



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

Jonata Jefferson Andrade

Flávio de Souza Barbosa

Michèle Cristina Resende Farage

Leonardo Goliatt da Fonseca

jonata.jefferson@ice.ufjf.br

{flavio.barbosa, michele.farage, leonardo.goliatt}@ufjf.edu.br

Universidade Federal de Juiz de Fora

Rua José Lourenço Kelmer, s/n - São Pedro, 36036-900, Minas Gerais, Juiz de Fora, Brasil

Anne-Lise Beaucour

Sophie Ortola

{anne-lise.beaucour, sophie.ortola}@u-cergy.fr

Universidade de Cergy-Pontoise

33 Boulevard du Port, 95011, Cergy-Pontoise, França

Abstract. *O objetivo deste trabalho é avaliar o desempenho de 4 métodos de inteligência computacional para prever a módulo de elasticidade (Módulo de Young) e a resistência à compressão 28 dias de concretos de agregados leves em função do fator água/cimento, volume de agregado leve, quantidade de cimento e densidade do agregado leve. Depois de uma verificação da capacidade de generalização dos métodos através do processo de validação cruzada e ajuste de parâmetros por meio de busca exaustiva, observou-se as Máquinas de Vetores Suporte apresentaram o melhor desempenho na aproximação das duas propriedades mecânicas.*

Keywords: *Concreto de agregados leves, Aprendizado de máquina, Inteligência computacional*

1 INTRODUÇÃO

Nas últimas décadas, o concreto armado convencional, de densidade normal, tem sido o material estrutural mais empregado no Brasil. A razão deste fato está associada as seguintes vantagens, das quais se destacam: (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 densidade 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 densidade normal. 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°C .

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, nenhum trabalho explorou diferentes técnicas preditivas, além das redes neurais artificiais (Alshihri et al., 2009), tais como Máquinas de Vetores Suporte, Máquinas de Aprendizado Extremo e Florestas de Árvores de Decisão. O desenvolvimento da metodologia proposta neste trabalho, portanto, vem complementar os estudos encontrados atualmente na literatura e preenche esta lacuna.

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 daquele adotado em projeto, conseguindo dessa forma reduzir a quantidades de ensaios laboratoriais que demanda tempo e também o retrabalho.

2 CONCRETO DE AGREGADO LEVE

O concreto feito com agregados leves é chamado de concreto leve. O concreto 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 densidade normal varia entre 2200 kg/m³ e 2600 kg/m³, e a do concreto estrutural leve entre 1350 kg/m³ e 1850 kg/m³ (Tabela 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 leve é maior que o concreto de densidade normal, porém, mesmo com este investimento adicional o custo final da estrutura pode ser menor devido ao dimensionamento mais econômico das fundações.

Tabela 1: Classificação dos concretos leves quanto à sua massa específica aproximada.

Classificação	Densidade (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

Em geral, para um determinado agregado leve, há uma relação entre o teor de cimento do concreto e a resistência à compressão. Ademais, como o cimento tem massa específica maior do que o agregado leve, para um dado agregado, a resistência à compressão do concreto aumenta com o aumento de sua massa específica. Neste sentido, a maioria das normas e especificações apresenta relações entre a massa específica do concreto leve e sua resistência à compressão.

O módulo de elasticidade 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 densidade usual, o módulo de elasticidade da pasta de cimento é, geralmente, bem menor que o módulo dos agregados. Por outro lado, no concreto estrutural 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 densidade 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 densidade usual.

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. Propriedades como a resistência à compressão não é determinada somente pela razão de água/cimento, mas também por outros materiais utilizados na mistura. O concreto de agregado leve contém outros ingredientes tais como o tipo e a quantidade de agregado leve, além de aditivos e da composição da argamassa. Os vários ingredientes, além da não linearidade das estruturas de concreto, podem complicar o cálculo da resistência à compressão, 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 a resistência à compressão é determinada, independentemente do tipo de concreto, não só pela razão de água para cimento, mas também por outros componentes (Chandra and Berntsson, 2003; Kockal and Ozturan, 2011; Faust, 2000). As equações empíricas apresentadas para a estimativa da resistência à compressão 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 com a mistura 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, dado 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 densidade 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 custo alto.

