

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS  
ESCOLA POLITÉCNICA E DE ARTES  
GRADUAÇÃO EM CIÊNCIA DE COMPUTAÇÃO

**Paulo Henrique Garcia Ribeiro**



**DESENVOLVIMENTO DE UM SISTEMA PARA RECONHECIMENTO DAS  
NECESSIDADES DO BEBÊ ATRAVÉS DO CHORO.**

Goiânia,  
2024

Paulo Henrique Garcia Ribeiro

*Desenvolvimento de um Sistema para Reconhecimento das Necessidades do Bebê através do Choro.*

Trabalho de Conclusão de Curso,  
apresentando à Escola Politécnica e de Artes,  
da Pontifícia Universidade Católica de Goiás,  
como parte dos requisitos para a obtenção do  
título de Bacharel em Ciência da Computação.  
Orientador: Prof. Gustavo Siqueira Vinhal

Goiânia,  
2024

Paulo Henrique Garcia Ribeiro

**DESENVOLVIMENTO DE UM SISTEMA PARA RECONHECIMENTO DAS  
NECESSIDADES DO BEBÊ ATRAVÉS DO CHORO.**

Trabalho de Conclusão de Curso aprovado em sua forma final pela Escola  
Politécnica e de  
Artes, da Pontifícia Universidade Católica de Goiás, para obtenção do título de  
Bacharel em Ciência da Computação, em: \_\_\_\_/\_\_\_\_/\_\_\_\_

Orientador: Prof. Gustavo Siqueira Vinhal

---

Banca Examinadora: Lucília Gomes Ribeiro

---

Banca Examinadora: Rafael Leal Martins

---

Coordenadora de Trabalho de Conclusão de Curso

GOIÂNIA,  
2024

## RESUMO

Este Trabalho aborda o desenvolvimento de um sistema para reconhecimento das necessidades do bebê através do choro. O estudo será composto pelas seguintes etapas: o uso de uma base de dados de choros de bebês categorizados por necessidades específicas (fome, dor, cansaço, desconforto e frustração); análise acústica detalhada dessas gravações utilizando técnicas como MFCCs (*Mel-Frequency Cepstral Coefficients*); desenvolvimento e treinamento de modelos de aprendizado de máquina para a classificação dos diferentes tipos de choro; implementação de uma aplicação prática que permita o *upload* de áudios para identificar, em tempo real, a necessidade do bebê; e teste e validação do sistema com novos dados para avaliar sua precisão e eficácia. O trabalho visa criar uma ferramenta inovadora para pais e cuidadores, melhorando significativamente o cuidado infantil ao possibilitar respostas rápidas e adequadas às necessidades dos bebês. Além disso, esse trabalho contribuirá para a pesquisa na área da análise de choro infantil, oferecendo uma base sólida para futuros desenvolvimentos tecnológicos.

Palavras-chave: Aprendizado de Máquina, MFCCs (*Mel-Frequency Cepstral Coefficients*), Choro Infantil

## **ABSTRACT**

This work addresses the development of a system for recognizing babies' needs based on their cries. The study will consist of the following stages: using a database of categorized baby cries for specific needs (hunger, pain, tiredness, discomfort, and frustration); detailed acoustic analysis of these recordings using techniques such as Mel-Frequency Cepstral Coefficients (MFCCs); development and training of machine learning models to classify different types of cries; implementation of a practical application that allows uploading audio to identify, in real time, the baby's needs; and testing and validation of the system with new data to assess its accuracy and effectiveness. The aim is to create an innovative tool for parents and caregivers, significantly improving infant care by enabling quick and appropriate responses to babies' needs. Additionally, this work will contribute to research in the field of infant cry analysis, providing a solid foundation for future technological developments.

Keywords: Machine Learning, MFCCs (Mel-Frequency Cepstral Coefficients), Infant Cry

## LISTA DE SIGLAS E ABREVIATURAS

|       |                                            |
|-------|--------------------------------------------|
| MFCCs | <i>Mel-Frequency Cepstral Coefficients</i> |
| CNNs  | <i>Convolutional Neural Networks</i>       |
| RNNs  | <i>Recurrent Neural Networks</i>           |
| PCD   | Pessoa com Deficiência                     |
| ML    | <i>Machine Learning</i>                    |

## LISTA DE FIGURAS

|                                                              |    |
|--------------------------------------------------------------|----|
| Figura 01 - Trecho de código da extração do MFCCs.....       | 18 |
| Figura 02 - Tratamento de Exceções.....                      | 19 |
| Figura 03 – Treinamento, predição e avaliação do modelo..... | 19 |
| Figura 04 – Trecho final do código.....                      | 20 |

