UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**Patrícia Arcelo de Arruda**

# Desenvolvimento de Modelos Explicáveis de Inteligência Artificial para Classificação de Sinistros de Trânsito em Rodovias Federais de Pernambuco

**RECIFE**
**2024**
**UNIVERSIDADE FEDERAL DE PERNAMBUCO**
**CENTRO DE INFORMÁTICA**
**CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**Patrícia Arcelo de Arruda**

# Desenvolvimento de Modelos Explicáveis de Inteligência Artificial para Classificação de Sinistros de Trânsito em Rodovias Federais de Pernambuco

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em de Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de bacharel em Ciência da Computação.

Orientador(a): Ricardo Bastos Cavalcante Prudêncio

**RECIFE**

**2024**

**UNIVERSIDADE FEDERAL DE PERNAMBUCO**

**CENTRO DE INFORMÁTICA**

**CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Arruda, Patrícia Arcelo de.
Desenvolvimento de modelos explicáveis de Inteligência Artificial para classificação de sinistros de trânsito em rodovias federais de Pernambuco / Patrícia Arcelo de Arruda. - Recife, 2024.
59 p. : il., tab.

Orientador(a): Ricardo Bastos Cavalcante Prudêncio
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Ciências da Computação - Bacharelado, 2024.
Inclui referências.

1. Inteligência Artificial Explicável (XAI). 2. Aprendizado de máquina. 3. Classificação de sinistros de trânsito. 4. Rodovias federais de Pernambuco. 5. Fatores de influência. I. Prudêncio, Ricardo Bastos Cavalcante. (Orientação). II. Título.

000 CDD (22.ed.)

**Patrícia Arcelo de Arruda**

# Desenvolvimento de Modelos Explicáveis de Inteligência Artificial para Classificação de Sinistros de Trânsito em Rodovias Federais de Pernambuco

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em de Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de bacharel em Ciência da Computação.

Aprovado em: 08 / 10 / 2024

Banca Examinadora:

Orientador

Ricardo Bastos Cavalcante Prudêncio

Doutor

Examinador

Paulo Salgado Gomes de Mattos Neto

Doutor

# AGRADECIMENTOS

Agradeço aos grandes amigos que fizeram parte desta jornada CIn, iniciada em nossa juventude há tantos anos, e àqueles que me incentivaram a embarcar no desejo de retomá-la. A caminhada foi desafiadora e gratificante. Em especial, dedico a conquista a minha esposa Maria Karla, companheira de todos os momentos, e a nossas filhas Maya e Liz.

# RESUMO

Promover a explicabilidade de modelos de inteligência artificial pode propiciar novos conhecimentos e a descoberta de padrões e relacionamentos do problema estudado. Com essa motivação, este trabalho busca contribuir para o entendimento dos fatores de influência dos sinistros de trânsito nas rodovias federais de Pernambuco. Mediante técnicas de aprendizado de máquina supervisionado, os acidentes serão classificados conforme a sua causa e conforme sua gravidade, aplicando-se métodos de Inteligência Artificial Explicável (XAI) para compreensão das predições. Pretende-se contribuir possibilitando *insights* relevantes para políticas preventivas voltadas à redução dos sinistros de trânsito.

**Palavras-chave**: Inteligência Artificial Explicável (XAI), aprendizado de máquina, modelos de classificação, causa e gravidade de acidentes de trânsito, rodovias federais de Pernambuco, fatores de influência.

# ABSTRACT

Promoting the explainability of artificial intelligence models can provide new knowledge and uncover patterns and relationships within the studied problem. With this motivation, this work aims to contribute to the understanding of the influencing factors of traffic accidents on the federal highways of Pernambuco. Through supervised machine learning techniques, accidents will be classified according to their cause and severity, applying eXplainable Artificial Intelligence (XAI) methods to understand the predictions. This contribution intends to provide insights relevant for preventive policies aimed at reducing traffic accidents.

**Keywords**: eXplainable Artificial Intelligence (XAI), machine learning, classification models, cause and severity of traffic accidents, federal highways of Pernambuco, influencing factors.

# LISTA DE FIGURAS

# LISTA DE TABELAS

# LISTA DE ABREVIATURAS E SIGLAS

| | |
|---|---|
| CSV | Comma-separated Values |
| CRISP-DM | Cross Industry Standard Process for Data Mining |
| DNIT | Departamento Nacional de Infraestrutura de Transportes |
| ICE | Individual Conditional Expectation |
| KNN | K-Nearest Neighbors |
| OMS | Organização Mundial de Saúde |
| PDP | Partial Dependence Plots |
| PFI | Permutation Feature Importance |
| Pnatrans | Plano Nacional de Redução de Mortes e Lesões no Trânsito |
| PRF | Polícia Rodoviária Federal |
| RF | Random Forest |
| SVM | Support Vector Machine |
| XAI | eXplainable Artificial Intelligence (Inteligência Artificial Explicável) |
| XGBoost | eXtreme Gradient Boosting |

# SUMÁRIO

# 1. INTRODUÇÃO

A redução dos impactos sociais causados por sinistros de trânsito mostra-se como um desafio global de alta relevância. Este trabalho busca contribuir com esse propósito, utilizando técnicas de Inteligência Artificial Explicável (XAI) para a análise de sinistros de trânsito em rodovias federais de Pernambuco.

Para isso, serão utilizados dados de acidentes de trânsito disponibilizados pela Polícia Rodoviária Federal[1], com um recorte para ocorrências em rodovias federais no estado de Pernambuco. Serão desenvolvidos modelos explicáveis de classificação dos sinistros, com técnicas de aprendizado de máquina supervisionado e de XAI, para a realização de trabalhos de predição da gravidade dos acidentes e de suas causas.

A partir dessas interpretações com foco na explicabilidade dos modelos, pretende-se identificar os principais fatores que se associam a uma maior ou menor severidade dos sinistros, bem como às categorias de causas de acidentalidade identificadas nos dados. Ao revelar essas conexões, o estudo visa contribuir para a formulação de estratégias mais eficazes de prevenção nessa temática.

## 1.1. Motivação

A busca por medidas eficazes de segurança viária reflete-se em um esforço mundial para salvar vidas e reduzir os ferimentos graves no trânsito. Em seu Plano Global para implementar a Década de Ação pela Segurança no Trânsito 2021-2030 [1], a Organização Mundial de Saúde (OMS) fixa diretrizes para se atingir a ambiciosa meta de reduzir em pelo menos 50% as mortes e lesões no trânsito.

No Brasil, alinhando-se aos compromissos globais, o Plano Nacional de Redução de Mortes e Lesões no Trânsito (Pnatrans) [2] segue desde 2018 como importante ferramenta que visa a um sistema seguro de mobilidade. Um importante enfoque nacional volta-se para as rodovias federais, que conectam todas as regiões do país e têm papel fundamental na infraestrutura de transporte do Brasil [3].

---

[1] Disponível em: <https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos/dados-abertos-da-prf>. Acesso em: 18 jun. 2024.

Nesse cenário, esforços preventivos podem ser beneficiados por estudos que propiciem uma melhor compreensão dos fatores de influência dos sinistros de trânsito, conforme se propõe neste trabalho mediante o emprego de técnicas de inteligência artificial explicável (XAI - do inglês, *eXplainable Artificial Intelligence*). Ao "explicar para descobrir" [4], novos conhecimentos podem ser obtidos, com a identificação de padrões e relacionamentos sobre o problema estudado.

## 1.2. Objetivos

O objetivo geral do trabalho é o desenvolvimento e a interpretação de modelos de XAI que possam classificar eficientemente os sinistros de trânsito nas rodovias federais de Pernambuco, considerando dois aspectos principais: a gravidade e as causas dos acidentes. Com a análise das predições, pretende-se identificar e compreender, a partir dos dados, os principais fatores de influência relacionados a tais aspectos de acidentalidade.

Como objetivos específicos, inicialmente pretende-se a coleta e o tratamento dos dados necessários ao desenvolvimento do trabalho. Em seguida, haverá a implementação dos modelos de classificação utilizando-se técnicas de aprendizado de máquina com o emprego de distintos algoritmos como *Random Forest* (RF), *eXtreme Gradient Boosting* (XGBoost), *Support Vector Machine* (SVM) e *K-Nearest Neighbors* (KNN). Parte-se então para a avaliação dos classificadores construídos, com base em métricas como acurácia, *precision*, *recall* e *f1-score*. Por fim, a interpretação dos modelos de melhor resultado será realizada com o uso de métodos globais de XAI, como *Permutation Feature Importance* (PFI), *Partial Dependence Plots* (PDP) e *Individual Conditional Expectations* (ICE), propiciando-se uma análise quanto aos principais fatores de influência dos sinistros.

# 2. CONTEXTO E REVISÃO DA LITERATURA

Visando a um melhor entendimento sobre o tema em estudo, nesta seção revisaremos trabalhos anteriores que empregaram técnicas de inteligência artificial voltadas à problemática dos acidentes de trânsito em rodovias federais. Este esforço pretende situar nossa investigação no contexto mais amplo das pesquisas sobre segurança viária, destacando nosso enfoque na análise de variáveis e fatores que influenciam a severidade e as causas desses sinistros.

Também abordaremos fundamentos essenciais acerca de técnicas que serão empregadas no desenvolvimento do projeto. Trataremos de algoritmos de classificação utilizados no aprendizado de máquina supervisionado e dos métodos de XAI voltadas à interpretação dos classificadores.

## 2.1. Trabalhos Relacionados

Os esforços científicos e tecnológicos voltados ao entendimento e predição da sinistralidade de trânsito, com vistas a contribuir com possíveis iniciativas para mitigar essa problemática, têm sido objeto de diversas pesquisas desenvolvidas nessa área. Com foco em iniciativas no ramo do aprendizado de máquina, com especial atenção a métodos de classificação, alguns trabalhos prévios foram objeto de nossa revisão.

Recentemente, DE OLIVEIRA *et al.* [5] se propuseram a empregar métodos de classificação tendo como atributo alvo a causa do acidente, utilizando-se de dados disponibilizados pela Polícia Rodoviária Federal referente aos anos de 2007 a 2017, com registros de acidentes em rodovias de todo o país. Nessa abordagem, optaram por empregar dois algoritmos baseados em árvore de decisão (*RandomTree* e *J48*) e concluíram que ambos tiveram resultados similares, apresentando uma taxa de classificação correta de 82% na base de predição. O artigo, contudo, não detalha informações sobre a quantidade de classes do atributo alvo, nem qual foi o desempenho dos classificadores para cada uma dessas classes.

Em uma abordagem similar, OLIVEIRA e BORGES [6] também realizam experimentos similares tomando como base registros de sinistros da PRF dos anos de 2017 a 2021, mas visando a classificação quanto ao atributo multiclasse que categoriza o tipo do acidente (colisão frontal, colisão lateral, capotamento, etc). O estudo se mostrou mais aprofundado, utilizando os algoritmos Árvore de Decisão, *Random Forest* e *K-Nearest Neighbors*, além de técnicas diversas como balanceamento de dados e validação cruzada.

Contudo, os resultados obtidos, baseados na métrica de acurácia, foram tímidos, não superando a taxa de 32%.

Já TAVEIRA *et al*. [7] focam especificamente em uma das principais rodovias do estado de Pernambuco (BR-101), entre os anos de 2017 e 2022, e experimentaram métodos de agrupamento, classificação e regressão. Quanto à classificação, o estudo buscou a predição das causas de acidentalidade, categorizadas em 4 classes, empregando os algoritmos de Árvore de Decisão, *Random Forest* e *eXtreme Gradient Boosting*, atingindo como melhor resultado uma acurácia de 81% e precisão de 57%.

Essas produções refletem uma variedade de abordagens e técnicas utilizadas para compreender a complexidade dos acidentes de trânsito em rodovias federais e ressaltam o potencial da inteligência artificial para aprimorar as estratégias de segurança e prevenção. Nos trabalhos identificados, contudo, não se verificou o emprego de técnicas de XAI para expansão dos experimentos com vistas à análise dos principais fatores de influência dos acidentes de trânsito, conforme pretendido no presente estudo.

## 2.2. Fundamentos e Ferramentas de Análise

A mineração de dados como aplicação do aprendizado de máquina mostra-se como uma importante ferramenta para o entendimento das relações existentes entre múltiplos atributos em um conjunto de dados, conforme destaca KOTSIANTIS [8]. Em busca da extração de conhecimentos acerca dessas possíveis relações no contexto dos acidentes de trânsito ora estudados, serão apresentados alguns fundamentos do aprendizado de máquina supervisionado, tratando-se em especial de técnicas de classificação e de explicabilidade (XAI) que serão empregadas neste trabalho.

### *2.2.1. Técnicas de Classificação*

No aprendizado supervisionado, a partir do treinamento com dados que possuam rótulos conhecidos, é possível desenvolver um classificador com capacidade de generalizar predições para novos dados [8], havendo uma gama de algoritmos que podem ser utilizados em tais tarefas. Dentre eles, vamos nos ater ao *Random Forest*, *eXtreme Gradient Boosting*, *K-Nearest Neighbors* e *Support Vector Machine*.

O *Random Forest*, conforme apontam FAWAGREH, GABER, e ELYAN [9], baseia-se na técnica de *ensemble learning*, através da qual múltiplos modelos individuais são

utilizados em conjunto para aprimorar a acurácia das predições, as quais são escolhidas com base nos votos de cada modelo integrante do *ensemble*. Neste caso, os modelos individuais que compõem o RF são árvores de decisão, cada uma construída a partir de uma amostra aleatória do conjunto de dados, as quais possuem poder de voto igualitário para a decisão do ensamble. Essa abordagem é denominada de *bagging* (*Bootstrap Aggregating*).

O processo de construção das árvores envolve ainda a seleção de subconjuntos de características em cada nó interno (nó de decisão) para determinar como dividir os dados. Alguns critérios podem ser utilizados nessa seleção para otimizar os resultados, como o índice gini (medida de heterogeneidade) a entropia (ganho de informação), e o *log loss* (perda logarítmica). Na implementação do classificador *RandomForestClassifier* disponível na biblioteca scikit-learn [10], esse parâmetro pode ser ajustado conforme necessidade, assim como diversos outros, dos quais citamos *n_estimators* (quantidade de árvores), max_depth (profundidade máxima de cada árvore), *min_samples_split* (mínimo de amostras para dividir um nó) e *min_samples_leaf* (mínimo de amostras em uma folha).

Similarmente, conforme lecionado por BENTÉJAC, CSÖRGŐ e MARTÍNEZ-MUÑOZ [11], o XGBoost também se configura como uma técnica de *ensemble* de árvores de decisão, porém se baseia no conceito de *boosting*, utilizando um treinamento iterativo em que cada novo modelo individual integrante do conjunto é construído para corrigir os erros dos modelos anteriores, minimizando a função de perda.

O algoritmo também adota mecanismos para reduzir a complexidade do modelo, estabelecendo, por exemplo, um limite de profundidade às árvores e um critério de divisão de nós internos que considera uma estratégia de pré-poda, onde é necessário um ganho mínimo da função de perda para permitir a divisão. Ainda, para combater o overfitting, há o parâmetro de taxa de aprendizado ou *shrinkage* e técnicas de randomização para aleatorização de amostras usadas no treinamento das árvores e para subamostragem de *features*. Esses e outros parâmetros podem ser utilizados na implementação do XGBoost disponibilizada para Python [12].

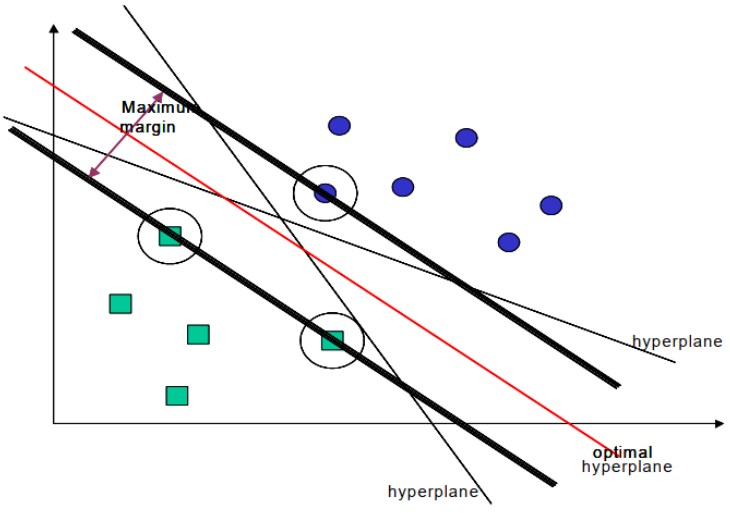
Quanto ao KNN, trata-se de um algoritmo de aprendizado com método estatístico baseado em instâncias, consideradas como pontos em um espaço *n*-dimensional (correspondente aos *n*-atributos do dataset), o qual é pautado no princípio de que observações similares (de uma mesma classe) situam-se próximas umas das outras, conforme explicado

por KOTSIANTIS [8]. Assim, ao se analisar uma nova instância não rotulada, sua classificação será determinada de acordo com a classe mais frequente entre os *k*-vizinhos rotulados mais próximos.

Várias métricas de distância podem ser utilizadas, como por exemplo a Euclidiana, de Minkowsky ou Manhattan, sendo esse um parâmetro regulável na implementação do algoritmo disponível na biblioteca scikit-learn [13], além de outras como *n_neighbors* (número dos k-vizinhos mais próximos) e *weights* (determina se o peso do voto dos vizinhos será uniforme ou conforme sua distância). Idealmente, a métrica de distância utilizada deve maximizar a distância para registros de classes distintas e minimizá-la para instâncias similares [8].

Já o algoritmo SVM, ainda segundo KOTSIANTIS [8], procura determinar um hiperplano que separe duas classes distintas em um espaço multidimensional elevado, de modo que a distância mínima entre os pontos das classes e esse hiperplano, denominada margem, seja maximizada, como mostrado na Figura 1. Para dados que não são linearmente separáveis, como comumente ocorre no mundo real, é empregada uma função de kernel para mapeá-los a um espaço de maior dimensão, onde essa separação linear poderá ocorrer.

