

ANÁLISE DE DESVIO DE CONDUÇÃO PARA VEÍCULOS DE MINERAÇÃO À CÉU ABERTO: UMA ANÁLISE EXPERIMENTAL BASEADA EM ALGORITMOS DE RECONHECIMENTO DE PADRÕES

DRIVING DEVIATION ANALYSIS FOR OPEN SKY MINING VEHICLES: AN EXPERIMENTAL ANALYSIS BASED ON PATTERN RECOGNITION ALGORITHMS

Danilo Saviotti Stangherlin Oliveira* E-mail: danilosaviotti@hotmail.com

Lucas Roberto Peregrino* E-mail: peregri_o903@gmail.com

Marcos Jordan Cota* E-mail: marcosjordancota@gmail.com

Renata Duarte Mellim* E-mail: renatamellim16@gmail.com

Diva de Souza e Silva Rodrigues* E-mail: divasouz@gmail.com

Flávio Henrique Batista de Souza* E-mail: flabasouza@yahoo.com.br

*Centro Universitário de Belo Horizonte (UNIBH), Belo Horizonte, MG, Brasil.

Resumo: Um dos setores mais relevantes da Indústria, a saber a mineração à céu aberto, tem uma das maiores demandas de atuação de processos otimizados por tecnologias para processos inteligentes. Durante o processo de mineração, alguns riscos são inerentes às atividades, o que pode condizer com uma real apreensão sobre os cuidados que os recursos humanos estão expostos. Assim, esta pesquisa foca no processo de condução veicular, com uma proposta de análise preditiva, com suporte de algoritmos de reconhecimento de padrões, para predizer possíveis desvios de condução para sua identificação e possibilitar tratativas preventivas, o que representa um investimento à proteção da vida dos condutores envolvidos. Pode-se encontrar estudos sobre o tema, que focam em análises preditivas do comportamento do motorista ou do veículo, mas o estudo proposto contribui com uma análise do processo produtivo. Durante os experimentos, algoritmos de reconhecimento de padrões baseados em Multilayer Perceptron (MLP), com uma variação de tipos e parâmetros com base na literatura, foram utilizados pra mensurar a capacidade de predizer o efeito Fadiga dos demais efeitos demonstrados em uma mineradora em operação. O processo contou com uma avaliação estatística tanto da base de dados fornecida, quanto os experimentos via MLP, com capacidades preditivas que chegaram a uma acurácia (baseada em AUC – Area Under the Curve, que se apresenta em valores entre 0 e 1) de 0.9742.

Palavras-chave: Desvio de Condução. Análise de Dado. Multilayer Perceptron. Fadiga.

Abstract: One of the most relevant sectors of the Industry, namely open pit mining, has one of the greatest demands on the performance of processes optimized by technologies for intelligent processes. During the mining process, some risks are inherent to the activities, which can be consistent with a real apprehension about the care that human resources are exposed to. Thus, this research focuses on the vehicle driving process, with a proposal for predictive analysis, supported by pattern recognition algorithms, to predict possible driving deviations for their identification and enable preventive treatments, which represents an investment in the protection of life. of the drivers involved. Studies on the subject can be found, which focus on predictive analyzes of driver or vehicle behavior, but the proposed research contributes to an analysis of the production process. During the experiments, pattern recognition algorithms based on Multilayer Perceptron (MLP), with a variation of types and parameters based on the literature, were used to measure the ability to predict the Fatigue effect of the other effects demonstrated in a mining company in operation. The process had a statistical evaluation of both the provided database and the experiments via MLP, with predictive capabilities that reached an accuracy (based on AUC - Area Under the Curve, which is presented in values between 0 and 1) of 0.9742.

Keywords: Driving Deviation. Data Analysis. Multilayer Perceptron. Fatigue.

1 INTRODUÇÃO

Uma análise estratégica através da gestão de não conformidades é uma das questões mais importantes em todo e qualquer tipo de negócio de risco. Dentre eles, a mineração a céu aberto pode ser referenciada. Assim, essa pesquisa demonstra um processo de avaliação baseado em Data Analytics, em conexão com reconhecimento de padrões, para avaliação de desvios de condução.

A mineração é uma atividade que consiste nos processos de pesquisa, exploração, extração e beneficiamento de minérios. Segundo Lamoso (2017), no território brasileiro, há uma distribuição geográfica diversificada, além de uma farta ocorrência de jazidas minerais, como ferro, caulim, bauxita, cobre e ouro, entre as quais se destacam as jazidas do Quadrilátero Ferrífero de Minas Gerais, há regiões nas quais a exploração mineral assume um protagonismo na produção do espaço econômico porque se torna uma atividade que mobiliza a força de trabalho, geração de renda, demanda pela implantação de infraestrutura material e imaterial, forma redes de poder político e econômico, cria e recria fluxos migratórios, gerando consumo de energia e recursos naturais, estabelece uma nova ordem na logística e nos preços das mercadorias e serviços, além de adensar e acelerar a formação de núcleos urbanos que lhe oferecem suporte.

O processo produtivo de uma mina, que pode ser realizada tanto a céu aberto quanto subterrâneo (nesta pesquisa, foca-se na primeira modalidade), há desafios a serem enfrentados. Um exemplo é o desvio de condução, que é um evento gerado pelo sistema de monitoramento do condutor que tem o objetivo de avaliar e monitorar o perfil do condutor durante a condução.

De acordo com Paranhos (2019), a sonolência durante a condução é uma das maiores causas de acidentes no Brasil. Cerca de 20% dos acidentes que ocorrem em estradas brasileiras são causadas por motoristas que dormiram ao volante. Situações deste tipo são cada vez mais frequentes, segundo dados da Associação Brasileira de Medicina de tráfego (ABRAMET), em pesquisa realizada com 495 pessoas no ano de 2017, identificou que cerca de 40% dos entrevistados já vivenciou situações de perigo por dirigir com sono.

Paralelo a este cenário, tecnologias da informação não somente evoluíram, mas permearam praticamente todos os cenários produtivos contemporâneos. Diante do acúmulo da informação, do desenvolvimento da capacidade computacional e da Indústria 4.0, as companhias têm enxergado com grande valor respostas e resultados nas análises de seus dados. Dentre as metodologias e técnicas de análise de dados, há o Data Analytics.

De acordo com Machado (2018), Malik *et al.* (2021) e Moreira *et al.* (2018), Data Analytics é o processo de analisar dados variados com um objetivo claro de encontrar padrões específico trazendo maior clareza em relação aos esforços. Seu foco é encontrar respostas com base nos dados, com a finalidade de solucionar gargalos que passam despercebidos no processo organizacional, auxiliando assim na tomada de decisões.

Além do Data Analytics, uma outra tecnologia tem auxiliado consideravelmente na utilização de dados de forma estratégica, que é a técnica de Machine Learning baseada em algoritmos de reconhecimento de padrões. De acordo com Campos *et al.* (2020), o processo de reconhecimento de padrões se define como uma tarefa de classificação ou categorização, com o acúmulo de informações catalogadas, uma rede neural, consegue “aprender” com base em um determinado padrão e antecipar quando o padrão pode ou vai acontecer. Além do papel importante, que é a prevenção, a tecnologia possibilita também uma tomada de decisão mais precisa e rápida quando diante de algum risco.

Em circunstâncias de emergência, caso a empresa tenha esse tipo de tecnologia implantada, possibilitaria ter uma resposta rápida e pronto atendimento, pois seria possível precisar em tempo real, quantas pessoas estavam no local, número de vítimas e viabilizar a tomada de decisões para um socorro mais eficiente. Pesquisadores do Laboratório de Ciência da Computação e Inteligência Artificial do MIT (Massachusetts Institute of Technology) e do Centro de Inteligência Artificial do Qatar, criaram um modelo capaz de aprender através de reconhecimento de padrões ocorridos no passado e prever acidentes de trânsito antes que eles aconteçam (MIT NEWS, 2021).

Assim, o objetivo geral desse trabalho é apresentar o desenvolvimento de uma análise experimental, baseada em uma base de dados de uma operação real de mineração a céu aberto de minas gerais, afim de proporcionar uma avaliação analítica dos desvios encontrados, além de um estudo preditivo baseado em variáveis

presentes no processo produtivo. Para tanto, como objetivos específicos, visa-se: realizar o processo de coleta da base de dados; desenvolver um processo de análise dos principais desvios de condução encontrados e; realizar uma análise experimental primária para avaliação do comportamento da base dados para o algoritmo de reconhecimento de padrões a saber o MLP (Multilayer Perceptron).

Esta pesquisa se justifica não só pelo o seu grau de risco que pode envolver a preservação da vida das pessoas visto que acidentes são comuns nesse ambiente e mudar a mentalidade sobre a tecnologia visto que a meta desse projeto que concerne na preservação da vida. Outro fator é baseado nos bilhões de dólares que podem encontrar nesse tipo de negócio. Segundo o Instituto Brasileiro de Mineração (IBRAM) a produção mineral brasileira em 2021, em toneladas, cresceu cerca de 7% em relação a 2020, passando de 1,073 bilhão de toneladas para 1,150 bilhão de toneladas estimadas. Em 2021, a variação de preços das commodities no mercado internacional embalou o faturamento do setor em 62%, na comparação com 2020, crescendo de R\$ 209 bilhões para R\$ 339 bilhões. De acordo com dados do IBRAM, em 2021, o estado de Minas Gerais apresentou o seu maior crescimento no faturamento: 87%, passando de 76,4 bilhões de reais em 2020 para 143 bilhões de reais. Com este resultado, MG corresponde por 42% do faturamento global da indústria da mineração brasileira no ano 2021 — esta participação era de apenas 37% em 2020.

