INTELIGÊNCIA ARTIFICIAL NA GESTÃO DA INVESTIGAÇÃO CRIMINAL

Luís Henrique Costa Ferreira *

RESUMO: Este artigo apresenta um estudo aplicado que investiga a utilização da inteligência artificial (IA), por meio de redes neurais artificiais e técnicas de aprendizado de máquina supervisionado, na gestão da investigação criminal. O problema central abordado consiste em como incorporar o conhecimento empírico do investigador criminal a sistemas computacionais capazes de analisar, de forma precisa e em grande escala, boletins de ocorrência relacionados ao crime de estelionato. O objetivo é identificar, nos relatos registrados, indícios que possibilitem a aplicação do Método de Investigação do Rastejamento. A metodologia adotada compreendeu o desenvolvimento de um modelo computacional baseado em rede neural artificial, utilizando o pacote *ORANGE Data Mining*, a partir de um corpus de 500 boletins de ocorrência classificados por um analista humano. Os resultados demonstraram acurácia, sensibilidade e precisão superiores a 70%, com desempenho satisfatório na comparação com a classificação humana. Conclui-se que o modelo desenvolvido é viável como ferramenta de apoio à decisão, com potencial de transformar conhecimento tácito em ativo institucional. Recomenda-se o aperfeiçoamento do modelo e a ampliação para outros tipos penais, considerando o avanço de tecnologias mais robustas de IA e suas possíveis aplicações na segurança pública.

Palavras-chave: inteligência artificial; rede neural artificial; investigação criminal; estelionato; segurança pública.

DOI: https://doi.org/10.36776/ribsp.v7i19.253

Recebido em 3 de novembro de 2024.	Aprovado em 12 de dezembro de 2024

* Polícia Civil da Bahia e Instituto Geográfico e Histórico da Bahia. CV lattes: http://lattes.cnpq.br/8590658991191685

ARTIFICIAL INTELLIGENCE IN CRIMINAL INVESTIGATION MANAGEMENT

ABSTRACT: This paper presents an applied study that investigates the use of artificial intelligence (AI), through artificial neural networks and supervised machine learning techniques, in the management of criminal investigation. The central issue addressed is how to incorporate the empirical knowledge of criminal investigators into computational systems capable of accurately analyzing, on a large scale, police reports related to fraud. The objective is to identify in these reports signs that indicate the applicability of the Crawl Investigation Method. The methodology involved the development of a computational model based on artificial neural networks, using the ORANGE Data Mining package and a corpus of 500 police reports manually classified by a human analyst. The results showed accuracy, sensitivity, and precision above 70%, with satisfactory performance when compared to human classification. It is concluded that the developed model is feasible as a decision-support tool, with the potential to turn tacit knowledge into an institutional asset. The study recommends model improvement and extension to other criminal types, considering the evolution of more robust AI technologies and their potential applications in public security.

Keywords: artificial intelligence; artificial neural network; criminal investigation; fraud; public security.

1. INTRODUÇÃO

O mês de abril do ano de 2023, ocorreu, o seminário "Inteligência Artificial no STF: a O avanço exponencial das tecnologias de inteligência artificial (IA) tem proporcionado transformações significativas em diversos setores da administração pública, inclusive na segurança pública. No campo da investigação criminal, onde predominam decisões complexas e volumes massivos de dados, a utilização de sistemas computacionais inteligentes surge como alternativa promissora para potencializar a eficiência, a celeridade e a qualidade das análises realizadas por investigadores. Nesse cenário, a possibilidade de incorporar o conhecimento tácito dos profissionais de polícia judiciária aos sistemas inteligentes representa um marco no aprimoramento da gestão da atividade investigativa.

A relevância do tema reside na crescente demanda por instrumentos que auxiliem a superação de gargalos estruturais, como o excesso de registros, a escassez de pessoal qualificado e a assimetria na distribuição de casos. Apesar da ampla digitalização dos boletins de ocorrência nas últimas décadas, observa-se que boa parte das informações neles contidas permanece subutilizada, especialmente aquelas inscritas nos campos de texto livre. Assim, uma lacuna importante se configura: como transformar esses relatos em dados interpretáveis por máquinas, capazes de auxiliar na identificação de padrões e no direcionamento estratégico das investigações criminais?

O presente estudo parte da seguinte problemática: seria possível aplicar redes neurais artificiais para simular, com acurácia aceitável, a classificação humana de boletins de ocorrência quanto à possibilidade de aplicação do Método de Investigação do Rastejamento em casos de estelionato? Duas hipóteses são levantadas: (i) que redes neurais, quando treinadas com base no julgamento de especialistas, são capazes de reproduzir padrões decisórios com precisão satisfatória; e (ii) que a aplicação da IA pode transformar o conhecimento empírico dos investigadores em um ativo institucional, promovendo sua replicação e continuidade.

O objetivo geral é analisar a viabilidade de utilização da inteligência artificial, especificamente por meio de redes neurais artificiais, como ferramenta de apoio à tomada de decisão na atividade de investigação criminal. Como objetivos específicos, busca-se: (a) estruturar um modelo supervisionado de classificação de boletins de ocorrência de estelionato, com base em conhecimento humano previamente consolidado; e (b) mensurar o desempenho do modelo quanto à sua capacidade de simular a classificação realizada por analistas humanos.

A metodologia adotada é de natureza aplicada, com abordagem quantitativa e experimental. O estudo desenvolveu um modelo computacional utilizando o *software ORANGE Data Mining*, baseado em um corpus de 500 boletins de ocorrência já classificados quanto à pertinência do uso do Método de Rastejamento. O modelo foi treinado, testado e avaliado quanto a critérios de acurácia, precisão e sensibilidade.

A principal contribuição da pesquisa reside em demonstrar que sistemas de inteligência artificial podem ser integrados ao processo decisório policial sem substituir o fator humano, mas ampliando sua capacidade de ação. Ao final, argumenta-se pela adoção institucional de modelos semelhantes, respeitadas as especificidades de cada tipo penal, como forma de modernizar e qualificar a gestão da investigação criminal.

O artigo está estruturado em cinco seções, além desta introdução: a segunda seção apresenta o referencial teórico sobre inteligência artificial e sua aplicação na segurança pública; a terceira detalha os procedimentos metodológicos; a quarta expõe os resultados obtidos e suas análises; e a quinta seção traz as conclusões, limitações e sugestões para estudos futuros.

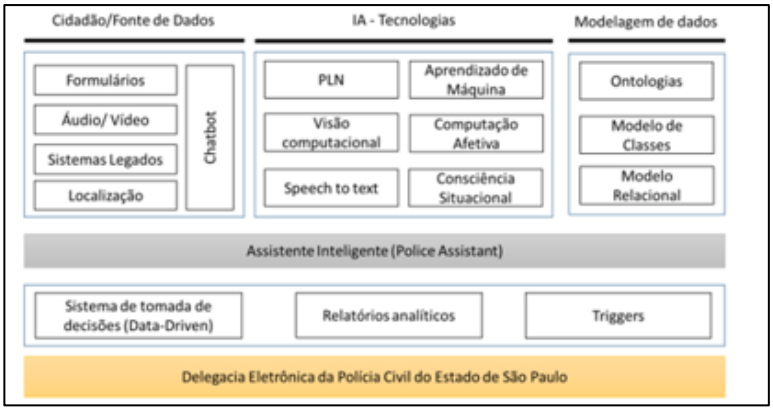
2. REFERENCIAL TEÓRICO

Rosa (2011, p 3) apresenta a Inteligência Artificial (IA) como sendo "o estudo de como fazer os computadores realizarem tarefas as quais, até o momento, os homens fazem melhor". Quais seriam essas tarefas? O mesmo autor responde que são as relacionadas "com o processamento simbólico, reconhecimento de imagens e tudo que envolva 'aprendizado'" (Rosa, 2011, p. 3).

Santos, Neto e Pereira (2020) após estudarem o Boletim de Ocorrência Eletrônico - BOE utilizado pela Polícia Civil do Estado de São Paulo, apresentaram trabalho propondo o uso da IA como ferramenta de apoio ao usuário no sentido de orientá-lo a fornecer informações mínimas para a tomada de decisão sobre a condução da investigação criminal. Para tanto, a ferramenta deve considerar o emprego de linguagem natural (Santos; Neto; Pereira, 2020).

Na Figura 1 está o modelo conceitual para o tratamento do BOE proposto por Santos, Neto e Pereira (2020):

Figura 1 – Modelo conceitual para o Boletim de Ocorrência Eletrônico – Polícia Civil, Estado de São Paulo



Fonte: Santos, Neto e Pereira (2020).

Isso nos leva as maneiras como um problema poderá ser abordado pela (Lima; Pinheiro; Santos, 2014):

a) Simbólica: baseada na hipótese de um sistema de símbolos definidos, em que há um conjunto de estruturas simbólicas e um conjunto de regras de manipulação dessas estruturas.

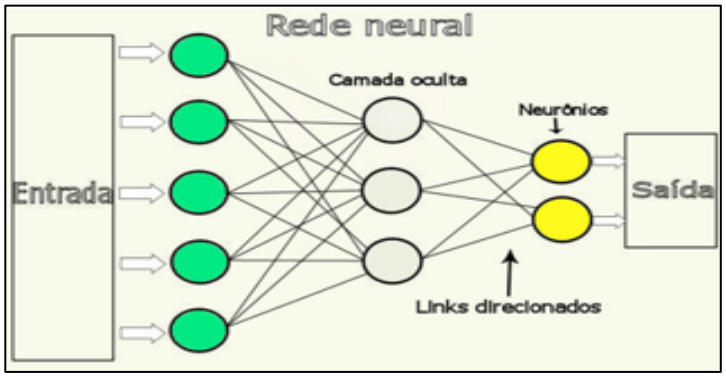
b) Conexionista: "baseia-se na hipótese de causa-efeito, segundo a qual um modelo suficientemente preciso de neurônios basta para reproduzir a inteligência humana. Essa abordagem trata de problemas imprecisos, mas que podem ser definidos através de exemplos".

c) Evolucionária: baseia-se na hipótese de que um sistema inteligente pode se modelado simulando a evolução de uma população.

d) Conjuntos difusos, Conjuntos aproximados: abordagens baseadas nas teorias dos conjuntos difusos e dos conjuntos aproximados, aplicados no tratamento de informações inconsistentes.

Na esfera da abordagem conexionista, surgem as redes neurais artificiais. As RNA são modelos computacionais conexionistas com capacidade de adaptar, aprender, generalizar, agrupar ou organizar dados. As unidades de processamento desse modelo são chamadas de neurônios e executam as tarefas simulando o comportamento de um neurônio biológico (Lima; Pinheiro; Santos, 2014). De forma geral, a operação de um neurônio da rede se resume em: I) Sinais são apresentados à entrada; II) cada sinal é multiplicado por um peso que indica sua influência na saída da unidade; III) é feita a soma ponderada dos sinais que produz um nível de atividade, então se esse nível excede um limite (*threshold*), IV) a unidade produz uma saída (Oliveira, 2011). A propósito, a Figura 2 exibe uma representação esquemática de um sistema de rede neural artificial.

Figura 2 – Representação esquemática de um sistema de rede neural artificial



Fonte: Oliveira (2011).

Na utilização de RNA no ambiente jurídico, Ferreira (2018), em trabalho de TCC, concluiu que os documentos jurídicos são computacionalmente separáveis. Ademais, Sousa (2016) apresenta as redes neurais artificiais como ferramenta cabível para a identificação de padrões e atividades que

envolvam *data mining*, classificação de documentos, previsões financeiras, organização e busca em bancos de dados multimídia, biometria e outras.

Os sistemas de computação neurais são dinâmicos e autoadaptáveis, ou seja, podem mudar para responder satisfatoriamente a novos estímulos e podem se autoajustar usando da experiência para modificar as respostas em determinadas situações. A adaptabilidade do sistema é dependente de três processos: O aprendizado, o treinamento e a generalização. O aprendizado é a autoadaptação do sistema, quando ele é ajustado para encontrar os resultados mais corretos. Com o treinamento, ou aprendizado de máquina, o sistema adapta-se e ajusta-se, é o modo pelo qual ele aprende a respeito da informação que necessitará para resolver o problema. O aprendizado pode ser supervisionado quando o sistema recebe um conjunto de padrões contendo os pares entrada e saída associados, ou não supervisionado quando recebe apenas os dados de entrada e deduz a saída. A generalização é a capacidade de responder de maneira adequada a padrões que não fizeram parte do conjunto de treinamento (Lima; Pinheiro; Santos, 2014).

A condução dos processos que resultarão no aprendizado, no treinamento e na generalização requer perícia e cuidados, uma vez que malconduzidos poderão gerar cenários de *overfitting* ou de *underfitting*. Os dois cenários são situações relacionadas ao aprendizado de máquina. O primeiro é o sobreajuste, quando o modelo é incapaz de generalizar. O segundo é o sub-ajuste, quando o modelo não atende as expectativas (Grus, 2016). A primeira situação é contornada com a seleção e preparo dos dados de treinamento, já a outra, impõe a substituição do modelo.