**Figura 1** - Otimização do hiperplano e maximização da margem no SVM.




Fonte: KOTSIANTIS [8].

No caso de problemas multiclasse, como o tratado no presente estudo, uma estratégia para a utilização do SVM é a de decompor a tarefa em vários problemas binários, conforme apontam ROCHA e GOLDENSTEIN [14], com abordagens one-versus-one (OVO), que

acarretará em um classificador binário para cada dupla de classes, ou one-versus-all (OVA), que leva a um classificador binário por classe.

Na implementação desse algoritmo constante da biblioteca scikit-learn denominada *Support Vector Classification* [15], adotou-se abordagem OVA, e alguns dos parâmetros ajustáveis permitem selecionar a função kernel (ex.: ‘linear’, ‘poly’, ‘rbf’ e ‘sigmoid), o grau da função (no caso de kernel polinomial) e seu coeficiente ‘gamma’, assim como o parâmetro de regularização ‘C’ para lidar com *overfitting*.

### *2.2.2. Métodos de XAI*

Seguindo-se para os fundamentos do ramo da Inteligência Artificial Explicável, tomando como foco seus métodos globais para a interpretação e compreensão de modelos complexos, trataremos do PFI, do PDP e do ICE, que são métodos diagnósticos baseados em permutação [4].

Inicialmente, cumpre esclarecer que os métodos globais de XAI fornecem uma visão geral para a interpretabilidade de um modelo de aprendizado de máquina. Diferente dos métodos locais, que explicam a decisão do modelo para uma instância específica, os métodos globais ajudam a entender o comportamento e a lógica subjacente do modelo como um todo [4]. Essa característica é especialmente relevante para o presente estudo, pois permitirá compreender como diferentes variáveis impactam as previsões dos modelos desenvolvidos, sinalizando assim os principais fatores de influência dos acidentes de trânsito e como se relacionam.

O PFI é um método de interpretação que avalia a importância das características de um modelo, medindo o impacto individual de cada característica (*feature*) no desempenho do modelo. Consoante explanado por LETRACHE e RAMDANI [16], para cada característica, seus valores no conjunto de dados são aleatoriamente embaralhados, mantendo-se os demais atributos inalterados. Após isso, diante do impacto dessas mudanças na performance do modelo, calcula-se e ranqueia-se a importância de suas *features*. Apesar de sua grande utilidade na interpretação do modelo, o PFI pode ter resultados enviesados se houver atributos com alta correlação, devendo tal ponto ser um aspecto de atenção.

PDP e ICE são técnicas agnósticas, ou seja, não são atreladas a um tipo específico de modelo de aprendizado de máquina, voltadas à interpretabilidade por visualização, conforme

apontado por ADADI e BERRADA [4]. Os gráficos de dependência parcial, ou PDPs, ilustram a relação entre uma ou mais características de entrada e a previsão do modelo, indicando uma média atribuída a essa relação parcial, calculada mediante o impacto da variação dos valores dessas características enquanto os demais atributos são mantidos constantes. Com isso, é possível observar padrões de mudanças nas tendências de predição do modelo à medida em que os valores das características de entrada são variados.

O ICE, por sua vez, é uma extensão dos PDPs que oferece uma visão mais detalhada de como as previsões do modelo variam de acordo com uma determinada característica, com resultados apresentados individualmente para cada instância do conjunto de dados [4]. Graficamente, o ICE mostra uma curva para cada observação do *dataset*, sendo particularmente útil para identificar heterogeneidades na resposta do modelo, que podem não ser visíveis em uma média global como o PDP.

Neste trabalho, dados extraídos a partir do PFI e gráficos PDP e ICE serão apresentados para os modelos de classificação de acidentes desenvolvidos para o presente estudo.

# 3. TRABALHO REALIZADO

## 3.1. Metodologia (CRISP-DM)

A metodologia utilizada para o desenvolvimento deste trabalho se alinha ao *Cross Industry Standard Process for Data Mining* (CRISP-DM), processo de mineração de dados amplamente aceito como um padrão na indústria, conforme apontado por SCHROER [17], o qual engloba as seguintes fases: Entendimento do Negócio, Entendimento dos Dados, Preparação dos Dados, Modelagem, Avaliação e Implantação. Excepcionando-se a fase de implantação, que não foi desenvolvida devido ao escopo acadêmico deste estudo, todas as demais etapas do CRISP-DM foram consideradas na implementação do projeto.

Primeiramente, para o entendimento do negócio, foram traçados os objetivos do projeto, conforme visto anteriormente, estudados trabalhos anteriores em temática similar, e realizada uma avaliação situacional para se verificar a disponibilidade de dados sobre acidentes de trânsito nas rodovias federais de Pernambuco, imprescindíveis ao desenvolvimento dos experimentos e análises.

Em seguida, procedeu-se à coleta, entendimento e preparação dos dados necessários, bem como à modelagem dos classificadores desenvolvidos, voltados à predição da gravidade e das causas dos sinistros de trânsito, conforme apresentado nos próximos subtópicos. Quanto à fase de avaliação, em que analisamos o desempenho dos modelos e interpretamos os conhecimentos obtidos a partir dos resultados, é tratada na seção "*3.3. Avaliação e Interpretação*".

Toda a implementação foi realizada na linguagem de programação Python, com o emprego das bibliotecas especializadas, dentre as quais destacam-se: Pandas [18] para manipulação e preparação de dados; Matplotlib para visualização de dados [19]; Scikit-learn [20] e Optuna [21] para a construção e avaliação de modelos de aprendizado de máquina; e MLflow [22] para o gerenciamento de experimentos e comparação de resultados.

Ao final, após seleção dos melhores modelos construídos, foram utilizadas técnicas de XAI para análise e interpretação das predições, de modo a proporcionar uma melhor compreensão acerca dos principais fatores que influenciam os acidentes estudados.

### *3.1.1. Aquisição e Entendimento dos Dados*

Os dados de acidentalidade necessários ao projeto são oriundos de registros da Polícia Rodoviária Federal (PRF) e foram obtidos em sítio eletrônico disponibilizado pela instituição[2]. São classificados como dados abertos, sendo portanto acessíveis ao público e de livre utilização, conforme estabelece o Decreto nº 8.777/2016, que institui a Política de Dados Abertos do Poder Executivo federal.

As informações obtidas estavam estruturadas em arquivos no formato CSV (*Comma-separated Values*), compartimentadas por ano e por recortes dos dados designados como "acidentes agrupados por ocorrência" e "acidentes agrupados por pessoa".

Inicialmente, com o objetivo de utilizar informações recentes para a análise, foram coletados dados referentes ao estado de Pernambuco dos anos de 2019 a 2023, totalizando-se 5 anos de registros de acidentes, com 13.606 ocorrências, e 32.326 pessoas envolvidas nos sinistros. Esses dados foram consolidados em nosso conjunto de dados principal, denominado, 'ocorrencias', e no conjunto auxiliar denominado 'envolvidos', cujas informações serão úteis para integrações a serem realizadas na fase de tratamento dos dados. Originalmente, os atributos constantes desses conjuntos de dados trazem as informações explicitadas nas Tabelas 1 e 2.

**Tabela 1** - Atributos originais do dataset principal 'ocorrencias'.

| Atributo Original | Descrição |
|---|---|
| id | Identificador único do registro do acidente em banco de dados |
| data_inversa | Data do acidente no formato aaaa-mm-dd |
| dia_semana | Dia da semana do acidente por extenso |
| horario | Hora em que o acidente ocorreu no formato hh:mm:ss |
| uf | Sigla da Unidade Federativa onde ocorreu o acidente |
| br | Número da rodovia federal onde ocorreu o acidente |
| km | Quilômetro da rodovia onde ocorreu o acidente |
| municipio | Município onde ocorreu o acidente |
| causa_acidente | Causa principal do acidente |
| tipo_acidente | Tipo de acidente (colisão frontal, atropelamento, capotamento, etc.) |
| classificacao_acidente | Gravidade do acidente (sem vítimas, vítimas feridas ou vítimas fatais) |
| fase_dia | Fase do dia em que ocorreu o acidente (dia, noite, etc.) |

[2] Disponível em: <https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos/dados-abertos-da-prf>. Acesso em: 18 jun. 2024.

| | |
|---|---|
| sentido_via | Sentido da via (crescente ou decrescente) |
| condicao_metereologica | Condições meteorológicas no momento do acidente |
| tipo_pista | Tipo de pista (simples, dupla, múltipla) |
| tracado_via | Traçado da via (reta, curva, cruzamento, etc.) |
| uso_solo | Indica se o acidente ocorreu em área urbana ou rural |
| pessoas | Número total de pessoas envolvidas no acidente |
| mortos | Número de mortos no acidente |
| feridos_leves | Número de pessoas com ferimentos leves |
| feridos_graves | Número de pessoas com ferimentos graves |
| ilesos | Número de pessoas ilesas |
| ignorados | Número de pessoas com estado desconhecido |
| feridos | Número total de feridos |
| veiculos | Número de veículos envolvidos no acidente |
| latitude | Latitude onde ocorreu o acidente |
| longitude | Longitude onde ocorreu o acidente |
| regional | Regional da PRF responsável pela ocorrência |
| delegacia | Delegacia da PRF responsável pela ocorrência |
| uop | Unidade Operacional da PRF responsável pela ocorrência |

Fonte: Elaborada pela autora, com base nos dados disponibilizados pela PRF.

**Tabela 2** - Atributos originais do dataset auxiliar ‘envolvidos’.

| Atributo Original | Descrição |
|---|---|
| pesid | Identificador único da pessoa envolvida no acidente em banco de dados |
| id_veiculo | Identificador do veículo envolvido no acidente |
| tipo_veiculo | Tipo de veículo envolvido (carro, moto, caminhão, etc.) |
| marca | Marca do veículo envolvido |
| ano_fabricacao_veiculo | Ano de fabricação do veículo envolvido |
| tipo_envolvido | Tipo de envolvimento da pessoa no acidente (condutor, pedestre) |
| estado_fisico | Condição física da pessoa envolvida após o acidente |
| idade | Idade da pessoa envolvida |
| sexo | Sexo da pessoa envolvida |
| ilesos | Indica se a pessoa envolvida saiu ilesa |
| feridos_leves | Indica se a pessoa envolvida sofreu ferimentos leves |
| feridos_graves | Indica se a pessoa envolvida sofreu ferimentos graves |
| mortos | Indica se a pessoa envolvida morreu no acidente |

*Obs.: Além desses atributos, o dataset ‘envolvidos’ também apresenta as demais colunas existentes no dataset ‘ocorrencias’.*

Fonte: Elaborada pela autora, com base nos dados disponibilizados pela PRF.

Posteriormente, para enriquecer nossa base de dados e permitir experimentações com modelos treinados com dados mais amplos, também foram agregados dados dos anos de 2017 e 2018, passando a um total de 7 anos de registros de acidentes, com 19.763 ocorrências e 47.243 pessoas envolvidas. Essa base estendida foi utilizada apenas em alguns experimentos, conforme indicado na seção “*3.1.3. Modelagem*”.

Na busca de outras informações úteis a serem agregadas a nosso conjunto de dados, também foram obtidos dados de equipamentos do tipo "controlador eletrônico de velocidade", usualmente conhecidos como radares, instalados nas rodovias federais de Pernambuco. Tratam-se igualmente de dados classificados como abertos, disponibilizados em sítio eletrônico do Departamento Nacional de Infraestrutura de Transportes (DNIT)[3]. Esses dados contabilizam 72 instâncias e foram carregados no dataset auxiliar denominado ‘radares’, contendo as informações apresentadas na Tabela 3.

[3] Disponível em: <https://servicos.dnit.gov.br/multas/informacoes/equipamentos-fiscalizacao>. Acesso em: 18 jun. 2024.

**Tabela 3** - Atributos originais do dataset auxiliar ‘radares’.

| Atributo | Descrição |
|---|---|
| Código do Equipamento | Identificador único do equipamento de radar |
| Equipamento | Tipo de equipamento de fiscalização |
| UF | Unidade Federativa onde o equipamento está instalado |
| Município | Município onde o equipamento está localizado |
| Rodovia | Rodovia federal onde o equipamento está instalado |
| Km | Quilômetro da rodovia onde o equipamento está localizado |
| Coordenadas | Coordenadas geográficas da localização do equipamento |
| Registro INMETRO | Número de registro do equipamento junto ao INMETRO |
| Nº de Série | Número de série do equipamento |

Fonte: Elaborada pela autora, com base nos dados disponibilizados pelo DNIT.

O entendimento inicial de todos esses dados obtidos foi essencial para determinar quais informações seriam relevantes para os objetivos do projeto, guiando as decisões acerca de seu tratamento e seleção de variáveis para utilização nas análises executadas nas fases subsequentes.

### *3.1.2. Preparação dos Dados*

Durante o pré-processamento dos dados foram executadas diversas tarefas visando à limpeza, transformação, integração e seleção das informações, de modo à melhor adequação do conjunto à subsequente modelagem dos classificadores.

Inicialmente, identificamos a consistência e unicidade dos ‘id’ de todas as instâncias constantes do dataset ‘ocorrencias’, assim como inspecionamos atributos chave, referentes ao tempo, local e circunstâncias do sinistro, findando por concluir pela inexistência de registros em duplicidade.

No tratamento de instâncias nulas desse conjunto de dados, detectamos a existência de alguns poucos registros para os quais o campo ‘br’ estava ausente. Neste caso, optamos por excluir esses registros, realizando a limpeza dos dados, pois indicam sinistros ocorridos em áreas que não se configuram como rodovias federais e, portanto, não se incluem no escopo do estudo.

Na transformação dos dados, a partir das colunas 'data_inversa' e 'horario', criamos o atributo ‘data_hora’ como uma representação numérica de um *datetime* expresso em

segundos. Também criamos os atributos ‘classes_8’[4] e ‘classes_condutor’[5], que representam distintos agrupamentos da coluna ‘causa_acidente’, a qual originalmente continha 86 possíveis valores. A criação destes atributos foi essencial para simplificar a análise e modelagem, dado o alto número de categorias distintas na variável original ‘causa_acidente’.

O atributo ‘classes_condutor’, em particular, foi escolhido como target para a predição da causa do acidente, conforme veremos na seção “*3.1.3. Modelagem*”, devido à sua relevância prática ao distinguir os sinistros cuja causa é de responsabilidade de algum dos condutores envolvidos, daqueles ocasionados por outros fatores diversos.

Também foram realizadas relevantes integrações de dados para enriquecer as informações constantes do conjunto de dados ‘ocorrencias’, o qual passou a conter os atributos representativos das categorias de veículos envolvidos nos acidentes (‘agricolas’, ‘automoveis’, ‘caminhoes’, ‘motocicletas’, ‘nao_motorizados’, ‘onibus’, ‘reboque’s e ‘veiculos_ignorados’), das faixas etárias de seus condutores (‘0-18’, ‘19-28’, ‘29-38’, ‘39-48’, ‘49-58’, ‘59-68’, ‘69-78’, ‘79+’, ‘idade_ignorada’), da quantidade de ‘pedestres’ atingidos no acidente e do número de radares existentes nas proximidades do sinistro em até 1km de distância (‘radar_1km’). Essas informações foram obtidas a partir de análises e tratamento dos atributos ‘tipo_veiculo’, ‘marca’, 'tipo_envolvido' e ‘idade’ constantes do conjunto de dados ‘envolvidos’, assim como dos atributos ‘UF’, ‘Rodovia’ e ‘Km’ do conjunto de dados ‘radares’.

Ao final, além das variáveis agregadas ao dataset ‘ocorrências’, foram selecionados os atributos que seriam relevantes para o processo de aprendizagem dos modelos, os quais indicassem informações como tempo e local do acidente (‘data_hora’, ‘dia_semana’, ‘fase_dia’, ‘br’, ‘km’, ‘municipio’, ‘sentido_via’), características viárias e climáticas (‘tipo_pista’, ‘uso_solo’[6], ‘condicao_metereologica’) e métricas de pessoas envolvidas no sinistro (‘mortos’, ‘feridos_leves’, ‘feridos_graves’, ‘ilesos’, ‘ignorados’). Outros atributos

[4] Categorias do atributo ‘classes_8’: Atenção ou Reação Ineficientes, Desrespeito a normas, Imprudência, Uso de Psicoativos, Estrutura Viária, Condições Veiculares, Fatores Naturais, Enfermidade ou Agressão.
[5] Categorias do atributo ‘classes_condutor’: Condutor, Outro.
[6] O atributo ‘uso_solo’ indica se a via é em área urbana ou rural.

irrelevantes (‘id’, ‘uf’, ‘regional’, ‘delegacia’, ‘uop’[7]) ou redundantes (‘veiculos’, ‘pessoas’, ‘feridos’, ‘latitude’, ‘longitude’[8]) foram excluídos.

Todos os atributos foram devidamente formatados como tipos numéricos ou categóricos, sendo estes codificados através de *Label Encoding* para viabilizar o posterior emprego dos algoritmos de *machine learning*. Essa é uma técnica de pré-processamento de dados usada para transformar entradas categóricas em números, onde cada rótulo único dentro de uma coluna é atribuído a um número inteiro sequencial.

Ao final, chegou-se à configuração de variáveis apresentada na Tabela 4, resultando em 13.569 instâncias. É importante ressaltar que na fase de modelagem, parte desses atributos ainda será descartada para melhor adequação aos trabalhos de classificação a serem realizados, quais sejam, o de predição da gravidade dos acidentes ou de predição da causa dos sinistros.

[7] Como todos os acidentes do conjunto de dados se restringem ao estado de Pernambuco, o atributo ‘uf’ era idêntico para todas as instâncias. Já os atributos ‘regional’, ‘delegacia’, ‘uop’ indicam unidades organizacionais da PRF para controle interno.

[8] Os atributos ‘veículos’, ‘pessoas’ e ‘feridos’ indicam os quantitativos totais desses elementos no acidente, sendo mero somatório de valores de outros atributos. Já ‘latitude’ e ‘longitude’ são dispensáveis diante da existência de variáveis que já indicam satisfatoriamente o local do acidente.