2 REFERENCIAL TEÓRICO

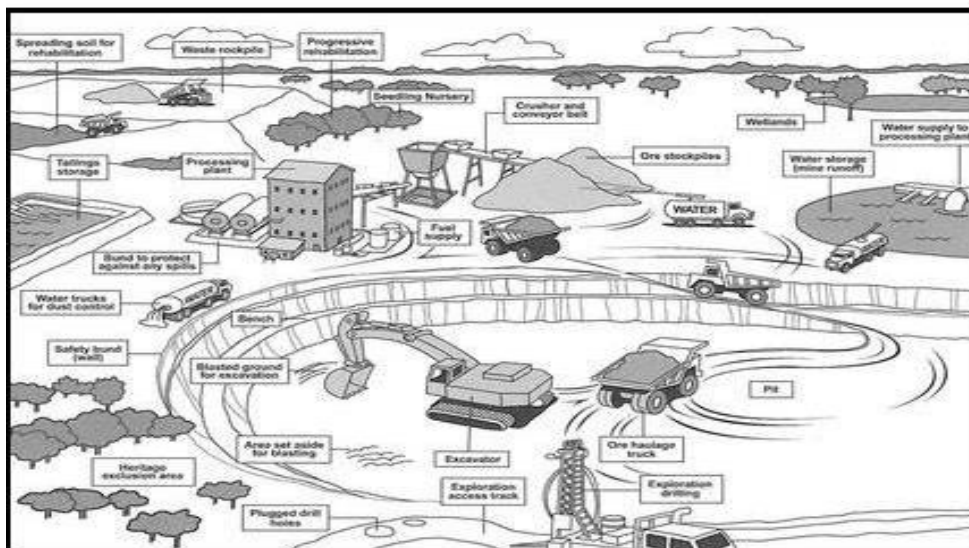
2.1 Mineração a céu aberto

Segundo Souza (2019), mineração a céu aberto é uma operação de mina em superfície pelos quais o minério, ou resíduo, é escavado a partir da superfície da terra. No processo, um poço cada vez mais profundo é formado até que se encerre a operação de mina, o minério de ferro é encontrado na natureza na forma de rochas, misturado a outros elementos. Por meio de diversos processos industriais com tecnologia de ponta, o minério é beneficiado para, posteriormente, ser vendido para as indústrias siderúrgicas e conseqüentemente ser utilizado na construção de casas, na fabricação de carros e na produção de eletrodomésticos.

De acordo com Baima (2018), para se dar o início ao processo de extração do recurso natural, primeiramente, faz-se necessário o desmatamento e a remoção do solo para tornar viável o acesso e a operação da lavra, o que provoca um efeito negativo na fauna e flora, além do impacto visual.

A Figura 1 demonstra uma exemplificação do processo produtivo de uma mina a céu aberto.

Figura 1 - Descrição processo de lavra de mineradora a céu aberto



Fonte: Adaptado de Souza (2019).

O Processo Produtivo resumido pode ser descrito como (SOUZA, 2019):

1° fase – Lavra: A primeira fase da mineração é a extração do minério bruto, que pode ser feita utilizando escavadeiras e tratores que raspam a rocha ou com uso de explosivos, quando o minério se encontra longe da superfície. Uma parte do que é retirado do solo é minério, o restante é estéril, material disposto em pilhas específicas.

2° fase – Transporte: As escavadeiras ou tratores carregam os caminhões fora-de-estrada para levar o minério até a britagem.

3° fase – Britagem: Os caminhões descarregam o minério bruto na britagem, que são quebrados na máquina de britagem até eles ficarem com cerca de 2 centímetros de diâmetro, o tamanho adequado para a separação.

4° fase – Correias Transportadoras: Depois do processo de britagem, o minério é transferido por correias transportadoras até a usina de beneficiamento.

5° fase – Usina: Ao chegar na usina, o minério passa por vários processos como: peneiramento, moagem, separação magnética, filtragem, flotação.

6° fase – Estocagem: Depois de processado, o minério segue pelas correias para ser estocado em pátios. As empilhadeiras e recuperadoras armazenam o minério em pilhas.

7° fase – Recuperação de Produtos: As pilhas de minério são movimentadas por máquinas recuperadoras e uma recuperadora-empilhadeira. Elas transportam o minério até os silos, onde é descarregado nos vagões dos trens de carga.

8° fase – Carregamento de Trens: Os trens de carga passam por baixo dos silos, para receber o minério, este processo leva em média duas horas e meia. Depois de carregado, o trem inicia o percurso para o porto.

9° fase – Viradores de vagões: Ao chegar no porto o trem é descarregado por meio dos viradores. O equipamento vira os vagões a 180 graus e descarrega os minérios em silos, sendo depois levados aos pátios de estocagem dos portos.

10° fase – Embarque em navios: O minério estocado no pátio do porto é colocado em esteiras e levado diretamente para os porões dos navios. Após o embarque, o minério é transportado para diversos países do mundo.

Segundo Meller (2018), a mineração está presente no dia a dia da população, atuando como base na maioria dos segmentos industriais, desempenhando um papel considerável na economia do país, tanto como geradora de empregos e impostos, como no desenvolvimento de cidades e regiões.

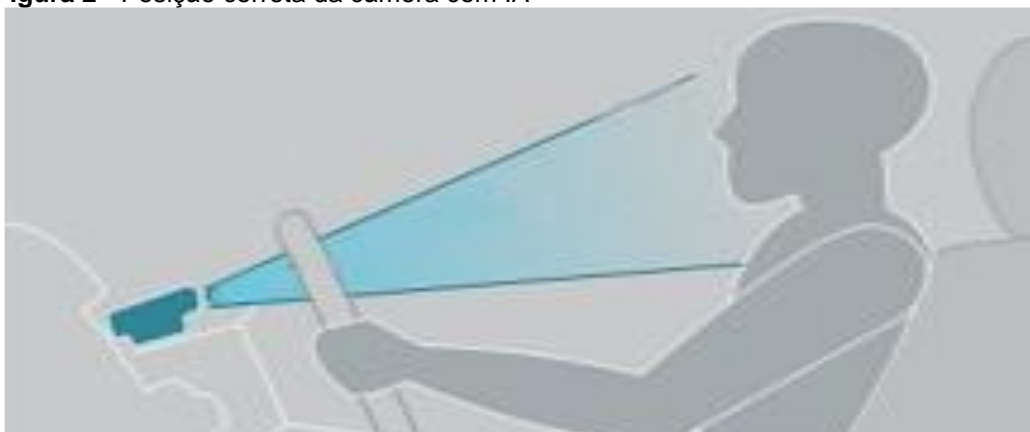
2.2 Desvios de condução

De acordo com Paranhos (2019) o avanço tecnológico na indústria automotiva faz com que os sistemas de auxílio ao condutor estejam cada vez mais presentes nos veículos atuais. Estes tipos de dispositivos visam garantir a segurança do usuário, buscando uma melhor sinergia entre homem e máquina em tempos em que as estradas são o maior causador de mortes no país.

Sobre esta prerrogativa, foi desenvolvido um sistema de monitoramento do condutor que funciona através de uma câmera com IA (Inteligência artificial) integrada, responsável por filmar e avaliar o perfil do condutor durante a condução, gerando e classificando automaticamente detecções de eventos de acordo com cada situação, através de evidências multimídia. Também possui tecnologia infravermelho, o que proporciona filmagens sem distorções de cores e imagens em alta qualidade.

A câmera, geralmente é instalada no painel do veículo, com ângulo frontal ao condutor como mostra a Figura 2.

Figura 2 - Posição correta da câmera com IA



Fonte: Arquivos Mineradora.

Como referência há as seguintes detecções de desvios (PARANHOS, 2019):

- **Detecção de Sonolência/Fadiga:** após o contato ON do veículo, inicia-se o monitoramento do perfil do condutor através da IA da câmera, avaliando os sinais de sonolência/fadiga durante todo o trajeto. Se detectada alterações, a câmera notifica o condutor através de alerta de voz e gera uma evidência multimídia que será enviada para a plataforma do sistema. Principais gatilhos: Fechamento dos olhos; Distanciamento das pálpebras; Falta de controle da parte superior do corpo; Cabeça pendendo.
- **Detecção de Cansaço:** Após o contato ON do veículo, inicia-se o monitoramento do condutor através da câmera. Se o condutor bocejar durante o trajeto, será notificado através de alerta de voz, gerando também uma evidência multimídia que será enviada para a plataforma do sistema. Principal gatilho: Bocejar durante a condução ou com o veículo ligado.
- **Detecção de uso de telefone:** A detecção de uso do telefone será registrada pela câmera após o contato ON do veículo, sempre que o condutor estiver com um objeto na mão, próximo ao rosto e/ou falando por ele, podendo ser rádio de comunicação ou telefone celular. Ao ser detectado o desvio, a câmera notifica o condutor através de alerta de voz e gera uma evidência multimídia que será enviada para a plataforma do sistema. Principais gatilhos: estar com aparelho de comunicação externa na mão, posicionado próximo ao rosto; falar ou utilizar o aparelho durante a condução do veículo e/ou o mesmo estando ligado.
- **Detecção de Distração:** A detecção de distração será registrada pela câmera após contato ON do veículo, sempre que o condutor desviar a

atenção da condução, olhando para qualquer direção com angulação para baixo ou para cima. Principal gatilho: olhar em outra direção se não a frontal para a via, com angulação para baixo ou para cima;

- **Motorista não visível:** A detecção de motorista não visível será registrada pela câmera após o contato ON do veículo, sempre que a mesma for movimentada ou obstruída, impedindo-a de focar a face do condutor. Ao ser detectado o desvio, a câmera notifica o condutor através de alerta de voz e gera uma evidência multimídia que será enviada para a plataforma do sistema. Principais gatilhos: a câmara com o ângulo desfocada da face do condutor durante a condução do veículo; a câmera for obstruída.

Recentemente reportado pelo G1 TV Globo (2022), o motorista do carro de transporte por aplicativo que levava o ex-BBB, admitiu que dormiu ao volante antes de bater o automóvel na traseira do caminhão que trafegava em um acesso à Marginal Pinheiros, entre as pontes Cidade Universitária e Eusébio Matoso, na região do Butantã, Zona Oeste de São Paulo.

De acordo com Fuga (2019), pesquisa divulgada pela Organização Mundial da Saúde (OMS), contabilizou, em 2009 cerca de 1,3 milhões de mortes por acidente de trânsito em 182 países, sendo a primeira faixa etária de 15 a 29 anos de idade. Esses acidentes, atualmente, representam um custo por ano entre 1% e 3% do PIB de cada país.

O Brasil aparece em quinto entre os países recordistas em mortes no trânsito. As principais causas de acidentes no Brasil são velocidades incompatíveis com o uso da via e ingestão de álcool pelo condutor, estima-se que circulam no Brasil cerca de trinta milhões de veículos, enquanto em 1960 era aproximadamente quinhentos mil.

