Identificação de Spoofing em GNSS de Aeronave Utilizando Machine Learning

Gabriel B. Pinheiro[1]

[1]Instituto Hardware BR
Brasília– DF – Brazil

gabrielbopi@gmail.com

Abstract. *The vulnerability of GNSS systems to spoofing attacks threatens aerospace security, especially in UAVs. This study develops a machine learning-based solution to detect and classify these interferences, analyzing 13 signal metrics from GNSS receivers. Using a dataset, the performance of eight algorithms was evaluated. The XGBoost model stood out with an accuracy exceeding 94%, demonstrating effectiveness in identifying sophisticated attacks and enabling robust defense in embedded hardware.*

Resumo. *A vulnerabilidade dos sistemas GNSS a ataques de Spoofing ameaça a segurança aeroespacial, especialmente em VANTs. Este estudo desenvolve uma solução baseada em Machine Learning para detetar e classificar essas interferências, analisando 13 métricas de sinal de receptores GNSS. Com um dataset, avaliou-se o desempenho de oito algoritmos. O modelo XGBoost destacou-se com acurácia superior a 94%, demonstrando eficácia na identificação de ataques sofisticados e viabilizando uma defesa robusta em hardware embarcado.*

1. Introdução

A integridade e a fiabilidade dos Sistemas Globais de Navegação por Satélite (GNSS), com destaque para o Sistema de Posicionamento Global (GPS), são fundamentais para a infraestrutura moderna. Estes sistemas são a espinha dorsal de operações críticas que variam desde a sincronização de redes de energia e telecomunicações até à navegação precisa na aviação comercial, militar e em Veículos Aéreos Não Tripulados (VANTs) [1, 2]. No entanto, a arquitetura aberta e não encriptada dos sinais civis de GNSS torna-os inerentemente vulneráveis a interferências intencionais, representando um risco significativo para a segurança e a eficiência operacional nestes setores.

1.1. Definições

Entre as ameaças mais sofisticadas e perigosas ao GNSS encontra-se o ataque de Spoofing (falsificação). Diferente do Jamming, que visa negar o serviço através da saturação do espectro com ruído, o Spoofing é um ataque de engano. O atacante transmite sinais falsos que mimetizam a estrutura dos sinais de satélite autênticos, com o objetivo de capturar os loops de rastreamento do receptor da vítima [3, 4].

Uma vez que o receptor "tranca" no sinal falso, o atacante pode manipular os dados de navegação de forma controlada. Isso pode resultar no cálculo de uma posição geográfica incorreta, numa velocidade falsa ou num desvio temporal, sem que o sistema de navegação ou o operador percebam a anomalia imediatamente [4]. Os efeitos nos dados recebidos podem variar desde desvios sutis na trajetória (ataques do tipo carry-off) até saltos abruptos de posição ou tempo, dependendo da sofisticação do

ataque [5]. Ataques mais complexos, como os realizados por receiver-spoofers, podem alinhar perfeitamente a fase do código e a fase da portadora com os sinais reais antes de assumir o controle, tornando a detecção extremamente desafiadora para receptores convencionais [3].

1.2. Justificativa e Relevância

A implementação de um sistema robusto de deteção de spoofing é uma necessidade premente, especialmente no setor aeroespacial. Na aviação comercial e militar, a dependência do GNSS para fases críticas do voo, como a aproximação e a aterragem (especialmente em sistemas baseados em desempenho, como o RNP), significa que informações de posição incorretas podem levar a consequências catastróficas, incluindo o desvio de rotas seguras, violação de espaços aéreos restritos ou colisões com o terreno [1, 2].

No contexto dos VANTs e drones, que frequentemente operam com elevado grau de autonomia, a vulnerabilidade é ainda mais crítica. Um ataque de spoofing bem-sucedido pode não apenas desviar o veículo da sua missão, mas também permitir que um atacante assuma o controle efetivo da aeronave, forçando-a a aterrar em local não autorizado ou a colidir intencionalmente [6]. A integração destes sistemas, desde o segmento espacial (satélites) até ao segmento do utilizador (aeronaves), exige uma cadeia de confiança ininterrupta. A proliferação de dispositivos de Rádio Definido por Software (SDR) de baixo custo democratizou o acesso a ferramentas de spoofing, aumentando a probabilidade de ataques tanto por atores estatais quanto por criminosos comuns, o que reforça a urgência de soluções de defesa acessíveis e eficazes [7].

