

# Utilizando o Modo de Dirigir do Motorista de Veículo Elétrico para o Planejamento e Roteirização de Viagem

Marcelo dos-Reis<sup>1</sup>, Fabiano Costa Teixeira<sup>1</sup>, Humberto T. Marques-Neto<sup>2</sup>

<sup>1</sup>Instituto de Ciências Exatas e Informática (ICEI)  
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)  
Campus de Poços de Caldas  
Poços de Caldas – MG – Brasil

<sup>2</sup>Programa de Pós-graduação em Informática (PPGInf)  
Instituto de Ciências Exatas e Informática (ICEI)  
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)  
Caixa Postal 30535-901 – Belo Horizonte – MG – Brasil

marceloreis@pucpcaldas.br, teixeira@pucpcaldas.br, humberto@pucminas.br

**Abstract.** *With the increasing adoption of electric vehicles worldwide, some limitations have emerged in their usage. The main limitations include low autonomy and a scarcity of charging points. In this work, we describe a software architecture for planning a stop at charging stations along a trip, by prediction of battery charge to be spent along the path. We describe the main components of this architecture and evaluate regression methods for the car consumption prediction module. We also use a real dataset built from an electric vehicle usage to validate the architecture concept and its viability analyzing multiple linear regression machine learning models. To further validate the architecture, we make a comparison between a simulated and a real trip.*

**Resumo.** *Com o aumento da adoção de veículos elétricos no mundo, começam a aparecer as limitações em sua utilização, tais como a baixa autonomia e a escassez de pontos de recarga. Este artigo apresenta uma arquitetura de software de uma aplicação capaz de prever a autonomia de carros elétricos e planejar a parada em postos de recarga ao longo de um trajeto de viagem. Além da descrição dos principais componentes da arquitetura, este trabalho também apresenta uma avaliação de métodos de regressão para o módulo de previsão de consumo do carro. Dados reais de um veículo elétrico foram coletados e utilizados para validar o conceito e a viabilidade da arquitetura, através da análise de modelos de aprendizado de máquina baseados em regressão linear múltipla. Ainda para validar a arquitetura, foi feita uma comparação entre uma viagem simulada e outra real.*

## 1. Introdução

O uso de carros elétricos no mundo tem crescido rapidamente na última década e essa frota pode representar 13% da frota mundial de carros em 2030. Isto representa um crescimento médio de 36% ao ano entre 2019 e 2030 [Gorner 2020]. Empresas, como a Ambev, Coca-Cola e JBS, vêm utilizando a eletrificação da frota de distribuição para atingir metas de redução da pegada de carbono. Espera-se que até 2040, cerca de 85% da

frota de carros por aplicativo será elétrica [McKinsey 2023]. A redução do impacto ambiental negativo per capita das cidades é um objetivo fundamental da ODS 11 (Objetivo de Desenvolvimento Sustentável 11 - Cidades e comunidades sustentáveis) e o planejamento de recarga dos veículos elétricos em viagens interurbanas pode ser um fator importante para incentivar mais pessoas a adotarem veículos elétricos.

Um problema comum na adoção em larga escala dos veículos elétricos é a baixa autonomia e os contratempos com os abastecimentos [Lebrouhi et al. 2021], tais como: (i) alta frequência de recargas, devido à capacidade limitada das baterias, (ii) longa duração dos abastecimentos, tipicamente de 30 minutos a 2 horas e (iii) picos de abastecimento, por existirem horários mais procurados para abastecimento, como intervalos de almoço [Wang et al. 2020]. No Brasil, também é importante destacar a oferta limitada de postos elétricos nas estradas e cidades brasileiras, em parte, devido ao fato de que somente em junho de 2018 a ANEEL publicou uma resolução permitindo a exploração comercial do abastecimento de veículos elétricos [Velandia Vargas et al. 2020]. Em outros países existem problemas similares com a infraestrutura de carregamento, como na Itália, com a ausência de investimento nas regiões ao sul desse país [Noussan 2020] e na Turquia, onde também existe deficiência na infraestrutura [Gönül et al. 2021].

Por esses motivos, a previsão correta da autonomia e a consequente recomendação de postos de recarga ao longo do trajeto é uma abordagem que pode contribuir com a mitigação de problemas que podem interferir no crescimento da adoção de carros elétricos. Destaca-se aqui que um dos fatores que pode influenciar nesta previsão é o modo de dirigir de cada motorista [Delnevo et al. 2019]. O modo de dirigir é um conceito complexo que pode ser definido como a maneira que o motorista opera o controle do veículo no contexto de direção e das condições externas.

Este artigo propõe uma arquitetura de software que faz uso de um classificador de tipos de motoristas e um modelo para previsão de autonomia. Dados de telemetria são utilizados para determinar e classificar os motoristas pelo seu modo de dirigir e para uma previsão de consumo de energia de cada trecho de uma viagem. Realizar uma previsão considerando o tipo de motorista e os dados de contexto do trecho poderia promover a recomendação de postos de recarga ao longo do percurso, com as devidas autonomias calculadas, para oferecer uma melhor experiência ao motorista ao longo de sua viagem. Ainda que existam aplicativos proprietários, como da Tesla e Volvo, este trabalho descreve uma arquitetura aberta, com API que pode ser utilizada por veículos mais populares e integrada a aplicativos já existentes, como *Waze* e *Google Maps*.

Enfim, a arquitetura proposta tem como objetivo orientar a construção de softwares capazes de realizar a previsão de autonomia de veículos elétricos e recomendar pontos de recarga. Assim, os objetivos deste trabalho foram: (a) avaliar se a arquitetura proposta atende ao propósito de estimar o gasto de energia ao longo de uma viagem, (b) identificar os melhores algoritmos de aprendizado de máquina na previsão da autonomia de veículos elétricos e (c) incorporar as evidências de dirigibilidade e de contexto de viagem nas estratégias de previsão de reabastecimento ao longo de uma viagem.