Em 1980 a taxa de mortalidade por acidente de trânsito era de 17 pessoas a cada 100 mil, já em 2011 houve acréscimo e a taxa chegou a 22,5. O mesmo estudo aponta que a taxa de internações no SUS por acidente de trânsito com carro caiu de 20.134 em 1998 para 20.134 em 2012 e que 67,9% dos óbitos por acidente de trânsito aconteceram no final de semana.

Em 2011, os gastos com tratamento pelo SUS atingiram R\$210,8 milhões, sendo que destes R\$25,1 relacionado a acidentes com carro. Também em 2011, o custo desses acidentes foi estimado em 10,6 bilhões e, desse valor, destaca-se 13,3% com atendimento hospitalar, 42,8% com perda de produção, 3,7% com processos judiciais e 2,4% com previdenciários dentre outros (Fuga, 2019).

A frota de veículos automotores no Brasil é crescente, alarmante e os acidentes de trânsito na mesma proporção, crescem e fazem vítimas. Os danos são desde o simples material para o conserto do veículo, às sequelas e o óbito.

Um dos objetivos do monitoramento de veículos é analisar o comportamento dos condutores durante a condução, para que seja possível detectar anomalias e prever comportamentos futuros. Na maioria das vezes, os veículos seguem trajetos definidos. O conhecimento desses padrões e o desenvolvimento de metodologias automáticas para extração de informações pertinentes aos deslocamentos são importantes, por exemplo, para ações de planejamento do sistema de transporte (SHINOHARA *et al.*, 2015).

Segundo Shinohara (2013), constatou-se que seus associados possuem sistemas de rastreamento que armazenam além do posicionamento, dados operacionais e de telemetria. Estes dados estão sendo utilizados para o gerenciamento de frotas, programas de manutenção de veículos, controle logístico e especialmente para a gestão de segurança. A utilização de equipamentos de posicionamento e de comunicação embarcados é comum em veículos comerciais de transporte de carga e, com o seu barateamento, está ocorrendo expansão para o setor de veículos particulares principalmente em aplicações de gerenciamento de segurança.

Segundo Shinohara *et al.* (2015) o entendimento do comportamento dos condutores durante o trajeto ao longo do tempo é uma das questões-chave na investigação dos padrões de comportamentos das pessoas. Sendo assim, o conhecimento desses padrões de condução é, ainda hoje, limitado pela disponibilidade de dados durante os deslocamentos.

2.3 Data Analytics

Tecnologias para análise de dados são desenvolvidas e podem ser aplicada para mitigar ou minimizar os problemas, como o Data Analytics. Juntamente com a quantidade sem precedentes de informações disponíveis em formato eletrônico através de sistemas, a análise de dados extrai informações relevantes entre a enorme massa de dados que são atualmente gerados (MIHAS, 2019).

A partir de 2008 esse conceito ganhou força por nascer a necessidade de segmentar, organizar e estruturar os dados. Através do poder dos softwares e hardwares, é possível colher informações de diferentes fontes na rapidez e volume

exigidos pelas condições de negócios contemporâneos. De acordo com Chaouchi et. Al (2019), três tipos básicos de abordagens de Data Analytics:

Análise Preditiva: quando a objetivo é antecipar os efeitos de uma decisão. Nesse tipo de análise, são organizados modelos estatísticos com base em regressão, pelos quais se colocam relações de causa e efeito. Desse jeito, uma empresa pode adiantar os resultados das suas atuações pela aplicação de fórmulas estatísticas.

Análise Prescritiva: quando a análise descritiva se debruça para entender o que está acontecendo agora. São muito os casos em que as empresas perdem dinheiro e recursos pelos temíveis gargalos no seu processo produtivo, muitos dos quais não ficam claros durante anos. Nesse caminho, essas análises podem ajudar com a resposta da pergunta “o que está acontecendo?”.

Análise Diagnóstica: serve para entender o contexto atual. A discordância, aqui, é que nela o propósito é esclarecer por que algo sucedeu e as razões para tal. Sendo assim, é indicada para companhias que já passaram por uma turbulência e que, no futuro, almejam ter formas para tratar com os mesmos problemas caso eles venham a se reproduzir.

Buscando entender os dados atuais para saber o que pode ou não ocorrer em um futuro próximo, oferece uma quantidade imensa de dados para que as mais variadas análises possam ser realizadas. Nesse contexto, o Big Data Analytics tem como principal objetivo processar dados de alto volume, velocidade, variedade e veracidade usando métodos inteligentes e técnicas computacionais (KAKHANI *et al.*, 2015).

Entretanto, a BDA refere-se às técnicas utilizadas para examinar, processar, descobrir e expor padrões ocultos subjacentes, relações interessantes e outras ideias relacionadas ao contexto (IQBAL *et al.*, 2020). Dados bem analisados podem efetivamente mudar a realidade de um negócio para melhor, com base em big data analytics que gestores apoiam suas decisões quando precisam mudar algo nas operações ou rotinas.

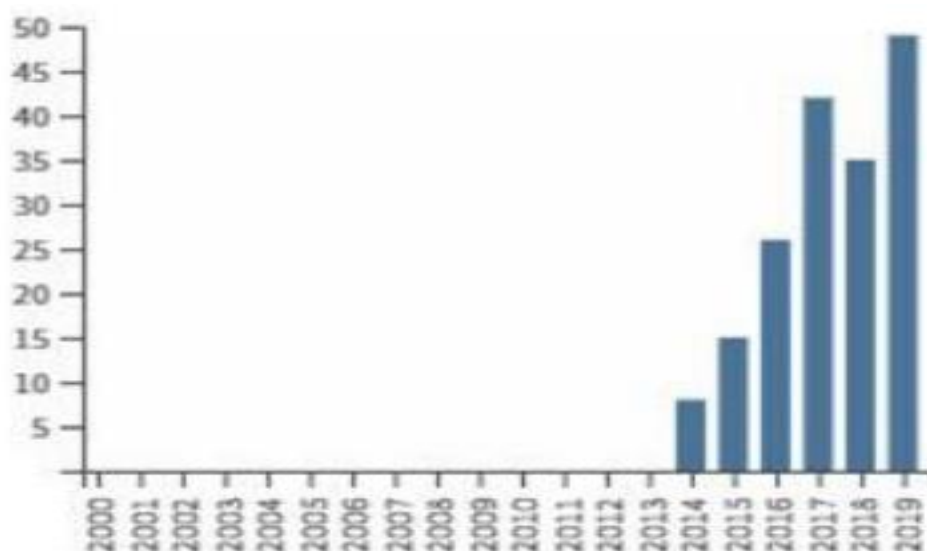
Após uma pesquisa constatou-se que o dado referente à artigo mais antigo presente na Web of Science é uma pesquisa de 2014 denominado “Big Data Analytics for Supply Chain Mangement”, publicado por Leveling.

No artigo, o autor trata soluções baseadas em Big Data dado ênfase na cadeia de suprimentos. Desde então, o total nos últimos 3 anos são responsáveis por 2594 citações, o que corresponde a 89% do total. A partir disso, é possível concluir a

relevância do tema e à proporção que tem tornado nos últimos anos e isso o torna um tema muito propício para o desenvolvimento de estudos.

Ao avaliar a Figura 3, é plausível o aumento deste tema no meio acadêmico, que a tendência é continuar crescendo. O crescimento é evidenciado pelo no último ano avaliado que em 2019, que ainda não chegou ao fim e já se tornou o ano com mais trabalhos publicados.

Figura 3 - Quantidade de citações referente ao tema



Fonte: Vale *et al.* (2019).

2.4 Reconhecimento de padrões

Segundo Campos (2000), o reconhecimento de padrões é uma área da ciência que tem como objetivo a separação de objetos dentro de um número de categorias ou classes. Esses objetos de estudo podem variar nos mais diversos campos de aplicações que vão desde imagens, sinais em forma de ondas (como voz, luz, rádio) ou qualquer tipo de medida que necessite ser classificada. Um sistema de reconhecimento de padrões é composto por um sensor que obtém observações a serem classificadas, um mecanismo de extração de que compila as informações das observações e um esquema de classificação das observações, que depende das características extraídas.

A forma como a é feita a classificação geralmente é baseada na disponibilidade do conjunto de padrões que foram previamente classificados, ou, o "conjunto de treinamento". O resultado do aprendizado é dito como um aprendizado supervisionado. O aprendizado pode também ser não supervisionado, de forma que

o sistema não recebe informações a dos padrões, estabelecendo então as classes dos padrões através de análise de padrões estatísticos.

De acordo com Campos *et al.* (2020) o Reconhecimento de padrões, consiste em uma tarefa de classificação ou categorização, onde as classes são definidas pelo projetista do sistema (classificação supervisionada) ou são “aprendidas” de acordo com a similaridade dos padrões (classificação não-supervisionada).

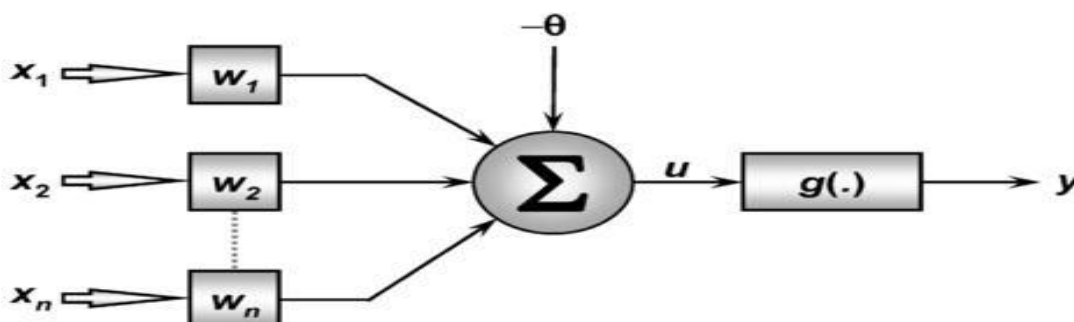
2.4.1 Multilayer Perceptron

Segundo Araújo *et al.* (2018) um Perceptron é a representação matemática de um neurônio humano conforme dado pelo modelo descrito na Figura 4. O Perceptron é uma formula matemática que descreve um neurônio humano, onde é definido como x_1 a x_n as entradas de informação.

