



Ferramenta para Anotação de Datasets de Áudio

Projeto Final de Programação

Daniel de Sousa Moraes

Mat.: 2012403

Orientador: Sérgio Colcher

Rio de Janeiro, 2021

Sumário

1	Especificação	2
1.1	Objetivo e Escopo	2
1.2	Definição de Requisitos	2
1.2.1	Requisito Funcionais	2
1.2.2	Requisito Não Funcionais	3
2	Projeto	3
2.1	Diagrama de Casos de Uso	3
2.2	Diagramas de Atividade	4
2.3	Arquitetura e Estrutura	6
3	Testes	7
4	Manual do Usuário	8
4.1	Download e Instalação	8
4.2	Uso da aplicação	8
4.2.1	Nova Anotação	8
4.2.2	Carregar Anotação	10
4.2.3	Playlist e Anotação	11
4.2.4	Concluir Anotação e Exportar CSV	12

1 Especificação

Este documento descreve a especificação e implementação de um projeto de software voltado para a anotação de datasets de áudios.

1.1 Objetivo e Escopo

O dataset é uma das mais importantes partes em qualquer projeto de *Machine Learning* (ML) ou *Deep Learning* (DL). A qualidade de um dataset reflete diretamente no sucesso de uma aplicação ML/DL. A popularização dessas aplicações tem aumentado significativamente os dados disponíveis. Porém, nas tarefas que envolvem áudio, assim como outras categorias de mídia, a existência de dados anotados adequadamente e de forma correta para a determinada tarefa ainda é um dos grandes desafios.

Assim, este trabalho propõe o desenvolvimento de uma ferramenta que auxilie na anotação de mídias de áudio para tarefas de ML/DL. A ferramenta deve permitir que o usuário anotador reproduza a mídia e escolha uma classe, dentre as opções, para classificar a determinada mídia.

1.2 Definição de Requisitos

Nesta seção, definimos os requisitos funcionais e não funcionais da ferramenta proposta.

1.2.1 Requisito Funcionais

Os requisitos funcionais identificados são:

- **RF01 - Criar Anotação de Datasets:** a ferramenta deve permitir a definição de nome do dataset, classes desejadas e a especificação dos caminhos das mídias de áudio pertencentes ao dataset;
- **RF02 - Anotar áudios:** a ferramenta deve permitir que o anotador reproduza o áudio e selecione a classe da qual aquela mídia pertence. Cada áudio poderá ser anotado com uma das classes anteriormente cadastradas;
- **RF03 - Salvar Anotação:** a ferramenta deve permitir que o anotador salve a anotação em qualquer momento, possibilitando a continuidade da anotação do dataset em outros momentos;
- **RF04 - Carregar Anotação:** a ferramenta deve permitir que o anotador carregue uma anotação iniciada anteriormente e salva, e continue a anotação;
- **RF05 - Exportar Datasets:** a ferramenta deve possibilitar a exportação da anotação final do dataset em um arquivo `.csv`.

1.2.2 Requisito Não Funcionais

Os seguintes requisitos não funcionais foram levantados:

- **RNF01 - Consistência:** a ferramenta deve assegurar a criação de datasets com todas as informações necessárias, garantindo o cadastro do nome, classes e dos arquivos de áudio necessários para a anotação;
- **RNF02 - Persistência dos Dados:** a ferramenta deve garantir que a anotação salva pelo anotador seja persistente e esteja disponível para continuidade da anotação;
- **RNF03 - Usabilidade:** a ferramenta deve ser de fácil uso;
- **RNF04 - Estabilidade:** a ferramenta deve considerar a utilização de datasets com grande número de mídias e garantir o funcionamento com manutenção do desempenho.
- **RNF05 - Desempenho:** a ferramenta deve prover uma operação com baixo consumo de memória e recursos da máquina.

