigem a realização da leitura do consumo por um funcionário da distribuidora. Estes equipamentos localizam-se próximos aos quadros de distribuição, dentro da unidade consumidora, permitindo que uma única linha elétrica ligue a entrada consumidora à diversas unidades, podendo tornar a instalação do sistema mais econômica, além de oferecer outros recursos de monitoramento, medição e controle do consumo.

Após a definição dos pontos de utilização e de comando (interruptores), os quais são inseridos no projeto de arquitetura, cabe ao projetista das IPE, a especificação do quadro de distribuição e do caminhamento e bitola de eletrodutos. Esses eletrodutos ligam a caixa de distribuição aos pontos de utilização e recebem condutores, fios e cabos, dimensionados para transportar a energia de maneira eficiente (SILVA, 2006).

Após a conclusão do projeto, é possível fazer um levantamento de todos os componentes da IPE, como quadros, caixas, eletrodutos, condutores, disjuntores, tomadas, interruptores, acabamentos e outros. No caso dos eletrodutos e condutores, o levantamento consiste na medição do comprimento de cada trecho do componente. Esta tarefa é considerada dispendiosa devido ao caminhamento desses componentes no projeto, com curvas, tanto no plano das paredes quanto no plano das lajes. Essa dificuldade faz com que, muitas vezes, a determinação da quantidade de material a ser comprada só seja feita com precisão em protótipos já em obra (“apartamentos modelo”). Quanto aos demais componentes, o levantamento pode ser considerado menos complexo e consiste em uma simples contagem.

O previsão de consumo de eletrodutos e condutores, sem um projeto executivo de IPE é tarefa complexa. Como visto na Seção 2.1, o TCPO não fornece dados satisfatórios para esse prognóstico.

Diversos fatores influenciam o consumo de materiais nas IPE, como:

- ❑ Sistema estrutural, já que a alvenaria estrutural limita cortes em paredes, podendo proporcionar maiores trajetórias de eletrodutos e condutores;
- ❑ Área e pé direito, os quais influenciam na trajetória de eletrodutos e condutores;
- ❑ Quantidade de pontos de utilização;
- ❑ Sistema de aquecimento de água, já que, o aquecimento a gás e solar reduzem os pontos de utilização quando comparado com chuveiros elétricos;

- ❑ Sistema de condicionamento de ar, quando presentes, aumentam o número de pontos de utilização;
- ❑ Data do projeto, já que, cada vez mais, as residências contam com um maior número de equipamentos cada vez mais potentes;
- ❑ Número de pontos de utilização e de comando, como tomadas, pontos de luz e interruptores (simples, paralelo, intermediário);
- ❑ Padrão do empreendimento. Nesse caso, há uma expectativa de que quanto mais alto o padrão do empreendimento, maior a disponibilidade de pontos de utilização e maior a carga prevista. Essa variável é subjetiva e a influencia de outras variáveis acabam sendo imperativa na estimativa de padrão do empreendimento. São elas:
 - presença de varanda *gourmet*;
 - presença de água quente em pias e lavatórios;
 - pavimentos tipo;
 - unidades por pavimento tipo
 - presença de lavabos e de banheiros de serviço.
- ❑ Número de condicionadores de ar. Além de ajudarem a definir o padrão do empreendimento, essa variável representa o aumento de pontos de utilização com TUE, aumentando o consumo de eletrodutos e condutores;
- ❑ Número de chuveiros. Determina o número de banheiros e, em caso de aquecimento de água por chuveiro elétrico, também resultam na elevação do consumo de eletrodutos e condutores.

A determinação do consumo de materiais de diferentes projetos é então dificultada frente ao grande número de variáveis relevantes e à complexidade para determinação do consumo de materiais das IPE em fases iniciais de um empreendimento em obra.

O foco deste trabalho consiste em prognosticar a quantidade de condutores e eletrodutos que ligam o quadro de distribuição da unidade até os pontos de utilização.

2.3 Redes Neurais Artificiais

Diante da complexidade apresentada para a realização de prognósticos de consumo de materiais de IPE, vislumbrou-se a possibilidade do uso de ferramentas matemáticas mais sofisticadas do que as médias, faixas de consumo e regressões lineares.

RNA são modelos computacionais inspirados no cérebro humano e que possuem a capacidade de aquisição e manutenção do conhecimento (informações). Em outras palavras,

RNA podem ser definidas como um conjunto de unidades de processamento (neurônios artificiais) que são interligadas por um grande número de interconexões (sinapses artificiais). RNA são capazes de modelar aspectos não-lineares de variáveis de entrada, sendo usada como ferramenta para (SILVA; SPATTI; FLAUZINO, 2010):

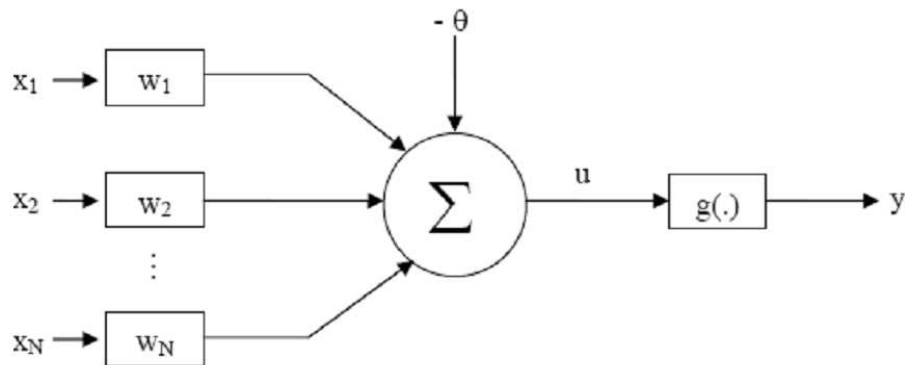
- ❑ Regressão, ou aproximação de função;
- ❑ Reconhecimento de padrões;
- ❑ Previsão;
- ❑ Otimização;
- ❑ Memórias associativas;
- ❑ Controle, entre outras.

Para este trabalho o interesse recai sobre sua utilização para problemas de regressão, no qual uma série de instâncias, representadas por cada uma das obras, apresentam uma série de características (atributos ou variáveis) e a partir delas é possível prever qual será o valor de uma característica desconhecida (classe, com valores reais e contínuos).

Pesquisas relacionadas a RNA têm mais de meio século de história e sua aplicação tem se expandido constantemente (SILVA; SPATTI; FLAUZINO, 2010).

O modelo geral do neurônio artificial pode ser ilustrado pela Figura 6.

Figura 6 – Modelo geral de neurônio artificial.



Fonte: Silva, Spatti e Flauzino (2010)

Na Figura 6 tem-se que:

- ❑ x_1, x_2, \dots, x_N : são as variáveis de entrada;
- ❑ w_1, w_2, \dots, w_N : são os pesos sinápticos obtidos durante o treinamento da rede;
- ❑ $-\Theta$: é o limiar de ativação do neurônio;
- ❑ u : é a saída do combinador linear;
- ❑ $g(\cdot)$: é a função de ativação e
- ❑ y : é o valor de saída do neurônio.

2.3.1 Arquitetura e topologia

As RNA possuem três partes, denominadas de camadas. Essas camadas podem ser divididas em (SILVA; SPATTI; FLAUZINO, 2010):

- ❑ Camada de entrada: é a camada responsável pela entrada de dados. Essas entradas devem ser normalizadas individualmente dentro dos seus respectivos valores mínimos e máximos, levando-se em consideração a variação dinâmica das funções de ativação utilizadas nos neurônios;
- ❑ Camada(s) escondida(s) ou oculta(s): camada que possui os neurônios responsáveis pela extração da informação requerida do processo e
- ❑ Camada de saída: camada também formada por neurônios e responsável pela produção do valor final do requerido processo.

Existem diversas arquiteturas de RNA, sendo estas, responsáveis por diferenciar as redes nas formas de ligação entre os neurônios e no direcionamento da informação entre eles. Dentre as arquiteturas mais conhecidas, pode-se citar (SILVA; SPATTI; FLAUZINO, 2010):

- ❑ Redes *Feed Forward* de camada simples: neste tipo de rede, tem-se uma camada de entrada e uma única camada de neurônios que é a própria camada de saída. São utilizadas em reconhecimento de padrões e a informação é sempre unidirecional sendo proveniente da camada de entrada em direção à camada de saída;
- ❑ Redes *Perceptron* Multi Camadas (PMC): esse tipo de rede difere da anterior pela presença de uma ou mais camadas escondidas de neurônios. São comumente usadas em aproximadores de funções, reconhecimento de padrões e identificação e controle. A informação é sempre unidirecional sendo proveniente da camada de entrada em direção à camada de saída;
- ❑ Redes Recorrentes: são redes que contém retroalimentação entre neurônios de camadas diferentes. As aplicações são previsão/estimação, séries temporais, otimização e sistemas dinâmicos e
- ❑ Redes com Estrutura Reticulada: consiste de uma disposição espacial de seus neurônios, normalmente formados por uma dimensão (*array*) ou duas dimensões (*grade*). São utilizadas em problemas de agrupamento (*clustering*), reconhecimento de padrões e grafos.

Ainda segundo Silva, Spatti e Flauzino (2010), diversas topologias são encontradas para cada uma das arquiteturas. Quanto às topologias, as diferenças entre elas residem no número de neurônios utilizados ou na utilização de diferentes funções de ativação. Cada uma dessas arquiteturas é idealizada para um objetivo específico e sua topologia depende do problema em questão.

2.3.2 Treinamento

O processo de treinamento é outro fator que diferencia as inúmeras RNA. O processo de treinamento constitui-se da aplicação de um algoritmo que visa ajustar os pesos e limites de ativação da rede de forma a aproximar suas respostas dos valores desejados. Com relação aos treinamentos, podem-se citar os seguintes tipos (SILVA; SPATTI; FLAUZINO, 2010):

- ❑ Supervisionado: a rede é treinada a partir de dados que contém disponíveis as saídas desejadas para as entradas da amostra. A aplicação do treinamento supervisionado depende apenas da disponibilidade desta tabela de atributos/valores, sendo que tudo se comporta como se houvesse um “professor” ensinando para a rede qual a resposta correta para cada amostra apresentada em suas entradas;
- ❑ Não supervisionado: não há uma saída específica em relação aos estímulos de entrada. A rede auto-organiza-se em relação às particularidades do conjunto de entrada e divide as instâncias em *clusters*;

- ❑ *Offline*: O ajuste de pesos e limiares é realizado após a apresentação de todo o conjunto de treinamento e
- ❑ *Online*: O ajuste de pesos e limiares é realizado após a apresentação de cada amostra de treinamento.

Para o treinamento de uma RNA, além da definição da topologia, existem alguns parâmetros que devem ser ajustados (SILVA; SPATTI; FLAUZINO, 2010) (HALL et al., 2009):

- ❑ *Taxa de aprendizagem*: determina a intensidade do ajuste dos pesos sinápticos a cada época. Taxas muito baixas podem exigir um grande número de épocas para atingir o erro mínimo, enquanto taxas muito altas podem pular vales com erros mínimos de uma época para a próxima;
- ❑ *Decaimento*: caso utilizado, a taxa de aprendizagem passa a ser dividida pelo número da época do treinamento, fazendo com que essa taxa diminua de acordo com o avanço do treinamento. Dessa forma evita-se que a RNA saia da situação de melhor desempenho. Porém, a utilização do decaimento pode fazer com que a taxa de aprendizagem fique tão pequena que o treinamento seja interrompido ao atingir o número máximo de épocas, antes alcançar a situação de melhor desempenho (sob-ajuste);
- ❑ *Momentum*: aumenta a intensidade dos ajustes nos pesos sinápticos quando o erro ainda está distante do mínimo, acelerando o processo de treinamento e
- ❑ *Número de épocas*: determina a quantidade máxima de iterações, com os ajustes dos pesos sinápticos, realizadas pelo algoritmo de treinamento. Um número baixo de épocas máximas pode impedir que o treinamento chegue ao melhor desempenho, enquanto um número grande pode levar ao sobreajuste.

Embora existam algoritmos para a busca dos melhores parâmetros de treinamento, número de camadas escondidas, bem como da quantidade de neurônios dessas camadas, a busca por parâmetros ótimos é usualmente feita de forma empírica (WITTEN; FRANK, 2005). Testa-se diversas configurações e avalia-se seus resultados. Essa é apontada como uma das ressalvas das RNA, quando comparada a outras ferramentas que apresentam menos parâmetros a serem testados (BASHIRI; GERANMAYEH, 2011).

O desempenho de uma RNA com um conjunto de dados pode ser avaliado por meio de uma série de indicadores. Os que interessam para o presente trabalho são (WITTEN; FRANK, 2005):

- Raiz do erro quadrático médio (Equação 1): é utilizada a raiz quadrada para facilitar a interpretação do indicador, conservando a mesma dimensionalidade da classe prevista. É útil para identificar se o modelo apresenta instâncias com erros muito grandes, já que o indicador é bastante afetado por elas;

$$RMSE = \sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}} \quad (1)$$

onde $RMSE$ é a raiz do erro quadrático médio, p é o valor previsto, a é o valor atual e n é o número de instâncias testadas.

- Erro absoluto médio (Equação 2): É pouco afetado por instâncias com erros extremos, sendo menos vulnerável à *outliers*. Não leva em consideração o sinal do erro, usando o seu módulo, evitando assim que um erro positivo anule um negativo;

$$MAE = \frac{|p_1 - a_1| + \dots + |p_n - a_n|}{n} \quad (2)$$

onde MAE é a erro absoluto médio, p é o valor previsto, a é o valor atual e n é o número de instâncias testadas.

- Coeficiente de correlação (Equação 3): mede a correlação estatística entre os valores atuais e os previstos. É um indicador diferente dos demais por não ser influenciado pela magnitude do erro, e

$$CC = \frac{S_{PA}}{\sqrt{S_P \cdot S_A}}, \text{ onde}$$

$$S_{PA} = \frac{\sum_i (p_i - \bar{p}) \cdot (a_i - \bar{a})}{n - 1},$$

$$S_P = \frac{\sum_i (p_i - \bar{p})^2}{n - 1}, \text{ e}$$

$$S_A = \frac{\sum_i (a_i - \bar{a})^2}{n - 1} \quad (3)$$

e onde CC é coeficiente de variação, p é o valor previsto, \bar{p} é a média dos valores previstos, a é o valor atual, \bar{a} é a média dos valores atuais e n é o número de instâncias testadas.

- Erro absoluto médio em relação a média dos valores atuais (Equação 4): por ser um indicador relativo, permite a comparação entre os modelos de prognóstico que trabalham com valores de ordem de grandeza diferentes. Também é um indicador comum nos trabalhos de consumo de materiais, permitindo a comparação com os demais trabalhos da área.

$$RE = \frac{MAE}{\bar{a}} \quad (4)$$

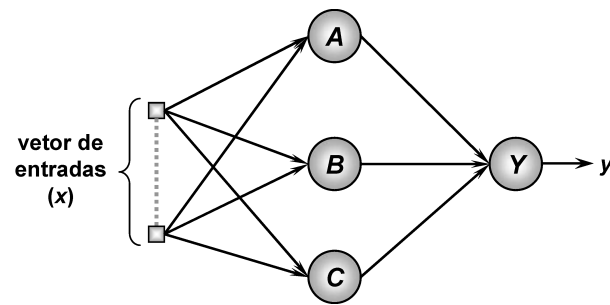
onde RE é o erro absoluto médio em relação a média dos valores atuais, MAE é o erro absoluto médio e \bar{a} é a média dos valores atuais.

2.3.3 Aproximador universal de funções

Para mapear o consumo de componentes de IPE, possibilitando então o prognóstico, foram utilizadas redes *Perceptron* Multicamadas (PMC), como aproximadores universais de funções (*fitting tool*). O aprendizado do PMC por meio de exemplos (entradas e saídas conhecidas previamente) torna-o uma opção de mapeamento de processos quando ferramentas convencionais não apresentam resultados satisfatórios ou então quando só apresentam resultados satisfatórios em condições demasiadamente restritas de operação (SILVA; SPATTI; FLAUZINO, 2010).

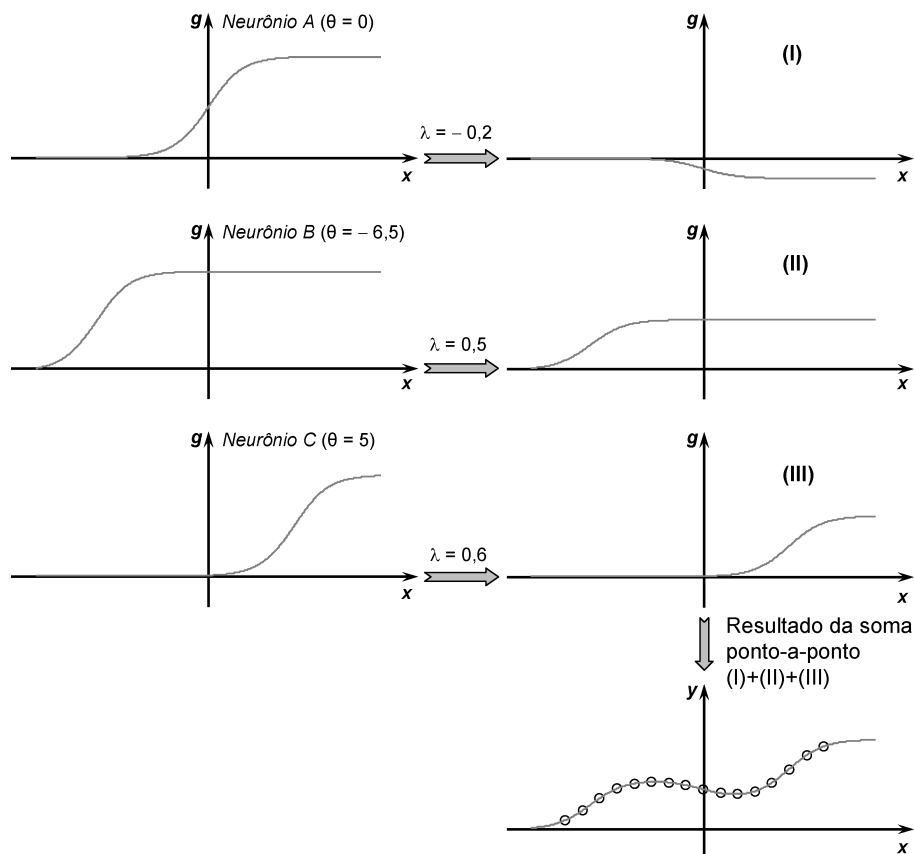
Mediante o teorema da aproximação universal (KOLMOGOROV, 1957) demonstra-se que um PMC com uma única camada escondida, como o apresentado na Figura 7, utilizando função de ativação sigmoideal (capaz de captar a não linearidade da função a ser aproximada) e função de ativação linear na camada de saída (capaz de realizar a combinação das funções sigmoideais da camada anterior), é capaz de aproximar uniformemente qualquer função contínua, dado um conjunto de treinamento significativo para representar a função (CYBENKO, 1989). A Figura 8 ilustra o funcionamento de um PMC, como aquele apresentado na Figura 7, aplicado à aproximação de funções, onde os limiares $\{\theta\}$ dos neurônios da camada escondida transladam a função de ativação em seus domínios e os pesos $\{\lambda\}$ da camada de saída fazem o escalonamento desses funções (SILVA; SPATTI; FLAUZINO, 2010).

Figura 7 – Exemplo de topologia utilizada para problemas de aproximação de função.



Fonte: Silva, Spatti e Flauzino (2010).

Figura 8 – Ilustração do funcionamento de um PMC, como aquele apresentado na Figura 7, aplicado a aproximação de funções, com a superposição de funções logísticas.



Fonte: Silva, Spatti e Flauzino (2010).

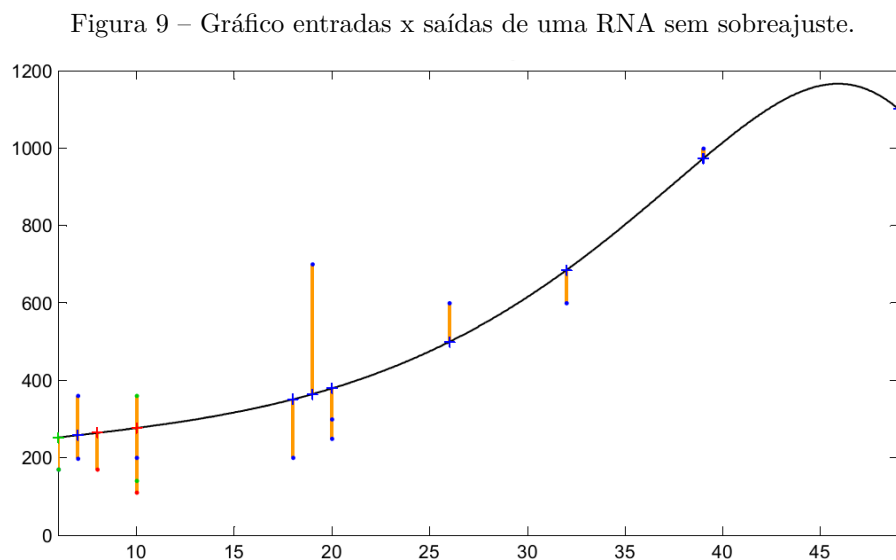
2.3.4 Sobreajuste

Embora possa surgir a ideia de aumentar indefinidamente o número de neurônios na camada escondida de um PMC para melhorar seu desempenho, sabe-se que isto pode levar ao sobreajuste. O sobreajuste ocorre quando o modelo passa a captar inclusive os ruídos dos dados, aproximando o erro no conjunto de treinamento a zero, porém,

apresentando erros elevados quando são apresentados novos dados. Esse é um problema típico de Sistemas Inteligentes (RYNKIEWICZ, 2012).

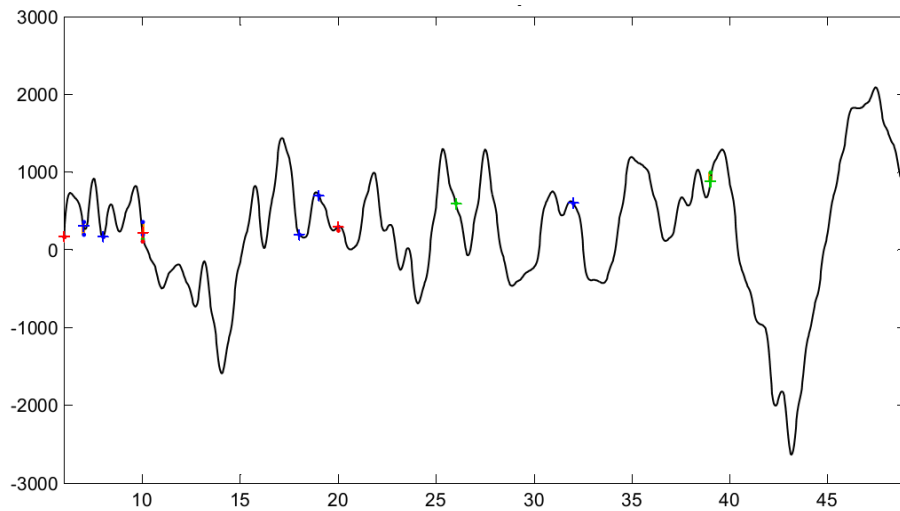
É possível treinar e testar a RNA em todo o conjunto de dados, porém, os índices de desempenho obtidos, não terão validade quando testados dados que não foram utilizados no treinamento. O modelo poderá ter um desempenho excelente trabalhando apenas com aqueles dados que utilizou para treinar e apresentar grandes erros quando trabalhando com dados inéditos.

A Figura 9 apresenta uma função gerada por RNA em que não houve sobreajuste, enquanto a Figura 10 apresenta uma outra função, gerada com os mesmos dados, porém utilizando uma quantidade muito maior de neurônios na camada escondida, resultando em sobreajuste. Nota-se que o erro no subconjunto de treinamento é praticamente nulo, porém, a função não tem representatividade da realidade.



Fonte: elaborado pelo autor.

Figura 10 – Gráfico entradas x saídas de uma RNA com sobreajuste.



Fonte: elaborado pelo autor.

Para prevenir o sobreajuste, o treinamento, etapa em que se define os valores dos pesos, deve ser feito em dados distintos daqueles nos quais a RNA será testada. Utilizar os mesmos dados no treinamento e no teste, dará uma falsa impressão de bom desempenho, que não será atingido quando o modelo for testado com dados inéditos (WITTEN; FRANK, 2005).

A análise de desempenho dos modelos de regressão realizados por RNA são, comumente, feitas da seguinte maneira: Seleciona-se aleatoriamente uma parcela dos dados disponíveis e utiliza-os como subconjunto de treinamento. A outra parcela é utilizada como subconjunto de teste, de forma a verificar qual rede possui melhor poder de generalização do problema (em geral aproximadamente 80% do banco de dados é usado para o conjunto de treinamento das topologias candidatas e 20% destes dados serão usado para o teste de desempenho). Existem duas desvantagens nessa técnica, a primeira delas é que ela reduz a quantidade de dados disponíveis para o treinamento da RNA, o que é significativo quando a quantidade de dados é limitada. A segunda é que os indicadores de desempenho da rede podem mudar muito de acordo com o subconjunto de teste utilizado, ou seja, ao treinar a rede dez vezes, com subconjuntos de treinamento e teste diferentes, os indicadores de desempenho podem apresentar valores bastante distintos, o que reduz a confiabilidade desses indicadores (WITTEN; FRANK, 2005).

Outra técnica utilizada é a validação cruzada. Nesse caso, todo o banco de dados é dividido em um número N de grupos, previamente determinados. O algoritmo de treinamento é executado varias vezes e em cada uma delas, um dos grupos será excluído do banco de dados de treinamento e será utilizado como subconjunto de teste, sendo que o treinamento será realizado com os dados restantes. A cada vez que o algoritmo de treinamento é executado, são calculados os índices de desempenho. Após todos os dados terem sido utilizados tanto para treinamento quanto para teste, é realizada uma média dos

indicadores de desempenho, que será o indicador final de desempenho para aquela RNA. A desvantagem da validação cruzada é que ela é mais lenta, o que pode ser relevante quando utilizada com grande quantidade de dados. Além disso, ela exige que todo o banco de dados seja carregado simultaneamente, o que exige memória computacional suficiente para tal tarefa (WITTEN; FRANK, 2005).

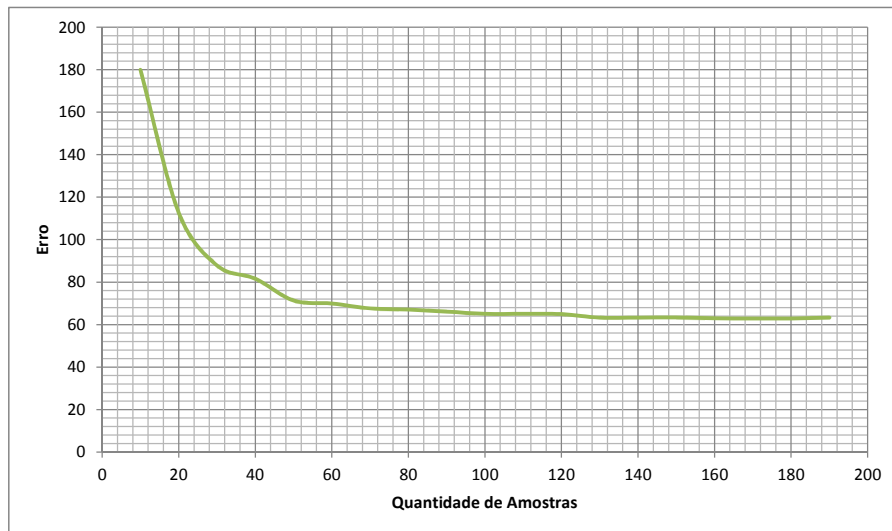
Há também o procedimento de parada antecipada (*early stopping*), que consiste em acrescentar um novo subconjunto de dados denominado de conjunto de validação. Este conjunto não é utilizado para a atualização de pesos sinápticos e limiares de ativação durante o treinamento, porém, o algoritmo de treinamento consulta o desempenho do PMC nesse subconjunto de dados e, caso o desempenho nesse subconjunto comece a reduzir de forma consistente, o processo de treinamento é interrompido. O uso da parada antecipada é útil para evitar o sobreajuste, porém, reduz a quantidade de dados utilizada para o treinamento, o que é significativo quando a quantidade de dados disponível é limitada (SUETAKE, 2012).

Neste trabalho optou-se pelo uso da validação cruzada, sem parada antecipada, para obter-se índices de desempenho mais confiáveis e para não reduzir a quantidade de dados para treinamento dos modelos.

2.3.5 Curvas de Aprendizagem

Uma das dificuldades nos problemas de regressão é saber qual a quantidade de amostras necessárias para o correto treinamento dos algoritmos, já que a obtenção desses dados é custosa. Uma forma de avaliar a quantidade necessária de amostras são as curvas de aprendizagem. As curvas de aprendizagem apresentam o desempenho do modelo de regressão em relação a quantidade de amostras disponibilizadas para seu treinamento. A Figura 11 apresenta um exemplo de curva de aprendizagem (WITTEN; FRANK, 2005).

Figura 11 – Exemplo de curva de aprendizado.



Fonte: elaborado pelo autor.

Observa-se nas curvas de aprendizagem uma região onde um pequeno aumento no número de amostras causa grande melhora no desempenho (grande inclinação do gráfico) e outra região onde o aumento do número de amostras praticamente não impacta o desempenho (assíntota da curva de aprendizagem). Uma forma de avaliar se a quantidade de amostras utilizadas para o treinamento de um modelo foi suficiente para que ele atingisse seu desempenho potencial é traçando uma curva de aprendizagem com os mesmos dados e condições de treinamento utilizadas para o treinamento daquele modelo. A curva de aprendizagem é levantada refazendo-se o treinamento do modelo, porém, com subgrupos de dados reduzidos aleatoriamente. O treinamento do modelo é refeito, por exemplo, com 10% dos dados disponíveis, com 15%, 20%, incrementando a quantidade de dados até que todo o conjunto de dados disponíveis seja utilizada. Os indicadores de desempenho desses treinamentos geram a curva de aprendizagem. O modelo terá atingido todo seu potencial de generalização quando a curva de aprendizagem atingir sua assíntota, quando isso não ocorrer, o aumento do número de amostras ainda poderia melhorar o desempenho do modelo (WITTEN; FRANK, 2005).

2.3.6 Aplicações de Redes Neurais Artificiais na área de gestão de recursos da construção civil

As RNA tem sido muito utilizadas para a estimativa de custos, apresentando resultados positivos e negativos na precisão das estimativas, porém, demonstrando maior capacidade de trabalhar com diversas variáveis quando comparada com ferramentas estatísticas convencionais (MARCHIORI, 2009).

Uma importante aplicação de RNA na área de gestão de recursos físicos é na previsão

de produtividade de mão de obra e de equipamentos. Essa aplicação foi proposta inicialmente por Moselhi, Hegazy e Fazio (1991) e desenvolvida em âmbito internacional para análise de produtividade de:

- ❑ Equipamentos de terraplenagem (KARSHENAS; FENG, 1992) (CHAO; SKIBNIEWSKI, 1994);
- ❑ Atividade de formas para fundação (PORTAS; ABOURIZK, 1997);
- ❑ Instalações de tubulações industriais e formas para concreto armado (KNOWLES, 1997);
- ❑ Formas, lançamento e acabamento de concreto (SONMEZ; ROWINGS, 1998);
- ❑ Soldagem e instalação de tubulações industriais (ABOURIZK; KNOWLES; HERMANN, 2001);
- ❑ Escavadoras, comparando a um modelo de regressão múltipla (TAM; TONG; TSE, 2002) e
- ❑ Atividades com interferência das condições ambientais locais (ABOURIZK; WALES, 1993).

Em âmbito nacional, Mori (2008) trabalhou na predição da produtividade da execução de alvenaria utilizando RNA. Suas variáveis de entrada envolvem condições climáticas, forma de contratação, características de projeto e da obra, chegando a uma precisão mediana de 70%.

2.3.7 Redes Neurais Artificiais para o prognóstico do consumo unitário de materiais

Um dos caminhos para se prognosticar o consumo de materiais levando-se em consideração a intensidade da influência de vários fatores, consiste em utilizar RNA. Tal proposição já foi objeto de estudo no âmbito da construção civil. Como exemplo, cita-se o trabalho realizado por Furusaka et al. (1999), no qual desenvolveu um modelo que utiliza RNA para prever a quantidade de concreto utilizado na construção de agências de correios, a partir de informações disponíveis nos estágios iniciais do projeto, como área construída e privativa, número de pavimentos, pé direito, comprimento e largura do empreendimento. O objetivo foi fornecer informações aos arquitetos que permitissem intervenções nessas fases iniciais com o objetivo de redução de custos. Nos testes dos modelos desenvolvidos o erro médio variou de 4,63 até 8,37% de acordo com os atributos utilizados como entrada.

As características das RNA que as tornam adequadas na predição do consumo de materiais são:

- ❑ Capacidade de treinamento e refinamento por meio da aquisição de novos dados, ou seja, o sistema é dinâmico, aumentando sua precisão com a inclusão de novos dados. Há a possibilidade inclusive de uma empresa específica trabalhar com um banco de dados com dados exclusivamente provenientes de seus projetos;
- ❑ Capacidade de processar grande número de variáveis não lineares (SILVA; SPATTI; FLAUZINO, 2010), permitindo o uso de diversas características dos projetos como variáveis de entrada;
- ❑ Possibilidade de trabalhar com valores absolutos, quando comparado com a metodologia atual que apresenta consumos unitários e intervalos de valores baseado em valores mínimos, máximos e medianas e
- ❑ Possibilidade de fazer a previsão com diferentes variáveis de entrada, com acordo com a disponibilidade, por exemplo, com variáveis disponíveis no estudo de viabilidade ou no estudo preliminar.

