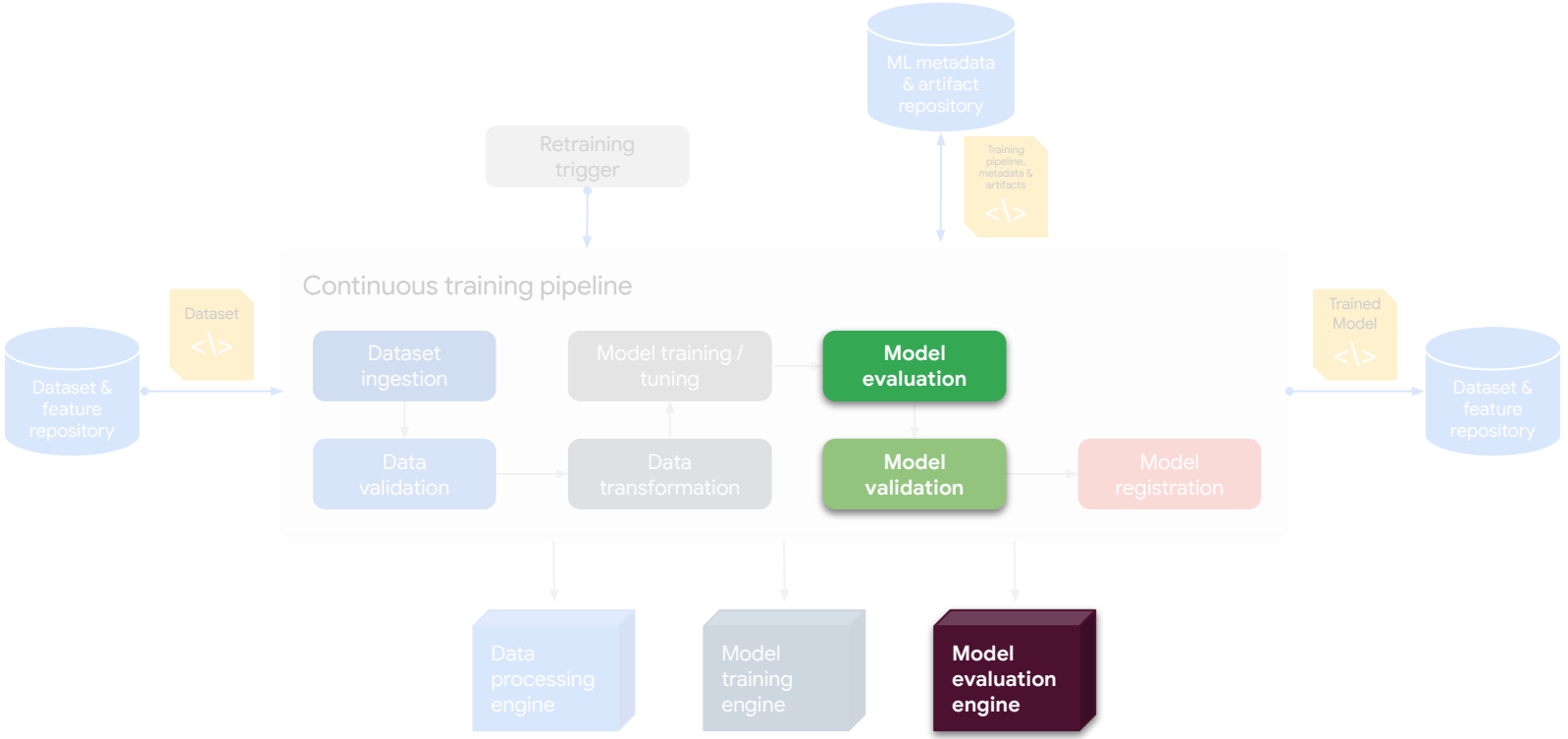
Continuous Training: Use Case Metrics

MLOps: Continuous Training



The MLOps **Personas**

ML Engineer

ML Researcher

Data Scientist

Data Engineer

