

Universidade Federal de Juiz de Fora  
Instituto de Ciências Exatas  
Programa de Pós-Graduação em Modelagem Computacional

**Jonata Jefferson Andrade**

**Métodos de inteligência computacional com otimização evolucionária para a estimativa de propriedades mecânicas do concreto de agregado leve**

Juiz de Fora

2017

**Jonata Jefferson Andrade**

**Métodos de inteligência computacional com otimização evolucionária para a estimativa de propriedades mecânicas do concreto de agregado leve**

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Mestre em Modelagem Computacional.

Orientador: Leonardo Goliatt da Fonseca

Juiz de Fora

2017

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Andrade, Jonata Jefferson .

Métodos de inteligência computacional com otimização evolucionária para a estimativa de propriedades mecânicas do concreto de agregado leve / Jonata Jefferson Andrade. -- 2017. 75 f.

Orientador: Leonardo Goliatt da Fonseca

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, ICE/Engenharia. Programa de Pós-Graduação em Modelagem Computacional, 2017.

1. Inteligência computacional. 2. Concreto de agregado leve. 3. Otimização. I. Fonseca, Leonardo Goliatt da, orient. II. Título.

Jonata Jefferson Andrade

**Métodos de inteligência computacional com otimização evolucionária para a estimativa de propriedades mecânicas do concreto de agregado leve**

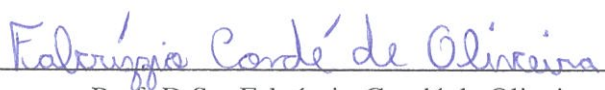
Dissertação apresentada ao Programa de Pós-graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Mestre em Modelagem Computacional.

Aprovada em:

BANCA EXAMINADORA

  
\_\_\_\_\_  
Prof. D.Sc. Leonardo Goliatt da Fonseca - Orientador  
Universidade Federal de Juiz de Fora

  
\_\_\_\_\_  
Prof. D.Sc. Michèle Cristina Resende Farage  
Universidade Federal de Juiz de Fora

  
\_\_\_\_\_  
Prof. D.Sc. Fabrício Condé de Oliveira  
Universidade Salgado de Oliveira

*Aos meus pais, irmãos, amigos e a minha querida noiva que não mediram esforços para me ajudar a alcançar mais esse objetivo.*

## **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 noiva 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, Tales e Lucas, pela amizade e companheirismo durante todos os momentos.

Aos professores do Programa de Pós-Graduação em Modelagem Computacional pelos seus ensinamentos e aos funcionários do curso, que durante esses anos, contribuíram de algum modo para o meu enriquecimento pessoal e profissional.

“Saber muito não lhe torna inteligente. A inteligência se traduz na forma que você recolhe, julga, maneja e, sobretudo, onde e como aplica esta informação.”

Carl Sagan

## RESUMO

No concreto de agregado leve, a resistência à compressão e o módulo de elasticidade são as propriedades mecânicas mais importantes e conseqüentemente as mais comumente analisadas. A relação entre os componentes do concreto de agregado leve e suas propriedades mecânicas é altamente não-linear, e o estabelecimento de um modelo de previsão abrangente de tais características é usualmente problemático. Existem trabalhos que buscam encontrar essa relação de formas empíricas. Há também trabalhos que buscam aplicar técnicas de inteligência computacional para prever essas propriedades a partir dos componentes do concreto. Prever com precisão as propriedades mecânicas do concreto de agregado leve é um problema crítico em projetos de engenharia que utilizam esse material. O objetivo desta dissertação é avaliar o desempenho de diferentes métodos de inteligência computacional para prever a módulo de elasticidade e a resistência à compressão aos 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. Para a escolha da melhor configuração de cada método, foi definida uma metodologia utilizando o algoritmo de otimização PSO (Particle Swarm Optimization). Por fim, é verificada a capacidade de generalização dos métodos através do processo de validação cruzada de modo a encontrar o método que apresenta o melhor desempenho na aproximação das duas propriedades mecânicas.

Palavras-chave: Inteligência computacional, Concreto de agregado leve, Otimização.



## **ABSTRACT**

In lightweight aggregate concrete, the compressive strength, the elastic modulus and specific weight are the most important properties and consequently the most commonly analyzed. The relationship between lightweight aggregate concrete components and their mechanical properties is highly nonlinear, and establishing a comprehensive predictive model of such characteristics is usually problematic. There are works that seek to find this relation of empirical forms. There are also works that seek to apply computational intelligence techniques to predict these properties from the concrete components. Accurately predicting the mechanical properties of lightweight aggregate concrete is a critical problem in engineering projects that use this material. The objective of this dissertation is to evaluate the performance of different computational intelligence methods to predict the elastic modulus and the compressive strength at 28 days of lightweight aggregates concrete as a function of water/cement factor, lightweight aggregate volume, cement quantity and density of the lightweight aggregate. In order to choose the best configuration of each method, a methodology was defined using the Particle Swarm Optimization (PSO) algorithm. Finally, the generalization of the methods through the cross validation process is verified in order to find the method that presents the best performance in the approximation of the two mechanical properties.

**Key-words:** Computational intelligence, Lightweight aggregate concrete, Optimization

## LISTA DE ILUSTRAÇÕES

Figura 1 – Barra prismática com cargas axiais nas extremidades. Adaptado de (Hibbeler, 2010). . . . .	19
Figura 2 – Deformações de um corpo de prova cilíndrico de um material com comportamento elástico submetido a uma carga de tração uniaxial (a) e uma carga de compressão uniaxial (b). Adaptado de (Callister and Rethwisch, 2011). . . . .	20
Figura 3 – Comportamento da curva tensão-deformação até a fratura do material para um ensaio de tração. No ponto $M$ temos a tensão máxima de tração (Resistência a tração) $RT$ . Adaptado de (Callister and Rethwisch, 2011). . . . .	20
Figura 4 – Diagrama tensão-deformação para mistura de concreto. Adaptado de (Hibbeler, 2010). . . . .	21
Figura 5 – Imagens de agregados de argila expandida (Bogas and Gomes, 2013). . . . .	23
Figura 6 – Estrutura interna de agregados de argila expandida. Retirado de (Ardakani and Yazdani, 2014) . . . . .	23
Figura 7 – Templo romano Pantheon, construído utilizando agregados leves de pedrapomes no concreto (Chandra and Berntsson, 2002). . . . .	24
Figura 8 – Transferência das forças de compressão no concreto normal e leve (Bogas and Gomes, 2013). . . . .	25
Figura 9 – Seção do agregado argila expandida (a) e seção do agregado xisto expandido em (b). Adaptado de (Ke et al., 2010). . . . .	32
Figura 10 – Gráficos de associação entre os valores dos atributos de entrada das amostras de teste utilizados e o módulo de elasticidade. . . . .	33
Figura 11 – Gráficos de associação entre os valores dos atributos de entrada das amostras de teste utilizadas e a resistência à compressão. . . . .	34
Figura 12 – Acima, um exemplo de uma rede com 4 variáveis de entrada, uma camada escondida com 5 neurônios e um neurônio na camada de saída. Abaixo o detalhamento de um único neurônio da rede. . . . .	36
Figura 13 – Modelos de MLP unidirecional simples (esquerda) e unidirecional totalmente conectada (direta) . . . . .	37
Figura 14 – Exemplos de função de ativação. Na esquerda, a função sinal que é descontínua, é geralmente usada para classificação. No centro uma função semi-linear e na direita a função sigmóide, que são funções contínuas usadas para regressão. . . . .	38
Figura 15 – 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). . . . .	41

Figura 16 – Exemplo de uma curva de regressão treinada pelo SVR e as respectivas variáveis de folga e o parâmetro de precisão $\epsilon$ , para uma entrada $x$ e saída $y$ (Drucker et al., 1997). . . . .	43
Figura 17 – Exemplo de atualização da posição e velocidade do PSO, para uma partícula $i$ na iteração $t$ no espaço de busca. . . . .	46
Figura 18 – Esquema da validação cruzada k-fold, para $k=5$ . . . . .	49
Figura 19 – Exemplo de mapeamento das variáveis discretas, sendo $P_j$ o vetor que representa um parâmetro $j$ com 4 opções e $aux_j$ a variável auxiliar associada. . .	51
Figura 20 – Boxplots do MAPE dos métodos para o Módulo de elasticidade . . . . .	57
Figura 21 – Boxplots do MAPE dos métodos para a Resistência à compressão . . . . .	58
Figura 22 – Gráficos de dispersão para a previsão da resistência à compressão pelos métodos testados que obtiveram desempenho mediano nas 30 execuções. . .	59
Figura 23 – Gráficos de dispersão para a previsão do módulo de elasticidade pelos métodos testados que obtiveram desempenho mediano nas 30 execuções. . . . .	60
Figura 24 – Previsão da Resistência à compressão pelo DT com desempenho mediano. . .	61
Figura 25 – Previsão do Módulo de elasticidade pelo DT com desempenho mediano. . .	61
Figura 26 – Previsão da Resistência à compressão pelo ELM com desempenho mediano. . .	62
Figura 27 – Previsão do Módulo de elasticidade pelo ELM com desempenho mediano. . .	62
Figura 28 – Previsão da Resistência à compressão pelo MLP com desempenho mediano. . .	63
Figura 29 – Previsão do Módulo de elasticidade pelo MLP com desempenho mediano. . .	63
Figura 30 – Previsão da Resistência à compressão pelo SVR com desempenho mediano. . .	64
Figura 31 – Previsão do Módulo de elasticidade pelo SVR com desempenho mediano. . .	64
Figura 32 – Gráficos de barras das importâncias (médias) das variáveis de entrada para previsão do módulo de elasticidade pela árvore de decisão. . . . .	66
Figura 33 – Gráficos de barras das importâncias (médias) das variáveis de entrada para previsão da resistência à compressão pela árvore de decisão. . . . .	67

## LISTA DE TABELAS

Tabela 1 – Classificação dos concretos leves a partir da peso específico. Dados de Neville (2015) . . . . .	24
Tabela 2 – Tipos de argamassa utilizadas nos experimentos . . . . .	32
Tabela 3 – Tipos de agregados utilizados nos experimentos. . . . .	32
Tabela 4 – Variação das entradas e saídas da base de dados experimental. . . . .	33
Tabela 5 – Funções de mapeamento comumente usados no ELM. . . . .	39
Tabela 6 – Funções Kernel utilizadas no SVR. . . . .	44
Tabela 7 – Variação dos parâmetros dos métodos que são usados na busca do PSO. . . . .	50
Tabela 8 – Configuração do PSO utilizado. . . . .	52
Tabela 9 – Resumo dos resultados das métricas de avaliação nas previsões do módulo de elasticidade ( $E_c$ ) para cada método. . . . .	55
Tabela 10 – Resumo dos resultados das métricas de avaliação nas previsões da resistência à compressão ( $f_c$ ) para cada método. . . . .	56
Tabela 11 – Melhores parâmetros encontrados pelo PSO para cada modelo na previsão do módulo de elasticidade. . . . .	65
Tabela 12 – Melhores parâmetros encontrados pelo PSO para cada modelo na previsão da resistência à compressão. . . . .	65
Tabela 13 – Tempo de execução médio e respectivo desvio padrão em segundos das 30 execuções de cada processo de busca do PSO, com validação e previsão das propriedades, para cada método. . . . .	68
Tabela 14 – Tempo de execução médio do processo de aprendizado, validação e previsão para cada método testado. . . . .	68

## LISTA DE ABREVIATURAS E SIGLAS

DT	Árvore de decisão
ELM	Máquina de Aprendizado Extremo
MLP	Rede neural Perceptron de Multiplas Camadas
RNA	Rede Neuronal Artificial
SVM	Máquina de vetores suporte
SVR	Máquina de vetores suporte para regressão
PSO	Algoritmo de otimização por enxame de partículas
MAPE	Erro percentual absoluto médio
MSE	Erro quadrático médio
RMSE	Raiz do erro quadrático médio
$R^2$	Coefficiente de determinação

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
<b>2</b>	<b>CONCEITOS FUNDAMENTAIS E REVISÃO BIBLIOGRÁFICA</b>	<b>18</b>
2.1	PROPRIEDADES MECÂNICAS DE MATERIAIS	18
2.1.1	Tensão e deformação	18
2.1.2	Tensão normal	18
2.1.3	Deformação elástica	19
2.2	CONCRETO	21
2.2.1	Concreto de agregado leve	22
2.2.2	Propriedades mecânicas do concreto de agregado leve	24
2.3	PREVISÃO DAS PROPRIEDADES MECÂNICAS DO CONCRETO	26
2.4	OTIMIZAÇÃO DE PARÂMETROS DE MÉTODOS DE PREVISÃO	28
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>31</b>
3.1	BASE DE DADOS EXPERIMENTAL	31
3.2	MÉTODOS DE REGRESSÃO	34
3.2.1	Redes Neurais	35
3.2.2	Máquina de Aprendizado Extremo	38
3.2.3	Árvores de decisão	39
3.2.4	Máquinas de Vetores Suporte	42
3.3	OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS	44
<b>4</b>	<b>PROPOSTA</b>	<b>48</b>
4.1	VALIDAÇÃO CRUZADA	48
4.2	MÉTRICAS DE AVALIAÇÃO	48
4.3	BUSCA DOS PARÂMETROS	50
4.4	CÓDIGO COMPUTACIONAL	52
<b>5</b>	<b>RESULTADOS</b>	<b>55</b>
5.1	ANÁLISE DAS MÉTRICAS DE AVALIAÇÃO	55
5.2	ANÁLISE DAS ESTIMATIVAS DOS MODELOS	58
5.3	OTIMIZAÇÃO DOS PARÂMETROS	65
5.4	IMPORTÂNCIA DOS ATRIBUTOS	66
5.5	ANÁLISE DO TEMPO DE EXECUÇÃO	68
<b>6</b>	<b>CONCLUSÃO</b>	<b>70</b>

**REFERÊNCIAS** ..... 72

## 1 INTRODUÇÃO

O concreto tem sido um dos materiais estruturais mais empregados em todo o mundo na construção civil. Como resultado do rápido crescimento econômico e da urbanização, a demanda e produção do cimento cresceram rapidamente nas últimas décadas (Gao et al., 2017; Shen et al., 2017; Igliński and Buczkowski, 2017). O concreto possui várias características únicas, que o faz extremamente útil em praticamente qualquer tipo de projeto da construção civil. Dentre estas características vantajosas de utilizar o concreto podemos citar:

- adaptação a qualquer tipo de forma permitindo liberdade à concepção arquitetônica;
- solução para se obter estruturas monolíticas e hiperestáticas;
- grande durabilidade e baixíssimo custo de manutenção e conservação;
- resistência a efeitos térmicos, atmosféricos e a desgastes mecânicos.

Por outro lado, um fator que deve ser levado em conta, e pode ser uma desvantagem do concreto, é o seu peso próprio. Dependendo dos materiais utilizados na mistura do concreto o seu peso próprio pode ser muito alto para determinadas aplicações, então buscar diminuir o peso final do concreto é um objetivo em comum dos pesquisadores da área de engenharia de estruturas.

Dentre os componentes básicos do concreto estão o cimento, a água, os agregados e aditivos em determinadas proporções. Os agregados são uns dos grandes responsáveis pelo peso do concreto. Normalmente cerca de 75% do volume de um elemento estrutural de concreto é constituído de agregados (Neville, 2013).

Nesse contexto, o uso de concretos com agregados de menor peso específico surge como a solução alternativa para o problema do alto peso próprio do material. O concreto feito com agregados leves, com baixo peso específico em comparação aos usuais, é chamado de concreto de agregado leve. O peso específico dos concretos com agregados convencionais varia entre 2200 kg/m<sup>3</sup> e 2600 kg/m<sup>3</sup>, e o do concreto de agregado leve estrutural entre 1350 kg/m<sup>3</sup> e 1850 kg/m<sup>3</sup>, podendo representar então uma redução de até aproximadamente 50% do seu peso dependendo dos materiais utilizados.

Além da redução do peso próprio, a baixa condutividade térmica do concreto de agregado leve proporciona um desempenho significativamente superior em situação de incêndio quando comparado aos concretos de peso específico convencional (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 artificiais em geral são estáveis a temperaturas elevadas, uma vez que usualmente são fabricados em temperaturas superiores a 1100°C.



Assim como em qualquer outro material utilizado na engenharia, é preciso conhecer bem as propriedades mecânicas do concreto de agregado leve antes de sua utilização. O usual é realizar ensaios laboratórios controlados para encontrar tais propriedades. No concreto de agregado leve, a resistência à compressão, o módulo de elasticidade e o peso específico são as propriedades mecânicas mais importantes e conseqüentemente as mais comumente analisadas. Se possível, prever essas propriedades com um baixo erro ao invés de serem realizados ensaios laboratórios ou até mesmo reduzir o número de ensaios é um objetivo em comum dos engenheiros e pesquisadores da área. Assim, pode-se propor diferentes metodologias e algoritmos de inteligência computacional que realizam essas tarefas.

Para realizar previsões de algum fenômeno biológico, químico, físico entre outros, os algoritmos de inteligência computacional mais utilizados precisam de uma base de dados que possuam os resultados reais desses fenômenos para que possam realizar o seu aprendizado. No caso desse trabalho, o interesse é em prever as propriedades mecânicas do concreto de agregado leve, assim, faz-se necessário a utilização de dados de experimentos de laboratório (ensaios mecânicos) da propriedade a ser prevista.

Esse tipo de aprendizado, quando se tem um conjunto de dados com diversas características e uma saída conhecida, é dito supervisionado. Esses métodos são chamados também de métodos de aprendizado de máquina. Um dos métodos mais utilizados na literatura para esse tipo de problema é a rede neural artificial. Como as propriedades mecânicas são valores reais, denominamos que a previsão é uma regressão. Os algoritmos de regressão utilizam esses dados de experimentos para realizar o treinamento e, após o treinamento, o algoritmo pode ser aplicado em dados de testes para que possa realizar as previsões.

Diferentes técnicas preditivas, como as redes neurais artificiais, máquinas de Vetores Suportes, árvores de Decisão entre outras, já são amplamente utilizadas na previsão de propriedades mecânicas do concreto convencional, no de alto desempenho, em concretos de agregados reciclados e outros (Chou et al., 2014; Behnood et al., 2017; de Melo and Banzhaf, 2017; González-Taboada et al., 2016; Chiew et al., 2017). Para o concreto de agregados leves, alguns trabalhos já foram desenvolvidos utilizando as redes neurais artificiais e as máquinas de vetores suporte, como em Alshihri et al. (2009) e Abd and Abd (2017), mas ainda há uma carência de trabalhos na literatura que explorem e comparem diferentes técnicas preditivas para prever as suas propriedades mecânicas.

As técnicas de previsão podem ser utilizadas também em conjunto com algum algoritmo de otimização, visando a otimização dos seus parâmetros, de modo a obter o melhor desempenho preditivo da técnica. Diversas técnicas para a otimização de parâmetros de métodos de predição já foram testadas em diversas aplicações na literatura, como algoritmos evolutivos (Lin et al., 2008; Oliveira et al., 2010; Huang and Dun, 2008), modelos Bayesianos (Snoek et al., 2012), entre outros. O desenvolvimento da metodologia proposta neste trabalho, vem complementar os estudos encontrados atualmente na literatura através do emprego de diferentes técnicas de

predição, acoplando a elas uma técnica de otimização.

Logo, o objetivo deste trabalho é implementar, otimizar 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. Os métodos de predição utilizados no trabalho serão: redes neuronais artificiais (RNA), máquinas de vetores suporte para regressão (SVR), máquinas de aprendizado extremo (ELM) e árvores de decisão (DT). O processo de otimização se dá através do algoritmo de otimização por enxame de partículas (PSO), no qual os parâmetros de cada método são escolhidos de forma a minimizar o erro de predição. Espera-se que o resultados obtidos pelos métodos sejam próximos ao do medido em laboratório, conseguindo dessa forma propor o uso desses métodos para que seja possível reduzir a quantidade de ensaios laboratoriais, que demanda tempo e também o retrabalho.

O texto deste trabalho está organizado como segue. O capítulo 2, apresenta uma revisão bibliográfica dos conceitos envolvidos no trabalho, como as propriedades mecânicas dos materiais, o concreto de agregado leve e os métodos utilizados para prever as propriedades do concreto. Os métodos de regressão, o algoritmo de otimização PSO e a base de dados utilizada são apresentados no capítulo 3. No capítulo 4 é descrita toda a metodologia de validação proposta do trabalho, todas as características utilizadas nos métodos de predição e na otimização dos parâmetros. O capítulo 5 descreve os resultados obtidos e uma breve discussão dos mesmos. E por fim, no capítulo 6, a conclusão do trabalho é apresentada em conjunto com os trabalhos futuros.

## 2 CONCEITOS FUNDAMENTAIS E REVISÃO BIBLIOGRÁFICA

### 2.1 PROPRIEDADES MECÂNICAS DE MATERIAIS

Em diversos tipos de projetos da engenharia, é preciso determinar e conhecer as propriedades dos materiais e componentes utilizados nos mesmos. As propriedades mecânicas são responsáveis pela caracterização do comportamento do material quando está submetido a solicitações mecânicas externas. Essas propriedades estão relacionadas com a capacidade do material de resistir, deformar e falhar ao ser exposto a esforços mecânicos.

As propriedades mecânicas dos materiais são, em geral, obtidas através de experimentos laboratoriais extremamente controlados que buscam reproduzir um ambiente o mais próximo possível das condições reais de aplicação do material. Na realização desses ensaios é levado em conta as condições de uso na que o material estará sujeito, definidas pelos tipos de esforços (tração, compressão, cisalhamento), magnitude, duração de aplicação, temperatura do ambiente e outros. A porção de material utilizada é chamada de corpo de prova, e suas características geométricas são definidas utilizando normas específicas para o tipo de ensaio utilizado.

Dentre as propriedades mecânicas, podemos citar as mais comumente analisadas, como a resistência mecânica (à tração ou à compressão), elasticidades, ductilidade, dureza, tenacidade e outras. Neste trabalho estaremos interessados na resistência à compressão e o módulo de elasticidade. A resistência à compressão está relacionada à capacidade do material resistir e distribuir as forças de compressão exercidas sobre o mesmo. Já o módulo de elasticidade está relacionado à predisposição do material em deformar quando está sujeito a esforços, e de após a solicitação das forças de voltar a sua forma original.