Após as análises utilizando dados reais coletados diretamente de um veículo elétrico, verificou-se quais métodos de regressão obtiveram resultados satisfatórios para prever a quantidade de energia de bateria consumida em trechos da viagem. O método

*Random Forest Regressor* obteve erro médio absoluto de 0,346 pontos percentuais (p.p.). Através do aplicativo de roteirização, é possível prever a autonomia de uma viagem projetada, utilizando a previsão de consumo trecho a trecho. Em um teste real, a energia gasta durante uma viagem de 221 km resultou em 74%, contra um gasto simulado de 71,89%. Em outro teste real que teve consumo de 68%, o modelo previu 68,31%. Além disso, a partir do cálculo das autonomias, pode-se planejar a viagem, calculando-se quais postos de recarga seriam possíveis de se alcançar antes da descarga total da bateria. Este planejamento, e uma possível integração com aplicativos de roteirização de entregas, pode fazer a diferença na adoção de veículos elétricos no transporte e distribuição de cargas.

O restante deste artigo está organizado como segue. Na Seção 2, os trabalhos relacionados são analisados quanto à verificação de como estão sendo solucionados os problemas de classificação do modo de dirigir e previsão de autonomia de veículos elétricos. A Seção 3 apresenta um resumo das principais tecnologias utilizadas no projeto. A arquitetura proposta é descrita na Seção 4, juntamente com o detalhamento do funcionamento dos principais módulos. A avaliação dos módulos e os resultados preliminares, com os dados atualmente coletados, são apresentados na Seção 5. E, por fim, a Seção 6 discorre sobre futuros trabalhos e a conclusão.

## **2. Trabalhos Relacionados**

Existem muitos estudos sobre a identificação do estilo de condução de um motorista. Esses estudos utilizam uma variedade de técnicas de aprendizado de máquina, incluindo redes neurais [Jia et al. 2020], árvores de decisão [Krishnamurthy et al. 2019], clustering [Fugiglando et al. 2019], entre outras. Geralmente, os motoristas são classificados em diferentes categorias com base em dados de telemetria, como pressão do pedal de freio, velocidade, aceleração e desaceleração [Chen et al. 2019a]. Alguns estudos também exploram o impacto do estilo de condução no consumo de combustível ou energia da bateria de um veículo [Ping et al. 2019].

Os dados de telemetria são geralmente coletados a partir de dispositivos instalados no veículo, como a rede CAN (*Controller Area Network*) [Chen et al. 2019a]. Técnicas de amostragem podem ser utilizadas para diminuir a quantidade de dados a serem processados [Fugiglando et al. 2019]. Alguns trabalhos comparam diferentes algoritmos de aprendizado de máquina para ver qual é mais eficaz na classificação dos motoristas [Chen et al. 2019c], enquanto outros detalham e avaliam as métricas mais comumente utilizadas [Chen et al. 2019a].

Além de classificar os motoristas em diferentes categorias, alguns estudos também tentam prever o desempenho de um motorista com base em seus dados de telemetria [Siami et al. 2020], utilizando, por exemplo, modelos de redes neurais para prever o consumo de combustível ou energia da bateria de um veículo com base no estilo de condução do motorista [Chen et al. 2019b]. Outros estudos tentam prever se perda ou ganho de altitude em um trajeto pode afetar a autonomia de um veículo elétrico [Bolovinou et al. 2014].

Em resumo, existem muitos trabalhos sobre a identificação do estilo de condução de um motorista, utilizando uma variedade de técnicas de aprendizado de máquina e dados de telemetria. Esses estudos têm como objetivo classificar os motoristas em diferentes categorias, bem como prever o desempenho e o consumo de recursos com base em seu

estilo de condução [Stanton and Allison 2020].

E por fim, existem trabalhos que utilizam dados de telemetria na predição de autonomia de veículos elétricos [Bailey et al. 2022]. Porém, nenhum desses trabalhos propõe uma arquitetura de software como a proposta aqui. O principal foco desse trabalho é apresentar tal arquitetura, integrando um algoritmo de classificação dos motoristas, que os agrupa por modo de dirigir usando *features* extraídas dos dados de telemetria, através de um hardware embarcado, com um algoritmo de regressão linear que leva em consideração tanto o tipo de motorista quanto os dados de contexto da viagem para determinar o consumo a cada trecho. Assim, é possível tanto calcular a autonomia do veículo quanto planejar as paradas em postos de recarga ao longo de uma viagem ou rota de entrega.

### **3. Tecnologias de Apoio**

Na arquitetura de software proposta, tecnologias foram utilizadas e integradas em módulos, como o uso de dados de telemetria em algoritmos de classificação e regressão para a previsão do consumo e autonomia. Todas essas tecnologias são descritas a seguir.

#### **3.1. Telemetria**

Com o aumento da frota, a utilização dos dados de telemetria dos veículos elétricos pode ser uma importante fonte de dados para determinar o modo de dirigir para prever como é o gasto energético e prever a autonomia do veículo. A rede CAN é um sistema de barramento serial de alta velocidade muito utilizado em sistemas embarcados, largamente utilizada nos veículos modernos a partir da adoção de vários microcontroladores, como injeção eletrônica, ABS e outros. Através desse barramento os dados de aceleração, velocidade, odômetro, pressão no pedal acelerador e outros dados dos veículos elétricos podem ser obtidos [Fugiglando et al. 2019].