Assim como nesses trabalhos para concreto de densidade normal, além das redes neurais, outras técnicas preditivas contidas na literatura, tais como Máquinas de Vetores Suporte, Máquinas de Aprendizado Extremo e Florestas de 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 indicará qual técnica possui a melhor generalização do problema real. Cada uma das técnicas são descritas a seguir.

3.1 Redes Neurais

Uma Rede Neural Artificial (RNA) pode ser caracterizada como um modelo computacional com propriedades particulares 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, chamados de neurônios, arranjados em múltiplas camadas, que se comunicam pela troca de mensagens através de suas conexões.

Perceptron multicamada (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. 1.

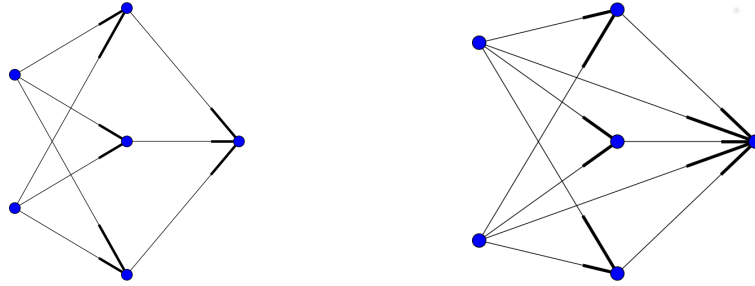


Figura 1: Modelos de MLP unidirecional simples (esquerda) e unidirecional totalmente conectada (direita)

Neste trabalho usamos a função de ativação sigmoideal, dada pela expressão $\phi = 1/(1 + \exp(x))$. A rede neural é treinada ajustando-se os pesos sinápticos com o algoritmo TNC (Nocedal and Wright, 2006).

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

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

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 $\mathbf{a}_i \in \mathbb{R}^d$ e $b_i > 0$. Neste trabalho, utilizaremos a função de ativação G como sendo a função de base radial (RBF, do inglês Radial Basis Function).

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

$$\begin{bmatrix} G(\mathbf{a}_1, b_1, \mathbf{x}_1) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_1) \\ \vdots & \vdots & \vdots \\ G(\mathbf{a}_1, b_1, \mathbf{x}_N) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_N) \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad (2)$$

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

$$\mathbf{w} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} = \mathbf{H}^\dagger \mathbf{y} \quad (3)$$

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. 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:

$$\hat{y}(x) = \sum_{m=1}^K \hat{c}_m I(x \in R_m) \quad (4)$$

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

$$\hat{c}_m = \text{média}(x \in R_m) \quad (5)$$

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

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 + \right] \quad (6)$$

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

$$\hat{c}_1 = \text{média}(y_i | x_i \in R_1(j,s)) \text{ e } \hat{c}_2 = \text{média}(y_i | x_i \in R_2(j,s)) \quad (7)$$

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

$$\hat{y}(x) = \sum_{j=1}^m w_j \phi_j(x) + b \quad (8)$$

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

$$L(x) = \begin{cases} 0 & \text{se } |y(x) - \hat{y}(x)| \leq \varepsilon \\ |y(x) - \hat{y}(x)| & \text{caso contrário} \end{cases} \quad (9)$$

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

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|_2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{sujeito a} \quad & \begin{cases} y_i - \hat{y}_i \leq \varepsilon + \xi_i^* \\ \hat{y}_i - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, n \end{cases} \end{aligned} \quad (10)$$

O problema de otimização pode ser transformado no problema dual

$$\hat{y}(x) = \sum_{i=1}^{n_{SV}} (\alpha_i - \alpha_i^*) K(x_i, x) + b, \text{ sujeito a } 0 \leq \alpha_i \leq C, 0 \leq \alpha_i^* \leq C \quad (11)$$

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

3.5 Generalização Dos Modelos De Predição

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 foi utilizado $k = 10$, resultando então em 9 divisões da base de dados.

Como cada modelo possuem diversos parâmetros a serem escolhidos e queremos 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 2 apresenta a variação dos parâmetros de cada modelo.

