



Técnicas de inteligência computacional para a previsão de propriedades mecânicas do concreto de agregado leve

Jonata Jefferson Andrade

JUIZ DE FORA

MARÇO, 2016

Técnicas de inteligência computacional para a previsão de propriedades mecânicas do concreto de agregado leve

JONATA JEFFERSON ANDRADE

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação
Orientador: Leonardo Goliatt da Fonseca

JUIZ DE FORA
MARÇO, 2016

TÉCNICAS DE INTELIGÊNCIA COMPUTACIONAL PARA A
PREVISÃO DE PROPRIEDADES MECÂNICAS DO CONCRETO DE
AGREGADO LEVE

Jonata Jefferson Andrade

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS
EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTEGRANTE
DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM
CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Leonardo Goliatt da Fonseca
Doutor em Modelagem Computacional

Luciana Conceição Dias Campos
Doutora em Engenharia Elétrica

Michèle Cristina Resende Farage
Doutora em Engenharia Civil

JUIZ DE FORA
04 DE MARÇO, 2016

Aos meus amigos e irmãos.

Aos pais, pelo apoio e sustento.

Resumo

No concreto de agregado leve, a resistência à compressão e o peso específico são as propriedades mais importantes e conseqüentemente as mais comumente analisadas. A maioria das normas apresentam relações entre a densidade do agregado leve e sua resistência à compressão. O módulo de elasticidade (módulo de Young) também possui grande importância para o concreto estrutural leve dada a sua influência sobre as deformações das peças. A relação entre os componentes do concreto e suas propriedades mecânicas é altamente não-linear, e o estabelecimento de um modelo matemático abrangente é usualmente problemático. Dessa forma, modelos mecanicistas das propriedades do concreto podem não ser suficientes para atender às exigências do projeto de concretos de agregados leves. Assim, prever com precisão as propriedades mecânicas do concreto é um problema crítico em construções de concreto de agregado leve. Até o presente momento, considerando o cenário de pesquisa de concretos de agregados leves, faz-se necessário avaliar o desempenho de diferentes técnicas preditivas além das redes neurais artificiais na previsão das propriedades do concreto de agregado leve. O objetivo deste trabalho é avaliar o desempenho de métodos de inteligência computacional, tais como, Redes Neurais Artificiais, Máquinas de Vetores Suporte, Máquinas de Aprendizado Extremo e Árvores de Decisão, para prever o módulo de Young e a resistência à compressão de concretos de agregado leve, de forma a obter um valor próximo ao real. Depois de uma verificação da capacidade de generalização dos métodos, através do processo de validação cruzada, observou-se que as Máquinas de Vetores Suporte apresentaram o melhor desempenho na aproximação das duas propriedades mecânicas.

Palavras-chave: Concreto de agregados leves, Aprendizado de máquina, Inteligência computacional.

Abstract

In the lightweight aggregate concrete, the compressive strength and density are the most important properties and, consequently, the most commonly analyzed. Most standards have relations between the density of lightweight concrete and its compressive strength. The modulus of elasticity (Young's modulus) also has great importance for the lightweight structural concrete given its influence on the deformation of the parts. The relationship between the specific components and their mechanical properties is highly nonlinear, and the establishment of a comprehensive mathematical model is usually problematic. Therefore, mechanistic models of concrete properties may not be sufficient to meet the requirements of lightweight aggregate concrete project. Thus accurately predict the mechanical properties of concrete is a critical problem in lightweight aggregate concrete buildings. Until now, considering the concrete research landscape of lightweight aggregates, it is necessary to evaluate the performance of different predictive techniques beyond the artificial neural network in the prediction of lightweight concrete properties. The objective of this study is to evaluate the performance of computational intelligence methods, namely, Artificial Neural Networks, Vector Support Machines, Extreme Learning Machines and Decision Trees to predict the Young's modulus and the compressive strength of lightweight aggregate concrete in order to obtain a value close to real. After verification of the generalizability of the methods, using the cross-validation process, it was observed that the Support Vector Machines showed the best performance in prediction the two mechanical properties.

Keywords: Lightweight aggregate concrete, Machine learning, Computational intelligence.

Agradecimentos

Em primeiro lugar, quero agradecer a Deus por ter me dado saúde e força para superar todas as dificuldades.

Aos meus pais, Carlos e Ângela, e meus irmãos, Tiago, Carlos e Otávio, pelo carinho, encorajamento e apoio. A minha querida namorada Fabiane que sempre me incentivou e apoiou em toda a minha vida acadêmica.

Ao professor Leonardo Goliatt pela orientação, amizade e principalmente, pela paciência, sem a qual este trabalho não se realizaria.

Aos meus amigos do curso pela amizade e companheirismo durante todos os momentos.

Aos professores dos Departamentos de Ciência da Computação e Mecânica Aplicada e Computacional pelos seus ensinamentos e aos funcionários do curso, que durante esses anos, contribuíram de algum modo para o nosso enriquecimento pessoal e profissional.

“A vida é como andar de bicicleta. Para manter o equilíbrio é preciso se manter em movimento.”

Albert Einstein

Sumário

Lista de Figuras	7
Lista de Tabelas	8
Lista de Abreviações	9
1 Introdução	10
1.1 Apresentação do tema	10
1.2 Justificativa	10
1.3 Objetivo	11
1.4 Organização do trabalho	11
2 Concreto leve	12
2.1 Características	12
2.2 Propriedades mecânicas	13
3 Previsão de propriedades do concreto	16
3.1 Redes Neurais	16
3.2 Máquinas De Aprendizado Extremo	18
3.3 Árvores de Decisão	19
3.4 Máquinas de Vetores Suporte	21
4 Generalização dos modelos de predição	23
4.1 Validação cruzada	23
4.2 Busca exaustiva	23
4.3 Métricas de avaliação	24
5 Experimentos e resultados	26
5.1 Base de dados	26
5.2 Código computacional	26
5.3 Resultados	27
6 Conclusões e trabalhos futuros	33

Lista de Figuras

2.1	Transferência das foças de compressão no concreto normal e leve (Bogas and Gomes, 2013).	14
3.1	À esquerda, um exemplo de uma rede com 4 variáveis de entrada, uma camada escondida com 5 neurônios e um neurônio na camada de saída. À direita o esquema de um único neurônio.	17
3.2	Modelos de MLP unidirecional simples (esquerda) e unidirecional totalmente conectada (direta)	17
3.3	À esquerda um exemplo da uma árvore gerada pelo DT e à direita uma perspectiva da superfície de regressão gerada por esse modelo, para entradas na forma $X = (X_1, X_2)$ e os pontos de divisão t_1, t_2, t_3, t_4 , com regiões R_1, R_2, R_3, R_4, R_5 (Hastie et al., 2009).	20
3.4	Exemplo dos parâmetros do SVR para uma entrada x e saída y (Drucker et al., 1997).	22
5.1	Esquema do algoritmo implementado.	27
5.2	Boxplots do MPE dos modelos para a Resistência à compressão	28
5.3	Boxplots do MPE dos modelos para o Módulo de Young	29
5.4	Previsão da Resistência à compressão pelo DT com desempenho mediano	30
5.5	Previsão do Módulo de Young pelo DT com desempenho mediano	30
5.6	Previsão da Resistência à compressão pelo ELM com desempenho mediano	31
5.7	Previsão do Módulo de Young pelo ELM com desempenho mediano	31
5.8	Previsão da Resistência à compressão pelo MLP com desempenho mediano	31
5.9	Previsão do Módulo de Young pelo MLP com desempenho mediano	32
5.10	Previsão da Resistência à compressão pelo SVR com desempenho mediano.	32
5.11	Previsão do Módulo de Young pelo SVR com desempenho mediano	32

Lista de Tabelas

2.1	Classificação dos concretos leves quanto à sua massa específica aproximada. . .	12
3.1	Funções de mapeamento comumente usados no ELM.	18
4.1	Parâmetros dos modelos usados na busca exaustiva com a validação cruzada. .	24
5.1	MPE, MSE, R^2 e seus respectivos desvios referentes a cada um dos modelos de predição para Resistência à compressão	28
5.2	MPE, MSE, R^2 e seus respectivos desvios referentes a cada um dos modelos de predição para o Módulo de Young	28
5.3	Melhores parâmetros encontrados pela busca exaustiva com a validação cruzada para cada modelo na previsão das duas propriedades	29

Lista de Abreviações

DT	Decision Trees
ELM	Extreme Learning Machine
MLP	Multilayer perceptron
RNA	Rede Neural Artificial
SVM	Support Vector Machines
SVR	Support Vector regression
MPE	Mean Percentual Error
MSE	Mean Squared Error

1 Introdução

1.1 Apresentação do tema

Nas últimas décadas, o concreto armado convencional, de peso específico normal, tem sido o material estrutural mais empregado no Brasil. A razão deste fato está associada às seguintes vantagens: (a) adaptação a qualquer tipo de forma permitindo liberdade à concepção arquitetônica; (b) solução para se obter estruturas monolíticas e hiperestáticas (c) grande durabilidade e baixíssimo custo de manutenção e conservação (d) resistência a efeitos térmicos, atmosféricos e a desgastes mecânicos. Por outro lado, a grande desvantagem do concreto armado convencional é o seu elevado peso próprio.

Nesse contexto, o uso de concretos de menor peso específico surge como a solução alternativa para esse problema. Além disso, a baixa condutividade térmica do concreto estrutural leve proporciona um desempenho significativamente superior em situação de incêndio quando comparado aos concretos de peso específico normal (Neville, 2015). A combinação de uma baixa condutividade térmica com um baixo coeficiente de dilatação térmica é benéfica no caso de exposição ao fogo, pois os agregados leves em geral são estáveis a temperaturas elevadas, uma vez que usualmente são fabricados em temperaturas superiores a $1100^{\circ}C$.

