

GUSTAVO ALVES DA CÂMARA

ALGORITMOS DE MACHINE LEARNING NA ANÁLISE ANTECIPADA DE CONFLITOS

Barreiras-Ba

09/2025

RESUMO

Este artigo investiga a aplicação de algoritmos de *machine learning* na análise antecipada de conflitos, abordando a incapacidade dos métodos tradicionais em lidar com a complexidade e o volume de dados contemporâneos. O objetivo geral é analisar criticamente a literatura sobre o tema, avaliando a eficácia das técnicas e os resultados de estudos de caso. A metodologia utilizada foi a pesquisa bibliográfica, com revisão sistemática de artigos científicos, livros e relatórios técnicos. Os resultados indicam que os modelos de *machine learning*, especialmente os de conjunto como *Random Forest*, demonstram acurácia preditiva superior aos métodos estatísticos clássicos. No entanto, a discussão revela que a eficácia desses modelos é limitada pela qualidade e pelos vieses dos dados de entrada, além de sua inerente incapacidade de prever eventos sem precedentes, os "Cisnes Negros". Conclui-se que o *machine learning* é uma poderosa ferramenta de auxílio à decisão, mas sua implementação responsável exige uma abordagem crítica, preferencialmente no modelo "Humano no Circuito", para mitigar riscos éticos e maximizar seu potencial na promoção da estabilidade global.

Palavras-chave: machine learning; previsão de conflitos; ética em algoritmos.

1. INTRODUÇÃO

A humanidade, em suas mais diversas civilizações, sempre buscou formas de obter respostas a questões que transcendem o alcance da compreensão humana ordinária. Essa tentativa de comunicação com uma ordem superior de conhecimento, conhecida como adivinhação, tinha como propósito não apenas a previsão de eventos futuros, mas fundamentalmente a orientação das ações humanas no presente (LOEWE; BLACKER, 1981). Na antiguidade, os oráculos funcionavam como um instrumento para decifrar a vontade divina e guiar decisões críticas (RABELLO, 2013). Em uma analogia contemporânea, a aplicação de algoritmos de *machine learning* (ML) na análise de dados representa uma nova forma de oráculo, um método que busca extrair, de volumes massivos de informação, padrões e tendências que orientem a tomada de decisão em cenários complexos, como a antecipação de conflitos.

A problemática central que esta pesquisa aborda é a crescente incapacidade dos

métodos tradicionais de análise em prever com a acurácia necessária a eclosão de conflitos armados e guerras civis (PERRY, 2013). Tais métodos, frequentemente limitados pela subjetividade e pela incapacidade de processar grandes volumes de dados de forma adaptativa, mostram-se insuficientes diante da complexidade do cenário geopolítico atual (SOUZA et al., 2018). Com a exponencial quantidade de dados digitais disponíveis — de indicadores econômicos a discursos em redes sociais —, explorar novas abordagens que utilizem a capacidade do *machine learning* torna-se não apenas uma oportunidade, mas uma necessidade para aprimorar a precisão das previsões e qualificar a tomada de decisão estratégica (ALPAYDIN, 2020).

O objetivo geral deste trabalho é analisar a literatura existente sobre o uso de algoritmos de *machine learning* na previsão de conflitos, complementado pela análise crítica das técnicas empregadas e dos resultados obtidos em estudos de caso relevantes.

Para alcançar o propósito estabelecido, foram definidos os seguintes objetivos específicos: realizar uma revisão abrangente da literatura sobre os algoritmos de ML aplicados à previsão de conflitos; identificar e analisar estudos de caso que demonstrem a aplicação prática do uso de ML; definir e avaliar o desempenho de diferentes algoritmos na predição de áreas que entraram em conflito; e, por fim, comparar os métodos tradicionais com as abordagens de *machine learning* de melhor aproveitamento.

