

BDM 3035 - Big Data Capstone Project

**MILESTONE REPORT 04 FOR  
SPEECH EMOTION RECOGNITION PROJECT**

SUBMITTED TO

**Instructor Meysam Effati**



SUBMITTED BY

**ADRIANA M. PENARANDA BARON (C0898944)**

**ARUNA GURUNG (C0896129)**

**CARLOS A. REY PINTO (C0868575)**

**HALDO J. SOMOZA SOLIS (C0904838)**

**PUJAN SHRESTHA (C0901167)**

July 20, 2024

**Table of Contents**

**I. INTRODUCTION ..... 3**

**II. PROGRESS REPORT ..... 3**

**III. MODIFIED TIMELINE TABLE ..... 5**

**IV. NEXT STEPS ..... 6**

**V. CHALLENGES FACED ..... 7**

**VI. LESSONS LEARNED ..... 7**

**VII. CONCLUSIONS ..... 8**

## **I. INTRODUCTION**

The Speech Emotion Recognition (SER) project focuses on creating a Python-based model that can detect human emotions from speech using the Librosa library and machine learning techniques. This application is particularly useful for call centers, where recognizing customer emotions can improve service quality and boost conversion rates. By leveraging the dataset, which includes emotionally labeled speech and song files, we will extract crucial audio features using Librosa, train a machine learning model, and assess its accuracy.

This project provides practical experience in audio processing and machine learning, showcasing the effective integration of these technologies in understanding human emotions. By carrying out this project, we will gain hands-on experience in audio processing, feature extraction, and machine learning model training, offering a thorough understanding of SER systems.

The GitHub link were is uploaded the milestones notebooks is:

[https://github.com/adripenaranda/Speech\\_Emotion\\_Recognition\\_Project](https://github.com/adripenaranda/Speech_Emotion_Recognition_Project)

## **II. PROGRESS REPORT**

### **SUMMARY OF TASKS COMPLETED**

In this step on the first place we checked if the dataset was balance to try to improve our model performance, we realized the data was very lightly unbalance but still we decided to applied SMOTE to balance class distributions and used GridSearchCV for hyperparameter tuning of an MLP classifier. However, the model performed better without SMOTE, so we finally decided to leave the same model we had before, just improving the hyperparameters alpha to 0,001 and hidden layers to 500 we could improve our accuracy to 79,96%.

Also We developed a function to generate and visualize spectrograms for different emotions and we analyzed them to gain insights.

### **KEY ACHIEVEMENTS AND MILESTONES REACHED**

These are the achievements and milestones reached

**Milestone 1:**

- Gather the dataset.
- Explore and understand the dataset structure.
- Extract audio features using Librosa.
- Split the data into training, validation, and test sets.
- Select appropriate machine learning model for emotion recognition.
- Implement initial model using data.
- Train model on the training data
- Test model on the validation set.
- Evaluate initial model performance and identify areas for improvement.

**Milestone 2:**

- Fine-tune model hyperparameters.
- Implement additional feature extraction techniques if needed.
- Retrain and test models with refined parameters.
- Perform cross-validation to ensure model robustness.
- Compare performance metrics and select the best model.

**Milestone 3:**

- Check data balance
- Finish tuning model
- Create plots of spectrograms for different emotions.
- Visualize spectrograms to analyze frequency patterns related to different emotions.
- Document findings from spectrogram analysis.
- Integrate spectrogram visualization into model evaluation if it is needed.

**Milestone 4:**

- Designed API architecture for emotion recognition
- Implemented API endpoints for model prediction
- Created an easy-to-use interface for the application.
- Connected the interface to the backend API.
- Tested the full workflow from audio input to emotion output
- Ensured data security and privacy protocols
- Documented API architecture and integration process

**DEVIATIONS FROM THE ORIGINAL PLAN AND REASONS FOR THESE CHANGES.**

We encountered issues with the model's performance. Despite thorough testing and integration, the model's accuracy and reliability in predicting emotions did not meet our expectations. We are investigating potential causes, including model optimization and data preprocessing techniques. The plan remains the same for the rest of the project.

### III. MODIFIED TIMELINE TABLE

After the development of this step we don't have to modify the timeline table for the upcoming steps.

#### MILESTONE 1

Name	Assignee	Due date	Priority
Extract audio features using Librosa.	AS	May 30	High
Split the data into training, validation, and test sets.	AS	Jun 2	High
Select appropriate machine learning model for emotion recognition.	AS	Jun 4	High
Implement initial model using the preprocessed data.	AS	5 days ago	High
Create Milestone Report 1	AS	Today	Urgent

#### MILESTONE 2

Nombre	Persona asignada	Fecha límite	Prioridad
Fine-tune model hyperparameters.	AS	jun. 14	Normal
Implement additional feature extraction techniques if needed.	AG	jun. 17	Normal
Retrain and test models with refined parameters.	AS	jun. 20	Normal
Perform cross-validation to ensure model robustness.	AS	jun. 24	Normal
Compare performance metrics and select the best model.	AS	Hace 5 días	Normal
Create Milestone Report 2	AS	Hace 2 días	Normal

#### MILESTONE 3

Nombre	Persona asignada	Fecha límite	Prioridad
Finish tuning model	AS	jul. 5	Normal
Check data balance	AS	jul. 6	Normal
Visualize spectrograms to analyze frequency patterns related to different emotions.	AS	jul. 8	Normal
Create plots of spectrograms for different emotions.	AS	jul. 8	Normal
Document findings from spectrogram analysis.	AS	jul. 8	Normal
Integrate spectrogram visualization into model evaluation if it is needed.	AS	jul. 8	Normal
Create Project Milestone Report 3	AS	Hace 2 días	Normal

MILESTONE 4

Big Data Caps...