Esses dados podem ser transmitidos pela Internet, por meio de um protocolo de comunicação. Ao utilizar o GPRS, disponível nas áreas de cobertura de 3G e 4G, consegue-se cobrir quase a totalidade das cidades brasileiras e 55.711km, ou 44,2%, de suas rodovias, segundo painel de dados de infraestrutura da ANATEL [ANATEL 2022].

#### **3.2. Classificação de Motoristas com Base no Modo de Dirigir**

Para classificar os motoristas pelo modo de dirigir, a proposta foi a utilização de um método não supervisionado de análise de agrupamento de dados (*clustering*). Métodos não supervisionados permitem determinar parâmetros de separação dos grupos a partir dos dados analisados [Arabie et al. 1996].

A acurácia dos métodos de classificação depende da quantidade de grupos e *features*. Adicionalmente, o número de características analisadas influencia na caracterização dos grupos, aumentando ou diminuindo o número de classes para diferenciação. Também, poucos grupos podem resultar em grupos muito genéricos, com poucas similaridades, mas muitos grupos podem tornar a aprendizagem do modelo muito complexa e computacionalmente custosa. Encontrar a quantidade certa de grupos, sem perder informação, já foi objeto de estudos [Fugiglando et al. 2019].

#### **3.3. Regressão Linear**

Algoritmos de regressão linear são utilizados para determinar o comportamento de um conjunto de dados e extrapolar valores de saída baseados nos dados de entrada. Busca-se

encontrar uma correlação entre as variáveis independentes e os resultados. Algoritmos de regressão linear múltipla são uma extensão do modelo de regressão linear simples, onde várias variáveis de predição são utilizadas para apenas um resultado de um conjunto contínuo [Eberly 2007].

Para prever o consumo de veículos elétricos, este trabalho propõe um algoritmo de regressão múltipla que considera dados de contexto de viagem, como o consumo e o tipo de motorista, trecho a trecho. A partir dessas informações, é possível estimar o consumo e a autonomia para o tipo de motorista em questão nos próximos trechos da viagem.

#### 4. Arquitetura

A arquitetura proposta é formada por vários componentes e processos trabalhando em conjunto para se fazer um planejamento de viagem. Alguns componentes são embarcados no veículo e outros são executados na nuvem, conforme Figura 1, onde as setas identificadas por números indicam a troca de informações entre os componentes. Todos os componentes são projetados para serem modulares e facilmente substituídos, conforme necessidade. Esta modularidade também permite que outras aplicações ou componentes se integrem com a plataforma, consumindo a API ou enviando dados de telemetria.

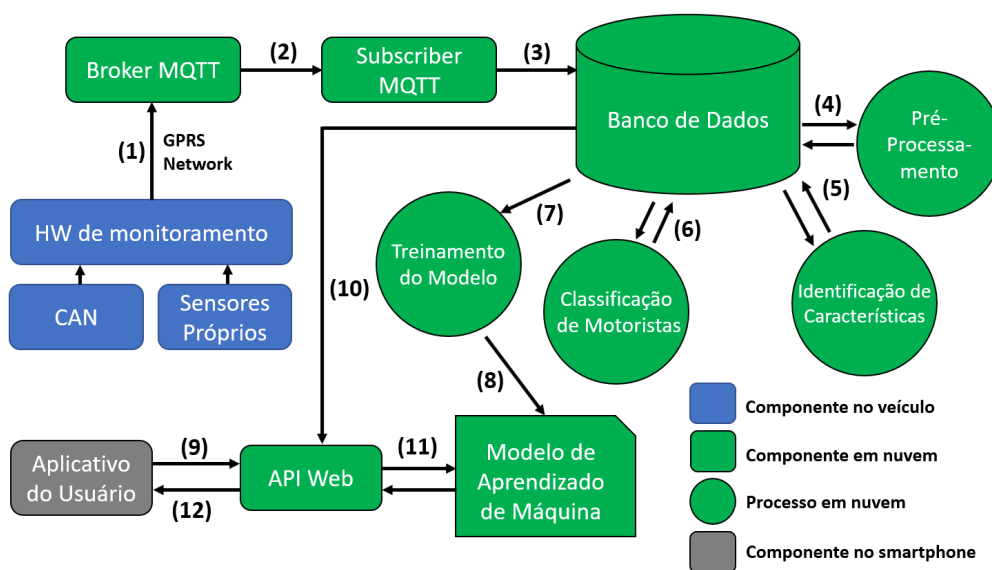


Figura 1. Arquitetura de software

O **hardware de monitoramento** é responsável por capturar os dados de telemetria através de sensores e da rede CAN. Esses dados são então enviados, via *publish* por GPRS (1), para um *Broker* MQTT, que é um servidor responsável por receber e transmitir mensagens entre dispositivos.

O *Subscriber* MQTT é um dispositivo que se inscreve para receber mensagens do *Broker* MQTT e, quando essas mensagens são recebidas (2), o *Subscriber* lê os dados e os grava em um banco de dados (3). No módulo de **pré-processamento**, os dados passam por um processo (4) onde são agrupados, normalizados, reamostrados e completados. As *features* são então extraídas (5) através do processo de **identificação de características**, que gera dois conjuntos de dados: um para classificar os motoristas e outro para analisar os

dados de cada viagem e extrair informações como veículo, motorista, velocidade média, variação de altitude, uso do ar-condicionado, modo de economia e consumo.

O processo de **classificação de motoristas** utiliza as *features* (6) para agrupar os motoristas em classes com o uso do algoritmo *K-Means*. Em seguida, um modelo de regressão é treinado (7) para prever o consumo de energia por trecho, levando em consideração o tipo de motorista e os dados do contexto da viagem. Esse modelo é chamado (11) por uma **API Web** que pode ser utilizada por um aplicativo de celular para estimar o consumo de energia para uma rota escolhida pelo usuário.

