**Projeto de Pesquisa e Planejamento de Atividades**

|  |  |  |
| --- | --- | --- |
| **Aluno**: Matheus Felipe Sozza | | **Data início curso**: 01/10/2021 |
| **Orientador**: Maurício Eloy | | **Defesa em:**  07/2022 |
| **Curso**: MBA Data Science e Analytics | **Modalidade**: Distância | Turma: 212 |

1. **Título do projeto** *(Inicial)*

Detector de motores a combustão por meio de análise sonora através de uma rede neural.

1. **Introdução**

Motores a combustão são sistemas muito presentes no cotidiano dado sua extensa aplicabilidade e o fato de ser uma tecnologia já há muito conhecida e cujo fenômeno base, a combustão, é dominada com robustez pela indústria.

Ao mesmo tempo, é sabido que apesar da relevância que esses dispositivos possuem, existem alguns pontos contra notáveis, como a poluição atmosférica, a poluição sonora, ou até mesmo o risco de sufocamento e envenenamento por dióxido de carbono no caso da operação desses dispositivos em locais fechados.

Fazendo uso dos padrões sonoros do motor é possível extrair informações a respeito da sua performance de operação, o que permite julgar se o mesmo está em funcionamento, operando corretamente, e até mesmo diagnosticar falhas em partes específicas, como citado em Wu. Z. et al., (2022). O sucesso no diagnóstico de falhas é altamente correlacionado à expertise do técnico mecânico responsável pela manutenção, o que resulta em baixo nível de sucesso na maneira em que se é realizado atualmente, havendo então a oportunidade de se melhorar esse processo por meio de algoritmos inovadores (Kemalkar e Bairagi, 2016).

Outra possível utilização de dados sonoros de motores é aquela voltada à geração de estatísticas de tráfego e planejamento de demanda, ou até mesmo no controle de semáforos baseado em situações especiais, como a presença de veículos de emergência (ambulâncias), como visto em Analytics Vidhya (2022). Dados extraídos num contexto de *Big-Data* têm sido cada vez mais utilizados no âmbito *das* cidades inteligentes (*Smart-Cities*), como no planejamento e gestão do tráfego das zonas rural e urbana (Zhao, Y. et al., 2018).

Mohammadi e Al-Fuqaha (2018) enunciam que, apesar das altas capacidades de coleta e armazenamento dos sistemas computacionais atuais, muito pouco se aproveita da informação ali contida, com métodos tradicionais baseados apenas em análise temporal e que negligenciam a presença de padrões valorosos contidos nos dados armazenados. Por outro lado, o uso de redes neurais profundas (DNN ou *Deep Neural Networks*) se mostra uma alternativa válida e promissora para a extração de informações valorosas do ponto de vista analítico.

Outro ponto enunciado por Mohammadi e Al-Fuqaha (2018) é o fato de que, uma vez que os dados são coletados e armazenados, é improvável que venham a ser reutilizados no futuro, encorajando o processamento imediato.

Quando se trata de padrões não observáveis ou perceptíveis a priori, a área de aprendizado de máquina se mostra promissora no aspecto de extração e geração de informação, agregando valor à dados antes não aproveitados.

1. **Objetivo**

O intuito do trabalho aqui apresentado é propor um algoritmo capaz de detectar o funcionamento de um motor a combustão apenas baseado na análise de sons do ambiente processados através de um classificador, num modelo de aprendizado de máquina supervisionado.

Essa solução em princípio se mostra simples e viável do ponto de vista econômico, dado que microfones são dispositivos comuns e que os atuais dispositivos computacionais, sejam móveis (como celulares) ou fixos (como os computadores) possuem boa capacidade de processamento.

Ao fim do projeto de pesquisa espera-se determinar, de maneira experimental, o quão preciso é um classificador de sons baseado em redes neurais aplicado ao contexto de sons urbanos, mais especificamente na detecção de motores a combustão, e se é possível aplicá-lo de maneira escalável e robusta.