Considerando o escopo deste trabalho, convém limitar a interpretação do que consiste em classificar um elemento. Um problema de classificação em termos de aprendizado máquina consiste em conceber um modelo computacional capaz de receber um estímulo (ou dado, ou registro) e prever o conjunto ao qual ele pertence. Nesse sentido frisa-se que classificação difere da clusterização pois esta procura semelhanças e diferenças num conjunto de dados e agrupa os registros semelhantes em segmentos ou *clusters*.

Para a condução da pesquisa utilizou-se o *ORANGE Data Mining*, que é um *software* de *open source*, voltado para *machine learning* e mineração de dados e desenvolvido em *script* Python adotando o conceito de *workflow*. O programa destina-se a programadores experientes e estudantes de mineração de dados (Demsar *et al.*, 2013). A escolha do pacote considerou o fato dele ser dotado de uma biblioteca projetada para simplificar a montagem de fluxos de trabalho para análise de dados, bem como fazer abordagens de mineração de dados a partir de uma combinação de componentes (*widgets*) (Demsar *et al.*, 2013). Já o *software* Nvivo, consiste em um programa de indexação e categorização de dados não estruturados, focado em pesquisas que baseadas na teoria fundamentada nos dados[1] (Freitas; Arruda; Falqueto, 2017), ele foi utilizado como ferramenta de apoio na análise e classificação do *corpus* conduzida pelo analista humano.

[1] *Grounded theory*.

Luiz Ribeiro (2012) apresenta a prática de investigação criminal denominada de Método do Rastejamento como o procedimento de seguir as pistas, ou indícios, encontrados em ação similar ao rastrear. O investigador criminal tendo encontrado indício, ou pista, deve segui-lo, rastreá-lo em busca de dados e/ou de outras pistas. Para o autor o método é bastante eficaz quando utilizado em investigação de homicídio, ao nosso ver o método pode ser o ponto de partida para qualquer investigação criminal.

3. ASPECTOS METODOLÓGICOS

O objetivo geral desta pesquisa, de caráter aplicada, foi estudar como levar o conhecimento e a capacidade de interpretação do investigador criminal para o computador e com isso analisar de maneira rápida e precisa grandes volumes de boletins de ocorrências de modo obter alertas capazes de subsidiar a gestão da investigação.

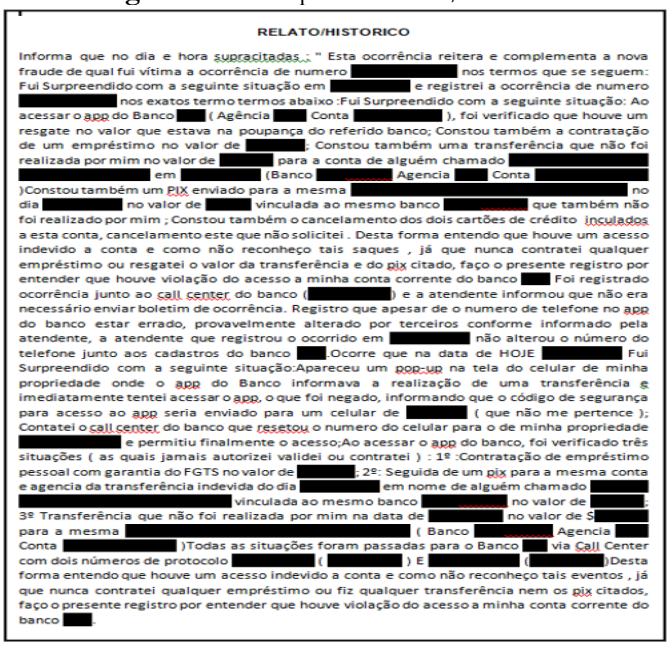
Como objetivo especifico foi estabelecido a concepção e teste de um modelo computacional capaz de, com eficiência, sensibilidade e precisão, processar linguagem natural e analisar boletins de ocorrências tipificados como estelionato classificando-os como: a) documentos onde estão expostos eventos nos quais é provável que exista linha telefônica para ser rastreada; b) documentos onde estão expostos eventos nos quais é provável que exista conta bancária para ser rastreada e c) documentos onde estão expostos eventos nos quais é provável que existam linha telefônica e conta bancária para serem rastreadas. Tudo com a finalidade de alertar o investigador criminal para a possibilidade de existência de indício, e/ou pista sobre o qual possa ser aplicado o Método de Investigação do Rastejamento.

Os documentos trabalhados foram Certidões de boletins de ocorrências classificados como estelionato. A Certidão em questão possui formato semiestruturado no qual podem constar as identificações, com qualificações, do comunicante, da vítima, do suspeito e do envolvido, dados do fato, como local e datas, as coisas relacionadas com o fato, identificação das unidades policiais onde foi efetuado registro e responsável pela apuração etc. Os documentos foram extraídos para arquivos digitais no formato PDF[2] com a estrutura passível de impressão direta.

Dentre os campos existentes na Certidão o de interesse para este estudo é o identificado pelo título Relato/Histórico, pois nele está a narrativa circunstanciada do fato ocorrido, redigida em linguagem natural por um agente do Estado ou diretamente pelo comunicante. Nas Figuras 3, 4 e 5 estão exibidos exemplos do campo Relato/Histórico conforme constam nos arquivos, exceto pela omissão dos dados interpretados como sensíveis.

[2] "PDF é a abreviação de *Portable Document Format* (formato portátil de documento). É um formato de arquivo versátil criado pela Adobe que proporciona uma maneira fácil e confiável de apresentar e compartilhar documentos em qualquer software, hardware ou sistema operacional usado pela pessoa que exibe o documento" (Adobe, 2024).

Figura 3 – Exemplo de Relato/Histórico

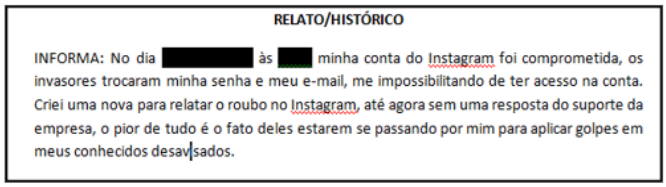


RELATO/HISTORICO

Informa que no dia e hora supracitadas : " Esta ocorrência reitera e complementa a nova fraude de qual fui vítima a ocorrência de numero [redacted] nos termos que se seguem: Fui Surpreendido com a seguinte situação em [redacted] e registrei a ocorrência de numero [redacted] nos exatos termo termos abaixo :Fui Surpreendido com a seguinte situação: Ao acessar o app do Banco [redacted] (Agência [redacted] Conta [redacted]), foi verificado que houve um resgate no valor que estava na poupança do referido banco; Constou também a contratação de um empréstimo no valor de [redacted]; Constou também uma transferência que não foi realizada por mim no valor de [redacted] para a conta de alguém chamado [redacted] em [redacted] (Banco [redacted] Agencia [redacted] Conta [redacted])Constou também um PIX enviado para a mesma [redacted] no dia [redacted] no valor de [redacted] vinculada ao mesmo banco [redacted] que também não foi realizado por mim ; Constou também o cancelamento dos dois cartões de crédito inculados a esta conta, cancelamento este que não solicitei . Desta forma entendo que houve um acesso indevido a conta e como não reconheço tais saques , já que nunca contratei qualquer empréstimo ou resgatei o valor da transferência e do pix citado, faço o presente registro por entender que houve violação do acesso a minha conta corrente do banco [redacted] Foi registrado ocorrência junto ao call center do banco ([redacted]) e a atendente informou que não era necessário enviar boletim de ocorrência. Registro que apesar de o numero de telefone no app do banco estar errado, provavelmente alterado por terceiros conforme informado pela atendente, a atendente que registrou o ocorrido em [redacted] não alterou o número do telefone junto aos cadastros do banco [redacted] .Ocorre que na data de HOJE [redacted] Fui Surpreendido com a seguinte situação:Apareceu um pop-up na tela do celular de minha propriedade onde o app do Banco informava a realização de uma transferência e imediatamente tentei acessar o app, o que foi negado, informando que o código de segurança para acesso ao app seria enviado para um celular de [redacted] (que não me pertence); Contatei o call center do banco que resetou o numero do celular para o de minha propriedade [redacted] e permitiu finalmente o acesso;Ao acessar o app do banco, foi verificado três situações (as quais jamais autorizei validei ou contratei) : 1º :Contratação de empréstimo pessoal com garantia do FGTS no valor de [redacted]; 2º: Seguida de um pix para a mesma conta e agencia da transferência indevida do dia [redacted] em nome de alguém chamado [redacted] vinculada ao mesmo banco [redacted] no valor de [redacted]; 3º Transferência que não foi realizada por mim na data de [redacted] no valor de S[redacted] para a mesma [redacted] (Banco [redacted] Agencia [redacted] Conta [redacted])Todas as situações foram passadas para o Banco [redacted] via Call Center com dois números de protocolo [redacted] ([redacted]) E [redacted] ([redacted])Desta forma entendo que houve um acesso indevido a conta e como não reconheço tais eventos , já que nunca contratei qualquer empréstimo ou fiz qualquer transferência nem os pix citados, faço o presente registro por entender que houve violação do acesso a minha conta corrente do banco [redacted]

Fonte: O Autor.

Figura 4 – Exemplo de Relato/Histórico

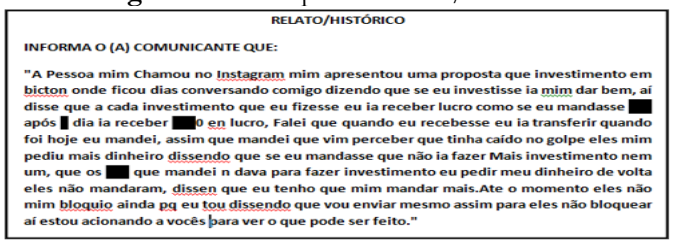


RELATO/HISTÓRICO

INFORMA: No dia [redacted] às [redacted] minha conta do Instagram foi comprometida, os invasores trocaram minha senha e meu e-mail, me impossibilitando de ter acesso na conta. Criei uma nova para relatar o roubo no Instagram, até agora sem uma resposta do suporte da empresa, o pior de tudo é o fato deles estarem se passando por mim para aplicar golpes em meus conhecidos desavisados.

Fonte: O Autor.

Figura 5 – Exemplo de Relato/Histórico



RELATO/HISTÓRICO

INFORMA O (A) COMUNICANTE QUE:

"A Pessoa mim Chamou no Instagram mim apresentou uma proposta que investimento em bicton onde ficou dias conversando comigo dizendo que se eu investisse ia mim dar bem, aí disse que a cada investimento que eu fizesse eu ia receber lucro como se eu mandasse [redacted] após [redacted] dia ia receber [redacted]0 en lucro, Falei que quando eu recebesse eu ia transferir quando foi hoje eu mandei, assim que mandei que vim perceber que tinha caído no golpe eles mim pediu mais dinheiro dissendo que se eu mandasse que não ia fazer Mais investimento nem um, que os [redacted] que mandei n dava para fazer investimento eu pedir meu dinheiro de volta eles não mandaram, dissen que eu tenho que mim mandar mais.Ate o momento eles não mim bloquio ainda pq eu tou dissendo que vou enviar mesmo assim para eles não bloquear aí estou acionando a vocês para ver o que pode ser feito."

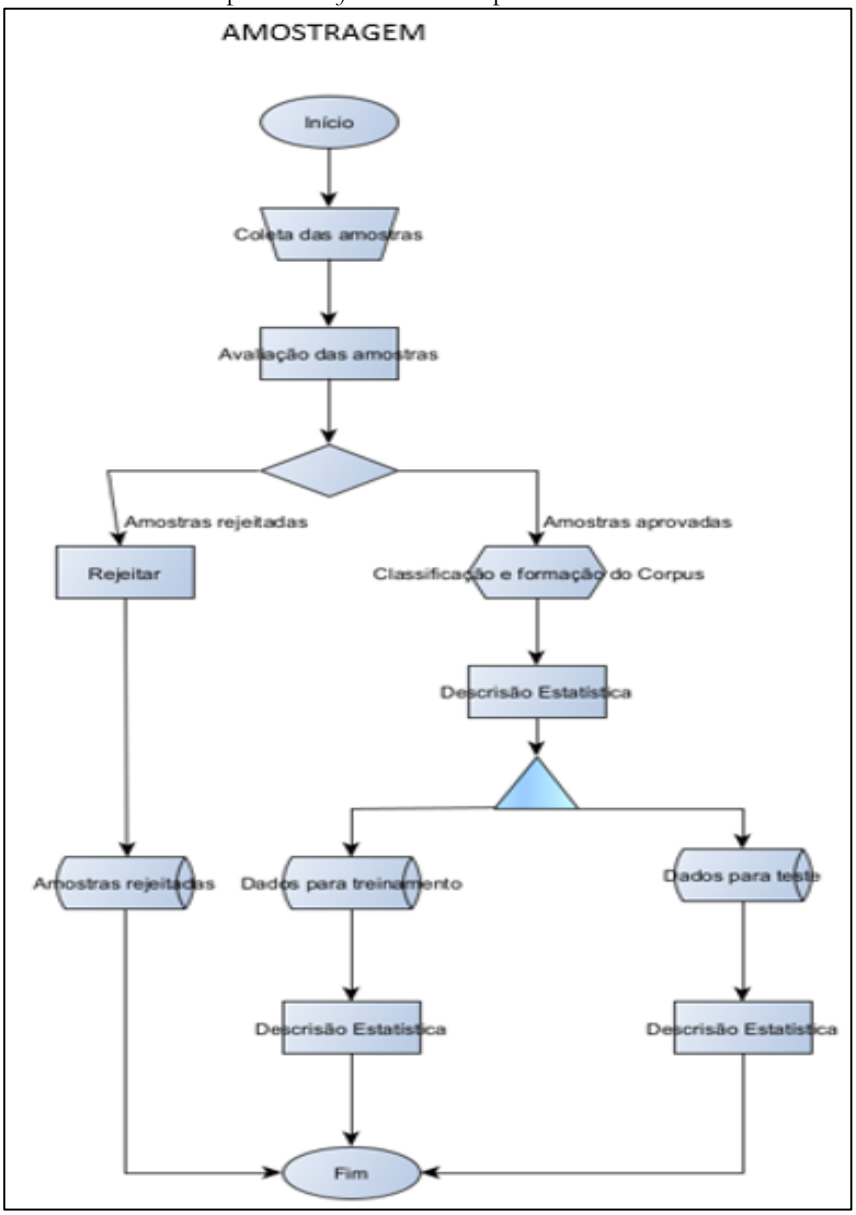
Fonte: O Autor.