1.3. Objetivos

O objetivo principal deste trabalho é o desenvolvimento e a avaliação de um algoritmo de Aprendizado de Máquina (*Machine Learning* - ML) capaz de identificar ataques de spoofing em receptores GNSS embarcados em aeronaves. O sistema proposto visa analisar as características do sinal recebido e classificar a sua integridade, distinguindo entre sinais legítimos e sinais falsificados, com foco específico na constelação GPS.

1.4. Proposta Técnica

A proposta técnica deste relatório baseia-se na implementação de algoritmos de ML que podem ser integrados nos sistemas de navegação de veículos aéreos como uma camada de software de segurança. Ao monitorizar continuamente métricas de qualidade do sinal e consistência dos dados, o algoritmo poderá reconhecer padrões anômalos associados ao spoofing, evitando que o sistema de navegação utilize dados corrompidos para o cálculo da posição e notificando imediatamente os operadores ou o piloto automático sobre a ameaça [8].

O desenvolvimento e treino destes modelos de ML dependem de dados representativos. Embora *datasets* com dados reais de ataques de spoofing sejam escassos na literatura aberta devido à natureza sensível e ilegal da transmissão de sinais falsos, existem recursos valiosos como o *Texas Spoofing Test Battery* (TEXBAT) [9] e os

dados disponibilizados por Aissou et al. [10]. Além disso, é possível gerar dados próprios através de métodos de simulação avançada utilizando dispositivos de SDR, como o USRP ou HackRF, configurados para emular cenários de ataque controlados [7, 11]. A criação e análise destes cenários exigem profissionais com conhecimento especializado em telecomunicações via satélite e processamento de sinais, destacando a interdisciplinaridade necessária para abordar este problema de segurança cibernética no domínio aeroespacial.

2. Abordagem

A metodologia deste trabalho foi estruturada para desenvolver e validar um sistema eficaz de deteção de ataques de *Spoofing* em sinais GNSS, com foco na aplicação em sistemas aeroespaciais. O processo envolve a análise de um conjunto de dados representativo, o pré-processamento dessas informações para adequação aos modelos de *Machine Learning* (ML) e a implementação de uma arquitetura híbrida que combina técnicas clássicas de ML com abordagens de *Deep Learning*.

2.1. Dataset e Engenharia de Atributos

Os dados utilizados neste estudo são provenientes de [2][1]. Este conjunto de dados foi escolhido pela sua relevância e qualidade, contendo sinais GPS reais misturados com ataques simulados de alta fidelidade, capturados através de um rádio definido por software (USRP).

O *dataset* é composto por 13 características (*features*) extraídas de 8 canais paralelos do receptor GPS, que monitoram diferentes estágios do processamento do sinal (aquisição, rastreamento e descodificação). As principais grandezas analisadas incluem:

- **Identificação e Tempo:** Número do satélite (PRN), Tempo do Receptor (RX) e Tempo da Semana (TOW).
- **Medições de Navegação:** *Pseudorange* (PD), que estima a distância ao satélite, e *Doppler Offset* (DO), que mede o desvio de frequência devido ao movimento relativo.
- **Métricas de Qualidade e Correlação:** Razão Portadora-Ruído (CN0), Fase da Portadora (CP), e as magnitudes dos correlatores *Early* (EC), *Late* (LC) e *Prompt* (PC). Adicionalmente, são considerados os componentes em fase (PIP) e quadratura (PQP) do correlator *Prompt*, essenciais para detectar distorções na forma de onda do sinal.

A variável alvo classifica cada amostra em uma de quatro categorias, permitindo uma abordagem de aprendizagem supervisionada multiclasse:

1. **Classe 0 (Legítimo):** Sinais autênticos sem interferência.

[1] Disponível publicamente em: https://data.mendeley.com/datasets/z7dj3yyzt8/3.

2. **Classe 1 (Ataque Simplista):** Sinais falsos não sincronizados com os reais, geralmente com alta potência e desvios Doppler anormais.
3. **Classe 2 (Ataque Intermédio):** O atacante conhece a posição aproximada da vítima e tenta alinhar a fase do código e o Doppler, mas introduz distorções nos correlatores durante a captura.
4. **Classe 3 (Ataque Sofisticado):** Ataques complexos que utilizam múltiplas antenas para manipular a direção de chegada ou introduzem desvios muito sutis, difíceis de detetar por métodos convencionais.

