**Speech Emotion Recognition (SER) Report**

**1. Introduction**

The Speech Emotion Recognition (SER) project aims to classify emotions from audio recordings using machine learning techniques. The dataset comprises emotional speech samples, and the model is trained to recognize specific emotions using extracted audio features.

**2. Dataset and Features**

* The dataset consists of .wav audio files labeled with different emotions.
* Emotion Mapping:
  + Neutral
  + Calm ✅ *(Used in classification)*
  + Happy ✅ *(Used in classification)*
  + Sad ✅ *(Used in classification)*
  + Angry
  + Fearful ✅ *(Used in classification)*
  + Disgust ✅ *(Used in classification)*
  + Surprised
* Extracted Features:
  + **MFCC (Mel-frequency cepstral coefficients)**: Captures timbral features.
  + **Chroma**: Represents tonal content.
  + **Mel-Spectrogram**: Captures frequency distribution over time.

**3. Data Preprocessing**

* Audio features are extracted using librosa.
* The dataset is split into training (75%) and testing (25%).

**4. Model Implementation**

* **Algorithm Used**: Multi-Layer Perceptron Classifier (MLP)
* **Hyperparameters**:
  + Hidden Layer Size: 300
  + Learning Rate: Adaptive
  + Max Iterations: 500
  + Batch Size: 256
  + Alpha: 0.01 (L2 regularization parameter)

**5. Model Performance**

* **Accuracy**: {:.2f}%. *(Replace with actual accuracy)*
* **Sample Predictions**:

| **Actual Emotion** | **Predicted Emotion** |
| --- | --- |
| {} | {} |
| {} | {} |
| {} | {} |
| {} | {} |
| {} | {} |

**6. Visualization**

To further analyze model performance, we can use the following visualizations:

* **Confusion Matrix** (Heatmap representation of predictions)
* **Feature Importance Plots** (Comparing MFCC, Chroma, and Mel contributions)
* **Waveform and Spectrogram Plots** (Visualizing emotional audio samples)

**7. Conclusion and Next Steps**

* The model successfully classifies emotions with an accuracy of {:.2f}%.
* Potential Improvements:
  + **More data augmentation** to enhance training diversity.
  + **Hyperparameter tuning** to improve accuracy.
  + **Deep learning approaches** like CNNs or LSTMs for better performance.