2 Projeto

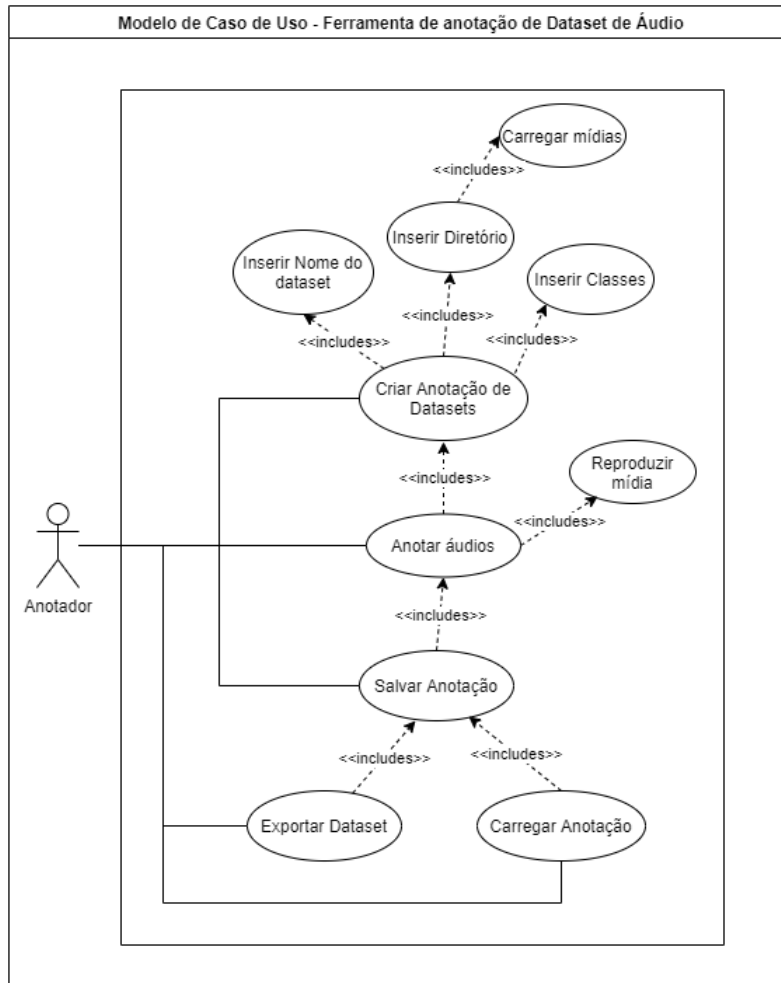
Esta seção detalha o desenvolvimento da ferramenta proposta, descrevendo como a modelagem da ferramenta, arquitetura e estruturação abrangem os requisitos definidos para a aplicação.

2.1 Diagrama de Casos de Uso

Baseado nos requisitos funcionais e não funcionais definidos na seção anterior, elaboramos o diagrama de casos de uso da ferramenta, mostrado na Figura 1. Ele possibilita a compreensão do comportamento geral da aplicação em termos de funcionalidades.

No contexto da ferramenta proposta temos como ator, o anotador. Este é o responsável por ouvir os arquivos de mídia e determinar a que classe cada um pertence. O anotador pode executar as seguintes tarefas: **Criar Anotação de Datasets**, o que inclui a inserção dos dados como nome, classes e o diretório contendo as mídias; **Anotar os áudios**, após a criação da anotação; ele pode também **Salvar Anotação** que está em andamento; **Carregar Anotação** uma anotação salva para dar continuidade; e ao final ele pode **Exportar o Dataset** como toda a anotação finalizada.