#### 4.1. Pré-processamento

O pré-processamento dos dados é executado uma vez ao dia, analisando os dados armazenados do dia anterior e salvando os dados processados no banco de dados. Para este processamento foram feitas algumas considerações: (i) cada viagem consiste de uma coleção de dados cronologicamente sequenciais, de um mesmo veículo e um mesmo motorista em um mesmo dia; (ii) o início e fim de cada viagem são demarcados quando os dados de velocidade forem zero e o pedal de freio e pedal do acelerador não estiverem pressionados, indicando carro em repouso; (iii) quando uma viagem compreender mais de um dia, é separada em mais de uma viagem; (iv) a cada período de um minuto sem dados, também é considerada uma nova viagem, dado não ser viável aferir os valores faltantes.

#### 4.2. Identificação de Características

Com a análise desses dados, são verificadas as melhores características (*features*) para identificar padrões de dirigibilidade, com respeito ao consumo instantâneo de bateria. Algumas métricas utilizadas na classificação de motoristas são: aceleração, inclinação, velocidade, posição dos pedais, entre outras. Para identificar esses padrões, como não é possível definir a quantidade de padrões e nem quais motoristas serão de cada classe à priori, são testadas diversas técnicas de aprendizagem não-supervisionada de classificação, como *K-means*, rede neurais profundas, redes neurais convolucionais, e outras.

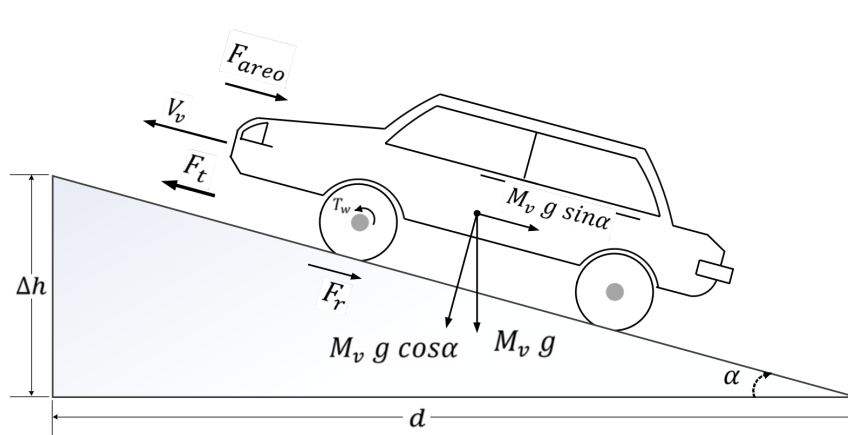


Figura 2. Forças atuando em um veículo em uma rampa - [Akl et al. 2019], Fig. 1

Outro conjunto de *features* é criado para utilização na estimativa de consumo, baseado no tipo de motorista e nos dados de contexto da viagem. Em [Akl et al. 2019], as forças que atuam em um carro em movimento são descritas, conforme Figura 2. Desta

forma, a força necessária para propelir o carro em um dado momento é calculada pela Equação 1, onde  $M_v$ ,  $g$  e  $\alpha$  são a massa do veículo, a aceleração gravitacional e a inclinação da estrada, respectivamente. A resistência do ar,  $F_{aero}$ , é correlacionada à velocidade [Hucho 1998] e a força gravitacional aplicada depende da inclinação.

$$F_t = F_{aero} + F_r + M_v g \sin \alpha \quad (1)$$

Adicionalmente, veículos elétricos tem um modo de economia, que funciona da maneira oposta ao modo esportivo em carros automáticos, gastando menos energia quando ligado. Além disso, o principal sistema auxiliar a drenar energia é o ar-condicionado, que rouba energia diretamente da bateria. A energia gasta é calcula pela Equação 2, onde  $E_t$  e  $E_a$  são, respectivamente, a energia necessária para gerar  $F_t$  por um período de tempo e a energia gasta em sistemas auxiliares neste mesmo período. Baseado nisso, algumas *features* que podem influenciar no consumo de energia foram selecionadas.

$$E_s = E_t + E_a \quad (2)$$

Essas *features* são descritas a seguir, juntamente com a forma como são calculadas, a cada trecho de dez quilômetros: (a) DIST: distância em quilômetros; (b) VEIC: identificação do veículo - cada veículo tem suas características e capacidade de carga da bateria; (c) COND: identificação da classe de motorista - para cada classe, existe uma forma de dirigir, mais ou menos econômica, dependendo ainda dos dados do percurso; (d) VARALT: diferença de altitude no trecho em metro por metro - calculada como a diferença de altitude no início e no fim do trecho dividido pelo seu tamanho em metros; (e) SPD: velocidade média no trecho, em metros por segundo - calculada como a distância percorrida no trecho pelo tempo gasto; (f) ARC: utilização do ar condicionado no trecho - 0 para desligado e 1 para ligado - valores intermediários indicam a porcentagem de utilização no trecho; (g) ECO: utilização do modo econômico no trecho - 0 para desligado e 1 para ligado - valores intermediários indicam a porcentagem de utilização no trecho; (h) CONS: percentual de bateria consumida no trecho, com precisão de 0,5 p.p., conforme informado pelo veículo.