Se em um neurônio humano os impulsos nervosos seriam a entrada de informação, em um perceptron são os valores que representam as entradas. Já os valores de w_1 a w_n são os pesos que essas entradas terão na função analisada e que posteriormente poderão ser ajustadas para obtenção de melhores resultados em nossa rede neural, como analogia ao neurônio os pesos são a representação matemática de nossas sinapses.

A representação visual Σ será o resultado do somatório de todas as entradas que culminara em nossa função de transferência, ou seja, o neurônio somente seria acionado se o valor obtido por nossa função for maior que o valor estipulado para o acionamento desse neurônio. Nas décadas de 60 e 70, acreditava-se que o uso de perceptrons para simulação de uma inteligência artificial seria inviável, tendo em vista que com um único neurônio não seria possível de resolver um problema simples como o da porta XOR.

Figura 4 - Modelo matemático de um neurônio

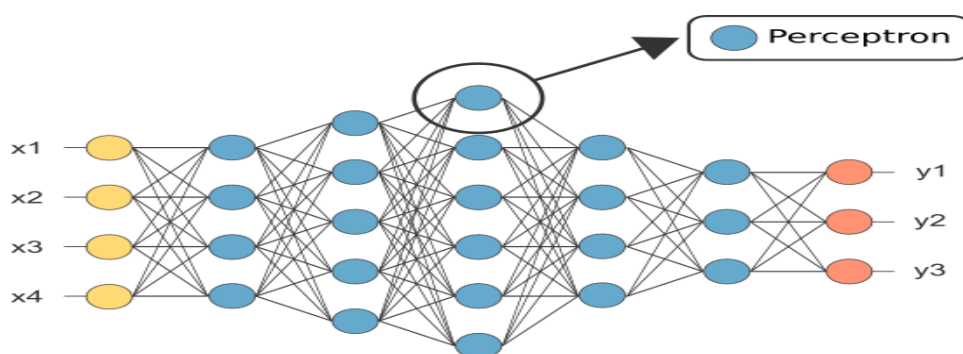


Fonte: Araújo *et al.* (2021).

Já na década de 80, Geoffrey Hinton, apresentou o conceito de Multilayer Perceptrons, onde ao invés de ter uma única rede de Perceptrons, passa-se a ter N camadas interconectadas de N neurônios, formando assim uma rede de multicamadas (Multilayer Perceptrons), possibilitando a resolução de diversos problemas computacionais.

Para que a rede multicamadas funcione, ela utiliza o algoritmo de backpropagation, onde os valores que são inseridos nos inputs layers (camada de entrada) da rede são propagados para as camadas subsequentes, a fim de se obter o resultado esperado no output layer (camada de saída), caso isso não ocorra, o algoritmo volta reajustando cada peso de forma aleatória a fim de ajustar a rede buscando uma melhor assertividade em seu resultado. A Figura 5 demonstra uma MLP.

Figura 5 - Representação de Multilayer Perceptrons



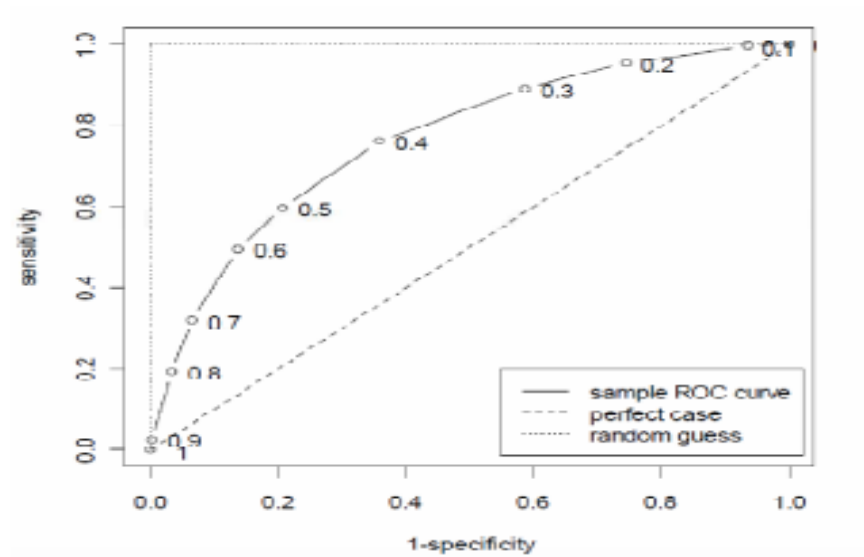
Fonte: Araújo *et al.* (2021).

2.5 Curva ROC e AUC

A curva ROC, é uma representação gráfica que ilustra o desempenho (ou performance) de um sistema classificador binário à medida que o seu limiar de discriminação varia. A curva ROC é também conhecida como curva de característica de operação relativa, porque o seu critério de mudança é resultado da operação de duas características (PV e PF).

A curva ROC é obtida pela representação de duas razões: $R_{PV} = \text{Positivos Verdadeiros} / \text{Positivos Totais}$ e $R_{PF} = \text{Positivos Falsos} / \text{Negativos Totais}$, para vários valores do limiar de classificação. A Figura 6 demonstra uma representação da Curva ROC.

Figura 6 - Curva ROC

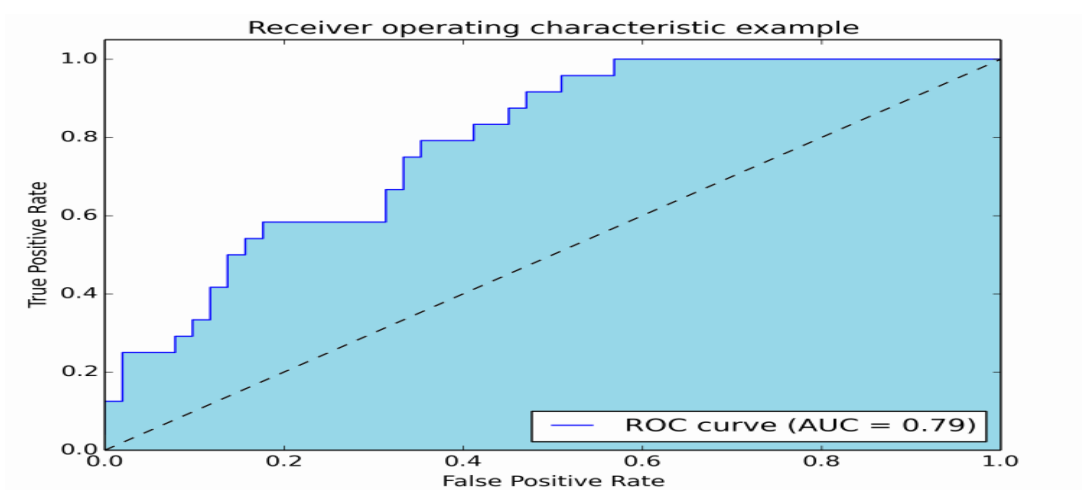


Fonte: Novais (2018).

Uma curva ROC traça RPV \times RPF em diferentes limiares de classificação conforme a figura 7 demonstra como exemplo.

Assim, na tentativa de simplificar a análise da ROC, a AUC (“Area Under the ROC Curve”) nada mais é que uma maneira de resumir a curva ROC em um único valor, agregando todos os limiares da ROC, calculando a “área sob a curva”.

Figura 7 - AUC



Fonte: Novais (2018).

2.6 Trabalhos Relacionados

Foram pesquisados trabalhos similares que utilizam sistemas de detecção de desvios através de sensores instalados nos veículos. Percebeu-se a aplicação de

vários métodos de medição e controle de resultados de forma a atuar somente na detecção, onde pode se observar semelhança entre eles (Tabela 1):

Tabela 1 - Trabalho relativos

Referência	Nome	Descrição	Diferença
Paranhos (2019)	Sistema de detecção de sonolência.	O artigo apresenta a proposta de um sistema de baixo custo para detecção de sonolência por meio de técnicas de processamento digital de imagens.	Sistema voltado apenas para detecção e não para prevenção.
Salles (2018)	Detecção automática de sonolência em condutores de veículos utilizando redes neurais artificiais.	O artigo propôs uma identificação de sonolência através de Redes Neurais Artificiais através de machine learning, obtendo 91,8% de acurácia.	Sistema dedicado apenas para detecção da sonolência em cenário diferente do proposto.
Berri (2019)	Sistema ADAS para identificação de distrações e perturbações do motorista na condução de veículos	Artigo apresenta um sistema que avalia informações provenientes de um sensor Kinect V2. que detecta situações perigosas com o objetivo de identificar quando o motorista está apto ou não para a direção.	Sistema que realiza detecção do situações de risco, que apenas tem atuação quando o risco acontece.
Porto et al (2018)	Gestão de fadiga de operadores de caminhões fora de estrada: estudo de caso em uma mina a céu aberto	Trabalho apresenta a efetividade da utilização do sistema de gestão de fadiga na frota de uma mineradora.	Sistema realiza apenas a detecção de fadiga quando a mesma acontece, atuando somente na correção e não na prevenção.

Fonte: Elaborados pelos Autores (2022).

3 METODOLOGIA

Esse trabalho propõe um estudo de caso. Segundo Gerring (2019), estudo de caso é um método de pesquisa científica que avalia um fenômeno em uma determinada situação e todas as variáveis que o influenciam. O estudo de caso será dividido em 3 etapas: (I) Apresentação da base de dados, (II) Avaliação do perfil da Base de Dados e (III) Experimentos com Algoritmos de Reconhecimento de Padrões com uma Discussão de possibilidades estratégicas.

Primeiramente, detalhou-se a base de dados utilizada, bem como quais as variáveis consideradas para o estudo de caso contidas na base além dos resultados esperados e obtidos.

As variáveis que serão consideradas da base de dados fornecida serão:

- **Evento:** Identifica o tipo do evento registrado pelo sensor do veículo (Detecção de distração, Detecção de Fadiga, Detecção de Cansaço, Motorista não visível e Detecção de uso do Telefone);
- **Data:** Registra a data em que o evento aconteceu;
- **Velocidade:** Identifica a velocidade em que o veículo estava trafegando no momento do registro do evento;
- **Direções:** Identifica o sentido em que o desvio ocorreu, podendo ser E – Leste, S – Sul, NE – Nordeste, SO – Sudoeste, O – Oeste, SE – Sudeste, N – Norte e NO – Noroeste;

A base de dados possui eventos registrados do dia 23 de março de 2021 até 25 de fevereiro de 2022, totalizando 8748 eventos registrados.

