

In-Domain



Fixed
Vocabulary

LM_{IN}

LM_{OUT}

CE_{IN}

CE_{OUT}

$$\text{DiffCE} = \text{CE}_{\text{IN}} - \text{CE}_{\text{OUT}}$$

Out-of-Domain

$\text{Train}_{\text{OUT}}$

Test_{OUT}

Bin 1

Bin 2

Bin 3

Test_{OUT}

Bin N - 2

Bin N - 1

Bin N

Test_{OUT}

$\max(\text{diffCE})$
Test set closest to
in-domain data

Extraction of:
Grammar
Phrase Table
Paraphrase Table

Domain-Adapted Paraphrase
Table