O desenvolvimento dos algoritmos baseou-se nos dados da planilha `GPS_Data_Simplified_2D_Feature_Map.xlsx`, que organiza os dados brutos dos canais numa estrutura tabular bidimensional, facilitando a ingestão pelos modelos de ML.

2.2. Ambiente de Desenvolvimento e Software

Todos os experimentos foram realizados utilizando a linguagem de programação **Python**, amplamente adotada na comunidade de ciência de dados devido ao seu rico ecossistema de bibliotecas. O código fonte está documentado em *Jupyter Notebooks* disponíveis no repositório do projeto[2]. As principais bibliotecas utilizadas foram:

- **Pandas:** Para manipulação, limpeza e estruturação dos dados tabulares.
- **Matplotlib e Seaborn:** Para visualização de dados, análise exploratória e plotagem de métricas de desempenho (como matrizes de confusão e curvas ROC).
- **Scikit-learn:** Para implementação de algoritmos clássicos de ML, pré-processamento (normalização `StandardScaler`) e métricas de avaliação.
- **XGBoost:** Para a implementação de modelos baseados em *Gradient Boosting*, conhecidos pela sua alta performance em dados tabulares estruturados.
- **TensorFlow/Keras:** Utilizado para a construção e treino da arquitetura de *Deep Learning* proposta.

2.3. Diferenciais Metodológicos

Este trabalho baseia-se na fundação estabelecida por pesquisas anteriores, nomeadamente o trabalho de referência [1], que explorou classificadores clássicos como *Random Forest* para detecção de spoofing. No entanto, esta proposta introduz algumas mudanças na metodologia visando aumentar a robustez e a precisão da detecção e, principalmente, o rigor na avaliação de performance dos modelos.

Primeiramente, foi dada ênfase ao tratamento do desequilíbrio de classes, uma característica comum em cenários de segurança onde ataques são eventos mais raros que o funcionamento normal. Foram aplicadas técnicas de balanceamento de dados, como SMOTE, para garantir que o modelo não enviese as suas previsões para a classe majoritária.

[2]https://github.com/gabrielbopi/ML_spoofing_detection

A principal inovação técnica reside na implementação de uma arquitetura de **Rede Neural Profunda (DNN)** híbrida, incorporando camadas **BiLSTM (Bidirectional Long Short-Term Memory)** e **BiGRU (Bidirectional Gated Recurrent Unit)**. Diferente dos modelos estáticos, esta arquitetura é capaz de aprender dependências temporais nos dados de sinal GNSS. Ao processar a sequência de dados em ambas as direções (passado e futuro dentro de uma janela temporal), o modelo BiLSTM-BiGRU, em tese, consegue capturar anomalias dinâmicas sutis, como variações graduais no *Doppler* ou inconsistências temporais nos correlatores, que são características de ataques de spoofing mais sofisticados (como o tipo *carry-off*).

Por fim, vale ressaltar que, por restrições de escopo, este estudo foca exclusivamente na detecção de *Spoofing*, não abordando ataques de *Jamming* (negação de serviço), concentrando-se na integridade e autenticidade da informação de navegação.

3. Resultados

A análise experimental foi conduzida para avaliar a eficácia de diferentes algoritmos na identificação de sinais de GNSS autênticos e falsificados. Esta seção detalha a análise exploratória, o pré-processamento dos dados e o desempenho comparativo dos modelos propostos.

3.1. Análise Exploratória de Dados

Inicialmente, foi realizada uma verificação da integridade do *dataset*. Não foram encontradas amostras com dados faltantes ou nulos, garantindo a qualidade da base de dados para o treino. A análise de correlação entre as 13 características extraídas foi visualizada através de um **Mapa de Calor de Matriz de Correlação**, mostrado na Figura 1.

Optou-se por não realizar o tratamento ou remoção de *outliers*. Em contextos de cibersegurança e deteção de anomalias, valores espúrios frequentemente representam a assinatura de um ataque (por exemplo, picos de potência em ataques de *Jamming* ou desvios de fase em *Spoofing*). Consequentemente, a normalização das *features* foi realizada utilizando o `StandardScaler`, em detrimento do `RobustScaler`, para preservar a influência destes valores extremos significativos.

