

Machine Learning para Prevenção de Acidentes Ferroviários

Dhiego Iannarelli*, Rafael Pinto, Vitor Braga e Vivian Gasperino

Diretoria de Tecnologia, Rua Emílio Bertolini nº 100, 82920-030, Curitiba/PR

dhiego.iannarelli@rumolog.com, rafael.pinto@rumolog.com
vitor.braga@rumolog.com e vivian.gasperino@rumolog.com

Resumo O objetivo foi desenvolver um modelo preditivo, em trecho próximo ao porto de São Francisco, no estado do Paraná, contemplando cerca de duzentos quilômetros de extensão de ferrovia. Foram avaliadas 651 viagens realizadas entre os anos de 2018 e 2019, utilizando-se de um total de 108 variáveis (entre via permanente, material rodante e integradoras), proporcionado um *dataset* de dimensão de 651 viagens x 108 variáveis. Em termos técnicos, visou-se obter como *output*, através de um modelo de *Machine Learning*, uma forma de se representar a similaridade de um trem prestes a sair para sua viagem *versus* eventos de acidentes já ocorridos no passado e devidamente registrados pela Companhia. Dessa forma, em posse desses importantes dados, se mensuraria o potencial risco de acidente ao longo de uma viagem, antes mesmo de seu início efetivo, demonstrando as principais variáveis críticas daquela viagem. Sendo assim, possibilitaria-se condições ainda mais assertivas nas tomadas de decisão, prezando pela segurança ferroviária e pessoal.

INTRODUÇÃO

É nítida a evolução por parte das companhias, em termos de dados, permeando desde sua geração, extração, análises, democratização, bem como, até sua utilização para tomada de decisão. Todas essas tratativas em função dos dados, muitas das vezes, eram consideradas impensáveis anos atrás, mas em contraste, possível atualmente, sobretudo pelos importantes avanços tecnológicos que vivenciamos dia após dia.

Soma-se a isso, e de maneira complementar, o crescimento da digitalização, a conectividade de equipamentos, materiais e ferramentas – *IoT*, que por sua vez, acabam proporcionando terreno ainda mais fértil para a utilização ímpar dos dados, pelas mais variadas companhias, independente de seus respectivos setores.

É também condição *sine qua non* ao crescimento das companhias, a necessidade

de geração de valor e a parceria inegociável para com a sociedade, de modo a existir um crescimento sustentável conjunto (benéfico e agregador para ambas as partes).

Ao coadunar a riqueza de dados e o crescimento sustentável conjunto de todos *stakeholders* envolvidos (sociedade, colaboradores, investidores, companhias e congêneres), pode-se imaginar um cenário extremamente frutífero, através por exemplo, do reconhecimento de padrões, entendimento de perfis da sociedade, dos clientes, sensibilidade em função de cenários micro e macroeconômicos, possíveis correlações e tendências, análise de riscos, entre outros. Dessa forma, obtendo-se como *output*, informações cada vez mais valiosas, concisas, estruturadas e direcionadoras (de maneira automatizada), para tomadas de decisões ainda mais assertivas em prol do crescimento sustentável conjunto.

Atualmente essas avaliações podem ser executadas manualmente por profissionais, entretanto, em função do grande volume e possíveis combinações entre os dados, pode-se obter as devidas respostas em tempos não condizentes com a velocidade que de fato se necessita, impactando diretamente na assertividade das tomadas de decisões.

Diante disto, a Rumo Logística, impulsionada pela área de Tecnologia e Inovação, vislumbrou um cenário similar de análise de dados atrelado à resolução e/ou mitigação de algum desafio enfrentado no dia a dia de seus negócios e que poderia proporcionar benefícios dos mais variados à todos os envolvidos - direta e indiretamente.

É de extrema importância ressaltar que a segurança é um valor inegociável para a Rumo, a qual tem investido significativamente em diversas frentes (a exemplo de estudos, projetos e campanhas informativas e educacionais). Portanto, a proposta central aqui descrita foi de utilizar da vasta gama de dados disponíveis em nossos sistemas, e, correlaciona-las em benefício da segurança ferroviária e pessoal.

