SLURP DATASET ANALYSIS

黃科維 方佑心 陳剛頡
指導老師: 陳縕儂

Introduction

In SLU (Spoken Language Understanding) task, there are two main approaches to training:

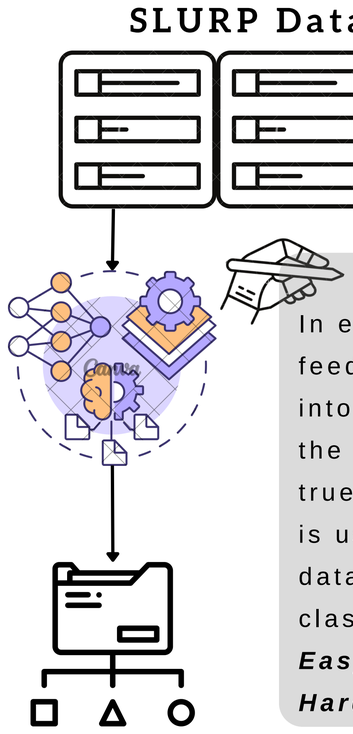
- **Pipeline method:** Use LM to run text-classification through ASR-translate text.
- **E2E method:** Train a end-to-end model from speech to labels.

The noise in dataset can be roughly divided into **ASR noise** and **Label noise**.

Contribution

Systematically identify ASR noise and label noise

Training Dynamics



Recording

In each training step, feed every instance into model to output the probability of the true label. Probability is used to divide dataset into three classes: ***Easy, Ambiguous and Hard.***

probability of the true label

$$\hat{\mu}_i = \frac{1}{E}\sum_{e=1}^{E} p_{\boldsymbol{\theta}^{(e)}}(y_i^* \mid \boldsymbol{x}_i)$$

Confidence
How certain is the model about this instance's prediction on the true label?

$$\hat{\sigma}_i = \sqrt{\frac{\sum_{e=1}^{E}\left(p_{\boldsymbol{\theta}^{(e)}}(y_i^* \mid \boldsymbol{x}_i) - \hat{\mu}_i\right)^2}{E}}$$