1.2 Justificativa

No concreto estrutural leve, a resistência à compressão, o módulo de elasticidade e a massa específica são as propriedades mais importantes e conseqüentemente as mais comumente analisadas. Prever com precisão essas propriedades é um problema crítico em construções de concreto de agregado leve. Até o presente momento, considerando o cenário de pesquisa de concretos de agregados leves, é preciso explorar diferentes técnicas preditivas, além das redes neurais artificiais (Alshihri et al., 2009), na previsão das propriedades do concreto de agregado leve, tais como Máquinas de Vetores Suporte, Máquinas de Aprendizado Extremo e Árvores de Decisão. O desenvolvimento da metodologia proposta neste trabalho, portanto, vem complementar os

estudos encontrados atualmente na literatura.

1.3 Objetivo

O objetivo deste trabalho é implementar e avaliar o desempenho de métodos de inteligência computacional para prever duas importantes propriedades mecânicas de concretos de agregado leve, o módulo de elasticidade e a resistência à compressão, de forma a obter um valor próximo ao medido em laboratório, conseguindo dessa forma reduzir a quantidades de ensaios laboratoriais que demanda tempo e também o retrabalho.

1.4 Organização do trabalho

O texto deste trabalho está organizado como segue nesta seção. O capítulo 2, apresenta o concreto de agregado leve, citando suas características, problemas e utilização e, em seguida, o capítulo 3 exhibe os métodos utilizados para prever as propriedades do concreto. No capítulo 4 é descrita a metodologia utilizada no trabalho, com os métodos e métricas utilizadas para calcular o poder de generalização dos modelos de predição. O capítulo 5 descreve como foi utilizada a base de dados, o código computacional e os resultados obtidos, e por fim, no capítulo 6 a conclusão do trabalho é apresentada.

2 Concreto leve

2.1 Características

O concreto feito com agregados leves, com baixo peso específico, é chamado de concreto de agregado leve. O concreto de agregado leve possui propriedades refratárias de isolamento térmico e seu peso específico é aproximadamente dois terços do peso específico do concreto feito com agregado normal. Nos Estados Unidos (ACI, 1999), por exemplo, o concreto leve estrutural é definido como o material que tem resistência à compressão superior a 17 MPa aos 28 dias e massa específica não superior a 1850 kg/m^3 . No Brasil, a norma NBR NM 35 (ABNT, 1995) estabelece as seguintes relações entre a resistência à compressão e massa específica pra concretos leves estruturais: (a) resistência à compressão superior a 28 MPa aos 28 dias e massa específica não superior a 1840 kg/m^3 ; (b) resistência à compressão superior a 21 MPa aos 28 dias e massa específica não superior a 1760 kg/m^3 ; (c) resistência à compressão superior a 17 MPa aos 28 dias e massa específica não superior a 1680 kg/m^3 .

A massa específica dos concretos de peso específico normal varia entre 2200 kg/m^3 e 2600 kg/m^3 , e a do concreto estrutural leve entre 1350 kg/m^3 e 1850 kg/m^3 (Tabela 2.1). Com isso, o uso do concreto estrutural leve pode representar uma redução significativa do peso próprio da estrutura. O custo de produção do concreto de agregado leve é maior que o concreto de peso específico normal, porém, mesmo com este investimento adicional o custo final da estrutura pode ser menor devido ao dimensionamento mais econômico das fundações.

Tabela 2.1: Classificação dos concretos leves quanto à sua massa específica aproximada.

Classificação	Peso específico (kg/m^3)	Agregado utilizado
Concreto de baixa massa específica	300 a 800	Vermiculita e Perlita
Concreto de moderada resistência	800 a 1350	Perlita, Pedra-pomes e Lava porosa
Concreto estrutural	1350 a 1850	Argila, folhelho, ardósia expandida em fornos rotativos, cinza volante sinterizada em grelhas, Argila expandida

2.2 Propriedades mecânicas

Em geral, para um determinado agregado leve, há uma relação entre o quantidade de argamassa do concreto e a resistência à compressão. Para um dado agregado, a resistência à compressão do concreto aumenta com o aumento de seu peso específico. Neste sentido, a maioria das normas e especificações apresenta relações entre a massa específica do agregado leve e sua resistência à compressão (Neville, 2015).

O módulo de elasticidade (módulo de Young) tem grande importância para o concreto estrutural leve dada a sua influência sobre as deformações das peças sujeitas à flexão, sobre a distribuição das forças internas e sobre a carga crítica no caso de peças sujeitas a flambagem. Em concretos de peso específico usual, o módulo de elasticidade da pasta de cimento é, geralmente, bem menor que o módulo dos agregados. Por outro lado, no concreto de agregado leve, os valores do módulo de elasticidade das partículas do agregado leve e da pasta de cimento são bastante próximos.

Ensaio de laboratório comprovam que, para um mesmo nível de resistência à compressão, o valor do módulo de elasticidade estático do concreto leve é consideravelmente menor que o valor obtido para um concreto de peso específico convencional. Essa diferença se deve ao menor valor do módulo de elasticidade do agregado leve em relação ao agregado convencional. Portanto, as deformações de uma estrutura fabricada com concreto leve serão maiores se comparadas a estruturas construídas com concreto de peso específico usual. Esse efeito representa uma desvantagem do uso do concreto leve e deve ser levado em conta em um projeto (da Silva, 2013).

A Fig. 2.1 apresenta uma comparação entre concreto de peso específico normal e o concreto de agregado leve quando estão sendo solicitados a uma força de compressão. É possível notar que no caso do concreto de agregado leve as linhas de forças se distribuem ao longo da argamassa, dado que o mesmo possui maior peso específico.

A relação entre os componentes do concreto e suas propriedades mecânicas é altamente não-linear, e o estabelecimento de um modelo matemático abrangente é usualmente problemático. O concreto de agregado leve possui vários ingredientes, tais como o tipo e a quantidade de agregado leve, além de aditivos e da composição da argamassa. Esses vários ingredientes, além da não linearidade das estruturas de concreto, podem complicar o cálculo

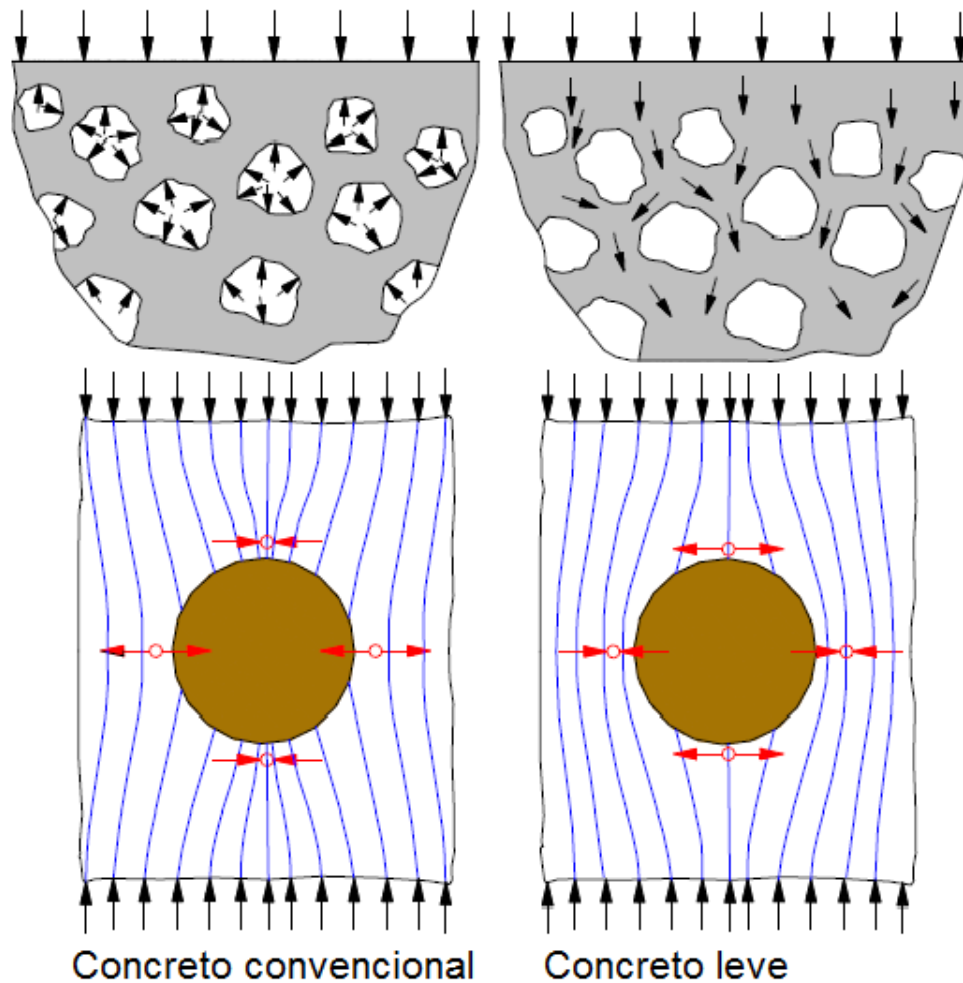


Figura 2.1: Transferência das forças de compressão no concreto normal e leve (Bogas and Gomes, 2013).

da resistência à compressão, do módulo de Young, entre outras propriedades. Dessa forma, modelos mecanicistas (analíticos) de propriedades do concreto podem não ser suficientes para atender as exigências do projeto de concretos de agregados leves.