Para o escopo deste estudo foi definido, por conveniência, um *corpus* composto por 500 documentos, coletados de maneira aleatória entre os registros de boletins de ocorrências, cadastrados no banco de dados do sistema Sinesp PPE[3] utilizado pela Polícia Civil da Bahia e tipificados como estelionato (artigo 171 do Código Penal Brasileiro). Cada BO gerou um arquivo digital no formato PDF.

[3] O Sinesp PPE – "Procedimentos Policiais Eletrônicos é uma solução disponibilizada pela Senasp/MJSP às Unidades da Federação que permitem o registro de ocorrências, despacho homologatório e a lavratura de procedimentos de polícia, além de fornecer recursos que permitem a gestão cartorária e compartilhamento/uso de dados e informações registrados pelos entes federados participantes" (Brasil, 2019).

A Figura 6 ilustra o processo de coleta e tratamento das amostras, que foi efetuado com o apoio do *software* NVivo para *Windows*.

Figura 6 – Processo de coleta e tratamento das amostras, efetuado com o apoio do *software* NVivo para *Windows*



Fonte: O Autor.

As amostras foram coletadas, de maneira aleatória entre registros efetuados nos anos[4] de 2021, 2022 e 2023. Foram coletadas 500 amostras.

[4] O Sistema Sinesp PPE iniciou operação no estado da Bahia no ano de 2021.

A avaliação das amostras teve como objetivo eliminar os elementos, que apesar de apresentarem tipificação como estelionato, o histórico do fato a ser apurado descreve uma situação, que a critério do avaliador, diverge do tipo penal, como por exemplo alguma desavença comercial, ou qualquer outro negócio jurídico não afeto ao Direito Penal. Situação não difícil de ocorrer quando o registro do BO é conduzido pelo próprio comunicante utilizando a delegacia virtual. Não foram encontrados documentos para serem eliminados nas amostras coletadas.

O corpus ficou composto pelos 500 documentos digitais selecionados após a avaliação. E o passo seguinte foi classificar os textos manualmente, com o apoio do software NVivo, segundo a narrativa contida no espaço destinado para o histórico do fato e de acordo com as seguintes dimensões:

a) local do registro: descreve o meio escolhido pelo comunicante para proceder o registro do boletim de ocorrência. Tem como atributos: a) Delegacia Territorial, para o BO registrado por um preposto do Estado com o comunicante fornecendo as informações dentro de uma unidade da Polícia Civil e b) Delegacia Virtual, atribuído quando o BO é registrado pelo próprio comunicante utilizando a página disponibilizada pela Polícia Civil na Internet.

b) ano do registro: contém o ano em que o registro do BO foi efetuado, assumindo os seguintes valores: 2021, 2022 ou 2023.

c) contato com os autores do golpe: contém informação sobre o meio/instrumento utilizado pelo autor do estelionato para sustentar o contato com a vítima e induzi-la ou mantê-la em erro, pode assumir um dos seguintes atributos: i) telefone, *Whatsapp* ou *Telegram*; ii) pessoalmente; iii) site da *internet*; iv) outras redes sociais utilizando *internet*; v) *email*; ou vi) não identificado no BO.

d) prejuízo ocasionado: descreve como o bem deixou a posse do comunicante e passou para posse do autor da fraude, pode assumir um dos seguintes rótulos: i) uma dívida foi constituída em desfavor da vítima; ii) dinheiro foi transferido entre contas bancárias; iii) dinheiro foi sacado do banco; iv) dinheiro ou coisa foi entregue diretamente pela vítima; v) coisa ou dinheiro foi enviado para um local; vi) rede social ou conta na *internet* sequestrada; ou vii) não identificado no BO.

O *corpus* foi dividido em dois conjuntos:

a) o conjunto de treinamento, destinado a treinar, com o uso de técnicas de aprendizado de máquina, o modelo de inteligência artificial;

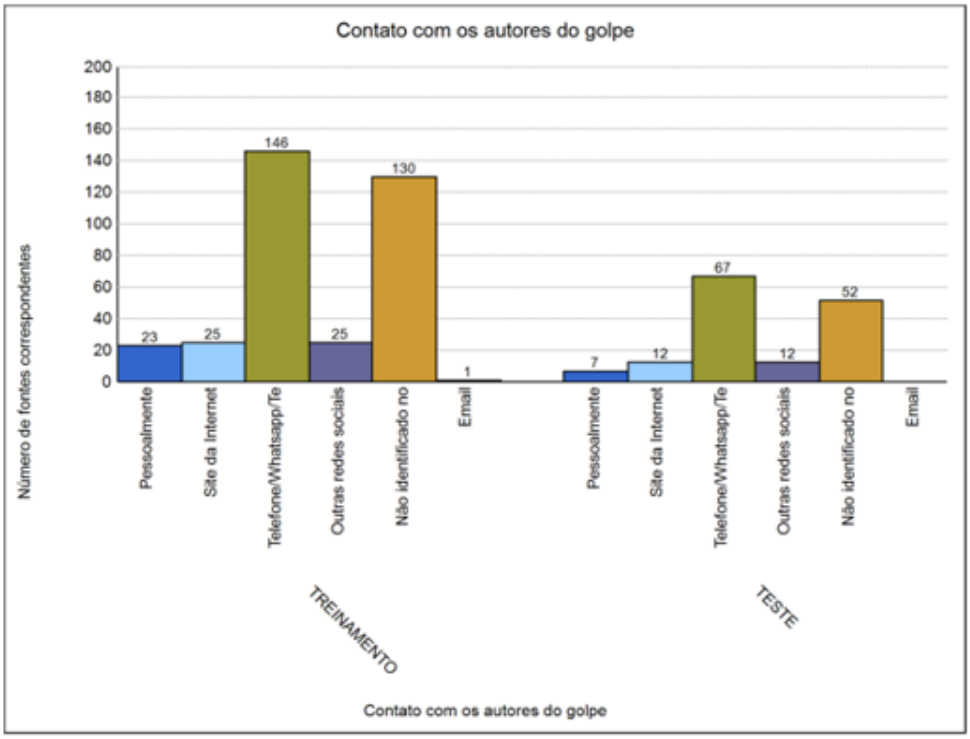
b) o conjunto de teste, que servirá para avaliar a desenvoltura do modelo.

A divisão foi aleatória e obedeceu ao critério 70% para Treinamento e 30% destinados ao teste, resultando em 350 e 150 documentos respectivamente.

As características dos conjuntos de treinamento e de teste estão expostas nos Gráficos 1, 2, 3 e 4.

O Gráfico 1 revela que os meios mais utilizados pelos criminosos foram o telefone e aplicativos de mensagens, como *WhatsApp*, com 146 ocorrências no conjunto de treinamento e 67 no de teste. Em seguida, destacam-se os casos em que o meio de contato não foi identificado (130 no treinamento e 52 no teste). O site da *internet* e outras redes sociais apresentaram frequência moderada e constante, enquanto os contatos pessoais e por e-mail foram pouco representativos. Esses dados indicam uma forte preferência por canais digitais e remotos, que facilitam o anonimato e o alcance das ações fraudulentas. A repetição dos padrões entre os conjuntos analisados sugere uniformidade na estratégia dos autores, evidenciando o potencial desses dados para alimentar sistemas de detecção e prevenção de golpes.

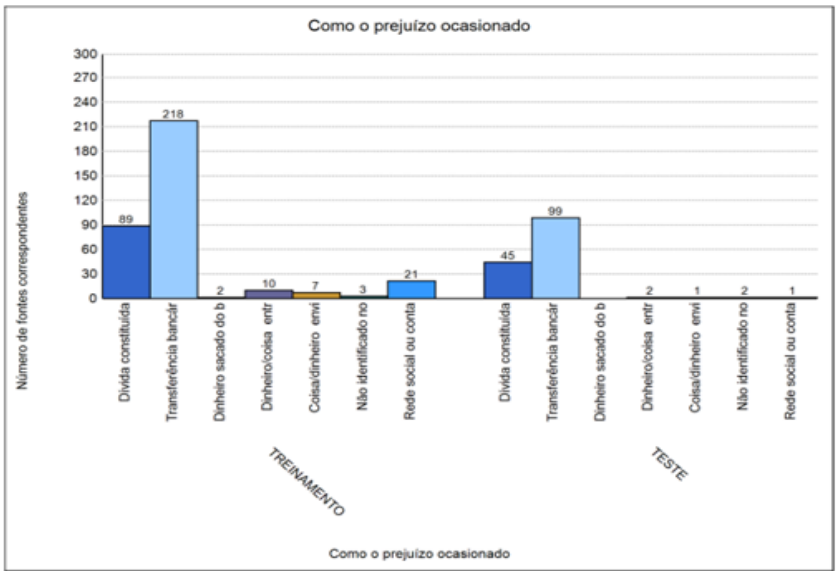
Gráfico 1 – Contato com os autores do golpe



Fonte: O Autor.

O Gráfico 2 demonstra que a principal forma de prejuízo ocasionado às vítimas foi a transferência bancária, com 218 ocorrências na fase de treinamento e 99 na de teste. A segunda forma mais comum foi a constituição de dívida, com 89 casos no treinamento e 45 no teste. As demais categorias, como saque de dinheiro, empréstimos não autorizados, cobranças indevidas e uso de redes sociais ou contas, apresentaram frequência residual, com menos de dez registros em cada fase. A predominância da transferência bancária evidencia a centralidade do sistema financeiro como vetor de prejuízo nas fraudes analisadas, exigindo maior controle e rastreabilidade dessas operações. A repetição do padrão nos dois conjuntos analisados reforça a uniformidade do golpe e sua previsibilidade, o que pode subsidiar ações preventivas e modelos preditivos.

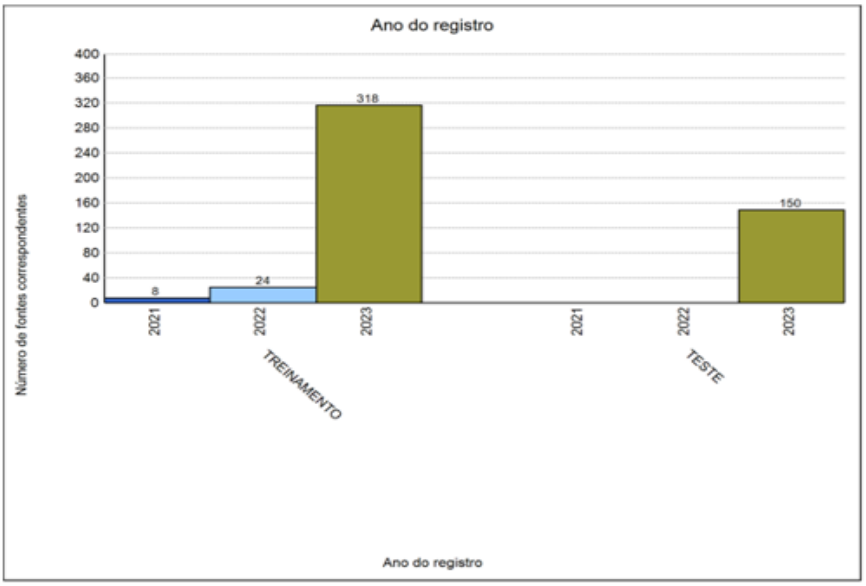
Gráfico 2 – Como o prejuízo ocasionado



Fonte: O Autor.

O Gráfico 3 revela uma clara concentração das ocorrências analisadas no ano de 2023, tanto no conjunto de treinamento (318 registros) quanto no de teste (150 registros). Os anos de 2021 e 2022 apresentaram volume significativamente inferior, com apenas 8 e 24 registros, respectivamente, todos no conjunto de treinamento. Essa distribuição indica que a maioria dos casos de golpes analisados é recente, refletindo um crescimento acentuado das ocorrências em 2023. Tal concentração temporal pode estar associada ao aumento da digitalização de serviços, ao maior acesso a canais online pelos criminosos ou à intensificação da notificação por parte das vítimas.

