

PERFIL DE CONDUÇÃO DE MOTORISTAS: ANÁLISE POR MEIO DE ALGORITMOS DE RECONHECIMENTO DE PADRÕES

WILSON BORBA DA ROCHA NETO*, JOSÉ MARIA PIRES DE MENEZES JR.*

**Departamento de Engenharia Elétrica, Universidade Federal do Piauí (UFPI) - Campus Universitário Ministro Petrônio Portella - Bairro Ininga - CEP 64049-550 - Teresina - PI - Brazil*

Emails: wborbaneto@gmail.com, josemenezesjr@ufpi.edu.br

Abstract— This work aims to present the development of an automotive telemetry system capable of capturing data of the electromechanical system of a vehicle, so that it can describe the behavior of the automobile and its driver. In this way, this work proposes to construct a data acquisition system through the OBD II diagnostic system, to construct primitives that allow to store this information remotely and, finally, to use algorithms of pattern recognition, in order to evidence models of conduction of a car, classifying drivers' driving profiles so that it is possible to suggest changes in behavior and prevent mechanical failures. The results obtained suggest that the attributes used in the classification allow a hit rate higher than 90 %, especially the ELM network in computational cost.

Keywords— Automotive Telemetry, Classification, OBD, Extreme Learning Machine.

Resumo— Este trabalho tem como finalidade apresentar o desenvolvimento de um sistema de telemetria automotiva capaz de capturar dados do sistema eletromecânico de um veículo, para que estes possam descrever o comportamento do automóvel e de seu condutor. Desta forma, este trabalho propõe construir um sistema de aquisição de dados através do sistema de diagnóstico OBD II, construir primitivas que permitam guardar essas informações remotamente e, por fim, empregar algoritmos de reconhecimento de padrões, com o intuito de evidenciar modelos de condução de um automóvel, classificando os perfis de condução dos motoristas para que seja possível sugerir mudanças de comportamento e prevenir falhas mecânicas. Os resultados obtidos sugerem que os atributos utilizados na classificação permitem uma taxa de acerto superior a 90 %, com destaque para a rede ELM em custo computacional.

Palavras-chave— Telemetria automotiva, Classificação, OBD, Máquina de aprendizado extremo.

1 Introdução

A telemetria é um sistema eletrônico de monitoramento construído para rastrear, medir ou enviar comandos a algum equipamento que esteja distante ou em constante movimento. Assim, a telemetria pode ser entendida como a medição, processamento de dados e tratamento de informações em um ponto distante do de captura. O monitoramento em tempo real das funções eletromecânicas de um veículo tornou-se possível graças ao avanço da eletrônica automotiva, iniciada pela Bosch em 1917, com o desenvolvimento da injeção eletrônica (Bosch, 2011).

Surgindo de um simples controle da injeção de combustível, a atual rede eletrônica automotiva é tão grande que, do ponto de vista da ciência da computação, o carro é um sistema em tempo real, complexo e distribuído, com até 60 microcontroladores que se comunicam via, aproximadamente, 2500 sinais de diferentes tipos (Steinbach et al., 2011). Esta evolução da eletrônica veicular permite a leitura de um grande volume de informações, as quais podem ser acessadas através da interface *On Board Diagnostic* (OBD), atualmente na versão OBD II. A TSBEC (2004) define o OBD II como um padrão com conexão e pinos especificados para cada função, protocolos eletrônicos, códigos de diagnósticos de problemas e lista de parâmetros a serem monitorados.

O OBD II é capaz de acessar informações

instantâneas que envolvem velocidade, aceleração, rotações por minuto, posição do pé do acelerador, carga no motor, consumo de combustível, temperatura do motor, distância percorrida, tensão na bateria e ainda realizar o diagnóstico de problemas dos sistemas eletromecânicos do automóvel (Szalay et al., 2015).

Por estar em constante movimento, a Telemetria é uma alternativa na captura de informações de um veículo. Esta auxilia em ações voltadas para manutenção, consumo de combustível e na redução de riscos de acidentes. Porém, apenas esta não toma ação nenhuma. Existe a necessidade de organizar dados, encontrar padrões, associações, mudanças e anomalias relevantes. Para tal, é necessário um conjunto de ferramentas e técnicas, que através do uso de algoritmos de aprendizagem ou classificadores baseados em redes neurais e estatísticas, são capazes de explorar um conjunto de dados, extraíndo ou ajudando a evidenciar padrões nestes e auxiliando na descoberta de conhecimento.