Mediante a aproximação funcional realizada pela RNA, o peso de cada variável no cálculo da quantidade de cada componente a ser prognosticada é definido.

Pode-se dizer que as RNA possuem como característica principal a adaptação de parâmetros internos a partir da apresentação de exemplos dos valores de entrada e de saída e, por meio de um processo de treinamento realiza a aquisição do seu conhecimento. Após a etapa de treinamento da rede, essa é capaz de operar estimando soluções que até então eram desconhecidas. Esta última habilidade ou característica da rede é chamada de habilidade de generalização (SILVA; SPATTI; FLAUZINO, 2010).

Este trabalho utilizará o conhecimento extraído por uma RNA para auxiliar na estimativa da quantidade de componentes de IPE, descobrindo influências e correlações entre diversas variáveis de um edifício residencial de múltiplos pavimentos e suas IPE, ou seja, mapear a função de consumo de componentes das IPE.

Dessa forma, uma nova proposta da estimativa de consumo de materiais é estabelecida, de forma a auxiliar na obtenção do quantitativo das IPE. As RNA utilizam cada variável introduzida no sistema e indica o consumo de forma objetiva, e não por meio de uma faixa como nos sistemas de prognósticos atuais. Além disso, após o modelo ter sido obtido, a inserção de novos dados e a repetição do processo de treinamento são simples, permitindo o refinamento da estimativa.

2.4 Seleção de Atributos

A mineração de dados apresenta ferramentas que podem melhorar a qualidade dos dados utilizados nos problemas de classificação e regressão, assim como melhorar o desempenho das tarefas de classificação e regressão. Uma dessas ferramentas é a seleção de

atributos. Nos dados reais, em geral, uma grande quantidade de atributos é existente e, muitas vezes, alguns são irrelevantes ou redundantes. O desempenho de classificadores e aproximadores de função, como as RNA, pode ser melhorado com a pré-seleção dos atributos, eliminando-se aqueles irrelevantes e/ou redundantes, que podem “atrapalhar” o treinamento de uma ferramenta de regressão (WITTEN; FRANK, 2005).

A redução de atributos resulta na redução de dimensionalidade do problema e tem por consequência fornecer ao algoritmo de treinamento um banco dados menor e mais compacto, que melhora o desempenho desses algoritmos. Ao descartar variáveis, também é possível que seja facilitado o trabalho de levantamento de dados, porém, isso só é válido para dados homogêneos, já que a inserção de novas amostras em um banco de dados pode fazer com que a relevância dos atributos mude (KOHAVI; SOMMERFIELD, 1995). Por exemplo, em um banco de dados com amostras unicamente de obras com chuveiro elétrico, o atributo “sistema de aquecimento de água” vai ser irrelevante, já que ele não varia dentro das amostras, porém, quando são inseridos dados de obras com aquecedores de passagem à gás, talvez esse atributo passe a ser relevante.

É importante para a melhoria do desempenho dos prognósticos que seja feita a pré-seleção dos atributos participantes do prognóstico de cada um dos componentes das IPE. Caso se tenha profundo conhecimento dos dados que serão tratados e do problema, é possível fazer a seleção de dados manualmente, porém, com a crescente quantidade de atributos disponíveis, a tarefa de seleção manual fica cada vez mais complexa. Além da melhora do desempenho dos modelos de regressão, a seleção de atributos melhora o entendimento do problema, ajudando a identificar as características relevantes dos dados de entrada (WITTEN; FRANK, 2005).

Uma das formas de selecionar atributos é o método *wrapper*, que consiste em utilizar um modelo de regressão ou classificação com todas as possíveis combinações de atributos e selecionar aquela que apresenta o melhor desempenho. Como o teste de todas as combinações possíveis (*exhaustive search*) pode demandar muito tempo de processamento, existem métodos de busca para reduzir a quantidade de subconjuntos testados. Por exemplo, começando o teste com todos os atributos e recursivamente remover aquele que mais melhora o desempenho do modelo, até que o desempenho comece a cair (*backward elimination*). Outro exemplo é iniciar a busca com o melhor atributo e recursivamente acrescentar um a um, os atributos que mais melhoram o desempenho do modelo, até que o acréscimo de um novo atributo piore esse desempenho (*forward selection*) (KOHAVI; SOMMERFIELD, 1995).

Existem também algoritmos que avaliam atributos individualmente e independentemente do algoritmo de treinamento, classificando-os de acordo com sua capacidade de prognosticar uma classe, permitindo eliminar atributos irrelevantes e identificar os mais relevantes. O tempo de processamento desses algoritmos, que não testam os subconjuntos possíveis com o algoritmo de treinamento do modelo, é muito menor (HALL et al., 2009).

Os algoritmos para seleção de atributos para problemas de classificação, que trabalham com classes discretas, são mais abundantes que aqueles disponíveis para problemas de regressão, que trabalham com variáveis contínuas (ROBNIK-SIKONJA; KONONENKO, 1997).

Um dos algoritmos para seleção de atributos utilizado no presente trabalho foi o *Relief F*, que analisa a capacidade de um atributo prognosticar uma classe por repetidamente comparar as instâncias do atributo verificando como elas se comportam em relação a variação da classe. Um atributo é considerado bom para o prognóstico de uma classe quando suas instâncias se diferenciam para amostras de classes diferentes e se assemelham para amostras da mesma classe. Dessa forma, o *Relief F* ordena os atributos de acordo com suas capacidades de prognosticar a classe (KIRA; RENDELL, 1992) (KONONENKO, 1994) (ROBNIK-SIKONJA; KONONENKO, 1997).

Os algoritmos que avaliam os atributos individualmente apresentam como dificuldade a definição da quantidade de atributos relevantes que deve ser escolhida. Em geral, isso pode ser feito por meio de experimentação.

Também existem algoritmos de seleção de atributos que, sem utilizar o algoritmo de treinamento do modelo de regressão, também rápidos quando comparados ao método *wrapper* e que, quando associados a métodos de busca, avaliam a capacidade de subconjuntos de atributos em prognosticar uma classe, de forma a eliminar tanto atributos irrelevantes, quanto atributos redundantes (HALL et al., 2009).

O algoritmo Seleção de Subconjunto de Atributos Baseada em Correlação para Aprendizado de Máquinas (CfsSubsetEval), que também foi utilizado no presente trabalho, avalia a capacidade de subconjuntos de atributos em prognosticar uma classe, levando em conta a correlação entre esse subconjunto e a classe a ser prognosticada, além da redundância dos atributos do subconjunto, ou seja, para esse algoritmo selecionar um atributo, este deve ser altamente correlacionado com a classe a ser prognosticada e ter baixa correlação com os demais atributos do subconjunto (HALL, 1998).

Embora do CfsSubsetEval apresente os atributos do subconjunto na ordem em que eles aparecem no banco de dados, é possível utilizar uma opção de ranqueamento, que apresenta os atributos na ordem em que eles são adicionados ao subconjunto, de forma a obter o melhor mérito com aquela quantidade de atributos, o que auxilia no entendimento da importância de cada um dos atributos.

É importante salientar que os algoritmos de seleção de atributos identificam relações de correlação, ou seja, situações que ocorrem simultaneamente. Correlações podem representar verdadeiras relações de causalidade ou apenas coincidências que ocorrem em conjuntos de dados específicos. Para um correto prognóstico, as relações de causalidade são mais relevantes, pois na causalidade, uma situação implica diretamente na outra, ou seja, quando uma característica varia, a outra necessariamente também varia. Um exemplo de relação de causalidade pode ser representado pela situação que se segue: A quantidade de um determinado condutor, segundo um algoritmo de seleção de atributos,

apresenta grande correlação com a quantidade de pontos de utilização e com o número de unidades por pavimento. Sempre que há variação na quantidade de pontos de utilização, a quantidade de condutores também sofre variação, havendo então uma relação de causalidade. Porém, aumentar um pavimento, dobrando o número de unidades, em nada interfere com a quantidade de condutores de cada unidade, ou seja, não há, nesse caso, relação de causalidade. Na verdade o número de unidades por pavimento é correlacionado com o padrão do empreendimento, que por sua vez se correlaciona com a quantidade de condutores, ou seja a relação é indireta. Isso faz com que seja muito importante a análise das saídas dos selecionadores de atributos. Um especialista em instalações prediais pode identificar que um ponto de utilização a mais implica invariavelmente em mais condutores e que a quantidade de unidades por pavimento é um indicativo do padrão do empreendimento, não ligado diretamente ao consumo de condutores.

Método de Pesquisa

O presente capítulo apresenta como os conceitos e ferramentas apresentados na revisão bibliográfica são utilizados para atingir os objetivos estabelecidos, desde a coleta e análise de dados, passando pela seleção de atributos, até a obtenção de modelos e avaliação de seus desempenhos.

3.1 Coleta de dados

A coleta de dados referentes às IPE e à outras características dos projetos, realizada para se atingir os objetivos deste trabalho, foi feita em construtoras, projetistas e instaladores. Esses agentes realizam o levantamento de quantitativo teórico de materiais a partir de projetos executivos de IPE visando a elaboração de orçamentos. As construtoras consultadas contratam o levantamento desse quantitativo do próprio projetista de IPE, ou de um profissional terceirizado.

Os quantitativos são constituídos pela quantidade teoricamente necessária de materiais, serviços e componentes necessárias para realização de uma certa parte do empreendimento, levantadas a partir dos projetos. Em geral, o edifício é dividido pelos seus pavimentos: sub solos, térreo, ático, cobertura e os pavimentos tipo. Os pavimentos tipo ainda podem ser considerados por inteiro, ou serem subdivididos em apartamentos, halls e escadarias. Esses quantitativos informam em tabelas, por exemplo, quantas unidades de interruptores simples são necessárias para cada pavimento tipo.

Para a coleta de dados, foi realizada inicialmente uma entrevista não estruturada com cada um dos agentes para explicação dos objetivos do trabalho e identificação de particularidades do levantamento do quantitativo, como, por exemplo, arredondamentos e consideração de perdas. A opção pela entrevista não estruturada foi feita para que cada um dos envolvidos tivesse mais liberdade para contribuir com o trabalho por meio de sugestões, dúvidas e questionamentos. Foram utilizados apenas os quantitativos referentes ao apartamento tipo, mais especificamente, os trechos que vão da caixa de distribuição até os pontos de utilização.

As variáveis utilizadas dos quantitativos são apresentadas a seguir:

- Construtora;
- Se o levantamento considera o hall junto com o apartamento tipo ou não;
- Número de caixas com tomadas. Nesse caso, mesmo que uma caixa contenha mais de uma tomada, ela é contada apenas uma vez, pois o mesmo condutor alimenta todas as tomadas da caixa;
- Número de interruptores simples, intermediários e paralelos. Nesse caso, cada interruptor é contado, mesmo que estejam na mesma caixa, pois eles são ligados a condutores distintos;
- Número de pontos de luz, que são levantados com o número de caixas octogonais, ou de fundo móvel 4x4x2;
- Quantidade de eletrodutos de cada diâmetro, que serão um dos itens a serem prognosticados no trabalho e
- Quantidade de condutores de cada bitola encontrada no levantamento de dados (1.5, 2.5, 4 e 6 mm^2).

Outras variáveis foram levantadas nas entrevistas com os agentes e também em materiais publicitários e são apresentadas a seguir:

- Projetista das instalações elétricas;
- Cidade onde o empreendimento será implantado;
- Sistema estrutural;
- Sistema de aquecimento de água;
- Sistema de condicionamento de ar;
- Se a cozinha possui água quente;
- Se os lavatórios possuem água quente;
- Ano do início do projeto;
- Área do apartamento tipo;
- Número de pavimentos tipo;
- Unidades por pavimento tipo;
- Número de dormitórios;

- Se possui varanda gourmet;
- Se possui banheiro de empregados;
- Se possui lavabo;
- Número de condicionadores de ar;
- Número de chuveiros e
- Pé direito do apartamento tipo.

Todas as variáveis levantadas são conhecidas na fase de estudo preliminar, quando há uma planta com indicação dos pontos de utilização das IPE. Durante a fase de estudo de viabilidade, não existe, necessariamente, uma planta com esses pontos, portanto, não foram considerados disponíveis para uso como entrada nos modelos de prognóstico na fase de estudo de viabilidade, dados referentes aos:

- Número de caixas com tomadas;
- Número de interruptores simples, intermediários e paralelos e
- Número de pontos de luz.

Todos os dados coletados foram inseridos em uma planilha eletrônica para a formação de um banco de dados, com cada obra constituindo uma instância e cada atributo uma característica das obras. Os casos de valores desconhecidos são representados por um ponto de interrogação (?). Esses casos ocorrem devido a indisponibilidade de dados sobre alguma característica das obras.

3.2 Análise preliminar

Análises preliminares foram feitas com base nas informações obtidas nas entrevistas, no conhecimento das IPE e nas observações dos dados durante o trabalho.

As análises preliminares são válidas para todos os modelos. Ou seja, realizam processamentos globais em todo o banco de dados. As análises visam identificar atributos que não devem ser utilizados pelos modelos, combinar dois ou mais atributos para criar um novo, ressaltar características das IPE estudadas ou mesmo dos levantamentos de quantitativos.

Análises específicas para cada um dos modelos de prognóstico serão apresentadas separadamente, na apresentação dos resultados de cada modelo, já que algumas considerações importantes para um modelo podem ser irrelevantes ou inválidas para outro.

3.3 Seleção de Atributos

Para cada um dos grupos de componentes estudados (eletrodutos, condutores de $1,5 \text{ mm}^2$, condutores de $2,5 \text{ mm}^2$ e condutores de 4 e 6 mm^2) uma preparação dos dados foi feita. Foram removidas as instâncias em que a quantidade do componente em estudo não era conhecida.

As quantidades de eletrodutos e condutores, além daquele que o modelo objetiva prognosticar, também são excluídas. Ou seja, para o modelo de prognóstico de condutores de classe $1,5 \text{ mm}^2$, são removidos os atributos referentes as quantidades de condutores de classe $2,5 \text{ mm}^2$, 4 e 6 mm^2 e a quantidade de eletrodutos.

Para a seleção dos atributos, foram utilizadas duas técnicas distintas. Para a seleção do melhor subconjunto de atributos para o prognóstico de cada classe, foi utilizado o algoritmo *CfsSubsetEval* com o método de busca *GreedyStepwise*, ambos com configurações padrão. Um novo banco de dados é elaborado com o conjunto de atributos selecionados e utilizado para a elaboração dos modelos.

Os atributos selecionados pelo *CfsSubsetEval* são apresentados na mesma ordem em que eles aparecem no banco de dados, por isso, com o objetivo de entender melhor as saídas do seletor de atributos, ele também foi utilizado com a opção *generateRanking* ativada no método de busca. Com essa opção o algoritmo mostra a correlação obtida para cada subconjunto de atributos, à medida que esses subconjuntos são construídos, ou seja, à medida que atributos são acrescentados ao subconjunto. Dessa forma, eles são ordenados de acordo com sua habilidade de melhorar ou piorar a capacidade do subconjunto de atributos em prever a classe.

O uso de um conjunto de atributos como entrada dos modelos de prognóstico é importante para o entendimento da relação entre o consumo de materiais e os atributos individualmente. Espera-se que, com esse conjunto ótimo de atributos, o modelo de prognóstico apresente melhor desempenho. Já o uso de um modelo com um único atributo de entrada (unidimensional), por outro lado, pode ser bastante simples de ser utilizado com ferramentas como a regressão linear, além de permitir a plotagem de gráficos, que facilitam ainda mais a tarefa do prognóstico. Dessa forma, foram elaborados também, modelos de regressão simplificados utilizando apenas um atributo como entrada, ou seja, unidimensionais.

É importante notar que, como os atributos são correlacionados, não é possível manter constante todos os atributos, com exceção de um, para plotagem de gráficos de prognóstico. Por exemplo, não é possível manter constante o atributo área e variar o número de tomadas, pois esses dois atributos são correlacionados. Não há um apartamento com área grande e número de tomadas pequenas. O exposto ratifica a importância dos modelos unidimensionais.

Para a seleção do melhor atributo individual foi utilizado o algoritmo *Relief F* com o método de busca *Ranker*, ambos com configurações padrões. A saída do algoritmo é

uma lista com todos os atributos em ordem decrescente de capacidade de prognosticar a classe. A seleção dos atributos para os modelos simplificados unidimensionais leva em conta a saída do *Relief F* e também uma análise da relação de causalidade do atributo com a classe a ser prognosticada, ou seja, se a alteração do valor do atributo interfere diretamente no valor da classe.

O tempo de processamento dos algoritmos utilizados para seleção de atributos é baixo se comparado ao método *wrapper* (poucos segundos no total contra dezenas de minutos para cada topologia candidata). Isso já é relevante para bancos de dados pequeno, como os utilizados, e mais ainda quando o trabalho envolve grande quantidade de dados.

Os dois algoritmos foram utilizados no software Weka (HALL et al., 2009). Existem diversos aplicativos voltados a seleção de atributos e também a modelos de regressão. A escolha do Weka deve-se por sua distribuição livre e pelo fato de concentrar ferramentas para tratamento dos dados, seleção de atributos, modelos de regressão, visualização dos dados, entre outros, tudo em um mesmo aplicativo, facilitando a operação quando comparado com o uso de aplicativos específicos para cada tarefa.

3.4 Modelos de Prognóstico

Após a seleção de atributos, as instâncias que apresentaram valores não conhecidos em mais da metade dos atributos foram removidas do banco de dados. Embora todos os algoritmos de regressão utilizados sejam capazes de trabalhar com valores não conhecidos, atributos com muitos valores não conhecidos podem piorar o desempenho do modelo. Durante o treinamento, o algoritmo de regressão considera a média (para variáveis numéricas) ou a moda (para variáveis discretas) dos atributos no lugar dos valores desconhecidos. Porém esses valores são indeterminados e caso o valor da classe seja um extremo para essa instância de valores desconhecidos, o treinamento será prejudicado.

Foram feitos modelos de regressão utilizando regressão linear e RNA, em ambos os casos no Weka. Além disso, foram realizados o pronóstico utilizando as informações da TCPO.

O algoritmo utilizado no Weka para regressão linear é o *LinearRegression* com configurações padrões e sem a opção de utilização de seletores de atributos, pois esse processamento já foi previamente realizado. Os atributos nominais são convertidos automaticamente para binários.

Para a RNA é utilizado o algoritmo *MultilayerPerceptron*, que é um PMC configurado para aproximação de funções, como demonstrado por Cybenko (1989) e apresentado na Subseção 2.3.3. A quantidade de camadas escondidas, a taxa de aprendizado, o *momentum* e a quantidade de iterações são escolhidas comparando-se a média do erro quadrático e do erro absoluto da validação cruzada com dez partições de cada topologia candidata. As demais configurações utilizadas são padrões. Esse algoritmo já normaliza os dados,

portanto, não é necessário realizar tal operação previamente. Os atributos nominais são convertidos automaticamente para numéricos.

Em ambos os casos os valores desconhecidos são substituídos pela média do conjunto, no caso de variáveis contínuas, ou pela moda, no caso de variáveis discretas.

Utilizou-se validação cruzada com dez partições para ambos os classificadores, obtendo-se assim indicadores de desempenho mais confiáveis do que com a divisão aleatória dos dados em um subconjunto de treinamento e outro de teste.

Para o prognóstico utilizando o TCPO, foram utilizados os índices disponíveis, conforme apresentado anteriormente na Tabela 1, ou seja, a quantidade de pontos de utilização de cada amostra é multiplicada pelo índice de consumo previsto:

- 15 m de eletroduto para cada interruptor, ponto de luz e ponto de tomada;
- 33 m de condutor de 1,5 mm^2 para cada ponto de luz e para cada interruptor e
- 33 m de condutor de 2,5, de 4 ou de 6 mm^2 para cada ponto de tomada, de condicionador de ar e chuveiro elétrico.

Para a obtenção de modelos mais robustos frente a possíveis *outliers*, após um primeiro modelo obtido, as instâncias foram ordenadas de acordo com o erro absoluto da validação cruzada. Dez por cento das instâncias, aquelas com maior erro, são eliminadas do banco de dados e o treinamento dos modelos é refeito. (WITTEN; FRANK, 2005)

Como indicadores de desempenho dos modelos foram utilizados a raiz do erro quadrático médio, o erro absoluto médio, o coeficiente de variação e a relação entre o erro absoluto médio e a média dos valores utilizados para elaboração dos modelos.

Para cada um dos componentes prognosticados foi elaborada uma tabela com os indicadores de desempenho de cada modelo, além de indicação de quais instâncias foram removidas do banco de dados, tanto aquelas com excesso de atributos com valores não conhecidos, quanto aquelas com os maiores erros absolutos durante a validação cruzada.

Para possibilitar o melhor entendimento do efeito de algumas ferramentas utilizadas no trabalho, a Tabela 16 - **Análise de desempenho dos modelos de prognóstico da quantidade de eletrodutos de $\frac{1}{2}$ e $\frac{3}{4}$** , apresenta, além dos resultados finais, alguns resultados intermediários:

- Com e sem a seleção de atributos e
- Com e sem a remoção das instâncias com maiores erros absolutos.

Os relatórios de saída do Weka para cada um dos modelos estão disponíveis nos Apêndices.

3.4.1 Curvas de aprendizagem

Para avaliar a quantidade de amostras utilizadas no treinamento dos modelos, foi levantada uma curva de aprendizagem para cada um deles. A avaliação foi feita observando se a curva de aprendizagem atingiu sua assíntota de forma clara, ou seja, se o modelo atingiu seu potencial máximo de desempenho. Neste ponto, o aumento do número de amostras não melhora mais o desempenho do modelo. Caso a curva não atinja a assíntota, o número de amostras utilizadas não foi suficiente para atingir o potencial de desempenho máximo do modelo e, nesse caso, o aumento do número de amostras melhoraria o desempenho do prognóstico.

3.4.2 Gráficos dos Modelos Simplificados

Para os modelos unidimensionais foram levantados gráficos que incluem os prognósticos feitos com regressão linear e com RNA.

Embora os modelos unidimensionais sejam simplificados para o problema estudado, já que usam um único atributo para prever uma classe que varia de acordo com diversos atributos, estes possuem extrema simplicidade de interpretação. Possibilitam também uma fácil visualização do comportamento do consumo com determinados atributos de entrada.

3.4.3 Teste em Dados Inéditos

Como forma de validar os modelos desenvolvidos, estes foram aplicados a dados inéditos, ou seja, que não participaram de seu treinamento ou de seus testes. O resultado dos prognósticos obtidos por meio dos modelos é comparado com o consumo teórico dos componentes estudados. Como indicadores de desempenho foram utilizados o erro absoluto e o erro relativo.

Desenvolvimento dos Modelos

Neste capítulo o método de pesquisa apresentado anteriormente é aplicado para a obtenção dos modelos de prognóstico e para sua avaliação.

4.1 Dados coletados

Foram coletados dados teóricos referentes à 59 empreendimentos, os quais são apresentados nas Tabela 2 à Tabela 7, onde cada linha (instância) representa uma obra e as colunas (atributos) representam características dessas obras. Nesse conjunto encontram-se obras de 12 construtoras, 6 projetistas e 7 cidades diferentes.

Os nomes das obras, construtoras e dos projetistas foram omitidos por solicitação dos envolvidos na coleta de dados.

Tabela 2 – Banco de dados. Obras 1 a 30. Parte 1 de 3.

Obra	Construtora	Projetista	Cidade	Hall	Sistema Estrutural	Sistema de Aquecimento de Água	Sistema de Condicionamento de Ar	Água Quente na Cozinha
1	A	A	Sertãozinho	Sim	Concreto Armado	Passagem à gás	Split	Não
2	A	E	Jaboticabal	Sim	Concreto Armado	Chuveiro Elétrico	Janela	Não
3	A	A	Jaboticabal	Não	Concreto Armado	Passagem à gás	Split	Não
4	A	A	Jaboticabal	Não	Concreto Armado	Passagem à gás	Split	Não
5	A	A	Ribeirão Preto	Não	Concreto Armado	Chuveiro Elétrico	Janela	Não
6	A	B	Ribeirão Preto	Sim	Concreto Armado	Passagem à gás	Split	Sim
7	A	B	Ribeirão Preto	Sim	Concreto Armado	Passagem à gás	Split	Sim
8	A	C	Ribeirão Preto	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não
9	A	C	Ribeirão Preto	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não
10	A	C	Ribeirão Preto	Sim	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não
11	A	C	Ribeirão Preto	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não
12	A	C	Ribeirão Preto	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não
13	A	C	Jaboticabal	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não
14	A	A	Ribeirão Preto	Não	Concreto Armado	Passagem à gás	Split	Sim
15	D	D	São Paulo	Sim	Concreto Armado	Aquecimento solar	Split	Sim
16	D	F	Diadema	Sim	Concreto Armado	Aquecimento solar	Split	Sim
17	D	D	Diadema	Sim	Alvenaria Estrutural	Chuveiro Elétrico	Ausente	Não
18	D	F	São Paulo	Sim	Concreto Armado	Passagem à gás	Split	Sim
19	D	D	Diadema	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não
20	D	D	São Bernardo do Campo	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não
21	D	F	São Paulo	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não
22	D	D	São Paulo	Sim	Concreto Armado	Passagem à gás	Split	Sim
23	E	D	Osasco	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não
24	F	?	Ribeirão Preto	?	?	?	Split	?
25	G	?	Ribeirão Preto	Sim	Concreto Armado	?	Janela	Não
26	H	?	Ribeirão Preto	?	Concreto Armado	?	?	?
27	C	C	Ribeirão Preto	?	Concreto Armado	?	?	?
28	B	?	Ribeirão Preto	?	Concreto Armado	?	Janela	Não
29	A	C	Ribeirão Preto	Não	Concreto Armado	Central	Split	Não
30	I	?	Ribeirão Preto	?	?	?	Janela	Não

Fonte: elaborado pelo autor.

Tabela 3 – Banco de dados. Obras 1 a 30. Parte 2 de 3.

Obra	Água Quente nos Lavatórios	Ano	Área	Pavimentos Tipo	Unidades por Pavimento	Varanda Gourmet	Dormitórios	WC Serviço	Lavabo	Condicionadores de Ar	Chuveiros	Tomadas
1	Sim	2009	102.28	20.00	4.00	Sim	3.00	Sim	Sim	4.00	3.00	28.00
2	Não	2012	80.03	20.00	4.00	Não	3.00	Não	Sim	2.00	2.00	29.00
3	Não	2008	106.89	14.00	4.00	Não	3.00	Não	Sim	4.00	2.00	30.00
4	Não	2006	193.76	19.00	2.00	Sim	4.00	Sim	Sim	6.00	4.00	43.00
5	Não	2011	87.84	24.00	4.00	Sim	3.00	Sim	Não	2.00	2.00	34.00
6	Sim	2010	209.31	20.00	2.00	Sim	4.00	Sim	Sim	7.00	5.00	43.00
7	Sim	2010	178.07	20.00	2.00	Sim	4.00	Sim	Sim	6.00	4.00	41.00
8	Não	2011	46.50	3.00	10.00	Não	1.00	Não	Não	1.00	1.00	16.00
9	Não	2011	46.50	3.00	12.00	Não	1.00	Não	Não	1.00	1.00	16.00
10	Não	2010	48.62	3.00	8.00	Não	1.00	Não	Não	1.00	1.00	16.50
11	Não	2011	38.75	2.00	12.00	Não	1.00	Não	Não	1.00	1.00	15.00
12	Não	2011	38.75	2.00	16.00	Não	1.00	Não	Não	1.00	1.00	15.00
13	Não	2008	63.62	5.00	8.00	Não	2.00	Não	Não	2.00	1.00	20.00
14	Sim	2007	232.79	23.00	2.00	Sim	5.00	Sim	Sim	5.00	4.00	60.00
15	Sim	2013	241.00	27.00	4.00	Sim	4.00	Sim	Sim	3.00	4.00	71.00
16	Sim	2013	197.00	26.00	4.00	Sim	4.00	Sim	Sim	3.00	4.00	51.00
17	Não	2010	61.22	10.00	8.00	Não	3.00	Não	Não	0.00	1.00	21.00
18	Sim	2011	160.00	24.00	2.00	Sim	5.00	Sim	Sim	2.00	4.00	50.00
19	Não	2012	54.69	19.00	8.00	Não	3.00	Não	Não	0.00	1.00	21.00
20	Não	2012	64.06	18.00	8.00	Não	3.00	Não	Não	0.00	1.00	26.00
21	Sim	2013	41.00	16.00	8.00	Sim	3.00	Não	Não	0.00	2.00	26.00
22	Sim	2012	223.00	22.00	2.00	Sim	4.00	Sim	Sim	6.00	5.00	75.00
23	Não	2013	62.83	15.00	8.00	Não	3.00	Não	Não	0.00	1.00	23.00
24	?	?	131.50	4.00	5.00	Não	3.00	Sim	Sim	4.00	?	?
25	?	2010	75.50	9.00	4.00	Não	2.00	Não	Não	?	2.00	?
26	?	2010	195.00	22.00	4.00	Sim	4.00	Sim	Sim	?	4.00	54.00
27	?	2009	167.65	24.00	4.00	Sim	4.00	Sim	Sim	?	?	?
28	Não	2008	142.00	23.00	4.00	Sim	3.00	Sim	Sim	3.00	4.00	35.00
29	Sim	2001	177.00	24.00	2.00	Sim	4.00	Sim	Sim	5.00	4.00	23.00
30	?	?	80.84	15.00	?	?	3.00	?	?	4.00	?	23.00

Fonte: elaborado pelo autor.

Tabela 4 – Banco de dados. Obras 1 a 30. Parte 3 de 3.

Obra	Luz + Interruptores + Tomada + Condicionadores de Ar + Chuveiros Elétricos	Tomada + Condicionadores de Ar + Chuveiros Elétricos	Condutor 2,5mm ² , 4mm ² e 6mm ²	Eletróduto 1/2" e 3/4"	Condutor 4mm ² e 6mm ²	Condutor 2,5mm ²	Condutor 1,5mm ²
1	75.00	33	720.00	420	120	600	200
2	74.00	33	1100.00	475	75	1025	250
3	95.00	34	850.00	290	100	750	300
4	138.00	50	1940.00	700	140	1800	1000
5	80.00	39	1050.00	450	50	1000	700
6	120.50	51	2660.00	800	160	2500	600
7	110.00	48	2660.00	700	160	2500	600
8	34.00	18	750.00	350	100	650	360
9	34.00	18	750.00	350	100	650	360
10	33.50	18.5	285.94	116	20	266	198
11	32.00	17	375.00	150	25	350	170
12	32.00	17	375.00	150	25	350	170
13	44.00	23	700.00	350	50	650	360
14	168.00	66	2150.00	750	250	1900	1100
15	161.00	75	2217.50	467	1200	1018	0
16	121.00	55	1081.50	298	201	881	15
17	41.00	22	107.50	178	30	78	200
18	114.00	53	1658.00	495	75	1583	20
19	36.00	21	232.00	192	0	232	170
20	45.00	26	390.00	176	133	257	110
21	45.00	26	390.00	181	133	257	110
22	174.00	82	2858.00	459	483	2375	45
23	42.00	23	235.50	108	0	236	141
24	?	?	1700.00	?	300	1400	1800
25	?	?	?	100	?	?	?
26	?	?	1751.00	?	909	842	1360
27	?	?	1050.00	?	30	1020	900
28	?	39	2200.00	700	100	2100	1700
29	?	29	1475.00	537	175	1300	820
30	?	?	632.00	249	56	576	272

Fonte: elaborado pelo autor.

Tabela 5 – Banco de dados. Obras 31 a 60. Parte 1 de 3.

Obra	Construtora	Projetista	Cidade	Hall	Sistema Estrutural	Sistema de Aquecimento de Água	Sistema de Condicionamento de Ar	Água Quente na Cozinha
31	I	?	Ribeirão Preto	?	?	?	Janela	Não
32	H	?	Ribeirão Preto	?	Concreto Armado	?	Janela	?
33	C	?	Ribeirão Preto	?	Concreto Armado	?	Split	?
34	J	?	Ribeirão Preto	?	Alvenaria Estrutural	Passagem à gás	Split	Não
35	J	?	Ribeirão Preto	?	Alvenaria Estrutural	Passagem à gás	Split	Não
36	H	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	Janela	Não
37	C	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	?	?
38	H	?	Ribeirão Preto	?	Concreto Armado	Chuveiro Elétrico	Janela	Não
39	H	?	Ribeirão Preto	Sim	?	Chuveiro Elétrico	Janela	Não
40	K	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	?	?
41	H	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	Split	Não
42	H	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	Split	Não
43	H	?	Ribeirão Preto	Não	Concreto Armado	Passagem à gás	Split	Não
44	?	?	Ribeirão Preto	?	?	Chuveiro Elétrico	Janela	Não
45	F	?	Ribeirão Preto	?	?	?	?	?
46	F	?	Ribeirão Preto	?	?	?	?	?
47	H	?	Ribeirão Preto	?	Concreto Armado	?	?	?
48	H	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	Split	Não
49	H	?	Ribeirão Preto	Não	Concreto Armado	Passagem à gás	Split	Não
50	H	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	Janela	Não
51	H	?	Ribeirão Preto	?	Concreto Armado	Passagem à gás	Split	Não
52	H	?	Ribeirão Preto	Não	Concreto Armado	Chuveiro elétrico	Janela	Não
53	H	?	Ribeirão Preto	Não	Concreto Armado	Passagem à gás	Janela	Não
54	H	?	Ribeirão Preto	?	Concreto Armado	Chuveiro Elétrico	Janela	Não
55	?	?	Ribeirão Preto	?	?	?	?	?
56	A	C	Ribeirão Preto	Sim	Concreto Armado	Chuveiro Elétrico	Janela	Não
57	A	C	Ribeirão Preto	Não	Concreto Armado	Central	Split	Sim
58	A	C	Ribeirão Preto	?	Concreto Armado	Central	Janela	Não
59	A	C	Ribeirão Preto	Sim	Concreto Armado	Central	?	?

