

INSTITUTO FEDERAL GOIANO
CAMPUS RIO VERDE
BACHARELADO EM ENGENHARIA CIVIL

LUCAS SOARES COSTA

**APLICAÇÃO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA NA ESTIMATIVA
DO FATOR DE SEGURANÇA DE TALUDES COM DADOS SIMULADOS PARA
DIFERENTES MÉTODOS DE EQUILÍBRIO LIMITE**

Trabalho de Curso

Rio Verde
2026

LUCAS SOARES COSTA

**APLICAÇÃO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA NA
ESTIMATIVA DO FATOR DE SEGURANÇA DE TALUDES COM DADOS
SIMULADOS PARA DIFERENTES MÉTODOS DE EQUILÍBRIO LIMITE**

Trabalho de Curso apresentado ao
Instituto Federal Goiano, como parte
dos requisitos necessários à obtenção
do título de Bacharel em Engenharia
Civil

Orientador: Charles Pereira Chaves

Rio Verde
2026

**Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema Integrado de Bibliotecas do IF Goiano - SIBi**

L933s Soares Costa, Lucas
Aplicação de Algoritmos de Aprendizado de Máquina na
Estimativa do Fator de Segurança de Taludes com Dados
Simulados para Diferentes Métodos de Equilíbrio Limite / Lucas
Soares Costa. Rio Verde 2026.

94f. il.

Orientador: Prof. Me. Charles Pereira Chaves.
Tcc (Bacharel) - Instituto Federal Goiano, curso de 0220084 -
Bacharelado em Engenharia Civil - Integral - Rio Verde (Campus
Rio Verde).

1. Aprendizado de Máquina. 2. Inteligência Artificial. 3.
Engenharia Geotécnica. 4. Estabilidade de Taludes. I. Título.

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR PRODUÇÕES TÉCNICO-CIENTÍFICAS NO REPOSITÓRIO INSTITUCIONAL DO IF GOIANO

Com base no disposto na Lei Federal nº 9.610, de 19 de fevereiro de 1998, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia Goiano a disponibilizar gratuitamente o documento em formato digital no Repositório Institucional do IF Goiano (RIIF Goiano), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IF Goiano.

IDENTIFICAÇÃO DA PRODUÇÃO TÉCNICO-CIENTÍFICA

- | | |
|--|---|
| <input type="checkbox"/> Tese (doutorado) | <input type="checkbox"/> Artigo científico |
| <input type="checkbox"/> Dissertação (mestrado) | <input type="checkbox"/> Capítulo de livro |
| <input type="checkbox"/> Monografia (especialização) | <input type="checkbox"/> Livro |
| <input checked="" type="checkbox"/> TCC (graduação) | <input type="checkbox"/> Trabalho apresentado em evento |

Produto técnico e educacional - Tipo:

Nome completo do autor:

LUCAS SOARES COSTA

Matrícula:

2021102200840129

Título do trabalho:

Aplicação de Algoritmos de Aprendizado de Máquina na Estimativa do Fator de Segurança de Taludes com Dados Simulados para Diferentes Métodos de Equilíbrio Limite

RESTRIÇÕES DE ACESSO AO DOCUMENTO

Documento confidencial: Não Sim, justifique:

Informe a data que poderá ser disponibilizado no RIIF Goiano: / /

O documento está sujeito a registro de patente? Sim Não

O documento pode vir a ser publicado como livro? Sim Não

DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O(a) referido(a) autor(a) declara:


- Que o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- Que obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autoria, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia Goiano os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- Que cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia Goiano.

Rio Verde – GO

Local

/ /

Data

Documento assinado digitalmente
 LUCAS SOARES COSTA
Data: 10/04/2026 12:03:18-0300
Verifique em <https://validar.iti.gov.br>

Assinatura do autor e/ou detentor dos direitos autorais

Ciente e de acordo:

Assinatura do(a) orientador(a)



Documento assinado digitalmente

CHARLES PEREIRA CHAVES

Data: 10/04/2026 17:39:19-0300

Verifique em <https://validar.iti.gov.br>

Regulamento de Trabalho de Curso (TC) – IF Goiano - Campus Rio Verde

ANEXO V - ATA DE DEFESA DE TRABALHO DE CURSO

Ao vigésimo terceiro dia do mês de março de dois mil e vinte e seis, às 14 horas, reuniu-se a Banca Examinadora composta por: Prof. Charles Pereira Chaves (orientador), Prof. Philippe Barbosa Silva e André Luis Brasil Cavalcante (UNB), para examinar o Trabalho de Curso (TC) intitulado “APLICAÇÃO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA NA ESTIMATIVA DO FATOR DE SEGURANÇA DE TALUDES COM DADOS SIMULADOS PARA DIFERENTES MÉTODOS DE EQUILÍBRIO LIMITE” de Lucas Soares Costa, estudante do curso de ENGENHARIA CIVIL do IF Goiano – Campus Rio Verde, sob Matrícula nº 2021102200840129. A palavra foi concedida ao(à) estudante para a apresentação oral do TC, em seguida houve arguição do candidato pelos membros da Banca Examinadora. Após tal etapa, a Banca Examinadora decidiu pela APROVAÇÃO COM RESSALVA do(a) estudante desde que seja apresentado o TC na forma escrita com as devidas correções à banca novamente até o dia 20/04/2026, caso não seja enviada as correções será considerado REPROVADO. Ao final da sessão pública de defesa foi lavrada a presente ata, que, foi assinada pelos membros da Banca Examinadora e Mediador de TC.

Rio Verde, 23 de março de 2026.

Charles Pereira Chaves

Membro 1 da Banca Examinadora

Philippe Barbosa Silva

Membro 2 da Banca Examinadora

André Luis Brasil Cavalcante*

Membro 3 da Banca Examinadora

*Neste ato o orientador assina pelo membro externo.

Documento assinado eletronicamente por:

- **Charles Pereira Chaves, COORDENADOR(A) DE CURSO - FUC1 - CCBEC-RV**, em 10/04/2026 10:56:30.
- **Philippe Barbosa Silva, PROFESSOR ENS BASICO TECN TECNOLOGICO**, em 10/04/2026 11:26:01.

Este documento foi emitido pelo SUAP em 10/04/2026. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifgoiano.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código Verificador: 809365

Código de Autenticação: 4ed9b25209



RESUMO

A estabilidade de taludes é um dos principais temas da engenharia geotécnica, sendo essencial para garantir a segurança e a funcionalidade de diversas obras civis. A análise da estabilidade envolve uma série de variáveis geotécnicas, geométricas e hidrológicas, sendo tradicionalmente conduzida por meio de métodos de equilíbrio limite, como os de Fellenius, Bishop e Janbu. Com a evolução tecnológica, ferramentas computacionais passaram a permitir análises mais sofisticadas e representativas.

Recentemente, a aplicação de técnicas de inteligência artificial, em especial o machine learning, tem se destacado como uma alternativa promissora para a predição da estabilidade de taludes, oferecendo maior eficiência na análise de dados e na modelagem de incertezas. Este trabalho propõe o desenvolvimento de um modelo preditivo baseado em algoritmos de machine learning, utilizando dados simulados por softwares geotécnicos. O objetivo é avaliar a capacidade dos modelos em estimar o fator de segurança de taludes e compará-los aos métodos tradicionais, visando contribuir para o avanço da inteligência artificial na engenharia geotécnica.

Palavras-chave: Estabilidade de taludes, engenharia geotécnica, equilíbrio limite, aprendizado de máquina, modelagem preditiva.

ABSTRACT

Slope stability is a fundamental aspect of geotechnical engineering and plays a crucial role in ensuring the safety and performance of civil and environmental engineering projects. Traditional slope stability assessments rely on limit equilibrium methods, such as those developed by Fellenius, Bishop, Janbu, and others, which evaluate the safety factor based on simplified assumptions about soil behavior. With the evolution of computational tools, geotechnical software has made it possible to conduct more realistic and complex simulations, enhancing the reliability of engineering designs.

Recently, artificial intelligence techniques, especially machine learning, have emerged as promising tools for improving slope stability analysis by better addressing uncertainties and reducing the limitations of conventional methods. This study proposes the development of a predictive model using machine learning algorithms trained on simulated data generated from geotechnical software. The objective is to evaluate the performance of different algorithms in estimating the safety factor of slopes and to compare their accuracy and efficiency with traditional methods. The research aims to contribute to the advancement of AI applications in geotechnical engineering and to promote safer, more efficient design practices.

Palavras-chave: Slope stability; Geotechnical engineering; Limit equilibrium; Machine learning; Predictive modeling.

LISTA DE FIGURAS

Figura 1: Diferença entre tensão efetiva, poropressão e tensão cisalhante	14
Figura 2: Empuxo em estruturas de contenção	15
Figura 3: Exemplo de divisão em fatias de uma superfície circular	16
Figura 4: Procura da superfície crítica	17
Figura 5: Exemplos de abordagens de aprendizado de máquina.	22
Figura 6: Separação de classes por fronteira de decisão (classificação) e ajuste de uma função para prever valores contínuos (regressão).	23
Figura 7: Comparação na generalização e ajustes de modelos.....	24
Figura 8: Relação entre a complexidade do classificador e a acurácia de generalização para diferentes quantidades de dados de treinamento	25
Figura 9 : Trade-off de explicabilidade e complexidade no aprendizado de máquina.	26
Figura 10: Estrutura de uma Árvore de decisão	28
Figura 11: Estrutura de uma Floresta Aleatória.....	29
Figura 12: Distância entre hiperplanos e vetores suporte.	31
Figura 13: Transformação do plano utilizando o truque de kernel.....	32
Figura 14: Representação esquemática de uma Rede Neural de Múltiplas Camadas	34
Figura 15: Representação das estratégias de Bagging e Boosting.....	36
Figura 16: Matriz de confusão para avaliação de modelos de classificação.	39
Figura 17: Exemplos de curvas ROC com diferentes níveis de desempenho de modelos de classificação binária.....	40
Figura 18: Histograma de distribuição do peso específico	49
Figura 19: Histograma de distribuição da coesão.	50
Figura 20: Histograma de distribuição do ângulo de atrito.	50
Figura 21: Histograma de distribuição da altura do talude.	51
Figura 22: Histograma de distribuição da inclinação do talude.	51
Figura 23: Histograma de distribuição da poropressão.....	52
Figura 24: Exemplo de talude com FS calculado.	53
Figura 25: Critérios de convergência adotados.	54
Figura 26 : Trecho do código Python referente à definição das variáveis de entrada e saída e à divisão do conjunto de dados em treinamento e teste.....	55
Figura 27: Definição da validação cruzada com 5 partições (K-Fold)	56
Figura 28 : Trecho do código referente à otimização dos modelos por meio de Grid Search com validação cruzada para seleção dos melhores hiperparâmetros	57
Figura 29: Trecho do código referente à aplicação do modelo ao conjunto de teste e ao cálculo das métricas.....	58
Figura 30: Representação esquemática geral do funcionamento do algoritmo de aprendizado de máquina para este trabalho.	59
Figura 31: Fluxograma da metodologia do trabalho	65
Figura 32 : Histograma dos fatores de segurança calculados.....	67
Figura 33 : Comparação do EQM entre os modelos de aprendizado de máquina e métodos de equilíbrio limite.....	68
Figura 34 : Comparação do R^2 entre os modelos de aprendizado de máquina e métodos de equilíbrio limite.....	69
Figura 35 : Comparação da Acurácia entre os modelos de aprendizado de máquina e métodos de equilíbrio limite.....	70
Figura 36 : Comparação do <i>CV-score</i> entre os modelos de aprendizado de máquina	

e métodos de equilíbrio limite.....	71
Figura 37 : Curvas ROC obtidas para Árvore de Decisão	72
Figura 38 : Curvas ROC obtidas para Floresta Aleatória	73
Figura 39: Curvas ROC obtidas para RNA.....	74
Figura 40: Curvas ROC obtidas para SVM	75
Figura 41: Matrizes de confusão obtidas para Árvore de Decisão	76
Figura 42: Matrizes de confusão obtidas para Floresta Aleatória.....	77
Figura 43: Matrizes de confusão obtidas para RNA	78
Figura 44: Matrizes de confusão para SVM	79
Figura 45: Importância das variáveis para Árvore de Decisão	80
Figura 46: Importância das variáveis para Árvore de Decisão	81
Figura 47 : Importância das Variáveis para RNA	82
Figura 48: Importância das Variáveis para SVM	83

LISTA DE TABELAS

Tabela 1: Fatores de segurança mínimos para escorregamentos.	16
Tabela 2: Comparação entre os diferentes Métodos de Equilíbrio Limite	19
Tabela 3: Comparação entre os artigos analisados.	44
Tabela 4: Valores para os parâmetros geotécnicos e geométricos.....	48
Tabela 5: Valores calculados para o fator de segurança	66

SUMÁRIO

RESUMO	3
ABSTRACT	4
LISTA DE FIGURAS	5
LISTA DE TABELAS	6
SUMÁRIO	7
1 INTRODUÇÃO	10
2 REVISÃO BIBLIOGRÁFICA	12
2.1 Análise de Estabilidade de Taludes	12
2.1.1 Forças atuantes em um talude	12
2.1.2 Fatores de segurança e superfícies de ruptura	15
2.1.3 Métodos de Equilíbrio Limite	18
2.2 Inteligência Artificial e Aprendizado de Máquina	20
2.2.1 Histórico da Inteligência Artificial	20
2.2.2 Aprendizado de Máquina (<i>Machine Learning</i>)	20
2.2.3 Exemplos de modelos de aprendizado de máquina	25
2.2.3.1 Árvore de Decisão	26
2.2.3.2 Floresta Aleatória	29
2.2.3.3 Máquinas de Vetores Suporte	31
2.2.3.4 Redes Neurais Artificiais	33
2.2.4 Aprendizado em conjunto (<i>Ensemble Learning</i>)	34
2.2.5 Padronização e Normalização de dados	36
2.2.6 Validação cruzada	37
2.3 Métodos de classificação e avaliação dos modelos	38
2.4 Estado da arte	41
3 OBJETIVOS	46
3.2 Objetivo geral	46
3.3 Objetivos específicos	46
4 METODOLOGIA	47
4.1 Levantamento estatístico dos parâmetros geotécnicos e geométricos	47
4.2 Cálculo do Fator de Segurança via Análise de Equilíbrio Limite	52
4.3 Construção do algoritmo para treinamento e teste dos modelos	54
4.3.1 Árvore de decisão	60
4.3.2 Floresta Aleatória	61
4.3.3 Máquinas de Vetores Suporte (SVM)	61
4.3.4 Redes Neurais Artificiais	62
4.4 Avaliação dos modelos	63

5	RESULTADOS E DISCUSSÕES.....	66
5.1	Resultados do fator de segurança obtidos pelo método de equilíbrio limite	66
5.2	Desempenho dos modelos de aprendizado de máquina	67
5.2.1	Análise do Erro Quadrático Médio.....	67
5.2.2	Análise do Coeficiente de Determinação (R^2).....	68
5.2.3	Análise da Acurácia	69
5.2.4	Análise do <i>CV-score</i>	70
5.2.5	Análise da Curva ROC	71
5.2.6	Análise da Matriz de Confusão.....	75
5.2.7	Análise da Importância das Variáveis	80
5.2.8	Desempenho dos modelos de aprendizado de máquina sob a perspectiva geotécnica	84
5.2.8.1	Interpretação física do desempenho dos modelos	84
5.2.8.2	Influência dos métodos de equilíbrio limite	85
5.3	Limitações do modelo e análise do uso do aprendizado de máquina para estabilidade de taludes.....	86
6	CONCLUSÕES.....	89
6.1	Desempenho dos modelos de aprendizado de máquina.....	89
6.2	Desempenho dos métodos de equilíbrio limite	90
6.3	Limitações e recomendações para trabalhos futuros	90
	REFERÊNCIAS.....	92

LISTA DE SIGLAS

ACC – Acurácia
ANN – Artificial Neural Network (Rede Neural Artificial)
AUC – Área sob a Curva ROC
BPSO – Binary Particle Swarm Optimization (Otimização por Enxame de Partículas Binária)
CNN – Convolutional Neural Network (Rede Neural Convolutacional)
DT – Decision Tree (Árvore de Decisão)
EMA – Erro Médio Absoluto
EQM – Erro Quadrático Médio
F1 – Escore F1
FEM – Finite Element Method (Método dos Elementos Finitos)
FS – Fator de Segurança
GB – Gradient Boosting
GBC – Gradient Boosting Classifier
ID – Índice de Desempenho
k-NN – k-Nearest Neighbors (k Vizinhos Mais Próximos)
LEM – Limit Equilibrium Method (Método de Equilíbrio Limite)
LR – Logistic Regression (Regressão Logística)
MLP – Multilayer Perceptron (Perceptron Multicamadas)
MLR - Multiple Linear Regression (Regressão Linear Múltipla)
NB – Naïve Bayes
PREC – Precisão
R – Coeficiente de Correlação
R² – Coeficiente de Determinação
REC – Revocação (Sensibilidade)
REQM – Raiz do Erro Quadrático Médio
RF – Random Forest (Floresta Aleatória)
ROC – Receiver Operating Characteristic
RSR – Razão entre a Raiz do Erro Quadrático Médio e o Desvio Padrão
SRM – Strength Reduction Method (Método da Redução da Resistência)
SVM – Support Vector Machine (Máquina de Vetores de Suporte)
SVR – Support Vector Regression (Regressão por Vetores de Suporte)
VAF – Variância Explicada
VC – Validação Cruzada
XGB – Extreme Gradient Boosting (XGBoost)

1 INTRODUÇÃO

A estabilidade de taludes constitui um dos pilares fundamentais no âmbito da engenharia geotécnica. A sua aplicação abrange diversas áreas da engenharia civil e ambiental, sendo indispensável para a segurança e o desempenho de uma vasta gama de empreendimentos. A relevância da análise de estabilidade de taludes manifesta-se notadamente em cenários que envolvem a interação entre obras humanas e o meio geológico, abrangendo desde intervenções em áreas urbanas até projetos de grande escala na indústria e na infraestrutura.

Nesse contexto, a análise da segurança de estabilidade de taludes é de grande importância, pois permite avaliar o risco de ruptura de maciços de solo ou rocha sob diversas condições de carregamento e ambientais. A análise da estabilidade geralmente se torna o fator determinante em projetos geotécnicos, sendo expressa por meio de um fator de segurança mínimo admissível ou pela estimativa da probabilidade de ruptura (GERSCOVICH, 2016).

A estabilidade de taludes é influenciada por uma série de fatores inter-relacionados, que envolvem desde as características intrínsecas dos materiais geológicos até condições externas. As análises de estabilidade de taludes construídos consideram aspectos geométricos, como altura, comprimento e inclinação, além de carregamentos aplicados, propriedades mecânicas dos solos que compõem a estrutura e a interferência da água, incluindo seu fluxo e a consequente elevação das pressões neutras no interior do maciço.

Sendo assim, devido a essa complexidade, historicamente se desenvolveram métodos empíricos e modelos matemáticos para se analisar a estabilidade de taludes, com o objetivo de estimar o risco de ruptura e garantir a segurança das obras geotécnicas. Os primeiros métodos adotados foram gráficos e empíricos, baseados em observações de campo e ensaios simplificados. Posteriormente, surgiram os chamados métodos de equilíbrio limite (LEM), amplamente utilizados até os dias atuais para a determinação do fator de segurança (FS) em taludes naturais e artificiais.

Dentre os métodos clássicos de equilíbrio limite, destacam-se o método de Fellenius (1936), conhecido como método da fatia ordinária, o de Bishop (1955), Janbu (1954), Spencer (1967) e Morgenstern-Price (1965), cada um com suas aplicações específicas. Esses métodos avaliam a segurança de taludes pelo equilíbrio de forças e momentos em superfícies potenciais de ruptura, assumindo, entretanto, algumas simplificações: consideram o solo como rígido-plástico,

desconsideram o caráter dinâmico da ruptura e admitem o fator de segurança constante ao longo da superfície de ruptura (MASSAD, 2010).

Com o avanço da tecnologia computacional, a engenharia geotécnica incorporou o uso de softwares especializados, que ampliaram as possibilidades de modelagem e análise. Esses programas passaram a permitir simulações mais realistas e precisas, considerando diferentes geometrias, carregamentos, condições de fluxo e modelos constitutivos mais avançados, e possibilitaram análises complexas em tempo reduzido, aumentando a confiabilidade dos projetos e possibilitando a consideração de diferentes cenários.

Mais recentemente, técnicas de inteligência artificial, como o machine learning, vêm sendo aplicadas na engenharia geotécnica, oferecendo recursos adicionais para lidar com as incertezas inerentes aos parâmetros de solo e às condições operacionais. O uso desses métodos, no contexto da geotecnia, se justifica pela capacidade de reduzir os erros associados aos procedimentos tradicionais, tratar melhor as incertezas presentes na análise de estabilidade e exigir menor capacidade computacional em relação a análises numéricas convencionais. Além disso, essas técnicas permitem mensurar a acurácia e a performance das previsões, otimizando a tomada de decisão e a definição de estratégias de estabilização mais eficientes (NANEHAKARAN et al, 2023).

Diante desse contexto, este trabalho propõe o desenvolvimento de um modelo preditivo para análise de estabilidade de taludes, utilizando técnicas de *machine learning* aplicadas a dados simulados em softwares geotécnicos especializados. A pesquisa busca avaliar a capacidade de diferentes algoritmos em estimar o fator de segurança de taludes, considerando a variabilidade dos parâmetros geotécnicos e geométricos. Além disso, pretende-se identificar as vantagens e limitações desses métodos em comparação aos procedimentos tradicionais, visando contribuir com o avanço do uso da inteligência artificial na geotecnia e colaborar para a superação dos desafios associados à implementação dessa ferramenta, que representa a nova fronteira tecnológica do século XXI e possui enorme potencial para aumentar a eficiência, a confiabilidade e a segurança dos projetos geotécnicos.

2 REVISÃO BIBLIOGRÁFICA

2.1 Análise de Estabilidade de Taludes

2.1.1 Forças atuantes em um talude

A estabilidade de taludes é um dos temas centrais da engenharia geotécnica, dado seu papel fundamental na segurança de empreendimentos urbanos, viários, industriais e de infraestrutura. Taludes são superfícies inclinadas naturais ou artificiais, e sua segurança depende do equilíbrio entre as tensões atuantes e a resistência ao cisalhamento dos materiais constituintes. A instabilidade ocorre quando as forças cisalhantes mobilizadas ultrapassam as tensões resistentes disponíveis, levando a movimentos de massa que podem assumir diferentes formas.

As forças cisalhantes mobilizadas, também denominadas tensões cisalhantes atuantes, são aquelas responsáveis por provocar o escorregamento ou a deformação da massa de solo ou rocha ao longo de uma superfície potencial de ruptura. Sua principal origem é a componente tangencial do peso próprio do material que compõe o talude, atuando paralelamente à superfície inclinada. O **peso específico do solo (γ)** que compõe o talude contribui diretamente para o aumento do peso total da estrutura e, conseqüentemente, para a elevação das forças mobilizantes. Além disso, fatores geométricos, como a altura do **talude (H)** e sua **inclinação (β)**, influenciam significativamente nesse processo, uma vez que maiores alturas aumentam o peso da massa instável e inclinações mais acentuadas elevam a componente tangencial desse peso, potencializando as tensões cisalhantes atuantes na estrutura. Outros fatores como sobrecargas aplicadas no topo (edificações, tráfego), pressão hidrostática em fendas de tração e forças sísmicas também contribuem para o aumento das forças desestabilizadoras (GERSCOVICH, 2016).

As tensões resistentes correspondem à capacidade do solo em suportar as tensões cisalhantes mobilizadas sem sofrer ruptura, sendo determinadas pelas propriedades mecânicas do material, como **coesão (c)** e **ângulo de atrito interno (ϕ)**, que definem essencialmente dois fatores: a resistência por atrito entre partículas sólidas e a resistência por coesão. O atrito resulta da interação física entre os grãos minerais e é proporcional à tensão normal atuante no plano de cisalhamento, enquanto a coesão é a força de atração entre as partículas do solo, e pode ser real — resultante de ligações químicas ou cimentações naturais — ou

aparente, proveniente da tensão de sucção capilar em solos não saturados (PINTO, 2006).

A resistência ao cisalhamento dos solos é influenciada pelo tipo de material. Solos arenosos apresentam resistência predominantemente por atrito entre grãos, enquanto solos argilosos têm significativa participação da coesão, tanto real quanto aparente. Ensaio de laboratório, como compressão triaxial e cisalhamento direto, são empregados para caracterizar a resistência de solos em diferentes condições de drenagem e tensões confinantes (PINTO, 2006).

O comportamento resistente dos solos é frequentemente descrito pelo critério de ruptura de Mohr-Coulomb, que estabelece que a ruptura ocorre quando a tensão cisalhante mobilizada atinge um valor limite, dependente da coesão e da tensão normal (σ) atuante no plano de cisalhamento, conforme a expressão:

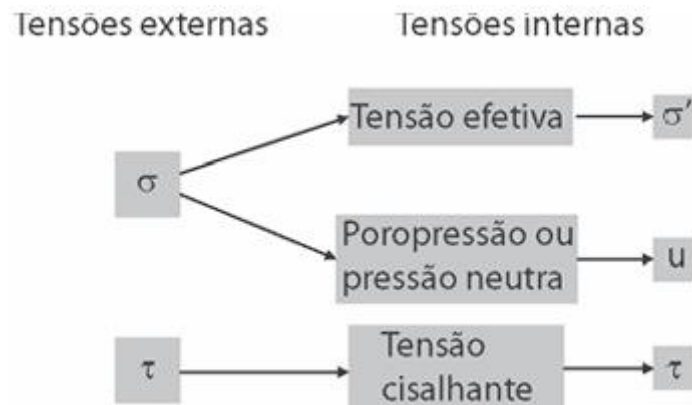
$$\tau = c + \sigma' \cdot \tan\varphi \quad \text{Eq. (1)}$$

Estudos numéricos confirmam a forte sensibilidade do fator de segurança às propriedades de resistência do solo e à geometria do talude. Por exemplo, Hulagabali et al. (2019) realizaram modelagens no software GEO-STUDIO variando coesão, ângulo de atrito interno, peso específico, altura e inclinação do talude; os autores observaram aumento do fator de segurança com o aumento da coesão e do ângulo de atrito, e redução do fator de segurança com o aumento da altura e da inclinação do talude. As variações do peso específico, por sua vez, mostraram influência desprezível no FS para o caso, indicando que os parâmetros de resistência (c , φ) dominam o comportamento de ruptura (Hulagabali et al., 2019).

O conceito de tensão efetiva (σ') é essencial para a análise da estabilidade, pois apenas a parcela das tensões normais transmitidas pelos contatos entre partículas sólidas contribui para a resistência ao cisalhamento, enquanto a porção suportada pela água, essa também conhecida como poropressão, ou pressão neutra (u), atua no sentido de reduzir essa resistência (GERSCOVICH, 2016).

A Figura 1 mostra que a tensão normal total aplicada ao solo (σ) se divide em tensão efetiva (σ') e poropressão (u), conforme o princípio de Terzaghi ($\sigma = \sigma' + u$). A tensão efetiva é transmitida pelo esqueleto do solo e controla sua resistência e deformação, enquanto a poropressão é suportada pela água nos vazios.

Figura 1: Diferença entre tensão efetiva, poropressão e tensão cisalhante



Fonte: Gresovich (2016)

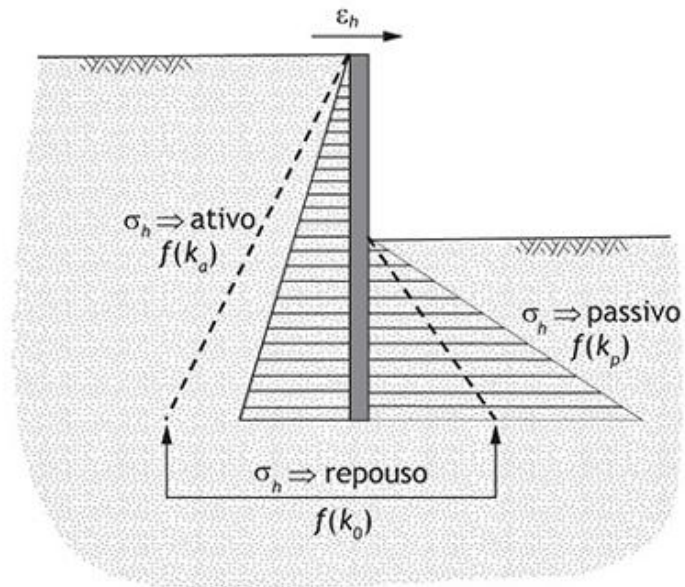
Por fim, a estabilidade de taludes também é condicionada pelos estados de empuxo atuantes no maciço, conforme representado na Figura 2. Conforme a teoria clássica da mecânica dos solos, o comportamento das tensões horizontais no solo pode ser classificado em três condições distintas: empuxo em repouso, empuxo ativo, e empuxo passivo. O coeficiente de empuxo (K) define a relação entre a tensão horizontal (σ_h) e a tensão vertical (σ_v) no solo:

$$K = \frac{\sigma_v}{\sigma_h} \quad \text{Eq. (2)}$$

O empuxo ativo ocorre quando o solo “empurra” a estrutura, ou seja, quando uma superfície de contenção ou da própria face do talude, tende a se afastar da massa de solo. Por outro lado, o empuxo passivo ocorre quando a estrutura é empurrada contra o solo, que reage resistindo a esse movimento (GRESOVICH et al., 2016). Essas duas condições são descritas pela Teoria de Rankine, na qual as tensões horizontais no solo, na condição ativa, são determinadas pelo coeficiente de empuxo ativo (K_a) enquanto, na condição passiva, são definidas pelo coeficiente de empuxo passivo (K_p).

Além disso, influência das tensões *in situ* também é relevante na estabilidade de taludes. Essas tensões existem no solo previamente, originadas pelo seu peso próprio, ou por adensamentos anteriores. O coeficiente de empuxo em repouso (K_0) descreve a razão entre as tensões horizontais e verticais em solos antes da aplicação de cargas adicionais.