## Sumário

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| 1. Introdução .....                                                      | 9  |
| 2. Fundamentos Teóricos .....                                            | 11 |
| 2.1 Desenvolvimento Infantil e Comunicação .....                         | 11 |
| 2.2 Tipos de Choro .....                                                 | 13 |
| 2.2.1 Choro Por Fome .....                                               | 13 |
| 2.2.2 Choro Por Cansaço ou desconforto.....                              | 14 |
| 2.2.3 Choro por Dor.....                                                 | 14 |
| 2.2.4 Por Cólica.....                                                    | 14 |
| 2.3. MFCCs.....                                                          | 15 |
| 2.4 Biblioteca Librosa .....                                             | 15 |
| 2.5 Biblioteca Flask.....                                                | 15 |
| 2.6 Biblioteca Pandas .....                                              | 15 |
| 2.7 Biblioteca Numpy .....                                               | 16 |
| 2.8 Modulo OS.....                                                       | 16 |
| 2.9 Biblioteca Joblib .....                                              | 16 |
| 2.10 Reconhecimento de voz .....                                         | 16 |
| 2.11 Sistemas de Reconhecimento de Padrões e Aprendizado de Máquina..... | 17 |
| 2.12 Randon Forest .....                                                 | 17 |
| 3. Metodologia .....                                                     | 17 |
| 3.3. Redução de Ruídos Usando o Audacity.....                            | 18 |
| 3.4 Extrair Características Espectrais MFCCs .....                       | 18 |
| 3.5 Tratamento de exceções para erro de diretório.....                   | 18 |
| 3.6 Treinamento, predição e avaliação do modelo.....                     | 19 |
| 3.7 Trecho final do código.....                                          | 20 |
| 4. Resultados .....                                                      | 20 |
| 5. Dificuldades encontradas.....                                         | 20 |
| 5.1 Variabilidade Individual.....                                        | 21 |
| 5 Conclusão .....                                                        | 21 |
| Referência .....                                                         | 23 |

## 1. Introdução

A comunicação entre pais e bebês é essencial para atender às necessidades básicas e promover o bem-estar infantil. Nos primeiros meses de vida, os bebês ainda não possuem a capacidade de utilizar a linguagem verbal, e o choro se torna a principal forma de expressão dos recém-nascidos, sendo sua linguagem inicial e primordial. Através do choro, os bebês comunicam uma variedade de necessidades e desconfortos, como fome, fralda suja, cansaço, dor, e necessidade de atenção. De acordo com um estudo feito pelo *The Hospital Clínic-IDIBAPS* e com a colaboração *The Swiss-Catalan startup Zoundream AG*, cada tipo de choro pode ter diferentes características, como variações de tom, intensidade e duração, que podem fornecer pistas sobre o que o bebê está tentando comunicar (*Computers in Biology and medicine*, 2023).

No entanto, compreender o significado por trás desse choro pode ser desafiador para os pais, especialmente aqueles sem experiência prévia, cuidadores de múltiplos filhos ao mesmo tempo, ou pais que PCD que por ter dificuldade de audição ou aqueles que não conseguem ouvir. A falta de compreensão pode levar a uma resposta inadequada às necessidades do bebê, resultando em estresse tanto para o bebê quanto para os cuidadores, e para os cuidadores PCD por não possuir audição ajudar a saber quando e porque o bebê está chorando. Esse desafio é ainda maior em situações em que os bebês apresentam choros atípicos ou condições de saúde que influenciam seu comportamento vocal. Para pais e cuidadores de crianças com deficiência, a interpretação dos choros pode ser ainda mais complicada devido a possíveis variações ou limitações na expressão vocal desses bebês. Portanto, estratégias eficazes para reconhecer e interpretar diferentes tipos de choro são essenciais para fornecer o cuidado adequado e reduzir o estresse para todos os envolvidos.

Nesse contexto, o presente trabalho propõe a criação de um sistema que o usuário carrega um arquivo de áudio do choro do bebê, o sistema analisa e fala para o usuário o porquê o bebê chora. Esse sistema utilizará técnicas avançadas de processamento de sinais de áudio e aprendizado de máquina para analisar os padrões acústicos do choro do bebê e classificá-los de acordo com a necessidade subjacente. O objetivo é criar uma ferramenta acessível e fácil de usar proporcionando aos pais e cuidadores uma assistência adicional na interpretação e resposta adequada às demandas do bebê.

O desenvolvimento desse sistema envolve várias etapas, começando com a coleta de dados de várias gravações de choro de bebês, que será coletada de um repositório online que foi criado com o intuito do estudo de dados. Em seguida, essas gravações passarão por um processo de pré-processamento para remover ruídos usando o programa grátis Audacity que é um programa para vários tipos de edição e manipulação de áudio e normalizar os dados de áudio. Depois haverá uma extração de características, como coeficientes cepstrais de frequência mel (MFCC).

O MFCCs serve para extrair as características de fala de um orador permitindo trabalhar essas características presentes para apresentá-las para o sistema. Os dados são encaminhados como um *input* sistema de reconhecimento de fala (Leonardo Manhães, 2018).