Tabela 6 – Banco de dados. Obras 31 a 60. Parte 2 de 3.

Obra	Água Quente nos Lavatórios	Ano	Área	Pavimentos Tipo	Unidades por Pavimento	Varanda Gourmet	Dormitórios	WC Serviço	Lavabo	Condicionadores de Ar	Chuveiros	Tomadas
31	?	?	71.44	15.00	?	?	2.00	?	?	3.00	?	22.00
32	?	?	98.00	15.00	4.00	Não	3.00	?	?	2.00	?	30.00
33	?	2011	302.00	28.00	2.00	Sim	4.00	Sim	Sim	?	5.00	63.00
34	Não	2008	101.10	16.00	?	Sim	4.00	Não	Não	?	3.00	?
35	Não	2008	77.96	16.00	?	Sim	3.00	Não	Não	?	?	?
36	Não	?	136.00	20.00	4.00	Sim	4.00	Sim	Sim	3.00	3.00	32.00
37	?	2008	174.57	26.00	4.00	Sim	3.00	Sim	Sim	6.00	4.00	32.00
38	Não	?	100.00	16.00	4.00	Sim	3.00	Não	Sim	2.00	2.00	29.00
39	Não	?	62.00	14.00	6.00	Não	3.00	Não	Não	1.00	2.00	27.00
40	?	?	128.00	24.00	2.00	Sim	4.00	Sim	Sim	3.00	4.00	36.00
41	Não	2009	146.00	26.00	4.00	Sim	5.00	Sim	Sim	6.00	4.00	37.00
42	Não	?	165.00	20.00	4.00	Sim	4.00	Sim	Sim	5.00	4.00	35.00
43	Não	2008	180.00	25.00	4.00	Sim	5.00	Sim	Sim	6.00	4.00	42.00
44	Não	?	45.00	3.00	?	Não	1.00	Não	Não	?	1.00	?
45	?	?	105.00	4.00	?	Não	4.00	Sim	Não	?	3.00	?
46	?	2008	117.00	4.00	?	Sim	3.00	Sim	Não	?	?	?
47	?	2011	140.00	12.00	4.00	Sim	5.00	Sim	Sim	?	4.00	?
48	Não	?	264.00	20.00	2.00	Sim	4.00	Sim	Sim	6.00	4.00	47.00
49	Não	?	188.00	26.00	2.00	Não	5.00	Sim	Sim	5.00	4.00	45.00
50	Não	?	155.00	18.00	4.00	Não	5.00	Sim	Sim	3.00	4.00	41.00
51	Não	?	238.00	22.00	2.00	Sim	5.00	Sim	Sim	6.00	4.00	59.00
52	Não	?	135.00	18.00	4.00	Não	5.00	Sim	Sim	5.00	3.00	38.00
53	Não	?	143.00	14.00	2.00	Não	5.00	Sim	Sim	4.00	3.00	38.00
54	Não	?	115.00	18.00	4.00	Não	4.00	Sim	Não	2.00	3.00	31.00
55	?	?	?	?	?	?	?	?	?	1.00	?	14.00
56	Não	2007	103.00	14.00	4.00	Não	3.00	Não	Sim	3.00	2.00	30.00
57	Sim	2002	181.00	22.00	2.00	Sim	5.00	Sim	Sim	5.00	4.00	43.00
58	Sim	2000	233.00	24.00	2.00	Sim	5.00	Sim	Sim	5.00	4.00	52.00
59	Não	1998	120.00	23.00	4.00	Não	3.00	Sim	Sim	4.00	3.00	?

Fonte: elaborado pelo autor.

Tabela 7 – Banco de dados. Obras 31 a 60. Parte 3 de 3.

Obra	Interuptores Simples	Interuptores Intermediários	Interuptores Paralelos	Pé Direito	Luz	Eletroduto 1/2"	Eletroduto 3/4"	Condutor 6mm ²	Condutor 4mm ²	Condutor 2,5mm ²	Condutor 1,5mm ²
31	11.00	0.00	4.00	?	11.00	133	116	0	58	500	265
32	9.00	0.00	10.00	?	?	?	?	60	10	750	700
33	26.00	2.00	20.00	?	?	?	?	200	0	2400	600
34	?	?	?	?	15.00	200	120	?	?	?	?
35	?	?	?	?	11.00	180	100	?	?	?	?
36	8.00	0.00	18.00	?	?	?	?	50	30	690	500
37	16.00	1.00	19.00	?	?	0	150	100	0	1600	1200
38	8.00	0.00	12.00	?	?	75	150	60	0	640	380
39	6.00	0.00	10.00	?	12.00	50	183	?	?	?	?
40	17.00	0.00	8.00	?	33.00	200	700	0	200	1500	1150
41	14.00	1.00	21.00	?	?	?	?	0	100	1300	1000
42	14.00	1.00	21.00	?	?	?	?	0	100	1300	1000
43	13.00	1.00	24.00	?	?	0	150	0	100	950	950
44	?	?	?	?	?	?	?	40	0	700	400
45	?	?	?	?	42.00	200	200	0	100	100	800
46	?	?	?	?	?	?	?	200	0	1800	100
47	?	?	?	?	?	?	?	0	100	1800	1000
48	18.00	2.00	36.00	?	32.00	0	400	100	100	2200	1700
49	12.00	3.00	25.00	?	?	?	?	100	75	1100	820
50	13.00	0.00	20.00	?	?	?	?	50	15	1000	700
51	14.00	1.00	32.00	?	?	?	?	0	300	1700	1200
52	9.00	0.00	18.00	?	?	?	?	100	50	900	550
53	10.00	1.00	17.00	?	?	?	?	100	50	850	700
54	7.00	1.00	17.00	?	?	?	?	100	100	1000	700
55	7.00	0.00	4.00	?	8.00	150	150	0	50	400	340
56	7.00	0.00	14.00	2.88	13.50	75	325	50	0	720	600
57	17.00	0.00	20.00	2.88	26.00	?	?	100	75	820	1200
58	?	?	?	2.88	25.00	?	?	40	100	1300	1500
59	?	?	?	?	?	?	?	0	130	900	1000

Fonte: elaborado pelo autor.

4.2 Análise preliminar

Nessa seção são apresentadas análises preliminares dos dados, as quais valerão para todos os modelos.

Uma observação importante sobre os dados levantados é que em todos os banheiros de serviço o chuveiro é elétrico, independentemente do tipo de sistema de aquecimento de água dos demais pontos do apartamento, portanto, isso terá influência quando for necessário considerar a quantidade de chuveiros elétricos do apartamento.

4.2.1 Eletrodutos de 1/2” e de 3/4”

Foi constatado, nas entrevistas com projetistas de IPE, a tendência de evitar o uso de eletrodutos de 1/2”, já que os usuários tendem a instalar cada vez mais equipamentos após a entrega das unidades habitacionais (como sistemas de automação, som, alarme, equipamentos mais potentes, entre outros). Dessa forma, muitas reclamações tem sido feitas devido à falta de espaço extra nos eletrodutos para as instalações adicionais. Os eletrodutos de 1/2” vem sendo substituídos pelos de 3/4” e deve-se esperar cada vez um consumo menor de eletrodutos de 1/2” e maior de 3/4”. Como essa tendência poderia interferir nos modelos de prognósticos, as quantidades de eletrodutos de 1/2” e de 3/4” foram mescladas de forma a gerar uma nova variável a ser prognosticada.

4.2.2 Condutores de 4 e 6 mm^2

De forma geral, nos projetos estudados, os circuitos de TUG utilizam condutores de 2,5 mm^2 , correspondente ao mínimo especificado pela NBR5410. No caso de ligações diretas para chuveiros, TUE para alguns condicionadores de ar, máquinas de lavar e secar roupa e outros equipamentos com alta potência, a bitola dos condutores depende muito da especificação dos equipamentos. Um empreendimento pode contar com condutores para os chuveiros elétricos de 4 mm^2 e um outro empreendimento, com condutores de 6 mm^2 , pois o chuveiro considerado no segundo projeto é mais potente que no primeiro. Para melhorar o desempenho dos modelos, os condutores com bitola maior que 2,5 mm^2 tiveram suas quantidades mescladas.

4.2.3 Ano, construtora, cidade e projetista

Visto que a função dos modelos de prognóstico é ser utilizada de forma genérica e não por empreendimentos de construtoras específicas, em determinadas cidades e projetadas por determinados profissionais, os atributos referentes a essas características (“Construtora”, “Cidade” e “Projetista”) foram eliminados do banco de dados, independentemente de terem ou não correlação com as classes a serem prognosticadas.

A metodologia proposta poderia ser aplicada de forma a gerar modelos onde esses atributos fossem considerados. Por exemplo, pode-se verificar se há relação clara entre o CUM e projetistas diversos, determinando qual o projeto mais econômico, porém, seria necessário um banco de dados específico para isso. A quantidade de dados coletados para esse trabalho é insuficiente para tal tarefa.

O atributo “Ano” foi removido do banco de dados por motivo análogo ao exposto anteriormente. Como o objetivo é fazer prognóstico, não é relevante considerar um atributo temporal como entrada dos modelos. Mesmo que haja uma correlação entre o tempo e os consumos, não há como saber se essa correlação iria se manter nos próximos anos.

4.2.4 Combinação de atributos

Algumas combinações de atributos foram realizadas, com a intenção de obter novos atributos de entrada para os modelos. Essa combinação de atributos é especialmente importante para o algoritmo de regressão linear, o qual apresenta pior desempenho quando trabalha com muitos atributos na entrada, e para os modelos simplificados, que são unidimensionais.

As combinações de atributos foram feitas manualmente, de acordo com as características físicas das IPE, buscando maior relação de causalidade com a classe a ser prognosticada. Após a combinação, os novos atributos foram incluídos no banco de dados para serem utilizados na seleção de atributos, ou seja, eles foram utilizados na elaboração de todos os modelos.

Os atributos combinados foram:

- “Interruptores_Simples” + “Interruptores_Paralelos” + “Interruptores_Intermediários”, que passou a ser nomeado “Interruptores”, e combinando o número total interruptores, é esperado que auxilie o prognóstico de condutores de $1,5 \text{ mm}^2$;
- “Luz” + “Tomada” + “Interruptores” + “Condicionadores_de_Ar” + “Chuveiros” apenas no caso do “Sistema_de_Aquecimento_de_Água” ser “Chuveiro_Elétrico”, que representam todos os pontos de comando e utilização, com a expectativa de que melhore o prognóstico dos eletrodutos;
- “Condutores de $2,5 \text{ mm}^2$, 4 mm^2 e 6 mm^2 ”, composto pela soma dos condutores com bitola maior que $1,5 \text{ mm}^2$, foi realizada para permitir o pronóstico dos condutores utilizados nos pontos de força, ver Subseção 4.5.2 e
- “Tomada” + “Condicionadores_de_Ar” + “Chuveiros” apenas no caso do “Sistema_de_Aquecimento_de_Água” ser “Chuveiro_Elétrico”, representando todos os pontos de força, para ser utilizado no prognóstico dos condutores que alimentam esses pontos.

Os atributos combinados participaram do processo de seleção de atributos de todas as classes a serem prognosticadas, independentemente delas terem sido motivadas por uma classe específica.

4.2.5 Banco de dados após análise preliminar

Após a análise preliminar, o banco de dados foi alterado de acordo com as considerações apresentadas nesta seção e pode ser observado nas Tabela 8 à Tabela 13.

Tabela 8 – Banco de dados após análise preliminar. Obras 1 a 30. Parte 1 de 3.

Obra	Hall	Sistema Estrutural	Sistema de Aquecimento de Água	Sistema de Condicionamento de Ar	Água Quente na Cozinha	Água Quente nos Lavatórios	Área	Pavimentos Tipo	Unidades por Pavimento	Varanda Gourmet
1	Sim	Concreto Armado	Passagem à gás	Split	Não	Sim	102.28	20.00	4.00	Sim
2	Sim	Concreto Armado	Chuveiro Elétrico	Janela	Não	Não	80.03	20.00	4.00	Não
3	Não	Concreto Armado	Passagem à gás	Split	Não	Não	106.89	14.00	4.00	Não
4	Não	Concreto Armado	Passagem à gás	Split	Não	Não	193.76	19.00	2.00	Sim
5	Não	Concreto Armado	Chuveiro Elétrico	Janela	Não	Não	87.84	24.00	4.00	Sim
6	Sim	Concreto Armado	Passagem à gás	Split	Sim	Sim	209.31	20.00	2.00	Sim
7	Sim	Concreto Armado	Passagem à gás	Split	Sim	Sim	178.07	20.00	2.00	Sim
8	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não	Não	46.50	3.00	10.00	Não
9	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não	Não	46.50	3.00	12.00	Não
10	Sim	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não	Não	48.62	3.00	8.00	Não
11	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não	Não	38.75	2.00	12.00	Não
12	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não	Não	38.75	2.00	16.00	Não
13	Não	Alvenaria Estrutural	Chuveiro Elétrico	Janela	Não	Não	63.62	5.00	8.00	Não
14	Não	Concreto Armado	Passagem à gás	Split	Sim	Sim	232.79	23.00	2.00	Sim
15	Sim	Concreto Armado	Aquecimento solar	Split	Sim	Sim	241.00	27.00	4.00	Sim
16	Sim	Concreto Armado	Aquecimento solar	Split	Sim	Sim	197.00	26.00	4.00	Sim
17	Sim	Alvenaria Estrutural	Chuveiro Elétrico	Ausente	Não	Não	61.22	10.00	8.00	Não
18	Sim	Concreto Armado	Passagem à gás	Split	Sim	Sim	160.00	24.00	2.00	Sim
19	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não	Não	54.69	19.00	8.00	Não
20	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não	Não	64.06	18.00	8.00	Não
21	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não	Sim	41.00	16.00	8.00	Sim
22	Sim	Concreto Armado	Passagem à gás	Split	Sim	Sim	223.00	22.00	2.00	Sim
23	Sim	Alvenaria Estrutural	Passagem à gás	Ausente	Não	Não	62.83	15.00	8.00	Não
24	?	?	?	Split	?	?	131.50	4.00	5.00	Não
25	Sim	Concreto Armado	?	Janela	Não	?	75.50	9.00	4.00	Não
26	?	Concreto Armado	?	?	?	?	195.00	22.00	4.00	Sim
27	?	Concreto Armado	?	?	?	?	167.65	24.00	4.00	Sim
28	?	Concreto Armado	?	Janela	Não	Não	142.00	23.00	4.00	Sim
29	Não	Concreto Armado	Central	Split	Não	Sim	177.00	24.00	2.00	Sim
30	?	?	?	Janela	Não	?	80.84	15.00	?	?

Fonte: elaborado pelo autor.

Tabela 9 – Banco de dados após análise preliminar. Obras 1 a 30. Parte 2 de 3.

Obra	Dormitórios	WC Serviço	Lavabo	Condicionadores de Ar	Chuveiros	Tomadas	Interruptor	Interruptores Simples	Interruptores Intermediários	Interruptores Paralelos	Pé Direito	Luz
1	3.00	Sim	Sim	4.00	3.00	28.00	24.00	10.00	0.00	14.00	2.70	18.00
2	3.00	Não	Sim	2.00	2.00	29.00	21.00	8.00	1.00	12.00	2.70	20.00
3	3.00	Não	Sim	4.00	2.00	30.00	41.00	19.00	0.00	22.00	2.76	20.00
4	4.00	Sim	Sim	6.00	4.00	43.00	49.00	22.00	1.00	26.00	2.72	39.00
5	3.00	Sim	Não	2.00	2.00	34.00	22.00	10.00	0.00	12.00	2.73	19.00
6	4.00	Sim	Sim	7.00	5.00	43.00	38.00	19.00	0.00	19.00	2.75	31.50
7	4.00	Sim	Sim	6.00	4.00	41.00	36.00	15.00	0.00	21.00	2.75	26.00
8	1.00	Não	Não	1.00	1.00	16.00	9.00	5.00	0.00	4.00	2.65	7.00
9	1.00	Não	Não	1.00	1.00	16.00	9.00	5.00	0.00	4.00	2.65	7.00
10	1.00	Não	Não	1.00	1.00	16.50	8.00	4.00	0.00	4.00	2.70	7.00
11	1.00	Não	Não	1.00	1.00	15.00	9.00	5.00	0.00	4.00	2.70	6.00
12	1.00	Não	Não	1.00	1.00	15.00	9.00	5.00	0.00	4.00	2.70	6.00
13	2.00	Não	Não	2.00	1.00	20.00	11.00	5.00	0.00	6.00	2.65	10.00
14	5.00	Sim	Sim	5.00	4.00	60.00	53.00	15.00	2.00	36.00	2.88	49.00
15	4.00	Sim	Sim	3.00	4.00	71.00	48.00	15.00	1.00	32.00	2.70	38.00
16	4.00	Sim	Sim	3.00	4.00	51.00	37.00	23.00	0.00	14.00	2.70	29.00
17	3.00	Não	Não	0.00	1.00	21.00	9.00	9.00	0.00	0.00	2.70	10.00
18	5.00	Sim	Sim	2.00	4.00	50.00	31.00	11.00	0.00	20.00	2.70	30.00
19	3.00	Não	Não	0.00	1.00	21.00	7.00	7.00	0.00	0.00	2.70	8.00
20	3.00	Não	Não	0.00	1.00	26.00	9.00	9.00	0.00	0.00	2.70	10.00
21	3.00	Não	Não	0.00	2.00	26.00	9.00	9.00	0.00	0.00	2.70	10.00
22	4.00	Sim	Sim	6.00	5.00	75.00	47.00	17.00	0.00	30.00	2.70	45.00
23	3.00	Não	Não	0.00	1.00	23.00	9.00	9.00	0.00	0.00	2.70	10.00
24	3.00	Sim	Sim	4.00	?	?	?	?	?	?	?	?
25	2.00	Não	Não	?	2.00	?	?	?	?	?	?	?
26	4.00	Sim	Sim	?	4.00	54.00	36.00	11.00	2.00	23.00	?	?
27	4.00	Sim	Sim	?	?	?	?	?	?	?	?	?
28	3.00	Sim	Sim	3.00	4.00	35.00	29.00	12.00	1.00	16.00	?	?
29	4.00	Sim	Sim	5.00	4.00	23.00	24.00	15.00	0.00	9.00	?	?
30	3.00	?	?	4.00	?	23.00	15.00	11.00	0.00	4.00	?	11.00

Fonte: elaborado pelo autor.

Tabela 10 – Banco de dados após análise preliminar. Obras 1 a 30. Parte 3 de 3.

Obra	Luz + Interruptores + Tomada + Condicionadores de Ar + Chuveiros Elétricos	Tomada + Condicionadores de Ar + Chuveiros Elétricos	Condutor 2,5mm ² , 4mm ² e 6mm ²	Eletroduto 1/2" e 3/4"	Condutor 4mm ² e 6mm ²	Condutor 2,5mm ²	Condutor 1,5mm ²
1	75.00	33	720.00	420	120	600	200
2	74.00	33	1100.00	475	75	1025	250
3	95.00	34	850.00	290	100	750	300
4	138.00	50	1940.00	700	140	1800	1000
5	80.00	39	1050.00	450	50	1000	700
6	120.50	51	2660.00	800	160	2500	600
7	110.00	48	2660.00	700	160	2500	600
8	34.00	18	750.00	350	100	650	360
9	34.00	18	750.00	350	100	650	360
10	33.50	18.5	285.94	116	20	266	198
11	32.00	17	375.00	150	25	350	170
12	32.00	17	375.00	150	25	350	170
13	44.00	23	700.00	350	50	650	360
14	168.00	66	2150.00	750	250	1900	1100
15	161.00	75	2217.50	467	1200	1018	0
16	121.00	55	1081.50	298	201	881	15
17	41.00	22	107.50	178	30	78	200
18	114.00	53	1658.00	495	75	1583	20
19	36.00	21	232.00	192	0	232	170
20	45.00	26	390.00	176	133	257	110
21	45.00	26	390.00	181	133	257	110
22	174.00	82	2858.00	459	483	2375	45
23	42.00	23	235.50	108	0	236	141
24	?	?	1700.00	?	300	1400	1800
25	?	?	?	100	?	?	?
26	?	?	1751.00	?	909	842	1360
27	?	?	1050.00	?	30	1020	900
28	?	39	2200.00	700	100	2100	1700
29	?	29	1475.00	537	175	1300	820
30	?	?	632.00	249	56	576	272

Fonte: elaborado pelo autor.

Tabela 11 – Banco de dados após análise preliminar. Obras 31 a 60. Parte 1 de 3.

Obra	Hall	Sistema Estrutural	Sistema de Aquecimento de Água	Sistema de Condicionamento de Ar	Água Quente na Cozinha	Água Quente nos Lavatórios	Área	Pavimentos Tipo	Unidades por Pavimento	Varanda Gourmet
31	?	?	?	Janela	Não	?	71.44	15.00	?	?
32	?	Concreto Armado	?	Janela	?	?	98.00	15.00	4.00	Não
33	?	Concreto Armado	?	Split	?	?	302.00	28.00	2.00	Sim
34	?	Alvenaria Estrutural	Passagem à gás	Split	Não	Não	101.10	16.00	?	Sim
35	?	Alvenaria Estrutural	Passagem à gás	Split	Não	Não	77.96	16.00	?	Sim
36	?	Concreto Armado	Passagem à gás	Janela	Não	Não	136.00	20.00	4.00	Sim
37	?	Concreto Armado	Passagem à gás	?	?	?	174.57	26.00	4.00	Sim
38	?	Concreto Armado	Chuveiro Elétrico	Janela	Não	Não	100.00	16.00	4.00	Sim
39	Sim	?	Chuveiro Elétrico	Janela	Não	Não	62.00	14.00	6.00	Não
40	?	Concreto Armado	Passagem à gás	?	?	?	128.00	24.00	2.00	Sim
41	?	Concreto Armado	Passagem à gás	Split	Não	Não	146.00	26.00	4.00	Sim
42	?	Concreto Armado	Passagem à gás	Split	Não	Não	165.00	20.00	4.00	Sim
43	Não	Concreto Armado	Passagem à gás	Split	Não	Não	180.00	25.00	4.00	Sim
44	?	?	Chuveiro Elétrico	Janela	Não	Não	45.00	3.00	?	Não
45	?	?	?	?	?	?	105.00	4.00	?	Não
46	?	?	?	?	?	?	117.00	4.00	?	Sim
47	?	Concreto Armado	?	?	?	?	140.00	12.00	4.00	Sim
48	?	Concreto Armado	Passagem à gás	Split	Não	Não	264.00	20.00	2.00	Sim
49	Não	Concreto Armado	Passagem à gás	Split	Não	Não	188.00	26.00	2.00	Não
50	?	Concreto Armado	Passagem à gás	Janela	Não	Não	155.00	18.00	4.00	Não
51	?	Concreto Armado	Passagem à gás	Split	Não	Não	238.00	22.00	2.00	Sim
52	Não	Concreto Armado	Chuveiro elétrico	Janela	Não	Não	135.00	18.00	4.00	Não
53	Não	Concreto Armado	Passagem à gás	Janela	Não	Não	143.00	14.00	2.00	Não
54	?	Concreto Armado	Chuveiro Elétrico	Janela	Não	Não	115.00	18.00	4.00	Não
55	?	?	?	?	?	?	?	?	?	?
56	Sim	Concreto Armado	Chuveiro Elétrico	Janela	Não	Não	103.00	14.00	4.00	Não
57	Não	Concreto Armado	Central	Split	Sim	Sim	181.00	22.00	2.00	Sim
58	?	Concreto Armado	Central	Janela	Não	Sim	233.00	24.00	2.00	Sim
59	Sim	Concreto Armado	Central	?	?	Não	120.00	23.00	4.00	Não

Fonte: elaborado pelo autor.

Tabela 12 – Banco de dados após análise preliminar. Obras 31 a 60. Parte 2 de 3.

Obra	Dormitórios	WC Serviço	Lavabo	Condicionadores de Ar	Chuveiros	Tomadas	Interruptor	Interruptores Simples	Interruptores Intermediários	Interruptores Paralelos	Pé Direito	Luz
31	2.00	?	?	3.00	?	22.00	15.00	11.00	0.00	4.00	?	11.00
32	3.00	?	?	2.00	?	30.00	19.00	9.00	0.00	10.00	?	?
33	4.00	Sim	Sim	?	5.00	63.00	48.00	26.00	2.00	20.00	?	?
34	4.00	Não	Não	?	3.00	?	?	?	?	?	?	15.00
35	3.00	Não	Não	?	?	?	?	?	?	?	?	11.00
36	4.00	Sim	Sim	3.00	3.00	32.00	26.00	8.00	0.00	18.00	?	?
37	3.00	Sim	Sim	6.00	4.00	32.00	36.00	16.00	1.00	19.00	?	?
38	3.00	Não	Sim	2.00	2.00	29.00	20.00	8.00	0.00	12.00	?	?
39	3.00	Não	Não	1.00	2.00	27.00	16.00	6.00	0.00	10.00	?	12.00
40	4.00	Sim	Sim	3.00	4.00	36.00	25.00	17.00	0.00	8.00	?	33.00
41	5.00	Sim	Sim	6.00	4.00	37.00	36.00	14.00	1.00	21.00	?	?
42	4.00	Sim	Sim	5.00	4.00	35.00	36.00	14.00	1.00	21.00	?	?
43	5.00	Sim	Sim	6.00	4.00	42.00	38.00	13.00	1.00	24.00	?	?
44	1.00	Não	Não	?	1.00	?	?	?	?	?	?	?
45	4.00	Sim	Não	?	3.00	?	?	?	?	?	?	42.00
46	3.00	Sim	Não	?	?	?	?	?	?	?	?	?
47	5.00	Sim	Sim	?	4.00	?	?	?	?	?	?	?
48	4.00	Sim	Sim	6.00	4.00	47.00	56.00	18.00	2.00	36.00	?	32.00
49	5.00	Sim	Sim	5.00	4.00	45.00	40.00	12.00	3.00	25.00	?	?
50	5.00	Sim	Sim	3.00	4.00	41.00	33.00	13.00	0.00	20.00	?	?
51	5.00	Sim	Sim	6.00	4.00	59.00	47.00	14.00	1.00	32.00	?	?
52	5.00	Sim	Sim	5.00	3.00	38.00	27.00	9.00	0.00	18.00	?	?
53	5.00	Sim	Sim	4.00	3.00	38.00	28.00	10.00	1.00	17.00	?	?
54	4.00	Sim	Não	2.00	3.00	31.00	25.00	7.00	1.00	17.00	?	?
55	?	?	?	1.00	?	14.00	11.00	7.00	0.00	4.00	?	8.00
56	3.00	Não	Sim	3.00	2.00	30.00	21.00	7.00	0.00	14.00	2.88	13.50
57	5.00	Sim	Sim	5.00	4.00	43.00	37.00	17.00	0.00	20.00	2.88	26.00
58	5.00	Sim	Sim	5.00	4.00	52.00	?	?	?	?	2.88	25.00
59	3.00	Sim	Sim	4.00	3.00	?	?	?	?	?	?	?

Fonte: elaborado pelo autor.

Tabela 13 – Banco de dados após análise preliminar. Obras 31 a 60. Parte 3 de 3.

Obra	Luz + Interruptores + Tomada + Condicionadores de Ar + Chuveiros Elétricos	Tomada + Condicionadores de Ar + Chuveiros Elétricos	Condutor 2,5mm ² , 4mm ² e 6mm ²	Eletroduto 1/2" e 3/4"	Condutor 4mm ² e 6mm ²	Condutor 2,5mm ²	Condutor 1,5mm ²
31	51.00	25.00	558.00	249	58	500	265
32	?	32	820.00	?	70	750	700
33	?	?	2600.00	?	200	2400	600
34	?	?	?	320	?	?	?
35	?	?	?	280	?	?	?
36	?	36	770.00	?	80	690	500
37	?	39	1700.00	150	100	1600	1200
38	?	33	700.00	225	60	640	380
39	58.00	30.00	?	233	?	?	?
40	98.00	40.00	1700.00	900	200	1500	1150
41	?	44	1400.00	?	100	1300	1000
42	?	41	1400.00	?	100	1300	1000
43	?	49	1050.00	150	100	950	950
44	?	?	740.00	?	40	700	400
45	?	?	200.00	400	100	100	800
46	?	?	2000.00	?	200	1800	100
47	?	?	1900.00	?	100	1800	1000
48	142.00	54.00	2400.00	400	200	2200	1700
49	?	51	1275.00	?	175	1100	820
50	?	45	1065.00	?	65	1000	700
51	?	66	2000.00	?	300	1700	1200
52	?	47	1050.00	?	150	900	550
53	?	43	1000.00	?	150	850	700
54	?	37	1200.00	?	200	1000	700
55	?	?	450.00	300	50	400	340
56	69.50	35.00	770.00	400	50	720	600
57	112.00	49.00	995.00	?	175	820	1200
58	?	58	1440.00	?	140	1300	1500
59	?	?	1030.00	?	130	900	1000

Fonte: elaborado pelo autor.

4.3 Eletrodutos

4.3.1 Preparação dos dados

As instâncias sem informações de quantidade de eletroduto foram eliminadas: 24, 26, 27, 32, 33, 36, 41, 42, 44, 46, 47, 49, 50, 51, 52, 53, 54, 57, 58 e 59.

4.3.2 Seleção de atributos

A Tabela 14 e Tabela 15 apresentam as saídas dos algoritmos de seleção de atributos para o prognóstico da quantidade de eletrodutos com e sem a utilização dos pontos de utilização, respectivamente.

Tabela 14 – Resultado dos algoritmos de seleção de atributos para eletrodutos **com** informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefFAAttributeEval / Ranker
melhor suconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Água Quente na Cozinha	0.661 Luz	0.12146 Pe Direito
Água Quente nos Lavatórios	0.68 Unidades por Pavimentos	0.11847 Luz
Unidades por Pavimentos	0.687 Chuveiros	0.10598 Água Quente na Cozinha
Varanda Gourmet	0.689 Lavabo	0.07641 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos
WC Servico	0.692 WC Servico	0.04860 Interruptores Intermediários
Lavabo	0.694 Varanda Gourmet	0.04612 Água Quente nos Lavatórios
Chuveiros	0.695 Água Quente nos Lavatórios	0.02284 Área
Pe Direito	0.695 Água Quente na Cozinha	0.02039 Interuptores Paralelos
Luz	0.695 Pe Direito	0.01924 Condicionadores de Ar
	0.695 Condicionadores de Ar	0.01729 Hall
	0.694 Sistema Estrutural	0.01641 Interruptores
	0.693 Hall	0.00884 Interruptores Simples
	0.692 Sistema de Condicionamento de Ar	0.00209 WC Servico
	0.690 Interruptores Intermediários	-0.00189 Tomadas
	0.687 Sistema de Aquecimento de Água	-0.01397 Sistema de Condicionamento de Ar
	0.684 Dormitorios	-0.01489 Dormitorios
	0.670 Interruptores Simples	-0.01840 Lavabo
	0.654 Pavimentos Tipo	-0.02508 Tomadas Condicionadores Chuveiros Eletricos
	0.640 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos	-0.02568 Chuveiros
	0.630 Interruptores Paralelos	-0.02626 Pavimentos Tipo
	0.619 Interruptores	-0.03701 Unidades por Pavimentos
	0.609 Área	-0.03876 Sistema de Aquecimento de Água
	0.603 Tomadas Condicionadores Chuveiros Eletricos	-0.04624 Sistema Estrutural
	0.598 Tomadas	-0.07011 Varanda Gourmet

Fonte: elaborado pelo autor.

Tabela 15 – Resultado dos algoritmos de seleção de atributos para eletrodutos **sem** informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefFAttributeEval / Ranker
melhor subconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Água Quente na Cozinha	0.649 WC Servico	0.123472 Pé Direito
WC Servico	0.670 Chuveiros	0.082781 Água Quente na Cozinha
Lavabo	0.674 Pé Direito	0.051497 Água Quente nos Lavatórios
Chuveiros	0.676 Água Quente na Cozinha	0.022673 Área
Pé Direito	0.679 Lavabo	0.016650 Condicionadores de Ar
	0.671 Água Quente nos Lavatórios	0.007456 Sistema de Condicionamento de Ar
	0.667 Sistema Estrutural	0.003344 Hall
	0.662 Unidades por Pavimentos	-0.000717 WC Servico
	0.659 Sistema de Condicionamento de Ar	-0.004497 Dormitórios
	0.656 Varanda Gourmet	-0.015749 Lavabo
	0.650 Condicionadores de Ar	-0.023740 Pavimentos Tipo
	0.647 Dormitórios	-0.035882 Chuveiros
	0.642 Hall	-0.044940 Unidades por Pavimentos
	0.636 Sistema de Aquecimento de Água	-0.045881 Sistema Estrutural
	0.591 Área	-0.048243 Sistema de Aquecimento de Água
	0.591 Pavimentos Tipo	-0.082300 Varanda Gourmet

Fonte: elaborado pelo autor.

O fato das variáveis “Água_Quente_na_Cozinha”, “Água_Quente_nos_Lavatórios”, “Varanda Gourmet”, “WC_de_Serviço”, “Lavabo” e mesmo “Pe_Direito” apresentarem grande relevância na seleção de atributos indica que o consumo de eletrodutos varia muito de acordo com o padrão do empreendimento.

O atributo que apresenta melhor classificação no selecionador de atributos *Relief F* e que possui relação de causalidade com a quantidade de eletrodutos é o “Luz”. O atributo “Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar” foi criado pois, representa todos os pontos de utilização e todos eles dependem de eletrodutos, sendo possível, dessa forma, obter uma relação entre quantidade de eletroduto consumido por ponto de utilização com melhor relação de causalidade. Os dois atributos foram testados como entrada do modelo simplificado unidimensional.

