

## **UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NA PREDIÇÃO DE ACIDENTES DE TRÂNSITO NA CIDADE DE CAMPO GRANDE – MS**

### ***THE USE OF ARTIFICIAL INTELLIGENCE IN PREDICTING TRAFFIC ACCIDENTS IN THE CITY OF CAMPO GRANDE – MS***

### ***EL USO DE LA INTELIGENCIA ARTIFICIAL EN LA PREDICCIÓN DE ACCIDENTES DE TRÁFICO EN LA CIUDAD DE CAMPO GRANDE – MS***

**Brenno Ysrael Maluf Luccas Correia<sup>1</sup>**  
**Cristiano Pereira Alves de Moraes<sup>2</sup>**  
**Ederson Roberto da Costa<sup>3</sup>**  
**Fabricao Almeida de Oliveira<sup>4</sup>**  
**Thays Martines Ferreira<sup>5</sup>**

**RESUMO:** O presente estudo traz uma investigação sistematizada com a aplicação de técnicas de inteligência artificial (IA) para buscar uma previsão de acidentes de trânsito na cidade de Campo Grande, no estado de Mato Grosso do Sul (MS). Foi utilizado algoritmo de aprendizado de máquina a partir de dados históricos de acidentes para identificar padrões e prever futuras ocorrências. O principal objetivo foi auxiliar na tomada de decisões para a melhoria da segurança viária da cidade de Campo Grande, MS. Utilizando uma metodologia que combinou revisão da literatura, a análise de dados estatísticos de acidentes da referida cidade, o estudo ofereceu uma visão aprofundada dos desafios enfrentados pelo poder público em identificar o período de maior incidência de acidentes, contribuindo para que possam utilizar o estudo na aplicação de ações pontuais mitigando tais fatos e direcionando a ações em regiões de maior incidência de acidente. A IA, com base em diversas variáveis e sua capacidade avançada de análise de dados, pode identificar padrões e prever incidentes. A utilização de condições meteorológicas, horários de pico e incidentes são dados históricos, os quais podem contribuir no modelo para o treinamento do algoritmo. Este trabalho analisou a aplicação

---

<sup>1</sup> Brenno Ysrael Maluf Luccas Correia é acadêmico do curso de Análise e Desenvolvimento de Sistemas do Instituto Avançado de Ensino Superior e Desenvolvimento Humano – INSTED. ORCID iD: <https://orcid.org/0009-0003-7375-5949>. E-mail: [brenno.maluf@gmail.com](mailto:brenno.maluf@gmail.com)

<sup>2</sup> Cristiano Pereira Alves de Moraes é acadêmico do curso de Análise e Desenvolvimento de Sistemas do Instituto Avançado de Ensino Superior e Desenvolvimento Humano – INSTED. ORCID iD: <https://orcid.org/0009-0001-8795-6114>. E-mail: [cpa.cristiano@gmail.com](mailto:cpa.cristiano@gmail.com)

<sup>3</sup> Ederson Roberto da Costa é professor do Instituto Avançado de Ensino Superior e Desenvolvimento Humano – INSTED. E-mail: [edersondacosta@hotmail.com](mailto:edersondacosta@hotmail.com). ORCID iD: <https://orcid.org/0009-0009-5617-4644>

<sup>4</sup> Fabrício Almeida de Oliveira é acadêmico do curso de Análise e Desenvolvimento de Sistemas do Instituto Avançado de Ensino Superior e Desenvolvimento Humano – INSTED. ORCID iD: <https://orcid.org/0009-0006-7283-3353>. E-mail: [devfasbalmeida@gmail.com](mailto:devfasbalmeida@gmail.com).

<sup>5</sup> Thays Martines Ferreira é acadêmica do curso de Análise e Desenvolvimento de Sistemas do Instituto Avançado de Ensino Superior e Desenvolvimento Humano – INSTED. ORCID iD: <https://orcid.org/0009-0003-4978-5347>. E-mail: [thaysmartines18@gmail.com](mailto:thaysmartines18@gmail.com).

de algoritmos de aprendizado de máquina para processar e interpretar grandes volumes de dados de acidentes de trânsito, com o objetivo de fornecer previsões precisas que auxiliem na implementação de medidas preventivas e na melhoria da segurança viária na cidade.

**PALAVRAS-CHAVE:** Conflitos Coletivos; Desafios Urbanístico; Inteligência Artificial.

**ABSTRACT:** The present study offers a systematic investigation through the application of artificial intelligence (AI) techniques to forecast traffic accidents in the city of Campo Grande, in the state of Mato Grosso do Sul (MS). A machine learning algorithm was used with historical accident data to identify patterns and predict future occurrences. The main goal was to assist decision-making to improve road safety in the city of Campo Grande, MS. Using a methodology that combined a literature review and the analysis of statistical accident data from the mentioned city, the study provided an in-depth view of the challenges faced by public authorities in identifying periods of highest accident incidence. This contribution can be utilized to apply targeted actions to mitigate such occurrences and direct interventions toward regions with higher accident rates. AI, based on various variables and its advanced data analysis capabilities, can identify patterns and predict incidents. The use of weather conditions, peak hours, and historical incidents are data that can contribute to the model for algorithm training. This work analyzed the application of machine learning algorithms to process and interpret large volumes of traffic accident data, with the goal of providing accurate forecasts that aid in the implementation of preventive measures and the improvement of road safety in the city.