Assim como diversas outras empresas, na Rumo, usava-se um monitoramento tradicional, avaliando-se variáveis de maneira isolada, enquanto que em casos de acidentes, geralmente, ocorre-se uma soma de fatores, ou seja, existe a possibilidade que as variáveis isoladas possam estar todas dentro dos limites estabelecidos como seguros, mas quando combinadas entre si, possam gerar um desequilíbrio e suceder de fato um acidente. Isto posto, se propôs o desenvolvimento de um sistema de predição, que permita avaliar a condição de um trem e as localidades a serem percorridas pelo mesmo, retornando à similaridade de um trem prestes a sair para sua viagem em relação a eventos de acidentes já ocorridos no passado e devidamente registrados pela Companhia.

Dessa forma, estar-se-ia de posse de informações que mensurariam o potencial risco de acidente ao longo de uma viagem, antes mesmo do início efetivo da mesma,

disponibilizando as principais variáveis críticas daquela viagem, para aquele trem.

Uma vez definido o escopo de análise, segmentou-se diversas etapas que poderiam evoluir, em boa parte do estudo, de maneira paralela, e, conseqüentemente, ter-se-iam validações e testes parciais.

Como primeira etapa, reuniu-se diversos especialistas dos mais variados campos de atuação, com o objetivo de melhor entender a respeito da segurança ferroviária e pessoal sob diferentes perspectivas, englobando por exemplo, os tipos de acidentes existentes, possíveis variáveis influenciadoras, sistemas que armazenam dados a respeito dessas ocorrências (seja dos acidentes seja das variáveis a serem possivelmente monitoradas), configuração do ambiente de arquitetura para coleta dos dados, definição de esforços, fluxo de processos existentes (restrições, correções, manutenções, planejamento de transporte, revestimento, viagem etc), entre outros.

Em seguida, foi definida a localidade a ser estudada, selecionando-a em função da quantidade substancial de dados à disposição para análise, bem como, das problemáticas enfrentadas por cada uma das localidades mapeadas (quantidade de acidentes e respectivas gravidades). Portanto, foi escolhido um trecho de cerca de duzentos quilômetros de extensão, em região próxima ao porto de São Francisco, no estado do Paraná. Nesta análise, foram selecionadas três causas primárias de acidentes a serem analisadas pelo estudo, sendo elas, amortecimento, bitola e nivelamento, em função do potencial impacto positivo gerado (crescimento sustentável conjunto), bem como, da complexidade em obtenção, análise e aplicabilidade dos dados em relação ao tempo de estudo proposto.

Diante das causas primárias de acidentes selecionadas (ainda de maneira conjunta com diversos especialistas, dos mais variados campos de atuação), foram elencadas variáveis e testes de hipóteses de acidentes, englobando não somente estas variáveis

elencadas, bem como, outras intrínsecas às hipóteses apontadas, de modo a verificar possíveis indícios de relevantes variáveis preditoras para um acidente ferroviário, como por exemplo, dormentação inservível, abertura de bitola e empeno.

Isto posto, houve uma estruturação visando a captura dos dados, a consolidação dos mesmos (de maneira automatizada, reduzindo tempo de trabalho e possíveis erros/equívocos de manipulação manual dos dados), o tratamento, as análises, as correlações, e, conseqüentemente, a geração de possíveis alertas de acidentes (em função de similaridade com acidentes passados).

A título de exemplo, analisando apenas oito sistemas que poderiam ser fontes úteis ao estudo, existiam mais de 200 tabelas. Fontes de informações estas, que são atualizadas periodicamente, reforçando a necessidade de tecnologias apropriadas para esta interação contínua.

DIAGNÓSTICO

O acidente ferroviário pode trazer diversas conseqüências, aos mais variados *stakeholders*, permeando desde a redução da qualidade de prestação do serviço até mesmo risco à própria vida dos envolvidos.

A ANTT (Agência Nacional de Transportes Terrestres) divulga recorrentemente, dados referentes à acidentes ferroviários. De acordo com a Agência, conforme Figura 1, é demonstrado a proporção de acidentes ferroviários, por tipologia, que ocorreram no Brasil, entre 2006 e 2013.

