FBK’s Multilingual Neural Machine Translation System for IWSLT 2017
Surafel M. Lakew1,2, Quintino F. Lotito2, Marco Turchi1, Matteo Negri1, Marcello Federico1
FBK-Fondazione Bruno Kessler1, University of Trento2, Trento, Italy
{lakew, lotito, turchi, negri, federico}@fbk.eu

Two Shared Task Participation
I. Multilingual Model: a translation task in a small data condition for 20 language directions.
II. Zero-Shot Model: a zero-shot translation (ZST) task in a similar small data condition (4 zero-shot & 16 non-zero-shot).

Overview
- 5 languages
- 20 possible directions
- ~ 200K available parallel sentences for each direction
- 5 artificial language tags (2en, 2de, 2nl, 2it, 2ro), used on the source side

Multilingual Setting
- 5 languages
- 20 possible directions
- ~ 200K available parallel sentences for each direction
- 5 artificial language tags (2en, 2de, 2nl, 2it, 2ro), used on the source side

Types of Models Trained
Single Language Pair:
- 20 single language pair models
Multilingual Model:
- utilized all the available dataset for 20 directions
Zero-Shot Model:
- excludes four pairs from the 20 direction training data, (Dutch<-->German, Italian<-->Romanian).

* All models are trained using the same parameter settings.

Zero-shot Translation Comparison

Results Comparing 20 Single-Pair, a Multilingual, and a Zero-Shot Model

Zero-shot Translation Methods

Conclusion
- The Multilingual system delivered better performance than the single language pair systems.
- The Zero-Shot model showed a comparable performance in the non-zero-shot directions with the Multilingual model.
- The Zero-Shot model outperformed the pivoting approach using Single-Pair models.