Figura 2: Empuxo em estruturas de contenção



Fonte: Gresovich *et al.* (2016)

2.1.2 Fatores de segurança e superfícies de ruptura

Conhecendo as forças atuantes e resistentes, o principal objetivo da análise de estabilidade de taludes é encontrar o **Fator de Segurança (FS)**, que representa a razão entre as resistências média ao cisalhamento do solo (τ_f) e as tensões cisalhantes mobilizadas ao longo da superfície potencial de ruptura (τ_d) (DAS, 2011)

$$F_s = \frac{\tau_f}{\tau_d} \quad \text{Eq. (3)}$$

Valores de FS inferiores a 1 indicam ruptura iminente, enquanto valores superiores representam condições de segurança satisfatórias. As normas brasileiras, como a NBR 11682 (ABNT, 2008), estabelecem valores mínimos recomendados de FS em função do risco à vida humana e a bens materiais, usualmente variando de 1,2 a 1,5, dependendo da aplicação de acordo com a Tabela 1 (GERSCOVICH, 2016).

Tabela 1: Fatores de segurança mínimos para escorregamentos.

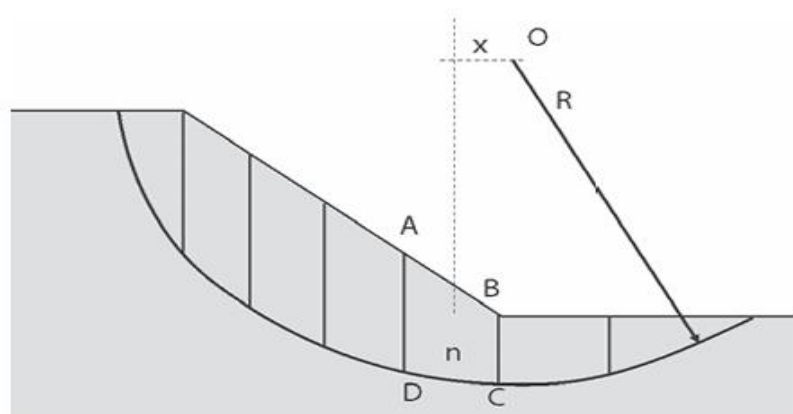
Nível de segurança contra danos materiais e ambientais	Nível de segurança contra danos a vidas humanas		
	Alto	Médio	Baixo
Alto	1,5	1,5	1,4
Médio	1,5	1,4	1,3
Baixo	1,4	1,3	1,2

Fonte: Gresovich (2016)

As análises de estabilidade consideram as possíveis superfícies de ruptura, que podem assumir formas circulares, planas, multiplanares ou mistas, a depender da heterogeneidade do maciço, da presença de descontinuidades e da geometria do talude (GERSCOVICH, 2016).

O método de equilíbrio limite (LEM) são os mais tradicionais e amplamente utilizados na prática geotécnica. Baseiam-se no princípio de que a ruptura ocorre ao longo de uma superfície potencial previamente definida e que os elementos do solo acima dessa superfície atingem simultaneamente o estado de equilíbrio limite ($FS = 1$). As análises podem ser feitas por procedimentos de massa, considerando toda a massa instável como um único bloco, ou pelo método das fatias (Figura 3), que subdivide a massa em fatias verticais e avalia o equilíbrio de forças e momentos para cada uma delas (GERSCOVICH, 2016).

Figura 3: Exemplo de divisão em fatias de uma superfície circular

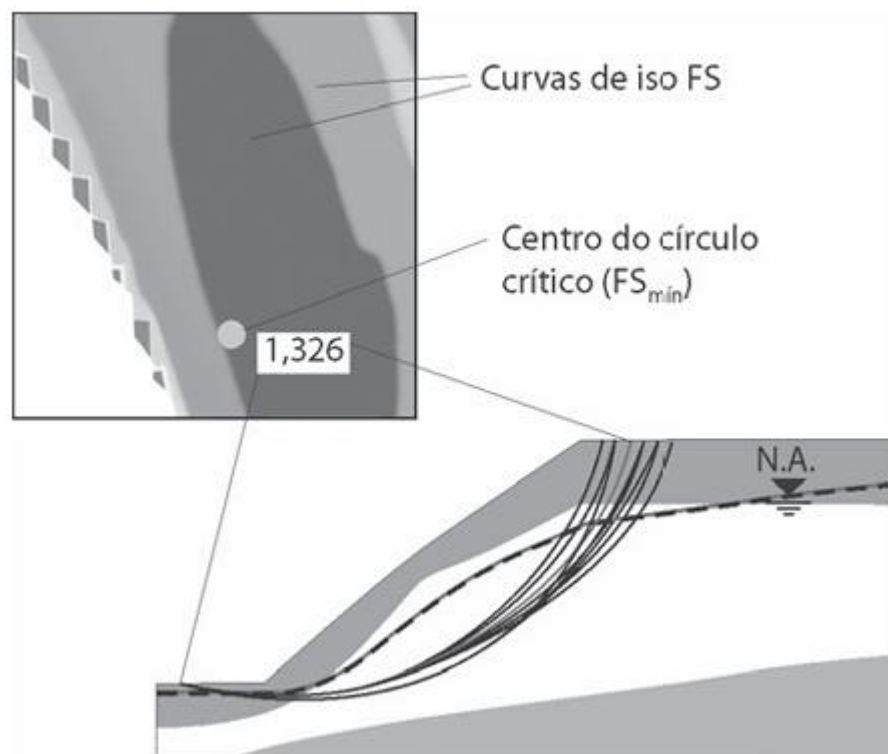


Fonte: Gresovich (2016)

Na aplicação dos métodos de equilíbrio limite, um aspecto essencial é a definição da superfície potencial de ruptura, ou seja, a trajetória ao longo da qual se admite que ocorrerá o escorregamento da massa instável. No entanto, como a geometria real da ruptura não é previamente conhecida, torna-se necessário avaliar diversas superfícies possíveis e identificar aquela que apresenta o menor fator de segurança, que é denominada superfície crítica de ruptura.

A superfície crítica é a configuração de ruptura mais provável, pois corresponde àquela em que o fator de segurança assume o valor mínimo dentre todas as alternativas analisadas. Essa busca é feita, geralmente, por meio de processos iterativos, avaliando-se múltiplas superfícies candidatas e comparando-se os fatores de segurança associados a cada uma delas (GRESOVICH, 2016). A Figura 4 ilustra esse procedimento, apresentando diferentes superfícies potenciais de ruptura e as curvas de iso-fator de segurança (iso FS), que indicam regiões com valores semelhantes de FS.

Figura 4: Procura da superfície crítica



Fonte: Gresovich (2016)

Tradicionalmente, os softwares de estabilidade de taludes automatizam esse processo, permitindo a análise de centenas ou milhares de superfícies potenciais, com diferentes formas e localizações, e fornecendo como resultado a superfície crítica e as superfícies com fatores de segurança próximos do mínimo admissível.

Diversos métodos de equilíbrio limite foram desenvolvidos ao longo do tempo para avaliar a estabilidade de taludes, diferenciando-se principalmente pela forma como consideram as forças atuantes entre fatias e os critérios de equilíbrio adotados. Os métodos mais consagrados na prática geotécnica são os de Fellenius (1936), Bishop Simplificado (1955), Janbu (1954), Spencer (1967) e Morgenstern-Price (1965).

Nos métodos de equilíbrio limite, o fator de segurança pode ser obtido satisfazendo apenas o equilíbrio de momentos (Bishop), apenas o equilíbrio de forças (Janbu), ou ambos simultaneamente, como no método de Morgenstern-Price. A principal diferença entre esses métodos está no tratamento das forças entre fatias, cuja consideração torna o sistema matematicamente indeterminado, exigindo a introdução de um parâmetro de escala (λ) para relacionar as forças normais e tangenciais entre fatias (ALOK *et al.*, 2024).

2.1.3 Métodos de Equilíbrio Limite

O método de Fellenius, também conhecido como método da fatia ordinária, é o mais simples entre os métodos de equilíbrio limite. Considera o equilíbrio de momentos em relação ao centro de uma superfície circular de ruptura, desconsiderando as forças de interação entre as fatias. Por essa razão, tende a fornecer resultados conservadores, ou seja, fatores de segurança mais baixos que os métodos mais completos (GERSCOVICH, 2016).

O método de Bishop Simplificado aprimorou o de Fellenius ao incluir o efeito das componentes normais das forças entre fatias no equilíbrio de momentos, embora ainda desconsidere as componentes horizontais dessas forças. Esse método é aplicável a superfícies circulares de ruptura e oferece boa precisão para taludes homogêneos e não sujeitos a grandes carregamentos externos. A expressão para o fator de segurança é obtida por iteração, pois depende do próprio valor de FS na resolução da equação de equilíbrio (GERSCOVICH, 2016).

O método de Janbu é mais versátil, pois permite a análise de superfícies não circulares, sendo baseado no equilíbrio de forças horizontais e verticais para cada fatia. Assim como Bishop, desconsidera os momentos entre fatias e as

componentes horizontais das forças de interação, adotando fatores corretivos para compensar essas omissões. Janbu propôs duas versões: a simplificada, mais conservadora e prática, e a generalizada, mais precisa, que utiliza fatores de correção para diferentes geometrias de ruptura (GERSCOVICH, 2016).

O método de Spencer é um método rigoroso, que considera simultaneamente o equilíbrio de forças e momentos em todas as fatias, assumindo uma relação constante entre as forças de interação horizontal entre as fatias. Essa característica permite a análise de superfícies circulares e não circulares com alto grau de precisão, sendo considerado um método exato dentro das hipóteses de equilíbrio limite. Por conta dessa completude, Spencer é amplamente empregado como referência para validar outros métodos (GERSCOVICH, 2016).

O método de Morgenstern-Price generaliza o método de Spencer ao permitir que a inclinação das forças de interação entre fatias varie de acordo com uma função arbitrária definida previamente, possibilitando ajustar o modelo às características específicas do talude. Assim como Spencer, considera o equilíbrio de forças e momentos em todas as fatias e é aplicável a superfícies de ruptura circulares e não circulares. Essa flexibilidade torna o método de Morgenstern-Price um dos mais completos e robustos disponíveis na engenharia geotécnica (GERSCOVICH, 2016).

Tabela 2: Comparação entre os diferentes Métodos de Equilíbrio Limite

Método	Complexidade	Forças Interslice	Equilíbrio de Forças	Equilíbrio de Momentos	Superfície
Fellenius	Baixa	Não	Vertical	Sim (sobre centro)	Circular
Bishop Simplificado	Média	Não	Vertical	Sim (sobre centro)	Circular
Janbu	Média	Sim (horizontais)	Vertical e Horizontal	Não	Circular e não circular
Spencer	Alta	Sim (com razão constante)	Vertical e Horizontal	Sim	Circular e não circular
Morgenstern-Price	Muito alta	Sim (com função definida)	Vertical e Horizontal	Sim	Circular e não circular

2.2 Inteligência Artificial e Aprendizado de Máquina

2.2.1 Histórico da Inteligência Artificial

A Inteligência Artificial (IA) é um termo geral que se refere à capacidade de uma máquina ou computador de executar atividades antes exclusivas à cognição humana, como: aprender, estabelecer padrões, correlacionar, tomar decisões, entre outras. Essa ideia surgiu quase simultaneamente à própria computação, com Alan Turing, matemático britânico e pioneiro nesse ramo da ciência, que em 1950 publicou o artigo “Computing Machinery and Intelligence” no periódico *Mind*, no qual propôs o famoso Teste de Turing — um método para avaliar se uma máquina poderia exibir comportamento inteligente indistinguível do humano (SMITH *et al.*, 2006).

O surgimento oficial da IA se deu em 1956, durante o Dartmouth Workshop, organizado por cientistas como John McCarthy, Marvin Minsky e Claude Shannon. Nesse evento, o termo “artificial intelligence” foi formalmente cunhado por McCarthy. Nas décadas seguintes, pesquisadores exploraram técnicas como a IA simbólica, baseada em regras lógicas para imitar o raciocínio humano, e as primeiras redes neurais artificiais, inspiradas no funcionamento do cérebro, mas limitadas pela capacidade de processamento computacional da época (SMITH *et al.*, 2006).

Devido à limitação tecnológica dos computadores e ao excesso de otimismo das previsões iniciais, a inteligência artificial entrou em um período de estagnação, conhecido como “Inverno da IA” (AI Winter), que ocorreu de forma mais aguda entre as décadas de 1970 a 1990, caracterizado pela queda de interesse e cortes de financiamentos governamentais e privados. Os relatórios ALPAC (1966) e Lighthill (1973) foram determinantes para esse cenário ao questionarem os avanços concretos e os custos excessivos da pesquisa em IA na época (SMITH *et al.*, 2006).

A partir do século XXI, com o surgimento do Big Data, da internet e da computação em nuvem, além do avanço exponencial do poder computacional, a IA teve seu ressurgimento. Atualmente, ela está presente em inúmeras aplicações cotidianas, como redes sociais, mecanismos de busca, além de áreas críticas como medicina, segurança, mobilidade urbana, engenharia e ciências ambientais.

2.2.2 Aprendizado de Máquina (*Machine Learning*)

O aprendizado de máquina (*machine learning*) pode ser entendido como a capacidade de um sistema computacional melhorar seu desempenho em uma

determinada tarefa por meio da experiência. Em termos formais, considera-se que um programa aprende a partir de uma experiência E, com respeito a uma classe de tarefas T e uma medida de desempenho P, se o seu desempenho em T, avaliado por P, melhora com a experiência (ALPAYDIN, 2004)

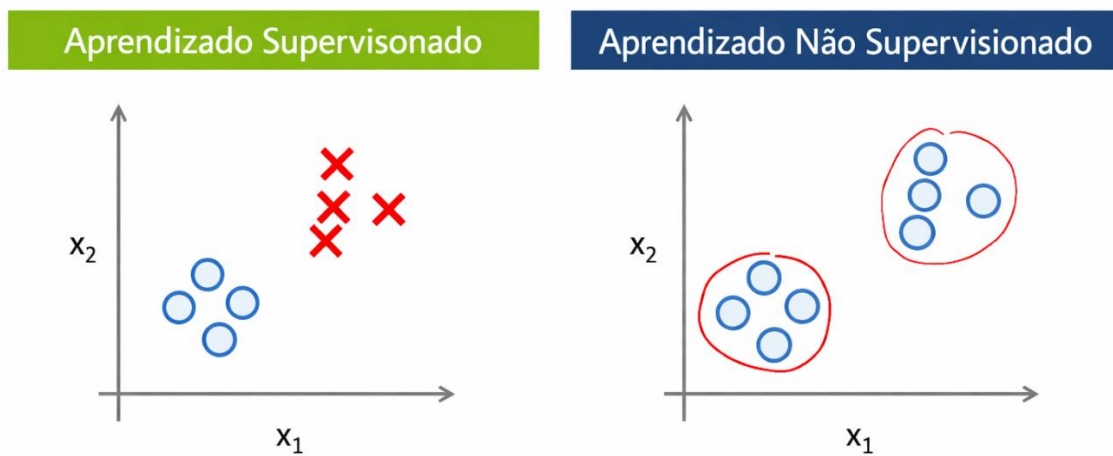
Dentro desse domínio, é possível classificar os métodos de aprendizado de máquina em diferentes categorias, a depender da natureza dos dados disponíveis e do objetivo do modelo, conforme. O aprendizado supervisionado é uma das abordagens mais utilizadas, caracterizando-se pela existência de um conjunto de dados rotulados, em que cada exemplo de entrada está associado a uma saída desejada. Dessa forma, o algoritmo aprende uma função que relaciona as variáveis independentes às variáveis dependentes, visando prever a saída correta para novas instâncias. (ALPAYDIN, 2004)

Em contraposição, no aprendizado não supervisionado, o conjunto de dados não apresenta rótulos ou categorias pré-definidas. O objetivo, nesse caso, é descobrir padrões, estruturas ou regularidades implícitas nos dados, por meio de métodos como agrupamento (*clustering*) e redução de dimensionalidade. Essa abordagem é especialmente relevante para a exploração de dados, segmentação de públicos e identificação de comportamentos anômalos (ALPAYDIN, 2004).

A Figura 5 ilustra essa distinção: no aprendizado supervisionado, os dados encontram-se previamente rotulados, permitindo a separação explícita entre classes, enquanto no aprendizado não supervisionado, os dados são agrupados com base em similaridades, formando *clusters* identificados pelo próprio algoritmo.

No contexto da engenharia geotécnica, essa característica é particularmente crítica, uma vez que os modelos precisam representar adequadamente comportamentos complexos do solo a partir de conjuntos de dados limitados. Assim, a escolha de algoritmos supervisionados deve considerar não apenas sua capacidade de ajuste aos dados históricos, mas principalmente sua robustez frente a novas condições de carregamento e variabilidade dos parâmetros do solo, o que impacta diretamente a confiabilidade da previsão do fator de segurança de taludes.

Figura 5: Exemplos de abordagens de aprendizado de máquina.



Fonte: Ernesto *et al.*(2018)

Os algoritmos de aprendizado de máquina também podem ser divididos em modelos de classificação e regressão, de acordo com a natureza da variável de resposta.

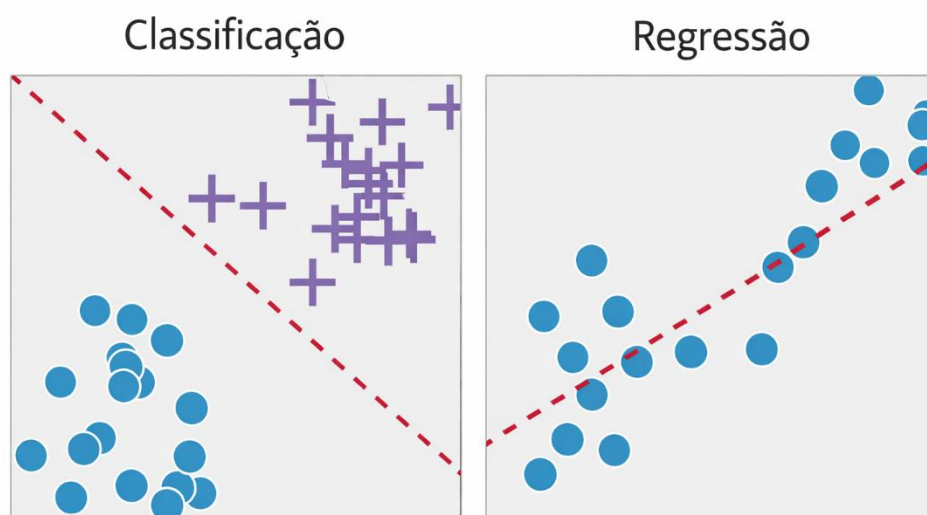
A classificação é um tipo de aprendizado de máquina no qual o objetivo consiste em atribuir cada exemplo de entrada a uma categoria previamente definida, através de uma função que separe as categorias de maneira satisfatória. Essas categorias são finitas e mutuamente exclusivas na maioria das aplicações, embora existam situações em que um mesmo exemplo possa pertencer a mais de uma categoria (MURPHY, 2022).

Já os modelos de regressão se diferem os modelos de classificação, em que o objetivo é escolher uma categoria, e buscam, por sua vez, estimar um valor numérico contínuo que represente adequadamente o fenômeno estudado. Nesse caso, o modelo aprende uma função que relaciona as entradas a um valor quantitativo. (MURPHY, 2022).

Em resumo, a escolha entre essas abordagens no estudo de taludes define a utilidade prática do modelo. Enquanto a classificação permite uma triagem rápida entre taludes 'estáveis' ou 'instáveis', a regressão, ao estimar valores contínuos como o Fator de Segurança (FS), oferece uma análise de sensibilidade superior. De acordo com os princípios de Alpaydin (2004), a regressão permite quantificar a margem de incerteza, sendo mais adequada para tomadas de decisão de engenharia onde o grau de proximidade do estado crítico é tão relevante quanto o

estado em si. A Figura 6 ilustra um problema de classificação, em que uma linha divisória separa dois grupos de pontos pertencentes a classes distintas e um problema de regressão, no qual uma reta ajustada aos dados representa a estimativa de uma variável contínua.

Figura 6: Separação de classes por fronteira de decisão (classificação) e ajuste de uma função para prever valores contínuos (regressão).



Fonte: Edell (2015)

Segundo Alpaydin (2004), o propósito fundamental do aprendizado de máquina não é reproduzir exatamente os dados utilizados no treinamento, mas sim produzir previsões corretas para novas instâncias não observadas, característica conhecida como capacidade de generalização. A generalização expressa o quão bem um modelo treinado consegue inferir resultados corretos para entradas externas ao conjunto de treinamento, sendo, portanto, um critério central para a avaliação do desempenho de modelos de aprendizado (ALPAYDIN, 2004).

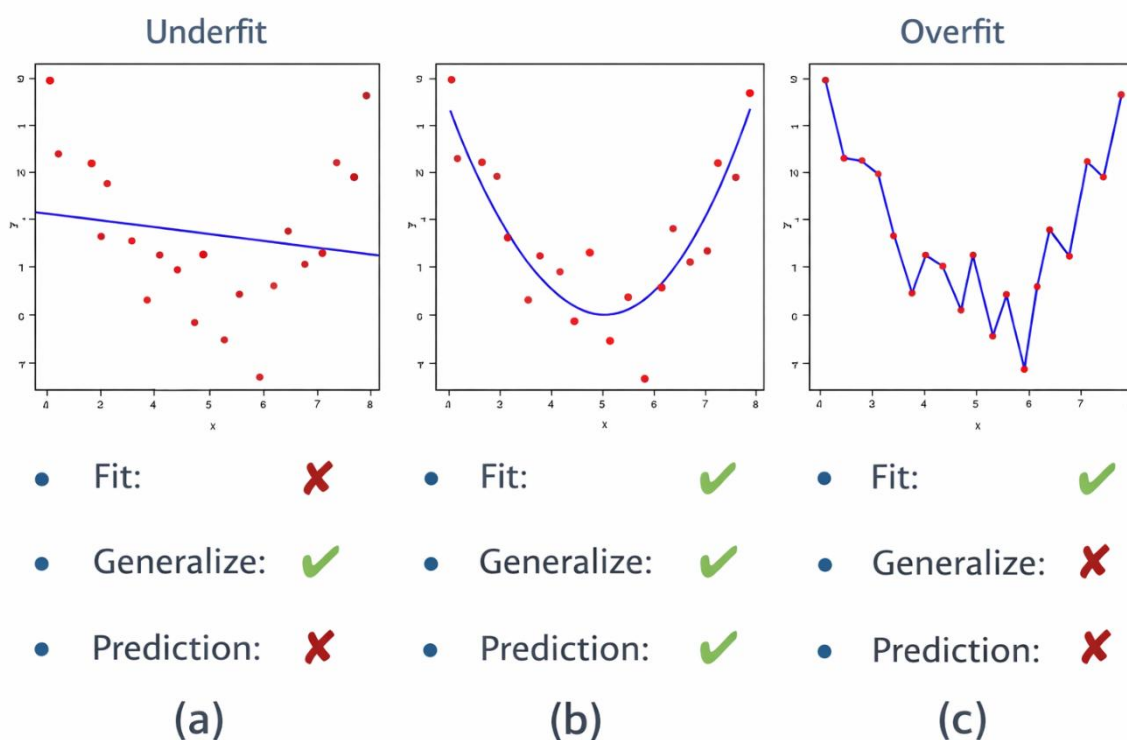
Para que se obtenha boa generalização, é necessário estabelecer um equilíbrio entre a complexidade do modelo e a complexidade da função subjacente aos dados utilizados no treinamento. Quando a hipótese adotada é excessivamente simples em relação ao fenômeno real, o modelo é incapaz de capturar os padrões fundamentais dos dados, caracterizando o fenômeno conhecido como underfitting (ALPAYDIN, 2004).

Por outro lado, quando a hipótese é muito complexa, o modelo pode se

ajustar não apenas à estrutura real dos dados, mas também aos ruídos, que são os valores destoantes do padrão do restante dos dados, resultando em *overfitting*. Nesse caso, o erro no conjunto de treinamento é reduzido, mas o erro de generalização aumenta, o que prejudica a previsão para novos dados (ALPAYDIN, 2004).

No caso de *underfitting*, o modelo é excessivamente simples e não consegue representar adequadamente o padrão dos dados, resultando em baixo ajuste e previsões imprecisas. Já no *overfitting*, o modelo é demasiadamente complexo, ajustando-se excessivamente aos dados de treinamento e apresentando baixa capacidade de generalização, conforme a Figura 7.

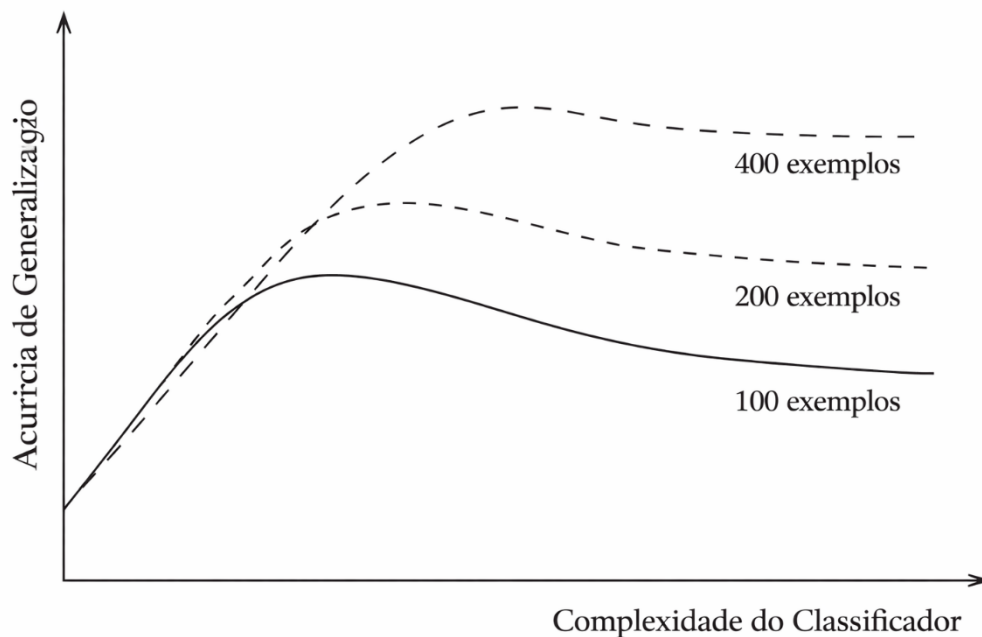
Figura 7: Comparação na generalização e ajustes de modelos. (a) Underfitting: o modelo é generalista demais e não captura o padrão dos dados; (b) Ajuste adequado: o modelo representa corretamente os dados e apresenta boa capacidade de generalização; (c) Overfitting: o modelo se ajusta excessivamente aos dados de treinamento, reduzindo a capacidade de generalização.



Fonte: Velazquez (2020)

A aplicação de modelos de ML em problemas geotécnicos enfrenta o desafio do *triple trade-off* proposto por Alpaydin (2004), apoiado no trabalho de Dietterich (2003),: a relação entre a complexidade do algoritmo, a quantidade de dados disponíveis e o erro de generalização (Figura 8). Em bancos de dados de encostas, que frequentemente possuem amostras limitadas, o aumento indiscriminado da complexidade do modelo não resulta necessariamente em maior precisão, mas sim em *overfitting*. Portanto, a seleção do modelo deve ser pautada visando esse equilíbrio, priorizando algoritmos que compensem a capacidade de capturar a não-linearidade dos solos sem comprometer a estabilidade preditiva em novos cenários.

Figura 8: Relação entre a complexidade do classificador e a acurácia de generalização para diferentes quantidades de dados de treinamento



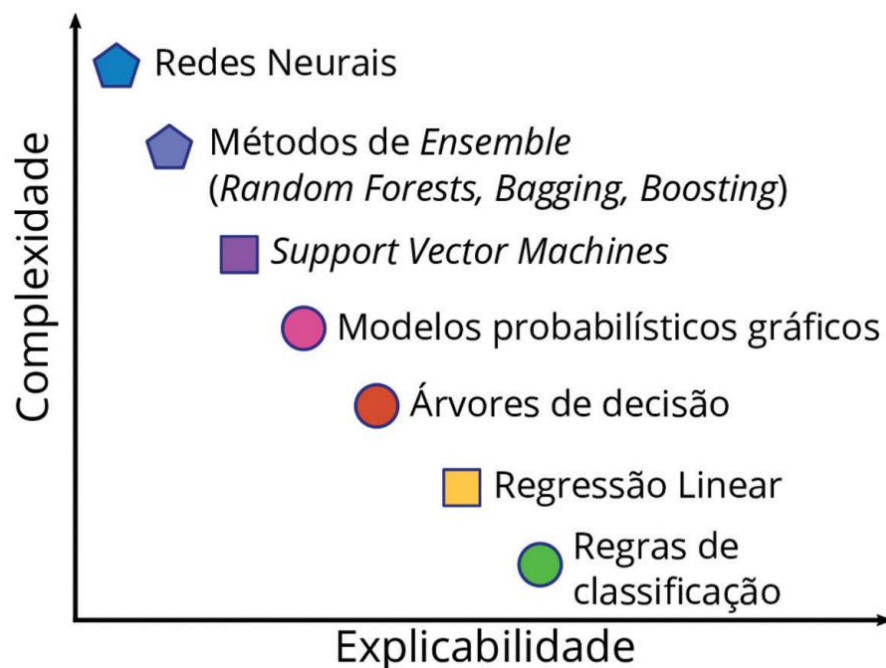
Fonte: Adaptado de Dietterich (2003).

2.2.3 Exemplos de modelos de aprendizado de máquina

Os diferentes tipos de modelos de aprendizado de máquina distinguem-se principalmente quanto à complexidade, capacidade de generalização e nível de interpretabilidade (Figura 9). Métodos mais simples tendem a oferecer maior

transparência na relação entre entradas e saídas, enquanto abordagens mais complexas apresentam maior flexibilidade para capturar interações não lineares e padrões sutis nos dados, frequentemente resultando em melhor desempenho preditivo (ALPADYN, 2004).

Figura 9 : Trade-off de explicabilidade e complexidade no aprendizado de máquina.



Fonte: Schleder e Fazzio (2021)

2.2.3.1 Árvore de Decisão

As Árvores de Decisão são algoritmos de aprendizado supervisionado amplamente empregados em problemas de classificação e regressão, nos quais o processo de tomada de decisão é estruturado de forma hierárquica, em uma representação em forma de árvore (Figura 10). Nessa estrutura, cada nó interno corresponde a um teste condicional aplicado a uma variável de entrada, os ramos representam os possíveis resultados desse teste e, ao final do processo, os nós folha armazenam a predição final do modelo. Uma das principais vantagens desse método reside em sua elevada interpretabilidade, uma vez que a lógica do modelo pode ser facilmente convertida em um conjunto de regras compreensíveis, mesmo por usuários não especialistas (ALPAYDIN, 2004).