#### 2.1.1 Tensão e deformação

Quando um objeto formado por um determinado material se deforma, associamos ao tipo de deformação uma grandeza chamada tensão. A tensão descreve a intensidade da força interna sobre um plano específico (área) que passa por um ponto (Hibbeler, 2010). A intensidade da força, ou força por unidade de área, que age perpendicularmente à área, é definida como tensão normal, que pode ser de tração ou compressão. A intensidade da força, ou força por unidade de área, que age tangente à área, é denominada tensão de cisalhamento.

#### 2.1.2 Tensão normal

Para elementos estruturais cilíndricos, com cargas axiais aplicadas às extremidades do elemento e com todas as seções transversais iguais (barra prismática), como visto na Figura 1, a tensão normal média nos pontos na área da seção transversal é dada por:

$$\sigma = P/A \quad (2.1)$$

onde  $P$  é força normal interna resultante, aplicada no centroide da área da seção transversal e  $A$  é área da seção transversal do elemento. A unidade de tensão no SI é o Pascal ( $Pa$ ).

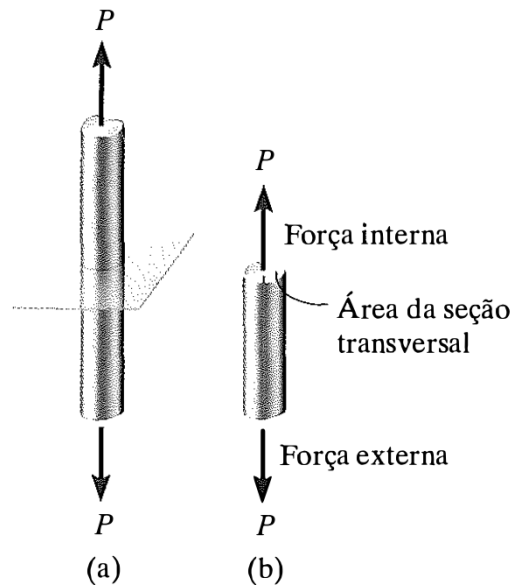


Figura 1 – Barra prismática com cargas axiais nas extremidades. Adaptado de (Hibbeler, 2010).

### 2.1.3 Deformação elástica

Para a maioria dos materiais, em regime de pequenas deformações, podemos considerar que a relação entre a tensão e a deformação é linear. O alongamento ou contração de um segmento de reta por unidade de comprimento é denominado deformação normal.

Assim para o caso de um corpo de prova prismático, como os da Figura 2, a relação pode ser escrita como:

$$\sigma = E\varepsilon \quad (2.2)$$

onde  $\varepsilon$  é a deformação normal (adimensional) e  $E$  é a constante de proporcionalidade denominado módulo de elasticidade ou módulo de Young. A unidade do módulo de elasticidade no SI é o  $GPa$  (Giga Pascal). Nesse caso, a deformação pode ser escrita como

$$\varepsilon = \frac{l - l_0}{l_0} = \frac{\Delta l}{l_0} \quad (2.3)$$

onde  $l_0$  é o comprimento do corpo de prova original,  $l$  é o comprimento com a aplicação do esforço e  $\Delta l$  é a variação do comprimento.

O ensaio que é utilizado para encontrar o módulo de elasticidade e a resistência mecânica é o de tração ou o de compressão. Na Figura 3 temos um diagrama de tensão-deformação de um ensaio de tração em um corpo de prova até o seu rompimento. Para o ensaio de compressão o comportamento geralmente é análogo, mudando apenas a forma como o corpo falha.

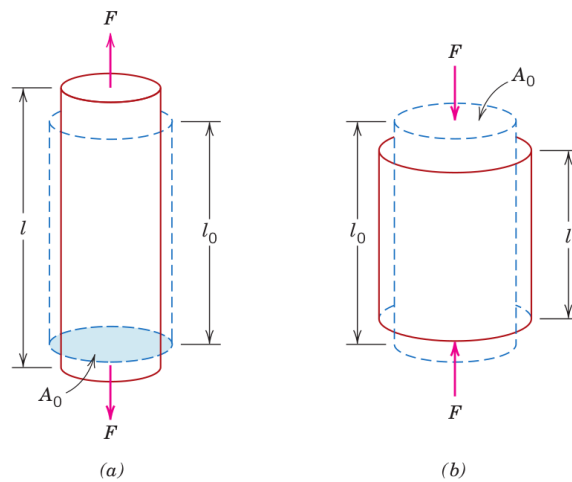


Figura 2 – Deformações de um corpo de prova cilíndrico de um material com comportamento elástico submetido a uma carga de tração uniaxial (a) e uma carga de compressão uniaxial (b). Adaptado de (Callister and Rethwisch, 2011).

No ensaio de tração ou compressão, um corpo de prova do material a ser estudado é submetido a esforços (de compressão ou tração) axiais nas extremidades do corpo. A partir do ensaio é possível montar um diagrama tensão-deformação, utilizando a equação 2.1 para encontrar os valores de tensão e utilizando um extensômetro podemos encontrar as deformações, dado a mudança do comprimento da barra, utilizando a equação 2.3. O valor de  $\varepsilon$  é dado pela inclinação da reta do gráfico para pequenas deformações.

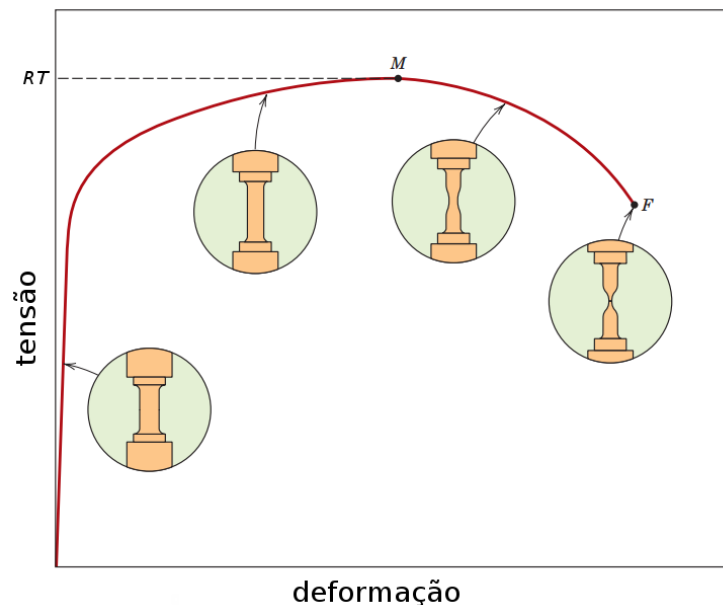


Figura 3 – Comportamento da curva tensão-deformação até a fratura do material para um ensaio de tração. No ponto  $M$  temos a tensão máxima de tração (Resistência a tração)  $RT$ . Adaptado de (Callister and Rethwisch, 2011).

Para materiais frágeis, como o concreto, a tensão máxima de compressão é maior que

a tensão máxima de tração. Assim, quando é preciso suportar esforços de tração, o concreto é projetado com barras ou hastes de aço para aumentar a sua resistência a tração. Um exemplo de diagrama de tensão deformação para o concreto pode ser observado na Figura 4, onde pode ser vista a diferença entre as tensões máximas de compressão e tração.

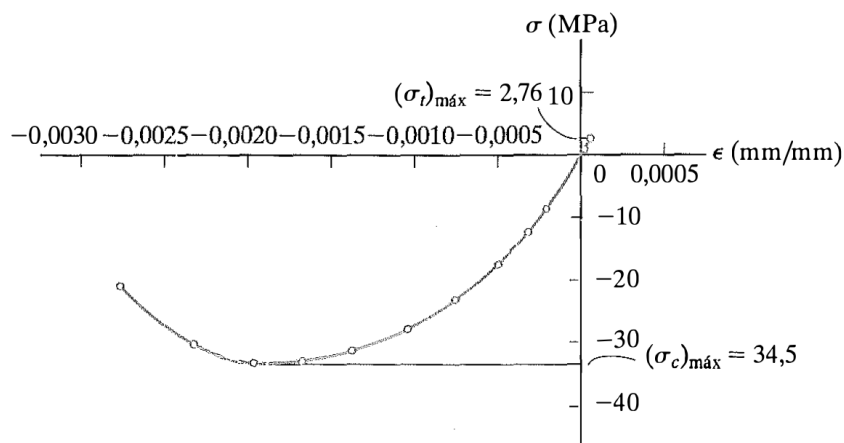


Figura 4 – Diagrama tensão-deformação para mistura de concreto. Adaptado de (Hibbeler, 2010).

## 2.2 CONCRETO

Um dos materiais mais empregados na construção civil é o concreto. O concreto pode ser definido, de forma bem genérica, como qualquer produto ou massa que possui em sua produção um meio cimentante (Neville, 2013). Na maioria das vezes, esse meio cimentante é o resultado da reação de um cimento com a água. Os componentes básicos de um concreto são: o cimento, água, agregados e aditivos.

Os agregados podem ser graúdos, geralmente pedras obtidas por britagem de rochas, e miúdos, em geral areia. A mistura de cimento e água é chamada pasta, e a mistura da pasta com o agregado miúdo é denominada argamassa. A adição dos agregados se dá para a redução de custo, visto que o custo da pasta é muito superior ao dos agregados, e também para melhorar a estabilidade e durabilidade. Normalmente, cerca de 75% do volume do concreto é constituído de agregados (Neville, 2013).

As propriedades mecânicas, térmicas e químicas dos agregados influenciam diretamente nas propriedades mecânicas do concreto no estado fresco e endurecido. O estado fresco é a fase após a mistura em que o concreto começa o seu endurecimento. No estado endurecido é desejável que o concreto tenha uma boa resistência a compressão, visto que na maioria das aplicações o material fica sujeito a esforços de compressão.

### 2.2.1 Concreto de agregado leve

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<sup>3</sup>. 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:

- resistência à compressão superior a 28 MPa aos 28 dias e massa específica não superior a 1840 kg/m<sup>3</sup>;
- resistência à compressão superior a 21 MPa aos 28 dias e massa específica não superior a 1760 kg/m<sup>3</sup>;
- resistência à compressão superior a 17 MPa aos 28 dias e massa específica não superior a 1680 kg/m<sup>3</sup>.

A partir da origem dos materiais utilizados, pode-se dividir os tipos de agregado leve em duas categorias distintas (Chandra and Berntsson, 2002):

1. Aqueles que são encontrados naturalmente e estão prontos para o uso somente com tratamento mecânico, isto é, britagem e peneiração.
2. Aqueles produzidos por processos industriais por tratamento térmico de materiais naturais ou de subprodutos industriais, resíduos, etc.

Diversos tipos agregados leves são utilizados hoje em dia, dentre as duas categorias apresentadas. A escolha do agregado a ser utilizado depende sempre da aplicação e as características mecânicas finais desejadas.

Os agregados mais comuns dentre os naturais fabricados por processos de tratamento térmico são: perlita, vermiculita, argila, xisto, ardósia, entre outros. Já os provenientes de subprodutos industriais são as cinzas volantes pulverizadas, escória de alto forno, resíduos industriais, lamas, etc. Estes são produzidos por um processo de expansão (inchaço) ou aglomeração. Um exemplo de agregado produzido por esse processo é mostrado na Figura 5, onde são apresentados exemplos de agregados leves formados por argila expandida.

Dentre os agregados naturais, que são aplicados apenas a tratamentos mecânicos, os mais utilizados são: as pedra-pomes, a escória e até materiais orgânicos provenientes dos rejeitos da agricultura.



Figura 5 – Imagens de agregados de argila expandida (Bogas and Gomes, 2013).

Na Figura 6 são apresentadas fotos de diversos agregados de argila expandida seccionados que são utilizados na confecção do concreto leve. É possível observar as porosidades dos agregados, que ocorre devido ao seu processo de fabricação, responsável então por um maior volume de preenchimento do concreto com um menor peso próprio final.

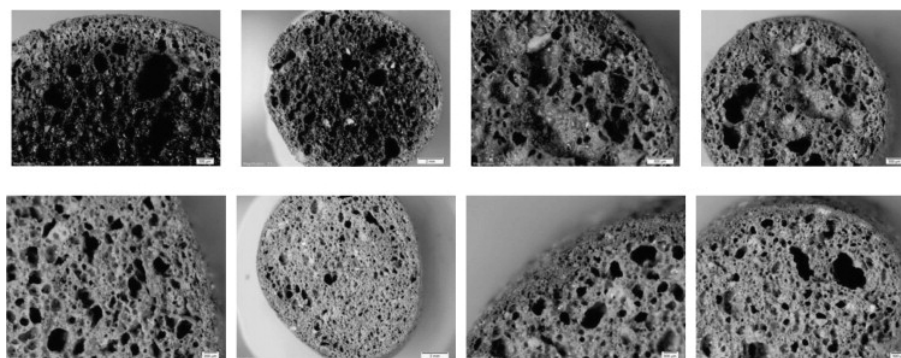


Figura 6 – Estrutura interna de agregados de argila expandida. Retirado de (Ardakani and Yazdani, 2014)

O peso específico dos concretos que utilizam agregados convencionais varia entre  $2200 \text{ kg/m}^3$  e  $2600 \text{ kg/m}^3$ , e a do concreto estrutural leve entre  $1350 \text{ kg/m}^3$  e  $1850 \text{ kg/m}^3$  (Neville, 2015), a classificação dos concretos leves é apresentada na 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 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.

O concreto de agregado leve é utilizado desde os tempos antigos em grandes aplicações da construção civil. Uma das principais aplicações antigas foi no templo Romano de *Pantheon* (Figura 7), construído entre os anos de 118 à 128 utilizando agregados leves de pedra-pomes e está em boas condições até os dias de hoje (Chandra and Berntsson, 2002). Além do Pantheon, tantas outras construções antigas gregas e romanas, que foram construídas utilizando de agregados leves de pedra-pomes, estão de pé até hoje. Esse fato indica a boa durabilidade dessas estruturas. Nos dias de hoje, grandes construções em diversas partes do mundo utilizam agregados leves para a confecção do concreto, dentre as aplicações pode-se citar: estádios de futebol, pontes,



Tabela 1 – Classificação dos concretos leves a partir da peso específico. Dados de Neville (2015)

Classificação	Peso específico (kg/m <sup>3</sup> )	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

galpões industriais, casas, túneis, estações de extração de petróleo e muitas outras.



Figura 7 – Templo romano Pantheon, construído utilizando agregados leves de pedra-pomes no concreto (Chandra and Berntsson, 2002).

### 2.2.2 Propriedades mecânicas do concreto de agregado leve

Para encontrar as propriedades mecânicas do concreto, é preciso realizar ensaios diferentes mecânicos. O mais utilizado é o de compressão uniaxial. Esses ensaios devem ser realizados em mais de um corpo de prova de concreto, com composição previamente definida, a fim de encontrar o valor final das propriedades através da média dos valores de cada ensaio. Através disso, é possível encontrar um valor mais confiável para a propriedade mecânica visto que, pode haver variações na sua composição na prática, como por exemplo o tamanho e quantidade dos agregados.

Em geral, para o concreto de agregado leve, há uma relação entre o fator água/cimento 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 a resistência à compressão do concreto (da Silva, 2013; Neville, 2015). Vários estudos mostraram também que a resistência à compressão é determinada, independentemente do tipo de concreto, não apenas

para a relação água-cimento, mas também para outros componentes (Kasperkiewicz et al., 1995; Faust, 2000; Ni and Wang, 2000; Papadakis and Tsimas, 2002; Chandra and Berntsson, 2003; Kockal and Ozturan, 2011; Duan et al., 2013b).

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 convencional, 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.

Ensaaios 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. 8 apresenta uma comparação entre concreto de peso específico normal e o concreto de agregado leve quando estão sujeitos 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 a argamassa possui maior rigidez em relação aos agregados.

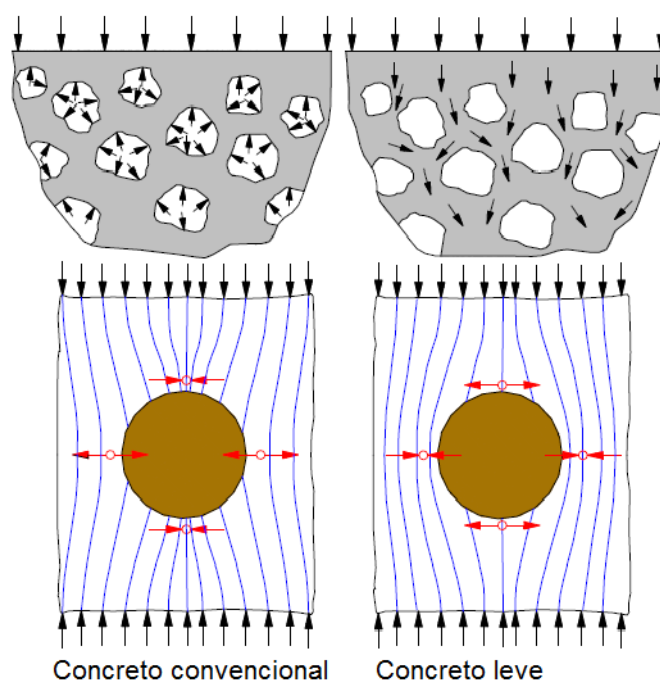


Figura 8 – Transferência das forças de compressão no concreto normal e leve (Bogas and Gomes, 2013).

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 pode ser constituído de diferentes tipos de agregado leves (Tabela 1), além das diversas variações na composição da argamassa, como o fator água-cimento e quantidade de cimento. Esses diversos tipos de agregados que podem ser utilizados, além da não linearidade das estruturas de concreto, complicam o cálculo analítico da resistência à compressão, do módulo de elasticidade, entre outras propriedades. Dessa forma, modelos analíticos de propriedades do concreto podem não ser suficientes para atender as exigências do projeto de concretos de agregados leves.

As equações empíricas apresentadas para a estimativa da resistência à compressão e do módulo de elasticidade são baseadas 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.

### 2.3 PREVISÃO DAS PROPRIEDADES MECÂNICAS 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. Esse estudos mostram que várias técnicas conseguem resultados muito promissores na previsão de propriedades mecânicas do concreto. Os métodos de previsão conseguem uma boa generalização dos dados das propriedades do concreto na maioria dos casos, fazendo com que as metodologias propostas possam trazer uma diminuição do trabalho em laboratório, o que demanda muito tempo e um alto custo. Alguns trabalhos desenvolvidos nesse contexto são citados a seguir em uma breve revisão bibliográfica.

Em (Kasperkiewicz et al., 1995) foi utilizado um tipo de rede neural artificial para prever a resistência a compressão à 28 dias de concretos de alto desempenho. A rede neural utilizada foi do tipo fuzzy-ARTMAP, que é usada para mapeamentos multidimensionais de um espaço em uma variável unidimensional. Foi utilizado para a validação da metodologia um conjunto de dados com 340 amostras obtidos na literatura, sendo utilizados como variáveis de entrada dos dados os seguintes componentes do concreto: cimento, sílica, super plastificante, água, agregado

miúdo e agregado graúdo. Para realizar a validação do método, os dados foram divididos em duas partes para realizar a validação do método, sendo 200 usados para o treinamento e 140 para testar o método. Os resultados obtidos mostraram que o problema pode ser modelado efetivamente pela rede neural, obtendo bons resultados de correlação entre os valores preditos e reais. Além disso, é mostrado que a rede neural pode ser ainda utilizada para otimização das propriedades da mistura do concreto visando um melhor valor de resistência a compressão do concreto projetado.

No trabalho de (Ni and Wang, 2000), foram desenvolvidos dois métodos para a predição da resistência a compressão do concreto, as redes neuronais artificiais e a programação genética. Os métodos utilizaram dados de treinamento obtidos de experimentos de laboratório, em períodos de cura de 28, 56 e 91 dias. Para realizar a validação foram utilizados dados retirados na literatura. A programação genética além de ser usada para prever a resistência a compressão, ela tem a capacidade de encontrar uma expressão matemática para a resistência a compressão. Após a execução dos métodos, concluiu-se que a rede neural com algoritmo de treinamento Levenberg-Marquardt obteve os melhores resultados, chegando em vários casos a um coeficiente de correlação  $R^2$  maior que 0.9.

Existem trabalhos também visando a previsão de propriedades de concreto de agregados reciclados, em (Duan et al., 2013c) é utilizada redes neuronais artificiais para a previsão do módulo de elasticidade, através de 324 conjuntos de dados coletados de 21 literaturas internacionais publicadas e também de 16 conjuntos de dados experimentais do próprio autor. Os dados possuíam 16 parâmetros de entrada, caracterizados pela mistura do concreto e 1 uma saída, o valor do módulo de elasticidade. Os resultados encontrados indicaram que rede consegue uma excelente generalização dos dados, encontrando um valor de coeficiente de correlação  $R^2$  mais que 0.99 nas bases de dados testadas.

Em um trabalho com metodologia semelhante, em (Duan et al., 2013a), foram utilizadas 146 bases de dados da literatura para a previsão da resistência a compressão para o concreto com agregados reciclados. Neste caso foram utilizados 14 parâmetros de entrada. Para a previsão da resistência a compressão, a rede neural também obteve ótimos resultados, com coeficiente de correlação  $R^2$  superior a 0.99. Em ambos os trabalhos, é interessante notar que a rede neural obteve esses resultados de um conjunto de dados de diferentes tipos e fontes de agregado reciclado, mostrando assim o seu alto poder de aprendizado e aplicabilidade nesses tipos de problemas.

Para a previsão da resistência a compressão do concreto de agregado leve, o trabalho de (Alshihri et al., 2009) utilizou dois modelos de rede neuronais artificiais. Foi utilizado dados de laboratório para resistência a compressão após 3, 7, 14 e 28 dias de cura em função de 8 variáveis de entrada: areia, fator água/cimento, agregado miúdo leve, agregado graúdo leve, fumaça de sílica utilizada em solução, fumo de sílica em adição ao cimento, super plastificante e o período de cura. Os modelo de rede utilizado foram a *feed-forward* com treinamento *back-propagation*

e a *cascade-correlation*. Os dados foram divididos em duas partes, 70% para o treinamento e 30% para o teste das redes. Além disso, visando um melhor treinamento da rede, foi feita uma normalização linear nos dados, fazendo com que os valores de entrada e saída fiquem normalizados entre 0 e 1. Com os resultados das previsões, o autor concluiu que os modelos de rede neural são suficientes para prever a resistência a compressão do concreto de agregado leve. Os resultados indicam que os modelos conseguem um erro absoluto médio menor que 5% nos dados de teste utilizados.