A justificativa para o desenvolvimento desta pesquisa reside na necessidade premente de aprimorar a capacidade de antecipação de conflitos, um avanço que pode contribuir diretamente para a prevenção de crises humanitárias e para a mitigação dos seus impactos sociais, econômicos e políticos. Adicionalmente, este estudo poderá fornecer subsídios valiosos para a formulação de políticas públicas e estratégias de segurança mais eficazes por parte de governos e organizações internacionais, oferecendo uma ferramenta analítica mais robusta para um mundo em constante transformação.

2. REFERENCIAL TEÓRICO

2.1. A Evolução da Análise Preditiva: Da Estatística ao *Machine Learning*

A capacidade de antecipar eventos futuros é um dos pilares da tomada de decisão

estratégica em qualquer campo do conhecimento humano. Tradicionalmente, essa tarefa era domínio dos métodos estatísticos, que se baseiam em modelos matemáticos para identificar tendências e realizar projeções a partir de dados históricos (SOUZA et al., 2018). Ferramentas como a regressão linear e as análises de séries temporais foram, e ainda são, fundamentais para a previsão de demanda, planejamento de recursos e análise de cenários econômicos (BISHOP, 2006). Contudo, a crescente complexidade e o volume massivo de dados — o chamado *Big Data* — do mundo contemporâneo começaram a expor as limitações desses métodos tradicionais, que assumem muitas vezes relações lineares e distribuições de dados predefinidas, mostrando-se menos eficazes para capturar os padrões não-lineares e as interações complexas inerentes a sistemas dinâmicos, como os que caracterizam as relações internacionais e a eclosão de conflitos (SOUZA et al., 2018).

Nesse contexto, os algoritmos de *Machine Learning* (ML) surgem como uma evolução paradigmática. Diferentemente da modelagem estatística clássica, onde as regras são explicitamente programadas, o ML permite que os sistemas "aprendam" diretamente a partir dos dados. Conforme definido por Arthur Samuel em 1959, trata-se de um "campo de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados" (IBM, [s.d.]). Essa capacidade de identificar padrões de forma autônoma, adaptar-se a novas informações e melhorar seu desempenho com a experiência torna os algoritmos de ML ferramentas exponencialmente mais poderosas para a análise preditiva em ambientes de alta complexidade e incerteza (ALPAYDIN, 2020).

2.2. Fundamentos dos Algoritmos de *Machine Learning*

O campo do *Machine Learning* é vasto, mas seus algoritmos podem ser categorizados em três tipos principais de aprendizado, cada um adequado a diferentes tipos de problemas e dados (DATAGEEKS, 2019). A compreensão dessas categorias é essencial para a aplicação correta em qualquer domínio, incluindo a análise de conflitos.

O **Aprendizado Supervisionado** é a abordagem mais comum e direta. Nela, o algoritmo é treinado com um conjunto de dados previamente rotulado (*labeled data*), ou seja, um conjunto de dados onde a "resposta correta" ou o resultado desejado já é conhecido (DATACAMP, 2024). O objetivo é que o modelo aprenda a mapear as

entradas (*features*) para as saídas correspondentes. As duas tarefas mais comuns são a **Regressão**, que busca prever um valor contínuo, e a **Classificação**, que visa atribuir uma categoria a um dado. Algoritmos populares nesta categoria incluem Regressão Linear, Máquinas de Vetores de Suporte (SVM) e Redes Neurais (IBM, [s.d.]).

O **Aprendizado Não Supervisionado**, por sua vez, lida com dados que não possuem rótulos predefinidos. O objetivo do algoritmo é explorar a estrutura inerente aos dados e extrair padrões significativos por conta própria (DATAGEEKS, 2019). A principal tarefa aqui é a **Clusterização** (ou agrupamento), que consiste em agrupar dados semelhantes em "clusters". Outra aplicação é a **Redução de Dimensionalidade**, que simplifica o número de variáveis em um conjunto de dados, mantendo as informações mais importantes (BISHOP, 2006).