Tabela 2: 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	RBF
	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.
	Algoritmo de treinamento	TNC (Nocedal and Wright, 2006)
SVR	Precisão da regressão (<i>epsilon</i>)	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}

De modo a avaliar o desempenho de previsão dos modelos, nós utilizamos 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.

$$R^2(y, \hat{y}) = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (12)$$

$$\text{MPE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{|y_i|} \times 100 \quad (13)$$

$$\text{MSE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (14)$$

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

Para realizar a busca exaustiva dos parâmetros é necessário utilizar como forma de medida uma métrica de avaliação, para isso, utilizamos 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 será dado pela média em cada uma das 30 repetições.

4 EXPERIMENTOS, RESULTADOS E DISCUSSÃO

A fim de avaliar o desempenho dos métodos de inteligência computacional na previsão de propriedades mecânicas de concretos de agregado leve, empregamos dados de ensaios laboratoriais do concreto com suas diferentes características. Os dados experimentais foram obtidos em um extenso programa experimental na Universidade de Cergy-Pontoise. A base de dados possui 75 amostras de teste com quatro parâmetros e duas saídas. As propriedades dos concretos leves coletadas foram o módulo de elasticidade (Módulo de Young) e a resistência à compressão 28 dias, em função do fator água/cimento, volume de agregado leve, quantidade de cimento e densidade do agregado leve. Para cada amostra, as duas propriedades possuem quatro medições, a fim de garantir uma maior confiabilidade dos dados. Se foi possível realizar as quatro medições, o valor final foi dado pela média aritmética das quatro medições. Nas amostras que não foram possível a realização das quatro medições (devido algum erro de medição, por exemplo) foi considerado apenas as medições feitas e o valor final também como sendo a média aritmética.

Para utilização das técnicas de previsão, foi desenvolvido um código computacional baseado na 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 da base de dados e em seguida aplica os algoritmos de predição 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.

Como os valores das métricas foram dados pela média de todas as execuções, também foi calculado 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 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. Na Tabela 3 e Tabela 4 é apresentado 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.

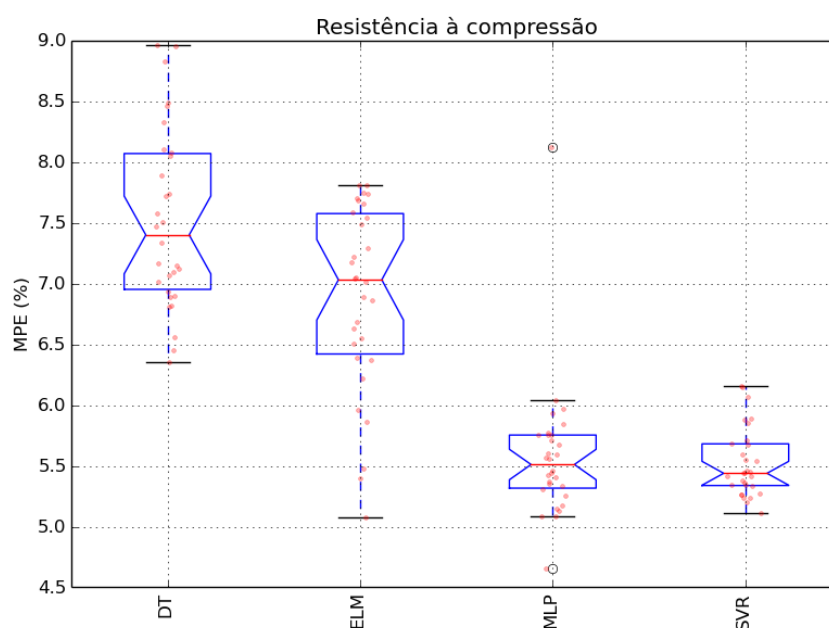
A Fig. 2 e Fig. 3 apresentam os erros médios percentuais (MPE) em gráfico de boxplots para cada um dos modelos de previsão nas 30 execuções. É possível observar nas figuras o valores de MPE nas 30 execuções (pontos em vermelho). Com os pontos, fica evidente a distribuição dos valores de MPE e os *outliers* para cada modelo nas execuções.