Pode-se identificar um motorista com um perfil pouco cuidadoso, que tenha um comportamento agressivo no trânsito, coletando apenas algumas informações do automóvel. Como exemplo, um motorista apressado ou não muito cuidadoso tem geralmente o perfil de freadas bruscas. Observar a frequência e intensidade das freadas já seria suficiente para a diferenciação de tipos de condução, mas dificilmente é possível definir o compor-

tamento de um condutor observado apenas uma única variável. Para montar padrões e classificar perfil de condução é necessário, então, um conjunto de observações. Quanto mais informações forem coletadas, melhor será a descoberta de conhecimento, sendo assim, por exemplo, analisar a intensidade das acelerações também seria importante para reconhecer um comportamento agressivo no trânsito.

Dada a importância do monitoramento, do diagnóstico e da evolução dos sistemas ao longo do tempo, o objetivo deste estudo é desenvolver um sistema embarcado microcontrolado integrado ao automóvel, capaz de conectar-se a um servidor web, utilizando a rede de telefonia móvel para o envio de informações. Este possibilita criação de um banco de dados online com as informações obtidas do OBD II, o qual será utilizado no algoritmo de reconhecimento de padrões. Os algoritmos utilizados na classificação dos perfis de condução são: classificador clássico K-vizinhos mais próximos (do inglês: K nearest neighbors, K-NN) e a rede neural Máquina de Aprendizado Extremo.

No tocante a construção do artigo, a seção 2 apresenta uma descrição detalhada do sistema OBD II. Uma breve explicação do sistema projetado para captura é abordada na seção 3. A seção 4 aborda a definição dos classificadores utilizados, introduzindo conceitos que facilitam a compreensão do proposto na seção 5, a qual apresenta os resultados alcançados. Conclusão e trabalhos futuros, trazidos na seção 6, fecham o documento.

2 OBD II

Para que se possa receber informações do veículo, este deve possuir uma rede de comunicação interna capaz de monitorar seus sistemas eletromecânicos e atender às requisições realizadas por um agente externo. Além disso, é necessário um sistema capaz de enviar tais requisições pelo terminal OBDII, as quais são comandos de diagnóstico ou parâmetros de informações, conhecidos como *Parameter Identification* (PID), padronizados pela SAE (2012) e ISO (2015).

Os PIDs são divididos em um total de 10 serviços. Cada um possui uma série de códigos que podem ser utilizados para atividades específicas (SAE, 2012). Neste trabalho, utiliza-se majoritariamente o serviço \$01, o qual contém os PIDs relacionados aos dados de funcionamento do automóvel. As respostas obtidas para cada PID, extraídas do sistema OBD II, são sequências de diferentes arranjos para cada um dos parâmetros de entrada. A Figura 1 exemplifica o processo de captura e decodificação para uma das informações utilizadas neste trabalho. Note que os símbolos H_1 e H_2 representam a sequência de hexadecimais recebida como resposta útil para a requisição.

De todas as informações disponibilizadas pelo

sistema OBD II, selecionaram-se aquelas que, de acordo com a opinião do especialista, poderiam ser utilizadas na avaliação da situação do automóvel e do perfil de condução do motorista (veja a Tabela 1). Além disso, os dados escolhidos também podem ser utilizados no cálculo do consumo de combustível, informação não disponibilizada pelo OBD II - como será abordado na Seção 3.

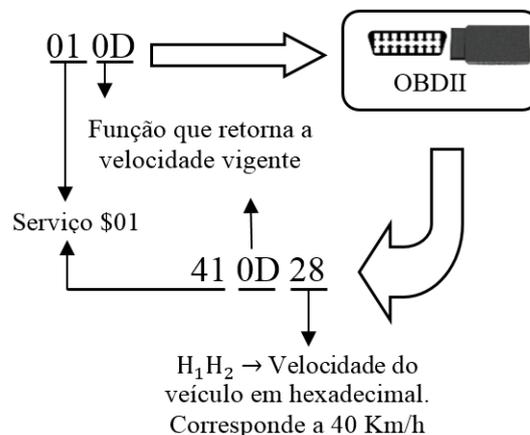


Figura 1: Exemplo de comunicação utilizando um PID

Tabela 1: Parâmetros escolhidos para captura

| PID | Informação Correspondente | Acrônimos |
|------|------------------------------------|------------|
| 010D | Velocidade(Km/h) | <i>Vel</i> |
| 010C | Rotações por minuto(RPM) | <i>RPM</i> |
| 0110 | Fluxo de massa ar/combustível(g/s) | <i>MAF</i> |
| 0111 | Abertura do pedal de aceleração(%) | <i>Acl</i> |
| 0104 | Carga no motor(%) | <i>CMo</i> |