Finalmente, o **Aprendizado por Reforço** se inspira na psicologia comportamental. Nesse modelo, um "agente" aprende a tomar decisões em um ambiente dinâmico por meio de um sistema de tentativa e erro, sendo recompensado por ações que o aproximam de um objetivo e penalizado por ações que o afastam (IBM, [s.d.]). Embora mais complexo, tem grande potencial para simular a tomada de decisão em cenários estratégicos.

2.3. O Desafio da Incerteza: O Cisne Negro e a Antifragilidade

Apesar do poder dos modelos preditivos, a história é marcada por eventos de alto impacto que eram considerados totalmente improváveis antes de sua ocorrência. Nassim Nicholas Taleb (2021) conceitua esses eventos como "Cisnes Negros": anomalias que residem fora do âmbito das expectativas regulares, possuem um impacto extremo e, somente após sua ocorrência, são racionalizados por retrospectiva, como se pudessem ter sido previstos. A crise financeira de 2008 ou o início abrupto de certos conflitos são exemplos de eventos que desafiam qualquer modelo baseado puramente em dados históricos, pois, por definição, não há precedentes em seus conjuntos de treinamento para antecipá-los (G1, 2021).

A teoria do Cisne Negro expõe uma falácia fundamental na análise de risco: a dependência excessiva da observação do passado para prever o futuro em sistemas complexos e não-lineares. Diante dessa limitação, Taleb (2012) propõe o conceito de **Antifragilidade**. Enquanto a robustez descreve um sistema que resiste a choques e permanece o mesmo, a antifragilidade refere-se a sistemas que se beneficiam do

caos, da volatilidade e dos estressores. Um sistema antifrágil não apenas sobrevive à incerteza, mas se fortalece e evolui a partir dela. No contexto da análise de conflitos, isso implica pensar em modelos e estratégias que não apenas tentem prever o imprevisível, mas que sejam construídos para se adaptar e até se beneficiar de eventos inesperados, reconhecendo a inevitabilidade da incerteza.

2.4. Os Riscos do Oráculo Digital: Algoritmos de Destruição em Massa

A crescente aplicação de algoritmos em decisões críticas — da concessão de crédito à sentença judicial — levanta profundas questões éticas. Cathy O’Neil (2020) cunha o termo "Algoritmos de Destruição em Massa" (*Weapons of Math Destruction* - WMDs) para descrever modelos matemáticos que são opacos, não regulados e inquestionáveis, e que causam danos em larga escala. Tais algoritmos, muitas vezes, codificam e amplificam vieses humanos, perpetuando a desigualdade sob um verniz de objetividade técnica. Um modelo de previsão de conflitos, por exemplo, se treinado com dados históricos enviesados, poderia perpetuar estereótipos sobre determinadas regiões ou etnias, levando a políticas de segurança discriminatórias e contraproducentes (O’NEIL, 2020).

Um WMD, segundo O’Neil (2020), possui três características principais: opacidade, escala e dano. A aplicação de *Machine Learning* na análise de segurança internacional deve, portanto, ser acompanhada de uma rigorosa reflexão ética, buscando a transparência dos modelos (*explainable AI*), a auditoria constante para a detecção de vieses e a criação de mecanismos de contestação e correção, a fim de garantir que essas poderosas ferramentas sirvam como instrumentos de estabilização, e não como catalisadores de injustiça.

3. METODOLOGIA

O presente trabalho caracteriza-se como uma pesquisa de natureza qualitativa, cujo método principal é a **pesquisa bibliográfica**. A abordagem adotada consiste em uma revisão sistemática e crítica da literatura científica e técnica sobre a aplicação de algoritmos de *Machine Learning* na análise e previsão de conflitos. Para tanto, foram utilizadas como fontes artigos científicos indexados em bases de dados como Google Scholar, Scielo e IEEE Xplore, livros-texto de referência nas áreas de Inteligência Artificial e Relações Internacionais, teses, dissertações e publicações técnicas de organizações especializadas no tema.