Pelos dados apresentados na Figura 1, verifica-se que mais de 50% dos casos de acidentes estão relacionados a material rodante e via permanente. E ainda, as demais causas/tipologias, são de difícil atuação direta (em termos de análise de dados), tais como, vandalismo, interferência de terceiros e falhas humanas (representando cerca de 25% dos acidentes reportados).

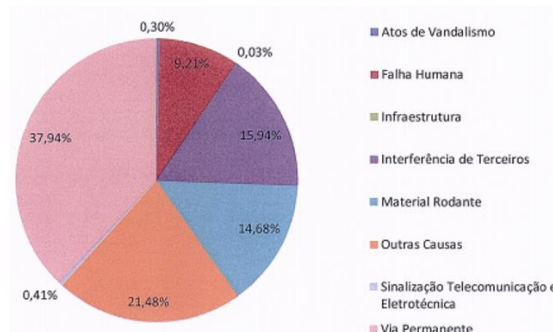


Figura 1 – Proporção por tipo de acidente ferroviário no Brasil entre 2006 e 2013

Na Rumo, a situação parecia ser similar, então um diagnóstico foi realizado através de um *design sprint*, técnica de imersão e prototipação muito utilizada (Banfield et al., 2015). Toda a fase de mapeamento, implementação e análise de resultados foi realizada em conjunto à Keyrus, um dos parceiros da Rumo em projetos de ciência de dados. Dentro da fase de mapeamento, foi gerado uma árvore de causas e variáveis para o histórico de acidentes, criada em conjunto com diversos especialistas da Rumo.

Ao fim da fase de mapeamento, a oportunidade a ser explorada, bem como, o escopo, estavam definidos. O estudo deveria analisar três causas primárias (amortecimento, bitola e nivelamento) através de diversos dados históricos, tais como, manutenção, condução e inspeções, e, confrontá-los com o histórico de acidentes devidamente registrados pela Companhia. Essas causas primárias, foram definidas por termos dados gerados diretamente pela empresa, possuem uma confiança aceitável e que representam significativamente o número de acidentes. Ademais, foi realizada uma matriz considerando o impacto operacional e as dificuldades percebidas em questão de predição. Todavia, qualquer variável poderia ser avaliada, sendo relacionada diretamente a essas causas ou não.

TEORIA

De antemão, é de fundamental importância a escolha de um algoritmo de *machine learning* que seja apropriado às grandezas a serem analisadas. Diante disso, foram testados e ponderados os prós e contras de alguns modelos que atendiam ao escopo proposto, entre eles, *XGBoost*, regressão linear, rede neural e *random forest*. O escolhido ao final dessa análise para continuidade do estudo, foi o *XGBoost*, o qual, costuma ter assertividade elevada e facilidade de implementação, porém, com interpretabilidade mediana, ou seja, dificuldade no entendimento do porquê das decisões do modelo (Chen e Guestrin, 2016).

A escolha pelo *XGBoost* foi pautada pelos resultados iniciais obtidos, o qual apresentou o maior valor em termos de *Precision*, entre as quatro opções testadas, assim como, o segundo maior *Recall*.

Como o acidente ferroviário é um evento raro (<1% dos eventos analisados), a interpretabilidade mediana do algoritmo é aceita, uma vez que sua eficiência é relativamente fácil de ser conferida e o

impacto de falsos positivos é relativamente pouco significativo na operação ferroviária.

Foram avaliados quatro pilares fundamentais que deveriam ser analisados diante do estudo, sendo: o possível *overfitting* do modelo, existência de classes minoritárias, métricas de resultados e interpretabilidade dos resultados.

Em relação ao *overfitting*, de acordo com a biblioteca *Scikit Learn*, pode-se obter resultados preditivos satisfatórios quando do teste com dados já conhecidos pelo modelo, porém, quando replicado para dados inéditos, pode-se obter resultados insatisfatórios. Para tal, se faz necessário segregar os dados em duas parcelas, uma para treinamento e uma para teste. Uma metodologia conhecida é o *k-fold*, ou seja, os dados são parcelados em k partes, sendo o modelo treinado em k-1 parcelas e testado na parte restante.

Em relação ao número de parcelas que irão compor a totalidade dos dados (k), deve-se ponderar, de modo a se levar em conta, os resultados esperados para o modelo *versus* os custos/esforços envolvidos para tal.

