

Universidade Federal de Itajubá - UNIFEI
IENTI01 - TINYML - Aprendizado de Máquina Aplicado para Dispositivos
IOT Embarcados
Turma 02

Proposta de Projeto Final
TinyML aplicado à Detecção da Doença
Pulmonar Obstrutiva Crônica

Ana Paula Bessa Medeiros - 2019004438 - UNIFEI
Letícia Tempone de Oliveira - 2019004910 - UNIFEI
Rafael Coelho Paes - 2019000081 - UNIFEI
Ricardo Magno do Carmo Junior - 2019018649 - UNIFEI

1 Introdução

No ano de 2019, a Doença Pulmonar Obstrutiva Crônica (DPOC) foi a terceira maior causa de morte no mundo, contando com cerca de 3.23 milhões de mortes ([Organização Mundial da Saúde, 2021](#)), como mostra o infográfico da Figura (1).

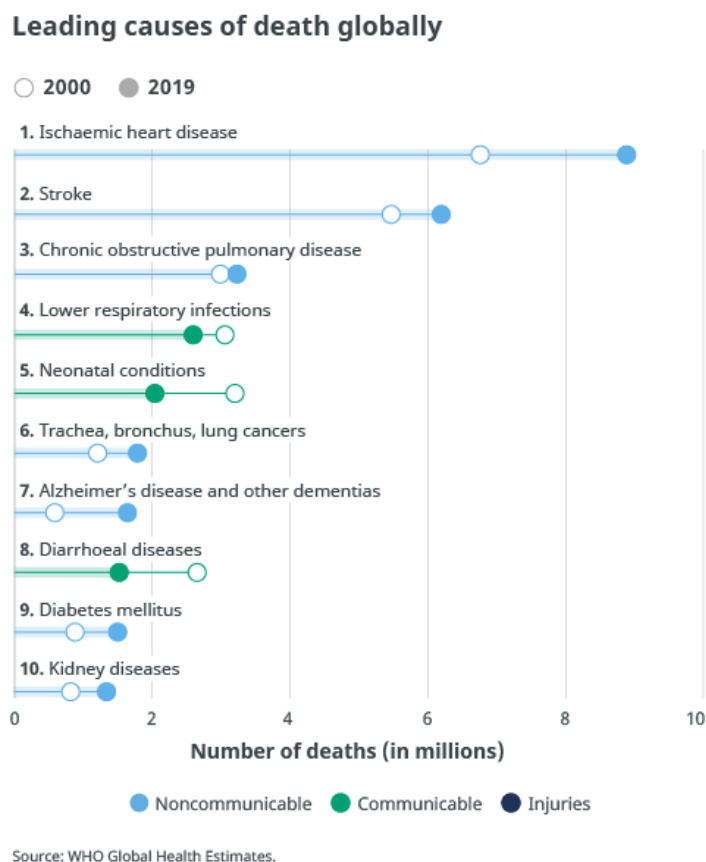


Figura 1 – Principais causas de morte no ano de 2019, comparados ao ano 2000

A DPOC nada mais é do que um conjunto de diversas doenças respiratórias, e está fortemente ligada à prática do tabagismo. Alguns dos sintomas são a **Bronquite Crônica** - responsável pelo estreitamento das vias aéreas e a **Enfisema** - que causa danos irreversíveis nos alvéolos pulmonares, estruturas fundamentais para o funcionamento do aparelho respiratório, causando perda progressiva da função pulmonar

Como o próprio nome diz, trata-se de uma doença crônica, portanto não há uma cura definitiva. Contudo, existe a possibilidade de tratamento, essencial para seu controle e consequentemente para a longevidade do portador da doença. Portanto, quanto antes o diagnóstico é feito, mais eficaz o resultado do tratamento será.

Entretanto, apresentar um diagnóstico precoce ainda é um desafio, visto que comumente a DPOC apenas é detectada em estados mais graves, nos quais sintomas já são marcantes. Tal fato faz os pacientes buscarem ajuda médica tardiamente e, consequentemente, iniciar o tratamento em situações mais avançadas.

Isto posto, uma grande possibilidade de avanço para a área são métodos que auxiliem no diagnóstico, principalmente o tornando mais rápido e eficaz.