Os procedimentos metodológicos foram estruturados em etapas sequenciais para garantir o rigor e a abrangência da análise. A primeira etapa consistiu no levantamento e fichamento do material bibliográfico, selecionando as obras mais relevantes a partir de palavras-chave como "previsão de conflitos", "*conflict prediction*", "*machine learning*", "cisne negro" e "ética em algoritmos". Na segunda etapa, foi realizada uma análise crítica do conteúdo, visando identificar os principais conceitos, os algoritmos mais recorrentes na literatura, os estudos de caso mais significativos e os resultados (positivos e negativos) reportados pelos pesquisadores. A terceira e última etapa metodológica é a síntese e a discussão comparativa dos dados levantados. Nesta fase, as informações coletadas são organizadas, comparadas e interpretadas à luz do referencial teórico adotado. O objetivo é contrastar a eficácia das diferentes abordagens de *Machine Learning* com os métodos tradicionais de previsão, destacando suas respectivas vantagens, desvantagens, desafios de implementação e, crucialmente, as implicações éticas decorrentes do uso de tais tecnologias para fins de segurança e governança, conforme discutido na fundamentação teórica.

Importa salientar que, por sua natureza teórico-bibliográfica, esta pesquisa não contempla a implementação prática ou o desenvolvimento de um novo modelo computacional. O escopo do trabalho está delimitado à análise crítica e à síntese do conhecimento já produzido na área, culminando na apresentação de uma discussão aprofundada que consolida os achados da literatura e aponta direções para futuras investigações no campo.

4. DESENVOLVIMENTO

4.1. O Ecossistema de Dados na Previsão de Conflitos

A transição de modelos teóricos para abordagens orientadas por dados (*data-driven*) representa uma mudança fundamental no campo da previsão de conflitos (SÁNCHEZ-MYERS et al., 2021). A capacidade computacional permite que os próprios dados revelem correlações e padrões preditivos que não seriam evidentes apenas pela análise humana. O sucesso de qualquer modelo de *Machine Learning* nesta área depende, primeiramente, da coleta e integração de dois tipos principais de dados: os estruturais e os de eventos.

Os **dados estruturais** referem-se a características macro e de lenta alteração de um

país ou região, como indicadores socioeconômicos, demográficos e institucionais. Fontes como o Banco Mundial e o *Polity IV* fornecem variáveis como PIB per capita, mortalidade infantil e nível de democracia. Esses dados são essenciais para modelar as condições subjacentes que podem tornar uma região suscetível a instabilidades (SÁNCHEZ-MYERS et al., 2021). O *Political Instability Task Force* (PITF) utiliza predominantemente dados estruturais para gerar previsões de médio prazo (um a dois anos).

Por outro lado, os **dados de eventos** capturam a dinâmica de curto prazo das interações políticas, registrando ocorrências específicas como protestos, acordos diplomáticos e ataques. Grandes bancos de dados como o GDELT (*Global Database of Events, Language, and Tone*) e o ACLED (*Armed Conflict Location & Event Data Project*) são cruciais para essa finalidade. A integração de dados de eventos permite a criação de modelos de previsão em tempo real ou de curtíssimo prazo, capazes de identificar gatilhos e a escalada de tensões que podem levar a um conflito iminente (PERRY, 2013).

4.2. Algoritmos em Ação: Uma Análise Comparativa

Uma vez que os dados são coletados e processados, a escolha do algoritmo de *Machine Learning* é o próximo passo crítico. A literatura sobre previsão de conflitos aponta para o uso recorrente de um conjunto de classificadores, cuja performance é frequentemente comparada para determinar o mais adequado para a tarefa. Entre os mais citados, destacam-se a Regressão Logística, as Máquinas de Vetores de Suporte (SVM), o Naïve Bayes e, com especial proeminência, os métodos de conjunto (*ensemble methods*) como o *Random Forest* (ALARIFI et al., 2020).

