

UNIVERSIDADE FEDERAL DE ALAGOAS  
CENTRO DE TECNOLOGIA  
GRADUAÇÃO EM ENGENHARIA DE PETRÓLEO

LUCAS VERAS DE SIQUEIRA

APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL PARA  
AVALIAÇÃO DE PARÂMETROS DE SOLO COM FOCO EM PROJETO  
GEOTÉCNICO DE POÇO

MACEIÓ  
2024

LUCAS VERAS DE SIQUEIRA

**APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL  
PARA AVALIAÇÃO DE PARÂMETROS DE SOLO COM FOCO EM  
PROJETO GEOTÉCNICO DE POÇO**

Trabalho de Conclusão de Curso apresentado ao curso de Engenharia de Petróleo, da Universidade Federal de Alagoas, realizado sob orientação do Prof. Dr. João Paulo Lima Santos e coorientação do Prof. Dr. Lucas Pereira de Gouveia.

MACEIÓ

2024

**Catálogo na Fonte**  
**Universidade Federal de Alagoas**  
**Biblioteca Central**  
**Divisão de Tratamento Técnico**

Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 – 1767

S618a Siqueira, Lucas Veras de.

Aplicação de técnicas de inteligência artificial para avaliação de parâmetros de solo com foco em projeto geotécnico de poço / Lucas Veras de Siqueira. – Maceió, 2024.

31 f. : il., grafs. e tabs. color.

Orientador: João Paulo Lima Santos.

Co-orientador: Lucas Pereira de Gouveia.

Monografia (Trabalho de conclusão de curso em Engenharia de Petróleo) – Universidade Federal de Alagoas. Centro de Tecnologia. Maceió, 2024.

Bibliografia: f. 30-31.

1. Inteligência artificial. 2. Krigagem. 2. Máquina de vetores de suporte. 3. Perfuração de poços. 4. Geoestatística. 5. Aprendizado do computador. I. Título.

CDU: 622.25

## Folha de Aprovação

LUCAS VERAS DE SIQUEIRA

Aplicação de técnicas de inteligência artificial para avaliação de parâmetros de solo com foco em projeto geotécnico de poço

Trabalho de Conclusão de Curso submetido à banca examinadora do curso de Engenharia de Petróleo da Universidade Federal de Alagoas e aprovada em 22 de março de 2024.

Documento assinado digitalmente  
 JOAO PAULO LIMA SANTOS  
Data: 11/12/2024 15:51:33-0300  
Verifique em <https://validar.iti.gov.br>

---

(Orientador(a) – Prof. Dr. João Paulo Lima Santos)

### Banca examinadora:

Documento assinado digitalmente  
 EDUARDO TOLEDO DE LIMA JUNIOR  
Data: 11/12/2024 16:55:14-0300  
Verifique em <https://validar.iti.gov.br>

---

(Examinador(a) Externo(a) – Prof. Dr. Eduardo Toledo de Lima Junior)

Documento assinado digitalmente  
 LUCAS PEREIRA DE GOUVEIA  
Data: 11/12/2024 16:16:35-0300  
Verifique em <https://validar.iti.gov.br>

---

(Examinador(a) Interno(a) – Prof. Dr. Lucas Pereira de Gouveia)

## RESUMO

Atualmente, a indústria de petróleo desempenha um papel crucial no fornecimento global de energia; no entanto, sua exploração enfrenta desafios complexos. Isso é particularmente evidente na construção de poços em ambientes offshore e na alta demanda por estudos prévios ao poço em locais extremos. O conhecimento preciso dos parâmetros do solo é imperativo para a construção segura e eficiente de poços de hidrocarbonetos. Porém, em ambientes complexos, realizar estudos de solo no ponto de interesse muitas vezes é inviável. Portanto, há uma necessidade de alternativas capazes de reduzir os custos associados ao longo do processo de construção, mitigar erros humanos e prever as caracterizações sem necessidade do estudo in situ. Dessa forma, este trabalho concentra-se na aplicação de técnicas de inteligência artificial e o uso de Krigagem para avaliar testes de solo e extrapolar dados de pontos geometricamente correlatos. A aplicação dessas técnicas permite obter uma estimativa dos parâmetros de interesse em pontos próximos, além de avaliar a influência da distância nesse processo. Nesse contexto, este estudo visa automatizar a aplicação de modelos de Inteligência Artificial, especificamente utilizando a técnica Support Vector Machine (SVM) em dados públicos de ensaios de cone (CPTu). A implementação será realizada na linguagem de programação Python, e os resultados obtidos serão comparados com técnicas de regressão tradicionais, ampliando a análise no contexto de estudos geotécnicos.

Palavras-chave: Inteligência Artificial, Parâmetros de Solo, Projeto Geotécnico, Perfuração de Poços, Krigagem, Cone Penetration Test (CPT), Support Vector Machine (SVM), Aprendizado de Máquina, Caracterização Geotécnica, Indústria de Óleo e Gás.

## **ABSTRACT**

Currently, the oil industry plays a crucial role in global energy supply; however, its exploration faces complex challenges. This is particularly evident in well construction in offshore environments and the high demand for pre-well studies in extreme locations. Precise knowledge of soil parameters is imperative for the safe and efficient construction of hydrocarbon wells. However, in complex environments, conducting soil studies at the point of interest is often unfeasible. Therefore, there is a need for alternatives capable of reducing associated costs throughout the construction process, mitigating human errors, and predicting characterizations without the need for in situ studies. Thus, this work focuses on the application of artificial intelligence techniques and the use of Kriging to evaluate soil tests and extrapolate data from geometrically correlated points. The application of these techniques allows for an estimation of parameters of interest in nearby points, as well as assessing the influence of distance in this process. In this context, this study aims to automate the application of Artificial Intelligence models, specifically using the Support Vector Machine (SVM) technique on public Cone Penetration Test (CPTu) data. Implementation will be carried out in the Python programming language, and the results obtained will be compared with traditional regression techniques, expanding the analysis in the context of geotechnical studies.

Key-words: Artificial Intelligence, Soil Parameters, Geotechnical Engineering, Oil Well Drilling, Kriging, Cone Penetration Test (CPT), Support Vector Machine (SVM), Machine Learning, Geotechnical Characterization, Oil and Gas Industry.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Desenho esquemático das fases de um poço de petróleo.....	11
Figura 2 - Piezocone utilizado nos testes de solo .....	13
Figura 3 - Semivariograma experimental básico .....	15
Figura 4 - Fluxograma do trabalho .....	19
Figura 5 - Coordenadas UTM dos furos escolhidos .....	20
Figura 6 - Resistência não drenada dos 6 furos conhecidos.....	21
Figura 7 - Previsão da resistência não drenada omitindo um furo conhecido.....	22
Figura 8 - Comparação do valor previsto e real .....	22
Figura 9 - Previsão da resistência não drenada para um furo escolhido.....	23
Figura 10 - Distribuição inicial dos dados CPTu no Python .....	24
Figura 11 - Previsão do SVM omitindo um furo conhecido .....	25
Figura 12 - Comparação valores reais e previstos (SVM).....	26
Figura 13 - Resistência não drenada para um ponto fictício com SVM.....	27

## SUMÁRIO

1	INTRODUÇÃO .....	7
2	OBJETIVOS .....	8
2.1	Objetivos específicos .....	8
3	JUSTIFICATIVA .....	8
4	REFERENCIAL TEÓRICO .....	9
4.1	Perfuração de Poços de Petróleo .....	9
4.2	Ensaio de penetração de cone .....	11
4.2.1	Resistência não drenada do solo ( $S_u$ ) .....	13
4.3	Krigagem .....	13
4.3.1	Krigagem Simples .....	14
4.3.2	Krigagem Ordinária .....	14
4.3.3	Semivariograma .....	15
4.4	Inteligência Artificial .....	16
4.4.1	Aprendizado de máquina .....	16
4.4.2	Support Vector Machine (SVM) .....	17
5	METODOLOGIA .....	18
5.1	Pré-processamento dos dados .....	19
5.2	Fonte dos Dados .....	20
5.3	Visualização do Problema .....	20
6	resultados e discussões .....	21
6.1	Uso da Krigagem .....	21
6.2	Implementação do SVM .....	23
7	Conclusões .....	27
	REFERÊNCIAS .....	30

## 1 INTRODUÇÃO

A indústria de óleo e gás desempenha um papel essencial em toda a evolução econômica da sociedade, principalmente em aspectos de fornecimento energético mundial. No entanto, essa responsabilidade traz consigo desafios e riscos significativos em todos os aspectos da exploração de hidrocarbonetos, exigindo abordagens e aplicações para o enfrentamento desses problemas.