Paralelamente, com avanço das análises computacionais e do poder de sensoriamento, técnicas modernas como Inteligência Artificial são utilizadas na saúde, na segurança, na economia, assim como em muitas outras áreas sociais. Não obstante, especificamente o *TinyML* demonstra-se uma poderosa tecnologia a ser agregada no diagnóstico de doenças.

O *TinyML* permite a implementação de algoritmos de aprendizado de máquina em dispositivos na ponta, em sistemas de baixo consumo de energia, com uso de sensores pequenos e de processamento até mesmo em microcontroladores. Sucintamente, essa vertente permite aplicar técnicas computacionais sofisticadas em dispositivos fisicamente pequenos, com baixíssimo consumo de energia e, no geral, em modo *always-on*.

Desse modo, com sua aplicação para investigar DPOC em pacientes, é possível atingir um sensoramento fácil, prático e eficiente, por meio de um dispositivo sempre ativo no paciente. Em razão do seu baixo uso de memória e de consumo de energia, tal dispositivo pode apresentar dimensões físicas pequenas e peso leve, podendo ser incluído no uso diário dos pacientes. Um exemplo nessa linha foi a tendência das pulseiras/relógios inteligentes.

2 Objetivos

Este trabalho final tem como objetivo principal desenvolver um modelo de TinyML capaz de detectar a DPOC através do som da respiração de uma pessoa de modo a auxiliar no diagnóstico da doença, evitando tratamentos tardios e a agravamento da doença.

Os objetivos específicos deste trabalho são dados pelos tópicos a seguir:

1. Aplicar os conhecimentos adquiridos na disciplina IESTI01 - TinyML - Aprendizado de Máquina Aplicado para Dispositivos IOT Embarcados para o desenvolvimento de um sistema detector de DPOC.
2. Investigar e compreender técnicas adicionais de *Machine Learning* para implementação em dispositivos de ponta.
3. Criar um *dataset* de áudios de respiração saudável e ajustar suas frequências.

3 Descrição do Projeto

3.1 Diagrama de Blocos

O diagrama de blocos contendo a ordem cronológica das etapas do projeto pode ser visualizado na Figura (2).

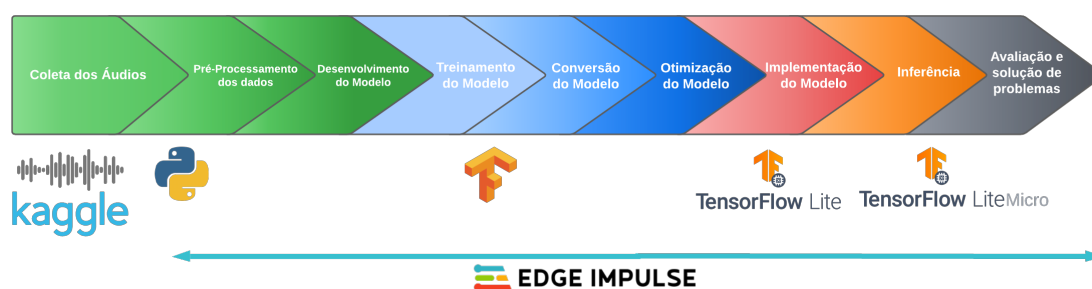


Figura 2 – Diagrama de Blocos

3.2 Hardware utilizado

Para a implementação do modelo a ser elaborado, será necessário utilizar o *Kit Didático Arduino* voltado para *Tiny Machine Learning*, que acompanha os itens:

- Arduino Nano 33 BLE Sense Board;
- Arduino Tiny Machine Learning Shield;

- Cabo USB-A para Micro USB.

A placa *Arduino Nano 33 BLE Sense*, possui, em sua estrutura, um microprocessador ARM Cortex-M4F 32-bit, que opera em até 64 MHz, com 1 MB de memória Flash e 256 KB de memória RAM.

3.3 Coleta de Dados

A princípio, os dados utilizados para o desenvolvimento do projeto serão recolhidos do *dataset Respiratory Sound Database*, encontrado no site *Kaggle*, uma comunidade online formada por cientistas de dados e profissionais de *Machine Learning*. Uma breve ilustração do site se encontra na Figura (3).