Vários estudos têm mostrado que propriedades como a resistência à compressão e o módulo de Young são determinadas, independentemente do tipo de concreto, não só pelo fator água/cimento, mas também por outros materiais utilizados na mistura (Chandra and Berntsson, 2003; Kockal and Ozturan, 2011; Faust, 2000). As equações empíricas apresentadas para a estimativa da resistência à compressão e do módulo de Young são baseados em testes de concreto confeccionado com materiais que em geral estão disponíveis no mercado. Dado que o concreto de agregado leve pode ser produzido de vários tipos agregados, que muitas vezes não estão disponíveis no mercado, a validade dessas relações, para o concreto com agregados leves é, portanto, passível de discussão.

Estruturas que utilizam o concreto de agregado leve como matéria prima precisam ser projetadas com uma alta precisão nos dias de hoje com o objetivo de evitar desperdício de material e ao mesmo tempo aumentar a sua resistência aos esforços externos. Calcular as propriedades mecânicas, como o módulo de elasticidade e a resistência a compressão, se torna uma tarefa importante, mas também complicada para os engenheiros, dada a não linearidade da relação entre os componentes que constituem o concreto e as suas propriedades. Conseguir de alguma maneira prever os valores das propriedades do concreto com baixo custo e alta confiabilidade é um objetivo em comum dos pesquisadores da área.

3 Previsão de propriedades do concreto

Muitos autores já se empenharam na pesquisa de vários métodos de inteligência computacional para a previsão das propriedades mecânicas do concreto de peso específico normal (Papadakis and Tsimas, 2002; Ni and Wang, 2000; Kasperkiewicz et al., 1995; Duan et al., 2013). Os modelos de previsão conseguem uma boa aproximação das propriedades do concreto na maioria dos casos, fazendo com que diminua o trabalho em laboratório, o que demanda muito tempo e um alto custo.

Assim como nesses trabalhos para concreto de peso específico normal, além das redes neurais, outras técnicas preditivas contidas na literatura, tais como Máquinas de Vetores Suporte, Máquinas de Aprendizado Extremo e Árvores de Decisão também podem ser utilizadas para a previsão das propriedades do concreto de agregado leve. Uma abordagem e análise criteriosa de cada técnica nos conjuntos de dados podem indicar qual técnica possui a melhor generalização do problema real, o que é feito neste trabalho. Cada uma das técnicas que são utilizadas no trabalho é descrita a seguir.

3.1 Redes Neurais

Uma Rede Neural Artificial (RNA) pode ser caracterizada como um modelo computacional com a habilidade de adaptação ou aprendizado, para generalizar, agrupar ou organizar dados (Kröse and van der Smagt, 1993). São inspiradas nas redes neurais do cérebro dos animais, na qual consiste em um conjunto de simples unidades de processamento, chamadas de neurônios, arranjados em múltiplas camadas, que se comunicam pela troca de mensagens através de suas conexões sinápticas.

Os principais componentes do modelo são: (1) um conjunto de unidades de processamento; (2) um estado de ativação z_k para cada unidade, que é equivalente à saída da unidade; (3) conexões entre as unidades, geralmente definidas por pesos sinápticos w_{jk} , e (4) uma função de ativação φ , que determina o novo nível de ativação de cada unidade de processamento $z_k = \varphi(s_k)$ baseado na saída efetiva $s_k = \sum w_{kj}x_j$ desta unidade. A Fig. 3.1 mostra uma rede

neuronal com os componentes citados acima.

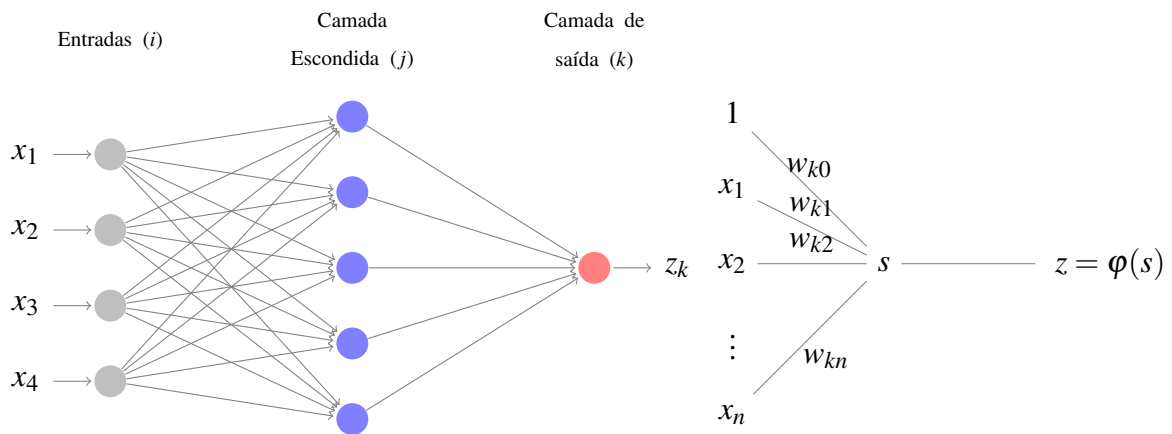


Figura 3.1: À esquerda, um exemplo de uma rede com 4 variáveis de entrada, uma camada escondida com 5 neurônios e um neurônio na camada de saída. À direita o esquema de um único neurônio.

A rede neural Perceptron multicamadas (MLP, do inglês Multilayer perceptron) é um dos modelos de RNA comumente utilizados para estudos de predição e será adotado nesse trabalho. A estrutura da rede MLP é constituída por uma camada de entrada, uma camada de saída, e pelo menos uma camada interna, composta por neurônios internos (Kolay and Baser, 2014).

As topologias da rede MLP que serão utilizadas nesse trabalho serão a unidirecional simples (feed-forward) e a unidirecional totalmente conectada (fully connected). A diferença entre as topologias podem ser observadas na Fig. 3.2.

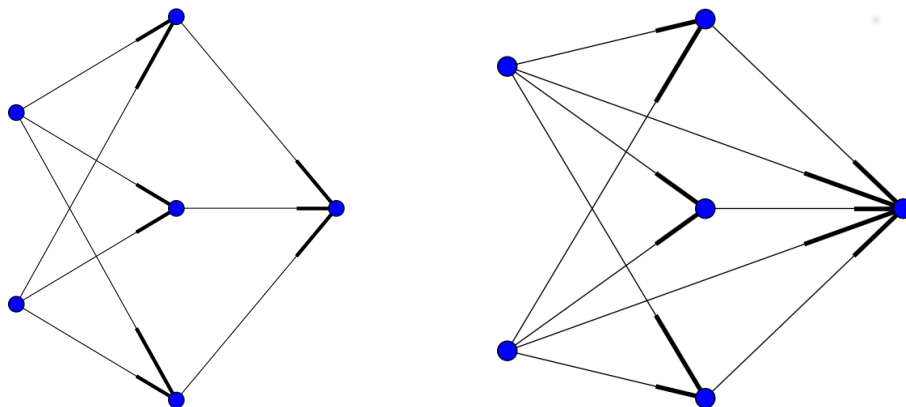


Figura 3.2: Modelos de MLP unidirecional simples (esquerda) e unidirecional totalmente conectada (direita)

Neste trabalho, foi usada a função de ativação sigmoidal, dada pela expressão $\varphi =$

$1/(1 + \exp(x))$. A rede neural é treinada ajustando-se os pesos sinápticos w_{kj} com o algoritmo de Newton Truncado (TNC) (Nocedal and Wright, 2006). O algoritmo de treinamento TNC foi escolhido neste trabalho pois gera melhores resultados do que outros algoritmos tradicionais de treinamento, como o *Backpropagation* (Radonjic et al., 2012).

3.2 Máquinas De Aprendizado Extremo

A Máquina de Aprendizado Extremo (ELM, Extreme Learning Machine) (bin Huang et al., 2006) é um caso especial das redes neurais artificiais, na qual possui uma única camada interna. No processo de aprendizagem do algoritmo Extreme Learning Machine, os pesos e os bias da camada interna são atribuídos aleatoriamente, e os pesos de saída são determinados analiticamente.

A função de saída da ELM é dado por

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^L w_i h_i(\mathbf{x}) = \mathbf{h}(\mathbf{x})\mathbf{w} \quad (3.1)$$

onde $\mathbf{x} = [x_1, \dots, x_n]^T$, $\mathbf{w} = [w_1, \dots, w_L]^T$ é o vetor de pesos de saída entre a camada interna dos L nós para os nós de saída, e $h_i(\mathbf{x})$ é a saída dos i -ésimo nó interno. O vetor de saída (linha) da camada interna em relação a entrada \mathbf{x} é chamado de mapeamento de características não-linear e é escrito como $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), \dots, h_L(\mathbf{x})]$, onde $h_i = G(\mathbf{a}_i, b_i, \mathbf{x})$ é a saída do i -ésimo nó interno com $\mathbf{a}_i \in \mathfrak{R}^d$ e $b_i > 0$. Funções de ativação comuns $G(\mathbf{a}, b, \mathbf{x})$ com pesos dos nós internos (\mathbf{a}, b) são apresentadas na Tabela 3.1.

Tabela 3.1: Funções de mapeamento comumente usados no ELM.