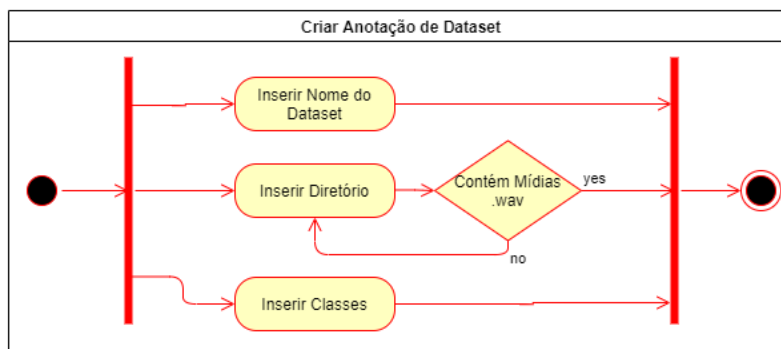
Figura 1: Diagrama de Casos de Uso



2.2 Diagramas de Atividade

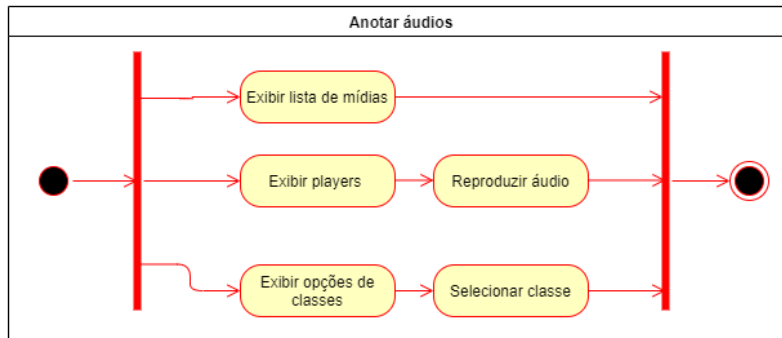
Os diagramas de atividades a seguir, descrevem os passos para a conclusão das principais funcionalidades da ferramenta. A primeira atividade é a de criar uma anotação de dataset, como mostra a Figura 2, consiste em fornecer o nome, o diretório das mídias e as classes desejadas.

Figura 2: Diagrama de Atividade - Criar Anotação do Dataset



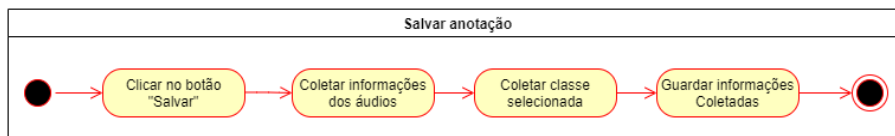
A segunda atividade é a de anotação do dataset, Figura 3. Ela consiste em exibir os arquivos de mídia inseridos pelo usuário e permitir que o usuário reproduza cada mídia e selecione uma das classe cadastradas para anotar aquela mídia.

Figura 3: Diagrama de Atividade - Anotar áudios



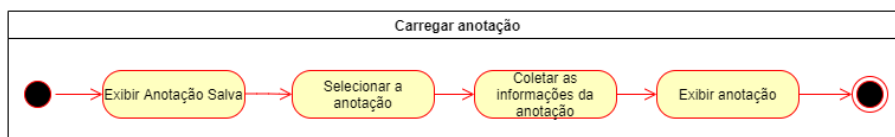
A ação de salvar uma anotação consiste na coleta das informações exibidas e modificadas pelo usuário até o momento e no envio para o armazenamento local da ferramenta. A Figura 4 mostra os passos.

Figura 4: Diagrama de Atividade - Salvar anotação



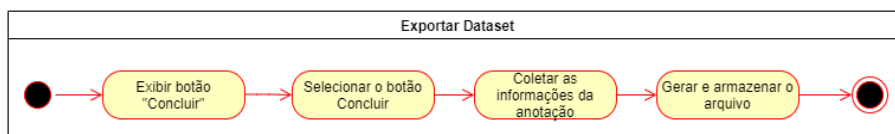
Para carregar uma anotação a ferramenta exibirá a anotação mais recente salva, permitindo que o usuário a selecione. A ferramenta então coleta as informações do armazenamento e redireciona o usuário para a área de anotação. É o que mostra a Figura 5.

Figura 5: Diagrama de Atividade - Carregar anotação



E por último, o usuário poderá exportar a anotação do dataset em um arquivo .csv. Para isso, o usuário deve selecionar o botão de conclusão da anotação, quando este estiver disponível na interface. O arquivo será gerado e armazenado em um diretório da máquina do usuário.