Por questões de compliance, os dados de identificação dos condutores e da mineradora a céu aberto de Minas Gerais foram classificados como confidenciais.

Através da base de dados foram definidos códigos para as variáveis utilizadas no processo de execução do algoritmo, conforme tabelas abaixo:

Tabela 2 - Códigos por tipo de evento

Eventos	Código
Detecção de Distração	1
Detecção de Fadiga	2
Detecção de Cansaço	3
Motorista não visível	4
Detecção de uso do telefone	5

Fonte: Elaborados pelos Autores (2022).

Tabela 3 - Códigos por direções

Código	Direções
1	E
2	N
3	NE
4	NO
5	O
6	S
7	SE
8	SO

Fonte: Elaborados pelos Autores (2022).

Após a definição dos códigos, foi dado o início a montagem da classificação dos dados que será usado para experimentação por reconhecimento de padrões, os dados foram lançados em uma planilha com 5 colunas sendo, Evento, Velocidade, Direções, Latitude, Longitude, Minutos e Rótulo (Label).

O primeiro processo foi a substituição das nomenclaturas dos eventos pelos códigos criados conforme tabela 2, em seguida foi realizada a substituição das nomenclaturas das direções pelos códigos criados conforme tabela 3.

Foi gerado um total de 4 base de dados sendo elas, detecção de fadiga com detecção de distração, detecção de fadiga com detecção de cansaço, detecção de fadiga com detecção de uso de telefone, detecção de fadiga com motorista não visível.

Como amostragem para cada experimento, foram definidos:

- 3674 eventos de detecção de fadiga contra 3661 eventos de detecção de distração,
- 3674 eventos de detecção de fadiga contra 806 eventos de detecção de cansaço,
- 3674 eventos de detecção de fadiga contra 323 eventos de detecção de uso de telefone,
- 3674 eventos de detecção de fadiga 284 eventos de detecção de motorista não visível.

As amostragens demonstradas já figuram um desafio, pois há casos de desproporções entre quantidades de eventos, qualificando alguns agrupamentos como desbalanceados.

Software R Studio só realiza a leitura dos dados através de códigos numéricos, e a coluna rótulo (Label) foi classificada com dados numéricos sendo 1 para o evento referência (Verdadeiro) que é a detecção de fadiga e 0 para o evento (Falso) que é a detecção dos demais eventos.

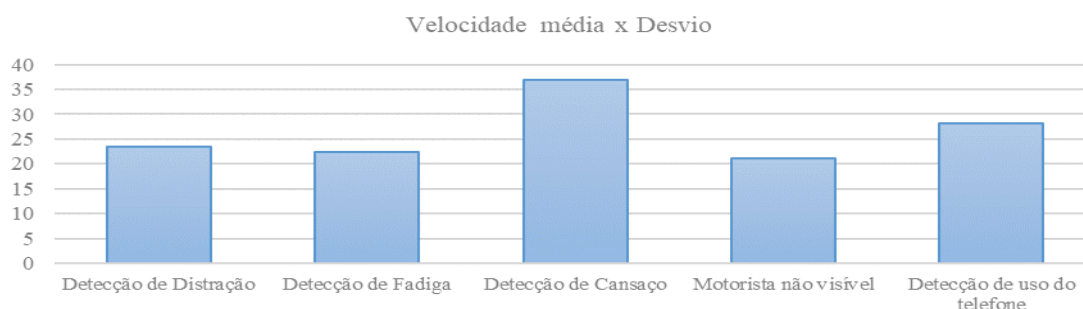
Para dar início a experimentação por reconhecimento de padrões, através do software R Studio, a planilha com a base de dados foi convertida em arquivo de Documento de Texto (txt), de forma que o software consiga suportar e reconhecer a base de dados.

4 RESULTADOS

4.1 Análise comportamental

Primeiramente foi considerada uma avaliação da velocidade média por desvio. Analisando as informações da base de dados, considerou-se para a primeira análise, as variáveis, Evento e Velocidade e realizou-se um cálculo a velocidade média em que os desvios eram detectados conforme disposto na figura 8.

Figura 8 - Gráfico de Velocidade Média por desvio



Fonte: Elaborados pelos Autores (2022).

É possível identificar que o desvio detectado com maior velocidade média é o de Detecção de Cansaço, que supera a velocidade média de 35 km/h.

Importante ressaltar que para essa análise, não foi considerado os desvios que a velocidade medida era igual a 0 (zero), isso se dá pois o principal motivo é que nessa velocidade não existe risco para o condutor além de outros fatores, onde o sensor pode estar captando um desvio incorreto, por exemplo, no momento em que o condutor está preenchendo uma ficha dentro do veículo com o mesmo ligado e olhando para baixo, o sensor acusa uma Detecção de Distração, contudo, o condutor está parado e fora de risco.

Em segundo lugar, foi desenvolvida uma avaliação de desvio por faixa de horário. Para essa análise, consideramos as variáveis Evento e Hora onde, Evento identifica o tipo de desvio detectado e a Hora representa o momento em que o desvio ocorreu, mas para critérios de melhorar a interpretação das informações, a variável Hora foi alterada seguindo o seguinte racional:

Foi considerado apenas a Hora em que o desvio ocorreu, desconsiderando o minuto, sendo assim, criamos uma faixa de horário para detecção dos desvios, se o desvio tivesse sido detectado as 11:38 por exemplo, seria considerado que ocorreu as 11:00, desconsiderando os minutos.

Após os ajustes citados acima, a tabela 4 foi construída representando a faixa de hora em que os desvios ocorrem ao longo do dia e quais os tipos de desvios ocorrem.

Tabela 4 - Desvio por faixa de horário

Faixa Horário	Detecção de Distração	Detecção de Fadiga	Detecção de Cansaço	Motorista não visível	Detecção de uso do telefone
00:00:00	0,011%	0,000%	0,000%	0,000%	0,000%
02:00:00	0,000%	0,000%	0,011%	0,000%	0,000%
03:00:00	0,103%	0,046%	0,011%	0,000%	0,000%
04:00:00	0,217%	0,023%	0,057%	0,000%	0,011%
05:00:00	2,641%	4,515%	0,389%	0,149%	0,103%
06:00:00	5,361%	5,498%	1,177%	0,412%	0,137%
07:00:00	3,041%	2,812%	0,320%	0,046%	0,091%
08:00:00	2,904%	2,824%	0,354%	0,274%	0,400%
09:00:00	3,875%	1,989%	0,503%	0,069%	0,286%
10:00:00	2,229%	2,229%	0,332%	0,149%	0,217%
11:00:00	2,549%	2,343%	0,412%	0,126%	0,206%
12:00:00	2,298%	1,955%	0,514%	0,251%	0,171%
13:00:00	2,538%	2,709%	0,709%	0,320%	0,332%
14:00:00	2,469%	3,178%	0,834%	0,160%	0,366%
15:00:00	2,904%	2,892%	0,514%	0,274%	0,412%
16:00:00	3,018%	1,818%	0,892%	0,251%	0,194%
17:00:00	3,406%	1,566%	1,303%	0,480%	0,354%
18:00:00	1,806%	2,732%	0,560%	0,183%	0,251%
19:00:00	0,274%	1,989%	0,251%	0,046%	0,137%
20:00:00	0,091%	0,617%	0,046%	0,046%	0,023%
21:00:00	0,091%	0,080%	0,011%	0,000%	0,000%
22:00:00	0,023%	0,183%	0,000%	0,011%	0,000%
23:00:00	0,000%	0,000%	0,011%	0,000%	0,000%

Fonte: Elaborados pelos Autores (2022).

Através da análise da tabela é possível identificar três períodos com maior índice de detecção de desvios, o primeiro é no horário das 06:00 horas, horário de início da jornada de trabalho da maioria dos condutores e muito mais propício para detecção de desvios como Fadiga e Distração. O Segundo período de alta de desvios é entre o horário de 12:00 e 14:00, horário pós almoço. O terceiro, é registrado entre as 17:00 e 18:00, horário de fim de jornada da maioria dos condutores que, assim como no início do dia, estão mais propícios a detecção de desvios.

4.2. Resultados via reconhecimento de padrões

Após inserir as bases de dados no software Studio R, obteve-se resultados diferentes de acordo com as variações dos parâmetros do software como os tipos de MLP. O teste se baseou no número de neurônios na camada oculta, variando em 3, 5 e 7, além da variação da reamostragem da validação alternando entre 35-65% (o que significa 65% da base para treino do algoritmo e 35% para validação) e 25-75%.

Foram utilizados 5 tipos de MLP diferentes: Standard Backpropagation, Backpropagation Momentum, Resilient Propagation, Backpropagation with Weight Decay e Quick Propagation, resultando em 120 testes para MLP.

Ao executar o algoritmo no software para cada uma das funções listadas, obtêm-se a Área sob a curva (Area Under Curve – ou AUC). Quanto mais próximo de 1 (um) o resultado, melhor conforme pode ser observado na tabela 5.