Sons urbanos costumam ser compostos por várias camadas e ruídos diversificados, em grande maioria de fonte antropomórfica, o que exige maior robustez da rede neural em identificar e classificar o mesmo tipo de som em diferentes situações.

1. **Material e Métodos**

A metodologia para obtenção, processamento e posterior classificação dos dados será baseada em programação na linguagem Python, de acordo com os passos a seguir:

1. **Dados:** Utilizar pequenos recortes de sons em formato *‘.mp3’* ou *‘.wav’* com durações de 500 milissegundos até 4 segundos, provenientes do *dataset* *UrbanSounds8k* (disponível gratuitamente na internet).
2. **Pré-processamento:** Preparar os recortes sonoros padronizando a taxa de amostragem, normalizando as amplitudes dos sinais, e padronizando a duração dos recortes de áudio para processamento no algoritmo.  
      
   Também é parte dessa etapa o processo de aumento de dados (*data-augmentation*) por meio da geração de amostras extra através de sobreposição das já existentes (*overlap)* etambém com a adição de ruído branco aleatório a uma mesma amostra. Essa etapa gera o aumento do número de amostras do conjunto de dados além de criar pequenas variações de uma mesma amostra de maneira a evitar o sobreajuste (*overfitting*) do modelo (Tang, S. et al., 2020).  
     
   Demais pré-processamentos necessários podem ser aplicados*,* apoiados sempre em bibliotecas Python como *Numpy* e *Pandas* (manipulação de dados e matrizes) e *Librosa* para manipulação dos áudios (McFee, B. et al, 2015).
3. **Extração de informações (*features*):** Decompor os sinais em frequências através da transformada rápida de Fourier (*FFT -* *Fast-Fourier Transform*) obtendo os respectivos espectogramas para cada recorte de áudio, de onde serão extraídos as principais informações (*features)* para alimentar a rede neural, por exemplo, os *MFCC* ou Coeficientes Cepstrais de Frequência Mel (*Mel-Frequency Cepstral Coefficients)*.   
     
   A utilização de espectogramas transforma o sinal sonoro unidimensional (intensidade sonora vs. tempo) em uma imagem 2D (intensidade por banda de frequência vs. tempo), o que auxilia no processo de separação do ruído de fundo e demais componentes indesejados do sinal sonoro principal (Wu. Z. et al., 2022), tornando o desafio de processar esses dados algo similar ao processamento de imagens.  
     
   Já a utilização dos *MFCCs* como informação é reconhecidamente um método efetivo para abordagens de processamento e classificação de sinais sonoros ambientais (J. Salamon, J. P. Bello, 2016).
4. **Processamento:** Treinar a rede neural convolucional (*CNN* *–* *Convolutional Neural Network*) com os dados decompostos e respectivas informaçõesextraídas, de maneira a buscar padrões de frequência e temporalidade que indiquem a presença de um motor a combustão ou não. Para tal, serão utilizadas como ferramentas as bibliotecas *Keras* e *Tensorflow* disponíveis em ambiente Python. Ou seja, será feito um processamento supervisionado dos dados.
5. **Validação e Métricas:** Testar e validar a solução obtida dividindo o conjunto de dados em conjunto de treino e conjunto de testes (*Train vs. Test),* realizando também validações cruzadas (*K-Fold Cross-Validation*).  
     
   As validações devem eventualmente ser repetidas com variações nos hiper parâmetros buscando uma resposta convergente tendendo a altos valores de acurácia de treino e habilidade de predição.  
     
   A análise se dará primordialmente com base na verificação das matrizes de confusão de cada um dos conjuntos treino-teste, não sendo avaliada apenas a acurácia ou quantidade de valores corretamente classificados, mas também como se comportam os valores classificados incorretamente, buscando entender se há vieses ou casos sensíveis no modelo proposto.
6. **Resultados Esperados**

Espera-se um algoritmo capaz de classificar os dados de maneira binária, atestando se o recorte de áudio em questão contém um motor a combustão em funcionamento ou não.

