

Comparação de modelos para classificação da gravidade de acidentes em rodovias federais do Paraná

Arthur M. P. Gabardo¹, Guilherme A. A. Schünemann¹, Luana S. Gallo¹

¹Programa de Pós-Graduação em Engenharia de Sistemas Eletrônicos (PPGESE)
Universidade Federal de Santa Catarina (UFSC) – Joinville, SC

{arthur.miguel, guilherme.arthur.a.schunemann}@grad.ufsc.br

Resumo. O Brasil, com sua vasta extensão territorial e uma das maiores malhas viárias do mundo, enfrenta desafios significativos relacionados à segurança no trânsito. Os acidentes de trânsito, em particular em rodovias federais, representam uma das principais causas de mortes no país, com implicações econômicas e sociais profundas. Este estudo compara o desempenho de diferentes algoritmos de classificação — kNN, SVM, árvore de decisão e rede neural — na predição da gravidade dos acidentes de trânsito. Utilizando uma base de dados aberta da PRF que abrange os anos de 2021 a 2023, foi realizado um extenso pré-processamento, codificação de variáveis categóricas e seleção de atributos. A técnica de SMOTE foi aplicada para lidar com o desbalanceamento das classes. A avaliação dos modelos foi conduzida por meio de métricas estatísticas como o F1-score. Os resultados demonstram que os modelos kNN e redes neurais foram os mais adequados para predição de acidentes fatais com score-F1 de 0,88 e 0,76, respectivamente.

1. Introdução

O Brasil é um país continental com um território de mais de 8,5 milhões de quilômetros quadrados, possuindo a quarta maior rede de rodovias do mundo, com 1,7 milhão de quilômetros. A malha viária brasileira, composta por rodovias federais, estaduais e municipais, públicas e concedidas, desempenha um papel crucial na integração regional com o transporte de pessoas e cargas, sendo fundamental para o desenvolvimento econômico do país. No entanto, a extensa dimensão territorial e as condições precárias das estradas contribuem para a alta incidência de acidentes de trânsito (Barros & Baggio, 2022).

Segundo a Confederação Nacional do Transporte, em 2023, os acidentes de trânsito no Brasil geraram um custo de R\$ 14 bilhões, com mais de 67 mil ocorrências registradas. No Paraná, a situação é mais preocupante, com o estado figurando como o terceiro no Brasil e o primeiro da região Sul em número de acidentes e óbitos. O estado registrou mais de 7 mil acidentes no ano, dos quais 7,8% foram fatais, gerando um impacto econômico de mais de R\$ 1,4 bilhão (Confederação Nacional do Transporte, 2024).

Esses números evidenciam a gravidade do problema e a necessidade da formulação de políticas públicas que visem a redução das fatalidades e o aumento da segurança nas estradas. Dado que as principais causas de acidentes são falhas humanas, infraestrutura inadequada e falta de fiscalização eficaz (Lobato, 2018), realizar uma análise detalhada dos dados coletados é essencial na tomada de decisões quanto a melhoria das vias e na promoção direcionada de campanhas para educação no trânsito. Diante deste contexto, técnicas de aprendizado de máquina e data mining têm ganhado relevância

por possibilitarem a identificação de padrões e tendências que não seriam facilmente detectáveis por meio de análises tradicionais (Amorim, 2019).

Neste trabalho é realizada uma comparação de diferentes modelos de classificação para prever a ocorrência de acidentes fatais em rodovias federais do Paraná, utilizando dados provenientes da base de dados aberta da Polícia Rodoviária Federal (2024). Os algoritmos, k-Nearest Neighbours (KNN), Support Vector Machine (SVM), Árvore de Decisão e Redes Neurais, são avaliados com base em métricas estatísticas de performance, como F1-score e AUC ROC. Além disso, uma busca em grade para determinar os melhores hiperparâmetros de cada modelo e técnicas de sobreamostragem são empregadas para lidar com o desbalanceamento da categoria alvo.

2. Trabalhos Relacionados

Nesta seção, serão apresentados estudos que exploram temas pertinentes ao projeto proposto e discutidos os principais objetivos e metodologias empregadas pelos autores. A análise se concentra em explorar as abordagens utilizadas nestes estudos e a relevância para o projeto desenvolvido.