Recentemente, a inteligência artificial (IA) emergiu como uma alternativa promissora em vários campos da engenharia, oferecendo soluções automatizadas e modernas para os desafios da indústria. Em particular, as análises de solo para início de poço tornaram-se uma área de foco crucial, dada a constante complexidade dos testes e de caracterização *in situ*.

Dois dos principais testes utilizados nesse contexto são o *Cone Penetration Test* (CPT) e o *Piezocone Penetration Test* (CPTu). O CPT fornece informações essenciais sobre a resistência do solo, coesão e fricção interna por meio da penetração de um cone no ambiente estudado. Por outro lado, o CPTu, ao incluir medições de pressão de poros, adiciona detalhes sobre a permeabilidade e a compressibilidade do terreno. Ambos desempenham um papel fundamental na caracterização completa do solo, influenciando as decisões no projeto de poços de hidrocarbonetos.

Este trabalho se propõe a explorar a aplicação da inteligência artificial na abordagem de problemas de engenharia geotécnica relacionados à caracterização do solo. Utilizando dados provenientes de OBERHOLLENZER (2020), que abrange mais de 1339 parâmetros de testes geotécnicos, como *CPT* e *CPTu*, realizados em diferentes locais da Áustria e Alemanha, buscamos desenvolver algoritmos capazes de extrapolar esses dados. Isso abrirá caminho para a previsão de parâmetros em locais próximos.

Além disso, nosso estudo concentra-se na comparação entre técnicas de interpolação geotécnica, especificamente o uso de inteligência artificial, como Support Vector Machine (SVM), e a abordagem tradicional de Krigagem. Ao aplicar esses métodos aos dados do teste CPTu, com ênfase em parâmetros como resistência ao cisalhamento não drenada, almeja-se avaliar a eficácia e validade de cada abordagem, especialmente em relação aos parâmetros de solo em poços próximos. Uma análise será realizada para investigar a influência da distância nos

resultados gerados por cada técnica, proporcionando *insights* cruciais para otimizar a caracterização geotécnico do solo e a tomada de decisões em projetos correlatos.

## **2 OBJETIVOS**

O objetivo deste trabalho consiste em desenvolver um modelo automatizado utilizando técnicas de IA para extrapolação de dados CPTu, empregando a técnica SVM. Além disso, busca-se realizar uma comparação de eficácia com a técnica tradicional de Krigagem.

### **2.1 Objetivos específicos**

- Treinar uma rede de IA utilizando o SVM para realizar a extrapolação de dados CPTu;
- Conduzir um estudo de caso empregando dados de base pública (OBERHOLLENZER, 2020);
- Avaliar e comparar a acurácia da técnica SVM em relação à abordagem de Krigagem.

## **3 JUSTIFICATIVA**

A introdução da inteligência artificial (IA) na engenharia de petróleo marca um avanço significativo em inovação e automação nos procedimentos de perfuração de poços. Essa abordagem revoluciona a investigação das propriedades do subsolo, abrindo novas perspectivas e impulsionando o conhecimento nesse campo vital.

Na geotecnia, que enfrenta dificuldades na análise de solos, especialmente em ambientes marinhos, a precisão na localização de estruturas muitas vezes dificulta a realização de análises diretas. Isso se deve à dificuldade de acesso ao ambiente de interesse, tornando impraticável a execução direta das análises nesse local, além da importância de evitar custos adicionais em todo o processo de perfuração. Portanto, a necessidade de extrapolação de dados de pontos próximos se torna crucial.

A aplicação de IA com treinamento de máquina emerge como uma solução promissora para simplificar e otimizar a caracterização do solo. Em comparação com métodos estabelecidos, como a Krigagem, a IA pode oferecer uma abordagem moderna e uma alternativa.

Além disso, a utilização de dados extrapolados pode resultar em economia significativa na construção de poços, eliminando despesas extras associadas a estudos de teste de solo em pontos específicos. Isso possibilita aproveitar dados de poços de correlação já existentes nas áreas de interesse, contribuindo para uma abordagem mais eficiente e econômica.

## **4 REFERENCIAL TEÓRICO**

Este capítulo constitui uma revisão bibliográfica abordando os temas presentes no trabalho, que incluem sistemas de perfuração de poços de hidrocarbonetos, estudos geotécnicos de solo, sistemas computacionais e de inteligência artificial aplicados à análise de áreas de interesse.

### **4.1 Perfuração de Poços de Petróleo**

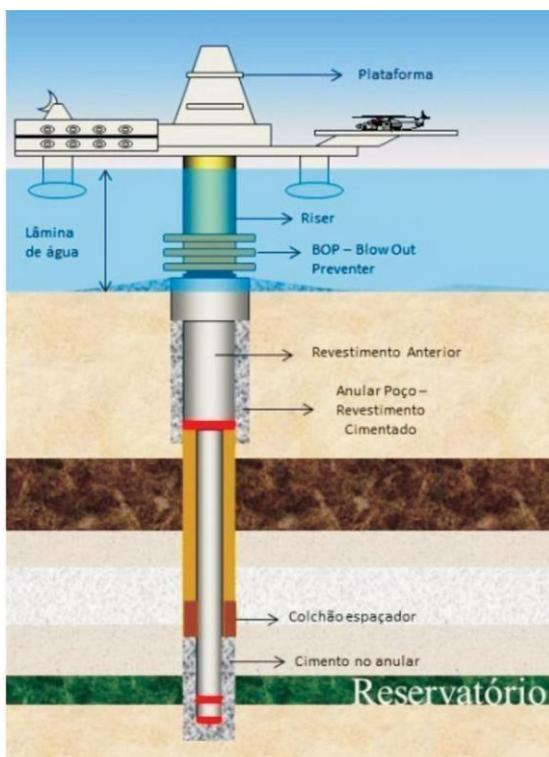
Para introdução referente aos estudos tratados neste trabalho, este tópico visa explicar os processos referentes à perfuração de poços de petróleo, assim como os equipamentos e métodos utilizados. Dentre os principais conceitos relevantes, temos:

- **Perfuração de poços:** O poço é feito através de uma sonda de perfuração, onde as rochas são perfuradas pela ação rotativa de uma broca na extremidade de uma coluna de perfuração. Os fragmentos da formação geológica são removidos através de um fluido de perfuração. Ao atingir determinada profundidade a coluna é retirada do poço e uma coluna de revestimento é descida. O anular entre os tubos de revestimento e as paredes do poço é cimentado com a finalidade de isolar as rochas atravessadas. Após isso, a coluna de perfuração é novamente descida para o prosseguimento da perfuração (THOMAS, 2004).
- **Colunas de perfuração:** Um dos principais componentes do sistema de rotação de uma sonda de perfuração. Ela transfere energia para a formação, em forma de rotação e peso aplicados sobre a broca, causando a trituração das mesmas. É constituída de tubos de aço conectados, onde a parte mais alta e longa é formada por tubos de perfuração (MONTEIRO, 2012).
- **Brocas:** São equipamentos que têm a função de promover a ruptura e desagregação das rochas ou formações (SILVA, 2012).

- Fluidos de perfuração: São misturas complexas de sólidos, líquidos e produtos químicos, cuja função é limpar o fundo do poço dos cascalhos gerados, além de exercer pressão hidrostática sobre as formações, resfriar e lubrificar a broca e coluna de perfuração (SILVA, 2012).
- Revestimento: São basicamente tubos de aço que são instalados em determinadas profundidades nos poços, destinados a isolar os fluidos no interior da coluna dos fluidos da formação, além de manter um espaço permanente no poço, permitir a produção de uma zona específica e fornecer um meio de fixar as válvulas de superfície necessárias e conexões para controlar e suportar os fluidos produzidos. (GATLIN, 1960).
- Cimentação: Quando instalada a coluna de revestimento, o espaço entre a coluna e a parede do poço é cimentado, visando uma melhor fixação da coluna e isolando as zonas permeáveis atravessadas pelo poço (SILVA, 2012).
- Completação: Fase que envolve a instalação de equipamentos como tubulações e válvulas, com principal intuito de isolar diferentes zonas do reservatório, controlar o fluxo de fluidos e suportar operações de produção.