Tabela 3: 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	7.529 (± 0.734)	23.594 (± 4.122)	0.918 (± 0.014)
ELM	6.883 (± 0.764)	16.370 (± 5.261)	0.943 (± 0.018)
MLP	5.577 (± 0.563)	11.5793 (± 3.560)	0.959 (± 0.012)
SVR	5.524 (± 0.281)	11.397 (± 1.242)	0.960 (± 0.004)

Tabela 4: 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	3.756 (± 0.406)	1.8e+06 ($\pm 5e+05$)	0.945 (± 0.016)
ELM	2.369 (± 0.206)	5.5e+05 ($\pm 1e+05$)	0.983 (± 0.003)
MLP	2.200 (± 0.155)	5e+05 ($\pm 7.9e+04$)	0.984 (± 0.002)
SVR	1.818 (± 0.048)	3.57e+05 ($\pm 1.1e+04$)	0.989 (± 0.00035)

**Figura 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 propriedades do concreto, esses parâmetros podem ser observados na Tabela 5. É importante notar que na maioria dos parâmetros dos modelos, não houve diferença entre os parâmetros ótimos na previsão da resistência à compressão e do módulo de Young. Isso mostra que um

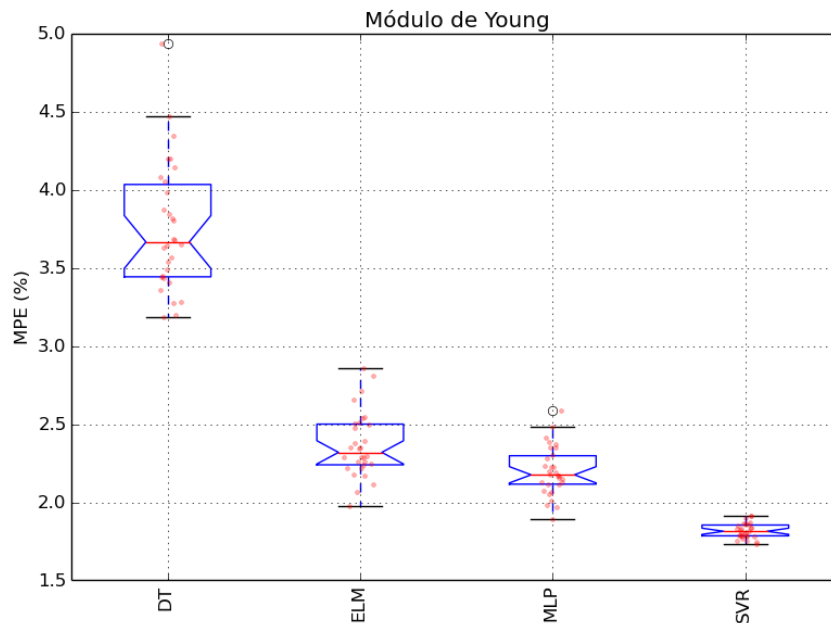


Figura 3: Boxplots do MPE dos modelos para a Módulo de Young

modelo que resolve bem a previsão de uma propriedade, também resolve bem a outra.

Tabela 5: 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	9
ELM	Neurônios ocultos	20	20
	RBF - largura	0.1	0.1
MLP	Topologia	unidirecional simples	unidirecional simples
	Camadas ocultas	uma camada com 10 neurônios	uma camada com 5 neurônios
SVR	Precisão (<i>epsilon</i>)	10^{-6}	10^{-2}
	Regularização (C)	10^6	10^6
	Grau polinomial	1	1
	RBF - <i>gamma</i>	0.1	0.1

5 CONCLUSÃO

Através dos resultados apresentados podemos verificar que os modelos na maioria dos casos conseguem uma boa aproximação dos valores reais. Os modelos MLP e SVR conseguiram os melhores resultados dentre os modelos testados. As métricas utilizadas mostram que os dois