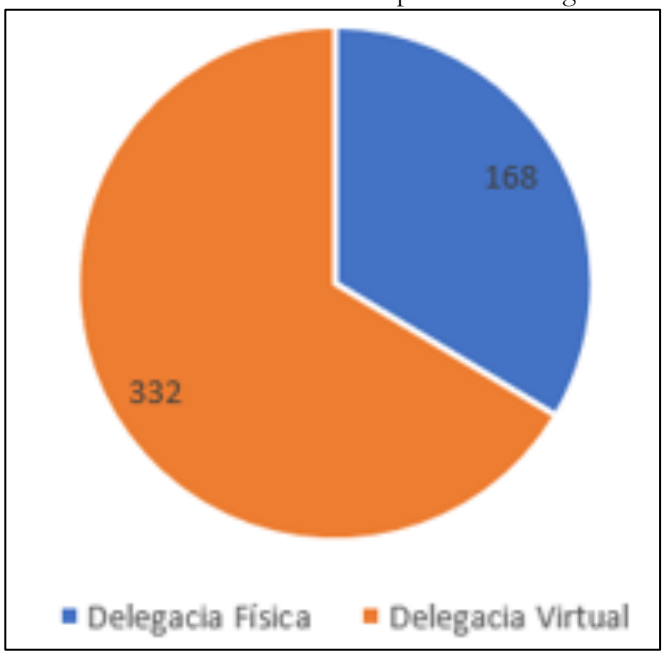
Gráfico 3 – Ano do registro



Fonte: O Autor.

O Gráfico 4 demonstra a distribuição dos registros conforme o local em que foram formalizados: delegacia física ou delegacia virtual. Observa-se que a maior parte dos registros foi realizada por meio da delegacia virtual, com 332 ocorrências (66,4%), enquanto as delegacias físicas concentraram 168 registros (33,6%). Essa predominância do meio virtual evidencia a consolidação das plataformas digitais como principal canal de comunicação entre o cidadão e os órgãos de segurança pública, especialmente em contextos que envolvem golpes eletrônicos. A acessibilidade, a praticidade e o anonimato proporcionados pela delegacia virtual podem estar associados à maior adesão por parte das vítimas.

Gráfico 4 – Número de fontes por local do registro



Fonte: O Autor.

A Figura 7 exibe a Nuvem de Palavras gerada a partir dos documentos incluídos no *corpus*. São as 500 palavras (com no mínimo 3 letras) mais frequentes no conjunto de documentos. Os nomes próprios e as informações consideradas sensíveis foram cobertas com tarjas negras. A nuvem de palavras gerada a partir dos boletins de ocorrência analisados evidencia os termos mais recorrentes nos registros, destacando a centralidade de palavras como "data", "nome", "registro", "ocorrência", "civil" e "delegacia". Esses termos refletem a estrutura padrão dos documentos policiais, os quais priorizam informações cadastrais, de identificação e descrição do fato. Além disso, aparecem com destaque palavras como "polícia", "estado", "bahia", "fato" e "documento", indicando elementos fundamentais para o enquadramento jurídico e administrativo das ocorrências. A presença de termos como "eletrônicos", "telefone" e "whatsapp" sugere a relevância dos meios digitais nos crimes registrados, reforçando o perfil cibernético dos golpes analisados.

Figura 7 – Nuvem de palavras do *corpus*

Fonte: O Autor.

A análise da nuvem de palavras (Figura 7) contribui para compreender a frequência e o contexto dos dados mais relevantes nos relatos, podendo subsidiar tanto a automatização de análises quanto a identificação de padrões narrativos nos registros criminais.

3.1 Concepção do Modelo de Inteligência Artificial

A pretensão da pesquisa foi desenvolver e estudar um modelo de classificação, baseado em conceitos de *machine learning* que, recebendo como estímulo Boletins de Ocorrências registrados com o uso de linguagem natural e tipificados como estelionato, identifique os eventos com possiblidade de uso do Método de Investigação Policial do Rastejamento. A análise preliminar do corpus, fez surgir dois elementos passiveis de rastreamento: a) Linha telefônica e b) Conta bancária. Identificar a possível presença de um desses elementos, ou dos dois, nas histórias registradas nos BO(s) fez com que se buscasse um classificador com boa especificidade.

Os BO(s) foram distribuídos, após avaliação pelo autor, nas seguintes classes:

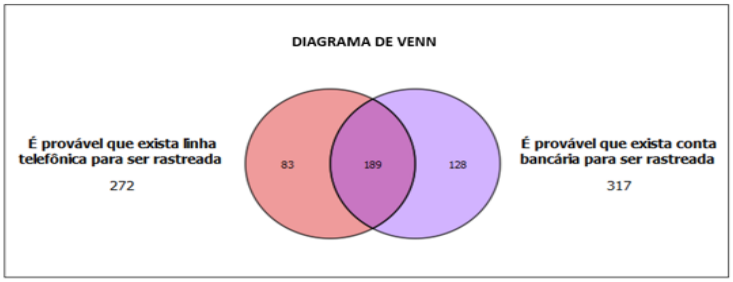
a) É provável que exista linha telefônica para ser rastreada, que recebeu os documentos classificados como Telefone/*Whastapp*/*Telegram* e compõem o Conjunto A.

b) É provável que exista conta bancária para ser rastreada, que recebeu os documentos classificados como Conta bancária e compõem o Conjunto B.

c) Outras situações excluindo telefone e conta bancária, recebeu todos os documentos que não foram incluídos nas classes A e B acima e compõem o Conjunto C.

No Diagrama 1[5] (Figura 8) representa a interseção[6] existente entre os Conjuntos A e B. 189 elementos pertencem a ambos os conjuntos.

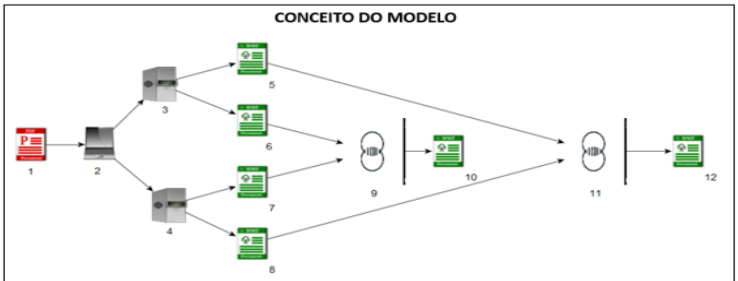
Figura 8 – Diagrama de Venn



Fonte: O Autor.

Sendo assim, considerando a interseção entre os conjuntos A e B, para facilitar o desenvolvimento da pesquisa, foram idealizados dois modelos para classificação dicotômica, ambos baseados em aprendizado de máquina supervisionado, capazes de processar linguagem natural e cujos resultados serão posteriormente integrados (Figura 9).

Figura 9 – Conceito do Modelo



Fonte: O Autor.

As seguintes saídas foram consideradas relevantes para a aplicação do Método de Investigação do Rastejamento, nos termos já expostos:

[5] O Diagrama 1 recebe a denominação Diagrama de Venn e consiste em uma das maneiras de representar um conjunto ou as relações entre dois ou mais conjuntos (Machado, 1988). Nele "os conjuntos são representados por regiões planas interiores a uma curva fechada e simples" (Antar Neto, 2009, p. 32).

[6] Em linguagem matemática o Diagrama 1 representa a operação entre conjuntos: $A \cap B = \{ x \parallel x \in A \text{ e } x \in B\}$, cuja leitura em português é: O conjunto da interseção dos Conjuntos A e B é composto pelo elemento x tal que x pertence ao Conjunto A e x pertence ao Conjunto B.

a) Documentos onde estão expostos eventos nos quais é provável que exista linha telefônica para ser rastreada.

b) Documentos onde estão expostos eventos nos quais é provável que exista conta bancária para ser rastreada.

c) Documentos onde estão expostos eventos nos quais é provável que existam linha telefônica e conta bancária para serem rastreadas.

Para finalizar a análise do corpus e a conceituação do modelo foi testada a hipótese de que as classes Transferência bancária e Telefone/Whatspp/Telegram, estejam associadas. A decisão foi aplicar o Teste Qui-Quadrado para associação aos dados dispotos na Tabela 1.

Tabela 1 – Linha de telefone X Conta bancária

Contato do autor do golpe	Destino do dinheiro subtraído		Total
	Conta bancária	Não conta bancária	
Uso de linha telefônica	189	83	272
Não uso de linha telefônica	128	100	228
Total	**317**	**183**	**500**

Fonte: O Autor.

A Tabela 1 apresenta a distribuição conjunta entre o tipo de contato utilizado pelo autor do golpe (uso ou não de linha telefônica) e o destino do dinheiro subtraído (conta bancária ou outros meios), totalizando 500 ocorrências analisadas. Observa-se que, nos casos em que houve o uso de linha telefônica, 189 ocorrências (69,49%) envolveram a transferência para conta bancária, enquanto 83 casos (30,51%) tiveram como destino meios que não envolveram contas bancárias. Já entre os casos sem uso de linha telefônica, 128 ocorrências (56,14%) resultaram em transferência para conta bancária, e 100 casos (43,86%) em outros destinos.

A análise permite inferir que há uma associação entre o uso de linha telefônica e a destinação do valor subtraído para contas bancárias, sugerindo que o uso do contato telefônico pode estar associado a estratégias mais organizadas de fraude, com o objetivo de induzir a vítima a transferir valores diretamente a contas controladas pelos criminosos. A distribuição dos dados, somando 317 ocorrências com destino bancário (63,4%) e 183 sem envolvimento direto de conta bancária (36,6%), evidencia que o sistema bancário segue sendo um canal predominante de escoamento dos recursos obtidos nos golpes, o que pode indicar a necessidade de aprimoramento dos mecanismos de rastreabilidade, bloqueio e cooperação entre instituições financeiras e órgãos de segurança pública.

Essa tabela pode, ainda, ser submetida a um teste estatístico de associação (como o teste do qui-quadrado de independência), de modo a avaliar a significância da relação entre as variáveis "tipo de contato" e "destino do dinheiro". Caso o valor de p resultante do teste seja inferior a 0,05, poderá ser afirmado, com 95% de confiança, que há associação estatisticamente significativa entre as variáveis analisadas.

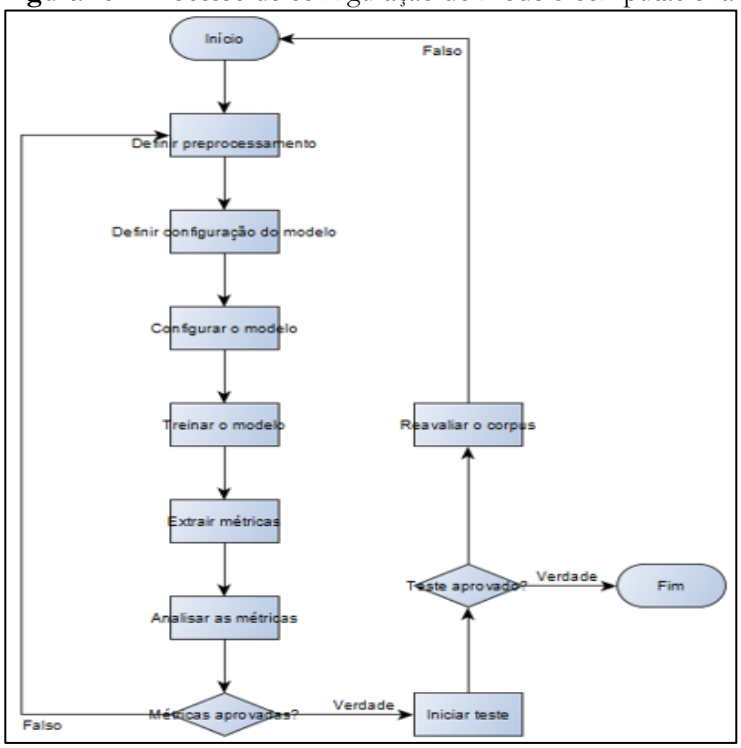
As hipóteses do teste Qui-Quadrado para associação são: H0: Não existe associação (variáveis são independentes) e H1: Existe associação. Para um nível de significância de 0,05 e grau de liberdade GL =1, os resultados do teste foram: Person = Razão de verossimilhança = 9,519 e P valor = 0,002. Logo um P valor < 0,05 leva a rejeição de H0, portanto há evidência de associação entre as classes.

3.2 O Modelo

Estabelecido o conceito do modelo computacional, conforme narrativa anterior, o método escolhido para implementá-lo foi a utilização de rede neural artificial construída com as ferramentas disponibilizadas pelo software ORANGE Data Mining, que fazem uso do algoritmo *Multi Layer Perceptron* – MLP com *backpropagation* do *sklearn*[7] e são capazes de satisfazer modelos computacionais lineares e não lineares. O hardware utilizado foi um Notebook Acer, Aspire 3, com 475 GB de RAM e equipado com o sistema operacional Windows 11.

O Modelo Computacional foi configurado em duas etapas, na primeira foi ensinado a classificar os documentos. Na segunda etapa o modelo treinado foi testado com estímulos não utilizados para o treinamento. As etapas 1 e 2 utilizaram, respectivamente o Conjunto de treino e o Conjunto teste. O fluxograma na Figura 10 exibe o processo para a configuração do modelo computacional.