Milestone 4 / List

Más

Favoritos

Espacios

Todo

Milestone 1

List

Project Understanding

Milestone 2

List

Milestone 3

List

Milestone 4

List

Ver todos los espacios

Crear espacio

Lista

Tablero

Cronograma

Tabla

Carga de trabajo

Vista

Grupo: Estado

Subtareas: Contraer todo

Columnas

Filtros

Modo Yo

Personas asignadas

Mostrar elementos cerrados

Ocultar

Busca en tareas...

BACKLOG

6

Añadir Tarea

Nombre	Persona asignada	Fecha límite	Prioridad	
Design the API architecture for emotion recognition.	AC	jul. 11	Normal	
Implement API endpoints for model prediction.	AP	jul. 12	Normal	
Develop a simple user interface (UI) for the application.	PS	vie.	Normal	
Integrate the UI with the backend API.	AC	jul. 16	Normal	
Test the full workflow from audio input to emotion output.	PS	jul. 18	Normal	
Create Project Milestone 4	CP	jul. 20	Normal	
+ Añadir Tarea				

MILESTONE 5

Big Data Caps...

Milestone 5 / List

Más

Favoritos

Espacios

Todo

Milestone 1

List

Milestone 2

List

Milestone 3

List

Milestone 4

List

Milestone 5

List

Ver todos los espacios

Invitar

Lista

Tablero

Calendario

Gantt

Equipo

Vista

Grupo: Estado

Subtareas: Contraer todo

Columnas

Filtros

Modo Yo

Personas asignadas

Mostrar elementos cerrados

Ocultar

Busca en tareas...

PENDIENTE

5

Añadir Tarea

Nombre	Persona asignada	Fecha límite	Prioridad	
Conduct comprehensive testing of the entire system.	PS			
Fix any bugs and optimize performance.	AP			
Prepare detailed documentation of the project.	AC			
Create user guides and technical documentation.	PS			
Finalize all deliverables for project submission.	CP			
+ Añadir Tarea				

IV. NEXT STEPS

UPCOMING STEPS AND ACTIVITIES PLANNED FOR THE PROJECT.

Brief Description of Tasks to be Undertaken

Milestone 5: July 30, 2024 - System Testing and Finalization

1. Perform thorough testing of the entire system

- Explanation:** Perform extensive testing on the entire system, including the UI, API, and emotion recognition model.
- Purpose:** Identify and address any issues to ensure the system operates smoothly and accurately under various conditions.

## 2. Fix Any Bugs and Optimize Performance

- **Explanation:** Resolve any bugs found during testing and make necessary optimizations to improve the model's accuracy and system performance.
- **Purpose:** Enhance the overall functionality and reliability of the emotion recognition system.

## 3. Prepare Detailed Documentation of the Project

- **Explanation:** Compile comprehensive documentation covering all aspects of the project, including architecture, implementation, and testing procedures.
- **Purpose:** Provide a clear and thorough record of the project for future reference and potential enhancements.

## 4. Create User Guides and Technical Documentation

- **Explanation:** Develop user guides and technical documentation to assist end-users and developers in using and understanding the system.
- **Purpose:** Ensure users can effectively interact with the system and developers can maintain and improve it.

## 5. Finalize All Deliverables for Project Submission

- **Explanation:** Complete and organize all project deliverables, including the final system, documentation, and user guides.
- **Purpose:** Ensure all required components are ready for submission, demonstrating the project's completion and readiness for deployment.

## V. CHALLENGES FACED

- **API Architecture Design:** Designing the API architecture was somewhat challenging. It required careful planning to ensure scalability and efficient handling of multiple requests.
- **User Interface Development:** Developing a user-friendly interface posed challenges. Ensuring the UI was both simple and intuitive while incorporating all necessary features required careful consideration.
- **System Testing and Validation:** We have been facing challenges in ensuring the model's performance. It needs careful attention to ensure the system met reliability standards.

## VI. LESSONS LEARNED

1. **API Architecture Design:** Good planning and a strong structure are important for a system that can grow and work well.
2. **User Interface Development:** A simple and easy-to-use interface is crucial for keeping users engaged.
3. **System Testing and Validation:** Thorough testing is necessary to make sure the system works reliably and meets user needs.

## VII. CONCLUSIONS

1. **API and UI Integration:** Successfully integrating the API with the user interface ensured a seamless flow from audio input to emotion prediction output. This integration demonstrated the importance of cohesive communication between front-end and back-end components.
2. **User-Friendly Design:** The development of a simple and intuitive user interface significantly enhanced user engagement. Ensuring ease of use while maintaining functionality proved vital in providing a positive user experience.
3. **Future Enhancements:** Future work should focus on further improving the system's efficiency and scalability. Exploring advanced model tuning techniques, enhancing feature extraction processes, and incorporating more sophisticated deep learning architectures could lead to even better performance and user satisfaction.