### 4.3. Classificação de Motoristas

Consiste em ler as *features*, de um período máximo de um ano, gerando histogramas de 10 classes, divididas igualmente, por motorista e veículo. A conversão dos valores em histogramas foi utilizada para diminuir o número de valores de cada *feature*, visando simplificar o processamento do modelo de classificação, sem perder informações importantes, conforme [Fugiglando et al. 2019]. Usando os histogramas de cada motorista como entrada, um classificador não-supervisionado agrupa os motoristas em classes, salvando as tuplas motorista e classe em uma tabela do banco de dados (6).

Segundo [Alpaydin 2020], um dos algoritmos de classificação mais utilizados é o *K-Means*, que é um método não supervisionado de classificação e agrupamento de dados que analisa a distribuição dos dados e identifica centróides para gerar grupos com características semelhantes.

### 4.4. Treinamento do Modelo

Utilizando os dados de *features* (7), substituindo a identificação do motorista por sua classe, um modelo que faz a previsão do consumo, trecho a trecho, é gerado. Este modelo

analisa, levando em consideração o tipo de motorista e os dados do contexto da viagem, a porcentagem de bateria gasta naquele trecho. Somando-se o gasto trecho a trecho, tem-se o valor percentual de bateria gasta no trajeto total. A validação e escolha do algoritmo de treinamento, bem como os dados utilizados, são discutidos na Seção 5.

Este modelo é salvo como um arquivo na nuvem (8) para utilização na previsão do consumo trecho a trecho e esse treinamento pode ser realizado diariamente ou em períodos mais espaçados de acordo com os recursos computacionais disponíveis.

#### **4.5. Modelo de Aprendizado de Máquina**

O modelo é chamado diretamente pela API (11) e recebe: (a) a classificação do motorista e (b) os dados da viagem, como por exemplo, variação de altitude, velocidade, uso do ar-condicionado e modo de economia, trecho a trecho. Este módulo carrega o modelo de regressão linear salvo a partir do módulo Treinamento do Modelo e, para cada trecho da entrada, o modelo prevê o consumo. Todos os trechos, com seus respectivos consumos são retornados à API (11).

#### **4.6. API Web**

A API é utilizada por um aplicativo de celular que traça a rota de acordo com o ponto inicial do usuário e o destino escolhido. Esse aplicativo pode utilizar uma API de rotas para determinar os dados dos trechos da forma mais precisa e com maior amostragem possível. Esta API recebe as informações de rota e a identificação do motorista (9).

A classe do motorista é determinada consultando-se a tabela no banco de dados (10), de acordo com a classificação de motoristas. Então o modelo é chamado, com os dados de classificação do motorista e de viagem, retornando a estimativa de consumo trecho a trecho (11). Esta estimativa é então repassada ao aplicativo, para o planejamento da viagem e recomendação das paradas (12).

#### **4.7. Aplicativo do Usuário**

Este aplicativo é o responsável por conectar o motorista ao veículo. Tem, também, a função de traçar a rota de viagem, a partir da origem e destino selecionados, e recomendar as paradas para recarga no trajeto.

Depois da rota traçada por uma API de roteamento de viagem, os dados dos trechos e do motorista são passados à API Web (9), que retorna o gasto e autonomia prevista (12), para subsidiar a recomendação de postos de recarga.

### **5. Avaliação da Arquitetura Proposta**

Para avaliação desta arquitetura, os dados de telemetria de um veículo elétrico, da marca JAC Motors, foram utilizados. Esses dados foram coletados desde abril de 2021, através de um P&D ANEEL, que adquiriu o veículo para a DME Distribuição S.A. (DMED).

Em parceria com a PUC Minas Poços de Caldas, foi desenvolvido um equipamento por pesquisadores do curso de Engenharia Elétrica, para capturar, através da rede CAN, os seguintes dados de telemetria: (a) velocidade instantânea; (b) SOC - % de bateria do veículo elétrico; (c) odômetro; (d) tensão no motor; (e) corrente no motor; (f) inclinação do pedal de aceleração; (g) inclinação do pedal de freio; (h) ar condicionado (ligado/desligado); (i) modo de economia (ligado/desligado); (j) autonomia prevista. Além



desses dados, o equipamento também fornece dados de GPS, como: (a) posicionamento (latitude e longitude); (b) inclinação do veículo elétrico (giroscópio); (c) aceleração (acelerômetro); (d) altitude; (e) data; (f) hora.

Esse equipamento recolhe os dados e envia ao *Broker* MQTT, através de conexão GRPS, onde são lidos e salvos em um banco de dados. A rotina de coleta é constante, gerando dados em uma frequência de dois ou três segundos entre cada coleta.

### 5.1. Conjunto de Dados

O automóvel elétrico utilizado está a serviço dos funcionários da DMED, que anotam data, hora, motorista e quilometragem a cada utilização. Esses dados são adicionados ao banco de dados para compor o conjunto de dados.

Condutor	Viagens	Deslocamento em Km
1	363	1.812,2
2	24	131,1
3	110	417,0
4	3	2,0
N/I	59	210,8

**Tabela 1. Viagens após o pré-processamento**

Para este artigo, o banco de dados é composto de 151.180 registros de telemetria, coletados entre abril e outubro de 2021. Após o pré-processamento, o conjunto de dados relevantes passou a ser de 99.822 registros, com 2.573 km de deslocamento registrados, divididos em 559 viagens identificadas, conforme detalhados na Tabela 1. A Tabela 2 mostra os dados sumarizados por modo de economia e utilização do ar-condicionado.