4.3.3 Modelos de prognóstico

Os relatórios emitidos pelo Weka referentes a todos os modelos elaborados para o prognóstico de eletrodutos, nos quais os detalhes de cada modelo podem ser observados estão nos apêndices e são organizados da seguinte forma:

- Modelos com pontos de utilização, para estudo preliminar: Subseção .1.1 e Subseção .1.2;

- Modelos sem pontos de utilização, para estudo de viabilidade: Subseção .1.3 e Subseção .1.4;
- Modelos simplificados unidimensionais, com “Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar” como entrada: Subseção .1.5 e Subseção .1.6 e
- Modelos simplificados unidimensionais, com “Luz” como entrada: Subseção .1.7 e Subseção .1.8.

O desempenho dos modelos obtidos são apresentados na Tabela 16. Ainda, nesta seção, são apresentados alguns dados intermediários da metodologia de forma a demonstrar e comprovar sua eficácia, como por exemplo, o desempenho dos modelos com e sem o uso de selecionadores de atributos e antes e depois da remoção dos 10% de atributos com maiores erros.

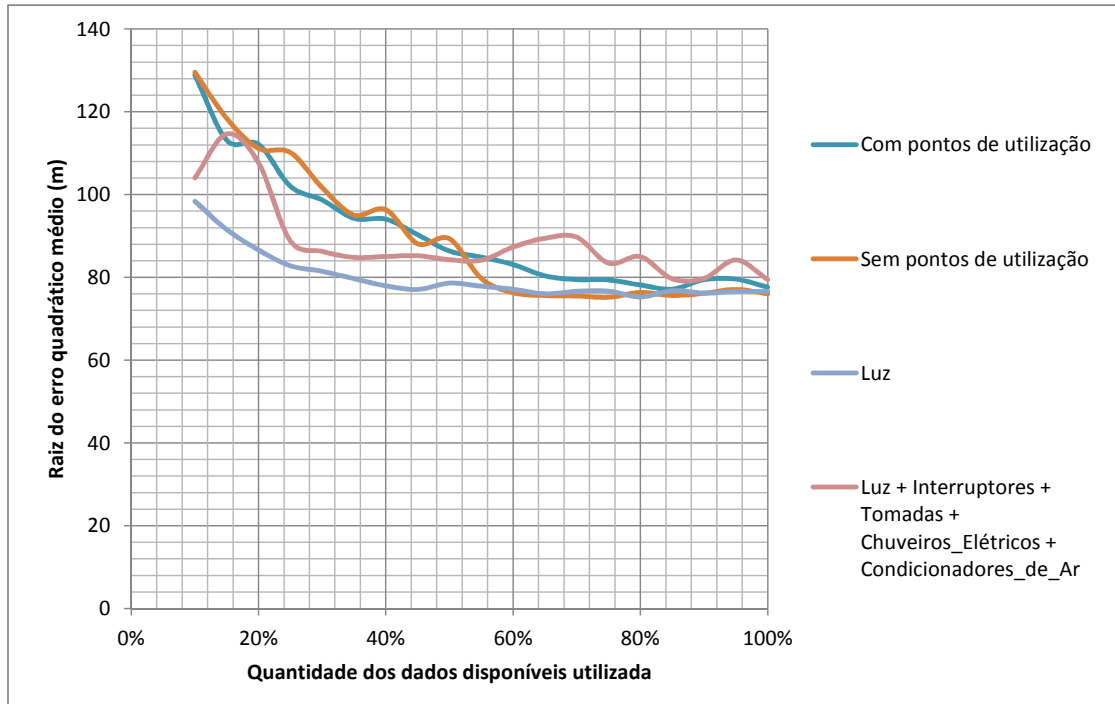
Tabela 16 – Análise de desempenho dos modelos de prognóstico da quantidade de eletrodutos de $\frac{1}{2}''$ e $\frac{3}{4}''$

		Desempenho								
Ferramenta	Com ou sem 10% de instâncias com maiores erros	Raiz do erro quadrático médio (m)	Erro absoluto médio (m)	Coefficiente de correlação	Erro absoluto médio / média	Instâncias eliminadas (+ de 50% de instâncias sem informação)	Instâncias eliminadas (10% com maior erro)			
Com pontos de utilização	TCPO	Com	913.79	746.88	0.68	197.7%	25, 26, 27, 34, 35, 36,	-		
		Sem	732.26	621.95	0.70	173.8%	37, 43, 44	14, 15, 22		
	Sem mineração de dados	Com	312.49	253.99	0.29	69.6%	45, 55	-		
		Sem	344.13	273.02	0.28	74.8%		15, 16, 22, 38		
	Com mineração de dados	Com	175.10	133.60	0.56	36.6%		-	-	
		Sem	142.87	110.80	0.65	31.5%		28, 37, 40, 43	28, 37, 40, 43	
	Sem pontos de utilização	Regressão Linear	Com	187.47	155.61	0.52	41.8%	30, 31, 55	-	
			Sem	153.08	123.26	0.68	34.6%		6, 37, 40, 43	
		Rede Neural	Com	170.50	133.57	0.60	35.9%		-	-
			Sem	128.87	109.37	0.76	29.4%		16, 37, 40, 43	16, 37, 40, 43
Regressão Linear		Com	249.44	185.55	0.47	50.0%	30, 31, 45, 55	-		
		Sem	210.26	168.37	0.51	47.5%		2, 18, 40, 43		
Rede Neural		Com	179.00	139.32	0.56	37.5%		-	-	
		Sem	194.72	156.27	0.73	41.3%		2, 18, 37, 43		
Modelo Unidimensional		Com mineração de dados	Com	160.85	130.59	0.61	35.8%	30, 31, 35, 37, 40, 45, 55	-	
			Sem	134.08	113.80	0.75	31.7%		2, 16, 28, 43	
	Regressão Linear	Com	153.80	126.79	0.65	34.8%	-		-	
		Sem	118.80	102.33	0.77	29.3%	4, 14, 16, 43			
	Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar	Com	167.53	135.75	0.63	34.7%	4, 25, 29, 30, 34, 35, 37, 38, 43, 45, 55	-		
		Sem	122.74	106.89	0.73	30.6%		6, 22, 40		
	Rede Neural	Com	161.27	128.40	0.67	32.8%		-	-	
		Sem	129.72	107.34	0.71	30.1%		6, 16, 40		
	Luz	Com	147.75	116.53	0.68	31.2%	25, 28, 29, 37, 38, 43	-		
		Sem	117.41	97.72	0.71	28.8%		6, 22, 40		
Regressão Linear	Com	142.32	113.31	0.71	30.3%	-		-		
	Sem	134.58	107.80	0.72	31.3%	6, 16, 40				

Fonte: elaborado pelo autor.

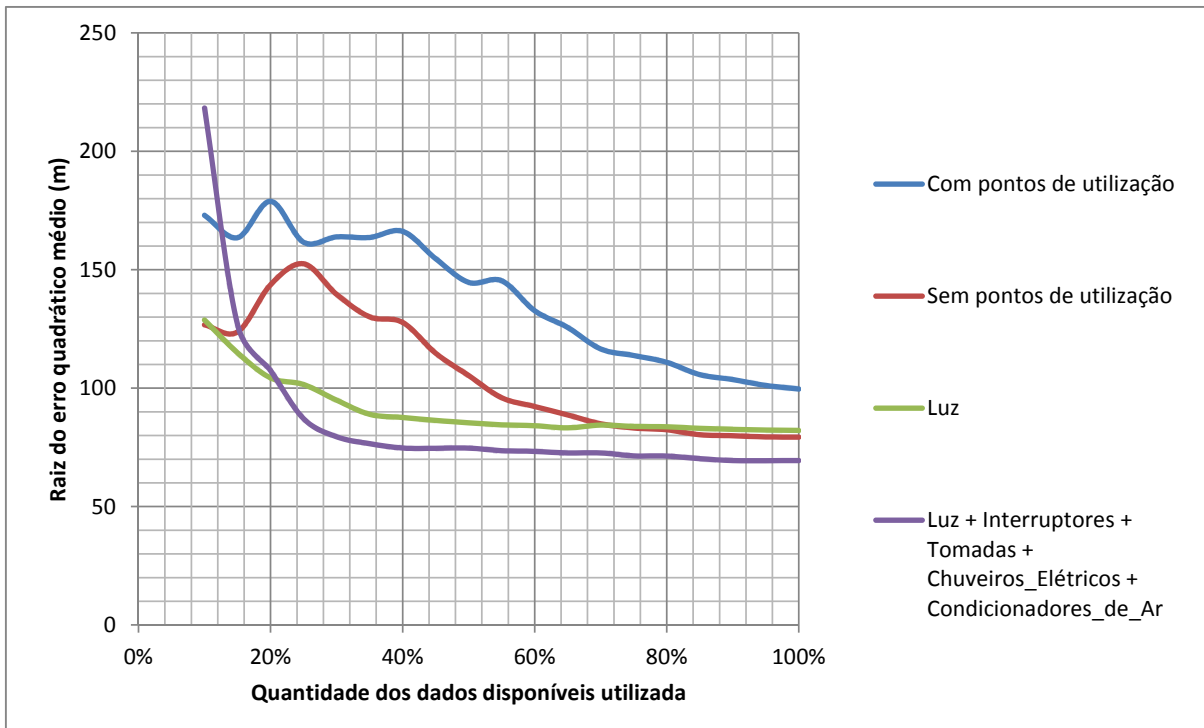
A Figura 12 e Figura 13 apresentam as curvas de aprendizagem dos modelos desenvolvidos para o prognóstico de eletrodutos.

Figura 12 – Curva de aprendizagem para os modelos que utilizam **RNA** para prognóstico de eletrodutos de $\frac{1}{2}$ " e $\frac{3}{4}$ ".



Fonte: elaborado pelo autor.

Figura 13 – Curva de aprendizagem para os modelos que utilizam **regressão linear** para prognóstico de eletrodutos de $\frac{1}{2}$ " e $\frac{3}{4}$ ".



Fonte: elaborado pelo autor.

Observando-se o desempenho dos modelos na Tabela 16 e as curvas de aprendizado nas Figura 12 e Figura 13 é possível observar que:

- ❑ O uso dos índices apresentados pela TCPO é leva a erros tão altos que pode ser considerado inadequado para o prognóstico da quantidade de eletrodutos;
- ❑ O uso do selecionador de atributos CfsSubsetEval apresentou melhor desempenho em todos os casos quando comparado à utilização de todos os atributos disponíveis. Essa melhora ocorre para as RNA mas é significativamente maior para as regressões lineares. Isso ratificando a dificuldade dos algoritmos de regressão em trabalharem com grande quantidade de atributos de entrada, principalmente no caso da regressão linear;
- ❑ A remoção de 10% dos atributos com maior erro absoluto, em geral, reduz levemente o erro absoluto médio, com um impacto mais significativo na raiz do erro quadrático médio. Porém, em dois casos houve aumento dos citados erros, o que não é impossível, já que, um novo modelo é treinado após a remoção dessas instâncias. Esta remoção visa deixar os modelos menos suscetíveis a *outliers*.
- ❑ De forma geral, tanto para regressão linear quanto para RNA, quanto maior o número de atributos, maior a necessidade de instâncias para que as curvas de apren-

dizagem dos modelos atinjam suas assíntotas, especialmente para as RNA. Na Figura 12 e Figura 13, fica claro que os modelos unidimensionais atingem suas assíntotas com menor quantidade de dados que os modelos multidimensionais;

- ❑ Nos modelos para uso na fase de estudo preliminar, aquele que utiliza RNA apresentou o melhor desempenho, com um erro absoluto médio de 109,37 m. Este valor representa 29,4% da quantidade média de eletrodutos das obras utilizadas no treinamento. Observando-se a curva de aprendizagem desse modelo, ele não atinge sua assíntota de forma clara, portanto, o aumento da quantidade de dados pode melhorar seu desempenho;
- ❑ Também no caso dos modelos para uso na fase de estudo de viabilidade, a RNA apresentou melhor desempenho com erro absoluto médio de 102,33 m. Este erro representa 29,3% da quantidade média de eletrodutos das obras utilizadas no treinamento;
- ❑ Quanto aos modelos unidimensionais, aquele que utilizou “Luz” como variável de entrada apresentou melhor desempenho se comparado com aquele que utilizou a variável “Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar”, ratificando a eficácia do selecionador de atributos *Relief F*;
- ❑ O melhor modelo unidimensional com a variável “Luz+Interruptores+Tomadas+Chuveiros Elétricos+Condicionadores de Ar” foi aquele que utilizou RNA e resultou um erro absoluto médio de 107,34 m, representando 30,1% da média de eletrodutos das obras utilizadas em seu treinamento. Embora o erro do modelo baseado em regressão linear seja um pouco menor, a média de consumo das obras utilizadas em seu treinamento também é menor. Este fato gera uma relação entre erro absoluto médio e média levemente maior. Além disso, o modelo ainda não atingiu todo seu potencial, já que sua curva de aprendizagem não chegou a assíntota de forma clara e
- ❑ A pequena diferença de desempenho nos modelos unidimensionais que utilizam RNA e regressão pode ser explicada pelo fato da regressão linear apresentar bom desempenho com baixo número de variáveis de entrada e com problemas que se aproximam da linearidade.

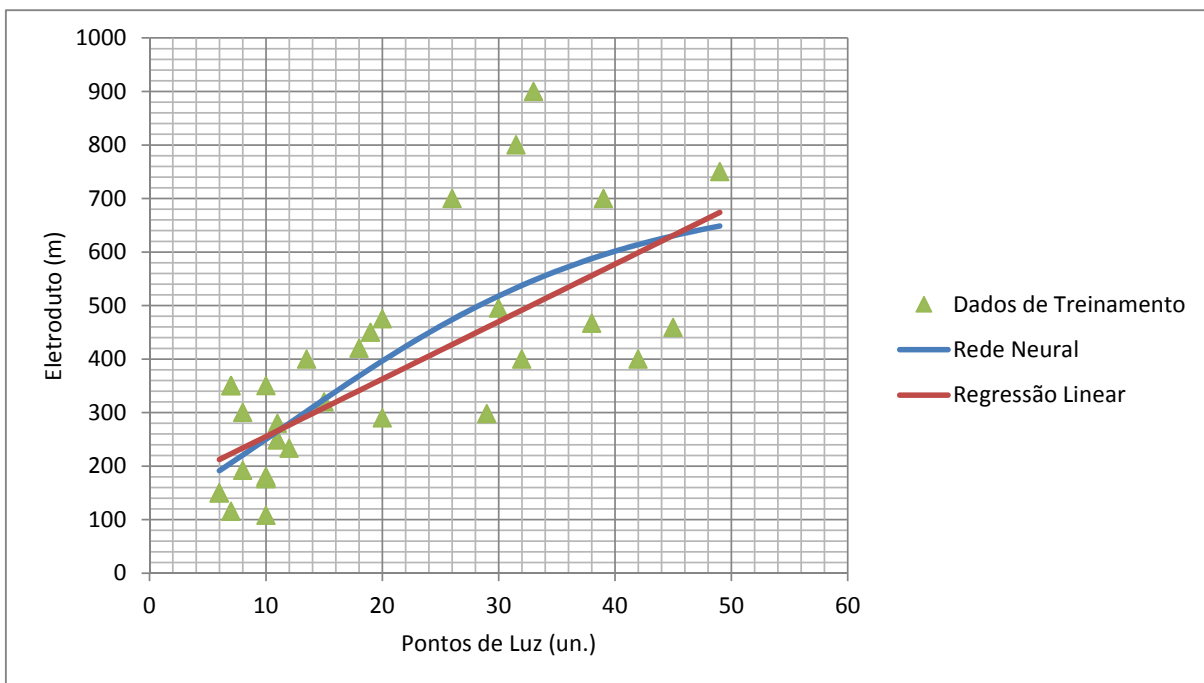
Observa-se que o erro do modelo unidimensional é menor que o erro do modelo multidimensional para uso no estudo de viabilidade que, por suas vez, é menor que o erro do modelo para uso no estudo preliminar, o qual conta com maior disponibilidade de informações. Isso ratifica a dificuldade de regressões em espaços multidimensionais e a importância do especialista na seleção e/ou combinação dos atributos (análise preliminar). O maior erro dos modelos multidimensionais não elimina sua utilidade uma vez que

estes conseguem avaliar a influência de várias características do empreendimento. Além disso, os modelos multidimensionais precisam de maior número de amostras para atingir seu desempenho potencial, por isso, com o aumento do banco de dados, espera-se que os modelos multidimensionais superem o desempenho dos unidimensionais.

4.3.4 Gráficos dos modelos simplificados unidimensionais

A Figura 14 apresenta o modelo simplificado, unidimensional, obtido por meio de RNA e de regressão linear, com a quantidade de pontos de luz como entrada e os dados utilizados no treinamento dos modelos.

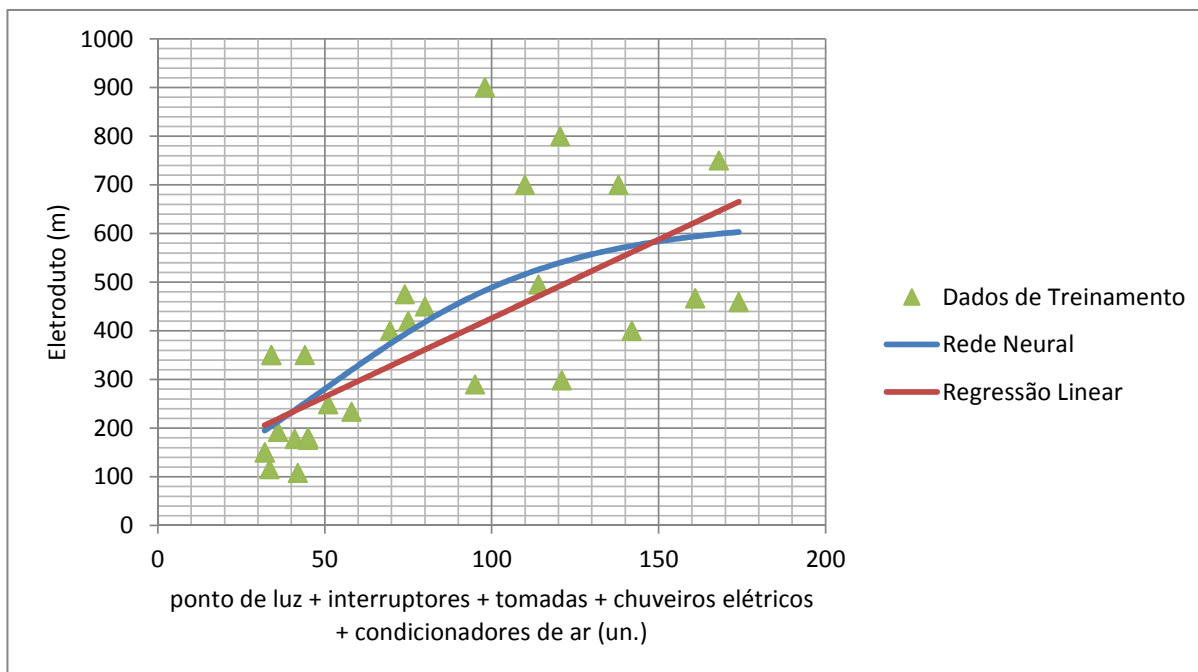
Figura 14 – Quantidade de eletroduto $\frac{1}{2}$ ” e $\frac{3}{4}$ ” por ponto de luz. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Já a Figura 15 apresenta o modelo simplificado, unidimensional, obtido por meio de RNA e de regressão linear, com a variável “Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar” como entrada e os dados utilizados para treinamento dos modelos.

Figura 15 – Quantidade de eletroduto $\frac{1}{2}$ " e $\frac{3}{4}$ " por ponto de luz + interruptores + tomadas + chuveiros elétricos + condicionadores de ar. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Analisando as Figuras 14 e 15 observa-se que o gráfico gerado pela RNA captura a não linearidade do comportamento real do consumo de eletrodutos. A partir de um certa quantidade de pontos, o consumo unitário de eletroduto começa a cair, o que é representado pela característica parabólica da função. Uma explicação possível para esse comportamento é que quando o número de pontos de utilização é alto, ocorre o compartilhamento de eletrodutos por condutores ligados a pontos de utilização diferentes. Por exemplo, em um cômodo com um grande número de tomadas, pode-se alimentar duas tomadas próximas por meio de condutores que chegam pelo mesmo eletroduto. Este fato indica pouco aumento no consumo de eletrodutos em relação ao mesmo cômodo com uma única tomada.

Analisando os dados de treinamento da Figura 14 e da Figura 15 observa-se que existe uma grande dispersão do consumo de eletrodutos em relação a quantidade de pontos de luz, e também com a quantidade total de pontos de utilização, o que dificulta o prognóstico, mesmo estes sendo os atributos com relação de causalidade com a classe mais bem classificados pelo selecionador de atributos.

Com o consumo unitário de eletrodutos sendo altamente dependente do padrão do empreendimento, uma possibilidade para a obtenção de melhores prognósticos seria a elaboração de bancos de dados distintos para obras de diversos padrões, porém, seria necessário uma quantidade de instâncias maior que a coletada para o presente trabalho.

4.4 Condutor de 1,5 mm^2

4.4.1 Preparação dos dados

Uma análise preliminar nos projetos e dados utilizados mostra que a grande maioria dos circuitos de iluminação utiliza a bitola mínima de 1,5 mm^2 especificada pela NBR5410. Em quatro casos, mais especificamente, nas amostras 15, 16, 18 e 45, isso não ocorreu, fazendo com que esses empreendimentos apresentem características distintas dos demais. Dessa forma, eles foram eliminados do conjunto de dados para não interferirem no desempenho dos prognósticos. Caso houvesse um maior número de amostras com essas características, poderia-se dividir os dados em dois conjuntos distintos. Porém, por tratar-se apenas de quatro amostras, o conjunto resultante seria muito pequeno, inviabilizando o treinamento dos modelos.

As instâncias sem informações de quantidade de condutores de 1,5 mm^2 foram eliminadas: 25, 34, 35 e 39.

4.4.2 Seleção de atributos

A Tabela 17 e Tabela 18 apresentam as saídas dos algoritmos de seleção de atributos para o prognóstico da quantidade de condutores de 1,5 mm^2 com e sem a utilização dos pontos de utilização, respectivamente.

Tabela 17 – Resultado dos algoritmos de seleção de atributos para condutores de 1,5 mm² com informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefFAttributeEval / Ranker
melhor suconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Hall	0.687 Área	0.13250 Interruptores Intermediários
Sistema Estrutural	0.698 Luz	0.12335 Pé Direito
Área	0.704 Tomadas Condicionadores Chuveiros Eletricos	0.09929 Luz
WC Servico	0.708 Interruptores Paralelos	0.09807 Água Quente na Cozinha
Lavabo	0.709 Hall	0.09356 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos
Condicionadores de Ar	0.709 Condicionadores de Ar	0.07320 Interruptores
Chuveiros	0.710 Chuveiros	0.07138 Chuveiros
Interruptores Intermediários	0.711 WC Servico	0.07063 Interruptores Simples
Interruptores Paralelos	0.711 Interruptores Intermediários	0.06762 Interruptores Paralelos
Pé Direito	0.711 Lavabo	0.06695 Água Quente nos Lavatórios
Luz	0.712 Sistema Estrutural	0.05287 Tomadas
Tomadas Condicionadores Chuveiros Eletricos	0.712 Pé Direito	0.04743 Tomadas Condicionadores Chuveiros Eletricos
	0.711 Sistema de Aquecimento de Água	0.03134 Sistema de Aquecimento de Água
	0.711 Varanda Gourmet	0.02397 Área
	0.711 Dormitórios	0.02355 Hall
	0.710 Água Quente na Cozinha	0.02014 Condicionadores de Ar
	0.710 Água Quente nos Lavatórios	0.01669 Varanda Gourmet
	0.709 Sistema de Condicionamento de Ar	0.00802 WC Servico
	0.709 Unidades por Pavimentos	0.00681 Pavimentos Tipo
	0.707 Tomadas	0.00343 Sistema Estrutural
	0.705 Interruptores	-0.00200 Dormitórios
	0.703 Pavimentos Tipo	-0.00956 Sistema de Condicionamento de Ar
	0.700 Interruptores Simples	-0.01200 Lavabo
	0.690 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos	-0.03151 Unidades por Pavimentos

Fonte: elaborado pelo autor.

Tabela 18 – Resultado dos algoritmos de seleção de atributos para condutores de $1,5 \text{ mm}^2$ **sem** informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefFAttributeEval / Ranker
melhor suconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Hall	0.687 Área	0.130451 Pé Direito
Sistema Estrutural	0.689 Condicionadores de Ar	0.081345 Chuveiros
Área	0.690 Hall	0.077336 Água Quente nos Lavatórios
Unidades por Pavimentos	0.692 Chuveiros	0.073863 Água Quente na Cozinha
WC Servico	0.693 Unidades por Pavimentos	0.040929 Área
Lavabo	0.694 WC Servico	0.028844 Hall
Condicionadores de Ar	0.694 Sistema Estrutural	0.026572 Sistema de Aquecimento de Água
Chuveiros	0.694 Lavabo	0.021151 WC Servico
Pé Direito	0.695 Pé Direito	0.014185 Condicionadores de Ar
	0.694 Dormitórios	0.013400 Sistema Estrutural
	0.694 Sistema de Aquecimento de Água	0.001522 Varanda Gourmet
	0.694 Varanda Gourmet	-0.000907 Sistema de Condicionamento de Ar
	0.694 Água Quente na Cozinha	-0.001512 Dormitórios
	0.693 Sistema de Condicionamento de Ar	-0.003089 Pavimentos Tipo
	0.692 Água Quente nos Lavatórios	-0.015727 Lavabo
	0.687 Pavimentos Tipo	-0.025240 Unidades por Pavimentos

Fonte: elaborado pelo autor.

Da mesma forma que nos prognósticos de eletrodutos, o selecionador de atributos indica a grande importância das características sem relação de causalidade com a classe a ser prognosticada, indicando que o consumo unitário de condutores de 1,5 mm² varia muito conforme o padrão do empreendimento.

Para o modelo simplificado unidimensional foram selecionados os atributos “Interruptores”, o qual foi criado para esse propósito, e “Luz”.

4.4.3 Modelos de prognóstico

Os relatórios emitidos pelo Weka referentes aos modelos elaborados para o prognóstico de condutores de 1,5 mm², no qual os detalhes de cada modelo podem ser observados, apresentam-se nos apêndices e são organizados da seguinte forma:

- ❑ Modelos com pontos de utilização, para estudo preliminar: Seções .2.1 e .2.2;
- ❑ Modelos sem pontos de utilização, para estudo de viabilidade: Seções .2.3 e .2.4;
- ❑ Modelos simplificados unidimensionais, com pontos de luz como entrada: Seções .2.5 e .2.6 e
- ❑ Modelos simplificados unidimensionais, com interruptores como entrada: Seções .2.7 e .2.8.

O desempenho dos modelos obtidos são apresentados na Tabela 19.

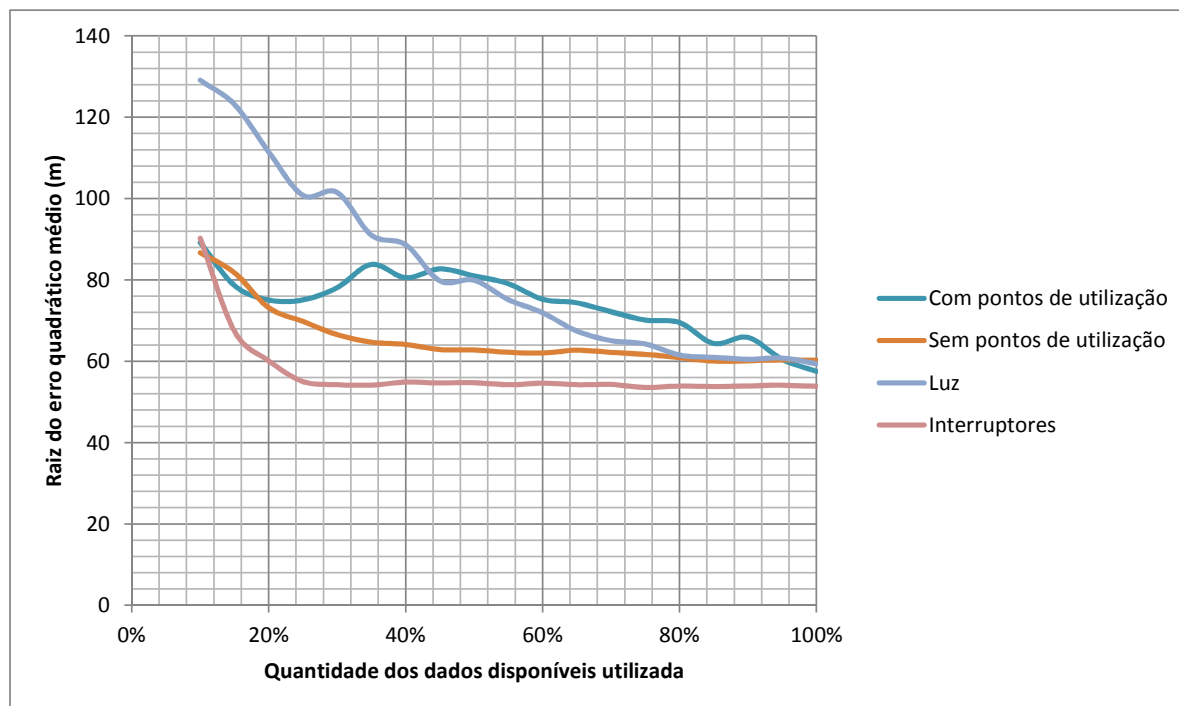
Tabela 19 – Análise de desempenho dos modelos de prognóstico da quantidade de condutores de 1,5 mm²

Ferramenta	Raiz do erro quadrático médio (m)	Erro absoluto médio (m)	Coefficiente de correlação	Erro absoluto médio / média	Instâncias eliminadas (+ de 50% de instâncias sem informação)	Instâncias eliminadas (10% com maior erro)
TCPO	763.52	644.43	0.85	150.5%	24, 26, 27, 28, 29, 32, 33, 36, 37, 38, 41, 42, 43, 44, 45, 46, 47, 49, 50, 51, 52, 53, 54, 58, 59	14, 15, 22
	318.0449	243.8263	0.7199	37.6%	24, 27, 30, 44, 45, 46, 47, 55	5, 14, 28, 33
Com pontos de utilização	230.5626	172.6879	0.8256	27.3%	47, 55	5, 28, 33, 48
	246.8892	203.8906	0.8151	29.3%	30, 31, 32, 44, 45, 46, 55	1, 24, 28, 33
Sem pontos de utilização	221.3982	175.1112	0.8208	26.2%	55	1, 24, 28, 48
	202.0915	156.099	0.8018	35.7%	24, 26, 27, 28, 29, 32, 33, 36, 37, 38, 41, 42, 43, 44, 46, 47, 49, 50, 51, 52, 53, 54, 59	45, 48, 58
Modelo unidimensional - Luz	174.8905	137.8699	0.8478	32.7%	51, 52, 53, 54, 59	48, 57, 58
	197.9613	159.2002	0.8689	26.9%	24, 27, 44, 45, 46, 47, 58, 59	28, 33, 40, 41
Modelo unidimensional - Interruptores	202.6542	157.9311	0.8365	26.5%	58, 59	3, 28, 33, 48

Fonte: elaborado pelo autor.

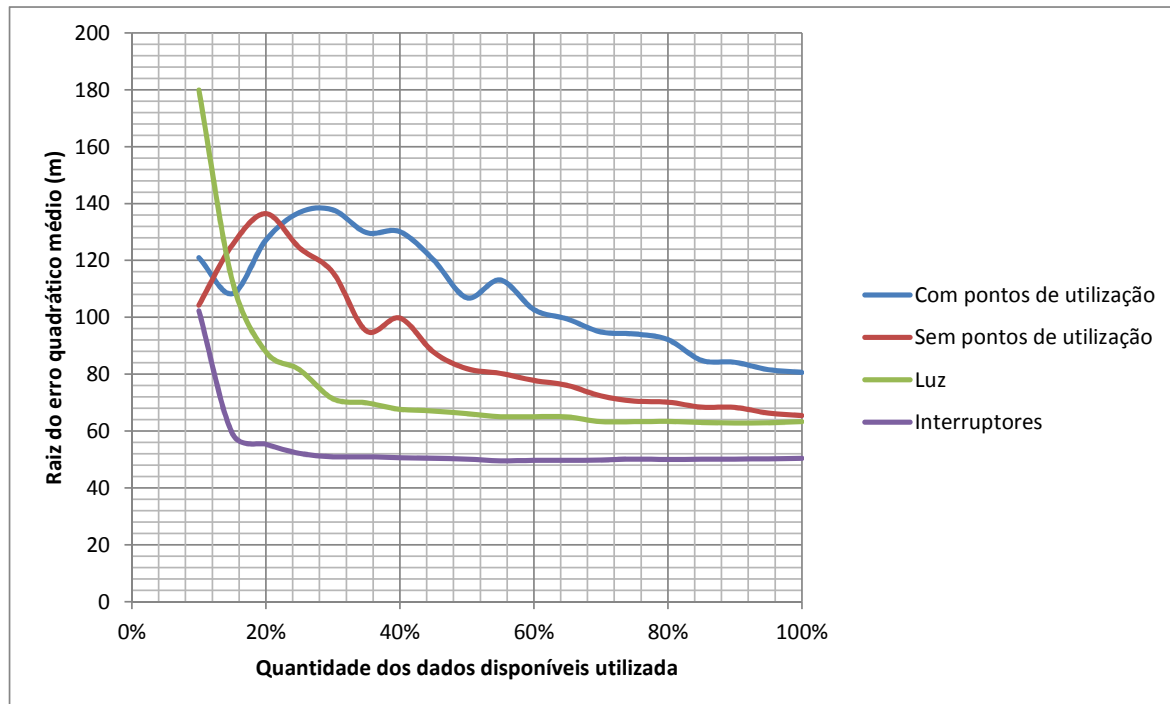
A Figura 16 e a Figura 17 apresentam as curvas de aprendizagem dos modelos desenvolvidos para o prognóstico de condutores de 1,5 mm².