O estudo conduzido por Cuenca et al. (2018) se destaca pela introdução de técnicas distintas para prever a severidade de acidentes, explorando metodologias como Boosting Trees, Deep Learning e Naive Bayes. Este trabalho amplia o espectro de técnicas aplicáveis à classificação, permitindo uma análise detalhada dos procedimentos utilizados na manipulação da base de dados, incluindo o balanceamento de dados e a seleção criteriosa de atributos.

O trabalho de Amorim (2019) aborda a aplicação de técnicas de aprendizado de máquina na classificação da severidade de acidentes em rodovias brasileiras. Com uma metodologia similar ao presente projeto, esse estudo utilizou uma base de dados extensa, abrangendo uma década de registros de acidentes em rodovias federais, e explorou diferentes algoritmos de classificação, como SVM, Redes Neurais e Random Forest.

A investigação de Jianfeng et al. (2019) propõe um modelo de classificação que combina a teoria dos conjuntos aproximados com o SVM, permitindo uma análise mais precisa e eficiente dos fatores que influenciam a gravidade dos acidentes de trânsito. A utilização da teoria dos conjuntos aproximados para simplificar e reconstruir os fatores de influência antes da aplicação do SVM melhora significativamente a acurácia e a velocidade de classificação, particularmente em cenários com grandes volumes de dados.

A pesquisa de Labib et al. (2019) utiliza aprendizado de máquina para analisar a intensidade e classificar a severidade de acidentes de trânsito em Bangladesh, país que também sofre com um alto índice de acidentes anualmente. As análises foram realizadas com diferentes técnicas de aprendizado supervisionado, como Árvore de Decisão, KNN, Naive Bayes e AdaBoost.

3. Metodologia

A proposta deste estudo é classificar a gravidade dos acidentes de trânsito – entre fatais e não fatais – por meio de uma análise abrangente dos dados de acidentes coletados pela PRF, entre os anos de 2021 e 2023 no estado do Paraná, Brasil. A metodologia é estruturada em cinco fases distintas: coleta e pré-processamento de dados, sobreamostragem da

base, seleção de atributos, treinamento e validação dos modelos. As seções a seguir detalham os processos e as decisões envolvidas em cada fase, bem como suas bases teóricas.

3.1. Coleta e Pré-processamento de Dados

O conjunto de dados analisado compreende registros de acidentes de trânsito de 2021 a 2023, provenientes de bancos de dados oficiais. Esses registros, agrupados por anos de ocorrência, incluem diversos atributos relacionados aos acidentes, como tipos de veículos, condições das vias e do tempo, dados dos envolvidos no acidente (número de pessoas, idade e sexo), bem como a classificação dos acidentes em “Sem Vítimas”, “Com Vítimas Feridas” e “Com Vítimas Fatais”.

Inicialmente, os dados compreendidos pelo intervalo pré-estabelecido (2021 – 2023) foram carregados em um único conjunto, dos quais foram mantidas somente as entradas referentes ao estado do Paraná. Ademais, entradas inválidas como valores implausíveis ou errôneos foram filtrados, por exemplo, anos de fabricação de veículos registrados como zero ou idades irreais (menores que 0 ou superiores a 130 anos). Observações com um ou mais atributos faltantes ou não informados foram removidas. Essa limpeza de dados garante que a análise seja baseada somente em registros confiáveis.

Após a remoção de entradas inválidas, as variáveis categóricas foram padronizadas para reduzir ruídos e melhorar a precisão dos modelos. Diferentes categorias dos atributos *tipo do veículo* e *tipo do acidente* foram agrupadas, por apresentarem semelhança na sua descrição ou apenas diferenças ortográficas. O agrupamento reduziu de 15 para seis o número de categorias do atributo *tipo do veículo* e de 17 para nove do atributo *tipo do acidente*.