Modo de Economia	Ar-Condicionado	Número de Registros
ligado	ligado	5.393
	desligado	62.027
desligado	ligado	1.503
	desligado	30.899

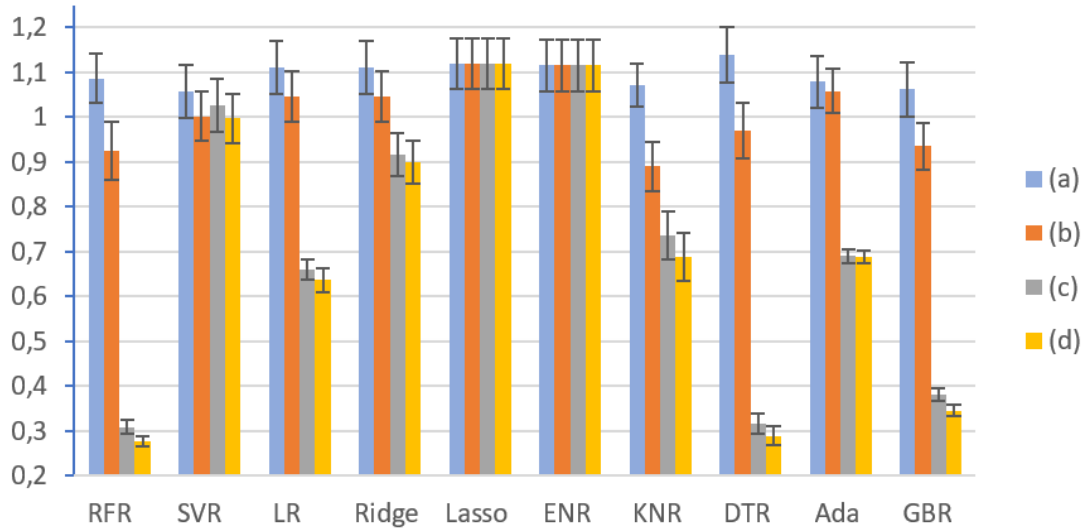
**Tabela 2. Dados após o pré-processamento**

Dos trechos de dez quilômetros, a fase de extração de *features* gerou 3.716 registros. Para extrair os trechos e maximizar o número de registros para melhorar o modelo, foram utilizadas janelas de 10 km para cada variação de 0,1 km de uma mesma viagem, gerando uma sobreposição de 99% de cada janela, conforme Figura 4. O tamanho da janela de 10 km foi escolhido para obter uma descarga de bateria relevante, e a variação de 0,1 km escolhida por ser a menor diferença de odômetro disponibilizada pelo sensor.

### 5.2. Resultados

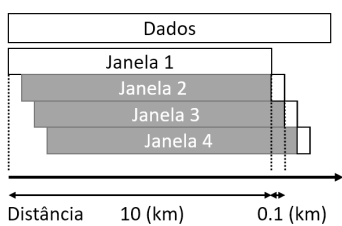
Para verificar se é possível utilizar essa arquitetura, aplicou-se dez técnicas de aprendizagem de máquina para regressão múltipla de função contínua: *Random Forest Regressor* (RFR), *Support Vector Regression* (SVR), *Linear Regression* (LR), *Ridge Regression* (Ridge), *Lasso Regression* (Lasso), *Elastic Net Regression* (ELR), *K-Neighbors Regressor*

(KNR), *Decision Tree Regressor* (DTR), *Ada Boost Regressor* (Ada) e *Gradient Boosting Regressor* (GBR). A avaliação inicial das técnicas se deu através de *10-fold cross validation*. O indicador utilizado nesta primeira avaliação foi o erro absoluto médio (MAE), com seu desvio padrão. Para cada técnica, foram avaliados os dados de quatro situações distintas: (a) sem informações de condutor ou contexto de viagem, (b) apenas com informações de condutor, (c) apenas com informações de contexto de viagem e (d) com informações de condutor e contexto de viagem. Os resultados são apresentados na Figura 3.

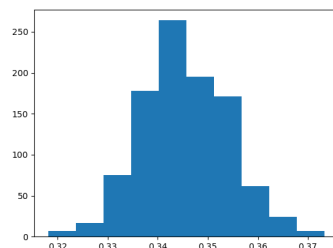


**Figura 3. Erro absoluto médio e desvio padrão**

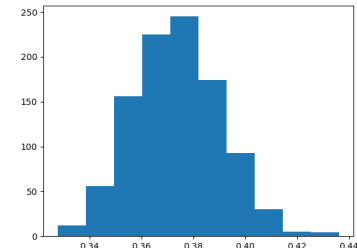
Nessa figura, verifica-se que os métodos RFR e DTR obtiveram melhores resultados com relação ao erro absoluto médio, indicando um erro de 0,277 com desvio padrão de 0,011 e 0,289, com desvio padrão de 0,022 respectivamente, para o caso onde todas as *features* foram utilizadas na regressão.



**Figura 4. Janela de 10 km e 99% de sobreposição**



**Figura 5. MAE Random Forest**



**Figura 6. MAE Decision Tree**

Ao analisar os resultados, observa-se também, que a inclusão dos dados de contexto de viagem diminui significativamente o erro absoluto médio, indicando que a utilização desses dados melhora a previsão realizada pelo modelo. Incluindo a identificação do condutor, nota-se também, uma diminuição no erro, porém de forma menos significativa. Se apenas os dados de velocidade média são utilizados, o modelo tem erros parecidos entre os diversos modelos, porém mais do que o dobro da precisão de 0,5 p.p. do medidor de carga da bateria.

A partir dessa análise, o método de reamostragem de Bootstrap foi aplicado no conjunto de dados "d", para medir o erro médio da população. Para essa etapa, foram testados os métodos que obtiveram melhor resultado na avaliação anterior (RFR e DTR). Os resultados são apresentados nas Figuras 5 e 6. Considerando o intervalo de confiança de 95%, os erros médios absolutos dos métodos podem ser vistos na Tabela 3. Os erros calculados ainda são menores do que a ordem de grandeza da medição da carga da bateria e indicam o método RFR como o melhor método para prever o consumo.