Nesse sentido, a construção da árvore ocorre a partir do conjunto de dados

de treinamento por meio de um processo iterativo de divisão recursiva do espaço de atributos, conhecido como *recursive partitioning*. Nesse processo, o conjunto de dados é sucessivamente particionado em subconjuntos cada vez mais homogêneos em relação à variável de saída. (GUIDO; MÜLLER, 2017)

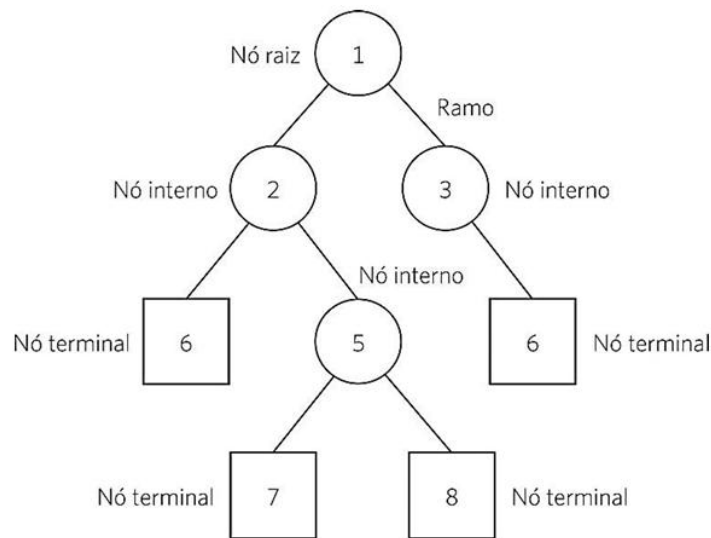
Um dos critérios mais utilizados para avaliar a qualidade das divisões é o índice Gini, que mede o grau de impureza de um nó. De acordo com Müller e Guido (2017), a busca por divisões que minimizem a impureza tende a gerar nós cada vez mais homogêneos, podendo resultar em folhas puras quando a árvore é totalmente desenvolvida.

Embora a minimização do índice Gini contribua para um excelente desempenho no conjunto de treinamento, Müller e Guido (2017) destacam que árvores excessivamente profundas tendem a memorizar os dados, tornando-se altamente sensíveis a ruídos e outliers, causando *overfitting*. Assim, aceitar níveis moderados de impureza nos nós pode levar a modelos mais simples e com melhor desempenho em dados não vistos.

Outro aspecto relevante na construção das Árvores de Decisão diz respeito ao uso de mecanismos de aleatoriedade. No contexto das implementações computacionais, o parâmetro “random_state” controla a inicialização do gerador de números aleatórios utilizado em decisões internas do algoritmo, como o desempate entre divisões com qualidade equivalente. (GUIDO; MÜLLER, 2017)

Müller e Guido (2017) ressaltam que, embora esse parâmetro não altere diretamente a complexidade do modelo, diferentes valores podem resultar em árvores estruturalmente distintas, com pequenas variações na precisão obtida. A fixação do “random_state” é, portanto, essencial para garantir a reprodutibilidade dos experimentos e permitir comparações consistentes entre modelos (GUIDO; MÜLLER, 2017).

Figura 10: Estrutura de uma Árvore de decisão



Fonte: Ernesto *et al.*(2018)

Segundo Sheppard (2017), esse método é particularmente eficiente para descobrir padrões ocultos em conjuntos de dados e estruturar o conhecimento adquirido de forma sequencial e hierárquica. Essa característica é especialmente vantajosa em aplicações de engenharia, nas quais a transparência das decisões é fundamental.

No entanto, uma das limitações das árvores de decisão é a tendência ao sobreajuste, especialmente quando muito profundas ou quando os dados de treinamento contêm ruídos e outliers. Assim, embora úteis para compreensão inicial dos dados e identificação de padrões, árvores de decisão isoladas podem não ser a melhor escolha para previsão precisa do fator de segurança em cenários geotécnicos complexos.

Para mitigar esse problema, técnicas de poda ou limitação da profundidade são aplicadas, visando simplificar a árvore e melhorar sua capacidade de generalização.

O processo de poda pode ocorrer de duas formas: pré-poda (*prepruning*) e pós-poda (*postpruning*). No *prepruning*, a construção da árvore é interrompida antecipadamente quando determinados critérios são atingidos, como o número mínimo de instâncias em um nó, prevenindo decisões baseadas em poucos dados e, conseqüentemente, reduzindo a variância. . (ALPAYDIN, 2004).

Já no *postpruning*, a árvore é inicialmente construída até atingir sua

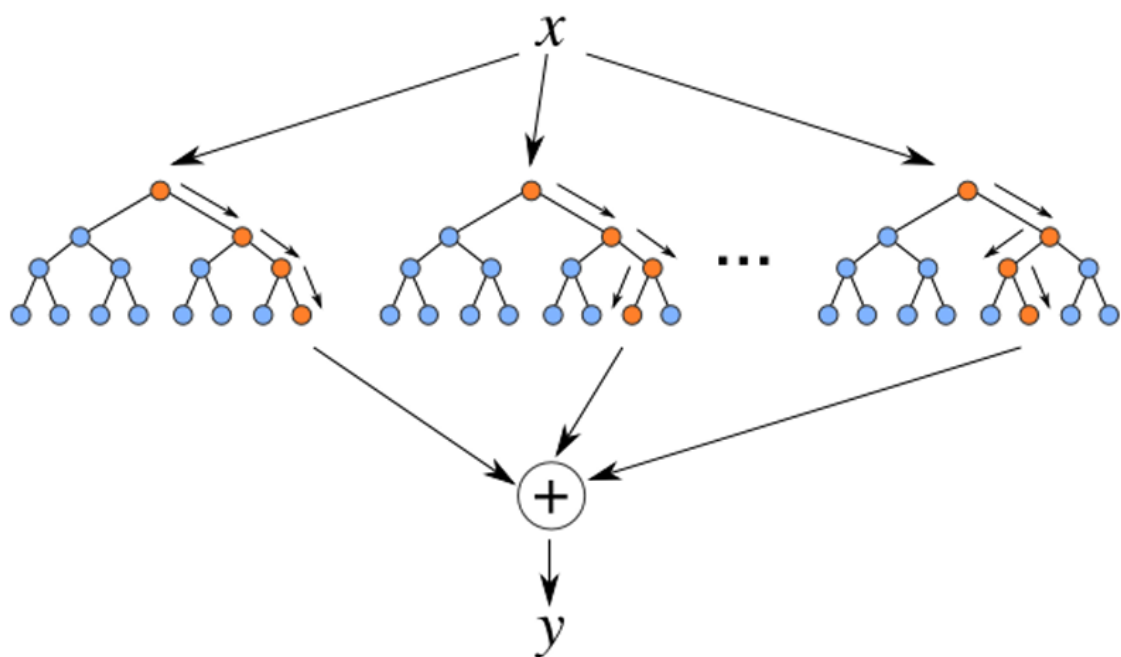
profundidade máxima, com todos os nós folha puros e erro de treinamento zero. Em seguida, utilizando-se um conjunto de dados separado, denominado *pruning set*, subárvores que não contribuem para a melhoria do desempenho são substituídas por nós folha, desde que essa simplificação não comprometa os resultados obtidos nesse conjunto de validação específico. Embora o *prepruning* seja mais rápido, o *postpruning* costuma gerar modelos mais precisos e com melhor capacidade de generalização. (ALPAYDIN, 2004).

2.2.3.2 Floresta Aleatória

A Floresta Aleatória é uma técnica de aprendizado supervisionado baseada na combinação de múltiplas árvores de decisão, conforme a estrutura representada na Figura 11. Cada árvore da floresta é construída a partir de uma amostra aleatória, com reposição, do conjunto de dados original e, em cada divisão de nó, uma seleção aleatória de atributos é considerada. (ALPAYDIN, 2004).

O algoritmo da Floresta Aleatória combina várias árvores de decisão construídas a partir de amostras aleatórias e subconjuntos de variáveis, buscando reduzir a correlação entre as árvores e aumentar a robustez do modelo (ALPAYDIN, 2004). Ao final, para realizar a predição, o algoritmo reúne as respostas de todas as árvores e aplica um processo de votação: no caso de classificação, vence a categoria mais votada entre as árvores; já na regressão, calcula-se a média das predições, obtendo assim um resultado final mais estável e preciso. Esse procedimento permite à floresta aleatória apresentar maior estabilidade, menor variância e melhor capacidade preditiva em comparação a árvores de decisão isoladas (Guido & Müller, 2016).

Figura 11: Estrutura de uma Floresta Aleatória.



Fonte: Ernesto *et al.*(2018)

De acordo com Sheppard (2017), o método se destaca por sua flexibilidade e aplicabilidade em problemas com dados mistos, alta dimensionalidade e conjuntos ruidosos. Contudo, essa abordagem apresenta como desvantagem o fato de as previsões resultarem da combinação de diversas árvores, tornando o modelo mais difícil de interpretar diretamente, diferentemente das árvores de decisão individuais.

Comparativamente, a floresta aleatória apresenta desempenho superior às árvores individuais em termos de generalização, porém com perda de interpretabilidade. No contexto da análise de estabilidade de taludes, essa troca é relevante: enquanto a precisão preditiva é essencial para segurança, a menor transparência do modelo pode dificultar a interpretação dos fatores que influenciam o comportamento do solo. Dessa forma, sua utilização deve equilibrar desempenho e capacidade de explicação dos resultados

No contexto da implementação computacional, o parâmetro “random_state” desempenha um papel fundamental ao controlar a inicialização do gerador de números aleatórios. Devido à natureza intrinsecamente aleatória da Floresta Aleatória, diferentes valores de “random_state” podem resultar em florestas estruturalmente distintas, com variações pontuais na precisão obtida, especialmente quando o número de árvores é reduzido. Assim, a fixação desse

parâmetro é essencial para garantir a reprodutibilidade dos resultados e permitir comparações consistentes entre experimentos (GUIDO; MÜLLER, 2017).

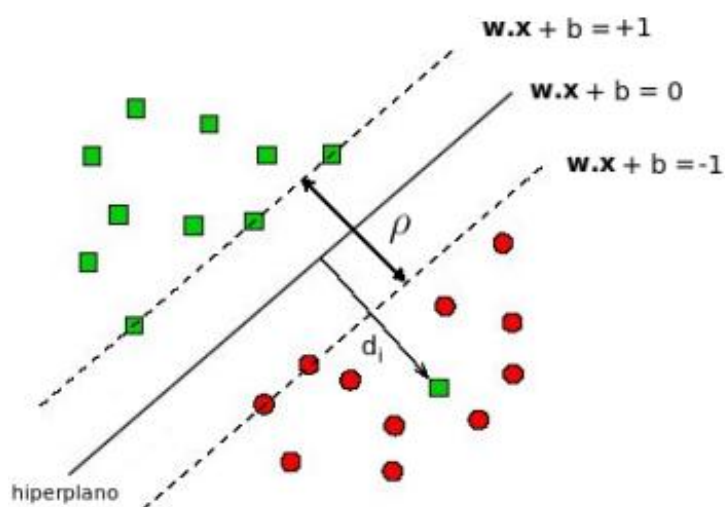
Müller e Guido (2017) destacam que, à medida que o número de árvores na floresta aumenta, o modelo torna-se progressivamente menos sensível à escolha do “random_state”, uma vez que o efeito da aleatoriedade é diluído pela agregação de um grande número de estimadores. Dessa forma, o uso de um número elevado de árvores contribui simultaneamente para a redução da variância e para o aumento da capacidade de generalização do modelo.

2.2.3.3 Máquinas de Vetores Suporte

As Máquinas de Vetores de Suporte (Support Vector Machines — SVMs) são algoritmos supervisionados voltados principalmente para tarefas de classificação, embora também possam ser aplicadas à regressão. Seu princípio de funcionamento consiste em identificar, dentre os diversos hiperplanos que podem separar as classes no espaço de atributos de um conjunto de dados, aquele que oferece a maior margem possível entre os exemplos mais próximos de cada classe, denominados vetores de suporte (CRISTIANINI; SHAWE-TAYLOR, 2000).

A Figura 12 apresenta o hiperplano de separação ($w \cdot x + b = 0$) e os hiperplanos marginais ($w \cdot x + b = \pm 1$), evidenciando a margem máxima entre as classes e destacando os vetores de suporte, que determinam a posição ótima do hiperplano.

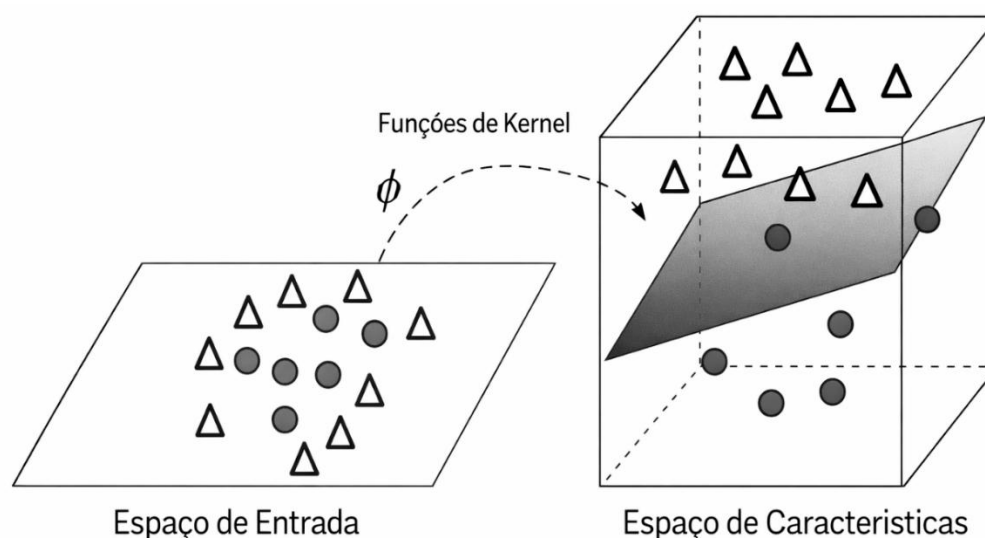
Figura 12: Distância entre hiperplanos e vetores suporte.



Fonte: Lorena e Carvalho (2007)

Esse conceito de maximização da margem confere ao modelo maior capacidade de generalização, uma vez que busca não apenas separar os dados de treinamento, mas fazê-lo com a maior distância possível entre as classes, reduzindo o risco de erro em novos exemplos. Quando o algoritmo, dentro de um hiperplano, não consegue encontrar um vetor suporte que separe os dados de maneira linear, as SVMs utilizam funções kernel para projetá-los em um espaço de maior dimensionalidade (Figura 13), onde a separação linear se torna viável, recurso conhecido como truque de kernel (ALPAYDIN, 2004).

Figura 13: Transformação do plano utilizando o truque de kernel



Fonte: Mezher (2022)

As máquinas de vetores de suporte são particularmente eficazes em conjuntos de dados de alta dimensionalidade e em situações em que as classes são separadas por margens bem definidas. Entretanto, sua aplicação pode se tornar computacionalmente custosa em bases muito extensas e requer a definição adequada de parâmetros como o tipo de kernel e a penalidade por erro, o que demanda processos de validação cuidadosos para ajuste e otimização do modelo.

Além disso, em comparação com modelos baseados em árvores, as SVMs tendem a ser menos interpretáveis, embora possam apresentar melhor desempenho em conjuntos de dados bem estruturados. Para a previsão do fator de segurança, sua eficácia está condicionada à qualidade da calibração do modelo e à representatividade dos dados geotécnicos disponíveis.

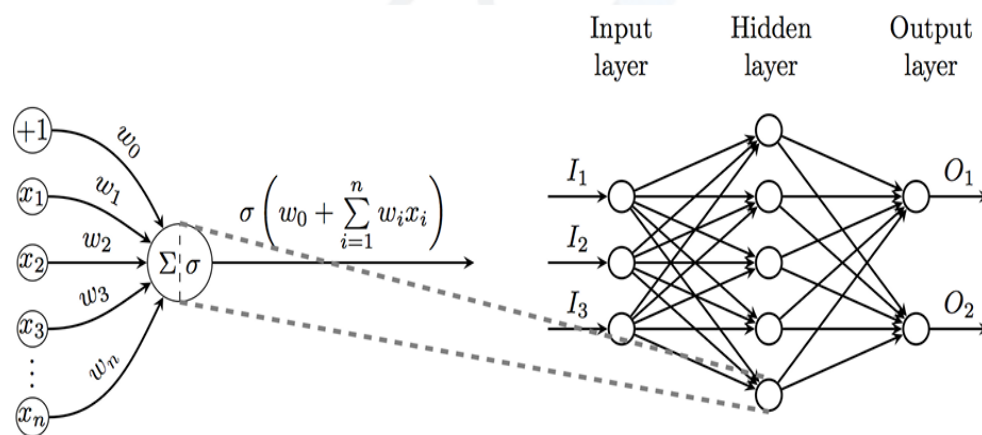
2.2.3.4 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são um modelo de aprendizado de máquina inspirados no funcionamento dos neurônios biológicos do cérebro, compostos por unidades de processamento interligadas, conhecidas como neurônios artificiais, organizadas em camadas. A estrutura de uma Rede Neural é dividida em camadas de entrada, camadas ocultas e camada de saída, conforme a Figura 14. Os dados são recebidos pela camada de entrada, enquanto uma ou mais camadas ocultas processam informações intermediárias e capturaram padrões complexos, e a camada de saída devolve o valor ou classificação previsto pelo modelo (Ernesto et al., 2018).

Cada conexão entre os neurônios possui um peso associado, que determina a influência daquele sinal na resposta final da rede. Os dados fornecidos na camada de entrada são transmitidos às camadas ocultas por meio dessas conexões ponderadas. Ao chegar a cada neurônio, os sinais são somados e processados por uma função de ativação, que define se e quanto daquele sinal será transmitido adiante. Essa função é essencial para introduzir não linearidade no modelo, permitindo que a rede aprenda e represente relações complexas entre as variáveis de entrada e saída (ALPAYDIN, 2004).

Durante o treinamento, a rede ajusta progressivamente esses pesos, de modo a minimizar a diferença entre as respostas produzidas pelo modelo e os valores esperados nos dados de treinamento. Para isso, calcula-se uma função de custo, que expressa a discrepância entre as previsões e os valores reais. Esse erro é então propagado de volta pelas conexões da rede, permitindo que os pesos sejam atualizados de acordo com sua contribuição no erro cometido, em um processo conhecido como retropropagação do erro. A cada iteração, os pesos são ajustados utilizando métodos de otimização, sendo o gradiente descendente um dos mais empregados para esse fim (GUIDO; MÜLLER, 2016).

Figura 14: Representação esquemática de uma Rede Neural de Múltiplas Camadas



Fonte: Ernesto *et al.*(2018)

As redes neurais destacam-se pela capacidade de modelar relações altamente não lineares e aprender representações complexas dos dados, o que as torna particularmente eficazes em tarefas de reconhecimento de padrões, análise de imagens, séries temporais e problemas preditivos de alta complexidade.

No entanto, esse desempenho elevado exige, em contrapartida, grandes volumes de dados para treinamento eficiente, além de recursos computacionais consideráveis. Outro ponto de atenção é o fato de que, apesar de sua flexibilidade, as redes neurais são consideradas modelos de difícil interpretação, pois as relações internas entre as variáveis e os pesos ajustados não são diretamente acessíveis ou explicáveis de maneira intuitiva (ALPAYDIN, 2004).

2.2.4 Aprendizado em conjunto (*Ensemble Learning*)

O aprendizado em conjunto (ensemble learning) é um paradigma do aprendizado de máquina que se baseia na combinação de múltiplos modelos preditivos simples, denominados modelos fracos (*weak learners*), com o objetivo de obter um desempenho superior ao de um único modelo altamente complexo. Em vez de buscar um modelo individual extremamente preciso, essa abordagem explora o fato de que um conjunto de modelos com baixa acurácia, quando adequadamente combinados, pode resultar em um meta-modelo com elevada capacidade preditiva (Burkov, 2020).

O métodos de *ensemble learning* podem ser classificados principalmente em duas grandes categorias: *bagging* e *boosting* (Figura 16).

Os métodos baseados em *bagging* (*bootstrap aggregating*) treinam seus modelos base de forma independente, utilizando diferentes subconjuntos do conjunto de dados. O *bagging* utiliza amostragem *bootstrap* com reposição para gerar múltiplos subconjuntos de treinamento, treinando um modelo em cada subconjunto e combinando suas previsões por média ou votação. (González *et al.*, 2020).

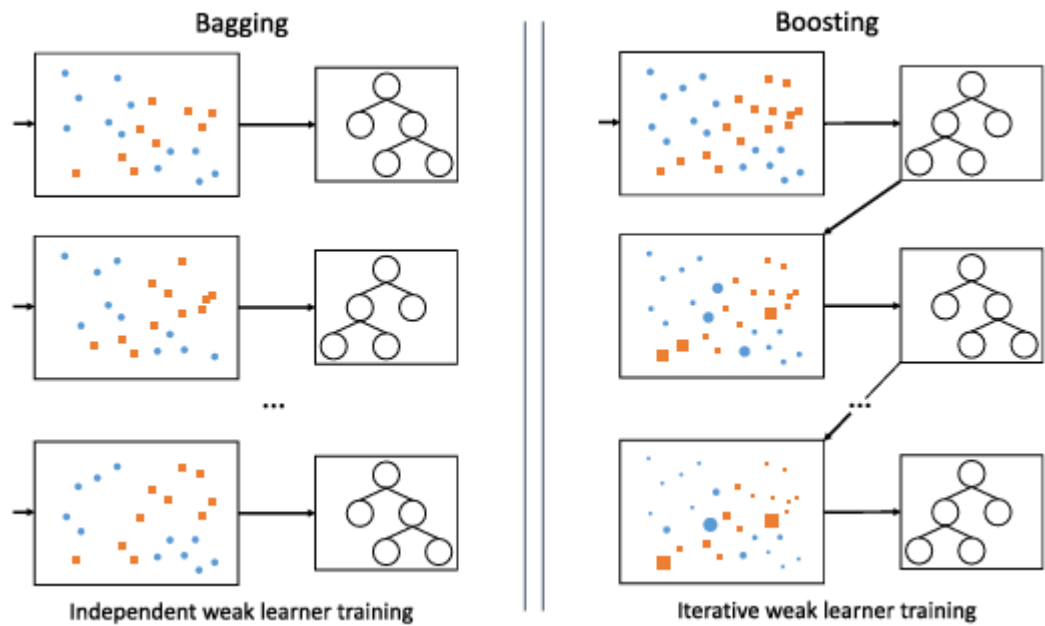
Segundo Alpaydn (2004), o *bagging* é particularmente eficaz quando aplicado a algoritmos instáveis, isto é, algoritmos cuja estrutura do modelo varia significativamente em resposta a pequenas mudanças no conjunto de treinamento, característica comum em modelos como árvores de decisão e redes neurais multicamadas. Portanto, essa é uma estratégia especialmente eficaz na redução da variância

Já os métodos de *boosting* adotam um treinamento sequencial e dependente entre os modelos. Cada novo modelo é ajustado com maior foco nos erros cometidos pelos modelos anteriores, atribuindo maior importância às amostras mal classificadas, buscando reduzir o viés do modelo e melhorar progressivamente a precisão (González *et al.*, 2020).

Alpaydin (2004) destaca que a principal diferença do *boosting* em relação ao *bagging* é que, em vez de depender apenas da aleatoriedade da amostragem, o primeiro busca ativamente construir aprendizes complementares. O conceito central é o de aprendiz fraco (*weak learner*), definido como um classificador com desempenho ligeiramente melhor que o acaso, e aprendiz forte (*strong learner*), que apresenta erro arbitrariamente pequeno.

Entre os principais métodos de *boosting* encontram-se o *AdaBoost*, um dos algoritmos mais clássicos dessa categoria, o *Gradient Boosting*, que utiliza a minimização iterativa de uma função de perda, e suas extensões modernas, como o *XGBoost* e o *LightGBM*, que incorporam técnicas de regularização e otimização computacional (González *et al.*, 2020).

Figura 15: Representação das estratégias de Bagging e Boosting.



Fonte: González *et al.* (2020)

2.2.5 Padronização e Normalização de dados

No contexto do aprendizado de máquina, o pré-processamento dos dados exerce papel fundamental no desempenho e na estabilidade dos modelos, especialmente quando as variáveis de entrada apresentam escalas distintas. Segundo Burkov (2019), muitas técnicas de aprendizado operam sobre vetores numéricos e podem ser influenciadas negativamente quando os atributos possuem ordens de grandeza muito diferentes, tornando necessária a aplicação de procedimentos de reescalonamento.

A normalização consiste em transformar os valores de um atributo numérico para um intervalo padrão, geralmente entre zero e um ou entre menos um e um. Esse procedimento preserva a distribuição relativa dos dados, mas garante que todos os atributos estejam contidos em uma faixa comparável. Burkov (2019) destaca que a normalização pode acelerar o processo de aprendizado, especialmente em algoritmos baseados em métodos iterativos de otimização, ao evitar que atributos com valores elevados dominem as atualizações dos parâmetros do modelo. Além disso, a normalização contribui para reduzir problemas numéricos associados ao uso de valores muito grandes ou muito pequenos.

A padronização, também conhecida como normalização do tipo *z-score*, é

outro método amplamente utilizado, no qual os valores dos atributos são transformados de modo a apresentar média igual a zero e desvio padrão igual a um. De acordo com Burkov (2019), a padronização é particularmente vantajosa quando os dados apresentam distribuição aproximadamente normal ou quando existem valores extremos, pois evita a compressão excessiva dos dados que pode ocorrer na normalização.

O pré-processamento de dados, especificamente a normalização, transcende a mera eficiência computacional mencionada na literatura técnica. Em problemas de estabilidade de taludes, variáveis como coesão (kPa) e ângulo de atrito possuem ordens de magnitude e unidades distintas. Sem a padronização, algoritmos baseados em gradiente ou distância podem conferir um peso desproporcional a variáveis com valores numéricos maiores, ignorando parâmetros geomecânicos cruciais. Assim, a normalização atua como uma ferramenta de equidade estatística, garantindo que cada característica do solo contribua proporcionalmente para a convergência do modelo.

2.2.6 Validação cruzada

A validação cruzada (*cross-validation*) é uma técnica mais comum e amplamente utilizada para estimar o erro de generalização de modelos de aprendizado de máquina, que estima diretamente o erro esperado em dados não utilizados no treinamento (Hastie, Tibshirani & Friedman, 2009)

Entre as diferentes abordagens existentes, destaca-se a validação cruzada K-fold, na qual o conjunto de dados é dividido em k subconjuntos aproximadamente iguais. Conforme descrito por Hastie *et al.* (2009), os dados são divididos em K partes aproximadamente do mesmo tamanho e, para cada iteração, o modelo é ajustado utilizando K-1 partes dos dados e o erro de predição é calculado na parte restante. Esse processo é repetido até que todos os subconjuntos tenham sido utilizados uma vez como conjunto de validação, sendo o desempenho final (*CV-score*) obtido pela média das métricas ao longo das iterações.

A escolha do número de partições envolve um compromisso entre viés e variância da estimativa do erro. De acordo com Hastie *et al.* (2009), embora o *cross-validation* apresente baixo viés, ele pode resultar em alta variância. Por esse motivo, os autores afirmam que valores típicos de K são 5 ou 10 e que a validação cruzada com cinco ou dez partições é recomendada como um bom compromisso.

2.3 Métodos de classificação e avaliação dos modelos

Para avaliar o desempenho de modelos preditivos e classificatórios, alguns métodos e métricas são essenciais para interpretar a eficácia das previsões em relação aos valores reais. Ernesto *et al.* (2018), destacam os seguintes:

- **Coeficiente de Determinação (R^2):**

Mede a proporção da variabilidade dos dados que é explicada pelo modelo. Ele compara a soma dos quadrados dos erros das previsões (diferença entre valores reais e previstos) com a soma dos quadrados das diferenças entre os valores reais e sua média, ou seja, a variabilidade total dos dados. Quanto mais próximo de 1 estiver o valor de R^2 , melhor será o desempenho do modelo (Ernesto *et al.*, 2018).

$$R^2 = 1 - \frac{\sum(y_i - \bar{y})^2}{\sum(y_i - \bar{y}^i)^2} \quad \text{Eq. (4)}$$

- **Erro Quadrático Médio (EQM):**

Representa a média dos quadrados das diferenças entre os valores previstos e os valores reais. Ele é obtido elevando ao quadrado a diferença entre cada valor real e o valor previsto, somando todos esses quadrados e dividindo pelo número total de amostras avaliadas. O EQM é uma métrica que penaliza mais fortemente erros maiores, sendo útil para identificar grandes discrepâncias nas previsões (Ernesto *et al.*, 2018).

$$EQM = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}|^2 \quad \text{Eq. (5)}$$

- **Acurácia:**

Refere-se à proporção total de previsões corretas realizadas pelo modelo em relação ao total de casos avaliados (Ernesto *et al.*, 2018).

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Eq. (1)}$$

Onde:

- TP (True Positive): Casos positivos corretamente previstos.
- TN (True Negative): Casos negativos corretamente previstos.

- FP (False Positive): Casos negativos erroneamente previstos como positivos.
- FN (False Negative): Casos positivos erroneamente previstos como negativos.

- **Matriz de confusão:**

A matriz de confusão (Figura 16) é uma tabela utilizada para avaliar o desempenho de um modelo de classificação, comparando as classes previstas pelo modelo com as classes reais dos dados. Em problemas de classificação binária, ela é composta por quatro elementos principais: verdadeiros positivos (VP), quando o modelo prevê corretamente a classe positiva; verdadeiros negativos (VN), quando prevê corretamente a classe negativa; falsos positivos (FP), quando o modelo prevê positivo para um caso que é negativo; e falsos negativos (FN), quando o resultado é negativo para um caso que é positivo. Esses elementos permitem analisar os acertos e erros do modelo e calcular métricas de desempenho (Ernesto et al., 2018).

Figura 16: Matriz de confusão para avaliação de modelos de classificação.