Devido à alta cardinalidade das variáveis categóricas no conjunto de dados, utilizou-se o *k-fold target encoding* para sua codificação. Essa técnica substitui cada categoria pela média da variável alvo em cada fold da validação cruzada, capturando a relação entre os atributos categóricos e o desfecho (Rodríguez et al., 2018). O *k-fold target encoding* é vantajoso nesse contexto ao explorar padrões complexos, minimizando o vazamento de dados na codificação e o overfitting. Foram feitos cinco folds na base de dados para codificação das variáveis categóricas, seguindo as recomendações apontadas na literatura (Pargent et al., 2022).

Em seguida, os valores dos atributos foram centralizados e reescalados para assegurar uma contribuição equilibrada no treinamento, evitando que atributos com maiores intervalos numéricos influenciassem desproporcionalmente os modelos. Essa etapa é essencial para integrar variáveis com diferentes escalas, mantendo o foco nas relações entre os atributos, em vez de suas magnitudes.

Uma decisão significativa durante o pré-processamento foi focar especificamente em casos que resultaram em fatalidades ou desfechos não fatais. A partir desta escolha dois subconjuntos foram gerados, um considerando apenas entradas com as classes “Sem Vítimas” ou “Com Vítimas Fatais”, e outro subconjunto onde entradas com a classe “Com Vítimas Feridas” foram agrupadas à classe “Sem Vítimas”, isolando a categoria-alvo “Com Vítimas Fatais”. A abordagem do segundo subconjunto resultou em uma base desbalanceada que foi tratada por meio da sobreamostragem da categoria-alvo.

3.2. Sobreamostragem

A fim de lidar com o desbalanceamento da classe alvo, foi empregada a técnica de *Synthetic Minority Over-Sampling Technique* (SMOTE), conforme proposto por Chawla et al. (2002). O SMOTE gera novas amostras sintéticas da classe minoritária (“Com Vítimas Fatais”) através da interpolação entre os pontos existentes, aumentando a representatividade desta classe no conjunto de dados. Ao invés de simplesmente replicar as instâncias existentes, essa técnica cria novas instâncias, que auxiliam a capacidade do modelo de generalizar. Após a aplicação do SMOTE, a proporção de registros “Com Vítimas Fatais” aumentou de 8,8% para 36,7%, proporcionando um conjunto de dados mais equilibrado para o treinamento dos modelos.

É importante destacar que, apesar do aumento na proporção da classe minoritária, a base resultante da aplicação da técnica deve manter uma distribuição das variáveis similar à da base original. Essa similaridade é crucial para garantir que o modelo treinado capture os padrões reais presentes nos dados, e não artefatos gerados pela sobreamostragem. A Figura 1 ilustra as distribuições de densidade das variáveis antes e depois da aplicação do SMOTE, evidenciando a consistência das distribuições e a eficácia do método em preservar as características originais dos dados.