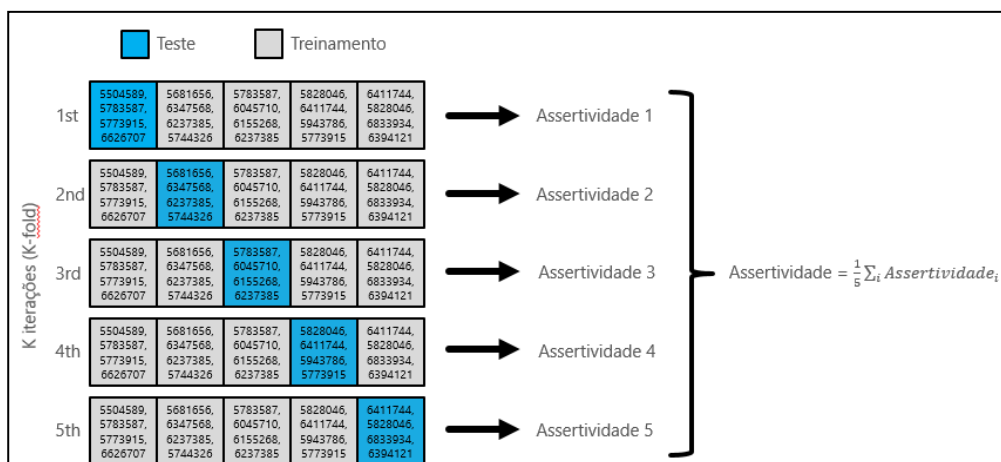


Figura 2 - Validação cruzada *K-fold*

No que se refere à existência de classes minoritárias, é fundamental atentar-se do cumprimento ou não da condicionante, uma vez que o modelo pode ser enviesado a prever em todas as circunstâncias a favor da classe majoritária, ou seja, em nosso cenário, as previsões poderiam ser classificadas como

“não acidentes”, e, apesar de fazer a escolha correta em parte substancial das vezes, não estaria contribuindo para o propósito deste estudo de avaliar/predizer os casos de acidentes - classes minoritárias. Ou seja, existiria um valor de Revocação elevada, porém, uma Precisão reduzida.

Para tanto, foi adotada a técnica *SMOTE* (*Synthetic Minority Oversampling Technique*), a qual, segundo *Chawla, N. V. et al., 2002*, tem o objetivo de criar amostras artificiais entre as classes minoritárias, através de uma conexão entre os dados já existentes.

É importante citar, que as amostras artificiais criadas, são utilizadas apenas na parte de dados selecionada para treino, ou seja, sem influência direta na parcela destinada a validação. Portanto, as métricas de performance do modelo não são afetadas pela aplicação da técnica em questão.

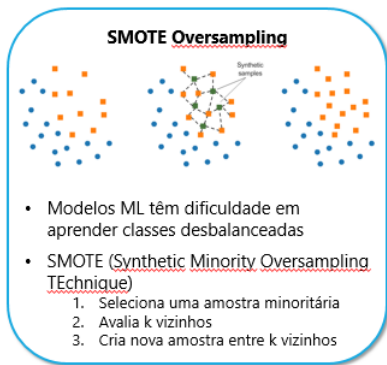


Figura 3 – Exemplificação de técnica *SMOTE*

Em relação às métricas de resultados do modelo preditivo, utilizou-se da Matriz de Confusão, que por sua vez, terá como escopo declarar a frequência de cada resultado, bem como, sua respectiva classificação na relação entre predição *versus* análise real.

Na prática, o funcionamento ocorre da seguinte maneira:

1. Verdadeiro Positivo: Ocorre quando as predições preveem o acontecimento e de fato ocorre junto ao cenário real estudado.
2. Falso Positivo: Ocorre quando as predições preveem o acontecimento e de fato não ocorre junto ao cenário real estudado.
3. Falso Negativo: Ocorre quando as predições preveem que não ocorrerá o acontecimento, porém, de fato, ocorre junto ao cenário real estudado.

4. Verdadeiro Negativo: Ocorre quando as predições preveem que não ocorrerá o acontecimento, e, de fato, não ocorre junto ao cenário real estudado.