**Tabela 4** - Atributos finais do dataset principal ‘ocorrencias’ após tratamento

| Atributo | Datatype | Descrição |
|---|---|---|
| data_hora | int64 | Data e hora do acidente (timestamp) |
| dia_semana | category | Dia da semana em que ocorreu o acidente. |
| br | category | Rodovia federal onde ocorreu o acidente. |
| km | float64 | Quilômetro da rodovia onde ocorreu o acidente. |
| municipio | category | Município onde ocorreu o acidente. |
| tipo_acidente | category | Tipo de acidente (colisão frontal, atropelamento, etc.). |
| fase_dia | category | Fase do dia em que ocorreu o acidente (dia, noite, etc.). |
| sentido_via | category | Sentido da via (crescente ou decrescente). |
| condicao_metereologica | category | Condições meteorológicas no momento do acidente. |
| tipo_pista | category | Tipo de pista (simples, dupla, múltipla). |
| uso_solo | category | Indica se o acidente ocorreu em área urbana ou rural. |
| mortos | int64 | Número de mortos no acidente. |
| feridos_leves | int64 | Número de pessoas com ferimentos leves. |
| feridos_graves | int64 | Número de pessoas com ferimentos graves. |
| ilesos | int64 | Número de pessoas ilesas. |
| ignorados | int64 | Número de pessoas com estado desconhecido. |
| agricolas | int64 | Número de veículos agrícolas envolvidos no acidente. |
| automoveis | int64 | Número de automóveis envolvidos no acidente. |
| caminhoes | int64 | Número de caminhões envolvidos no acidente. |
| motocicletas | int64 | Número de motocicletas envolvidas no acidente. |
| nao_motorizados | int64 | Número de veículos não motorizados envolvidos no acidente. |
| onibus | int64 | Número de ônibus envolvidos no acidente. |
| reboques | int64 | Número de reboques envolvidos no acidente. |
| veiculos_ignorados | int64 | Número de veículos com tipo ignorado envolvidos no acidente. |
| 0-18 | int64 | Número de condutores na faixa etária de 0 a 18 anos. |
| 19-28 | int64 | Número de condutores na faixa etária de 19 a 28 anos. |
| 29-38 | int64 | Número de condutores na faixa etária de 29 a 38 anos. |
| 39-48 | int64 | Número de condutores na faixa etária de 39 a 48 anos. |
| 49-58 | int64 | Número de condutores na faixa etária de 49 a 58 anos. |
| 59-68 | int64 | Número de condutores na faixa etária de 59 a 68 anos. |
| 69-78 | int64 | Número de condutores na faixa etária de 69 a 78 anos. |
| 79+ | int64 | Número de condutores na faixa etária acima de 79 anos. |
| idade_ignorada | int64 | Número de condutores com idade ignorada. |
| pedestres | int64 | Número de pedestres envolvidos no acidente. |
| radar_1km | int64 | Número de radares em um raio de 1 km do local do acidente. |
| classes_8 | category | Causa do acidente (agrupamento em 8 categorias). |
| **classificacao_acidente** | **category** | **Gravidade do acidente (sem vítimas, feridos ou fatais).** |
| **classes_condutor** | **category** | **Classe binária da causa do acidente (Condutor ou Outros).** |

Fonte: Elaborada pela autora, após tratamento dos dados.

### *3.1.3. Modelagem*

A fase de modelagem do projeto envolveu a seleção de algoritmos de aprendizado de máquina supervisionado para a classificação dos acidentes, assim como a construção e otimização dos modelos com base em seu desempenho. Seguimos dois enfoques distintos, sendo o primeiro voltado ao trabalho de predição com base na severidade dos sinistros, e o segundo voltado à predição da causa dos acidentes.

#### *3.1.3.1. Predição da Gravidade dos Acidentes*

Nesta primeira abordagem, com foco na predição da gravidade dos sinistros de trânsito, selecionamos como *target* a variável 'classificacao_acidente'. Esse atributo multiclasse possui três categorias distintas que remetem à existência de vítimas conforme gravidade dos danos físicos sustentados pelas pessoas envolvidas no sinistro, quais sejam: ‘sem vítimas’, ‘com vítimas feridas’ e ‘com vítimas fatais’. Consideramos que cada uma dessas categorias representa, respectivamente, as classes de gravidade baixa, gravidade moderada e gravidade alta do acidente.

Com o recorte de dados inicialmente utilizado, que engloba os sinistros registrados entre os anos de 2019 e 2023, a distribuição verificada dentre as três classes é mostrada na Tabela 5, onde se nota significativo desbalanceamento com a categoria ‘com vítimas feridas’ (gravidade moderada) sendo disparadamente majoritária. Essa situação será considerada no desenvolvimento dos modelos.

**Tabela 5** - Informações do *target* ‘classificacao_acidente’

| Valor Numérico | Valor Categórico | Gravidade do Acidente | Quantidade de Instâncias | Percentual do Total de Instâncias |
|---|---|---|---|---|
| 0 | Sem Vítimas | Baixa | 2621 | 19,32% |
| 1 | Com Vítimas Feridas | Moderada | 9597 | 70,73% |
| 2 | Com Vítimas Fatais | Alta | 1351 | 9,96% |
| *Base de dados: Acidentes de 2019 a 2023* | | | | |

Fonte: Elaborada pela autora, conforme dados tratados.

Para possibilitar adequadamente o treinamento dos classificadores, a seleção de features foi refinada com base no *target* escolhido, excluindo-se atributos que estivessem diretamente associados à variável alvo, quais sejam: 'mortos', 'feridos_leves', 'feridos_graves', 'ilesos', 'ignorados' (v. Tabela 4). Também foi retirado o atributo ‘classes_condutor’, pois o

propósito dessa variável é sua utilização como *target* na segunda abordagem a ser realizada na modelagem, com foco na causa dos acidentes.

Quanto aos algoritmos empregados, optamos por trabalhar com RF, SVM e KNN, implementados com a biblioteca scikit-learn, e com o XGBoost, que tem sua própria biblioteca. O objetivo foi explorar diferentes técnicas e identificar aquela que melhor se adapta ao conjunto de dados em estudo.

Construímos então 6 modelos distintos para a predição da gravidade dos acidentes, os quais denominamos de G1 a G6. Para todos eles a otimização de hiperparâmetros durante o treinamento foi realizada através do Optuna, com meta de melhora da acurácia geral, e todos os experimentos e resultados foram registrados utilizando MLFlow, garantindo a rastreabilidade e a reprodutibilidade das experimentações.

Para todos os modelos, dividimos a base de dados em conjuntos de treinamento (50%), validação (25%) e teste (25%), de modo a permitir adequadamente o desenvolvimento, ajuste de parâmetros e avaliação final dos modelos. Os classificadores foram então treinados usando o conjunto de treinamento, enquanto o conjunto de validação foi utilizado para ajustar hiperparâmetros e realizar otimizações. O conjunto de teste foi utilizado para a avaliação final dos modelos, proporcionando uma medida de desempenho dos classificadores em dados não vistos anteriormente.

Na construção do modelo G1 utilizamos o algoritmo RF e seu treinamento foi realizado sem um prévio balanceamento das distintas classes constantes no conjunto de dados. As instâncias do conjunto compreenderam acidentes ocorridos entre os anos de 2019 a 2023.

Já no modelo G2, seguimos utilizando o RF, mas realizamos o balanceamento de dados para as distintas classes nos recortes de treinamento e validação, utilizando método de *random oversampling*, que replica aleatoriamente instâncias das classes minoritárias até que todas as classes tenham o mesmo número de registros. O objetivo é melhorar o equilíbrio no desempenho dos modelos ao equalizar a distribuição das classes. Como o método se mostrou eficiente para a consecução de tal objetivo, passamos a utilizar o balanceamento de dados para todos os modelos subsequentes.

Para os modelos G3, G4 e G5, empregamos respectivamente os algoritmos XGBoost, SVM e KNN. No caso do G5, aplicamos o método de codificação *One-Hot Encoding* para os

atributos categóricos, transformando-se cada categoria única dentro de um atributo categórico em uma nova variável binária (0 ou 1). Optamos por essa abordagem para o classificador KNN para evitar distorções no cálculo da distância entre os pontos de dados, ao retirar as discrepâncias que poderiam ser causadas, neste caso, pela codificação simples do *Label Encoding*, que introduz uma ordem artificial entre as categorias.

Por fim, para o modelo G6 tornamos a utilizar o RF, por ter apresentado melhores resultados que os demais algoritmos, e optamos por ampliar a base de dados para incluir registros de acidentes dos anos de 2017 a 2023, de modo a aumentar a quantidade de dados disponíveis durante o treinamento do classificador, com vistas a obter uma melhor generalização e performance.

Na Tabela 6, é possível observar resumidamente as informações de modelagem de todos esses classificadores, sendo apresentados inclusive os melhores hiperparâmetros obtidos para cada um deles.

**Tabela 6** - Modelagem de G1 a G6 (gravidade do acidente).

(continua)

| Modelo | Algoritmo | Base de Dados | Divisão dos Dados | Técnicas Utilizadas | Melhores Parâmetros |
|---|---|---|---|---|---|
| G1 | Random Forest | 2019 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | - | *criterion*: gini, *n_estimators*: 294, *max_depth*: 41, *min_samples_split*: 10, *min_samples_leaf*: 1 |
| G2 | Random Forest | 2019 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | Balanceamento de dados | *criterion*: log_loss, *n_estimators*: 121, *max_depth*: 45, *min_samples_split*: 8, *min_samples_leaf*: 8 |
| G3 | XGBoost | 2019 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | Balanceamento de dados | *objective*: multi, *lambda*: 2.79e-05, *alpha*: 0.0066, *colsample_bytree*: 0.744, *subsample*: 0.724, *learning_rate*: 0.072, *n_estimators*: 427, *max_depth*: 14, *min_child_weight*: 74 |

**Tabela 6** - Modelagem de G1 a G6 (gravidade do acidente).

(conclusão)

| Modelo | Algoritmo | Base de Dados | Divisão dos Dados | Técnicas Utilizadas | Melhores Parâmetros |
|---|---|---|---|---|---|
| G4 | SVM | 2019 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | Balanceamento de dados | *C*: 0.0589, *kernel*: rbf, *gamma*: scale |
| G5 | KNN | 2019 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | Balanceamento de dados OneHotEncoding | *metric*: manhattan, *weights*: uniform, *n_neighbors*: 96 |
| G6 | Random Forest | 2017 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | Base de dados ampliada Balanceamento de dados | *criterion*: log_loss, *n_estimators*: 298, *max_depth*: 80, *min_samples_split*: 4, *min_samples_leaf*: 7 |

Fonte: Elaborada pela autora, conforme dados tratados.

Na seção “3.3. Avaliação e Interpretação”, trataremos da avaliação dos resultados obtidos para cada um desses modelos, selecionando o que melhor se adequa à tarefa pretendida pelo estudo.

#### *3.1.3.2. Predição da Causa dos Acidentes*

Na abordagem voltada à predição da causa dos acidentes, selecionamos como *target* a variável 'classes_condutor’, um atributo binário cujas categorias indicam se a causa do sinistro está relacionada a algum condutor envolvido (valor: ‘condutor’) ou a outros fatores diversos (valor: ‘outros’). Com o recorte de dados inicialmente utilizado, que engloba os sinistros registrados entre os anos de 2019 e 2023, a distribuição verificada dentre as duas classes é mostrada na Tabela 7, onde se nota um forte desbalanceamento, tendo-se a classe ‘condutor’ como majoritária. Essa situação será considerada no desenvolvimento dos modelos.

**Tabela 7** - Informações do target ‘classes_condutor’.

| Valor Numérico | Valor Categórico | Causa do Acidente | Quantidade de Instâncias | Percentual do Total de Instâncias |
|---|---|---|---|---|
| 0 | Condutor | Relacionada a algum condutor envolvido | 10044 | 74,02% |
| 1 | Outros | Relacionada a outros fatores | 3525 | 25,98% |
| *Base de dados: Acidentes de 2019 a 2023* | | | | |

Fonte: Elaborada pela autora, conforme dados tratados.

A partir do conjunto de dados obtido ao final da fase de preparação de dados, refinamos a seleção de atributos para esse *target*, excluindo-se a variável ‘classes_8’. Isso porque tal *feature* guarda redundância com os dados da variável alvo, uma vez que também corresponde a um agrupamento das causas do acidente, contendo porém 8 categorias. Todos os demais atributos foram mantidos (v. Tabela 4).

Quanto aos algoritmos empregados, optamos por trabalhar apenas com o *Random Forest*, uma vez que nos experimentos anteriores, realizados na abordagem voltada à predição da severidade dos acidentes, esse algoritmo demonstrou ter melhor adaptação ao nosso conjunto de dados. O detalhamento quanto à performance de cada modelo será visto na seção “*3.3. Avaliação e Interpretação*”.

Construímos então 3 modelos RF distintos para a predição da causa dos acidentes, os quais denominamos de C1 a C3, variando em cada um deles aspectos referentes à utilização ou não de balanceamento de dados, ou ao tamanho do conjunto de dados utilizado. A otimização de hiperparâmetros foi igualmente realizada através do Optuna, com meta de melhora da acurácia geral, e os experimentos foram registrados utilizando o MLFlow. Similarmente aos experimentos anteriores, para estes modelos também particionamos a base de dados em conjuntos de treinamento (50%), validação (25%) e teste (25%), de modo a permitir adequadamente o desenvolvimento, ajuste de parâmetros e avaliação final dos modelos.

No treinamento do modelo C1, utilizamos conjunto de dados compreendendo acidentes ocorridos entre os anos de 2019 a 2023 e optamos por não realizar o balanceamento de suas distintas classes. Para o modelo C2, seguimos com o mesmo conjunto de dados, mas realizamos o balanceamento das classes nos conjuntos de treinamento e validação, utilizando o método *random oversampling*, com o objetivo de melhorar o equilíbrio no desempenho dos classificadores ao equalizar a distribuição das classes.

Quanto ao modelo C3, mantivemos o balanceamento de classes e ampliamos a base de dados para incluir registros de acidentes dos anos de 2017 a 2023, aumentando a quantidade de instâncias disponíveis durante o treinamento do classificador, com vistas a obter uma melhor generalização e performance.

Na Tabela 8, é possível observar resumidamente as informações de modelagem desses classificadores, sendo apresentados ainda os melhores hiperparâmetros obtidos para cada um deles.

**Tabela 8** - Modelagem de C1 a C7 (causa do acidente).

| Modelo | Algoritmo | Base de Dados | Divisão dos Dados | Técnicas Utilizadas | Melhores Parâmetros |
|---|---|---|---|---|---|
| C1 | Random Forest | 2019 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | - | *criterion*: entropy, *n_estimators*: 90, *max_depth*: 25, *min_samples_split*: 2, *min_samples_leaf*: 5 |
| C2 | Random Forest | 2019 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | Balanceamento dos dados | *criterion*: gini, *n_estimators*: 186, *max_depth*: 10, *min_samples_split*: 9, *min_samples_leaf*: 3 |
| C3 | Random Forest | 2017 a 2023 | Teste 25%, Validação 25%, Treinamento 50% | Base de dados ampliada Balanceamento de dados | *criterion*: log_loss, *n_estimators*: 125, *max_depth*: 47, *min_samples_split*: 5, *min_samples_leaf*: 8 |

Fonte: Elaborada pela autora, conforme dados tratados.

Na seção "*3.3.2.1. Resultados dos Modelos Desenvolvidos*", trataremos da avaliação dos resultados obtidos para cada um desses modelos, selecionando o que melhor se adequa à tarefa pretendida pelo estudo.

## 3.2. Avaliação e Interpretação

Após o desenvolvimento dos modelos de classificação, passaremos à avaliação de sua eficácia e selecionaremos os de melhor desempenho para, em seguida, interpretar os resultados obtidos. A avaliação é essencial para assegurar a precisão e a confiabilidade das previsões geradas, enquanto a interpretação possibilitará melhor entendimento dos fatores que mais influenciam os resultados.

Será avaliado não apenas o comportamento em geral dos modelos, mas também como eles se portam frente às diferentes categorias de dados, tendo-se em vista especialmente o cenário de desbalanceamento de classes verificado no problema estudado. Dessa forma, as métricas de desempenho utilizadas serão a acurácia, *precision*, *recall* e *f1-score*.

Identificados os melhores modelos, tanto para classificação da causa do acidente quanto de sua gravidade, faremos sua interpretação com o emprego de métodos de XAI. O enfoque será em explicações globais auferidas a partir de técnicas como PFI, PDP e ICE, permitindo extrair conclusões sobre fatores de relevância para os sinistros de trânsito aprendidos pelos modelos.

### *3.2.1. Predição da Gravidade do acidente*

Nesta seção, trataremos dos modelos de classificação desenvolvidos especificamente para prever a gravidade dos acidentes nas rodovias federais de Pernambuco, tomando como *target* o atributo ‘classificacao_acidente’.

Esses classificadores foram treinados com 4 (quatro) algoritmos de aprendizado de máquina, quais sejam, *Random Forest*, XGBoost, SVM e KNN, fazendo-se uso de distintas abordagens. Em alguns casos, realizamos o balanceamento dos dados de treinamento e validação para lidar com desequilíbrios na distribuição das classes, em outros ampliamos o conjunto de dados para obtermos um treinamento mais amplo.

#### *3.2.1.1. Resultados dos Modelos Desenvolvidos*

Os resultados são apresentados em tabelas que mostram as métricas de desempenho para cada modelo testado, os quais, como visto, são nomeados de G1 a G6.

Na Tabela 9, indicam-se os algoritmos de treinamento utilizado, qual intervalo de ocorrências o conjunto de dados engloba, se houve prévio balanceamento de dados (treinamento e validação), e os resultados obtidos para as métricas de acurácia geral (proporção de predições corretas no total de casos) e *precision*, *recall* e *f1-score* médios.

Na Tabela 10, são detalhadas para cada modelo as métricas de *precision* (proporção de verdadeiros positivos no total de predições positivas), *recall* (proporção de verdadeiros positivos no total de casos realmente positivos) e *f1-score* (média harmônica entre *precision* e *recall*) obtidas para cada uma das 3 classes de gravidade existentes no conjunto de dados.