3 Sistema de Telemetria

Para capturar e salvar as informações em um servidor utilizando o OBD implementou-se um sistema de telemetria embarcado no veículo. Este é capaz de organizar em um banco de dados online todas as informações destacadas na Tabela 1. Além disso, o uso do GPS permite a identificação e localização das capturas, o que adiciona ainda mais variáveis ao processamento.

Apesar de já existirem opções no mercado que supram essa necessidade, optou-se por projetar um sistema próprio, já que o controle do fluxo de dados e modificações do sistema se dão de forma livre em servidores e protótipos próprios. O esquemático do sistema, que permite a visualização de suas funções e componentes, pode ser visto na Figura 2.

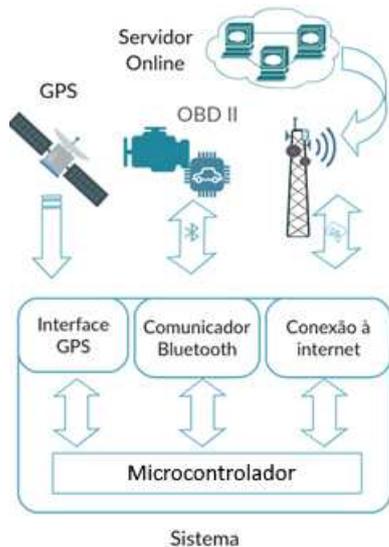


Figura 2: Sistema de telemetria desenvolvido.

O sistema de telemetria desenvolvido realiza uma captura a cada 0,5 segundos. Cada captura consiste em um vetor, ilustrado na Figura 3, composto pelos dados do veículo e a data/local nos quais estes foram capturados. Além das informações capturadas, acrescenta-se ao vetor o consumo de combustível obtido através de uma fórmula experimental (veja a Equação 1) baseada na análise dimensional das variáveis e de correções realizadas por testes empíricos.

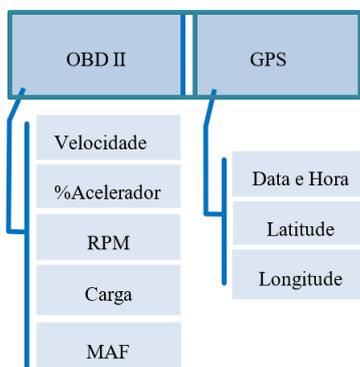


Figura 3: Vetor de dados enviado ao servidor.

$$C\left(\frac{L}{Km}\right) = \frac{3600 \times MAF}{\frac{Ar}{Combustivel} \times Densidade \times Vel} \quad (1)$$

Toda a informação enviada ao servidor é disponibilizada no banco de dados, para posterior processamento e implementação do algoritmo de classificação. Os arquivos de cada veículo são salvos separadamente e identificados pelo número de série do sistema de telemetria correspondente. Além disso, para uma análise mais didática, as in-

formações são disponibilizadas em gráficos e tabelas em um endereço eletrônico, permitindo que o usuário do sistema acompanhe as estatísticas das capturas realizadas em seu veículo

4 Classificação

O reconhecimento de padrão envolve, essencialmente, as seguintes etapas de processamento: pré-processamento, extração de parâmetros e classificação. O pré-processamento, tem o objetivo de eliminar dados desnecessários fazendo com que a entrada apresente apenas dados relevantes para o reconhecimento do objeto em análise.

A extração de parâmetros, também denominada etapa de seleção de atributos, consiste em análise dos dados de entrada a fim de extrair e derivar informações úteis para o processo de reconhecimento. O estágio final do reconhecimento de padrão é a classificação, ou etapa de tomada de decisão, que analisa as características extraídas dos dados de entrada e declara o objeto em análise como pertencente a uma determinada categoria (MATOS, 2008). Classificações de padrões investiga o desenvolvimento de técnicas que permitem a construção de sistemas automáticos de aprendizado. Normalmente ele é dado em duas fases: treinamento, onde é aprendida uma descrição geral do conceito; e teste, que utiliza a descrição aprendida para categorização de novos dados passados ao sistema.