A Figura 1 demonstra um esquemático de um poço de hidrocarbonetos, com foco em mostrar as fases de revestimento.

**Figura 1 - Desenho esquemático das fases de um poço de petróleo**



Fonte: Nelson, 2006.

## 4.2 Ensaios de penetração de cone

O estudo preliminar do solo, antecedendo a perfuração de um poço, tem como objetivo adquirir informações geológicas detalhadas sobre a área específica de início da instalação do poço, responsável por fornecer parâmetros utilizados na instalação de revestimento condutor e de superfície. Nesse contexto, emprega-se o ensaio de penetração de cone, denominado CPT (*Cone Penetration Test*), e sua variação, o CPTu (*Cone Penetration Test with Pore Pressure Measurement*).

Conforme delineado por Lunne, Robertson e Powell (1997), no *Cone Penetration Test* (CPT), um cone na ponta de uma série de hastes é gradualmente introduzido no solo a uma taxa constante, realizando medições contínuas da resistência à penetração desse cone. Durante esse processo, o sistema automaticamente registra dados ao longo da profundidade, obtendo parâmetros fundamentais como a resistência de ponta ( $q_c$ ), indicativa da resistência do solo à deformação, e a resistência de atrito lateral ( $f_s$ ), que revela o nível de coesão do solo. A correlação entre essas medidas desempenha um papel crucial na identificação do

ambiente geotécnico estudado, fornecendo *insights* para o planejamento e execução das operações de perfuração e completção de poços.

O CPTu, similar ao CPT, adiciona um resultado crucial: a medição da pressão dinâmica dos poros ( $u$ ). Esse indicador extra não apenas contribui para uma caracterização mais refinada do local, mas também desempenha um papel na avaliação da resistência não drenada do solo ( $Su$ ), que seria a capacidade desse terreno de suportar carga sem permitir a dissipação de pressão de poros. Esses dados são fundamentais para compreender as necessidades de suporte da coluna de perfuração e outros carregamentos inerentes ao processo de perfuração.

Nos estudos CPTu obtidos, foram utilizados os seguintes parâmetros:

- $q_c$  (Resistência de ponta): É um indicador da resistência da camada de solo que o cone encontra em sua trajetória
- $q_t$  (Resistência total): Representa a resistência global encontrada durante o ensaio, utilizada para avaliar a resistência do solo à penetração.
- $Bq$  (Pressão intersticial): Descreve a pressão da água no interior dos poros de um solo.
- $SIGMAV0$  (Pressão vertical efetiva): Indicadora da pressão vertical do ponto estudado, levando em consideração a pressão do solo e da água subterrânea.
- $U0$  (Velocidade de penetração do cone): Parâmetro importante para analisar a taxa de penetração.
- $U2$  (Velocidade de rotação do cone): Relevante para o registro da velocidade angular.
- $FS$  (Resistência do atrito lateral): Resistência do solo ao redor do cone, em forma de atrito.
- $Su$  (Resistência não drenada): Refere-se à resistência do solo quando a água não tem a oportunidade de drenar do sistema durante o ensaio.
- $NKT$  (Fator do cone): Leva em conta o tamanho e a forma do cone utilizado no ensaio, ajuda na interpretação dos resultados e na conversão de medidas.

- *SIGMAV0EF* (Pressão vertical com contribuição da água subterrânea): Medida importante para entender o efeito da água no comportamento do solo.

A Figura 2 mostra o Piezocone utilizado nos testes CPTu in situ.

**Figura 2 - Piezocone utilizado nos testes de solo**



Fonte: Geositu, 2023.

#### **4.2.1 Resistência não drenada do solo (*Su*)**

A resistência não drenada refere-se à capacidade do solo de resistir a deformações sem permitir a drenagem de água. Em ambientes petrolíferos, onde a estabilidade do solo é vital, essa propriedade se torna fundamental para avaliar os riscos geotécnicos associados às atividades de perfuração.

Segundo Eichelberger et al. (2018), a resistência não drenada é entendida como resistência ao cisalhamento oferecida pelo solo quando é solicitado rapidamente, sem que haja condição de drenagem da água intersticial. Portanto, esse parâmetro é de foco central e será o principal elemento considerado nas demonstrações e aplicações apresentadas ao longo deste trabalho.

#### **4.3 Krigagem**

A Krigagem consiste em um método estatístico utilizado na geoestatística para estimar valores desconhecidos em locais não amostrados, com base em dados observados em pontos próximos. Segundo Carvalho & Assad (2005), o método usa

a dependência espacial entre amostras vizinhas, para estimar valores em qualquer posição dentro do campo, sem tendência e com variância mínima.

Logo, a Krigagem utiliza funções matemáticas, nas quais acrescenta-se pesos maiores nas posições mais próximas aos pontos amostrais, e pesos menores nas posições mais distantes, para criar os novos pontos interpolados com base nessas combinações lineares de dados (JAKOB, 2002).

Tal método é muito utilizado para previsão de um valor específico de uma variável regionalizada em um local, e para o cálculo médio de uma variável regionalizada para um volume maior de dados e informações, caso em que foi utilizado nesse trabalho. A metodologia visa uma atribuição de pesos ótimas às amostras fornecidas pelos problemas, fornecendo um melhor estimado sem vies (TENÓRIO et al., 2019).

Segundo Landim e Sturaro (2002), a Krigagem pode ser dividida em dois principais tipos: Krigagem Simples e Krigagem Ordinária.

#### **4.3.1 Krigagem Simples**

Trata-se da estratégia utilizada quando a média é assumida como estatisticamente constante para toda a área estudada.

Em geral, é aplicável a fenômenos que não apresentam uma tendência devido a suposição de estacionariedade.

A metodologia envolve a construção do semivariograma, que descreve a variabilidade espacial dos dados. A modelagem do semivariograma é crucial, influenciando diretamente a precisão das estimativas. A Krigagem simples, baseada na ponderação dos valores amostrados conforme sua proximidade e correlação espacial, é então aplicada para estimar valores em locais não amostrados.

#### **4.3.2 Krigagem Ordinária**

Nesse tipo, é considerada a média como flutuante ou móvel por toda a área estudada. Logo, quando os valores de uma variável regionalizada apresentam média constante, porém desconhecida, essa metodologia é aplicada.

A construção do semivariograma desempenha um papel crucial, permitindo a modelagem da variabilidade espacial em conjunto com a tendência global. A ponderação dos pontos amostrados, essencial para a Krigagem ordinária, agora incorpora informações sobre a tendência global, resultando em estimativas mais refinadas em locais não amostrados.

### 4.3.3 Semivariograma

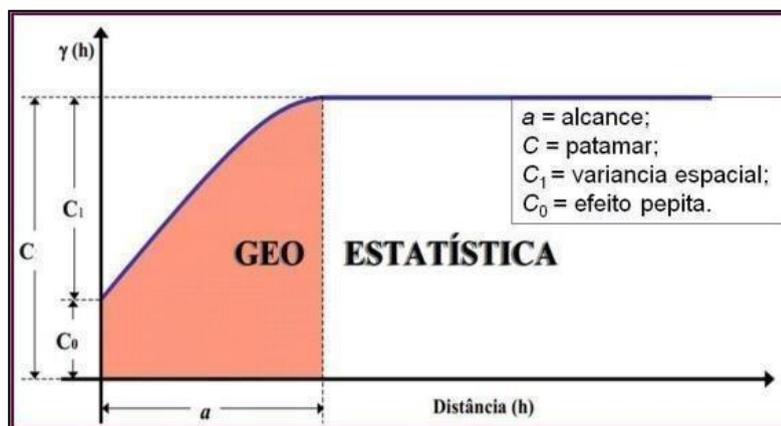
Uma ferramenta essencial para o funcionamento do método é o conceito de construção de um semivariograma, que envolve o cálculo das semivariâncias entre pares de pontos em relação à distância entre eles.

É visto como parte central dos estudos geoestatísticos, capaz de descrever qualitativamente e quantitativamente a variação espacial, além de ser o ponto chave na determinação do preditor geoestatístico Krigagem (Mello, 2005).