O *Random Forest*, por exemplo, é frequentemente favorecido por sua robustez e capacidade de lidar com interações complexas e não-lineares, além de sua menor suscetibilidade a *overfitting*. Estudos comparativos que buscam prever conflitos em tempo real avaliam esses algoritmos com base em métricas como Precisão, Revocação e a Curva ROC e sua AUC (Área Sob a Curva), que é uma medida robusta da capacidade geral de discriminação do modelo (ALARIFI et al., 2020). Estudos demonstram que modelos como *Random Forest* e Redes Neurais tendem a superar os modelos de regressão mais simples em termos de poder preditivo (ALARIFI et al., 2020).

4.3. Estudo de Caso: Modelando a Instabilidade na América Central

Para ilustrar a aplicação prática desses conceitos, a análise do trabalho de Perry (2013) sobre a previsão de instabilidade política na América Central serve como um estudo de caso exemplar. O objetivo do projeto era desenvolver um modelo de *Machine Learning* capaz de prever a ocorrência de "eventos materiais de instabilidade" com um mês de antecedência. A pesquisa utilizou uma combinação de dados estruturais do Banco Mundial e dados de eventos provenientes do banco de dados GDELT (PERRY, 2013).

O processo metodológico envolveu uma criteriosa engenharia de atributos, na qual os dados brutos de eventos foram transformados em variáveis preditivas, como a média do "tom" da mídia e a contagem de eventos de conflito e cooperação (PERRY, 2013). O algoritmo escolhido para a tarefa foi o *Random Forest*. O desempenho do modelo foi avaliado utilizando a métrica AUC, alcançando um valor de 0.84, o que indica uma capacidade de discriminação consideravelmente boa (PERRY, 2013). Este caso demonstra como a fusão de diferentes fontes de dados e a aplicação de um algoritmo robusto podem gerar um sistema de alerta precoce com validade empírica.

4.4. O Elo Mais Fraco: A Criticidade da Qualidade dos Dados

A máxima "lixo entra, lixo sai" (*garbage in, garbage out*) é um princípio fundamental no campo do *Machine Learning*. A performance de qualquer algoritmo é intrinsecamente limitada pela qualidade dos dados utilizados para seu treinamento (GONG et al., 2023). A qualidade dos dados é um conceito multidimensional, que pode ser avaliado a partir de atributos como **acurácia**, **completude**, **consistência** e **pontualidade** (ANOMALO, 2024).

Problemas comuns como dados faltantes, outliers e *data drift* podem degradar severamente a capacidade preditiva de um modelo (GONG et al., 2023). Um banco de dados de eventos que sub-representa sistematicamente incidentes em áreas rurais, por exemplo, introduzirá um viés geográfico no modelo. Portanto, um pipeline de *Machine Learning* robusto para a previsão de conflitos exige um investimento maciço em validação, limpeza e monitoramento contínuo dos dados (ANOMALO, 2024).

4.5. Análise Comparativa de Abordagens: Previsão de Curto vs. Médio Prazo

A análise da literatura sobre a aplicação de *Machine Learning* na previsão de

conflitos revela a existência de duas abordagens principais que operam em diferentes horizontes temporais e utilizam tipos distintos de dados. A distinção fundamental reside na natureza da pergunta que cada modelo busca responder: se o foco é identificar **gatilhos** iminentes ou **vulnerabilidades** sistêmicas. Os estudos de caso de Perry (2013) e do *Political Instability Task Force* (PITF) ilustram perfeitamente essa dicotomia.