Em (Abd and Abd, 2017), foram utilizadas a máquina de vetores suporte para regressão (SVR) e a regressão não linear multivariada para a previsão da resistência a compressão do concreto celular leve. O concreto celular é uma nova geração de concreto leve. Foram utilizados dados experimentais da resistência à compressão após 28 dias de cura. A base de dados possuía 150 dados experimentais, tendo como entrada seis características constitutivas do concreto celular. Foram utilizadas diferentes funções kernel no SVR, e os resultados indicaram que a função de base radial (RBF) obtiveram os melhores resultados e também foram superiores à regressão não linear multivariada. O coeficiente de correlação entre as previsões e os dados experimentais chegaram a 0.99 com o uso do kernel RBF, já o erro absoluto médio foi inferior a 2%. Os resultados indicam que o SVR é um algoritmo muito promissor para a previsão de propriedades mecânicas do concreto e que deve ser ainda mais explorado para diferentes tipos de concreto.

Incentivado por esses trabalhos mencionados, além das redes neurais, que é o método mais amplamente utilizado, como pode-se observar em tantos outros trabalhos publicados na literatura, outras técnicas preditivas de aprendizado de máquina, tais como as definidas neste trabalho (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 criteriosa de cada técnica em um conjunto de dados experimental, utilizando um algoritmo de otimização para encontrar os parâmetros ótimos de cada método, podem indicar qual técnica possui a melhor generalização do problema real, o que é feito neste trabalho.

## 2.4 OTIMIZAÇÃO DE PARÂMETROS DE MÉTODOS DE PREVISÃO

A maioria das técnicas de inteligência computacional para aprendizado supervisionado necessitam de escolha de parâmetros para realizar o treinamento. O ajuste desses parâmetros não é uma tarefa simples, e muitas das vezes é feito de forma empírica. Desenvolver técnicas automáticas então é uma tarefa comum em trabalhos que buscam utilizar as técnicas de previsão em diversas aplicações, visando a otimização dos parâmetros dos métodos, de modo a obter o melhor desempenho preditivo da técnica.

A técnica mais simples é a de *grid-search* (busca em grade) (Lerman, 1980), onde é feita uma busca exaustiva em um conjunto de parâmetros a serem testados. A grande desvantagem dessa abordagem é a necessidade de se testar todas as combinações existentes entre os parâmetros,

o que pode acarretar em um alto custo computacional. Além disso, só é possível testar valores discretos, o que pode não ser o suficiente para representar os parâmetros que possuem valores contínuos.

Então, para melhorar o desempenho desse processo de busca, é comum a utilização de algoritmos de otimização heurísticos, que muitas das vezes possui desempenho superior ao do *grid-search*, visto que o número de testes pode ser muito inferior ao número total de combinações do *grid-search* e além disso a busca pode ser muito mais refinada, podendo ser realizada de forma contínua em parâmetros que são números reais. Diversas técnicas para a otimização de parâmetros de métodos de predição já foram testadas em diversas aplicações na literatura, a seguir, são citados algumas dessas aplicações, que servem de inspiração para a metodologia proposta neste trabalho.

Em (Snoek et al., 2012), é desenvolvida uma metodologia para a otimização dos parâmetros em três métodos de aprendizado de máquina: *Latent Dirichlet Allocation* (LDA), Máquina de vetores suporte (SVM) e Redes neurais convolucionais (RNC). A metodologia proposta utiliza a estrutura da otimização Bayesiana, na qual o desempenho da generalização do algoritmo de aprendizagem é modelado como amostra de um processo Gaussiano (GP). Diversas abordagens e algoritmos são desenvolvidos utilizando a abordagem proposta e são testados em comparação a técnica de *grid-search* nos três métodos de predição. Para todos os casos o *grid-search* obtém um desempenho inferior aos algoritmos desenvolvidos, mostrando assim que a técnica é pouco competitiva com abordagens mais inteligentes. Além disso, em comparação a um perito humano, os algoritmos mostram que podem alcançar ou ultrapassar o nível de otimização do perito.

Algoritmos evolutivos também são amplamente testados para a otimização de parâmetros de métodos de previsão. No trabalho de (Oliveira et al., 2010) é testado um algoritmo genético (GA) para a otimização de parâmetros de alguns métodos de aprendizado e também para a seleção de características do problema de esforço de desenvolvimento de *software*. São utilizadas as redes neurais com função de base radial (RBF), redes neurais perceptron de múltiplas camadas (MLP), máquina de vetor suporte para a regressão (SVR) e árvores de decisão. O trabalho tinha como objetivo então o uso do algoritmo genético para simultaneamente selecionar um subconjunto de características de entrada ideal para o problema e otimizar os parâmetros dos métodos de regressão. Em todos os conjuntos de dados testados pelo autor mostraram que o método proposto baseado no algoritmo genético foi capaz de melhorar o desempenho dos métodos de aprendizagem. Com o processo de otimização, mostra-se que os métodos superam alguns métodos relatados na literatura recentemente para a estimativa do esforço de software, evidenciando que o processo de otimização pode impactar positivamente nos resultados de previsão das técnicas.

Em (Huang and Dun, 2008) é proposto um novo modelo híbrido que utiliza a otimização por enxames de partículas (PSO) e máquinas de vetores de suporte (SVM) para melhorar a precisão da classificação através da otimização de parâmetros do SVM e da seleção de

características dos dados experimentais testados. O sistema híbrido proposto também foi implementado através de uma arquitetura distribuída usando a tecnologia do serviço *web* para reduzir o tempo computacional. Os resultados mostram que o PSO foi capaz de realizar a otimização dos parâmetros e características e melhorar o desempenho do SVR. Além disso, mostra-se que a arquitetura distribuída proposta através do serviço *web* pode ser facilmente implementada em um ambiente de computação heterogêneo para diminuir o tempo computacional da metodologia.

Já no trabalho de (Lin et al., 2008), também é utilizado um algoritmo de otimização por enxames de partículas (PSO) para determinação de parâmetros e seleção de características de máquinas de vetor de suporte (SVM). Vários conjuntos de dados públicos são utilizados para calcular a taxa de precisão da classificação e assim poder avaliar a abordagem do PSO em conjunto com o SVM desenvolvida. Os resultados são comparados com a otimização por busca exaustiva (*grid-search*) e com o algoritmo genético. A comparação dos resultados obtidos com as de outras abordagens demonstra que a abordagem PSO em conjunto com o SVM desenvolvida pelo autor possui uma melhor ou equivalente precisão de classificação do que outros métodos testados.

Dado os presentes trabalhos mencionados e muitos outros existentes na literatura nos dias de hoje, foi escolhido o algoritmo de otimização por enxames de partículas (PSO) para a utilização neste trabalho, com o objetivo de otimizar os parâmetros dos métodos testados. O PSO possui desempenho superior ou competitivo a diversas técnicas já testadas na literatura, e também possui a vantagem de ser um algoritmo de simples implementação, o que facilita a replicabilidade da metodologia proposta neste trabalho.

### 3 MATERIAIS E MÉTODOS

Neste capítulo serão detalhados a base de dados e os métodos computacionais que foram utilizados para a realização do processo de previsão das propriedades mecânicas do concreto de agregado leve. Primeiramente será apresentada a base de dados experimental utilizada e, em seguida, serão especificados os métodos de regressão utilizados para prever sobre os dados experimentais e apresentado o método de otimização evolucionária utilizado na busca de parâmetros dos métodos de regressão.

#### 3.1 BASE DE DADOS EXPERIMENTAL

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 por (Ke, 2008).

A base de dados possui amostras de teste de mais de 300 corpos de prova cilíndricos, com quatro parâmetros e duas saídas. Os corpos de prova foram utilizados em ensaios de compressão a fim de se obter as propriedades mecânicas dos concretos de agregado leve. Foram medidos o módulo de elasticidade ( $E_c$ ) e a resistência à compressão ( $f_c$ ) aos 28 dias, em função do fator água/cimento, volume de agregado leve, quantidade de cimento e o peso específico do agregado leve.

Os dados experimentais envolvem três tipos de argamassa ('a', 'b' e 'c'), como pode ser visto na tabela 2, para os cinco tipos diferentes de agregados e cinco níveis de volume de agregado ( $V_a = 0\%$  ou  $12.5\%$  ou  $25\%$  ou  $37.5\%$  ou  $45\%$ ). O cimento utilizado nas argamassas foi o Portland CEM I 52,5R.

Para cada corpo de prova, as duas propriedades possuem até quatro medições, a fim de garantir uma maior confiabilidade nos valores das propriedades obtidos. Assim, para montar uma base de dados de teste para ser utilizado nos métodos de previsão, o valor final de cada propriedade foi dado pela média aritmética das quatro medições, gerando um total de 75 amostras.

Os tipos de agregados leves testados nos experimentos eram três compostos por argila expandida e dois por xisto expandido. Na figura 9 é possível observar dois agregados leve seccionados utilizados nos experimentos. Os cinco tipos de agregados utilizados eram caracterizados também pelos seus diâmetros mínimo ( $d_{min}$ ) e máximo ( $d_{max}$ ), e pelo peso específico aparente ( $\gamma_{ap}$ ), conforme resumido na 3.

Na tabela 4 é apresentado um resumo da variação dos valores e as respectivas unidades de medida dos atributos de entrada e das propriedades mecânicas da base de dados experimental.

A fim de auxiliar a análise dos dados utilizados, são apresentados gráficos de associação



Tabela 2 – Tipos de argamassa utilizadas nos experimentos

Argamassa	Cimento ( $kg/m^3$ )	agua/cimento	Areia ( $kg/m^3$ )	$V_a$ (%)
a	753,90	0.446	1055.46	0.0
	659,66		923.53	12.5
	565,42		791.59	25.0
	471,19		659.66	37.5
	414,64		580.50	45.0
b	811,38	0.350	1135.94	0.0
	709,96		993.95	12.5
	608,54		851.95	25.0
	507,11		709.96	37.5
	446,26		624.77	45.0
c	826,45	0.290	1074.38	0.0
	723,14		940.08	12.5
	619,83		805.79	25.0
	516,53		671.49	37.5
	454,55		590.91	45.0

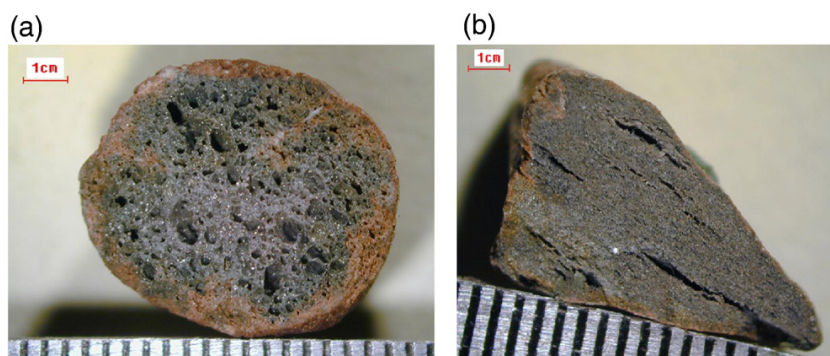


Figura 9 – Seção do agregado argila expandida (a) e seção do agregado xisto expandido em (b). Adaptado de (Ke et al., 2010).

Tabela 3 – Tipos de agregados utilizados nos experimentos.

Tipo	$d_{min}(mm)$	$d_{max}(mm)$	$\gamma_{ap}(kg/m^3)$	Material
1	0	4	650	Argila Expandida
2	4	10	550	Argila Expandida
3	4	10	430	Argila Expandida
4	4	10	520	Xisto Expandida
5	4	8	750	Xisto Expandida

entre os valores dos atributos de entrada utilizados e as propriedades mecânicas, na Figura 10 para o módulo de elasticidade e na Figura 11 para a resistência a compressão. Os pontos em azul representam cada amostra de teste dos dados experimentais para cada propriedade mecânica medida.

Na metodologia proposta neste trabalho, foram utilizados para a realização das previsões

Tabela 4 – Variação das entradas e saídas da base de dados experimental.

Variáveis	Mínimo	Máximo
Água/cimento	0.29	0.446
Quantidade de cimento	415kg/m <sup>3</sup>	826kg/m <sup>3</sup>
Volume do agregado	0%	45%
Peso específico do agregado	$3.18 \times 10^6$ kg/m <sup>3</sup>	$3.60 \times 10^7$ kg/m <sup>3</sup>
Módulo de elasticidade	$1,512 \times 10^4$ MPa	$3,5757 \times 10^4$ MPa
Resistência à compressão	24,10 MPa	88,190 MPa

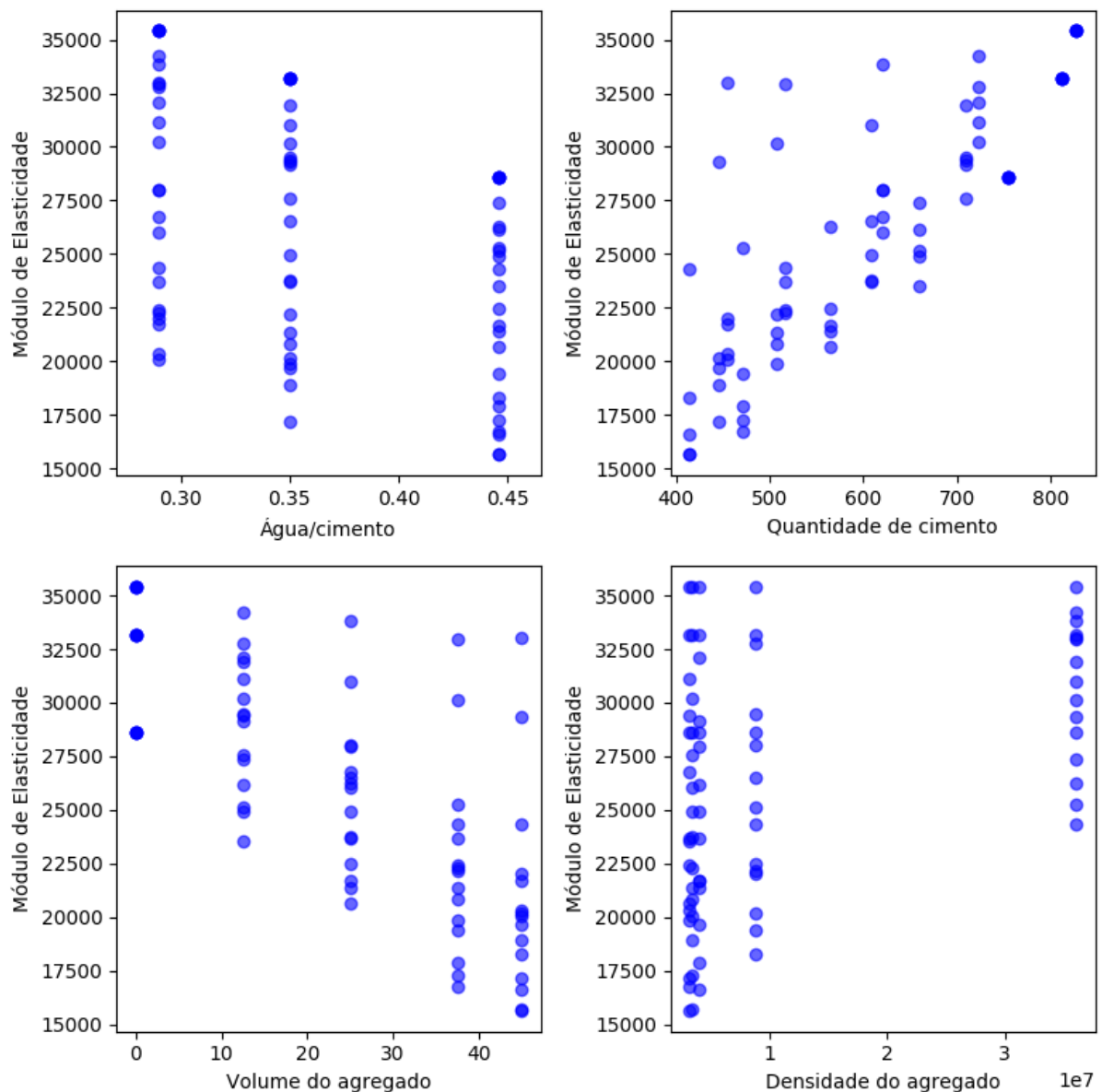


Figura 10 – Gráficos de associação entre os valores dos atributos de entrada das amostras de teste utilizados e o módulo de elasticidade.

os quatro atributos medidos dos dados experimentais do concreto (o fator água/cimento, o volume de agregado leve, a quantidade de cimento e o peso específico do agregado leve). Em

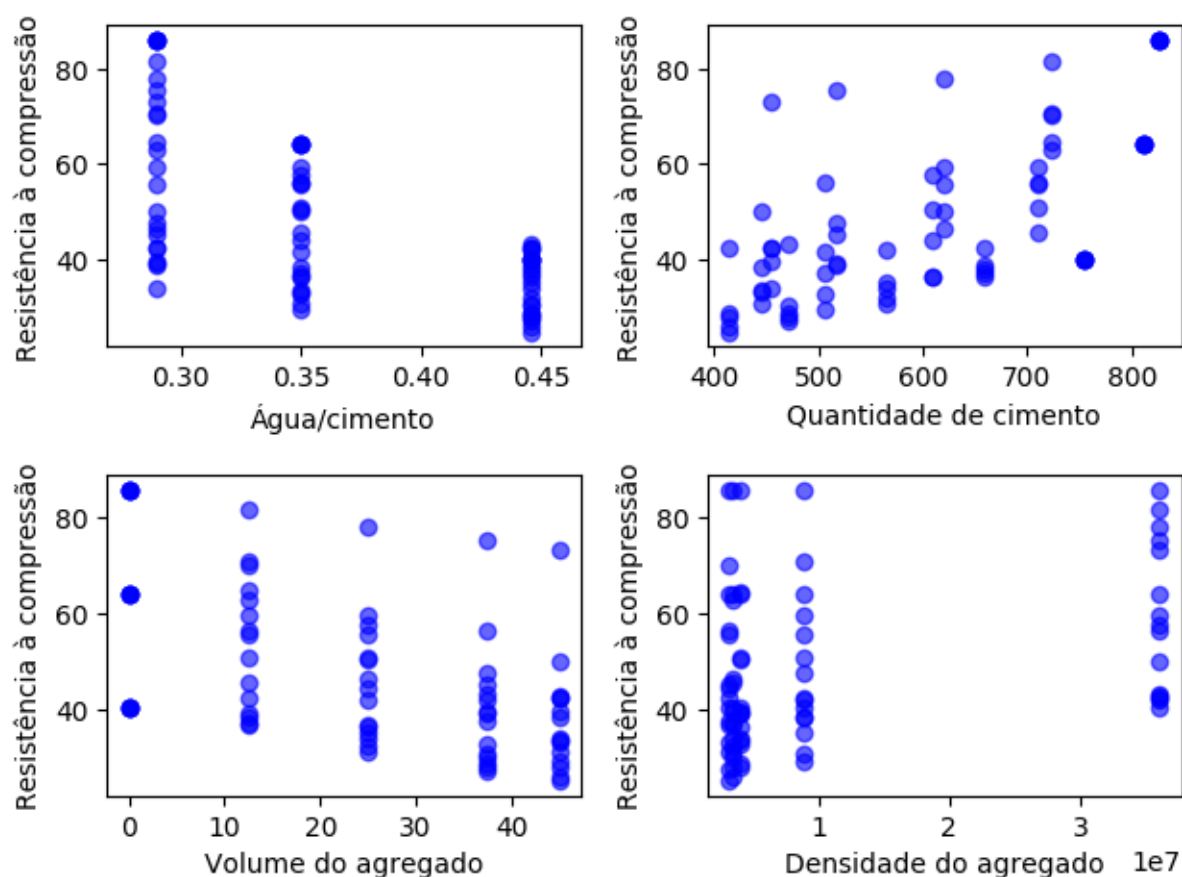


Figura 11 – Gráficos de associação entre os valores dos atributos de entrada das amostras de teste utilizadas e a resistência à compressão.

trabalhos que visam a previsão de propriedades mecânicas do concreto, a maioria das vezes são utilizadas dezenas de atributos de entrada. A proposta deste trabalho é justamente ir contra essa característica, visando buscar um modelo de previsão que obtenha um bom desempenho mesmo com um conjunto mais simples de dados de entrada, mas que ao mesmo tempo esses dados representem de forma satisfatória as características dos concretos testados. Isso facilita o processo de previsão e aplicação dessa metodologia na prática, pois é necessário especificar poucos parâmetros para encontrar as previsões das propriedades.

### 3.2 MÉTODOS DE REGRESSÃO

A inteligência computacional é uma grande área de pesquisa que abrange diversas áreas de estudos, focadas no desenvolvimento de técnicas computacionais com o objetivo de aprender sobre dados. O problema mais típico a ser resolvido é o de previsão. Os algoritmos que resolvem esses problemas de previsão estão contidos na subárea intitulada de aprendizado de máquina.

A partir de um conjunto de dados, que podem possuir diversas características (atributos ou variáveis) e uma ou mais saídas mensuradas à partir das características, é possível treinar

algoritmos para que possam desenvolver a capacidade de aprender sobre os dados e assim no futuro possam prever as saídas para novos dados. Esse tipo de aprendizado é dito supervisionado, pois é necessário conhecer previamente as saídas de um conjunto de dados de teste. As saídas desses dados podem ser discretas (categóricas) ou contínuas (quantitativas), definindo assim dois tipos de problemas de previsão, a classificação e a regressão, respectivamente. Neste trabalho serão utilizados métodos de regressão, visto que as saídas da base de dados são contínuas. Muitos algoritmos de regressão são desenvolvidos a partir de modificações ou adaptações de ideias iniciais para resolver o problema de classificação.

Formalmente, os métodos de aprendizado de máquina para regressão são algoritmos de aprendizado supervisionado que buscam resolver o problema de encontrar um modelo de predição  $f$ , de forma que a partir de uma entrada de dados  $X$  e uma saída  $Y$ , a relação  $f(X) \approx Y$  exista, e que  $f(X) - Y \rightarrow 0$ . Para isso, é comum nos algoritmos de regressão a resolução de um subproblema de minimização da função  $J(X, Y) = f(X) - Y$ , ou alguma variação dela, sendo  $J$  definida muitas vezes por função de perda. Essa função de perda sempre está associada a alguma medida de erro do método, com o objetivo de minimizá-lo.

Após o processo de aprendizado de um método de previsão, podemos definir então que o método modela os dados reais, que podem ser representados por uma função  $f$ , com uma função aproximada  $\hat{f}$ . Logo, podemos chamar o resultado do aprendizado do método como um modelo de regressão. No texto deste trabalho, quando for citada a palavra “modelo” está sendo citado então o resultado de um método de regressão.