Figura 6: Diagrama de Atividade - Exportar Dataset

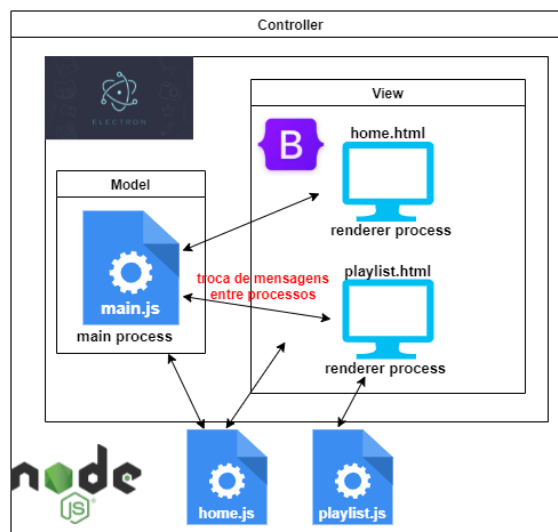


2.3 Arquitetura e Estrutura

O processo de desenvolvimento da ferramenta inicia com a utilização dos *frameworks*: **Electron**¹, para desenvolvimento da aplicação desktop; **Node.js**², para integração do ambiente de execução; e do **Bootstrap**³, para estrutura visual da aplicação.

A utilização dessas ferramentas permite o desenvolvimento do projeto seguindo o padrão de projeto *Model-View-Controller*, onde cada uma exerce uma função. O **Electron** exerce tanto a função de *Model* como de *View*, junto com o **Bootstrap**. O **Node.js** é o responsável pela função *Controller*, integrando as duas primeiras, como mostra a Figura 7.

Figura 7: Arquitetura da Ferramenta



A estruturação final de arquivos da ferramenta é composta da seguinte forma:

- **arquivo package.json:** arquivo de configuração do *framework* Electron contendo informações de dependências de pacotes do programa e informações do autor;
- **diretório assets:** um diretório com arquivos de mídia utilizadas pelo programa;
- **diretório out:** diretório de destino do arquivo **.csv** que será exportado pelo programa;
- **diretório src:** contém o arquivo **main.js** responsável pela inicialização da aplicação e processo principal. Contém também três subdiretórios com os códigos fonte das telas (**html**) e da estruturação visual (**css**), além do diretório com os scripts de controle (**js**).

¹<https://www.electronjs.org/>

²<https://nodejs.org/pt-br/>

³<https://getbootstrap.com/>

A Figura 8 ilustra essa estrutura da ferramenta a partir do diretório raiz AUDIODT_ANOTATOR.

Figura 8: Estrutura de Arquivos

```
AUDIOTD_ANOTATOR
|
| .gitignore
| package-lock.json
| package.json
| README.md
|
|+---assets
|   |
|   | arrow-left.svg
|   | arrow-right.svg
|   |
|+---out
|   |
|   | DataSetTeste.csv
|   |
|+---src
|   |
|   | main.js
|   |
|   |+---css
|   |   |
|   |   | main.css
|   |   |
|   |+---html
|   |   |
|   |   | home.html
|   |   | playlist.html
|   |   |
|   |+---js
|   |   |
|   |   | home.js
|   |   | playlist.js
```

3 Testes

Para assegurar o funcionamento e verificar a qualidade da ferramenta, utilizamos os *frameworks* de teste Spectron⁴ e Jest⁵. O teste valida a inicialização da interface da aplicação como mostra a Figura 9.

Figura 9: Log de Testes Spectron e Jest

```
C:\Windows\System32\cmd.exe
Microsoft Windows [versão 10.0.19043.1110]
(c) Microsoft Corporation. Todos os direitos reservados.

D:\Documentos\dev\audioidt_annotator>npm run test

> audioidtanotator@0.0.0 test D:\Documentos\dev\audioidt_annotator
> jest

PASS test/test.js (8.098 s)
  ✓ should launch app (7 ms)
  ✓ Displays App window (2 ms)

Test Suites: 1 passed, 1 total
Tests: 2 passed, 2 total
Snapshots: 0 total
Time: 8.198 s, estimated 9 s
Ran all test suites.

D:\Documentos\dev\audioidt_annotator>
```

⁴<https://www.electronjs.org/spectron>

⁵<https://jestjs.io/pt-BR/>

4 Manual do Usuário

Nesta seção descrevemos o processo de instalação e utilização da ferramenta. Antes de fazer o download da ferramenta, o usuário deve instalar e configurar a ferramenta **Node.js**⁶. A instalação da ferramenta **Git**⁷ é opcional.