A primeira abordagem, exemplificada pelo estudo de caso sobre a América Central, concentra-se na **previsão de curto prazo baseada em dados de eventos**. Este método utiliza fontes de dados de alta frequência, como o monitoramento de notícias globais, para capturar a dinâmica de interações políticas em tempo quase real. O objetivo é identificar padrões em eventos recentes que possam servir como gatilhos para a eclosão de instabilidade em um futuro próximo. Modelos desta natureza funcionam como um barômetro tático, sendo mais eficazes para sistemas de alerta precoce (PERRY, 2013).

Em contraste, a segunda abordagem, personificada pelo trabalho do PITF, foca na **previsão de médio prazo baseada em dados estruturais**. Este método utiliza variáveis de evolução lenta que descrevem as condições fundamentais de um país, como mortalidade infantil e tipo de regime político, para construir modelos que identifiquem as vulnerabilidades estruturais em um horizonte de um a dois anos. Essa abordagem funciona como um diagnóstico estratégico, mais adequada para o planejamento de longo prazo e a alocação de recursos para a prevenção de conflitos (SÁNCHEZ-MYERS et al., 2021).

A comparação entre os dois casos de uso evidencia que não há uma metodologia única ou superior. Enquanto os modelos baseados em eventos são essenciais para o monitoramento tático, os modelos estruturais são fundamentais para o planejamento estratégico. Uma arquitetura de previsão de conflitos verdadeiramente robusta, portanto, integraria ambas as abordagens para obter uma visão completa, combinando a análise de vulnerabilidades com o monitoramento contínuo de gatilhos.

5. RESULTADOS OBTIDOS

A presente seção compila os principais resultados obtidos a partir da revisão sistemática da literatura, focando na eficácia, nas metodologias e nos desafios da

aplicação de algoritmos de *Machine Learning* (ML) na previsão de conflitos.

5.1. Síntese do Desempenho dos Algoritmos na Literatura

O resultado mais proeminente da análise bibliográfica é a constatação de que os modelos de *Machine Learning* superam consistentemente os métodos estatísticos tradicionais em acurácia. Estudos comparativos, como o realizado por Alarifi et al. (2020), quantificam essa superioridade. A tabela abaixo sintetiza o desempenho de diferentes classificadores, utilizando a métrica AUC.

Tabela 1 – Comparativo de Desempenho de Classificadores de ML (AUC).

Classificador	AUC
Floresta Aleatória (Random Forest)	0.95
Máquina de Vetores de Suporte (SVM)	0.92
Regressão Logística	0.89
Naïve Bayes	0.86

Fonte: Adaptado de Alarifi et al. (2020).

Além do desempenho geral, a literatura revela quais fatores são mais influentes. Para modelos de curto prazo, a **média do "tom" da mídia** e a **frequência de protestos** são cruciais (PERRY, 2013). Já para modelos de médio prazo, fatores estruturais como **histórico de instabilidade**, **mortalidade infantil** e o **tipo de regime político** demonstram maior peso preditivo (SÁNCHEZ-MYERS et al., 2021). A literatura aponta que a combinação de dados estruturais com dados de eventos tende a gerar os modelos mais precisos.

5.2. Discussão dos Resultados

Os resultados obtidos respondem adequadamente aos objetivos específicos. A evidência coletada permite afirmar que a aplicação de ML é uma abordagem metodologicamente válida e empiricamente eficaz. Contudo, uma discussão aprofundada exige que estes resultados promissores sejam confrontados com os limites da previsibilidade e os riscos éticos. O primeiro aspecto negativo é a inerente incapacidade desses modelos de prever eventos do tipo **Cisne Negro** (TALEB, 2021). Conflitos que emergem de dinâmicas totalmente novas, sem precedentes no conjunto de treinamento, permanecerão imprevisíveis (G1, 2021).