**KEYWORDS:** Collective Conflicts; Urban Planning Challenges; Artificial Intelligence.

**RESUMEN:** El presente estudio presenta una investigación sistematizada con la aplicación de técnicas de inteligencia artificial (IA) para prever accidentes de tráfico en la ciudad de Campo Grande, en el estado de Mato Grosso do Sul (MS). Se utilizó un algoritmo de aprendizaje automático con datos históricos de accidentes para identificar patrones y predecir futuras ocurrencias. El objetivo principal fue apoyar la toma de decisiones para mejorar la seguridad vial en la ciudad de Campo Grande, MS. Utilizando una metodología que combinó la revisión de la literatura y el análisis de datos estadísticos de accidentes de dicha ciudad, el estudio ofreció una visión profunda de los desafíos que enfrentan las autoridades públicas para identificar los períodos de mayor incidencia de accidentes. Esto contribuye a la aplicación de acciones puntuales para mitigar tales hechos y a la focalización de intervenciones en las regiones con mayor tasa de accidentes. La IA, basada en diversas variables y su avanzada capacidad de análisis de datos, puede identificar patrones y predecir incidentes. El uso de condiciones meteorológicas, horas pico y datos históricos de incidentes puede contribuir al modelo para el entrenamiento del algoritmo. Este trabajo analizó la aplicación de algoritmos de aprendizaje automático para procesar e interpretar grandes volúmenes de datos de accidentes de tráfico, con el objetivo de proporcionar predicciones precisas que ayuden en la implementación de medidas preventivas y la mejora de la seguridad vial en la ciudad.

**PALABRAS CLAVE:** Conflictos Colectivos; Desafíos Urbanísticos; Inteligencia Artificial.

## **INTRODUÇÃO**

Com o surgimento dos primeiros automóveis em Portugal, remontando ao século XIX (Educação de Trânsito, s.d.), veio ao longo dos anos o desafio do aumento do número de veículos em circulação, as autoridades perceberam a importância do transporte motorizado para que houvesse o progresso do país. Contudo, também passaram a reconhecer os novos desafios de segurança associados ao uso de automóveis, especialmente em relação à gravidade dos acidentes no trânsito regional, evidenciando que a acidentalidade não é uma questão recente.

Dessa maneira, há de se compreender que a segurança no trânsito é uma das preocupações mais prementes em âmbito global (Organização Mundial da Saúde [OMS], 2021). As vítimas de acidentes de trânsito tornam a acidentalidade um grave problema de saúde pública, considerando que os sistemas de trânsito regional são considerados complexos e arriscados, com os quais as pessoas interagem diariamente (LEAL, 2016).

Contudo, essa problemática não se resume apenas à perda trágica de vidas humanas ou aos danos físicos e psicológicos irreparáveis que afetam as vítimas e suas famílias. Também representa um obstáculo significativo ao desenvolvimento econômico dos países. Além das perdas humanas e do sofrimento, há uma destruição de riqueza que, segundo estudos realizados para a Comissão Europeia, superou em Portugal, nesses últimos seis anos, a marca de 12,4 bilhões de euros, correspondendo, em média, a cerca de 1,24% do PIB (Resolução do Conselho de Ministros n.º 85/2017, de 19 de junho, 2017).

Entretanto, os riscos associados à acidentalidade no trânsito regional podem ser prevenidos ou, ao menos, atenuados. A adoção de tecnologias e a implementação de políticas de trânsito cada vez mais rigorosas têm mostrado resultados positivos, embora esses avanços ainda não atendam completamente às expectativas (MARCILLO, CARAGUAY, & HERNÁNDEZ-ALVAREZ, 2022).

A adoção de sistemas fundamentados em inteligência artificial (IA) para levantamento de informações na segurança no trânsito já se consolidou,

apresentando resultados bastante significativos (MARCILLO et al., 2022). Em particular, os modelos preditivos relacionados à acidentalidade no trânsito fornecem uma base científica robusta para a avaliação e escolha de medidas de segurança, além de facilitar o processo de tomada de decisões com custos relativamente baixos (YANNIS et al., 2016).

No município de Campo Grande, localizado no estado de Mato Grosso do Sul, em 2023, foram contabilizadas cerca de 35 ocorrências de acidentes diariamente (O ESTADO, 2023). A maior parte desses incidentes de trânsito é atribuída a ações humanas, ou seja, a maioria das situações resulta de infrações cometidas por motoristas ou até mesmo pedestres. Com o intuito de promover a conscientização sobre os acidentes e fatalidades no trânsito, foi lançada a campanha Maio Amarelo, que leva os órgãos públicos a intensificarem as iniciativas educativas tanto nas áreas urbanas quanto nas estradas do Mato Grosso do Sul (O ESTADO, 2023).