4.1 Download e Instalação

Para download da ferramenta, o usuário pode acessar o link do repositório no GitHub: https://github.com/danielsm/audiodt_annotator. Ou fazer o download utilizando a ferramenta Git com o seguinte comando:

```
$ git clone https://github.com/danielsm/audiodt_annotator.git
```

Com a ferramenta baixada, o usuário deve abrir o diretório raiz da ferramenta, (audiodt_annotator/). Uma vez neste diretório, são necessários apenas dois comandos para instalação e inicialização da aplicação:

```
$ npm install  
$ npm start
```

4.2 Uso da aplicação

A tela inicial da aplicação possui duas abas: **Nova Anotação** e **Carregar Anotação**. A primeira possui um formulário para criação de um novo dataset. A segunda aba mostra a anotação mais recente feita, permitindo que o usuário carregue e continue essa anotação. As subseções a seguir descrevem cada uma das abas.

4.2.1 Nova Anotação

A aba **Nova Anotação** contém um formulário com três campos de entrada e três botões, como mostra a Figura 10. A sequência de preenchimento desse formulário pode ser como a seguir:

1. No campo **Nome do Dataset** o usuário pode digitar o nome desejado para o dataset;
2. Para preencher o segundo campo, **Diretório dos Áudios**, o usuário precisa clicar no botão **Escolha o diretório**. Este irá exibir uma janela de gerenciamento de arquivos do sistema operacional, onde o usuário poderá selecionar um diretório que contenha os arquivos **.wav** pretendidos;

Caso o usuário selecione um diretório que não possui arquivos de áudio **.wav**, a ferramenta mostra um alerta informando ser necessário a existência dos arquivos de áudio no diretório selecionado. Como demonstra a Figura 11.