		Valor Verdadeiro (confirmado por análise)	
		positivos	negativos
Valor Previsto (predito pelo teste)	positivos	VP Verdadeiro Positivo	FP Falso Positivo
	negativos	FN Falso Negativo	VN Verdadeiro Negativo

Fonte: Ernesto et al.(2018)

- **Curva ROC:**

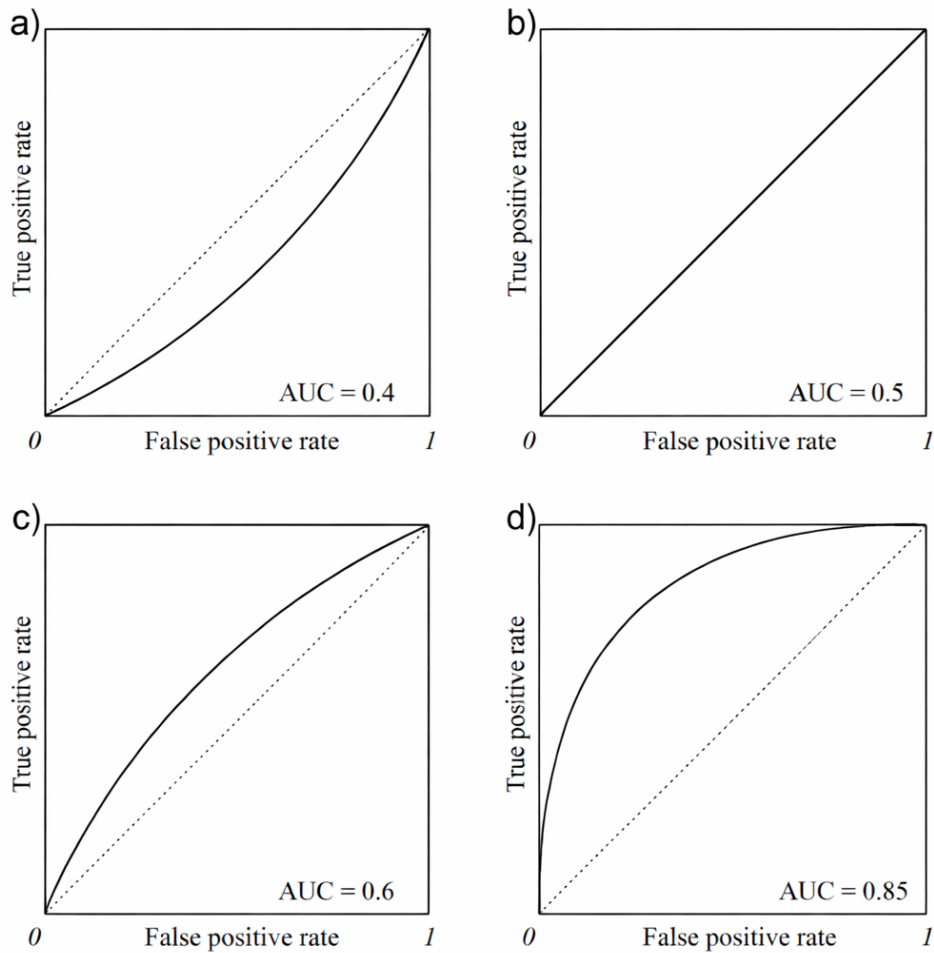
A curva ROC (*Receiver Operating Characteristic*) é uma ferramenta gráfica

utilizada para avaliar o desempenho de modelos de classificação binária que produzem um valor contínuo de decisão ou um escore de confiança. Essa curva é construída a partir da relação entre a taxa de verdadeiros positivos e a taxa de falsos positivos, calculadas para diferentes limiares de decisão (Burkov, 2020).

A área sob a curva ROC, denominada AUC (*Area Under the Curve*), fornece uma medida global da capacidade discriminatória do modelo, indicando o quão bem ele consegue distinguir entre as classes positiva e negativa. Conforme a Figura 17, valores de AUC próximos de 1 indicam excelente desempenho, enquanto valores próximos de 0,5 correspondem a um comportamento equivalente ao de um classificador aleatório (Burkov, 2020).

Figura 17: Exemplos de curvas ROC com diferentes níveis de desempenho de modelos de classificação binária: (a) Curva ROC abaixo da linha aleatória, indicando baixo desempenho do modelo; (b) Curva ROC coincidente com a linha diagonal, representando classificação aleatória; (c) Curva ROC levemente acima da linha diagonal, indicando desempenho moderado; (d) Curva ROC bem acima da linha

diagonal, indicando bom desempenho do modelo.



Fonte: Burkov (2019)

2.4 Estado da arte

Aryal e Acharya (2022) desenvolveram um framework de previsão da estabilidade de taludes com base em 4208 conjuntos de dados gerados no software Slide, utilizando Bishop, Fellenius, Janbu, Morgenstern-Price e Spencer. Para a etapa de regressão, foram construídos modelos específicos de Regressão Linear Múltipla (MLR) e de Rede Neural Artificial (RNA) para cada um dos cinco métodos. Os valores de FS calculados por cada método serviram como variável-alvo independente, permitindo comparar o desempenho preditivo conforme o método determinístico adotado. Na regressão, todos os modelos MLR apresentaram R^2 superior a 0,90, enquanto a RNA superou a MLR em todos os casos, com melhor desempenho para o método de Spencer ($R^2 = 0,966$ no treinamento e $R^2 = 0,973$ no teste).

Já na etapa de classificação (Árvore de Decisão, SVC, Logistic Regression, Naïve Bayes, Floresta Aleatória, Gradient Boosting e RNA), os valores de FS foram utilizados apenas para rotular os dados como estáveis ou instáveis. O Decision Tree apresentou o melhor desempenho global, alcançando acurácia próxima de 99%, além de elevados valores de precisão, recall e F1-score, com baixo número de falsos positivos e falsos negativos na matriz de confusão. Os demais modelos também apresentaram desempenho consistente, todos com acurácia superior a 90%, destacando-se SVC e RNA com resultados próximos ao melhor modelo, enquanto Floresta Aleatória, Gradient Boosting, Logistic Regression e Naïve Bayes mantiveram desempenho satisfatório, porém ligeiramente inferior ao da Árvore de Decisão.

Nanehkaran *et al.* (2023) realizaram uma análise comparativa de quatro métodos de aprendizado de máquina para previsão do fator de segurança de taludes: Multilayer Perceptron (MLP), Árvore de Decisão, Support Vector Machines (SVM) e Floresta Aleatória. Utilizando uma base de 100 taludes da região do Irã, os modelos foram calibrados e validados contra o método de Janbu e o software GeoStudio. Os melhores resultados foram obtidos pelo MLP (acurácia = 0,901), seguido de SVM (acurácia = 0,873). O MLP também obteve a menor perda média (0,29), confirmando-se como o modelo mais robusto neste estudo.

Karir *et al.* (2022) aplicaram diferentes algoritmos de aprendizado de máquina para previsão de fatores de segurança em taludes de mineração. Entre os métodos utilizados destacam-se Floresta Aleatória, Gradient Boosting (GB), XGBoost (XGB) e Support Vector Regression (SVR). Os autores concluíram que o XGBoost apresentou o melhor desempenho geral, devido à sua capacidade de regularização, validação cruzada interna e bom manejo de dados ausentes. O XGB atingiu valores de R^2 superiores a 0,97, superando os demais métodos em termos de erro residual e estabilidade dos resultados.

Yadav *et al.* (2025) investigaram a aplicação de técnicas de aprendizado de máquina na previsão da estabilidade de taludes, adotando métodos de ensemble, como bagging e boosting, associados a Árvore de Decisão, Floresta Aleatória e SVM. Também incorporaram a técnica de redução de dimensionalidade Kernel Principal Component Analysis (KPCA) para avaliar a robustez dos modelos frente à diminuição do número de variáveis. A base de dados utilizada foi composta por 125 registros, contendo sete parâmetros relacionados às propriedades geotécnicas e geométricas dos taludes. A avaliação foi realizada por meio de divisões treino-teste (70:30 e 80:20) e validação cruzada 10-fold, sendo o desempenho médio das

partições representado pelo *CV-score*, utilizado para medir a capacidade de generalização.

Nos resultados, os métodos ensemble superaram os modelos individuais. O bagging com Árvore de Decisão e o boosting com Floresta Aleatória alcançaram acurácia superior a 90%, com ganhos médios de 8–10%. O *bagging* com Árvore de Decisão também apresentou *CV-score* próximo de 90%, indicando maior robustez. Na regressão, bagging e Lasso-Lars obtiveram os melhores valores de R^2 , mantendo ganhos médios semelhantes.

Azmoon, Biniyaz e Liu (2021) realizaram um estudo comparativo entre métodos tradicionais de equilíbrio limite e técnicas de *deep learning* para análise da estabilidade de taludes. Foram implementados quatro métodos clássicos de equilíbrio limite — Bishop Simplificado, Fellenius, Janbu Simplificado e Janbu Corrigido (JCM) — cujos fatores de segurança foram calculados por código próprio e validados no software Slide2. A média dos FS obtidos pelos quatro métodos foi adotada como valor de referência para rotular um banco sintético composto por 45.000 imagens de taludes homogêneos, geradas com variações geométricas e geotécnicas.

Dois modelos baseados em Redes Neurais Convolucionais (CNN) foram treinados: um modelo de classificação multiclases (nove faixas de FS) e um modelo de regressão (previsão contínua do FS). Nos testes com 9.000 imagens independentes, o modelo de classificação atingiu acurácia de 80,9% e erro absoluto médio de 0,0195, enquanto o modelo de regressão apresentou erro absoluto médio de 0,0265 e desempenho equivalente a 97,73% de acerto direto. Considerando tolerância de $\pm 0,1$ no FS (critério aceitável em engenharia geotécnica), ambos os modelos superaram 99,6% de precisão. Considerando tolerância de $\pm 0,1$ no FS, ambos superaram 99,6% de precisão e foram cerca de 18 vezes mais rápidos que os métodos de equilíbrio limite. (AZMOON; BINIYAZ; LIU, 2021).

Sheode, Anuragi e Dharavath (2025) propuseram um modelo híbrido BPSO (Otimização por Enxame de Partículas Binária) com SVM para previsão da estabilidade de taludes, reformulando o problema como uma tarefa de classificação binária (estável/instável). A base de dados foi composta por 108 casos de taludes, contendo sete parâmetros de entrada. Os dados foram divididos em 70% para treinamento e 30% para validação, com normalização entre 0 e 1. Os hiperparâmetros da SVM foram otimizados via Grid Search com validação cruzada, enquanto o BPSO foi empregado para seleção de atributos e ajuste fino do modelo, utilizando 50 partículas e 500 iterações.

Os resultados indicaram que o modelo híbrido BPSO-SVM superou a SVM convencional em termos de capacidade discriminatória, alcançando AUC = 1,00 para os kernels RBF e polinomial, enquanto a SVM isolada apresentou AUC máximo de 0,96. O modelo híbrido também reduziu o custo da função objetivo e selecionou subconjuntos relevantes de variáveis, demonstrando maior eficiência e robustez na classificação da estabilidade dos taludes. Os autores concluem que a integração entre técnicas de otimização evolutiva e SVM constitui uma ferramenta computacional promissora para aplicações em engenharia geotécnica.

Tun, Zeng e Jamil (2025) realizaram um estudo comparativo para previsão do fator de segurança de taludes utilizando cinco algoritmos de aprendizado de máquina, com uma base de dados foi composta por 1500 instâncias de taludes, com parâmetros geotécnicos e geométricos — coesão, ângulo de atrito, peso específico, altura do talude e inclinação — sendo o FS calculado por Método de Elementos Finitos no software FLAC/SLOPE. Os dados foram divididos em 80% para treinamento e 20% para teste, com ajuste de hiperparâmetros via grid search e random search. Os resultados indicaram que o MLP apresentou o melhor desempenho preditivo (MAE = 0,03607; RMSE = 0,05289; R² = 0,99011; AUC = 0,99418), seguido por SVM e k-NN, que também demonstraram elevada precisão. Os modelos DT e RF apresentaram desempenho inferior, porém satisfatório. Os autores concluíram que técnicas de aprendizado de máquina, especialmente o MLP, constituem ferramentas eficientes e confiáveis para previsão da estabilidade de taludes, podendo complementar métodos tradicionais baseados em análises numéricas.

Tabela 3: Comparação entre os artigos analisados.

Artigo	Aryal (2022)	Nanehkaran (2023)	Karir (2022)	Yadav (2025)	Azmoon (2021)	Sheode (2025)	Tun (2025)
Base total (n)	4208	100	906	125	45.000	108	1500
Divisão treino/teste (n)	80 / 20	70 / 30	80 / 20	80 / 20	80 / 20	70 / 30	80 / 20
Fonte do FS	LEM (Bishop, Morgenstern, Fellenius, Spencer, Janbu)	LEM (Janbu)	FEM	Não mencionado	LEM (Bishop, Fellenius, Janbu Corrigido, Janbu Simplificado)	Não mencionado	FEM

Validação cruzada (VC)	✓	X	✓	✓	X	✓	X
Grid Search / Random Search	✓	—	✓	✓	X	✓	✓
Modelos Aanalizados	9 (MLR; RNA; DT; SVC; LR; GBC; NB; RF)	4 (MLP; SVM; DT; RF)	4 (RF; GB; XGB; SVR)	3 (DT; RF; SVM)	1 (CNN)	2 (BPSO; SVM)	5 (MLP; SVM; k-NN; DT; RF)
Ensemble	Boosting	X	Boosting	Bagging / Boosting	X	X	X
Métricas adotadas	R ² , ACC, PREC; REC; F1	ACC; PREC	R ² , REQM	PREC; RECALL F1; ACC, CV-score; AUC; R ²	EMA; ACC	AUC	EMA, EQM, REQM, R ² , RSR, EPMA, VAF, ID, AUC
Melhor modelo	RNA	MLP	XGBoost	Boosting com RF	CNN	BPSO-SVM	MLP
Melhor desempenho	R ² = 0,973	ACC = 0,901 PREC = 0,90	R ² > 0,97	ACC > 90%	ACC ≈ 97–99% EMA = 0,026	AUC = 1,00	R ² = 0,990; AUC = 0,994; REQM = 0,05

3 OBJETIVOS

3.2 Objetivo geral.

Avaliar o desempenho de algoritmos de aprendizado de máquina — Redes Neurais Artificiais, Árvore de Decisão, Floresta Aleatória e Máquinas de Vetores de Suporte — na estimativa do fator de segurança de taludes, utilizando parâmetros geotécnicos e geométricos simulados, com base em métricas de erro e validação cruzada, comparando seus resultados aos métodos de equilíbrio limite.

3.3 Objetivos específicos

Entre os objetivos específicos, estão:

- Comparar o desempenho dos algoritmos de aprendizado de máquina na estimativa do fator de segurança de taludes, considerando diferentes métodos de equilíbrio limite.
- Avaliar a capacidade dos modelos em prever o fator de segurança como variável contínua (regressão), por meio de métricas de erro e coeficiente de determinação.
- Avaliar a capacidade dos modelos em classificar taludes como estáveis ou instáveis, com base em métricas de desempenho classificatório.
- Analisar a capacidade de generalização dos modelos, verificando a ocorrência de *overfitting* por meio da validação cruzada.
- Investigar a influência da complexidade dos métodos de equilíbrio limite no desempenho dos algoritmos de aprendizado de máquina.
- Analisar a importância das variáveis de entrada na estimativa do fator de segurança para cada modelo.
- Avaliar a viabilidade do uso de algoritmos de aprendizado de máquina como ferramenta auxiliar na análise de estabilidade de taludes.

4 METODOLOGIA

Este trabalho caracteriza-se como um estudo quantitativo, aplicado e exploratório, com abordagem computacional para previsão do fator de segurança de taludes a partir de parâmetros geotécnicos e geométricos, utilizando algoritmos de aprendizado de máquina supervisionado. A metodologia adotada compreende as seguintes etapas

4.1 Levantamento estatístico dos parâmetros geotécnicos e geométricos

O levantamento estatístico dos parâmetros geotécnicos e geométricos foi realizado a partir de um conjunto de dados secundários disponibilizado na plataforma Kaggle, intitulado *Slope Stability Analysis Dataset*. Trata-se de uma base de dados de natureza sintética, desenvolvida para representar diferentes cenários geotécnicos e geométricos de taludes, com variáveis definidas dentro de faixas realistas comumente adotadas na prática da engenharia geotécnica.

O dataset original é composto por 10.000 amostras independentes, nas quais cada amostra representa um cenário de talude caracterizado por um conjunto de variáveis de entrada e seus respectivos resultados. As variáveis de entrada incluem parâmetros geotécnicos — peso específico do solo (γ), coesão (c), ângulo de atrito interno (ϕ) e grau de saturação (ru) — e parâmetros geométricos — altura do talude (H) e ângulo de inclinação (β). A variável de saída corresponde ao fator de segurança (FS), calculado por diferentes métodos de equilíbrio limite.

Para cada amostra, o fator de segurança foi obtido por meio de quatro métodos de equilíbrio limite: Bishop Simplificado, Janbu, Spencer e Morgenstern–Price, resultando em quatro valores de FS associados a um mesmo conjunto de variáveis de entrada.

Devido ao elevado custo computacional associado ao cálculo do fator de segurança para todas as amostras, optou-se por utilizar 5.463 amostras neste estudo. A determinação do FS foi realizada individualmente para cada amostra no software geotécnico HYRCAN, exigindo processamento iterativo para busca da superfície crítica de ruptura. Considerando a aplicação dos quatro métodos de equilíbrio limite, o tempo total de processamento foi da ordem de três semanas de computação contínua, o que evidencia o alto custo computacional do procedimento.

Dessa forma, a utilização de uma amostra reduzida foi adotada como estratégia para viabilizar o estudo, mantendo, contudo, um conjunto de dados

estatisticamente representativo. O número de amostras utilizadas é compatível com estudos semelhantes da literatura que aplicam técnicas de aprendizado de máquina na análise de estabilidade de taludes, nos quais bases de dados de magnitude similar são empregadas para treinamento e validação dos modelos.

Assim, o conjunto de dados adotado apresenta dimensão suficiente para garantir a robustez das análises, ao mesmo tempo em que torna o custo computacional do problema viável.

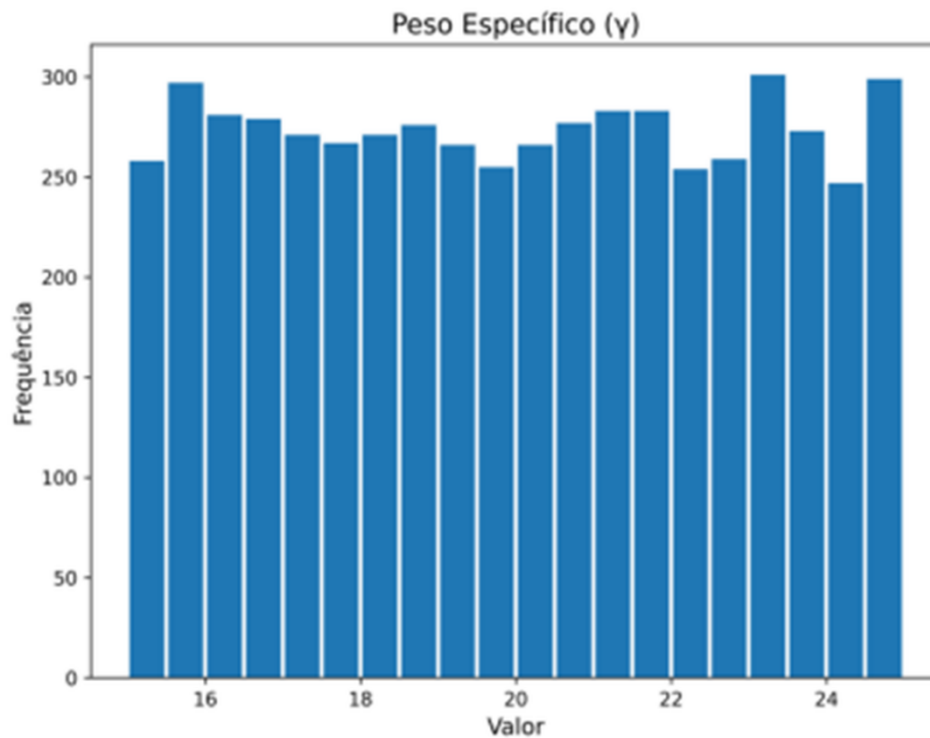
Para a caracterização estatística dos dados, foram calculadas medidas descritivas clássicas como valor mínimo, valor máximo, média, e desvio padrão (Tabela 4), assim como a distribuição em histogramas (Figura 19), para representar a dispersão, tendência central e a representatividade dos parâmetros estudados.

Tabela 4: Valores para os parâmetros geotécnicos e geométricos.

Variável	Mínimo	Máximo	Média	Desvio Padrão
Peso específico (kN/m ³)	15,00	24,99	20,00	2,90
Coesão (kPa)	5,01	49,99	26,94	13,04
Ângulo de atrito interno (°)	20,00	45,00	32,82	7,19
Altura do talude (m)	5,00	50,00	28,34	12,79
Ângulo do talude (°)	10,00	59,99	33,95	14,17
Grau de saturação (<i>ru</i>)	0,00	1,00	0,51	0,29

O peso específico distribui-se de forma homogênea no intervalo entre aproximadamente 15 e 25 kN/m³, garantindo representação equilibrada de solos com diferentes densidades.

Figura 18: Histograma de distribuição do peso específico



A coesão, variando entre cerca de 5 e 50 kPa, também apresenta frequência relativamente constante ao longo do intervalo, assegurando ampla variação da resistência coesiva no conjunto analisado (Figura 19). Já os valores do ângulo de atrito compreendem entre 20° e 45°, com distribuição relativamente constante dentro desse intervalo (Figura 20).

Figura 19: Histograma de distribuição da coesão.

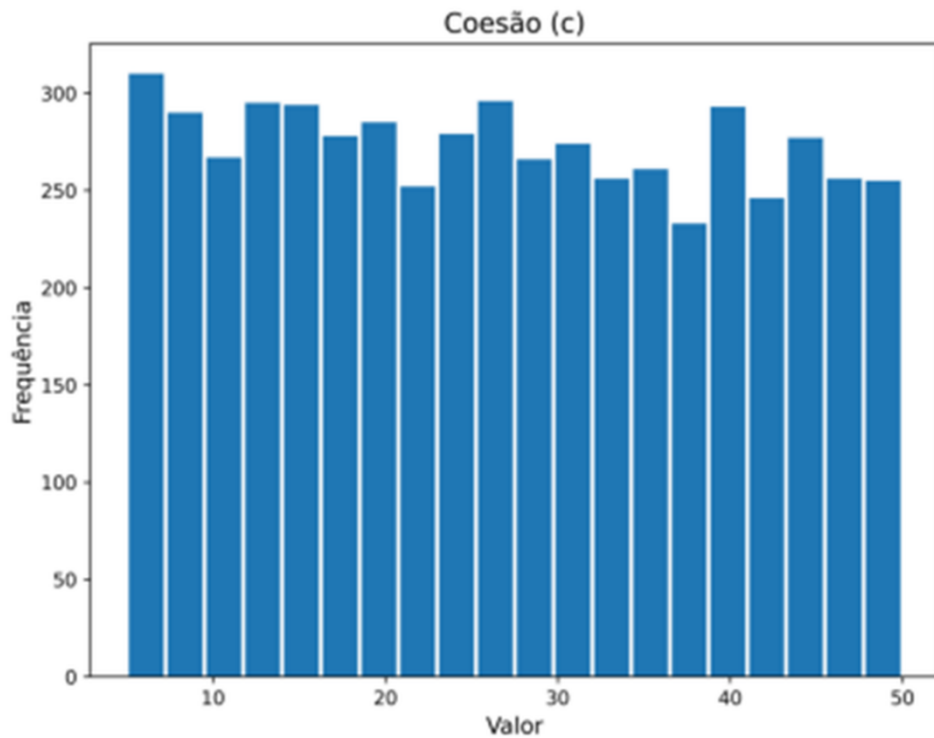
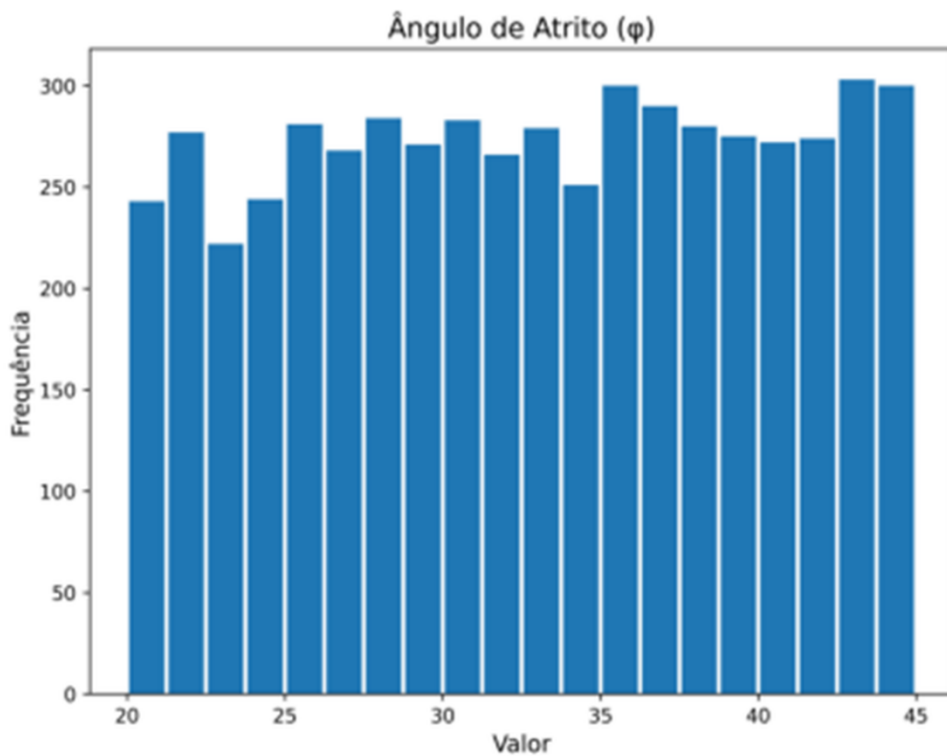


Figura 20: Histograma de distribuição do ângulo de atrito.



No que se refere às variáveis geométricas, a altura do talude (Figura 21) variando de aproximadamente 5 m a 50 m, e a inclinação (Figura 22) situada entre

cerca de 10° e 60°, apresentam distribuições equilibradas, permitindo que o modelo considere condições geométricas distintas.

Figura 21: Histograma de distribuição da altura do talude.

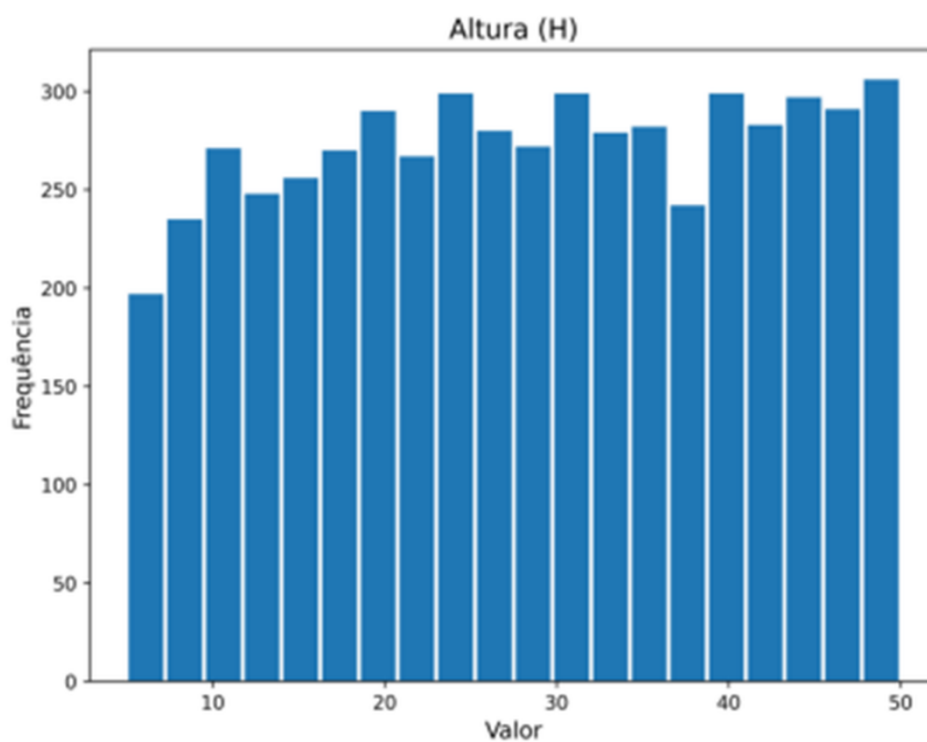
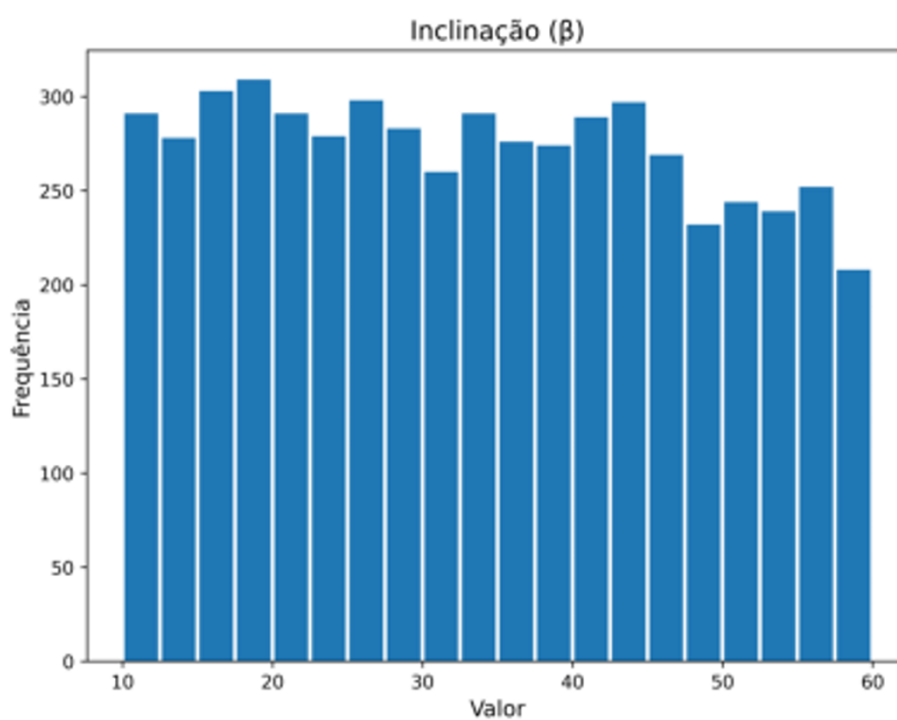
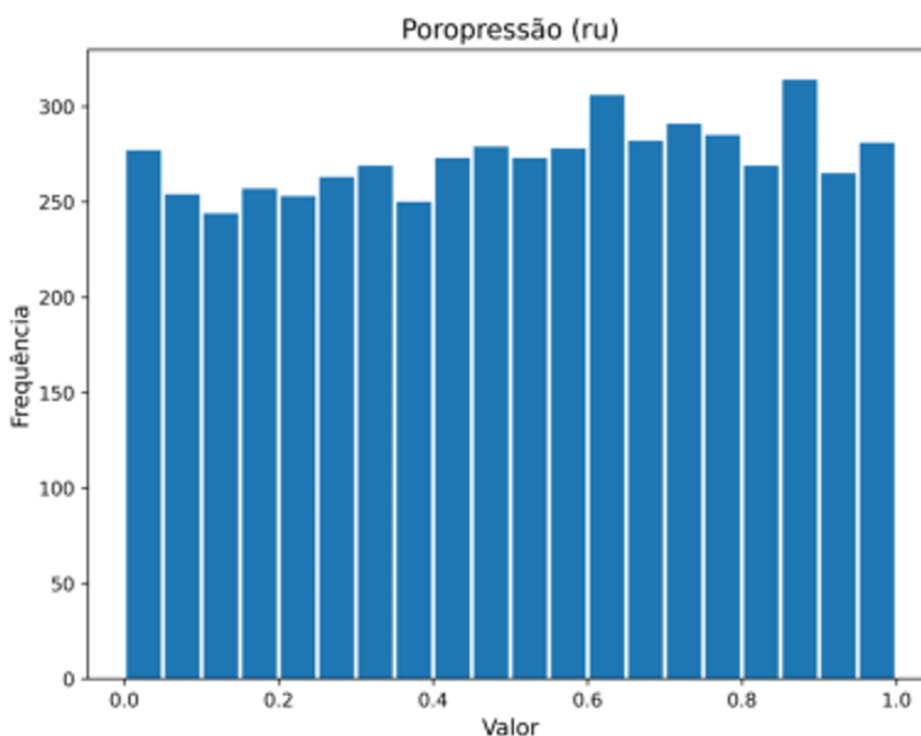


Figura 22: Histograma de distribuição da inclinação do talude.