A seguir, serão detalhados os quatro métodos de regressão utilizados neste trabalho para o problema de previsão das propriedades mecânicas do concreto de agregado leve a partir de informações dos constituintes do concreto medidos em laboratório. Os métodos foram escolhidos com base em sua vasta aplicabilidade já testada na literatura, sendo sempre os mais utilizados e/ou que apresentam melhores resultados. Outro motivo é que os métodos apresentados a seguir, como são comumente utilizados, podem ser encontrados em diversas bibliotecas de aprendizado de máquina de várias linguagens de programação, facilitando assim a replicabilidade da proposta deste trabalho.

### 3.2.1 Redes Neurais

Uma Rede Neuronal 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 neuronais 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, denominados neurônios;
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  $w_{jk}$ ;
4. uma função de ativação  $\varphi$ , que determina o novo estado de ativação de cada unidade de processamento  $z_k = \varphi(s_k)$  em função da combinação linear  $S_k = \sum w_{kj}x_j$  desta unidade;
5. e uma saída  $d$  final de toda a rede.

A Figura 12 mostra uma rede neuronal com os componentes citados acima e com o detalhamento de um neurônio isoladamente. O peso  $w_0$  é chamado de *bias*, podendo também ser representado por  $b$ , usado para aumentar os graus de liberdade e o poder de generalização de cada neurônio.

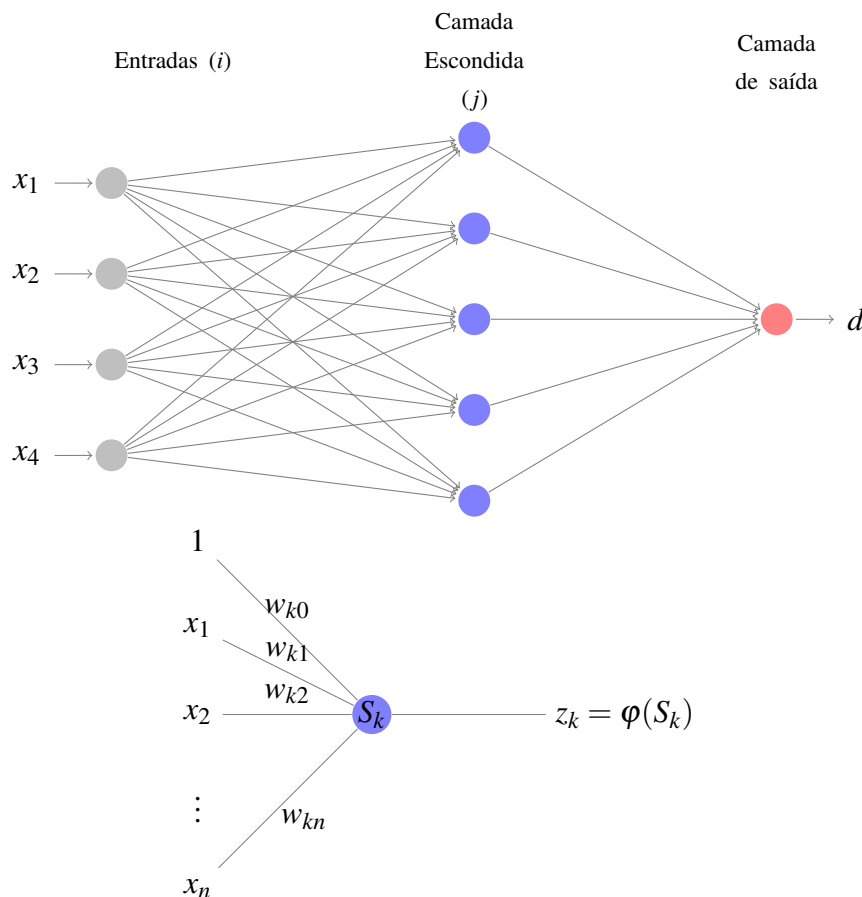


Figura 12 – Acima, um exemplo de uma rede com 4 variáveis de entrada, uma camada escondida com 5 neurônios e um neurônio na camada de saída. Abaixo o detalhamento de um único neurônio da rede.

As informações da rede como apresentada na Figura 12 é propagada de forma unidirecional, da entrada para a saída. Para o caso de apenas uma camada interna e um neurônio, esse tipo

de rede começou a ser denominada na literatura como perceptron (Rosenblatt, 1958), quando aplicada a problemas de classificação. E para problemas de regressão, foi criado o modelo ADALINE, que é uma adaptação do perceptron para saídas contínuas (Widrow and Hoff, 1960).

A Rede Neuronal Perceptron de 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). Por possuir mais de uma camada interna, a MLP possui a capacidade de resolver problemas não lineares.

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 Figura 13.

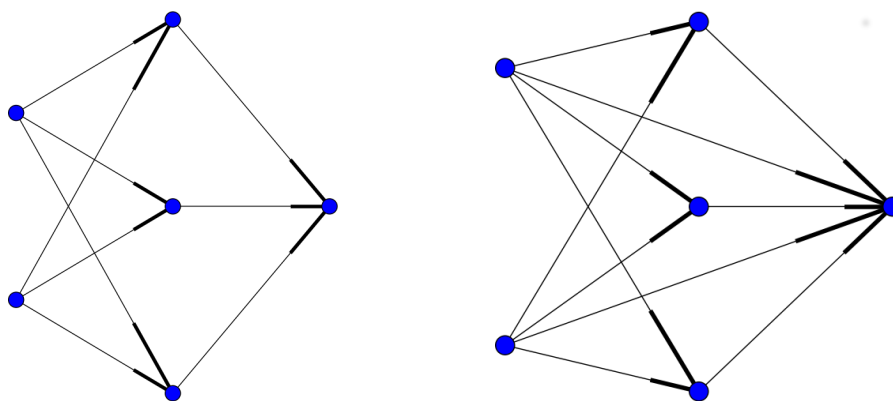


Figura 13 – Modelos de MLP unidirecional simples (esquerda) e unidirecional totalmente conectada (direita)

A função de ativação é responsável pela saída de cada neurônio. As funções podem ser contínuas ou descontínuas. Na Figura 14 são apresentadas graficamente alguns tipos de funções de ativação utilizadas. Neste trabalho, foi usada a função de ativação não linear sigmoideal, dada pela expressão

$$\varphi(S) = \frac{1}{1 + \exp(S)}$$

onde  $S$  é o resultado da combinação linear das entradas com os pesos.

A rede neural é treinada ajustando-se os pesos sinápticos  $w$  a partir de uma função denominada função de perda. Essa função representa o erro que a rede está cometendo para um determinado conjunto de pesos  $w$ . A função de perda pode ser escrita como:

$$J = \frac{1}{N} \sum_{i=1}^N (y_i - d_i)^2 \quad (3.1)$$

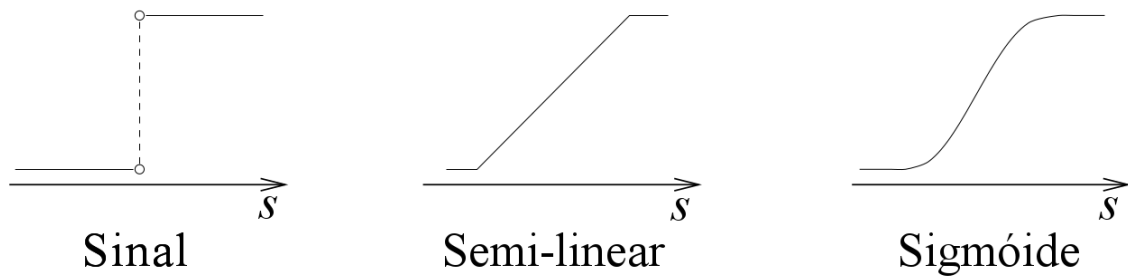


Figura 14 – Exemplos de função de ativação. Na esquerda, a função sinal que é descontínua, é geralmente usada para classificação. No centro uma função semi-linear e na direita a função sigmóide, que são funções contínuas usadas para regressão.

onde  $N$  o número de amostras de treinamento,  $y_i$  é o valor alvo de cada amostra e  $d_i$  é o valor predito pela rede. Essa função representa a média do erro quadrático  $((y_i - d_i)^2)$  na previsão do conjunto de dados de teste. Pode se utilizar diferentes técnicas para otimizar essa função de perda  $J$ . Neste trabalho foi utilizado o algoritmo de Newton Truncado (TNC) (Nocedal and Wright, 2006). O algoritmo de treinamento TNC foi escolhido pois gera melhores resultados em comparação a outros algoritmos tradicionais de treinamento (Radonjic et al., 2012).

### 3.2.2 Máquina de Aprendizado Extremo

A Máquina de Aprendizado Extremo (ELM, Extreme Learning Machine) (bin Huang et al., 2006) é um caso especial das redes neuronais artificiais, na qual possui apenas uma única camada interna. O seu diferencial está no processo de aprendizagem, onde os pesos e o bias da camada interna são atribuídos aleatoriamente, e os pesos de saída são determinados analiticamente.

A função de saída da ELM é dado por

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^L w_i h_i(\mathbf{x}) = \mathbf{h}(\mathbf{x}) \mathbf{w} \quad (3.2)$$

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  $\mathbf{a}_i \in \mathcal{R}^d$  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 5.

Como os parâmetros  $(\mathbf{a}, b)$  são gerados aleatoriamente no ELM, ao invés de ser explicitamente treinados, isso leva a um aumento de eficiência em relação ao tempo de execução, comparado com redes neuronais *backpropagation* tradicionais (Huang et al., 2015).

Tabela 5 – Funções de mapeamento comumente usados no ELM.

Nome	Função
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	$G(\mathbf{a}, b, \mathbf{x}) = 1/(\ \mathbf{a} - \mathbf{x}\  + b^2)^{1/2}$

Os pesos que ligam a camada interna e a camada de saída, denotado por  $\mathbf{w}$ , são encontrados através da minimização do erro quadrático:

$$\min_{\mathbf{w} \in \mathfrak{R}^L} \|\mathbf{H}\mathbf{w} - \mathbf{y}\| \quad (3.3)$$

Para  $N$  pares distintos  $(\mathbf{x}_i, y_i), \mathbf{x}_i \in \mathfrak{R}^n, y_i \in \mathfrak{R}^n, i = 1, \dots, N$ , a solução do problema de encontrar os pesos é dado pelos pesos que minimizam o erro, assim isso pode ser resolvido escrevendo o sistema

$$\begin{bmatrix} G(\mathbf{a}_1, b_1, \mathbf{x}_1) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_1) \\ \vdots & \vdots & \vdots \\ G(\mathbf{a}_1, b_1, \mathbf{x}_N) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_N) \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad (3.4)$$

ou forma compacta  $\mathbf{H}\mathbf{w} = \mathbf{y}$ , onde  $\mathbf{y}$  é o dado de treinamento. A solução ótima então pode ser escrita como

$$\mathbf{w} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} = \mathbf{H}^\dagger \mathbf{y} \quad (3.5)$$

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

O ELM é um algoritmo de aprendizado de máquina rápido, que fornece baixos erros de treinamento, assim como boa capacidade de generalização nos conjuntos de dados de teste de diversos tipos de problemas (Huang et al., 2012).

### 3.2.3 Árvores de decisão

Árvores de decisão (DT, do inglês *Decision Trees*) (Breiman et al., 1984) são ferramentas computacionais usadas na implementação de sistemas especialistas e em problemas de classificação e regressão. O método toma como entrada uma situação descrita por um conjunto de atributos e retorna uma decisão, podendo ser discretos (classificação) ou contínuos (regressão), através de diversos testes em uma estrutura de dados representada por uma árvore (Russell and



Norvig, 2003). Os atributos de entrada podem ser discretos ou contínuos, sendo na maioria das vezes mais eficiente para atributos discretos.

As árvores de decisão executam uma sequência de testes nos atributos de entrada, onde cada um desses testes são representados por um nó interno de uma árvore. Para problemas de regressão, os testes são quantitativos, utilizando os valores de saída definidos pelos atributos de entrada em comparação à um valor de divisão. Normalmente o teste é definido comparando se um valor é menor ou igual a outro. A partir de cada teste, é definido então os ramos da árvore, onde cada nó folha da árvore especifica o valor de retorno para um determinado dado de entrada. O valor de retorno de um determinado nó folha é dado pela média do valor de saída de todos dados de teste que chegam ao nó.

Para definir qual atributo preditivo é utilizado em cada nó da árvore, é preciso utilizar um critério de seleção. Existem diferentes tipos de critérios de seleção, sendo esta escolha uma das variações entre os diversos algoritmos de indução de árvores de decisão. A maioria dos algoritmos de indução de árvores de decisão trabalha com funções de divisão uni-variável, ou seja, cada nó interno da árvore é dividido de acordo com um único atributo de entrada dos dados. Nesse caso, o algoritmo tenta encontrar o melhor atributo para realizar essa divisão. Os critérios de seleção para a melhor divisão são baseados em diferentes medidas, tais como impureza (entropia), distância e dependência. O critério mais utilizado é o de impureza e será o utilizado neste trabalho.

A maior parte dos algoritmos de indução busca dividir os dados de um nó pai de forma a minimizar o grau de impureza ou entropia dos nós-filhos. Quanto menor o grau de impureza, mais desbalanceada é a distribuição de classes. Em um determinado nó, a impureza é nula se todos os exemplos nele pertencerem à mesma classe.

A impureza de um nó é dado por uma função, que depende do problema a ser tratado, sendo ele de classificação ou regressão. Neste trabalho foi utilizado o algoritmo CART (Breiman et al., 1984). O CART constrói árvores binárias usando as variáveis e um limiar que produzem o maior ganho de informações em cada nó. Como neste trabalho envolve um problema de regressão, a impureza é definida através do erro quadrático médio (MSE). O MSE pode ser definido como

$$\text{MSE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3.6)$$

onde  $\hat{y}$  é a saída estimada,  $y$  é a saída correta e  $N$  o número de amostras.

Assim, para construir a árvore de decisão 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:

$$\hat{y}(x) = \sum_{m=1}^K \hat{c}_m I(y|x \in R_m) \quad (3.7)$$

Adotando como critério de minimização o erro quadrático médio, verifica-se que o melhor  $c_m$  é apenas a média de  $y_i$  na região  $R_m$  (Fig. 15):

$$\hat{c}_m = \text{média}(y|x \in R_m) \quad (3.8)$$

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

$$\min_{j,s} \left[ \min_{c_1} \sum_{x_i \in R_1(j,s)} \frac{1}{N} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} \frac{1}{N} (y_i - c_2)^2 \right] \quad (3.9)$$

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

$$\begin{aligned} \hat{c}_1 &= \text{média}(y_i | x_i \in R_1(j,s)) \\ \hat{c}_2 &= \text{média}(y_i | x_i \in R_2(j,s)) \end{aligned} \quad (3.10)$$

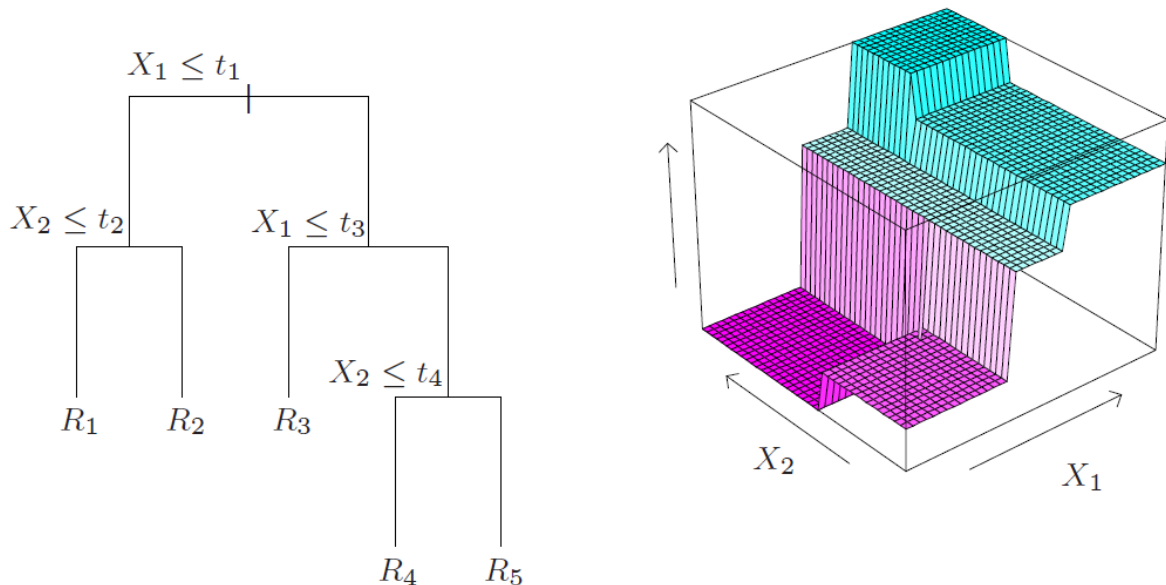


Figura 15 – À esquerda um exemplo de uma árvore gerada pelo DT e à direita uma perspectiva da superfície de regressão gerada por esse modelo, para entradas na forma  $X = (X_1, X_2)$  e os pontos de divisão  $t_1, t_2, t_3, t_4$ , com regiões  $R_1, R_2, R_3, R_4, R_5$  (Hastie et al., 2009).

Apesar de serem mais indicadas para problemas de classificação, as árvores de decisão foram utilizadas neste trabalho para regressão pois através delas é possível encontrar valores que determinam a importância de cada variável de entrada na previsão da saída. Esses valores são chamados de importâncias (ou relevância) das variáveis. Isso é possível pois se pode associar um valor real a cada atributo dos dados de entrada, que representa o quanto ele é importante na

decisões de partições da árvore. Essa capacidade justifica o uso da mesma para problemas de regressão como um método auxiliar para a análise dos dados utilizados.

Para calcular esses valores de importância, para cada nó interno que se divide no atributo  $i$ , é calculado a redução de erro desse nó multiplicada pelo número de amostras que foram encaminhadas para o nó. Assim a soma dessas multiplicações é chamado de importância da variável  $i$ . A redução do erro depende da medida de impureza utilizada, neste caso o MSE. A redução do erro é a impureza do conjunto de amostras que é encaminhado para o nó interno em questão menos a soma das impurezas das duas partições criadas pela divisão.

### 3.2.4 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 (Wang, 2005). 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 (Vapnik, 2013).

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

$$\hat{y}(x) = (w \cdot x) + b \quad (3.11)$$

onde  $\hat{y}$  é o valor aproximado de  $y$  em  $x$ ,  $w$  representa um vetor de pesos,  $b$  é o bias e  $(\cdot)$  é o produto interno.

SVR executa uma regressão linear utilizando uma  $\varepsilon$ -insensível função de perda

$$L(x) = \begin{cases} 0 & \text{se } |y(x) - \hat{y}(x)| \leq \varepsilon \\ |y(x) - \hat{y}(x)| & \text{caso contrário} \end{cases} \quad (3.12)$$

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

$$\begin{aligned} \min & \quad \frac{1}{2} \|w\|_2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{sujeito a} & \quad \begin{cases} y_i - \hat{y}_i \leq \varepsilon + \xi_i^* \\ \hat{y}_i - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, n \end{cases} \end{aligned} \quad (3.13)$$

O parâmetro de regularização,  $C$ , determina o trade-off entre a complexidade do modelo e do grau para o qual o maior desvio de  $\varepsilon$  é tolerado (Smola and Schölkopf, 2004).

Na figura 16 podemos observar um exemplo dos parâmetros do SVR para uma dada entrada  $x$  e os seus respectivos valores observados e preditos  $y$ .

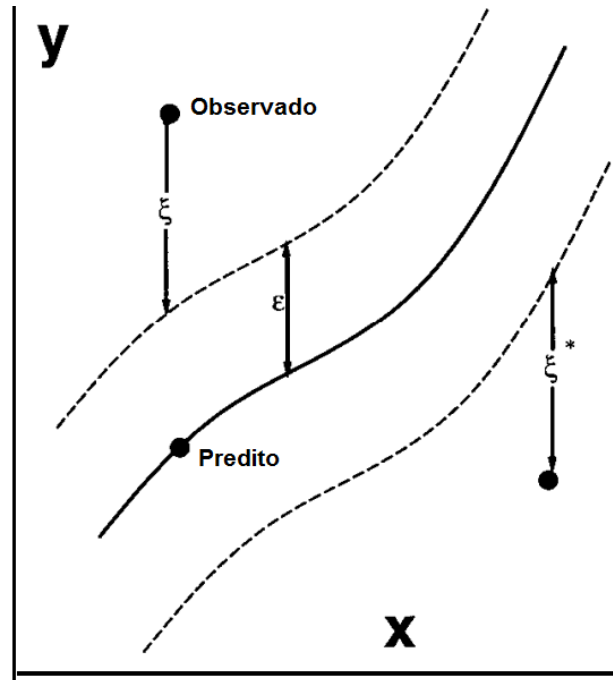


Figura 16 – Exemplo de uma curva de regressão treinada pelo SVR e as respectivas variáveis de folga e o parâmetro de precisão  $\epsilon$ , para uma entrada  $x$  e saída  $y$  (Drucker et al., 1997).

O problema de otimização não é resolvido em sua forma primal, ele pode ser transformado no problema dual

$$\begin{aligned} \max \quad & -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) (x_i \cdot x_j) - \epsilon \sum_{i=1}^n (\alpha_i - \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \\ \text{sujeito a} \quad & \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ 0 \geq \alpha_i \geq C, 0 \geq \alpha_i^* \geq C \end{cases} \end{aligned} \quad (3.14)$$

onde  $w$  pode ser escrito como  $w = \sum_{i,j=1}^n (\alpha_i - \alpha_i^*) x_i$  e a saída estimada do SVR linear é dada por

$$\hat{y}(x) = \sum_{i,j=1}^n (\alpha_i - \alpha_i^*) (x_i \cdot x) + b. \quad (3.15)$$

A regressão não linear pode ser realizada substituindo o produto interno ( $w \cdot x$ ) na Eq. (3.11) por  $(w \cdot \phi(x))$ , onde  $\phi(x)$  é uma transformação não linear. Assim definindo o produto interno  $K(x, x') = \phi(x) \cdot \phi(x')$  como sendo o uma função Kernel, podemos reescrever o problema

como:

$$\begin{aligned} \max \quad & -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) - \varepsilon \sum_{i=1}^n (\alpha_i - \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \\ \text{sujeito a} \quad & \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ 0 \geq \alpha_i \geq C, 0 \geq \alpha_i^* \geq C \end{cases} \end{aligned} \quad (3.16)$$

agora temos que  $w = \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)\phi(x_i)$  e a saída estimada do SVR não linear é dada por

$$\hat{y}(x) = \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)K(x_i, x) + b. \quad (3.17)$$

Além do próprio kernel linear, também serão utilizados neste trabalho outras três funções comumente utilizadas para o SVR, essas função estão definidas na Tabela 6.