**Tabela 9** - Modelos de predição de gravidade dos acidentes (resultados gerais).

| Modelo | Algoritmo | Dados | Balancea mento | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|---|---|---|
| G1 | Random Forest | 2019 a 2023 | Não | **0.76** | **0.72** | 0.48 | 0.52 |
| G2 | Random Forest | 2019 a 2023 | Sim | 0.64 | 0.56 | **0.68** | 0.58 |
| G3 | XGBoost | 2019 a 2023 | Sim | 0.63 | 0.56 | **0.68** | 0.57 |
| G4 | SVM | 2019 a 2023 | Sim | 0.53 | 0.51 | 0.61 | 0.48 |
| G5 | KNN | 2019 a 2023 | Sim | 0.57 | 0.52 | 0.64 | 0.52 |
| G6 | Random Forest | 2017 a 2023 | Sim | 0.66 | 0.58 | 0.65 | **0.59** |

Fonte: Elaborada pela autora, com base nos experimentos.

**Tabela 10** - Modelos de predição de gravidade dos acidentes (resultados por classe)

| Modelo | Class 0 Precision | Class 1 Precision | Class 2 Precision | Class 0 Recall | Class 1 Recall | Class 2 Recall | Class 0 F1-Score | Class 1 F1-Score | Class 2 F1-Score |
|---|---|---|---|---|---|---|---|---|---|
| G1 | **0.73** | 0.77 | **0.67** | 0.36 | **0.96** | 0.13 | 0.48 | **0.85** | 0.22 |
| G2 | 0.47 | 0.89 | 0.32 | 0.81 | 0.59 | 0.65 | 0.60 | 0.71 | 0.43 |
| G3 | 0.50 | 0.88 | 0.28 | 0.78 | 0.59 | 0.67 | 0.61 | 0.70 | 0.40 |
| G4 | 0.38 | **0.90** | 0.24 | **0.85** | 0.43 | 0.56 | 0.52 | 0.59 | 0.34 |
| G5 | 0.23 | 0.88 | 0.46 | 0.69 | 0.51 | **0.72** | 0.34 | 0.65 | **0.56** |
| G6 | 0.51 | 0.86 | 0.38 | 0.82 | 0.63 | 0.49 | **0.63** | 0.72 | 0.43 |

Fonte: Elaborada pela autora, com base nos experimentos.

O modelo G1, empregando o algoritmo *Random Forest* sem balanceamento de dados demonstrou uma acurácia geral de 76%, a melhor entre todos os modelos testados. Contudo, essa métrica não reflete integralmente o desempenho do classificador, devido à significativa discrepância observada na performance entre as diferentes classes.

A análise dos resultados aponta que, enquanto a Classe 1 (gravidade moderada) alcançou um recall excepcionalmente alto de 96%, indicando uma eficácia notável do modelo em identificar casos dessa categoria, as Classes 0 (gravidade leve) e 2 (gravidade alta) não tiveram o mesmo desempenho, ficando respectivamente com 36% e 13% de recall. Tal discrepância reflete-se também na disparidade do f1-score, que resultou em 85% para a classe de melhor performance (Classe 1) e em apenas 22% para a de pior desempenho (Classe 2). Isso sugere que o modelo tende a superestimar a classe de gravidade moderada, possivelmente devido ao desequilíbrio nos dados de treinamento, onde essa classe é preponderante.

No modelo G2, seguimos com o algoritmo *Random Forest*, mas para tentar reduzir o desequilíbrio dos resultados, realizamos o balanceamento dos dados nos conjuntos de treinamento e validação. A acurácia geral caiu para 64%, menor do que a observada no

modelo G1, mas houve evidente melhora de desempenho para as classes minoritárias, equilibrando melhor a capacidade de previsão entre as diferentes categorias de gravidade dos acidentes. Para tais classes, quais sejam, Classe 0 (gravidade leve) e Classe 2 (gravidade alta), o recall passou a ser respectivamente de 81.25% e 64.67%, refletindo em f1-scores de 60% e 43% e, consequentemente, indicando melhora do modelo ao identificar acidentes dessas categorias. Em contrapartida, o recall da Classe 1 (gravidade moderada), majoritária, reduziu-se para 58.74%, com f1-score de 71%, possivelmente devido ao balanceamento que fortaleceu o reconhecimento das outras classes.

É importante notar que apesar de a Classe 2 (gravidade alta) ser a menos frequente em nosso conjunto de dados, correspondendo a apenas 9,96% das instâncias (v. Tabela 5), sua relevância para a problemática dos acidentes de trânsito é inegável, uma vez que reflete sinistros que resultam em óbito. Trata-se, portanto, de uma classe prioritária para as análises, elevando-se a importância de se obter um modelo que tenha bom desempenho na identificação de sinistros dessa categoria.

Assim, a perda na acurácia geral (64%) do modelo G2 é um *trade-off* aceitável em troca da melhora que ele apresenta no recall médio (68%), se mostrando um classificador com desempenho mais equilibrado que o G1, o que é especialmente importante no contexto de desbalanceamento de classes existente no problema estudado. Desse modo, para os experimentos com os modelos subsequentes foi seguida a estratégia de se manter o balanceamento de dados nos conjuntos de treinamento e validação.

No modelos G3, G4 e G5, foram utilizados respectivamente os algoritmos XGBoost, SVM e KNN, utilizando-se para este último a estratégia de OneHotEncoding para codificar os dados categóricos. Todavia, o desempenho geral de todos eles foi inferior ao modelo G2, com acurácia caindo para 63% no G4 (XGBoost), 53% no G5 (SVM) e de 57% no G6 (KNN), sem que houvesse nenhuma contrapartida de aprimoramento no equilíbrio das predições para diferentes classes, visto que as métricas de recall, precision e f1-score médios não se elevaram.

No último experimento, para o modelo G6, tornamos a utilizar o *Random Forest*, algoritmo que mostrou ser o de melhor performance para o problema, e seguimos a mesma estratégia de balanceamento dos dados de treinamento e validação. Mas, além disso,

ampliamos a base de dados utilizada no estudo, passando a englobar os acidentes ocorridos de 2017 a 2023, para propiciar um treinamento mais amplo do classificador.

A acurácia geral do classificador G6 (RF) atingiu 66%, superando discretamente o resultado do modelo G2 (64%). Sua recall média de 65% representa um pequena queda nesta métrica, mas em compensação houve também leve aumento na precisão, que atingiu 56%, resultando em um f1-score médio de 59%, levemente melhor que o G2 (58%).

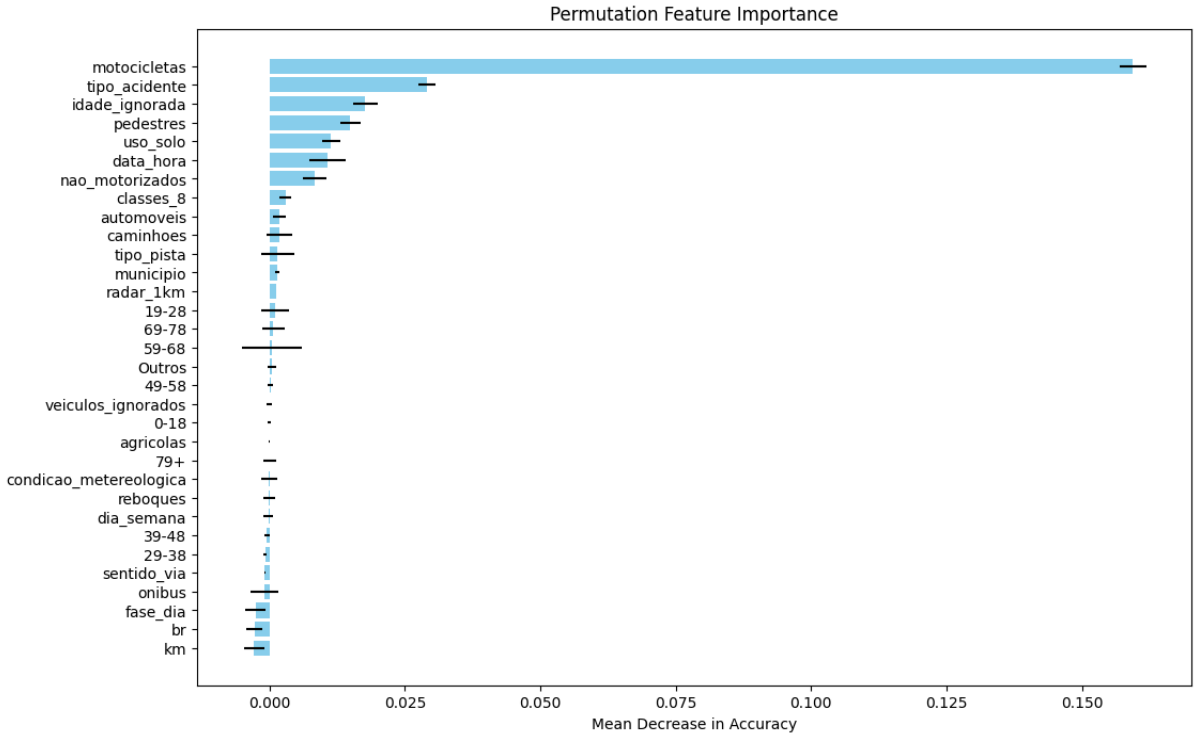
No detalhamento das métricas por classe, conforme Tabela 10, o G6 permanece com um similar equilíbrio de resultados para as distintas classes, quando comparado ao modelo G2, mas se mostra ligeiramente superior em termos de desempenho geral, como visto. Desse modo, para o estudo ora analisado reputamos o classificador G6 como o de melhor resultado.

#### *3.2.1.2. Interpretações com Métodos de XAI*

O modelo G6, um classificador RF, foi escolhido como o mais adequado com base em sua performance global e melhor equilíbrio entre as predições para diferentes classes. Nesta seção, para proporcionar uma interpretação compreensível das decisões do modelo, trabalharemos com métodos de XAI que facilitam a compreensão das relações entre variáveis e como elas afetam influenciam nas predições.

Na Figura 2, são apresentados os resultados do PFI, indicando a importância de cada atributo na determinação da categoria de gravidade dos acidentes de trânsito em rodovias federais de Pernambuco, conforme aprendido pelo modelo G6. O gráfico de barras horizontais mostra a importância relativa das variáveis, indicada pela média da diminuição da acurácia (*Mean Decrease in Accuracy - MDA*) quando essas variáveis são aleatoriamente permutadas. As barras de erro indicam a variação na importância quando o cálculo é repetido, ajudando a entender a consistência da influência de cada variável. Neste trabalho vamos analisar as quatro variáveis mais influentes no modelo.

**Figura 2** - PFI aplicado ao modelo G6 (classificador de gravidade).

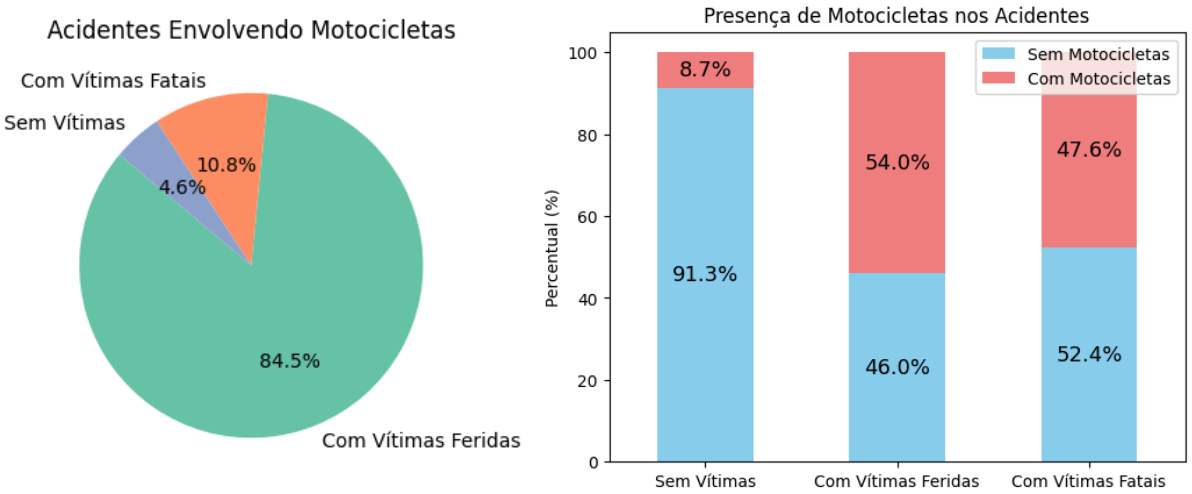



Fonte: Elaborado pela autora durante experimentos.

De acordo com o gráfico apresentado, a variável "motocicletas" tem grande destaque como a mais influente, com um alto MDA que supera 0.15, indicando que o envolvimento desse tipo de veículo nos acidentes tende a ser um forte preditor de sua gravidade. Isso pode refletir uma maior vulnerabilidade dos motociclistas em rodovias, assim como a gravidade potencial dos acidentes que as envolvem.

Diante desse achado, passamos a explorar a distribuição das ocorrências, conforme sua gravidade, para os acidentes que envolvem motocicletas, conforme mostrado na Figura 3. No gráfico à esquerda, a figura indica que 84,5% desses eventos tem gravidade moderada (com vítimas feridas), enquanto 10,8% tem gravidade alta (com vítimas fatais), e apenas a minoria de 4,6% tem gravidade leve (sem vítimas). Em um outro recorte dos dados, mostrado à direita da figura, verifica-se que de todos os acidentes com vítimas fatais, 47,6% envolvem motocicletas, enquanto nos sinistros com vítimas feridas 54% envolvem tais tipos de veículos.

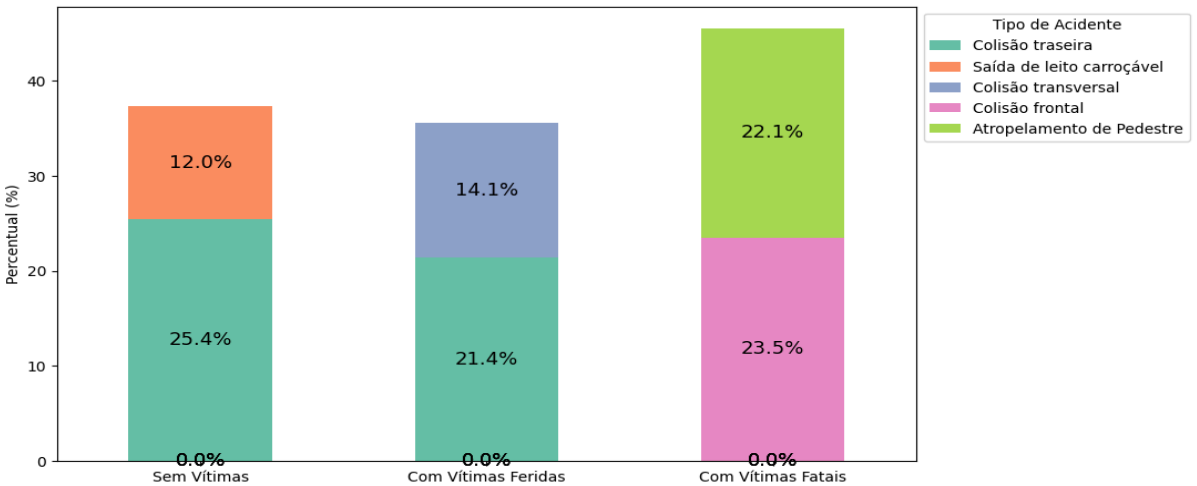
**Figura 3** - Gráficos exploratórios do atributo ‘motocicletas’ relacionado à gravidade.




Fonte: Elaborado pela autora durante experimentos.

Seguindo na análise do PFI, o segundo atributo mais influente nas predições do modelo é "tipo_acidente", porém em menor proporção que a variável ‘motocicletas’. Esse resultado sugere que as distintas características de cada tipo de acidente afetam sua gravidade. Na Figura 4, é possível perceber algumas dessas distinções, ao observarmos os dois tipos de acidentes mais frequentes em cada categoria de gravidade do sinistro.

**Figura 4** - Gráfico dos principais tipos de acidente para cada categoria de gravidade.




Fonte: Elaborado pela autora durante experimentos.

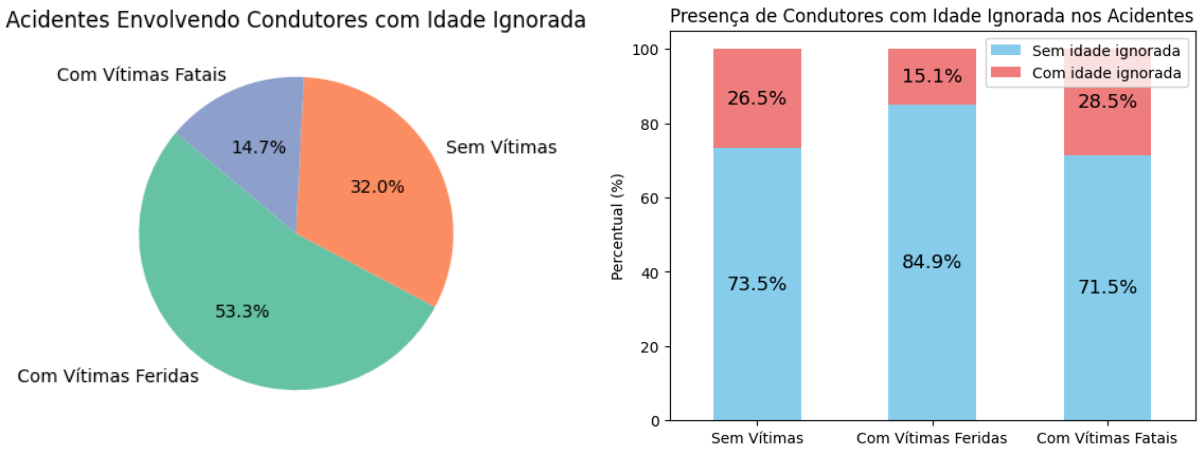
Do gráfico, constata-se que nos sinistros de gravidade alta (com vítimas fatais) os tipos de acidentes mais frequentes são a colisão frontal, respondendo por 23,5% das ocorrências dessa categoria, e o atropelamento de pedestres, que corresponde a 22,1%. Diferentemente do que ocorre nas ocorrências de gravidade leve (sem vítimas), onde os principais tipos de acidentes são a colisão traseira (25,4%) e a saída de leito carroçável (12%).