A poropressão (ru), com valores entre 0 e 1, também se distribui de maneira uniforme, representando adequadamente diferentes condições de saturação e influência da pressão neutra no maciço.

Figura 23: Histograma de distribuição da poropressão.



De modo geral, não se observa concentração excessiva de valores em faixas específicas, nem assimetrias marcantes, o que evidencia a ausência de viés amostral significativo. Essa característica é fundamental para o treinamento de modelos de aprendizado de máquina, pois favorece a capacidade de generalização e reduz o risco de especialização excessiva em faixas restritas de valores, contribuindo para maior robustez nas previsões do fator de segurança.

4.2 Cálculo do Fator de Segurança via Análise de Equilíbrio Limite

O cálculo do fator de segurança foi realizado por meio da análise de equilíbrio limite, utilizando os métodos clássicos de Janbu, Bishop Simplificado, Morgenstern–Price e Spencer amplamente empregados na avaliação da estabilidade de taludes.

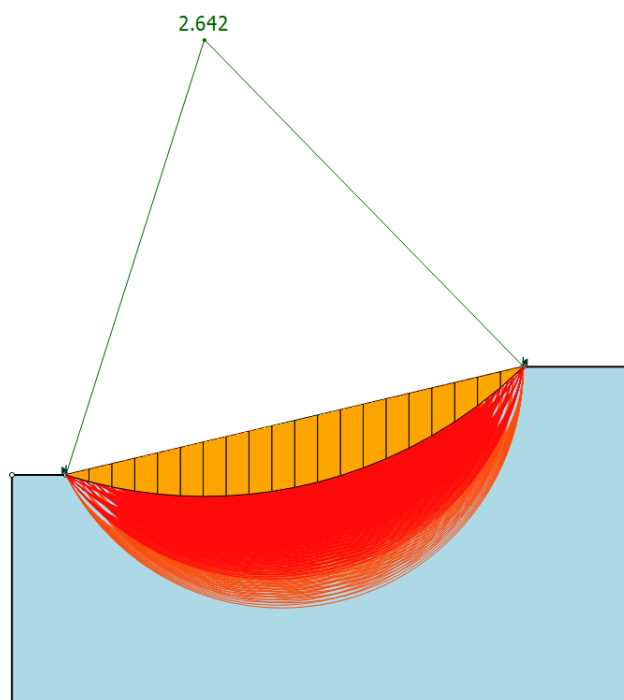
As análises foram executadas com o auxílio do software HYRCAN, uma ferramenta computacional específica para análises de estabilidade de taludes baseada em métodos de equilíbrio limite. O HYRCAN é um software de uso gratuito,

o que possibilita sua ampla utilização em contextos acadêmicos e de pesquisa.

Considerando o conjunto de dados composto por 5.000 amostras, os cálculos foram realizados de forma automatizada, por meio da utilização de rotinas de programação (scripting) integradas ao software. Essa abordagem permitiu a execução sistemática das análises para todas as amostras, garantindo maior eficiência na geração dos resultados.

A Figura 24 apresenta um modelo típico de talude analisado no estudo. O modelo consiste em um maciço homogêneo com superfície potencial de ruptura circular, dividida em fatias verticais para aplicação dos métodos de equilíbrio limite. As superfícies críticas foram determinadas por meio de busca automática, avaliando múltiplas geometrias e o fator de segurança indicado corresponde à superfície crítica identificada pelo algoritmo de busca, considerando o equilíbrio global do maciço.

Figura 24: Exemplo de talude com FS calculado.



Fonte: HYRCAN

Foram realizados cálculos utilizando 20 fatias verticais do maciço, valor adotado como compromisso entre precisão numérica e tempo para o cálculo. O número de fatias influencia diretamente a representação das forças atuantes ao longo da superfície de ruptura, pois valores muito baixos podem reduzir a precisão e valores elevados aumentam o tempo de processamento sem ganhos

significativos para a análise.

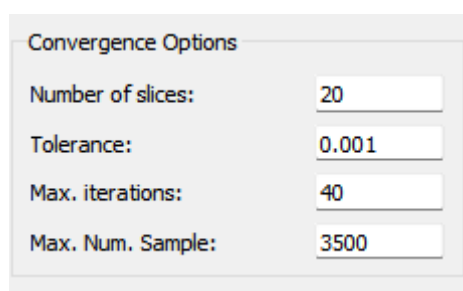
Os critérios de convergência utilizados no processo iterativo (Figura 25) foram definidos da seguinte forma:

Tolerância (0,001): corresponde ao critério de parada do processo iterativo para o cálculo do fator de segurança. A solução é considerada convergida quando a diferença entre duas iterações sucessivas do fator de segurança é inferior a 0,001. Esse valor garante estabilidade numérica adequada sem comprometer a eficiência computacional.

Número máximo de iterações (40): estabelece o limite superior de iterações permitidas para que o algoritmo alcance a convergência. Caso o critério de tolerância não seja satisfeito dentro desse limite, o processo é interrompido, evitando ciclos excessivos de cálculo.

Número máximo de amostras (3500): define o total máximo de superfícies potenciais avaliadas durante o processo de busca da superfície crítica. Esse parâmetro controla a abrangência da varredura.

Figura 25: Critérios de convergência adotados.



Convergence Options	
Number of slices:	20
Tolerance:	0.001
Max. iterations:	40
Max. Num. Sample:	3500

Fonte: HYRCAN

4.3 Construção do algoritmo para treinamento e teste dos modelos

A construção do algoritmo foi feita na linguagem de programação Python, utilizando bibliotecas amplamente utilizadas na área de aprendizado de máquina e computação científica. A biblioteca *pandas* foi empregada para a leitura, organização e manipulação dos dados, enquanto *NumPy* foi utilizada para operações numéricas e tratamento de matrizes. A construção, o treinamento e a aplicação dos algoritmos foram conduzidos com o auxílio da biblioteca *scikit-learn*, responsável por fornecer ferramentas para divisão de dados, definição de modelos e controle do processo de aprendizado.

O conjunto de dados, contendo os parâmetros geotécnicos e geométricos

dos taludes (γ , c , ϕ , H , β e ru) e seus respectivos fatores de segurança calculados pelos métodos de Bishop, Janbu, Spencer e Morgenstern-Price, inicialmente dividido em dois subconjuntos: 80% destinados ao treinamento e 20% reservados exclusivamente para teste. Essa separação foi realizada de forma aleatória, com o parâmetro `random_state = 42`, garantindo reprodutibilidade.

Foram definidos dois tipos de variáveis de saída: `y_fs`, correspondente aos valores contínuos do fator de segurança (utilizado para problemas de regressão), e `y_class`, correspondente à variável categórica de estabilidade (0 = instável; 1 = estável), empregada nas análises de classificação, conforme a seção do código apresentada na Figura 26.

Inicialmente, os dados foram organizados em variáveis de entrada, compostas pelos parâmetros geotécnicos e geométricos do talude, e variável de saída representada pelo fator de segurança (FS). A partir dessa estrutura, o problema foi abordado sob duas perspectivas distintas: regressão e classificação.

Na abordagem de regressão, os modelos foram treinados para estimar diretamente o valor contínuo do fator de segurança, preservando a natureza original da variável alvo. Já na abordagem de classificação, foi realizada a discretização do FS, considerando taludes estáveis aqueles com $FS \geq 1,5$ e instáveis aqueles com $FS < 1,5$, conforme critérios usuais de segurança geotécnica.

Essa separação metodológica permitiu avaliar o desempenho dos modelos tanto na estimativa quantitativa do fator de segurança quanto na capacidade de discriminar a condição de estabilidade dos taludes.

Figura 26 : Trecho do código Python referente à definição das variáveis de entrada e saída e à divisão do conjunto de dados em treinamento e teste

```

labels_variaveis = {
    'gamma': 'Peso específico',
    'c': 'Coesão',
    'phi': 'Ângulo de atrito interno',
    'H': 'Altura do talude',
    'beta': 'Ângulo do talude',
    'ru': 'Poropressão'
}
X = df[['gamma', 'c', 'phi', 'H', 'beta', 'ru']]
y_fs = df['FS']
y_class = df['Safety'] # 0 = instavel | 1 = estavel

X_train, X_test, y_fs_train, y_fs_test, y_class_train, y_class_test = train_test_split(
    X, y_fs, y_class,
    test_size=0.20,
    random_state=42
)
```

Durante o treinamento, para evitar o *overfitting* e selecionar a configuração mais adequada do modelo, foi aplicado o método validação cruzada do tipo k-fold ($k = 5$), conforme a Figura 27. Nesse procedimento, o conjunto de treinamento é subdividido em cinco partes (folds), sendo que, a cada iteração, quatro partes são utilizadas para o treinamento e uma para validação, alternando-se sucessivamente até que todas tenham sido utilizadas como conjunto de validação.

Figura 27: Definição da validação cruzada com 5 partições (K-Fold)

```
cv = KFold(n_splits=5, shuffle=True, random_state=42)
FS_LIMIAR = 1.5
metricas = []
```

Também para evitar *overfitting*, para cada modelo o algoritmo testa diferentes combinações de parâmetros definidas para cada um, por meio da técnica de grid search, onde se calcula o R^2 em cada iteração, permitindo a seleção da configuração que apresenta o menor erro médio. Ao final, são selecionados automaticamente o modelo com melhor desempenho, seus respectivos parâmetros ideais e o valor da métrica obtida durante a validação, conforme mostrado na Figura 28.

Figura 28 : Trecho do código referente à otimização dos modelos por meio de Grid

Search com validação cruzada para seleção dos melhores hiperparâmetros

```
for nome, modelo in modelos.items():

    grid = GridSearchCV(
        estimator=modelo,
        param_grid=param_grids[nome],
        cv=cv,
        scoring='r2',
        n_jobs=-1,
        verbose=1
    )

    grid.fit(X_train, y_fs_train)

    best_model = grid.best_estimator_
    best_params = grid.best_params_
    best_cv_r2 = -grid.best_score_

    print("Melhores parâmetros:", best_params)
    print("Melhor r2 (CV):", best_cv_r2)
```

Somente após essa etapa o modelo final é aplicado ao conjunto de teste (Figura 29). Inicialmente, são geradas as previsões contínuas do fator de segurança para os dados de teste. Em seguida, o desempenho do modelo é quantificado por meio de métricas de regressão, como o erro quadrático médio e o coeficiente de determinação que indicam, respectivamente, a magnitude média dos erros e o grau de explicação da variabilidade dos valores reais pelo modelo.

Paralelamente à estimativa contínua do fator de segurança, os resultados do modelo são também utilizados para uma análise de classificação binária da estabilidade dos taludes, na qual cada cenário é categorizado como estável ($FS \geq 1,5$) ou instável ($FS < 1,5$). A partir dessa categorização, calcula-se a acurácia do modelo, permitindo avaliar seu desempenho também sob a perspectiva de decisão prática de estabilidade, ampliando a análise comparativa entre os algoritmos.

Figura 29: Trecho do código referente à aplicação do modelo ao conjunto de teste e ao cálculo das métricas

```
y_fs_pred = best_model.predict(X_test)

mse = mean_squared_error(y_fs_test, y_fs_pred)
r2 = r2_score(y_fs_test, y_fs_pred)

# Classificação derivada
y_class_pred = (y_fs_pred >= FS_LIMIAR).astype(int)

acc = accuracy_score(y_class_test, y_class_pred)
```

Figura 30: Representação esquemática geral do funcionamento do algoritmo de aprendizado de máquina para este trabalho.



A seleção dos algoritmos utilizados neste estudo foi fundamentada tanto na literatura clássica de aprendizado de máquina quanto em trabalhos recentes aplicados à engenharia geotécnica, especialmente na análise de estabilidade de taludes. A escolha buscou contemplar modelos com diferentes níveis de complexidade, explicabilidade e capacidade de generalização, permitindo uma avaliação comparativa robusta do desempenho preditivo.

Do ponto de vista teórico, conforme discutido por Alpaydin (2004) e Guido e Müller (2016), não existe um algoritmo universalmente superior para todos os problemas, sendo recomendada a utilização de abordagens distintas que capturem diferentes padrões nos dados. Nesse sentido, foram selecionados quatro algoritmos amplamente consolidados: Árvore de Decisão, Floresta Aleatória, Máquinas de Vetores de Suporte (SVM) e Redes Neurais Artificiais (RNA).

A escolha dos algoritmos foi feita] baseada em evidências da literatura e na necessidade de comparar diferentes paradigmas de aprendizado: modelos interpretáveis (Árvore de Decisão), modelos ensemble (Floresta Aleatória), modelos baseados em margens (SVM) e modelos altamente não lineares (RNA).

4.3.1 Árvore de decisão

A Árvore de Decisão foi escolhida por sua alta interpretabilidade e capacidade de representar relações não lineares de forma hierárquica. Em estudos geotécnicos, modelos baseados em árvores têm sido aplicados com sucesso na identificação de padrões em dados de estabilidade, como demonstrado por Aryal et al. (2022), que destacam sua eficiência na interpretação dos fatores que influenciam o fator de segurança.

No projeto, a Árvore de Decisão, durante o processo de treinamento, realiza divisões sucessivas dos conjuntos de dados de treino com base nos parâmetros geotécnicos e geométricos dos taludes. Em cada etapa, o algoritmo seleciona uma variável e um limiar de decisão, formulando perguntas do tipo “altura do talude > 20 m”, cujas respostas (“sim” ou “não”) dividem o conjunto de dados em dois ramos. Esse processo é repetido até que se atinjam os nós terminais, nos quais é atribuído um valor estimado do fator de segurança para cada grupo de amostras.

A profundidade máxima da árvore foi definida como hiperparâmetro variável no processo de *grid search*, assumindo os valores 3, 5, 7, 9 e 12. A variação desse parâmetro teve como objetivo avaliar diferentes níveis de complexidade do modelo e selecionar a configuração com melhor desempenho por meio da validação cruzada. O critério adotado para a definição das divisões foi o Índice de Gini, responsável por mensurar a pureza dos subconjuntos formados em cada ramificação.

Na fase de teste, o modelo correspondente à melhor combinação de hiperparâmetros identificada pelo processo de busca em grade com validação cruzada é aplicado ao conjunto de dados reservado para teste, percorrendo os

ramos da árvore até que cada amostra atinja um nó terminal, resultando na estimativa do fator de segurança correspondente.

4.3.2 Floresta Aleatória

A Floresta Aleatória foi incluída como uma extensão da Árvore de Decisão, utilizando aprendizado em conjunto (ensemble learning) para reduzir a variância e aumentar a robustez das previsões. Trabalhos recentes, como os de Nanehkaran et al. (2023) e Yadav et al. (2025), indicam que a Floresta Aleatória apresenta elevado desempenho na previsão do fator de segurança de taludes, sendo frequentemente superior a modelos individuais devido à sua resistência a ruídos e outliers.

A Floresta Aleatória realiza o processo de treinamento criando Árvores de Decisão a partir dos conjuntos de dados de treino. Cada árvore é construída de forma independente, utilizando uma amostra aleatória dos dados e um subconjunto aleatório dos parâmetros disponíveis em cada divisão. Assim como na Árvore de Decisão, as divisões dentro de cada árvore são feitas com base nos parâmetros dos taludes, criando sucessivos ramos até atingir o nó terminal, que fornece um fator de segurança médio associado aos dados daquele grupo.

Na Floresta Aleatória, a profundidade máxima das árvores e o número de árvores do conjunto foram definidos como hiperparâmetros variáveis no processo de busca em grade. Foram avaliadas diferentes combinações de profundidade máxima (5, 8, 10 e sem limite definido) e número de árvores (100, 200, 300 e 500), com o objetivo de analisar o impacto da complexidade individual das árvores e da robustez do conjunto sobre o desempenho preditivo.

Para definir as divisões dentro de cada árvore, assim como no modelo anterior, também foi utilizado o Índice de Gini, que avalia a homogeneidade dos grupos formados.

Na fase de teste, o modelo correspondente à melhor combinação de hiperparâmetros identificada no Grid Search foi aplicado ao conjunto de dados de teste. Cada árvore do conjunto gera uma estimativa individual do fator de segurança, sendo o valor final obtido pela média das previsões, característica fundamental do método de agregação da Floresta Aleatória.

4.3.3 Máquinas de Vetores Suporte (SVM)

As Máquinas de Vetores de Suporte (SVM) foram selecionadas devido à sua eficácia em problemas de alta dimensionalidade e sua capacidade de modelar relações não lineares por meio do uso de funções kernel. No contexto da engenharia geotécnica, estudos como os de Nanekaran (2023) e Yadav (2025) demonstram que a SVM apresenta desempenho consistente na estimativa de parâmetros geotécnicos e estabilidade de taludes, especialmente em conjuntos de dados complexos e não lineares.

A Máquina de Vetores de Suporte foi aplicada para realizar regressão, estimando valores contínuos de fator de segurança a partir dos conjuntos de dados de treino. Durante o treinamento, o algoritmo mapeia os dados para um espaço de maior dimensão utilizando o *kernel radial* (RBF), o que permite modelar relações não lineares entre os parâmetros dos taludes e o fator de segurança. O modelo SVM foi implementado utilizando um pipeline contendo normalização dos dados (StandardScaler) e o algoritmo SVR. A normalização foi necessária devido à sensibilidade do modelo à escala das variáveis.

Os hiperparâmetros do modelo foram definidos por meio de busca em grade com validação cruzada. Foram avaliados os seguintes valores:

- Parâmetro de regularização (C): 1, 10, 50 e 100
- Margem de tolerância (ϵ): 0,01; 0,05; 0,1 e 0,2

O parâmetro C controla o nível de penalização aplicado aos erros de previsão, sendo que valores maiores tendem a reduzir o erro no conjunto de treinamento, mas podem aumentar o risco de *overfitting*. O parâmetro ϵ define a largura da zona de insensibilidade em torno da função de regressão, dentro da qual os erros não são penalizados.

A combinação que apresentou o maior R^2 médio durante a validação cruzada foi selecionada como configuração ótima do modelo. Na fase de teste, essa configuração final foi aplicada ao conjunto de dados de treino, produzindo as estimativas utilizadas para o cálculo das métricas de desempenho.

4.3.4 Redes Neurais Artificiais

As Redes Neurais Artificiais foram incluídas devido à sua elevada capacidade de modelar relações altamente não lineares entre variáveis de entrada e saída. Segundo Alpaydin (2004), esses modelos são particularmente eficazes quando o fenômeno estudado envolve interações complexas, como é o caso da estabilidade de taludes. Diversos estudos recentes, como os de Nanekaran et al.

(2023) e Dhruva Karir (2022), evidenciam o alto desempenho das redes neurais na previsão do fator de segurança, frequentemente superando métodos tradicionais e outros algoritmos de aprendizado de máquina.

A Rede Neural Artificial foi aplicada para estimar valores contínuos do fator de segurança a partir do conjunto de treinamento (80% da amostra total). A arquitetura da rede, representada pelo número de neurônios nas camadas ocultas, e o número máximo de iterações foram definidos como hiperparâmetros variáveis no processo de busca em grade.

Foram avaliadas diferentes configurações de camadas ocultas — (10), (10,10), (50), (50,50) e (100) — bem como diferentes limites de iteração (500, 1000 e 1500), com o objetivo de analisar o impacto da complexidade estrutural e do tempo de treinamento sobre o desempenho do modelo. A Rede Neural Artificial foi implementada por meio do modelo MLPRegressor, também inserido em um pipeline com normalização dos dados. A arquitetura adotada foi do tipo perceptron multicamada (MLP).

Durante o treinamento, os dados são processados pela rede por meio da combinação linear das entradas com pesos ajustáveis, seguida da aplicação de funções de ativação. O erro entre o valor previsto e o valor real do fator de segurança é calculado e propagado retroativamente pelo método de retropropagação, permitindo a atualização iterativa dos pesos para minimizar o erro.

Na fase de teste, o modelo correspondente à melhor arquitetura identificada no Grid Search foi aplicado ao conjunto de dados de treino, utilizando os pesos ajustados durante o treinamento para gerar as estimativas finais do fator de segurança..

4.4 Avaliação dos modelos

Essa etapa teve como objetivo verificar a capacidade dos modelos em estimar corretamente o valor contínuo do fator de segurança. Para essa finalidade, foram utilizadas métricas que comparam diretamente os valores previstos pelos modelos com os valores reais obtidos pelos métodos de equilíbrio limite.

- Coeficiente de determinação (R^2): indica o quanto da variabilidade dos valores reais do fator de segurança é explicada pelas previsões do modelo.
- Erro quadrático médio (MSE): quantifica a magnitude média dos erros cometidos pelo modelo, fornecendo uma medida direta da diferença entre os

valores previstos e os valores reais.

- *CV-Score*: Média do R^2 cv-obtido nas cinco divisões da validação cruzada (K-Fold = 5) de cada modelo

Além da análise contínua do fator de segurança, foi realizada uma avaliação complementar sob a ótica da classificação da estabilidade dos taludes, na qual os resultados foram convertidos em classes discretas. Essa abordagem permite interpretar o desempenho dos modelos em termos de tomada de decisão, aspecto relevante em aplicações de engenharia.

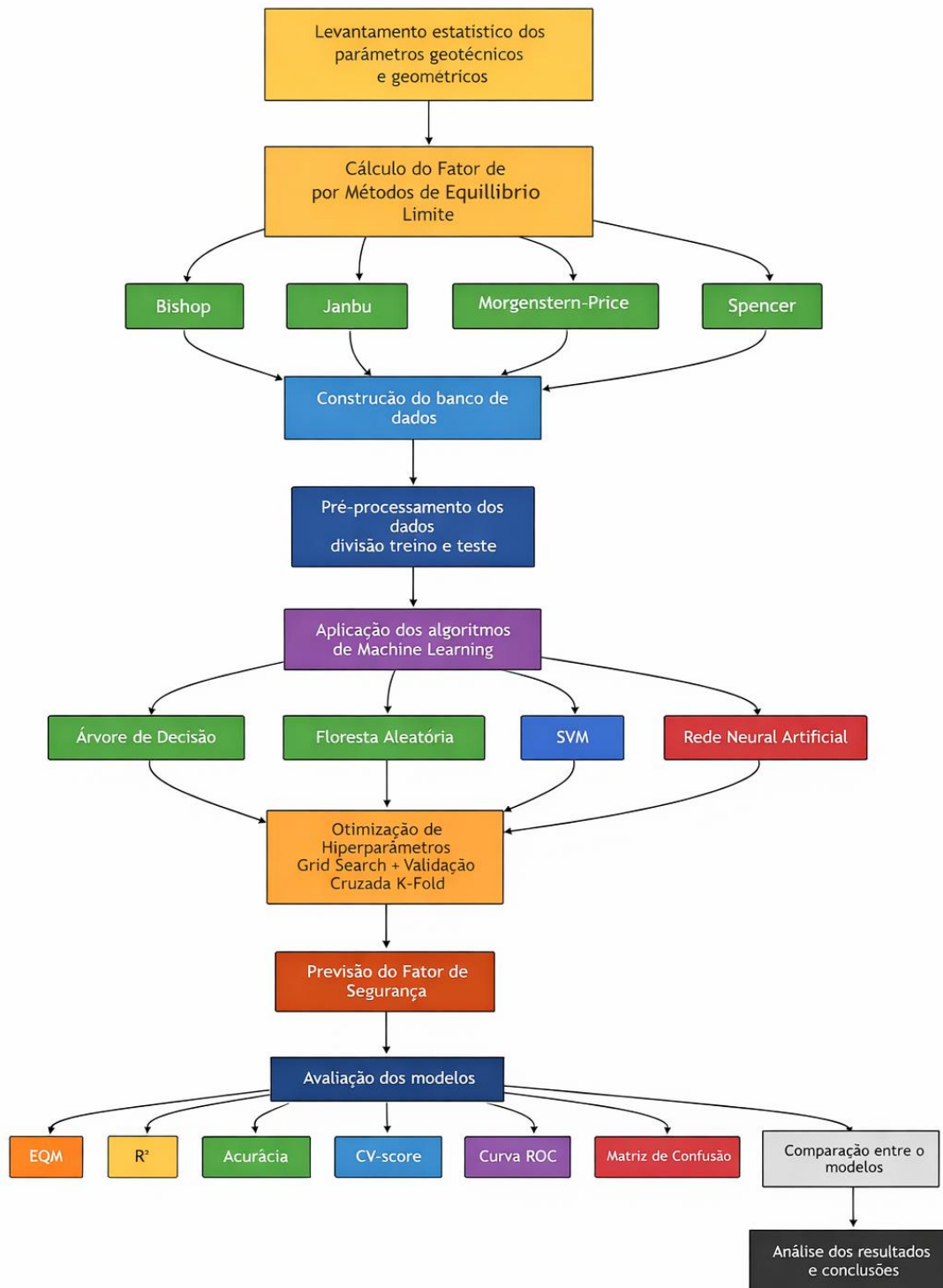
- *Acurácia*: Utilizada para quantificar a proporção de classificações corretas, considerando simultaneamente os casos estáveis e instáveis.
- *Matriz de confusão*: Empregada para detalhar o desempenho do modelo na identificação correta de cada classe, permitindo a análise dos acertos e erros de classificação.
- *Curva ROC*: Utilizada para avaliar a capacidade discriminatória dos modelos ao longo de diferentes limiares de decisão, fornecendo uma visão global do desempenho na distinção entre taludes estáveis e instáveis.

De forma complementar, foram construídos gráficos destinados à interpretação do papel das variáveis de entrada no desempenho dos modelos para identificar quais parâmetros exercem maior influência na capacidade preditiva dos modelos, por meio do método de importância por permutação (Permutation Importance), aplicado sobre o conjunto de teste. Esse método consiste em medir a variação no desempenho do modelo após a aleatorização dos valores de cada variável de entrada, mantendo as demais constantes.

O procedimento adotado foi realizado utilizando a função `permutation_importance` da biblioteca `scikit-learn`, com 10 repetições para cada variável (`n_repeats=10`) e `random_state=42`, garantindo a reprodutibilidade dos resultados. Como métrica de avaliação, foi utilizado o coeficiente de determinação (R^2).

A lógica do método baseia-se no fato de que, ao embaralhar os valores de uma variável relevante, ocorre uma degradação significativa no desempenho do modelo. Assim, quanto maior a redução no valor de R^2 após a permutação, maior é a importância dessa variável para a predição do fator de segurança.

Figura 31: Fluxograma da metodologia do trabalho



5 RESULTADOS E DISCUSSÕES

5.1 Resultados do fator de segurança obtidos pelo método de equilíbrio limite

Os resultados obtidos no cálculo por meio do HYRCAN (Tabela 5) obtiveram fatores de segurança que variaram entre os diferentes métodos de equilíbrio limite.

A média dos fatores de segurança evidencia diferenças mais marcantes entre os métodos. O método de Spencer apresentou a maior média (1,54), indicando tendência geral a estimar fatores de segurança mais elevados. Em seguida, aparecem Morgenstern–Price (1,23) e Bishop (1,19), com valores bastante próximos entre si, enquanto Janbu apresentou a menor média (1,08), sugerindo comportamento mais conservador na estimativa do fator de segurança.

Quanto à dispersão dos resultados, os desvios padrão variaram entre 0,96 (Janbu) e 1,06 (Spencer), indicando que todos os métodos apresentam variabilidade semelhante nas estimativas. Isso sugere que as diferenças observadas entre os métodos estão associadas principalmente ao deslocamento da média e não a diferenças substanciais na dispersão dos dados.