Também conhecida como variograma, consiste em um gráfico da semivariâncias em função da distância, relacionando-as.

A Figura 3 ilustra um semivariograma experimental com características próximas do ideal.

**Figura 3 - Semivariograma experimental básico**



Fonte: Giovanini (2022).

Onde, segundo Giovanini (2022), temos:

- Alcance ( $a$ ): A distância onde o variograma atinge o nível máximo de variabilidade.
- Patamar ( $C$ ): Nível de variabilidade onde o variograma se estabiliza, igual à variância amostral.
- Variância espacial ( $C_1$ ): Diferença espacial entre os valores de uma variável tomada em dois pontos separados por distâncias cada vez maiores.
- Efeito Pepita ( $C_0$ ): Segundo Santos e Oliveira (2016), trata-se de fatores como erros de amostragens, medidas ou fatores que causam descontinuidade na origem do variograma.

## **4.4 Inteligência Artificial**

Segundo Russel e Norvig (2013), a Inteligência Artificial (IA) consiste na habilidade dos sistemas cibernéticos em reproduzir funções cognitivas humanas. Essa capacidade se evidencia na resolução de problemas através do aprendizado respaldado pela percepção. Em outras palavras, a IA busca imitar as habilidades mentais humanas, especialmente na abordagem e solução de desafios, utilizando a aprendizagem baseada na interpretação de informações. A indústria petrolífera emprega essas técnicas visando a diminuição do tempo na resolução de problemas e redução global dos custos associados.

A utilização da IA tem como objetivo otimizar desafios inerentes aos diversos estágios de desenvolvimento humano. Nesse contexto, Ludermir (2021) propõe diferentes categorias de funcionamento da inteligência artificial, baseadas na natureza dos problemas abordados:

- IA Focada: Consiste em algoritmos especializados em resolver problemas em uma área e ou problema específico. Nesse tipo, os sistemas armazenam uma grande quantidade de dados e, com isso, os algoritmos são capazes de realizar tarefas complexas, mas sempre focando no objetivo para o qual foram criadas.
- IA Generalizada: Em geral, os algoritmos usam técnicas de aprendizado de máquina como ferramenta, é visto como o nível atual da IA.
- IA Superinteligente: São algoritmos vistos ainda como inexistentes, que serão mais capazes que humanos em praticamente todas as tarefas, é um campo previsional.

### **4.4.1 Aprendizado de máquina**

Como ressaltado anteriormente, dentro da categoria de IA Generalizada, temos o aprendizado de máquina (AM), uma abordagem fundamental que se subdivide em aprendizado supervisionado e não supervisionado. No aprendizado supervisionado, os algoritmos são informados sobre os resultados esperados, enquanto no aprendizado não supervisionado, esses resultados não são explicitamente fornecidos. Além disso, variáveis de diferentes naturezas, como binárias, contínuas ou categóricas, podem ser manipuladas para caracterizar o problema (ROSA, 2020).

Essa abordagem tem uma ampla gama de aplicações que abrangem diversos setores, desde diagnósticos médicos até recomendações de produtos online. Sua capacidade de automatizar a tomada de decisões com base em dados tem implicações significativas, principalmente em previsão e classificação.

Para garantir seu funcionamento eficaz, é vital atender a requisitos específicos. Isso inclui a necessidade de conjuntos de dados de alta qualidade, livres de ruídos, e a implementação de processos adequados de pré-processamento para lidar com desafios como ausência de dados, inconsistências e redundâncias. A interpretação precisa dos resultados também é crucial, exigindo uma compreensão profunda do contexto. Além disso, é essencial reconhecer os riscos, especialmente quando os dados apresentam imperfeições significativas, o que pode afetar a confiabilidade dos resultados.

Segundo Silva (2021), a etapa de pré-processamento de dados é crucial na utilização de algoritmos de *machine learning*, e geralmente incorpora três principais fases:

- Limpeza dos dados: É necessário realizar uma espécie de limpeza em toda a reunião de dados para se certificar que não há dados em branco ou faltantes.
- Transformação dos dados: Os dados passam por processos de transformação, sendo convertidos para formatos adequados para a utilização em modelos, como a normalização.
- Redução dos dados: Em algumas situações, a enorme quantidade de dados pode tornar o processo lento e menos eficaz. Portanto, frequentemente são empregadas técnicas de redução do tamanho total, buscando maior eficiência, sem comprometer a qualidade dos dados.

#### **4.4.2 Support Vector Machine (SVM)**

O Support Vector Machine (SVM) se trata de um algoritmo de aprendizado de máquina, do tipo IA Generalizada, que pode ser utilizado para desafios de classificação ou regressão, onde, seu maior foco é na classificação de um *dataset*.

O algoritmo opera plotando cada dado como um ponto em um espaço n-dimensional, onde o valor do dado é tratado como uma coordenada. O objetivo é encontrar um hiperplano, que dá sentido à disposição dos dados e permite realizar previsões precisas.

Ao invés de simplesmente separar dados em categorias, como na classificação, o SVM em problemas de regressão busca identificar um hiperplano que melhor se adapte aos dados, minimizando a margem de erro entre as previsões e os valores reais. Essa versatilidade do SVM em lidar com problemas de regressão é especialmente valiosa em cenários onde a relação entre as variáveis é complexa, permitindo uma modelagem precisa e flexível para a previsão de valores contínuos. A escolha para esse algoritmo foi baseada em outros trabalhos com a mesma linha de pesquisa, regressão de dados.

## 5 METODOLOGIA

A metodologia deste trabalho consiste em utilizar os dados CPTu disponibilizados pelo *Institute of Soil Mechanics, Foundation Engineering and Computational Geotechnics*, provenientes de localidades específicas na Áustria. Essa abordagem é combinada com a linguagem de programação Python e técnicas de aprendizado de máquina. O propósito é realizar previsões utilizando um ponto já conhecido, validando esses resultados. Adicionalmente, busca-se fornecer previsões para novos pontos próximos aos dados de entrada, empregando o algoritmo *Support Vector Machine* para atender a ambos os objetivos.

Adicionalmente, os dados são aplicados no *software SESTSOLOS* que utiliza a metodologia de Krigagem Ordinária tanto para o ponto a ser validado quanto para o ponto a ser previsto. Isso permite realizar uma comparação entre os resultados obtidos pelos dois modelos.

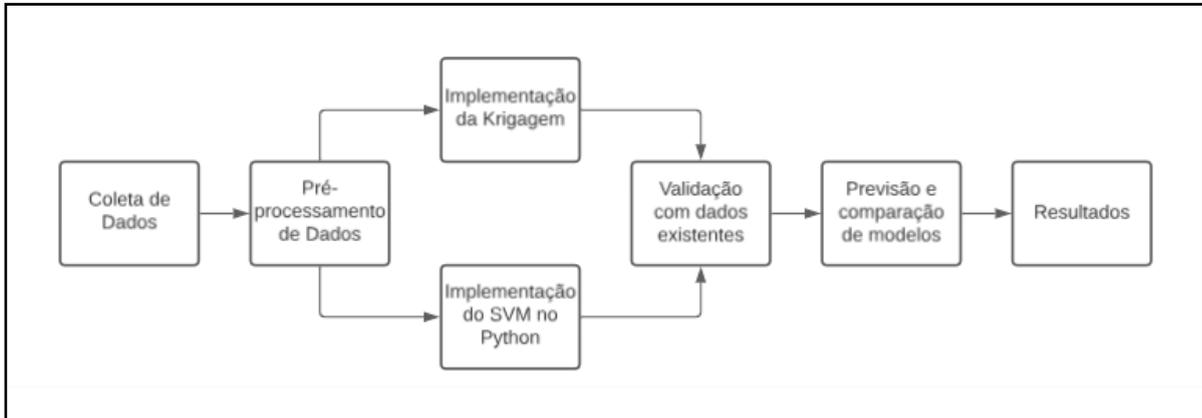
Para as comparações feitas na validação de dados e previsões, é utilizado a fórmula de erro quadrático médio, dada como:

$$EQM = \frac{1}{n} \sum_{i=1}^n (Y_i - Y'_i)^2$$

Onde  $n$  é o número total de observações,  $Y_i$  é o valor observado da  $i$ -ésima observação e  $Y'_i$  é o valor previsto da  $i$ -ésima observação.

A metodologia segue o fluxograma da Figura 4.