Como pode ser visto na Figura 2, a distribuição das classes no conjunto de dados original revelou um desequilíbrio substancial, com a classe "Legítima" (0) a representar a vasta maioria das amostras, conforme ilustrado na figura de distribuição de classes do projeto.

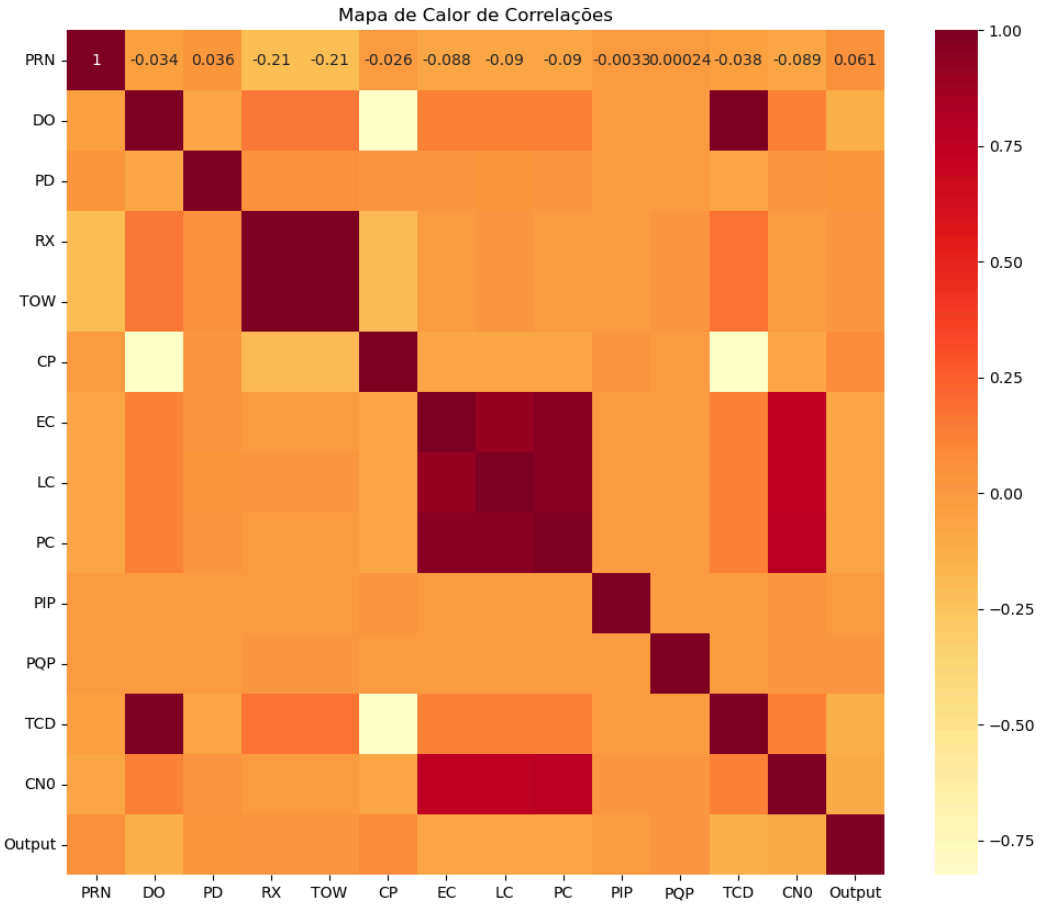

Figura 1. Matriz de correlação das features em formato de mapa de calor.