Figura 16 – Curva de aprendizagem para os modelos que utilizam **RNA** para prognóstico de condutores de 1,5 mm².



Fonte: elaborado pelo autor.

Figura 17 – Curva de aprendizagem para os modelos que utilizam **regressão linear** para prognóstico de condutor de $1,5 \text{ mm}^2$.



Fonte: elaborado pelo autor.

Pela observação dos desempenhos dos modelos na Tabela 19 e as curvas de aprendizado nas Figura 16 e Figura 17, é possível constatar que:

- ❑ A TCPO também é inadequada para o pronóstico da quantidade de condutores de $1,5 \text{ mm}^2$;
- ❑ Nos modelos para uso na fase de estudo preliminar, aquele que utiliza a quantidade de pontos de utilização, a RNA apresentou o melhor desempenho, com um erro absoluto médio de 172,69 m. Este valor representa 27,3% da quantidade média de condutor de $1,5 \text{ mm}^2$ das obras utilizadas no treinamento. Esse modelo claramente não atingiu a assíntota de sua curva de aprendizagem, portanto, o aumento do número de amostras ainda deve melhorar seu desempenho;
- ❑ Também no caso dos modelos para uso na fase de estudo de viabilidade, a RNA apresentou melhor desempenho com erro absoluto médio de 175,11 m. Este erro representa 26,2% da quantidade média de de $1,5 \text{ mm}^2$ das obras utilizadas no treinamento;
- ❑ Quanto aos modelos unidimensionais, aquele que utilizou “Luz” como variável de entrada apresentou menor erro absoluto médio, porém, aquele que utilizou a variável

“Interruptores” apresentou menor relação entre erro absoluto médio e a média da quantidade de condutor de 1,5 mm²;

- ❑ O melhor modelo unidimensional com a variável “Luz” foi aquele que utilizou RNA, com erro absoluto médio de 137,87 m, representando 32,7% da média de condutor de 1,5 mm² das obras utilizadas em seu treinamento;
- ❑ O melhor modelo unidimensional com a variável “Interruptores” foi aquele que utilizou RNA, com erro absoluto médio de 157,93 m, representando 26,5% da média de condutor de 1,5 mm² das obras utilizadas em seu treinamento e
- ❑ A diferença de desempenho entre os modelos que utilizam RNA e regressão linear é muito menor nos modelos unidimensionais que nos modelos que contam com vários atributos como entrada, ratificando novamente a capacidade das RNA de trabalhar com diversas variáveis.

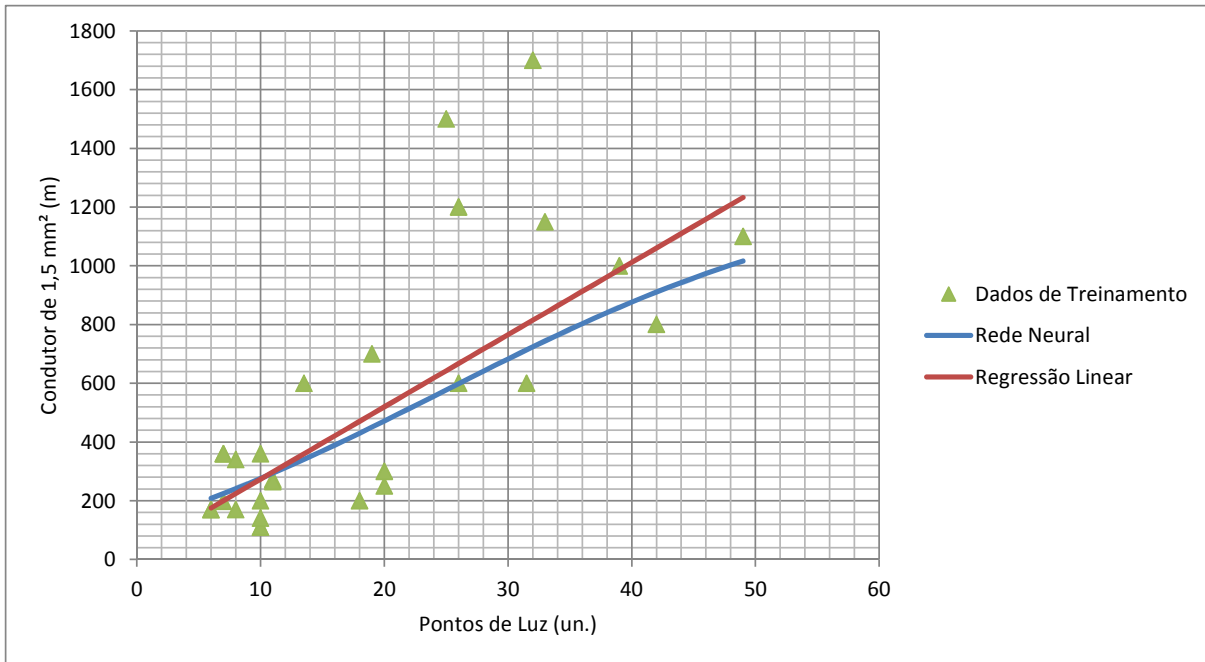
Observa-se que o modelo unidimensional com o atributo “Interruptores” utilizado como entrada apresentou erro similar aos modelos mais complexos. Da mesma forma como observado no prognóstico de eletrodutos, fica ratificada novamente a dificuldade do uso de muitas entradas nos prognósticos e a importância da seleção e/ou combinação dos atributos. Aponta-se novamente que os modelos mais complexos continuam válidos para avaliação do impacto de características referentes aos atributos de entrada.

4.4.4 Gráficos dos modelos simplificados unidimensionais

A Figura 18 apresenta os modelos simplificados, unidimensionais, obtidos por meio de RNA e de regressão linear, com a quantidade de pontos de luz utilizada como entrada nos dados de treinamento dos modelos.

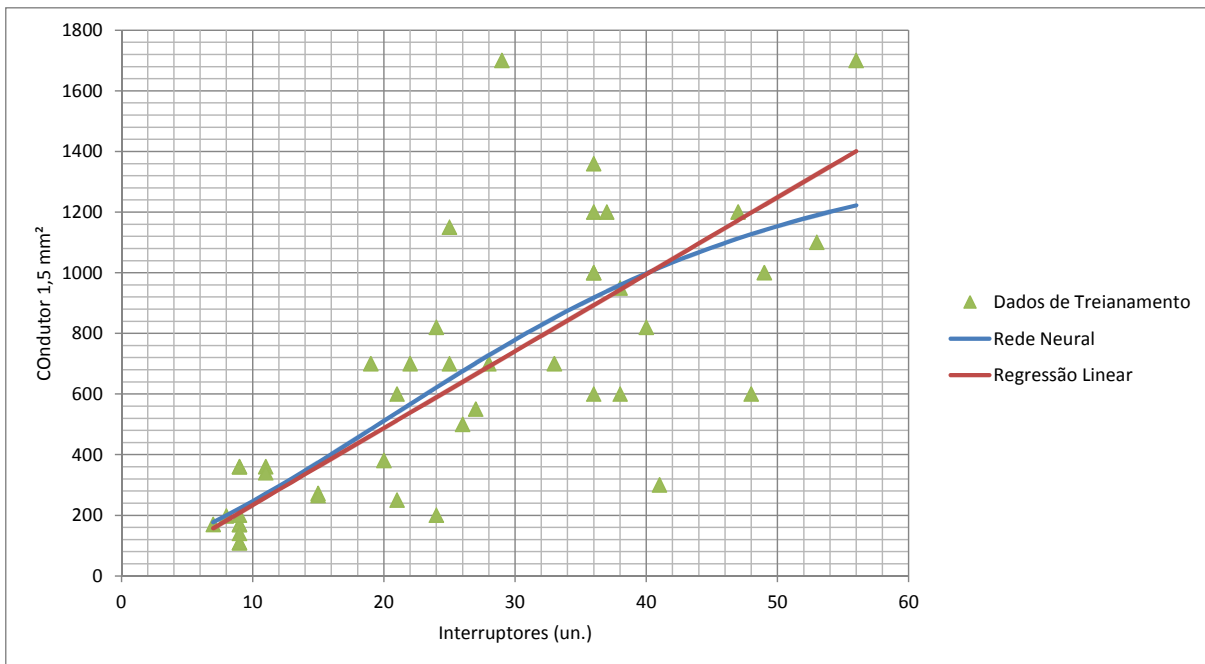
Já a Figura 19 apresenta os modelos simplificados, unidimensionais, obtidos por meio de RNA e de regressão linear, com a variável “Interruptores” como entrada e os dados de treinamento dos modelos.

Figura 18 – Quantidade de condutor de 1,5 mm² por ponto de luz. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Figura 19 – Quantidade de condutor de 1,5 mm² por “Interruptor”. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Analisando os gráficos produzidos pela RNA nas Figuras 18 e 19 observa-se uma característica parabolóide do prognóstico a partir de um certa quantidade de pontos, assim como pode-se observar no caso de prognóstico de eletroduto. No caso do modelo que utiliza os pontos de luz como entrada, o compartilhamento de pontos poderia explicar esse fenômeno, já que podem haver pontos de luz proximamente localizados e ligados aos mesmos interruptores, compartilhando a mesma fiação. Nesses casos o ponto de luz adicional incrementa pouco a quantidade de condutor utilizada. No caso dos interruptores essa explicação não é válida, já que em geral os interruptores paralelos e intermediários não ficam próximos e seu uso incrementa a quantidade de condutores. Dessa forma, pode-se afirmar que o consumo unitário de condutor de 1,5 mm² por interruptores possui característica mais linear.

Os dados de treinamento apresentado pela Figura 18 e pela Figura 19 mostram que os dados coletados sobre consumo de condutor de 1,5 mm² relacionados aos pontos de luz ou interruptores apresentam grande dispersão. De forma análoga aos eletrodutos, o uso de subgrupos levando em conta o padrão dos empreendimentos poderia levar a modelos com melhor desempenho, contudo, isso exigiria um banco de dados maior.

4.5 Condutor de 2,5 mm²

4.5.1 Preparação dos dados

Os empreendimentos em que os circuitos de iluminação utilizaram condutor de 2,5 mm² foram removidos para os modelos de prognóstico desses condutores, da mesma forma que nos modelos de 1,5 mm², já que essas obras teriam uma quantidade maior de condutores com bitola de 2,5 mm² que as demais, piorando o desempenho dos modelos.

As instâncias sem informações de quantidade de condutores de 2,5 mm² foram eliminadas: 25, 34, 35 e 39.

4.5.2 Seleção de atributos

A Tabela 20 e Tabela 21 apresentam as saídas dos algoritmos de seleção de atributos para o prognóstico da quantidade de condutores de 2,5 mm² com e sem a utilização dos pontos de utilização, respectivamente.

Tabela 20 – Resultado dos algoritmos de seleção de atributos para condutores de 2,5 mm² **com** informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefAttributeEval / Ranker
melhor suconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Água Quente na Cozinha	0.767 Área	0.13220 Água Quente na Cozinha
Água Quente nos Lavatórios	0.792 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos	0.08803 Interruptores Intermediários
Área	0.795 Interruptores Simples	0.06682 Condicionadores de Ar
Varanda Gourmet	0.796 Condicionadores de Ar	0.05315 Interruptores
WC Servico	0.797 Chuveiros	0.05166 Água Quente nos Lavatórios
Condicionadores de Ar	0.798 Interruptores	0.05045 Interruptores Simples
Chuveiros	0.798 Varanda Gourmet	0.04924 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos
Interruptores Simples	0.798 Água Quente na Cozinha	0.04737 Tomadas Condicionadores Chuveiros Eletricos
Interruptores	0.798 WC Servico	0.04586 Pé Direito
Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos	0.798 Água Quente nos Lavatórios	0.04460 Tomadas
	0.798 Pé Direito	0.04259 Interruptores Paralelos
	0.798 Sistema Estrutural	0.04060 Luz
	0.798 Sistema de Condicionamento de Ar	0.02988 Área
	0.798 Sistema de Aquecimento de Água	0.02487 Hall
	0.798 Hall	0.01440 Sistema de Aquecimento de Água
	0.798 Lavabo	0.01428 Chuveiros
	0.797 Interruptores Intermediários	0.00627 Unidades por Pavimentos
	0.796 Unidades por Pavimentos	0.00604 Sistema Estrutural
	0.795 Dormitórios	-0.00158 Pavimentos Tipo
	0.792 Interruptores Paralelos	-0.00465 Dormitórios
	0.789 Tomadas	-0.01402 Lavabo
	0.786 Luz	-0.02198 Varanda Gourmet
	0.782 Pavimentos Tipo	-0.02457 Sistema de Condicionamento de Ar
	0.778 Tomadas Condicionadores Chuveiros Eletricos	-0.06235 WC Servico

Fonte: elaborado pelo autor.

Tabela 21 – Resultado dos algoritmos de seleção de atributos para condutores de 2,5 mm² **sem** informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefAttributeEval / Ranker
melhor suconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Hall	0.687 Área	0.130451 Pé Direito
Sistema Estrutural	0.689 Condicionadores de Ar	0.081345 Chuveiros
Área	0.690 Hall	0.077336 Água Quente nos Lavatórios
Unidades por Pavimentos	0.692 Chuveiros	0.073863 Água Quente na Cozinha
WC Servico	0.693 Unidades por Pavimentos	0.040929 Área
Lavabo	0.694 WC Servico	0.028844 Hall
Condicionadores de Ar	0.694 Sistema Estrutural	0.026572 Sistema de Aquecimento de Água
Chuveiros	0.694 Lavabo	0.021151 WC Servico
Pé Direito	0.695 Pé Direito	0.014185 Condicionadores de Ar
	0.694 Dormitórios	0.013400 Sistema Estrutural
	0.694 Sistema de Aquecimento de Água	0.001522 Varanda Gourmet
	0.694 Varanda Gourmet	-0.000907 Sistema de Condicionamento de Ar
	0.694 Água Quente na Cozinha	-0.001512 Dormitórios
	0.693 Sistema de Condicionamento de Ar	-0.003089 Pavimentos Tipo
	0.692 Água Quente nos Lavatórios	-0.015727 Lavabo
	0.687 Pavimentos Tipo	-0.025240 Unidades por Pavimentos

Fonte: elaborado pelo autor.

Novamente, assim como no prognóstico de eletrodutos e de condutores de 1,5 mm², variáveis sem relação de causalidade aparecem com muita importância nos selecionadores de atributos. A importância dada à quantidade de interruptores intermediários, a presença de varanda *gourmet* e ao aquecimento de água em cozinhas e lavatórios demonstra claramente a grande influência de padrão do empreendimento no consumo de condutores de 2,5 mm².

Para o modelo simplificado unidimensional foram selecionados os atributos “Tomadas” e “Área”, porém, nenhum dos dois está no topo da saída do algoritmo de seleção de atributos *Relief F*. “Tomadas” é o décimo primeiro atributo em ordem de importância e “Área” o décimo quarto, porém, considera-se os mais adequados em termos de causalidade.

Observando-se as saídas dos seletores de atributo, aponta-se uma deficiência dos dados disponíveis. Existem TUE que utilizam condutores de bitola superior à 2,5 mm² e o levantamento não identifica essas tomadas. Este fato pode prejudicar o desempenho tando dos modelos relacionados aos condutores de 2,5 mm² quando aos de 4 e 6 mm². Desta forma, foi criada uma nova classe, combinando as quantidade de condutores de 2,5, 4 e 6 mm² para ser usada em um modelo simplificado. Neste modelo simplificado, a entrada utilizada foi um atributo que combina a quantidade de tomadas, de chuveiros elétricos e de condicionadores de ar. Resumindo, foi criada uma classe para ser prognosticada, a qual combina todos os condutores de circuitos de força e um atributo que combina todos os pontos de utilização desses circuitos. O desempenho dos modelos é apresentado na Subseção 4.5.3 e sua visualização em Subseção 4.5.4.

Para a comparação de desempenho, no prognóstico que utiliza a TCPO, também foram utilizados como entrada a combinação das quantidades de tomadas, condicionadores de ar e chuveiros elétricos e o resultado foi comparada com a quantidade dos condutores de 2,5, 4 e 6 mm².

4.5.3 Modelos de prognóstico

Os relatórios emitidos pelo Weka referentes aos modelos elaborados para o prognóstico de condutores de 2,5 mm², onde os detalhes de cada modelo podem ser observados, estão nos apêndices, organizados da seguinte forma:

- ❑ Modelos com pontos de utilização, para estudo preliminar: Seções .3.1 e .3.2;
- ❑ Modelos sem pontos de utilização, para estudo de viabilidade: Seções .3.3 e .3.4;
- ❑ Modelos simplificados unidimensionais, com tomadas como entrada: Seções .3.5 e .3.6;
- ❑ Modelos simplificados unidimensionais, com área como entrada: Seções .3.7 e .3.8 e
- ❑ Modelos simplificados unidimensionais, com área como entrada: Seções .3.7 e .3.8.

Os relatórios referentes aos modelos unidimensionais simplificados para prognóstico da quantidade de condutores de $2,5 \text{ mm}^2$, 4 mm^2 e 6 mm^2 com tomadas, chuveiros elétricos e condicionadores de ar como entrada estão nos apêndices, nos itens: .4.1 à .4.2.

O desempenho dos modelos obtidos são apresentados na Tabela 22 e na Tabela 23.

Tabela 22 – Análise de desempenho dos modelos de prognóstico da quantidade de condutores de 2,5 mm²

Ferramenta	Raiz do erro quadrático médio (m)	Erro absoluto médio (m)	Coefficiente de correlação	Erro absoluto médio / média	Instâncias eliminadas (+ de 50% de instâncias sem informação)	Instâncias eliminadas (10% com maior erro)
Com pontos de utilização	Regressão Linear	306.0955	0.846	26.6%	24, 27, 45, 46, 47, 55	7, 26, 28, 33, 57
	Rede Neural	244.281	0.899	20.9%		6, 7, 26, 28, 57
Sem pontos de utilização	Regressão Linear	315.019	0.857	24.1%	30, 31, 45, 46, 55	7, 24, 28, 43, 57
	Rede Neural	301.3471	0.843	23.6%		6, 7, 28, 43, 57
Modelo unidimensional - Tomadas	Regressão Linear	336.2188	0.7964	30.0%	24, 27, 44, 45, 46, 47,	6, 7, 26, 28
	Rede Neural	384.5298	0.7237	32.2%	59	6, 7, 28, 29
Modelo unidimensional - Área	Regressão Linear	303.6553	0.827	26.0%	55	6, 7, 28, 45, 46
	Rede Neural	305.9923	0.8241	26.4%		6, 7, 28, 45, 46

Fonte: elaborado pelo autor.

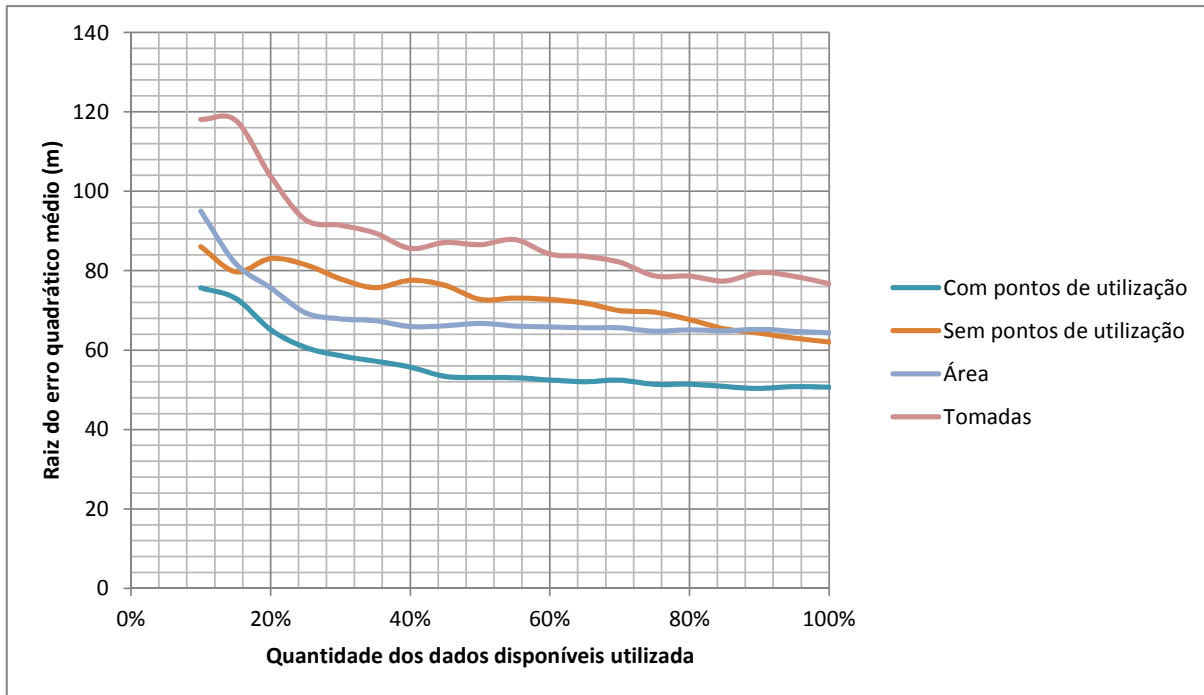
Tabela 23 – Análise de desempenho dos modelos de prognóstico da quantidade de condutores de 2,5, 4 e 6 mm²

Ferramenta	Desempenho					
	Raiz do erro quadrático médio (m)	Erro absoluto médio (m)	Coefficiente de correlação	Erro absoluto médio / média	Instâncias eliminadas (+ de 50% de instâncias sem informação)	Instâncias eliminadas (10% com maior erro)
TCPO, considerando Condutores de 2,5, 4 e 6 mm ² por Tomadas + Chuveiros Elétricos + Condicionadores de Ar	365.31	320.13	0.831	31.9%	24, 25, 26, 27, 30, 33, 34, 35, 44, 46, 47, 55, 59	6, 7, 28, 57
Modelo unidimensional - Tomadas + Chuveiros Elétricos + Condicionadores de Ar	Regressão Linear	246.996	0.827	26.0%	55	6, 7, 28, 45, 46
	Rede Neural	305.9923	0.8241	26.4%		6, 7, 28, 45, 46

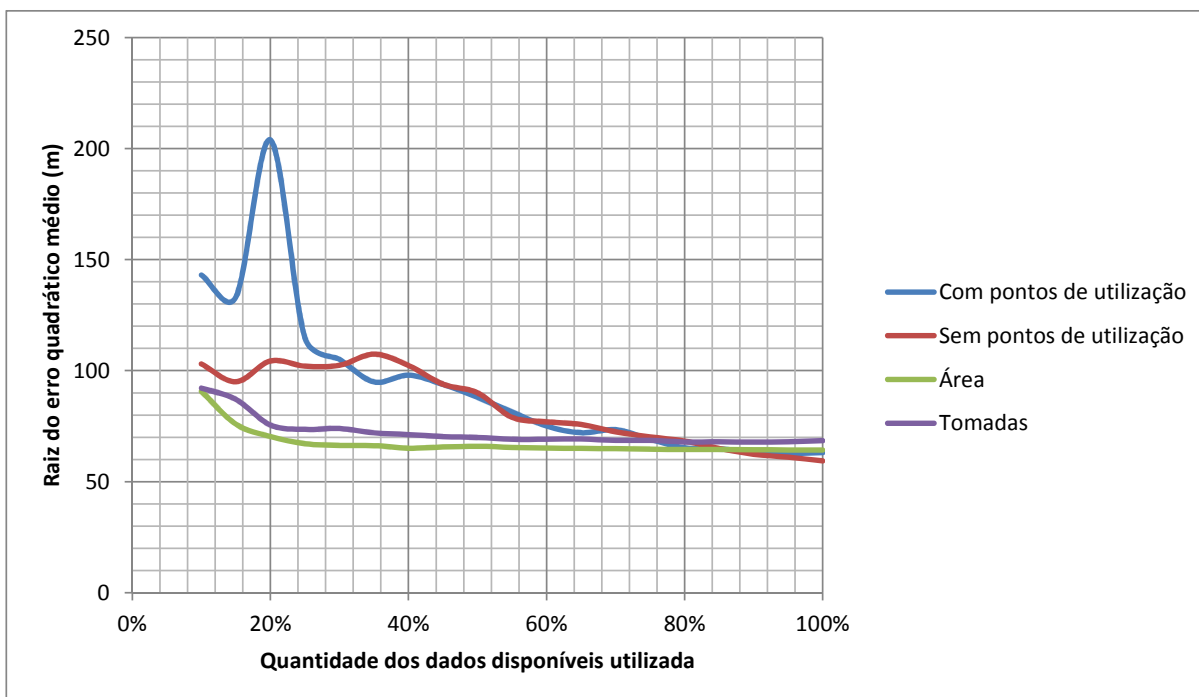
Fonte: elaborado pelo autor.

A Figura 20 e a Figura 21 apresentam as curvas de aprendizagem dos modelos desenvolvidos para o prognóstico de condutores de $2,5 \text{ mm}^2$, enquanto a Figura 22 apresenta as curvas de aprendizagem dos modelos desenvolvidos para o prognóstico de condutores de $2,5$, 4 e 6 mm^2 .

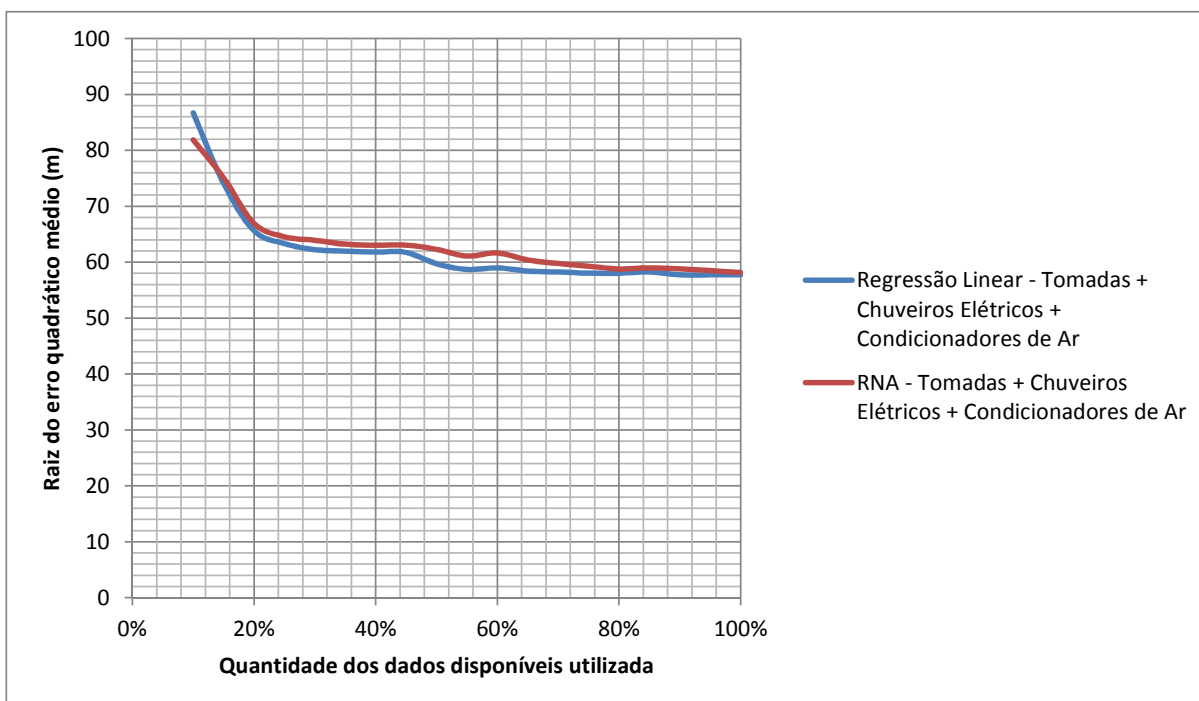
Figura 20 – Curva de aprendizagem para os modelos que utilizam **RNA** para prognóstico de condutores de $2,5 \text{ mm}^2$.



Fonte: elaborado pelo autor.

Figura 21 – Curva de aprendizagem para os modelos que utilizam **regressão linear** para prognóstico de condutor de 2,5 mm².

Fonte: elaborado pelo autor.

Figura 22 – Curva de aprendizagem para os modelos que utilizam **RNA** e **regressão linear** para prognóstico de condutores de 2,5, 4 e 6 mm².

Fonte: elaborado pelo autor.

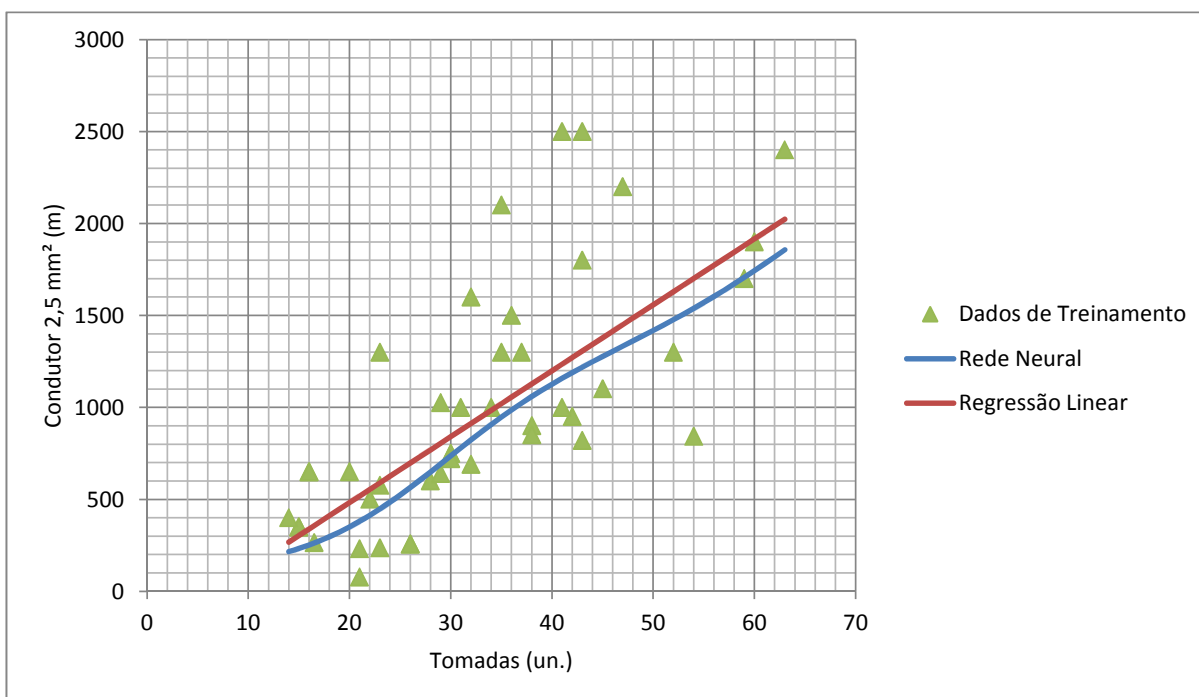
Observando o desempenho dos modelos na Tabela 22 e na Tabela 23, e as curvas de aprendizado nas Figura 20, Figura 21 e Figura 22 verifica-se que:

- ❑ No caso dos condutores de 2,5, 4 e 6 mm^2 , o desempenho da TCPO foi pior que os modelos propostos para o prognóstico de 2,5 mm^2 , porém, seu desempenho não foi tão ruim quanto nos prognósticos de eletrodutos e do condutor de 1,5 mm^2 . Lembrando que foram necessárias adaptações para seu uso, que originalmente só previa condutores de 1,5 mm^2 ;
- ❑ No caso do modelo para uso na fase de estudo preliminar (com a quantidade de pontos de utilização) o modelo utilizando RNA apresentou o melhor desempenho, com um erro absoluto médio de 193,37 m, valor que representa 20,9% da quantidade média de condutor de 2,5 mm^2 das obras utilizadas no treinamento;
- ❑ Também no caso do modelo para uso na fase de estudo de viabilidade a RNA apresentou melhor desempenho com erro absoluto médio de 229,81 m, que representa 23,6% da quantidade média de condutor de 2,5 mm^2 das obras utilizadas no treinamento;
- ❑ Quanto aos modelos unidimensionais, aquele que utilizou “Área” como variável de entrada apresentou menor erro absoluto médio. O modelo que utiliza “Tomadas” continua útil, visto que apresenta melhor relação de causalidade, possibilitando simular a inclusão e remoção de tomadas no projeto;
- ❑ O melhor modelo unidimensional com a variável “Tomadas” foi aquele que utilizou regressão linear, com erro absoluto médio de 274,13 m, representando 30,0% da média de condutor de 2,5 mm^2 das obras utilizadas em seu treinamento;
- ❑ O melhor modelo unidimensional com a variável “Área” foi aquele que utilizou regressão linear, com erro absoluto médio de 247,00 m, representando 26,0% da média de condutor de 2,5 mm^2 das obras utilizadas em seu treinamento e
- ❑ O uso do atributo “Tomadas + Condicionadores + Chuveiros_Eletricos” resultou em um modelo melhor, possibilitando a simulação da inclusão e remoção dos pontos de utilização e apresentando um menor erro absoluto médio, 262,14 m, representando 25,7 % da média de condutores de 2,5, de 4 e de 6 mm^2 das obras utilizadas em seu treinamento. Para o uso do modelo, seria necessário levantar o número de pontos de utilização que utilizam cada uma das bitolas e manter essa proporcionalidade na quantidade de condutores.