Tabela 6 – Funções Kernel utilizadas no SVR.

Nome	Função
Sigmóide	$K(x_i, x) = \tanh(\gamma(x \cdot x'))$
RBF	$K(x_i, x) = e^{-\gamma x-x' ^2}$
Polinomial	$K(x_i, x) = \gamma(x \cdot x')^d$
Linear	$K(x_i, x) = x \cdot x'$

Na Tabela 6, o parâmetro  $\gamma$  é um coeficiente a ser definido para o kernel RBF, sigmóide e polinomial e o  $d$  é o parâmetro que representa o grau do polinômio do kernel polinomial. Esses parâmetros serão utilizados como entrada do método, e são definidos muitas vezes por testes de tentativa e erro ou por processos de otimização.

### 3.3 OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS

Problemas de otimização estão presentes em diversas áreas do conhecimento. Um problema de otimização pode ser definido como a busca de soluções otimizadas, caracterizadas por variáveis de projeto e avaliadas em uma função objetivo. Para resolver esses problemas existem diversos métodos na literatura, podendo eles serem classificados como métodos exatos ou heurísticos. Os métodos exatos buscam encontrar a solução exata do problema. Já os métodos heurísticos buscam encontrar uma solução aproximada do problema. Muitos métodos heurísticos são inspirados em fenômenos encontrados na natureza, como o comportamento e/ou evolução de animais.

Dentre as diversas heurísticas encontradas na literatura, temos o algoritmo de otimização por enxame de partículas (*Particle Swarm Optimization* - PSO, em inglês). É um algoritmo

amplamente utilizado nos dias de hoje e foi desenvolvido por (Eberhart and Kennedy, 1995) para resolução de diversos problemas de otimização práticos e é considerado um dos algoritmos evolucionários mais simples.

O PSO é um algoritmo estocástico fundamentado em um conceito muito simples, e que pode ser implementado poucas linhas de código. O método é inspirado no comportamento de grandes populações de animais, como por exemplo o de pássaros. Cada indivíduo dessa população “sobrevoa” o espaço de busca com uma velocidade que é ajustada dinamicamente de acordo com sua própria experiência voo e a experiência de voo de seus companheiros da população (Shi and Eberhart, 1999).

Diferentemente de outros algoritmos evolucionários tradicionais, como os Algoritmos Genéticos, o PSO não utiliza operadores de seleção na população. Ele também é o único algoritmo evolutivo que não utiliza a ideia da sobrevivência dos indivíduos mais aptos da população, que é dado pela remoção de alguns indivíduos da população com menor aptidão (Eberhart and Shi, 1998).

No algoritmo, o conjunto de todas as partículas é definido como população. Cada partícula representa uma possível solução de um problema de otimização, sendo representado pela sua posição e velocidade no espaço de busca e seu valor de avaliação associado ao problema (*fitness*).

O valor *fitness* é obtido através de uma função objetivo, que é caracterizada pelo problema a ser resolvido, muitas vezes associado a algum conceito de custo. O espaço de busca é definido a partir de limites inferiores e superiores, também definidos a partir do problema a ser resolvido.

O processo de aprendizado e inteligência do algoritmo é dado pela capacidade de memória das partículas. Cada partícula guarda a melhor posição encontrada no espaço de busca até o momento, essa posição é chamada de *pBest*. Existe também um valor que caracteriza a melhor posição da população até o momento, chamado de *gBest*, que também é de conhecimento de cada partícula. Assim, as partículas podem utilizar essas informações, que são locais e globais, para mudar sua posição no espaço de busca, a fim de se obter cada vez melhores soluções e com a capacidade de exploração do espaço.

Resumidamente, a otimização do PSO se dá a partir da mudança de velocidade de cada partícula em direção a sua melhor posição (*pBest*) e a melhor global (*gBest*). As posições  $x$  e velocidades  $v$  de cada partícula  $i$  são atualizadas a cada iteração  $t$  do algoritmo, dado pelas seguintes regras:

$$v_i(t) = v_i(t-1) + \phi_p \cdot r_1(x_{pBest} - x_i(t-1)) + \phi_g \cdot r_2(x_{gBest} - x_i(t-1)) \quad (3.18)$$

$$x_i(t) = x_i(t-1) + v_i(t) \quad (3.19)$$

onde  $\phi_p$  e  $\phi_g$  são constantes que controlam a movimentação na direção da melhor posição da partícula e da melhor posição global, respectivamente e  $r_1$  e  $r_2$  são variáveis randômicas com

distribuição uniforme, com valores de 0 a 1. As variáveis  $r_1$  e  $r_2$ , que são atualizadas a cada iteração do algoritmo, são responsáveis pela diversidade de soluções e uma exploração mais eficiente do espaço de busca. O vetor de velocidade da iteração anterior  $v_i(t-1)$  representa a inércia da partícula.

Para exemplificar o processo de busca, na Figura 17 é apresentado um esquema da atualização da posição e velocidade de uma partícula  $i$  através das equações 3.18 e 3.19, para valores de  $\phi_p$ ,  $\phi_g$ ,  $r_1$  e  $r_2$  iguais a 1.

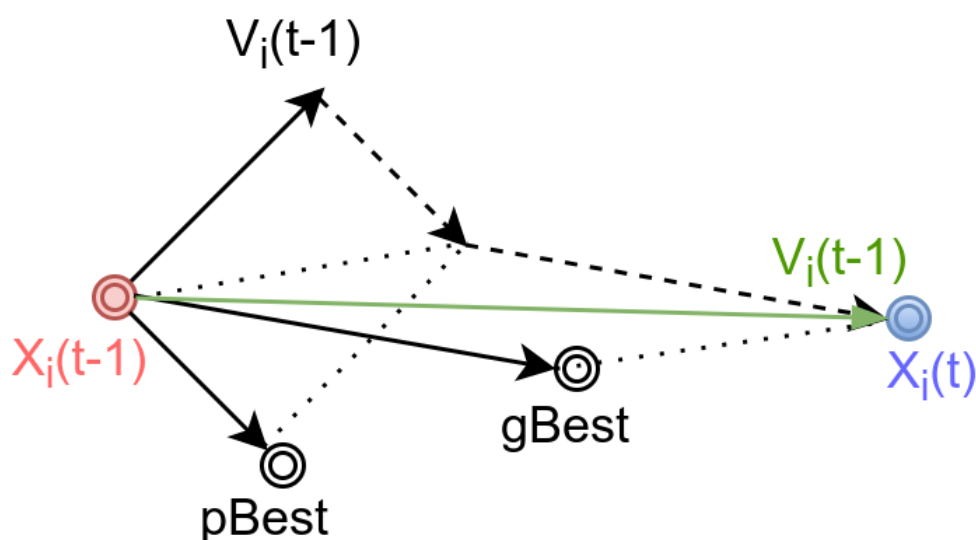


Figura 17 – Exemplo de atualização da posição e velocidade do PSO, para uma partícula  $i$  na iteração  $t$  no espaço de busca.

O critério de parada do algoritmo pode ser definido como um número máximo de iterações ( $iter_{max}$ ) ou até mesmo um valor de tolerância para a variação do valor *fitness* associado à partícula  $gBest$ . Neste trabalho será utilizado um número máximo de iterações como critério de parada. Sendo assim, definindo o número total da população  $NP$ , podemos escrever o pseudocódigo do PSO como visto no Algoritmo 1.

---

**Algoritmo 1:** Pseudocódigo do PSO

---

Inicializa  $x_i$  e  $v_i$  randomicamente;  
**para cada**  $i \in \{0, 1, 2, \dots, NP\}$  **faça**  
    inicializar randomicamente  $x_i$ ;  
    inicializar randomicamente  $v_i$ ;  
**fim para cada**  
**repita**  
    **para cada**  $i \in \{0, 1, 2, \dots, NP\}$  **faça**  
        calcular  $fitness$ ;  
        atualizar  $v_i$ ; (Eq. 3.18)  
        atualizar  $x_i$ ; (Eq. 3.19)  
         $pBest_i \leftarrow$  melhor entre  $pBest_i$  e  $x_i$ ;  
    **fim para cada**  
     $gBest_i \leftarrow$  melhor entre  $gBest_i$  e todos  $pBest_i$ ;  
**até** até satisfazer condição de parada;  
**Resultado:**  $gBest$

---



## 4 PROPOSTA

A metodologia proposta neste trabalho para prever as propriedades mecânicas do concreto de agregado leve a partir dos dados experimentais utilizados é constituída de três conceitos fundamentais:

- busca de parâmetros dos métodos utilizando uma técnica de otimização, no caso o PSO;
- métricas de avaliação estatísticas para medir quantitativamente o poder de generalização dos métodos;
- e a técnica de validação cruzada para separar a base de dados em grupos de treinamento e teste de forma a obter um resultado final mais confiável.

A seguir, será detalhado cada um desses conceitos e sua utilização dentro da metodologia proposta, e também como foi implementado o código computacional de previsão das propriedades.

### 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. Um exemplo esquemático da divisão da base de dados para  $k = 5$  é apresentado na Figura 18.

### 4.2 MÉTRICAS DE AVALIAÇÃO

De modo a avaliar o desempenho de modelos de previsão, faz-se uso de métricas de avaliação estatísticas. Essas métricas têm como objetivo quantificar a capacidade de generalização de um método através do erro. Dependendo do problema e sua aplicação, diferentes métricas

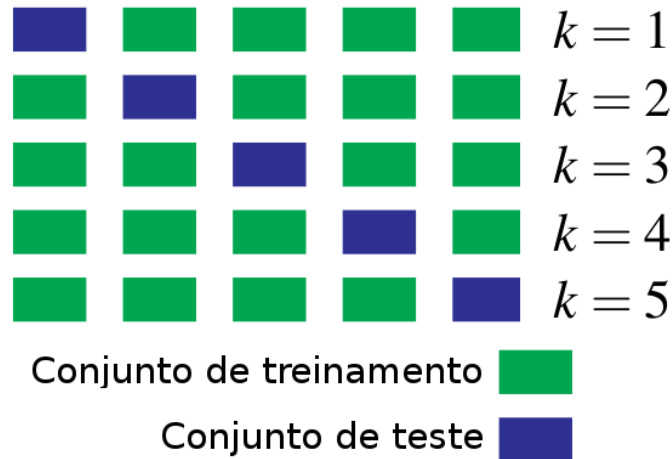


Figura 18 – Esquema da validação cruzada k-fold, para k=5.

podem ser definidas, mas sempre sendo uma função dependente da saída estimada pelo método e o seu respectivo valor experimental. A maioria das métricas usuais retornam um valor real a ser interpretado pelo especialista como bom ou ruim para a devida aplicação.

Neste trabalho foram utilizadas três conhecidas métricas de avaliação a fim de comparar os métodos: o coeficiente de determinação  $R^2$ , o erro percentual absoluto médio (MAPE, do inglês Mean Percentual Error) e a raiz do erro quadrático médio (RMSE, do inglês Root Mean Squared Error). Essas métricas são amplamente utilizadas na literatura e sendo assim de fácil entendimento e uso para comparações com outros trabalhos.

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 percentual absoluto médio (MAPE) é definido como

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

sendo que o valor do MAPE é adimensional e é dado em porcentagem (%).

A raiz do erro quadrático médio (RMSE) é dada por

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

onde o valor de RMSE possui a mesma dimensão da saída do problema, pode ser interpretado como o desvio padrão da amostra de erros entre valores previstos e valores observados.

### 4.3 BUSCA DOS PARÂMETROS

Como cada método de previsão possui diversos parâmetros a serem definidos e é preciso realizar uma comparação entre os métodos, foi utilizado o algoritmo de otimização PSO para encontrar os melhores parâmetros possíveis de cada método de predição. O problema de otimização dos parâmetros dos métodos não é simples e encontrar uma solução exata (ótima) é inviável. Mas com o uso do PSO temos uma boa aproximação da solução desses problema. E assim, a comparação entre os métodos se dará em termos dos melhores modelos de previsão encontrado pelo PSO.

Devido À quantidade de parâmetros que podem ser testados, foi preciso definir intervalos de variação para os parâmetros contínuos e algumas opções para os parâmetros discretos. Os valores utilizado foram definidos pelos valores usuais encontrados na literatura e a partir de testes preliminares.

O PSO realiza a busca por uma solução aproximada do problema de otimização dos parâmetros dos métodos, sejam eles discretos e/ou contínuos, no espaço de busca definido conforme a variação dos parâmetros apresentados na Tabela 7.

Tabela 7 – Variação dos parâmetros dos métodos que são usados na busca do PSO.

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}
MLP	Topologia	{unidirecional simples, unidirecional totalmente conectada}
	Camadas ocultas	uma camada com 4 ou 8 neurônios, duas camadas com 4 ou 8 neurônios cada uma e três camadas com 4 ou 8 neurônios cada uma
SVR	Precisão da regressão ( <i>epsilon</i> - $\epsilon$ )	$[10^{-6}, 10^{-1}]$
	Parâmetro de regularização (C)	$[1, 10^6]$
	Tipo do Kernel	{RBF, polinomial, sigmóide, linear}
	Grau polinomial (Kernel: polinomial)	{1, 2, 3, 4, 5}
	Coefficiente do Kernel ( $\gamma$ )	$[1, 10^{-6}]$

Em sua implementação clássica, que foi utilizada neste trabalho, o PSO é capaz de otimizar apenas variáveis contínuas. Portanto, para que possa ser utilizado variáveis discretas no processo de busca do PSO, sem que seja necessária uma mudança na implementação do algoritmo em si, normalmente é definido uma metodologia de conversão (mapeamento) de variáveis discretas reais para contínuas auxiliares.

Cada parâmetro discreto  $j$  de um método pode ser representado como um vetor  $P_j$  de  $n$  posições, sendo  $n$  o número de opções para o parâmetro  $j$ . Assim, para que o PSO consiga otimizar os parâmetros discretos, foram criadas variáveis auxiliares  $aux_j$  que mapeiam as posições do vetor de cada parâmetro discreto. Essas variáveis auxiliares recebem valores reais do processo de otimização do PSO. Portanto, para fazer o mapeamento com os parâmetro, os seus valores são convertidos para o número inteiro mais próximo e associado a posição do vetor do parâmetro  $P$ . Por exemplo, se uma variável auxiliar  $aux_j = 2.3$ , o seu valor será convertido para 2 e assim será utilizado o valor  $P_j[2]$  do vetor na propriedade  $j$  do método de predição que está sendo otimizado. Esse exemplo de mapeamento é ilustrado na Figura 19.

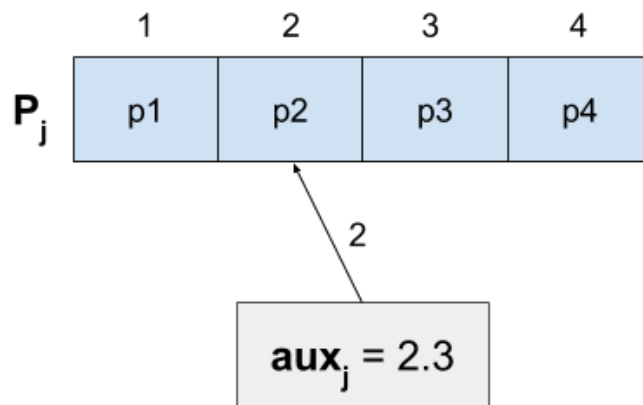


Figura 19 – Exemplo de mapeamento das variáveis discretas, sendo  $P_j$  o vetor que representa um parâmetro  $j$  com 4 opções e  $aux_j$  a variável auxiliar associada.

É preciso definir uma função objetivo  $F$  para ser otimizada no PSO, para construí-la foi preciso utilizar a métrica de avaliação MSE. Para cada conjunto de parâmetros discretos  $PD$  e um conjunto de parâmetros  $PC$  contínuos de um método  $M$ , podemos obter o desempenho do método utilizando subconjuntos de parâmetros contidos a esses conjuntos e utilizando a métrica MSE como medida de erro, definindo assim a função a ser otimizada. A função  $F$  é definida pela expressão

$$F(pc, pd) = MSE(y, M(pc, pd)) \quad (4.4)$$

onde  $y$  são os valores reais da variável a ser predita e  $M(pc, pd)$  são os valores preditos pelo método  $M$  com os parâmetros discretos  $pd \subset PD$  e os parâmetros contínuos  $pc \subset PC$ . Através da Tabela 7 é definido os conjuntos de possíveis valores para  $PC$  e  $PD$ , sendo que os valores dos parâmetros contínuos são formados por intervalos, definidos através de um valor mínimo e máximo.

Portanto, podemos escrever o problema de otimização dos parâmetros a ser resolvido

como descrito na equação 4.5.

$$\begin{aligned} \min_{pc, pd} \quad & F(pc, pd) \\ \text{sujeito a} \quad & pc \subset PC \\ & pd \subset PD \end{aligned} \quad (4.5)$$

Para cada avaliação de um método no PSO, foi utilizado a validação cruzada k-fold, com  $k = 5$ , para a obtenção o valor de MSE final. As principais configurações do PSO utilizado são apresentadas na Tabela 8.

Tabela 8 – Configuração do PSO utilizado.

Parâmetros	Valor
População ( $NP$ )	25
Máximo de iterações ( $iter_{max}$ )	100
$\phi_p$	0.4
$\phi_g$	0.6
Métrica da função objetivo	MSE

#### 4.4 CÓDIGO COMPUTACIONAL

Para utilização da metodologia apresentada, foi desenvolvido um código computacional utilizando a biblioteca de aprendizado de máquina *Scikit-learn* (Pedregosa et al., 2011) para utilização dos métodos de regressão e a biblioteca de análise de dados *Pandas* (McKinney, 2010) para realizar a leitura e preparação da base de dados. O *Scikit-learn* é uma biblioteca amplamente utilizada nos dias de hoje para aprendizado supervisionado e de suporte contínuo e é desenvolvida na linguagem de programação Python. Já o *Pandas* também é implementado em *Python* e é a biblioteca mais utilizada para análise e leitura de dados no *Python*.

Resumidamente, o código realiza a leitura e preparação da base de dados e, em seguida aplica os algoritmos de predição na base de dados utilizando as técnicas de validação cruzada e busca de parâmetros via PSO. Assim, no fim da execução pode ser avaliado o poder de generalização dos métodos através das métricas e os seus respectivos parâmetros ótimos.

Após a leitura da base de dados foi realizada uma normalização linear nos dados de entrada e saída experimentais. A maioria dos métodos de regressão alcançam melhores resultados de aprendizado em base de dados normalizadas. A normalização visa transformar os valores das entradas com a mesma variação de valores, ou seja o valor mínimo e máximo de cada parâmetro de entrada são os mesmo. Os limites mínimo e máximo da normalização utilizada foi igual a 0 e 1, respectivamente. A normalização de uma variável  $X$  é obtida através da expressão:

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4.6)$$

onde  $X_{min}$  é o valor mínimo da variável,  $X_{max}$  é o valor máximo da variável e  $X_n$  é o valor normalizado de  $X$ . Para as saídas  $Y$  (propriedades mecânicas) o processo é análogo:

$$Y_n = \frac{Y - Y_{min}}{Y_{max} - Y_{min}} \quad (4.7)$$

onde  $Y_{min}$  é o valor mínimo da saída,  $Y_{max}$  é o valor máximo da saída e  $X_n$  é o valor normalizado da propriedade  $Y$ .

De forma a garantir uma maior confiabilidade das métricas utilizadas, o processo de validação deve ser repetido diversas vezes, pois os métodos de predição, juntamente com o algoritmo de otimização PSO, possuem aleatoriedade nos seus resultados, necessitando assim de uma amostragem dos seus resultados para que possamos realizar análises estatísticas. Portanto, toda a validação foi realizada em 30 execuções independentes para cada método e para cada propriedade mecânica. Logo, os valores finais de cada métrica para os resultados de cada modelo de previsão encontrado pelo PSO foram dados pela média das 30 repetições.

Como isso, 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 uma boa confiabilidade nas respostas.

Para a Máquina de vetores suporte SVR e a rede neural MLP é preciso definir critérios de parada, pois os dois métodos são iterativos, com número de iterações não-determinístico e os parâmetros utilizados na busca do PSO não garantiam a parada dos métodos. Para o SVR foi definido um número máximo de iterações igual a 50000. Para o MLP definimos também um número máximo de iterações do algoritmo de otimização TNC seguindo a regra: o valor máximo entre 100 e 10 vezes a quantidade de pesos sinápticos.

Um pseudo-código resumido da proposta de validação e previsão das propriedades mecânicas do concreto de agregado leve implementada é apresentado no Algoritmo 2.

---

**Algoritmo 2:** Pseudocódigo da metodologia implementada

---

Leitura da base de dados;

Normalização dos dados;

**repita**

**para cada** *método de previsão* **faça**

**para cada** *propriedade mecânica* **faça**

            Busca de parâmetros: PSO + validação cruzada;

            Previsão da propriedade pelo melhor modelo encontrado;

            Métricas de Avaliação das previsões;

**fim para cada**

**fim para cada**

**até** 30 iterações;

Média e desvio padrão das métricas;

**Resultado:** Previsões e métricas para cada método

---

## 5 RESULTADOS

Neste capítulo serão apresentados os resultados encontrados pela abordagem proposta neste trabalho. De início será apresentada uma discussão dos erros de previsões, utilizando as métricas de avaliação. Em seguida, são apresentados gráficos para auxiliar a análise das previsões dos métodos. Com a ajuda das árvores de decisão foi possível realizar uma análise de importância dos atributos utilizados nos dados experimentais. Por fim, são apresentados os tempos de execução dos experimentos realizados.

Foi utilizada para execução do código computacional uma máquina com sistema operacional 64-bits *Ubuntu 16.04 (linux)*, com 8GB de memória RAM e processador *Intel Core i7 5500U* de 2,4 GHz com 2 núcleos e 4 *threads*. O código foi executado em paralelo nas 4 *threads* disponíveis a fim de se obter o melhor desempenho computacional disponível.