A partir disso, observando os dados, é de extrema importância investigar como a IA pode contribuir para a prevenção e alerta em relação à acidentalidade no trânsito, utilizando o modelo de algoritmo de previsão de dados ou séries temporais para como estudo de caso.

A previsão de dados ou séries temporais possibilita estimar informações desconhecidas com base em um conjunto de dados já disponíveis. Para esse fim, foram sugeridas abordagens tanto lineares quanto não lineares para efetuar as previsões. Em geral, as metodologias lineares pressupõem que os dados seguem uma determinada distribuição estatística e, com isso, definem parâmetros para ajustar um modelo a esses dados. No entanto, muitos dados temporais do mundo real envolvem fenômenos naturais que podem exibir comportamentos não lineares.

## **REFERENCIAL TEÓRICO**

A inteligência artificial (IA) busca estabelecer modelos que permitam às máquinas executarem tarefas cognitivas que, para os seres humanos, são realizadas de forma intuitiva e direta. Por exemplo, enquanto uma criança pode

facilmente reconhecer a presença de um indivíduo em uma fotografia, um computador requer um processo de aprendizado específico para alcançar esse nível de percepção (HAYKIN, 2007). Dentro desse campo, destaca-se a aprendizagem de máquina (Machine Learning - ML), uma área fundamental que visa desenvolver a capacidade de aprendizado das máquinas, permitindo que elas adquiram novas habilidades e conhecimentos, além de aprimorar a forma como organizam as informações previamente assimiladas (BATISTA et al., 2003).

O emprego de sistemas baseados em inteligência artificial (IA) no campo da segurança viária já se consolidou como uma prática relevante, evidenciando resultados expressivos (MARCILLO et al., 2022). Em particular, os modelos preditivos de incidentes no trânsito têm fornecido uma fundamentação científica robusta para a avaliação e seleção de estratégias de segurança viária, além de apoiar os processos decisórios de maneira eficiente e com custos relativamente baixos (YANNIS et al., 2016).

A modelagem estatística pode ser entendida como o processo de simplificar a realidade para representar, explicar ou prever uma ou mais variáveis aleatórias de interesse a partir de outras variáveis. Ao aplicar diferentes modelos estatísticos aos dados a serem analisados, os especialistas conseguem interpretar as informações de maneira mais direta e objetiva. Em vez de lidar com dados brutos, a modelagem facilita a identificação de conexões entre variáveis, possibilitando a criação de previsões baseadas nos dados disponíveis (STOBIERSKI, 2019).

Com o avanço da Inteligência Artificial, a modelagem estatística passou a ser aplicada a conjuntos de dados de volumes muito maiores. Isso também permitiu o uso de técnicas e algoritmos capazes de aprender com diferentes tipos de informação, aperfeiçoando-se continuamente com o tempo – um conceito que hoje é conhecido como Aprendizado de Máquina.

Após o processo de treinamento de um modelo de aprendizado de máquina, é fundamental aplicar métricas que possibilitem a avaliação de seu desempenho. Há diversas abordagens para mensurar a eficácia dos resultados

alcançados, entre as quais se destacam a matriz de confusão e a acurácia. Essas métricas fornecem insights valiosos sobre a capacidade do modelo em classificar corretamente os dados, permitindo não apenas a identificação de acertos, mas também a análise de erros, o que é crucial para o aprimoramento contínuo do sistema.

O emprego de novas fontes de dados, aliado à implementação de técnicas algorítmicas com capacidade de aprendizado — conhecido como Aprendizado de Máquina — trouxe uma série de benefícios. Entre esses, destaca-se a maior automatização dos processos de modelagem e a capacidade de autoaprendizado, o que eleva significativamente o poder preditivo dos modelos (MANAGEMENT SOLUTIONS, 2018). Por outro lado, a predição, ou seja, os modelos preditivos, permite a correlação entre informações oriundas de diferentes fontes (MARCILLO et al., 2022, p. 2). Enquanto um Modelo Estatístico envolve o uso de métodos estatísticos para construir uma representação dos dados e realizar análises que identifiquem relações entre variáveis e seus efeitos, um modelo de Aprendizado de Máquina utiliza abordagens matemáticas e estatísticas para interpretar os dados e gerar previsões. Nesse último, o computador é "treinado" para aprender padrões a partir de grandes volumes de dados.

A previsão, no contexto de sistemas da informação, envolve a utilização de dados passados para antecipar eventos futuros, considerando um conjunto específico de condições. Quando aplicamos essa ideia ao estudo de séries temporais, lidamos com variáveis que são registradas ao longo do tempo, como dados de acidentes diários, maior incidência de morte ou lesões corporais. A série temporal é, portanto, uma sequência de observações organizadas cronologicamente, onde o tempo é um fator crucial na análise e no comportamento das variáveis.