4.1 Definição de classes e atributos

O sistema OBD possui dezenas de informações que podem ser capturadas de um veículo. Dessa forma tornou-se necessário especificar e analisar quais destas seriam as mais importantes para o processo de classificação do comportamento do condutor. Inicialmente essa escolha foi feita baseada no senso comum, utilizando a lógica para decidir quais informações teriam mais impacto no perfil do condutor - conforme visto na Seção 2.

Na Figura 4 é apresentada a relação entre as variáveis escolhidas. Nela pode-se observar a relação importante entre o sensor MAF e os atributos *Acl* e *RPM*, que se justifica, pois, a maior entrada de ar no motor é diretamente ligada a maior pressão no pedal do acelerador e a maior rotação do motor. Pode-se verificar também a baixa correlação entre *Vel* e *Acl*, pois essa relação tem influência principalmente do câmbio.

Optou-se pela utilização de três classes: Agressiva, Moderada e Leve. Para validar funcionamento do sistema de telemetria e compor o banco de dados, o automóvel percorreu um circuito de teste controlado, isto é, um circuito onde foi possível preservar as mesmas características de via para cada modo de condução. O percurso escolhido (i) foi realizado em um perímetro urbano,

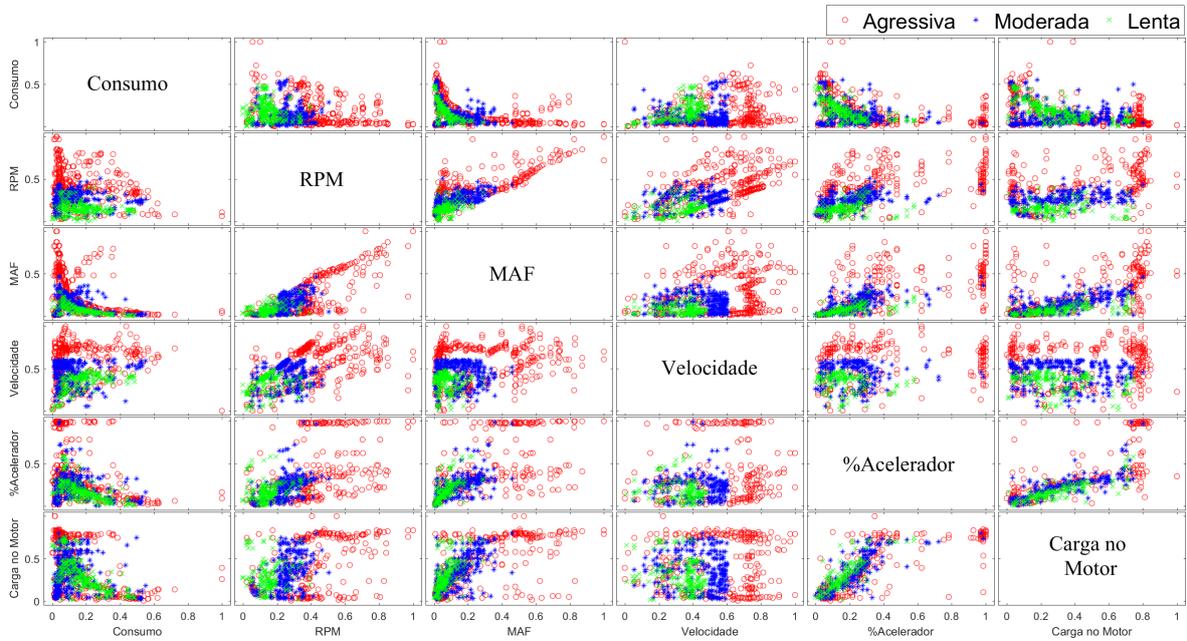


Figura 4: Relação de correlação entre as variáveis envolvidas

em pista asfaltada sem interferência de outros veículos; (ii) com o tempo ensolarado e alta visibilidade; (iii) possui um redutor de velocidade e uma curva acentuada; (iv) e tem extensão total de 1,5 quilômetros. O motorista que realizou os testes foi orientado a percorrê-lo, em diferentes testes, de maneira (i) calma, respeitando o desenvolvimento do veículo e com aceleração sutil; (ii) normal, como dirigiria respeitando o limite de velocidade urbano; (iii) e agressiva, com uma aceleração alta e com o objetivo de terminar o percurso o mais rápido possível.