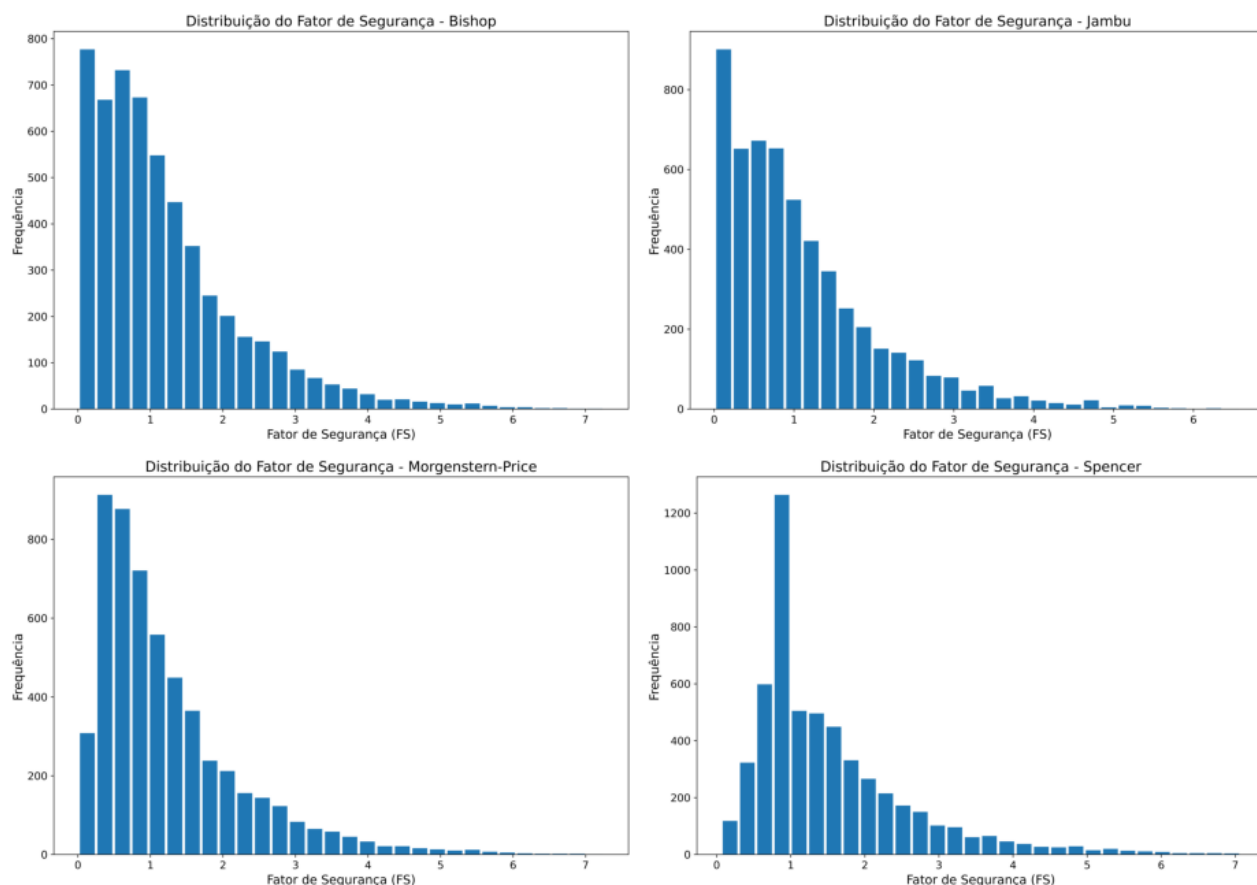
Tabela 5: Valores calculados para o fator de segurança

Método	Mínimo	Máximo	Média	Desvio padrão
FS – Bishop	0,01	7,25	1,19	1,03
FS – Janbu	0,01	6,58	1,08	0,96
FS – Morgenstern–Price	0,01	7,25	1,23	1,00
FS – Spencer	0,06	7,07	1,54	1,06

Observando os histogramas (Figura 32) observa-se uma distribuição assimétrica à esquerda, com grande concentração de valores entre 0 e aproximadamente 2,0 e uma cauda longa que se estende até valores superiores a 6 ou 7, indicando que a maioria dos cenários apresenta fatores de segurança baixos ou moderados, enquanto poucos casos atingem valores elevados. No caso de Janbu e Bishop, a concentração de frequências é mais intensa abaixo do limiar de 1,5, com maior proporção de taludes classificados como instáveis. Já em Spencer, nota-se um deslocamento da distribuição para valores ligeiramente mais altos, aumentando a quantidade de observações acima do limite de estabilidade. Morgenstern–Price apresenta comportamento muito semelhante ao de Bishop,

porém com leve tendência a valores mais elevados, coerente com sua média ligeiramente superior.

Figura 32 : Histograma dos fatores de segurança calculados



5.2 Desempenho dos modelos de aprendizado de máquina

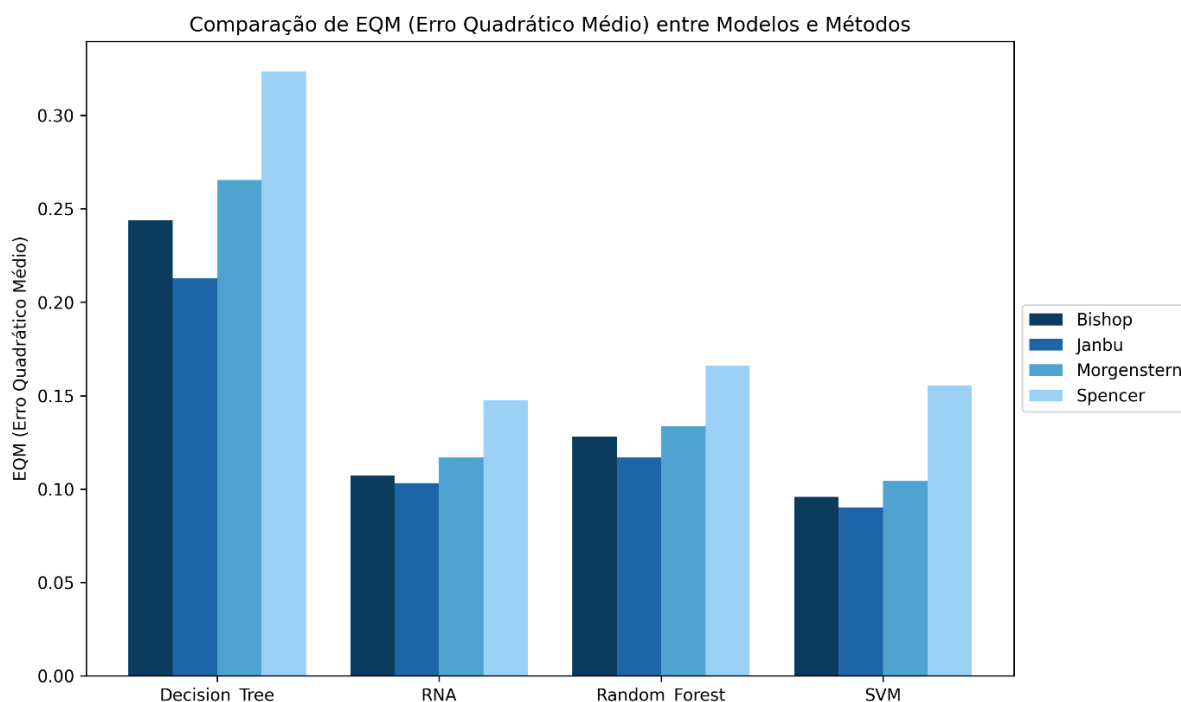
A avaliação do desempenho dos modelos de aprendizado de máquina para a estimativa do Fator de Segurança revelou padrões distintos que variam conforme a complexidade matemática do Método de Equilíbrio Limite adotado.

5.2.1 Análise do Erro Quadrático Médio

O Erro Quadrático Médio (EQM) permitiu identificar que o modelo SVM, especialmente quando aplicado ao método de Janbu e Bishop, atingiu um patamar de erro extremamente reduzido (0,090 e 0,096), o que evidencia uma alta capacidade de regressão. Ao observar o gráfico de barras comparativo de erros ,

percebe-se que, enquanto SVM e RNA mantêm erros baixos e constantes entre os diferentes métodos, a Árvore de Decisão apresenta picos de erro significativos, sinalizando uma dificuldade em lidar com a continuidade dos valores de FS.

Figura 33 : Comparação do EQM entre os modelos de aprendizado de máquina e métodos de equilíbrio limite.

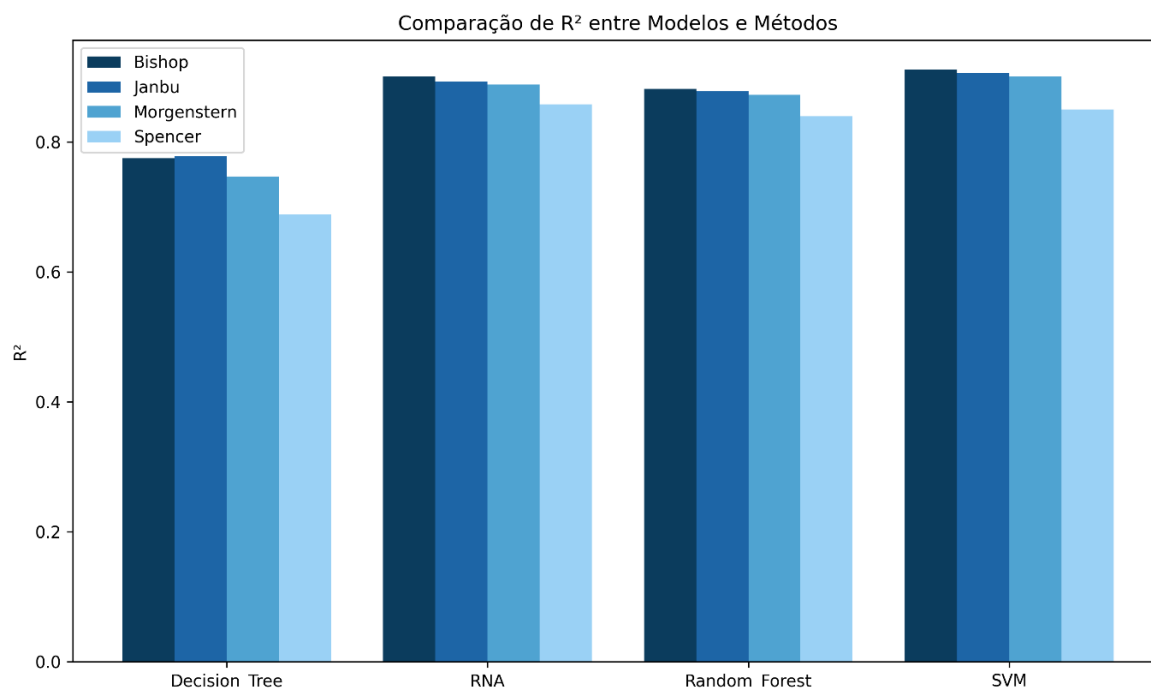


5.2.2 Análise do Coeficiente de Determinação (R^2)

O Coeficiente de Determinação (R^2) complementa a análise. Para os métodos de Janbu, Bishop e Morgenstern-Price, os valores de R^2 superaram a marca de 0,90 nos modelos SVM e RNA, com o maior valor de 0,911 para o modelo de SVM com Bishop, indicando que as variáveis de entrada são altamente explicativas para o comportamento do talude nestas formulações.

No entanto, o gráfico (Figura 33) também revela uma queda de desempenho no método de Spencer para todos os modelos, sugerindo que a maior complexidade deste método introduz uma dispersão que desafia o ajuste perfeito dos modelos, embora os resultados ainda se mantenham em níveis estatisticamente aceitáveis (em torno de 0,85 para RNA, SVM e Floresta Aleatória). Da mesma forma, os resultados para a Árvore de decisão também foram menos precisos para todos os métodos, em comparação aos outros três modelos.

Figura 34 : Comparação do R^2 entre os modelos de aprendizado de máquina e métodos de equilíbrio limite.

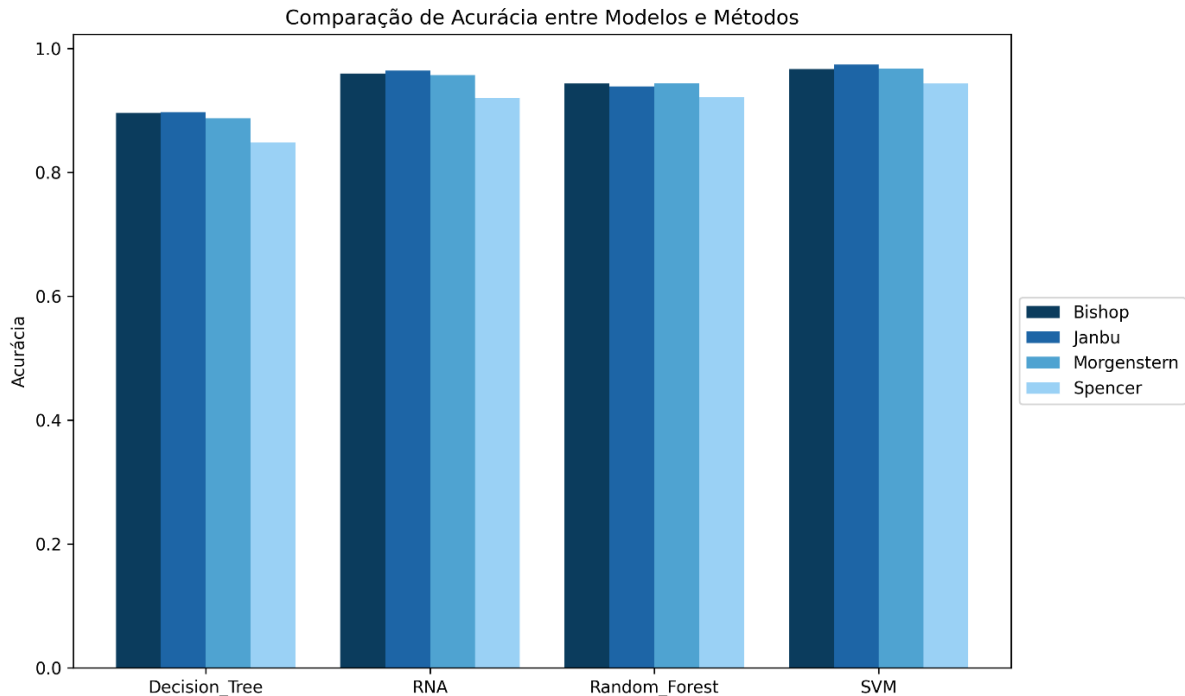


5.2.3 Análise da Acurácia

A Acurácia foi utilizada para avaliar a eficácia dos modelos sob uma perspectiva de classificação binária, ou seja, na capacidade de distinguir corretamente entre taludes estáveis e instáveis. Os resultados demonstraram que, com exceção dos modelos de Árvore de Decisão, todos os demais atingiram uma Acurácia maior que 90%, com destaque para o SVM, que alcançou uma precisão de 97,43% para Janbu, e 96,71% no cenário de Bishop.

Esses resultados mostram que, embora a predição exata do valor numérico do Fator de Segurança possa apresentar variações maiores entre os modelos, conforme evidenciado pelos resultados do EQM e R^2 , a sensibilidade dos algoritmos para identificar o estado crítico na classificação binária de segurança é bastante satisfatória.

Figura 35 : Comparação da Acurácia entre os modelos de aprendizado de máquina e métodos de equilíbrio limite.



5.2.4 Análise do CV-score

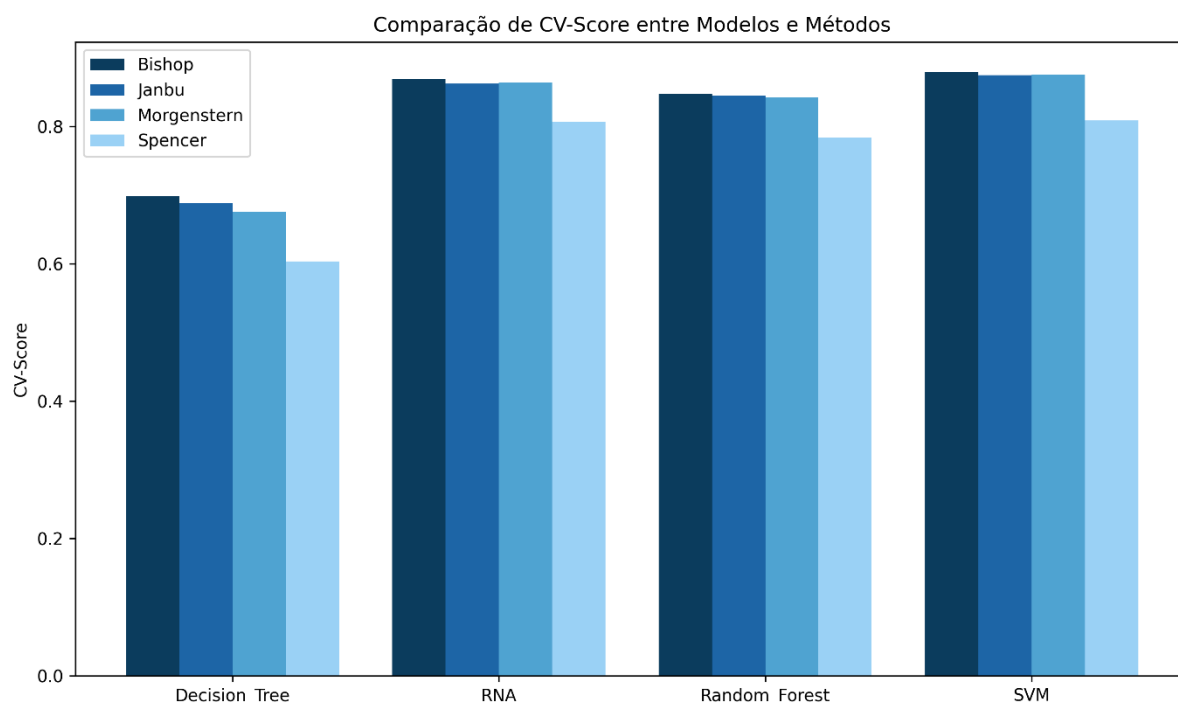
Por fim, o *CV-score* (validação cruzada) funcionou como o principal indicador de robustez e generalização dos modelos, permitindo aferir se o desempenho observado seria replicável em novos conjuntos de dados. Conforme ilustrado no gráfico de estabilidade, o SVM demonstrou a maior consistência global, mantendo scores elevados mesmo em métodos rigorosos. Em contrapartida, a Árvore de Decisão apresentou a maior volatilidade nos resultados de validação, provando ser o algoritmo mais sensível a pequenas variações na amostra.

Além disso, a comparação direta entre os valores de R^2 obtidos no conjunto de teste e o *CV-score* evidencia diferenças que permitem avaliar a estabilidade dos modelos. De modo geral, observa-se que o *CV-score* é sistematicamente inferior ao R^2 em todos os casos analisados, o que é esperado, uma vez que a validação cruzada tende a fornecer uma estimativa mais conservadora do desempenho do modelo. As maiores diferenças negativas entre R^2 e *CV-score* são observadas na Árvore de Decisão para todos os métodos, especialmente para o método Spencer, em que o R^2 de aproximadamente 0,689 reduz-se para um *CV-score* de cerca de

0,604, indicando maior variabilidade entre as partições da validação cruzada e menor capacidade de generalização.

Por outro lado, os modelos SVM, RNA e Floresta Aleatória apresentam diferenças menores entre R^2 e *CV-score*, o que sugere maior estabilidade e melhor capacidade de generalização. Destaca-se particularmente o desempenho do SVM, que apresenta diferenças pequenas entre R^2 e *CV-score* em todos os métodos, como observado nos métodos Bishop, Morgenstern e Janbu, em que o R^2 permanece acima de aproximadamente 0,90 enquanto o *CV-score* se mantém próximo de 0,87–0,88. Esses resultados indicam que modelos mais robustos e capazes de capturar relações não lineares nos dados apresentam maior consistência entre desempenho observado e desempenho validado, reforçando sua adequação para a predição do fator de segurança em análises de estabilidade de taludes.

Figura 36 : Comparação do *CV-score* entre os modelos de aprendizado de máquina e métodos de equilíbrio limite.



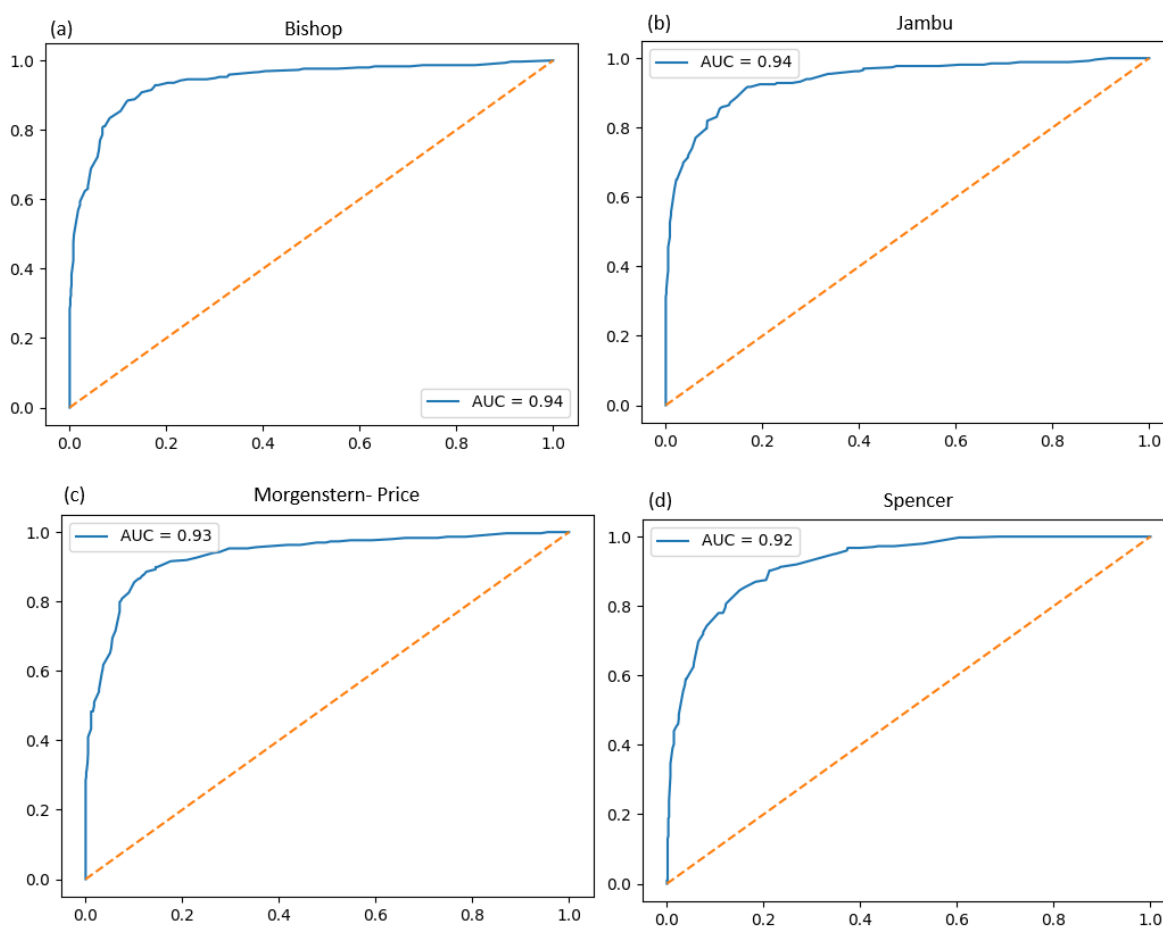
5.2.5 Análise da Curva ROC

Assim como na acurácia, a análise das curvas ROC obtidas para os

diferentes modelos e métodos de equilíbrio limite demonstra que todos os algoritmos apresentaram elevada capacidade de discriminação entre taludes estáveis e instáveis. Em todos os casos, as curvas situam-se significativamente acima da linha diagonal de referência, indicando desempenho muito superior ao de um classificador aleatório. Os valores de AUC variaram aproximadamente entre 0,92 e 0,98, mostrando um desempenho excelente a excepcional na tarefa de classificação binária.

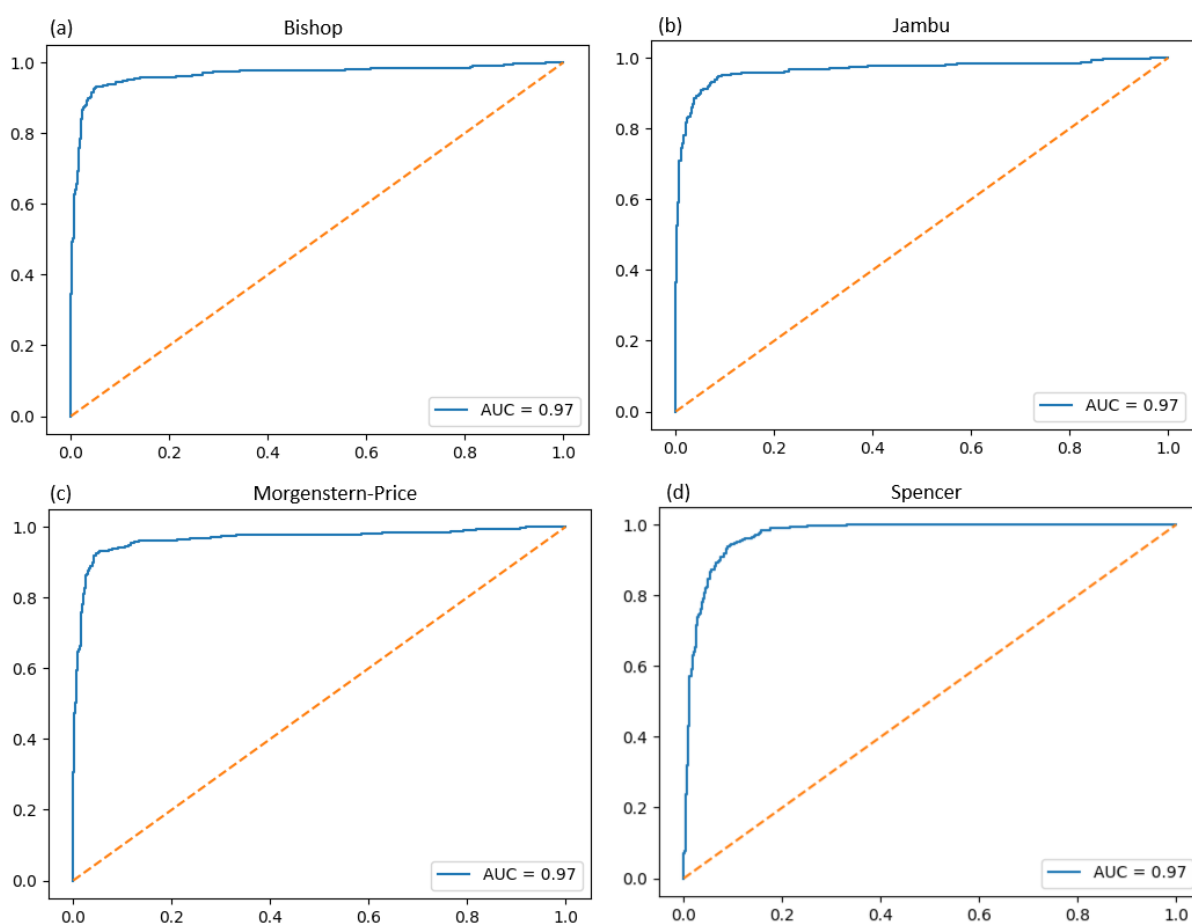
Referente ao modelo de Árvore de Decisão, observa-se bom desempenho em todos os métodos de equilíbrio limite, com valores de AUC entre aproximadamente 0,92 e 0,94. Embora as curvas indiquem capacidade satisfatória de discriminação entre as classes, os valores são ligeiramente inferiores aos obtidos pelos demais modelos, sugerindo maior sensibilidade às variações dos dados e menor robustez em comparação com algoritmos mais complexos.

Figura 37 : Curvas ROC obtidas para Árvore de Decisão: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



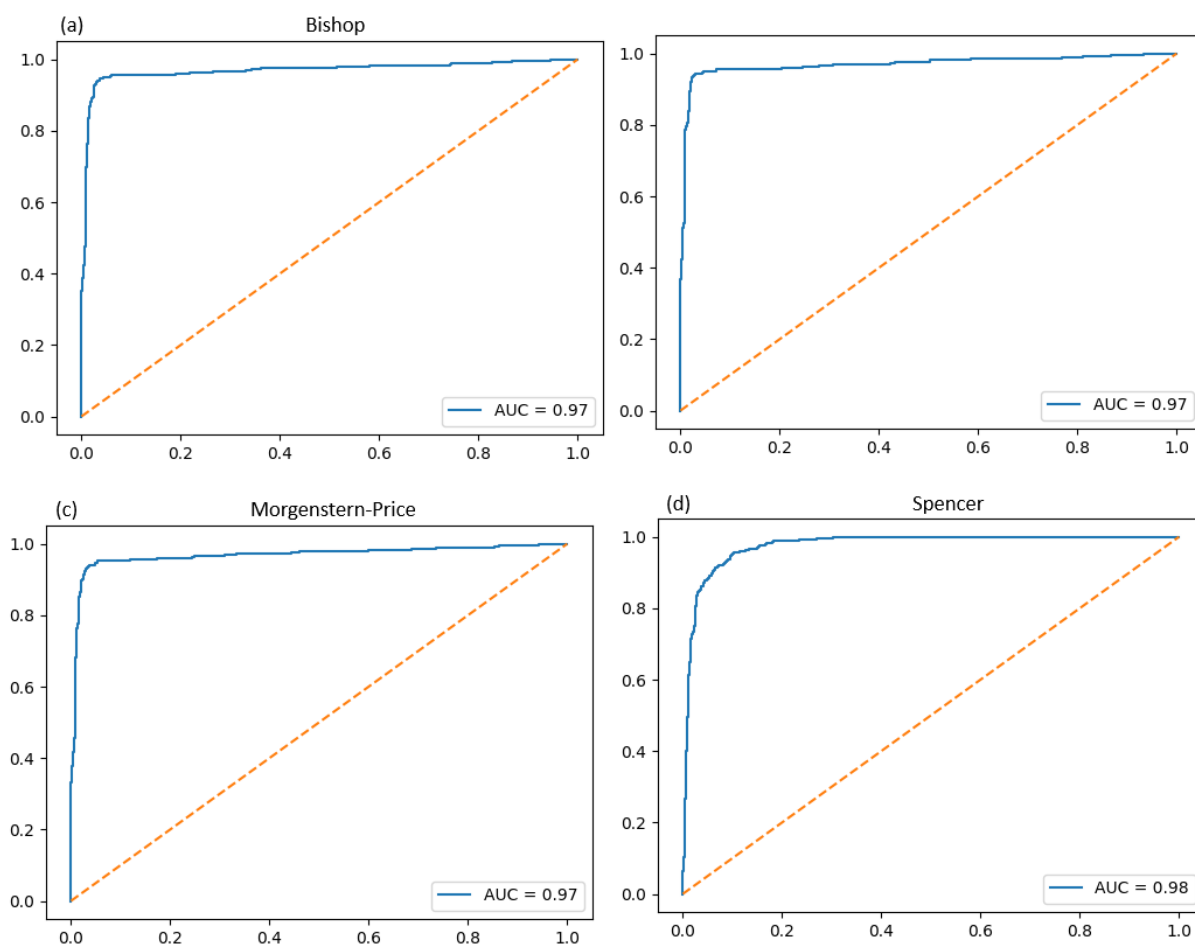
O modelo de Floresta Aleatória, apresenta curvas ROC com AUC próximas de 0,97 para todos os métodos analisados. Esse resultado demonstra elevado desempenho e grande estabilidade do modelo, indicando que a abordagem baseada em ensemble de árvores de decisão melhora significativamente a capacidade de separação entre taludes estáveis e instáveis em relação à árvore de decisão individual.