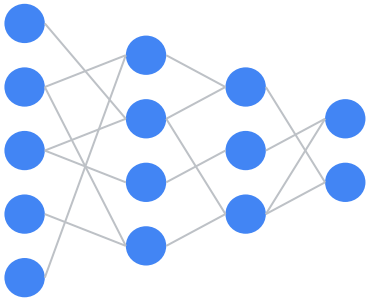
Software Engineer

DevOps

Business Analyst

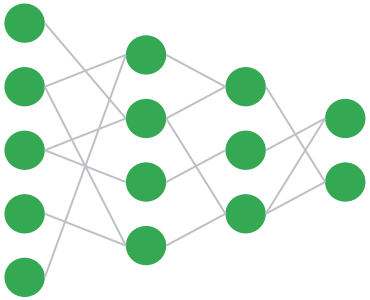
Test Set Accuracy

V1



Accuracy: **84%**

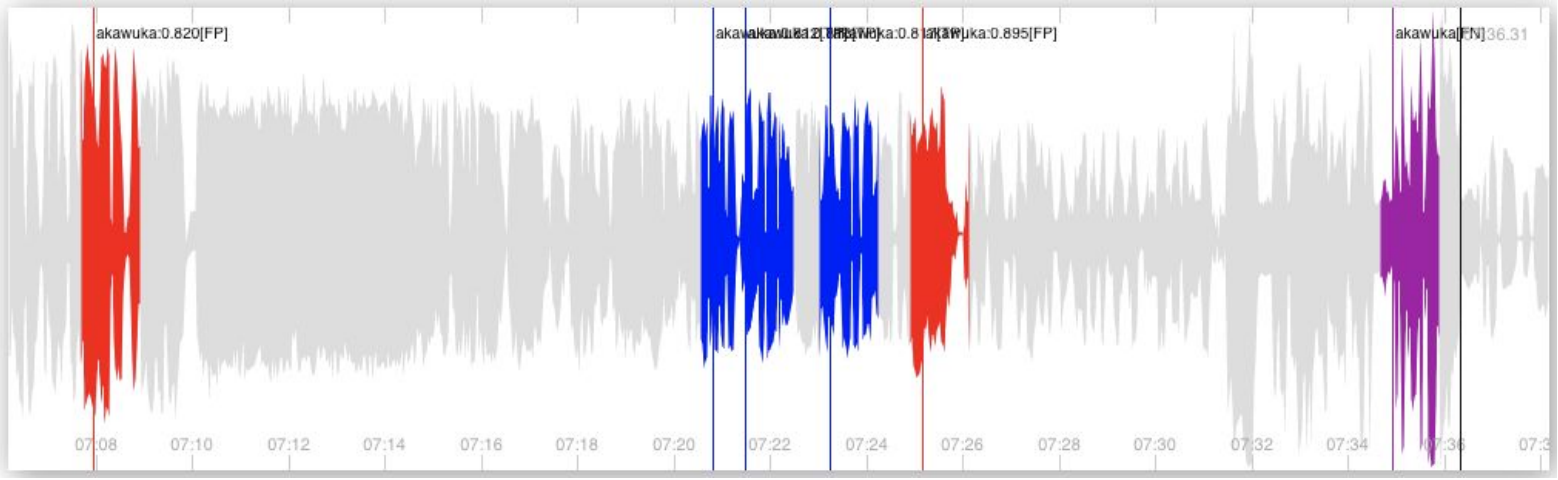
V2