		Valor Verdadeiro (confirmado pela análise)	
		Positivos	Negativos
Valor Previsto (predito pelo teste)	Positivos	VP Verdadeiro Positivo	FP Falso Positivo
	Negativos	FN Falso Negativo	VN Verdadeiro Negativo

Figura 4 – Matriz de Confusão

Diante disso, é possível calcular duas métricas fundamentais, sendo elas:

$$\text{Precisão} = \frac{\text{Verdadeiro Positivo}}{\text{Verdadeiro Positivo} + \text{Falso Positivo}} \quad (1)$$

$$\text{Revocação} = \frac{\text{Verdadeiro Positivo}}{\text{Verdadeiro Positivo} + \text{Falso Negativo}} \quad (2)$$

No tocante à interpretabilidade dos resultados, como já mencionado, o *XGBoost* demonstra dificuldade em entender, a razão das decisões do modelo, o que suscitou a aplicação do *LIME* (*Local Interpretable Model-Agnostic Explanation*). De acordo com *Ribeiro, M. T. et al., 2016*, trata-se de um algoritmo capaz de proporcionar explicabilidade às variáveis utilizadas no modelo.

Pode-se entender a técnica como responsável pela perturbação das variáveis dispostas como entrada, medindo a variação do resultado do modelo em função da alteração de cada uma dessas variáveis. Dessa maneira, é possível verificar o nível de sensibilidade de cada uma das variáveis no modelo, e, conseqüentemente, qual o impacto/representatividade de cada uma delas, diante do resultado alcançado. Ou seja,



é possível, de certa maneira, entender o porquê dos resultados gerados pelo modelo preditivo (interpretabilidade).

APLICAÇÃO

Como exposto, foi realizado um modelo preditivo no trecho próximo ao porto de São Francisco, no estado do Paraná, englobando uma extensão ferroviária de aproximadamente duzentos quilômetros.

Neste cenário, foram avaliadas 651 viagens realizadas entre os anos de 2018 e 2019, utilizando-se de 104 variáveis extraídas (entre via permanente e material rodante), resultando em um *dataset* de dimensão de 651 viagens x 108 variáveis (104 extraídas e 4 integradoras).

As Figuras 5 e 6, apresentam um exemplo do conjunto de variáveis aplicadas ao modelo desenvolvido e suas respectivas descrições. É válido lembrar, que houve uma preocupação em encontrar variáveis além daquelas já existentes (variáveis que já estavam prontas a serem extraídas dos sistemas), através de uma ótica dentre as próprias variáveis, obtendo por exemplo, valor mínimo, máximo, média *etc.* Além disso, buscou-se eventos ferroviários específicos, que agrupam mais de uma variável, como é o caso de eventos como o empeno, da velocidade máxima permitida, do desbalanceamento *etc.*

Nome das Features	Descrição	Descrição
os	Chave	W Ordem de Serviço
equipamento	Chave	ID do Equipamento
data	Chave	Data (do acidente ou do trem em data aleatória)
niv_val_max,min,median,std	Via	Nivelamento direito
niv_val_min,max,min,median,std	Via	Nivelamento esquerdo
niv_val_max,min,median,std	Via	Nivelamento transversal
ali_val_max,min,median,std	Via	Alinhamento direito
ali_val_min,max,min,median,std	Via	Alinhamento esquerdo
bi_val_max,min,median,std	Via	Valor Bitola
empino_max,min,median,std	Via	Empino
velocidade	Via	Velocidade no momento da data (Com ou sem restrição)
acidente	Acidente	Se houve acidente - Variável alvo (0 ou 1)
desbal_diff_vazios	Via/Vagão	Diferença máxima do desbalanceamento de um vagon vazio da composição com o limiar de 15%
accel_diff_vazios	Via/Vagão	Diferença máxima do aceleração transversal de um vagon vazio da composição com o limiar de 0.423m/s²
desbal_diff_carreg	Via/Vagão	Diferença máxima do desbalanceamento de um vagon carregado da composição com o limiar de 15%
accel_diff_carreg	Via/Vagão	Diferença máxima do aceleração transversal de um vagon carregado da composição com o limiar de 0.423m/s²
vel_maxima	Via	Velocidade máxima projetada
cs	Via	Flag se é curva (0 ou 1)
raio	Via	Raio do equipamento
rampa	Via	Rampa do equipamento
peso_geometrico	Via	Peso Geométrico do equipamento
suspencao_diff	Via	Diferença máxima da super-elevação real e projetada
restricao_T	Via	Flag restrição temperatura
restricao_V	Via	Flag restrição Via
num_HFD	Vagão	Quantidade de vagões HFD da composição
num_HFE	Vagão	Quantidade de vagões HFE da composição

