

Efficient Training and Labeling for Instrument Recognition using Active Learning



NEW YORK UNIVERSITY



Ana Elisa Mendez Mendez¹, Yu Wang¹, and Juan Pablo Bello^{1,2}

¹Music and Audio Research Laboratory, New York University

²Center for Urban Science and Progress, New York University

Please address correspondence to: anaelismendez@nyu.edu

wangyu@nyu.edu

Summary

Supervised learning is great! But labeling data...



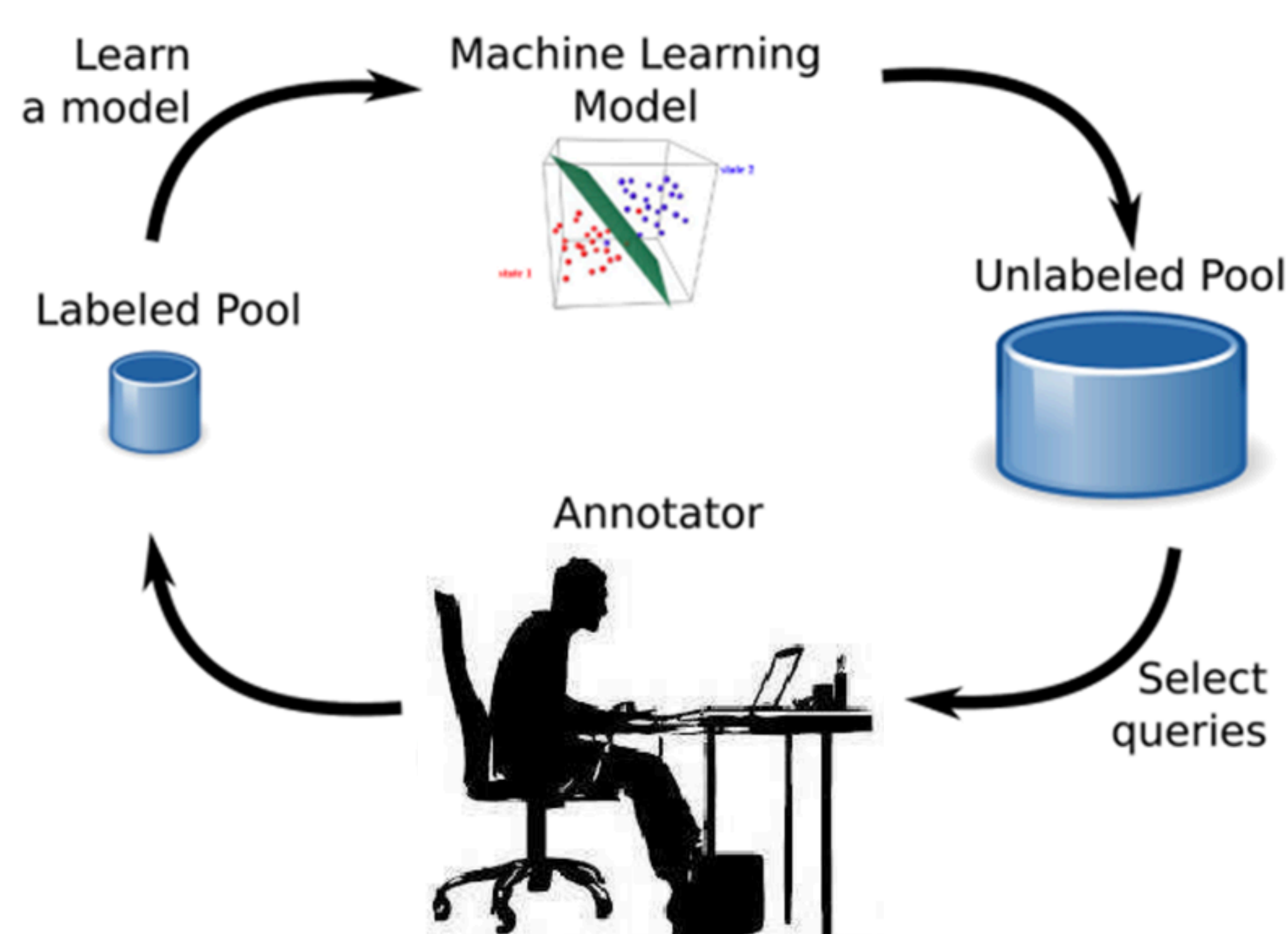
Active learning:
Good model performance
with much less labeled data!

Active Learning

► A machine learning method that **queries for labels of the most informative instances** to increase model performance.