Sigmóide	$G(\mathbf{a}, b, \mathbf{x}) = \frac{1}{1 + \exp(\mathbf{a} \cdot \mathbf{x} + b)}$
Tangente Hiperbolica	$G(\mathbf{a}, b, \mathbf{x}) = \frac{1 - \exp(\mathbf{a} \cdot \mathbf{x} + b)}{1 + \exp(\mathbf{a} \cdot \mathbf{x} + b)}$
Gaussiana	$G(\mathbf{a}, b, \mathbf{x}) = \exp(-(\mathbf{a} \cdot \mathbf{x} + b)^2)$
Multiquádrica	$G(\mathbf{a}, b, \mathbf{x}) = (\ \mathbf{a} - \mathbf{x}\ + b^2)^{1/2}$
Multiquádrica inversa	$G(\mathbf{a}, b, \mathbf{x}) = 1/(\ \mathbf{a} - \mathbf{x}\ + b^2)^{1/2}$

Como, no ELM, os parâmetros (\mathbf{a}, b) são gerados aleatoriamente, em vez de ser explicitamente treinados, isso leva a uma eficiência notável em comparação com redes neurais *backpropagation* tradicionais (Huang et al., 2015).

Os pesos que ligam a camada interna e a camada de saída, denotado por \mathbf{w} , são encontrados através da resolução do seguinte sistema

$$\begin{bmatrix} G(\mathbf{a}_1, b_1, \mathbf{x}_1) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_1) \\ \vdots & \vdots & \vdots \\ G(\mathbf{a}_1, b_1, \mathbf{x}_N) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_N) \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad (3.2)$$

ou forma compacta $\mathbf{H}\mathbf{w} = \mathbf{y}$, onde \mathbf{y} é o dado de treinamento. A solução ótima é dada por

$$\mathbf{w} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} = \mathbf{H}^\dagger \mathbf{y} \quad (3.3)$$

onde \mathbf{H}^\dagger denota a matriz generalizada inversa de Moore-Penrose da matriz \mathbf{H} .

O ELM é um algoritmo de treinamento rápido que fornece baixos erros de treinamento, assim como boa capacidade de generalização.

3.3 Árvores de Decisão

Árvores de decisão (DT, do inglês Decision Trees) (Russell and Norvig, 2003) são ferramentas computacionais usadas na implementação de sistemas especialistas e em problemas de classificação e regressão. As árvores de decisão tomam como entrada uma situação descrita por um conjunto de atributos e retorna uma decisão, que é o valor predito para o valor de entrada. Os atributos de entrada podem ser discretos ou contínuos. A árvore de decisão executa uma sequência de testes e no final chega a uma decisão. Cada nó interno da árvore corresponde a um teste do valor de uma das propriedades de entrada, e os ramos deste nó são identificados com os possíveis valores do teste. Cada nó folha da árvore especifica o valor de retorno para um determinado dado de entrada.

Para construir a árvore o algoritmo precisa decidir automaticamente quais são as variáveis e os pontos de divisão, e também qual a topologia (forma) que a árvore deverá ter (Hastie et al., 2009). Com os dados consistindo de p variáveis e uma saída, para cada entrada, ou seja, (x_i, y_i) $i = 1, 2, \dots, N$, com $x_i = (v_{i1}, v_{i2}, \dots, v_{ip})$, suponha primeiramente que temos um partição

em K regiões R_1, R_2, \dots, R_K , e a saída é dada como uma constante c_m em cada região:

$$\hat{y}(x) = \sum_{m=1}^K \hat{c}_m I(x \in R_m) \quad (3.4)$$

Adotando como critério de minimização a soma dos quadrados $(y_i - f(x_i))^2$, verifica-se que o melhor c_m é apenas a média de y_i na região R_m (Fig. 3.3):

$$\hat{c}_m = \text{média}(x \in R_m) \quad (3.5)$$

Então buscamos a variável de divisão j e o ponto de divisão s que resolva:

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right] \quad (3.6)$$

Para qualquer escolha de j e s , a minimização interna é resolvida por:

$$\hat{c}_1 = \text{média}(y_i | x_i \in R_1(j,s)) \text{ e } \hat{c}_2 = \text{média}(y_i | x_i \in R_2(j,s)) \quad (3.7)$$

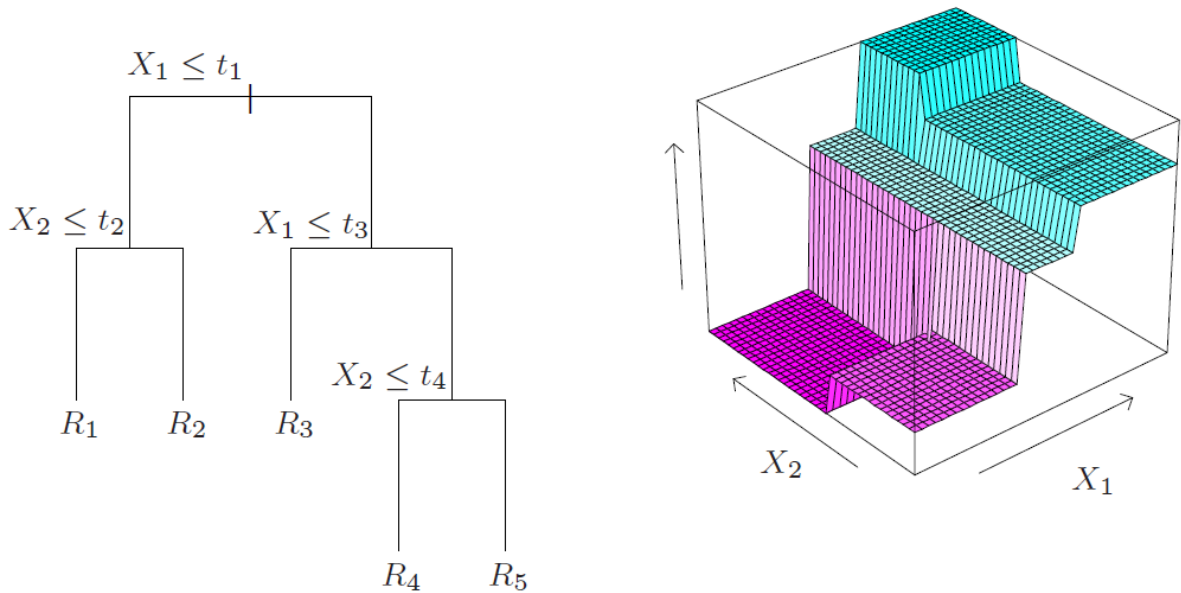


Figura 3.3: À esquerda um exemplo de uma árvore gerada pelo DT e à direita uma perspectiva da superfície de regressão gerada por esse modelo, para entradas na forma $X = (X_1, X_2)$ e os pontos de divisão t_1, t_2, t_3, t_4 , com regiões R_1, R_2, R_3, R_4, R_5 (Hastie et al., 2009).

3.4 Máquinas de Vetores Suporte

As Máquinas de Vetores Suporte (SVM, do inglês Support Vector Machines) (Vapnik and Kotz, 2006) consistem em um algoritmo que provê o estado da arte para um amplo domínio de aplicações, como reconhecimento de escrita, reconhecimento de objetos, identificação de fala, detecção de face e categorização de texto. O SVM é baseado na teoria do aprendizado estatístico, e consiste na minimização do risco estrutural associado ao erro de generalização, que é uma minimização dos erros sobre a base de dados.

A versão do SVM para análise de regressão é chamada de Vetores Suporte para Regressão (SVR, do inglês Support Vector regression). SVR mapeia os vetores de entrada $x = (x_1, \dots, x_n)$ em um espaço de maior dimensão onde a máquina linear é construída pela minimização de um funcional regularizado. A máquina linear pode ser escrita como

$$\hat{y}(x) = \sum_{j=1}^m w_j \phi_j(x) + b \quad (3.8)$$

onde \hat{y} é o valor aproximado de y em x , w representa um vetor de pesos, b é o bias e $\phi_j(x)$ é um conjunto de transformações não lineares.

SVR executa uma regressão linear utilizando uma ε -insensível função de perda

$$L(x) = \begin{cases} 0 & \text{se } |y(x) - \hat{y}(x)| \leq \varepsilon \\ |y(x) - \hat{y}(x)| & \text{caso contrário} \end{cases} \quad (3.9)$$

enquanto tenta se reduzir a complexidade do modelo, minimizando a norma do vetor de ponderação, $\|w\|_2$. Incluindo as variáveis de folga não negativas ξ_i e ξ_i^* , $i = 1, \dots, n$ o problema de otimização pode ser formulado como

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|_2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{sujeito a} \quad & \begin{cases} y_i - \hat{y}_i \leq \varepsilon + \xi_i^* \\ \hat{y}_i - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n \end{cases} \end{aligned} \quad (3.10)$$

O parâmetro de regularização, C , determina o trade-off entre a complexidade do mo-

delo e do grau para o qual o maior desvio de ε é tolerado.

Na figura 3.4 podemos observar um exemplo dos parâmetros do SVR para uma dada entrada x e os seus respectivos valores observados e preditos y .

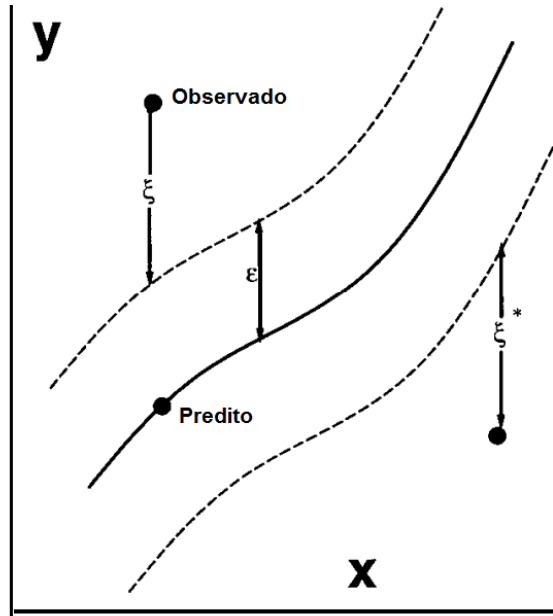


Figura 3.4: Exemplo dos parâmetros do SVR para uma entrada x e saída y (Drucker et al., 1997).