Com essas características, modelos de aprendizado de máquina, incluindo redes neurais convulsionais (CNNs) e redes neurais recorrentes (RNNs), serão treinados para reconhecer e categorizar os diferentes tipos de choro. Além disso, a interface do usuário será projetada para ser intuitiva e fornecer feedback em tempo real, permitindo que os pais respondam rapidamente às necessidades do bebê.

A implementação de um sistema desse tipo tem o potencial de revolucionar o cuidado infantil, proporcionando uma camada extra de suporte para pais, especialmente os de primeira viagem e os pais com necessidades especiais, ajudando-os a construir uma conexão mais forte e eficaz com seus bebês. Ao melhorar a comunicação entre pais e bebês, esperamos contribuir

para o desenvolvimento saudável e o bem-estar emocional dos recém-nascidos, aliviando o estresse dos cuidadores e promovendo um ambiente familiar mais harmonioso.

## **2. Fundamentos Teóricos**

Compreender os diferentes tipos de choro do bebê é fundamental para atender às suas necessidades e emoções. Embora alguns pais possam rapidamente identificar o que o bebê precisa pelo choro, muitos pais de primeira viagem podem precisar de orientação para decodificar essa forma primordial de comunicação. O choro pode variar em tom, intensidade e ritmo, indicando diferentes necessidades como fome, desconforto, cansaço ou necessidade de atenção. Aprender a reconhecer essas variações ajuda os pais a responderem de maneira mais eficaz e rápida, promovendo o bem-estar do bebê e fortalecendo o vínculo entre pais e filhos (*Computers in Biology and medicine*, 2023).

### **2.1 Desenvolvimento Infantil e Comunicação**

O choro do bebê é a primeira tentativa adequada que os seres humanos utilizam para se comunicar durante seus primeiros meses de vida [1–3]. Como resultado, os pais tendem a interpretar o choro do bebê como um sinal de alerta ou necessidade. No entanto, novos cuidadores frequentemente se sentem confusos sobre o significado do choro e podem ser incapazes de acalmar o bebê [4], o que pode resultar em uma variedade de sentimentos mistos [5]. Considerando que os bebês choram, em média, entre 1,5 e 3 horas por dia [4], o impacto do choro do bebê nos pais pode variar entre experiências de ansiedade, depressão, desamparo, raiva e frustração em resposta ao choro, afetando negativamente o vínculo e a percepção parental do bebê [1,6]. Os pais podem até experimentar pensamentos violentos em relação ao recém-nascido e, posteriormente, sentimento de culpa e vergonha [4]. Considerando que a resposta parental é crucial para o novo relacionamento diádico, desvios e/ou mal-entendidos da mensagem do choro podem comprometer o cuidado do bebê e seu processo de desenvolvimento neuropsicológico futuro [1]. Portanto,

é importante encontrar um significado nos primeiros choros para garantir o bem-estar e a saúde do bebê.

Pesquisas [4] sobre características espectrográficas do som do choro vêm sendo conduzidas desde a década de 1960. Análises adicionais [4] investigaram como o choro a duração e a intensidade podem variar conforme o contexto. Tanto abordagens no domínio do tempo quanto no domínio da frequência levaram à conclusão de que os sons do choro transmitem um nível de angústia ou urgência da necessidade do recém-nascido, oferecendo algumas pistas sobre a causa específica do choro [4]. Além disso, juntamente com informações contextuais, os sons do choro podem ser altamente informativos, facilitando uma interpretação precisa (embora complexa) para o cuidador ou profissional de saúde [4].

Nas últimas décadas, o campo da análise automatizada do choro infantil avançou significativamente, atraindo a atenção de um número crescente de pesquisadores, clínicos e cientistas da computação [7–9]. Além disso, o advento das técnicas de Deep Learning (DL) facilitou o desenvolvimento de algoritmos de Inteligência Artificial mais confiáveis e precisos para a análise de choros infantis [10–17].

Embora avanços significativos tenham sido alcançados na pesquisa sobre o choro infantil, a interpretação dos choros ainda permanece parcialmente sem resposta e subjetiva. Há um debate contínuo no campo sobre a detecção e classificação de diferentes tipos de choro. Enquanto alguns estudos mostraram resultados promissores, outros permanecem inconclusivos [12–16]. Portanto, é evidente que futuras pesquisas devem se concentrar na coleta de dados multimodais para entender os choros infantis, avaliando simultaneamente diversas medidas dos recém-nascidos. A análise paralela do choro infantil, juntamente com medidas neurofisiológicas como espectroscopia no infravermelho próximo (NIRS), eletroencefalografia (EEG) e sinais comportamentais simultâneos (extraídos de expressões faciais e movimentos corporais usando gravações de vídeo), surge como uma abordagem pioneira promissora para fornecer evidências científicas e objetivas para apoiar e validar a literatura sobre acústica do choro e sua análise [4].