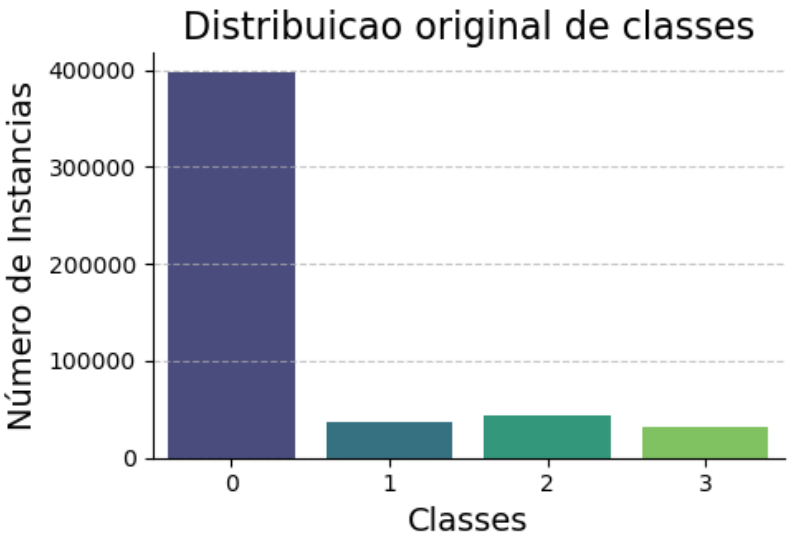

Figura 2. Número de amostras para cada classe de ataque no *dataset*, há muito menos instâncias para as classes de ataque (1, 2 e 3).

3.2. Estratégia de Balanceamento de Dados

Para mitigar o enviesamento dos classificadores em direção à classe majoritária, foram realizados experimentos de balanceamento aplicados apenas aos dados de treino (70% do total). Diferente de abordagens anteriores que utilizaram *Random Undersampling* [1], este trabalho optou pela técnica de *Oversampling* **SMOTE** (Synthetic Minority Over-sampling Technique).

- **Número original de amostras de treino:** 510.530
- **Amostras após Undersampling (descartado):** 89.180
- **Amostras após SMOTE (adotado):** 1.114.688

Esta expansão do *dataset* permitiu que as classes de ataque (Simplista, Intermédio e Sofisticado) tivessem representatividade melhor frente à classe Legítima durante o treino, melhorando a capacidade de generalização do modelo para ameaças raras.

3.3. Avaliação dos Modelos e Hiperparâmetros

Foram avaliados 8 modelos de classificação, expandindo o trabalho de referência com a inclusão de uma arquitetura de *Deep Learning* baseada em Redes Neurais Recorrentes (RNN). Os modelos testados foram:

1. Regressão Logística
2. K-Nearest Neighbors (KNN)
3. Gaussian Naive Bayes
4. Support Vector Machine (SVM)
5. Decision Tree
6. Random Forest
7. Gradient Boosting (XGBoost)
8. **RNN Profunda (BiLSTM-BiGRU)** - *Inovação proposta*

Para os classificadores clássicos, foi realizada uma otimização de hiperparâmetros utilizando **Grid Search** com validação cruzada estratificada em 5 lotes (*splits*). A métrica de *score* principal para seleção foi a **acurácia (accuracy)**, dada a importância crítica de maximizar a taxa global de deteção correta em sistemas de segurança aeroespacial.

Alguns dos melhores hiperparâmetros encontrados incluíram:

- *Regressão Logística:* `{'C': 10, 'solver': 'lbfgs'}`
- *SVM:* `{'C': 1, 'gamma': 0.1, 'kernel': 'rbf'}`
- *Decision Tree:* `{'criterion': 'entropy', 'max_features': 10, 'min_samples_leaf': 10}`

3.4. Arquitetura da Rede Neural

A arquitetura da DNN proposta combina camadas bidirecionais para capturar dependências temporais complexas nos sinais de rádio. A rede é composta por:

1. Camada de Entrada.
2. **Camada Bidirecional GRU** (64 unidades): Para capturar dependências de curto prazo.
3. **Camada Bidirecional LSTM** (32 unidades): Para aprender sequências de longo prazo nos dados de sinal.
4. Camadas Densas (*Dense*) com ativação *Relu* (32 unidades) e *Softmax* (4 unidades) para classificação final.

Sua arquitetura é mostrada na Tabela 1.

Tabela 1. Detalhamento de shapes e parâmetros de cada camada da DNN.

Layer (type)	Output Shape	Param #
bidirectional (Bidirectional)	(None, 13, 128)	25,728
bidirectional_1 (Bidirectional)	(None, 64)	41,216
dense (Dense)	(None, 32)	2,080
dense_1 (Dense)	(None, 4)	132

5. Análise dos resultados e Discussão

Os resultados obtidos no conjunto de teste demonstraram uma superioridade marcante dos modelos baseados em árvores de decisão. A, Figura 3, comparativa de pontuações (acurácia) revela o seguinte panorama:

- **Modelos Lineares e Probabilísticos:** A *Regressão Logística* (~43%) e o *Naive Bayes* (~44%) apresentaram desempenho insatisfatório, indicando que a fronteira de decisão entre sinais legítimos e *spoofing* sofisticado não é linear.
- **Modelos de Ensemble:** O **XGBoost** e o **Random Forest** atingiram as melhores performances, com acurácias no conjunto de teste superiores a **93%**. Estes modelos demonstraram robustez no tratamento das *features* não-lineares dos correlatores.
- **Rede Neural (RNN):** O modelo BiLSTM-BiGRU obteve uma acurácia competitiva (~86%), embora inferior aos modelos de *boosting*.

