Machine Translation English to German

By Yue Zhao and Cindy Lin

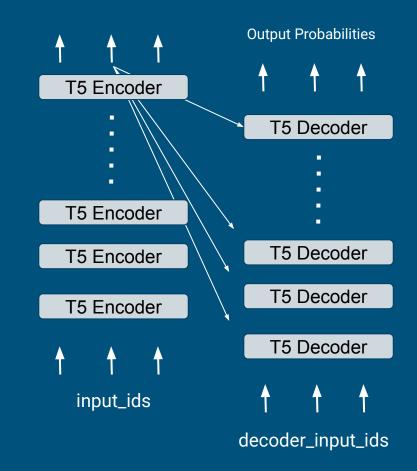
INTRODUCTION

Task: Machine Translation from English to German

Experimental setup:	
Data Selection: WMT14 English - German	Preprocessing:
 Size: 1.4GB Training set: 800 thousands bilingual sentence pairs Test set: 2000 examples Validation set: 2000 examples 	 First model: Pre-trained BERT tokenizer + target tokenizer in hw5 Second model: pre-trained T5 tokenizer
Architecture:	Metric:
 First model: Pre-trained BERT encoder + transformer decoder Second model: Fully pre-trained T5-base model + fine tuning 	Bleu, No human evaluation

Input and Output example: I don't speak German-> Ich spreche kein Deutsch

Output Probabilities MODELS Softmax Linear Add & Norm Feed Forward Add & Norm MHA Pre-trained **BERT** Add & Norm Model Masked MHA **Output Embedding** Inputs Outputs(shifted right)



RESULTS

Hyperparameters: Learning rate: 2e-4 Batch size: 24

After 318 runs on GCP, we got these results:

