

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)

Aug 03, 2024

TABLE OF CONTENTS

I.	INTRODUCTION.....	3
II.	PROGRESS REPORT.....	3
III.	MODIFIED TIMELINE TABLE.....	4
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

II. PROGRESS REPORT

SUMMARY OF TASKS COMPLETED

During the development of the Speech Emotion Recognition (SER) project, significant progress was made across various milestones. We implemented the initial setup and data preprocessing stages using the Librosa library to extract crucial audio features from the labeled speech and song files and trained a machine learning model to recognize emotions from speech. Iterative testing and validation were conducted to ensure the model's accuracy and reliability, followed by the integration of system components to ensure seamless interaction between the audio processing module and the emotion recognition model. Comprehensive testing of the entire system was performed, bugs were identified and fixed, and the system was optimized for better performance. All deliverables, including project code, trained model, and associated files, were finalized for submission. Detailed documentation was prepared, covering every aspect of the project from the initial setup and data preprocessing to model training, testing, and optimization.

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

Milestone 5:

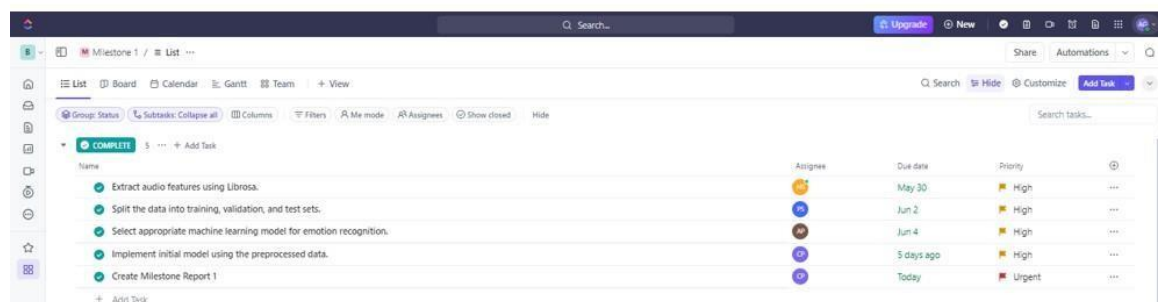
- Conduct comprehensive testing of the entire system.
- Fix any bugs and optimize performance.
- Finalize all deliverables for project submission.
- Prepare detailed documentation of the project.

DEVIATIONS FROM THE ORIGINAL PLAN AND REASONS FOR THESE CHANGES.

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



The screenshot shows a project management interface with a task list for Milestone 1. The tasks are listed in a table with columns for Name, Assignee, Due date, and Priority. The tasks are:

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

MILESTONE 2

Milestone 2 / Sub-tasks

CompartirAutomatizaciones

ListaTableroCronogramaTablaCarga de trabajo+ Vista

BuscarOcultarPersonalizarAñadir Tarea

Grupo: EstadoSubtareas: Contraer todoColumnasFiltrosModo YoPersonas asignadasMostrar elementos cerradosOcultar

Busca en tareas...

SHIPPED6+ Añadir Tarea

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

MILESTONE 3

Subir de plantNuevo

¿Quieres activar las notificaciones en el navegador?

Big Data Caps...Milestone 3 / Sub-tasks

CompartirAutomatizaciones

ListaTableroCronogramaTablaCarga de trabajo+ Vista

BuscarOcultarPersonalizarAñadir Tarea

Grupo: EstadoSubtareas: Contraer todoColumnasFiltrosModo YoPersonas asignadasMostrar elementos cerradosOcultar

Busca en tareas...

SHIPPED7+ Añadir Tarea

Nombre	Persona asignada	Fecha límite	Prioridad	
Finish tuning model	AG	jul. 5	P3	
Check data balance	AP	jul. 6	P3	
Visualize spectrograms to analyze frequency patterns related to different emotions.	AG	jul. 8	P3	
Create plots of spectrograms for different emotions.	HS	jul. 8	P3	
Document findings from spectrogram analysis.	AG	jul. 8	P3	
Integrate spectrogram visualization into model evaluation if it is needed.	PS	jul. 8	P3	
Create Project Milestone Report 3	CP	Hace 2 días	P3	

MILESTONE 4

Milestone 4 / Sub-tasks

CompartirAutomatizaciones

ListaTableroCronogramaTablaCarga de trabajo+ Vista

BuscarOcultarPersonalizarAñadir Tarea

Grupo: EstadoSubtareas: Contraer todoColumnasFiltrosModo YoAsignadosMostrar elementos cerradosOcultar

Busca en tareas...