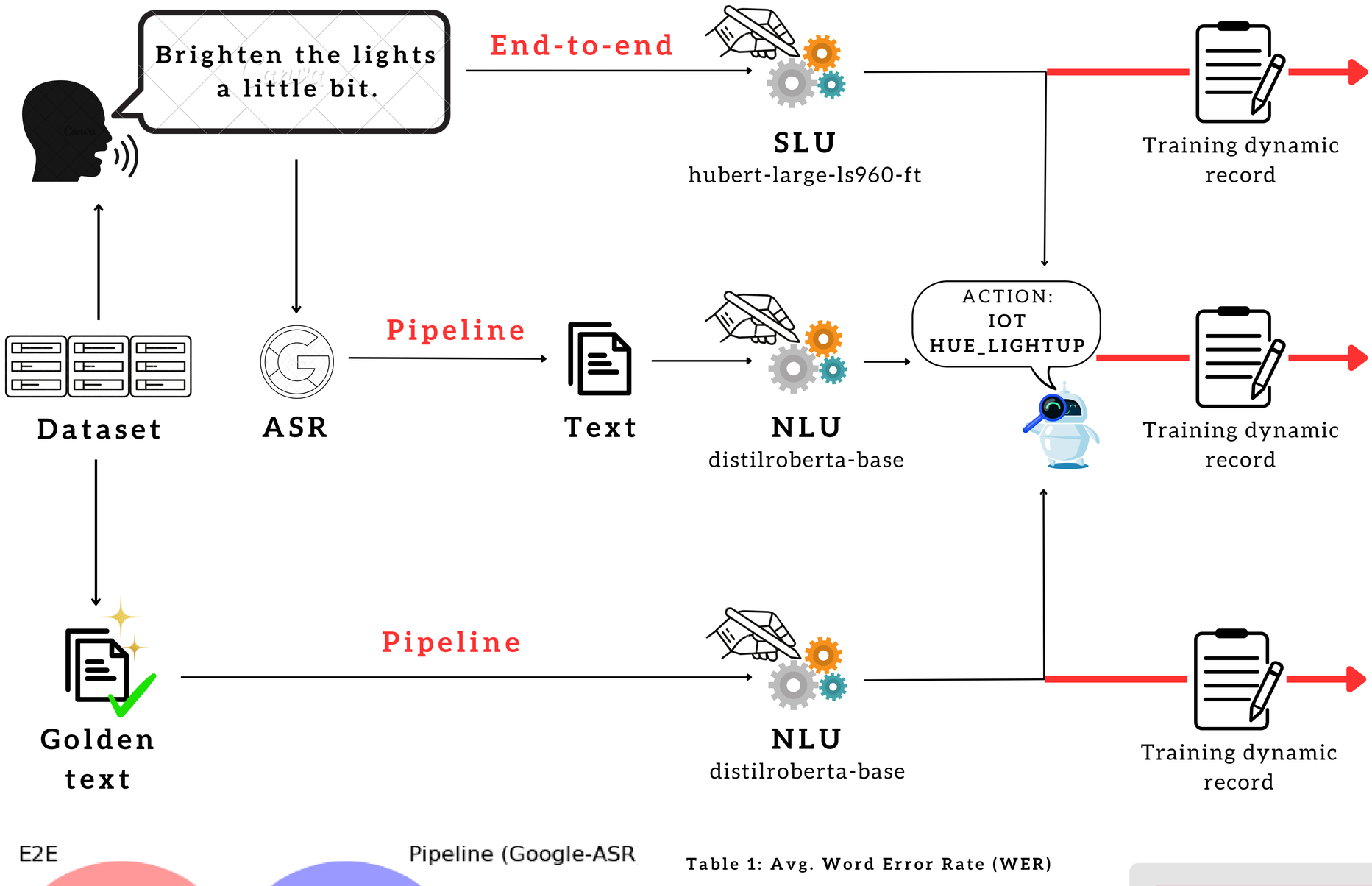
Variability
How challenging is this instance for the model?

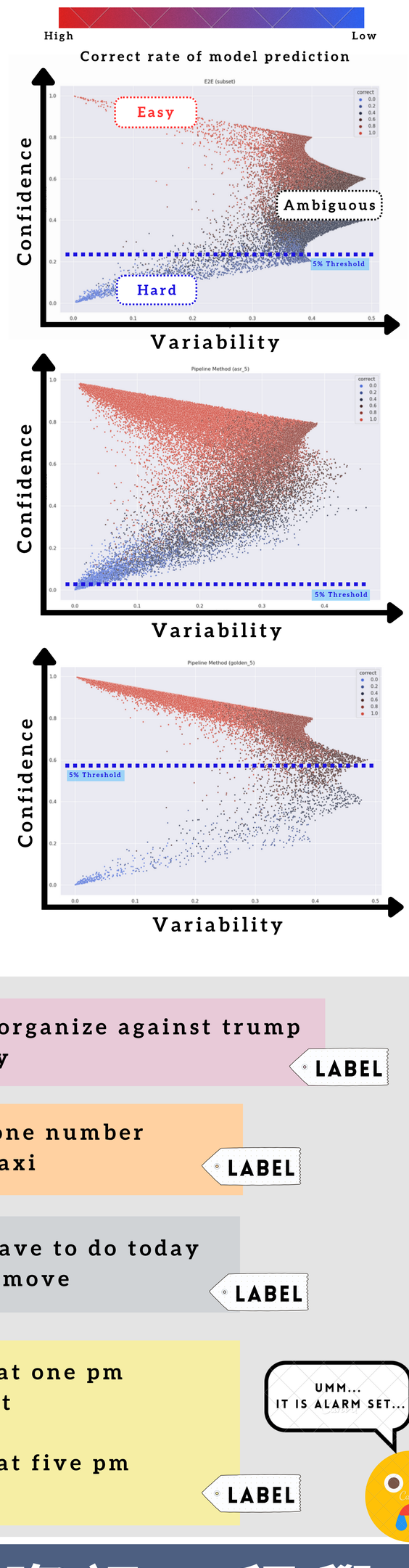
What is Easy, Ambiguous and Hard instances?

