Title: Report on Audio Feature-Based Age Prediction

1. Introduction:

This section provides an overview of the aim of the report, which is to develop a machine learning model capable of predicting the age of individuals based on features extracted from audio signals, specifically speech. It sets the context for the subsequent discussion by highlighting the importance and potential applications of such a model, such as in voice-controlled systems or demographic analysis.

2. Data Collection and Preprocessing:

- **Data collection:** Mention the sources of the datasets used for training and testing the model, which are "cv-valid-test.csv" and "truncated_train.csv". Briefly describe the contents of these datasets.
- **Feature extraction:** Explain the process of extracting various features from audio signals, including pitch, intensity, duration, formants, and spectral features like centroid, bandwidth, etc. Describe each feature extraction technique briefly.
- **Data cleaning:** Outline the steps taken to clean the data, including handling missing values (dropping NaNs) and normalization techniques used (z-score and min-max scaling).

3. Model Training:

- **Linear Regression Model:** Describe the implementation of the linear regression model using the normal equation approach. Explain how the model is trained using the extracted audio features and the target variable, which is age.
- **Evaluation Metrics:** Discuss the evaluation metrics used to assess the model's performance, such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2) score.

4. Model Performance Assessment:

Calculate and interpret the evaluation metrics obtained from the trained model. Discuss what each metric signifies in terms of the model's accuracy and predictive capability.

Assess the model's performance across different age groups by calculating MAE for each group. Interpret the results to understand how well the model performs for different age ranges.

5. Comparison with Library Implementation:

Describe the process of repeating the model training using scikit-learn library's linear regression implementation. Highlight any differences in implementation or results compared to the custom implementation. Visualize the actual vs. predicted ages using graphical representations and discuss any observations or insights gained from the comparison.

6. Conclusion:

Summarize the key findings and implications of the study. Emphasize the effectiveness of the developed model in predicting age based on audio features. Discuss potential applications and future directions for further improvement or refinement of the model.

7. Recommendations:

Provide suggestions for future research or development based on the insights gained from the study. This could include exploring additional features, optimizing model parameters, or deploying the model in practical applications.Highlight any areas where further investigation or experimentation could yield valuable insights or improvements in the model's performance.

8. Comparison with Library Implementation:

To ensure the robustness and correctness of the developed model, a comparison was made with the implementation provided by the scikit-learn library.

- **Custom Implementation:** The model was trained using a linear regression approach implemented without using external libraries. This involved deriving coefficients using the normal equation and making predictions based on the extracted audio features.
- **Library Implementation:** The same dataset was used to train a linear regression model using the scikit-learn library. This involved initializing and training the model using the Linear Regression class provided by scikit-learn and making predictions on the test data.
- **Visualization:** Actual vs. predicted ages were visualized using scatter plots for both custom and library implementations. The plots provided a visual comparison of the predictions made by each implementation method.
- **Performance Evaluation:** Evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2)

score were calculated for both implementations. This facilitated a quantitative comparison of the performance of the custom and library-based models.

- **Observations:** Any notable similarities or differences in performance between the two implementations were observed and discussed. Insights gained from the comparison were used to validate the correctness of the custom implementation and assess its effectiveness relative to a standard library-based approach.

9. Screenshot:

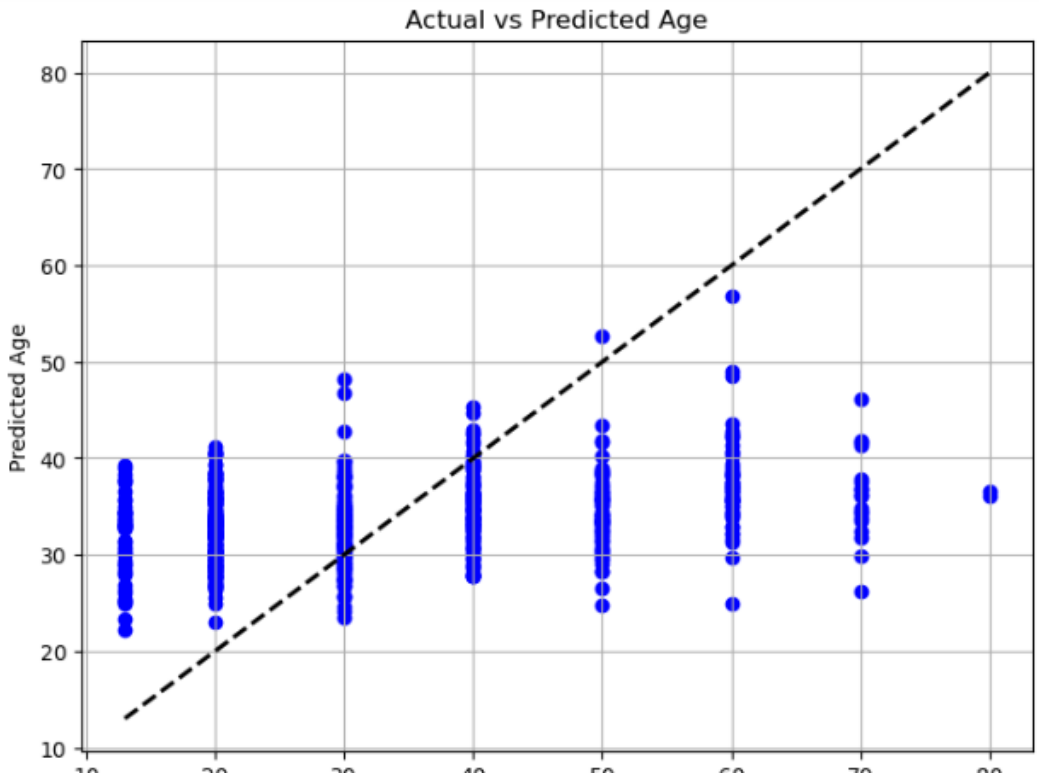
- **Merging dataset of both files with data cleaning and scaling and normalization**