⁶<https://nodejs.org/pt-br/>

⁷<https://git-scm.com/>

Figura 10: Tela Inicial - Aba Nova Anotação

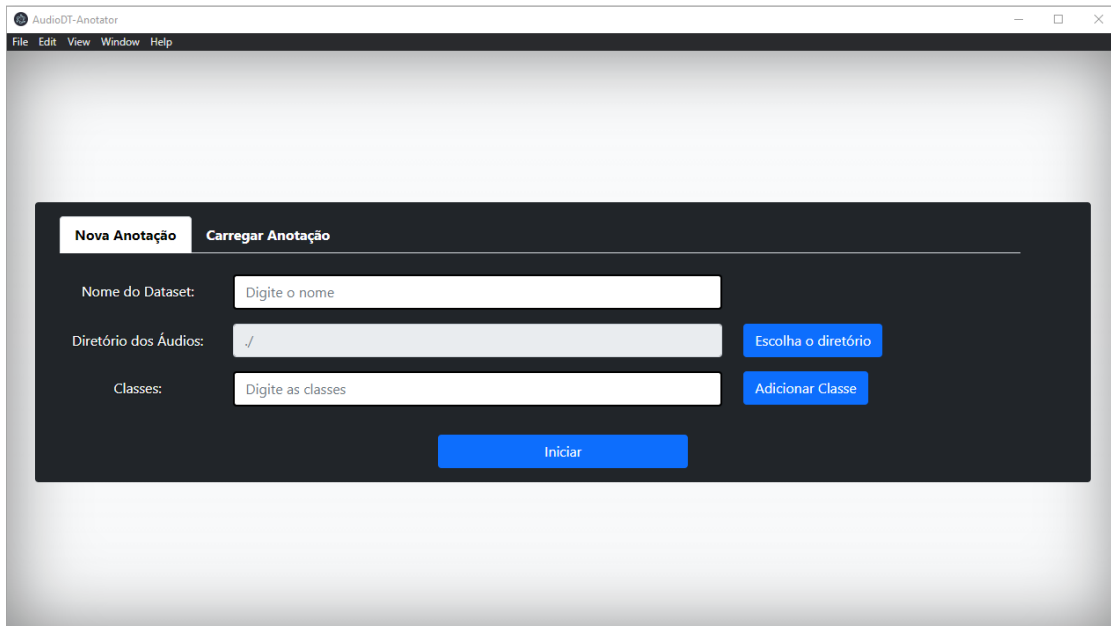
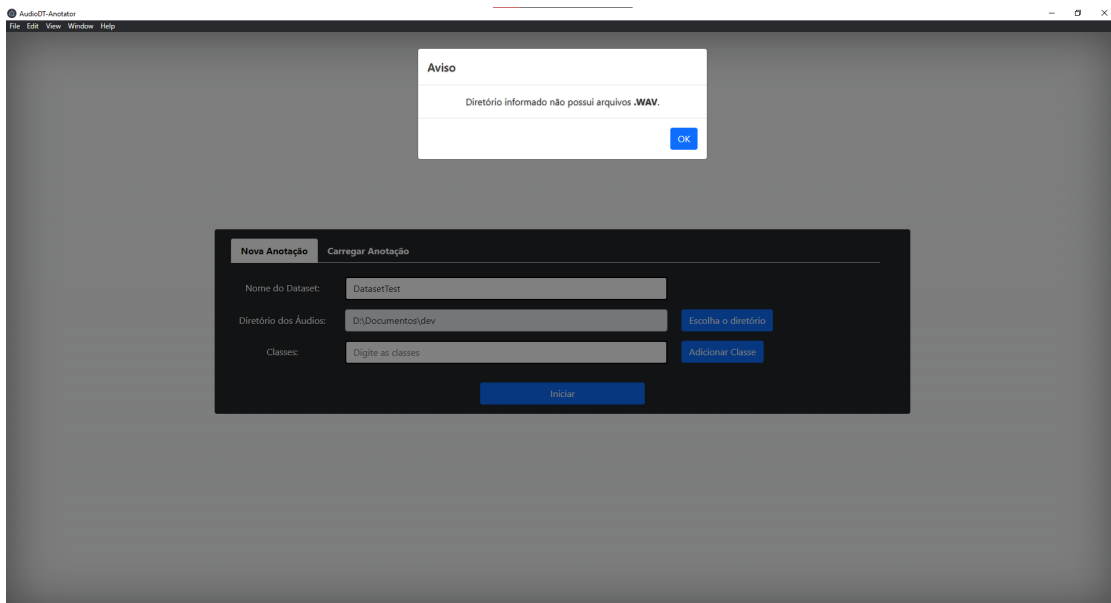


Figura 11: Tela Inicial - Alerta de diretório sem áudios



3. No campo **Classes** o usuário deve digitar as classes que deseja utilizar na anotação dos áudios. A cada classe adicionada, uma lista de classes é atualizada logo abaixo do campo de entrada, como mostra a Figura 12. Cada classe pode ser excluída clicando no botão vermelho situado no lado esquerdo;
4. Após o preenchimento de todos os campos, o usuário deve clicar no botão **Iniciar**, que o levará para a tela de anotação do dataset. Caso algum dos campos não tenham sido preenchido e o usuário tenha clicado no botão **Iniciar**, o formulário exibe mensagens informando quais campos ainda necessitam preenchimento. A Figura 13 exemplifica as mensagens para os três campos, mas cada uma também pode ser exibida isoladamente.

Figura 12: Tela Inicial - Adição e Listagem das Classes

The screenshot shows the 'AudioDT-Anotator' application window. The 'Nova Anotação' tab is active. The form contains the following elements:

- Nome do Dataset:** Text input field containing 'DataSetTeste'.
- Diretório dos Áudios:** Text input field containing 'D:\Documentos\audios' and a blue button labeled 'Escolha o diretório'.
- Classes:** Text input field containing 'teste3' and a blue button labeled 'Adicionar Classe'.
- Class List:** A list showing 'teste1' and 'teste2', each with a red square icon containing a white 'X'.
- Iniciar:** A large blue button at the bottom center.