O objetivo principal deste estudo é caracterizar diferentes sons de choro de recém-nascidos com base em características acústicas, sinais neurofisiológicos e comportamentais. O objetivo secundário é determinar as características mais relevantes que permitem a distinção de diferentes razões para o choro com base em Machine Learning (ML) dentro de um conjunto de dados multimodal. Por fim, pretendemos demonstrar que abordagens de DL, como a apresentada neste artigo (ou seja, o Acoustic MultiStage Interpreter (AMSI [18])), desenvolvido pelo nosso grupo de pesquisa, são ferramentas automáticas, eficazes e confiáveis para interpretar os choros infantis e avaliar o bem-estar dos recém-nascidos.

Portanto, hipotetizamos que o que está ocorrendo em nível acústico pode ser reforçado com a análise de padrões comportamentais e neurofisiológicos cerebrais. Além disso, ao analisar as características sonoras dos choros infantis, o AMSI [18] é capaz de identificar biomarcadores acústicos valiosos que indicam o estado de bem-estar do recém-nascido. Através dessa abordagem de análise multimodal, obtemos uma melhor compreensão e uma interpretação mais precisa do complexo fenômeno do choro infantil. Isso pode ter implicações significativas para o vínculo e implicar importantes resultados para a saúde dos recém-nascidos com base em uma compreensão confiável da primeira tentativa comunicativa humana (*Computers in Biology and medicine*, 2023).

## **2.2 Tipos de Choro**

De acordo com os estudos feitos pelo *The Hospital Clínic-IDIBAPS* e com a colaboração *The Swiss-Catalan startup Zoundream AG*, foi descoberto que os choros têm significado e esses choros têm certas características (*Computers in Biology and medicine*, 2023).

### **2.2.1 Choro Por Fome**

Esse tipo de choro agudo, rítmico e repetitivo, acompanhado de sinais como a procura ativa pelo peito e movimentos de sucção, indica que o bebê

está com fome. É crucial agir rapidamente e oferecer o leite ao bebê sem demora. Se a alimentação for atrasada, o bebê pode ficar ansioso e mamar rapidamente, engolindo ar em excesso. Isso pode levar à formação de gases e aumentar o desconforto, piorando o choro. Portanto, é importante atender prontamente à necessidade de alimentação do bebê para evitar o desconforto adicional (*Computers in Biology and medicine*, 2023).

### **2.2.2 Choro Por Cansaço ou desconforto**

É importante prestar atenção a um choro manhoso, nasal e contínuo que aumenta em intensidade quando o bebê está exausto. Esse tipo de choro geralmente indica a necessidade de uma soneca, como se o bebê estivesse dizendo: "Preciso dormir, por favor!". Ele pode ser acompanhado por bocejos ou repetidos toques nos olhos. Esse mesmo padrão de choro também pode indicar desconforto, como uma fralda suja, calor ou desejo de mudar de posição (*Computers in Biology and medicine*, 2023).

### **2.2.3 Choro por Dor**

O choro de dor é um choro repentino e estridente, semelhante ao choro de uma criança mais velha ou de um adulto que se machucam. Geralmente, é caracterizado por longos gritos seguidos, interrompidos por pausas para recuperar o fôlego, e recomeça em seguida. Esse tipo de choro indica que o bebê está experimentando algum tipo de desconforto físico ou dor (*Computers in Biology and medicine*, 2023).

### **2.2.4 Por Cólica**

Quando um bebê está sofrendo de cólicas, é comum que seu choro seja mais intenso e até mesmo acompanhado por gritos, indicando fortes dores. Além do choro, o bebê ficará extremamente inquieto, se movendo de forma contínua devido ao desconforto causado pelas cólicas (*Computers in Biology and medicine*, 2023).

## 2.3. MFCCs

O método prevalecente em sistemas modernos de reconhecimento de voz usado para extrair características espectrais é MFCC. Trata-se de uma das técnicas de extração de características acústicas de sinais de voz mais populares usadas no reconhecimento da fala com base no domínio da frequência, sendo considerada mais precisa que técnicas que concatenam dados no domínio do tempo (TIWARI, 2010). Essa ferramenta utiliza uma escala chamada Mel, que se baseia na escala do ouvido humano (LOGAN, 2000).

## 2.4 Biblioteca Librosa

Librosa é um pacote Python para análise de música e áudio. Ele fornece os blocos de construção necessários para criar sistemas de recuperação de informação musical (Librosa, 2015).

## 2.5 Biblioteca Flask

*Flask* é um *framework web*, um módulo Python que permite desenvolver aplicações web com facilidade. Ele possui um núcleo pequeno e fácil de estender: é um *microframework* que não inclui um ORM (*Object Relational Manager*) ou recursos similares (Flask Documentation Team. Flask).