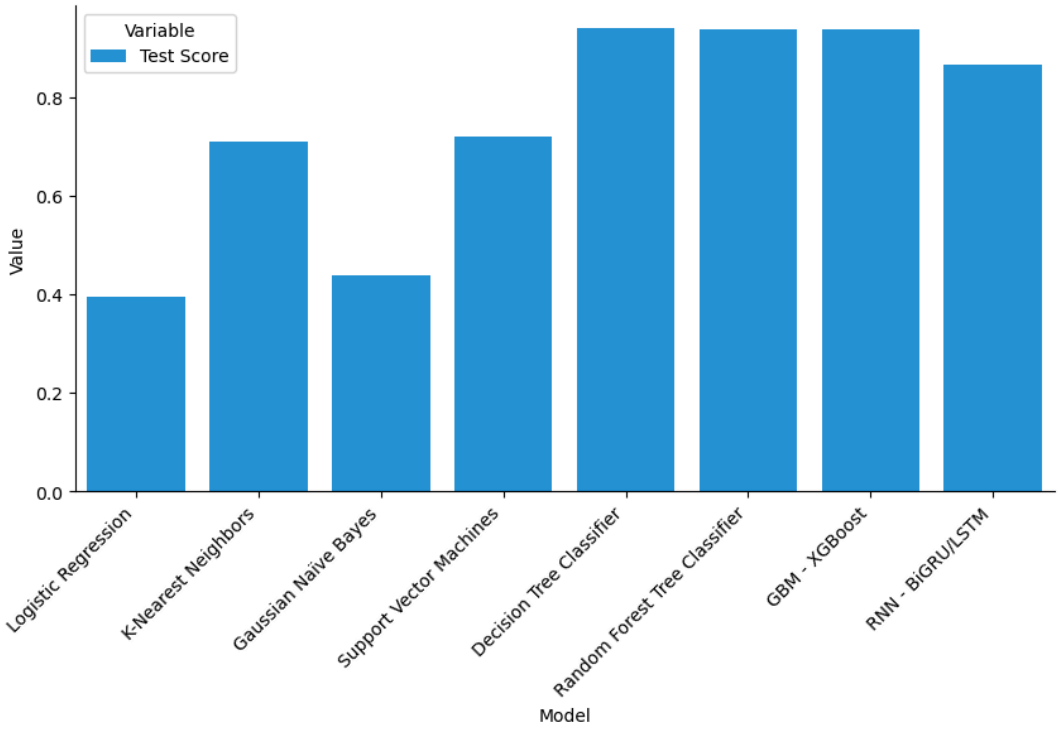

Figura 3. Comparação de pontuação (acurácia) para cada método.

A análise da **Matriz de Confusão** do modelo de RNN e XGBoost (Figura 3) confirma que a maior dificuldade reside na distinção entre a *Classe 0* (Legítimo) e a *Classe 2* (Ataque Intermédio), onde os sinais são temporalmente alinhados. No entanto, as curvas **ROC-AUC (Figura 4)** apresentaram áreas próximas de 1.00 para a maioria das classes nos modelos de topo, validando a eficácia do sistema proposto na detecção de ataques com baixa taxa de falsos positivos.