► Training framework: In each training iteration,

1. Train the model with current labeled data.
2. Search in the unlabeled data pool and select the query using **least confident uncertainty sampling strategy**.
3. A **human annotator** listens to the queried audio example, labeling it, and adds it to the labeled data pool.



In this work:

- Task: **instrument recognition**
- Model: binary random forest classifier
- Initial labeled pool: **two data points**, one positive/one negative

Data

► OpenMIC-2018 dataset

- 20,000 examples of 10-second excerpts
- Partially labeled for presence/absence of 20 instrument classes
- Target instrument: **guitar**
- 1,650 labeled guitar examples, 1,137 positives, 513 negatives. The rest of the examples are used to create the unlabeled data pool

► SONYC dataset

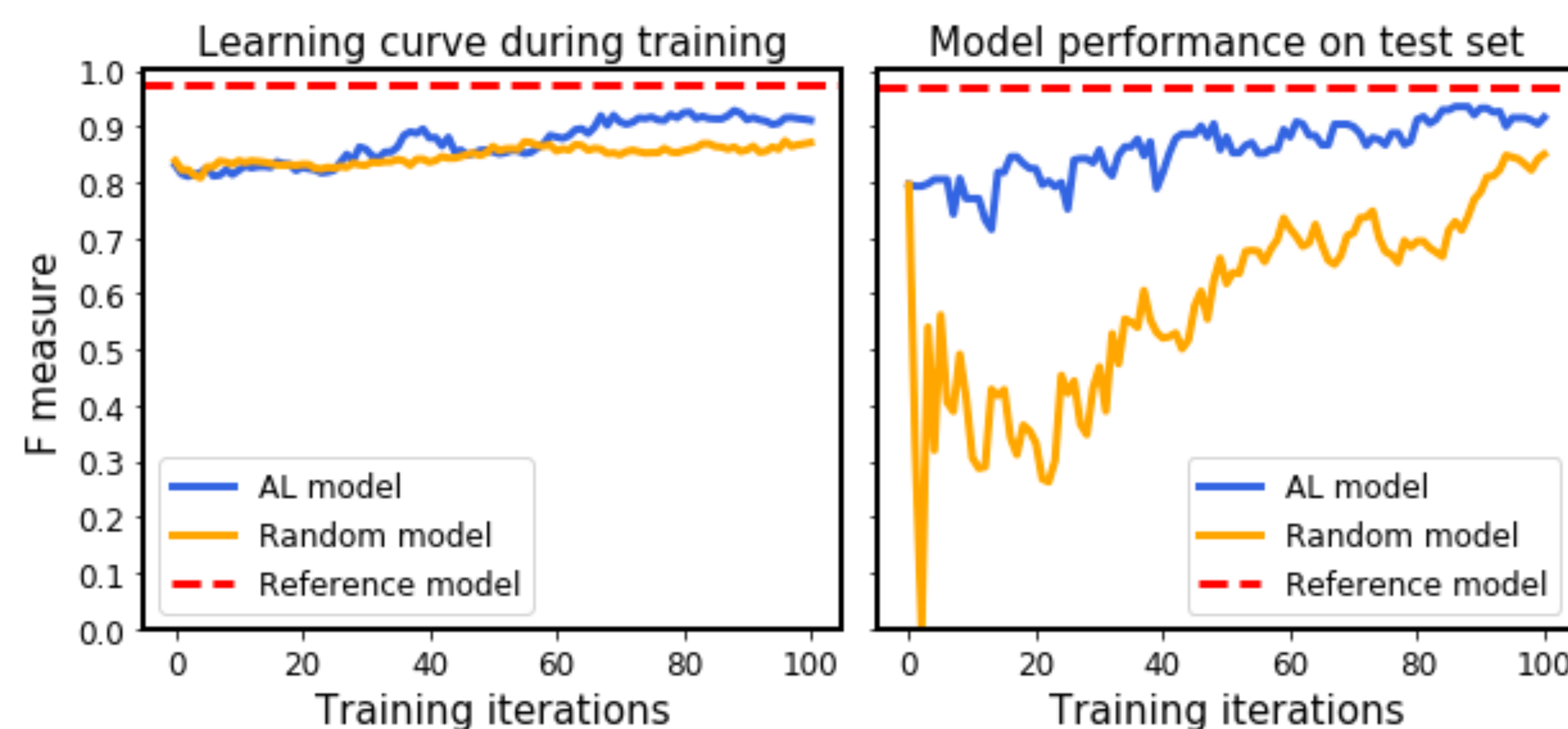
- Urban sound dataset collected from sensor network in NYC
- Target: **interference noise**

