**Model Report**

**Emotion Detection from Audio File**

**Introduction**

This report outlines the development and evaluation of a machine learning model designed to detect emotions from audio file that also contains some conditions to result output. The model aims to classify audio samples into various emotional states while ensuring that the inputs are specifically from female voices speaking in English.

**Background**

Emotion recognition from voice is a crucial component in various applications, including mental health monitoring, customer service, and human-computer interaction. Existing models often face challenges in accurately identifying emotions due to variations in gender, language, and recording quality.

**Learning Objective**

The primary objective of this project was to build and evaluate a robust model capable of detecting emotions in female voices with high accuracy. The model also had to reject non-female voices and non-English language recordings to ensure specificity.

**Activities and Tasks**

1. **Data Collection**: Gathered a diverse dataset of female voice recordings labeled with emotions such as happy, sad, angry, fearful etc.
2. **Preprocessing**: Applied noise reduction, normalization, and feature extraction (MFCCs) to prepare the audio data for model training.
3. **Model Training**: Developed and trained separate models for gender, language, and emotion detection.
4. **Integration**: Combined the models into a single pipeline to process input audio, check for gender and language, and then predict the emotion.
5. **GUI Development**: Created a user-friendly interface using Tkinter to upload audio files and display the detected emotions.

**Skills and Competencies**

* **Machine Learning**: Understanding of supervised learning, model training, and evaluation.
* **Audio Processing**: Experience with audio feature extraction using libraries like Librosa.
* **Programming**: Proficiency in Python, including libraries such as TensorFlow/Keras, Scikit-learn, and Tkinter.
* **User Interface Design**: Ability to design and implement a graphical user interface.

**Feedback and Evidence**

* **Model Performance**: The model achieved an accuracy of approximately 85% on the test dataset.
* **User Testing**: Feedback from users indicated that the GUI was intuitive and the model's predictions were generally accurate for female voices in English.
* **Error Analysis**: Cases where the model misclassified emotions were analyzed to identify patterns and potential improvements.

**Challenges and Solutions**

* **Data Imbalance**: Addressed by augmenting underrepresented classes in the dataset.
* **Noise in Recordings**: Implemented noise reduction techniques to improve the quality of feature extraction.
* **Gender and Language Detection**: Integrated robust pre-trained models for gender and language detection to filter out non-female and non-English recordings effectively.

**Outcomes and Impact**

* **Accuracy**: The model demonstrated strong performance in correctly identifying emotions from female voice recordings.
* **User Adoption**: Positive feedback from initial users suggested the model could be valuable in real-world applications.
* **Future Work**: Plans to expand the dataset, improve model robustness, and explore additional features for further enhancement.

**Conclusion**

This project successfully developed a machine learning model that accurately detects emotions from female voice recordings while filtering out non-female and non-English inputs. The integration of a user-friendly GUI facilitates easy adoption and usage. Future efforts will focus on refining the model and exploring broader applications.