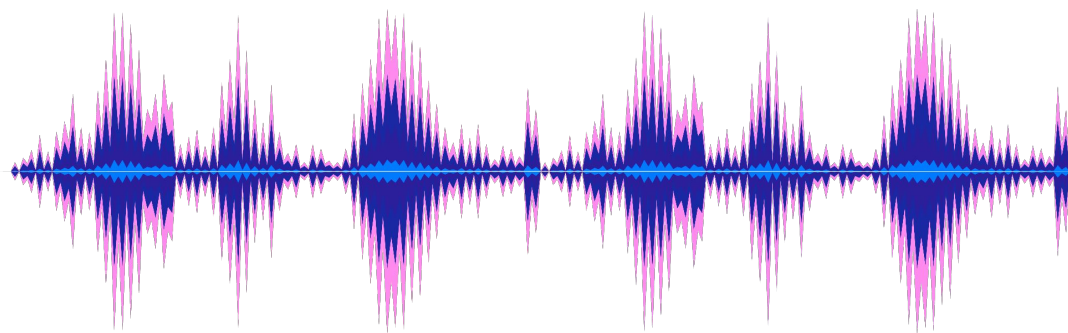
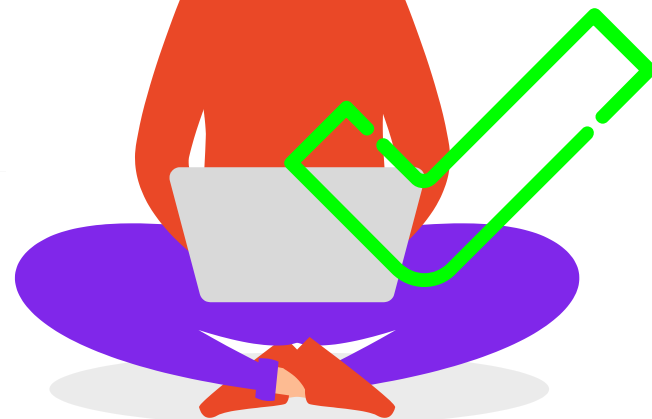


Voice ID

Verifying clients with voice



Why use voice?

- Added security level
- Accessibility
- Convenient
- Fraud prevention

HSBC UK saw a 50% drop in telephone fraud after implementing voice biometrics



Stakeholder

Current ID verification:

- Mobile App: PIN, password, FaceID
- Web App: password, 2 factor authentication with SMS/Push