Experiments

► Input representation: **mean and std of VGGish features**

► Baseline models: random model and reference model

OpenMIC-2018 results:

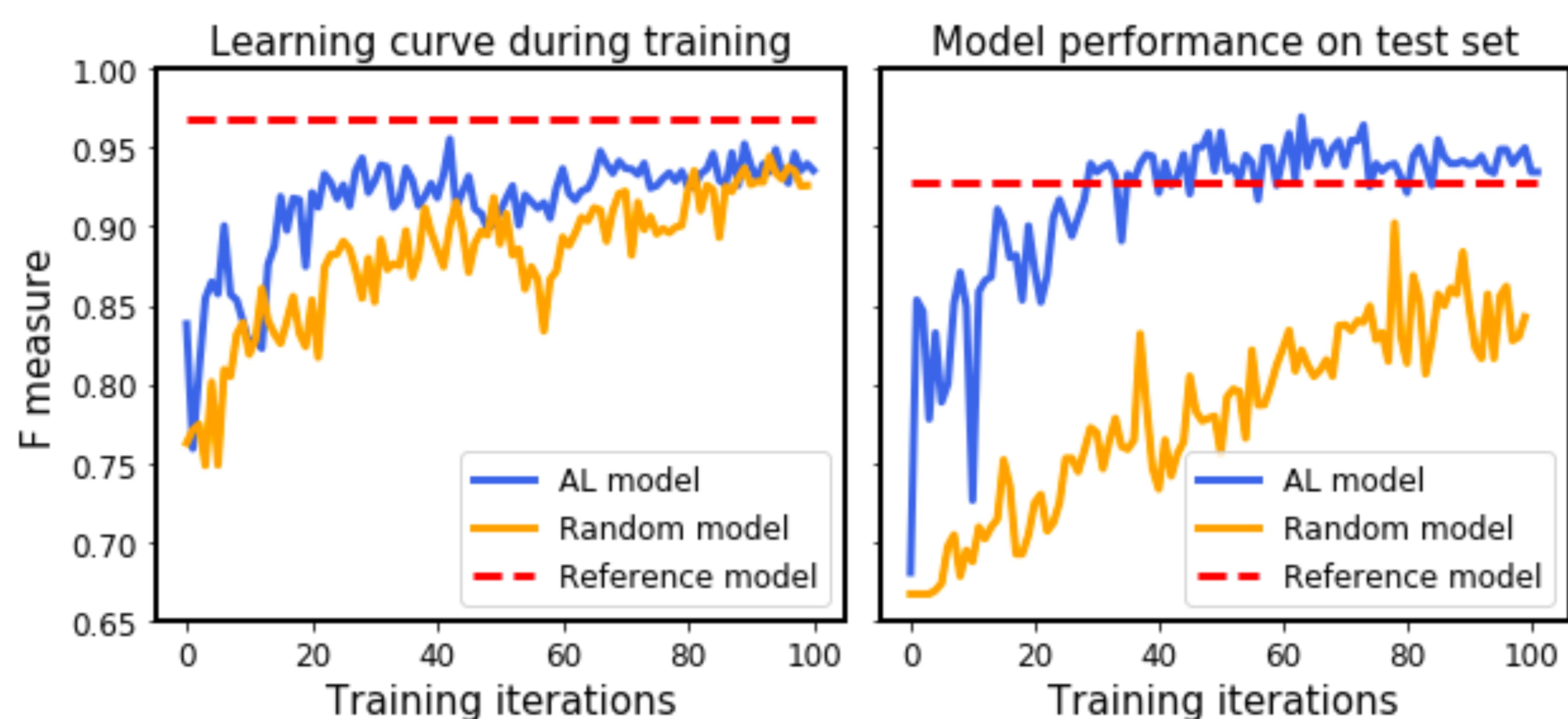


► AL model: **0.94** F-measure, trained on **88** examples.

► Random model: **0.73** F-measure, trained on **88** examples.

► Reference model: **0.97** F-measure, trained on **955** examples.

SONYC results:



► Queries returned from AL is much more balanced than those from random sampling.

	Positive examples in queries	
	OpenMic-2018	SONYC
AL	47%	47%
Random	36%	8%

Discussion and Future Work

► AL provides a more efficient training and labeling process when building a classifier on large unlabeled datasets.

► AL can be a useful tool for:

1. MIR tasks that involve **supervised learning + unlabeled/partially labeled datasets**
2. Building /labeling dataset
3. Crowdsourcing

► Key: **well defined and measurable "informativeness"**