Em cada uma das situações propostas ao motorista o tempo necessário para percorrer o percurso definido é diferente. Por conta disso, a fim de coletar uma mesma quantidade de capturas para cada classe, selecionou-se um número de dados que representasse ao menos uma volta completa no percurso escolhido. Para isso, os dados foram selecionados aleatoriamente, até que se conseguisse o número desejado. Dessa forma, possibilitou-se que todas as classes apresentassem os mesmos eventos, como partida do carro, curvas e paradas, e que estes fossem mantidos em todas as classes.

Como uma forma de conhecer as classes utilizadas nesse projeto, optou-se por representar num gráfico de uma maneira bidimensional dois atributos coletados, que, nesse caso, foram RPM e Consumo. Assim, chegou-se ao gráfico apresentado na Figura 5, o qual possui a representação de pontos instantâneos de cada captura. Verifica-se uma certa separação entre as classes, com a formação de clusters, mesmo utilizando apenas dois atributos.

Deve-se destacar que a separação entre as

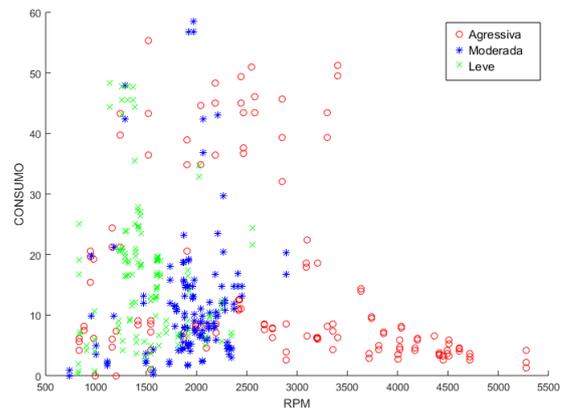


Figura 5: Relação de correlação entre as variáveis envolvidas

classes é de menor intensidade em pontos próximos à origem. Essa região descreve condições de parada, passagem por quebra-molas e curvas; comportamento esse, presente em todos os modos de condução. Porém essa região é mais densamente ocupada pela classe Leve, pois nesse modo de condução o veículo prioriza baixo RPM.

4.2 Classificação por K-NN

Este trabalho faz uso do classificador de K-vizinhos mais próximos, um dos mais simples e utilizados em tarefas de reconhecimento de padrões, pois permite uma busca por similaridade entre os dados de fácil compreensão. Ele foi proposto por Fukunaga (1975) é um classificador sub-ótimo que cria fronteiras de decisões complexas a partir de um conjunto de padrões

de treinamento com classes conhecidas a priori (MATOS, 2008). Mais precisamente, para um dado vetor característico no conjunto teste, os vetores de K mais próximos no conjunto de treino são selecionados (de acordo com alguma medida de distância, sendo que neste trabalho é utilizado a distância euclidiana) e o vetor de características de teste é atribuído ao rótulo mais representado na classe dos K vizinhos (Miche et al., 2010). Por conta dessas características, o K-NN mostra-se uma ferramenta útil na classificação de clusters de dados, como os apresentados na Figura 5, justificando sua escolha para este trabalho.

Classificar um exemplo desconhecido com o algoritmo K-NN pode ser resumido nos seguintes passos:

- Calcular a distância entre o exemplo desconhecido e os outros exemplos do conjunto de treinamento;
- Identificar os K vizinhos mais próximos;
- Utilizar o rótulo da classe de vizinhos mais próximos para determinar o rótulo de classe do exemplo desconhecido em uma espécie de votação.