\$1,824

US Dollar account

USD ▾

Balance: \$1,840



\$1 = €0.9038

EUR ▾

Balance: €1,250

+ 90

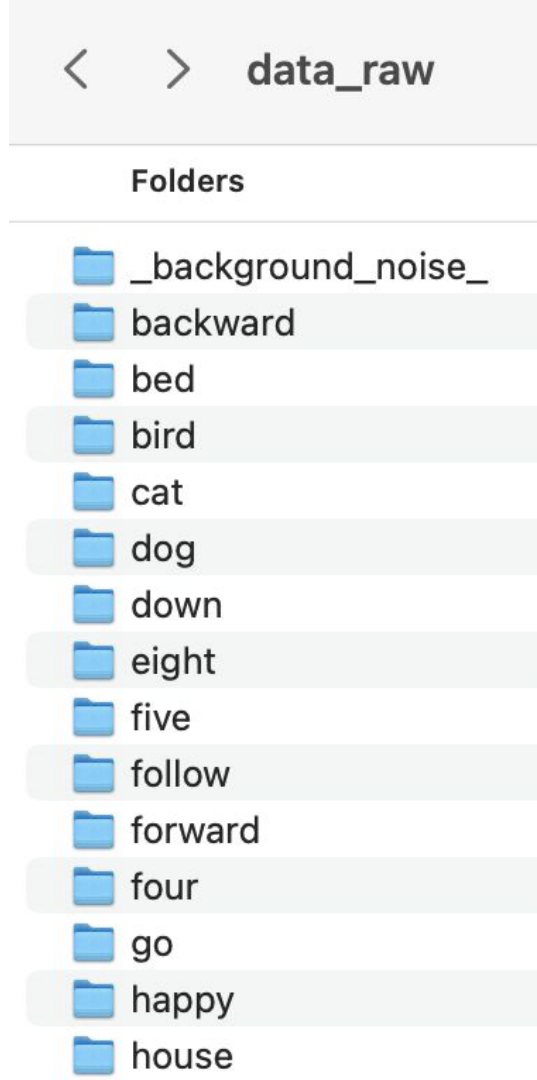
Exchange

Local

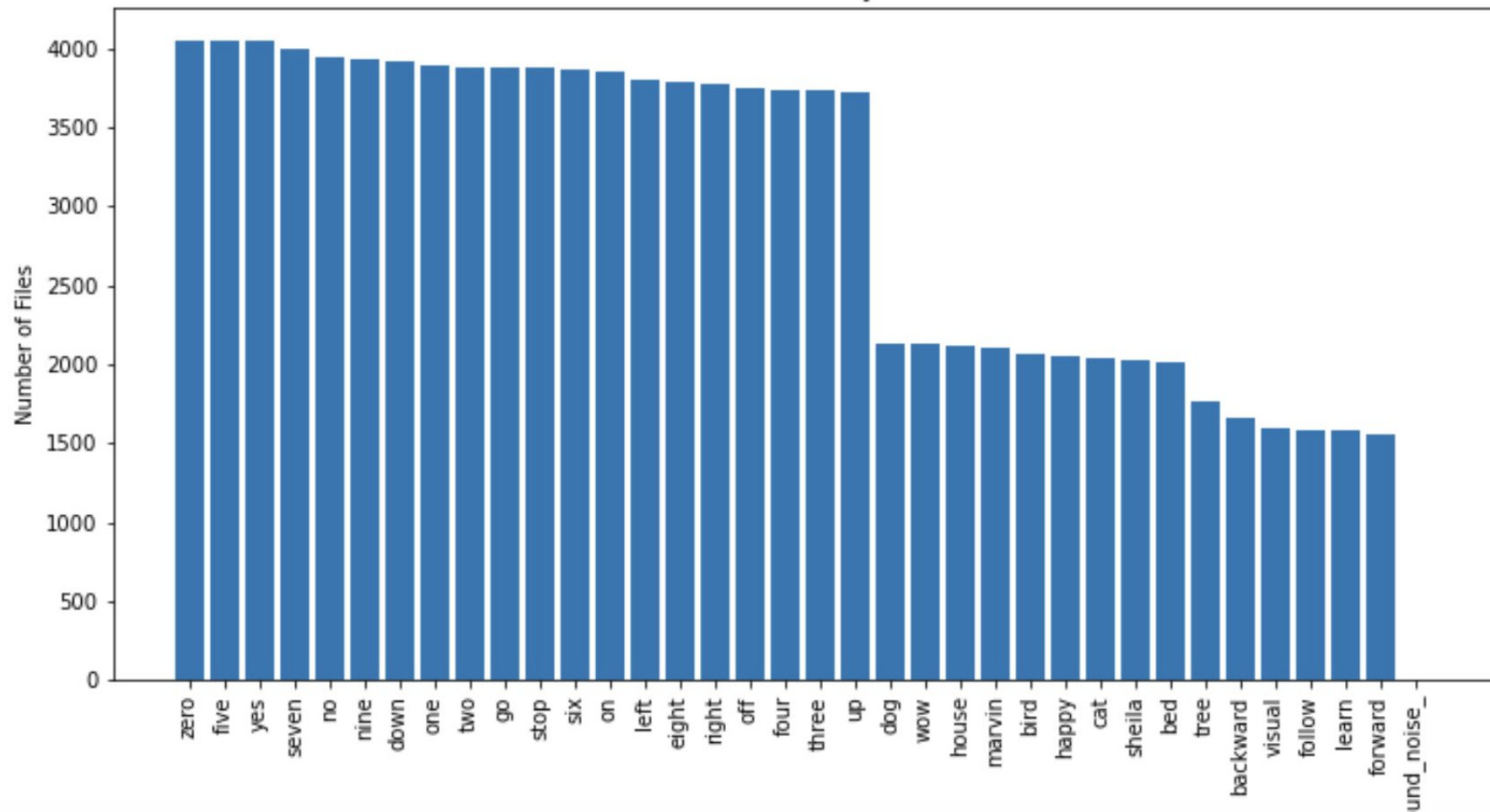
IBA

Data: Google Command Speech Data set

- 105,000 one second audio files
- .WAV files
- 35 unique words
- by 2625 users



