Meow-by-Meow: An app to understand why your cat is yelling at you
Team 28: Brady Ali,William Craig, Zach Hafen-Saavedra, Tantrik Mukerji, Jinjing Yi

GitHub: https://github.com/jinjinglunayi/meow-by-meow\

Interface: https://meow-by-meow.streamlit.app

Overview: Cat vocalizations (“meows”) are typically directed at humans, rather than other cats. Cat meows therefore present an opportunity for computational audio analysis to improve relationships between cats and their owners. In our analysis we developed an interface for users to upload audio recordings and have them classified as “comfortable”, “uncomfortable”, or “hungry”. Our classification leverages machine learning models trained on preprocessed and augmented data from the CatMeows dataset.

Stakeholders: Cat owners and enthusiasts, veterinarians, animal psychologists, and in-general researchers and industries invested in audio-classification. Our project is a prototype for a monetized app providing the same services.

Approach:

We developed two models to predict the situation (“comfortable”, “uncomfortable”, “hungry”), breed, and sex from a publicly available dataset of recordings of cat-meows. We increase the size of our dataset using standard techniques of data augmentation. We first convert the .wav file recordings into spectrograms images and classify the images using the following models

- **K-Nearest-Neighbors (KNN):** Treats the flattened spectrogram as a point in Euclidean space and predicts the label for that point using the K-nearest neighbors to that point.

- **Convolution Neural Networks (CNN):** A regularized type of neural network that learns labels through the use of filters that enhance specific features of the input. These models have seen a considerable amount of use in computer vision and were a natural choice for classifying the spectrogram images.

Using these models, we then develop an interface for users to upload and classify personal recordings of “meows.”

Results & Strategies
Using both models, we obtain an accuracy of over 90% on validation data. The optimal choice of K for the KNN was K=1. Our CNN model obtained similar results with the use of augmented data which was augmented through the use of time shifts, stretches, random noise and frequency and time-masking.

Future Iterations:
In future iterations, we hope to train and update our models on user-supplied recordings and potentially generalize our methods and interface to other audio-classification tasks such as dog barks, baby cries, etc.