LT2216/LT2816 V26
Dialogsystem
Lab: 4
Adib Wahid Quader

Ans to the Ques of Part A-VG

Answer. After training and deploying my Azure Custom Speech model (with a custom endpoint in `dm4.ts`), I tested the following target names:

- **Panthera Leo**
- **Bin Quader**
- **Sultan Mehmed**
- **Kazi Nazrul Islam**
- **Khalid Bin Walid**

Below is a comparison between the new custom model and the previous baseline system.

Endpoint ID: 0eb97935-6149-4e57-b2a7-f4bc9d8ec8b7

1. High-Confidence Improvements

Panthera Leo
Recognized as: *"Panthera, Leo"*
Confidence: 0.9321

Sultan Mehmed
Recognized as: *"Sultan Mehmed"*
Confidence: 0.9071

These results show strong improvement compared to the baseline model, where scientific and historical names previously received lower confidence or noticeable distortions.

The new model handles:

- Binomial scientific names
- Multi-word historical names

with significantly higher reliability.

2. Partial or Distorted Recognitions

Bin Quader
Recognized as:

- *"BIN card there"* – 0.7313
- *"Bin Cuddler"* – 0.4050

This shows mixed performance:

- One output had medium confidence (though incorrect).
- One output was low confidence and correctly flagged as unreliable.

Compared to the old model, which often produced very low confidence scores for such names, the custom model shows slightly improved stability but still struggles with phonetic similarity.

Kazi Nazrul Islam
Recognized as: *"Call Gino's rule Islam"*
Confidence: 0.6850

This remains a challenging multi-word foreign name. However, the confidence is moderate rather than extremely low, suggesting improved acoustic modeling compared to the baseline.

Khalid Bin Walid
Recognized as: *"Harley bin Walid"*
Confidence: 0.6008

The surname portion (*"bin Walid"*) is partially preserved, which indicates that:

- The custom model better captures recurring name patterns.
- The first name is still phonetically confused.

3. Overall Comparison with Baseline

Improvements Observed:

- Higher confidence for scientific and historical names.
- Better handling of multi-word expressions.
- Clearer distinction between high- and low-confidence cases.
- Improved stability in repeated recognition attempts.

Remaining Limitations:

- Phonetically complex South Asian and Arabic-origin names are still prone to distortion.
- Some incorrect recognitions still receive moderate confidence scores.

4. Conclusion

The Azure Custom Speech model demonstrates noticeable improvement over the baseline ASR system.

Names like **Panthera Leo** and **Sultan Mehmed** are now recognized with very high confidence (> 0.9), which was not consistently the case before.

However, complex multi-word foreign names such as **Kazi Nazrul Islam**, **Bin Quader**, and **Khalid Bin Walid** still present challenges, although confidence calibration has improved.

Overall, domain adaptation through Custom Speech has increased robustness, especially for structured and recurring name patterns.