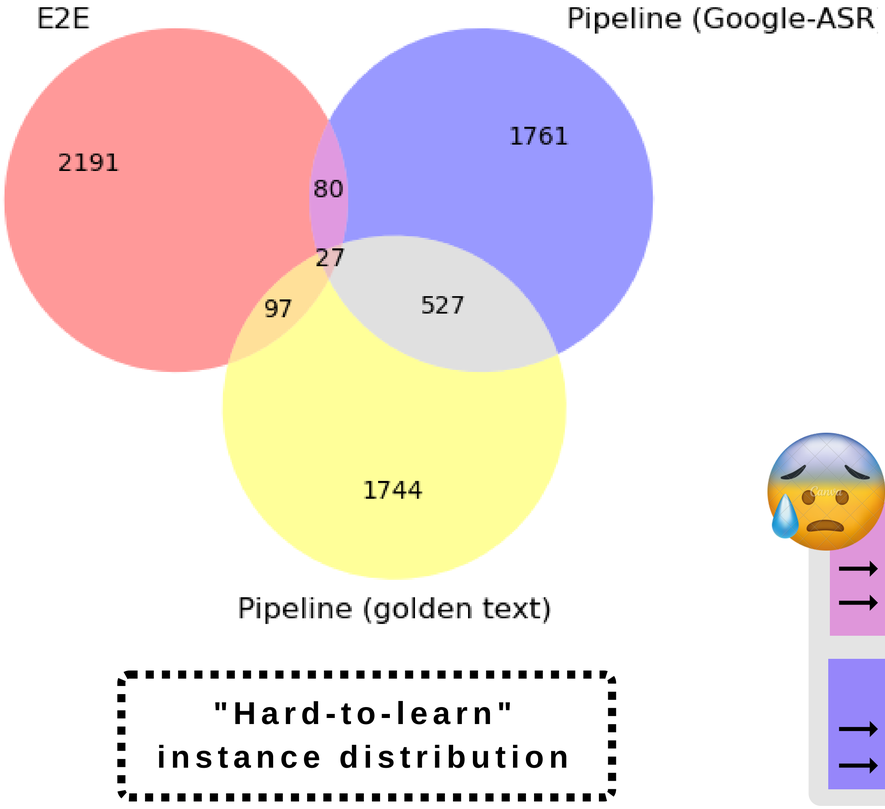
- Easy: High confidence. Help model to converge.
- Ambiguous: High variability. Useful for high performance.
- Hard: Low confidence. Noise usually appears here.

Framework

EACH APPROACH IS TRAINING ON 5 EPOCH,
AND TAKE HARD-TO-LEARN DATA WITH A CONFIDENCE LEVEL OF THE BOTTOM 5%.







"Hard-to-learn" instance distribution

Table 1: Avg. Word Error Rate (WER)

	Avg. WER		Avg. WER
purple	0.81	gray	0.46
blue	0.78	yellow	0.21
pink	0.36	red	0.22
orange	0.23		

could you delete this event
→ paducah police station
→ WER: 1.0
ASR WER 0.81

minimize the light
→ mameyes delights
→ WER: 1.0
ASR WER 0.77

how do we organize against trump
→ email_query
LABEL

call the phone number
→ transport_taxi
LABEL

what do i have to do today
→ calendar_remove
LABEL

remind me at one pm
→ calendar_set

remind me at five pm
→ alarm_set
LABEL

UMM... IT IS ALARM SET...

國立臺灣大學資訊工程學系