Figura 10 – Processo de configuração do modelo computacional



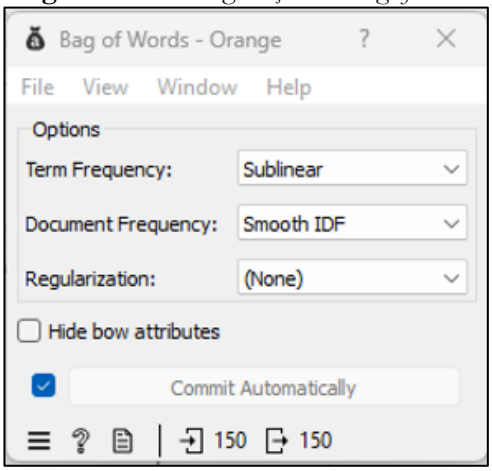
Fonte: O Autor.

[7]A *scikit-learn* é uma biblioteca de aprendizado de máquina de código aberto para a linguagem de programação Python (Wikipedia, 2022).

Para que um modelo de aprendizado de máquina trabalhe com documentos de texto é fundamental que estes sejam submetidos a alguma espécie de pré-processamento. Quando o objetivo é a classificação de textos o pré-processamento pode envolver transformações de palavras e/ou sentença, remoção de palavras e elementos que possam gerar ruídos e outras técnicas isoladas, ou combinadas. Neste trabalho, em específico, o pré-processamento do corpus gerou um arquivo denominado *bag of words*. O pré-processamento consistiu em, entre outras transformações: Remoção de palavras irrelevantes; remoção de determinadas classes de caracteres, (como números, caracteres especiais etc.) e identificação e remoção de e-mail e URL[8]. Foi retirada a maioria das palavras utilizadas na composição do formulário padrão e títulos dos campos. Ficou estabelecido que apenas as palavras mais frequentes no conjunto de textos iriam compor o *bag of words*. Essas palavras formaram os elementos de entrada do classificador. A Figura 11 exibe a configuração do *widget bag of word* e a Figura 12 exibe a Nuvem de Palavras após o pré-processamento. As palavras consideradas desnecessárias, ou passiveis de representar ruído, foram incluídas no arquivo *stopword* para serem excluídas durante o pré-processamento.

A configuração do módulo *Bag of Words* no *software Orange* (Figura 11) foi ajustada com a finalidade de otimizar a representação vetorial dos documentos textuais utilizados na análise. A frequência dos termos foi tratada com a técnica Sublinear TF, que aplica uma transformação logarítmica à contagem de palavras, reduzindo o impacto de termos com alta repetição. Para o cálculo da frequência inversa de documentos, utilizou-se o método *Smooth* IDF, que suaviza os valores e evita divisões por zero em documentos com termos raros. Não foi aplicada regularização, conforme indicado na opção Regularization: (None), mantendo os vetores originais gerados pelo modelo TF-IDF.

Figura 11 – Configuração do *bag of words*



Fonte: O Autor.

[8] O *Uniform Resource Locator* (URL), é um termo técnico (e anglicismo de tecnologia da informação) que foi traduzido para a língua portuguesa como "localizador uniforme de recursos". Um URL se refere ao endereço de rede no qual se encontra algum recurso informático, como por exemplo um arquivo de computador ou um dispositivo periférico (impressora, equipamento multifuncional, unidade de rede etc.). Essa rede pode ser a Internet, uma rede corporativa (como uma intranet) etc (Wikipédia, 2023).

A nuvem de palavras gerada a partir dos relatos textuais contidos nos boletins de ocorrência (Figura 12) evidencia a predominância de termos relacionados a práticas de fraude por meios digitais. Palavras como "whatsapp", "valor", "email", "mensagens", "pix", "redes sociais", "banco", "falsa" e "receber" destacam-se como os termos mais recorrentes, indicando que os golpes analisados envolvem majoritariamente o uso de aplicativos de comunicação, transferências financeiras via PIX e interação por e-mails ou redes sociais. A presença de termos como "conta", "informados", "utilização" e "sms" reforça o padrão de engenharia social associado à coleta de dados sensíveis para fins fraudulentos. A estrutura da nuvem revela a centralidade dos meios digitais como ambiente privilegiado para execução dos golpes, reforçando a necessidade de políticas públicas voltadas à prevenção e educação digital das vítimas, bem como ao rastreamento de movimentações financeiras em ambientes virtuais.

Figura 12 – Nuvem de palavras – pré-processamento

Fonte: O Autor.

O Modelo Computacional escolhido por decisão do pesquisador, entre os disponíveis no catálogo do pacote ORANGE Data Mining, foi rede neural artificial. E definir uma RNA *perceptron* multicamadas, com *backpropagation*, significa, de maneira bastante simplória, estabelecer os seguintes parâmetros:

a) Quantidade de neurônios de entrada.

b) Quantidade de camadas ocultas e suas respectivas quantidade de neurônios.

c) Quantidade de neurônios de saída.

d) A função de ativação a ser aplicada a cada camada.

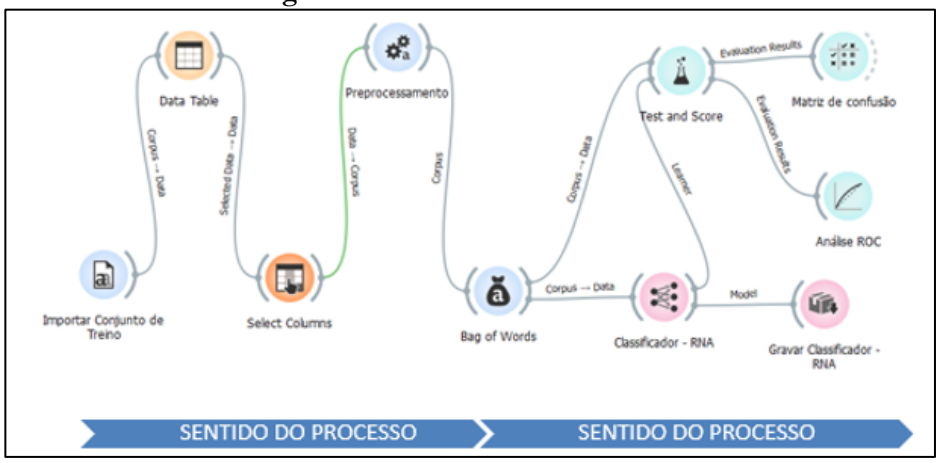
e) O otimizador ALFA.

f) O número de interações, ou épocas, para o treino.

Para a configuração e treino do modelo classificador foram utilizados os recursos do pacote ORANGE Data Mining, especificamente os *widgets* exposto na Figura 13, que exibe o fluxo de trabalho

desenhado na tela do aplicativo (canva) de acordo com as orientações colhidas no sítio na internet disponibilizado pelos gestores do software (UNIVERSITY OF LJUBLJANA, [s.d.]).

Figura 13 – Workflow - Classificador

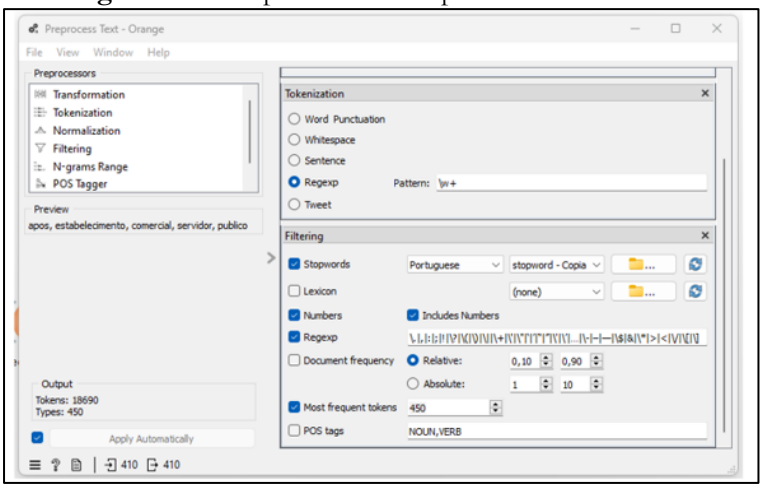


Fonte: O Autor.

O *workflow* da Figura 13 foi utilizado para treinar dois classificadores: O Classificador A, destinado a identificar os documentos a serem incluídos na categoria "É provável que exista linha telefônica para ser rastreada" e o Classificador B, destinado a identificar os documentos a serem rotulados com a classe "É provável que exista conta bancária para ser rastreada".

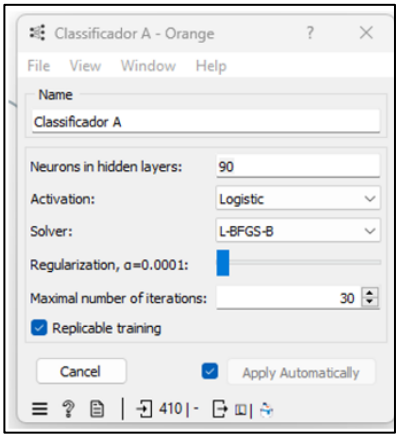
Os parâmetros adotados, em ambas as aplicações, foram obtidos após sucessivos ajustes nos quais maiores precisão e sensibilidade foram privilegiadas em detrimento das outras métricas. Imagens das telas com os parâmetros para o pré-processamento utilizados no Classificador A e no Classificador B estão expostas nas Figuras 14 a 17, respectivamente:

Figura 14 – Pré-processamento para o Classificador A



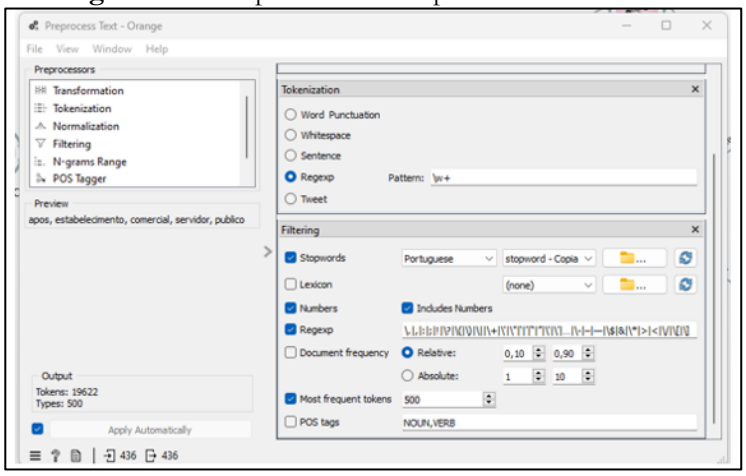
Fonte: O Autor.

Figura 15 – Parâmetros do Classificador A



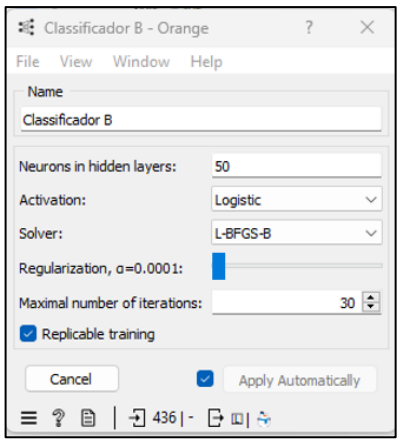
Fonte: O Autor.

Figura 16 – Pré-processamento para o Classificador B



Fonte: O Autor.

Figura 17 – Parâmetros do Classificador B



Fonte: O Autor.

Foram utilizados dois conjuntos de documentos para o treino, todos com 350 documentos cada:

a) Conjunto Telefone: Com os documentos classificados em duas categorias: a) documentos onde estão expostos eventos nos quais é provável que exista linha telefônica para ser rastreada e b) documentos onde estão expostos eventos nos quais não é provável que exista linha telefônica para ser rastreada. Foi utilizado para treinar o Classificador A.

b) Conjunto Conta: Com os documentos classificados em dois grupos: a) documentos onde estão expostos eventos nos quais é provável que exista conta bancária para ser rastreada e b) documentos onde estão expostos eventos nos quais é provável que exista conta bancária para ser rastreada. Foi utilizado para treinar o Classificador B.

Para evitar qualquer viés durante o treino os conjuntos de dados utilizados foram balanceados com relação as categorias, com a duplicação de documentos até que as quantidades em cada categoria fossem as mesmas para cada conjunto de treino.

A partir das pretensões esperadas para o modelo as métricas destinadas a avaliar a operações dos classificadores levaram em consideração as seguintes possibilidades de respostas:

a) Verdadeiro Positivo – PP, quando o documento foi classificado corretamente nas categorias de interesse.

b) Falso Positivo – PN, quando o documento foi rotulado incorretamente nas categorias de interesse.

c) Verdadeiro Negativo – NN, quando o documento foi classificado corretamente fora das categorias de interesse.

Falso Negativo – NP, quando o documento foi classificado incorretamente fora das categorias de interesse.