Figura 38 : Curvas ROC obtidas para Floresta Aleatória: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



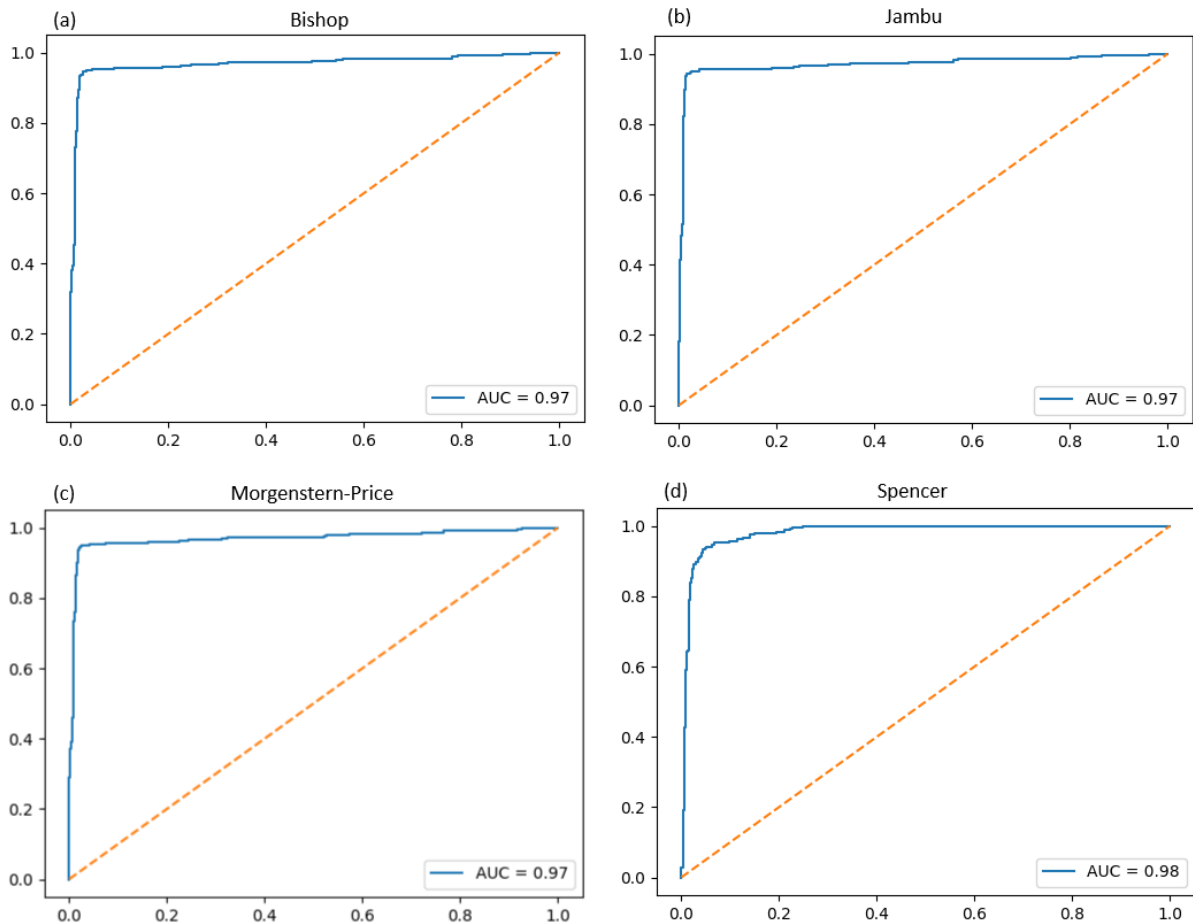
Para a Rede Neural Artificial, as curvas ROC também apresentam AUC próximas de 0,97, indicando um desempenho elevado. O comportamento observado é consistente entre os diferentes métodos de equilíbrio limite, evidenciando que as redes neurais foram capazes de prever satisfatoriamente a classificação binária para todos os métodos de equilíbrio limite, mesmo com as diferenças de abordagem entre eles.

Figura 39: Curvas ROC obtidas para RNA: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Por fim, o modelo SVM, mostra os maiores valores de AUC entre os modelos analisados, variando aproximadamente entre 0,97 e 0,98. Esses resultados indicam excelente capacidade de discriminação e reforçam a elevada capacidade de generalização do SVM na classificação da estabilidade de taludes.

Figura 40: Curvas ROC obtidas para SVM: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



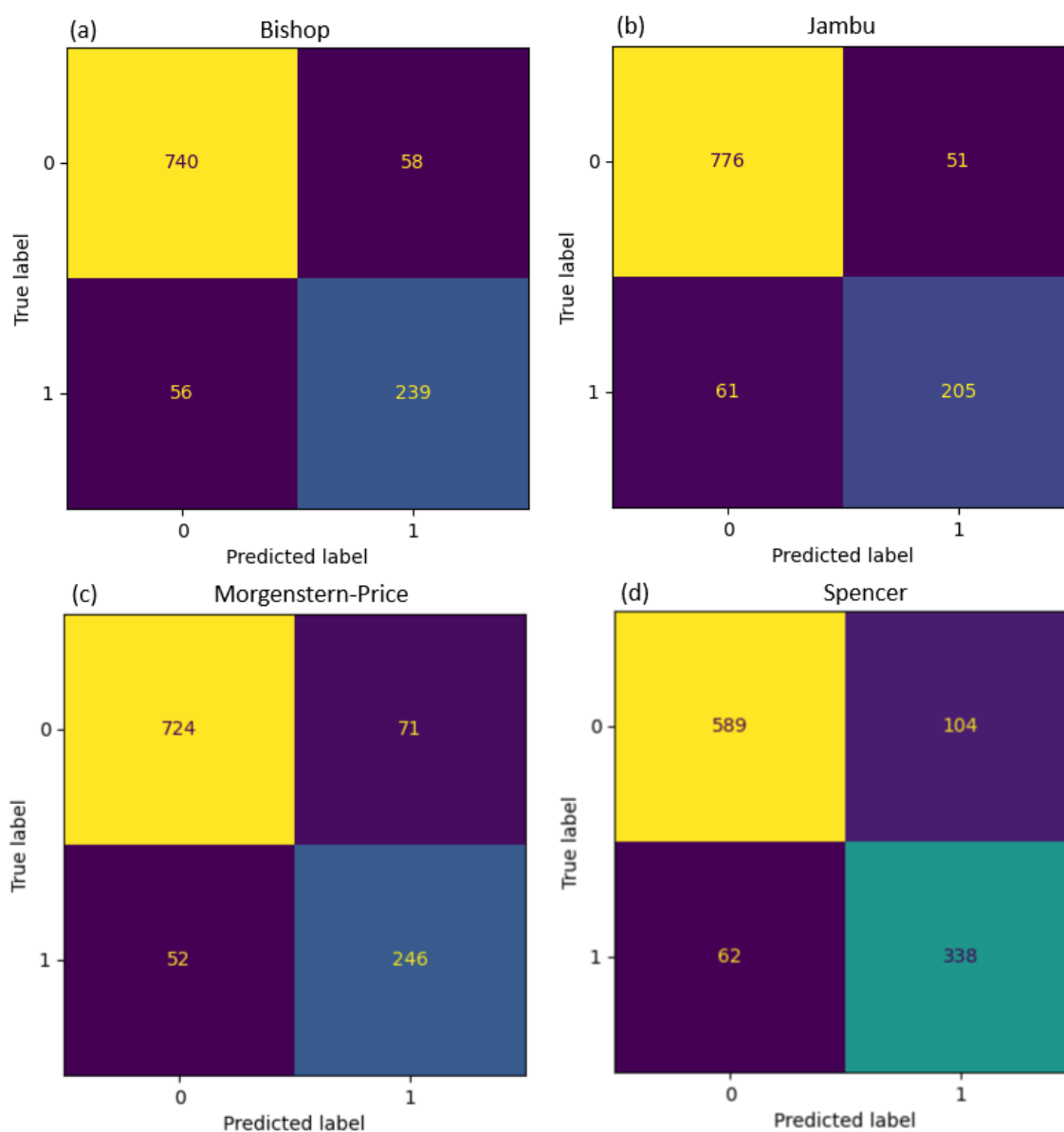
5.2.6 Análise da Matriz de Confusão

Da mesma forma, a análise das matrizes de confusão (Figura 26) permite avaliar comportamento dos modelos na classificação binária dos taludes, identificando separadamente os verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN), o que complementa as métricas anteriores, permitindo compreender o tipo de erro predominante em cada modelo e método de equilíbrio limite.

Para o modelo Árvore de Decisão (Figura 26), observa-se maior dispersão nos erros quando comparado aos demais algoritmos. No método de Bishop, foram obtidos 740 verdadeiros negativos e 239 verdadeiros positivos, com 58 falsos positivos e 56 falsos negativos. Para o método de Janbu, o número de erros

permaneceu elevado, com 51 falsos positivos e 61 falsos negativos. No método de Morgenstern-Price, foram observados 71 falsos positivos e 52 falsos negativos, enquanto no método de Spencer ocorreu o maior número de erros, com 104 falsos positivos e 62 falsos negativos. Esses resultados indicam que a árvore de decisão apresenta maior sensibilidade às variações do conjunto de dados, resultando em maior taxa de classificações incorretas.

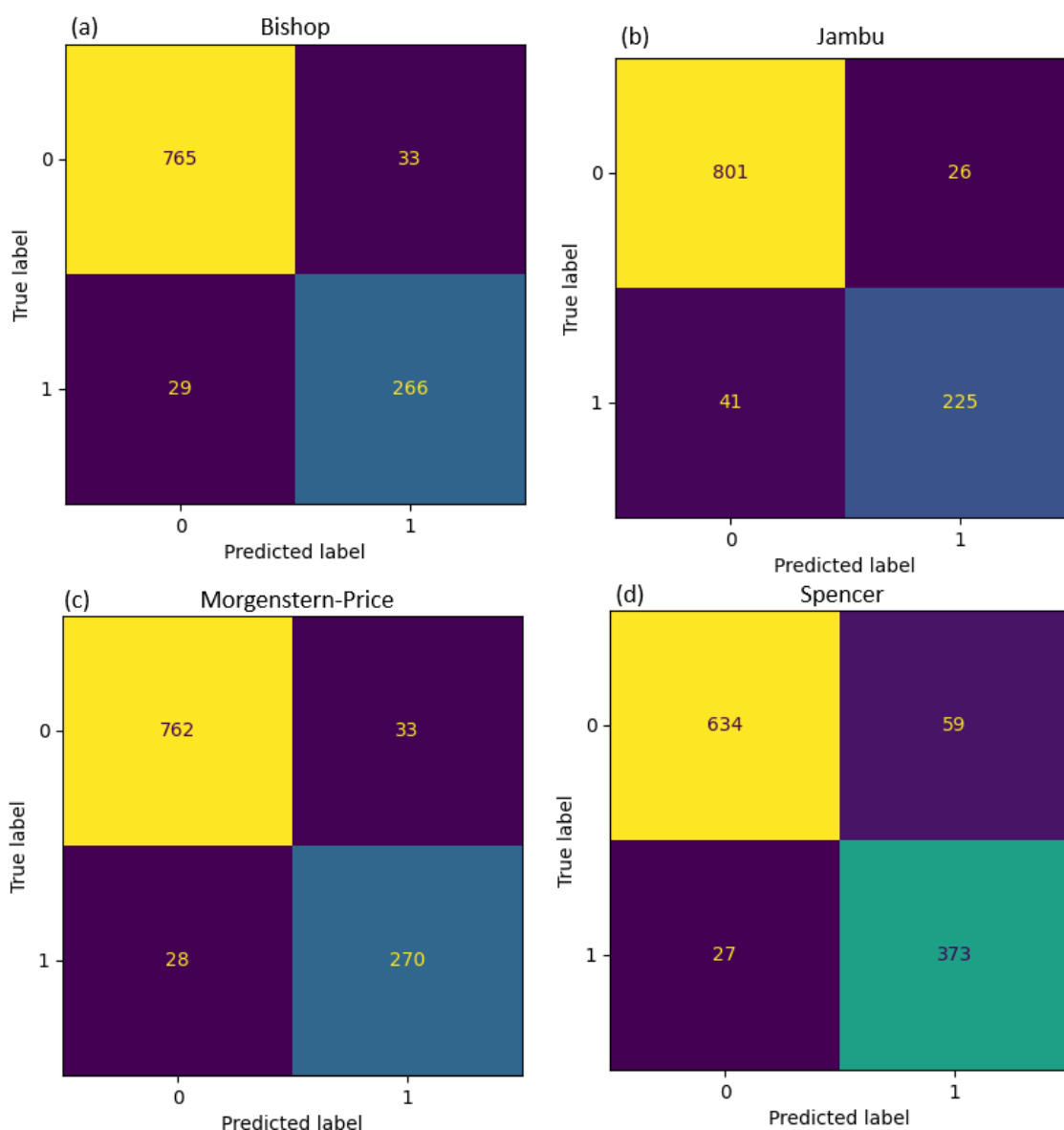
Figura 41: Matrizes de confusão obtidas para Árvore de Decisão: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Para o modelo de Floresta Aleatória, observa-se uma redução significativa

no número de erros. No método de Bishop, foram registrados 765 verdadeiros negativos e 266 verdadeiros positivos, com apenas 33 falsos positivos e 29 falsos negativos. No método de Janbu, o modelo apresentou 26 falsos positivos e 41 falsos negativos. Para o método de Morgenstern–Price, foram observados 33 falsos positivos e 28 falsos negativos, enquanto no método de Spencer ocorreu um leve aumento nos erros, com 59 falsos positivos e 27 falsos negativos. De forma geral, a Floresta Aleatória apresentou bom equilíbrio entre verdadeiros positivos e verdadeiros negativos, mantendo baixa taxa de erros.

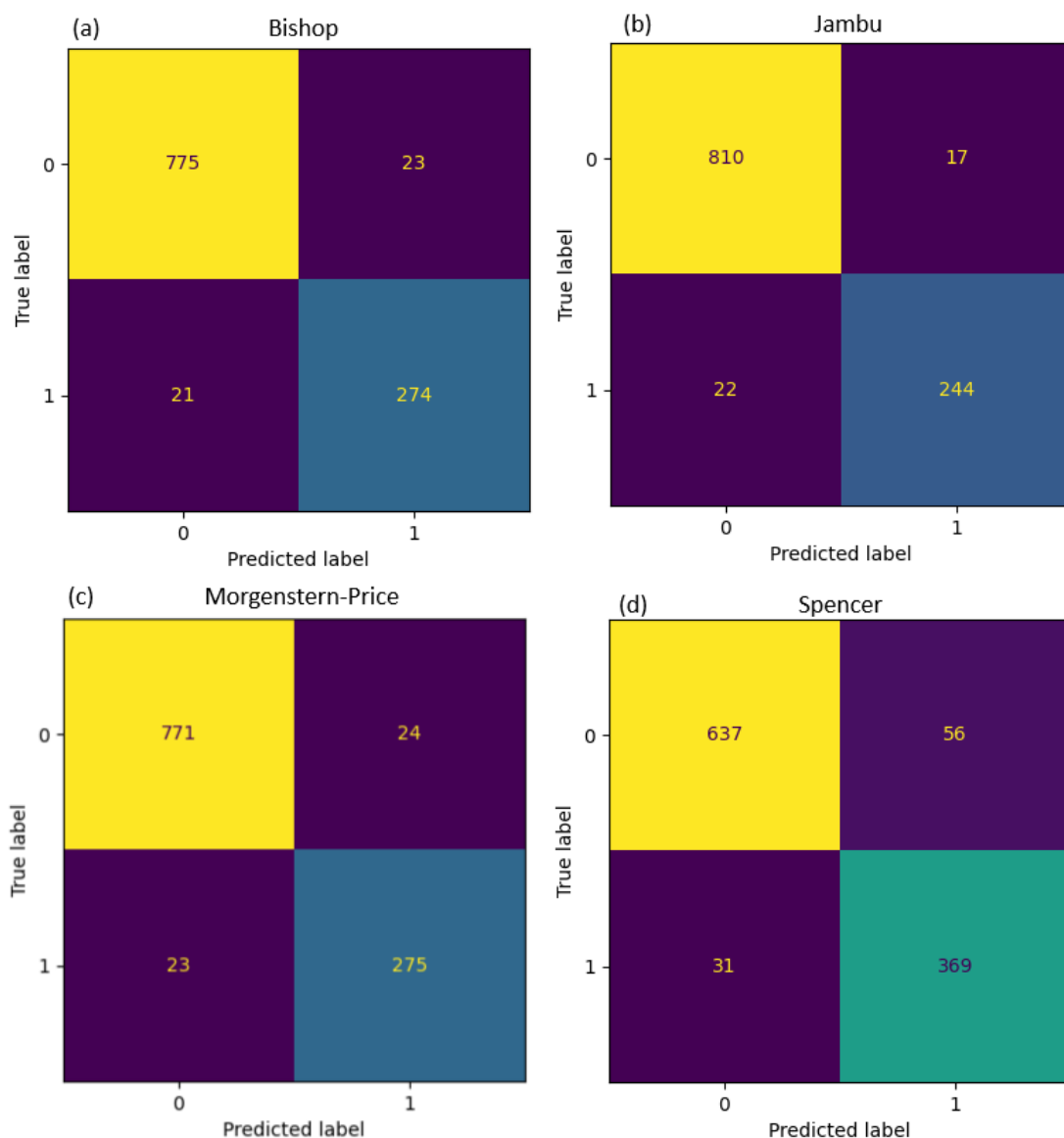
Figura 42: Matrizes de confusão obtidas para Floresta Aleatória: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Para a RNA, o desempenho foi ainda mais consistente, com redução

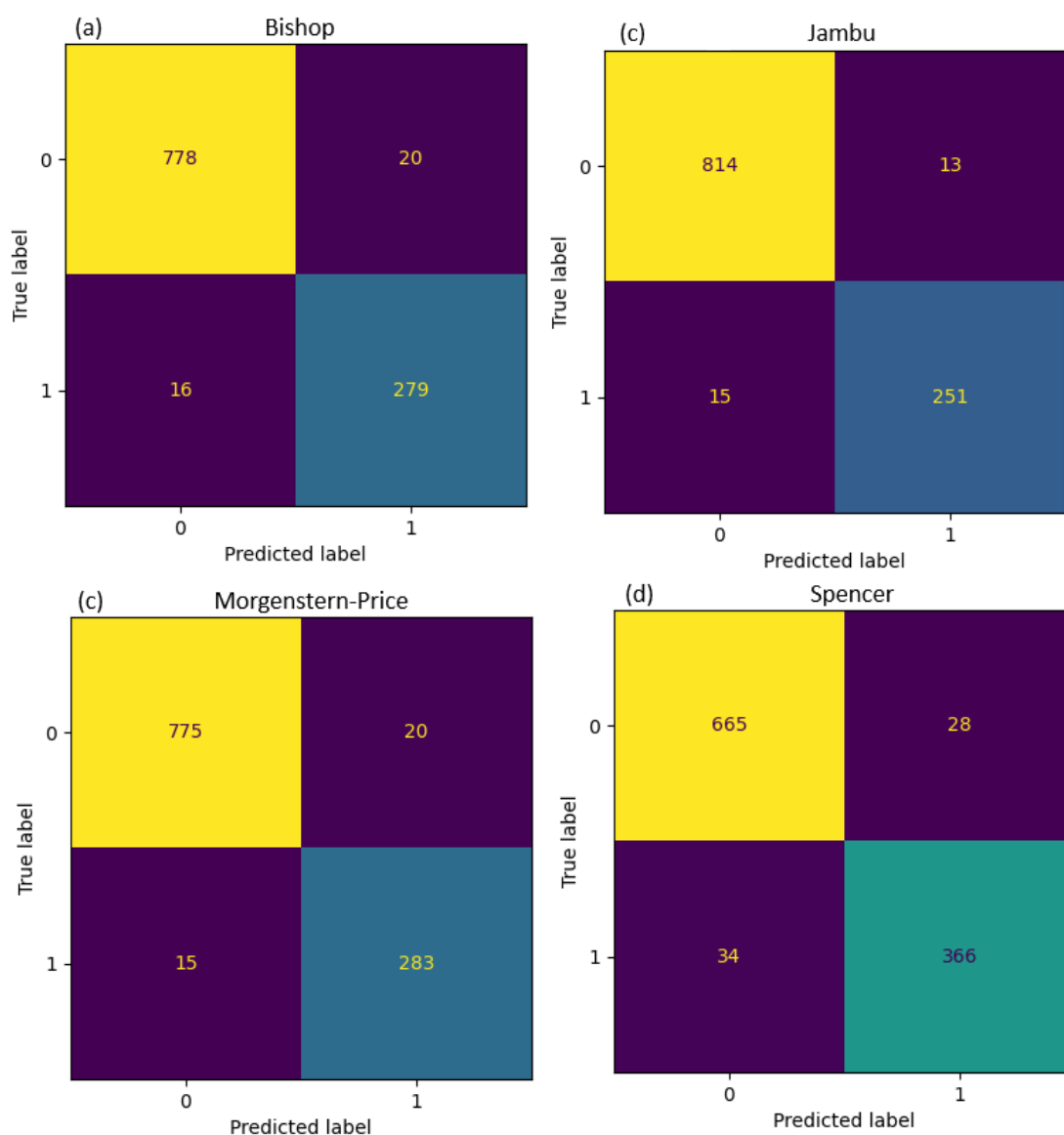
adicional nos erros de classificação. No método de Bishop, foram observados apenas 23 falsos positivos e 21 falsos negativos, com elevado número de classificações corretas. Para o método de Janbu, os erros permaneceram baixos, com 17 falsos positivos e 22 falsos negativos. No método de Morgenstern–Price, foram registrados 24 falsos positivos e 23 falsos negativos, enquanto no método de Spencer ocorreu aumento moderado nos erros, com 56 falsos positivos e 31 falsos negativos. Apesar desse aumento, a RNA manteve elevada capacidade de discriminação entre taludes estáveis e instáveis.

Figura 43: Matrizes de confusão obtidas para RNA: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Para o modelo de SVM, observa-se o melhor desempenho geral entre os algoritmos avaliados. No método de Bishop, foram obtidos 778 verdadeiros negativos e 279 verdadeiros positivos, com apenas 20 falsos positivos e 16 falsos negativos. Para o método de Janbu, o modelo apresentou desempenho ainda mais expressivo, com apenas 13 falsos positivos e 15 falsos negativos, configurando o menor número total de erros entre todos os cenários analisados. No método de Morgenstern–Price, os erros permaneceram reduzidos, com 20 falsos positivos e 15 falsos negativos. Já no método de Spencer observou-se leve aumento nos erros, com 28 falsos positivos e 34 falsos negativos, embora o desempenho geral ainda permaneça superior ao dos demais modelos.

Figura 44: Matrizes de confusão para SVM: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.

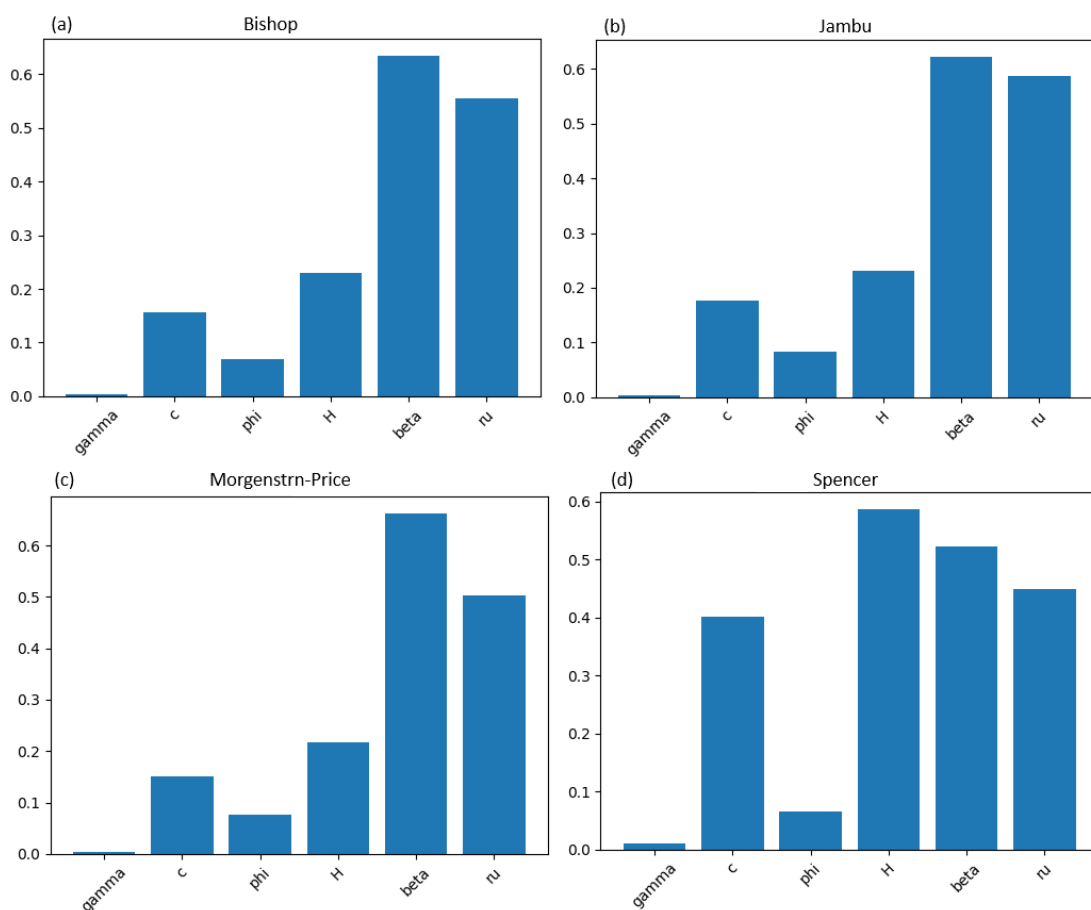


5.2.7 Análise da Importância das Variáveis

A análise da importância das variáveis evidencia diferenças estruturais relevantes entre os métodos de equilíbrio limite, ao mesmo tempo em que revela padrões consistentes entre os modelos de aprendizado de máquina.

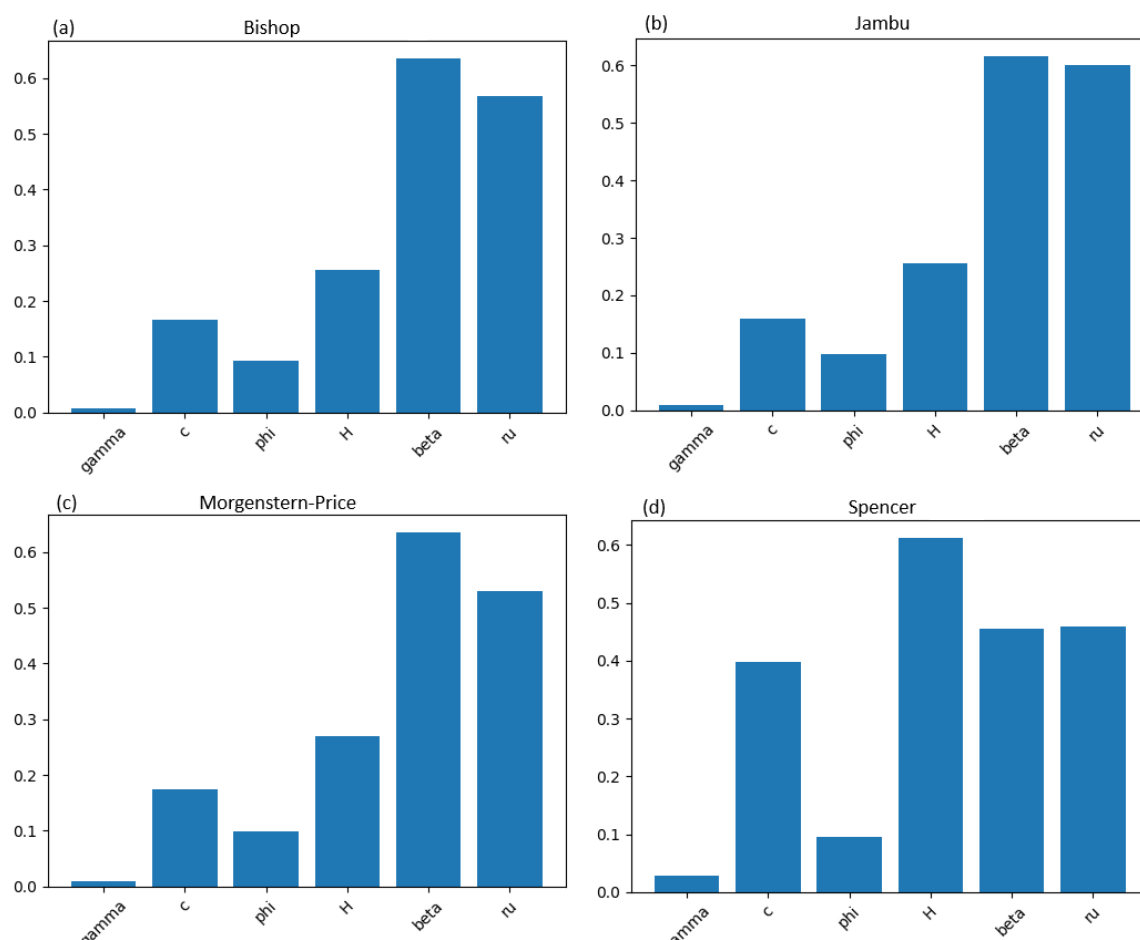
Para a Árvore de Decisão, observa-se uma forte concentração da importância em poucas variáveis dominantes. Nos métodos de Bishop, Janbu e Morgenstern–Price, o ângulo do talude (β) aparece como a variável mais influente, seguido pelo grau de saturação (ru). A altura do talude (H) apresenta importância intermediária, enquanto a coesão (c) e o ângulo de atrito interno (ϕ) possuem influência menor. O peso específico do solo (γ) apresenta contribuição praticamente nula. No método de Spencer, entretanto, ocorre uma mudança no padrão observado, com maior relevância da altura do talude (H), seguida por β e ru , enquanto a coesão passa a apresentar importância relativamente mais elevada em comparação aos demais métodos.