A regressão não linear pode ser realizada utilizando uma função kernel, K (Cherkassky and Ma, 2004). Neste trabalho será utilizada a função de base radial (RBF, do inglês Radial Basis Function).

O problema de otimização pode ser transformado no problema dual

$$\hat{y}(x) = \sum_{i=1}^{n_{SV}} (\alpha_i - \alpha_i^*) K(x_i, x) + b, \text{ sujeito a } 0 \geq \alpha_i \geq C, 0 \geq \alpha_i^* \geq C \quad (3.11)$$

onde n_{SV} é número de vetores suportes e $K(x_i, x) = \sum_{i=1}^m \phi_i(x_i) \phi_i(x)$.

4 Generalização dos modelos de predição

4.1 Validação cruzada

Para testar o poder de generalização de uma técnica de predição, faz-se necessário o uso de uma técnica estatística de amostragem. O método de validação cruzada com k -fold (Hastie et al., 2009) realiza esta tarefa, de modo a avaliar o desempenho do modelo de predição para um novo conjunto de dados com as mesmas características. Para realizar a validação cruzada é preciso que o conjunto de treinamento original seja dividido de forma aleatória em k subconjuntos. Para cada base de dados, dos k subconjuntos obtidos na divisão, um deles é separado para a validação do modelo já treinado nos $k - 1$ subconjuntos restantes. O processo é então repetido k vezes para o modelo treinado, de modo que todos os k subconjuntos sejam utilizados exatamente uma vez como dado de teste para validação do modelo.

Utilizar o parâmetro k adequado é um dos desafios existentes no método, dado que um valor muito alto para k indica um conjunto de teste pequeno, que muitas vezes podem não conter as mesmas características da base original. Na literatura o valor de k é geralmente escolhido ente 5 e 10. Neste trabalho, após testes preliminares, foi escolhido $k = 8$.

4.2 Busca exaustiva

Como cada modelo possuem diversos parâmetros a serem escolhidos e é preciso realizar uma comparação entre os métodos, foi utilizado a técnica de busca exaustiva (*Grid Search*) (Pedregosa et al., 2011) para que possa ser encontrado os parâmetros ótimos de cada modelo de predição. A Tabela 4.1 apresenta a variação dos parâmetros utilizados no trabalho para cada modelo.

Tabela 4.1: Parâmetros dos modelos usados na busca exaustiva com a validação cruzada.

Modelo	Parâmetros	Variação
DT	Profundidade máxima	sem restrição, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 50
ELM	Função de ativação	sigmóide, tangente hiperbólica, gaussiana, multiquádrica, multiquádrica inversa
	Neurônios ocultos	5, 10, 20, 30, 50, 100, 150, 200
	RBF - largura	0, 0.1, 0.2, ..., 0.9, 1
MLP	Topologia	unidirecional simples, unidirecional totalmente conectada
	Camadas ocultas	uma camada com 5 ou 10 neurônios, duas camadas com 5 ou 10 neurônios cada uma e três camadas com 5 ou 10 neurônios cada uma.
SVR	Precisão da regressão (<i>epsilon</i> - ϵ)	10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6}
	Parâmetro de regularização (C)	1, 10, 10^2 , 10^3 , 10^4 , 10^5 , 10^6
	Grau polinomial	1, 2, 3, 4, 5
	Tipo do Kernel	RBF
	RBF - <i>gamma</i>	1, 0.1, 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6}

4.3 Métricas de avaliação

De modo a avaliar o desempenho de previsão dos modelos, foram utilizadas neste trabalho três métricas de avaliação: o coeficiente de determinação R^2 , o erro percentual médio (MPE, do inglês Mean Percentual Error) e o erro quadrático médio (MSE, do inglês Mean Squared Error).

Sendo \hat{y} é a saída estimada, y é a saída correta, \bar{y} é a média das saídas corretas e N o número de amostras, podemos definir as métricas como a seguir:

O coeficiente de determinação R^2 é escrito como

$$R^2(y, \hat{y}) = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (4.1)$$

sendo que o valor do R^2 varia entre 0 e 1, e quanto mais próximo de 1, mais representativo é o modelo;

O erro médio percentual (MPE) é definido como

$$\text{MPE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{|y_i|} \times 100 \quad (4.2)$$

O erro quadrático médio (MSE) é dado por

$$\text{MSE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (4.3)$$

Para realizar a busca exaustiva dos parâmetros é necessário utilizar como forma de medida uma métrica de avaliação, para isso, foi utilizado o erro quadrático médio (MSE).

De forma a garantir uma maior confiabilidade das métricas utilizadas, o processo de validação foi repetido 30 vezes para cada modelo e os valores finais de cada métrica para cada modelo serão dados pela média das 30 repetições.

5 Experimentos e resultados

5.1 Base de dados

A fim de avaliar o desempenho dos métodos de inteligência computacional na previsão de propriedades mecânicas de concretos de agregado leve, foram empregados dados de ensaios laboratoriais de concretos com composições variadas. Os dados experimentais foram obtidos em um extenso programa experimental na Universidade de Cergy-Pontoise (Ke, 2008). A base de dados possui 75 amostras de teste com quatro parâmetros e duas saídas. As propriedades dos concretos de agregado leve coletadas foram o Módulo de Young (módulo de elasticidade) e a resistência à compressão aos 28 dias, em função do fator água/cimento, volume de agregado leve, quantidade de cimento e o peso específico do agregado leve. A variação das propriedades se deu da seguinte forma: fator água/cimento entre 2.90×10^{-1} e 4.46×10^{-1} , quantidade de cimento entre $4.15 \times 10^2 (kg/m^3)$ e $8.26 \times 10^2 (kg/m^3)$, volume do agregado entre $0.00 (m^3)$ e $4.50 \times 10^1 (m^3)$ e o peso específico do agregado entre $3.18 \times 10^6 (kg/m^3)$ e $3.60 \times 10^7 (kg/m^3)$. Para cada amostra, as duas propriedades possuem até quatro medições, a fim de garantir uma maior confiabilidade dos dados. O valor final foi dado pela média aritmética das quatro medições. Em algumas amostras, houve um número menor de medições, e o valor final também é dado pela média aritmética.

5.2 Código computacional

Para utilização das técnicas de previsão, foi desenvolvido um código computacional utilizando a biblioteca de aprendizado de máquina Scikit-learn (Pedregosa et al., 2011). Essa conhecida biblioteca foi desenvolvida na linguagem de programação Python, linguagem na qual foi implementado o código. O código desenvolvido realiza a leitura e preparação da base de dados e em seguida aplica os algoritmos de previsão na base utilizando as técnicas de validação cruzada e busca exaustiva para que possa ser avaliado o poder de generalização dos métodos e os seus respectivos parâmetros ótimos. Assim, no fim da execução é possível a comparação dos

métodos. Um esquema resumido do algoritmo é apresentado na Fig. 5.1.

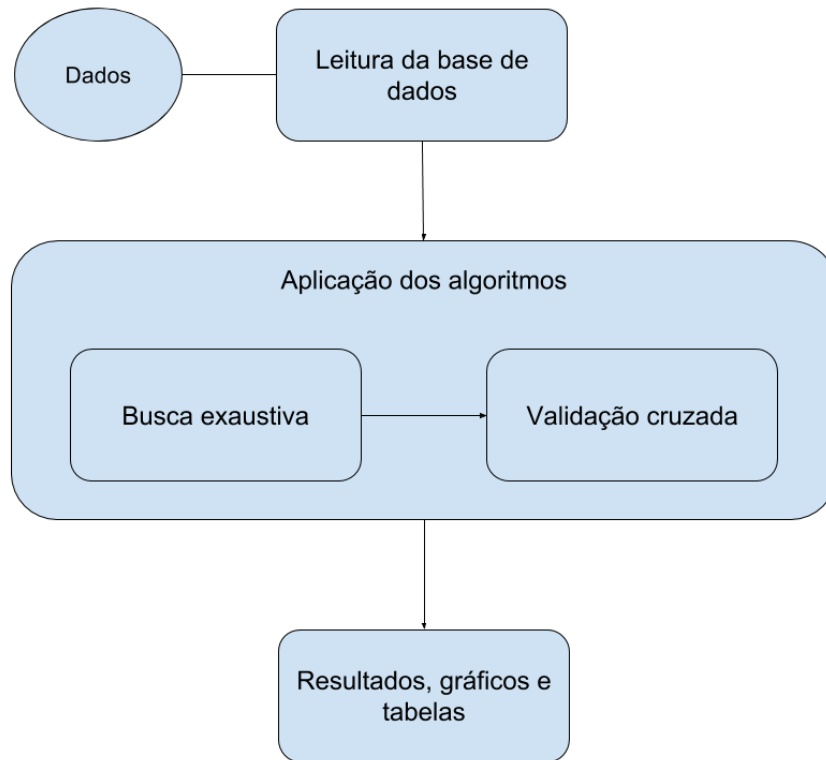


Figura 5.1: Esquema do algoritmo implementado.

Como os valores das métricas foram dados pela média de todas as execuções, também foram calculados os valores de desvio padrão para cada métrica, de modo a identificar os modelos que possuem uma menor variação na previsão, ou seja, se o desvio de um método é o menor de todos na previsão de um certo conjunto de dados, podemos concluir que esse método é o mais estável na previsão desse conjunto de dados. Logo, um método que possui em média o melhor desempenho mas um desvio muito alto, pode não ser o mais interessante na previsão do conjunto de dados em questão, pois o mesmo não possui confiabilidade nas respostas.