## 2.6 Biblioteca Pandas

Pandas é um módulo de análise de dados para a linguagem de programação Python. É de código aberto e licenciado sob a licença BSD. Pandas é utilizado em uma ampla gama de áreas, incluindo academia, finanças, economia, estatística, análise de dados, entre outras (Python Basics, 2021).

- Um objeto *DataFrame* rápido e eficiente para manipulação de dados com indexação integrada (Pandas, 2008).
- Ferramentas para leitura e escrita de dados entre estruturas de dados na memória e diferentes formatos: arquivos CSV e de texto, Microsoft Excel, bancos de dados SQL e o formato rápido HDF5 (Pandas, 2008).

- Alinhamento inteligente de dados e tratamento integrado de dados ausentes: obtenha alinhamento automático baseado em rótulos em cálculos e manipule facilmente dados desordenados em uma forma ordenada (Pandas, 2008).

## 2.7 Biblioteca Numpy

É uma biblioteca Python que fornece um objeto de array multidimensional, vários objetos derivados (como arrays mascarados e matrizes) e uma variedade de rotinas para operações rápidas em arrays, incluindo operações matemáticas, lógicas, manipulação de forma, ordenação, seleção, E/S, transformadas de Fourier discretas, álgebra linear básica, operações estatísticas básicas, simulação aleatória e muito mais (Numpy, 2008).

## 2.8 Modulo OS

Este módulo fornece uma maneira portátil de usar funcionalidades dependentes do sistema operacional. Se você apenas deseja ler ou escrever um arquivo, veja a função `'open()'`. Se você deseja manipular caminhos, veja o módulo `'os.path'`. Para ler todas as linhas em todos os arquivos na linha de comando, veja o módulo `'fileinput'`. Para criar arquivos e diretórios temporários, veja o módulo `'tempfile'`, e para manipulação de arquivos e diretórios em um nível mais alto, veja o módulo `'shutil'` (Python, 2024).

## 2.9 Biblioteca Joblib

Joblib é otimizado para ser rápido e robusto em dados grandes, em particular, e possui otimizações específicas para arrays numpy. É licenciado sob BSD (Joblib, 2021).

## 2.10 Reconhecimento de voz

O reconhecimento de fala, também conhecido como ASR (*Automatic Speech Recognition*), reconhecimento de fala por computador ou fala para texto, é um recurso que permite que um programa processe a fala humana em um formato escrito. Embora seja comumente confundido com o reconhecimento de voz, o reconhecimento de fala se concentra na conversão

da fala de um formato verbal para um de texto, enquanto o reconhecimento de voz busca apenas identificar a voz de um usuário individual (IBM).

## **2.11 Sistemas de Reconhecimento de Padrões e Aprendizado de Máquina**

Também conhecido pelo nome em inglês machine learning (ML), aprendizado de máquina é um subcampo da ciência da computação e da inteligência artificial que foca no desenvolvimento de modelos e algoritmos capazes de reconhecer padrões a partir de uma amostragem de dados. Tudo isso para tornar softwares capazes de fazer previsões e tomar decisões.

## **2.12 Randon Forest**

Randon Forest e uma árvore de decisão, uma estrutura de dados que representa uma sequência de decisões e suas consequências.

O Random Forest é um algoritmo de machine learning que usa vários subconjuntos de dados de treinamento para construir uma série de árvores de decisão. A ideia é treinar o sistema para que ele consiga, com base nesse aprendizado, prever o que acontece quando novos dados entram.

Depois de criadas, as árvores são utilizadas pelo sistema para fazer previsões sobre novas amostras de dados. As previsões das várias árvores (que compõe a “floresta” que dá nome ao algoritmo) são analisadas em conjunto para afinar a precisão dos resultados obtidos.

O sistema pode ser utilizado para classificações e regressões, nas linguagens R e Python. A classificação é uma análise que atribui uma categoria a uma variável: por exemplo, se usada em algum banco, ela pode classificar um cliente como potencial bom ou mau pagador. Já a regressão atribui um valor numérico à variável de saída, como preço, custo, altura (Ebac, 2024).

## **3. Metodologia**

### **3.2. Coleta de dados**

A primeira etapa para dar início a esse projeto será ter acesso ao áudio dos choros de bebês, que nesse caso usarei uma base de dados já existente criada pelo Gabor Veres disponibilizada no repositório online GitHub. Esse repositório contém o choro de vários bebês de várias idades, dos sexos masculino e feminino. Nessa base de dados cada arquivo de áudio está

devidamente organizado com cada áudio contendo as informações do bebê e o porquê ele estava chorando.