**Figura 4 - Fluxograma do trabalho**



Fonte: Autor (2024).

## 5.1 Pré-processamento dos dados

Na fase de pré-processamento dos dados, procederemos à remoção de valores ausentes e o escalonamento dos demais para otimizar a aplicação do SVM. Esse procedimento visa colocar todos os dados em escalas próximas, possibilitando que o modelo os interprete como igualmente relevantes.

Além disso, faremos o uso de algumas correlações para determinar os valores não calculados de algumas variáveis essenciais para aplicação na Krigagem, seriam elas o fator de cone ( $Nkt$ ) e a resistência não drenada ( $Su$ ). Segundo Mayne (2018), podemos utilizar as seguintes correlações:

$$Nkt = 10.5 - 4.6 * \ln(Bq + 0.1)$$

$$Su = \frac{qt - SIGMAvo}{Nkt}$$

A escolha dos algoritmos baseou-se em sua eficácia em problemas de regressão no contexto de aprendizado de máquina. Essas técnicas demonstram excelentes resultados ao separar dados em classes distintas, utilizando critérios de otimização matemática para problemas associados a regressão numérica. Além disso, sua capacidade de lidar com dados em alta dimensão destaca-se, sendo

particularmente relevante para problemas na indústria de petróleo e gás, onde a coleta de informações em tempo real por meio de numerosos sensores é comum.

## 5.2 Fonte dos Dados

Os dados CPTu deste trabalho foram disponibilizados através do *Database* do *Institute of Soil Mechanics, Foundation Engineering and Computational Geotechnics*, onde, fornece 1339 testes de solo (CPT e CPTu) em diferentes ambientes da Áustria e Alemanha, executados pela empresa *Premstaller Geotechnik ZT GmbH*, tais dados podem ser obtidos através de Oberhollenzer (2020).

Para o propósito deste trabalho em específico, foram utilizados dados do ambiente de uma distância relativamente próxima da localidade de Flachgau, uma porção de Salzburgo na Áustria, no total com 6 pontos de interesse coletados, onde um será utilizado como validação.

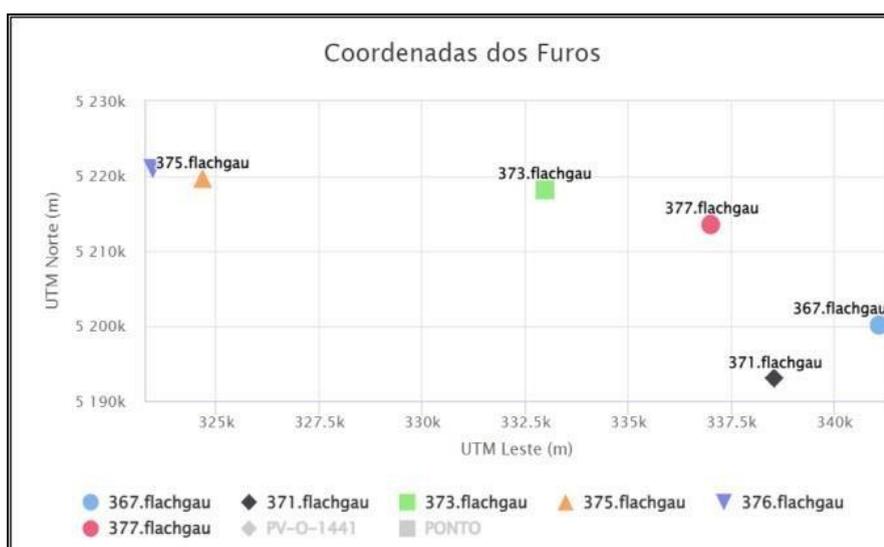
Todos os mais de 60000 dados foram organizados em um arquivo Excel para posterior implementação para o processo de krigagem e consequente utilização do SVM na linguagem Python.

Ademais, o *dataset* fornecido conta com todas as variáveis especificadas no tópico de Ensaio de penetração de cone.

## 5.3 Visualização do Problema

Primeiramente é importante visualizar a disposição espacial dos pontos coletados para o estudo deste trabalho, como demonstra a Figura 5.

**Figura 5 - Coordenadas UTM dos furos escolhidos**



Fonte: Autor (2024).

Logo, é possível observar a relativa proximidade entre os pontos escolhidos, com uma variação de coordenadas entre 5190000 m e 5225000 m UTM Norte e 320000 m e 342500 m UTM Leste.

## 6 RESULTADOS E DISCUSSÕES

Nesta seção, examina-se de perto as análises comparativas entre as duas metodologias, destacando padrões, tendências e avaliando a eficácia de cada abordagem.

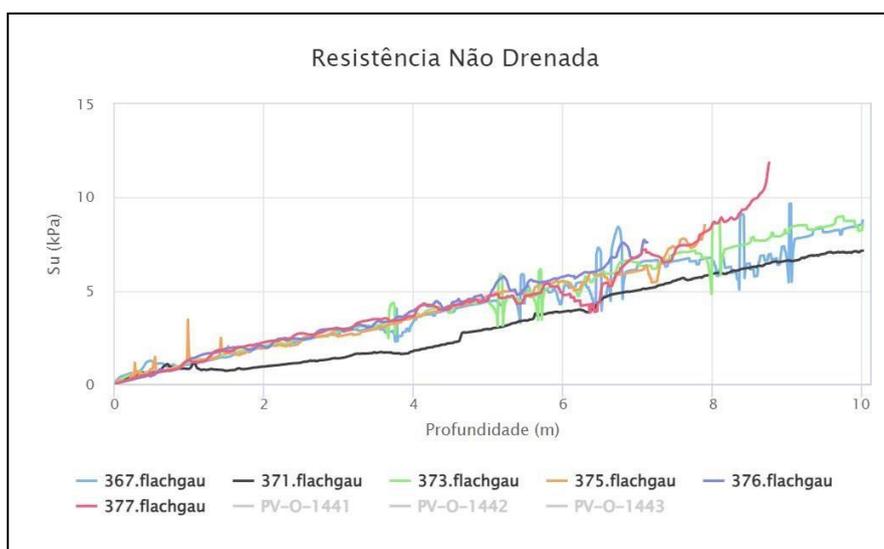
A compreensão detalhada desses resultados é essencial para informar as práticas futuras na caracterização do solo em projetos geotécnicos, equilibrando as inovações proporcionadas pela inteligência artificial com a confiabilidade e praticidade das técnicas tradicionais.

### 6.1 Uso da Krigagem

Primeiramente, se fez uma espécie de validação dos dados ao se omitir um ponto conhecido, e implementar a Krigagem para a previsão das características nesse ponto. O furo virtual escolhido foi o denominado “373.flachgau” com coordenadas de 5218223.39 m UTM Norte e 332977.22 m UTM Leste.

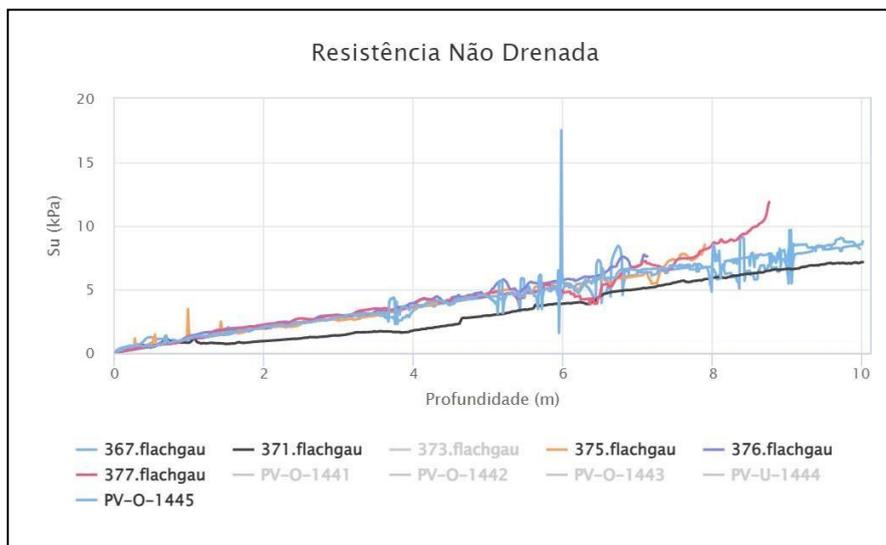
As Figuras 6 e 7 abaixo demonstram a previsão feita pela krigagem ordinária em comparação com o ponto real, utilizando a resistência não drenada como parâmetro previsto.