O arquivo baixado contendo *dataset*, com um tamanho de aproximadamente 4GB, possui um total de 920 arquivos de áudio no formato *.wav*, compreendendo gravações de 126 pacientes de variadas faixas etárias: crianças, adultos e idosos. No total, há uma soma de 5.5 horas de gravação, contendo 6898 ciclos respiratórios, que podem incluir sons saudáveis ou sons contendo chiados (classificados como *wheezes*) ou crepitações (classificados como *crackles*). Alguns dos arquivos de áudios, de modo a simular condições em um ambiente real, possuem ruídos sonoros de fundo.

Além dos arquivos de áudio, o *dataset* acompanha outros arquivos que auxiliam na manipulação e análise do mesmo:

- *demographic_info.txt* - Informações sobre o paciente, como número de identificação, idade, sexo, índice de massa corporal e, para caso de pacientes crianças, peso e altura.
- *patient_diagnosis.csv* - Tabela contendo o **Diagnóstico** de cada um dos 126 pacientes, que podem ser: *Saudável (Healthy)*, *DPOC (COPD)*, *Infecção do trato respiratório inferior (LRTI)*, *Infecção do trato respiratório superior (URTI)*.
- *filename_format.txt* - Uma breve explicação sobre o formato de nomeação dos arquivos: O nome de cada arquivo compreende informações como o número do paciente, a região do corpo em que a gravação foi realizada, o modo de aquisição e o dispositivo utilizado para gravar.
- Os arquivos de áudio possuem uma nomenclatura que possibilita identificar o número do paciente, região de gravação e o dispositivo de gravação:
 - Microfone AKGC417L (AKGC417L);
 - Estetoscópio 3M Littmann Classic II SE (LittC2SE);
 - Estetoscópio 3M Litmmann 3200 Electronic (Litt3200).
- Cada arquivo de áudio *.wav* compreende um arquivo *.txt* de mesmo nome, onde está registrado a quantidade de chiados e crepitações do respectivo áudio, juntamente com o momento temporizado em que estes ocorreram.

Além deste *dataset*, para contribuir com a qualidade do modelo, serão gravados, pela equipe, três vezes mais arquivos de áudio saudáveis em comparação com os áudios com DPOC utilizados do *dataset*, utilizando o microfone digital **MP34DT05** presente na placa *Arduino Nano 33 BLE Sense*.

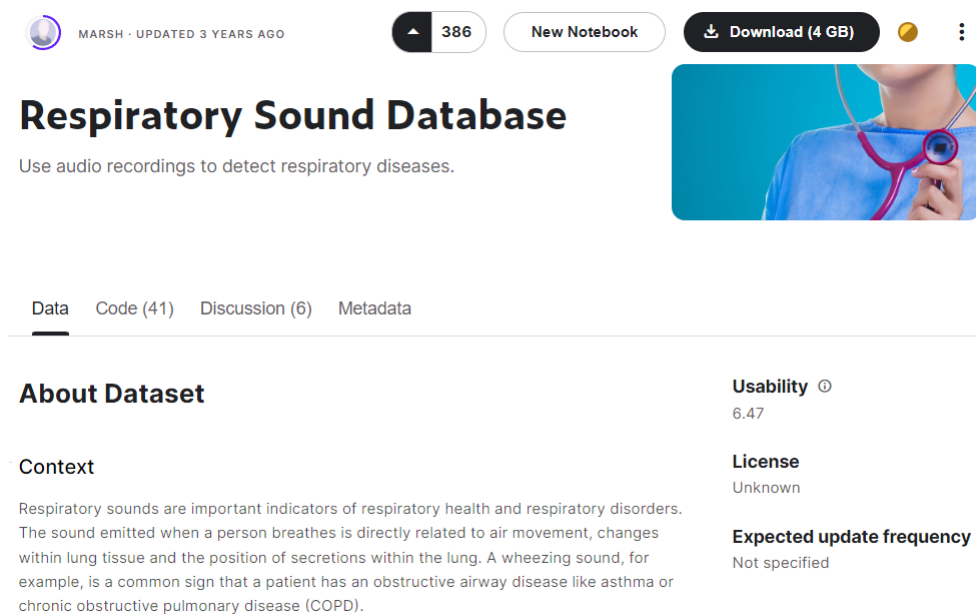


Figura 3 – Página do *dataset* utilizado no site Kaggle

3.4 Pré-Processamento

Tendo em vista os arquivos presentes no *dataset*, é nítido que será necessário efetuar uma pré-seleção nos dados. No caso, a doença a ser identificada no projeto será restrita à DPOC, portanto se faz essencial recolher apenas os dados que correspondem à esta doença para a primeira parte do *dataset*.