Já nas ocorrências com gravidade moderada (com vítimas feridas) o principal tipo de acidente também é a colisão traseira (21,4%), mas em seguida, distintamente das demais categorias, aparece a colisão transversal (14%).

A terceira variável mais influente mostrada pelo PFI é a "idade_ignorada" (v. Figura 2), atributo que indica a quantidade de condutores envolvidos no acidente, cuja idade não consta do registro da ocorrência. A falta de documentação da idade do condutor indica uma lacuna nas informações disponíveis nos laudos de acidente, cujo conhecimento seria relevante para o aprendizado do classificador desenvolvido, de modo a propiciar predições mais precisas.

O PFI mostra, contudo, que o modelo está associando a presença ou não de tal lacuna nos dados à maior ou menor gravidade do sinistro, mas é possível também que os resultados não sejam uma decorrência direta da variável, e sim de uma possível interação com outro atributo. Na Figura 5 (à direita), demonstra-se que, nos acidentes de gravidade alta (com vítimas fatais), 28,5% dos registros contêm um ou mais condutores envolvidos cuja idade é desconhecida, similarmente ao que ocorre nos de gravidade leve (sem vítimas), cujo percentual é de 26,5%. Já nos de gravidade moderada esse índice cai para 15,1%.

**Figura 5** - Gráficos exploratórios do atributo 'idade_ignorada' relacionado à gravidade.



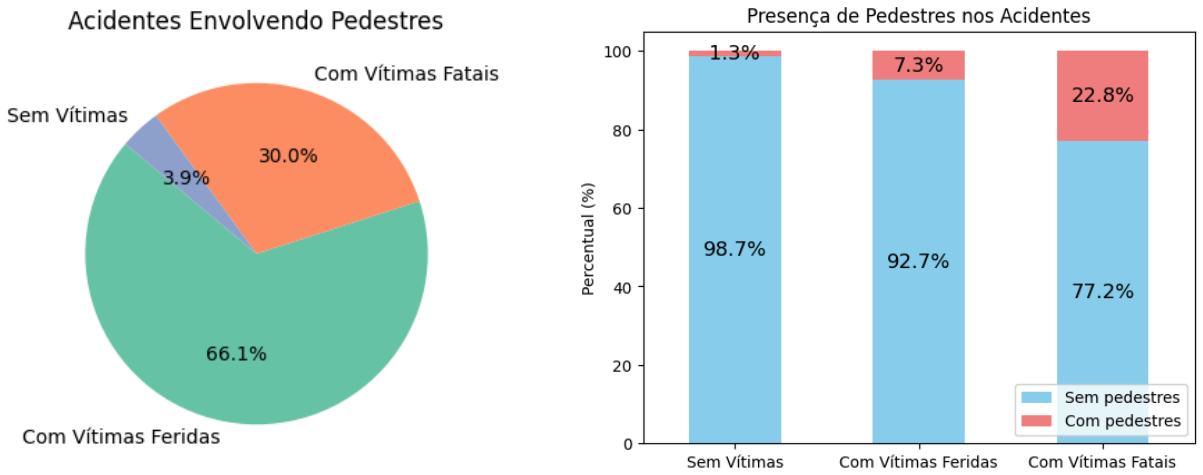

Fonte: Elaborado pela autora durante experimentos.

Constata-se ainda (à esquerda da Figura 5) que, dos registros em que há lacuna quanto à idade de condutor envolvido, a grande maioria (53.3%) é de gravidade moderada (com vítimas feridas), seguida pela categoria de gravidade leve (32%) e por fim a de gravidade alta (14,7%).

A variável 'pedestres', identificada como a quarta mais influente na análise de PFI, captura a quantidade de transeuntes envolvidos no acidente, tendo o modelo identificado haver relação desse atributo com a gravidade do sinistro. Intuitivamente, imagina-se que acidentes envolvendo pedestres têm maior probabilidade de resultar em vítimas feridas ou fatais devido à vulnerabilidade física dos pedestres comparada a ocupantes de veículos motorizados.

Na Figura 6 (à esquerda), demonstra-se que, de fato, a grande maioria dos acidentes que envolvem pedestres têm gravidade moderada (66,1%) ou alta (30%). Ademais, de todas as ocorrências com vítimas fatais, 22,8% envolvem pedestres (Figura 6, à direita).

**Figura 6** - Gráficos exploratórios do atributo ‘pedestres’ relacionado à gravidade.

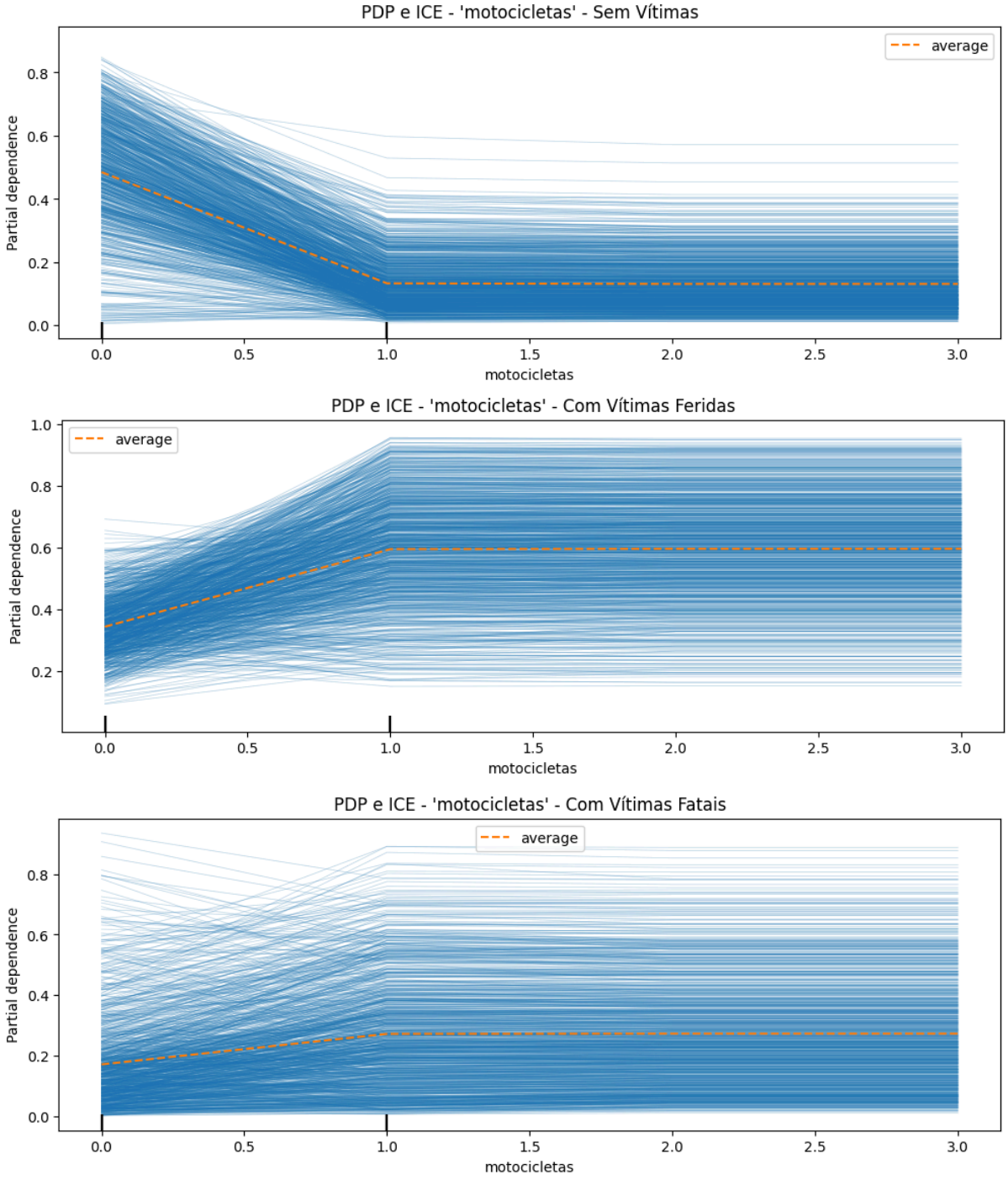



Fonte: Elaborado pela autora durante experimentos.

Aprofundando-se as análises de interpretabilidade mediante métodos de XAI, com foco nesses quatro atributos identificados como mais influentes pelo PFI - *'motocicletas', 'tipo_acidente', 'idade_ignorada' e 'pedestres'* - veremos a seguir os resultados obtidos com as técnicas de PDP e ICE.

Nos gráficos a serem apresentados a seguir, a linha pontilhada laranja representa o PDP, indicando o efeito médio da variável na previsão do modelo ao longo de um intervalo de valores. Já as linhas azuis representam o ICE, mostrando as trajetórias de previsão para as instâncias individuais.

**Figura 7** - PDP e ICE para a variável ‘motocicletas’ (classificador de gravidade).




Fonte: Elaborado pela autora durante experimentos.

Conforme Figura 7.1 (gráfico superior), no caso do atributo ‘motocicletas’, o PDP mostra uma tendência de decréscimo na previsão para a classe de gravidade baixa (sem vítimas) quando há uma motocicleta envolvida no acidente, estabilizando-se para quantitativos maiores. Isso sugere que, segundo o modelo, a presença de uma motocicleta no acidente leva a uma redução na probabilidade de predição para essa classe (sem vítimas), mas que aumentos adicionais no número de motocicletas não alteram significativamente a

probabilidade. Isso pode ser interpretado como um indicativo de que o risco inicial associado à presença de motocicletas é o fator crítico, com poucas variações adicionais na gravidade uma vez que esse limiar é ultrapassado.

Ainda na Figura 7.1, observa-se que o comportamento para as instâncias individuais mostrado pelo ICE indica um padrão similar à média na maioria dos casos, mas para ocorrências que já apresentam uma baixa probabilidade de predição para essa classe, mesmo quando não há motocicletas envolvidas (por influência de outros fatores), a amplitude do decréscimo tende a ser reduzida. Isso também se observa para pouquíssimas instâncias que já iniciam com uma alta probabilidade de classificação nessa categoria.

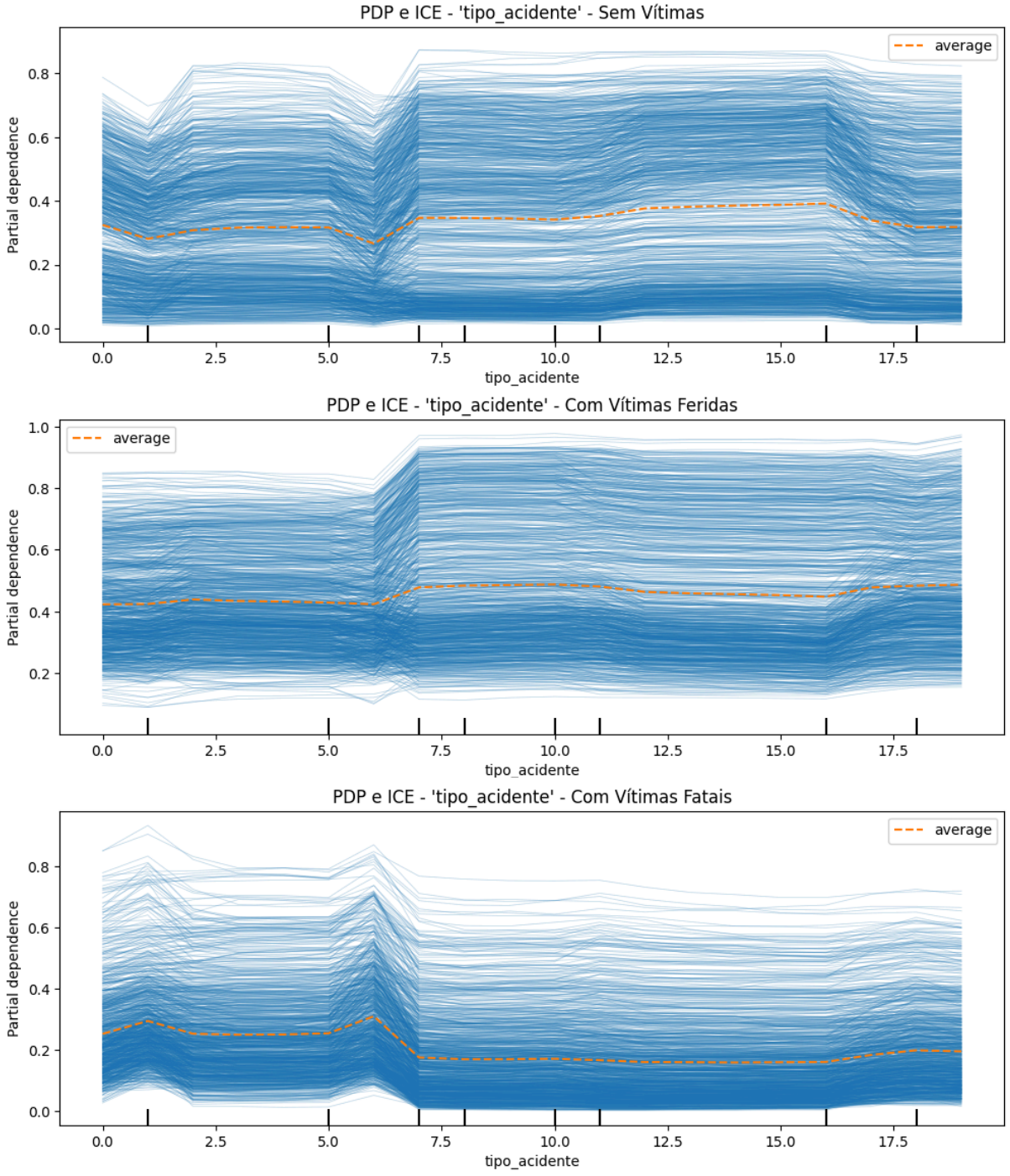
Nas Figuras 7.2 (gráfico central) e 7.3 (gráfico inferior), demonstra-se que para as classes de gravidade moderada com vítimas feridas) e de gravidade alta (vítimas fatais) a tendência do PDP inverte-se: há crescimento na probabilidade de predição quando há uma motocicleta envolvida, também estabilizando-se para quantitativos maiores desse veículo. As curvas nesse caso, contudo, são mais sutis, principalmente para a classe de gravidade alta (vítimas fatais), onde a variação do PDP é bastante discreta. Isso indica que, para essas classes, a presença de motocicletas ainda se mostra como um fator crítico para a predição, mas o impacto da variável ‘motocicletas’ é mitigado por outros atributos de influência nos acidentes.

Quanto ao ICE para essas categorias (Figuras 7.2 e 7.3), também segue-se majoritariamente padrão de comportamento da média do PDP, porém há mais variabilidade, inclusive com algumas instâncias individuais mostrando tendência inversa, de decréscimo da probabilidade de predição. Nesses casos, outros fatores, não capturados por este atributo, são mais decisivos na determinação da gravidade dos acidentes.

Em relação ao atributo ‘tipo_acidente’, o PDP mostra na Figura 8.1 (gráfico superior) que para as predições da classe de gravidade baixa (sem vítimas), há queda na probabilidade principalmente para os acidentes do tipo 6 (colisão frontal), mas com pequena baixa também para os tipos 1 (atropelamento de pedestre) e 17 (queda de ocupante de veículo)[9]. Comportamento oposto ocorre para as predições da classe de gravidade alta (com vítimas fatais, conforme Figura 8.3 (gráfico inferior).

[9] A identificação categórica correspondente a cada código da variável ‘tipo_acidente’ pode ser observada na Figura 9, na qual é apresentado um gráfico de distribuição dos sinistros conforme tal atributo.

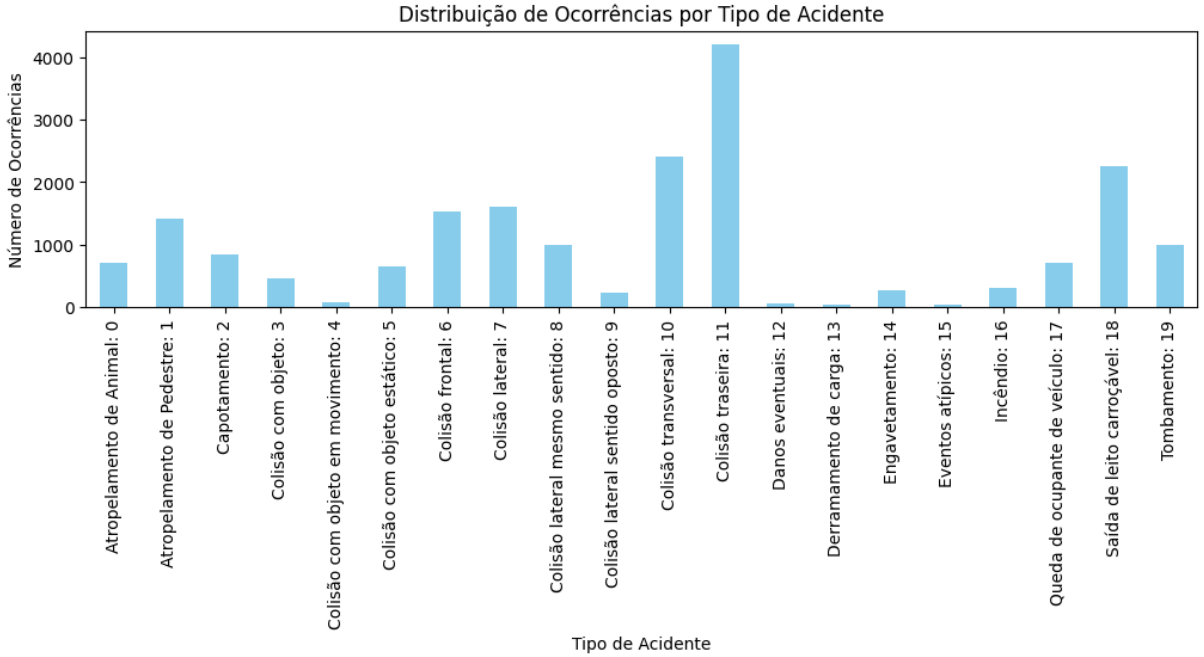
**Figura 8** - PDP e ICE para a variável ‘tipo_acidente’ (classificador de gravidade).