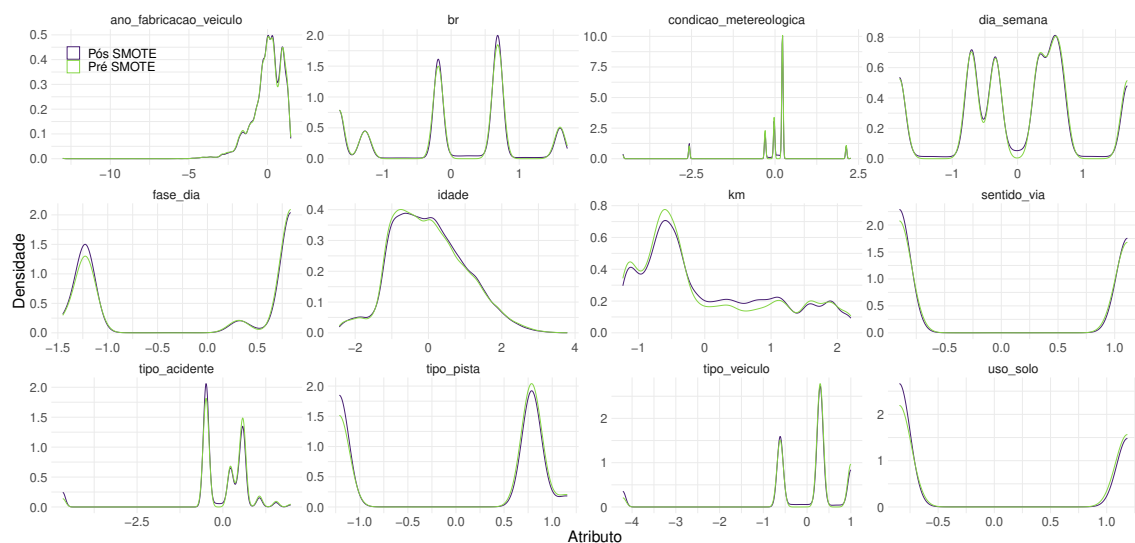


Figura 1. Densidades da distribuição de variáveis da base pré e pós SMOTE.

3.3. Seleção de Atributos

A seleção de atributos mais relevantes para a classificação foi realizada utilizando a técnica *Random Forest - Recursive Feature Elimination* (RF-RFE), conforme descrito por Archer and Kimes (2008). O método remove os atributos menos relevantes com base em suas importâncias relativas, calculadas por meio dos modelos de Random Forest. Esse processo continua de maneira iterativa até que um subconjunto otimizado de atributos seja obtido com base em uma métrica estatística de performance (tipicamente acurácia ou Kappa de Cohen), contendo apenas aqueles que contribuem significativamente para a previsão da categoria-alvo (Gregorutti et al., 2017).

A aplicação do RF-RFE reduz a dimensionalidade do conjunto de dados e elimina atributos com alta correlação, mantendo a precisão dos modelos, melhorando a eficiência

computacional e facilitando a interpretação dos resultados, mudanças que permitem o foco nos fatores mais críticos. A técnica foi aplicada tanto na base reduzida (com a classe “Com Vítimas Feridas” removida), quanto na base resultante do SMOTE. A Tabela 1 apresenta os atributos selecionados para cada uma dessas bases, destacando as variáveis que demonstraram maior relevância para a classificação dos acidentes.

Tabela 1. Atributos selecionados pelo RF-RFE.

Base de Dados	Atributos Selecionados
Base reduzida	“tipo_acidente”, “tipo_veiculo”, “tipo_pista”, “km”, “uso_solo”, “condicao_metereologica”, “dia_semana”, “br”, “fase_dia”
Base SMOTEd	“tipo_acidente”, “tipo_veiculo”, “tipo_pista”, “km”, “uso_solo”, “condicao_metereologica”, “dia_semana”, “br”, “fase_dia”, “ano_fabricacao_veiculo”, “idade”, “sentido_via”

3.4. Modelos de Classificação

O problema de classificação consiste no processo de prever uma categoria ou classe para uma determinada amostra com base em seus atributos. O objetivo desse processo é desenvolver um modelo que, após ser treinado com um conjunto de dados rotulado, seja capaz de generalizar suas previsões para novos dados. No caso deste estudo, o problema de classificação envolve distinguir entre acidentes de trânsito fatais e não fatais, utilizando uma série de variáveis preditoras relacionadas às condições dos acidentes, características dos veículos e envolvidos. A seguir, são descritos os principais modelos de classificação utilizados nesta análise.

O *k-Nearest-Neighbors* (KNN) é um dos métodos de classificação mais simples e intuitivos, baseado na ideia de que amostras similares tendem a pertencer à mesma classe. Ao classificar uma nova observação, o algoritmo identifica as k amostras mais próximas no espaço de atributos e atribui a classe mais comum entre essas vizinhas (Ripley, 2007). Sua eficácia depende da escolha adequada de valores de k .

O *Support Vector Machine* (SVM) é um método de classificação que busca encontrar um hiperplano de separação ótimo, maximizando a margem entre as amostras de diferentes classes (Boser et al., 1992). No caso deste estudo, utilizou-se um SVM com kernel linear, o que significa que o modelo procura uma fronteira de decisão linear no espaço dos atributos.

As *Árvores de Decisão* são modelos que utilizam uma estrutura hierárquica de decisões baseadas em estruturas condicionais para classificar as amostras. A construção da árvore envolve a divisão recursiva do conjunto de dados com base no atributo que mais separa as classes, resultando em uma série de nós que representam os critérios de decisão e folhas que correspondem às classes finais (Breiman, 2017).