Para desenvolver modelos preditivos eficazes em séries temporais, os algoritmos de regressão se destacam como uma ferramenta robusta. Estes algoritmos permitem modelar a relação entre uma variável dependente e uma ou

mais variáveis independentes, possibilitando a identificação de padrões e tendências ao longo do tempo. Ao contrário de métodos de regressão tradicionais, os modelos para séries temporais precisam lidar com dependências temporais entre as observações, o que exige a adaptação de técnicas específicas para capturar essas dinâmicas.

Assim, a previsão em séries temporais se torna uma prática essencial para diversos campos, como finanças, saúde e logística, oferecendo insights que podem orientar tomadas de decisão mais informadas. No desenvolvimento de um artigo científico sobre o tema, o foco está em explorar como algoritmos de regressão, como o ARIMA, o SARIMA e métodos de aprendizado de máquina como LSTM, podem ser aplicados para modelar esses dados temporais, fornecendo previsões mais precisas e consistentes em diferentes contextos.

Tal sistema fundamenta-se no estudo detalhado de dados históricos, com o objetivo de identificar padrões recorrentes ao longo do tempo. Esses padrões permitem modelar e prever o comportamento futuro da variável de interesse. Um modelo de séries temporais parte da premissa de que os padrões observados ao longo do tempo tendem a se repetir, tornando esse método adequado para variáveis que exibem comportamentos sazonais ou tendências definidas.

Outro quesito importante a se destacar é o papel das bibliotecas computacionais de sistemas para banco de dados no processamento e análise dos dados, que se dá de forma linear e em inflação no decorrer dos tempos em que o sistema é alimentado. Dessa forma, quando se trata de manipulação, visualização e modelagem de grandes volumes de dados, algumas bibliotecas se destacam pela sua eficiência e simplicidade de uso, como Pandas, Matplotlib e NumPy.

O uso de Pandas na análise de dados tem se consolidado como uma ferramenta indispensável para cientistas e analistas, especialmente devido à sua habilidade de manipular dados de maneira eficiente e flexível. Sua principal estrutura de dados, o DataFrame, é uma tabela bidimensional, onde os dados são organizados em linhas e colunas rotuladas, permitindo uma abordagem

intuitiva para a manipulação de grandes volumes de informações. O conceito de DataFrame, inspirado nas séries de dados estatísticos, possibilita o tratamento de diferentes tipos de dados e suas respectivas operações, como filtragem, seleção, e modificação de valores.

No contexto específico das séries temporais, que são dados sequenciais organizados cronologicamente, a biblioteca Pandas se destaca por sua capacidade de manipulação precisa de datas e períodos de tempo. Segundo Hamilton (2020), a manipulação de séries temporais em Pandas permite "um controle granular sobre a granularidade temporal", ou seja, a biblioteca é capaz de agrupar, resamplear e fatiar os dados temporalmente para extrair informações valiosas sobre a dinâmica da variável observada ao longo do tempo.

Já NumPy oferece suporte a operações matemáticas e algébricas de alta performance, especialmente com arrays multidimensionais, que são estruturas essenciais para a realização de cálculos e transformações complexas em grandes conjuntos de dados. A biblioteca oferece funções otimizadas para manipulação de matrizes, o que torna os cálculos mais rápidos e eficientes, especialmente em cenários que envolvem séries temporais e algoritmos de previsão.

Neste sentido, a visualização dos dados é uma etapa imprescindível para compreender os padrões e tendências dentro de séries temporais. A biblioteca Matplotlib permite a criação de gráficos complexos e customizáveis que facilitam essa análise visual. Com ela, são visíveis histogramas para visualizar distribuições e scatter plots para explorar correlações entre variáveis.

Essas bibliotecas, quando utilizadas em conjunto, fornecem uma base sólida para a análise de séries temporais, desde a coleta e manipulação dos dados até a sua visualização e modelagem.

Aguiar (2021) utilizou a predição para acidentes, observou-se a amostra inicial de 288 interseções semaforizadas controladas pelo CTAFOR - Controle de Tráfego em Área de Fortaleza, posteriormente foi reduzida para 119 interseções devido a problemas no armazenamento dos dados de fluxo veicular. Após análise por imagens aéreas do Google Earth, 18 interseções foram excluídas da amostra,

resultando em 101 interseções semaforizadas utilizadas no processo de modelagem. Foi utilizado o SIAT-FOR, um banco de dados georreferenciado que compila informações desde o ano 2000, incluindo data, hora, local, tipo de acidente, veículos envolvidos e severidade.

O volume de tráfego foi estimado com base nos dados de fluxo veicular armazenados pelo CTAFOR, utilizando o sistema SCOOT, que estima e armazena fluxos horários em intervalos de 15 minutos. Ao final, o VDMA - Volume Diário Médio Anual foi ajustado com um fator de expansão obtido através da análise do padrão de variação espaço-temporal do tráfego.