Accuracy: **90%**

Precision & Recall

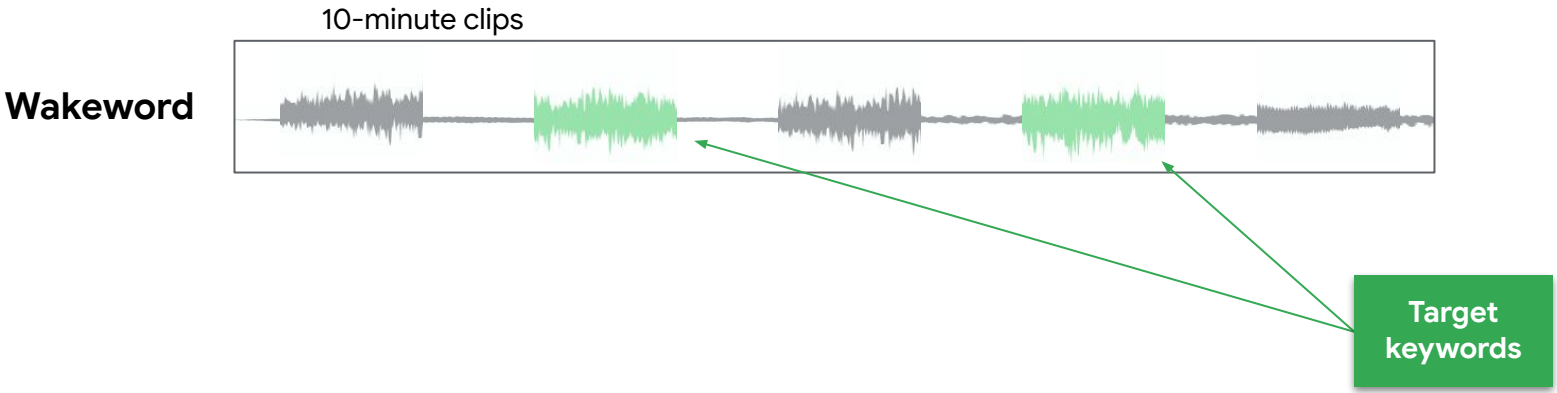
False Positive
True Positive
False Negative



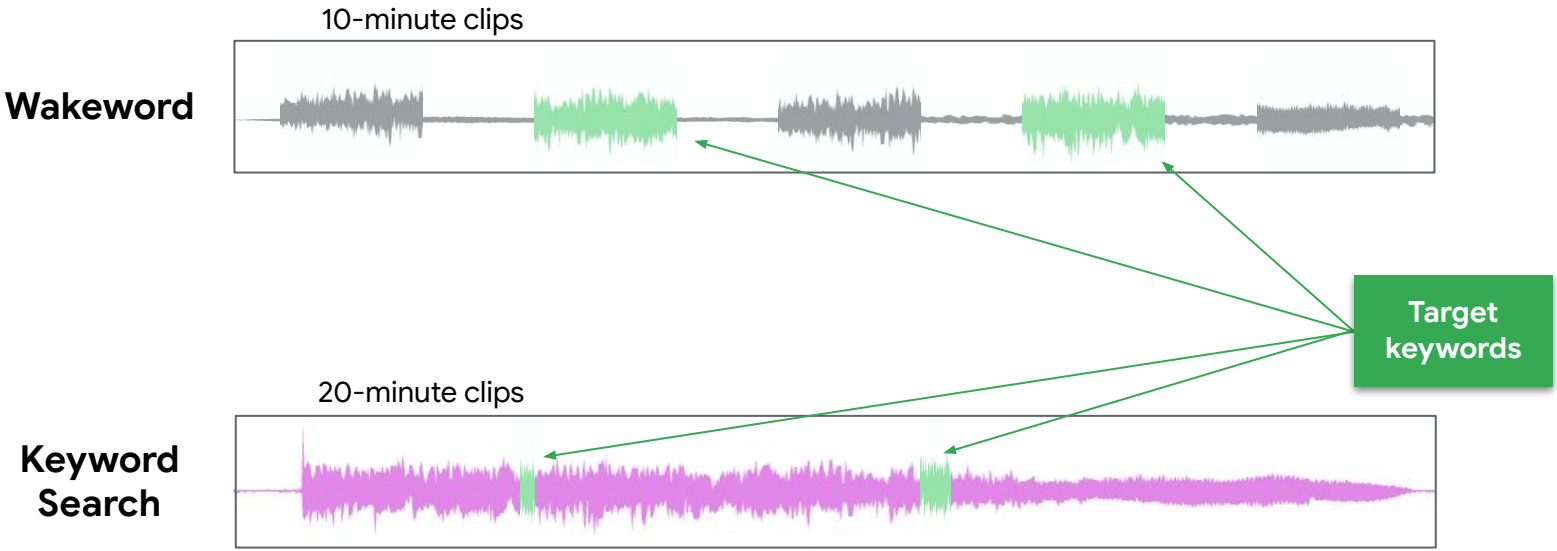
$$\text{Precision:} \quad \frac{TP}{TP + FN}$$