5.3 Resultados

Na Tabela 5.1 e Tabela 5.2 é apresentada a comparação do desempenho dos métodos utilizando os valores encontrados do MPE, MSE, R^2 e seus respectivos desvios nas 30 execuções para a resistência à compressão e o Módulo de Young, respectivamente.

As Fig. 5.2 e Fig. 5.3 apresentam os erros médios percentuais (MPE) em gráfico de boxplots para cada um dos modelos de previsão nas 30 execuções. O MPE é uma métrica

Tabela 5.1: MPE, MSE, R^2 e seus respectivos desvios referentes a cada um dos modelos de predição para Resistência à compressão

Modelo	MPE	MSE	R^2
DT	7.486 (± 0.803)	24.387 (± 4.598)	0.915 (± 0.016)
ELM	6.796 (± 1.116)	15.468 (± 5.246)	0.946 (± 0.018)
MLP	5.705 (± 0.705)	13.181 (± 4.819)	0.954 (± 0.017)
SVR	5.923 (± 0.234)	12.192 (± 0.851)	0.957 (± 0.003)

Tabela 5.2: MPE, MSE, R^2 e seus respectivos desvios referentes a cada um dos modelos de predição para o Módulo de Young

Modelo	MPE	MSE	R^2
DT	4.054 (± 0.410)	2e+06 ($\pm 5.2e+05$)	0.939 (± 0.016)
ELM	2.377 (± 0.243)	5.7e+05 ($\pm 1.3e+05$)	0.982 (± 0.004)
MLP	2.160 (± 0.191)	4.8e+05 ($\pm 1e+05$)	0.985 (± 0.003)
SVR	1.856 (± 0.080)	3.7e+05 ($\pm 2.6e+04$)	0.988 (± 0.00081)

importante, pois é levada em conta na maioria dos projetos de engenharia, e possui fácil entendimento e aplicação. É possível também observar nas figuras o valores de MPE em cada uma das 30 execuções, através dos pontos em vermelho. Com os pontos, fica evidente a distribuição dos valores de MPE e os *outliers* para cada modelo nas execuções.

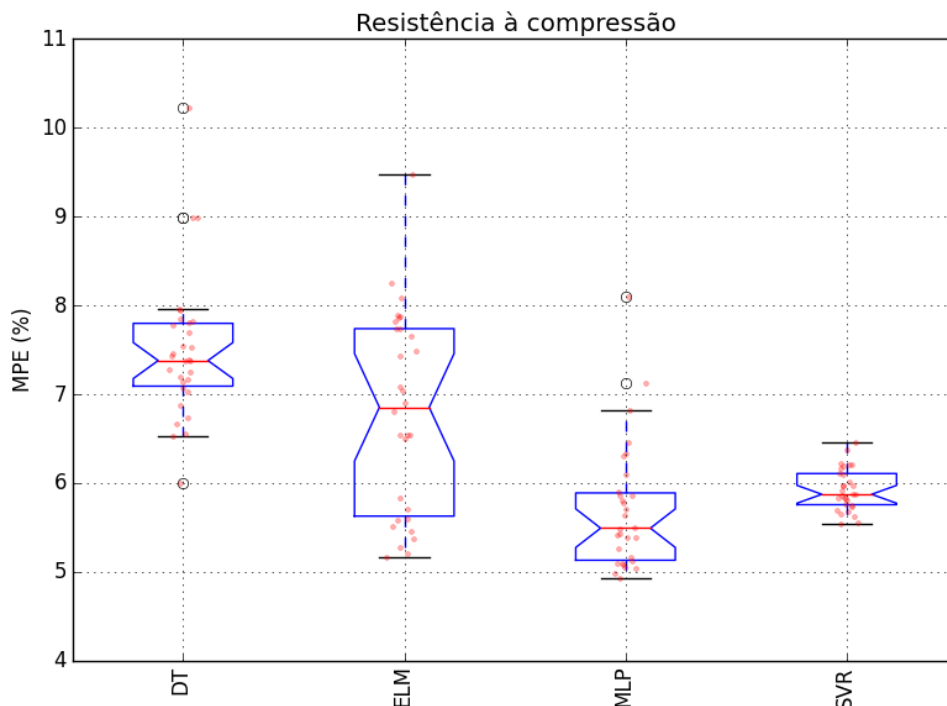


Figura 5.2: Boxplots do MPE dos modelos para a Resistência à compressão

Para cada modelo, a busca exaustiva, utilizando como métrica o MSE, encontrou os parâmetros ótimos que maximizam o poder generalização dos modelos para a previsão das

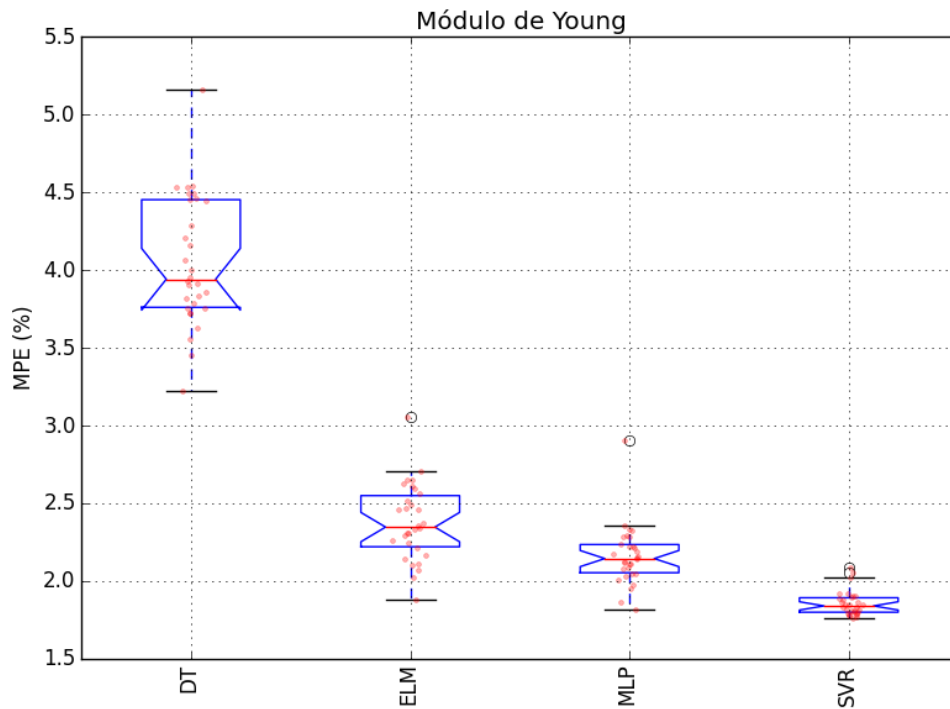


Figura 5.3: Boxplots do MPE dos modelos para o Módulo de Young

propriedades do concreto. Esses parâmetros podem ser observados na Tabela 5.3. É importante notar que na maioria dos parâmetros dos modelos, houve diferença entre os parâmetros ótimos na previsão da resistência à compressão e do módulo de Young. Isso mostra que um modelo que resolve bem a previsão de uma propriedade, pode não ser o mais adequado para resolver outra, mostrando então que a busca exaustiva possui grande importância na metodologia utilizada.

Tabela 5.3: Melhores parâmetros encontrados pela busca exaustiva com a validação cruzada para cada modelo na previsão das duas propriedades

Modelo	Parâmetros	Resistência à compressão	Módulo de Young
DT	Profundidade máxima	sem restrição	6
ELM	Função de ativação	tangente hiperbólica	multiquádrica
	Neurônios ocultos	20	20
	RBF - largura	0.2	0.0
MLP	Topologia	unidirecional simples	unidirecional totalmente conectada
	Camadas ocultas	uma camada com 5 neurônios	uma camada com 5 neurônios
SVR	Precisão (<i>epsilon</i> - ϵ)	10^{-6}	10^{-2}
	Regularização (C)	10^4	10^6
	Grau polinomial	1	1
	RBF - <i>gamma</i>	0.1	0.1

Podemos observar nos gráficos das figuras 5.4 a 5.11 as relações entre os valores reais

e estimados para as 75 amostras de teste, utilizando a busca exaustiva e validação cruzada, para cada modelo que possui um desempenho mediano na previsão das propriedades nas 30 execuções. Para isso, para um dado modelo e propriedade, foram utilizados os valores do erro percentual médio (MPE) das execuções, e assim, foi usado o modelo que possui valor mediano de MPE dentre todas as execuções para criar o gráfico. Nas figuras são identificadas as previsões das amostras de teste que ficaram dentro ou fora do intervalo confiança de 95% para todos métodos e cada propriedade, utilizando também o MPE. É importante ressaltar que, como esses resultados são referentes apenas aos modelos que possuem valores de MPE medianos, os valores de R^2 e MPE são diferentes dos apresentados na Tabela 5.1 e Tabela 5.2, pois na tabela são apresentados os valores médios das métricas em todas as execuções.