Number of Files by Folder



WAV file

File: data_raw/backward/0a2t

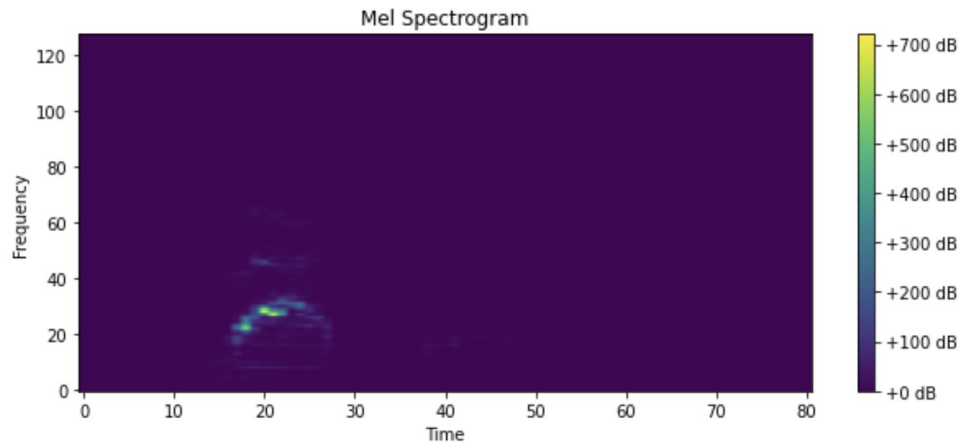
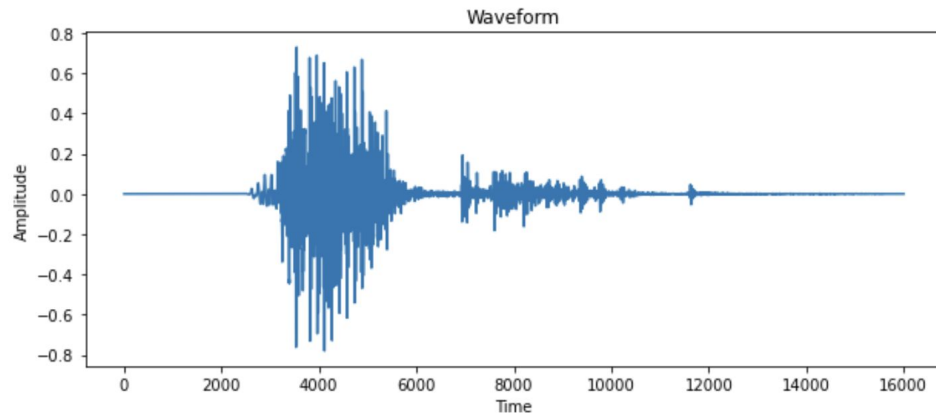
Sample Rate: 16000 Hz

Bit Depth: 16 bits

Number of Channels: 1

Duration: 1.00 seconds

File Size: 32044 bytes



Data preprocessing

WAV -> Waveform -> Resample to 8000 Hz -> tensor (1, 8000)