### 3.3. Redução de Ruídos Usando o Audacity

Para que na hora do treinamento da desse modelo ser o mais preciso possível, será necessário remover os ruídos de fundo dos arquivos de áudio dos choros dos bebês, para fazer isso será usado o software gratuito e de código aberto Audacity. Ele é amplamente utilizado para gravar e editar sons em diversas plataformas, como Windows, macOS e Linux. O Audacity oferece uma variedade de funcionalidades, incluindo:

- Edição de Áudio: Oferece ferramentas para cortar, copiar, colar e excluir segmentos de áudio. É possível trabalhar com múltiplas faixas de áudio simultaneamente.
- Análise de Áudio: Fornece ferramentas de análise de espectro, incluindo um visualizador de espectro de frequências, análise de envelopes, e outros gráficos para ajudar na análise detalhada dos sinais de áudio.

### 3.4 Extrair Características Espectrais MFCCs

Para extrair os MFCCs foi criado um código em python para a extração. A imagem a seguir mostra o trecho do código onde é feito a extração do MFCCs, depois dessa extração ele cria um dataset.

Figura 01 – Trecho de código da extração do MFCCs

```
19 # Função para extrair MFCCs de um arquivo de áudio
20 def extract_mfcc(file_path, n_mfcc=13, max_pad_len=400):
21     try:
22         audio, sample_rate = librosa.load(file_path, res_type='kaiser_fast')
23         mfccs = librosa.feature.mfcc(y=audio, sr=sample_rate, n_mfcc=n_mfcc)
24         pad_width = max_pad_len - mfccs.shape[1]
25         mfccs = np.pad(mfccs, pad_width=((0, 0), (0, pad_width)), mode='constant')
26         return mfccs
27     except Exception as e:
28         print(f"Error encountered while parsing file: {file_path}. Error: {e}")
29         return None
30
```

Fonte: Autoria Própria.

### 3.5 Tratamento de exceções para erro de diretório

Para avisar o usuário caso ele tenha digitado o diretório do arquivo errado, ele seja avisado do erro.

Figura 02 – Tratamento de Exceções

```
31 # Verificar se o diretório existe
32 if not os.path.exists(data_dir):
33     print(f"O diretório especificado não existe: {data_dir}")
34 else:
35     # Criação do dataset
36     data = []
37     labels = []
38     max_pad_len = 400
39
40     for label in os.listdir(data_dir):
41         label_dir = os.path.join(data_dir, label)
42         if os.path.isdir(label_dir):
43             for file in os.listdir(label_dir):
44                 file_path = os.path.join(label_dir, file)
45                 mfccs = extract_mfcc(file_path, max_pad_len=max_pad_len)
46                 if mfccs is not None:
47                     data.append(mfccs)
48                     labels.append(label)
```

Fonte: Autoria Própria.

### 3.6 Treinamento, predição e avaliação do modelo

Nessa parte do código serão chamadas as funções para dividir os dados para treino e teste do modelo. Depois irá fazer o treinamento usando o Random Forest para aprender as características de cada choro e fazer a predição.

Figura 03 – Treinamento, predição e avaliação do modelo

```
# Dividir os dados em treino e teste
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Treinamento do modelo Random Forest
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Predição e avaliação do modelo
y_pred = model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
```

Fonte: Autoria Própria.

### 3.7 Trecho final do código

O trecho final do código ele vai imprimir na tela para o usuário digitar o caminho onde se encontra o arquivo de áudio se caso a pasta estiver vazia ira aparecer uma mensagem senão ele executo todo o código de extração do MFCCs e do treinamento usando Random Forest e imprime na tela os resultados da predição e da precisão.

Figura 04 – Trecho final do código

```
# Imprimir a precisão e as predições
print(f"Accuracy: {accuracy * 100:.2f}%")
print("Predictions:")
for i, prediction in enumerate(y_pred):
    print(f"Audio {i + 1}: Predicted={prediction}, Actual={y_test[i]}")

# Função para prever a classe de um novo arquivo de áudio
def predict_new_audio(file_path):
    mfccs = extract_mfcc(file_path, max_pad_len=max_pad_len)
    if mfccs is not None:
        mfccs = mfccs.reshape(1, -1) # Reshape para (1, n_features)
        prediction = model.predict(mfccs)
        return prediction[0]
    else:
        print("Erro ao extrair MFCCs do novo áudio.")
        return None

# Solicitar ao usuário para fornecer o caminho do novo arquivo de áudio
new_audio_path = input("Por favor, insira o caminho do novo arquivo de áudio: ")
if os.path.exists(new_audio_path):
    prediction = predict_new_audio(new_audio_path)
    if prediction is not None:
        print(f"A previsão para o novo áudio é: {prediction}")
    else:
        print(f"O arquivo especificado não existe: {new_audio_path}")
```

Fonte: Autoria Própria.