Fonte: Elaborado pela autora durante experimentos.

Isso sugere que esses tipos de acidentes - *colisão frontal, atropelamento de pedestre e queda de ocupante de veículo* -, segundo aprendido pelo modelo, costumam ter mais chances de resultarem em vítimas fatais. acidentes de colisão frontal estão menos associados a situações onde não há vítimas.

**Figura 9** - Distribuição de acidentes conforme ‘tipo_acidente’ (com nome e código).




Fonte: Elaborado pela autora durante experimentos.

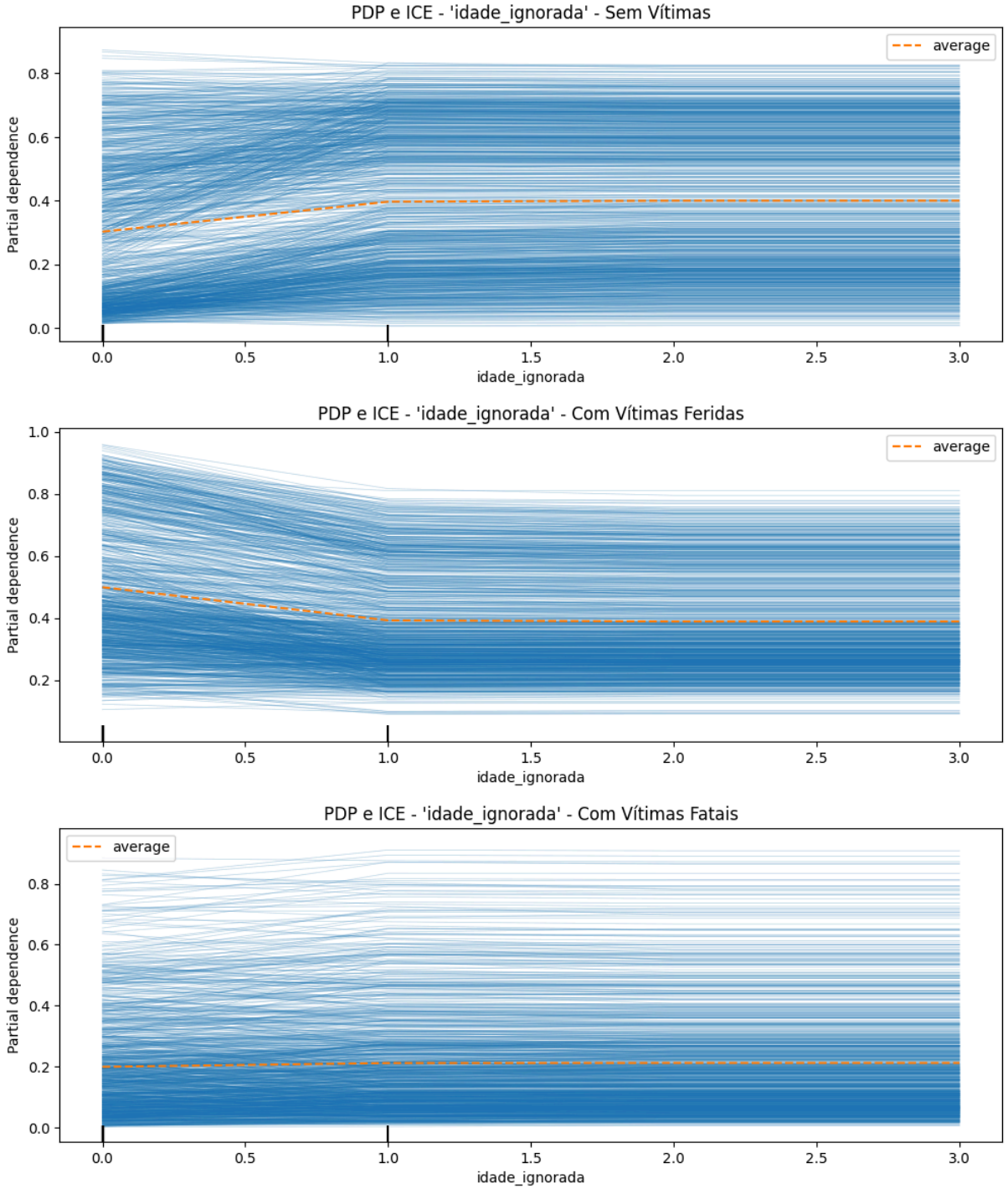
Quanto à classe de gravidade moderada (com vítimas feridas), o PDP mostra um leve incremento na probabilidade de predição para o tipo de acidente de valor 7 (colisão lateral), como visto na Figura 8.2 (gráfico central), indicando maior associação desses sinistros com a ocorrência de ferimentos nas pessoas envolvidas. Quanto ao ICE, de um modo geral para as três classes preditivas, observa-se que as instâncias individuais tendem a seguir padrão similar à média.

O terceiro atributo em ordem de importância apontada pelo PFI é a variável ‘idade_ignorada’. Vale ressaltar que esse atributo resulta de lacuna de dados nos registros dos sinistros, a qual mostrou-se de relevância para o aprendizado do modelo, indicando que eventual saneamento desse aspecto na elaboração dos laudos de acidentes poderia refletir em melhores resultados do classificador.

De qualquer modo, as Figuras 10.1 (gráfico superior) e 10.2 (gráfico central) mostram que o PDP do atributo ‘idade_ignorada’ indica elevação na probabilidade de predição para a classe de gravidade leve (sem vítimas) e redução para a classe de gravidade moderada (com vítimas feridas), quando o valor do atributo se eleva para 1, sendo que, após isso, aumentos subsequentes são indiferentes. A variação do PDP, contudo, é discreta, indicando que o atributo tem impacto sutil nas predições do modelo. Já nas predições para a classe de gravidade alta (com vítimas fatais), o PDP visto na Figura 10.3 (gráfico inferior) é

praticamente plano, sinalizando que para essa categoria a variável tem pouca relevância na média, não sendo um fator crítico para a predição.

**Figura 10** - PDP e ICE para a variável ‘idade_ignorada’ (classificador de gravidade).

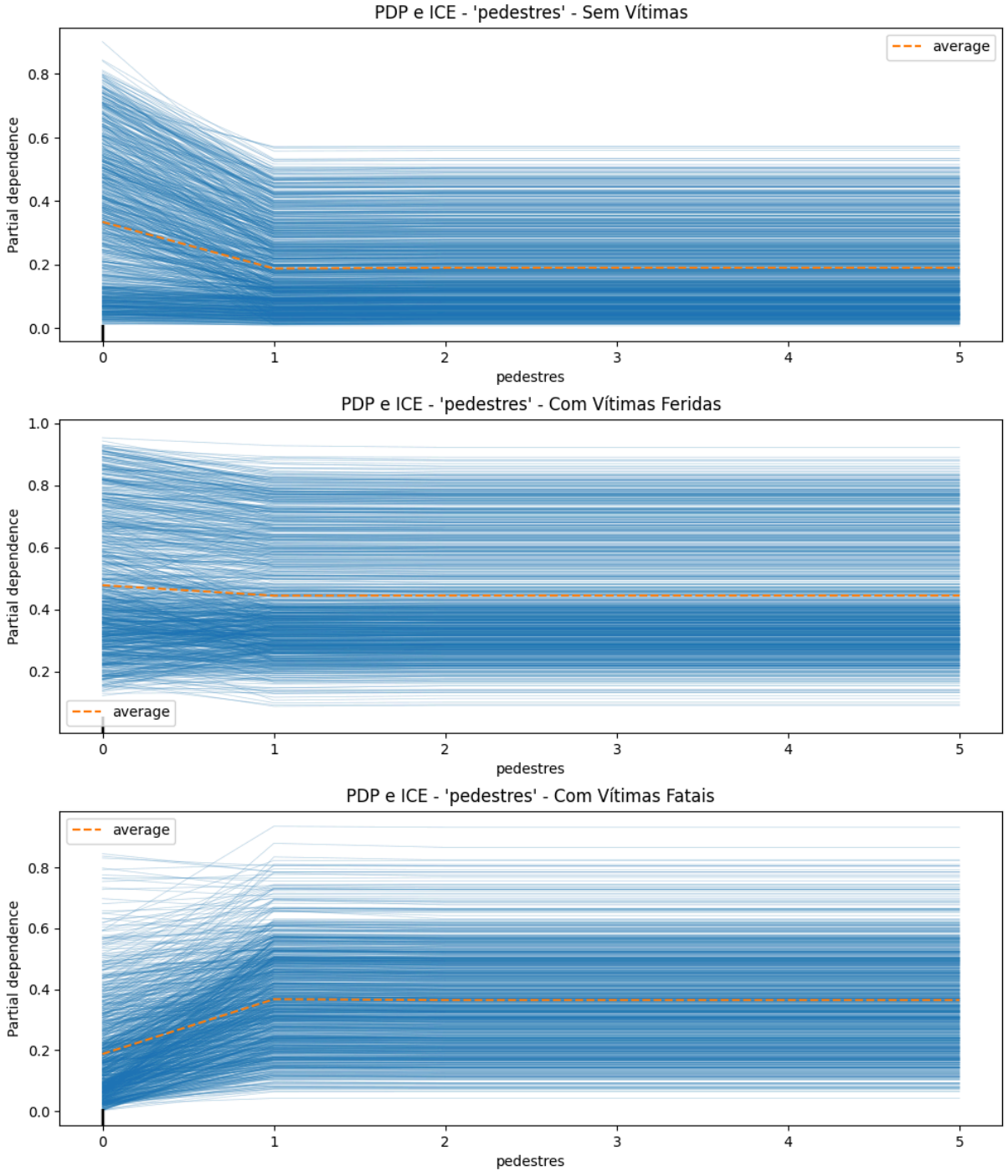



Fonte: Elaborado pela autora durante experimentos.

Quanto ao ICE, demonstra que as predição para as instâncias individuais são mais uniformes quanto à classe de gravidade moderada, havendo maior variabilidade, e portanto maior influência de outros fatores, nas probabilidades das classes de gravidade leve e alta.

A última variável objeto de nossa interpretação é o atributo ‘pedestres’. As Figuras 11.1 (gráfico superior) e 11.2 (gráfico inferior) mostram, respectivamente, o PDP e ICE para as predições das classes de gravidade leve (sem vítimas) e gravidade alta (com vítimas fatais), indicando que essa *feature* tem impacto inverso para as classificações em cada uma dessas categorias, em proporção bastante similar.

**Figura 11** - PDP e ICE para a variável ‘pedestres’ (classificador de gravidade).




Fonte: Elaborado pela autora durante experimentos.

Com efeito, os gráficos mostram que o envolvimento de um pedestre reduz a probabilidade de que o acidente seja “sem vítimas”, ao mesmo tempo em que aumenta as chances do sinistro resultar em letalidade. Já para a classe de gravidade moderada (com vítimas feridas) o gráfico é praticamente plano, apresentando apenas um decréscimo ínfimo da probabilidade nessas condições.

Quanto ao ICE, mostra-se majoritariamente uniforme para a classe de baixa gravidade, mas para a categoria de alta gravidade e, principalmente, para a de gravidade moderada, há notável presença de tendências distintas para parte das instâncias individuais. Isso indica que, para estes casos, há maior interferência de outros fatores nas predições.

### *3.2.2. Predição da Causa do Acidente*

Nesta seção, trataremos dos modelos de classificação desenvolvidos para prever a causa dos acidentes nas rodovias federais de Pernambuco, os quais tomam como *target* o atributo binário ‘classes_condutor’. Como visto na modelagem, a classe 0 indica que o sinistro foi causado por um ‘condutor’ envolvido, enquanto a classe 1 aponta que a causa se atribui a ‘outros’ fatores.

Neste caso, o algoritmo utilizado para o treinamento dos classificados foi o *Random Forest*, opção adotada diante dos resultados obtidos nos experimentos relativos à predição da gravidade do acidente, que demonstraram ser essa a técnica de aprendizado de máquina de melhor performance para nosso conjunto de dados, frente aos demais algoritmos empregados (*XGBoost, SVM e KNN*).

Também foram utilizadas diferentes abordagens, ora fazendo-se uso do balanceamento dos dados de treinamento e validação, ora ampliando-se o conjunto de dados para auferir um treinamento mais amplo do modelo.

#### *3.2.2.1. Resultados dos Modelos Desenvolvidos*

Os resultados são apresentados em tabelas que mostram as métricas de desempenho para cada modelo testado, os quais denominados de C1 a C3. Na Tabela 11, indica-se o algoritmos de treinamento utilizado, qual intervalo de ocorrências o conjunto de dados engloba, se houve prévio balanceamento de dados (treinamento e validação), e os resultados obtidos para as métricas de acurácia geral e *precision*, *recall* e *f1-score* médios.

Na Tabela 12, são detalhadas para cada modelo as métricas de *precision*, *recall* e *f1-score* obtidas para cada uma das 3 classes de gravidade existentes no conjunto de dados.

**Tabela 11** - Modelos de predição da causa dos acidentes (resultados gerais).

| Modelo | Algoritmo | Dados | Balancea-mento | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|---|---|---|
| C1 | Random Forest | 2019 a 2023 | Não | **0.82** | **0.80** | 0.71 | 0.74 |
| C2 | Random Forest | 2019 a 2023 | Sim | 0.80 | 0.74 | 0.77 | 0.75 |
| C3 | Random Forest | 2017 a 2023 | Sim | 0.81 | 0.75 | **0.77** | **0.76** |

Fonte: Elaborada pela autora, com base nos experimentos.

**Tabela 12** - Modelos de predição da causa dos acidentes (resultados por classe).

| Modelo | Algoritmo | Class 0 Precision | Class 1 Precision | Class 0 Recall | Class 1 Recall | Class 0 F1-Score | Class 1 F1-Score |
|---|---|---|---|---|---|---|---|
| C1 | Random Forest | 0.84 | **0.77** | **0.95** | 0.48 | **0.89** | 0.59 |
| C2 | Random Forest | **0.89** | 0.59 | 0.82 | **0.72** | 0.86 | **0.65** |
| C3 | Random Forest | 0.88 | 0.61 | 0.85 | 0.69 | 0.87 | **0.65** |

Fonte: Elaborada pela autora, com base nos experimentos.

O modelo C1, treinado a partir de um conjunto de dados compreendendo acidentes de 2019 a 2023, sem que houvesse o balanceamento entre as classes, atingiu a melhor acurácia observada, com um valor 82%. Contudo, ao avaliarmos os resultados por classe, chama atenção o baixo *recall* apresentado para a classe 1 (minoritária), que atingiu 48%, indicando que o classificador não tem bom desempenho para identificar casos reais dessa categoria. Essa performance também puxa para baixo, consequentemente, o *f1-score* da classe, que foi 59%, apesar de o valor de *precision* ter sido razoável (77%).

Para os modelos seguintes, passamos a realizar o balanceamento de dados nos conjuntos de treinamento e validação, de modo a aprimorar o equilíbrio no desempenho entre as classes. No classificador C2 utilizamos a base de dados correspondente aos acidentes registrados entre 2019 e 2023, enquanto que no C3 esse conjunto foi ampliado, englobando os anos de 2017 a 2023.

Em ambos os modelos houve um *tradoff* em relação aos resultados do C1, com ligeira baixa na acurácia geral, mas melhor equilíbrio nos resultados entre as classes. O modelo C2 atingiu 80% de acurácia, enquanto o C3 marcou 81% para essa métrica, se mostrando ligeiramente mais eficaz.

Quanto aos indicadores por classe, houve significativa melhora no *recall* da classe minoritária (classe 1), que chegou a 72% e 69%, respectivamente, para os modelos C2 e C3. Houve também melhora no *f1-score* dessa categoria, que passou para 65% em ambos os modelos, apesar de ter havido piora nos valores de *precision* (atingiu 59% para C2 e 61% para C3). Portanto, esses modelos demonstram ter atingido uma melhor sensibilidade para a classe minoritária, capturando uma proporção maior de casos positivos e conseguindo melhor equilíbrio em relação à performance dessa classe.

Assim, os modelos C2 e C3 apresentaram desempenhos bastante próximos em termos de equilíbrio entre as classes, contudo C3 sobreleva-se sutilmente nos resultados globais, pois sua acurácia (81%), *precision* médio (75%) e *f1-score* médio (76%) superam C2 em 1 ponto percentual.