As medidas de avaliação selecionadas entre as disponibilizadas no *widget* TEST & SCORE foram (Microsoft, 2023):

a) Acurácia – AC: "É o percentual de previsões que coincidem exatamente com os rótulos de classes verdadeiros, quanto mais próximo de 1 melhor".

b) Precisão – P, "É a capacidade de o modelo evitar rotular documentos negativos como positivos, quanto mais próximo de 1 melhor".

c) Sensibilidade – S: "É a capacidade de um modelo para detectar todas as amostras positivas. Quanto mais próximo de 1 melhor".

d) Escore F1 – F1: "É a média harmônica entre Precisão e Sensibilidade, quanto mais próximo de 1 melhor".

e) Coeficiente de Correlação de Mathews - CCM: "Medida de precisão que varia no intervalo fechado [-1, 1], sendo que 1 indica previsão perfeita, 0 (zero) previsão aleatória e -1 previsão inversa. Quanto mais próximo de 1 melhor".

f) Área sob a Curva ROC - AROC: "É a área embaixo da Curva de Característica Operacional do Receptor - ROC. Quanto mais próxima de 1 melhor".

g) Matriz de Confusão - MC: "Apresenta uma imagem de como um modelo de *machine learning* está cometendo erros sistemáticos em suas previsões de classificação. A matriz de confusão de um modelo bom terá a maioria das amostras distribuídas na diagonal".

As métricas obtidas e aprovadas para encerrar o treino e aceitar os modelos estão exibidas na Tabela 2 e nos Gráficos 5 e 6:

Tabela 2 – Métricas de treinamento

Classificador	Categoria	Métricas					
		Área sob ROC	Acurácia	Precisão	Sensibilidade	F1	CC Mathews
A	É provável que exista linha telefônica para ser rastreada	0,828	0,756	0,727	0,818	0,770	0,516
	Não é provável que exista linha telefônica para ser rastreada	0,828	0,756	0,793	0,694	0,740	0,516
B	É provável que exista conta bancária para ser rastreada	0,915	0,858	0,847	0,872	0,859	0,716
	Não é provável que exista conta bancária para ser rastreada	0,915	0,858	0,869	0,844	0,856	0,716

Fonte: O Autor.

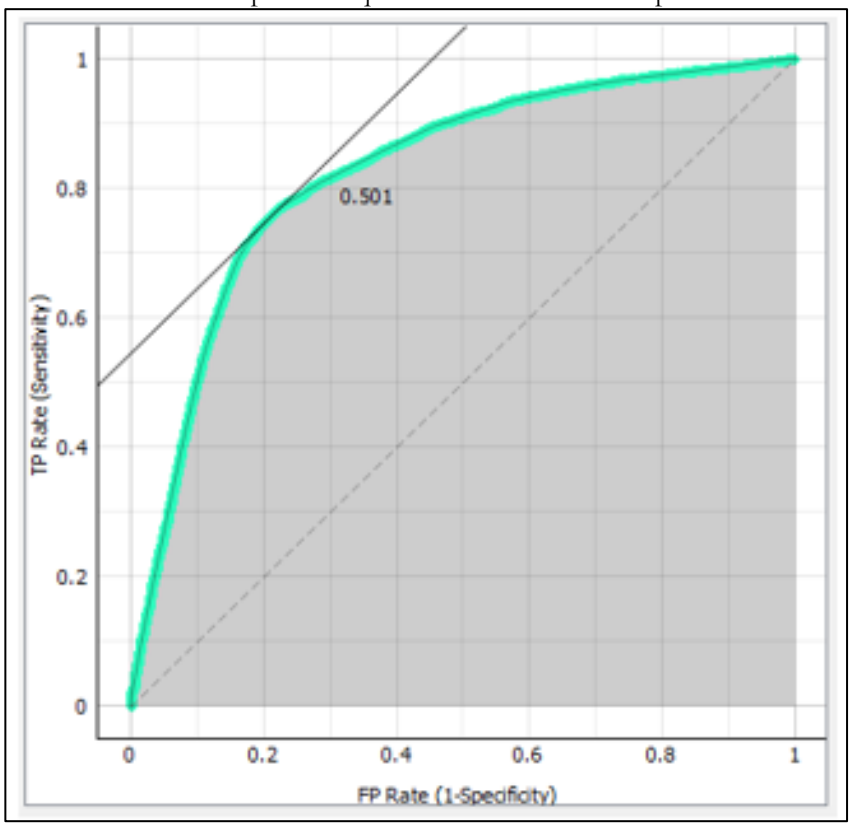
A Tabela 2 apresenta as métricas de desempenho obtidas no treinamento de dois classificadores aplicados a diferentes categorias preditivas. O Classificador A, relacionado à previsão da existência de linha telefônica rastreável, obteve Área sob a Curva ROC (AUC) de 0,828 para ambas as classes ("provável" e "não provável"), o que indica uma capacidade discriminativa moderadamente alta. A acurácia geral foi de 0,756, e os valores de precisão (0,727 e 0,793) e sensibilidade (0,818 e 0,694) revelam uma troca equilibrada entre falsos positivos e falsos negativos, com leve ênfase na capacidade de identificar corretamente os casos positivos. O valor de F1-score (0,770 e 0,740) e o coeficiente de correlação de Matthews (0,516) sugerem que o modelo apresenta desempenho consistente, porém moderado para esta tarefa.

Já o Classificador B, voltado à previsão da existência de conta bancária rastreável, apresentou desempenho superior em todos os indicadores. A AUC de 0,915 demonstra excelente capacidade discriminativa, e os valores de acurácia (0,858), precisão (0,847 e 0,869) e sensibilidade (0,872

e 0,844) indicam uma alta capacidade preditiva e balanceamento entre classes. Os F1-scores (0,859 e 0,856) confirmam a robustez do modelo, e o coeficiente de Matthews (0,716) reforça a confiabilidade da classificação, mesmo considerando possíveis desbalanceamentos nas classes.

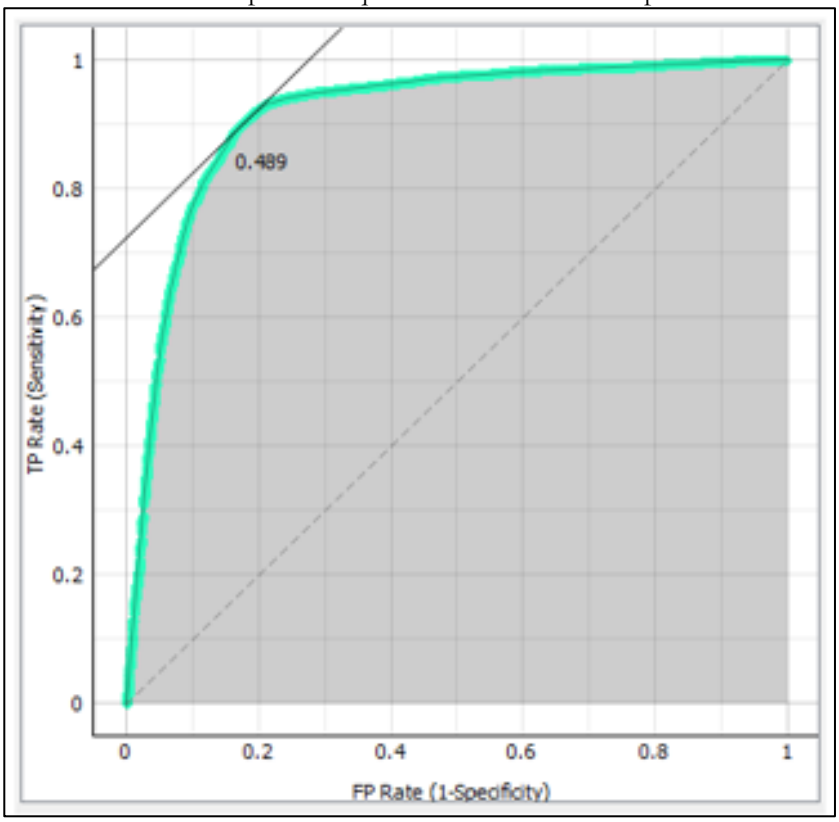
Em síntese, os resultados demonstram que o classificador B superou o classificador A em todas as métricas, sendo mais eficaz na predição de casos relacionados a contas bancárias do que de linhas telefônicas rastreáveis. Isso pode indicar que os padrões linguísticos ou estruturais associados à presença de contas bancárias são mais consistentes ou informativos para o modelo, o que deve ser considerado na seleção de atributos relevantes e na priorização de investigações automatizadas.

Gráfico 5 – ROC – é provável que exista linha telefônica para ser rastreada



Fonte: O Autor.

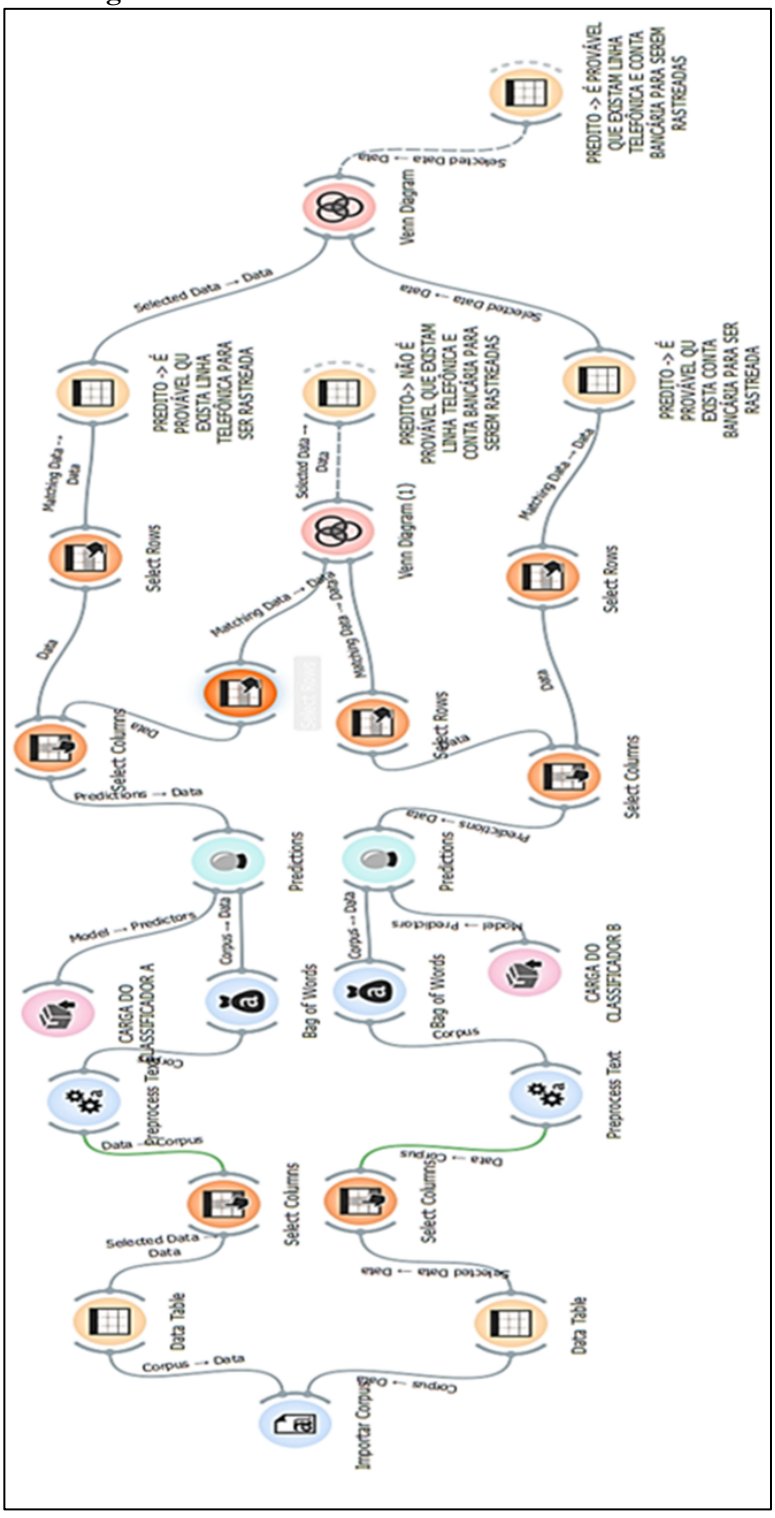
Gráfico 6 – ROC – é provável que exista conta bancária para ser rastreada



Fonte: O Autor.

Concebidos os Classificadores A e B, a etapa seguinte foi estruturar o modelo computadorizado para atender ao conceito estabelecido da Figura 5 anterior. A imagem da Figura 18 exibe o *workflow* projetado para classificar os Boletins de ocorrências de estelionato.