Outro exemplo relevante de aplicação de tecnologia em rodovias é o sistema MOPREVIS, que foi desenvolvido com base em metodologias modernas e avanços em Inteligência Artificial (IA) e Ciência de Dados (INFANTE et al., 2022a, p. 178). Esse sistema é utilizado principalmente para a análise de acidentes rodoviários, utilizando dados oficiais do Comando Territorial de Setúbal, especialmente através do Boletim Estatístico de Acidentes de Viação (BEAV), elaborado pela GNR para cada acidente registrado (INFANTE et al., 2022a, p. 8). Os dados analisados englobam 28.102 registros de acidentes rodoviários ocorridos entre 2016 e 2019, provenientes de diversas fontes de informação. Com base nesses dados, foram selecionadas várias variáveis e construídas novas, proporcionando um conjunto robusto para a análise.

A interface digital do MOPREVIS permite não só a visualização de dados históricos e presentes, mas também a predição de eventos futuros. Os usuários podem selecionar intervalos de tempo específicos para análise e aplicar uma série de filtros para visualizar os dados em diversas formas, como gráficos e infográficos. Ao todo, a ferramenta oferece 218 opções de representações gráficas (Infante et al., 2022a, p. 166), além de relatórios e mapas detalhados elaborados pela equipe do projeto (INFANTE et al., 2022a, p. 177).

A inovação mais significativa do sistema, no entanto, é sua funcionalidade preditiva, que integra análises estatísticas, espaciais e de IA. Com isso, o MOPREVIS pode prever zonas de alto risco para acidentes em

determinados trechos de rodovias e momentos específicos do dia (INFATE, et al., 2023b, p. 2). Embora ainda esteja em fase experimental, o sistema já foi testado em condições reais pela GNR no Comando Territorial de Setúbal.

Por outro lado, os modelos de previsão baseados em relações causais consideram que as variações da variável dependente podem ser explicadas pela influência de outras variáveis independentes. Ao invés de focar exclusivamente na passagem do tempo, os modelos causais exploram relações de causa e efeito entre diferentes fatores, buscando descobrir como essas interações afetam o comportamento da variável de interesse.

## **DESENVOLVIMENTO**

Para a execução do projeto, fez-se necessário a elaboração de um estudo comparativo e de predição de acidentes com base em dados governamentais, levantando todas as informações a partir de dados disponibilizados pelo Ministério dos Transportes, pelo Registro Nacional de Sinistros e Estatísticas de Trânsito (BRASIL, 2024).

Buscou-se desenvolver um dataset robusto que contemplasse as informações históricas de sinistros ocorridos a partir de 2018. O acesso a dados oficiais, disponibilizados por órgãos como a Polícia Rodoviária Federal, o Departamento Nacional de Trânsito (DENATRAN) e outras fontes governamentais, é um ponto de partida essencial para garantir a qualidade e a confiabilidade da base de dados.

Além destes dados estatísticos na construção do dataset, houve a seleção cuidadosa de variáveis relevantes, que não se limitam a dados geoespaciais, características do local do acidente (como condições da via, iluminação, sinalização), tipo de veículo, dados sobre as vítimas e condutores, além de fatores temporais como a data e hora do acidente.

Em um segundo momento, fundamentou-se o projeto na aplicação de técnicas estatísticas de limpeza e tratamento dos dados, como a remoção de duplicidades e inconsistências, além da normalização de variáveis numéricas e

categóricas. A utilização de ferramentas como *Pandas* e *NumPy* foi crucial para essa etapa, permitindo manipular grandes volumes de dados de forma eficiente.

Após a estruturação do dataset, se tornou possível avançar para a modelagem preditiva utilizando algoritmos de *machine learning*, como regressão logística, *random forest* e redes neurais. Tais modelos foram treinados para identificar padrões de comportamento em diferentes cenários e prever a probabilidade de acidentes futuros com base nas condições observadas no passado. Buscou-se não apenas prever a ocorrência de acidentes, mas também fornecer uma base comparativa que permita avaliar a evolução dos sinistros ao longo dos anos, identificando tendências e possíveis intervenções que possam ser adotadas para mitigar os riscos em áreas de maior incidência.

## **RESULTADOS**

Esta seção apresenta os resultados obtidos com a aplicação do dataset no algoritmo de predição, buscando avaliar o desempenho do algoritmo utilizado na predição de acidentes de trânsito.

Os gráficos apresentam parâmetros nas classes municipais, estaduais e nacionais de acidentes ocorridos ao longo dos últimos seis anos (2018-2023) sobre dados preditivos, conforme apresentado até aqui, neste contexto, foi aplicado o método de decomposição baseado em regressão múltipla descrito anteriormente. Para tanto, atribuiu-se uma variável dummy positiva para cada período mensal. A Figura 1, detalha a quantidade de acidentes em cada período na cidade de Campo Grande MS, a linha vermelha representa os valores reais adquiridos no dataset do Governo, já a linha azul demonstra a predição apresentada pelo modelo de IA.