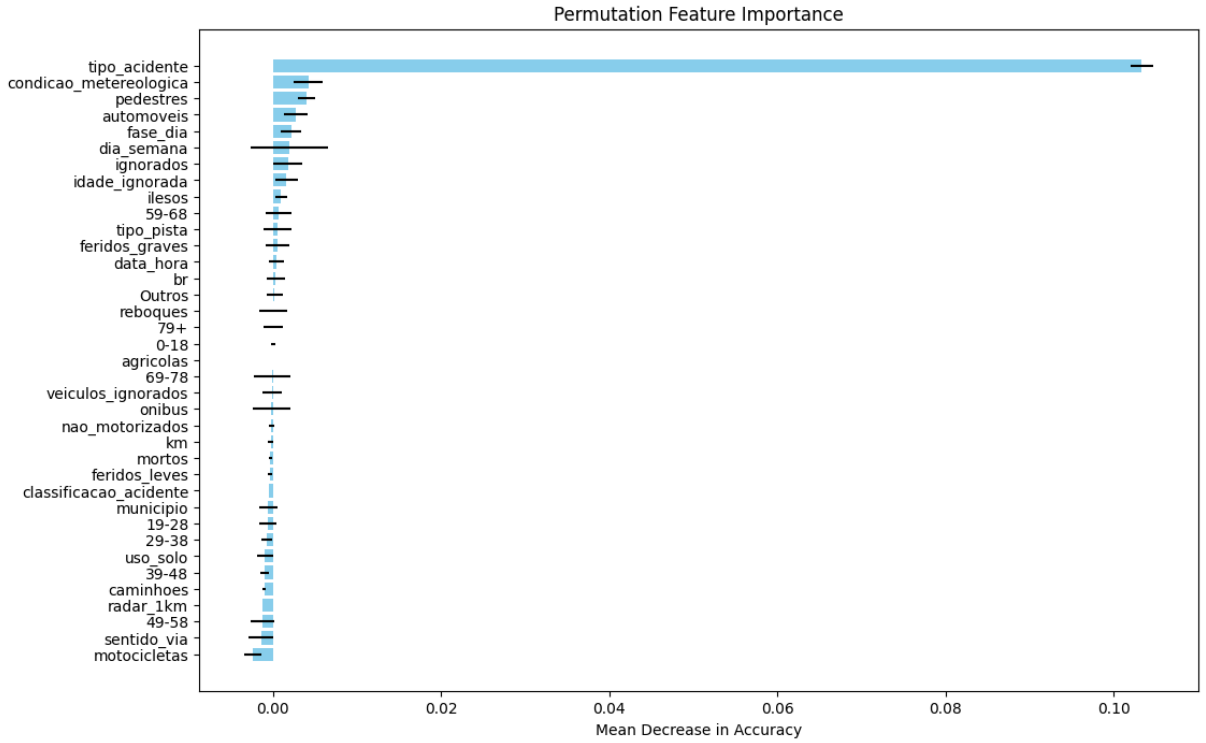
Desse modo, após nossas análises, consideramos que o modelo C3 apresentou a melhor performance para a tarefa de predição da causa dos acidentes frente ao conjunto de dados utilizado neste estudo, com base em sua performance global e melhor equilíbrio entre as predições para diferentes classes.

#### *3.2.2.2. Interpretações com Métodos de XAI*

Nesta seção, nosso modelo de melhor desempenho (classificador C3, do tipo *Random Forest*, conforme Tabelas 11 e 12) será explorado mediante técnicas de XAI, de modo a interpretarmos quais fatores mais influenciam em suas predições.

Na Figura 12, são apresentados os resultados do PFI, indicando a importância de cada atributo na determinação da causa (*condutores* ou *outros*) dos acidentes de trânsito em rodovias federais de Pernambuco, conforme aprendido pelo modelo C3. O gráfico de barras horizontais mostra a importância relativa das variáveis, indicada pela média da diminuição da acurácia (*Mean Decrease in Accuracy*) quando essas variáveis são aleatoriamente permutadas. As barras de erro indicam a variação na importância quando o cálculo é repetido, ajudando a entender a consistência da influência de cada variável. Neste trabalho vamos analisar as quatro variáveis mais influentes no modelo.

**Figura 12** - PFI aplicado ao modelo C3 (classificador de causa).

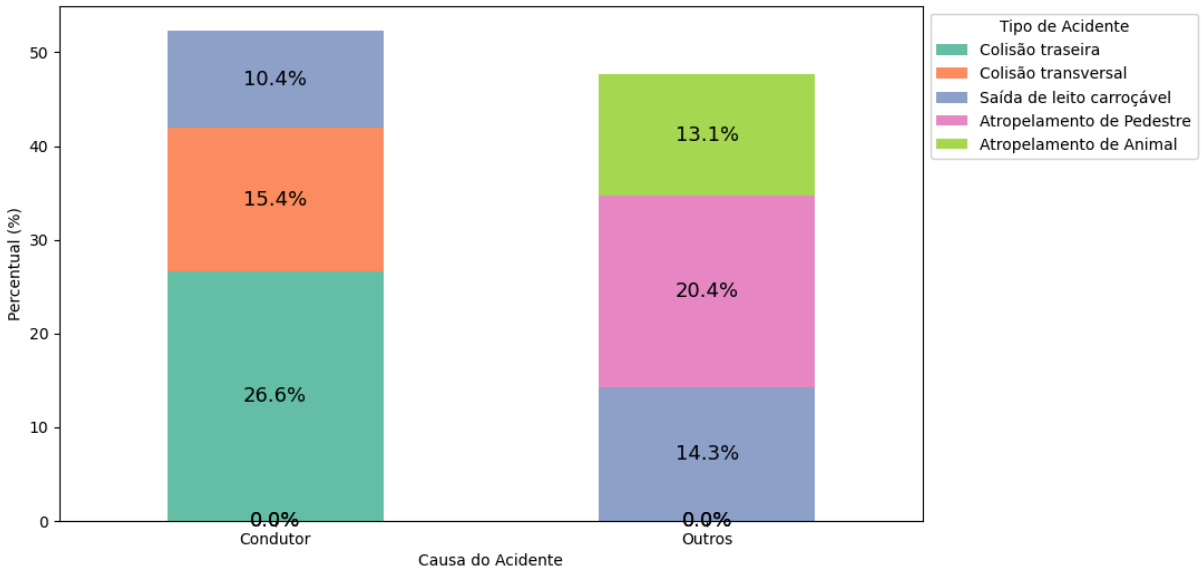



Fonte: Elaborado pela autora durante experimentos.

Similarmente ao que ocorreu com o modelo preditor da gravidade dos acidentes, este classificador da causa dos sinistros também apresenta um atributo cuja influência sobre a predição se sobressai a todos os demais. Neste caso, com um MDA superior a 0.10, trata-se da variável ‘tipo_acidente’, sinalizando que distintos tipos de sinistros têm maior associação a alguma das causas de acidentes passíveis de classificação.

Essa indicação do PFI se alinha a observações que podem ser obtidas da exploração dos dados, uma vez que algumas espécies de sinistros se mostram mais presentes de acordo com a causa da ocorrência. Com efeito, na Figura 13, observamos que os acidentes causados pelo condutor são em sua maioria do tipo ‘colisão traseira’ (26,6%), seguido por ‘colisão transversal’ (15,4%) e ‘saída de leito carroçável’ (10,4%). Já os sinistros provocados por outros fatores tem sua maior parcela nos ‘atropelamentos de pedestres’ (20,4%), seguido por ‘saída de leito carroçável’ (14,3%) e ‘atropelamento de animal’ (13,1%).

**Figura 13** - Gráfico dos principais tipos de acidente para cada categoria de causa do sinistro.



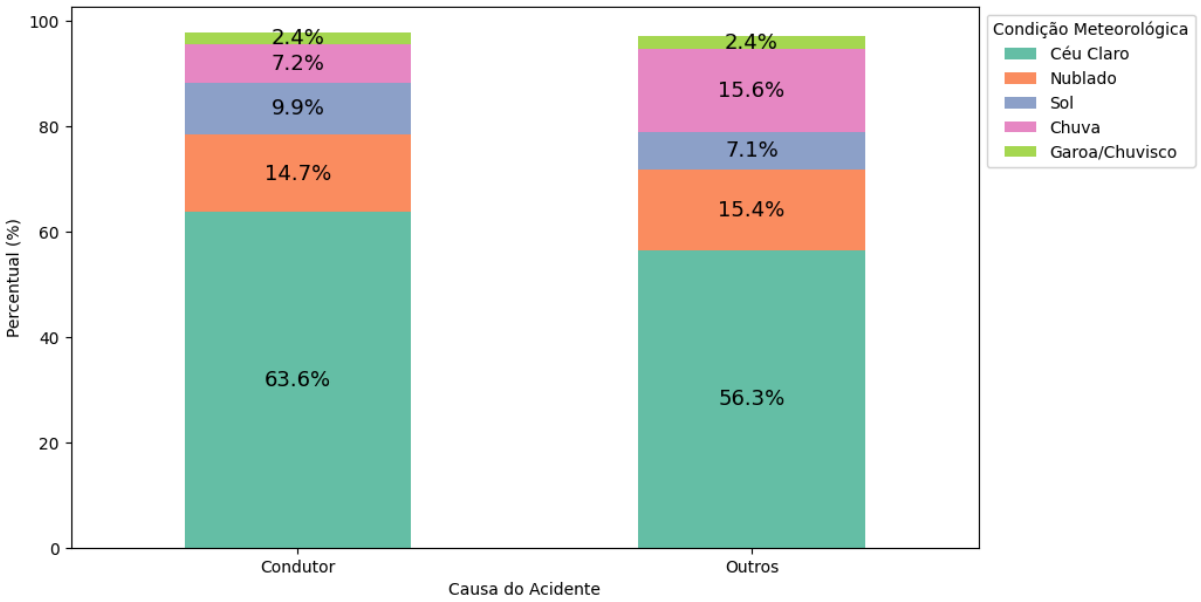

Fonte: Elaborado pela autora durante experimentos.

Os próximos atributos de maior influência mostrados pelo PFI (Figura 12), em ordem, são ‘condição_metereologica’[10], ‘pedestres’ e ‘automoveis’. Todavia, seu impacto se mostra bem mais discreto, uma vez que o resultado da permutação refletiu em valores de MDA inferiores a 0.01.

A variável ‘condicao_metereologica’ traz uma representação das condições climáticas durante os acidentes, sendo possível inferir do PFI que algumas condições meteorológicas podem estar mais atreladas a uma determinada causa do acidente. Ao observarmos as cinco condições meteorológicas mais frequentes em cada classe de acidentes, conforme Figura 14, podemos destacar que, embora a maioria dos eventos de ambas as categorias causais tenha se passado sob “céu claro” ou com tempo “nublado”, no caso de acidentes ocasionados por “outros fatores” o clima de “chuva” (15,6%) se fez mais presente do que naqueles provocados pelo “condutor” (7,2%).

[10] A grafia equivocada “metereologica” já consta do nome original contido nos *datasets* disponibilizados pela Polícia Rodoviária Federal. Portanto, mantivemos o nome do atributo exatamente como consta nos *datasets*.

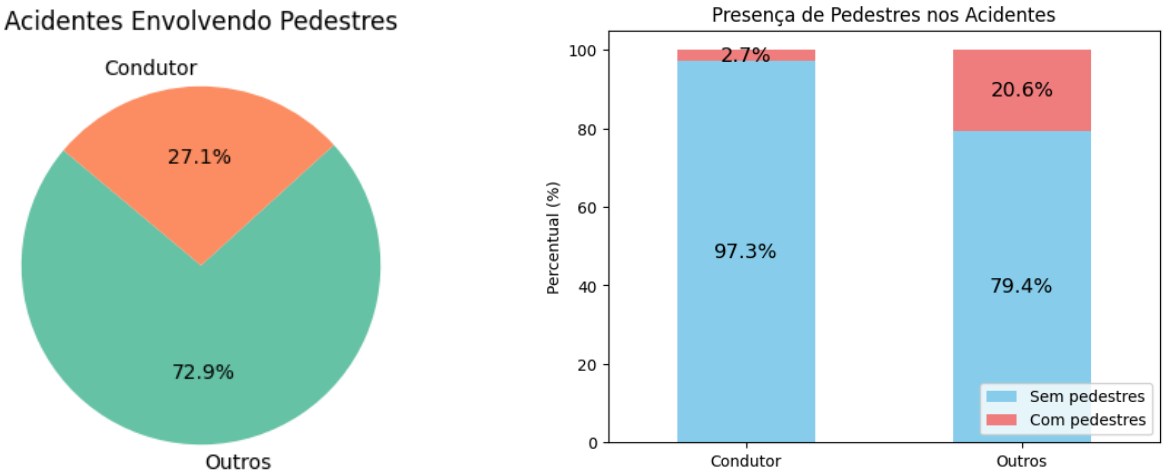
**Figura 14** - Gráfico das principais condições meteorológicas para cada causa do sinistro.




Fonte: Elaborado pela autora durante experimentos.

Quanto ao atributo 'pedestres', que indica a quantidade de transeuntes envolvidos no acidente, a Figura 15 mostra que a grande maioria dos acidentes que envolvem pedestres são causados por ‘outros fatores’ (72,9%). Além disso, de todos os acidentes causados pelo ‘condutor’, apenas 2,7% envolvem pedestres, enquanto que 20,6% daqueles ocasionados por outras circunstâncias.

**Figura 15** - Gráficos exploratórios do atributo ‘pedestres’ relacionado à causa do acidente.

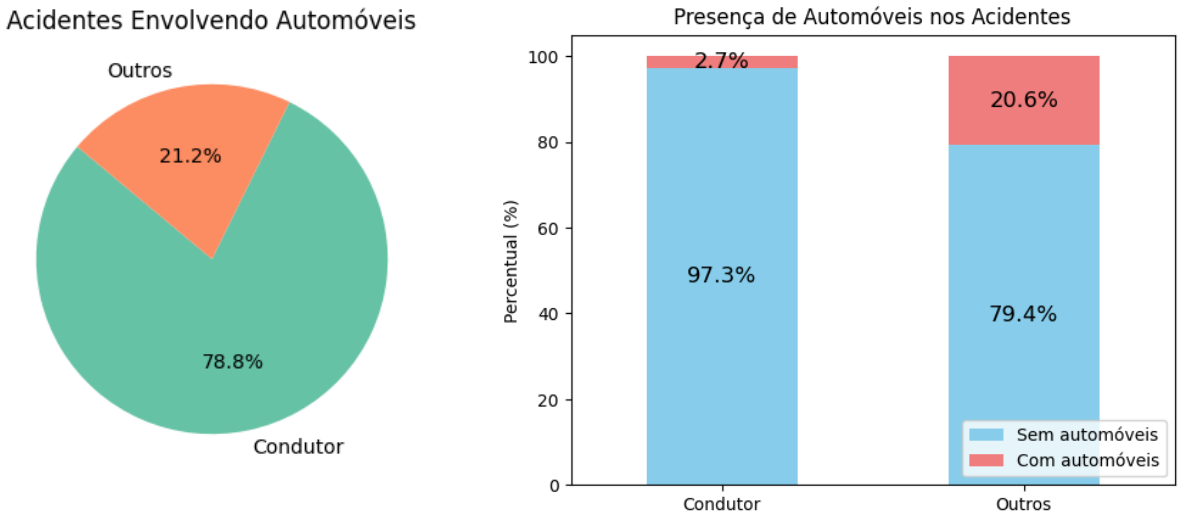



Fonte: Elaborado pela autora durante experimentos.

Já a variável ‘automóveis’, quarta mais influente, expressa a quantidade de veículos dessa categoria envolvidos no acidente, indicando que se trata de um atributo com impacto nas predições. A Figura 16 mostra que 78.8% dos acidentes em que há envolvimento de

automóveis têm causa classificada como ‘condutor’, sinalizando que a presença desse tipo de veículo estaria mais associada a essa causalidade. Mas, apesar disso, tais sinistros correspondem a apenas 2.7% de todos os acidentes pertencentes à classe ‘condutor’, enquanto que no caso da classe ‘outros’, ou seja, para acidentes provocados por outros fatores, em 20,6% das ocorrências há o envolvimento de automóveis. Diante desse cenário de leitura menos clara, a aplicação do PDP, que será realizada mais à frente, trará um melhor entendimento de como essa variável impacta na predição do modelo.

**Figura 16** - Gráficos exploratórios do atributo ‘automóveis’ relacionado à causa do acidente.




Fonte: Elaborado pela autora durante experimentos.

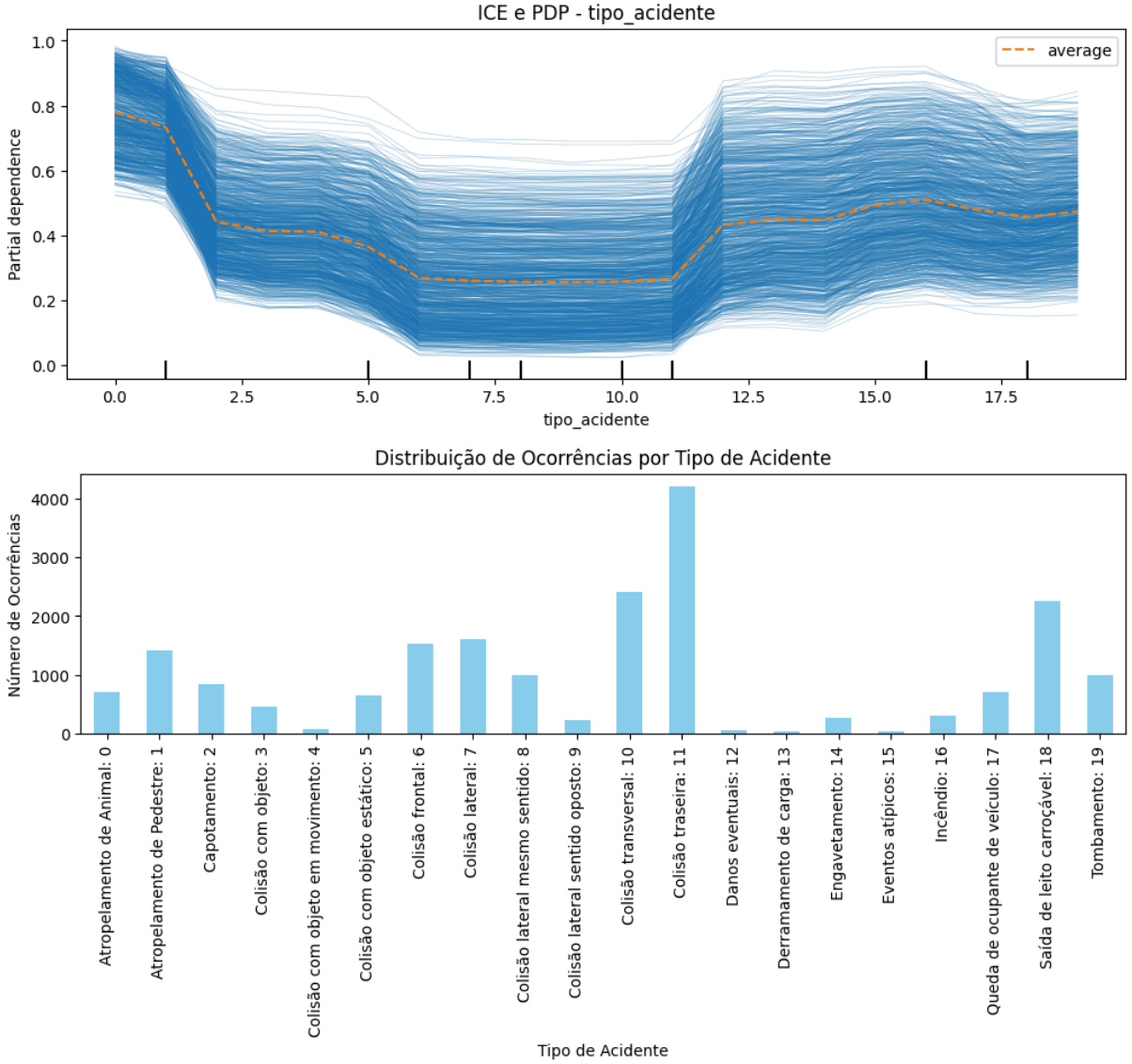
Seguindo-se com a aplicação de métodos de XAI, passamos ao emprego das técnicas de PDP e ICE. Relembramos que, para o entendimento dos gráficos correspondentes, deve-se considerar que a linha pontilhada laranja indica o efeito médio da variável na previsão do modelo, representando o PDP, enquanto as linhas azuis mostram as trajetórias de previsão para cada instância individual, representando o ICE.