$$\text{Recall:} \quad \frac{TP}{TP + FP}$$

Streaming Accuracy Scenarios



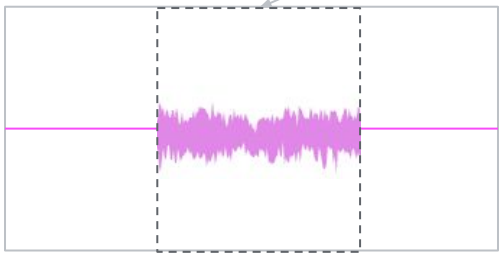
Streaming Accuracy Scenarios



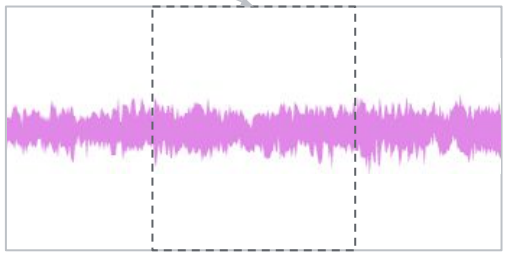
Extracting keywords with audio **context**

Keyword in Common Voice sentence

Improves keyword search accuracy

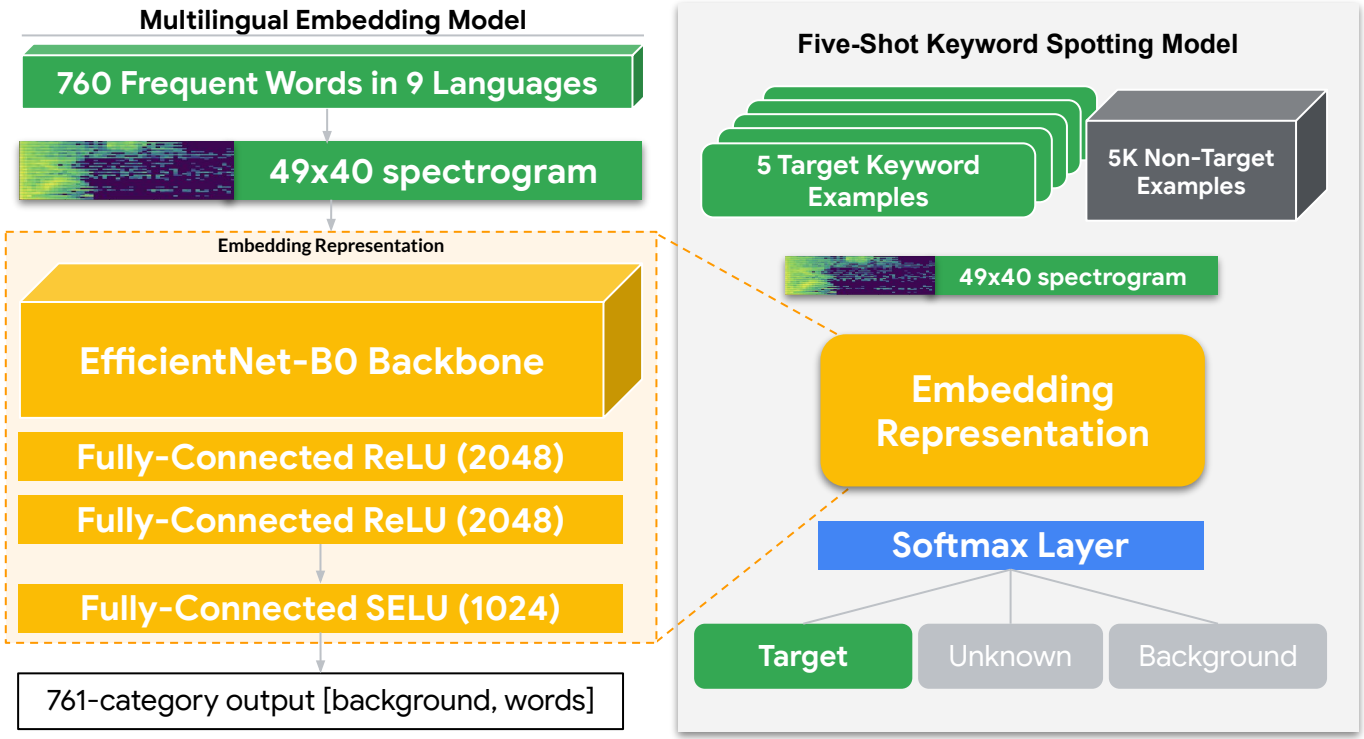


1-second **silence-padded** extraction

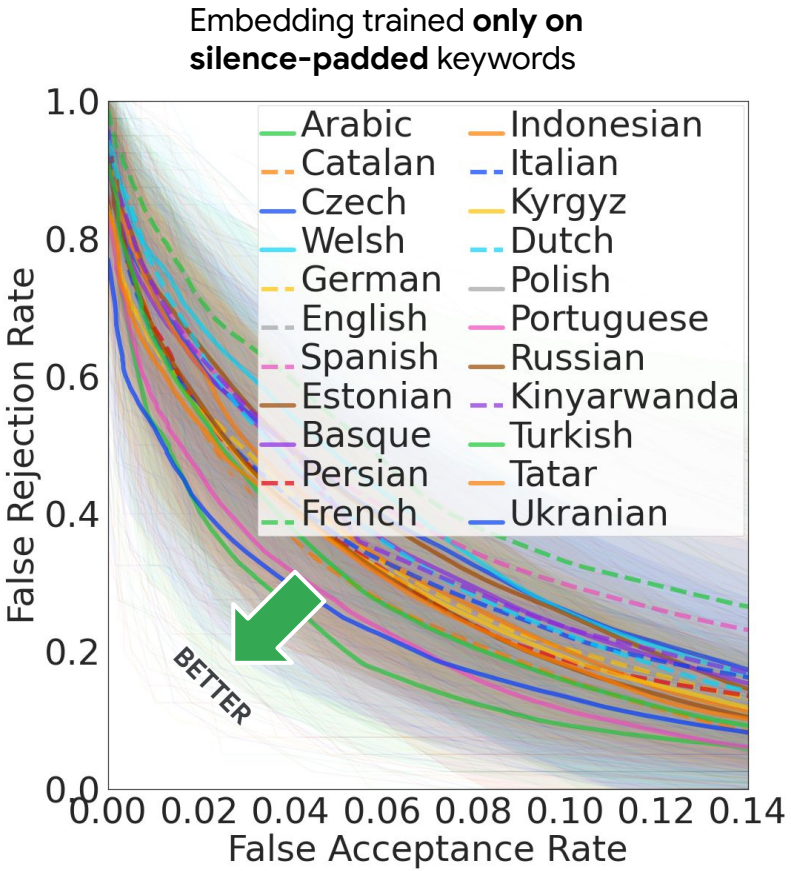