Figura 5 - Variáveis consideradas – Parte 1

Nome das Features	Descrição	Descrição
num_HFE	Vagão	Quantidade de vagões HFE da composição
num_HPE	Vagão	Quantidade de vagões HPE da composição
num_vazios	Vagão	Quantidade de vagões vazios na composição
num_carreg	Vagão	Quantidade de vagões carregados na composição
mas_cinco_domingos	Via	Frequência de 5 ou mais domingos inerevis
num_locom_frente	Vagão	Número de locomotivas na frente da composição
num_locom_back	Vagão	Número de locomotivas atrás da composição
tamanho_composicao	Vagão	Tamanho da composição
kms_ultima_revisao_geral_max,min,median,std	Vagão	Kms desde a última revisão geral
kms_ultima_revisao_regul_max,min,median,std	Vagão	Kms desde a última revisão regul
dias_ultima_revisao_geral_max,min,median,std	Vagão	Dias desde a última revisão geral
dias_ultima_revisao_regul_min,median,std	Vagão	Dias desde a última revisão regul
num_vagoes_sem_rg	Vagão	Número de vagões sem revisão geral
num_vagoes_sem_rf	Vagão	Número de vagões sem revisão regul
num_manutencoes_max,min,median,std_acoplament	Vagão	Número de manutenções de acoplamento
num_manutencoes_max,min,median,std_freio	Vagão	Número de manutenções de freio
num_manutencoes_max,min,median,std_rodeteiro	Vagão	Número de manutenções de rodeteiro
num_manutencoes_max,min,median,std_sist_mecani	Vagão	Número de manutenções de sistema mecânico
num_manutencoes_max,min,median,std_truque	Vagão	Número de manutenções de truque
num_aviarias_max,mean_acidente	Vagão	Número de avarias de acidente registradas
num_aviarias_max,mean_descarrilado	Vagão	Número de avarias de descarrilamento registradas
num_aviarias_max,mean_freio	Vagão	Número de avarias de freio registradas
num_aviarias_max,mean_retencao_incorreta	Vagão	Número de retenções incorretas registradas
num_aviarias_max,mean_revisao	Vagão	Número de revisões registradas
num_aviarias_max,mean_rodeteiro	Vagão	Número de avarias de rodeteiro registradas
num_aviarias_max,mean_truque	Vagão	Número de avarias de truque registradas
velocidade_marco_diff	Via/Vagão	Diferença da velocidade máxima calculada para a velocidade atual da via. Número negativo está acima da velocidade máxima calculada
num_vagoes_acima_vel_max	Via/Vagão	Número de vagões com velocidade acima da velocidade máxima calculada para o vagon

Figura 6 - Variáveis consideradas – Parte 2

Em relação ao *overfitting* do modelo, foram separados 80% dos dados para treinamento (520 viagens) e 20% para validação (131 viagens), conforme discorrido anteriormente na técnica de *k-fold*.

No que se refere à existência de classes minoritárias, dada a singela quantidade de dados para a análise em questão (651 viagens de maneira aleatória), onde apenas 18 deles, seriam a variável alvo do modelo, (acidente), foi realizado um *oversampling* através da técnica *SMOTE (Synthetic Minority Oversampling TEchnique)*, discorrido anteriormente.

Portanto, na Tabela abaixo, nota-se que para os dados direcionados para treino, foram criados 492 acidentes sintéticos (totalizando o mesmo número de não acidentes), não mais existindo a divergência numérica entre classes minoritárias e majoritárias. É válido lembrar que as métricas de assertividade são baseadas no conjunto de validação, portanto, não sendo influenciadas pela criação de acidentes sintéticos.