SHIPPED6+ Añadir Tarea

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

MILESTONE 5

Milestone 5 / Sub-tasks

CompartirAutomatizaciones

ListaTableroCalendarioGanttEquipo+ Vista

BuscarOcultarPersonalizarAñadir Tarea

Grupo: EstadoSubtareas: Contraer todoColumnasFiltrosModo YoAsignadosMostrar elementos cerradosOcultar

Busca en tareas...

COMPLETADO4+ Añadir Tarea

Nombre	Persona asignada	Fecha límite	Prioridad	
✓ Conduct comprehensive testing of the entire system.	HS	jul. 20	🚩	⋮
✓ Fix any bugs and optimize performance.	AP	jul. 22	🚩	⋮
✓ Prepare detailed documentation of the project.	AG	jul. 23	🚩	⋮
✓ Finalize all deliverables for project submission.	CP	jul. 26	🚩	⋮

+ Añadir TareaCalcular

IV. NEXT STEPS

UPCOMING STEPS AND ACTIVITIES PLANNED FOR THE PROJECT.

To successfully conclude the Speech Emotion Recognition (SER) project, the following tasks are scheduled to be undertaken:

Final Report: Aug 10, 2024

1. REVIEW ALL PROJECT COMPONENTS AND ENSURE COMPLETENESS.

- Explanation:**

The initial step involves a thorough review of all project components, including the code, datasets, trained models, and documentation. This review ensures that each part of the project is complete, functional, and ready for submission.
- Purpose:**

The goal of this task is to identify and address any gaps or issues that may exist within the project. Ensuring completeness and accuracy of all components will contribute to the overall quality and reliability of the final deliverables.

2. PREPARE AND DELIVER A PRESENTATION OF THE PROJECT FINDINGS.

- Explanation:**

This step involves preparing a comprehensive presentation that summarizes the project’s objectives, methodologies, results, and key findings. The presentation will be designed to effectively communicate the value and outcomes of the project to stakeholders.
- Purpose:**

The purpose of delivering this presentation is to provide a clear and concise overview of the project to interested parties, including peers, supervisors, and potential users of the system. It serves as an opportunity to highlight the project’s achievements and the practical implications of the findings.

3. CREATE AND SUBMIT THE FINAL PROJECT REPORT

- **Explanation:**

The final task is to compile a detailed project report that encapsulates the entire project lifecycle. This report will include sections on the introduction, literature review, methodology, results, discussion, and conclusions, along with appendices for supplementary materials.

- **Purpose:**

The final project report aims to provide a comprehensive record of the project, documenting all aspects from conception to completion. Submitting this report fulfills the academic and project requirements, ensuring that the work is formally recognized and accessible for future reference.

V. CHALLENGES FACED

Several challenges were encountered throughout the project:

- **Data Quality:** Ensuring the quality of the audio data was a challenge. Some audio files had poor recording quality, which impacted feature extraction and model training.
- **Feature Extraction:** Extracting meaningful features from the audio data using Librosa required careful tuning and selection of parameters. Finding the optimal feature set for emotion recognition was time-consuming.
- **Model Accuracy:** Achieving a high level of accuracy in emotion recognition proved challenging due to the complexity of human emotions and their subtle variations in speech.
- **System Integration:** Integrating various system components, including audio processing, feature extraction, the machine learning model and the UI was also challenged, and required meticulous planning and execution.

VI. LESSONS LEARNED

The project provided valuable insights and lessons:

- **Importance of Data Quality:** High-quality, well-labeled data is crucial for the success of machine learning projects, especially in tasks like emotion recognition from speech. Investing time in data preprocessing and cleaning is essential.
- **Feature Engineering:** Effective feature engineering can significantly impact model performance. Exploring different feature extraction techniques and selecting the most relevant features is key to building a robust model.
- **Iterative Testing:** Iterative testing and validation help in identifying and addressing issues early in the development process. Regular testing ensures that the system performs reliably under various conditions.
- **Comprehensive Documentation:** Detailed documentation is vital for project transparency and reproducibility. Documenting each step, from data preprocessing to model training and testing, helps in maintaining clarity and ease of understanding for future reference.

VII. CONCLUSIONS

The Speech Emotion Recognition (SER) project successfully developed a Python-based system capable of detecting human emotions from speech. Despite challenges related to data quality, feature extraction, and model accuracy, the project achieved its goals through iterative testing and optimization. Key lessons learned include the importance of high-quality data, effective feature engineering, and comprehensive documentation. The final system demonstrates the potential for practical applications in call centers and other customer service environments, where recognizing customer emotions can improve service quality and increase conversion rates. The project provides a solid foundation for future research and development in the field of emotion recognition from speech.