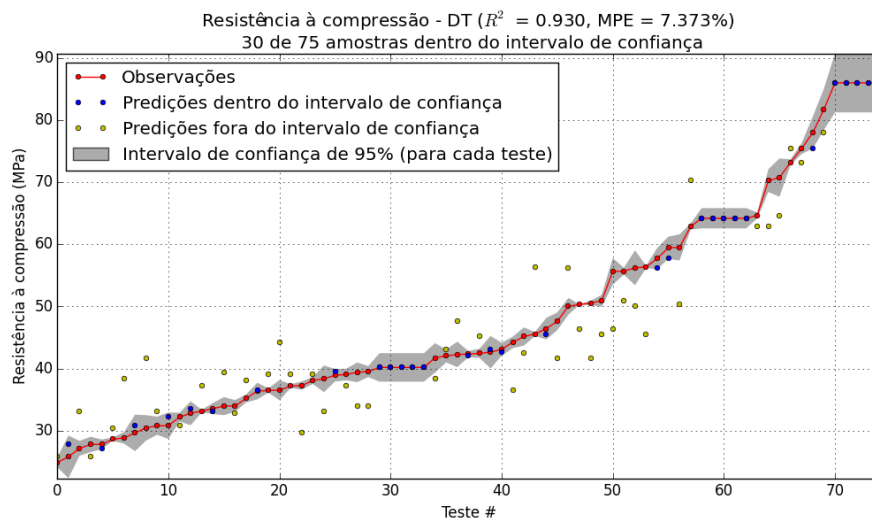


Figura 5.4: Previsão da Resistência à compressão pelo DT com desempenho mediano

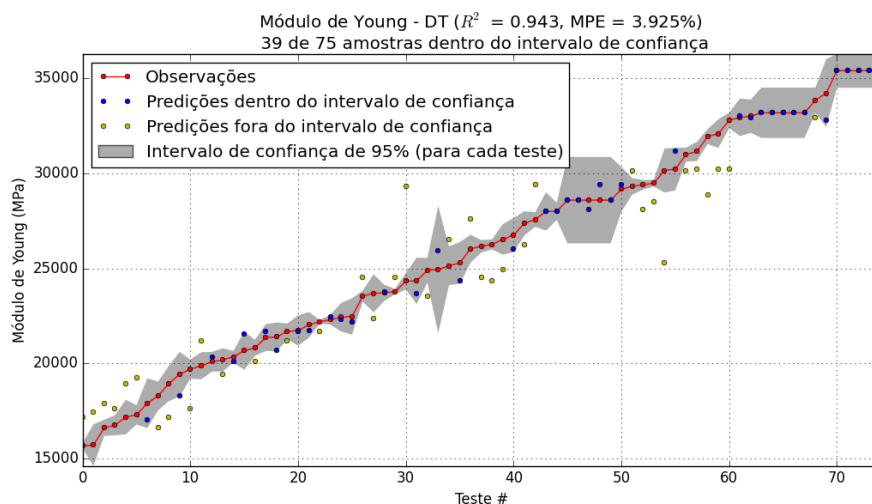


Figura 5.5: Previsão do Módulo de Young pelo DT com desempenho mediano

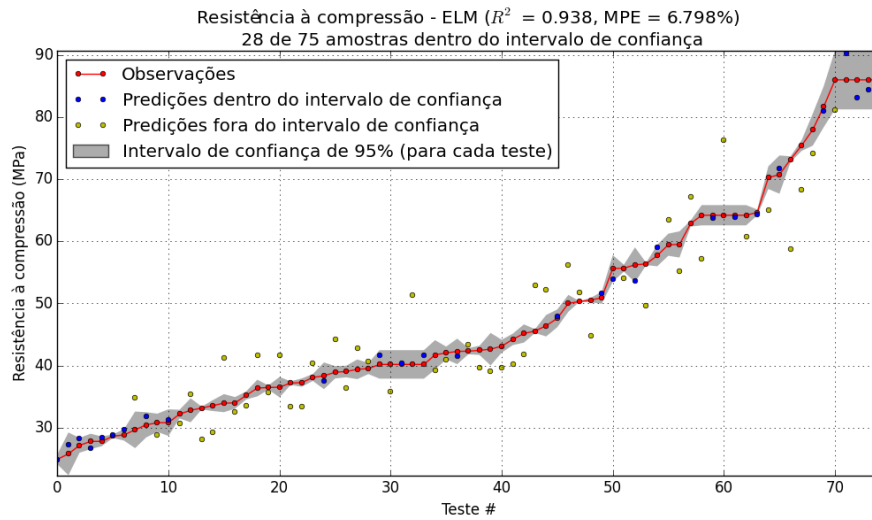


Figura 5.6: Previsão da Resistência à compressão pelo ELM com desempenho mediano

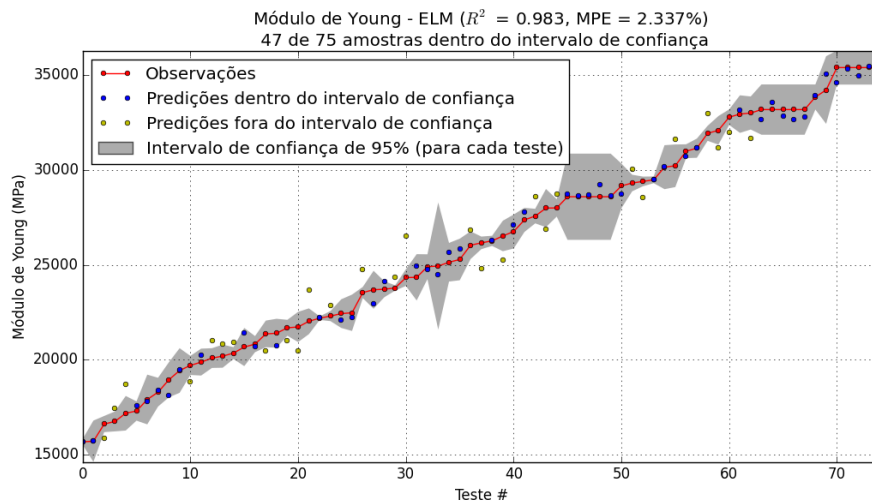


Figura 5.7: Previsão do Módulo de Young pelo ELM com desempenho mediano

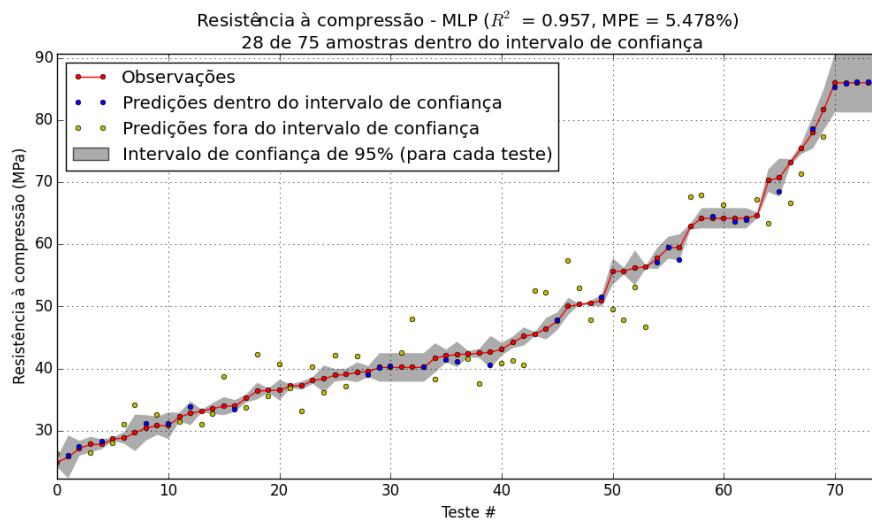


Figura 5.8: Previsão da Resistência à compressão pelo MLP com desempenho mediano

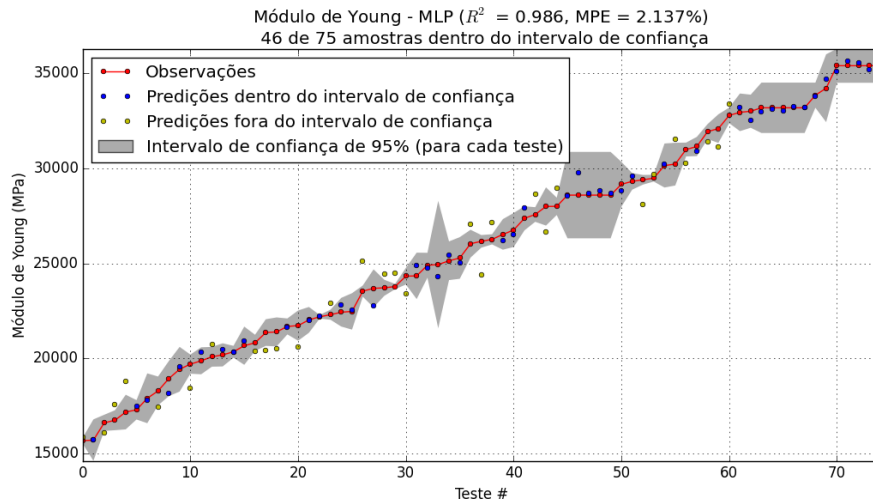


Figura 5.9: Previsão do Módulo de Young pelo MLP com desempenho mediano

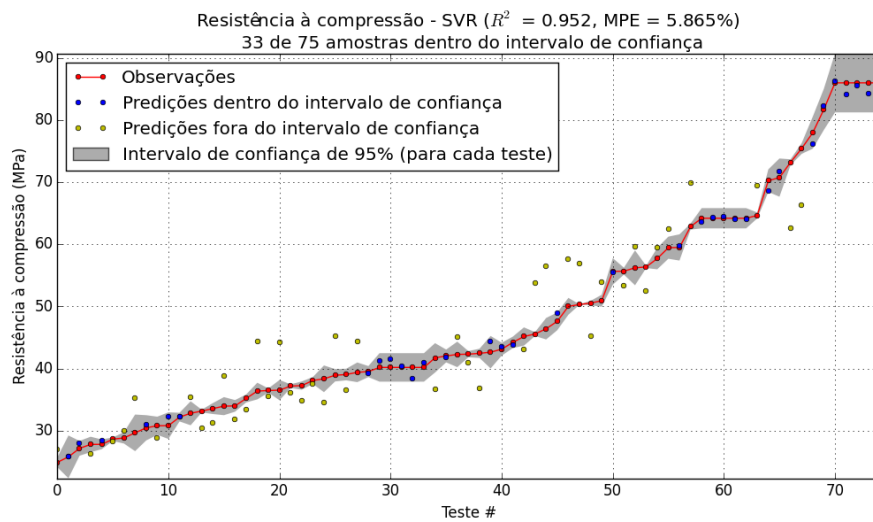


Figura 5.10: Previsão da Resistência à compressão pelo SVR com desempenho mediano.