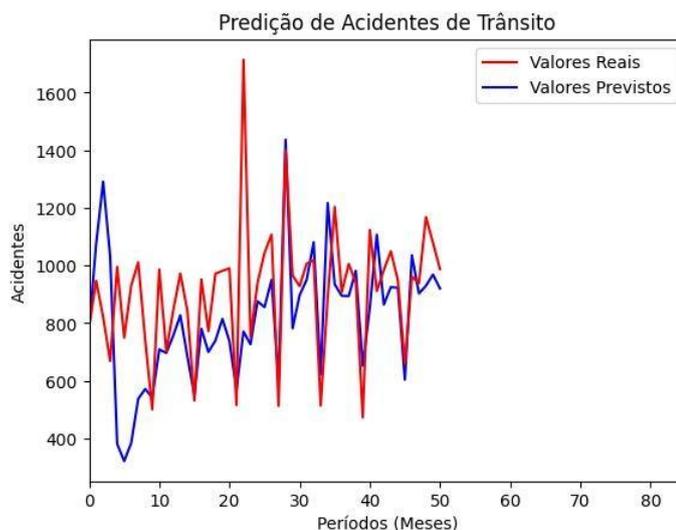
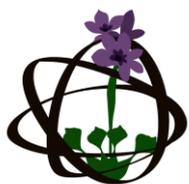


Figura 1: Predição de Acidentes na cidade de Campo Grande - MS

Na Figura 1 é possível perceber que o modelo de IA esteve bem próximo de prever os valores, principalmente entre os períodos de 30 a 50 meses, uma hipótese está no fato dos dados para treinamento serem menores.

Já a Figura 2, demonstra os acidentes no estado de Mato Grosso do Sul por períodos, no mesmo formato dos períodos da Figura 1.

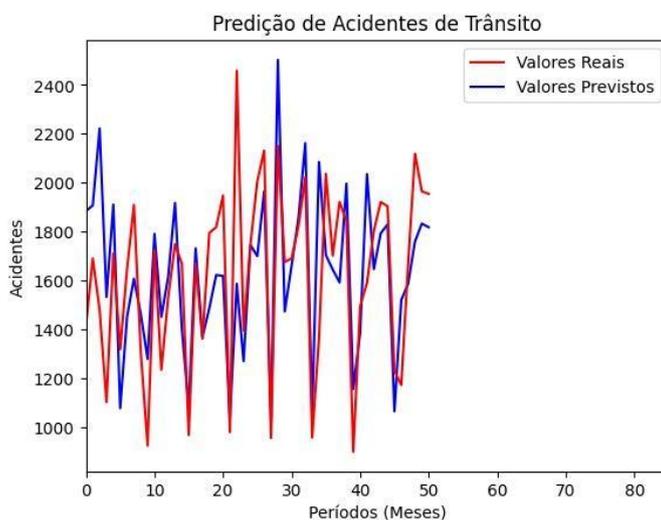


Figura 2: Predição de Acidentes no estado de Mato Grosso do Sul.

Na Figura 2 é possível perceber que o modelo de IA já possui uma maior regularidade com relação às previsões, uma hipótese está no fato dos dados para treinamento possuírem uma quantidade mais significativa. Por fim, a

Figura 3 exibe os dados analisados de todo o Brasil, com o intuito de prever os acidentes.

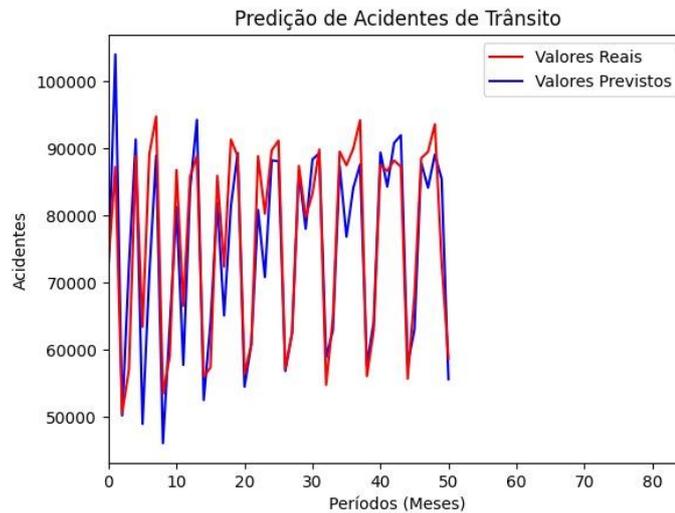


Figura 3: Predição de Acidentes no Brasil.

Na Figura 3 é possível perceber que com o dataset maior o modelo de IA conseguiu prever em grande parte dos períodos.

Para uma análise da acurácia do modelo foi utilizado a MAPE (Mean Absolute Percentage Error - Erro Percentual Absoluto Médio) que é utilizado para medir a magnitude média do erro produzido por um determinado modelo, bem como a distância da predição e do valor real. Por exemplo, um valor MAPE de 10% significa que entre o valor real e a predição existe uma diferença de 10%. Assim, significa que o modelo está em média 10% fora do valor real. Uma previsão com MAPE de 0% significa que a previsão é a mesma que a real. Já o contrário significa uma previsão menos precisa (DE MYTTENAERE, 2016). A Tabela 1 Detalha a MAPE utilizada no modelo, detalhando os resultados na predição dos acidentes municipais, estaduais e nacionais.