**Figura 6 - Resistência não drenada dos 6 furos conhecidos**



Fonte: Autor (2023).

**Figura 7 - Previsão da resistência não drenada omitindo um furo conhecido**



Fonte: Autor (2023).

Para melhor visualização, a figura 8 compara diretamente os dados previstos e os dados reais.

**Figura 8 - Comparação do valor previsto e real**



Fonte: Autor (2024).

Neste exemplo, é possível calcular o erro quadrático médio para comparação da previsão feita e do dado real, utilizando a média dos valores previstos e reais.

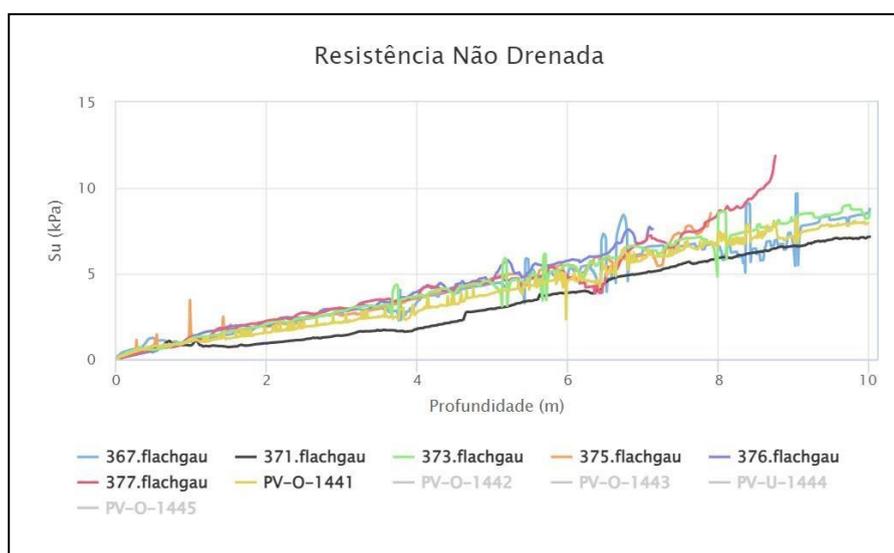
Com isso, chegamos a um erro quadrático médio de aproximadamente **0,0286**.

O resultado obtido, com um erro de aproximadamente 0,0286, está dentro do limite considerado muito bom na literatura geotécnica, que geralmente estabelece um limite de 0,05. Além disso, é visto como dentro do esperado, pela eficácia do método

da Krigagem em situações comuns como essa. Considerando apenas o resultado encontrado neste estudo, podemos afirmar que a previsão da resistência não drenada do solo utilizando a Krigagem ordinária foi bastante precisa. Isso sugere que o modelo desenvolvido com base nos dados CPTu disponibilizados e implementado utilizando a metodologia proposta é eficaz na previsão desse parâmetro geotécnico essencial. Portanto, podemos concluir que o resultado é adequado e oferece uma estimativa confiável da resistência do solo, o que é fundamental para a tomada de decisões em projetos geotécnicos.

Feito isso, é possível realizar a aplicação de Krigagem para um ponto virtual aleatório escolhido, para a previsão. O ponto escolhido foi o com coordenadas 5200000 m UTM Norte e 300000 m UTM Leste, gerando a previsão que demonstra a Figura 9 abaixo.

**Figura 9 - Previsão da resistência não drenada para um furo escolhido**



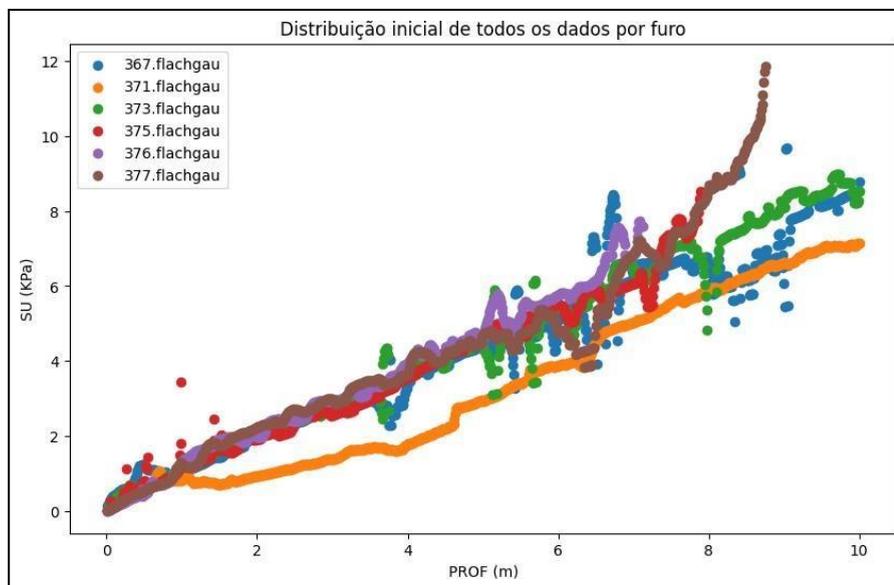
Fonte: Autor (2023).

## 6.2 Implementação do SVM

Primeiramente, para utilizar o Support Vector Machine (SVM) na linguagem de programação Python e realizar a implementação na análise dos dados CPTu, foi essencial conduzir um abrangente pré-processamento desses dados. Esse procedimento envolveu a remoção de valores ausentes, o escalonamento dos dados e todas as etapas relacionadas à consideração dos parâmetros previamente explicados.

Com isso, a Figura 10 detalha a distribuição inicial dos dados CPTu, com enfoque na resistência não drenada.

**Figura 10 - Distribuição inicial dos dados CPTu no Python**



Fonte: Autor (2024).

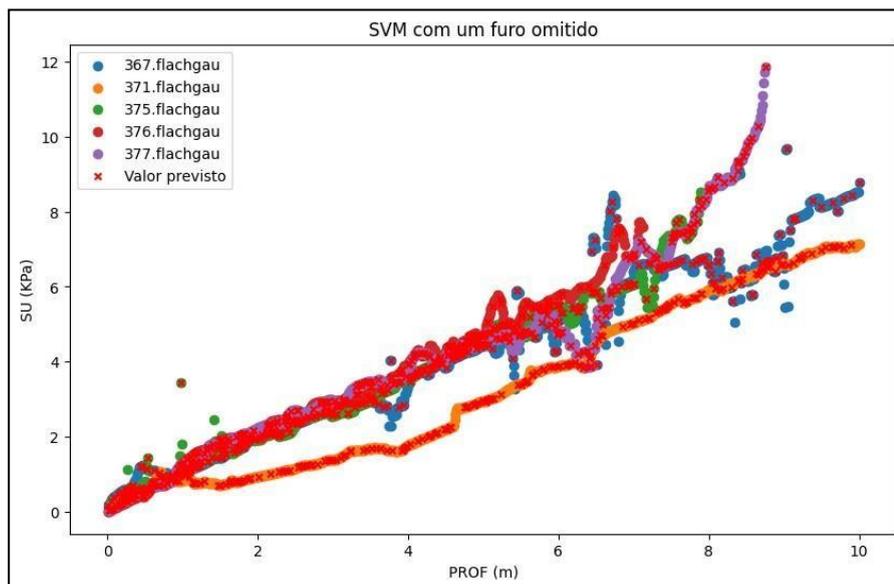
Na implementação do Support Vector Machine (SVM) para previsão da resistência não drenada do solo, foram cuidadosamente considerados diversos detalhes. Em relação aos hiper parâmetros utilizados, optou-se por um kernel RBF (*Radial Basis Function*) e um parâmetro de regularização C igual a 1.0, com um termo de tolerância de erro (*epsilon*) fixado em 0.2. A escolha desses hiper parâmetros foi baseada em experimentação e validação cruzada para garantir um equilíbrio entre o ajuste aos dados de treinamento e a capacidade de generalização para novos dados.