Figura 13: Tela Inicial - Mensagens de Validação do formulário

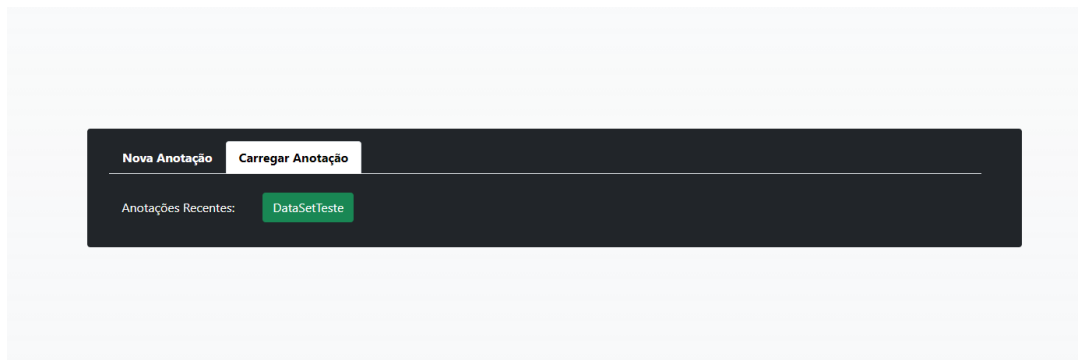
The screenshot shows the 'AudioDT-Anotator' application window with validation messages displayed below the input fields in the 'Nova Anotação' tab:

- Nome do Dataset:** The input field has a red border. Below it, the text 'Digite o nome' and 'O nome do Dataset é obrigatório' is displayed in red.
- Diretório dos Áudios:** The input field has a red border. Below it, the text 'Um diretório precisa ser selecionado' is displayed in red.
- Classes:** The input field has a red border. Below it, the text 'Pelo menos uma classe precisa ser adicionada' is displayed in red.

4.2.2 Carregar Anotação

A aba **Carregar Anotação** lista a anotação mais recente salva pelo usuário. Ela cria um botão com o nome do dataset adicionado para aquela anotação, como mostra a Figura 14, e permite que usuário selecione essa anotação e carregue-a para a continuidade da anotação. Ao clicar nesse botão o usuário é direcionado para a tela de anotação.

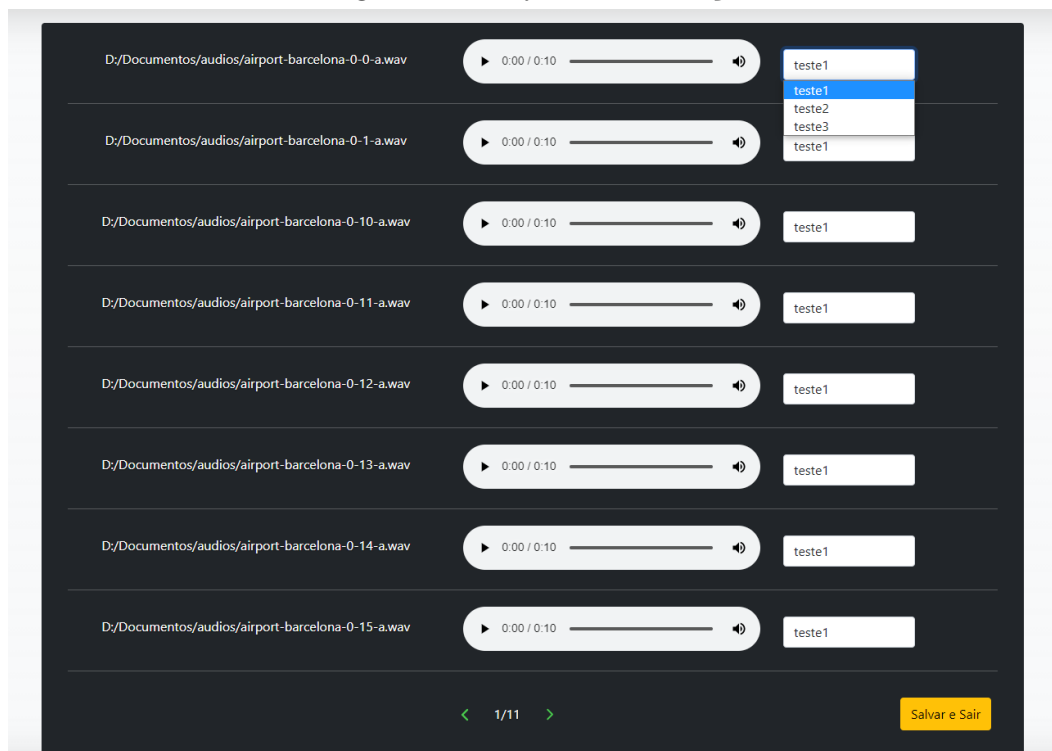
Figura 14: Tela Inicial - Aba Carregar Anotação