Tabela 1 Resultado MAPE dos modelos

Dataset	Valor MAPE
Dados Campo Grande	16,55 %
Dados Mato Grosso do Sul	14,61 %
Dados Brasil	6,58 %

Observou-se que, ao utilizar o tempo como variável preditiva e considerar a presença de sazonalidade em subperíodos mensais em um modelo aditivo, foi alcançado um MAPE abaixo de 20%. Um melhor resultado se deu nos dados nacionais, percebe-se que os resultados foram melhorando à medida em que os dados foram aumentando.

## **CONCLUSÃO**

Neste trabalho, foi realizada uma análise detalhada da estrutura de elaboração do sistema de monitoramento de possíveis acidentes ocorridos a partir de dados reais. Utilizando a regressão com o algoritmo de regressão com séries temporais. O objetivo foi prever locais e incidência de acidentes, com ênfase em aspectos temporais, como data e hora, para reduzir a margem de erro nas previsões. Para aprimorar a precisão dos resultados e lidar com a variabilidade nas distâncias, foi implementada uma abordagem de agrupamento dos dados por meio da biblioteca pandas, matplotlib e numpy, utilizando o método participativo para determinar o número ideal de grupos. Isso permitiu que as funções de regressão fossem aplicadas de maneira mais eficaz a todos, considerando a predição como um todo ao final do processo.

Essa constatação sugere que a regressão de dados com base nas diversas esferas de poder (municipal, estadual e federal), pode produzir resultados mais precisos e menos suscetíveis a erros quando aplicada uma metodologia de amostragem para dividir a base de dados em subconjuntos, permitindo uma análise mais robusta.

Durante a execução deste projeto, diversas abordagens relevantes foram consideradas, mas foram descartadas para manter o foco da pesquisa. Contudo, a inclusão dessas abordagens pode ser considerada em investigações futuras. Um dos aspectos não abordados foi a análise de acidentes levando em conta não apenas atributos temporais, mas também as causas dos acidentes, permitindo investigar potenciais correlações entre os locais e suas respectivas causas.

Em futuras pesquisas, seria interessante explorar diferentes métodos de regressão e realizar uma comparação dos resultados obtidos através da análise por grupos. Outra proposta para investigação futura é o desenvolvimento de um sistema de apoio à decisão, que utilize dados de acidentes previstos para calcular possíveis índices e prever o acidente em sim nas áreas urbanas, contribuindo assim para a diminuição de incidentes e a melhoria da segurança viária.

## REFERÊNCIAS

AGUIAR, Wellington Sousa, et al. "**Uso da Inteligência Artificial Para Predição de Acidentes no Trânsito.**" Disponível em: (2021).

ASSOCIAÇÃO BRASILEIRA DE PREVENÇÃO DOS ACIDENTES DE TRÂNSITO. **Estatísticas nacionais de acidentes de trânsito.** 2020. Disponível em: [http://vias-seguras.com/layout/set/print/os\\_acidentes/estatisticas/estatisticas\\_nacionais](http://vias-seguras.com/layout/set/print/os_acidentes/estatisticas/estatisticas_nacionais). Acesso em 07 out. 2024.

Batista, G. E. d. A. P. et al. (2003). "**Pré-processamento de dados em aprendizado de máquina supervisionado**". Phd thesis, Universidade de São Paulo, ICMC - Instituto de Ciências Matemáticas e de Computação. Disponível em: <http://www.teses.usp.br/teses/disponiveis/55/55134/tde06102003-160219/publico/TeseDoutorado.pdf>. Acesso em 07 out. 2024.

BONIN, Morvana. **Regressão Linear Simples: um pouco sobre regressão linear.** 2019. Disponível em: <https://tableless.com.br/regressao-linear-simples/>. Acesso em 07 out. 2024..

BRASIL. Ministério dos Transportes. **Registro Nacional de Sinistros e Estatísticas de Trânsito**, Atualizado em 23/07/2024. Disponível em:

<https://www.gov.br/transportes/pt-br/assuntos/transito/arquivos-senatran/docs/reanaest>. Acesso em: 11 out. 2024.

DE MYTTENAERE, Arnaud et al. **Mean absolute percentage error for regression models. Neurocomputing**, v. 192, p. 38-48, 2016.

ECORODOVIAS. **Pelo menos, 90% dos acidentes de trânsito poderiam ser evitados.** 2020. Disponível em: <https://g1.globo.com/sp/campinas-regiao/especial-publicitario/ecorodovias/noticia/2020/09/23/pelo-menos-90percent-dos-acidentes-de-transito-poderiam-ser-evitados.ghtml>. Acesso em 07 out. 2024.

HAYKIN, S. (2007). **"Redes neurais: princípios e prática"**. Bookman Editora.

