**Intro**

* Hi, I’m **Lukas Morrison**, and I’m here with **Justin Nguyen**
* We did **Speaker Recognition** for our **Final Lab Project**
* We decided to do the **MATLAB Variation** of this project

**Flow Chart**

* Here is an outline of the **Project Flow**
* We have our **Data Set**, including **Training** and **Testing** samples
* These are both run through **Feature Extraction**; the extracted metric is the **Cepstrum** of the samples
* The **Training Set** is used by the **GMM Training** process, and it calculates a **Mean**, **Variance**, and **Weight** for each sample
* These parameters makeup the **Statistical Models** of the training samples
* With these models, the **GMM Classification** process compares **Test Samples** to the **Statistical Models** and outputs **Similarity Scores**

**Data Set**

* For our **Dataset**, we collected **Training** and **Testing** speech samples
* For the **Training** set: we had **4 Subjects**, and collected **40 second Speech Samples** from each
* For the **Testing** set: we had **6 subjects** (**4 subjects** were **Included** in the training database, and **2 subjects** were **not Included**). We collected **12 second Speech Samples** from each
* We used the **Same Sample Script** across all subjects for training and testing samples for consistency of the training

**Feature Extraction**

* For **Each Sample** (from the **Training** and **Testing** sets) we performed **Feature Extraction**
* For this, we used the **MelCepst** **Function**
* It took our loaded **Speech Samples** and the **Sample Rate** as **Inputs**, which in our case were **.wav** files recorded at **8kb/s**.
* We were then provided with the **Cepstrum** of the signal as the **Output**

**Notes**

Did Branden’s group present?