Além disso, foi realizada uma verificação de *overfitting* para avaliar a capacidade do modelo de generalizar para dados não vistos. Essa análise envolveu a divisão do conjunto de dados em conjuntos de treinamento e teste, seguida pela monitorização da métrica do erro de treinamento ao longo do processo de treinamento. A ausência de *overfitting* foi garantida pelo ajuste adequado dos hiperparâmetros e pela validação do modelo em um conjunto de dados de teste separado.

Esses detalhes são cruciais para garantir a confiabilidade e a eficácia do modelo desenvolvido e fornecem uma base sólida para a discussão dos resultados obtidos.

Dessa forma, após o pré-processamento, conseguimos implementar com sucesso o Support Vector Machine (SVM) para a previsão do ponto oculto designado como "373.flachgau". O gráfico resultante dessa previsão pode ser visualizado na Figura 11.

**Figura 11 - Previsão do SVM omitindo um furo conhecido**



Fonte: Autor (2024).

Com isso, é possível selecionar ponto a ponto da previsão feita e conseguir o ponto correlato à resistência não drenada em determinada profundidade.

Dessa maneira, foi possível calcular um erro quadrático médio de aproximadamente **0,0973**.

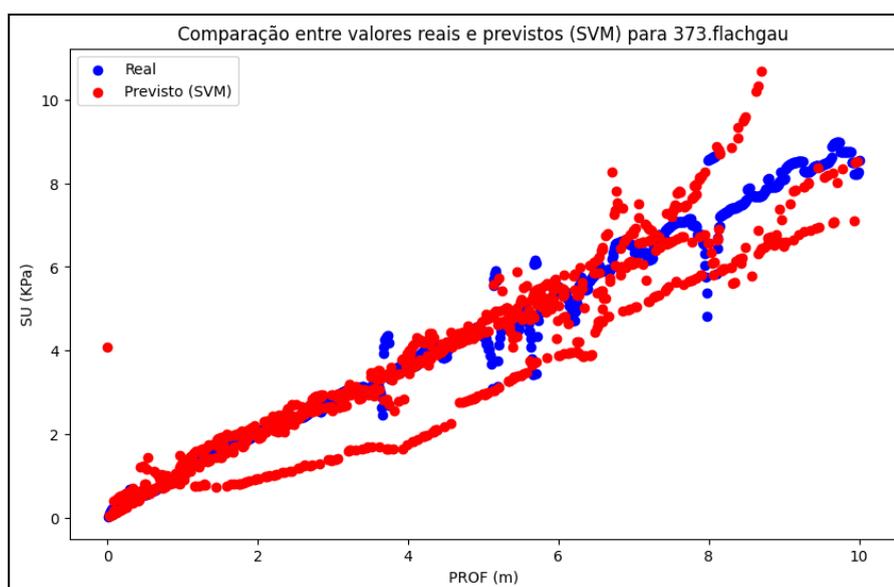
Inicialmente, a expectativa era de que o SVM oferecesse uma melhoria significativa na precisão das previsões em comparação com a krigagem. No entanto, os resultados revelaram um erro quadrático médio de aproximadamente 0,0973, o que pode ser considerado aceitável, mas não demonstra uma clara superioridade sobre a técnica anteriormente utilizada.

Essa disparidade nos resultados entre o SVM e a krigagem pode ser atribuída a diversas razões, que merecem uma análise detalhada. Uma possível explicação pode estar relacionada à sensibilidade dos hiperparâmetros escolhidos para o SVM, como o kernel RBF e os parâmetros de regularização  $C$  e  $\epsilon$ . Apesar dos esforços em otimizar esses hiperparâmetros, é possível que não tenham sido escolhidos os valores ideais para o conjunto de dados específico, resultando em um desempenho subótimo do modelo.

Além disso, as características intrínsecas dos dados CPTu, como a complexidade das relações entre os parâmetros e a variabilidade espacial do solo, podem influenciar significativamente o desempenho do SVM. É importante considerar que o SVM é uma técnica de aprendizado supervisionado que pressupõe relações lineares ou não lineares entre os atributos de entrada e a variável de saída. Se essas relações forem complexas ou não lineares, o SVM pode não ser capaz de capturá-las efetivamente, levando a previsões menos precisas.

Para fins de melhor comparação, a figura 12 apresenta somente a previsão feita utilizando SVM e os dados reais do “373.flachgau”

**Figura 12 - Comparação valores reais e previstos (SVM)**



Fonte: Autor (2024).

Logo, é possível observar que a zona final apresentou um ajuste pior comparada a zona inicial e intermediária, essa observação é fundamental para compreender as limitações de cada técnica em diferentes contextos geotécnicos. Enquanto a krigagem é uma técnica de interpolação espacial que se baseia em modelos geoestatísticos para estimar valores em locais não amostrados, o SVM é uma abordagem de aprendizado de máquina que busca identificar padrões nos dados e fazer previsões com base nesses padrões identificados.

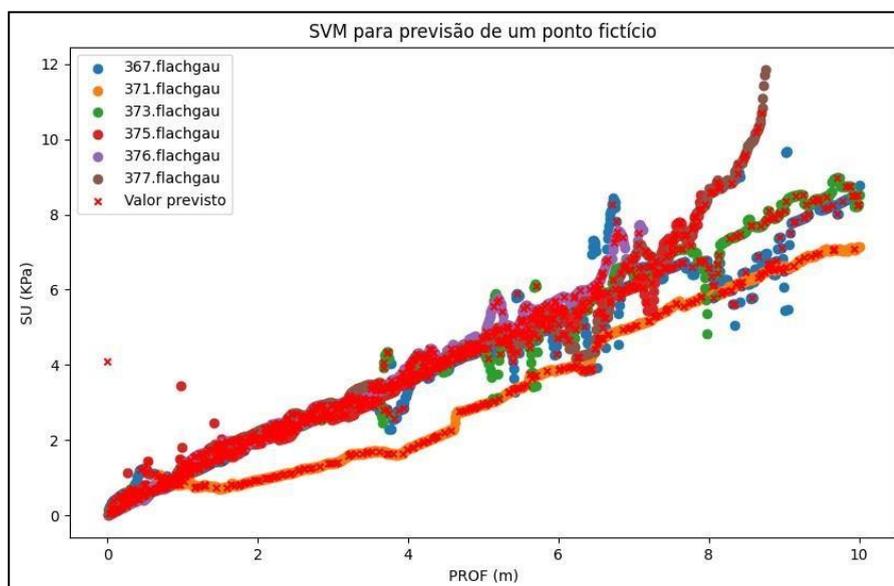
A zona final, que geralmente corresponde a áreas com menor densidade amostral ou maior variabilidade dos dados, pode representar um desafio para ambas as técnicas. No entanto, a figura comparativa revela que o SVM pode ter uma

capacidade limitada de generalização nessas regiões, resultando em previsões menos precisas ou ajustes menos satisfatórios em comparação com a Krigagem.

Essa discrepância no ajuste entre as duas técnicas ressalta a importância de considerar a variabilidade espacial dos dados e a adequação do modelo escolhido para o contexto específico do estudo. Enquanto a Krigagem pode oferecer uma estimativa mais suave e gradual dos valores em regiões com poucos dados, o SVM pode ser mais sensível a *outliers* e variações abruptas nos dados, o que pode afetar sua capacidade de generalização em áreas com características distintas das amostradas.

Seguindo a abordagem anterior utilizada na Krigagem, aplicamos o SVM a um ponto virtual aleatório, mantendo as coordenadas em 5200000 m UTM Norte e 300000 m UTM Leste para efeitos de comparação. A previsão gerada para esse ponto é ilustrada na Figura 13.

**Figura 13 - Resistência não drenada para um ponto fictício com SVM**



Fonte: Autor (2024).

## 7 CONCLUSÕES

No âmbito da indústria de óleo e gás, onde a segurança e eficiência na exploração de hidrocarbonetos são fundamentais, a caracterização precisa do solo desempenha um papel crucial. Neste estudo, exploramos a aplicação da inteligência artificial, em particular o algoritmo Support Vector Machine (SVM), na previsão de

parâmetros geotécnicos a partir de dados do Cone Penetration Test (CPTu) em ambientes específicos da Áustria e Alemanha.