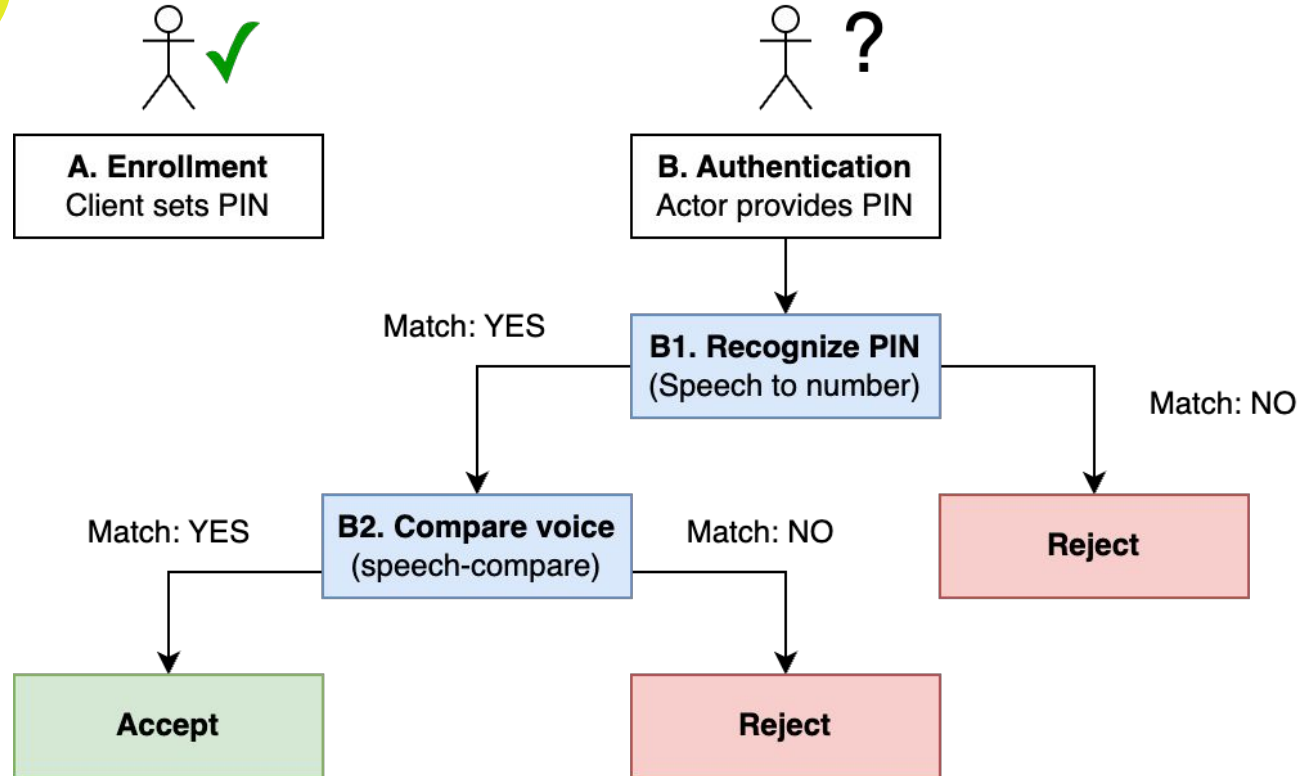
Dataset 1: Audio, Label (0-9)

Dataset 2: Audio1, Audio2, Label 1/0 (same user/not)