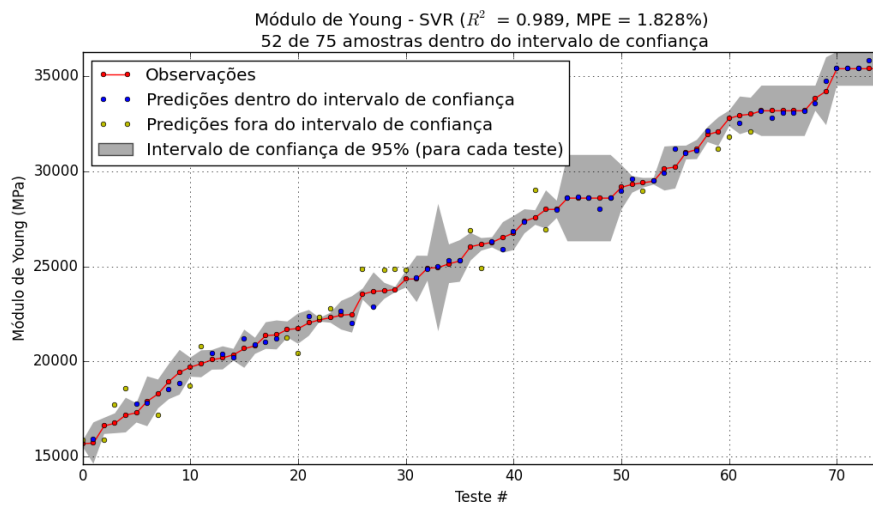


Figura 5.11: Previsão do Módulo de Young pelo SVR com desempenho mediano

6 Conclusões e trabalhos futuros

Através dos resultados apresentados podemos verificar que os modelos conseguem em média uma boa aproximação dos valores reais, com valores de erro bem baixos. Os modelos MLP e SVR conseguiram os melhores resultados dentre os modelos testados, as médias das métricas utilizadas mostram que os dois modelos possuem quase o mesmo poder de generalização, diferenciando apenas nos valores de desvio padrão, nas duas propriedades preditas.

Considerando os valores médios das 30 execuções, o MLP obteve o melhor desempenho na previsão da resistência a compressão considerando o MPE, mas para as outras métricas o SVR possui desempenho superior. O SVR foi melhor na previsão do módulo de Young em todas as métricas utilizadas (Tabela 5.1 e Tabela 5.2). Com relação aos desvios, o SVR obteve os menores valores em todas as métricas nas duas propriedades preditas.

Se levarmos em conta apenas o melhor modelo encontrado nas 30 execuções, ou seja, no melhor caso, o MLP obteve o melhor resultado na previsão da resistência a compressão e o SVR obteve o melhor resultado para o módulo de Young, isso pode ser verificado nos boxplots, na Fig. 5.2 e Fig. 5.3.

Para a quantidade de dados dentro do intervalo de confiança, o SVR obteve vantagem na previsão da resistência a compressão e do módulo de Young. Para a resistência a compressão houve 44% dos dados dentro do intervalo de confiança e para o módulo de Young houve 69,3% dos dados dentro do intervalo. Isso mostra que o SVR possui uma grande confiabilidade dos resultados para prever o módulo de Young.

Pelos valores dos desvios dos resultados nas execuções, pode-se perceber que é mais difícil e menos confiável prever a resistência a compressão que o módulo de Young. Isso também pode ser observado pela quantidade de previsões dentro do intervalo de confiança de 95%.

Por fim, podemos concluir que o SVR possui o melhor poder de generalização, pois possui o melhor desempenho para previsão das propriedades do concreto leve na maioria das métricas e também, o mesmo possui uma maior confiabilidade nas suas previsões, dados os menores desvios para todas as métricas utilizadas e a maior quantidade de dados dentro do in-

tervalo de confiança de 95%. Assim, o SVR mostrou então que pode ser utilizado na previsão dos valores reais das propriedades do concreto de agregado leve, dado o seu excelente desempenho. Isso mostra que o estudo de diferentes técnicas de inteligência computacional, além das redes neurais, para a previsão de propriedades mecânicas do concreto de agregado leve é importante, dado que essas técnicas podem conseguir resultados altamente satisfatórios e até mesmo melhores que os das redes neurais, como no caso do SVR apresentado neste trabalho.

Para trabalhos futuros, deseja-se aplicar outras técnicas de inteligência computacional, como a programação genética cartesiana (Peter and Thomson, 2000), a fim de descrever de forma matemática a relação entre os constituintes do concreto de agregado leve e as suas propriedades mecânicas.

Referências Bibliográficas

- ABNT (1995). *Agregados leves para concreto estrutural: especificação. NBR NM 35*. Associação Brasileira de Normas Técnicas.
- ACI (1999). *Guide for structural lightweight aggregate concrete. ACI 213R-87*. American Concrete Institute.
- Alshihri, M. M., Azmy, A. M., and El-Bisy, M. S. (2009). Neural networks for predicting compressive strength of structural light weight concrete. *Construction and Building Materials*, 23(6):2214 – 2219.
- bin Huang, G., yu Zhu, Q., and kheong Siew, C. (2006). Extreme learning machine: A new learning scheme of feedforward neural networks. pages 985–990.
- Bogas, J. A. and Gomes, A. (2013). Compressive behavior and failure modes of structural lightweight aggregate concrete â characterization and strength prediction. *Materials Design*, 46:832 – 841.
- Chandra, S. and Berntsson, L. (2003). *Lightweight aggregate concrete*. Science, technology and applications. Noyes Publications - William Andrew Publishing, USA.
- Cherkassky, V. and Ma, Y. (2004). Practical selection of svm parameters and noise estimation for svm regression. *Neural Networks*, 17(1):113 – 126.
- da Silva, M. D. (2013). *Estudo comparativo entre a utilização dos concretos convencional e leve nos elementos horizontais das estruturas de edifícios*. PhD thesis, Universidade Federal de Minas Gerais.
- Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., and Vapnik, V. (1997). Support vector regression machines. In *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 9*, pages 155–161. MIT Press.
- Duan, Z., Kou, S., and Poon, C. (2013). Using artificial neural networks for predicting the elastic modulus of recycled aggregate concrete. *Construction and Building Materials*, 44(0):524 – 532.
- Faust, T. (2000). Properties of different matrixes and lwas and their influences on the behavior of structural lwac. *Second int proceedings second international symposium on structural lightweight aggregate concrete*, pages 502 – 511.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer, New York.
- Huang, G., Huang, G.-B., Song, S., and You, K. (2015). Trends in extreme learning machines: A review. *Neural Networks*, 61:32 – 48.
- Kasperkiewicz, J., Racz, J., and Dubrawski, A. (1995). Hpc strength prediction using artificial neural network. *Journal of Computing in Civil Engineering*, 9(4):279–284.

- Ke, Y. (2008). *Characterization of the mechanical behavior of lightweight aggregate concretes: Experiment and modelling*. PhD thesis, PhD. Thesis, Université de Cergy-Pontoise.
- Kockal, N. U. and Ozturan, T. (2011). Strength and elastic properties of structural lightweight concretes. *Materials and Design*, 32(4):2396 – 2403.
- Kolay, E. and Baser, T. (2014). Estimating of the Dry Unit Weight of Compacted Soils Using General Linear Model and Multi-layer Perceptron Neural Networks. *Applied Soft Computing*, 18:223–231.
- Kröse, B. J. A. and van der Smagt, P. P. (1993). *An Introduction to Neural Networks*. University of Amsterdam, Amsterdam.
- Neville, A. (2015). *Propriedades do Concreto - 5ª Edição*. Bookman.
- Ni, H.-G. and Wang, J.-Z. (2000). Prediction of compressive strength of concrete by neural networks. *Cement and Concrete Research*, 30(8):1245 – 1250.
- Nocedal, J. and Wright, S. J. (2006). Numerical optimization, second edition. *Numerical optimization*, pages 497–528.
- Papadakis, V. and Tsimas, S. (2002). Supplementary cementing materials in concrete: Part i: efficiency and design. *Cement and Concrete Research*, 32(10):1525 – 1532.
- Pedregosa, F., Varoquax, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. D. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Peter, J. M. and Thomson, P. (2000). *Cartesian Genetic Programming*. Springer, New York.
- Radonjic, A., Zlatko, H., and Vladeta, C. (2012). Application of different learning algorithms for the prediction of power of inland pushboats. *PROCEEDINGS OF INTERNATIONAL CONFERENCE ON TRAFFIC AND TRANSPORT ENGINEERING (ICTTE)*, pages 29–39.
- Russell, S. J. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Pearson Education, 2 edition.
- Vapnik, V. and Kotz, S. (2006). *Estimation of Dependences Based on Empirical Data: Empirical Inference Science (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.