## 4. Resultados

Criação de um sistema para a análise de choros de bebê foi feita com sucesso. Assim que o código e rodado ele irá fazer extrair os MFCCs, criar um Database e treinar esse modelo. Depois disso o software irá pedir para o usuário digitar o caminho do arquivo do áudio que assim que for carregado será feito a extração do MFCCs, será feita a predição de acordo com os dados do treinamento e irá calcular a precisão dessa predição. E por fim ele vai imprimir na tela o resultado da predição e a precisão do modelo.

## 5. Dificuldades encontradas

As dificuldades encontradas durante o desenvolvimento desse trabalho foi a dificuldade de encontrar uma base de dados dos choros dos bebês que supria todas as necessidades do sistema. A base encontrada tinha os requisitos necessários pra o desenvolvimento do sistema, porém por causa de conter poucos áudios o treinamento do modelo ficou impreciso, mas durante os teste ele conseguiu acertar todas as predições que mandei executar.

O modelo *RandomForest* alcançou uma precisão satisfatória nos dados de teste, indicando que as características extraídas dos áudios de choro são eficazes para diferenciar os tipos de choro. No entanto, a precisão pode variar dependendo da qualidade e quantidade de dados disponíveis.

### **5.1 Variabilidade Individual**

Cada bebê é único, e o mesmo tipo de choro pode ter significados diferentes. Isso representa um desafio para a generalização do modelo.

## **5 Conclusão**

O desenvolvimento de um sistema para reconhecimento das necessidades do bebê através do choro representa uma iniciativa extremamente promissora, com potencial para transformar significativamente o cuidado infantil. Este trabalho evidenciou a viabilidade de utilizar técnicas de aprendizado de máquina para identificar diferentes tipos de choro e associá-los a necessidades específicas, fornecendo uma ferramenta inovadora e eficaz para pais e cuidadores.

Através da coleta e análise detalhada de dados acústicos dos choros dos bebês, é possível treinar um modelo de aprendizado de máquina que irá demonstrar uma precisão satisfatória na classificação dos tipos de choro desde que essa base de dados seja grande e os arquivos de áudios estejam sem ruídos para garantir uma melhor análise. Este avanço tecnológico pode auxiliar os cuidadores sejam eles PCD ou não a responderem de maneira mais rápida e apropriada às necessidades dos bebês, contribuindo para seu bem-estar e desenvolvimento saudável.

No entanto, o estudo também identificou vários desafios que precisam ser abordados para aprimorar a eficácia e a robustez do sistema. Um dos principais desafios é a variabilidade individual entre os bebês. Cada bebê possui um padrão de choro único, e a mesma necessidade pode ser expressa de maneiras diferentes por bebês distintos. Esse fator torna a generalização do modelo mais complexa e requer um volume maior de dados para treinar o sistema de forma abrangente.

Além disso, a presença de ruídos de fundo nas gravações de áudio pode interferir na precisão do reconhecimento do choro. Embora tenham sido implementadas técnicas de redução de ruído, a qualidade das gravações ainda é um fator crucial que pode impactar o desempenho do sistema. Portanto, melhorias contínuas nas técnicas de processamento de áudio são necessárias para aumentar a robustez e a confiabilidade das predições.

A aplicação prática de um sistema de reconhecimento de choro também levanta questões relacionadas à integração tecnológica no cotidiano dos cuidadores. A facilidade de uso, a acessibilidade e a aceitação do sistema por parte dos pais são fatores críticos para seu sucesso. A colaboração com profissionais de saúde e especialistas em desenvolvimento infantil é essencial para garantir que a ferramenta seja não apenas tecnicamente eficaz, mas também prática e útil na vida real.

Para o futuro seria ideal ter uma coleta de dados bem robusta para que se consiga o máximo de arquivos de áudio dos choros dos bebês. Desenvolver uma interface intuitiva e de fácil uso para facilitar o uso e disponibilizar versões desse sistema para celulares Android e IOS e para computadores com sistemas operacionais Windows, macOS e Linux.

Em conclusão, o sistema para reconhecimento das necessidades do bebê através do choro tem o potencial de revolucionar o cuidado infantil, oferecendo uma ferramenta poderosa para entender e atender as necessidades dos bebês de maneira mais precisa e eficaz. Com pesquisas contínuas e aprimoramentos técnicos, este sistema pode se tornar um recurso indispensável para pais e cuidadores em todo o mundo, promovendo um

ambiente mais seguro e responsivo para o crescimento e desenvolvimento das crianças.

## Referência

AMORIM, Katia de Souza. **Linguagem, Comunicação e significação em bebês**. 2012. Tese (Livre-Docente) - Faculdade de filosofia, ciencias e letras, [S. l.], 2012. Disponível em: <https://www.teses.usp.br/teses/disponiveis/livredocencia/59/tde-03052019-103233/publico/KATIADESOUZAAMORIM.pdf>. Acesso em: 9 jun. 2023.

BLOG, Kinedu. **Conheça os diferentes tipos de choro do bebê**. [S. l.], 15 nov. 2019. Disponível em: <https://blog-pt.kinedu.com/choro-do-bebe/>. Acesso em: 9 jun. 2023.