Tabela 5 - Resultados AUC por tipo de MLP (continua)

Tipo de MLP	Neurônios	Reamostragem	AUC (Fadiga vs Distração)	AUC (Fadiga vs Motorista)	AUC (Fadiga vs Cansaco)	AUC (Fadiga vs Telefone)
Standard	3		0,9652	0,8301	0,8716	0,8006
	5	65/35	0,9721	0,8901	0,8647	0,8203
	7		0,9654	0,8719	0,8901	0,8521
	3		0,9657	0,795	0,8786	0,8313
	5	75/25	0,9681	0,792	0,8896	0,7971
	7		0,9651	0,7725	0,8702	0,7687
Momentum	3		0,9727	0,7762	0,8843	0,8328
	5	65/35	0,9715	0,8633	0,8719	0,8291
	7		0,9651	0,8746	0,8819	0,8415
	3		0,9662	0,7943	0,8634	0,8027
	5	75/25	0,973	0,7869	0,8858	0,831
	7		0,9679	0,8852	0,872	0,8112
Resilient	3		0,9709	0,8197	0,8775	0,819
	5	65/35	0,9612	0,8085	0,8839	0,8339
	7		0,9636	0,869	0,888	0,854
	3		0,9712	0,7974	0,8527	0,7835
	5	75/25	0,9598	0,8372	0,8802	0,8051
	7		0,9625	0,8945	0,8589	0,8228
Weight Decay	3		0,9661	0,8145	0,8751	0,8025
	5	65/35	0,964	0,8346	0,8648	0,7907
	7		0,9742	0,8886	0,8871	0,8134
	3		0,9742	0,8089	0,876	0,8227
	5	75/25	0,9648	0,7909	0,8677	0,839
	7		0,9706	0,8166	0,8802	0,8273
Quick propagation	3		0,9628	0,8265	0,8443	0,75
	5	65/35	0,9684	0,9251	0,8347	0,7723
	7		0,9691	0,8552	0,8511	0,7825
	3		0,9635	0,7781	0,8611	0,7826
	5	75/25	0,9658	0,8656	0,8581	0,7448

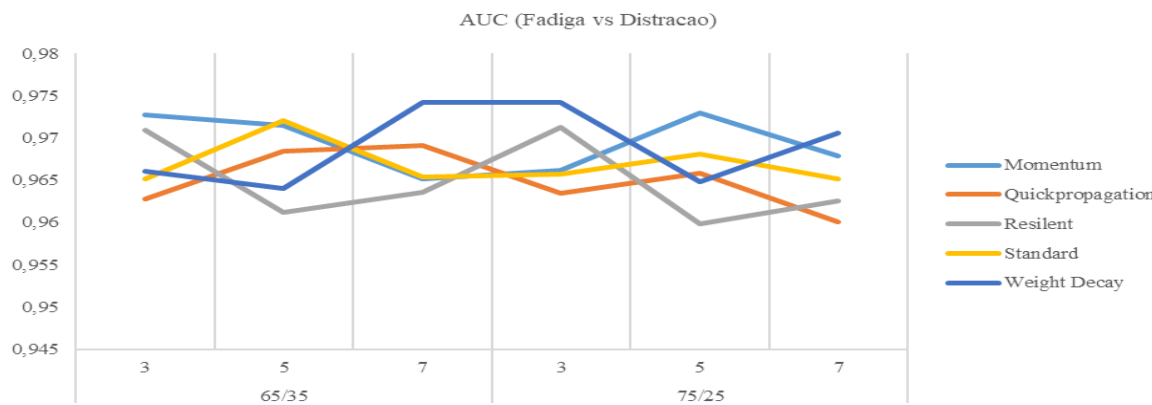
7	0,96	0,8928	0,8633	0,769
---	------	--------	--------	-------

Fonte: Elaborados pelos Autores (2022).

Observando o comparativo da Detecção de Fadiga com Detecção de Distração, é possível identificar que o tipo de MLP Quick Propagation, destacado na cor laranja na figura 9, foi o que menos teve variação, quando alternado a quantidade de neurônios na base de treinamento e validação de 65-35%, já o MLP Standard Propagation, destacado na cor amarela na figura 9, foi o que mais se manteve estável (variando menos de 1%) quando alternado a quantidade de neurônios na base de treinamento e validação 75-25%.

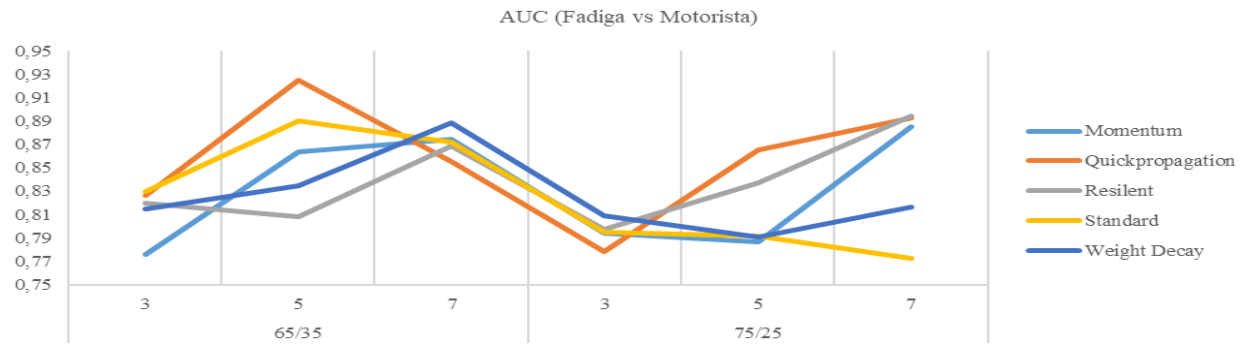
No comparativo de Detecção de Fadiga com Detecção de Motorista não visível, conforme é possível observar na Figura 10, na base de treinamento e validação 65-35%, o MLP Resilient Propagation foi o que demonstrou menor variação quando comparado os demais na variação da quantidade de neurônios no processamento do software.

Figura 9 - Resultados AUC Fadiga vs Distração



Fonte: Elaborados pelos Autores (2022).

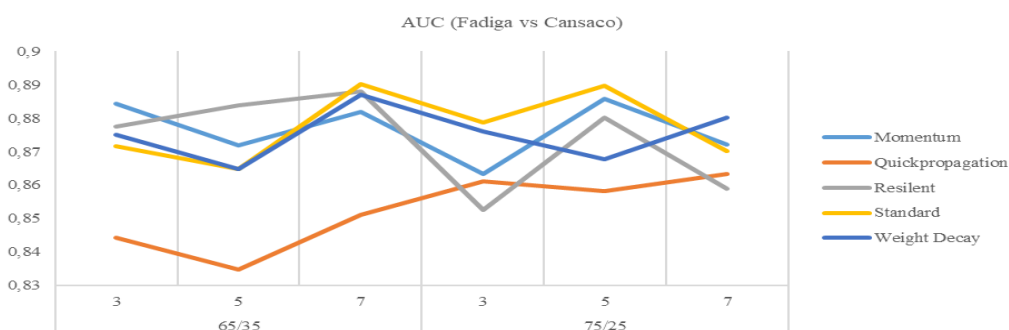
Figura 10 - Resultados AUC Fadiga vs Motorista



Fonte: Elaborados pelos Autores (2022).

No comparativo de Detecção de Fadiga com Detecção de Cansaço, o MLP que mais se manteve estável na primeira na base de treinamento e validação, de 65-35%, foi o Resilient Propagation, destacado na cor cinza na Figura 11, quando alternado a quantidade de neurônios. O MLP Quick Propagation, foi o que mais se manteve estável na base de treinamento e validação 75-25% quando alternado a quantidade de neurônios, apesar de ser o MLP com a menor AUC obtida nessa base. Porém os valores mais altos foram encontrados pelo modelo Standard, chegando a aproximadamente 0,89 e superando as demais.

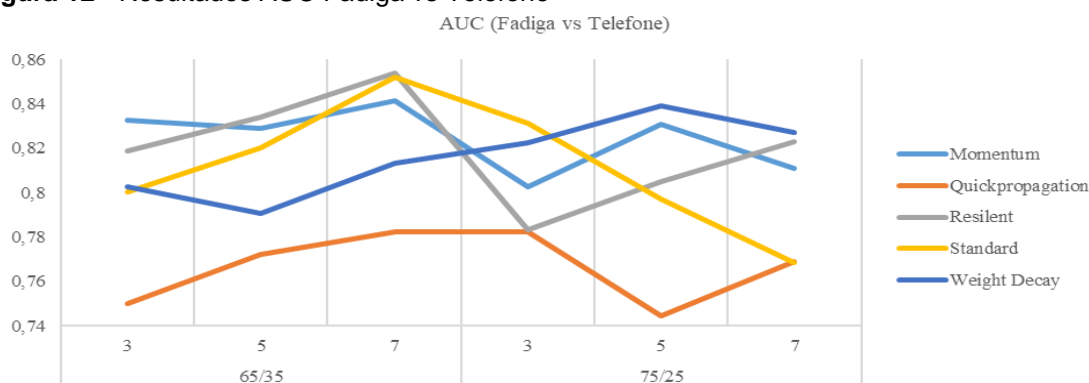
Figura 11 - Resultados AUC Fadiga vs Cansaço



Fonte: Elaborados pelos Autores (2022).

No último comparativo, o de Detecção de Fadiga com a Detecção do uso de celular, o MLP Back Propagation Momentum foi o que menos teve alteração na base de treinamento e validação de 65-35% durante a variação de neurônios e na base de treinamento e validação 75-25% o MLP Back Propagation Weight Decay foi o que mais se manteve estável, como pode ser observado na figura 12. A estabilidade do referido tipo não garantiu o melhor resultado entre as demais, onde a Resilient chegou a aproximadamente 0,86.

Figura 12 - Resultados AUC Fadiga vs Telefone



Fonte: Elaborados pelos Autores (2022).

4.3 Discussão de possibilidades estratégicas

Para cada um dos testes realizados, foram obtidos resultados diferentes conforme a base de dados era alternada, mudando as configurações da quantidade de neurônios, tipo de MLP utilizados e a proporção da quantidade de amostras a serem comparadas, na tabela 6 é possível identificar qual o MLP obteve o melhor resultado em cada comparativo.

Tabela 6 - Comparativo das Melhores AUCs por MLP

Base	Reamostragem	MLP	AUC
Fadiga vs Distracao	65/35	Weight Decay	0,9742
Fadiga vs Motorista	65/35	Quick Propagation	0,9251
Fadiga vs Cansaco	65/35	Standard	0,8901
Fadiga vs Telefone	65/35	Standard	0,8521

Fonte: Elaborados pelos Autores (2022).

Como pode ser observado na tabela 6, os melhores resultados de AUC representam uma capacidade superior à 85% de acurácia para a análise da Fadiga, o que figura um resultado expressivo, mesmo frente aos momentos de desbalanceamento de amostras.

Existem diversas possibilidades de aplicação para o reconhecimento de padrões, um deles, é a aplicação do modelo em transportadoras cujo os motoristas precisam percorrer grandes distâncias durante sua jornada de trabalho e precisam ficar atentos ao trânsito a diversos fatores externos para garantir a sua segurança e dos terceiros que estão a sua volta durante o trajeto da viagem.