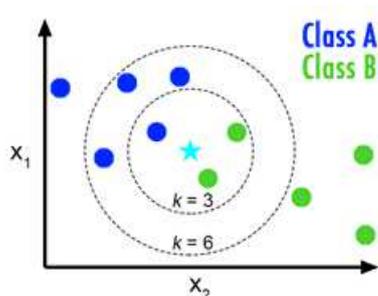


Figura 6: Classificação com o K-NN

4.3 Máquina de aprendizado extremo

Um novo paradigma de projeto de redes neurais vem sendo proposto sob a alcunha de Projeções Aleatórias (Random Projections) (Huang et al., 2006). O interesse neste paradigma por parte da comunidade científica vem aumentando devido à forma singular de projeto de tais arquiteturas, uma vez que esta não requer o treinamento dos neurônios não-lineares da camada oculta, chamada neste contexto de projeções. Apenas os neurônios da camada de saída têm seus pesos ajustados, o que pode ser realizado por métodos recursivos e não-recursivos. Estas arquiteturas também fazem uso geralmente de uma grande quantidade de neurônios ocultos, a fim de compensar a inicialização aleatória dos pesos dos neurônios ocultos.

Dentro do arcabouço geral das redes baseadas em projeções aleatórias, destaca-se a rede Máquina de Aprendizado Extremo (do inglês: Ex-

treme Machine Learning, ELM) (Huang et al., 2006), uma arquitetura feedforward. A rede ELM utiliza uma simples camada de uma rede neural feedforward, sem realimentação. Os pesos da camada oculta são iniciados aleatoriamente e não precisam ser atualizados, ficando a cargo apenas do cálculo dos pesos da camada de saída.

A Figura 7 ilustra a rede ELM com a entrada $x(n)$ definida de acordo com os atributos de entrada, com a matriz W representando os pesos da camada não-linear que permanecem fixos durante o treinamento e a matriz m representando os pesos dos neurônio de saída. Por simplicidade, os neurônios da camada de saída tiveram seus pesos ajustados pelo método dos mínimos quadrados, também conhecido como método da pseudoinversa.

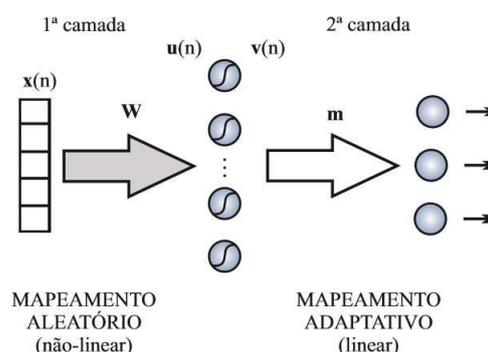


Figura 7: Dados de 2 atributos capturados divididos em 3 classes.

5 Resultados

O intuito deste trabalho é projetar um sistema capaz de classificar o modo de condução de determinado veículo com a utilização do algoritmo de classificação K-NN e da rede ELM, duas ferramentas de classificação com princípios e complexidade distintos. Além disso, é feita uma análise e comparação dos resultados obtidos por estas duas ferramentas.

5.1 Algoritmo de classificação

Com as classes e atributos definidos, foram coletados dados após o carro percorrer o trajeto algumas vezes. Ao fim, foram feitas 600 capturas/amostras para cada uma das três classes descritas anteriormente: Agressiva, Moderada e Leve. A descrição do conjunto de dados está detalhada na Tabela 2. O método de validação holdout foi escolhido para verificar o desempenho da classificação. Assim os blocos foram postos de forma aleatória e separados em matrizes de treino e matrizes de teste. Optou-se por uma divisão de 80%/20%, ou seja, os primeiros 80% das amostras foram utilizados para o treinamento e o restante foi utilizado para os testes.

Tabela 2: Descrição do conjunto de dados

| | |
|---------------------|---|
| Tipo | Númerico (quantitativo) |
| Classes | 3 (Agressiva, Moderada e Leve) |
| Atributos | 7 (RPM, MAF, velocidade, aceleração, temperatura, carga no motor e consumo) |
| Quantidade de dados | 1.800 (600 para cada classe) |

5.2 Resultados da classificação

Cada rodada de treino/teste foi repetida 1000 vezes, já que as matrizes são embaralhadas a cada rodada, permitindo diferentes resultados de classificação. Assim, obteve-se resultados médios de acertos e erros do classificador.

Primeiramente foram feitos testes de variação do número de vizinhos mais próximos do classificador K-NN. O gráfico apresentado na Figura 8 mostra o efeito de tal variação na porcentagem de acerto da classificação. Pode-se verificar que a partir de 5 vizinhos o classificador perde desempenho, não gerando melhor taxa de classificação com o aumento do número de vizinhos. O valor máximo de taxa de acerto se deu com 1 vizinho, aproximadamente 99% de taxa de acerto.