Random Forest Regressor	0,329 a 0,363
Decision Tree Regressor	0,342 a 0,407

**Tabela 3. Erro absoluto médio com intervalo de confiança de 95%**

Posteriormente, para testar a acurácia da previsão, foi utilizado o OpenStreetMap<sup>1</sup> para traçar uma rota entre a garagem em Poços de Caldas e uma estação de recarga em Jundiá, a 221 km de distância, e outra rota retornando de um posto de recarga em Limeira para a mesma garagem, com 170 km. Usando a API, vários trechos de 1 km, com a velocidade média e a variação de altitude, foram extraídos e aplicados ao modelo para estimar a quantidade de bateria gasta. Os resultados podem ser vistos nas Tabelas 4 e 5.

Motorista	Modo de Economia	Ar-Condicionado	Gasto de Bateria
#1	ligado	ligado	75,96%
		desligado	75,47%
	desligado	ligado	67,36%
		desligado	65,77%
#3	ligado	ligado	77,34%
		desligado	71,89%
	desligado	ligado	69,41%
		desligado	63,37%

**Tabela 4. Consumo de bateria simulado para a rota de ida**

Motorista	Modo de Economia	Ar-Condicionado	Gasto de Bateria
#1	ligado	ligado	71,11%
		desligado	69,81%
	desligado	ligado	60,98%
		desligado	58,57%
#3	ligado	ligado	73,30%
		desligado	68,31%
	desligado	ligado	63,89%
		desligado	58,21%

**Tabela 5. Consumo de bateria simulado para a rota de retorno**

Por fim, depois da simulação realizada, o motorista número 3 fez a viagem, seguindo a rota traçada anteriormente, com o ar-condicionado desligado e o modo de economia ligado. O carro iniciou a viagem com 96% de bateria e chegou ao destino com

<sup>1</sup>Disponível em <https://www.openstreetmap.org/>

22%, gastando um total de 74%. A simulação calculou um gasto de 71,89%, resultando em uma diferença de 2,11 p.p.. Para o mesmo trajeto e mesmas condições, a simulação com o motorista número 1 resultou em 75,47%. Para a outra rota, com ar desligado e modo de economia ligado, o motorista 3 saiu com 85% de bateria e chegou com 17%, gastando 68% de bateria. Outra vez, comparando com o simulado, que calculou 68,31%, o resultado foi muito bom, com uma diferença de apenas 0,31 p.p. Assim, entende-se que esses resultados podem contribuir com o planejamento de viagem de cada motorista.

## 6. Conclusão

Com a utilização de carros elétricos no mundo, os problemas com a baixa autonomia e contratempos com pontos de recarga irão aumentar. A correta previsão de autonomia e a recomendação destes pontos de recarga podem melhorar a experiência dos usuários de veículos elétricos e contribuir com a eletrificação da logística de distribuição.

A arquitetura apresentada se propõe a contribuir com a diminuição desse problema utilizando dados de telemetria para classificar motoristas de acordo com seu modo de dirigir e dados dos trechos da viagem planejada para prever consumo e autonomia, com informações que podem ser utilizadas para recomendar postos de recarga ao longo da viagem. Os modelos de classificação de motoristas e de previsão de consumo são implementados através de técnicas de aprendizagem de máquina.

Após a análise dos dados, conclui-se que é possível utilizar o método de regressão de *Random Forest* para prever a quantidade de bateria gasta em um determinado trecho, quando o condutor e os dados de contexto são conhecidos. Através de aplicativos de roteirização, é possível conseguir as informações como variação de altitude e velocidade média em determinado trecho. As informações de motorista e utilização do ar-condicionado ou modo de economia, assim como a origem e o destino da viagem, devem ser fornecidas no aplicativo. Entende-se que o objetivo principal deste trabalho, verificar se a arquitetura proposta é viável, foi atingido.

Verificou-se também que a qualidade da previsão depende da precisão dos dados. Como o valor da carga da bateria no veículo tem variação de 0,5 p.p., o valor de erro absoluto médio de 0,346 p.p. é um valor aceitável, mas que necessita ser melhorado para melhores previsões. No cálculo de valores gastos em trajetos conhecidos, o erro de 2,11 p.p e 0,31 p.p não foram grandes, gerando uma boa predição de autonomia. Aumentar a quantidade de dados analisados e a precisão do gasto de bateria por quilômetro deverá aumentar a acurácia da predição, já que podemos ver que o modelo não prevê muito bem com o modo de economia desligado. Isto pode estar relacionado ao *overfitting*, já que a quantidade de dados com o modo de economia ligado é mais do que o dobro dos dados com o modo de economia desligado e necessita de verificação.

Este trabalho não verificou se o *K-Means* seria o melhor algoritmo para classificar os motoristas pelo modo de dirigir e, tampouco, chegou a analisar a quantidade ótima de classes de motoristas a utilizar, devido ao número restrito de motoristas nos dados coletados. Também não houve nenhum teste quanto à janela de dados usada para treinar o modelo de regressão, por conta de os dados coletados não abrangerem um período muito grande. Estes seriam objeto de análises futuras. Ainda como trabalhos futuros seria necessário verificar como o sistema se comportará com os dados fornecidos pelas APIs de roteirização. Para isso, depois de planejado o trajeto e da autonomia calculada

pelo modelo, deve-se verificar se a autonomia prevista foi condizente com a distância percorrida até o ponto de recarga recomendado. Outro ponto a ser verificado é se as subidas e descidas dentro de um trecho têm grandes impactos no consumo. Também pode ser interessante verificar se outros dados de contexto, como temperatura, horários da viagem e coordenadas, por exemplo, influenciam nas previsões de consumo.