O segundo aspecto negativo reside no risco de que esses sistemas se tornem **Algoritmos de Destruição em Massa (WMDs)** (O'NEIL, 2020). A análise da

literatura revela que os dados sobre conflitos são particularmente suscetíveis a vieses sistêmicos (GONG et al., 2023). Entre os principais problemas identificados, destacam-se: (1) o **viés da cobertura da mídia**, onde bases de dados de eventos podem super-representar conflitos em áreas de interesse geopolítico (SÁNCHEZ-MYERS et al., 2021); (2) a **censura como dado faltante**, pois em regimes autoritários a supressão da livre imprensa cria uma ausência artificial de eventos de protesto (O'NEIL, 2020); e (3) a **defasagem temporal dos dados estruturais**, que são atualizados com baixa frequência (SÁNCHEZ-MYERS et al., 2021).

5.3. Resposta ao Problema de Pesquisa

O problema enunciado na introdução questionava a capacidade dos métodos tradicionais de prever com precisão a ocorrência de conflitos. Com base nos resultados, a pesquisa oferece uma resposta clara: **sim, os algoritmos de *Machine Learning* constituem uma ferramenta comprovadamente mais eficaz e com maior acurácia para a análise antecipada de conflitos em comparação aos métodos tradicionais.**

Esta afirmação, contudo, deve ser qualificada. O sucesso da abordagem de ML está condicionado a três fatores cruciais: (1) o acesso a dados de alta qualidade; (2) a escolha criteriosa de algoritmos e métricas; e (3) a implementação de salvaguardas éticas robustas para garantir a transparência e a justiça dos modelos (O'NEIL, 2020; GONG et al., 2023). Portanto, o ML não deve ser visto como um oráculo infalível, mas como um poderoso instrumento de auxílio à decisão.

6. CONSIDERAÇÕES FINAIS

Este trabalho se propôs a investigar a aplicação de algoritmos de *Machine Learning* como uma ferramenta avançada para a análise antecipada de conflitos. A pesquisa atingiu seus objetivos, demonstrando que o ML oferece um potencial preditivo significativamente superior. Contudo, a conclusão mais profunda desta análise não reside na celebração de uma solução tecnológica, mas no reconhecimento de suas complexas interações com o mundo real. O estudo revela que a eficácia de um modelo preditivo enfrenta um "paradoxo da ação": a lacuna entre a capacidade técnica de prever um conflito e a vontade ou capacidade política de preveni-lo (SÁNCHEZ-MYERS et al., 2021). Um alerta, por mais preciso, não se traduz

automaticamente em intervenção, sendo frequentemente neutralizado por questões de soberania ou interesses geopolíticos.

Diante desse cenário, a pesquisa aponta que o paradigma mais promissor não é o do oráculo autônomo, mas o do "Humano no Circuito" (*Human-in-the-Loop*). Os resultados indicam que o maior valor do *Machine Learning* reside em sua capacidade de aumentar a cognição do analista humano, e não em substituí-la. A tecnologia deve servir como um poderoso instrumento para sinalizar riscos, mas a interpretação final e a decisão ética devem permanecer no domínio da responsabilidade humana. Este modelo de simbiose é a salvaguarda mais eficaz contra a transformação de sistemas preditivos em "Algoritmos de Destruição em Massa" (O'NEIL, 2020).

Por fim, esta investigação nos convida a uma reflexão de longo prazo sobre a trajetória da Inteligência Artificial na esfera da segurança global. Atualmente, utilizamos a IA como uma ferramenta para analisar e prever conflitos *humanos*. No entanto, à medida que avançamos para sistemas cada vez mais autônomos, o desafio se transformará radicalmente. A questão deixará de ser apenas "como podemos prever conflitos?" para se tornar "como podemos garantir que futuras superinteligências permaneçam alinhadas com os valores humanos de paz e estabilidade?" (BOSTROM, 2018; TEGMARK, 2017).

Como sugestão para pesquisas futuras, portanto, além da exploração de modelos híbridos e da busca por uma IA Explicável (XAI), torna-se imperativo o aprofundamento nos estudos sobre o "problema do alinhamento" da IA. A tarefa fundamental para as próximas décadas será não apenas aperfeiçoar nossos oráculos digitais, mas garantir que as inteligências que criamos sejam fundamentalmente programadas para a cooperação e a preservação da estabilidade global.