### 5.1 ANÁLISE DAS MÉTRICAS DE AVALIAÇÃO

Após a execução dos métodos de predição, foram obtidos os resultados das previsões na base de dados experimental utilizada. Os resultados das previsões em relação às métricas de avaliação testadas são apresentados na Tabela 9 e na Tabela 10, para o módulo de elasticidade e a resistência à compressão, respectivamente. A primeira coluna apresenta o método de previsão utilizado e a partir da segunda coluna são apresentados os valores máximos, mínimos, médios e os valores de desvio padrão para as métricas MAPE, RMSE e  $R^2$ , respectivamente, nas 30 execuções.

Tabela 9 – Resumo dos resultados das métricas de avaliação nas previsões do módulo de elasticidade ( $E_c$ ) para cada método.

Método		MAPE	$R^2$	RMSE
DT	Média	3.953	0.950	1259.751
	Desvio	0.337	0.009	106.113
	Mínimo	3.577	0.921	1102.322
	Máximo	4.933	0.962	1591.573
ELM	Média	2.892	0.967	984.514
	Desvio	0.674	0.022	286.948
	Mínimo	2.089	0.881	643.172
	Máximo	5.589	0.987	1944.635
MLP	Média	2.582	0.968	889.683
	Desvio	0.674	0.060	487.243
	Mínimo	2.083	0.654	649.027
	Máximo	5.508	0.987	3321.562
SVR	Média	<b>2.231</b>	<b>0.984</b>	<b>701.193</b>
	Desvio	<b>0.267</b>	<b>0.003</b>	<b>73.501</b>
	Mínimo	<b>1.804</b>	<b>0.976</b>	<b>582.147</b>
	Máximo	<b>2.928</b>	<b>0.989</b>	<b>877.322</b>



Tabela 10 – Resumo dos resultados das métricas de avaliação nas previsões da resistência à compressão ( $f_c$ ) para cada método.

Método		MAPE	$R^2$	RMSE
DT	Média	8.698	0.896	5.437
	Desvio	<b>0.559</b>	<b>0.019</b>	<b>0.465</b>
	Mínimo	8.107	0.850	5.032
	Máximo	9.909	0.912	6.553
ELM	Média	6.900	<b>0.940</b>	4.083
	Desvio	0.982	0.025	0.761
	Mínimo	5.496	<b>0.861</b>	3.114
	Máximo	<b>9.285</b>	0.966	<b>6.319</b>
MLP	Média	<b>6.325</b>	0.931	4.182
	Desvio	1.334	0.062	1.529
	Mínimo	<b>4.727</b>	0.719	<b>2.873</b>
	Máximo	11.667	<b>0.971</b>	8.976
SVR	Média	6.592	<b>0.940</b>	<b>3.974</b>
	Desvio	0.983	0.050	1.148
	Mínimo	5.730	0.687	3.201
	Máximo	9.954	0.964	9.466

Os valores em negrito na Tabela 9 e Tabela 10 representam os melhores valores das métricas dos métodos, para cada propriedade mecânica predita. Ressalta-se que, para o MAPE e RMSE o objetivo é encontrar o menor valor possível. Para a métrica  $R^2$  busca-se sempre o maior valor próximo a 1. Assim o melhor valor encontrado para o  $R^2$  é o valor máximo em cada tabela e para o MAPE e RMSE são os valores mínimos da tabela.

Para o módulo de elasticidade os melhores valores de todas as métricas foram obtidos pelo método SVR, destacando-se pelo desvio padrão muito baixo em comparação aos outros métodos. Para a resistência à compressão os melhores valores de média das métricas foram obtidos pelo SVR e a ELM para o  $R^2$ , pelo SVR para a RMSE, e pelo MLP para o MAPE. O método ELM obteve bons desempenhos nos valores máximos das métricas, ou seja, no pior caso a ELM obteve sempre melhores resultados. As árvores de decisão (DT) não obtiveram bons valores para o desvio padrão, sendo até os melhores desvios para a resistência a compressão, mostrando assim uma boa estabilidade das previsões do método.

Com o objetivo de auxiliar a interpretação dos resultados das previsões dos métodos, as Figuras 21 e 20 apresentam os gráficos em *boxplots* para cada um dos métodos e para cada propriedade predita, sendo utilizada para sua geração a métrica MAPE nas 30 execuções. O MAPE é uma métrica importante, pois é levada em conta na maioria dos projetos de engenharia, e possui fácil entendimento e aplicação, dado que os valores são apresentados em porcentagem, pois a métrica calcula a diferença percentual média em relação a um valor esperado. É possível também observar nas figuras os valores de MAPE em cada uma das 30 execuções, através dos pontos em vermelho. Com os pontos, fica evidente a distribuição dos valores de MAPE e os

*outliers* para cada modelo nas execuções.

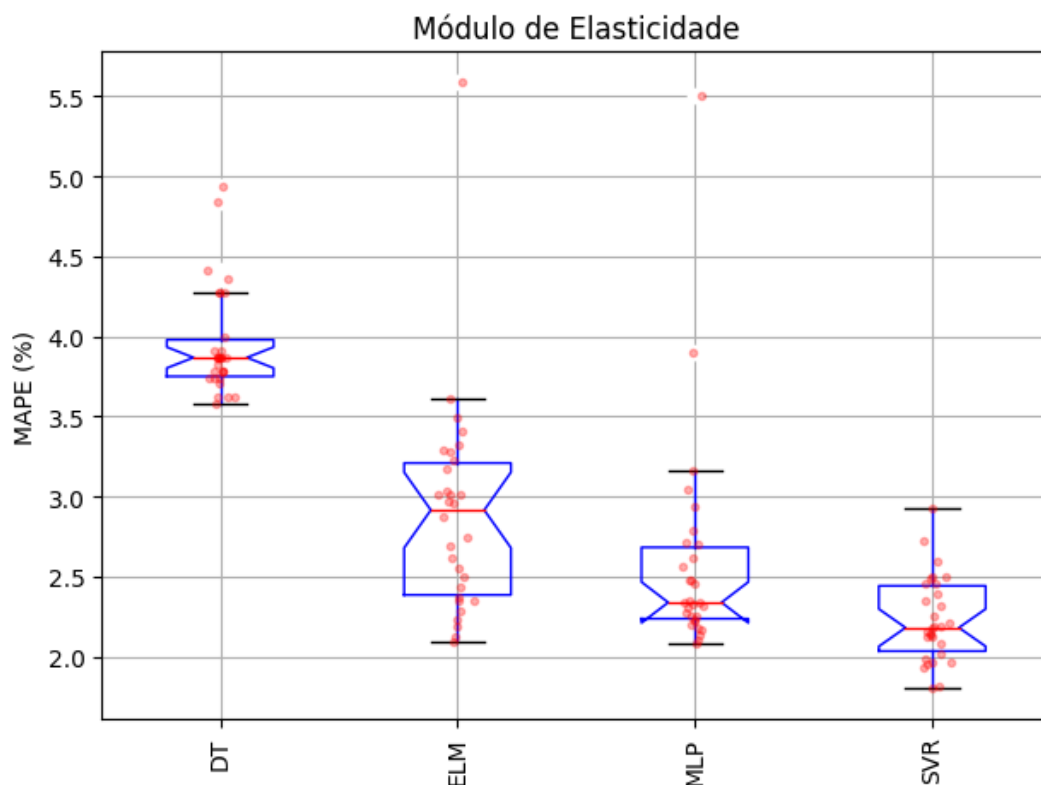


Figura 20 – Boxplots do MAPE dos métodos para o Módulo de elasticidade

Através dos *boxplots* apresentados é possível observar graficamente a variabilidade do desempenho dos métodos. A DT obteve a menor amplitude interquartil, confirmando a baixa variabilidade dos seus resultados já indicada pelo desvio padrão. Também é possível notar a proximidade dos resultados do MLP e o SVR em relação a mediana e aos 50% dos dados que ficaram dentro do intervalo interquartil, mostrando assim um poder de generalização próximo entre os métodos.

Com relação aos *outliers*, pode-se observar o seguinte: (i) para o módulo de elasticidade, o MLP e o ELM produziram resultados com valores máximos de MAPE muito acima do valor máximo mostrado pelo *boxplot*; (ii) para a resistência a compressão, o SVR e o ELM produziram também valores máximos de MAPE muito acima do máximo do *boxplot*. Sendo assim, estes resultados acima do máximo apresentado pelo *boxplot* são classificados como *outliers*. Isso representa para o caso deste trabalho, que no processo de otimização utilizado, o PSO encontrou e convergiu possivelmente para um mínimo local, caracterizado por uma combinação de parâmetros do método, com um desempenho muito pior em comparação aos outros resultados, o que não representa o poder de generalização real do método. Pode-se considerar que esses valores não são representativos para a amostra dos métodos, e é possível desconsiderá-los das análises de desempenho dos mesmos. Sendo assim, se forem desconsiderados os *outliers*, diferentemente do apresentado na Tabela 10, utilizando como base a métrica MAPE, podemos considerar que no

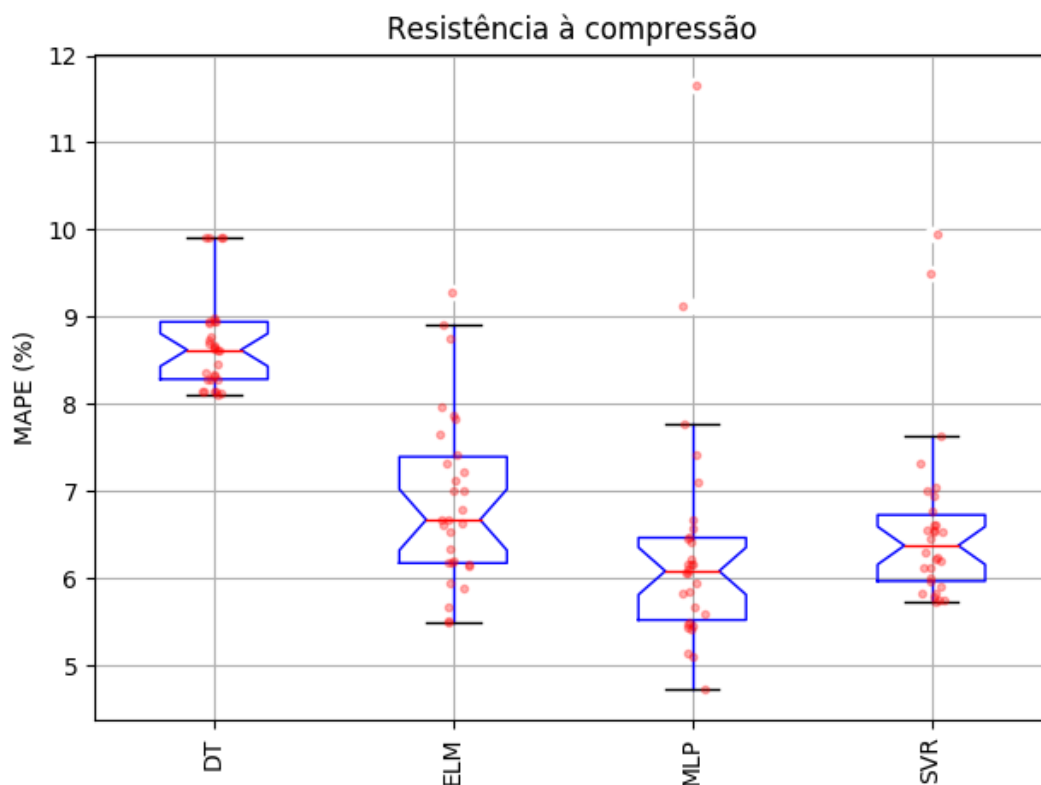


Figura 21 – Boxplots do MAPE dos métodos para a Resistência à compressão

pior caso (valor máximo de MAPE) da previsão da resistência a compressão, o melhor resultado obtido foi do SVR e não da ELM.

## 5.2 ANÁLISE DAS ESTIMATIVAS DOS MODELOS

É possível representar graficamente os resultados das previsões comparando com os valores medidos experimentalmente das propriedades. Esses gráficos são chamados de gráficos ou diagramas de dispersão. Esses gráficos são apresentados na Figura 22 para a resistência à compressão, e na Figura 23 para o módulo de elasticidade para cada método testado.

Para construir os gráficos, foram utilizados os valores preditos dos métodos que obtiveram um desempenho mediano dentre todas as execuções. Para encontrar esse modelo com desempenho mediano, para um dado método e propriedade, foram utilizados os valores do erro percentual absoluto médio (MAPE) das execuções, e assim, foi usado o modelo que apresentou valor mediano de MAPE dentre todas as execuções. Isso foi feito de modo a ser realizada uma melhor comparação entre os métodos, dado que se escolhida a média ou o melhor dos resultados dos modelos, haveria uma maior interferência dos *outliers*.

Nos gráficos de dispersão, se os pontos tendem a formar uma reta de bisseção, como mostrada em vermelho nos gráficos, então podemos concluir que os dados possuem alta correlação. Percebe-se que para os dados de previsão do módulo de elasticidade é notável graficamente

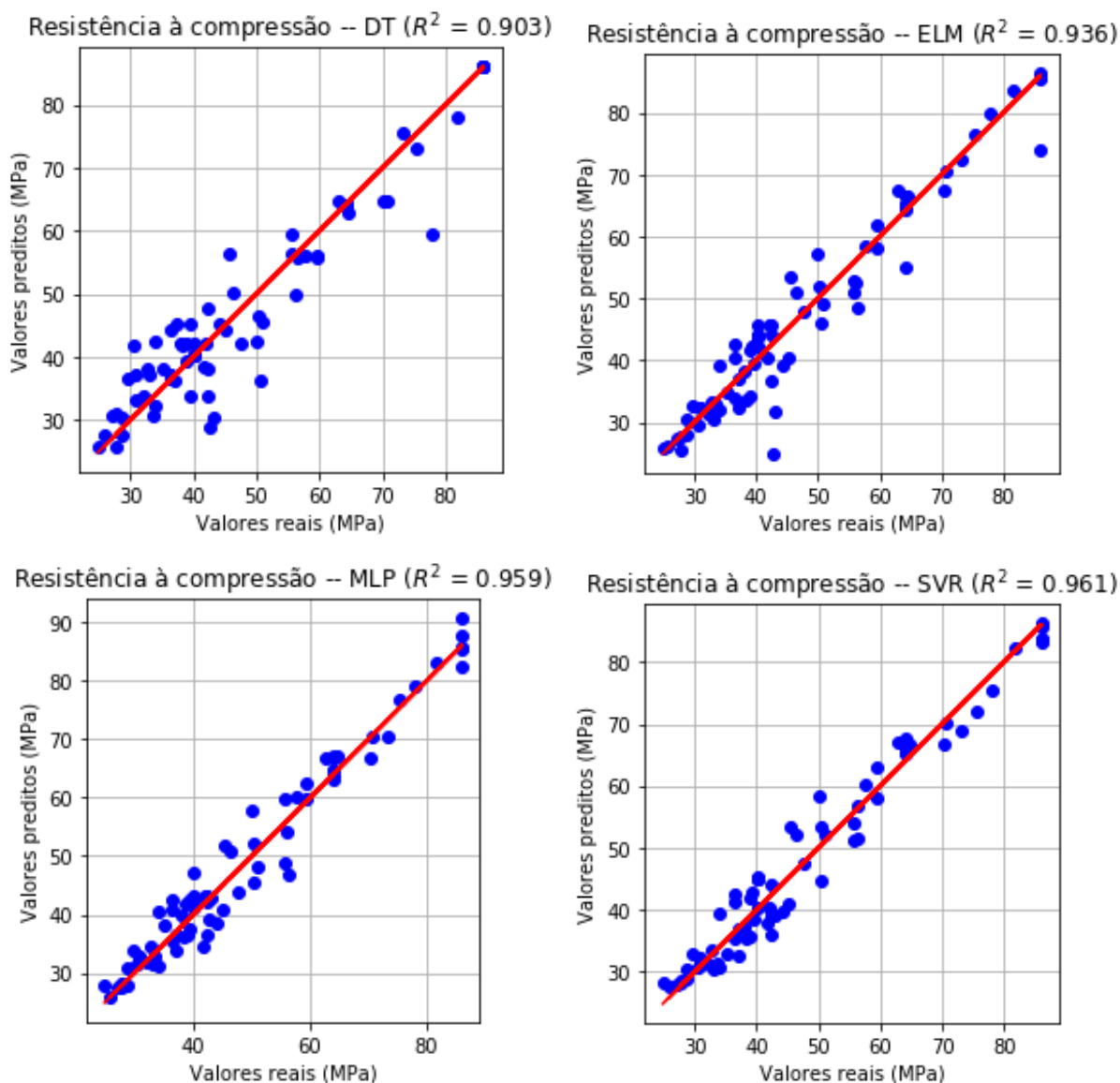


Figura 22 – Gráficos de dispersão para a previsão da resistência à compressão pelos métodos testados que obtiveram desempenho mediano nas 30 execuções.

uma maior correlação entre os dados em comparação a previsão da resistência a compressão, o que é também indicado quantitativamente pelo valores de  $R^2$  mais próximos de 1.

As previsões das propriedades de cada método são mostradas nos gráficos das Figuras 24 a 31. Nos gráficos é possível observar as relações entre os valores reais e estimados para as 75 amostras e os seus respectivos erros. Os resultados são para o modelo que possuiu um desempenho mediano na previsão de cada propriedade nas 30 execuções, utilizando a métrica MAPE. As amostras foram ordenadas a partir do menor para o maior valor da propriedade mecânica, para facilitar a visualização dos mesmos.

Nos gráficos de previsão das Figuras 24 a 31, assim como já foi mostrado nos resultados das métricas de avaliação apresentadas nas Tabelas 9 e 10, fica evidente visualmente a diferença

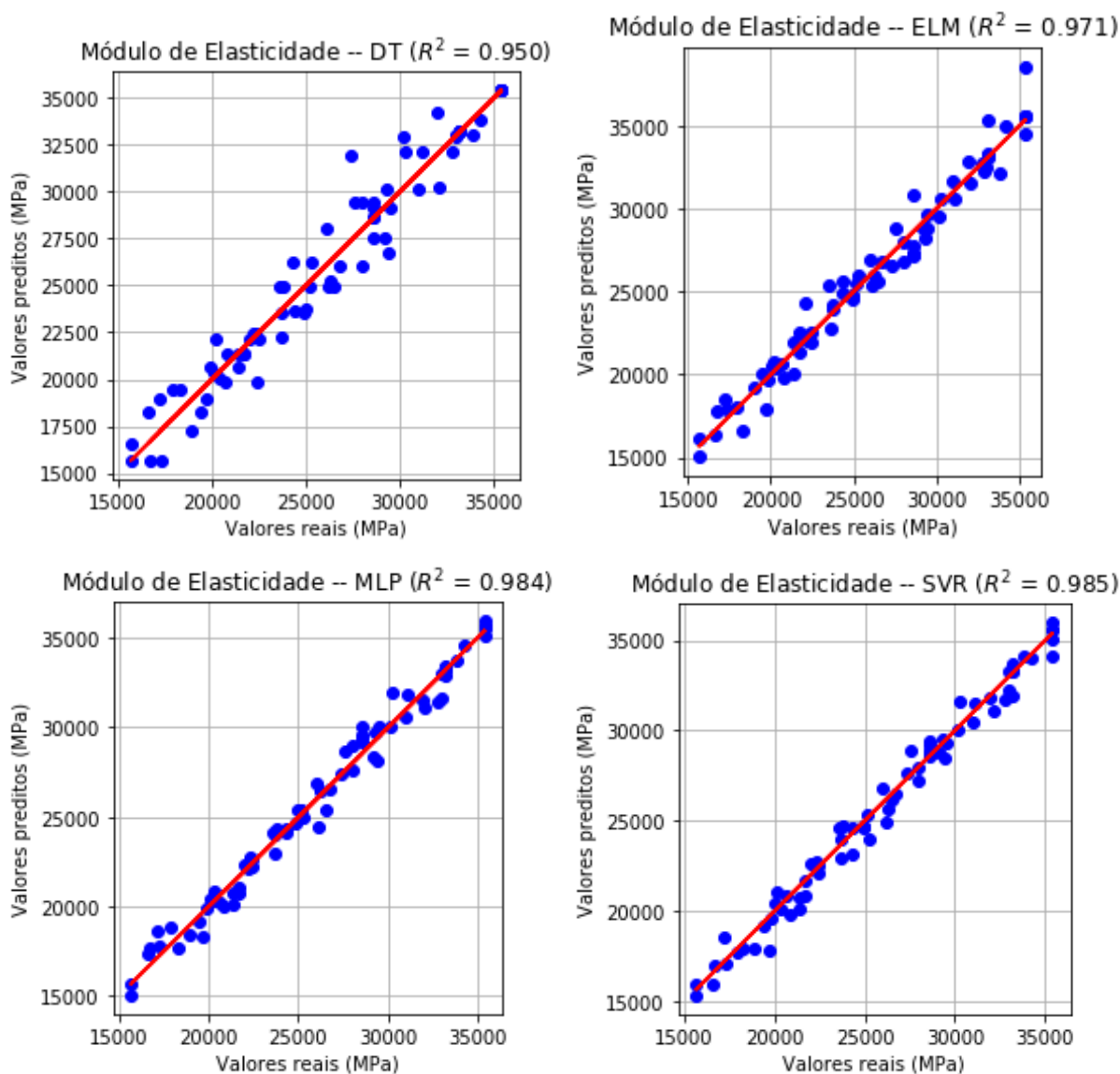


Figura 23 – Gráficos de dispersão para a previsão do módulo de elasticidade pelos métodos testados que obtiveram desempenho mediano nas 30 execuções.

de desempenho dos métodos entre as propriedades mecânicas previstas, a partir dos erros de previsão (barras em verde), visto que para todos os métodos o desempenho foi melhor na previsão do módulo de elasticidade.

É importante ressaltar que os resultados apresentados nesses gráficos de previsão, e também nos gráficos de dispersão, são referentes apenas aos modelos que possuem valores de MAPE medianos. Assim, os valores de  $R^2$  e MAPE indicados nas figuras são diferentes dos apresentados na Tabela 9 e na Tabela 10, pois nas tabelas são exibidos os valores médios das métricas em todas as execuções independentes.