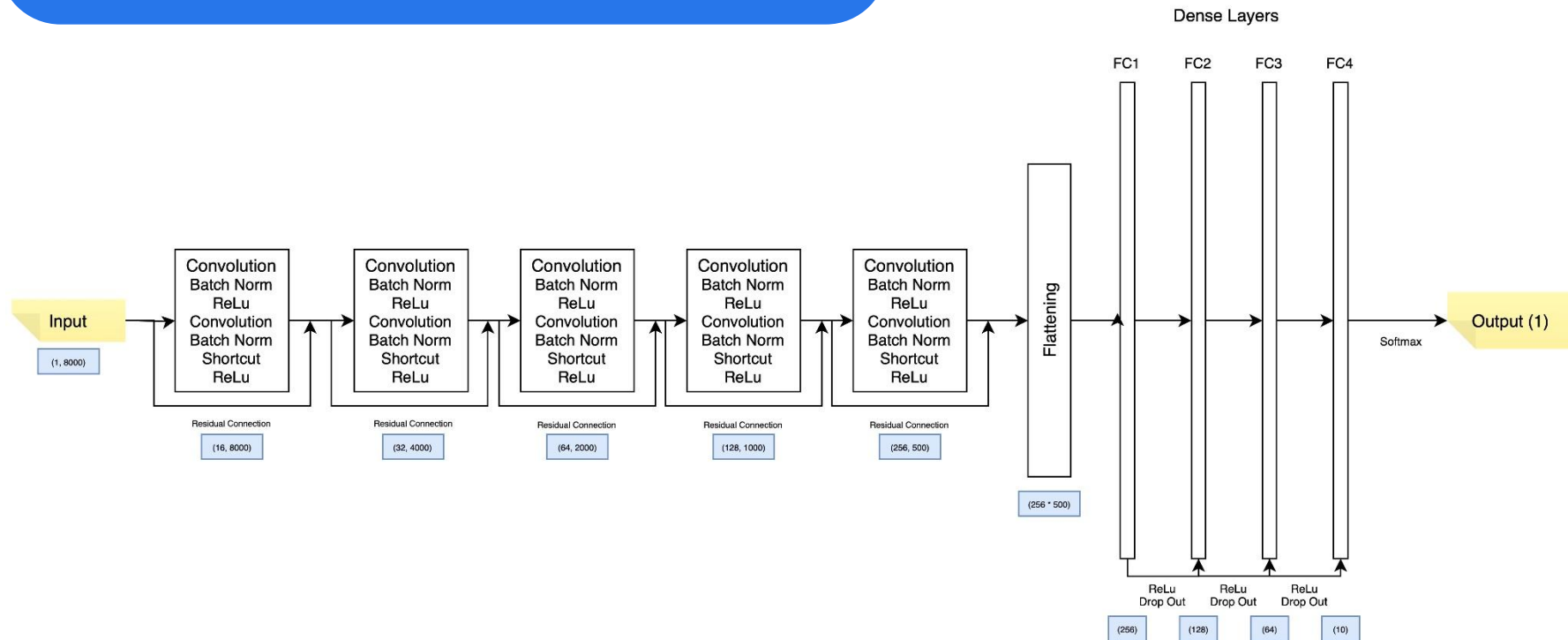
Modeling

CNN with BatchNorm, MaxPooling, ReLu, Residual;
Fully Connected, ReLu, DropOut

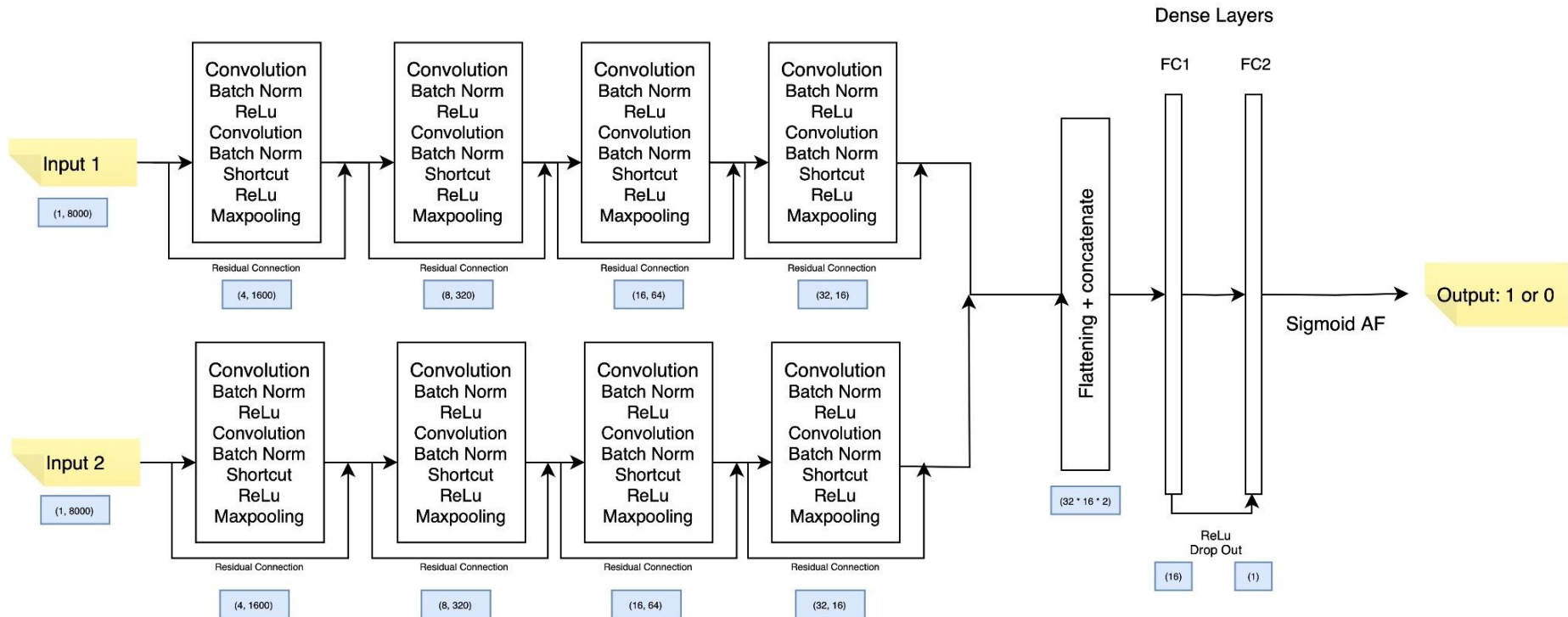
Flow



Step 1: Verifying PIN with Voice



Step 2: Verifying ID with voice



PIN verification Model

BCEWithLogitsLoss

Accuracy 87%

Train Loss: 0.5 -> 0.28

Test Loss: 0.44 -> 0.17

Hyperparameters:

dataset_limit = 38 000

batch_size = 32

lr = 0.0005

epochs=50

Voice verification Model

BCELoss

Accuracy 89%

Training Loss: 0.65 -> 0.42

Test Loss: 0.59 -> 0.29

Hyperparameters:

dataset_limit = 30 000

batch_size=64

lr=0.0005

epochs=50

weight_decay=0.0005

DEMO:

Client - Dolgor; PIN - 4163

Gradio

127.0.0.1:7860

Audio Input 1

Audio Input 2

Audio Input 3

Audio Input 4

PIN

1144

Save

Audio Input 5

Audio Input 6

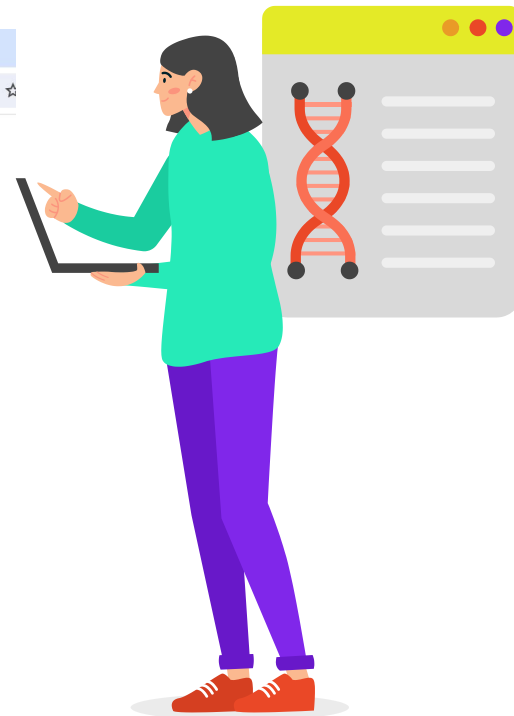
Audio Input 7

Audio Input 8

Verdict

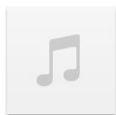
1144. Same voice: ✓✓✓✓

Verify

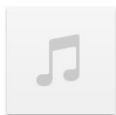


Next Steps

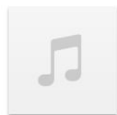
- Create a demo for system flow
- Add more training data
- Add more accents



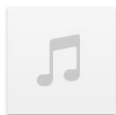
0a196374_nohas
h_2.wav



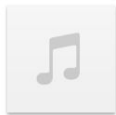
0a196374_nohas
h_3.wav



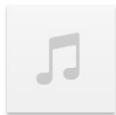
0a196374_nohas
h_4.wav



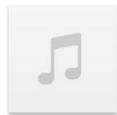
0ab3b47d_nohas
h_0.wav



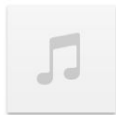
0b7ee1a0_nohash
_3.wav



0b7ee1a0_nohash
_4.wav



0b09edd3_nohas
h_0.wav



0b09edd3_nohas
h_1.wav

Recommendations

- Train model using user data
- Test system
- Implement the system

Verdict

4163. Same voice: ✓✓✓✓

Thank you!

github.com/dolgorp