## Agradecimentos

Este trabalho foi desenvolvido com o apoio da FAPEMIG (PPM-XII), DME Distribuição S.A. (DMED) e do Programa de Pesquisa e Desenvolvimento da Agência Nacional de Energia Elétrica (ANEEL).

## Referências

- Akl, M. M., Ahmed, A. A., and Rashad, E. E. M. (2019). A wide component sizing and performance assessment of electric drivetrains for electric vehicles. In *2019 21st International Middle East Power Systems Conference (MEPCON)*, pages 834–839. IEEE.
- Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.
- ANATEL (2022). Anatel - infraestrutura. <https://informacoes.anatel.gov.br/paineis/infraestrutura/>, Acessado em 07/01/2022.
- Arabie, P., Hubert, L., and De Soete, G. (1996). *Clustering and classification*. World Scientific.
- Bailey, C., Jones, B., Clark, M., Buck, R., and Harper, M. (2022). Electric vehicle autonomy: Realtime dynamic route planning and range estimation software. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, pages 2696–2701.
- Bolovinou, A., Bakas, I., Amditis, A., Mastrandrea, F., and Vinciotti, W. (2014). Online prediction of an electric vehicle remaining range based on regression analysis. In *2014 IEEE International Electric Vehicle Conference (IEVC)*, pages 1–8. IEEE.
- Chen, M.-S., Hwang, C.-P., Ho, T.-Y., Wang, H.-F., Shih, C.-M., Chen, H.-Y., and Liu, W. K. (2019a). Driving behaviors analysis based on feature selection and statistical approach: a preliminary study. *The Journal of Supercomputing*, 75(4):2007–2026.
- Chen, S., Sun, W., Li, Y., and Shi, L. (2019b). On the relationship between energy consumption and driving behavior of electric vehicles based on statistical features. In *2019 Chinese Control Conference (CCC)*, pages 3782–3787.
- Chen, W., Lin, Y., and Chen, W. (2019c). Comparisons of machine learning algorithms for driving behavior recognition using in-vehicle can bus data. In *2019 Joint 8th International Conference on Informatics, Electronics Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision Pattern Recognition (icIVPR)*, pages 268–273.
- Delnevo, G., Di Lena, P., Mirri, S., Prandi, C., and Salomoni, P. (2019). On combining big data and machine learning to support eco-driving behaviours. *Journal of Big Data*, 6(1):64.
- Eberly, L. E. (2007). Multiple linear regression. *Topics in Biostatistics*, pages 165–187.

- Fugiglando, U., Massaro, E., Santi, P., Milardo, S., Abida, K., Stahlmann, R., Netter, F., and Ratti, C. (2019). Driving behavior analysis through can bus data in an uncontrolled environment. *IEEE Transactions on Intelligent Transportation Systems*, 20(2):737–748.
- Gönül, Ö., Duman, A. C., and Güler, Ö. (2021). Electric vehicles and charging infrastructure in turkey: An overview. *Renewable and Sustainable Energy Reviews*, 143:110913.
- Gorner, M. (2020). Electric vehicles: Analysis - tracking report: June 2020. Disponível em: <https://www.iea.org/reports/electric-vehicles>. Acesso em: 19/07/2020.
- Hucho, W. (1998). Commercial vehicles. *Aerodynamics of Road Vehicles, fourth ed., SAE International, Warrendale, Pa., United States*, pages 131–237.
- Jia, S., Hui, F., Li, S., Zhao, X., and Khattak, A. J. (2020). Long short-term memory and convolutional neural network for abnormal driving behaviour recognition. *IET Intelligent Transport Systems*, 14(5):306–312.
- Krishnamurthy, V., Nezafati, K., Bae, J., Gursoy, E., Zhong, M., and Singh, V. (2019). Classification of driving behavior events utilizing kinematic classification and machine learning for down sampled time series data. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 3789–3796.
- Lebrouhi, B., Khattari, Y., Lamrani, B., Maaroufi, M., Zeraouli, Y., and Kousksou, T. (2021). Key challenges for a large-scale development of battery electric vehicles: A comprehensive review. *Journal of Energy Storage*, 44:103273.
- McKinsey (2023). O futuro da mobilidade no brasil: uma rota para eletrificação. <https://www.mckinsey.com.br/our-insights/all-insights/o-futuro-da-mobilidade-no-brasil>, Acessado em 27/01/2023.
- Noussan, M. (2020). *Public EV charging in Italy: policies, business models and future scenarios*. MDPI.
- Ping, P., Qin, W., Xu, Y., Miyajima, C., and Takeda, K. (2019). Impact of driver behavior on fuel consumption: Classification, evaluation and prediction using machine learning. *IEEE Access*, 7:78515–78532.
- Siarni, M., Naderpour, M., and Lu, J. (2020). A mobile telematics pattern recognition framework for driving behavior extraction. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–14.
- Stanton, N. A. and Allison, C. K. (2020). Driving towards a greener future: an application of cognitive work analysis to promote fuel-efficient driving. *Cognition, Technology & Work*, 22(1):125–142.
- Velandia Vargas, J. E., Seabra, J. E. A., Cavaliero, C. K. N., Walter, A. C. S., Souza, S. P., and Falco, D. G. (2020). The new neighbor across the street: An outlook for battery electric vehicles adoption in brazil. *World Electric Vehicle Journal*, 11(3).
- Wang, G., Zhang, Y., Fang, Z., Wang, S., Zhang, F., and Zhang, D. (2020). Faircharge: A data-driven fairness-aware charging recommendation system for large-scale electric taxi fleets. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(1).