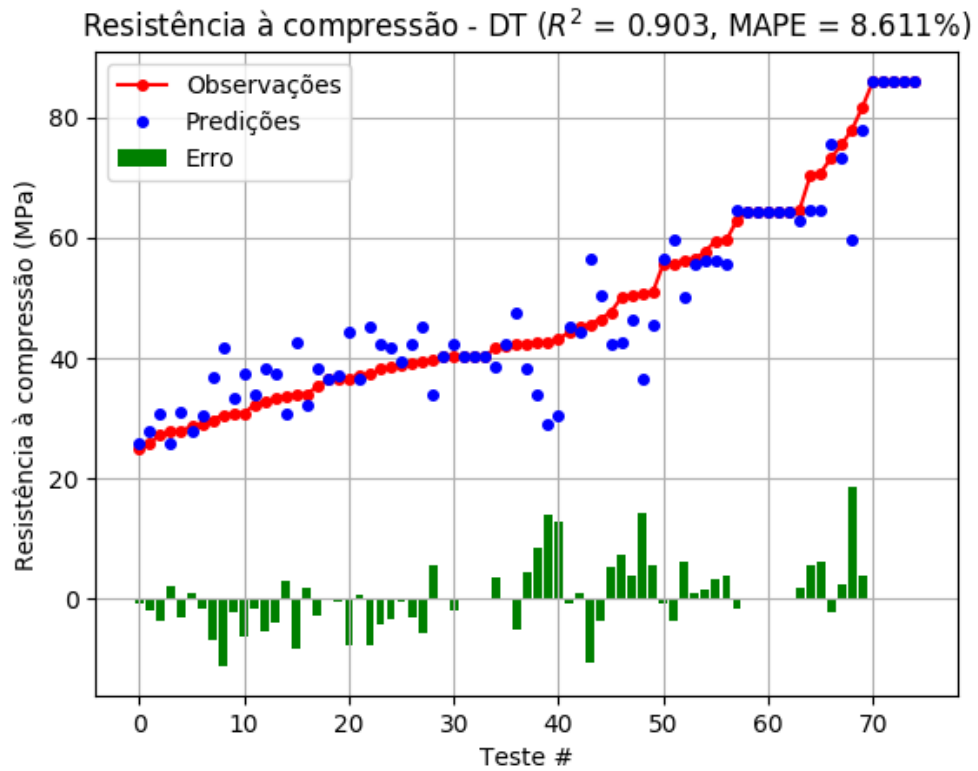


Figura 24 – Previsão da Resistência à compressão pelo DT com desempenho mediano.

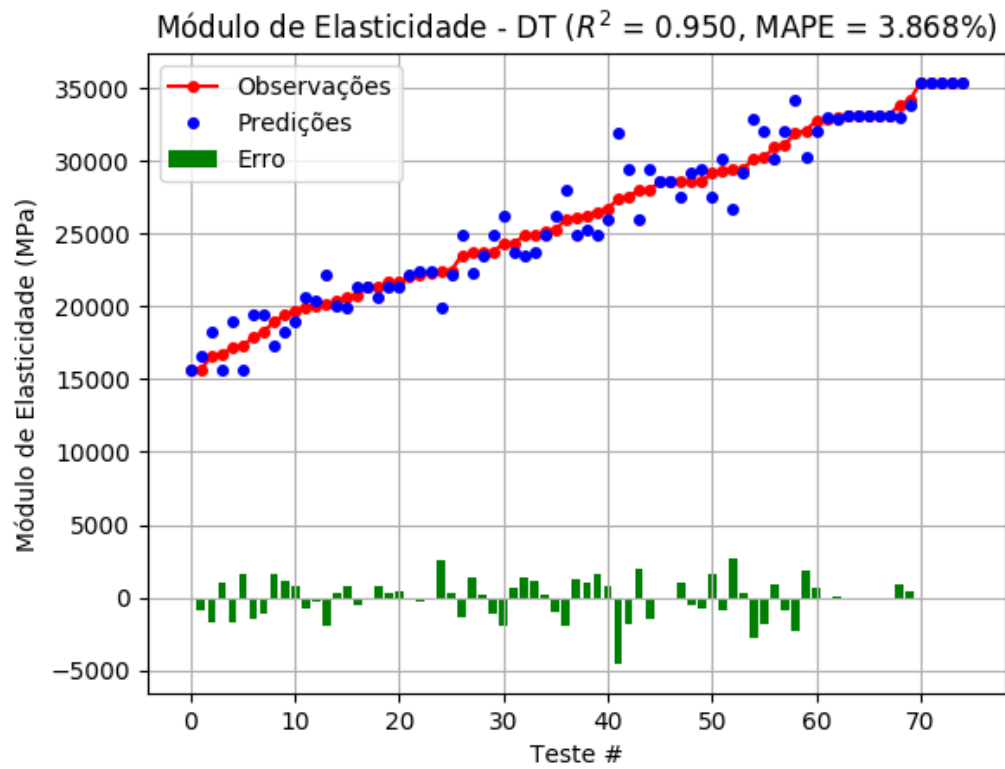


Figura 25 – Previsão do Módulo de elasticidade pelo DT com desempenho mediano.

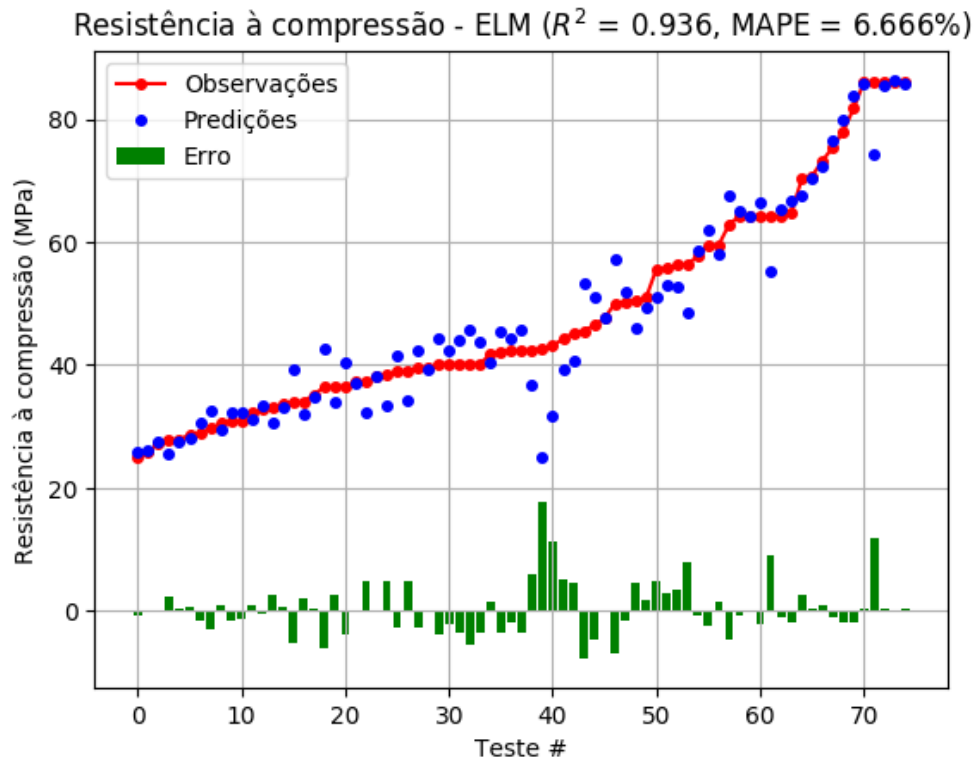


Figura 26 – Previsão da Resistência à compressão pelo ELM com desempenho mediano.

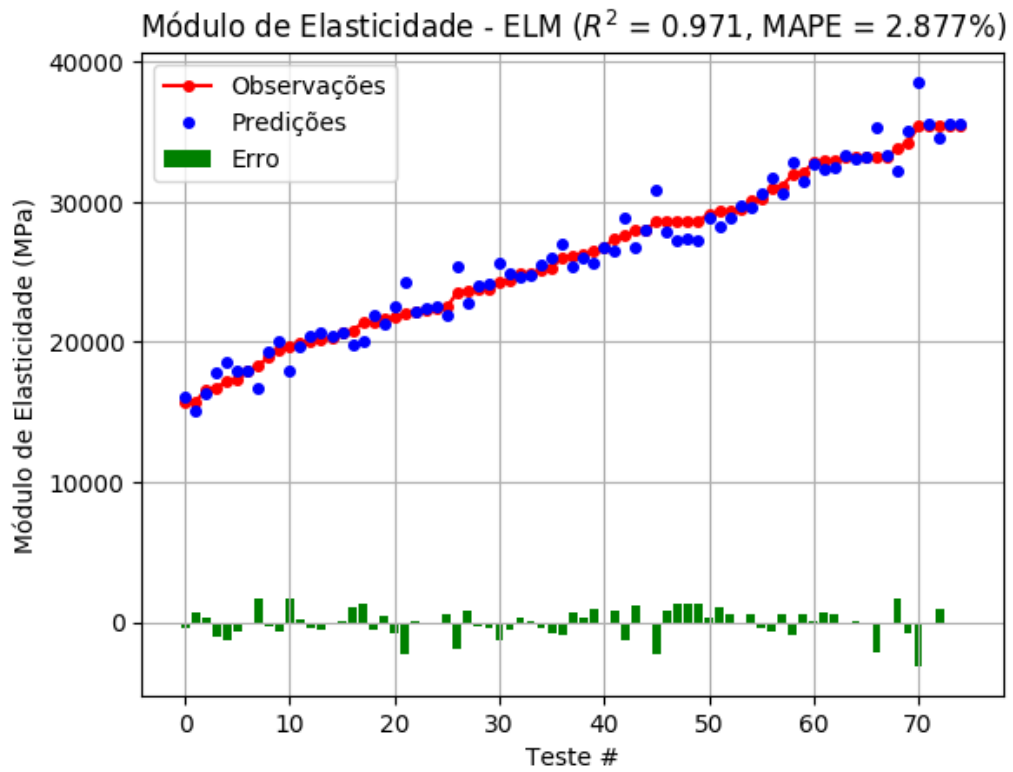


Figura 27 – Previsão do Módulo de elasticidade pelo ELM com desempenho mediano.

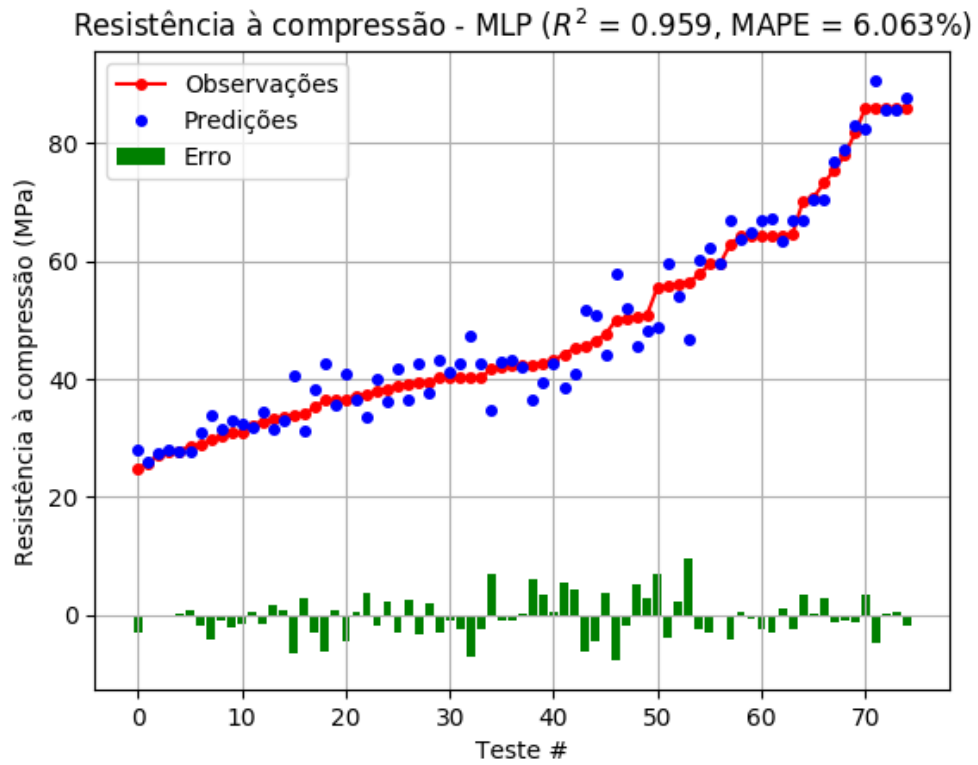


Figura 28 – Previsão da Resistência à compressão pelo MLP com desempenho mediano.

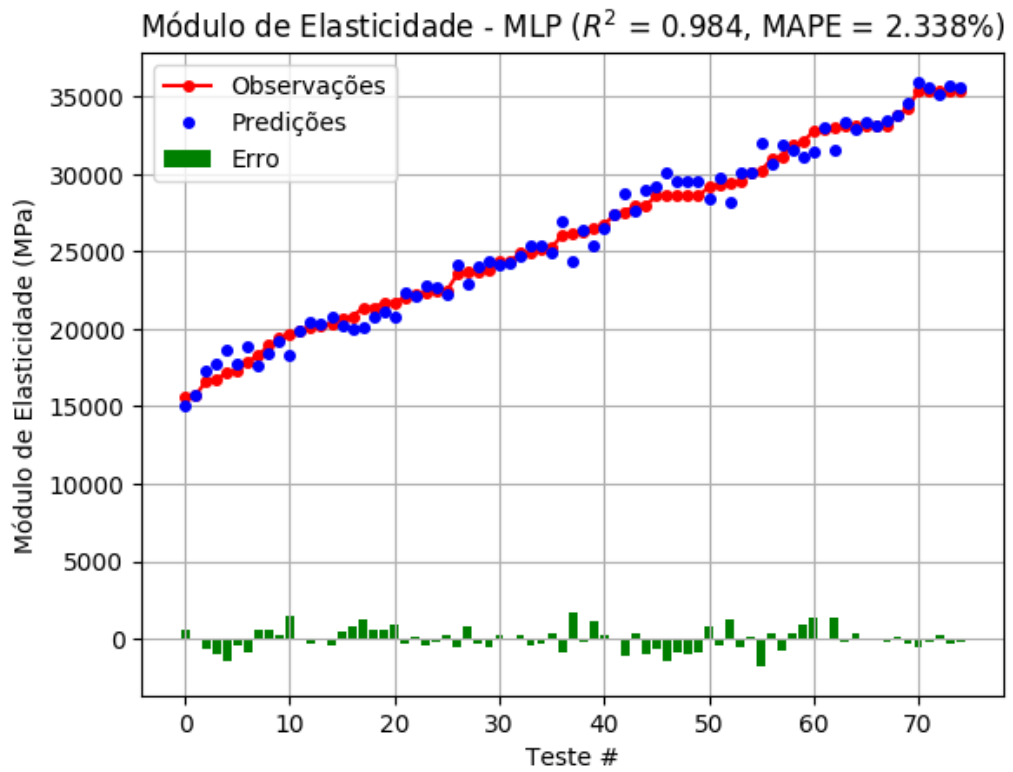


Figura 29 – Previsão do Módulo de elasticidade pelo MLP com desempenho mediano.



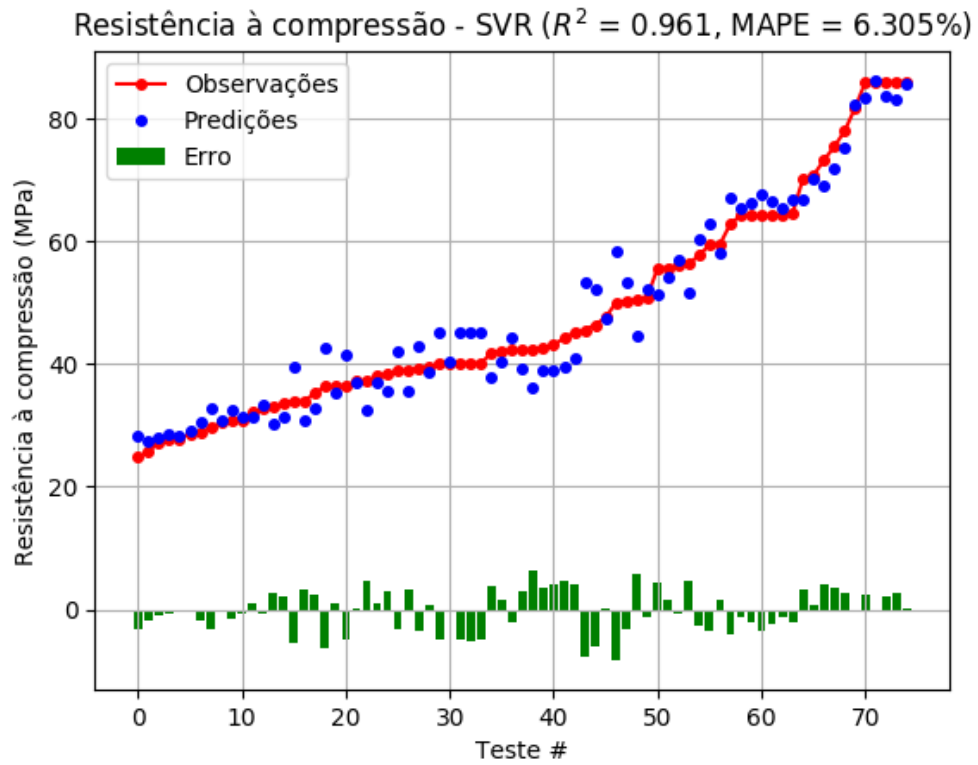


Figura 30 – Previsão da Resistência à compressão pelo SVR com desempenho mediano.

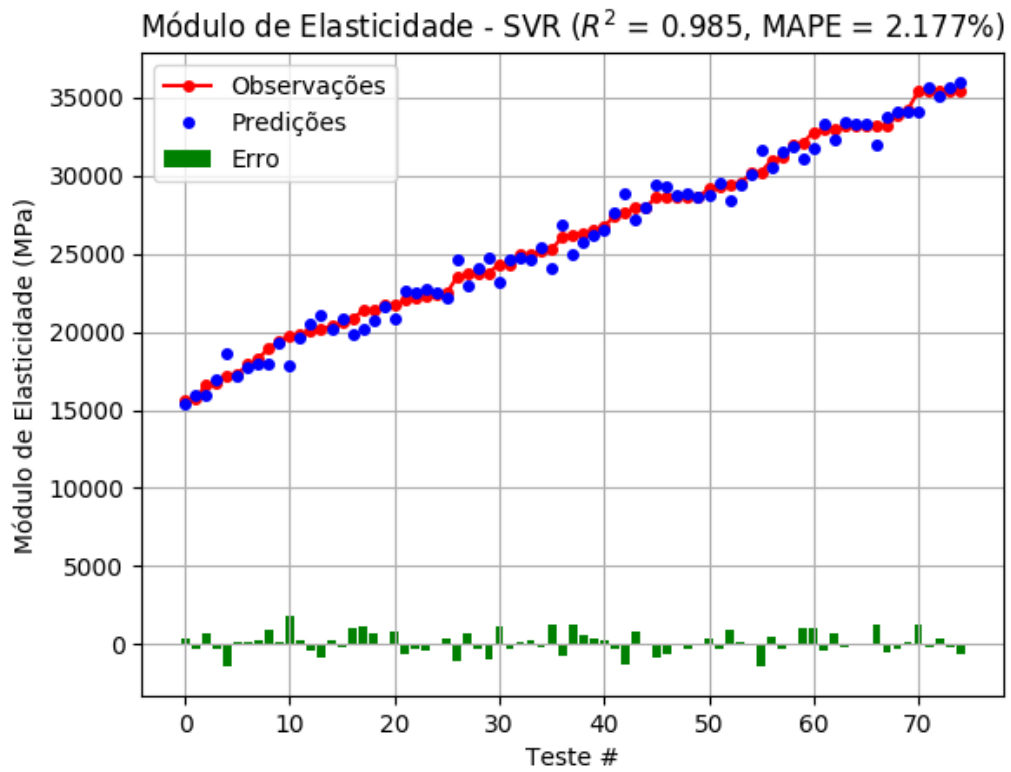


Figura 31 – Previsão do Módulo de elasticidade pelo SVR com desempenho mediano.

### 5.3 OTIMIZAÇÃO DOS PARÂMETROS

Para cada método de predição testado, o algoritmo de otimização PSO encontrou os parâmetros que têm o objetivo de maximizar o poder generalização desses métodos para a previsão das propriedades do concreto em cada uma das trinta execuções independentes. Os parâmetros da melhor configuração de cada método, levando em consideração a métrica MSE utilizada na função de otimização para o PSO, podem ser observados na Tabela 11 e na Tabela 12, para o módulo de elasticidade e a resistência à compressão, respectivamente.

Tabela 11 – Melhores parâmetros encontrados pelo PSO para cada modelo na previsão do módulo de elasticidade.

Modelo	Parâmetros	Valor
DT	Profundidade máxima	10
ELM	Função de ativação	multiquádrica
	Neurônios ocultos	20
MLP	Topologia	unidirecional totalmente conectada
	Camadas ocultas	uma camada com 4 neurônios
SVR	Precisão ( <i>epsilon</i> - $\epsilon$ )	0.001
	Regularização (C)	391.9
	Tipo do Kernel	RBF
	RBF - <i>gamma</i> - $\gamma$	0.0489

Tabela 12 – Melhores parâmetros encontrados pelo PSO para cada modelo na previsão da resistência à compressão.

Modelo	Parâmetros	Valor
DT	Profundidade máxima	9
ELM	Função de ativação	tangente hiperbólica
	Neurônios ocultos	20
MLP	Topologia	unidirecional simples
	Camadas ocultas	uma camada com 4 neurônios
SVR	Precisão ( <i>epsilon</i> - $\epsilon$ )	0.0179
	Regularização (C)	185.1
	Tipo do Kernel	RBF
	RBF - <i>gamma</i>	0.101

É possível notar que em alguns dos parâmetros dos métodos houve diferença entre os valores ótimos dos mesmos na previsão da resistência à compressão e do módulo de elasticidade. Isso mostra que um modelo que estima bem uma propriedade pode não ser o mais adequado para estimar outra, mostrando então que a otimização dos parâmetros é um processo importante na metodologia de previsão de propriedades mecânicas do concreto de agregado leve. Em contrapartida, alguns parâmetros importantes e difíceis de serem escolhidos sem um processo de otimização foram iguais na previsão das duas propriedades, como o kernel do SVR e o número

de neurônios e camadas na MLP. Esse resultados indicam que pode ser utilizada no processo de busca uma menor quantidade de parâmetros, o que pode deixar o processo mais rápido e eficaz.