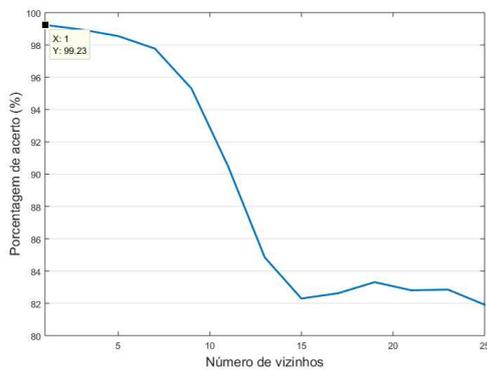


Figura 8: Variação do número de vizinhos na taxa de acerto do classificador do KNN.

Na Figura 9 encontram-se o resultado comparativo da rede ELM variando a quantidade de neurônios na primeira camada oculta. Pode-se destacar o melhor desempenho da rede ELM com valores altos na quantidade de neurônios. O melhor resultado se deu com quase 5.000 neurônios. Esse teste sugere que a taxa de acerto poderia melhorar com o aumento do número de neurônios, mas isso foi rejeitado pelo custo computacional elevado.

A seguir, nas Tabelas 3 e 4, temos a matriz de confusão dos classificadores KNN e da rede ELM. Nestas, os termos Agr, Mod e Lnt correspondem a Agresiva, Moderada e Lenta, respectivamente. Essa matriz informa a quantidade de acerto e erro

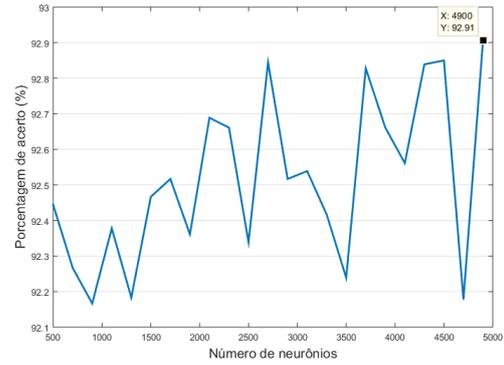


Figura 9: Variação do número de neurônios a taxa de acerto do classificador ELM.

de classificação para cada classe tratada nesse trabalho. Pode-se observar a menor taxa de acerto para a classe Agressiva, pois essa classe de perfil de condução possui condições de parada, passagem por quebra-molas e cruzamentos; comportamento esse, presente em todos os modos de condução. As outras classes possuem taxas de acerto acima de 95 %, demonstrando que os atributos selecionados conseguem separar melhor essas classes.

Tabela 3: Matriz de confusão do classificador K-NN

| | | Classe Verdadeira | | |
|-----------------|-----|-------------------|---------|--------|
| | | Agr | Mod | Lnt |
| Classe Estimada | Agr | 99.16% | 0% | 0.84% |
| | Mod | 0.84% | 100.00% | 0% |
| | Lnt | 0% | 0% | 99.16% |

Tabela 4: Matriz de confusão da rede ELM

| | | Classe Verdadeira | | |
|-----------------|-----|-------------------|--------|--------|
| | | Agr | Mod | Lnt |
| Classe Estimada | Agr | 90.00% | 0.83% | 0.83% |
| | Mod | 2.50% | 95.00% | 5.00% |
| | Lnt | 7.50% | 4.16% | 94.16% |

Por fim, a título de comparação, na Tabela 5 são mostrados os resultados dos desempenhos dos classificadores aqui discutidos, destacando as estatísticas de média (Mn), mediana (Md) e uma medida de dispersão, a variância (Var). Pode-se observar novamente aqui o melhor desempenho do classificador KNN, inclusive com uma menor variação na taxa de acerto. Esse resultado pode ser explicado pela talvez natureza linear de separação dos dados. De outro lado, a rede ELM possui melhor desempenho em custo computacional, importante nessa aplicação, onde é necessário embarcar o classificador num hardware de recursos limitados. Sistemas de telemetria fornecem uma série de dados e informações sobre vários aspectos do funcionamento dos veículos e do comportamento do

motorista aos gestores de uma frota. Esses dados e informações diversas podem ser transformados em indicadores importantes, que se aplicados corretamente na gestão de frotas, proporcionam uma série de benefícios.