Descrição	Dataset	Divisão de Treino (80%)	Oversampling	Divisão de Teste (20%)
Não Acidentes	633	506	506	127
Acidentes	18	14	506*	4
Total	651	520	1.012	131

* Sendo 492 sintéticos e 14 dados históricos

Tabela 1 - Oversampling

ANÁLISE DOS RESULTADOS

No caso de predição de acidentes ferroviários, a “Precisão” demonstra que quanto maior o seu valor, maior será a eficiência operacional na liberação/saída dos trens para execução de seus trajetos. Isto porque, quanto menor a quantidade de “Falso Positivo” (entenda-se a parada de trens sem a real necessidade, uma vez que o modelo interpretou como risco de acidente, mas no conjunto real, não houve), menor a quantidade de trens que serão interrompidos de maneira errônea.

Por outro lado, em relação à “Revocação”, quanto maior o seu valor, maior será a quantidade correta de predição de acidentes ferroviários. Isso porque, quanto menor a quantidade de “Falso Negativo” (entenda-se menor quantidade de predição de não acidente, mas que de fato ocorreu), menor a quantidade de trens que poderão ser identificados de maneira errônea como sem risco de acidente.

A Figura 8, mais a frente, apresenta os resultados obtidos em uma das rodadas de testes, avaliando dados históricos de acidentes.

Como demonstrado, observa-se que foram analisadas 651 viagens, dentre os quais, 80% ou 520 viagens, foram utilizadas para treinamento do modelo e 20% ou 131 trens serviram de validação. Dentre os dados de validação, historicamente, é possível afirmar que apenas 4 dentre os 131 trens foram envolvidos em acidentes. Diante dos resultados obtidos em uma das rodadas de testes, é possível verificar que houveram 2 “Verdadeiro Positivo” e 1 “Falso Positivo”. Em nosso melhor modelo, os resultados obtidos foram de captura média de 40% dos acidentes e precisão média de 70% nos alarmes.

Através da técnica *LIME*, foi possível entender quais variáveis foram responsáveis em maior ou menor grau na decisão do modelo na classificação entre acidente e não acidente ferroviário. Na Figura 9, mais a frente, é possível observar o resultado de uma análise específica utilizando a técnica mencionada para o caso aqui estudado, no qual o modelo

classificou corretamente como acidente ferroviário – probabilidade de predição de 27% para não acidente e de 73% para acidente.

Nota-se que para este exemplo em específico (levando-se em consideração, por exemplo, a localidade, o período de análise, os dados históricos estudados e as composições trafegadas), as variáveis em cor laranja à direita da Figura 9, são aquelas que impulsionaram o modelo na resposta afirmativa de possibilidade de um acidente ferroviário – 73%, entre elas: máximo de dias desde a última revisão geral e peso geométrico do equipamento. Por outro lado, as variáveis em cor azul, são aquelas que direcionaram o modelo na resposta negativa de possibilidade de um acidente ferroviário – 27%, entre elas: desvio padrão do nivelamento transversal e máximo empeno.

CONCLUSÕES

O trabalho apresentou uma metodologia para desenvolvimento de um algoritmo de *Machine Learning* capaz de prever similaridade de novas viagens ferroviárias em relação a acidentes ferroviários históricos registrados pela Companhia, através, sobretudo, de dados históricos de via permanente e material rodante. Apresentou também, os resultados obtidos pela sua aplicação em dados históricos de um determinado trecho da ferrovia da Rumo Logística.

O modelo preditivo desenvolvido, apresentou resultado promissor no que tange tanto à predição quanto à explicabilidade de dados. Em nosso melhor modelo, os resultados obtidos foram de captura média de 40% dos acidentes e precisão média de 70% nos alarmes.

O melhor cenário que poderíamos observar é o conjunto de informações sendo analisada vagão a vagão, localidade a localidade. Entretanto, por questões sistêmicas dos dados utilizados, não foi possível alcançar este nível de detalhe, sendo assim, os resultados do estudo referem-se à predição de acidentes com um trem dentro de um intervalo de espaço e tempo (neste caso,

cerca de 200 quilômetros, a partir da formação de um trem com origem e destino já determinados).