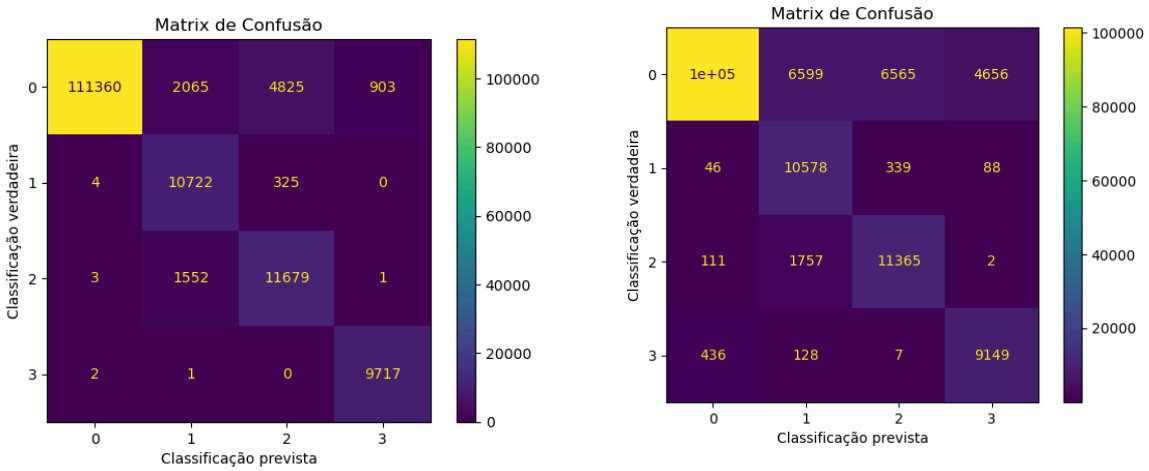

Figura 4. Matriz de confusão do método XGBoost, à esquerda, e do DNN, à direita.

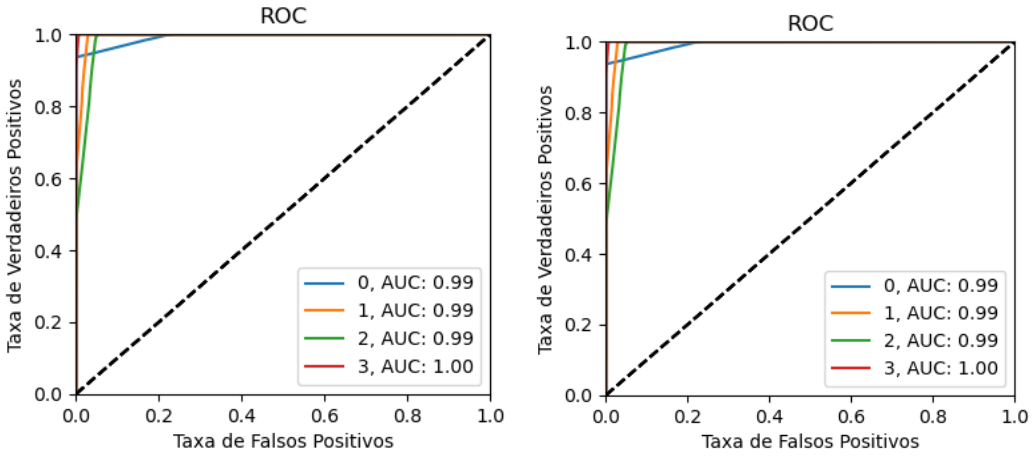

Figura 5. Curvas de ROC para os métodos XGBoost e de DNN .

Com base nestes resultados, a proposta técnica para implementação em VANTs e aeronaves tripuladas deve considerar que o modelo **XGBoost** é o candidato ideal para uma primeira fase de implementação. Ele oferece o melhor equilíbrio entre precisão e custo computacional, permitindo a execução em processadores de bordo com restrições de energia e capacidade de processamento, comuns em drones e aviônicos. A considerável taxa de acerto na *Classe 3 (Ataques Sofisticados)* valida que o sistema é capaz de detectar muitas tentativas de *spoofing* onde o atacante tenta sincronizar perfeitamente com o sinal do satélite, uma das ameaças mais perigosas para a navegação autônoma.

Com base nas evidências empíricas obtidas, a proposta técnica final para a indústria aeroespacial consiste na implementação de um **Módulo de Integridade GNSS Baseado em ML**, operando em duas camadas:

1. **Detecção em Tempo Real (Borda):** Integração de um modelo **XGBoost** leve diretamente no firmware dos receptores GNSS ou no computador de missão da aeronave/VANT. Este modelo processará as métricas de qualidade do sinal (como CN0 e saídas dos correlatores) a cada época de navegação, fornecendo um "sinal de confiança" instantâneo.
2. **Monitoramento e Análise (Solo/Nuvem):** Implementação de uma camada secundária baseada em **Redes Neurais Recorrentes** para análise pós-processada ou monitoramento de frota, capaz de identificar tendências de ataques lentos (*carry-off*) que possam escapar à detecção instantânea.