4.5.4 Gráficos dos modelos simplificados unidimensionais

A Figura 23 apresenta o modelo simplificado, unidimensional, obtido por meio de RNA e de regressão linear, com a quantidade de tomadas como entrada e os dados utilizados para o treinamento dos modelos.

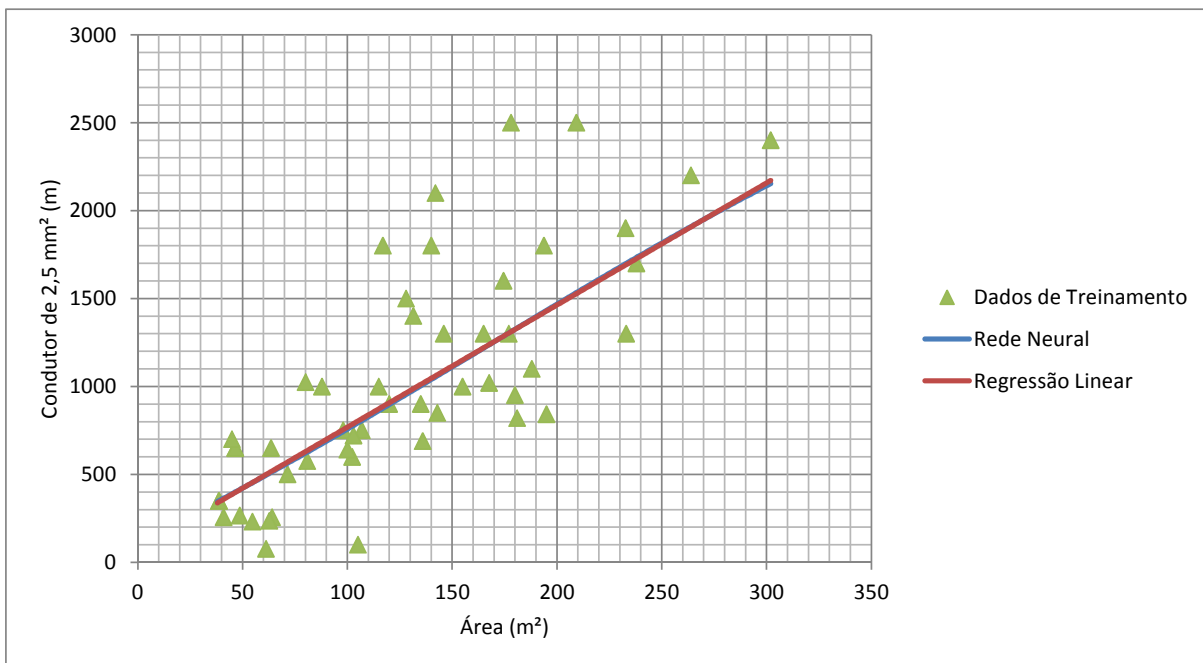
Figura 23 – Quantidade de condutor de 2,5 mm² por tomadas. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Já a Figura 24 apresenta o modelo simplificado, unidimensional, obtido por meio de RNA e de regressão linear, com a variável “Área” como entrada e os dados utilizados para o treinamento dos modelos.

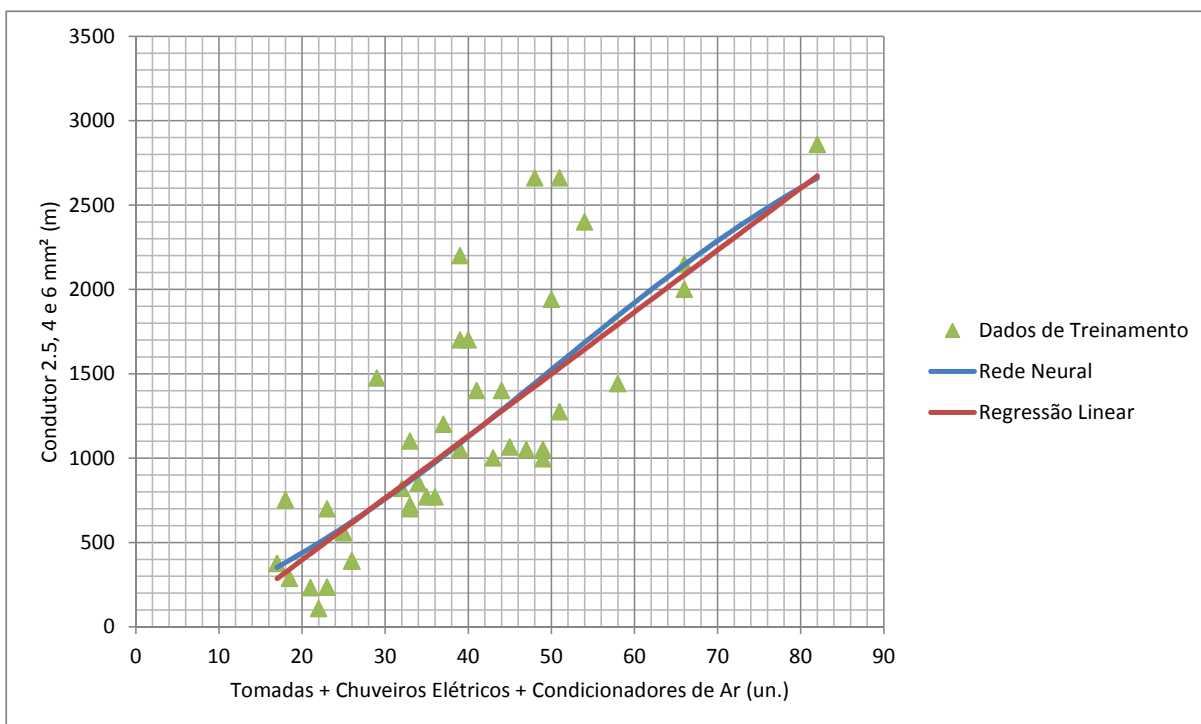
Figura 24 – Quantidade de condutor de 2,5 mm^2 por “Área”. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Os modelos simplificados unidimensionais para prognóstico da quantidade de condutores de 2,5, 4 e 6 mm^2 , utilizando RNA e regressão linear, são apresentados na Figura 25, assim como os dados utilizados para treinamento.

Figura 25 – Quantidade de condutor de 2,5, 4 e 6 mm² por “Tomadas + Chuveiros Elétricos + Condicionadores de Ar”. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Analisando os gráficos apresentados nas Figuras 23, 24 e 25 é possível observar que a distribuição dos dados de treinamento é bastante linear, e que a linha referente à RNA se aproximou muito da linha referente à regressão linear, o que explica o desempenho próximo obtido nos modelos simplificados unidimensionais.

4.6 Condutores de 4 e 6 mm^2

4.6.1 Preparação dos dados

As instâncias sem informações de quantidade de condutores de 4 e 6 mm^2 foram eliminadas: 25, 34, 35 e 39.

Duas instâncias apresentam quantidade zero de condutores de 4 e 6 mm^2 . Essas instâncias, 19 e 23, foram eliminadas do banco de dados. São obras que não usam chuveiro elétrico, não apresentam condicionadores de ar e não devem ter nenhuma tomada de uso específico com potência que exija condutores acima de 2,5 mm^2 . Em casos similares, não é necessário usar um modelo de prognóstico, bastando uma análise das especificações do projeto para a conclusão de que não serão utilizados condutores de 4 e 6 mm^2 .

4.6.2 Seleção de atributos

A Tabela 24 apresenta as saídas dos algoritmos de seleção de atributos para o prognóstico da quantidade de condutores de 4 e 6 mm^2 com a utilização dos pontos de utilização.

Tabela 24 – Resultado dos algoritmos de seleção de atributos para condutores de 4 e 6 mm² com informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefFAttributeEval / Ranker
melhor suconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Hall	0.633 Tomadas	0.16059 Sistema de Aquecimento de Água
Água Quente na Cozinha	0.635 Hall	0.09485 Tomadas Condicionadores Chuveiros Eletricos
Água Quente nos Lavatórios	0.637 Interruptores Intermediários	0.09219 Condicionadores de Ar
Tomadas	0.637 Água Quente nos Lavatórios	0.06812 Tomadas
Interruptores Intermediários	0.637 Água Quente na Cozinha	0.04855 Interruptores Intermediários
	0.637 Pé Direito	0.03468 Água Quente na Cozinha
	0.636 Varanda Gourmet	0.01637 Água Quente nos Lavatórios
	0.633 Sistema de Condicionamento de Ar	0.01520 Interruptores Paralelos
	0.631 WC Servico	0.00663 Área
	0.628 Lavabo	-0.01964 Interruptores Simples
	0.625 Sistema de Aquecimento de Água	-0.02071 Interruptores
	0.621 Sistema Estrutural	-0.03618 Luz
	0.612 Chuveiros	-0.03773 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos
	0.599 Dormitórios	-0.03897 Pavimentos Tipo
	0.580 Tomadas Condicionadores Chuveiros Eletricos	-0.04839 Hall
	0.572 Interruptores Paralelos	-0.05145 Dormitórios
	0.562 Condicionadores de Ar	-0.05651 Unidades por Pavimentos
	0.549 Luz	-0.05721 Chuveiros
	0.540 Pavimentos Tipo	-0.07397 Sistema Estrutural
	0.532 Luz Interruptores Tomada Condicionadores de Ar Chuveiros Elétricos	-0.07950 WC Servico
	0.527 Interruptores	-0.08142 Pé Direito
	0.523 Área	-0.08874 Sistema de Condicionamento de Ar
	0.520 Unidades por Pavimentos	-0.09778 Lavabo
	0.515 Interruptores Simples	-0.11438 Varanda Gourmet

Fonte: elaborado pelo autor.

A Tabela 25 apresenta as saídas dos algoritmos de seleção de atributos para o prognóstico da quantidade de condutores de 4 e 6 mm² sem a utilização dos pontos de utilização.

Tabela 25 – Resultado dos algoritmos de seleção de atributos para condutores de 4 e 6 mm^2 sem informações de pontos de utilização.

Algoritmo / Método de Busca		
CfsSubsetEval / Greedy Stepwise	CfsSubsetEval / Greedy Stepwise com Ranker	ReliefFAttributeEval / Ranker
melhor subconjunto	mérito do subconjunto e atributo adicionado	mérito do respectivo atributo
Hall	0.476 Área	0.18070 Sistema de Aquecimento de Água
Água Quente na Cozinha	0.477 Água Quente na Cozinha	0.07781 Condicionadores de Ar
Água Quente nos Lavatórios	0.478 Hall	0.06751 Água Quente na Cozinha
Área	0.478 Água Quente nos Lavatórios	0.03491 Área
	0.478 Sistema de Condicionamento de Ar	-0.00719 Água Quente nos Lavatórios
	0.478 Pé Direito	-0.02558 Chuveiros
	0.478 Varanda Gourmet	-0.03638 Pavimentos Tipo
	0.478 WC Servico	-0.04746 Hall
	0.477 Sistema Estrutural	-0.05370 Dormitórios
	0.477 Sistema de Aquecimento de Água	-0.05906 Unidades por Pavimentos
	0.476 Lavabo	-0.07376 Sistema Estrutural
	0.475 Chuveiros	-0.07694 WC Servico
	0.473 Dormitórios	-0.08292 Sistema de Condicionamento de Ar
	0.469 Condicionadores de Ar	-0.08509 Lavabo
	0.464 Pavimentos Tipo	-0.08809 Pé Direito
	0.460 Unidades por Pavimentos	-0.10831 Varanda Gourmet

Fonte: elaborado pelo autor.

Novamente, a quantidade de condutores de 4 e 6 mm^2 é altamente correlacionada com atributos que fazem referência ao padrão do empreendimento, como presença de água quente nos lavatórios e na cozinha.

Para o modelo simplificado unidimensional foi selecionado o atributo “Área”. Conforme explicado na Subseção 4.5.2, o modelo simplificado unidimensional com melhor relação de causalidade foi feito com a combinação das quantidades de condutores de 2,5, 4 e 6 mm^2 .

4.6.3 Modelos de prognóstico

Os relatórios emitidos pelo Weka referentes aos modelos elaborados para o prognóstico de condutores de 4 e 6 mm^2 , onde os detalhes de cada modelo podem ser observados, estão nos apêndices, organizados da seguinte forma:

- ❑ Modelos com pontos de utilização, para estudo preliminar: Seções .5.1 e .5.2;
- ❑ Modelos sem pontos de utilização, para estudo de viabilidade: Seções .5.3 e .5.4;
- ❑ Modelos simplificados unidimensionais, com “Área” como entrada: Seções .5.5 e .5.6.

Os modelos simplificados unidimensionais para prognóstico da quantidade de condutores de 2,5 mm^2 , 4 mm^2 e 6 mm^2 com tomadas, chuveiros elétricos e condicionadores de ar

como entrada foram apresentados na Subseção 4.5.3, bem como o resultado do prognóstico utilizando a TCPO.

O desempenho dos modelos obtidos são apresentados na Tabela 26.

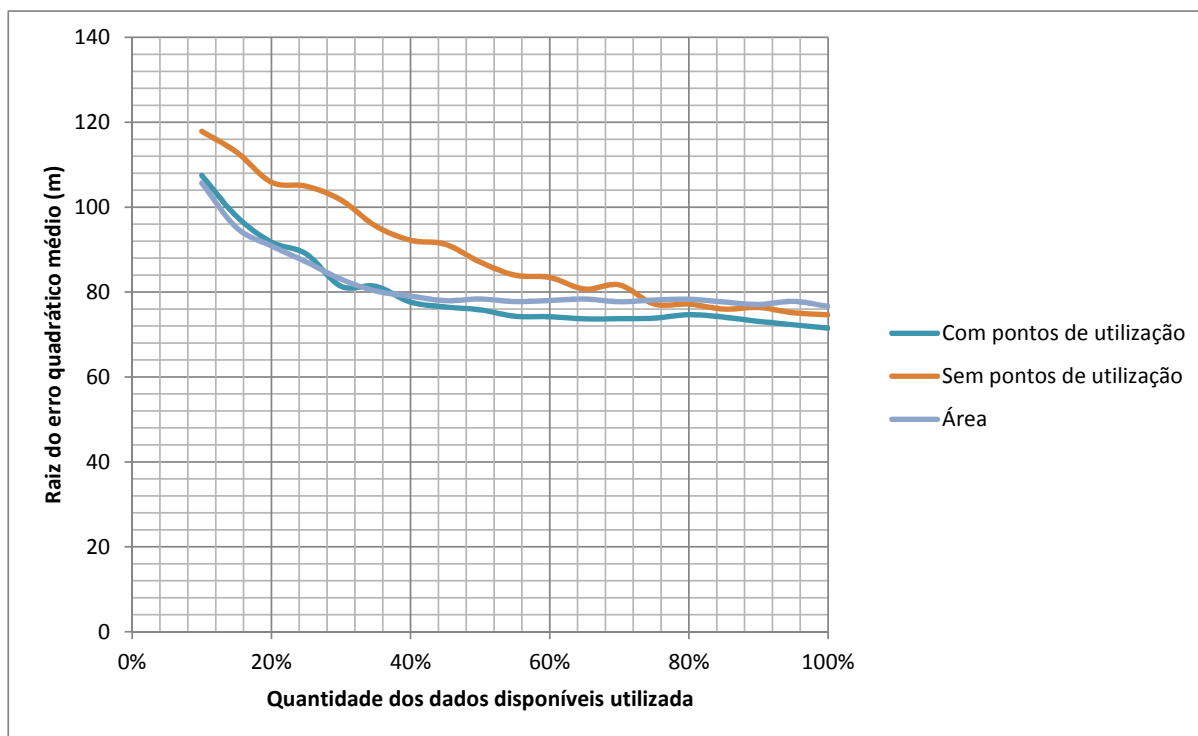
Tabela 26 – Análise de desempenho dos modelos de prognóstico da quantidade de condutores de 4 e 6 mm²

		Desempenho						
Ferramenta		Raiz do erro quadrático médio (m)	Erro absoluto médio (m)	Coefficiente de correlação	Erro absoluto médio / média	Instâncias eliminadas (+ de 50% de instâncias sem informação)	Instâncias eliminadas (10% com maior erro)	
Com pontos de utilização	Regressão Linear	61.5847	45.2794	0.7313	35.7%	24, 26, 27, 30, 31, 32,	8, 14, 15, 18,	
	Rede Neural	35.4842	27.8549	0.7965	24.5%	33, 37, 40, 44, 45, 46, 47, 55, 58, 59	15, 18, 22, 51	
Sem pontos de utilização	Regressão Linear	70.372	50.5643	0.6404	41.4%	24, 26, 27, 32, 33, 37,	6, 7, 15, 18	
	Rede Neural	40.7274	34.6721	0.7067	31.5%	40, 45, 46, 47, 55	15, 18, 22, 51	
Modelo unidimensional - Área	Regressão Linear	54.159	43.565	0.5966	36.4%		15, 22, 26, 27, 33	
	Rede Neural	47.0866	39.0579	0.6673	33.2%	55	15, 22, 24, 26, 27	

Fonte: elaborado pelo autor.

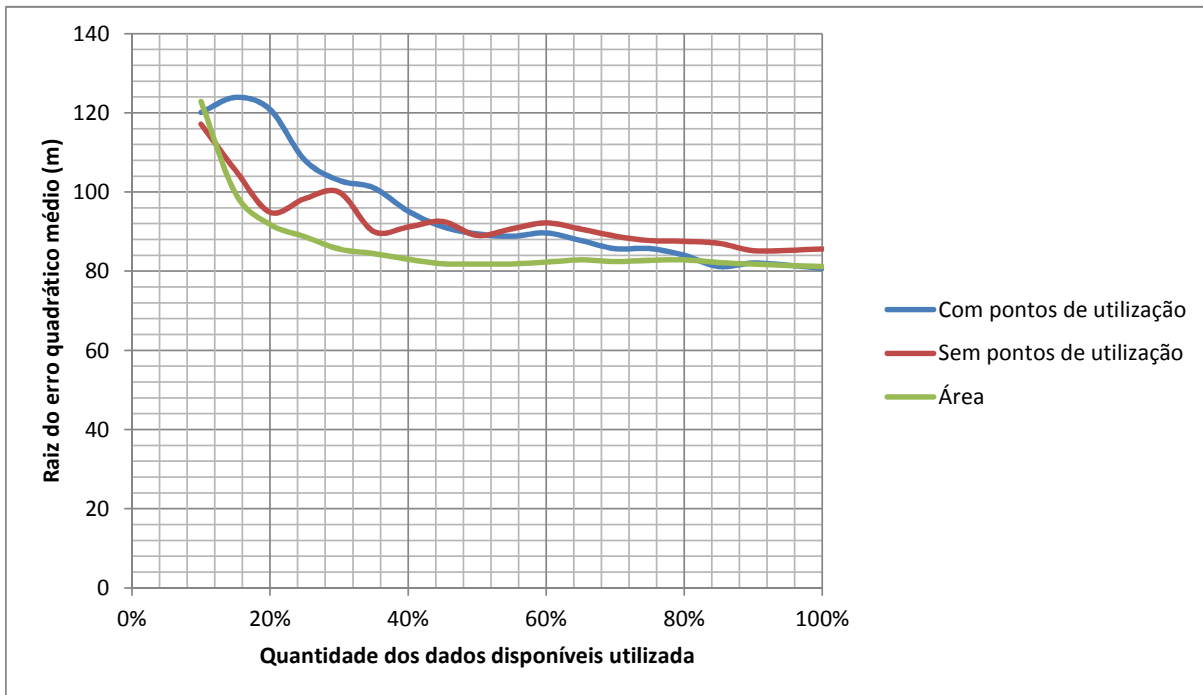
A Figura 26 e a Figura 27 apresentam as curvas de aprendizagem dos modelos desenvolvidos para o prognóstico de condutores de 4 e 6 mm².

Figura 26 – Curva de aprendizagem para os modelos que utilizam **RNA** para prognóstico de condutores de 4 e 6 mm².



Fonte: elaborado pelo autor.

Figura 27 – Curva de aprendizagem para os modelos que utilizam **regressão linear** para prognóstico de condutor de 4 e 6 mm^2 .



Fonte: elaborado pelo autor.

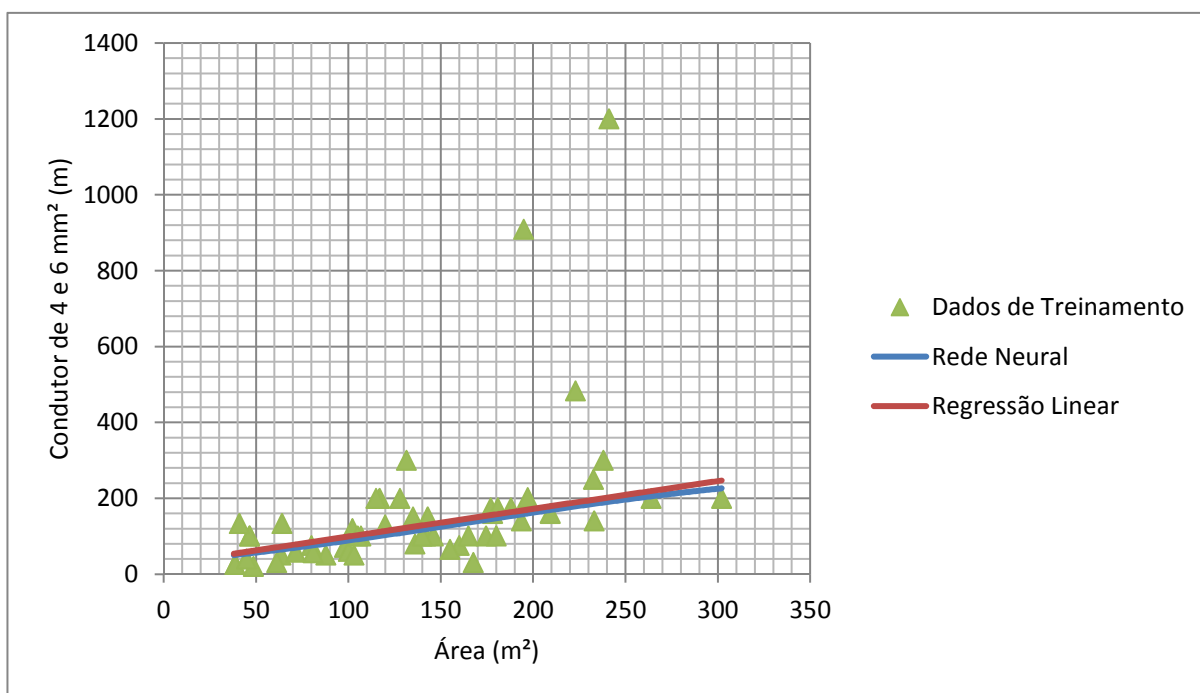
Observando o desempenho dos modelos na Tabela 26 é possível observar que:

- ❑ No caso do modelo para uso na fase de estudo preliminar (com a quantidade de pontos de utilização) o modelo utilizando RNA apresentou o melhor desempenho, com um erro absoluto médio de 27,9 m, valor que representa 24,5% da quantidade média de condutor de 4 e 6 mm^2 das obras utilizadas no treinamento;
- ❑ Também no caso do modelo para uso na fase de estudo de viabilidade a RNA apresentou melhor desempenho com erro absoluto médio de 34,7 m, que representa 31,5% da quantidade média de condutor de 4 e 6 mm^2 das obras utilizadas no treinamento;
- ❑ Quanto ao modelo unidimensional com a variável “Área”, o melhor desempenho foi apresentado por aquele que utilizou RNA, com erro absoluto médio de 39,1 m, representando 33,2% da média de condutor de 4 e 6 mm^2 das obras utilizadas em seu treinamento;
- ❑ A diferença de desempenho entre os modelos que utilizam RNA e regressão linear é significativa nos modelos que utilizam diversas variáveis como entrada, porém, é muito pequena no modelo unidimensional, ratificando novamente a capacidade das RNA de trabalhar com diversas variáveis.

4.6.4 Gráficos dos modelos simplificados unidimensionais

A Figura 28 apresenta o modelo simplificado, unidimensional, obtido por meio de RNA e de regressão linear, com a área como entrada e os dados utilizados para treinamento dos modelos.

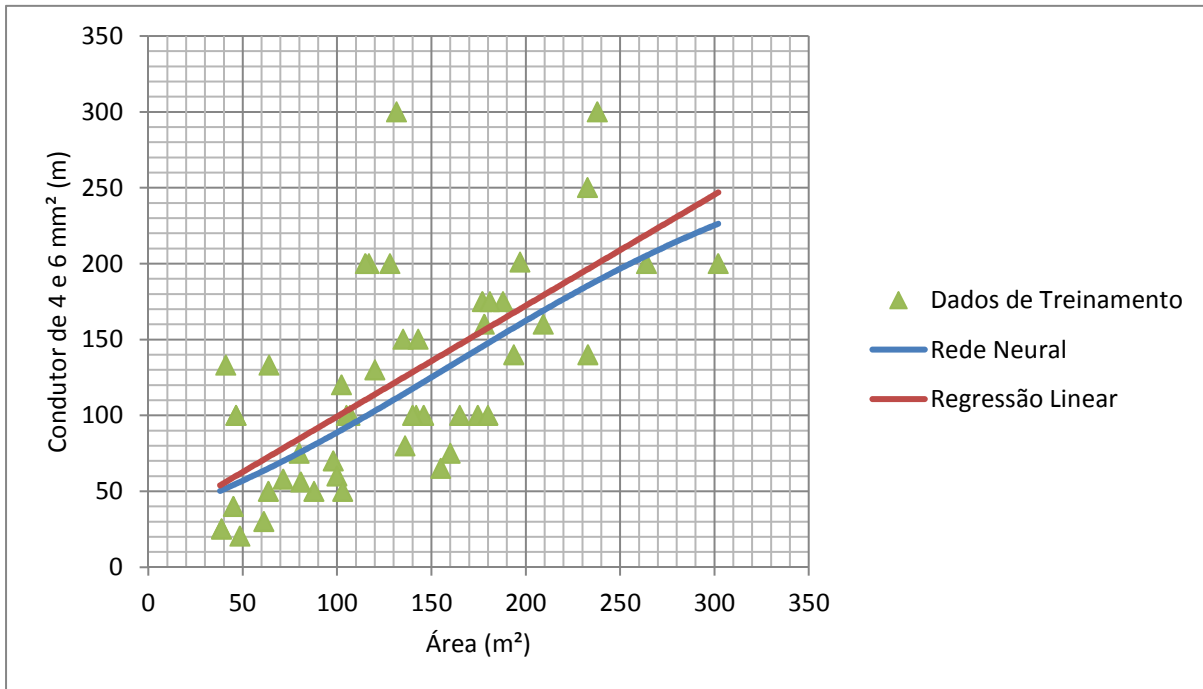
Figura 28 – Quantidade de condutor de 4 e 6 mm² por área. Prognóstico utilizando RNA e regressão linear e dados de treinamento.



Fonte: elaborado pelo autor.

Para facilitar a visualização dos dados, a Figura 29 também apresenta o modelo simplificado, unidimensional, obtido por meio de RNA e de regressão linear, com a área como entrada e os dados utilizados para treinamento dos modelos. Porém, esta figura representa o modelo com a remoção dos dados de treinamento referente as instâncias 15, 22, 26 e 27, que foram removidos tanto no modelo que usa regressão linear, quanto no modelo que usa RNA por fazerem parte do grupo dos prognósticos com maior erro absoluto.

Figura 29 – Quantidade de condutor de 4 e 6 mm² por área. Prognóstico utilizando RNA e regressão linear e dados de treinamento. Sem dados de treinamento referente às instâncias 15, 22, 26 e 27.



Fonte: elaborado pelo autor.

Analisando o gráfico produzido pela RNA nas Figuras 28 e 29 ve-se que a técnica de remoção dos 10% de atributos que apresentaram os maiores erros absolutos na validação cruzada foi eficaz evitando a influência de *outliers* (atributos 15, 22, 26 e 27, visualmente identificáveis).

Também é possível ver como os gráficos produzidos pela RNA e pela regressão linear ficaram próximos, explicando o desempenho parecido.

4.7 Teste em Dados Inéditos

A Tabela 27 e a Tabela 28 apresentam amostras inéditas para o teste dos modelos propostos.

Devido a quantidade de instâncias limitadas que foram coletadas, visto as curvas de aprendizagem de alguns modelos não atingiram suas assíntotas, foram analisadas apenas duas obras inéditas, optando-se por melhorar o treinamento dos modelos em detrimento da validade do teste com dados inéditos. O teste em questão pode ser muito suscetível à *outliers* e um resultado com desempenho pior que os apresentados na validação cruzada dos modelos seria difícil de ser analisado. De qualquer forma, o teste ilustra o uso dos modelos.

Tabela 27 – Amostras inéditas. Parte 1 de 2.

Obra	Hall	Sistema Estrutural	Sistema de Condicionamento de Ar	Água Quente na Cozinha	Água Quente nos Lavatórios	Área	Unidades por Pavimento	Varanda Gourmet	WC Serviço	Lavabo	Condicionadores de Ar	Chuveiros	Tomadas
Amostra Inédita 1	Não	Concreto Armado	Split	Não	Não	91.14	4.00	Sim	Não	Sim	4.00	2.00	30.00
Amostra Inédita 2	Não	Concreto Armado	Split	Sim	Sim	123.23	4.00	Sim	Sim	Não	5.00	4.00	32.00

Fonte: elaborado pelo autor.

Tabela 28 – Amostras inéditas. Parte 2 de 2.

Obra	Interruptor	Interruptores Simples	Interruptores Intermediários	Interruptores Paralelos	Pé Direito	Luz	Luz + Interruptores + Tomada + Condicionadores de Ar + Chuveiros Elétricos	Tomada + Condicionadores de Ar + Chuveiros Elétricos	Condutor 2,5mm ² , 4mm ² e 6mm ²	Eletroduto 1/2" e 3/4"	Condutor 4mm ² e 6mm ²	Condutor 2,5mm ²	Condutor 1,5mm ²
Amostra Inédita 1	28.00	11.00	1.00	16.00	2.73	22.00	84.00	34.00	1015.00	401	115	900	550
Amostra Inédita 2	32.00	18.00	1.00	13.00	2.73	30.00	100.00	38.00	1570.00	550	170	1400	710

Fonte: elaborado pelo autor.

A Tabela 29 apresenta os prognósticos realizados em obras inéditas utilizando os modelos apresentados no trabalho, além dos indicadores de desempenho desses prognósticos.

Tabela 29 – Teste em dados inéditos

Classe	Atributos	Ferramenta	Amostra Inédita 1			Amostra Inédita 2		
			Prognóstico (m)	Erro Absoluto (m)	Erro Relativo	Prognóstico (m)	Erro Absoluto (m)	Erro Relativo
Eletroduto	Com Pontos de Utilização	Regressão Linear	358.755	-43	-10.61%	429.106	-121	-21.98%
		RNA	384.252	-17	-4.25%	508.455	-42	-7.55%
	Sem Pontos de Utilização	Regressão Linear	269.824	-131	-32.77%	573.751	24	4.32%
		RNA	338.926	-62	-15.55%	468.593	-81	-14.80%
	Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar	Regressão Linear	374.214	-27	-6.75%	425.939	-124	-22.56%
		RNA	433.937	33	8.13%	488.891	-61	-11.11%
Conductor 1.5 mm ²	Luz	Regressão Linear	384.153	-17	-4.28%	470.057	-80	-14.54%
		RNA	423.53	22	5.54%	517.974	-32	-5.82%
	Com Pontos de Utilização	Regressão Linear	440.822	-109	-19.85%	413.634	-296	-41.74%
		RNA	347.305	-203	-36.85%	827.413	117	16.54%
	Sem Pontos de Utilização	Regressão Linear	310.72	-239	-43.51%	709.863	0	-0.02%
		RNA	427.619	-122	-22.25%	843.316	133	18.78%
Luz	Regressão Linear	568.783	19	3.42%	765.475	55	7.81%	
	RNA	513.46	-37	-6.64%	682.754	-27	-3.84%	
Conductor 2.5 mm ²	Interruptores	Regressão Linear	690.414	140	25.53%	791.892	82	11.53%
		RNA	728.012	178	32.37%	827.712	118	16.58%
	Com Pontos de Utilização	Regressão Linear	1012.054	112	12.45%	1644.8	245	17.49%
		RNA	788.752	-111	-12.36%	1619.856	220	15.70%
	Sem Pontos de Utilização	Regressão Linear	1032.421	132	14.71%	1761.168	361	25.80%
		RNA	840.481	-60	-6.61%	1804.252	404	28.88%
Tomadas	Regressão Linear	840.609	-59	-6.60%	912.297	-488	-34.84%	
	RNA	736.261	-164	-18.19%	823.101	-577	-41.21%	
Área	Regressão Linear	707.217	-193	-21.42%	930.022	-470	-33.57%	
	RNA	696.496	-204	-22.61%	920.063	-480	-34.28%	
Conductor 2.5 + 4 + 6	Tomadas + Chuveiros Elétricos + Condicionadores de Ar	Regressão Linear	910.439	-105	-6.68%	1497.995	-72	-3.16%
		RNA	900.41	-115	-7.32%	1524.458	-46	-2.00%
	Com Pontos de Utilização	Regressão Linear	94.512	-20	-17.82%	139.735	-30	-17.80%
		RNA	139.735	25	21.51%	193.154	23	13.62%
	Sem Pontos de Utilização	Regressão Linear	73.033	-42	-36.49%	201.668	32	18.63%
		RNA	89.535	-25	-22.14%	175.419	5	3.19%
Área	Regressão Linear	92.763	-22	-19.34%	116.219	-54	-31.64%	
	RNA	82.741	-32	-28.05%	105.26	-65	-38.08%	

Fonte: elaborado pelo autor.