1-second **context-padded** extraction

Few-Shot Keyword Spotting

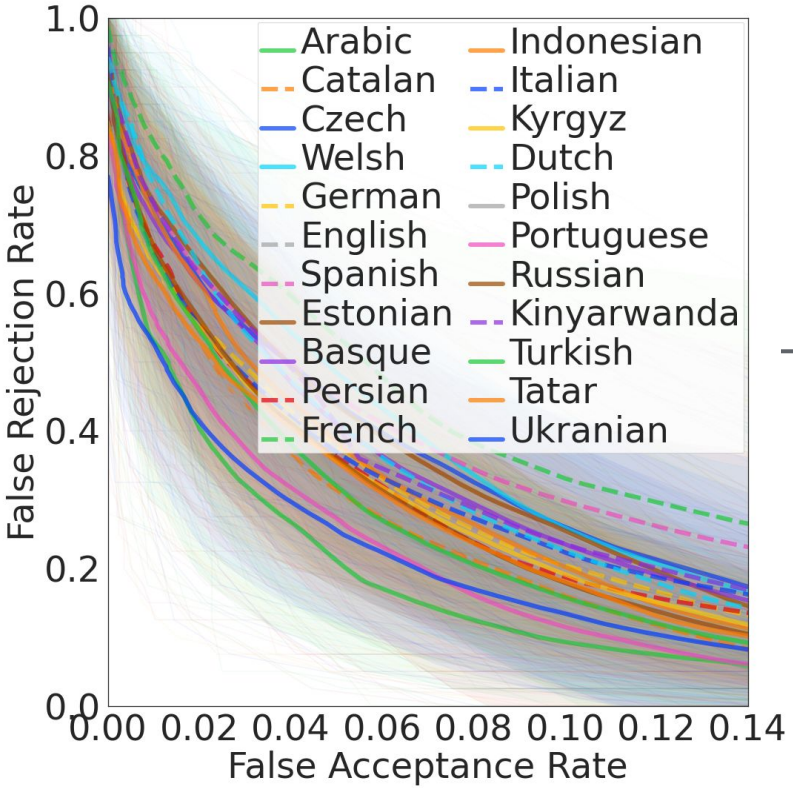


Streaming Accuracy on Keyword Search

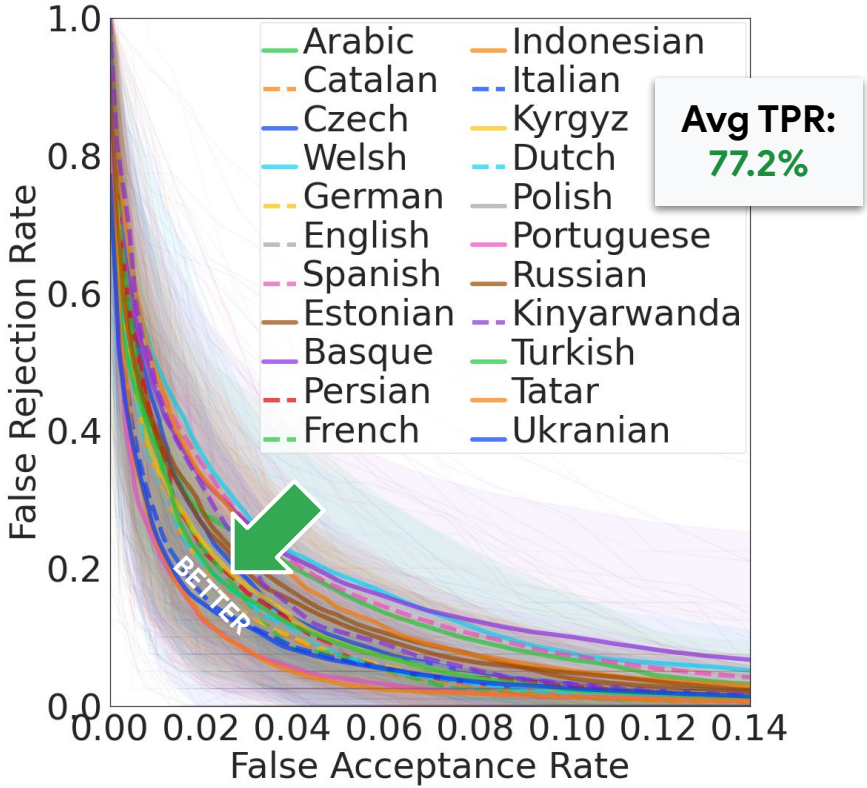


Streaming Accuracy on Keyword Search

Embedding trained **only on silence-padded** keywords



Embedding trained on **silence + context-padded** keywords



Continuous Training Metrics

Continuous training metrics **differ** from standard metrics

Continuous Training Metrics

Continuous training metrics **differ** from standard metrics

- Example goal: lift for a specific **challenging class**
  - Are **other classes negatively impacted**?
  - Might prefer to **accept a slight drop in global accuracy** to avoid catastrophic performance on an undersampled class
- Improve model latency through NAS?
- Evaluate model in additional scenarios