As *Redes Neurais* são modelos inspirados no funcionamento do cérebro humano, compostas por neurônios organizados em camadas. Para este estudo, foi implementada uma rede neural com uma única camada escondida e função de ativação sigmoide, o que torna o modelo uma aproximação não linear simples, mas ainda poderosa, para problemas de classificação (Ripley, 2007). A presença de uma camada escondida permite que a rede

capture interações complexas entre as variáveis de entrada, sendo adequada para cenários onde os padrões de decisão não são linearmente separáveis.

3.5. Validação

A validação dos modelos de classificação foi realizada utilizando a técnica de *Leave-Group Out Cross-Validation* (LGOCV) com 5 folds, aplicada ao conjunto de dados previamente dividido em 80% para treino e 20% para teste. O LGOCV é uma variação da validação cruzada onde, em cada iteração, um grupo de amostras é mantido para teste, enquanto o restante é usado para treinamento. Esse método é particularmente eficaz em garantir que o modelo seja validado em diferentes subconjuntos dos dados, reduzindo a variabilidade dos resultados e proporcionando uma avaliação mais robusta da performance dos modelos (Arlot & Celisse, 2010).

Além disso, foi realizada uma busca em grade (*grid-search*) para a otimização dos hiperparâmetros dos modelos, conforme os valores especificados na Tabela 2. Essa busca sistemática permitiu identificar a melhor combinação de parâmetros para maximizar o desempenho dos modelos em termos das métricas analisadas.

Tabela 2. Valores adotados para os hiperparâmetros durante busca em grade.

Modelo	Hiperparâmetro	Valor
k-Nearest Neighbors	k	[3, 5, 7, ..., 21]
Support Vector Machine	Custo (C)	[0.1, 0.2, ..., 1.5]
Árvore de Decisão	Profundidade máxima	[2, 3, 4, ..., 25]
Rede Neural	Núm. neurônios	[11, 13, 15, ..., 51]
	Decaimento	[10^{-4} , 10^{-3} , ..., 0.1]

As métricas utilizadas para avaliar a performance dos modelos foram acurácia, precisão, revocação, F1-score e AUC-ROC. A acurácia mede a proporção de previsões corretas em relação ao total de previsões feitas, sendo uma métrica global de performance. A precisão e a revocação são métricas focadas na classe positiva, sendo a precisão a fração de verdadeiros positivos entre todas as previsões positivas, e a revocação, a fração de verdadeiros positivos entre todos os casos reais positivos. O F1-score é a média harmônica entre precisão e revocação, proporcionando um equilíbrio entre essas duas métricas. Por fim, a AUC-ROC avalia a capacidade do modelo em distinguir entre as classes ao longo de diferentes limiares de decisão, oferecendo uma visão mais completa da performance do modelo.

4. Resultados

Ao utilizar a base de dados reduzida, na qual foram excluídas as entradas com a classe “Com Vítimas Ferida”, os resultados indicaram que apenas os modelos KNN e rede neural (nnet) conseguiram efetivamente classificar acidentes com vítimas fatais. Apesar da simplificação da base, os demais modelos testados (SVM e árvore de decisão) apresentaram dificuldades significativas em separar adequadamente as classes, resultando em baixa performance. As curvas ROC comparando todos os modelos (Figura 2), destacam essa disparidade de desempenho, com as áreas sob a curva (AUC) para kNN e nnet superando

as dos demais modelos. A ineficiência dos modelos SVM e árvore de decisão ficam claros ao comparar a taxa de verdadeiros positivos inferior ou próxima a 0,5 como demonstrado na Tabela 3.

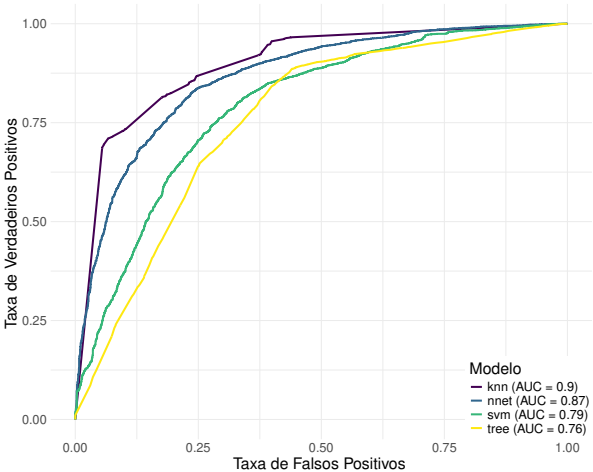


Figura 2. Curva característica ROC, modelos em base reduzida.