REFERÊNCIAS

- ALARIFI, Abdulrahman et al. Real-time conflict prediction - a comparative study of machine learning classifiers. **International Journal of Advanced Computer Science and Applications**, v. 11, n. 1, 2020.
- ALPAYDIN, Ethem. **Introduction to Machine Learning**. 4. ed. Cambridge: MIT Press, 2020.
- ANOMALO. **Data Quality in Machine Learning: Best Practices and Techniques**. 2024. Disponível em: <https://www.anomalo.com/blog/data-quality-in-machine-learning-best-practices-and-techniques>. Acesso em: 07 set. 2025.
- BISHOP, Christopher M. **Pattern Recognition and Machine Learning**. New York: Springer, 2006.
- BOSTROM, Nick. **Superinteligência: Caminhos, perigos, estratégias**. Tradução de B. S. T. V. Pires. Rio de Janeiro: Darkside, 2018.
- DATA CAMP. **8 modelos de aprendizado de máquina explicados em 20 minutos**. 2024. Disponível em: <https://www.datacamp.com/pt/blog/machine-learning-models-explained>. Acesso em: 07 set. 2025.
- DATAGEEKS. **Principais algoritmos de Machine Learning**. 2019. Disponível em: <https://www.datageeks.com.br/algoritmos-de-machine-learning/>. Acesso em: 07 set. 2025.
- G1. **Como ciência tenta prever os eventos 'cisnes negros'**. G1, 04 out. 2021. Disponível em: <https://g1.globo.com/ciencia/noticia/2021/10/04/como-ciencia-tenta-prever-os-eventos-cisnes-negros.g.html>. Acesso em: 07 set. 2025.
- GONG, Youdi et al. A survey on dataset quality in machine learning. **Information and Software Technology**, v. 162, 2023. DOI: <https://doi.org/10.1016/j.infsof.2023.107268>.
- IBM. **O que é um Algoritmo de Aprendizado de Máquina?**. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/topics/machine-learning-algorithms>. Acesso em: 07 set. 2025.
- LOEWE, Michael; BLACKER, Carmen (Ed.). **Divination and Oracles**. London: George Allen & Unwin, 1981.
- O'NEIL, Cathy. **Algoritmos de Destruição em Massa**. Tradução de Rafael Abraham. São Paulo: Editora Rua do Sabão, 2020.
- PERRY, Chris. Machine learning and conflict prediction: a use case. **Stability: International Journal of Security & Development**, v. 2, n. 3, p. 1-18, 2013. DOI: <http://dx.doi.org/10.5334/sta.cr>.
- RABELLO, Ivonete de Souza. **O Futuro no Passado: Estudo sobre os Oráculos na Obra de Heródoto**. 2013. Tese (Doutorado em Letras Clássicas) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2013.
- SÁNCHEZ-MYERS, A., et al. **Predictive Technologies in Conflict Prevention: A Primer on Machine Learning and Early Warning**. New York: United Nations University, Centre for Policy Research, 2021.
- SOUZA, Leandro de, et al. Análise comparativa entre métodos estatísticos tradicionais e baseados em inteligência artificial no problema de previsão de demanda. **Revista Produção Online**, v. 18, n. 2, p. 703-727, 2018.
- TALEB, Nassim Nicholas. **A Lógica do Cisne Negro**. 3. ed. Tradução de Marcelo Schild. Rio de Janeiro: Best Business, 2021.
- TALEB, Nassim Nicholas. **Antifrágil: Coisas que se Beneficiam com o Caos**. Tradução de Pedro M. S. P. Elói. Rio de Janeiro: Objetiva, 2012.
- TEGMARK, Max. **Vida 3.0: O ser humano na era da inteligência artificial**. Tradução de Érico Assis. São Paulo: Benvirá, 2017.