Ao comparar os resultados obtidos com a abordagem tradicional de Krigagem, com foco na resistência não drenada do solo ( $S_u$ ), identificamos padrões que fornecem insights valiosos para a engenharia geotécnica. De maneira significativa, observamos que, apesar do potencial inovador da inteligência artificial, a Krigagem apresentou uma eficácia superior, evidenciada por um padrão de erro menor na previsão dos parâmetros geotécnicos. Na estratégia de omissão de um furo conhecido, a técnica consolidada apresentou um erro quadrático médio de 0,0286 em comparação à 0,0973 com o SVM.

A Krigagem, já consolidada na geoestatística, demonstrou ser não apenas mais eficiente, mas também mais fácil de ser utilizada em comparação com o SVM. Sua simplicidade aliada à precisão na interpolação geotécnica a tornam uma escolha prática e robusta para a previsão de parâmetros em estudos geotécnicos.

Embora a inteligência artificial, representada pelo SVM, tenha seu mérito em diversas aplicações, sua implementação bem-sucedida demanda cuidados específicos, como a qualidade dos conjuntos de dados, pré-processamento adequado e interpretação precisa dos resultados.

Em suma, este estudo reforça que, no contexto específico de previsão de parâmetros geotécnicos, a Krigagem continua a ser uma alternativa confiável, eficiente e de fácil aplicação quando comparada a abordagens mais recentes baseadas em inteligência artificial. Essa conclusão não apenas ressalta a importância de escolher a metodologia mais adequada para cada cenário geotécnico, mas também destaca o papel contínuo das técnicas tradicionais na otimização dos processos de construção de poços.

Ressalta-se a importância de uma abordagem integrada que combine métodos tradicionais e técnicas de inteligência artificial para obter resultados mais abrangentes e precisos na caracterização do solo e na previsão de parâmetros geotécnicos. Ao capitalizar os pontos fortes de cada abordagem e mitigar suas limitações individuais, uma estratégia híbrida poderia oferecer uma visão mais completa do ambiente geotécnico, contribuindo para decisões mais fundamentadas na indústria de óleo e gás.

Há oportunidades significativas para aprimorar ainda mais nossa compreensão e capacidade preditiva. Recomenda-se, primeiramente, a expansão do conjunto de dados, incorporando uma variedade mais ampla de condições geográficas e parâmetros geotécnicos. Além disso, explorar a aplicação de outras técnicas de inteligência artificial, como redes neurais e algoritmos de aprendizado profundo, pode oferecer uma visão mais abrangente. Sugere-se, ainda, a realização de validações experimentais para garantir a robustez dos resultados em cenários do mundo real. A inclusão de outros parâmetros geotécnicos relevantes, análises de sensibilidade, avaliação de incertezas e a integração de dados multifontes são aspectos fundamentais para futuras pesquisas. Por fim, a aplicação dessas metodologias em diferentes regiões geográficas permitirá uma avaliação mais abrangente da sua eficácia. Essas recomendações visam não apenas enriquecer a pesquisa atual, mas também contribuir para avanços significativos na engenharia geotécnica aplicada à indústria de óleo e gás.

## REFERÊNCIAS

- CARVALHO, J. R. P.; ASSAD, E. D. **Análise espacial da precipitação pluviométrica no Estado de São Paulo**: Comparação de métodos de interpolação. *Engenharia Agrícola*, v. 25, n. 2, p. 377-384, 2005.
- CARVALHO, João Vítor & Eichelberger, Luiza & Oliveira, Camila & Bastos, Cezar & Alves, Antonio. (2019). **Análise da resistência não drenada de diferentes solos artificiais obtida por ensaios de laboratório**.
- DRUCKER H, BURGESS CJ, KAUMFMAN L, SMOLA AJ, VAPNIK V. **Support vector regression machines In Advances in neural information processing systems** 1997. p 155-161.
- EICHELBERGER, L. V. et al. **Avaliação da resistência não drenada de dois solos artificiais por diferentes ensaios de laboratório**. agosto de 2018.
- GATLIN, Carl, and R.L. Slobod. "The Alcohol Slug Process for Increasing Oil Recovery." *Trans.* 219 (1960): 46–53. doi: <https://doi.org/10.2118/1356-G>
- GIOVANINI, Adenilson. Semivariograma. Disponível em: <https://adenilsongiovanini.com.br/blog/semivariograma/>. 2022.
- JAKOB, Alberto Augusto Eichman. **A Krigagem como Método de Análise de Dados Demográficos.Campinas**, 2002.
- LUDERMIR, Teresa Bernarda. **Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências**. Recife, 2021.
- LUNNE, T., Robertson, P.K. and Powell, J.J.M. (1997) **Cone Penetration Testing in Geotechnical Practice**. Blackie Academic & Professional, Londres, 312 p.
- MAYNE, P. W., & Peuchen, J. (2018). **Evaluation of CPTU Nkt Cone Factor for Undrained Strength of Clays**. Georgia Institute of Technology, Atlanta, Georgia, USA; Fugro, Netherlands.
- MEHTA, R. **Big Data Analytics with Java**. Birmingham: Packt Publishing Ltd, 2017.
- MELLO, J. M. de et al. **Ajuste e seleção de modelos espaciais de semivariograma visando à estimativa volumétrica de Eucalyptus grandis**. *Scientia Forestalis*, n. 69, p. 25-37, 2005. Acesso em: 20 mar. 2024.
- MONTEIRO, Hugo Leonardo Salomão. **Análise de desempenho de diferentes leis de controle de vibrações torcionais em colunas de perfuração de poços de petróleo**. 2012. Dissertação (Mestrado em Dinâmica das Máquinas e Sistemas) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2012. doi:10.11606/D.18.2012.tde-26112012-084201. Acesso em: 2024-03-20.
- NORVIG, P., Russell S., **Inteligência Artificial**, editora Elsevier Brasil, 3ª edição, Rio de Janeiro, 2013, 1056p.

NELSON, Erik B.; Guillot, Dominique. **Well Cementing. Second Edition.** Sugar Lands: Schlumberger, 2006.

OBERHOLLENZER. Simon, et al., 2020. **Cone penetration test dataset Premstaller Geotechnik.**

OLIVEIRA, I. A. d., Campos, M. C. C., Marques Junior, J., Aquino, R. E. d., Teixeira, D. d. B., Silva, D. M. P. d., 2015, **Use of scaled semivariograms in the planning sample of soil chemical Properties in Southern amazonas**, Brazil. Revista Brasileira de Ciência do Solo 39(1), 31-39.

Remota y Sistemas de Información Espacial, Bol. Resumos Expandidos, 68-73.

ROSA, R. L. **Classificação de eventos indesejáveis na produção de petróleo offshore com aplicação de técnicas de inteligência artificial.** [Trabalho de Conclusão de Curso]. Universidade Federal Fluminense, Niterói, RJ, 2020.

SILVA, D. F. B. F. da. (2021). **Pré-processamento de Dados e Comparação entre Algoritmos de Machine Learning para a Análise Preditiva de Falhas em Linhas de Produção para o Controle** [Master's thesis, Instituto Superior de Engenharia do Porto]. Repositório Científico do Instituto Politécnico do Porto.  
<http://hdl.handle.net/10400.22/18266>

SILVA, F. S.; SOARES, G. J. P.; ANDRADE, R. S. **Brocas de Perfuração.** v. 2, 2012

SOUSA, B. V. P., Santos, F. C. de O. N., Rodrigues, F. C., dos Santos Filho, M. C., Souza, M. V. C. S., Feitosa, P. H., Dias, Y. M. S., & Silva, M. de J. (2012).

**Perfurações de poços de petróleo métodos e equipamentos utilizados.**

Caderno De Graduação - Ciências Exatas E Tecnológicas - UNIT - SERGIPE, 1(1), 103–108. Recuperado de

<https://periodicos.set.edu.br/cadernoexatas/article/view/212>

STURARO, J. R. & LANDIM, P. M. B.; (1995) - **Geoestatística indicativa aplicada à análise espacial.**

TENÓRIO, J. K. F., Várady Filho, C. A. F., Silva, P. H. M., Santos, J. P. L., Dias, R., & Cutrim, F. S. (2019). **Estimativa de resistência não drenada do solo marinho utilizando técnicas geoestatísticas.** Universidade Federal de Alagoas – UFAL.

THOMAS, J.E., **Fundamentos de Engenharia de Petróleo**, editora Interciência, 2ª edição, Rio de Janeiro, 2004, 271p.

VII Simpósio Latinoamericano de Percepción Remota. La Sociedad Latinoamericana de Percepción