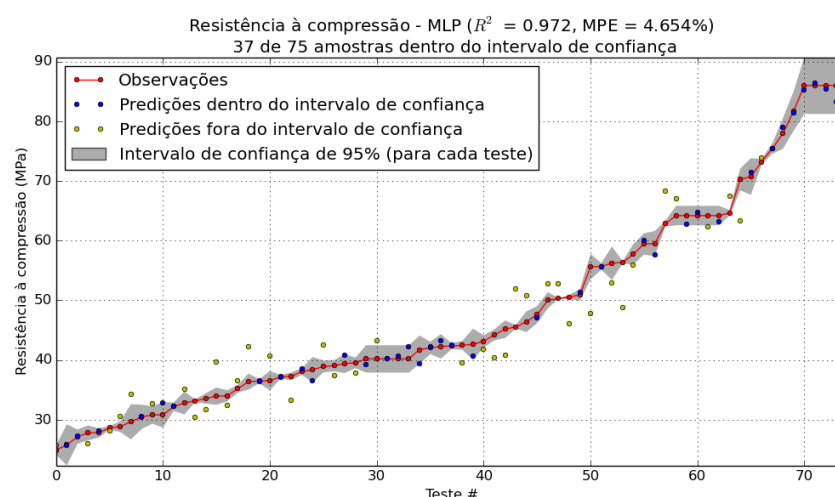


Figura 4: Previsão da Resistência à compressão pelo melhor modelo MLP encontrado pela busca exaustiva

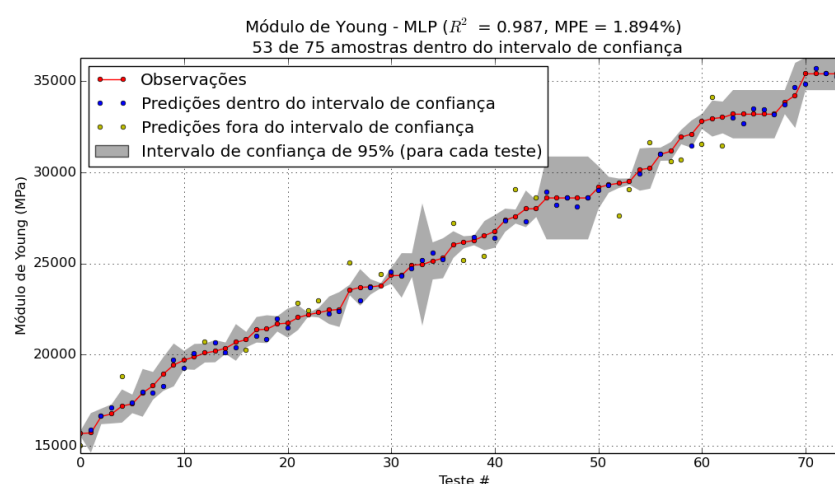


Figura 5: Previsão do Módulo de Young pelo melhor modelo MLP encontrado pela busca exaustiva

modelos possuem quase o mesmo poder de generalização, diferenciando apenas nos valores de desvio padrão.

Se levarmos em conta apenas o melhor modelo encontrado nas 30 execuções, o MLP obteve o melhor resultado na previsão da resistência a compressão (Fig. 2) e o SVR obteve o melhor resultado para o módulo de Young (Fig. 3). Mas considerando os valores médios das 30 execuções o SVR foi melhor na previsão das duas propriedades, principalmente para o módulo de Young (Tabela 3 e Tabela 4).

Podemos observar as 75 amostras de teste e as suas respectivas observações e previsões encontrados pela busca exaustiva nas 30 execuções para o melhor modelo MLP na Fig. 4 para resistência a compressão e Fig. 5 para o módulo de Young e para o melhor modelo SVR na Fig. 6 para resistência a compressão e Fig. 7 para o módulo de Young. É importante ressaltar que, como esses resultados são referentes apenas aos modelos MLP e SVR com maior poder de generalização, os valores de R^2 e MPE são diferentes, e melhores, dos apresentados na Tabela 3 e Tabela 4, pois na tabela são apresentados os valores médios das métricas em todas