Métrica ¹	knn	tree	svm	nnet
ACC	0,91	0,77	0,74	0,84
PPV	0,86	0,74	0,72	0,80
TPR	0,90	0,56	0,47	0,73
SPC	0,92	0,89	0,89	0,90
F1	0,88	0,64	0,57	0,76
AUC	0.90	0,76	0,79	0,87

Tabela 3. Métricas estatísticas, modelos em base reduzida.

A análise das matrizes de confusão, inseridas na Figura 3, para os modelos kNN e nnet revela que, embora tenham sido capazes de identificar alguns acidentes fatais, ambos apresentaram um número elevado de falsos negativos, indicando uma baixa revocação. A baixa prevalência de acidentes fatais na base reduzida pode ter contribuído para esses resultados, visto que os modelos podem ter se inclinado a classificar a maioria dos casos como não fatais. Esse fenômeno é comum em bases desbalanceadas, onde a classe majoritária tende a dominar as previsões, prejudicando a detecção da classe minoritária.

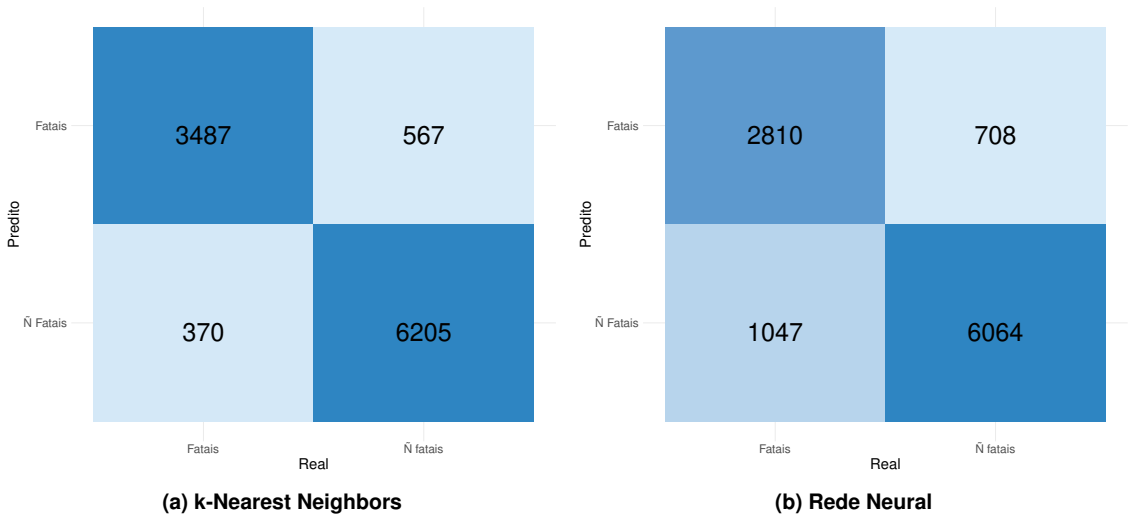


Figura 3. Matrizes de confusão dos modelos em base reduzida.

¹ACC = Acurácia, PPV = Precisão, TPR = Revocação, SPC = Especificidade, F1 = score-F1, AUC = área abaixo da curva característica ROC.

Após a aplicação da técnica de sobreamostragem SMOTE, os resultados demonstrados na Tabela 4 indicam uma melhora na acurácia e revocação dos modelos, entretanto, uma piora na precisão e no score-F1. Este desbalanço é demonstrado na Figura 4, onde o kNN apresentou uma melhora na área sob a curva, enquanto o SVM obteve uma piora nos resultados.