#### 5.4 IMPORTÂNCIA DOS ATRIBUTOS

Apesar de obterem resultados inferiores em comparação aos métodos testados, as árvores de decisão podem ser utilizadas para realizar uma análise de importância de atributos dos dados experimentais do concreto de agregado leve. Após a execução do algoritmo e realização do treinamento, é possível encontrar as importâncias das variáveis (atributos) de entrada em relação a uma propriedade de saída dos dados, no caso deste trabalho a relevância de cada variável em relação a uma propriedade mecânica. Como a importância dos atributos é obtida ao longo do processo de construção da árvore de decisão, a sua obtenção não representa nenhum custo computacional adicional no processo de treinamento do método.

Nos gráficos de barras da Figura 32 e da Figura 33 são apresentadas as importâncias de cada variável na previsão pela árvore de decisão do módulo de elasticidade e a resistência a compressão, respectivamente. Os valores variam de 0 a 1, sendo que, quanto mais próximo de 1 maior a relevância dos valores da variável em relação à propriedade mecânica. Como cada uma das execuções independentes pode gerar um valor de importância ligeiramente diferente, os valores finais apresentados de importância foram obtidos através da média dos mesmos nas 30 execuções da árvores de decisão para cada propriedade.

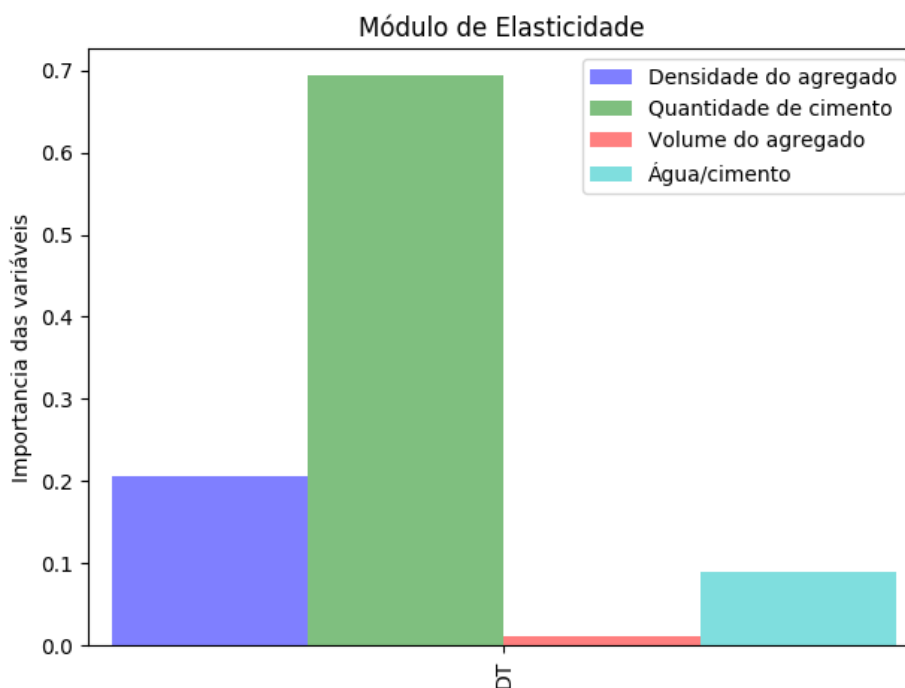


Figura 32 – Gráficos de barras das importâncias (médias) das variáveis de entrada para previsão do módulo de elasticidade pela árvore de decisão.

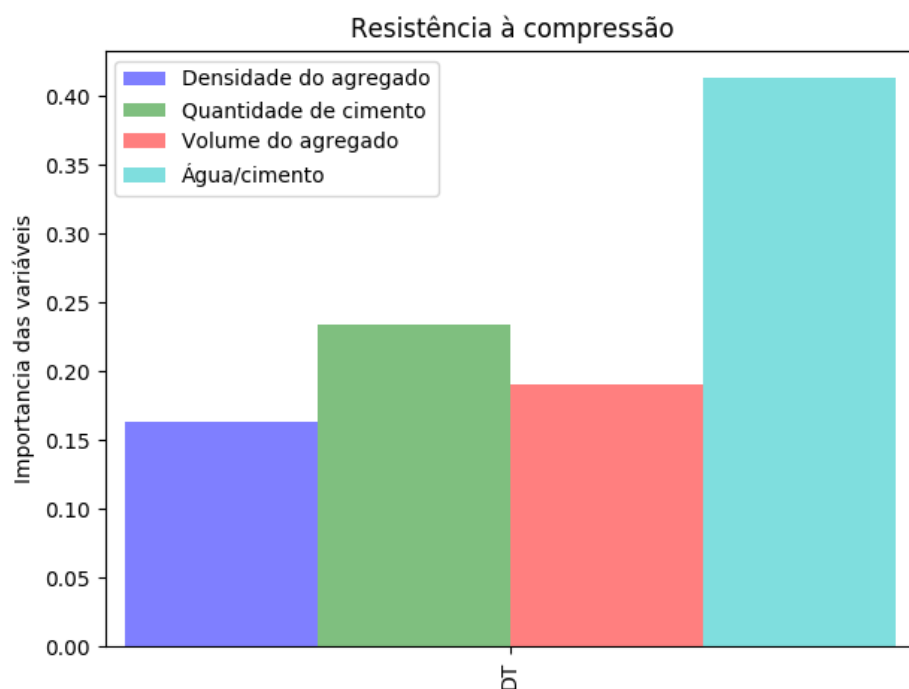


Figura 33 – Gráficos de barras das importâncias (médias) das variáveis de entrada para previsão da resistência à compressão pela árvore de decisão.

Na prática então, essa importância ou relevância pode ser interpretada basicamente como: se há uma alta variação da propriedade mecânica quando uma variável de entrada sofre uma alteração, dizemos que essa variável possui alta relevância nos dados. Para o módulo de elasticidade, foi encontrada uma maior importância para a quantidade de cimento, e a menor importância foi para o volume do agregado, o qual praticamente não influencia a previsão. O fator água/cimento também pouco influencia na previsão do módulo de elasticidade. Para a resistência a compressão, o fator água/cimento foi o mais predominante e a densidade do agregado possuiu a menor importância, mas com valor próximo aos volume do agregado. A quantidade de cimento obteve a segunda maior importância para a resistência a compressão. É interessante notar também que a densidade do agregado possui praticamente a mesma influência na previsão das duas propriedades.

Assim, esses resultados mostram que o módulo de elasticidade sofre uma influência predominante da quantidade de cimento na mistura. Já para resistência a compressão, há uma dependência mais equilibrada de todos às quatro variáveis utilizadas nos experimentos, mas o fator de água cimento é superior aos demais. Os resultados da análise reforçam que o cimento é um dos grandes responsáveis pela resistência de um elemento de concreto (Neville, 2013, 2015), visto que a quantidade de cimento e fator água/cimento são as duas medidas associadas ao cimento na base de dados.

## 5.5 ANÁLISE DO TEMPO DE EXECUÇÃO

A execução de todas as etapas da proposta deste trabalho, composta pela busca de parâmetros, validação e previsão, ocorreu em um período total de aproximadamente 2 horas e 50 minutos. Já os tempos de execução médio, nas trinta execuções independentes, para cada método de todo o processo citado anteriormente são apresentados na Tabela 13.

Tabela 13 – Tempo de execução médio e respectivo desvio padrão em segundos das 30 execuções de cada processo de busca do PSO, com validação e previsão das propriedades, para cada método.

Método	Módulo de Elasticidade	Resistência à compressão
DT	4.1s ( $\pm 1.1$ )	3.9s ( $\pm 1.5$ )
ELM	7.6s ( $\pm 1.9$ )	8.1s ( $\pm 1.7$ )
MLP	110.3s ( $\pm 47.1$ )	134.5s ( $\pm 170.9$ )
SVR	46.7s ( $\pm 76.4$ )	19.5s ( $\pm 41.3$ )

Dada a configuração utilizada no PSO (Tabela 8), foram utilizados 2500 modelos para avaliação na função objetivo que utilizava a métrica MSE. Pois, para cada método, em cada uma das 100 iterações, eram testados 25 (população) modelos de previsão. Assim, temos que em média a execução do processo de aprendizado, validação e previsão de um método individualmente são aproximadamente os apresentados na Tabela 14.

Tabela 14 – Tempo de execução médio do processo de aprendizado, validação e previsão para cada método testado.

Método	Módulo de Elasticidade	Resistência à compressão
DT	0.002s	0.002s
ELM	0.003s	0.003s
MLP	0.044s	0.054s
SVR	0.019s	0.008s

Para todo o processo de busca, a árvore de decisão obteve o menor tempo médio de execução, seguida pela ELM que também obteve um baixo tempo de execução. O MLP e o SVR obtiveram um tempo médio e desvio padrão muito alto em comparação aos outros dois métodos testados. Isso ocorre devido as diferentes configurações utilizadas no processo de busca, como o número de camadas para o MLP e o parâmetro de precisão  $\epsilon$  do SVR, que podem levar a uma modificação drástica no tempo de aprendizado desse métodos. Em contraste, o ELM e a DT possuem comportamento mais estáveis, em relação ao tempo de execução, independentemente

dos valores dos parâmetros testados. Isso mostra que o SVR e o MLP são muito sensíveis aos valores empregados em seus parâmetros em relação ao tempo de execução.

Em comparação aos processos de buscas exaustivas, que são amplamente utilizados na literatura, o PSO se mostrou bastante eficaz, dado que o mesmo possui a capacidade de realizar a busca de forma inteligente, em parâmetros com contínuos e com um número de avaliações baixo. Para a busca exaustiva, seria necessário realizar um refinamento no espaço de busca discreto de forma a obter resultados próximos aos valores encontrados por um método de busca contínua. Então, isso pode acarretar em um número de avaliações muito superior ao de um método evolutivo, como o PSO.

Já em relação aos tempos de execução do aprendizado e validação de cada método, pode-se perceber que os tempos de execução de um método isoladamente são extremamente pequenos, mesmos para o MLP e SVR, visto que em média o tempo é menor que 0.1 segundos. Isso mostra que com uma boa configuração do método previamente definido para um conjunto de dados, a execução do mesmo é muito rápida, como pode ser verificados nos resultados mostrados na Tabela 14. Assim, justifica-se o uso do processo de busca de parâmetros, visto que para um conjunto de dados com características parecidas com ao do utilizado neste trabalho, pode ser utilizado as configurações dos métodos encontradas nas Tabelas 11 e 12, não sendo necessário repetir o processo de busca.

## 6 CONCLUSÃO

O trabalho proposto apresentou uma abordagem de uso de diferentes técnicas de inteligência computacional combinadas com um algoritmo evolutivo para a previsão de propriedades mecânicas do concreto de agregado leve. Quatro técnicas de regressão foram utilizadas para que fosse possível a previsão das propriedades, as Redes Neurais Artificiais de múltiplas camadas (MLP), as Máquinas de Vetores Suporte (SVR), as Máquinas de Aprendizado Extremo (ELM) e as Árvores de Decisão (DT), aliadas a um algoritmo de otimização evolucionário, o algoritmo de otimização por enxame de partículas (PSO). O otimizador teve a função de encontrar o melhor modelo de previsão possível para a realização das previsões. O processo foi validado a partir da técnica de validação cruzada k-fold, utilizando métricas de avaliação de erros. Todo o processo foi repetido em trinta execuções independentes a fim de obter uma análise estatística dos resultados.

Após a execução da abordagem proposta foram apresentados os resultados de média e desvio padrão das métricas de avaliação testadas para cada método e cada propriedade mecânica. Verificou-se que os modelos encontrados para cada método conseguem em média uma boa aproximação dos valores reais, com valores médios do MAPE (erro percentual absoluto médio) no pior caso abaixo de 4% para a previsão do módulo de elasticidade e abaixo de 9% para a previsão da resistência a compressão. Já com relação ao coeficiente de correlação  $R^2$ , em média, os métodos obtiveram valores acima de 0.95 para o módulo de elasticidade e valores acima de 0.89 para a resistência à compressão. Esses valores encontrados pelas métricas condizem com resultados de outros trabalhos que visam obter previsões de propriedades do concreto encontrado na literatura, mostrando que a proposta, apesar de aplicada em um base de dados diferente, conseguiu obter resultados satisfatórios através dos métodos testados.

Dentre todos os métodos testados, como pode ser visto na Tabela 9 e Tabela 10, o SVR obteve os melhores resultados para a previsão do módulo de elasticidade. Já para a resistência à compressão, em média, o SVR obteve os melhores resultados com a métrica  $R^2$  e com o RMSE, para a métrica MAPE a rede MLP obteve o melhor resultado, mas com resultado próximo ao do SVR. O desvio padrão das métricas indicam que o SVR produziu os melhores resultados na previsão do módulo de elasticidade e também baixos valores para a resistência à compressão, para o qual a DT obteve o menor desvio padrão.

Considerando o melhor modelo encontrado pelo processo de busca nas 30 execuções para cada métrica, o MLP obteve o melhor resultado na previsão da resistência à compressão e o SVR obteve o melhor resultado para o módulo de elasticidade, em todas as métricas. Mostrando então que no melhor caso, o comportamento do MLP e do SVR foi parecido com o da média das métricas.

Por fim, pode-se concluir que os objetivos propostos por esse trabalho foram atingidos. O processo de busca através do algoritmo PSO se mostrou eficaz e produziu modelos otimizados

a partir de um processo de busca inteligente, sem que seja preciso realizar processos exaustivos de otimização, que podem acarretar em um grande custo computacional. A partir dos resultados dos métodos testados, conclui-se que o SVR possui em média o melhor poder de generalização na previsão das duas propriedades mecânicas testadas, e também o mesmo possui uma boa confiabilidade nas suas previsões, dado os valores de desvio padrão das métricas nas execuções da proposta. Outra vantagem do uso do SVR também é que o método possui um tempo de execução menor em relação ao da MLP, que obteve na maioria das métricas o segundo melhor resultado, considerando os valores médios. Sendo assim, para uma aplicação real de previsão de propriedades mecânicas do concreto leve, com características parecidas com as avaliadas neste trabalho, os resultados deste trabalho indicam que o SVR é o método mais apropriado para ser utilizado. Além disso, pode ser realizado o treinamento do método sem necessidade de otimização dos seus parâmetros, visto que os parâmetros apresentados nas Tabelas 11 e 12 podem ser reutilizadas.

Para a continuação do trabalho apresentado, é possível desenvolver trabalhos adicionais futuros. Dentre esses trabalhos pode-se citar os seguintes:

- Utilização de métodos combinados (*Ensemble*) visando um melhor desempenho de previsão, visto que as técnicas de *Ensemble* conseguem melhores resultados em comparação aos métodos separados (Breiman, 1996; Mendes-Moreira et al., 2012; Bonifacio, 2017).
- Utilizar outros tipos de composições do concreto de agregado leve que representem uma maior gama de tipos/classes de concreto, a fim de gerar modelos de regressão mais precisos.
- Utilizar técnicas de Regressão simbólica, como a Programação Genética (Koza, 1992) e a Programação Genética Cartesiana (Miller et al., 1997), para encontrar expressão analíticas alternativas às normas vigentes, para o cálculo das propriedades mecânicas do concreto de agregado leve.
- Realizar um análise de sensibilidade para identificar os componentes mais importantes para cada modelo.
- Aplicar um procedimento de seleção de característica no processo de otimização de modo que o algoritmo evolutivo encontre também as características mais relevantes.



## REFERÊNCIAS

- Abd, A. M. and Abd, S. M. (2017). Modelling the strength of lightweight foamed concrete using support vector machine (svm). *Case Studies in Construction Materials*, 6:8 – 15.
- 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.
- Ardakani, A. and Yazdani, M. (2014). The relation between particle density and static elastic moduli of lightweight expanded clay aggregates. *Applied Clay Science*, 93:28 – 34.
- Behnood, A., Behnood, V., Gharehveran, M. M., and Alyamac, K. E. (2017). Prediction of the compressive strength of normal and high-performance concretes using m5p model tree algorithm. *Construction and Building Materials*, 142:199 – 207.
- 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 – characterization and strength prediction. *Materials Design*, 46:832 – 841.
- Bonifacio, A. L. (2017). *Estratégia computacional para avaliação de propriedades mecânicas de concreto de agregado leve*. PhD thesis, Programa de Pós-graduação em Modelagem Computacional - UFJF, Juiz de Fora, Brasil.
- Breiman, L. (1996). Stacked regressions. *Machine Learning*, 24(1):49–64.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- Callister, W. D. and Rethwisch, D. G. (2011). *Materials science and engineering*, volume 5. John Wiley & Sons NY.
- Chandra, S. and Berntsson, L. (2002). 2 - production of lightweight aggregates and its properties. In Chandra, S., , and Berntsson, L., editors, *Lightweight Aggregate Concrete*, pages 21 – 65. William Andrew Publishing, Norwich, NY.

- Chandra, S. and Berntsson, L. (2003). *Lightweight aggregate concrete*. Science, technology and applications. Noyes Publications - William Andrew Publishing, USA.
- Chiew, F. H., Ng, C. K., Chai, K. C., and Tay, K. M. (2017). A fuzzy adaptive resonance theory-based model for mix proportion estimation of high-performance concrete. *Computer-Aided Civil and Infrastructure Engineering*, 32(9):772–786.
- Chou, J.-S., Tsai, C.-F., Pham, A.-D., and Lu, Y.-H. (2014). Machine learning in concrete strength simulations: Multi-nation data analytics. *Construction and Building Materials*, 73:771 – 780.
- 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.
- de Melo, V. V. and Banzhaf, W. (2017). Improving the prediction of material properties of concrete using kaizen programming with simulated annealing. *Neurocomputing*, 246:25 – 44. Brazilian Conference on Intelligent Systems 2015.
- 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. (2013a). Prediction of compressive strength of recycled aggregate concrete using artificial neural networks . *Construction and Building Materials*, 40(0):1200 – 1206.
- Duan, Z., Kou, S., and Poon, C. (2013b). Using artificial neural networks for predicting the elastic modulus of recycled aggregate concrete. *Construction and Building Materials*, 44(0):524 – 532.
- Duan, Z., Kou, S., and Poon, C. (2013c). Using artificial neural networks for predicting the elastic modulus of recycled aggregate concrete . *Construction and Building Materials*, 44(0):524 – 532.
- Eberhart, R. and Kennedy, J. (1995). A new optimizer using particle swarm theory. In *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, pages 39–43. IEEE.
- Eberhart, R. C. and Shi, Y. (1998). Comparison between genetic algorithms and particle swarm optimization. In *international conference on evolutionary programming*, pages 611–616. Springer.

- 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.
- Gao, T., Shen, L., Shen, M., Liu, L., Chen, F., and Gao, L. (2017). Evolution and projection of {CO<sub>2</sub>} emissions for china's cement industry from 1980 to 2020. *Renewable and Sustainable Energy Reviews*, 74:522 – 537.
- González-Taboada, I., González-Fonteboa, B., Martínez-Abella, F., and Pérez-Ordóñez, J. L. (2016). Prediction of the mechanical properties of structural recycled concrete using multivariable regression and genetic programming. *Construction and Building Materials*, 106:480 – 499.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer, New York.
- Hibbeler, R. (2010). *Resistência dos materiais*. Pearson Education do Brasil.
- Huang, C.-L. and Dun, J.-F. (2008). A distributed pso–svm hybrid system with feature selection and parameter optimization. *Applied Soft Computing*, 8(4):1381 – 1391. Soft Computing for Dynamic Data Mining.
- Huang, G., Huang, G.-B., Song, S., and You, K. (2015). Trends in extreme learning machines: A review. *Neural Networks*, 61:32 – 48.
- Huang, G.-B., Zhou, H., Ding, X., and Zhang, R. (2012). Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(2):513–529.
- Igliński, B. and Buczkowski, R. (2017). Development of cement industry in poland – history, current state, ecological aspects. a review. *Journal of Cleaner Production*, 141:702 – 720.
- 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.
- Ke, Y., Ortola, S., Beaucour, A., and Dumontet, H. (2010). Identification of microstructural characteristics in lightweight aggregate concretes by micromechanical modelling including the interfacial transition zone (itz). *Cement and Concrete Research*, 40(11):1590 – 1600.
- 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.
- Koza, J. R. (1992). *Genetic programming: on the programming of computers by means of natural selection*, volume 1. MIT press.
- Kröse, B. J. A. and van der Smagt, P. P. (1993). *An Introduction to Neural Networks*. University of Amsterdam, Amsterdam.
- Lerman, P. (1980). Fitting segmented regression models by grid search. *Applied Statistics*, pages 77–84.
- Lin, S.-W., Ying, K.-C., Chen, S.-C., and Lee, Z.-J. (2008). Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert Systems with Applications*, 35(4):1817 – 1824.
- McKinney, W. (2010). Data structures for statistical computing in python. In van der Walt, S. and Millman, J., editors, *Proceedings of the 9th Python in Science Conference*, pages 51 – 56.
- Mendes-Moreira, J., Soares, C., Jorge, A. M., and Sousa, J. F. D. (2012). Ensemble approaches for regression: A survey. *ACM Computing Surveys (CSUR)*, 45(1):10.
- Miller, J. F., Thomson, P., and Fogarty, T. (1997). Designing electronic circuits using evolutionary algorithms. arithmetic circuits: A case study.
- Neville, A. (2013). *Tecnologia do Concreto - 2ed.* Bookman Editora.
- 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.
- Oliveira, A. L., Braga, P. L., Lima, R. M., and Cornélio, M. L. (2010). Ga-based method for feature selection and parameters optimization for machine learning regression applied to software effort estimation. *Information and Software Technology*, 52(11):1155 – 1166. Special Section on Best Papers PROMISE 2009.
- 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.
- 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.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386.
- Russell, S. J. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Pearson Education, 2 edition.
- Shen, W., Liu, Y., Yan, B., Wang, J., He, P., Zhou, C., Huo, X., Zhang, W., Xu, G., and Ding, Q. (2017). Cement industry of china: Driving force, environment impact and sustainable development. *Renewable and Sustainable Energy Reviews*, 75:618 – 628.
- Shi, Y. and Eberhart, R. C. (1999). Empirical study of particle swarm optimization. In *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, volume 3, pages 1945–1950. IEEE.
- Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222.
- Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 25*, pages 2951–2959. Curran Associates, Inc.
- Vapnik, V. (2013). *The Nature of Statistical Learning Theory*. Information Science and Statistics. Springer New York.
- 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.
- Wang, L. (2005). *Support vector machines: theory and applications*, volume 177. Springer Science & Business Media.
- Widrow, B. and Hoff, M. E. (1960). Adaptive switching circuits. Technical report, STANFORD UNIV CA STANFORD ELECTRONICS LABS.