No presente caso, como o *target* da predição é binário, os resultados sinalizam a probabilidade atrelada à classe positiva, ou seja, àquela que corresponde ao valor ‘1’. No conjunto de dados sob estudo, essa classificação indica que a causa do acidente se atribui a ‘outros fatores’, não sendo causados, portanto, pelo ‘condutor’. Assim, um aumento no valor do PDP indica um aumento na probabilidade de a instância pertencer à classe 1 (‘outros’), e consequentemente uma redução, na mesma proporção, de pertencer à classe 0 (‘condutor’).

Na Figura 17 (gráfico superior), vemos os resultados obtidos para a variável ‘tipo_acidente’, que apresenta picos para os tipos 0 (atropelamento de animal) e 1

(atropelamento de pedestre)[11], indicando maior probabilidade de serem causados por ‘outros fatores’ (classe 1). Já as maiores quedas no PDP abarcam os tipos 6 a 11 (colisões frontal, laterais, transversal e traseira), os quais estariam, dessa forma, mais tendentes a serem classificados sob a causa ‘condutor’ (classe 0). Também nota-se a partir do ICE que as instâncias individuais apresentam grande uniformidade nas predições, seguindo em grande parte o comportamento observado na média.

**Figura 17** - PDP e ICE para a variável ‘tipo_acidente’ (classificador de causa).



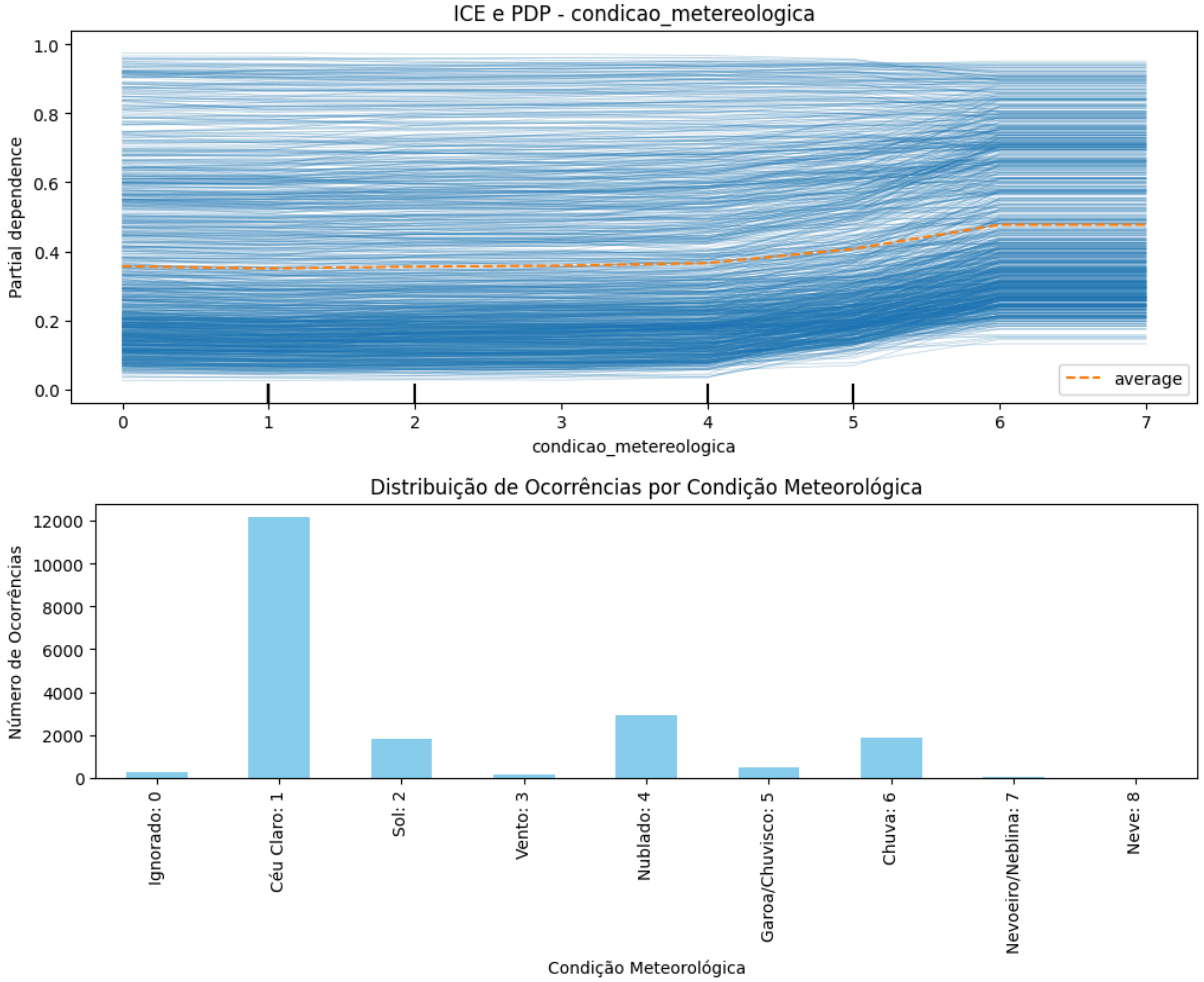

Fonte: Elaborado pela autora durante experimentos.

Quanto à variável ‘condicao_metereologica’, a Figura 18 (gráfico superior) mostra constância do PDP para os tipos 0 a 4 (ignorado, céu claro, sol, vento e nublado)[12], e uma

[11] Vide Figura 17, gráfico inferior, para identificar os códigos da variável ‘tipo_acidente’.
[12] Vide Figura 18, gráfico inferior, para identificar os códigos da variável ‘condicao_metereologica’.

discreta elevação para os tipos 5 a 7 (garoa/chuvisco, chuva e nevoeiro/neblina). Estes últimos apresentam então probabilidade ligeiramente maior de pertencerem à classe ‘outros’ e, consequentemente, de não serem provocados pelo condutor. O ICE novamente se mostra majoritariamente uniforme, mas com uma flutuação mais presente para instâncias individuais entre os valores 5 e 6 da variável (garoa/chuvisco, chuva), indicando que, nesses casos, outros fatores podem estar impactando nas predições.

**Figura 18** - PDP e ICE para a variável ‘condicao_metereologica’ (classificador de causa).

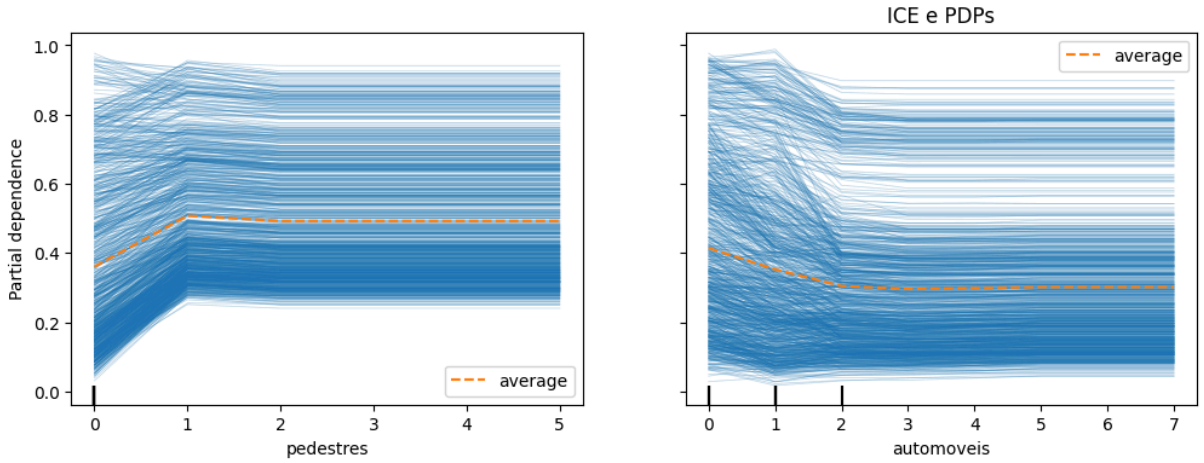



Fonte: Elaborado pela autora durante experimentos.

Na Figura 19, constam os gráficos PDP e ICE para os atributos ‘pedestres’ (à esquerda) e ‘automoveis’ (à direita). A dependência parcial plotada para ‘pedestres’ indica que quando a variável aumenta de 0 para 1, há elevação na probabilidade de o acidente ter sido causado por ‘outros’ fatores[13]. Desse modo, *contrario sensu*, acidentes que não envolvem pedestres estariam mais associados à causa ‘condutores’.

[13] Aumentos subsequentes no quantitativo de pedestres não acrescentam impacto à predição, mantendo-se o PDP inalterado.

**Figura 19** - PDP e ICE para as variáveis ‘pedestres’ e ‘automoveis’ (classificador de causa).




Fonte: Elaborado pela autora durante experimentos.

Quanto ao PDP da variável ‘automoveis’ (gráfico à direita), apresenta decréscimo à medida em que a quantidade desse tipo de veículo aumenta, até atingir 2 automóveis, após isso o valor da dependência parcial se estabiliza. Isso indica que, segundo aprendido pelo modelo, acidentes sem o envolvimento de automóveis são mais propensos a terem sido causados por ‘outros’ fatores, enquanto aqueles com um ou mais automóveis são mais inclinados a terem sido provocados pelo ‘condutor’.

Porém, na faixa em que há variação da dependência parcial, para o atributo ‘pedestres’ e também, de maneira mais intensa, para o atributo ‘automóveis’, o ICE mostra muita flutuação das instâncias individuais, que em alguns casos adotam trajetória oposta ao observado na média. Isso indica a existência de interações complexas com outras variáveis no modelo, as quais impactam a predição.

# 4. CONCLUSÃO

Nossa proposta de emprego de técnicas de aprendizado de máquina supervisionado e de métodos de XAI, voltada à problemática dos acidentes de trânsito nas rodovias federais de Pernambuco, possibilitou a identificação de quais fatores, dentre aqueles existentes nos conjuntos de dados trabalhados, detinham maior influência para a caracterização da severidade e causa dos sinistros.

Os resultados obtidos indicam que fatores como a presença de motocicletas, o tipo de acidente verificado, e o envolvimento de pedestres na ocorrência são significativos na determinação da gravidade do acidente. Também pode-se observar que a considerável existência de lacunas nos dados quanto à idade dos condutores envolvidos mostrou-se relevante para o aprendizado do modelo, sinalizando a importância de que os registros de acidentes sejam aprimorados nesse aspecto, possibilitando uma melhor elucidação de sua associação à severidade dos sinistros.

Quanto aos fatores relevantes para a causalidade dos acidentes, consideradas nas categorias sob responsabilidade do condutor ou de outros fatores, destacaram-se a identificação do tipo de sinistro, as condições meteorológicas, a presença de pedestres e a quantidade de veículos categorizados como automóveis envolvidos.

Informações desse tipo podem ser úteis para orientar a tomada de decisão quanto a políticas públicas nessa temática,tais como campanhas de conscientização focadas em motociclistas, reforço na fiscalização durante condições climáticas adversas, ou aprimoramentos viários para a segurança de pedestres, por exemplo.

Ademais, vislumbra-se que outros estudos que venham a explorar esse tipo de explicabilidade possam se favorecer, e propiciar mais contribuições, mediante a exploração de outras variáveis que ampliem o horizonte de dados analisados, tais como indicadores socioeconômicos e de infraestrutura das vias. A evolução nessa temática, com o emprego da tecnologia em prol da mitigação de riscos à Segurança Viária e à vida das pessoas, mostra-se de relevante importância.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

[1] WHO, U. N. Global Plan for the Decade of Action for Road Safety 2021– 2030. 2021.

[2] BRASIL. Ministério dos Transportes. **Você conhece o Pnatrans?** [Brasília]: Ministério dos Transportes, 02 fev. 2024. Disponível em: https://www.gov.br/transportes/pt-br/assuntos/transito/pnatrans. Acesso em: 29 abr. 2024.

[3] BRASIL. Ministério da Infraestrutura. **Rodovias Federais.** [Brasília]: Ministério da Infraestrutura, 28 mai. 2019. Disponível em: http://antigo.infraestrutura.gov.br/rodovias-brasileiras.html. Acesso em: 29 abr. 2024.

[4] ADADI, Amina; BERRADA, Mohammed. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). **IEEE access**, v. 6, p. 52138- 52160, 2018.

[5] DE OLIVEIRA, Maxuel Pereira *et al*. Uso de mineração de dados e tecnologia preditiva na prevenção de acidentes de trânsito no brasil. *In*: XIII SIMMEC, 2018, Vitória. **Anais eletrônicos**. Disponível em: https://doity.com.br/anais/xiiisimmec2018/trabalho/74037. Acesso em: 15 abr. 2024.

[6] OLIVEIRA, Luis E.; BORGES, André P. Comparação de modelos de classificação de categorias de acidentes nas rodovias federais. *In*: XI Escola Regional de Informática de Goiás, 2023, Goiânia. **Anais eletrônicos**. Porto Alegre: SBC, 2023. Disponível em: https://doi.org/10.5753/erigo.2023.237309. Acesso em: 15 abr. 2024.

[7] TAVEIRA, Julio *et al*. Aplicando Técnicas de Aprendizado de Máquina sobre Dados de Acidentes de Trânsito na BR 101 em Pernambuco. *In*: VI Seminário Internacional: Ciência, Tecnologia e Inovação em Segurança Pública, 2024, Florianópolis. **Anais**. Disponível em: https://www.researchgate.net/publication/380312551_Aplicando_Tecnicas_de_Aprendizado_de_Maquina_sobre_Dados_de_Acidentes_de_Transito_na_BR_101_em_Pernambuco_Applying_Machine_Learning_Techniques_to_Traffic_Crash_Data_on_BR_101_in_Pernambuco. Acesso em: 8 de jul. 2024.

[8] KOTSIANTIS, Sotiris B. *et al*. Supervised machine learning: A review of classification techniques. **Emerging artificial intelligence applications in computer engineering**, v. 160, n. 1, p. 3-24, 2007.

[9] FAWAGREH, Khaled; GABER, Mohamed Medhat; ELYAN, Eyad. Random forests: from early developments to recent advancements. **Systems Science & Control Engineering: An Open Access Journal**, v. 2, n. 1, p. 602-609, 2014.

[10] SCIKIT-LEARN DEVELOPERS. API Reference, RandomForestClassifier. Disponível em:https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier. Acesso em: 18 jun. 2024.

[11] BENTÉJAC, C.; CSÖRGŐ, A.; MARTÍNEZ-MUÑOZ, G. A comparative analysis of gradient boosting algorithms. **Artificial Intelligence Review,** v. 54, n. 3, p. 1937-1967, 2020. DOI: 10.1007/s10462-020-09896-5.

[12] XGBOOST DEVELOPERS. XGBoost Parameters. Disponível em: https://xgboost.readthedocs.io/en/stable/parameter.html. Acesso em: 18 jun. 2024.

[13] SCIKIT-LEARN DEVELOPERS. API Reference, KNeighborsClassifier. Disponível em: https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier. Acesso em: 18 jun. 2024.

[14] ROCHA, Anderson; GOLDENSTEIN, Siome Klein. Multiclass from binary: Expanding one-versus-all, one-versus-one and ecoc-based approaches. **IEEE transactions on neural networks and learning systems**, v. 25, n. 2, p. 289-302, 2013.

[15] SCIKIT-LEARN DEVELOPERS. API Reference, SVC. Disponível em: https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC. Acesso em: 18 jun. 2024.

[16] LETRACHE, Khadija; RAMDANI, Mohammed. Explainable Artificial Intelligence: A Review and Case Study on Model-Agnostic Methods. *In*: 2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA). IEEE, 2023. p. 1-8.

[17] SCHRÖER, Christoph; KRUSE, Felix; GÓMEZ, Jorge Marx. A systematic literature review on applying CRISP-DM process model. **Procedia Computer Science**, v. 181, p. 526-534, 2021.

[18] PANDAS DEVELOPERS. Pandas: API Reference. Disponível em: https://pandas.pydata.org/docs/reference/index.html. Acesso em: 25 jun. 2024.

[19] HUNTER, John et al. Matplotlib: API Reference. Disponível em: https://matplotlib.org/stable/api/index.html. Acesso em: 25 jun. 2024.

[20] SCIKIT-LEARN DEVELOPERS. Scikit-learn: Machine Learning in Python. Disponível em: https://scikit-learn.org/stable/. Acesso em: 25 jun. 2024.

[21] OPTUNA DEVELOPERS. Optuna: API Reference. Disponível em: https://optuna.readthedocs.io/en/stable/reference/index.html. Acesso em: 25 jun. 2024.

[22] MLFLOW DEVELOPERS. Mlflow: Documentation, Python API. Disponível em: https://mlflow.org/docs/latest/python_api/index.html. Acesso em: 25 jun. 2024.