Devido ao avanço das reivindicações dos caminhoneiros, além da exposição que existe quando acidentes acontecem, as empresas estão se preocupando cada vez mais com a prevenção de acidentes e com a segurança de seus motoristas, com isso, o reconhecimento de padrões ganha importância pois, diferente das muitas propostas já existentes no mercado, que trabalham com avaliação pós incidente, o reconhecimento de padrões através de uma rede neural permite atuar podendo prever onde um evento de detecção de fadiga tem chances de ocorrer com mais frequência por exemplo possibilitando que a empresa tome uma decisão estratégica de mudar uma rota do veículo ou o horário de jornada do motorista para evitar o risco.

5 CONCLUSÃO

Durante o desenvolvimento, o primeiro desafio foi entender a base de dados e montar uma planilha com as variáveis que seria utilizada no estudo, depois, foi preciso identificar cada variável por código para que o software fizesse o reconhecimento, em seguida iniciou-se a montagem da classificação dos dados que foi usado para experimentação por reconhecimento de padrões sendo lançado uma base com 5 variáveis de referência.

Com as amostragens definidas e a base de dados montada, o primeiro estudo foi identificar a velocidade média por desvio, onde foi possível identificar que o evento de detecção de cansaço obteve a maior média, sendo 37 km/h, lembrando que não consideramos os eventos com velocidade igual a 0 km/h.

Em seguida, foi realizado o estudo do desvio por faixa horária, ou seja, qual o desvio com maior reincidência e o horário detectado, consideramos o início e fim da jornada, sendo assim através da análise foi possível identificar que nos horários das 06:00 e 17:00 os eventos de Detecção de Distração e Detecção de Fadiga apresentaram maiores ocorrências.

Com relação ao reconhecimento de padrões, como já mencionado, a acurácia conseguiu superar a faixa de 85%, o que valida a capacidade preditiva dos algoritmos, mesmo frente ao desbalanceamento de amostras.

A proposta de análise estratégica da gestão dos desvios do sistema de sonolência foi bem-sucedida, apesar da dificuldade e desbalanceamento de amostras em algumas situações, em geral a análise funcionou de maneira satisfatória nas condições a qual o software foi submetido. Figura-se como a maior contribuição deste trabalho, comparando-se com o que se analisou na literatura, a capacidade preditiva de desvios de acordo com os dados do processo, ou seja, os dados do comportamento da operação, servindo como um complemento para as tecnologias que hoje focam na análise comportamental do motorista. Visa-se, para estudos futuros, a experimentação com as demais ocorrências como referência (Motorista, Telefone, Distração e Cansaço) a fim de vislumbrar uma análise multivariada.

REFERÊNCIAS

ARAUJO, Igor Baesse; RODRIGUES, Diva Silva e Souza; SOUZA, Flávio Henrique Batista de; CARVALHO, Luiz Melk; MUSTIFAGA, Yuri R. Structure and Experiments to Recognition and Classification of Fake News about Citizens in Brazilian

Government Positions. *In*: INTERNATIONAL CONFERENCE ON RESEARCH IN ENGINEERING, SCIENCE & TECHNOLOGY, 11., 2021, Budapest. DOI: <https://doi.org/10.33422/11th.restconf.2021.03.15>

BAIMA, Leonel Bittencourt. **Diagnóstico ambiental e de processo da lavra de rocha ornamental com vistas a aplicação de técnicas de produção mais limpa.** Instituto federal de Educação, Ciência Tecnologia do Rio Grande no Norte, 2018.

BERRI, Rafael Alcesti. **Sistema ADAS para identificação de distrações e perturbações do motorista na condução de veículos.** Universidade de São Paulo, 2019. DOI: <https://doi.org/10.11606/t.55.2019.tde-10062019-082306>

CAMPOS, Mailon *et al.* **Inteligência artificial com aprendizado continuado aplicada ao reconhecimento de padrões.** Instituto Federal do Mato Grosso, 2020. DOI: <https://doi.org/10.34117/bjdv6n5-001>

CAMPOS, Othon, S. F. **Data analytics transparente para descoberta de padrões e anomalias na realização de convênios e contratos de repasse federais.** Universidade Federal de Sergipe, 2018.

CAMPOS, Teofilo Emidio de. **Reconhecimento de Padrões.** Disponível em: <http://www.vision.ime.usp.br/~teo/publications/qualificacao/node5.html#:~:text=Reconhecimento%20de%20padr%C3%B5es%20%C3%A9%20a,s%C3%A3o%20seus%20nomes%20ou%20identifica%C3%A7%C3%B5es>. Acesso em: 11 abr. 2022.

CHAOUCHI, Mohamed, *et al.* **Análise Preditiva para Leigos.** 2 ed. Rio de Janeiro: Editora Alta Books, 2019

FARIAS, Andros Rafael *et al.* **Gestão de fadiga de operadores de caminhões fora de estrada:** estudo de caso em uma mina a céu aberto. Universidade Estadual do Pará, 2019.

FUGA, Bruno Augusto Sampaio. **Acidentes de trânsito:** responsabilidade civil e danos recorrentes. 3 ed. Londrina/PR: Editora Thoth, 2019.

GERRING, John. **Pesquisa de estudo de caso:** princípios práticos. 1 ed. Petrópolis, RJ: Editora Vozes, 2019.

GORDON, Rachel. **Deep learning helps predict traffic crashes before they happen.** Disponível em: <https://news.mit.edu/2021/deep-learning-helps-predict-traffic-crashes-1012>. Acesso em: 25 abr. 2022.

IQBAL, R. *et al.* Big data analytics: Computational intelligence techniques and application areas. **Technological Forecasting And Social Change**, [s.l.], v. 153, n. 1, p. 1-13, 2020. DOI: <https://doi.org/10.1016/j.techfore.2018.03.024>

KAKHANI, M. K.; KAKHANI, S.; BIRADAR, S. R. Research issues in big data analytics. **International Journal of Application or Innovation in Engineering & Management**, v. 2, n. 8, p. 228-232, 2015.

LAMOSO, L. P. Os territórios da mineração sob a lógica da acumulação financeira no capitalismo contemporâneo. **Geosp – Espaço e Tempo (Online)**, v. 21, n. 3, p. 718-736, dez. 2017. ISSN 2179-0892. DOI: <https://doi.org/10.11606/issn.2179-0892.geosp.2017.123788>

LEVELING, Jens; EDELBROCK, Matthias; OTTO, Boris. **Big data analytics for supply chain management**. In: IEEE INTERNATIONAL CONFERENCE ON INDUSTRIAL ENGINEERING AND ENGINEERING MANAGEMENT, IEEE, 2014. p. 918-922. DOI: <https://doi.org/10.1109/ieem.2014.7058772>

MELLER, Guilherme Semprebom *et al.* **Mineração x meio ambiente: mitigação de impactos ambientais causados pela mineração a céu aberto de basalto**. Faculdade satc, 2018. DOI: <https://doi.org/10.1109/ieem.2014.7058772>

MACHADO, Felipe Nery Rodrigues. **Big data o futuro dos dados e aplicações**. Saraiva Educação SA, 2018.

MALIK, Hasmad; FATEMA, Nuzhat; ALZUBI, Jafar A. (Ed.). **AI and machine learning paradigms for health monitoring system: intelligent data analytics**. Springer, 2021. DOI: <https://doi.org/10.1007/978-981-33-4412-9>

MIHAS, Paul. **Qualitative data analysis**. Oxford Research Encyclopedia of Education. 2019.

MOREIRA, João; CARVALHO, Andre; HORVATH, Tomás. **A general introduction to data analytics**. John Wiley & Sons, 2018.

NOVAIS, Gabriel D. L. *et al.* **Reconhecimento de padrões em situações de risco no combate à crimes em transportes coletivos**. Centro Universitário de Belo Horizonte, 2018.

PARANHOS, Vinícius Kerber. **Sistema de Detecção de Sonolência**. Universidade de Passo Fundo, Trabalho final de graduação (Engenheiro Eletricista). Curso de Engenharia Elétrica. 2019. DOI: <https://doi.org/10.14393/19834071.2012.17680>

PORTO, Pedro. H. E. *et al.* **Gestão de fadiga de operadores de caminhões fora de estrada: estudo de caso em uma mina a céu aberto**. Universidade Federal de Ouro Preto, 2018. DOI: <https://doi.org/10.22533/at.ed.24621181110>

SALLES, Arthur França. **Detecção automática de sonolência em condutores de veículos utilizando redes neurais artificiais**. Universidade Federal do Mato Grosso, 2018. DOI: <https://doi.org/10.11606/t.98.2012.tde-04062012-083028>

SOUZA, Fabiane. **Big data analytics como ferramenta de adaptação do total quality management na indústria 4.0, aplicado a uma empresa multinacional do ramo automobilístico.** UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANA, 2021.

SOUZA, Flávio, H. B. *et al.* **On modelling and simulating open pit mine through stochastic timed petri nets.** Universidade Federal de Minas Gerais, 2019.