Esta arquitetura híbrida garante uma defesa resiliente, minimizando a latência na tomada de decisão crítica a bordo, ao mesmo tempo que mantém uma capacidade de adaptação contínua contra novas ameaças. A solução proposta não exige alterações no hardware da constelação de satélites, tornando-a uma atualização de segurança custo-efetiva e escalável para a frota aérea existente.

4. Conclusão

Este estudo demonstrou que o uso de Inteligência Artificial é não apenas viável, mas altamente eficaz na proteção de sistemas aeroespaciais contra ataques de *Spoofing* GNSS. Através de uma abordagem metodológica rigorosa, que incluiu o balanceamento de dados e a avaliação de múltiplos classificadores, identificou-se que modelos de aprendizado de máquina podem distinguir entre sinais legítimos e maliciosos com elevada precisão.

Os resultados evidenciaram que modelos de *ensemble* baseados em árvores de decisão, nomeadamente o **XGBoost**, oferecem o desempenho mais robusto, com uma acurácia superior a **93%** e uma capacidade notável de identificar até mesmo os ataques mais sofisticados (Classe 3). A comparação com modelos lineares, que obtiveram resultados insatisfatórios, reforça a natureza não-linear das anomalias introduzidas por ataques de falsificação nos correlatores do receptor.

Adicionalmente, a exploração de arquiteturas de *Deep Learning* (BiLSTM-BiGRU) mostrou-se promissora, sugerindo um caminho para futuras investigações focadas na análise temporal contínua de sinais, o que poderia complementar a detecção estática oferecida pelos modelos de árvores.

Em suma, os experimentos validam que a utilização de Aprendizado de Máquina, especificamente métodos de *ensemble* treinados com dados balanceados via SMOTE, constitui uma barreira de defesa viável e com uma certa eficácia para proteger ativos aeroespaciais contra a manipulação de GNSS.

Referências

[1] Ghanbarzadeh, A., Soleimani, M., & Soleimani, H. "GNSS/GPS Spoofing and Jamming Identification Using Machine Learning and Deep Learning". School of Electrical Engineering, Iran University of Science and Technology (2025).

[2] Aissou, G., Benouadah, S., El Alami, H., & Kaabouch, N. "A DATASET for GPS Spoofing Detection on Autonomous Vehicles". University of North Dakota, Mendeley Data.

[3] Psiaki, M. L., & Humphreys, T. E. "GNSS Spoofing and Detection". Proceedings of the IEEE, vol. 104, no. 6 (2016).

[4] Jafarnia-Jahromi, A., Broumandan, A., Nielsen, J., & Lachapelle, G. "GPS Vulnerability to Spoofing Threats and a Review of Antispoofing Techniques". International Journal of Navigation and Observation, vol. 2012 (2012).

[5] Van der Merwe, J. R., Zubizarreta, X., Lukčin, I., Rügamer, A., & Felber, W. "Classification of Spoofing Attack Types". European Navigation Conference (ENC) (2018).

[6] Kerns, A. J., Shepard, D. P., Bhatti, J. A., & Humphreys, T. E. "Unmanned Aircraft Capture and Control Via GPS Spoofing". Journal of Field Robotics, vol. 31, no. 4 (2014).

[7] Islam, S., Bhuiyan, M. Z. H., Liaquat, M., Pääkkönen, I., & Kaasalainen, S. "An open GNSS spoofing data repository: characterization and impact analysis with FGI-GSRx open-source software-defined receiver". GPS Solutions (2024).

[8] Shafique, A., Mehmood, A., & Elhadef, M. "Detecting Signal Spoofing Attack in UAVs Using Machine Learning Models". IEEE Access, vol. 9 (2021).

[9] Humphreys, T. E., Bhatti, J. A., Shepard, D. P., & Wesson, K. D. "The Texas Spoofing Test Battery: Toward a Standard for Evaluating GPS Signal Authentication Techniques". Proceedings of the 25th International Technical Meeting of the Satellite Division of The Institute of Navigation (2012).

[10] Aissou, G., et al. "A DATASET for GPS Spoofing Detection on Autonomous Vehicles". University of North Dakota, Mendeley Data.

[11] Zelinka, J., Kost, O., & Hrúz, M. "Deep Sequence-to-Sequence Models for GNSS Spoofing Detection". arXiv:2510.19890v1 (2025).