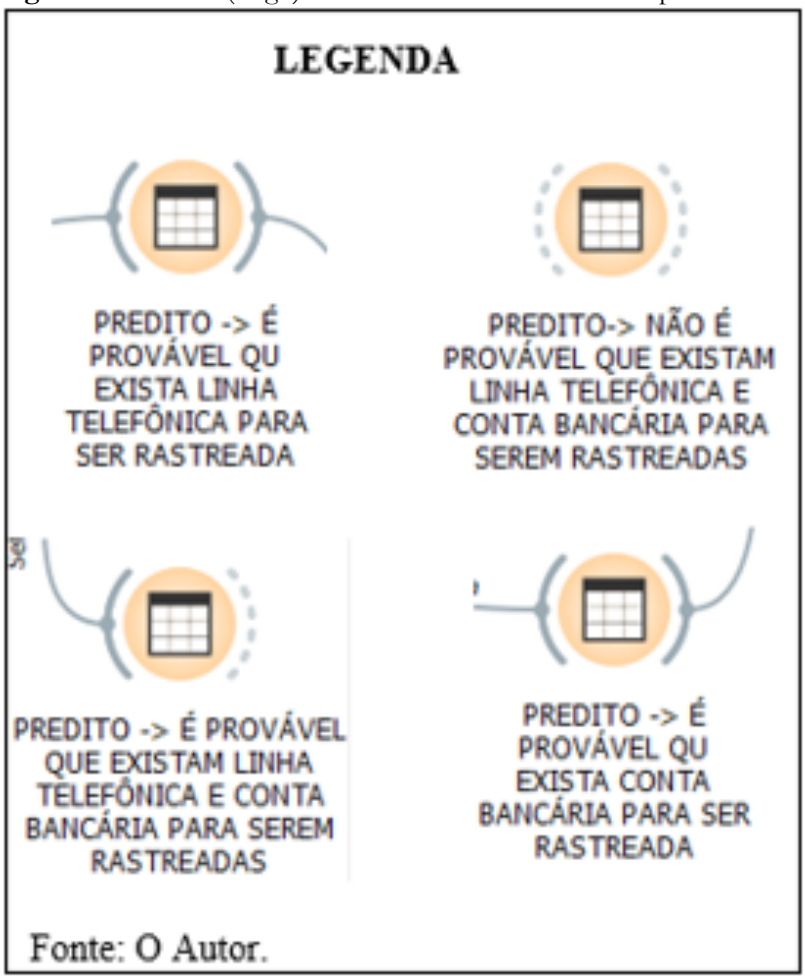
Figura 18 – Workflow do modelo de IA Classificador



Fonte: O Autor.

Na Figura 19 estão exibidos os ícones (*widget*) com as saídas do Modelo Computadorizado, cada um é uma tabela (arquivo digital) com os nomes de identificações dos Boletins de ocorrências (números dos registros) devidamente classificados gravados nas linhas.

Figura 19 – Ícones (*widget*) com as saídas do Modelo Computadorizado



Fonte: O Autor.

Para mensurar a eficiência do Modelo os resultados obtidos com a submissão do Conjunto de Teste[9] foram comparados com os resultados esperados[10]. Os Diagramas[11] de 2 a 5 representam as interseções dos dois conjuntos de resultados. Nas Tabelas 3 e 4 estão exibidas as Matrizes de Confusão.

[9] Dois conjuntos de teste correspondentes aos dois conjuntos de treino, com 150 elementos cada.

[10] Resultados esperados são os obtidos pelas classificações do Conjunto de Teste conduzidas por um analista humano (o Autor), eles formam o Conjunto CLASSIFICAÇÃO HUMANA. A classificação conduzida pelo modelo compõe o Conjunto CLASSIFICAÇÃO DO MODELO DE IA.

[11] Ver notas em rodapés 12 e 13 para comentários sobre os diagramas.

Figura 20 – Diagrama 2

Fonte: O Autor.

Figura 21 – Diagrama 3

Figura 22 – Diagrama 4

Figura 23 – Diagrama 5

A Tabela 3 representa a matriz de confusão do classificador aplicado à predição da existência de linha telefônica rastreável, com base na comparação entre as decisões da Inteligência Artificial (IA) e a classificação humana. O modelo identificou corretamente 44 casos verdadeiros positivos (PP) e 67 verdadeiros negativos (NN), totalizando 111 acertos em 150 instâncias, o que resulta em uma acurácia de 74%. Houve 23 falsos negativos (NP), em que a IA classificou como negativo casos que o humano considerou positivos, e 16 falsos positivos (PN), em que a IA classificou como positivo casos que o humano considerou negativos.

A precisão da classe positiva é de aproximadamente 73,3% (44 / [44+16]), indicando que, entre as previsões positivas feitas pela IA, 73,3% estavam corretas. Já a sensibilidade (ou recall) da mesma classe foi de 65,7% (44 / [44+23]), revelando que o modelo conseguiu identificar cerca de dois terços dos casos efetivamente positivos. O F1-score, que harmoniza precisão e sensibilidade, foi de 69,3%, refletindo desempenho moderado do classificador para essa categoria. Esses valores indicam que, embora o modelo seja funcional, ele apresenta limitações na identificação de todos os casos positivos, sugerindo espaço para aprimoramento no reconhecimento de padrões associados à presença de linha telefônica rastreável nos registros.

Tabela 3 – É provável que exista linha telefônica para ser rastreada

	IA – Positivo	**IA – Negativo**	**Total (Σ)**
HUMANO – Positivo	44	23	67
HUMANO – Negativo	16	67	83
Total (Σ)	**60**	**90**	**150**

Fonte: O Autor.

A Tabela 4 apresenta a matriz de confusão relativa à predição da existência de conta bancária rastreável, comparando os resultados do classificador com a categorização humana. O modelo identificou corretamente 85 verdadeiros positivos (PP) e 39 verdadeiros negativos (NN), totalizando 124 acertos em 150 instâncias, o que resulta em uma acurácia de 82,7%. Os erros foram relativamente baixos: 12 falsos positivos (PN) e 14 falsos negativos (NP).

A precisão do modelo foi de 87,6% (85 / [85 + 12]), indicando que a grande maioria das previsões positivas da IA estavam corretas. A sensibilidade foi de 85,9% (85 / [85 + 14]), evidenciando que o modelo conseguiu identificar a maior parte dos casos realmente positivos. O F1-score, que considera conjuntamente precisão e sensibilidade, foi de 86,7%, o que demonstra excelente equilíbrio entre os acertos positivos e a minimização de erros. O desempenho do classificador nesta categoria é significativamente superior ao observado na predição da existência de linha telefônica, o que indica maior consistência nos padrões textuais e estruturais relacionados à presença de contas bancárias nas ocorrências analisadas.

Tabela 4 – É provável que exista conta bancária para ser rastreada

	IA – Positivo	**IA – Negativo**	**Total (Σ)**
HUMANO – Positivo	85	14	99
HUMANO – Negativo	12	39	51
Total (Σ)	**97**	**53**	**150**

Fonte: O Autor.

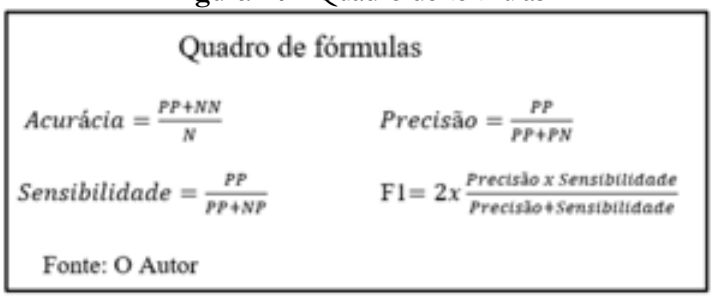
As métricas, exibidas na Tabela 6, foram calculadas com uso da Tabela 5 como gabarito e aplicando as fórmulas do Quadro 4 aos dados das Matrizes de Confusão.

Tabela 5 – Gabarito para Métricas

	IA – Positivo	**IA – Negativo**	**Total (Σ)**
HUMANO – Positivo	PP	NP	PP + NP
HUMANO – Negativo	PN	NN	PN + NN
Total (Σ)	PP + PN	NP + NN	N

Fonte: O Autor.

Figura 24 – Quadro de fórmulas



Fonte: O Autor.

Tabela 6 – Métricas

Categoria	**Acurácia**	**Precisão**	**Sensibilidade**	**F1**
É provável que exista linha telefônica para ser rastreada	0,740	0,733	0,657	0,693
É provável que exista conta bancária para ser rastreada	0,827	0,876	0,859	0,867

Fonte: O Autor.

A Tabela 6 apresenta as métricas de desempenho consolidadas dos classificadores treinados para prever a existência de linha telefônica e de conta bancária rastreável. Os resultados demonstram que o classificador voltado à conta bancária obteve desempenho superior em todos os indicadores. Com acurácia de 82,7%, precisão de 87,6%, sensibilidade de 85,9% e F1-score de 86,7%, o modelo mostra-se robusto e equilibrado, conseguindo tanto identificar corretamente os casos positivos quanto minimizar os erros de classificação.

Em contraste, o classificador associado à previsão da linha telefônica rastreável apresentou métricas mais modestas, com acurácia de 74,0%, precisão de 73,3%, sensibilidade de 65,7% e F1-score de 69,3%. Esses valores indicam maior dificuldade do modelo em identificar corretamente os casos positivos, com sensibilidade abaixo de 70%, o que pode comprometer a efetividade operacional quando aplicada em contexto investigativo.

A comparação entre as duas categorias sugere que os padrões linguísticos ou contextuais associados à presença de conta bancária são mais consistentes e previsíveis para o modelo de *machine learning*, ao passo que a presença de linha telefônica apresenta maior ambiguidade ou variação nos dados textuais. Esses achados podem orientar melhorias no pré-processamento, seleção de atributos e estratégias de reforço supervisionado para aprimorar o desempenho do classificador com menor performance.

4. RESULTADOS

As métricas foram calculadas considerando os rótulos "é provável que exista linha telefônica para ser rastreada" e "é provável que exista conta bancária para ser rastreada" como resultados positivos, os demais rótulos foram vistos como negativos.

Acreditamos que seja mais interessante perseguir uma pista falsa, ou que leve a lugar algum, do que desprezar um indício que depois se confirme importante e tenha sido perdido. Com essa convicção as métricas sensibilidade e precisão, e, por consequência, o Escore F1, foram os critérios mais fortes para a aceitação dos resultados e aprovação do modelo. Isso, evidentemente, após ficar demonstrado que não se está diante de um classificador ingênuo.

Os Gráficos 5 e 6, com as Análises ROC dos Classificadores A e B exibem resultados que indicam que eles se saíram melhor do que um classificador aleatório, ou seja, não atuaram como ingênuos. As áreas sob as curvas ROC para todas as classes são bastante próximas de 1. O treinamento foi encerrado com todas as métricas de avaliação acima de 0,5 e a maioria próxima de 1. As acurácias indicam que mais de 70% das classificações, chegando a um patamar acima de 80% para o Classificador B, foram coincidentes com as classificações conduzidas pelo classificador humano.

As precisões encontradas repetem os indicadores das acurácias em magnitude, pois acima de 70% dos documentos identificados como positivos receberam a classificação igual a atribuída pelo classificador humano. A sensibilidade para a identificação de positivos foi acima de 80%, ou seja, para cada 100 documentos positivos no conjunto de teste, pelo menos 80 serão rotulados corretamente. Os dados demonstram que o treinamento foi encerrado com a escolha de um modelo computacional de IA de rendimento aceitável para a finalidade deste estudo.

A avaliação prática do modelo considerou como ele se comportou com o Conjunto de teste. Para tanto as métricas foram calculadas comparando as classificações obtidas pela IA com as classificações conduzidas pelo analista humano e levaram as seguintes conclusões:

a) Ocorreu coincidência acima de 65% dos documentos rotulados com "é provável que exista linha telefônica para ser rastreada" e acima de 80% para os documentos classificados com "é provável que não exista linha telefônica para ser rastreada".

b) Ocorreu coincidência acima de 85% dos documentos rotulados com "é provável que exista conta bancária para ser rastreada" e acima de 70% para os documentos classificados com "é provável que não exista conta bancária para ser rastreada".

Logo, para os limites do estudo, não foram encontradas justificativas para rejeitar o modelo computacional, vez que superou a classificação ingênua e não exibiu sinais de *overfitting* ou de *underfitting*.

5. CONSIDERAÇÕES FINAIS

A presente pesquisa teve por objetivo avaliar a viabilidade de aplicação da inteligência artificial, por meio de redes neurais artificiais, como instrumento de apoio à gestão da investigação criminal. O estudo partiu da premissa de que o conhecimento empírico do investigador pode ser modelado computacionalmente e transformado em um ativo institucional. Os resultados alcançados confirmam tal hipótese, demonstrando que modelos de aprendizado supervisionado são capazes de reproduzir, com acurácia satisfatória, padrões decisórios anteriormente realizados por analistas humanos na classificação de boletins de ocorrência relativos ao crime de estelionato.

O modelo computacional construído apresentou desempenho robusto, especialmente na predição da presença de conta bancária rastreável, com métricas de precisão, sensibilidade e F1 superiores a 85%, evidenciando sua capacidade de apoiar a tomada de decisão policial com elevado grau de confiabilidade. Ainda que o classificador voltado à identificação de linhas telefônicas rastreáveis tenha obtido resultados mais modestos, com F1 em torno de 69%, suas métricas superaram significativamente o desempenho de um classificador aleatório, demonstrando aplicabilidade prática, desde que associado a processos de revisão e aprimoramento contínuo.

A utilização de técnicas de aprendizado de máquina supervisionado permitiu captar e replicar padrões linguísticos presentes nos relatos dos boletins de ocorrência. A construção do modelo contemplou etapas rigorosas de preparação e pré-processamento dos dados, validação cruzada e avaliação estatística dos resultados, evitando cenários de *overfitting* e *underfitting*. Com base nesses cuidados metodológicos, conclui-se que o modelo atendeu aos critérios de precisão, especificidade e sensibilidade esperados para uma solução de apoio à análise investigativa.