INFANTE, P., Nogueira, V., Manuel, P. R., Góis, P., Afonso, A., Santos, D., Clemente, R. (2022a). **A Sinistralidade Rodoviária no Distrito de Setúbal [versão PDF]**. Évora: Imprensa da Universidade de Évora.

KOCH, Márcio. **"Visão computacional para reconhecimento de faces aplicado na identificação e autenticação de usuários na web"**. 2012. Universidade Regional de Blumenau, Blumenau, SC, 2012

LUDERMIR, Teresa Bernarda. **"Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências"**. Estudos Avançados, v. 35, p. 85-94, 2021.

LUZ, Filipe. **Algoritmo KNN para classificação.** 2019. Disponível em: <https://inferir.com.br/artigos/algoritmo-knn-para-classificacao/>. Acesso em: 30 abr. 2021. SARAGIOTTO, Daniela. Mortes no Trânsito: Tráfego brasileiro mata 1 pessoa a cada 15 minutos. Mobilidade. Estadão, São Paulo, 15 set. 2020. Disponível em: <https://mobilidade.estadao.com.br/mobilidade-com-seguranca/mortes-no-transito-brasileiro-mata-1-pessoa-a-cada-15-minutos/>. Acesso em 07 out. 2024.

MAKRIDAKIS, S. (2017). **"The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms"**, 90, 46–60. <https://doi.org/10.1016/j.futures.2017.03.006>. Acesso em 07 out. 2024.

MONARD, M. C. e Baranauskas, J. A. (2003). **"Conceitos sobre aprendizado de máquina. Sistemas inteligentes-Fundamentos e aplicações"**, 1(1):32.

PENNER, Giovanni Chaves, **Práticas em Saneamento, Clima e Meio Ambiente. Capítulo 5 Aula prática: Detecção de mudanças a partir de série temporal de imagens orbitais utilizando os métodos diferença de NDVI e RCEN modificada.** Disponível em

[https://www.researchgate.net/profile/Helio-Almeida/publication/348799917\\_Aula\\_pratica\\_Sistema\\_de\\_recalque\\_Levantamento\\_da\\_curva\\_de\\_uma\\_bomba\\_centrifuga/links/60228db392851c4ed55d2392/Aula-pratica-Sistema-de-recalque-Levantamento-da-curva-de-uma-bomba-centrifuga.pdf#page=49](https://www.researchgate.net/profile/Helio-Almeida/publication/348799917_Aula_pratica_Sistema_de_recalque_Levantamento_da_curva_de_uma_bomba_centrifuga/links/60228db392851c4ed55d2392/Aula-pratica-Sistema-de-recalque-Levantamento-da-curva-de-uma-bomba-centrifuga.pdf#page=49) acesso em 07 out. 2024.

RELVAS, Tânia Regina Sordi, **Previsões Com Séries Temporais: Cálculo De Coeficientes sazonais Com Regressão Linear**. São Paulo, SP, disponível em: <https://anaiscbc.emnuvens.com.br/anais/article/view/3263/3263> acesso em 07 out. 2024.

SAMMUT, C. e Webb, G. I. (2017). **"Encyclopedia of machine learning and data mining"**. Springer Publishing Company, Incorporated, Boston, MA, 2nd edition.

SANTOS, João Pedro Silva et al. **"Evolução da Inteligência Artificial"**. In: Anais do Congresso Nacional Universidade, EAD e Software Livre. 2020. p. (2).

SILVA JÚNIOR, Valter Eduardo da. **"Uma nova abordagem do Real AdaBoost resistente a overfitting para classificação de dados binários"**. 2016. Dissertação de Mestrado. Universidade Federal de Pernambuco.

SKANSI, S. (2018). **"Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence"**. Springer.

TAMBELLI, Clarice Nassar. **Smart Cities: uma breve investigação crítica sobre os limites de uma narrativa contemporânea sobre cidades e tecnologia. Cidades inteligentes em perspectivas**, 2019.

TEMPORAL, Jessica. **Como definir o número de clusters para o seu KMeans**. 2019. Disponível em: <https://medium.com/pizzadedados/kmeans-e-metodo-do-cotovelo-94ded9fdf3a9>. Acesso em: 11 dez. 2021. VIAS SEGURAS (Brasil).

VALENTE, Pedro Miguel Monteiro. **Prevenção e alerta da sinistralidade rodoviária com o contributo da inteligência artificial**. 2023. Disponível em: [https://comum.rcaap.pt/bitstream/10400.26/46173/1/Pedro%20Monteiro%20Valente\\_Preven%c3%a7%c3%a3o%20e%20alerta%20da%20sinistralidade%20rodovi%c3%a1ria%20%28...%29.pdf](https://comum.rcaap.pt/bitstream/10400.26/46173/1/Pedro%20Monteiro%20Valente_Preven%c3%a7%c3%a3o%20e%20alerta%20da%20sinistralidade%20rodovi%c3%a1ria%20%28...%29.pdf) Acesso 11 out. 2024.