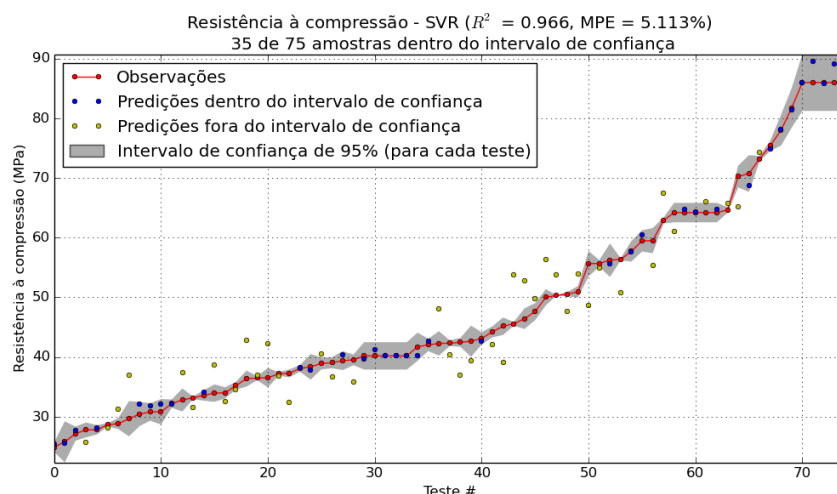


Figura 6: Previsão da Resistência à compressão pelo melhor modelo SVR encontrado pela busca exaustiva

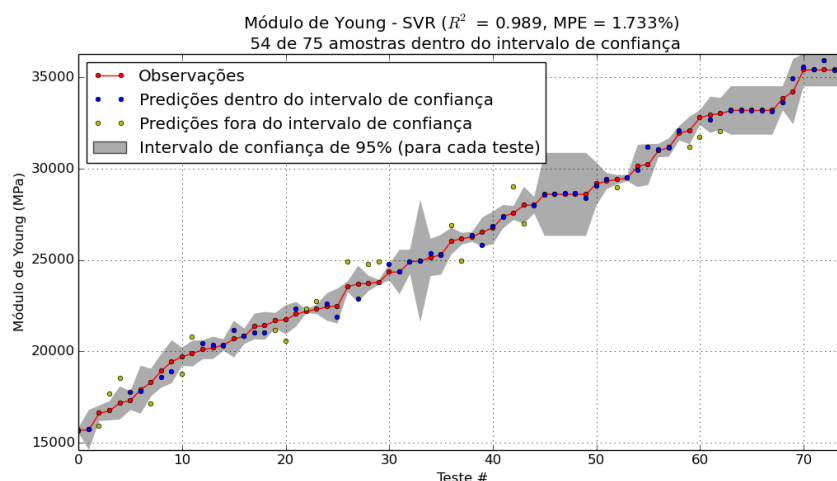


Figura 7: Previsão do Módulo de Young pelo melhor modelo SVR encontrado pela busca exaustiva

as execuções. Nas figuras são identificadas as previsões das amostras de teste que ficaram dentro ou fora do intervalo confiança de 95% para os dois métodos, calculado considerando as 4 medições das propriedades de cada amostra de teste presente na base de dados. Levando em conta a quantidade de dados dentro do intervalo de confiança, o MLP obteve uma pequena vantagem na previsão da resistência a compressão e o SVR um pequena vantagem na previsão do módulo de Young.

Por fim, podemos concluir que o SVR possui a melhor poder de generalização, dado que o mesmo possui os menores desvios e em média os melhores valores para as métricas utilizadas. Isso mostra que o estudo de outras técnicas de inteligência computacional, além das redes neurais, para a previsão de propriedades mecânicas do concreto de agregado leve fez-se necessário, dado que essas técnicas conseguem resultados altamente satisfatórios e até mesmo melhores que os das redes neurais, como no caso do SVR apresentado neste trabalho.

AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - PQ 306712/2012-1), à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG - PCE-01337-15, APQ-03900-10, APQ-01659-12 e PPM-00020/14), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) através do Programa Nacional de Cooperação Acadêmica (Procad 071/2013) e à Universidade Federal de Juiz de Fora, pelo apoio financeiro e de recursos para o desenvolvimento do projeto de pesquisa.

REFERÊNCIAS

- 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.
- Chandra, S. and Berntsson, L. (2003). *Lightweight aggregate concrete*. Science, technology and applications. Noyes Publications - William Andrew Publishing, USA.
- 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.
- 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.
- 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.
- 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.
- 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.