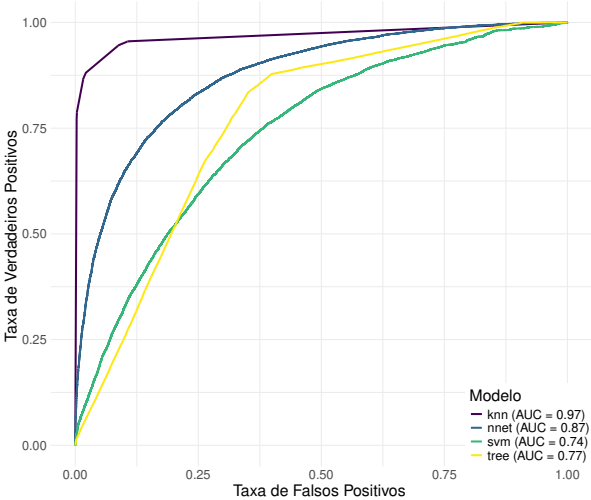


Figura 4. Curva característica ROC, modelos em base com sobreamostragem.

Métrica	knn	tree	svm	nnet
ACC	0,95	0,85	0,86	0,85
PPV	0,64	0,28	0,27	0,32
TPR	0,99	0,49	0,39	0,61
SPC	0,94	0,88	0,90	0,87
F1	0,78	0,36	0,32	0,41
AUC	0.97	0,77	0,74	0,87

Tabela 4. Métricas estatísticas, modelos em base com sobreamostragem.

As matrizes de confusão (Figura 5) para os modelos kNN e nnet reforçam a melhora na revocação e na piora precisão de ambos. Portanto, embora o SMOTE tenha aumentado a capacidade dos modelos de determinar se um acidente possui vítimas fatais, também houve um aumento na taxa de falsos positivos, que de maneira geral reduziu o score-F1 de todos os modelos estudados.

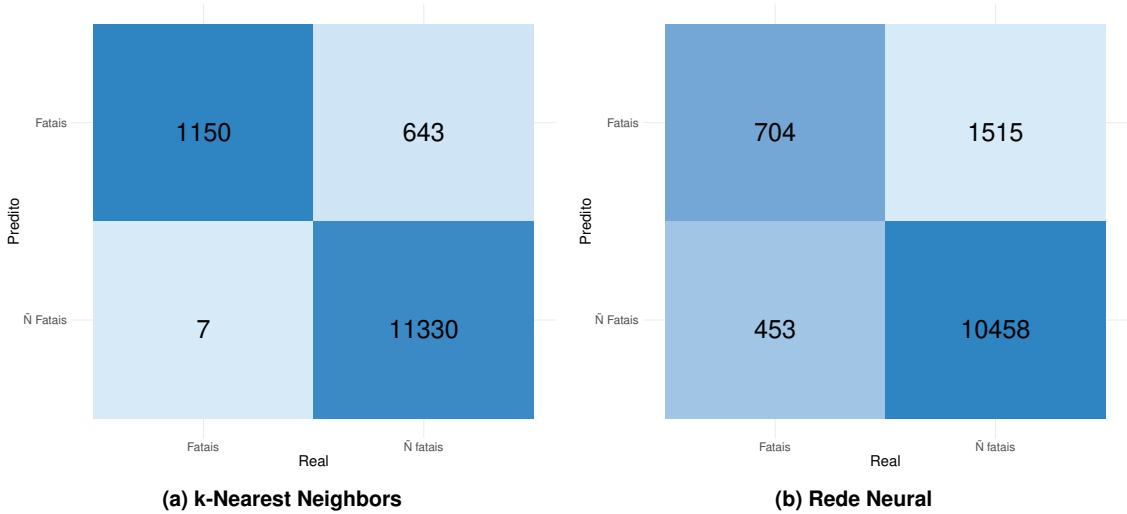


Figura 5. Matrizes de confusão dos modelos em base com sobreamostragem.