	filename	text	up_votes	down_votes	age	gender	accent	Pitch	Intensity	Duration	...	Spectral_Centroid	Spectral_Bandwidth	Spec
0	cv-valid-test/sample-000003.mp3	down below in the darkness were hundreds of pe...	4	0	20	male	us	0.000002	0.003974	0.000205	...	1.311659e-06	0.000002	
1	cv-valid-test/sample-000005.mp3	down below in the darkness were hundreds of pe...	4	1	20	male	us	0.000002	0.018136	0.000112	...	1.027567e-06	0.000002	
2	cv-valid-test/sample-000008.mp3	this was the strangest of all things that ever...	1	0	30	male	england	0.000003	0.038540	0.000161	...	1.367586e-06	0.000002	
3	cv-valid-test/sample-000009.mp3	it was glaringly hot not a cloud in the sky no...	3	0	50	male	us	0.000004	0.013836	0.000105	...	1.102629e-06	0.000002	
4	cv-valid-test/sample-000014.mp3	follow the instructions here	1	0	20	male	scotland	0.000003	0.011396	0.000040	...	1.318474e-06	0.000002	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
6137	cv-valid-train/sample-014993.mp3	but the marketplace was empty and he was far f...	1	0	40	male	us	0.000002	0.005700	0.000198	...	1.463317e-06	0.000003	

- **Prediction age:**

```
Predicted Age: [33.33828715 28.90352391 31.10928312 40.52161747 29.96859734 40.65671432
 30.99154684 37.20688287 33.53015575 26.65084842 33.37732872 33.11162634
 33.32158467 33.28099163 33.34143472 28.47946058 28.71032058 36.26254613
 37.74108708 33.92146907 27.04718096 33.25963979 29.88037956 34.27491861
 34.79068589 26.74275407 32.61283745 36.07912348 36.25955144 32.79886216
 37.28436721 32.89466248 43.13313136 36.19691419 36.11629748 34.18861855
 36.3377973  34.15252566 27.3642817  31.08391008 30.2011526  25.65951051
 33.44403485 42.72189591 38.41281632 33.3883052  33.11835164 34.63046374
 34.20991685 28.59107183 29.74117165 33.36150811 34.82287159 38.8467833
 33.37317732 30.22542035 33.15974869 56.32976698 37.08421002 31.34479094
 38.31416565 34.1320758  33.60768308 33.38642225 32.80491132 31.77891689
 28.79802832 33.89262535 35.75330326 32.047647   25.00362068 34.93890141
 27.64882322 29.3259712  36.49908775 38.19640489 35.57933587 32.09708096
 31.92817077 27.81508206 37.13674688 31.80894492 34.74649201 33.59727406
 30.43540843 32.12404105 33.16637839 28.29041741 32.91535113 31.91354723
 34.74780569 40.40167856 31.6553441  37.01587805 33.01352227 36.75884723
 33.66531323 35.02196444 33.51203521 36.90821201 29.1334891  38.68034639
 36.44207306 31.51131451 27.75784286 36.31399996 36.16236431 36.40941952
 36.46266639 31.74284461 34.75540892 32.13713346 38.49053303 24.76390848
 30.78719354 37.48284935 36.80186254 34.14511076 35.49823166 27.76209695
 40.12840235 36.14623119 33.64186571 28.90649851 34.07769601 35.70388249
 32.84459321 38.08774489 28.60415173 32.30503958 33.04961324 40.42115112
 34.53870083 33.11054395 32.61823181 32.35745119 32.49865122 29.68179596
 34.73954699 29.88675595 39.1057403  31.3459949  32.73197625 27.12204468
```

- **Graph:**



- **Accuracy:**

```
Mean Absolute Error: 11.810366492063203
Mean Squared Error: 208.0043412284495
Root Mean Squared Error: 14.422355606087708
R-squared: 0.109774430370291
```

- **Assessment:**

```
MAE for teens: 18.719953175525067
MAE for twenties: 12.607557744166945
MAE for thirties: 3.759596069217707
MAE for fourties: 5.751792733526149
MAE for fifties: 15.593549117804043
MAE for sixties: 22.542878723944074
MAE for seventies: 34.032661210266006
MAE for eighties: 43.66378739026646
```