Espera-se também que a qualidade da classificação nessa primeira abordagem experimental e exploratória possua uma performance que supere, ao menos, uma classificação aleatória, abrindo caminho para explorações mais direcionadas e que levem à uma melhor performance em trabalhos futuros.

Os resultados serão comparados com outros experimentos realizados no conjunto de dados *UrbanSounds8k* de maneira a classificar o quão efetivo, tanto quantitativa quanto qualitativamente, foi a abordagem proposta nesse estudo. Para tal, observaremos as matrizes de confusão resultantes das classificações para cada um dos conjuntos de validação cruzada, avaliando também a sensibilidade dos resultados com a variação de alguns hiper parâmetros.

1. **Cronograma de Atividades**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Atividades planejadas** | **Mês** | | | | | | | | | |
| **1 (Out/22)** | **2** | **3** | **4** | **5 (fev/22)** | **6** | **7** | **8** | **9** | **10 (Jul/23)** |
| Revisão bibliográfica + Definição da pesquisa | X | X |  |  |  |  |  |  |  |  |
| Preparo dos *Datasets* | X | X | X |  |  |  |  |  |  |  |
| Preparo inicial do algoritmo |  | X | X |  |  |  |  |  |  |  |
| Confecção do algoritmo e iterações sobre o mesmo |  | **X** | **X** | **X** |  |  |  |  |  |  |
| Extrair resultados preliminares |  |  | **X** | **X** |  |  |  |  |  |  |
| Refinar resultados finais |  |  |  | **X** | **X** |  |  |  |  |  |
| Elaborar Conclusões |  |  |  |  | **X** |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Escrever monografia nos moldes do manual |  |  |  |  | **X** | **X** |  |  |  |  |
| Rever monografia com o orientador |  |  |  |  |  | **X** |  |  |  |  |
| Submeter monografia no sistema |  |  |  |  |  |  | **X** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Entregar apresentação de defesa |  |  |  |  |  |  | **X** | **X** | **X** |  |
| Realizar defesa do TCC |  |  |  |  |  |  |  |  |  | **X** |
|  |  |  |  |  |  |  |  |  |  |  |

Projeto de Pesquisa; Resultados Preliminares; Entrega do Trabalho de Conclusão de Curso; Entrega da Apresentação da Defesa

1. **Referências Bibliográficas**

Analytics Vidhya. 2022. Vehicle Sound Classification Using Deep Learning. Disponível em <https://www.analyticsvidhya.com/blog/2022/01/vehicle-sound-classification-using-deep-learning>. Acesso em 3 de outubro de 2022.

J. Salamon; J. P. Bello. 2016. Deep Convolutional Neural Networks and Data

Augmentation for Environmental Sound Classification. IEEE Signal Processing Letters, vol. 24, no. 3, pp. 279-283, March 2017.

Kemalkar, A.K.; Bairagi, V.k.. 2016. Engine Fault Diagnosis Using

Sound Analysis. International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT): 943-946

McFee, B.; Colin, R; Dawen, L; Daniel, P. 2015. Librosa: Audio and music signal analysis in python. In Proceedings of the 14th python in science conference: 18-25.

Mohammadi, M.; Al-Fuqaha, A. 2018. Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges. IEEE Communications Magazine, vol. 56, no. 2: 94-101.

Tang, S.; Yuan, S.; Zhu, Y. 2020. Data Preprocessing Techniques in Convolutional Neural Network based on Fault Diagnosis towards Rotating Machinery. IEEE Access, vol. 8: 149487-149496

Wu, Z.; Wan, Z.; Ge, D. et al. 2022. Car engine sounds recognition based on deformable feature map residual network. Sci Rep 12, 2744 (2022).

Zhao, Y; Zhang, H; An, L; et al. 2018. Improving the approaches of traffic demand forecasting in the big data era. Cities, Volume 82: 19-26.