Embora com baixo número de amostras, os indicadores de desempenho apresentados na Tabela 29 são condizentes, com valores próximos ou menores que os obtidos na validação cruzada dos modelos propostos. Análises mais detalhadas sobre o desempenho de cada conjunto de atributos e ferramentas não são convenientes, pois, com o baixo número de amostras, podem não ser significativos.

Considerações Finais

O presente trabalho foi concluído com êxito, permitindo a aplicação do método proposto com a geração de modelos para prognóstico do consumo de eletrodutos e condutores de IPE.

Analisando os resultados obtidos, observa-se que o conjunto de dados obtidos pode ser considerado genérico, uma vez que abrange várias construtoras, projetistas, cidades e obras de diversos padrões e não apenas obras com características similares. Essa condição faz com que os resultados do trabalho possam ser úteis para um maior número de usuários, porém, devido a maior dispersão dos dados, os resultados apresentam desempenho pior que o esperado caso fossem utilizados dados pertencentes a subconjuntos de obras com características similares. O levantamento de um banco de dados suficientemente significativo é um dos maiores desafios para o prognóstico do consumo de materiais de IPE, especialmente por que esse consumo apresenta grande dispersão e varia de acordo com uma série de características do empreendimento, especialmente aquelas relacionadas ao seu padrão. Observando as curvas de aprendizado apresentadas, observa-se que os modelos multidimensionais que utilizam RNA ainda tem potencial de melhora de desempenho com o aumento do número de amostras.

Com um banco de dados maior, poderia ser possível classificar os empreendimentos em grupos distintos, com classes de padrão por exemplo, e então treinar modelos específicos para cada um desses grupos. Trabalhando com dispersões menores é esperado que os modelos apresentassem melhor desempenho.

As RNA apresentam desempenho significativamente melhor que a regressão linear nos modelos multidimensionais, ratificando sua capacidade de trabalhar com diversas variáveis de entrada, o que é útil para entender o impacto das decisões em cada característica significativa de um projeto, especialmente no caso das IPE que variam muito de acordo com as características de outros sistemas do edifício, como o tipo de aquecimento de água ou o condicionamento de ar, por exemplo.

Nos modelos unidimensionais, quando a relação entre a variável de entrada e a classe é próxima de linear, o desempenho das RNA e da regressão linear é similar, podendo ser

observado nos gráficos desses modelos a proximidade das linhas produzidas por ambas as ferramentas. Nos casos em que existem não linearidades nessa relação, as RNA conseguem representá-las.

Os desempenhos dos modelos desenvolvidos são melhores que aqueles obtidos por meio da TCPO, que é o método corrente, além de possibilitar o uso de diversas variáveis, embasando o processo de projeto. Além disso, havendo disponibilidade de dados para a criação de um banco de dados com obras que apresentem características específicas, relacionadas ao padrão, por exemplo, a aplicação da metodologia proposta deve levar a um desempenho significativamente superior, uma vez que os selecionadores de atributos indicaram grande correlação com atributos relacionados ao padrão dos empreendimentos. Espera-se então, que, reduzindo-se a dispersão de alguns atributos, o desempenho dos prognósticos deve melhorar.

As características dos empreendimentos disponíveis para levantamento também são importantes, visto que a não disponibilidade da bitola dos condutores das tomadas inviabilizou a elaboração dos modelos unidimensionais específicos para condutores de $2,5 \text{ mm}^2$ e de 4 e 6 mm^2 .

Como observado no Capítulo 4, a seleção de atributos melhora o desempenho dos modelos de regressão, além de reduzir a quantidade de características a serem levantadas e auxiliar na interpretação do problema em questão e reduzir o tempo de processamento necessário.

A análise dos atributos feita por especialistas também é importante. Enquanto os seletores de atributos levam em conta apenas a correlação entre atributos e classes, busca-se também relações de causalidade. Diversas vezes os atributos selecionados pelos algoritmos apresentam o melhor desempenho para a realização da regressão, porém, outro, com maior relação de causalidade, é mais adequado para os trabalhos de prognóstico. É o caso dos modelos simplificados unidimensionais para eletrodutos, em que a quantidade de pontos de luz foi apresentada pelos seletores como melhor que o atributo combinado “Luz + Interruptores + Chuveiros Elétricos + Condicionadores de Ar”, porém, devido a maior relação de causalidade, este atributo combinado foi utilizado e é considerado melhor que a quantidade de pontos de luz, mesmo com desempenho pior. Com o aumento do número de instâncias no banco de dados é esperado que os atributos com relação de causalidade sempre estejam bem ranqueados nos selecionadores de atributos, porém, com a dificuldade de obtenção de dados em grandes quantidades, a análise de atributos feita por especialista é fundamental.

O processo de remoção dos dados que apresentam os maiores erros no treinamento mostrou-se eficaz para evitar falhas devido a *outliers*. Os gráficos dos prognósticos de condutores de 4 e 6 mm^2 demonstra claramente que os *outliers* (visualmente identificáveis) não influenciaram os modelos.

Para o uso como ferramenta de prognóstico, onde não é necessário conhecer a influên-

cia de cada variável, os modelos simplificados permitem a plotagem de gráficos, facilmente utilizáveis, e, em alguns casos, com o conjunto de dados utilizados neste trabalho, os modelos unidimensionais apresentaram desempenho similar aos modelos multidimensionais. É esperado que o incremento do banco de dados aumente a diferença de desempenho dos modelos unidimensionais e multidimensionais.

Além disso, as curvas de aprendizagem mostram que o número de instâncias necessárias para o treinamento de um modelo unidimensional utilizando regressão linear é baixo, o que permite o uso do método por construtoras de médio e grande porte.

Espera-se que os modelos desenvolvidos possam ser utilizados por agentes da indústria da construção que precisem fazer prognósticos de consumo de eletrodutos e condutores, como departamentos de orçamentos de construtoras, por exemplo, mas, principalmente, que o método apresentado seja aplicado por empresas que tenham disponibilidade de dados, gerando modelos mais precisos e adequados a realidade de cada um deles.

5.1 Sugestões para trabalhos futuros

As sugestões para trabalhos futuros são:

- ❑ Aumento do banco de dados de forma a permitir o uso de amostras com características similares, de forma a diminuir a dispersão dos dados;
- ❑ Classificação das amostras de acordo com padrão, já que os seletores de atributos demonstraram que características relacionadas ao padrão são relevantes, como essa classificação depende de especificidades de mercados, regiões e, também muda muito com o tempo, sugere-se o teste de outros sistemas inteligentes para realizar essa classificação, como sistemas de inferência nebulosa;
- ❑ Estudar a influência da incorporação do *Building Information Modeling* (BIM) pelas empresas de projeto, já que se espera que isso aumente a disponibilidade de dados, bem como sua precisão, que é um ponto de preocupação no levantamento de quantitativos de eletrodutos e condutores, dada a complexidade da tarefa; Aplicação do método para prognóstico do consumo de materiais em IPE em pavimentos atípicos e
- ❑ Aplicação do método para o prognóstico de consumo de materiais em outros serviços, incluindo aqueles que já contam com prognósticos utilizando outros métodos, permitindo a comparação desses métodos com o proposto no presente trabalho.

Referências

ABNT. **NBR 5410 - Instalações elétricas de baixa tensão**. 2004.

ABOURIZK, S.; KNOWLES, P.; HERMANN, U. R. Estimating labor production rates for industrial construction activities. **Journal of construction engineering and management**, 2001. American Society of Civil Engineers, v. 127, n. 6, p. 502–511, 2001.

ABOURIZK, S. M.; WALES, R. J. Incorporating weather effects in project simulation. In: ASCE. **Computing in Civil and Building Engineering**. [S.l.], 1993. p. 1190–1197.

ANDRADE, A. C. **Método para quantificação das perdas de materiais em obras de construção de edifícios: superestrutura e alvenaria**. 1999. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, 1999.

ANDRADE, A. C. et al. Subsídios para o processo de orçamentação de revestimento cerâmicos. In: **II Simpósio Brasileiro de Gestão da Qualidade e Organização do Trabalho no Ambiente Construído**. Fortaleza: [s.n.], 2001.

ARAÚJO, L. O. C. **Método para previsão e controle da produtividade da mão-de-obra na execução de fôrmas, armação, concretagem e alvenaria**. 385 p. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, São Paulo, 2000.

ARAÚJO, N. M. C.; ALVES, G. d. S. Composições de custos unitários: Tpo x apropriação in loco. In: **V CONNEPI-2010**. [S.l.: s.n.], 2010.

BASHIRI, M.; GERANMAYEH, A. F. Tuning the parameters of an artificial neural network using central composite design and genetic algorithm. **Scientia Iranica**, 2011. v. 18, n. 6, p. 1600 – 1608, 2011.

BOLWIJN, P.; KUMPE, T. Manufacturing in the 1990s – Productivity, flexibility and innovation. **Long Range Planning**, 1990. v. 23, n. 4, p. 44–57, ago. 1990. ISSN 00246301. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/002463019090151S>>.

CARDOSO, F. F. Estratégias empresariais e novas formas de racionalização da produção no setor de edificações no Brasil e na França. Parte 1: o ambiente do setor e as estratégias. In: **Estudos econômicos da Construção 2**. São Paulo: [s.n.], 1996.

- CARDOSO, R. **Orçamento de obras em foco: um novo olhar sobre a engenharia de custos**. Pini, 2009. ISBN 9788572662161. Disponível em: <<http://books.google.com.br/books?id=WSeiQAAACAAJ>>.
- CARRARO, F. **Produtividade da mão-de-obra no serviço de alvenaria**. Tese (Doutorado) — Escola Politécnica da Universidade de São Paulo, São Paulo, 1998.
- CHAO, L.-C.; SKIBNIEWSKI, M. J. Estimating construction productivity: Neural-network-based approach. **Journal of Computing in Civil Engineering**, 1994. American Society of Civil Engineers, v. 8, n. 2, p. 234–251, 1994.
- CHWIF, L.; MEDINA, A. C. **Modelagem e simulação de eventos discretos**. [S.l.]: Afonso C. Medina, 2006.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of control, signals and systems**, 1989. Springer, v. 2, n. 4, p. 303–314, 1989.
- FURUSAKA, S. et al. New Approximate Estimation System of Building Using Neural Network. In: **CIB JOINT W55/W65 TRIENNIAL SYMPOSIUM ON CUSTOMER SATISFACTION: A FOCUS FOR RESEARCH & PRACTICE**. Cape Town: [s.n.], 1999.
- GALA, P.; CASTELO, A. M. **"Preparem-se para o segundo semestre", alerta SindusCon-SP**. 2009. Disponível em: <<http://www.sindusconsp.com.br/msg.asp?id=2414>>.
- GONCALVES, C. M. M. **Método para gestão do custo da construção no projeto de edificações**. 182 p. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, 2001.
- HALL, M. et al. The weka data mining software: an update. **ACM SIGKDD Explorations Newsletter**, 2009. ACM, v. 11, n. 1, p. 10–18, 2009.
- HALL, M. A. **Correlation-based Feature Subset Selection for Machine Learning**. Tese (Doutorado) — University of Waikato, Hamilton, New Zealand, 1998.
- IBGE. **Pesquisa anual da indústria da construção 2009**. Rio de Janeiro: Ministério do Planejamento, Orçamento e Gestão, 2009. 98 p.
- KÄRKKÄINEN, H.; JUSSILA, J. J.; LEINO, M. Learning from and with customers with social media: a model for social customer learning. **Mednarodna revija za management, znanje in izobraževanje**, 2012. n. 1, p. 5–25, 2012.
- KARSHENAS, S.; FENG, X. Application of neural networks in earthmoving equipment production estimating. In: ASCE. **Computing in Civil Engineering and Geographic Information Systems Symposium**. [S.l.], 1992. p. 841–847.
- KIRA, K.; RENDELL, L. A. A practical approach to feature selection. In: SLEEMAN, D. H.; EDWARDS, P. (Ed.). **Ninth International Workshop on Machine Learning**. [S.l.]: Morgan Kaufmann, 1992. p. 249–256.
- KNOWLES, R. P. **Predicting construction labour productivity using neural networks**. 215 p. Dissertação (Mestrado) — University of Alberta, 1997.

- KOHAVI, R.; SOMMERFIELD, D. Feature subset selection using the wrapper method: Overfitting and dynamic search space topology. In: **KDD**. [S.l.: s.n.], 1995. p. 192–197.
- KOLMOGOROV, A. N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. In: **Dokl. Akad. Nauk SSSR**. [S.l.: s.n.], 1957. v. 114, n. 5, p. 953–956.
- KONONENKO, I. Estimating attributes: Analysis and extensions of relief. In: BERGADANO, F.; RAEDT, L. D. (Ed.). **European Conference on Machine Learning**. [S.l.]: Springer, 1994. p. 171–182.
- LIBRAIS, C. F. **Método Prático para Estudo da Produtividade da Mão-de-Obra no Serviço de Revestimento Interno de Paredes e Pisos com Placas Cerâmicas**. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, São Paulo, 2001.
- MARCHIORI, F. F. **Desenvolvimento de um Método para Elaboração de Redes de Composições de Custo para Orçamentação de Obras de Edificações**. 237 p. Tese (Doutorado) — Escola Politécnica, Universidade de São Paulo, 2009.
- MASCARO, J. L. **O custo das decisões arquitetônicas**. 2a ed. rev. ed. Porto Alegre: [s.n.], 1998. 180 p.
- MELHADO, S. B.; AGOPYAN, V. O conceito de projeto na construção de edifícios: diretrizes para sua elaboração e controle. **Boletim Técnico BT/PCC/139**. São Paulo: **Escola Politécnica da Universidade de São Paulo**, 1995. 1995.
- MONTEIRO FILHA, D. C. et al. Construção civil no Brasil: investimentos e desafios. **BNDES: Perspectivas do Investimento 2010- 2013**, 2010. 2010. Disponível em: <http://www.bndes.gov.br/SiteBNDES/export/sites/default/bndes_pt/Galerias/Arquivos/conhecimento/liv_perspectivas/09_Perspectivas_do_Investimento_2010_13_CONSTRUCAO_CIVIL.pdf>.
- MOREIRA, D. A. **Medida da produtividade na empresa moderna**. São Paulo: Pioneira, 1991. 154 p.
- MORI, L. M. **Sistema de informação gerencial para previsão de produtividade do trabalho na alvenaria de elevação**. 232 p. Tese (Doutorado) — Universidade Federal de Santa Catarina, 2008.
- MOSELHI, O.; HEGAZY, T.; FAZIO, P. Neural networks as tools in construction. **Journal of Construction Engineering and Management**, 1991. American Society of Civil Engineers, v. 117, n. 4, p. 606–625, 1991.
- OLIVEIRA, R.; BELCHIOR, M.; JUNGLES, A. Análise das variáveis que influenciam na rentabilidade de empreendimentos residenciais: estudo de caso. In: **VII Seminário Internacional da LARES**. Campinas: [s.n.], 2007.
- Padilha Junior, M. A.; OLIVEIRA, F. L. d.; ARAÚJO, N. M. C. TCPO X realidade : um estudo comparativo em empresas de edificações verticais de João Pessoa - PB. In: **XXVII Encontro Nacional de Engenharia de Produção (ENEGEP)**. Foz do Iguaçu: [s.n.], 2007. p. 1–10.

- PALIARI, J. C. **Metodologia para coleta e análise de informações sobre perdas e consumos de materiais e componentes nos canteiros de obras de edifícios**. 473 p. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, 1999.
- PALIARI, J. C. **Método para prognóstico da produtividade da mão-de-obra e consumo unitário de materiais: sistemas prediais hidráulicos**. 621 p. Tese (Doutorado) — Escola Politécnica da Universidade de São Paulo, 2008.
- PICCHI, F. A. **Sistema de Qualidade: uso em empresas de construção de edifícios**. 462 p. Tese (Doutorado) — Escola Politécnica da Universidade de São Paulo, 1993.
- PINI, E. **TCPO, Tabelas de Composição de Preços para Orçamentos**. 2. ed. São Paulo: Pini, 1962.
- _____. **TCPO, Tabelas de Composição de Preços para Orçamentos**. 6. ed. São Paulo: Pini, 1977.
- _____. **TCPO, Tabelas de Composição de Preços para Orçamentos**. 7. ed. São Paulo: Pini, 1980.
- _____. **TCPO, Tabelas de Composição de Preços para Orçamentos**. 10. ed. São Paulo: Pini, 1996.
- _____. **TCPO, Tabelas de Composição de Preços para Orçamentos**. 10. ed. São Paulo: Pini, 2003.
- _____. **TCPO, Tabelas de Composição de Preços para Orçamentos**. 13. ed. São Paulo: Pini, 2008. ISBN 978-85-7266-201-7.
- PORTAS, J.; ABOURIZK, S. Neural network model for estimating construction productivity. **Journal of Construction Engineering and Management**, 1997. American Society of Civil Engineers, v. 123, n. 4, p. 399–410, 1997.
- ROBNIK-SIKONJA, M.; KONONENKO, I. An adaptation of relief for attribute estimation in regression. In: FISHER, D. H. (Ed.). **Fourteenth International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 1997. p. 296–304.
- RYNKIEWICZ, J. General bound of overfitting for {MLP} regression models. **Neurocomputing**, 2012. v. 90, n. 0, p. 106 – 110, 2012. Advances in artificial neural networks, machine learning, and computational intelligence (ESANN 2011).
- SHANNON, R. E. Introduction to the art and science of simulation. In: IEEE. **Simulation Conference Proceedings, 1998. Winter**. [S.l.], 1998. v. 1, p. 7–14.
- SILVA, A. L. d. **O uso do paradigma restrições geométricas para apoio a projetos de instalações elétricas**. Tese (Doutorado) — Universidade Federal de Uberlândia, 2006.
- SILVA, I. N. d.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010. 379 p.
- SONMEZ, R.; ROWINGS, J. E. Construction labor productivity modeling with neural networks. **Journal of Construction Engineering and Management**, 1998. American Society of Civil Engineers, v. 124, n. 6, p. 498–504, 1998.

SOUZA, J. R. A. d.; MORENO, H. **Guia EM da NBR 5410**. São Paulo: Editora Aranda, 2001. 289 p.

SOUZA, U. E. L. d. **Metodologia para o estudo da produtividade da mão-de-obra no serviço de fôrmas para estruturas de concreto armado**. 1996. 280 p. Tese (Doutorado) — Escola Politécnica da Universidade de São Paulo, São Paulo, 1996.

SOUZA, U. E. L. d. **Método para a previsão da produtividade da mão-de-obra e do consumo unitário de materiais para os serviços de fôrmas, armação, concretagem, alvenaria, revestimentos com argamassa, contrapiso, revestimentos com gesso e revestimentos cerâmicos**. :: Tese (Livre Docência) — Escola Politécnica da Universidade de São Paulo, São Paulo, 2001.

STOY, C.; SCHAULCHER, H.-R. Residential building projects: Building cost indicators and drivers. **Journal of construction engineering and management**, 2007. American Society of Civil Engineers, v. 133, n. 2, p. 139–145, 2007.

SUETAKE, M. **Sistemas Inteligentes para Monitoramento e Diagnósticos de Falhas em Motores de Indução Trifásicos**. Tese (Doutorado) — Universidade de São Paulo, 2012.

TAM, C.; TONG, T. K. L.; TSE, S. L. Artificial neural networks model for predicting excavator productivity. **Engineering Construction and Architectural Management**, 2002. Wiley Online Library, v. 9, n. 5-6, p. 446–452, 2002.

WITTEN, I. H.; FRANK, E. **Data Mining: Practical machine learning tools and techniques**. [S.l.]: Morgan Kaufmann, 2005.

Apêndices

.1 Relatórios do Weka - Eletrodutos

.1.1 Com Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 11-

Instances: 33

Attributes: 10

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Unidades_por_Pavimentos

Varanda_Gourmet

WC_Servico

Lavabo

Chuveiros

Pe_Direito

Luz

Eletroduto_Meia_Tres_Quartos

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Eletroduto_Meia_Tres_Quartos =

3.0111 * Agua_Quente_na_Cozinha=Sim +
-123.3005 * Agua_Quente_nos_Lavatórios=Sim +
2.4131 * Unidades_por_Pavimentos +
14.5487 * Varanda_Gourmet=Sim +
220.6214 * WC_Servico=Sim +
102.5369 * Lavabo=Sim +
38.225 * Chuveiros +
520.9469 * Pe_Direito +
-0.4868 * Luz +
-1255.9081

Time taken to build model: 0 seconds

==== Cross-validation ====
 ==== Summary ====

Correlation coefficient 0.675
 Mean absolute error 123.2646
 Root mean squared error 153.0834
 Relative absolute error 73.6963%
 Root relative squared error 73.3084%
 Total Number of Instances 33

.1.2 Com Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 8.0E-4 -M 0.0 -N 1000 -V 0
 -S 0 -E 20 -H 4

Relation: Banco de Dados Mestrado 12-

Instances: 34

Attributes: 25

Hall

Sistema_Estrutural

Sistema_de_Aquecimento_de_Agua

Sistema_de_Condicionamento_de_Ar

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Area

Pavimentos_Tipo

Unidades_por_Pavimentos

Varanda_Gourmet

Dormitorios

WC_Servico

Lavabo

Condicionadores_de_Ar

Chuveiros

Tomadas

Interuptores_Simples

Interuptores_Intermediarios

Interuptores_Paralelos

Interuptores

Pe_Direito
Luz
Tomadas_Condicionadores_Chuveiros_Eletricos
Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos
Eletroduto_Meia_Tres_Quartos
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Linear Node 0

Inputs Weights
Threshold 0.3153352806128212
Node 1 0.24703291578769232
Node 2 -0.39448051583758637
Node 3 -0.47586900059139714
Node 4 -0.49223769119572447

Sigmoid Node 1

Inputs Weights
Threshold -0.01239133915767141
Attrib Hall 0.07120155080451529
Attrib Sistema_Estrutural -0.03211464406943281
Attrib Sistema_de_Aquecimento_de_Agua=Passagem à gás -0.04341497808980629
Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro Elétrico 0.015529914656344157
Attrib Sistema_de_Aquecimento_de_Agua=Aquecimento solar -0.02112177216069931
Attrib Sistema_de_Aquecimento_de_Agua=Central -0.015097394877449438
Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro elétrico -0.04172136335123555
Attrib Sistema_de_Condicionamento_de_Ar=Split 0.033829046020310435
Attrib Sistema_de_Condicionamento_de_Ar=Janela 0.026982682405230288
Attrib Sistema_de_Condicionamento_de_Ar=Ausênte -0.057355320874908314
Attrib Agua_Quente_na_Cozinha 0.08656917000400768
Attrib Agua_Quente_nos_Lavatórios -0.06708346081096644
Attrib Area 0.02361489294539206
Attrib Pavimentos_Tipo -0.03906776782105785
Attrib Unidades_por_Pavimentos -0.005139914594320463
Attrib Varanda_Gourmet 0.004974303827793896
Attrib Dormitorios 0.047718759848963135
Attrib WC_Servico -0.057049338526717785
Attrib Lavabo -0.02900295537076491

Attrib Condicionadores_de_Ar -0.002493376262635647
 Attrib Chuveiros 0.022528353052819194
 Attrib Tomadas 0.019156837089732773
 Attrib Interruptores_Simples -0.0409435044768388
 Attrib Interruptores_Intermediarios -0.04809810758940637
 Attrib Interruptores_Paralelos -5.518360041048483E-4
 Attrib Interruptores 0.0437428557108765
 Attrib Pe_Direito 0.015785744883284544
 Attrib Luz 0.05498834814338206
 Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -0.02257099547593254
 Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos -1.98745634507982731

4

Sigmoid Node 2

Inputs Weights

Threshold -0.049121228590932015
 Attrib Hall -0.01479644758041151
 Attrib Sistema_Estrutural 0.14838636645731695
 Attrib Sistema_de_Aquecimento_de_Agua=Passagem à gás -0.007459747513410543
 Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro Elétrico -0.04132240454058373
 Attrib Sistema_de_Aquecimento_de_Agua=Aquecimento solar 0.08958608240230662
 Attrib Sistema_de_Aquecimento_de_Agua=Central -0.023171728376440187
 Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro elétrico -0.008522454362671072
 Attrib Sistema_de_Condicionamento_de_Ar=Split -0.1565187655557197
 Attrib Sistema_de_Condicionamento_de_Ar=Janela 0.025541117981291443
 Attrib Sistema_de_Condicionamento_de_Ar=Ausente 0.0954537875277186
 Attrib Agua_Quente_na_Cozinha -0.16927067938281648
 Attrib Agua_Quente_nos_Lavatórios 0.10110636004720588
 Attrib Area -0.10698015928632791
 Attrib Pavimentos_Tipo -0.09887890617269494
 Attrib Unidades_por_Pavimentos 0.10050494038526385
 Attrib Varanda_Gourmet 0.0792273449964051
 Attrib Dormitorios -0.021124266565859893
 Attrib WC_Servico 0.14839576440377658
 Attrib Lavabo 0.11922313132988122
 Attrib Condicionadores_de_Ar -0.09720214252351109
 Attrib Chuveiros -0.12375868537272951
 Attrib Tomadas -0.00841688076840091
 Attrib Interruptores_Simples -0.05714090129179726

Attrib Interruptores_Intermediarios 0.001755337659215511
Attrib Interruptores_Paralelos -0.055372014123738586
Attrib Interruptores -0.038414490237563485
Attrib Pe_Direito -0.005432253740571556
Attrib Luz -0.06802375945451865
Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -0.046300812088978544
Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos -0.028064005503

Sigmoid Node 3

Inputs Weights
Threshold 0.039811381387840135
Attrib Hall -0.04363501311566422
Attrib Sistema_Estrutural 0.172822837977801
Attrib Sistema_de_Aquecimento_de_Agua=Passagem à gás 0.027207129336525235
Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro Elétrico -0.007770954647827931
Attrib Sistema_de_Aquecimento_de_Agua=Aquecimento solar 0.1451074352352145
Attrib Sistema_de_Aquecimento_de_Agua=Central 0.012912331290942736
Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro elétrico -0.024824454231093897
Attrib Sistema_de_Condicionamento_de_Ar=Split -0.11513481779696381
Attrib Sistema_de_Condicionamento_de_Ar=Janela -0.032213472922302554
Attrib Sistema_de_Condicionamento_de_Ar=Ausente 0.06156230705558908
Attrib Agua_Quente_na_Cozinha -0.1336154296967235
Attrib Agua_Quente_nos_Lavatórios 0.12458272053178839
Attrib Area -0.0620829965858747
Attrib Pavimentos_Tipo -0.05199514932649703
Attrib Unidades_por_Pavimentos 0.0770350719798797
Attrib Varanda_Gourmet 0.0786120469578655
Attrib Dormitorios -0.030091949825820006
Attrib WC_Servico 0.2129088924925664
Attrib Lavabo 0.20235364703869635
Attrib Condicionadores_de_Ar -0.12626405028554366
Attrib Chuveiros -0.11598908794308892
Attrib Tomadas -0.025032470326840858
Attrib Interruptores_Simples -0.02597946481166423
Attrib Interruptores_Intermediarios -0.038792867067613424
Attrib Interruptores_Paralelos -0.09902892266515578
Attrib Interruptores -0.1138453092774357
Attrib Pe_Direito -0.06037945695268068
Attrib Luz -0.09943125130774355

Attrib Tomadas_Condicionadores_Chuveiros_Eletricos 3.696537816730869E-4

Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos -0.09009557203113794

Sigmoid Node 4

Inputs Weights

Threshold 0.018767499331518424

Attrib Hall -0.10784988772171802

Attrib Sistema_Estrutural 0.13969737567999682

Attrib Sistema_de_Aquecimento_de_Agua=Passagem à gás -0.02737456020416608

Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro Elétrico -0.00938857812382689

Attrib Sistema_de_Aquecimento_de_Agua=Aquecimento solar 0.11979045711711647

Attrib Sistema_de_Aquecimento_de_Agua=Central -0.006161628121588341

Attrib Sistema_de_Aquecimento_de_Agua=Chuveiro elétrico 0.044818473924230676

Attrib Sistema_de_Condicionamento_de_Ar=Split -0.08955923559440342

Attrib Sistema_de_Condicionamento_de_Ar=Janela 0.054472831885049214

Attrib Sistema_de_Condicionamento_de_Ar=Ausente 0.07400021488137158

Attrib Agua_Quente_na_Cozinha -0.17836560697581957

Attrib Agua_Quente_nos_Lavatórios 0.11773267678136295

Attrib Area -0.15077238207489507

Attrib Pavimentos_Tipo -0.06727289822888041

Attrib Unidades_por_Pavimentos 0.11924794205433854

Attrib Varanda_Gourmet 0.15071942269071917

Attrib Dormitorios -0.03515563216000029

Attrib WC_Servico 0.21337310426076478

Attrib Lavabo 0.14560152807750468

Attrib Condicionadores_de_Ar -0.11865851437109398

Attrib Chuveiros -0.05794065619481351

Attrib Tomadas -0.02289002132638023

Attrib Interruptores_Simples -0.02977426241229871

Attrib Interruptores_Intermediarios -0.029330298226135976

Attrib Interruptores_Paralelos -0.09010662145070678

Attrib Interruptores -0.11846061348928663

Attrib Pe_Direito -0.023776756945530247

Attrib Luz -0.040964571032696455

Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -0.08355110808717277

Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos -0.05480973029954530

Class

Input

Node 0

Time taken to build model: 0.09 seconds

==== Cross-validation ==== Summary ====

Correlation coefficient 0.6481

Mean absolute error 110.8035

Root mean squared error 142.8671

Relative absolute error 70.4207%

Root relative squared error 73.4265%

Total Number of Instances 34

.1.3 Sem Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 11-

Instances: 28

Attributes: 6

Agua_Quente_na_Cozinha

WC_Servico

Lavabo

Chuveiros

Pe_Direito

Eletroduto_Meia_Tres_Quartos

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Eletroduto_Meia_Tres_Quartos =

63.3576 * Agua_Quente_na_Cozinha=Sim +

208.2247 * WC_Servico=Sim +

18.2802 * Lavabo=Sim +

25.3126 * Chuveiros +

681.8613 * Pe_Direito +

-1660.5627

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7465

Mean absolute error 113.7952

Root mean squared error 134.081

Relative absolute error 66.5388%

Root relative squared error 63.9141%

Total Number of Instances 28

.1.4 Sem Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.002 -M 0.0 -N 700 -V 0 -S
0 -E 20 -H 5

Relation: Banco de Dados Mestrado 11-

Instances: 28

Attributes: 6

Agua_Quente_na_Cozinha

WC_Servico

Lavabo

Chuveiros

Pe_Direito

Eletroduto_Meia_Tres_Quartos

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.3111367851640782

Node 1 0.08677774447698196

Node 2 -0.5724837090276761

Node 3 -0.5685556897091627

Node 4 -0.448715127283515

Node 5 0.5914797223747481

Sigmoid Node 1

Inputs Weights

Threshold -0.047467851983156055

Attrib Agua_Quente_na_Cozinha 0.017151482928223046

Attrib WC_Servico -0.0035943835254276875

Attrib Lavabo 0.014491135947344502

Attrib Chuveiros -0.05534367581952203

Attrib Pe_Direito 0.03608060119044642

Sigmoid Node 2

Inputs Weights

Threshold -0.03653349697179755

Attrib Agua_Quente_na_Cozinha -0.19625664799917814

Attrib WC_Servico 0.4209109015884156

Attrib Lavabo 0.362058528302616

Attrib Chuveiros -0.2752845866805611

Attrib Pe_Direito -0.006646494992066444

Sigmoid Node 3

Inputs Weights

Threshold -0.0638346150801974

Attrib Agua_Quente_na_Cozinha -0.16708966443923057

Attrib WC_Servico 0.4296603020679471

Attrib Lavabo 0.37729831682828313

Attrib Chuveiros -0.2573375014533077

Attrib Pe_Direito -0.0253691589017184

Sigmoid Node 4

Inputs Weights

Threshold -0.05197622192623365

Attrib Agua_Quente_na_Cozinha -0.13427345154691642

Attrib WC_Servico 0.40194569566583793

Attrib Lavabo 0.27127656470820366

Attrib Chuveiros -0.2104862508212297

Attrib Pe_Direito -0.072327814542253

Sigmoid Node 5

Inputs Weights

Threshold -0.01438614933769798

Attrib Agua_Quente_na_Cozinha 0.1458167315023385

Attrib WC_Servico -0.28285486708178664

Attrib Lavabo -0.2622291456145776

Attrib Chuveiros 0.20585038413354567

Attrib Pe_Direito -0.00255579701607783

Class

Input

Node 0

Time taken to build model: 0.02 seconds

==== Cross-validation ==== Summary ====

Correlation coefficient 0.7711

Mean absolute error 102.3263

Root mean squared error 118.7984

Relative absolute error 66.8087%

Root relative squared error 63.1469%

Total Number of Instances 28

.1.5 Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8

Relation: Banco de Dados Mestrado 11-

Instances: 25

Attributes: 2

Eletroduto_Meia_Tres_Quartos

a2+a3-a1

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Eletroduto_Meia_Tres_Quartos =

3.2328 * a2+a3-a1 +
102.6576

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7342

Mean absolute error 106.8899

Root mean squared error 122.7379

Relative absolute error 72.3786%

Root relative squared error 66.4734%

Total Number of Instances 25

.1.6 Luz + Interruptores + Tomadas + Chuveiros Elétricos + Condicionadores de Ar - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.06 -M 0.0 -N 600 -V 0 -S 0
-E 20 -H 6

Relation: Banco de Dados Mestrado 11-

Instances: 25

Attributes: 2

Eletroduto_Meia_Tres_Quartos

a2+a3-a1

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.25447668159014947

Node 1 0.22686205675971327

Node 2 0.17523287727630607

Node 3 0.11479231825163146
Node 4 -0.7756627615073278
Node 5 -1.3594458725929914
Node 6 0.2804027554701506