4.2.3 Playlist e Anotação

Após criar uma nova anotação ou carregar uma anotação recente, o usuário é direcionado para a tela de anotação dos áudios, mostrada na Figura 15. Nessa tela a visualização dos áudios é dividida em páginas. A página contém uma lista de áudios (esquerda) com *players* (centro) que possibilitam ao usuário reproduzir o áudio, podendo navegar na duração e aumentar o volume de reprodução.

Figura 15: PlayList e Anotação



Além disso, para cada áudio há um elemento de seleção (direita) que permite o usuário escolher a que classe previamente cadastrada aquele áudio pertence, como exemplifica a Figura 15, onde um elemento de seleção exibe três opções de classe para um áudio.

Na parte inferior central das páginas, a ferramenta exibe a indicação da página atual sendo anotada em relação ao total de páginas. E também duas setas de

navegação que podem ser usadas para avançar para a próxima página ou retornar para a página anterior.

Ainda na parte inferior, cada página possui também um botão amarelo, **Salvar e Sair**, que possibilita que o usuário salve a anotação naquele ponto e seja direcionado para a página inicial da aplicação.

4.2.4 Concluir Anotação e Exportar CSV

Na última página da anotação, a ferramenta exibe, além dos elementos descritos até agora, um botão **Concluir e Exportar CSV**, Figura 16. Ao clicar neste botão o usuário finaliza a anotação do dataset e exporta todo dataset anotado em um arquivo **.csv** que recebe o nome do dataset adicionado na tela inicial. Após o clique no botão a ferramenta exibe uma mensagem de confirmação da exportação do arquivo no diretório **out/**, como mostra a Figura 17.

Figura 16: Botão Concluir e Exportar CSV

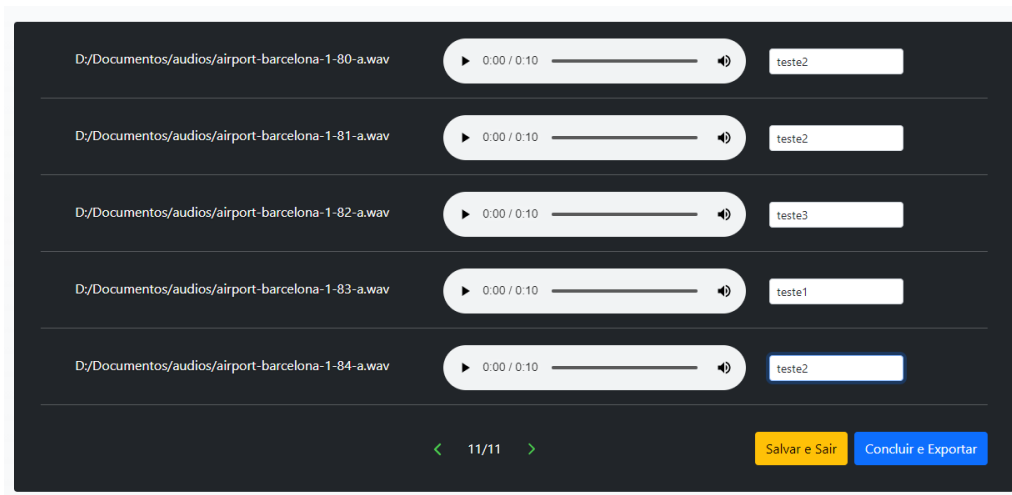


Figura 17: Mensagem de confirmação da exportação do CSV