BORNSTEIN, Marc H *et al.* **Neurobiology of culturally common maternal responses to infant cry**. [S. l.], 23 out. 2017. Disponível em: <https://www.pnas.org/doi/10.1073/pnas.1712022114>. Acesso em: 9 jun. 2023.

AZEVEDO, Rian de matos. **Análise e pré-processamento de sinais de áudio aplicados a treinamento e reconhecimento de fala**. [S. l.], 3 dez. 2023. Disponível em: <https://medium.com/@rianmatos.azevedo/pré-processamento-e-análise-de-sinais-de-áudio-aplicados-a-reconhecimento-de-fala-27e8fd650b99#:~:text=Pré-processamento%3A%20O%20áudio%20capturado,reduzir%20o%20tamanho%20dos%20dados>. Acesso em: 26 maio 2024.

DEEP Learning CNN – Redes Neurais Convolucionais no Deep Learning. [S. l.], 4 ago. 2023. Disponível em: [https://awari.com.br/deep-learning-cnn-redes-neurais-convolucionais-no-deep-learning/?utm\\_source=blog&utm\\_campaign=projeto+blog&utm\\_medium=Deep%20Learning%20CNN%20-%20Redes%20Neurais%20Convolucionais%20no%20Deep%20Learning](https://awari.com.br/deep-learning-cnn-redes-neurais-convolucionais-no-deep-learning/?utm_source=blog&utm_campaign=projeto+blog&utm_medium=Deep%20Learning%20CNN%20-%20Redes%20Neurais%20Convolucionais%20no%20Deep%20Learning). Acesso em: 27 maio 2024.

VERES, Gabor. donateacry-corpus: Repositório GitHub.2015. Disponível em: <https://github.com/gveres/donateacry-corpus>. Acesso em: 18 fev. 2024.

TIWARI, V. MFCC and its Applications in Speaker Recognition. In: International Journal on Emerging Technologies. 2010. ISSN: 0975-8364.

LOGAN, B. Mel Frequency Cepstral Coefficients for Music Modeling. In: International Symposium of Music Information Retrieval ISMIR, 2000.

McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. "librosa: Audio and music signal analysis in python." In Proceedings of the 14th python in science conference, pp. 18-25. 2015.

Flask Documentation Team. Flask - Web Development, One Drop at a Time. Flask Documentation. Disponível em: <https://flask.palletsprojects.com/>. Acesso em: 18 jun. 2024.

BARCELONA, Clinic. A study identifies the cause of infant crying in order to improve the infant-parent relationship. In: **A study identifies the cause of infant crying in order to improve the infant-parent relationship**. [S. l.], 19 dez. 2023. Disponível em: <https://www.clinicbarcelona.org/en/news/a-study-identifies-the-cause-of-infant-crying-in-order-to-improve-the-infant-parent-relationship?from=research>. Acesso em: 14 maio 2024.

LAGUNA, Ana *et al.* Multi-modal analysis of infant cry types characterization: Acoustics, body language and brain signals. **Computers in Biology and Medicine**, [S. l.], p. 1-10, 19 dez. 2023.

MACHADO, Mateus Lichfett. Implementação de um sistema de reconhecimento automático de voz utilizando as técnicas MFCC e Quantização Vetorial com atributos dinâmicos, de normalização e detecção de voz ativa. 2016. 149 f. Dissertação (Mestrado em Engenharia Mecânica) - Universidade Federal de Uberlândia, Uberlândia, 2016.

WHAT is pandas. [S. l.], 2021. Disponível em: <https://pythonbasics.org/what-is-pandas/>. Acesso em: 15 maio 2024.

NUMPY documentation. [S. l.], 2024. Disponível em: <https://numpy.org/doc/stable/>. Acesso em: 23 jun. 2024.

OS MISCELLANEOUS operating system interfaces. [S. l.], 21 jun. 2024. Disponível em: <https://docs.python.org/3/library/os.html>. Acesso em: 21 jun. 2024.

JOBLIB: running Python functions as pipelines jobs. [S. l.], 2021. Disponível em: <https://joblib.readthedocs.io/en/stable/>. Acesso em: 23 maio 2024.

O QUE é Random Forest? [S. l.], 3 fev. 2024. Disponível em: <https://ebaconline.com.br/blog/random-forest-seo>. Acesso em: 20 mar. 2024.

AUDACITY. 3.5.1. [S. l.], 1 fev. 2024. Disponível em: <https://www.audacityteam.org/download/>. Acesso em: 7 fev. 2024.

RIBEIRO, Paulo Henrique G. **Infant Cry**. [S. l.], 22 jun. 2024. Disponível em: <https://github.com/PauloH290199/Infant-Cry/tree/main>. Acesso em: 21 jun. 2024.