Para a aplicação desejada, de modo a utilizar os recursos disponíveis da placa (Microfone), se torna interessante lidar apenas com os dados gravados com o **Microfone AKGC417L**, uma vez que os sons captados por este, diferente de um estetoscópio profissional, podem se assemelhar muito mais aos sons captados pelo microfone da placa, tornando a implementação mais precisa.

Também faz-se necessário converter a frequência dos arquivos de áudio do *dataset*, uma vez que a gravação dos arquivos de áudio utilizando a placa será feita em 16kHz, portanto, os arquivos do *dataset* serão convertidos para essa mesma frequência para evitar erros no treinamento do modelo.

3.5 Preparação do Modelo e Otimizações

Visa-se testar diferentes arquiteturas de redes neurais, variando a quantidade de camadas, de neurônios e de técnicas de otimização do modelo. Para definir em si o modelo a ser usado, a análise será baseada entre um equilíbrio de boa performance de predição, ou seja, baixo erro. Mas também, mantendo a otimização de tempo e de uso da memória. Para assim, tentar garantir uma boa inferência, mas sem comprometer as características necessárias para sua implementação em dispositivos de hardware do *TinyML*.

3.6 Deploy

Após o treinamento e validação do modelo, espera-se fazer o *deployment* do mesmo na placa *Arduino Nano 33 BLE Sense* utilizando as técnicas de *Pruning* e *Quantização* que permitirão a implementação do sistema de reconhecimento em dispositivos *Tiny*. Como adicional, espera-se adicionar uma lógica ao código que trabalha acendendo um determinado LED para cada tipo de classificação durante a inferência do dispositivo embarcado.

4 Referências e inspirações

- [ICBHI 2017 Challenge - Respiratory Sound Database](#)
- [Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease](#)
- [A dataset of lung sounds recorded from the chest wall using an electronic stethoscope](#)

5 Problemas previstos e possíveis soluções

O *dataset* que será utilizado no trabalho apresenta poucos arquivos de áudio, principalmente arquivos de áudio de respiração saudável, o que pode comprometer a *accuracy* do modelo na hora de ser testado em prática pelo microfone da plaquinha, visto que terá tido poucos áudios como base. Dessa maneira, uma solução para este problema, será expandir os arquivos de áudio de respiração saudável, tendo em vista que todos os membros do trabalho não apresentam nenhuma doença respiratória, através do microfone da plaquinha será possível gravar novas respirações e adicioná-las ao dataset de respirações saudáveis após sua normalização.

Um outro problema que surge é sobre a qualidade dos áudios. Primeiro, os dados prontos coletados de pessoas que possuem DPOC, tanto os das pessoas que não possuem, são de 20s, dessa maneira, é possível que eles apresentem tempos de silêncio entre uma respiração e outra, assim como ruídos e barulhos indesejados, o que pode acabar comprometendo a qualidade do modelo desenvolvido. Com isso, uma possível solução seria recortar os áudios somente nas partes de interesse.

Além disso, ao tentarmos expandir o *dataset* de respiração saudável através do microfone da placa do nosso kit, os áudios podem apresentar má qualidade. Para lidar com isto, será necessário utilizar ferramentas que façam o processamento dos arquivos de áudio, tais como o MFCC e o MFE disponíveis no *Edge Impulse*.

Por fim, os dados pegos na internet apresentam uma frequência diferente da frequência de 16kHz da que a placa é capaz de trabalhar e que também será a frequência dos áudios capturados pelo seu microfone. Assim, para que estes apresentem a mesma frequência, será necessário fazer uso de uma biblioteca especial capaz de alterar a frequência dos arquivos de áudio que no caso pretende-se utilizar a biblioteca *librosa*.

Referências

Organização Mundial da Saúde. *Chronic obstructive pulmonary disease (COPD)*. [S.l.], 2021. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>>.