Figura 45: Importância das variáveis para Árvore de Decisão: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Para o modelo de Floresta Aleatória, o padrão geral permanece semelhante, porém com distribuição de importância ligeiramente mais equilibrada entre as variáveis. Nos métodos de Bishop, Janbu e Morgenstern–Price, β e γ continuam sendo as variáveis predominantes, seguidas por H . As variáveis c e ϕ apresentam influência intermediária, enquanto γ permanece com contribuição muito reduzida. No método de Spencer, observa-se novamente maior relevância da altura do talude (H), seguida por c , β e γ , evidenciando maior sensibilidade desse método às propriedades resistentes do solo.

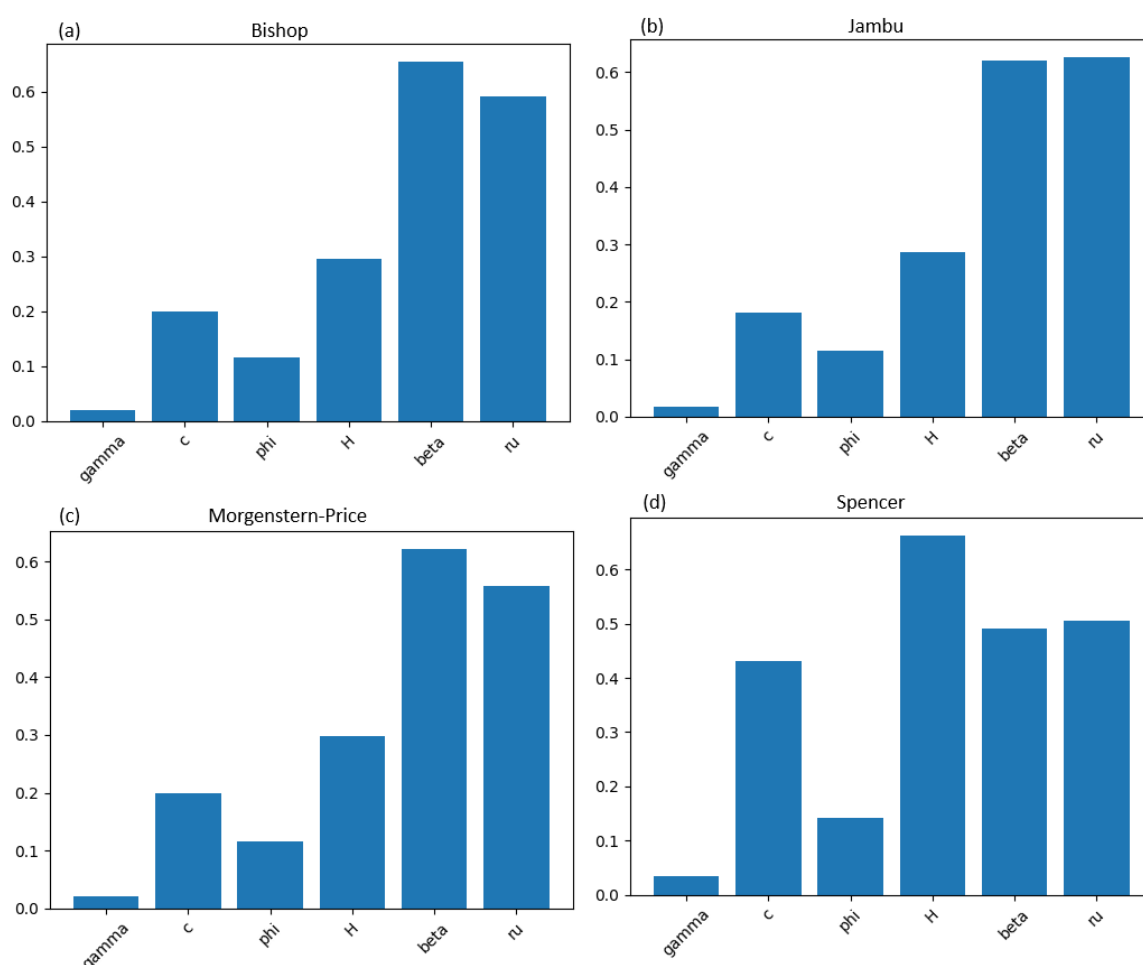
Figura 46: Importância das variáveis para Árvore de Decisão: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Para a RNA, observa-se distribuição mais equilibrada da importância das variáveis, refletindo a capacidade do modelo em capturar relações não lineares mais complexas entre os parâmetros de entrada. Nos métodos de Bishop, Janbu e Morgenstern–Price, β e γ continuam apresentando maior relevância, porém H

passa a assumir papel mais significativo na explicação do fator de segurança. As variáveis c e ϕ mantêm importância intermediária, enquanto γ permanece com contribuição mínima. No método de Spencer, novamente se destaca a maior influência de H , seguida por β e ru , indicando que a geometria e a altura do talude exercem papel mais expressivo nesse método.

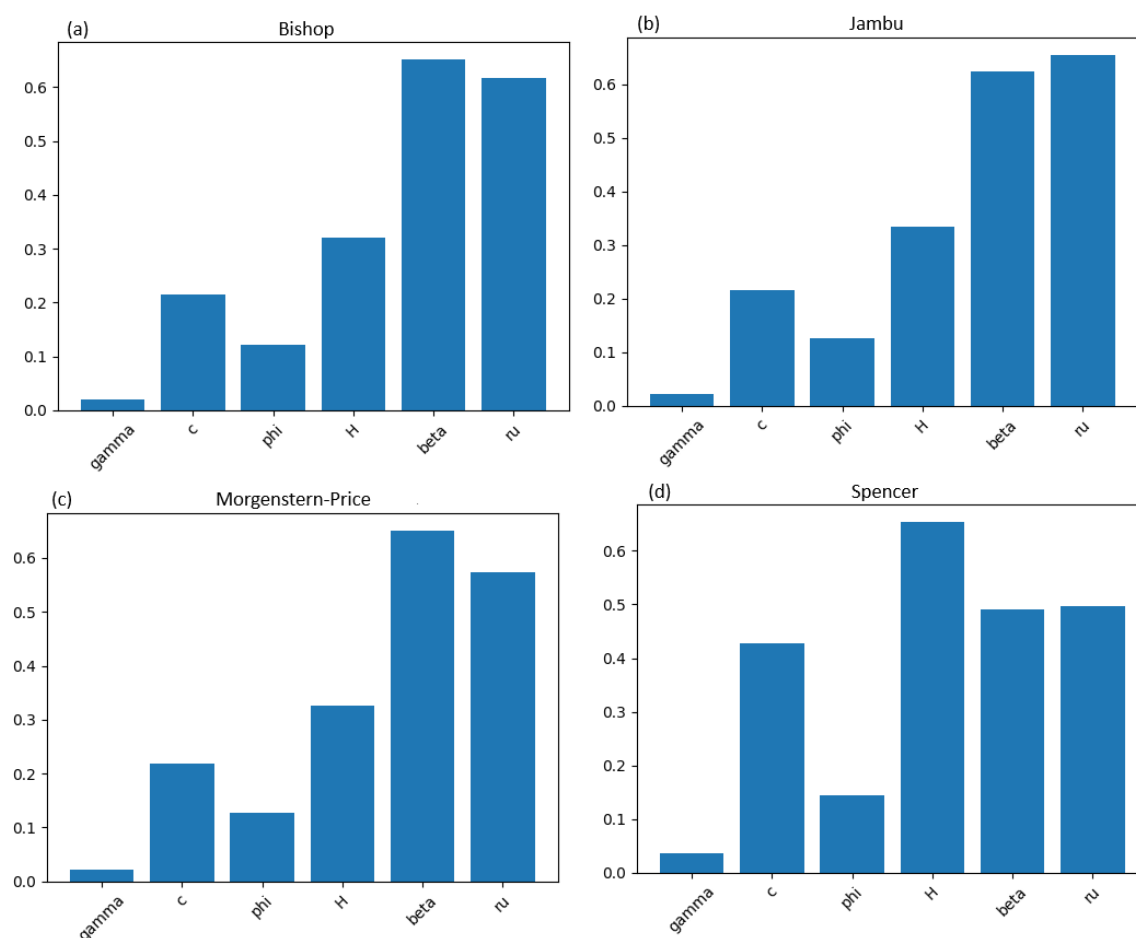
Figura 47 : Importância das Variáveis para RNA: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Para o modelo de SVM, observa-se comportamento semelhante ao da RNA, com distribuição relativamente equilibrada entre as variáveis mais relevantes. Nos métodos de Bishop, Janbu e Morgenstern-Price, β e ru apresentam as maiores contribuições, seguidas por H . As variáveis c e ϕ apresentam importância intermediária, enquanto γ mantém contribuição muito baixa. No método de Spencer, destaca-se novamente a predominância de H como variável mais influente, seguida por β e ru , evidenciando maior sensibilidade do método à

geometria e à altura do talude.

Figura 48: Importância das Variáveis para SVM: (a) Bishop; (b) Janbu; (c) Morgenstern-Price; (d) Spencer.



Esses resultados estão em consonância com estudos da literatura, como o de Hulagabali et al. (2019), que demonstraram, por meio de simulações numéricas, que o FS é altamente sensível às propriedades de resistência do solo (coesão e ângulo de atrito interno) e à geometria do talude, especialmente altura e inclinação. Da mesma forma, os autores observaram que o peso específico exerce influência secundária no comportamento de ruptura, corroborando os achados obtidos neste trabalho.

A predominância das variáveis geométricas (β e H) e do grau de saturação (ru) pode ser explicada pelo fato de que esses parâmetros atuam diretamente no equilíbrio entre forças solicitantes e resistentes. O aumento da inclinação e da altura do talude eleva as tensões atuantes e favorece a instabilidade, enquanto o aumento do grau de saturação reduz a resistência efetiva do solo, diminuindo o FS. Já os

parâmetros de resistência (c e φ), embora fundamentais do ponto de vista físico, apresentaram menor destaque relativo no conjunto analisado, possivelmente em função da faixa de variação adotada e da interação com as variáveis geométricas.

Além disso, observa-se que modelos mais complexos, como SVM e RNA, apresentam distribuição de importância mais equilibrada entre as variáveis, enquanto Árvore de Decisão e Floresta Aleatória tendem a concentrar maior relevância em poucos parâmetros dominantes, o que está de acordo com o comportamento observado anteriormente nas métricas de desempenho e nas matrizes de confusão.

5.2.8 Desempenho dos modelos de aprendizado de máquina sob a perspectiva geotécnica

A avaliação do desempenho dos modelos de aprendizado de máquina na estimativa do Fator de Segurança (FS) não deve ser interpretada apenas sob a ótica das métricas estatísticas, mas também à luz do comportamento físico do problema geotécnico envolvido. O FS é uma grandeza diretamente relacionada ao equilíbrio entre forças resistentes e solicitantes ao longo de uma superfície potencial de ruptura, sendo fortemente influenciado por parâmetros como inclinação do talude, resistência ao cisalhamento do solo e condições de saturação.

5.2.8.1 Interpretação física do desempenho dos modelos

Os modelos mais robustos, como a Máquina de Vetores de Suporte (SVM) e a Rede Neural Artificial (RNA), apresentaram melhor desempenho geral, o que pode ser explicado pela natureza não linear do problema geotécnico. A relação entre os parâmetros de entrada e o fator de segurança não é linear, especialmente devido à interação entre resistência ao cisalhamento (c e φ), geometria do talude (H e β) e poropressão (ru).

A SVM, ao utilizar funções kernel, é capaz de mapear essas relações não lineares em espaços de maior dimensão, capturando com maior precisão o comportamento físico do sistema. De forma semelhante, as redes neurais artificiais conseguem modelar interações complexas entre variáveis, o que é coerente com o fato de que a estabilidade de taludes resulta da combinação simultânea de múltiplos fatores geotécnicos.

Por outro lado, a Árvore de Decisão apresentou desempenho inferior, o que

pode ser interpretado a partir de sua estrutura baseada em regras discretas. Esse tipo de modelo tende a fragmentar o espaço de variáveis, dificultando a representação de fenômenos contínuos, como a variação progressiva do fator de segurança em função de pequenas alterações nos parâmetros do solo. Na prática geotécnica, o FS não varia de forma abrupta, mas sim contínua, o que explica a limitação desse modelo.

A Floresta Aleatória, por sua vez, apresentou desempenho intermediário, combinando múltiplas árvores e reduzindo a variabilidade dos resultados. Esse comportamento reflete uma melhor capacidade de capturar tendências gerais do problema, embora ainda limitada na representação de interações altamente não lineares.

5.2.8.2 Influência dos métodos de equilíbrio limite

Os resultados também evidenciam que o desempenho dos modelos está diretamente relacionado às características mecânicas dos métodos de equilíbrio limite utilizados na geração do fator de segurança.

Nos métodos de Bishop e Janbu, observou-se melhor desempenho dos modelos. Isso pode ser explicado pelo fato de que esses métodos adotam simplificações no equilíbrio de forças, resultando em funções de resposta mais suaves e previsíveis. Do ponto de vista físico, essas simplificações reduzem a sensibilidade a variações locais e produzem um comportamento mais regular do FS, facilitando o aprendizado pelos algoritmos.

Em contraste, o método de Spencer, e em menor grau, o método de Morgenstern-Price apresentaram maior dificuldade de ajuste para todos os modelos. Esse comportamento está associado ao fato de que métodos rigorosos satisfazem simultaneamente o equilíbrio de forças e momentos, exigindo soluções iterativas mais complexas. Como consequência, pequenas variações nos parâmetros podem resultar em mudanças mais significativas no FS, aumentando a variabilidade dos dados.

Essas diferenças estão diretamente associadas às bases mecânicas dos métodos de equilíbrio limite. Conforme discutido por Duncan, Wright e Brandon (2014), o cálculo do fator de segurança constitui um problema estaticamente indeterminado, exigindo a adoção de hipóteses simplificadoras distintas em cada método. Métodos como Bishop tendem a produzir respostas mais estáveis e menos sensíveis a variações locais, favorecendo o aprendizado dos modelos. Por outro

lado, métodos mais completos, como Spencer e Morgenstern–Price, envolvem processos iterativos mais complexos e maior sensibilidade às forças entre fatias, o que pode introduzir maior dispersão nos resultados.

5.3 Limitações do modelo e análise do uso do aprendizado de máquina para estabilidade de taludes

A aplicação de algoritmos de inteligência artificial demonstrou ser uma alternativa robusta para a análise de estabilidade de taludes, apresentando elevada capacidade de generalização e precisão na estimativa do fator de segurança (FS), para o conjunto de dados utilizado. Os resultados indicaram que modelos como Máquina de Vetores

de Suporte (SVM) e Redes Neurais Artificiais (RNA) alcançaram os melhores desempenhos, com valores de coeficiente de determinação (R^2) maiores que 90%, Erro Quadrático Médio menor que 0,01, Acurácia maior que 96% e Área da Curva ROC próximos de 0,97 e 0,98, o que evidencia uma excelente capacidade de prever o valor numérico do FS e um desempenho ainda mais significativo ao distinguir entre taludes estáveis e instáveis.

Apesar do sucesso dos modelos, o trabalho apresenta limitações importantes, como a dependência exclusiva de dados simulados, que podem não captar integralmente a heterogeneidade e as incertezas de maciços de solo reais. Além disso, a análise focou em geometrias de taludes relativamente simples e condições de saturação uniformes, o que pode limitar a aplicabilidade direta em cenários geológicos complexos ou sob regimes de chuva variáveis.

Embora o uso de dados simulados seja uma prática comum em estudos de engenharia geotécnica, especialmente devido à escassez de bases de dados reais, essa abordagem apresenta restrições quanto à representatividade e à capacidade de generalização dos modelos de aprendizado de máquina.

Os dados sintéticos são gerados a partir de distribuições previamente definidas e hipóteses simplificadoras, não incorporando ruídos experimentais, incertezas de medição e variabilidade natural típicos de dados obtidos em campo. Conforme destacado por Caballar e Stryker (2022), a qualidade dos dados sintéticos está diretamente relacionada às premissas adotadas em sua geração, podendo não refletir completamente a complexidade dos sistemas reais. No contexto geotécnico, isso implica que fatores como heterogeneidade do solo, anisotropia e variações hidrogeológicas podem não estar adequadamente

representados. Além disso, a variabilidade dos dados sintéticos tende a ser limitada pelos intervalos paramétricos definidos durante sua construção.

A homogeneidade dos dados também pode introduzir viés nos modelos de aprendizado de máquina. Algoritmos treinados em bases com baixa variabilidade tendem a apresentar desempenho elevado em validações internas, mas podem sofrer degradação significativa quando aplicados a dados reais. Conforme discutido por Dhruva Karir et al. (2022), o desempenho de modelos de machine learning em estabilidade de taludes depende fortemente da qualidade e diversidade dos dados de entrada, podendo variar significativamente em diferentes condições geotécnicas.

Outro aspecto crítico refere-se à capacidade de extrapolação dos modelos. Como os dados utilizados neste estudo são limitados a um domínio sintético específico, a aplicação dos modelos a cenários fora dessas faixas pode resultar em previsões não confiáveis. Em situações reais, onde a variabilidade dos parâmetros geotécnicos é significativamente maior, essa limitação pode comprometer a aplicabilidade prática dos modelos desenvolvidos.

Adicionalmente, destaca-se a ausência de validação com dados reais como uma limitação importante do presente estudo. Trabalhos da literatura frequentemente utilizam bases de dados reais, ainda que com número reduzido de amostras, para validar a robustez dos modelos. Por exemplo, Yaser A. Nanehkaran et al. (2022) utilizaram dados de taludes reais para avaliar modelos preditivos, enquanto Yadav et al. (2025) empregaram conjuntos reais para validação de modelos de aprendizado de máquina aplicados à estabilidade de taludes. Esses estudos evidenciam a importância da validação externa para garantir maior confiabilidade dos resultados.

Diante dessas limitações, recomenda-se que trabalhos futuros incorporem dados reais de campo ou de laboratório, possibilitando a calibração e validação dos modelos desenvolvidos. A combinação de dados sintéticos e reais pode contribuir para aumentar a variabilidade do conjunto de dados, reduzir vieses e melhorar a capacidade de generalização dos modelos.

Recomenda-se também a exploração de técnicas de *Deep Learning* para lidar com geometrias de ruptura não circulares e a análise da influência de diferentes funções de *kernel* no SVM para melhorar a predição em métodos rigorosos como Spencer e Morgenstern-Price. A integração desses modelos em plataformas de suporte à decisão pode otimizar projetos geotécnicos, permitindo avaliações de segurança mais rápidas e confiáveis.

Por fim, ressalta-se que os resultados obtidos neste estudo devem ser interpretados dentro do contexto do conjunto de dados utilizado, sendo mais adequados para análise comparativa entre algoritmos do que para aplicação direta em projetos geotécnicos reais sem validação adicional.

6 CONCLUSÕES

O presente trabalho permitiu avaliar a eficácia da aplicação de algoritmos de aprendizado de máquina (*machine learning*) na estimativa do fator de segurança (FS) de taludes, consolidando estas técnicas como ferramentas complementares robustas para a engenharia geotécnica. A utilização de dados simulados por softwares especializados, como o HYRCAN, constituiu uma base sólida para o treino de modelos preditivos, desde que assegurada uma distribuição equilibrada dos parâmetros geotécnicos e geométricos, o que evitou enviesamentos e promoveu a capacidade de generalização dos algoritmos.

Do ponto de vista metodológico, destaca-se que os modelos de aprendizado de máquina desenvolvidos neste trabalho foram treinados com dados gerados a partir de métodos de equilíbrio limite, de modo que suas previsões refletem, essencialmente, os padrões e hipóteses desses próprios métodos. Assim, o elevado desempenho observado não implica necessariamente na descoberta de novas relações físicas do fenômeno de ruptura, mas sim na capacidade dos algoritmos de reproduzir, de forma eficiente, a lógica implícita nos dados simulados.

Nesse sentido, a principal contribuição do aprendizado de máquina neste estudo está associada ao ganho operacional, especialmente na redução do tempo computacional e na viabilização de análises rápidas para múltiplos cenários. Entretanto, suas limitações estão diretamente ligadas às simplificações dos métodos que originaram os dados, o que indica que sua aplicação na engenharia deve ocorrer de forma complementar e não substitutiva, sempre acompanhada de análise crítica e validação geotécnica.

6.1 Desempenho dos modelos de aprendizado de máquina

Os resultados evidenciaram uma distinção clara no desempenho dos diferentes algoritmos testados. A Máquina de Vetores de Suporte (SVM), obtiveram o melhor desempenho tanto em capacidade de previsão, quanto na capacidade de generalização, alcançando coeficientes de determinação superiores a 0,90 e erros quadráticos médios inferiores a 0,01, e *CV-score* de até 0,879. Na tarefa de classificação binária entre taludes estáveis e instáveis, estes modelos atingiram uma acurácia superior a 96% e valores de AUC próximos de 0,98, demonstrando uma capacidade excepcional de distinção, mesmo perante a complexidade dos dados.

Por outro lado, embora a Árvore de Decisão ofereça maior transparência e facilidade de interpretação, revelou-se o modelo mais sensível e menos estável, apresentando um maior número de erros, com erro quadrático médio entre 0,21 e 0,32 e coeficiente de determinação entre 0,68 e 0,77, assim com *CV-score* entre 0,60 e 0,70, indicando uma menor capacidade de generalização.

6.2 Desempenho dos métodos de equilíbrio limite

Ao avaliar a importância das variáveis de entrada dos modelos, observou-se que a inclinação do talude e o grau de saturação (ru) foram as variáveis com maior impacto sobre o FS nos métodos Bishop, Janbu e Morgenstern–Price, seguidas pela altura, coesão e ângulo de atrito; o peso específico (γ) mostrou influência muito menor no conjunto estudado

Uma das principais contribuições foi a análise da influência do método de equilíbrio limite utilizado na qualidade da predição. Verificou-se que modelos treinados com dados do método de Bishop apresentou métricas de desempenho superiores, seguido por Janbu. Em contrapartida, o método de Spencer apresentou um maior nível de ruído e variabilidade no Fator de Segurança, o que resultou numa redução da capacidade de generalização dos algoritmos e em maiores erros na regressão contínua. Esta discrepância se explica no rigor mecânico dos métodos completos, que exigem soluções iterativas e são mais sensíveis a condições de contorno, tornando a função de resposta mais difícil de ser aprendida pelas máquinas.

Ao avaliar a importância das variáveis de entrada dos modelos, observou-se que a inclinação do talude e o grau de saturação (ru) foram as variáveis com maior impacto sobre o Fator de Segurança nos métodos Bishop, Janbu e Morgenstern–Price, seguidas pela altura, coesão e ângulo de atrito; o peso específico (γ) mostrou influência muito menor no conjunto estudado.

6.3 Limitações e recomendações para trabalhos futuros

Apesar dos resultados promissores, este estudo apresenta limitações importantes relacionadas principalmente à natureza dos dados utilizados. A utilização de dados sintéticos implica menor variabilidade e ausência de heterogeneidades típicas de condições reais, como estratificação do solo, anisotropia e variações hidrogeológicas. Essa limitação pode introduzir vies nos

modelos e reduzir sua capacidade de generalização quando aplicados a cenários reais mais complexos.

Dessa forma, conclui-se que o uso do aprendizado de máquina representa uma abordagem promissora para a engenharia geotécnica, especialmente como ferramenta complementar de apoio à tomada de decisão. No entanto, sua aplicação prática deve ser realizada com cautela, sendo indispensável a validação com dados reais.

Para trabalhos futuros, recomenda-se a ampliação do banco de dados com a inclusão de dados de campo e resultados experimentais, permitindo maior variabilidade e representatividade dos cenários analisados. Sugere-se também a investigação de geometrias mais complexas e condições geotécnicas heterogêneas, bem como a incorporação de modelos constitutivos mais avançados. Por fim, recomenda-se a exploração de técnicas de aprendizado mais sofisticadas, como modelos de Deep Learning e métodos ensemble avançados, avaliando o equilíbrio entre aumento de complexidade computacional e ganhos efetivos de desempenho.

REFERÊNCIAS

ALOK, Aman et al. **A Generalized Limit Equilibrium-Based Platform Incorporating Simplified Bishop, Janbu and Morgenstern–Price Methods for Soil Slope Stability Problems**. *Advances in Civil Engineering*, 2024.

ALPAYDIN, Ethem. **Introduction to machine learning**. Cambridge, Massachusetts; London, England: The MIT Press, 2004.

ARYAL, Milan; ACHARYA, Indra Prasad. **Machine learning based unified framework for slope stability prediction**. Proceedings of 12th IOE Graduate Conference, v. 12, p. [s.p.], out. 2022. ISSN 2350-8914 (Online), 2350-8906 (Print).

AZMOON, Behnam; BINIYAZ, Aynaz; LIU, Zhen (Leo). **Evaluation of deep learning against conventional limit equilibrium methods for slope stability analysis**. *Applied Sciences*, Basel, v. 11, n. 13, p. 6060, 2021. DOI: <https://doi.org/10.3390/app11136060>.

BURKOV, Andriy. **The Hundred-Page Machine Learning Book**. [S.l.]: Andriy Burkov, 2019.

CABALLAR, Rina Diane; STRYKER, Cole. **Oito melhores práticas para geração de dados sintéticos**. IBM Think, 2022. Disponível em: www.ibm.com/br-pt/think/insights/synthetic-data-eneration. Acesso em: 28 mar. 2026.

CRISTIANINI, Nello; SHAWE-TAYLOR, John. **An introduction to support vector machines and other kernel-based learning methods**. Cambridge: Cambridge University Press, 2000.

DANZIGER, Bernadete Ragoni; GERSCOVICH, Denise M. S.; SARAMAGO, Robson. **Conteções: teoria e aplicações em obras**. São Paulo: Oficina de Textos, 2016.

DAS, Braja M. **Fundamentos de engenharia geotécnica**. Tradução EZ2Translate; revisão técnica Leonardo R. Miranda. São Paulo: Cengage Learning, 2011.

DIETTERICH, Thomas G. **Machine learning**. Corvallis: Oregon State University, Department of Computer Science. Disponível em: <https://web.engr.oregonstate.edu/~tgd/publications/nature-ecs-machine-learning.pdf>. Acesso em: 8 mar. 2026.

DUNCAN, J. Michael; WRIGHT, Stephen G.; BRANDON, Thomas L. **Soil strength and slope stability**. Publicado por John Wiley & Sons, Inc., Hoboken, EUA, 2014

EDELL, Laura. *Is Machine Learning the New EPM Black? Musings on ML*,

Deep Learning & AI, 9 dez. 2015. Disponível em: <https://scorecardstreet.wordpress.com/2015/12/09/is-machine-learning-the-new-epm-black/>. Acesso em: 25 jan. 2026.

ERNESTO, Leonardo; HENRIQUE, Samuel; POLO, Felipe Maia; FALEIROS, Matheus. **Guia de estudos: Data Science e Machine Learning**. Ribeirão Preto: Grupo de Estudos em Data Science Neuron/USP, 2018. Disponível em: <https://pdfcoffee.com/machine-learning--pdf-free.html>. Acesso em: 10 mar. 2026.

GERSCOVICH, Denise M. S. **Estabilidade de taludes**. 2. ed. São Paulo: Oficina de Textos, 2016.

GUIDO, Sarah; MÜLLER, Andreas C. **Introduction to machine learning with Python**. Sebastopol, CA: O'Reilly Media, 2016.

GONZÁLEZ, S.; GARCÍA, S.; DEL SER, J.; ROKACH, L.; HERRERA, F. **A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities**. *Information Fusion*, v. 64, p. 205–237, 2020. DOI: 10.1016/j.inffus.2020.07.007.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The elements of statistical learning: data mining, inference, and prediction**. 2. ed. New York: Springer, 2009. 767 p. ISBN 978-0387848570.

KARIR, Dhruva et al. **Stability prediction of a natural and man-made slope using various machine learning algorithms**. *Transportation Geotechnics*, v. 32, 100745, 2022. Disponível em: <https://doi.org/10.1016/j.trgeo.2022.100745>. Acesso em: maio 2025.

LORENA, Ana Carolina; CARVALHO, André C. P. L. F. de. **Uma introdução às Support Vector Machines**. *Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 43-68, dez. 2007.

MASSAD, Faïçal. **Curso básico de geotecnia – com exercícios resolvidos**. 2. ed. São Paulo: Oficina de Textos, 2010. 216 p. ISBN 978-85-86238-97-0.

NANEHKARAN, Yaser A. et al. **Comparative Analysis for Slope Stability by Using Machine Learning Methods**. *Applied Sciences*, v. 13, n. 3, p. 1555, 2023. Disponível em: <https://doi.org/10.3390/app13031555>. Acesso em: maio 2025.

PINTO, Carlos de Sousa. **Curso básico de medicina dos solos** em 16 aulas. 3. ed. São Paulo: Oficina de Textos, 2006.

SCHLEDER, Gabriel R.; FAZZIO, Adalberto. **Machine learning na física**,

química e ciência de materiais: descoberta e design de materiais. *Revista Brasileira de Ensino de Física*, São Paulo, v. 43, supl. 1, e20200407, 2021. DOI: <https://doi.org/10.1590/1806-9126-RBEF-2020-0407>

SHEODE, Mohit; ANURAGI, Saurabh K.; DHARAVATH, Kishan. **Slope stability analysis using hybrid BPSO-SVM machine learning techniques.** *Journal of Management & Technology*, v. 25, n. 5, p. 54–76, 2025.

SHEPPARD, Clinton. **Tree-based machine learning algorithms: decision trees, random forests, and boosting.** [S.l.]: Createspace Independent Pub, 2017. 152 p.

SMITH, C.; McGUIRE, B.; HUANG, T.; YANG, G. **The history of artificial intelligence.** University of Washington: CSEP 590A - History of Computing, 2006. Disponível em: <https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf>. Acesso em: junho 2025.

TUN, Soe Hlaing; ZENG, Changnv; JAMIL, Farhad. **Prediction of slope stability based on five machine learning techniques approaches: a comparative study.** *Multiscale and Multidisciplinary Modeling, Experiments and Design*, v. 8, p. 224, 2025. DOI: <https://doi.org/10.1007/s41939-025-00808-0>

VELASQUEZ, Luis Henrique. **Uma visão geral sobre machine learning: Classificação.** *Statplace*, 13 ago. 2020. Disponível em: <https://statplace.com.br/blog/uma-visao-geral-sobre-machine-learning/>. Acesso em: 25 jan. 2026.

YADAV, Devendra Kumar et al. **Enhanced slope stability prediction using ensemble machine learning techniques.** *Scientific Reports*, v. 15, n. 7302, 2025. Disponível em: <https://doi.org/10.1038/s41598-025-90539-6>. Acesso em: maio 2025.