Além do desempenho técnico, a pesquisa contribui com uma abordagem metodológica replicável e escalável, que pode ser aplicada a outros tipos penais, regiões e contextos institucionais. A arquitetura modular e aberta do modelo, aliada ao uso de ferramentas como o *ORANGE Data Mining* e o NVivo, reforça sua adaptabilidade a diferentes fluxos de trabalho investigativo. A integração entre sistemas computacionais e conhecimento tácito representa um avanço estratégico para a modernização da segurança pública, especialmente diante da crescente complexidade dos crimes e da massificação dos registros digitais.

Contudo, como toda solução tecnológica aplicada à atividade estatal, a adoção de sistemas baseados em IA requer prudência e governança. O uso de classificadores computacionais deve respeitar os limites da função investigativa, atuando como suporte - e não como substituto - à análise humana. Aspectos éticos, jurídicos e operacionais, como a rastreabilidade das decisões algorítmicas, a mitigação de vieses e a transparência institucional, devem ser continuamente monitorados.

Entre as limitações do estudo, destaca-se o recorte temático restrito ao crime de estelionato e à base de dados oriunda de uma única unidade federativa, o que limita a generalização dos resultados. Ademais, o desempenho do modelo depende da qualidade e da representatividade do corpus utilizado no treinamento, exigindo estratégias de manutenção e atualização periódica.

Como propostas para trabalhos futuros, sugere-se: (a) a ampliação do modelo para outros crimes complexos, como extorsão, lavagem de dinheiro e associação criminosa; (b) a integração com bases relacionais e outros tipos de dados, como imagens, vínculos processuais e registros financeiros; (c) a construção de painéis de governança e controle sobre os resultados gerados pelos modelos; e (d) o desenvolvimento de normativas institucionais que estabeleçam protocolos éticos e técnicos para o uso de inteligência artificial na investigação criminal.

Em síntese, a pesquisa demonstrou que é possível aplicar a inteligência artificial como aliada estratégica da investigação criminal, respeitando os parâmetros técnicos e institucionais necessários. Ao transformar a experiência acumulada dos investigadores em um modelo analítico eficiente e reutilizável, abre-se caminho para a construção de uma governança investigativa mais inteligente, eficiente e responsiva às demandas contemporâneas da segurança pública.

REFERÊNCIAS

ADOBE. **O que é um PDF? Portable Document Format.** Adobe Acrobat. Disponível em: https://www.adobe.com/br/acrobat/about-adobe-pdf.html. Acesso em: 20 abr. 2024.

ALMEIDA, M. G. de A. *et al.* **Utilização de machine learning para classificação de crimes de morte no Estado de São Paulo**. Conferências IADIS Ibero-Americanas Computação Aplicada e WWW/Internet 2022, 2022, p. 103-111. Disponível em: https://www.iadisportal.org/digital-library/utiliza%C3%A7%C3%A3o-de-machine-learning-para-classifica%C3%A7%C3%A3o-de-crimes-de-morte-no-estado-de-s%C3%A3o-paulo. Acesso em: 23 abr. 2024.

ANTAR NETO, A. **Noções de matemática; conjuntos e funções**. Fortaleza: Vestseller, 2009.

ASCOM-PCSC. **Polícia Civil investe nas áreas de tecnologia e de Inteligência para dar agilidade nas investigações -** ACN - Agência Catarinense de Notícias. Disponível em: https://estado.sc.gov.br/noticias/policia-civil-investe-nas-areas-de-tecnologia-e-de-inteligencia-para-dar-agilidade-nas-investigacoes/. Acesso em: 14 nov. 2023.

BRASIL. **PPE -** Ministério da Justiça e Segurança Pública. Disponível em: https://www.gov.br/mj/pt-br/assuntos/sua-seguranca/seguranca-publica/sinesp-1/sinesp-ppe/ppe. Acesso em: 20 abr. 2024.

CNJ, C. N. de J. **Soluções de inteligência artificial promovem celeridade para o Poder Judiciário - Portal CNJ**. Disponível em: https://www.cnj.jus.br/solucoes-de-inteligencia-artificial-promovem-celeridade-para-o-poder-judiciario/. Acesso em: 14 nov. 2023.

CNJ, C. N. de J. **Justiça do Pará e STJ assinam acordo de Inovação e Inteligência Artificial - Portal CNJ**. Disponível em: https://www.cnj.jus.br/justica-do-para-e-stj-assinam-acordo-de-inovacao-e-inteligencia-artificial/. Acesso em: 14 nov. 2023.

DEMSAR, J. *et al.* Orange: Data Mining Toolbox in Python. **Journal of Machine Learning Research**, p. 2349-2353, 2013.

FERREIRA, L. H. C. A Prestação do Serviço de Investigação Criminal: um estudo para a aplicação da gestão por processo - Business Process Management (BPM). **Revista Brasileira de Ciências Policiais - RBCP**, p. 13-42, 2019.

FERREIRA, L. H. C.; FERREIRA, N. J. C. **Investigação criminal: Um estudo metodológico**. São Paulo: Sicurezza, 2013.

FERREIRA, M. H. P. **Classificação de peças processuais jurídicas: inteligência artificial no Direito**. 2018. Trabalho de Conclusão de Curso (Engenharia de Software) - Universidade de Brasília, DF, 2018. Disponível em: https://bdm.unb.br/handle/10483/21570. Acesso em: 15 nov. 2023.

FREITAS, L. D. C. de; ARRUDA, J. A. de A.; FALQUETO, J. M. Z. Uso do software Nvivo® em investigação qualitativa: ferramenta para pesquisa nas ciências sociais. In: **Congresso Ibero-Americano em Investigação Qualitativa**, 6., 2017. Atas do 6º Congresso Ibero-Americano em Investigação Qualitativa, [S.I.], v. 3, 2017. p. 621-626. Disponível em:

https://www.researchgate.net/publication/344563051_Uso_do_software_NvivoR_em_investigacao_qualitativa_ferramenta_para_pesquisa_nas_ciencias_sociais. Acesso em: 30 nov. 2023.

GREENE, J. R. **Administração do trabalho policial: Questões e análises**. São Paulo: EdUSP, 2007.

GRUS, J. **Data science do zero**. Rio de Janeiro: Alta Books, 2016.

JULIÃO, A. **Novo centro vai usar dados e inteligência artificial para tomada de decisões em segurança pública**. Agência FAPESP, 2023. Disponível em: https://agencia.fapesp.br/novo-centro-vai-usar-dados-e-inteligencia-artificial-para-tomada-de-decisoes-em-seguranca-publica/50055. Acesso em: 23 abr. 2024.

DIAS JUNIOR, J. A. **Os contos e os vigários**: uma história da trapaça no Brasil. São Paulo: Leya, 2010.

KREMER, G. R. **Algoritmos de aprendizado de máquina aplicados a dados públicos para obtenção de insights em segurança pública**. 2023. Trabalho de Conclusão de Curso (Ciência da Computação), UFRGS. 2023. Disponível em: https://lume.ufrgs.br/handle/10183/259327. Acesso em: 7 dez. 2023.

LIMA, I.; PINHEIRO, C.; SANTOS, F. **Inteligência artificial**. Rio de Janeiro: Elsevier Editora Ltda, 2014.

MACHADO, A. DOS S. **Matemática: conjuntos e funções**. São Paulo: Editora Atal, 1988.

MARINATTO, L. **Pela primeira vez, número de casos de estelionato supera o de roubos no estado do Rio**. Extra On Line, 6 abr. 2022. Disponível em: https://extra.globo.com/casos-de-policia/pela-primeira-vez-numero-de-casos-de-estelionato-supera-de-roubos-no-estado-do-rio-25460603.html#:~:text=O%20total%20de%20estelionatos%20no,foi%20de%20130%2C6%25. Acesso em: 23 abr. 2024.

MICROSOFT. **Avaliar os resultados de experimentos do AutoML -** Azure Machine Learning | Microsoft Learn. Disponível em: https://learn.microsoft.com/pt-br/azure/machine-learning/how-to-understand-automated-ml?view=azureml-api-2. Acesso em: 30 nov. 2023.

OLIVEIRA, J. **Informática | PETNews - Redes Neurais Artificiais**. Disponível em: http://www.dsc.ufcg.edu.br/~pet/jornal/setembro2011/materias/informatica.html. Acesso em: 16 nov. 2023.

PADULA, A. J. A. et al. **Segurança pública e inteligência artificial: um estudo georreferenciado para o Distrito Federal**. **Texto para Discussão**, p. 7-28, 2017. Disponível em: https://www.academia.edu/38556540/SEGURAN%C3%87A_P%C3%9ABLICA_E_INTELIG%C3%8ANCIA_ARTIFICIAL_UM_ESTUDO_GEORREFERENCIADO_PARA_O_DISTRITO_FEDERAL. Acesso em: 16 nov. 2023.

PEREIRA, E. da S. **Teoria da investigação criminal**: uma introdução jurídico-científica. Coimbra: Almedina, 2010.

PREFEITURA DE DUQUE DE CAXIAS. **Inteligência artificial auxilia na segurança pública de Duque de Caxias |** Prefeitura de Duque de Caxias | G1. Disponível em: https://g1.globo.com/rj/rio-de-janeiro/especial-publicitario/prefeitura-de-duque-de-caxias/noticia/2023/04/05/inteligencia-artificial-auxilia-na-seguranca-publica-de-duque-de-caxias.ghtml . Acesso em: 14 nov. 2023.

REDAÇÃO CONJUR. **Presidente do STJ exalta ganho de celeridade com uso de IA na corte -** Consultor Jurídico - CONJUR. Disponível em: https://www.conjur.com.br/2023-ago-21/presidente-stj-exalta-ganho-celeridade-uso-ia-corte/ . Acesso em: 14 nov. 2023.

RIBEIRO, L. J. **Investigação criminal: homicídio**. Brasília: Fábrica do Livro, 2012.

RONNIE, C. T. N.; GONÇALVES, A. L.; BARCELOS, B. O. Machine learning na segurança pública: uma análise de possíveis problemas mecânicos em viaturas policiais. **Anais do Congresso Internacional de Conhecimento e Inovação – ciki**, v. 1, n. 1, 15 fev. 2022. v. 1, n. 1Disponível em: https://proceeding.ciki.ufsc.br/index.php/ciki/article/view/1290 . Acesso em: 23 abr. 2024.

ROSA, J. L. G. **Fundamentos da inteligência artificial**. Rio de Janeiro: LTC, 2011.

SANTOS, J. E. L. DOS; NETO, M. F.; PEREIRA, F. D. **Boletim de ocorrência eletrônico no estado de São Paulo: Inteligência artificial como proposta de inovação**. Revista Jurídica UNICURITIBA, v. 3, n. 60, p. 426–446, 10 ago. 2020. Disponível em: https://revista.unicuritiba.edu.br/index.php/RevJur/article/view/4189 . Acesso em: 23 abr. 2024.

SOUSA, R. M. DE. **Inteligência computacional aplicada ao controle externo: classificação de padrões utilizando redes neurais artificiais**. Revista do TCU, p. 36–43, 2016. Disponível em: https://revista.tcu.gov.br/ojs/index.php/RTCU/article/view/1323 . Acesso em: 23 abr. 2024.

SSPAL, E. DE A. **Centro Integrado da Segurança Pública usa Inteligência Artificial para avançar na redução da violência no Estado – SSP AL**. Disponível em: http://seguranca.al.gov.br/noticia/2023/06/27/centro-integrado-da-seguranca-publica-usa-inteligencia-artificial-para-avancar-na-reducao-da-violencia-no-estado/ . Acesso em: 14 nov. 2023.

STF, S. T. F. **Presidente do STF abre seminário sobre uso da inteligência artificial na Corte**. Disponível em: https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=505755&ori=1 . Acesso em: 14 nov. 2023a.

STF, S. T. F.-. **STF amplia emprego de Inteligência Artificial**. Disponível em: https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=508710&ori=1 . Acesso em: 14 nov. 2023b.

STF, S. T. F.-. **STF faz chamamento público para projetos de inteligência artificial que automatizem resumos de processos**. Disponível em: https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=518467&ori=1 . Acesso em: 14 nov. 2023c.

STJ, S. T. DE J.-. **Inteligência artificial está presente em metade dos tribunais brasileiros, aponta estudo inédito**. Disponível em: https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/09032021-Inteligencia-

artificial-esta-presente-em-metade-dos-tribunais-brasileiros--aponta-estudo-inedito.aspx . Acesso em: 14 nov. 2023.

TJBA, T. DE J. B.-. **PJBA assina termo de cooperação com o STJ para uso de inteligência artificial nos fluxos das ações judiciais**. Disponível em: https://www.tjba.jus.br/portal/pjba-assina-termo-de-cooperacao-com-o-stj-para-uso-de-inteligencia-artificial-nos-fluxos-das-acoes-judiciais/ . Acesso em: 14 nov. 2023.

UNIVERSITY OF LJUBLJANA. **Orange Data Mining - Widget Catalog**. Disponível em: https://orangedatamining.com/widget-catalog/ . Acesso em: 8 dez. 2023.

WIKIPÉDIA. **scikit-learn – Wikipédia, a enciclopédia livre**. Disponível em: https://pt.wikipedia.org/wiki/Scikit-learn . Acesso em: 29 nov. 2023.

WIKIPÉDIA. **URL – Wikipédia, a enciclopédia livre**. Disponível em: https://pt.wikipedia.org/wiki/URL . Acesso em: 20 abr. 2024.