Como próximos passos, o algoritmo está em fase de expansão para outras localidades da malha da Rumo, porém, ainda está em desenvolvimento, a precisão de vagão a vagão, no menor intervalo de espaço possível, bem como, diferentes técnicas de explicabilidade e abordagens de modelo (como exemplo, *random forest*, rede neural, regressão logística e programação genética).

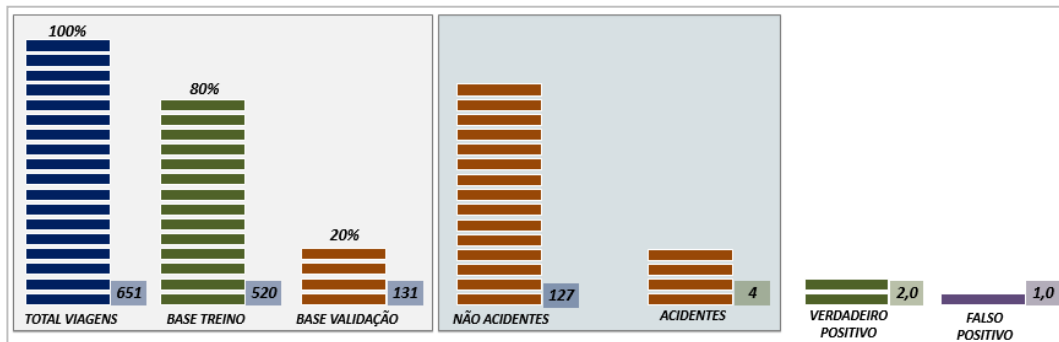


Figura 8 - Resultado do modelo preditivo de acidentes ferroviários

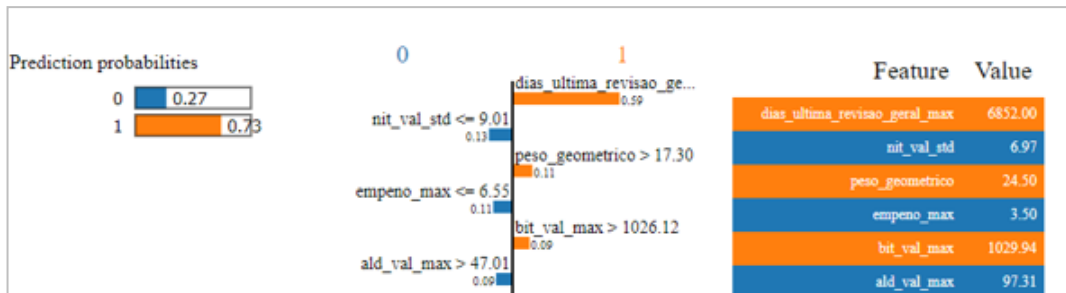


Figura 9 - Exemplo adaptado de interpretabilidade do modelo (técnica LIME)

REFERÊNCIAS BIBLIOGRÁFICAS

ANTT. Acidentes Ferroviários Ocorridos no Subsistema Ferroviário Federal no ano de 2013. Disponível em: <http://www.antt.gov.br/backend/galeria/arquivos/versao_fiscal_relatorio_01201gerofsufer_1.pdf>. Acesso em: 10/6/2020.

BANFIELD, R.; LOMBARDO, C. T.; WAX, T. Design Sprint: A Practical Guidebook for Building Great Digital Products. Sebastopol, CA, Estados Unidos: O'Reilly, 2015.

CHEN, T.; GUESTRIN, C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Anais . p.785–794, 2016. New York, NY, USA: ACM. Disponível em: <<https://dl.acm.org/doi/10.1145/2939672.2939785>>.

Chawla, N.V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. Journal Of Artificial Intelligence Research, Volume 16, pages 321-357, 2002.

Scikit Learn. Cross-validation: evaluating estimator performance. Disponível em: https://scikit-learn.org/stable/modules/cross_validation.html. Acesso em: 22 jun. 2020.

Ribeiro, M.T., Singh, S., Guestrin, C. Why should I trust you? Explaining the Predictions of Any Classifier. Disponível em: <https://arxiv.org/abs/1602.04938>. Acesso em: 22 jun.2020.