5. Conclusão

Este estudo comparou quatro algoritmos de classificação para prever a gravidade de acidentes de trânsito em rodovias federais no estado do Paraná, utilizando dados fornecidos pela PRF entre 2021 e 2023. Após um extenso pré-processamento, incluindo a remoção de dados inválidos e a codificação de variáveis categóricas, aplicou-se a técnica de sobre-amostragem SMOTE para lidar com o desbalanceamento da classe-alvo e a eliminação recursiva de atributos (RF-RFE) para determinação das variáveis mais significativas da base.

O aumento da representatividade da classe minoritária (acidentes fatais) resultou em uma melhora na acurácia e revocação dos modelos, mas uma piora no score-F1. O kNN aplicado na base reduzida obteve o melhor desempenho, com $\text{score-F1} = 0,88$, seguido da rede neural. Os modelos de SVM e árvore de decisões não demonstraram desempenho satisfatório em classificar os acidentes fatais, com taxas de positivos verdadeiros inferiores a 0,5.

É importante destacar que, apesar dos resultados positivos de alguns modelos, aprimoramentos podem ser feitos com a inclusão de novas variáveis descritivas que influenciam a gravidade dos acidentes, como o fluxo de veículos e a presença de infraestruturas de segurança. Pesquisas futuras podem explorar a integração dessas variáveis, além de buscar modelos de explicabilidade para obter melhores classificadores e informações a respeito da dinâmica dos acidentes.

Agradecimentos

Os autores agradecem ao Núcleo de Simulação e Otimização de Sistemas Dinâmicos pelos recursos computacionais disponibilizados para realização do trabalho.

Disponibilidade dos Dados

Todos os dados e códigos estão disponíveis para consulta em: https://github.com/arthur-miguel/accident_classification

Referências

- Amorim, B. d. S. P. (2019). *Uso de aprendizado de máquina para classificação de risco de acidentes em rodovias*. (Mestrado em Ciência da Computação). Universidade Federal de Campina Grande (UFCG), Paraíba.
- Archer, K. J., & Kimes, R. V. (2008). Empirical characterization of random forest variable importance measures. *Computational statistics & data analysis*, 52(4), 2249–2260.
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40–79.
- Barros, P. H., & Baggio, I. (2022). Uma análise espacial da malha rodoviária brasileira: relações com o desenvolvimento econômico regional. *Crescimento e desenvolvimento numa perspectiva interdisciplinar: ensaios sobre o crescimento econômico brasileiro*, 1, 360–372.
- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on computational learning theory* (pp. 144–152).

- Breiman, L. (2017). *Classification and regression trees*. Routledge.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321–357.
- Confederação Nacional do Transporte. (2024). *Painel CNT: Acidentes Rodoviários*. Brasília, Brasil. (Disponível em: <https://www.cnt.org.br/pesquisas>)
- Cuenca, L. G., Puertas, E., Aliane, N., & Andres, J. F. (2018). Traffic accidents classification and injury severity prediction. In *2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE)* (p. 52-57).
- Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, 27, 659–678.
- Jianfeng, X., Hongyu, G., Jian, T., Liu, L., & Haizhu, L. (2019). A classification and recognition model for the severity of road traffic accident. *Advances in Mechanical Engineering*, 11(5).
- Labib, F., Rifat, A. S., Hossain, M. M., Das, A. K., & Nawrine, F. (2019). Road accident analysis and prediction of accident severity by using machine learning in bangladesh. In *International Conference on Smart Computing and Communications* (p. 1-5).
- Lobato, C. R. V. (2018). *Análise dos acidentes nas rodovias federais concedidas no brasil* (Mestrado em Geotecnia E Transportes). Universidade Federal de Minas Gerais (UFMG), Belo Horizonte.
- Pargent, F., Pfisterer, F., Thomas, J., & Bischl, B. (2022). Regularized target encoding outperforms traditional methods in supervised machine learning with high cardinality features. *Computational Statistics*, 37(5), 2671–2692.
- Polícia Rodoviária Federal. (2024). *Dados abertos*. Brasília, Brasil. (Disponível em: <https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos>)
- Ripley, B. D. (2007). *Pattern recognition and neural networks*. Cambridge university press.
- Rodríguez, P., Bautista, M. A., Gonzalez, J., & Escalera, S. (2018). Beyond one-hot encoding: Lower dimensional target embedding. *Image and Vision Computing*, 75, 21–31.