Sigmoid Node 1

Inputs Weights
Threshold -0.395506551784242
Attrib $a_2+a_3-a_1$ 0.0437788033003582

Sigmoid Node 2

Inputs Weights
Threshold -0.39806340689362996
Attrib $a_2+a_3-a_1$ -0.01879648999233607

Sigmoid Node 3

Inputs Weights
Threshold -0.3848032397651372
Attrib $a_2+a_3-a_1$ -0.12815575837549414

Sigmoid Node 4

Inputs Weights
Threshold -1.3669031848657704
Attrib $a_2+a_3-a_1$ -1.5026813137348043

Sigmoid Node 5

Inputs Weights
Threshold -1.626150445080617
Attrib $a_2+a_3-a_1$ -2.2777724364061407

Sigmoid Node 6

Inputs Weights
Threshold -0.11322197784067953
Attrib $a_2+a_3-a_1$ 0.2606012539399707

Class

Input
Node 0

Time taken to build model: 0.02 seconds

==== Cross-validation ====
==== Summary ====

Correlation coefficient 0.7141
Mean absolute error 107.3371
Root mean squared error 129.717
Relative absolute error 69.9531%
Root relative squared error 69.3439%
Total Number of Instances 25

.1.7 Luz - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8
Relation: Banco de Dados Mestrado 12-
Instances: 30
Attributes: 2
Luz
Eletroduto_Meia_Tres_Quartos
Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Eletroduto_Meia_Tres_Quartos =

10.7381 * Luz +
147.9154

Time taken to build model: 0 seconds

==== Cross-validation ====
==== Summary ====

Correlation coefficient 0.7117
Mean absolute error 97.7248

Root mean squared error 117.4088
Relative absolute error 72.6495%
Root relative squared error 68.4029%
Total Number of Instances 30

.1.8 Luz - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.05 -M 0.01 -N 400 -V 0 -S
0 -E 20 -H 4

Relation: Banco de Dados Mestrado 12-
Instances: 31
Attributes: 2
Luz
Eletroduto_Meia_Tres_Quartos
Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights
Threshold 0.5275540925691
Node 1 0.1240174227670933
Node 2 -0.9370156097276627
Node 3 -1.1784763783968657
Node 4 -0.18613045619033736

Sigmoid Node 1

Inputs Weights
Threshold -0.2503928105592945
Attrib Luz -0.004300145225709967

Sigmoid Node 2

Inputs Weights
Threshold -0.9729785566748643
Attrib Luz -1.3286183095465711

Sigmoid Node 3

Inputs Weights

Threshold -1.1850663284000285
Attrib Luz -1.6414761375692672

Sigmoid Node 4

Inputs Weights
Threshold -0.24751669594260461
Attrib Luz -0.3658497312694085

Class

Input
Node 0

Time taken to build model: 0.01 seconds

==== Cross-validation ====
==== Summary ====

Correlation coefficient 0.7164
Mean absolute error 107.7966
Root mean squared error 134.5803
Relative absolute error 68.7345%
Root relative squared error 66.9286%
Total Number of Instances 31

.2 Relatórios do Weka - Condutores de 1,5 mm²

.2.1 Com Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 40

Attributes: 13

Hall

Sistema_Estrutural

Area

WC_Servico

Lavabo

Condicionadores_de_Ar
 Chuveiros
 Interuptores_Intermediarios
 Interuptores_Paralelos
 Pe_Direito
 Luz
 Tomadas_Condicionadores_Chuveiros_Eletricos
 Condutor_Um_Meio
 Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Um_Meio =

134.8 * Hall=Não +
 -318.1906 * Sistema_Estrutural=Concreto Armado +
 5.7855 * Area +
 190.655 * WC_Servico=Sim +
 295.0561 * Lavabo=Sim +
 -66.3801 * Condicionadores_de_Ar +
 30.1098 * Chuveiros +
 113.1746 * Interuptores_Intermediarios +
 0.9766 * Interuptores_Paralelos +
 1658.4429 * Pe_Direito +
 5.3504 * Luz +
 -8.8848 * Tomadas_Condicionadores_Chuveiros_Eletricos +
 -4464.8099

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7199

Mean absolute error 243.8263

Root mean squared error 318.0449

Relative absolute error 65.9386%

Root relative squared error 72.6884%

Total Number of Instances 40

.2.2 Com Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.04 -M 0.0 -N 800 -V 0 -S 0
-E 20 -H 2

Relation: Banco de Dados Mestrado 12-

Instances: 39

Attributes: 13

Hall

Sistema_Estrutural

Area

WC_Servico

Lavabo

Condicionadores_de_Ar

Chuveiros

Interuptores_Intermediarios

Interuptores_Paralelos

Pe_Direito

Luz

Tomadas_Condicionadores_Chuveiros_Eletricos

Condutor_Um_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.07073674103108066

Node 1 -2.010834346477277

Node 2 1.3882082461036964

Sigmoid Node 1

Inputs Weights

Threshold -0.12610149130783888

Attrib Hall 0.036485985168199384

Attrib Sistema_Estrutural 0.9239549834657216
Attrib Area -0.5366500176056301
Attrib WC_Servico 0.05366124446328722
Attrib Lavabo 0.17915182526358028
Attrib Condicionadores_de_Ar 0.7404025920525781
Attrib Chuveiros -1.160174644782679
Attrib Interruptores_Intermediarios -2.3531502814271215
Attrib Interruptores_Paralelos 0.7178651196307596
Attrib Pe_Direito -1.8660905248050776
Attrib Luz -1.3779518934133559
Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -0.2455280224702625

Sigmoid Node 2

Inputs Weights

Threshold 0.004504152819774943
Attrib Hall 0.04463223334484407
Attrib Sistema_Estrutural -0.0902419442971512
Attrib Area 0.43600164956237036
Attrib WC_Servico 0.002482190135247206
Attrib Lavabo -0.018254792488566637
Attrib Condicionadores_de_Ar 0.29908948251780826
Attrib Chuveiros -0.07410737233059723
Attrib Interruptores_Intermediarios -1.1788970417990439
Attrib Interruptores_Paralelos -0.2590148590813333
Attrib Pe_Direito -0.19122173865725126
Attrib Luz -0.38451557793857927
Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -0.15810031454395485

Class

Input

Node 0

Time taken to build model: 0.03 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8256

Mean absolute error 172.6879

Root mean squared error 230.5626
Relative absolute error 49.6667%
Root relative squared error 57.2748%
Total Number of Instances 39

.2.3 Sem Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 40

Attributes: 10

Hall

Sistema_Estrutural

Area

Unidades_por_Pavimentos

WC_Servico

Lavabo

Condicionadores_de_Ar

Chuveiros

Pe_Direito

Condutor_Um_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Um_Meio =

63.7397 * Hall=Não +

78.0725 * Sistema_Estrutural=Concreto Armado +

5.1319 * Area +

22.6569 * Unidades_por_Pavimentos +

273.691 * WC_Servico=Sim +

16.8523 * Lavabo=Sim +

-42.9158 * Condicionadores_de_Ar +

10.2672 * Chuveiros +

616.1976 * Pe_Direito +

-1937.3814

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8151

Mean absolute error 203.8906

Root mean squared error 246.8892

Relative absolute error 56.978%

Root relative squared error 58.295%

Total Number of Instances 40

.2.4 Sem Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.005 -M 0.0 -N 500 -V 0 -S
0 -E 20 -H 3

Relation: Banco de Dados Mestrado 12-

Instances: 40

Attributes: 10

Hall

Sistema_Estrutural

Area

Unidades_por_Pavimentos

WC_Servico

Lavabo

Condicionadores_de_Ar

Chuveiros

Pe_Direito

Condutor_Um_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.6145745749781776

Node 1 -0.8625016644146914
Node 2 -0.8328458572087594
Node 3 -0.0013187530842024532

Sigmoid Node 1

Inputs Weights
Threshold 0.0777745887910059
Attrib Hall -0.0863261188899573
Attrib Sistema_Estrutural 0.11689901043776704
Attrib Area -0.2762223786945478
Attrib Unidades_por_Pavimentos -0.019365153359320434
Attrib WC_Servico 0.6100332098716016
Attrib Lavabo 0.09262459737158524
Attrib Condicionadores_de_Ar -0.1987086877403747
Attrib Chuveiros -0.3409170234410791
Attrib Pe_Direito -0.45895186729994364

Sigmoid Node 2

Inputs Weights
Threshold 0.07685501838480732
Attrib Hall -0.15256603014679507
Attrib Sistema_Estrutural 0.1306997445342761
Attrib Area -0.2501911894812446
Attrib Unidades_por_Pavimentos 0.014983930775976711
Attrib WC_Servico 0.6061668633415
Attrib Lavabo 0.11274862738129787
Attrib Condicionadores_de_Ar -0.23652122222192837
Attrib Chuveiros -0.33762181476316544
Attrib Pe_Direito -0.37343775721468847

Sigmoid Node 3

Inputs Weights
Threshold 0.03608547413618582
Attrib Hall -0.04116964841945705
Attrib Sistema_Estrutural 0.0640131927553672
Attrib Area -0.02303875438158672
Attrib Unidades_por_Pavimentos 0.011332757212626719
Attrib WC_Servico 0.13176105150146336
Attrib Lavabo 0.08125799169801422

Attrib Condicionadores_de_Ar -0.06547839543517636
 Attrib Chuveiros -0.06468248248971668
 Attrib Pe_Direito -0.03174841346209911

Class

Input

Node 0

Time taken to build model: 0.02 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8208

Mean absolute error 175.1112

Root mean squared error 221.3982

Relative absolute error 52.2513%

Root relative squared error 56.2628%

Total Number of Instances 40

.2.5 Luz - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 25

Attributes: 2

Luz

Condutor_Um_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Um_Meio =

24.5864 * Luz +

27.8819

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8018

Mean absolute error 156.099

Root mean squared error 202.0915

Relative absolute error 55.3105%

Root relative squared error 58.9171%

Total Number of Instances 25

.2.6 Luz - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.005 -M 0.08 -N 600 -V 0 -
S 0 -E 20 -H 4

Relation: Banco de Dados Mestrado 12-

Instances: 25

Attributes: 2

Luz

Condutor_Um_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.8076497818731389

Node 1 -1.1864007731271151

Node 2 -1.116458482672681

Node 3 -0.2317084789370203

Node 4 0.7081931561511934

Sigmoid Node 1

Inputs Weights

Threshold -0.19534231246964756

Attrib Luz -1.3676699711635338

Sigmoid Node 2

Inputs Weights

Threshold -0.15485304072198922

Attrib Luz -1.3096753250535393

Sigmoid Node 3

Inputs Weights

Threshold -0.007776739301268208

Attrib Luz -0.479707841388301

Sigmoid Node 4

Inputs Weights

Threshold 0.02696301392808247

Attrib Luz 0.4595580483255979

Class

Input

Node 0

Time taken to build model: 0.01 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8478

Mean absolute error 137.8699

Root mean squared error 174.8905

Relative absolute error 53.0619%

Root relative squared error 55.6384%

Total Number of Instances 25

.2.7 Interruptores - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 39

Attributes: 2
Interuptores
Condutor_Um_Meio
Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Um_Meio =

25.3693 * Interuptores +
-19.926

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8689
Mean absolute error 159.2002
Root mean squared error 197.9613
Relative absolute error 47.22%
Root relative squared error 48.9474%
Total Number of Instances 39

.2.8 Interruptores - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.009 -M 0.0 -N 1000 -V 0 -
S 0 -E 20 -H 3

Relation: Banco de Dados Mestrado 12-
Instances: 39
Attributes: 2
Interuptores
Condutor_Um_Meio
Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.5271969283954563

Node 1 -0.7127749836096481

Node 2 0.7487960847517515

Node 3 -1.6521049575156728

Sigmoid Node 1

Inputs Weights

Threshold -0.16698022965669596

Attrib Interruptores -0.8611079462620344

Sigmoid Node 2

Inputs Weights

Threshold 0.0981450172152992

Attrib Interruptores 0.6243131834269402

Sigmoid Node 3

Inputs Weights

Threshold -0.7696358313470401

Attrib Interruptores -1.8162164119128559

Class

Input

Node 0

Time taken to build model: 0.03 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8365

Mean absolute error 157.9311

Root mean squared error 202.6542

Relative absolute error 48.9992%

Root relative squared error 54.3041%

Total Number of Instances 39

.3 Relatórios do Weka - Condutores de 2,5 mm²

.3.1 Com Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 40

Attributes: 11

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Area

Varanda_Gourmet

WC_Servico

Condicionadores_de_Ar

Chuveiros

Interruptores_Simples

Interruptores

Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos

Condutor_Dois_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Dois_Meio =

568.91 * Agua_Quente_na_Cozinha=Sim +

-290.8661 * Agua_Quente_nos_Lavatórios=Sim +

3.6331 * Area +

109.9602 * Varanda_Gourmet=Sim +

-132.0843 * WC_Servico=Sim +

65.6341 * Condicionadores_de_Ar +

97.4203 * Chuveiros +

9.0751 * Interruptores_Simples +

-9.4195 * Interruptores +

5.2422 * Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos +

-162.8374

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.846

Mean absolute error 245.7629

Root mean squared error 306.0955

Relative absolute error 55.2715%

Root relative squared error 53.7124%

Total Number of Instances 40

.3.2 Com Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.007 -M 0.0 -N 500 -V 0 -S
0 -E 20 -H 4

Relation: Banco de Dados Mestrado 12-

Instances: 40

Attributes: 11

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Area

Varanda_Gourmet

WC_Servico

Condicionadores_de_Ar

Chuveiros

Interuptores_Simples

Interuptores

Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos

Condutor_Dois_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 1.1046374811185653
Node 1 -0.6926494109802241
Node 2 -0.9387391992832337
Node 3 -0.27039543988761805
Node 4 -0.16689476503536968

Sigmoid Node 1

Inputs Weights

Threshold 0.09244058765885049

Attrib Agua_Quente_na_Cozinha -0.38502246580411975

Attrib Agua_Quente_nos_Lavatórios -0.07194151766100138

Attrib Area -0.5042199519060752

Attrib Varanda_Gourmet 0.13656057308128927

Attrib WC_Servico 0.09982210325220449

Attrib Condicionadores_de_Ar -0.09186697704467436

Attrib Chuveiros -0.201806857587095

Attrib Interuptores_Simples -0.3532435022754614

Attrib Interuptores -0.39803964930243546

Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos -0.299833415495

Sigmoid Node 2

Inputs Weights

Threshold 0.1974467850233931

Attrib Agua_Quente_na_Cozinha -0.4770741301159636

Attrib Agua_Quente_nos_Lavatórios -0.05126823262835383

Attrib Area -0.573533138069927

Attrib Varanda_Gourmet 0.18400095242828757

Attrib WC_Servico 0.05615769096083703

Attrib Condicionadores_de_Ar -0.07797308770418078

Attrib Chuveiros -0.2702639176478779

Attrib Interuptores_Simples -0.4685767636001361

Attrib Interuptores -0.5376265162144617

Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chuveiros Elétricos -
0.2893497268859466

Sigmoid Node 3

Inputs Weights

Threshold 0.025367268854135115

Attrib Agua_Quente_na_Cozinha -0.11484158960923786

Attrib Agua_Quente_nos_Lavatórios -0.10378977784706155
 Attrib Area -0.2755333961566694
 Attrib Varanda_Gourmet 0.08949177586116515
 Attrib WC_Servico 0.05144309172480046
 Attrib Condicionadores_de_Ar -0.09987100471781714
 Attrib Chuveiros -0.20259519137090637
 Attrib Interruptores_Simples -0.17944458498658122
 Attrib Interruptores -0.26148996709092687
 Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chiveiros Elétricos -0.17324720843012154

Sigmoid Node 4

Inputs Weights

Threshold -0.014864472530888498

Attrib Agua_Quente_na_Cozinha -0.062293424539568376

Attrib Agua_Quente_nos_Lavatórios -0.05101144688034086

Attrib Area -0.19177395251348076

Attrib Varanda_Gourmet 0.11644767303319563

Attrib WC_Servico 0.07690407314560332

Attrib Condicionadores_de_Ar -0.1152654965435095

Attrib Chuveiros -0.14652405888567543

Attrib Interruptores_Simples -0.2032204502700801

Attrib Interruptores -0.2530576650516598

Attrib Luz_Interruptores_Tomada_Condicionadores de Ar_Chiveiros Elétricos -0.1136704839963641

Class

Input

Node 0

Time taken to build model: 0.03 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8987

Mean absolute error 193.386

Root mean squared error 244.281

Relative absolute error 42.8914%

Root relative squared error 42.6161%

Total Number of Instances 40

.3.3 Sem Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 41

Attributes: 11

Sistema_Estrutural

Sistema_de_Condicionamento_de_Ar

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Area

Unidades_por_Pavimentos

Varanda_Gourmet

WC_Servico

Condicionadores_de_Ar

Chuveiros

Condutor_Dois_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Dois_Meio =

-381.0064 * Sistema_Estrutural=Concreto Armado +
389.8728 * Sistema_de_Condicionamento_de_Ar=Janela,Split +
24.2704 * Sistema_de_Condicionamento_de_Ar=Split +
732.4614 * Agua_Quente_na_Cozinha=Sim +
-382.02 * Agua_Quente_nos_Lavatórios=Sim +
2.3504 * Area +
-31.9704 * Unidades_por_Pavimentos +
229.9207 * Varanda_Gourmet=Sim +
-268.5035 * WC_Servico=Sim +
23.1967 * Condicionadores_de_Ar +
274.0938 * Chuveiros +

42.0517

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.857

Mean absolute error 245.2993

Root mean squared error 315.019

Relative absolute error 51.4251%

Root relative squared error 51.9857%

Total Number of Instances 41

.3.4 Sem Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.007 -M 0.0 -N 1400 -V 0 -
S 0 -E 20 -H 3

Relation: Banco de Dados Mestrado 12-

Instances: 41

Attributes: 11

Sistema_Estrutural

Sistema_de_Condicionamento_de_Ar

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Area

Unidades_por_Pavimentos

Varanda_Gourmet

WC_Servico

Condicionadores_de_Ar

Chuveiros

Condutor_Dois_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 1.3416720145425152

Node 1 -0.35653430090828364

Node 2 -1.3621848395902507

Node 3 -0.8185017857485453

Sigmoid Node 1

Inputs Weights

Threshold -0.02397339879708539

Attrib Sistema_Estrutural 0.10756017201518085

Attrib Sistema_de_Condicionamento_de_Ar=Split -0.019747743993351198

Attrib Sistema_de_Condicionamento_de_Ar=Janela -0.15497976194421303

Attrib Sistema_de_Condicionamento_de_Ar=Ausente 0.262327172950214

Attrib Agua_Quente_na_Cozinha -0.18589006176064024

Attrib Agua_Quente_nos_Lavatórios -0.19304744821034248

Attrib Area -0.2644308072094966

Attrib Unidades_por_Pavimentos 0.20102681880237522

Attrib Varanda_Gourmet 0.053448634089784536

Attrib WC_Servico 0.05836153199543895

Attrib Condicionadores_de_Ar -0.13067224808537692

Attrib Chuveiros -0.20700687563618936

Sigmoid Node 2

Inputs Weights

Threshold 0.3952002791077168

Attrib Sistema_Estrutural -0.5309182385271952

Attrib Sistema_de_Condicionamento_de_Ar=Split -0.46062040550355227

Attrib Sistema_de_Condicionamento_de_Ar=Janela 0.04848813907717404

Attrib Sistema_de_Condicionamento_de_Ar=Ausente 0.07600582623720738

Attrib Agua_Quente_na_Cozinha -0.8382609507003079

Attrib Agua_Quente_nos_Lavatórios -0.04619201353973609

Attrib Area -0.8787930625553658

Attrib Unidades_por_Pavimentos 0.10057771313085483

Attrib Varanda_Gourmet 0.4063200142707124

Attrib WC_Servico -0.008506475947552019

Attrib Condicionadores_de_Ar -0.297399800670327

Attrib Chuveiros -0.7213439832228962

Sigmoid Node 3

Inputs Weights

Threshold 0.048642279654398414
 Attrib Sistema_Estrutural 0.019405503914240535
 Attrib Sistema_de_Condicionamento_de_Ar=Split -0.05552683342445248
 Attrib Sistema_de_Condicionamento_de_Ar=Janela -0.33805472495285344
 Attrib Sistema_de_Condicionamento_de_Ar=Ausente 0.33757673398649096
 Attrib Agua_Quente_na_Cozinha -0.22977085293540347
 Attrib Agua_Quente_nos_Lavatórios -0.38008791467312947
 Attrib Area -0.44674220981707186
 Attrib Unididades_por_Pavimentos 0.27950598983998726
 Attrib Varanda_Gourmet -0.08597781254262121
 Attrib WC_Servico -0.09622146702863058
 Attrib Condicionadores_de_Ar -0.10887638714405534
 Attrib Chuveiros -0.43357967714406154

Class

Input

Node 0

Time taken to build model: 0.06 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.843

Mean absolute error 229.8081

Root mean squared error 301.3471

Relative absolute error 51.4005%

Root relative squared error 53.8858%

Total Number of Instances 41

.3.5 Tomadas - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 40

Attributes: 2

Tomadas

Condutor_Dois_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Dois_Meio =

35.844 * Tomadas +
-234.7119

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7964

Mean absolute error 274.1256

Root mean squared error 336.2188

Relative absolute error 60.7302%

Root relative squared error 59.0037%

Total Number of Instances 40

.3.6 Tomadas - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.1 -M 0.05 -N 1500 -V 0 -S
0 -E 20 -H 5

Relation: Banco de Dados Mestrado 12-

Instances: 40

Attributes: 2

Tomadas

Condutor_Dois_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold -0.03526715711846969

Node 1 -0.04034280970650437

Node 2 1.5775008363179517

Node 3 -0.11629712518227099

Node 4 -0.1285942787596329

Node 5 -0.8825145718576859

Sigmoid Node 1

Inputs Weights

Threshold -1.5090632800785324

Attrib Tomadas -0.6638045128861613

Sigmoid Node 2

Inputs Weights

Threshold -2.5873828213069996

Attrib Tomadas 2.1165500710716163

Sigmoid Node 3

Inputs Weights

Threshold -1.6233075108134944

Attrib Tomadas -0.33524905088099827

Sigmoid Node 4

Inputs Weights

Threshold -1.611252719942744

Attrib Tomadas -0.4296271231225327

Sigmoid Node 5

Inputs Weights

Threshold -1.4534602174879634

Attrib Tomadas -3.626737927352451

Class

Input

Node 0

Time taken to build model: 0.05 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7237

Mean absolute error 290.7913

Root mean squared error 384.5298

Relative absolute error 66.1911

Root relative squared error 68.0259

Total Number of Instances 40

.3.7 Área - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 45

Attributes: 2

Area

Condutor_Dois_Meio

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Dois_Meio =

6.9431 * Area +

74.42

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.827

Mean absolute error 246.996

Root mean squared error 303.6553

Relative absolute error 57.0095%
Root relative squared error 55.2651%
Total Number of Instances 45

.3.8 Área - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.008 -M 0.01 -N 800 -V 0 -
S 0 -E 20 -H 9

Relation: Banco de Dados Mestrado 12-
Instances: 45
Attributes: 2
Area
Conductor_Dois_Meio
Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights
Threshold 0.38345665650103206
Node 1 0.6170694117228981
Node 2 -0.3705200584667328
Node 3 -0.5014530321933772
Node 4 -0.9392077432764403
Node 5 -0.17032051698608802
Node 6 -0.803653454499996
Node 7 -0.04167984073449471
Node 8 0.6726165933504153
Node 9 0.7348762460535672

Sigmoid Node 1

Inputs Weights
Threshold 0.02884758764522018
Attrib Area 0.49846724739909726

Sigmoid Node 2

Inputs Weights
Threshold -0.0014821043765229742

Attrib Area -0.4728432945094732

Sigmoid Node 3

Inputs Weights

Threshold -0.0015109894032826407

Attrib Area -0.5934090285387291

Sigmoid Node 4

Inputs Weights

Threshold -0.01495988665469965

Attrib Area -0.9716189771402377

Sigmoid Node 5

Inputs Weights

Threshold 0.013910540871481044

Attrib Area -0.2794479641934454

Sigmoid Node 6

Inputs Weights

Threshold -0.05577292867416493

Attrib Area -0.8562827245892689

Sigmoid Node 7

Inputs Weights

Threshold -0.059786334617698986

Attrib Area -0.14547933008344577

Sigmoid Node 8

Inputs Weights

Threshold 0.05124266537244408

Attrib Area 0.5489148103861419

Sigmoid Node 9

Inputs Weights

Threshold 0.026483680784340904

Attrib Area 0.6078196359222615

Class

Input

Node 0

Time taken to build model: 0.06 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8241

Mean absolute error 251.1887

Root mean squared error 305.9923

Relative absolute error 57.9772

Root relative squared error 55.6905

Total Number of Instances 45

.4 Relatórios do Weka - Condutores de 2,5, 4 e 6 mm²

.4.1 Tomadas + Chuveiros Elétricos + Condicionadores de Ar - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 37

Attributes: 2

Tomadas_Condicionadores_Chuveiros_Eletricos

Condutor_Dois_Meio_Quatro_Seis

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Dois_Meio_Quatro_Seis =

36.7223 * Tomadas_Condicionadores_Chuveiros_Eletricos +
-338.1178

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.8671

Mean absolute error 262.1428

Root mean squared error 316.4533

Relative absolute error 53.0107%

Root relative squared error 49.283%

Total Number of Instances 37

.4.2 Tomadas + Chuveiros Elétricos + Condicionadores de Ar - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.008 -M 0.0 -N 1200 -V 0 -
S 0 -E 20 -H 5

Relation: Banco de Dados Mestrado 12-

Instances: 37

Attributes: 2

Tomadas_Condicionadores_Chuveiros_Eletricos

Condutor_Dois_Meio_Quatro_Seis

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.5105209855735511

Node 1 -0.706234969109549

Node 2 0.8512364966783394

Node 3 -0.12144212544651488

Node 4 0.5121612200021572

Node 5 -1.5551561597480321

Sigmoid Node 1

Inputs Weights

Threshold -0.022208763778259272

Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -0.8469966622991668

Sigmoid Node 2

Inputs Weights

Threshold 0.06562393934322772

Attrib Tomadas_Condicionadores_Chuveiros_Eletricos 0.6695133409934902

Sigmoid Node 3

Inputs Weights

Threshold -0.0313692069044816

Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -0.282644509625366

Sigmoid Node 4

Inputs Weights

Threshold 0.04070333382946831

Attrib Tomadas_Condicionadores_Chuveiros_Eletricos 0.34697816227970885

Sigmoid Node 5

Inputs Weights

Threshold 0.023575157108046255

Attrib Tomadas_Condicionadores_Chuveiros_Eletricos -1.5617058513599464

Class

Input

Node 0

Time taken to build model: 0.03 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.8602

Mean absolute error 266.8204

Root mean squared error 323.9593

Relative absolute error 53.9566%

Root relative squared error 50.4519%

Total Number of Instances 37

.5 Relatórios do Weka - Condutores de 4 e 6 mm²

.5.1 Com Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 33

Attributes: 6

Hall

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Tomadas

Interruptores_Intermediarios

Condutor_Quatro_Seis

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Quatro_Seis =

-7.4566 * Hall=Sim +

-45.204 * Agua_Quente_na_Cozinha=Sim +

79.2508 * Agua_Quente_nos_Lavatórios=Sim +

5.5882 * Tomadas +

1.9099 * Interruptores_Intermediarios +

-75.0426

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7313

Mean absolute error 45.2794

Root mean squared error 61.5847

Relative absolute error 67.6128%

Root relative squared error 66.0201%

Total Number of Instances 33

.5.2 Com Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.008 -M 0.05 -N 500 -V 0 -
S 0 -E 20 -H 3

Relation: Banco de Dados Mestrado 12-

Instances: 33

Attributes: 6

Hall

Agua_Quente_na_Cozinha

Agua_Quente_nos_Lavatórios

Tomadas

Interuptores_Intermediarios

Condutor_Quatro_Seis

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.5535936336611978

Node 1 -1.0122582594455092

Node 2 -0.9218310871935184

Node 3 1.0582625956863876

Sigmoid Node 1

Inputs Weights

Threshold -0.2797032457865429

Attrib Hall -0.12040306022351478

Attrib Agua_Quente_na_Cozinha -0.005253144272335214

Attrib Agua_Quente_nos_Lavatórios 0.49440514185741635

Attrib Tomadas -0.6144410032226836

Attrib Interuptores_Intermediarios -0.5072006808523715

Sigmoid Node 2

Inputs Weights

Threshold -0.25863595968491687
Attrib Hall -0.13781704961050048
Attrib Agua_Quente_na_Cozinha -0.011933624755223046
Attrib Agua_Quente_nos_Lavatórios 0.503084072801331
Attrib Tomadas -0.49007780711670534
Attrib Interruptores_Intermediarios -0.5275805634800714

Sigmoid Node 3

Inputs Weights
Threshold 0.21660050375503345
Attrib Hall 0.130532092990998
Attrib Agua_Quente_na_Cozinha -0.03894324462284908
Attrib Agua_Quente_nos_Lavatórios -0.3934880019337386
Attrib Tomadas 0.48998259144833
Attrib Interruptores_Intermediarios 0.49380850754194305

Class

Input
Node 0

Time taken to build model: 0.01 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7965

Mean absolute error 27.8549
Root mean squared error 35.4842
Relative absolute error 53.871%
Root relative squared error 58.0259%
Total Number of Instances 33

.5.3 Sem Pontos de Utilização - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 1 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-
Instances: 38
Attributes: 5

Hall
 Agua_Quente_na_Cozinha
 Agua_Quente_nos_Lavatórios
 Area
 Condutor_Quatro_Seis
 Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Quatro_Seis =

24.4831 * Hall=Sim +
 87.6746 * Agua_Quente_na_Cozinha=Sim +
 17.0927 * Agua_Quente_nos_Lavatórios=Sim +
 0.7438 * Area +
 5.2452

Time taken to build model: 0 seconds

==== Cross-validation ====
 ==== Summary ====

Correlation coefficient 0.6404
 Mean absolute error 50.5643
 Root mean squared error 70.372
 Relative absolute error 77.3111%
 Root relative squared error 77.2582%
 Total Number of Instances 38

.5.4 Sem Pontos de Utilização - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.005 -M 0.0 -N 500 -V 0 -S
 0 -E 20 -H 3

Relation: Banco de Dados Mestrado 12-
 Instances: 38
 Attributes: 5

Hall
Agua_Quente_na_Cozinha
Agua_Quente_nos_Lavatórios
Area
Condutor_Quatro_Seis
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Linear Node 0

Inputs Weights
Threshold 0.46354351131780885
Node 1 -1.3686946490993008
Node 2 -0.275066631349695
Node 3 0.6490857792733015

Sigmoid Node 1

Inputs Weights
Threshold -0.23737910442155954
Attrib Hall -0.25933236231782325
Attrib Agua_Quente_na_Cozinha -0.36295422893457124
Attrib Agua_Quente_nos_Lavatórios 0.3733820821967811
Attrib Area -1.0961109469862353

Sigmoid Node 2

Inputs Weights
Threshold -0.01147242681059524
Attrib Hall -0.06483632233270131
Attrib Agua_Quente_na_Cozinha -0.1329954337177402
Attrib Agua_Quente_nos_Lavatórios 0.12736248006544607
Attrib Area -0.33721196169227124

Sigmoid Node 3

Inputs Weights
Threshold 0.07950491769117453
Attrib Hall 0.06285500901367222
Attrib Agua_Quente_na_Cozinha 0.2065180719608757
Attrib Agua_Quente_nos_Lavatórios -0.1471511185213155

Attrib Area 0.43764804789504735

Class

Input

Node 0

Time taken to build model: 0.02 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.7067

Mean absolute error 34.6721

Root mean squared error 40.7274

Relative absolute error 68.8305%

Root relative squared error 68.5952%

Total Number of Instances 38

.5.5 Área - Regressão linear

==== Run information ====

Scheme:weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8

Relation: Banco de Dados Mestrado 12-

Instances: 47

Attributes: 2

Area

Condutor_Quatro_Seis

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Regression Model

Condutor_Quatro_Seis =

0.731 * Area +

26.1429

Time taken to build model: 0 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.5966

Mean absolute error 43.565

Root mean squared error 54.159

Relative absolute error 77.6017%

Root relative squared error 77.9601%

Total Number of Instances 47

.5.6 Área - Rede Neural Artificial

==== Run information ====

Scheme:weka.classifiers.functions.MultilayerPerceptron -L 0.015 -M 0.0 -N 500 -V 0 -S
0 -E 20 -H 4

Relation: Banco de Dados Mestrado 12-

Instances: 47

Attributes: 2

Area

Condutor_Quatro_Seis

Test mode:10-fold cross-validation

==== Classifier model (full training set) ====

Linear Node 0

Inputs Weights

Threshold 0.7383901176311449

Node 1 -0.4288561168165125

Node 2 -1.307070895807445

Node 3 0.49532388653066367

Node 4 -0.613239943022913

Sigmoid Node 1

Inputs Weights

Threshold -0.06359269960497364

Attrib Area -0.6175662770661622

Sigmoid Node 2

Inputs Weights

Threshold -0.09008548967286169

Attrib Area -1.4687459409008743

Sigmoid Node 3

Inputs Weights

Threshold 4.045857769834906E-5

Attrib Area 0.31989835014250134

Sigmoid Node 4

Inputs Weights

Threshold -0.07170771955619538

Attrib Area -0.7985391690614911

Class

Input

Node 0

Time taken to build model: 0.02 seconds

==== Cross-validation ====

==== Summary ====

Correlation coefficient 0.6673

Mean absolute error 39.0579

Root mean squared error 47.0866

Relative absolute error 72.4038%

Root relative squared error 72.1255%

Total Number of Instances 47