Tabela 5: Resultados dos classificadores

| Tipo de Classificador | Mn | Md | Var | Custo computacional |
|-----------------------|-------|-------|------|---------------------|
| KNN | 99,28 | 99,44 | 0,22 | 196 s |
| ELM | 92,90 | 92,91 | 2,53 | 48 s |

Por fim, deve-se destacar que o sistema permite guardar e analisar diversas variáveis geradas pelos veículos, possibilitando o emprego de técnicas de aprendizagem de máquinas para reconhecimento de padrões e predição de falhas ou defeitos no veículo, não somente no reconhecimento do comportamento do condutor.

6 Conclusão

No reconhecimento do comportamento do condutor, observou-se que a utilização de um classificador K-NN apresentou resultados promissores, com altas taxas de acerto. Porém, sua implementação, com características lineares, acarreta em um grande uso de tempo de processamento, o que não é interessante em sistemas embarcados com limitada performance computacional. Buscando-se eliminar esse ponto negativo, a utilização da rede ELM apresentou um tempo de processamento quatro vezes menor do que o classificador K-NN, com o custo de uma taxa de acerto um pouco menor, mas ainda consideravelmente alta.

Determinado motorista tende a apresentar comportamentos idênticos na ocorrência de determinado evento, como curvas, acelerações e desacelerações, pois estas são ações já mecanizadas devido a sua ocorrência constante. Apesar disso, observa-se semelhança de comportamentos, em eventos específicos, nas diferentes classes - como abordado na seção 4.1 - o que diminui a eficácia da classificação. Nessas situações menos comuns, a utilização de um único vizinho no classificador K-NN, consegue identificar melhor a conduta do motorista. Na Figura 5, por exemplo, verificamos que há dados coletados da classe Agressiva localizados perto da origem e próximos de dados predominantemente das outras duas classes. Nesse caso, o aumento do número de vizinhos de um dado Agressivo aí localizado diminuirá o acerto do classificador K-NN.

Esses desvios de classificação podem ser evitados caso essas situações sejam identificadas e aplicadas em processos de classificação específicos para cada uma, o que deixa o processo como um todo mais preciso. Além disso, para identificação de padrões, quanto maior for a quantidade de

dados e situações capturadas, mais refinado será o processo. Por isso, os trabalhos futuros envolvem aumentar abrangência das capturas, com um maior número de veículos, motoristas e percursos distintos, possibilitando a criação de uma lógica que permita a identificação de diversas situações pelas quais o motorista passa - como engarrafamentos, curvas e buracos na via - melhorando assim a eficácia da classificação.

Referências

- Bosch (2011). *Bosch Automotive: A product History*, Bosch.
- Fukunaga, K. and Narendra, P. M. (1975). A branch and bound algorithm for computing k-nearest neighbors, *IEEE transactions on computers* **100**(7): 750–753.
- Huang, G.-B., Zhu, Q.-Y. and Siew, C.-K. (2006). Extreme learning machine: theory and applications, *Neurocomputing* **70**(1-3): 489–501.
- ISO (2015). ISO 15031:road vehicles – communication between vehicle and external equipment for emissions-related diagnostics – part 5: Emissions-related diagnostic services.
- MATOS, F. M. d. S. (2008). Reconhecimento de faces utilizando a transformada cosseno discreta, *Masther Thesis*.
- Miche, Y., Schrauwen, B. and Lendasse, A. (2010). Machine learning techniques based on random projections., *ESANN*.
- SAE (2012). SAE J1979: E/E diagnostic test modes.
- Scaringella, N., Zoia, G. and Mlynek, D. (2006). Automatic genre classification of music content: a survey, *IEEE Signal Processing Magazine* **23**(2): 133–141.
- Steinbach, T., Korf, F. and Schmidt, T. C. (2011). Real-time ethernet for automotive applications: A solution for future in-car networks, *Consumer Electronics-Berlin (ICCE-Berlin), 2011 IEEE International Conference on*, IEEE, pp. 216–220.
- Szalay, Z., Kánya, Z., Lengyel, L., Ekler, P., Ujj, T., Balogh, T. and Charaf, H. (2015). Ict in road vehicles-reliable vehicle sensor information from obd versus can, *Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015 International Conference on*, IEEE, pp. 469–476.
- TSBEC (2004). On-board diagnostics ii (obdii) and light-duty vehicle emission related inspection (i/m) programs.