SHINOHARA, E. J. *et al.* **Caracterização de padrões do comportamento espacial de veículos (monitorados) em movimento – PCEVM.** *In:* INTERNATIONAL TRANSP. PLANNING SOCIETY. 2015. DOI: <https://doi.org/10.1590/2238-1031.jtl.v9n3a3>

SHINOHARA, E. J. **Modelagem do comportamento espaço-temporal de veículo rastreado.** Universidade de São Paulo. 2013. DOI: <https://doi.org/10.1590/2238-1031.jtl.v9n3a3>

TV GLOBO. **Vídeo mostra momento do acidente com ex-BBB Rodrigo Mussi em SP.** Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2022/04/04/video-mostra-momento-do-acidente-com-ex-bbb-rodrigo-mussi-em-sp.ghtml>. Acesso em: 14 abr 2022. DOI: <https://doi.org/10.47749/t/unicamp.2015.955858>

VALE, J.G.S. *et al.* Revisão de Literatura: Big Data Analytics na Engenharia de Produção. *In:* IX CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO, 2019. DOI: <https://doi.org/10.29327/xsimep.472546>



Artigo recebido em: 08/07/2022 e aceito para publicação em: 27/12/2022
DOI: <https://doi.org/10.14488/1676-1901.v22i3.4698>

APÊNDICE 1 - Resultados do experimento

	berlin52	kroA100	kroA150	kroA200	kroB100	kroB150	kroB200	kroC100	kroD100	kroE100	lin105	pr76	pr107	pr124	USGS
51%	14212	69765	116248	152346	73026	112162	145860	78179	70429	77816	46939	283833	205783	290825	12189000
52%	14078	75206	108948	153195	72064	116469	148587	68376	68208	71562	49433	267890	225484	289005	12108000
53%	15159	73112	97453	148279	69213	109155	145480	69263	69224	75911	50107	243820	218931	257818	12068000
54%	14483	71304	110274	149681	68823	107633	143397	67091	65847	71374	46457	248193	201295	292642	11678000
55%	14645	69401	106460	139433	72918	103992	139182	65169	71938	67187	49304	256180	194587	285836	12704000
56%	13006	65191	100314	145335	63513	99151	140268	69214	63706	75949	45603	222616	198024	255747	12690000
57%	14523	54779	97948	134107	72091	102590	131060	64545	64352	69587	48554	266870	189983	241060	11433000
58%	13909	68010	106465	120483	66774	95527	131300	67939	65220	64433	47330	248049	186313	265654	11020000
59%	11855	60665	90821	113056	62995	93957	119503	62753	57395	65303	45843	222897	193860	253788	11182000
60%	11629	62746	93811	123599	68533	99055	130505	62454	61436	65857	45457	213377	162336	249163	10451000
61%	11639	63252	91827	126269	60278	91895	126376	61511	65401	64972	46492	233447	182549	235512	10110000
62%	12902	59653	95217	126822	56704	87078	117989	58807	57840	57151	39829	232164	156687	212370	10956000
63%	13528	59240	90332	116781	58382	89820	118328	54268	53111	61928	38013	215954	173002	223635	10054000
64%	11964	62932	90205	122484	55529	88510	106882	54698	55808	53242	42407	237977	173282	217354	9892500
65%	12642	51532	84763	99697	56425	82833	108795	58699	55185	61831	42336	221038	138916	189473	9628300
66%	11802	52715	76705	113890	59312	78335	106570	57886	50633	53426	38760	214057	153896	214556	10330000
67%	11495	47515	83798	112680	55177	83083	104640	55158	55413	47386	38723	199791	146070	200239	9788100
68%	12765	50948	82838	104884	54585	84376	100377	50402	40715	55691	37436	199133	129015	193440	9902300
69%	11407	54188	70492	97594	51468	77460	95554	52378	50602	53467	30435	201976	138488	207159	9482100
70%	11800	48927	75091	95051	51809	73870	84659	49162	49681	50615	34345	200701	146460	183168	8573400
71%	11768	52534	72078	99590	51325	70672	104518	48334	45779	45842	32106	184581	135489	184667	8408000
72%	11508	48048	70567	89631	49479	64120	89938	49693	49904	44555	30225	177412	125129	177937	9165200
73%	11528	47510	66755	85781	46133	71654	89587	48918	48353	49524	33438	187366	119604	172008	7574400
74%	10309	49361	66686	92710	48733	64115	85303	42403	43334	49620	32343	190270	111362	170188	8859100
75%	11475	37775	68913	90636	41468	64226	82152	40678	40243	47238	33723	178293	112999	164520	7969600
76%	9955	43939	65429	85347	48022	67936	82301	46926	41170	46520	28172	196288	112630	156581	7864300
77%	9963	35176	63397	83144	43756	63429	74376	40063	42312	37923	30545	180994	85159	144302	8082400
78%	10570	41497	60958	81499	42328	63819	74241	40910	34702	42270	28789	172850	112214	136583	8134600
79%	10270	43934	59743	74310	42239	56494	68418	39767	41574	40213	27988	167913	103576	141776	7338800
80%	10435	39285	57567	71463	38576	55533	65943	38766	36765	40489	30357	173202	86479	141359	7605300
81%	11026	37125	57368	76164	37734	56653	72596	37740	39693	42134	26680	170993	93735	124444	7634800
82%	9510	36619	57090	67193	39305	52828	69419	34866	38794	31040	25800	161510	77862	130979	6788200
83%	9139	35128	55051	64664	35091	53582	65357	34704	35820	39753	25494	145930	76829	122345	6781900
84%	10034	34304	47988	61976	33764	52413	64177	33606	37264	32550	25183	168231	71640	106697	7003000
85%	9422	33721	50116	62850	33202	45065	54687	32942	34942	32206	24083	156477	72147	107870	6894400
86%	9678	29568	43209	57410	31110	47303	63600	31442	33726	31289	22967	154045	73398	114594	5859700
87%	9125	32042	42543	54767	33522	45376	56176	33869	34383	28282	22159	147642	64942	107711	6404400
88%	8932	31450	40470	52934	31018	43018	58878	31304	29720	31297	23212	153216	63948	96064	5903200
89%	8743	30261	42708	45386	32142	44049	51156	26340	32500	29601	20405	153519	58465	99277	5939400
90%	9074	30823	42138	46447	27145	40197	48389	27952	26437	30719	21892	149224	57048	92561	5791600
Mínimo	8743	29568	40470	45386	27145	40197	48389	26340	26437	28282	20405	145930	57048	92561	5791600

Fonte: Autores.

APÊNDICE 2 - Resultados do tempo de processamento

	berlin52	kroA100	kroA150	kroA200	kroB100	kroB150	kroB200	kroC100	kroD100	kroE100	lin105	pr76	pr107	pr124	UFGS
51%	0,9868	0,2634	1,0754	1,9468	0,0013	0	0	0	0	0	0	0	0	0	0
52%	1,2531	0,3181	1,3879	2,2353	0,0029	0	0	0	0	0	0	0	0	0	0
53%	1,0388	0,2762	1,1285	2,0622	0,0031	0	0	0	0	0	0	0	0	0	0
54%	1,0676	0,2841	1,1527	2,1222	0,0028	0	0	0	0	0	0	0	0	0	0
55%	1,2127	0,3102	1,1279	2,2945	0,0034	0	0	0	0	0	0	0	0	0	0
56%	1,0118	0,2696	1,1031	2,0042	0,0018	0	0	0	0	0	0	0	0	0	0
57%	1,1019	0,2770	1,2485	2,0896	0,0016	0	0	0	0	0	0	0	0	0	0
58%	1,5857	0,3472	1,3814	2,1864	0,0019	0	0	0	0	0	0	0	0	0	0
59%	1,0842	0,2885	1,1791	2,4229	0,0028	0	0	0	0	0	0	0	0	0	0
60%	1,7926	0,4748	1,8122	4,1562	0,0040	0	0	0	0	0	0	0	0	0	0
61%	1,7286	0,2682	1,0736	2,9463	0,0027	0	0	0	0	0	0	0	0	0	0
62%	0,9692	0,2585	1,0569	1,9447	0,0052	0	0	0	0	0	0	0	0	0	0
63%	0,9875	0,2674	1,0790	1,9592	0,0018	0	0	0	0	0	0	0	0	0	0
64%	1,0154	0,2728	1,0984	2,0013	0,0014	0	0	0	0	0	0	0	0	0	0
65%	1,1174	0,2850	1,0981	1,9750	0,0017	0	0	0	0	0	0	0	0	0	0
66%	0,9604	0,2550	1,0470	1,9176	0,0015	0	0	0	0	0	0	0	0	0	0
67%	1,2655	0,2973	1,2061	1,9868	0,0017	0	0	0	0	0	0	0	0	0	0
68%	0,9872	0,2630	1,0762	1,9510	0,0020	0	0	0	0	0	0	0	0	0	0
69%	1,1538	0,3097	1,1245	1,9422	0,0022	0	0	0	0	0	0	0	0	0	0
70%	0,9674	0,2588	1,0536	1,9175	0,0015	0	0	0	0	0	0	0	0	0	0
71%	1,0372	0,2750	1,1126	2,028	0,0014	0	0	0	0	0	0	0	0	0	0
72%	1,0228	0,3090	2,3403	3,4661	0,0017	0	0	0	0	0	0	0	0	0	0
73%	1,076	0,2790	1,1419	2,1701	0,0020	0	0	0	0	0	0	0	0	0	0
74%	0,9977	0,2659	1,0854	1,9818	0,0014	0	0	0	0	0	0	0	0	0	0
75%	1,0609	0,2850	1,1486	2,0969	0,0015	0	0	0	0	0	0	0	0	0	0
76%	1,0218	0,2741	1,1066	2,0181	0,0014	0	0	0	0	0	0	0	0	0	0
77%	1,4096	0,4136	6,0741	2,0291	0,0019	0	0	0	0	0	0	0	0	0	0
78%	1,0032	0,2684	1,0922	1,987	0,0009	0	0	0	0	0	0	0	0	0	0
79%	1,0139	0,2699	1,1018	2,0236	0,0018	0	0	0	0	0	0	0	0	0	0
80%	1,0187	0,2726	1,1155	2,0149	0,0013	0	0	0	0	0	0	0	0	0	0
81%	0,9685	0,2591	1,0595	1,9174	0,0016	0	0	0	0	0	0	0	0	0	0
82%	0,9856	0,2632	1,0751	1,9630	0,0011	0	0	0	0	0	0	0	0	0	0
83%	0,9845	0,2610	1,0902	1,9574	0,0011	0	0	0	0	0	0	0	0	0	0
84%	0,9658	0,2524	1,0440	1,8841	0,0020	0	0	0	0	0	0	0	0	0	0
85%	0,9744	0,2559	1,0504	1,9170	0,0018	0	0	0	0	0	0	0	0	0	0
86%	0,9691	0,2571	1,0555	1,9303	0,0014	0	0	0	0	0	0	0	0	0	0
87%	1,0134	0,2626	1,0758	1,9518	0,0018	0	0	0	0	0	0	0	0	0	0
88%	0,9948	0,2658	1,0839	1,9721	0,0017	0	0	0	0	0	0	0	0	0	0
89%	1,5245	0,4757	1,9486	2,2747	0,0018	0	0	0	0	0	0	0	0	0	0
90%	1,2908	0,4901	1,6337	2,8471	0,0028	0	0	0	0	0	0	0	0	0	0

Fonte: Autores.